Energy Minimising Coarse Spaces for Multiscale Elliptic PDEs

Robert Scheichl

Bath Institute for Complex Systems
Department of Mathematical Sciences
University of Bath

in collaboration with

Ivan Graham & Jan Van lent (both Bath)

LMS Symposium on *Computational Linear Algebra for PDEs*,

Durham, Monday, July 21st 2008

Model Problem

- Elliptic PDE in 2D or 3D bounded domain Ω

$$-\nabla \cdot (\alpha \nabla u) = f \qquad + \quad u = 0 \text{ on } \partial\Omega$$

Model Problem

- Elliptic PDE in 2D or 3D bounded domain Ω

$$-\nabla \cdot (\alpha \nabla u) = f \qquad + \quad u = 0 \;\text{ on }\; \partial\Omega$$

- **Highly variable (discontinuous)** coefficients $\alpha(x)$

Model Problem

- Elliptic PDE in 2D or 3D bounded domain Ω

$$-\nabla \cdot (\alpha \nabla u) = f \qquad + \quad u = 0 \;\text{ on }\; \partial\Omega$$

- **Highly variable (discontinuous)** coefficients $\alpha(x)$
- FE discretisation (p.w. linears V^h) on mesh $\mathcal{T}^h$:

$$A\,\mathbf{u} = \mathbf{b} \qquad (n \times n \text{ linear system})$$

Model Problem

- Elliptic PDE in 2D or 3D bounded domain Ω

$$-\nabla \cdot (\alpha \nabla u) = f \qquad + \quad u = 0 \text{ on } \partial\Omega$$

- **Highly variable (discontinuous)** coefficients $\alpha(x)$
- FE discretisation (p.w. linears V^h) on mesh $\mathcal{T}^h$:

$$A\,\mathbf{u} = \mathbf{b} \qquad (n \times n \text{ linear system})$$

- **Aim:** Find efficient & robust preconditioner for A
(i.e. independent of variations in n and in $\alpha(x)$)

Model Problem

- Elliptic PDE in 2D or 3D bounded domain Ω

$$-\nabla \cdot (\alpha \nabla u) = f \qquad + \quad u = 0 \ \text{ on } \ \partial\Omega$$

- **Highly variable (discontinuous)** coefficients $\alpha(x)$
- FE discretisation (p.w. linears V^h) on mesh $\mathcal{T}^h$:

$$A\,\mathbf{u} = \mathbf{b} \qquad (n \times n \text{ linear system})$$

- **Aim:** Find efficient & robust preconditioner for A (i.e. independent of variations in n and in $\alpha(x)$)
- **Key Question:** **Choice of coarse space !**

Motivation: Groundwater Flow

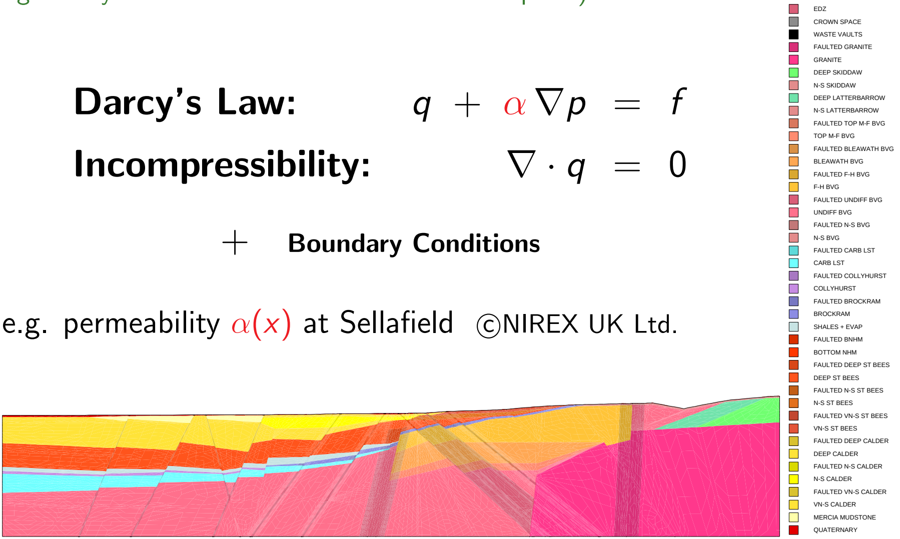
(e.g. safety assessment for radioactive waste disposal)

Darcy's Law: $\qquad q + {\color{red}\alpha} \nabla p = f$

Incompressibility: $\qquad \nabla \cdot q = 0$

$+$ **Boundary Conditions**

e.g. permeability ${\color{red}\alpha(x)}$ at Sellafield ©NIREX UK Ltd.



Relevant also for preconditioning of

- **Mixed FE discretisations** (Raviart-Thomas):

$$\begin{pmatrix} M & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$$

Relevant also for preconditioning of

- **Mixed FE discretisations** (Raviart-Thomas):

$$\begin{pmatrix} M & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$$

  - **Divergence-free reduction** $\longrightarrow$ standard pw.lin. FE for $-\nabla \cdot \left(\frac{1}{\alpha} \nabla \mathring{q}\right) = \mathring{g}$ (in 2D) **[Cliffe, Graham, RS, Stals; JCP 00]**

Relevant also for preconditioning of

- **Mixed FE discretisations** (Raviart-Thomas):

$$\begin{pmatrix} M & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$$

  - **Divergence-free reduction** $\longrightarrow$ standard pw.lin. FE for $-\nabla \cdot \left(\frac{1}{\alpha} \nabla \mathring{q}\right) = \mathring{g}$ (in 2D) **[Cliffe, Graham, RS, Stals; JCP 00]**
  - Scaled (2,2)-block in **Rusten–Winther** preconditioner: $BD_M^{-1}B^T \approx -\nabla^h \cdot (\alpha \nabla^h)$ **[Powell, Silvester; SIMAX 04]**

Relevant also for preconditioning of

- **Mixed FE discretisations** (Raviart-Thomas):

$$\begin{pmatrix} M & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix}$$

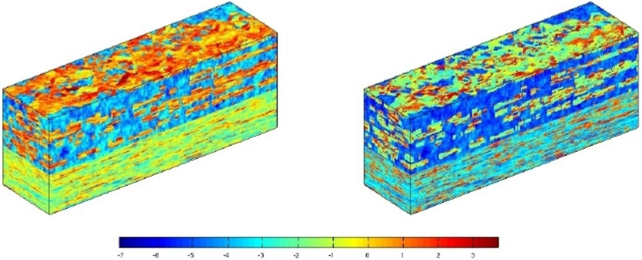
  - **Divergence-free reduction** $\longrightarrow$ standard pw.lin. FE for $-\nabla \cdot \left(\frac{1}{\alpha} \nabla \mathring{q}\right) = \mathring{g}$ (in 2D) **[Cliffe, Graham, RS, Stals; JCP 00]**
  - Scaled (2,2)-block in **Rusten–Winther** preconditioner: $BD_M^{-1}B^T \approx -\nabla^h \cdot (\alpha \nabla^h)$ **[Powell, Silvester; SIMAX 04]**

- **Two-phase porous media flow** (e.g. oil recovery):

$$\begin{pmatrix} A_{pp} & A_{ps} \\ A_{sp} & A_{ss} \end{pmatrix} \begin{pmatrix} \delta p \\ \delta s \end{pmatrix} = \mathrm{RHS}$$

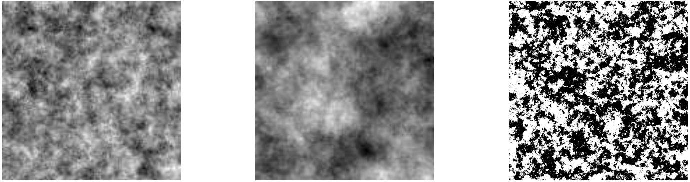
with $A_{pp} \approx -\nabla^h \cdot (\alpha \eta(s) \nabla^h) + \ldots$ a diffusion-type operator.

Heterogeneous multiscale deterministic media



Society of Petroleum Engineers (SPE) Benchmark SPE10

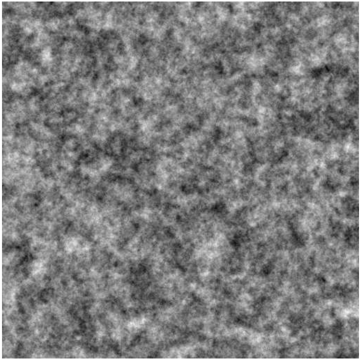
Multiscale stochastic media ($\lambda = 5h, 20h$ & clipped)



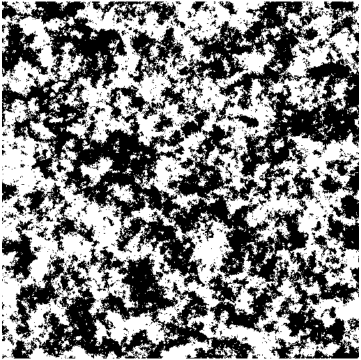
Stochastic Modelling of Heterogeneity

Log-normal random field $\alpha(x)$ (here with $n = 512^2$, $\lambda = 1/64$ and $\sigma^2 = 8$)

Log-normal Realisation



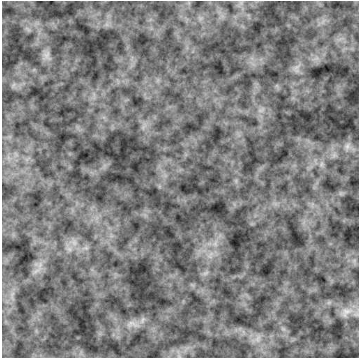
"Clipped" Realisation



Stochastic Modelling of Heterogeneity

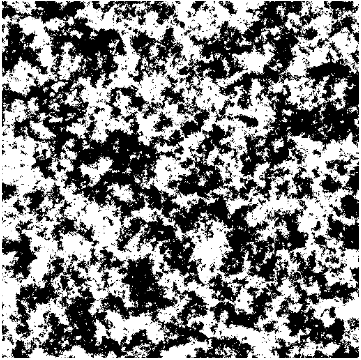
Log-normal random field $\alpha(x)$ (here with $n = 512^2$, $\lambda = 1/64$ and $\sigma^2 = 8$)

Log-normal Realisation



$$\max_{\tau,\tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}} \right) = O(10^{10}) \nearrow$$

"Clipped" Realisation



$$\max_{\tau,\tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}} \right) = O(10^{5}) \nearrow$$

Variance σ^2 determines **"contrast"** **!**

(**Correlation length** λ determines **"roughness"**)

Difficulties

- Requires very fine mesh resolution: $h \ll L$
- Complicated variation of $\alpha(x)$ on many scales
  (hard to **resolve** by (geometric) **coarse** mesh)
- A very large and very ill-conditioned, i.e.

$$\kappa(A) \lesssim \max_{\tau,\tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}} \right) h^{-2}$$

Difficulties

- Requires very fine mesh resolution: $h \ll L$
- Complicated variation of $\alpha(x)$ on many scales
  (hard to **resolve** by (geometric) **coarse** mesh)
- A very large and very ill-conditioned, i.e.

$$\kappa(A) \lesssim \max_{\tau,\tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}} \right) h^{-2}$$

Meaning of $\lesssim$

Difficulties

- Requires very fine mesh resolution: $h \ll L$
- Complicated variation of $\alpha(x)$ on many scales
  (hard to **resolve** by (geometric) **coarse** mesh)
- A very large and very ill-conditioned, i.e.

$$\kappa(A) \lesssim \max_{\tau,\tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}} \right) h^{-2}$$

Meaning of $\lesssim$

Goals

- Solver that is efficient, scalable, parallelisable AND
  - **robust** w.r.t. problem size n and mesh resolution h
  - as well as **robust** w.r.t. coefficients α **!!**
- Underpinning theory $\Longrightarrow$
  "Handle" for choice of components **!**

Possible Methods & Existing Theory

- Standard **Domain Decomposition** and **Multigrid** robust if coarse grid(s) **resolve(s)** coefficients
  [Chan, Mathew, Acta Numerica, 94], **[J. Xu, Zhu, Preprint, 07]**

Possible Methods & Existing Theory

- Standard **Domain Decomposition** and **Multigrid** robust if coarse grid(s) **resolve(s)** coefficients
  [Chan, Mathew, Acta Numerica, 94], **[J. Xu, Zhu, Preprint, 07]**
- Otherwise: **Coefficient-dependent** coarse space
  [Sarkis, Num Math, 97], **[Graham, Lechner, RS, Num Math, 07]**, **[RS, Vainikko, Comp. 07]**, **[Van lent, RS, Graham, submitted 08]**

Possible Methods & Existing Theory

- Standard **Domain Decomposition** and **Multigrid** robust if coarse grid(s) **resolve(s)** coefficients
  [Chan, Mathew, Acta Numerica, 94], **[J. Xu, Zhu, Preprint, 07]**
- Otherwise: **Coefficient-dependent** coarse space
  [Sarkis, Num Math, 97], **[Graham, Lechner, RS, Num Math, 07]**, **[RS, Vainikko, Comp. 07]**, **[Van lent, RS, Graham, submitted 08]**
- **FETI** also robust (with **correct scaling**)
  [Pechstein, RS, submitted, 08]

Possible Methods & Existing Theory

- Standard **Domain Decomposition** and **Multigrid** robust if coarse grid(s) **resolve(s)** coefficients
  [Chan, Mathew, Acta Numerica, 94], **[J. Xu, Zhu, Preprint, 07]**
- Otherwise: **Coefficient-dependent** coarse space
  [Sarkis, Num Math, 97], **[Graham, Lechner, RS, Num Math, 07]**, **[RS, Vainikko, Comp. 07]**, **[Van lent, RS, Graham, submitted 08]**
- **FETI** also robust (with **correct scaling**)
  [Pechstein, RS, submitted, 08]
- **Practically** most successful: **Algebraic Multigrid**
  No theory explaining coefficient robustness for standard AMG**!**

  Some first steps in **[Aksoylu, Graham, Klie, RS, Comp.Visual.Sci. 08]**

Possible Methods & Existing Theory

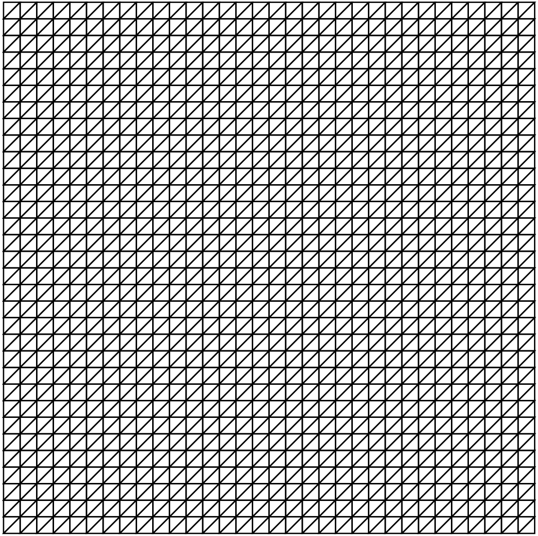
- Standard **Domain Decomposition** and **Multigrid** robust if coarse grid(s) **resolve(s)** coefficients
  [Chan, Mathew, Acta Numerica, 94], **[J. Xu, Zhu, Preprint, 07]**
- Otherwise: **Coefficient-dependent coarse space**
  [Sarkis, Num Math, 97], **[Graham, Lechner, RS, Num Math, 07]**,
  [RS, Vainikko, Comp. 07], [Van lent, RS, Graham, submitted 08]
- **FETI** also robust (with **correct scaling**)
  [Pechstein, RS, submitted, 08]
- **Practically** most successful: **Algebraic Multigrid**
  No theory explaining coefficient robustness for standard AMG**!**
  Some first steps in **[Aksoylu, Graham, Klie, RS, Comp.Visual.Sci. 08]**

Two-Level Overlapping Schwarz

Domain Decomposition

Two-Level Overlapping Schwarz Method

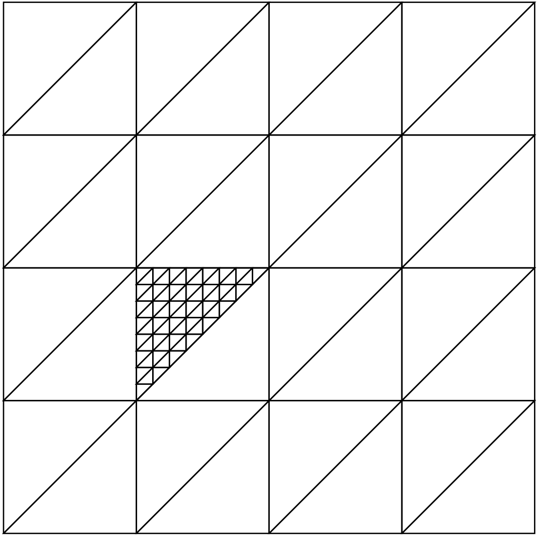
- Fine grid



Domain Decomposition

Two-Level Overlapping Schwarz Method

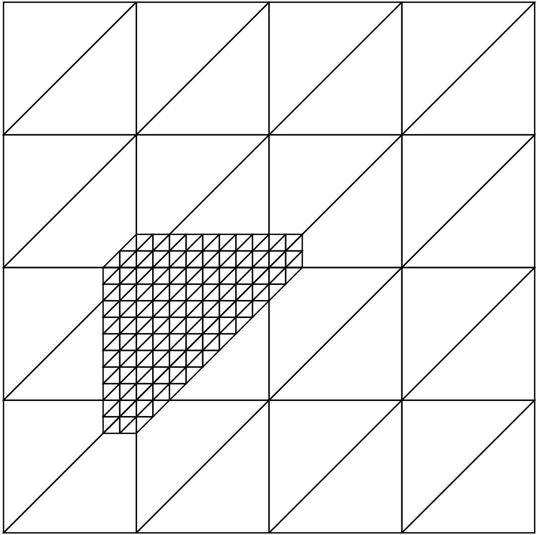
- Fine grid
- Subdomains



Domain Decomposition

Two-Level Overlapping Schwarz Method

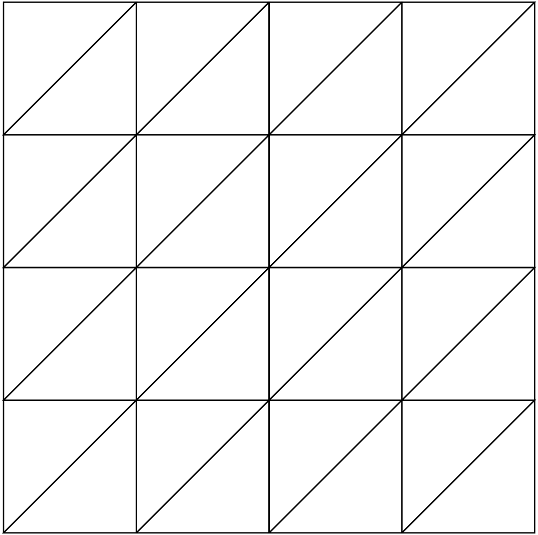
- Fine grid
- Subdomains
- Overlap



Domain Decomposition

Two-Level Overlapping Schwarz Method

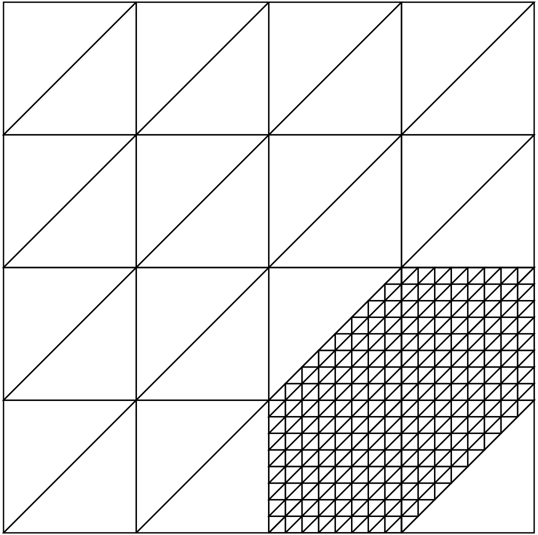
- Fine grid
- Subdomains
- Overlap
- Coarse "grid"



Domain Decomposition

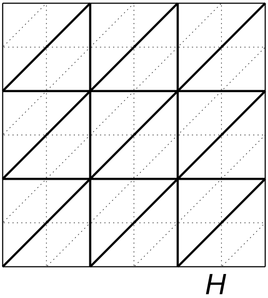
Two-Level Overlapping Schwarz Method

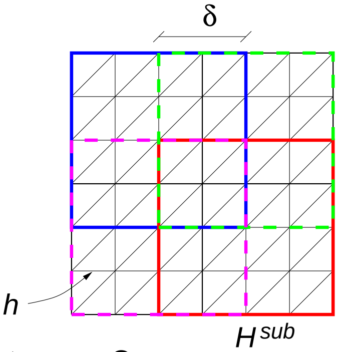
- Fine grid
- Subdomains
- Overlap
- Coarse "grid"
- Prolongation R_0^T (coarse space)



Two-Level Overlapping Additive Schwarz

$$M_{\mathrm{AS}} \;:=\; \underset{\substack{\uparrow\\ \textbf{coarse solve}\\ (A_0 \;:=\; R_0\,A\,R_0^T)}}{R_0^T\,A_0^{-1}\,R_0} \;+\; \sum_{i=1,\dots,S} \underset{\substack{\uparrow\\ \textbf{local solves}\\ (A_i \;:=\; R_i\,A\,R_i^T)}}{R_i^T\,A_i^{-1}\,R_i}$$

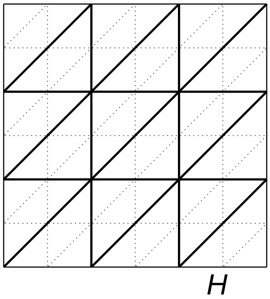


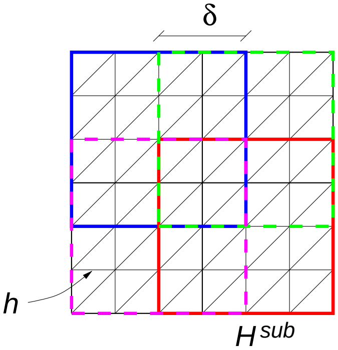


R_i simple injection into Ω_i, $i = 1, \dots, S$.

Two-Level Overlapping Additive Schwarz

$$M_{\mathrm{AS}} \;:=\; \underset{\substack{\uparrow \\ \textbf{coarse solve} \\ (A_0 \,:=\, R_0\, A\, R_0^T)}}{R_0^T\, A_0^{-1}\, R_0} \;+\; \sum_{i=1,\ldots,S} \underset{\substack{\uparrow \\ \textbf{local solves} \\ (A_i \,:=\, R_i\, A\, R_i^T)}}{R_i^T\, A_i^{-1}\, R_i}$$





But how to choose R_0^T **?**

Choice of prolongation R_0^T (abstract)

- Select a set of linearly independent FE functions (**coarse basis**)

$$\{\Psi_j \in V^h : j = 1, \ldots, N\}$$

s.t. $\sum_j \Psi_j \equiv 1$ and $\mathrm{supp}(\Psi_j) \subset \overline{\omega}_j$ (given)

(Assume ω_j overlaps at most η other supports $\longrightarrow$ A_0 sparse!)

Choice of prolongation R_0^T (abstract)

- Select a set of linearly independent FE functions (**coarse basis**)

$$\{\Psi_j \in V^h : j = 1, \ldots, N\}$$

s.t. $\sum_j \Psi_j \equiv 1$ and $\mathrm{supp}(\Psi_j) \subset \overline{\omega}_j$ (given)

(Assume ω_j overlaps at most η other supports $\longrightarrow$ A_0 sparse!)

- **Prolongation matrix**

$$R_0^T = [\mathbf{z}_1, \ldots, \mathbf{z}_N] \qquad \in \mathbb{R}^{n \times N}.$$

where $\mathbf{z}_j$ is coefficient vector corresponding to Ψ_j.

Coefficient-explicit Conditioning Analysis

Theorem (RS, Vainikko, Computing, 2007)

Let $\omega_j = \mathrm{supp}(\Psi_j)$. *Assume that each* $\omega_j \subset \Omega_i$ *for some* Ω_i *and that* $\{\omega_j\}$ *is a shape regular, uniformly overlapping, finite covering of* Ω. *Then*

$$\kappa(M_{\mathrm{AS}}A) \;\lesssim\; \gamma(\alpha)\left(1 + \max_j \frac{H_j}{\delta_j}\right)$$

where $H_j = \mathrm{diam}(\omega_j)$, $\delta_j = \mathit{overlap}(\omega_j)$ *and*

$$\gamma(\alpha) \;=\; \max_j \; \delta_j^2 \; \left\| \alpha |\nabla \Psi_j|^2 \right\|_{L_\infty(\Omega)},$$

the **"coarse space robustness indicator"**.

Coefficient-explicit Conditioning Analysis

Theorem (RS, Vainikko, Computing, 2007)

Let $\omega_j = \mathrm{supp}(\Psi_j)$. *Assume that each* $\omega_j \subset \Omega_i$ *for some* Ω_i *and that* $\{\omega_j\}$ *is a shape regular, uniformly overlapping, finite covering of* Ω. *Then*

$$\kappa(M_{\mathrm{AS}}A) \ \lesssim \ \gamma(\alpha) \left(1 + \max_j \frac{H_j}{\delta_j}\right)$$

where $H_j = \mathrm{diam}(\omega_j)$, $\delta_j = \mathit{overlap}(\omega_j)$ *and*

$$\gamma(\alpha) \ = \ \max_j \ \delta_j^2 \ \left\| \alpha |\nabla \Psi_j|^2 \right\|_{L_\infty(\Omega)},$$

the **"coarse space robustness indicator"**.

$\boxed{\text{\textbf{``Handle''} for choosing } \{\Psi_j\} \text{ w.r.t. } \alpha}$ $\longrightarrow$ **Energy minimisation!**

- **Robustness** possible even for **large variation** of coefficients inside coarse elements.

- **Robustness** possible even for **large variation** of coefficients inside coarse elements.
- **Previous theory** only for **resolved** coefficients!

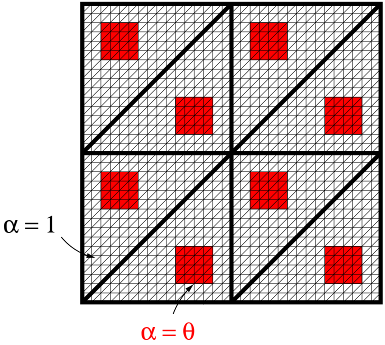
- **Robustness** possible even for **large variation** of coefficients inside coarse elements.
- **Previous theory** only for **resolved** coefficients!
- Importance of **energy minimisation** already noted in **AMG literature**, e.g.
  - [Mandel, Brezina, Vanek, Computing, 1999]
  - [Wan, Chan, Smith, SISC, 2000]
  - [Xu, Zikatanov, Comput Visual Sci, 2004]

- **Robustness** possible even for **large variation** of coefficients inside coarse elements.
- **Previous theory** only for **resolved** coefficients!
- Importance of **energy minimisation** already noted in **AMG literature**, e.g.
  - **[Mandel, Brezina, Vanek, Computing, 1999]**
  - **[Wan, Chan, Smith, SISC, 2000]**
  - **[Xu, Zikatanov, Comput Visual Sci, 2004]**
- Corollaries for **multiplicative** and **hybrid variants**:
  - **[Graham, RS, NMPDE, 2007]**

- **Robustness** possible even for **large variation** of coefficients inside coarse elements.
- **Previous theory** only for **resolved** coefficients!
- Importance of **energy minimisation** already noted in **AMG literature**, e.g.
  - [Mandel, Brezina, Vanek, Computing, 1999]
  - [Wan, Chan, Smith, SISC, 2000]
  - [Xu, Zikatanov, Comput Visual Sci, 2004]
- Corollaries for **multiplicative** and **hybrid variants**:
  - [Graham, RS, NMPDE, 2007]
- **Assumption** $\omega_j \subset \Omega_i$ **violated**: 2 indicators
  - [Graham, Lechner, RS, Numer Math, 2007]

Example: Standard pw. linear coarsening

(binary medium with $h = 1/512$ and $H = 8h$)

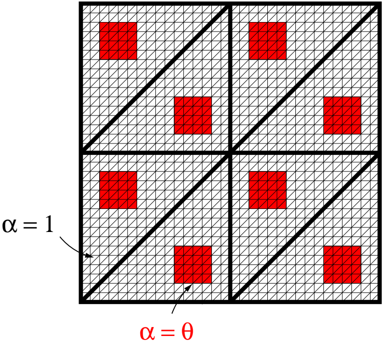


θ	$\kappa(M_{AS}A)$	$\gamma(\alpha)$
10^0	5.2	2
10^1	9.1	20
10^2	58.1	200
10^3	471	2000
10^4	1821	2.0(+4)
10^5	2561*	2.0(+5)

*same as one-level method (robust!)

Example: Standard pw. linear coarsening

(binary medium with $h = 1/512$ and $H = 8h$)



θ	$\kappa(M_{AS}A)$	$\gamma(\alpha)$
10^0	5.2	2
10^1	9.1	20
10^2	58.1	200
10^3	471	2000
10^4	1821	2.0(+4)
10^5	2561*	2.0(+5)

*same as one-level method (robust!)

$\longrightarrow$ **Coefficient-dependent Coarse Spaces**

Energy Minimising Coarse Spaces

Motivation – A bit of Linear Algebra

For simplicity assume $\mathcal{T}_h$ quasi-uniform, $\delta_j \sim H_j \sim H$ (generous overlap)

and α pw. constant $\quad \Longrightarrow \quad \alpha|\nabla\Psi_j|^2$ **also pw. constant w.r.t. $\tau \in \mathcal{T}_h$.**

Motivation – A bit of Linear Algebra

For simplicity assume $\mathcal{T}_h$ quasi-uniform, $\delta_j \sim H_j \sim H$ (generous overlap) and α pw. constant $\implies$ $\alpha|\nabla\Psi_j|^2$ also pw. constant w.r.t. $\tau \in \mathcal{T}_h$.

- Let Y be matrix with $Y_{\tau,j} := |\tau|^{1/2}\alpha_\tau^{1/2}|\nabla\Psi_j^\tau|$ and note that $Y^TY = R_0AR_0^T$.

Motivation – A bit of Linear Algebra

For simplicity assume $\mathcal{T}_h$ quasi-uniform, $\delta_j \sim H_j \sim H$ (generous overlap) and α pw. constant $\implies$ $\alpha|\nabla\Psi_j|^2$ also pw. constant w.r.t. $\tau \in \mathcal{T}_h$.

- Let Y be matrix with $Y_{\tau,j} := |\tau|^{1/2}\alpha_\tau^{1/2}|\nabla\Psi_j^\tau|$ and note that $Y^T Y = R_0 A R_0^T$.
- $\kappa(M_{AS}A) \lesssim H^2 \max_j \|\alpha|\nabla\Psi_j|^2\|_{L_\infty(\Omega)} \lesssim \frac{H^2}{h^d}\|Y\|_{\max}^2$

  where $\|B\|_{\max} := \max_{i,j}|B_{i,j}|$.

Motivation – A bit of Linear Algebra

For simplicity assume $\mathcal{T}_h$ quasi-uniform, $\delta_j \sim H_j \sim H$ (generous overlap) and α pw. constant $\implies$ $\alpha|\nabla\Psi_j|^2$ also pw. constant w.r.t. $\tau \in \mathcal{T}_h$.

- Let Y be matrix with $Y_{\tau,j} := |\tau|^{1/2}\alpha_\tau^{1/2}|\nabla\Psi_j^\tau|$ and note that $Y^TY = R_0AR_0^T$.
- $\kappa(M_{AS}A) \lesssim H^2 \max_j \|\alpha|\nabla\Psi_j|^2\|_{L_\infty(\Omega)} \lesssim \dfrac{H^2}{h^d}\|Y\|_{\max}^2$

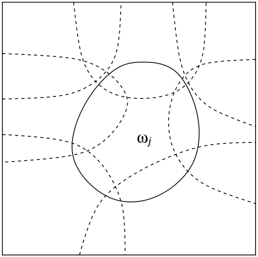
  where $\|B\|_{\max} := \max_{i,j}|B_{i,j}|$.
- **Idea:** Replace $\|Y\|_{\max}^2$ by Frobenius norm

  $$\|Y\|_F^2 = \mathrm{tr}(Y^TY) = \mathrm{tr}(R_0AR_0^T) = \sum_j \mathbf{z}_j^T A\,\mathbf{z}_j$$

  i.e. the **energy of the coarse space basis**.

Energy Minimising Coarse Space – Method

[Wan, Chan, Smith, SISC 2000]



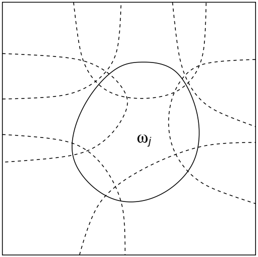
Choose **supports** $\{\omega_j\}$ for $\{\Psi_j\}$. Then solve **constrained minimisation problem** for $\{\mathbf{z}_j\}$:

$$\min \sum_j \mathbf{z}_j^T A\, \mathbf{z}_j \quad \text{s.t.} \quad \sum_j \mathbf{z}_j = \mathbf{1} \quad (\textbf{POU})$$

and $z_{j,i} \neq 0$ only if node $x_i \in \omega_j$ (**sparsity**).

Energy Minimising Coarse Space – Method

[Wan, Chan, Smith, SISC 2000]



Choose **supports** $\{\omega_j\}$ for $\{\Psi_j\}$. Then solve **constrained minimisation problem** for $\{\mathbf{z}_j\}$:

$$\min \sum_j \mathbf{z}_j^T A \, \mathbf{z}_j \quad \text{s.t.} \quad \sum_j \mathbf{z}_j = \mathbf{1} \quad (\textbf{POU})$$

and $z_{j,i} \neq 0$ only if node $x_i \in \omega_j$ (**sparsity**).

This problem has the **unique solution**

$$\mathbf{z}_j = \mathcal{R}_j^T \mathcal{A}_j^{-1} \mathcal{R}_j \boldsymbol{\lambda}$$

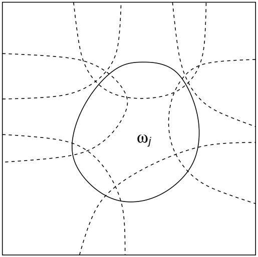
with $\mathcal{R}_j : \mathbb{R}^n \to \omega_j^h$ (**injection**) where $\boldsymbol{\lambda}$ is the solution of

$$B \, \boldsymbol{\lambda} = \mathbf{1} \qquad (\textbf{Lagrange multiplier system})$$

and $B := \sum_j \mathcal{R}_j^T \mathcal{A}_j^{-1} \mathcal{R}_j$ (**1-level Schwarz operator w.r.t.** $\{\omega_j\}$)

Energy Minimising Coarse Space – Method

[Wan, Chan, Smith, SISC 2000]



Choose **supports** $\{\omega_j\}$ for $\{\Psi_j\}$. Then solve **constrained minimisation problem** for $\{\mathbf{z}_j\}$:

$$\min \sum_j \mathbf{z}_j^T A\, \mathbf{z}_j \quad \text{s.t.} \quad \sum_j \mathbf{z}_j = \mathbf{1} \quad (\textbf{POU})$$

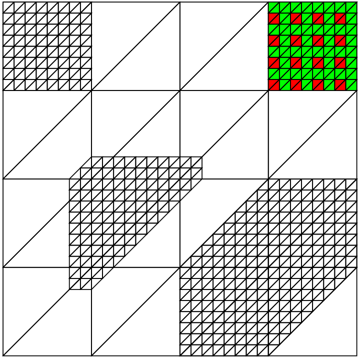
and $z_{j,i} \neq 0$ only if node $x_i \in \omega_j$ (**sparsity**).

Method/Questions

- Select a set of supports (patches) $\{\omega_j\}$. **How? Important?**
- Solve (global) problem $B\boldsymbol{\lambda} = \mathbf{1}$. **How? Cost?**
- Solve (local) problems $\mathcal{A}_j \mathbf{y}_j = \mathcal{R}_j \boldsymbol{\lambda}$. Set $\mathbf{z}_j = \mathcal{R}_j^T \mathbf{y}_j$. **Cheap!**

Example: Fine Scale Binary Medium

PCG-Iterations to solve $A\mathbf{u} = \mathbf{b}$ with preconditioner M_{AS} ($\varepsilon = 10^{-6}$)



$\alpha = 1$ and $\alpha = \theta$

$H = 8h$

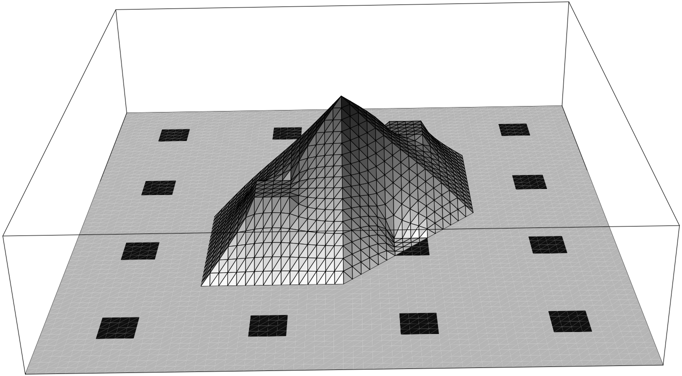
$n = 256^2$

θ	1-level	linear	minim.
10^0	79	13	13
10^2	80	63	14
10^4	84	113	14
10^6	87	115	14

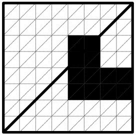
$\theta = 10^6$

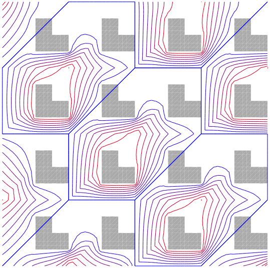
n	1-level	linear	minim.
32^2	15	19	13
64^2	23	33	14
128^2	45	59	14
256^2	87	115	14

A Typical Basis Function



Example: L-shaped inclusion

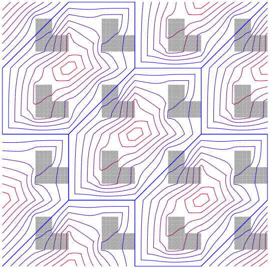




Minimising

$n = 256^2$

θ	linear	multiFE	minim.
10^0	13	13	13
10^2	64	46	29
10^4	167	157	32
10^6	179	162	32



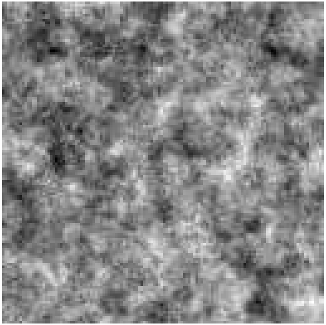
Multiscale FE

$\theta = 10^6$

n	linear	multiFE	minim.
32^2	29	26	21
64^2	48	48	26
128^2	89	88	29
256^2	179	162	32

Example: Lognormal Random Field

PCG-Iterations to solve $A\mathbf{u} = \mathbf{b}$ with preconditioner M_{AS} ($\varepsilon = 10^{-6}$)

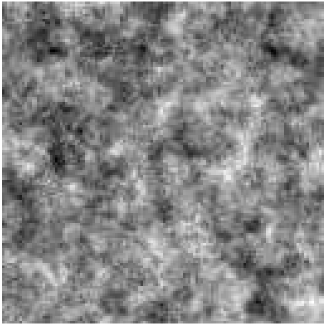


$n = 256^2$ and $\lambda = 4h$

σ^2	$\max_{\tau,\tau'} \frac{\alpha_\tau}{\alpha_{\tau'}}$	linear	multiFE	minim.
0	1.0	22	22	23
4	3.3(+7)	65	46	44
8	5.2(+10)	121	65	62
12	1.6(+13)	199	86	81
				$\uparrow$
				Not robust!

Example: Lognormal Random Field

PCG-Iterations to solve $A\mathbf{u} = \mathbf{b}$ with preconditioner M_{AS} ($\varepsilon = 10^{-6}$)



$n = 256^2$ and $\lambda = 4h$

σ^2	$\max_{\tau,\tau'} \frac{\alpha_\tau}{\alpha_{\tau'}}$	linear	multiFE	minim.
0	1.0	22	22	23
4	3.3(+7)	65	46	44
8	5.2(+10)	121	65	62
12	1.6(+13)	199	86	81

$\uparrow$

Not robust!

Need to adapt supports $\{\omega_j\}$ (see below)!

How to solve $B\boldsymbol{\lambda} = \mathbf{1}$?

Use **structure** of B (i.e. $B \approx A^{-1}$ and **“almost local”**):

How to solve $B\boldsymbol{\lambda} = \mathbf{1}$?

Use **structure** of B (i.e. $B \approx A^{-1}$ and **“almost local”**):

- **Idea 1:** precondition B with A **[Wan, Chan, Smith, 00]**

  Problem: $\kappa(AB) = \kappa(BA)$ **!** (i.e. like **1-level prec. for** A)

How to solve $B\boldsymbol{\lambda} = \mathbf{1}$?

Use **structure** of B (i.e. $B \approx A^{-1}$ and **"almost local"**):

- **Idea 1:** precondition B with A **[Wan, Chan, Smith, 00]**

  Problem: $\kappa(AB) = \kappa(BA)$ **!** (i.e. like **1-level prec. for** A)

- **Idea 2:** precondition B with $D = \mathrm{diag}(B)$ **[Xu, Zikatanov, 04]**

  Problem: only **robust** (w.r.t. h and α) for $H \sim \delta \sim h$ **!** (MG)

How to solve $B\boldsymbol{\lambda} = \mathbf{1}$?

Use **structure** of B (i.e. $B \approx A^{-1}$ and **"almost local"**):

- **Idea 1:** precondition B with A **[Wan, Chan, Smith, 00]**

  Problem: $\kappa(AB) = \kappa(BA)$ **!** (i.e. like **1-level prec. for** A)

- **Idea 2:** precondition B with $D = \mathrm{diag}(B)$ **[Xu, Zikatanov, 04]**

  Problem: only **robust** (w.r.t. h and α) for $H \sim \delta \sim h$ **!** (MG)

- **New Idea:** $\boxed{\text{1-level}}$ **additive Schwarz preconditioner** for B

$$C := \sum_j \mathcal{R}_j^T B_j^{-1} \mathcal{R}_j \quad \text{where} \quad B_j := \mathcal{R}_j B \mathcal{R}_j^T$$

Sounds impractical as B_j is **dense**. $\boxed{\textbf{But } B_j^{-1} \textbf{ is sparse(ish)!}}$

Implement using Linear Algebra "Trick"

- Consider ω_j overlapping only 2 neighbours ω_k & ω_l.
- **Local block**
$$B_j = A_j^{-1} + \hat{I}_{jk} A_k^{-1} \hat{I}_{jk}^T + \hat{I}_{jl} A_l^{-1} \hat{I}_{jl}^T$$

Implement using Linear Algebra "Trick"

- Consider ω_j overlapping only 2 neighbours ω_k & ω_l.
- **Local block**

$$B_j = A_j^{-1} + \hat{I}_{jk} A_k^{-1} \hat{I}_{jk}^T + \hat{I}_{jl} A_l^{-1} \hat{I}_{jl}^T$$

- Invert via **Sherman-Morisson-Woodbury** formula:

$$B_j^{-1} \; := \; A_j - A_j \begin{bmatrix} \hat{I}_{jk} & \hat{I}_{jl} \end{bmatrix} G_j^{-1} \begin{bmatrix} \hat{I}_{jk}^T \\ \hat{I}_{jl}^T \end{bmatrix} A_j$$

where

$$G_j = \begin{bmatrix} A_k & \\ & A_l \end{bmatrix} + \begin{bmatrix} \hat{I}_{jk}^T \\ \hat{I}_{jl}^T \end{bmatrix} A_j \begin{bmatrix} \hat{I}_{jk} & \hat{I}_{jl} \end{bmatrix}$$

Implement using Linear Algebra "Trick"

- Consider ω_j overlapping only 2 neighbours ω_k & ω_l.
- **Local block**

$$B_j = A_j^{-1} + \hat{I}_{jk} A_k^{-1} \hat{I}_{jk}^T + \hat{I}_{jl} A_l^{-1} \hat{I}_{jl}^T$$

- Invert via **Sherman-Morisson-Woodbury** formula:

$$B_j^{-1} := A_j - A_j \begin{bmatrix} \hat{I}_{jk} & \hat{I}_{jl} \end{bmatrix} G_j^{-1} \begin{bmatrix} \hat{I}_{jk}^T \\ \hat{I}_{jl}^T \end{bmatrix} A_j$$

where

$$G_j = \begin{bmatrix} A_k & \\ & A_l \end{bmatrix} + \begin{bmatrix} \hat{I}_{jk}^T \\ \hat{I}_{jl}^T \end{bmatrix} A_j \begin{bmatrix} \hat{I}_{jk} & \hat{I}_{jl} \end{bmatrix}$$

- **Cost:** 1 sparse (local) solve & some sparse MatVec

Robustness of Preconditioner for B

(Preconditioned CG iterations for $B\boldsymbol{\lambda} = \mathbf{1}$; tolerance $\varepsilon = 10^{-6}$)

Binary medium

θ	A	D	C
10^0	53	43	10
10^2	70	108	10
10^4	71	119	9
10^6	71	37	9

n	A	D	C
32^2	18	33	10
64^2	31	37	10
128^2	52	38	10
256^2	71	37	9

Log-normal field

σ^2	A	D	C
0	38	44	10
2	96	94	13
4	138	164	14
8	200^+	200^+	15
12	200^+	200^+	16
16	200^+	200^+	16

Robustness of Preconditioner for B

(Preconditioned CG iterations for $B\boldsymbol{\lambda} = \mathbf{1}$; tolerance $\varepsilon = 10^{-6}$)

Binary medium

θ	A	D	C
10^0	53	43	10
10^2	70	108	10
10^4	71	119	9
10^6	71	37	9

n	A	D	C
32^2	18	33	10
64^2	31	37	10
128^2	52	38	10
256^2	71	37	9

Log-normal field

σ^2	A	D	C
0	38	44	10
2	96	94	13
4	138	164	14
8	200^+	200^+	15
12	200^+	200^+	16
16	200^+	200^+	16

In practice sufficient to solve $B\boldsymbol{\lambda} = \mathbf{1}$ **inaccurately!**

Energy Minimisation – Comments

- Coarse grid construction is **scalable**, i.e. $O(n)$
  (more costly than other coarsening strategies, but **"best"!**)

Energy Minimisation – Comments

- Coarse grid construction is **scalable**, i.e. $O(n)$
  (more costly than other coarsening strategies, but **“best”!**)

- **Very important: Choice of supports** $\{\omega_j\}$**!?**
  (e.g. aggregation (strong connections), compatible relaxation)

Energy Minimisation – Comments

- Coarse grid construction is **scalable**, i.e. $O(n)$
  (more costly than other coarsening strategies, but **"best"!**)
- **Very important: Choice of supports** $\{\omega_j\}$**!?**
  (e.g. aggregation (strong connections), compatible relaxation)
- **Analysis** of $\kappa(CB)$ possible (via Schwarz theory)
  (collaboration with L. Zikatanov (Penn State))

Energy Minimisation – Comments

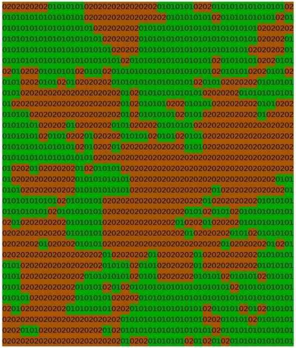
- Coarse grid construction is **scalable**, i.e. $O(n)$
  (more costly than other coarsening strategies, but **"best"!**)
- **Very important: Choice of supports** $\{\omega_j\}$**!?**
  (e.g. aggregation (strong connections), compatible relaxation)
- **Analysis** of $\kappa(CB)$ possible (via Schwarz theory)
  (collaboration with L. Zikatanov (Penn State))
- Also interesting for **Linear Elasticity** and **Maxwell**
  [Mandel et al, Comp. 99], **[Musy, Nicolas, Perrussel, SISC 07]**

Energy Minimisation – Comments

- Coarse grid construction is **scalable**, i.e. $O(n)$
  (more costly than other coarsening strategies, but **"best"!**)
- **Very important: Choice of supports** $\{\omega_j\}$**!?**
  (e.g. aggregation (strong connections), compatible relaxation)
- **Analysis** of $\kappa(CB)$ possible (via Schwarz theory)
  (collaboration with L. Zikatanov (Penn State))
- Also interesting for **Linear Elasticity** and **Maxwell**
  [Mandel et al, Comp. 99], **[Musy, Nicolas, Perrussel, SISC 07]**
- Potential also for **numerical homogenisation & upscaling** (especially if there is **no scale separation**!)

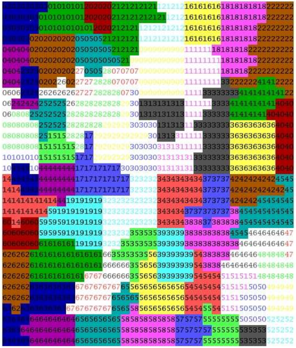
Choosing Supports via Aggregation

(based on strong connections in matrix – like in AMG)



Coefficient field (clipped)

$(n = 32^2,\ \lambda = \frac{1}{8},\ \max \frac{\alpha_\tau}{\alpha_{\tau'}} \approx 10^3)$

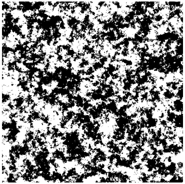


Aggregates (ADOUG)

(radius $r = 2$, threshold $\beta = 0.67$)

Example: Clipped random field

($n = 256^2$ and $\lambda = 1/64$; aggregation with radius $r = 2$ and $\beta = 0.67$)



AAMG ... Aggregation–type Algebraic Multigrid [Bastian]
DOUG ... Classical Additive Schwarz with linear coarsening
UMFPACK ... Sparse direct solver [Davies & Duff]

σ^2	$\max_{\tau,\tau'} \frac{\alpha_\tau}{\alpha_{\tau'}}$	**CG–Iterations ($\varepsilon = 10^{-6}$)** M_{AS}	AAMG	DOUG	**CPU-time (in secs)** M_{AS}	AAMG	UMFPACK
2	$1.5 * 10^1$	24	14	32	2.12	1.35	1.85
4	$2.2 * 10^2$	27	29	89	2.14	2.44	1.70
6	$3.3 * 10^3$	29	25	296	2.34	2.10	1.33
8	$4.9 * 10^4$	26	25	498	2.41	2.10	4.88

Energy minimisation for minimal overlap (no solve needed!)

Summary

- Solvers for **heterogeneous multiscale** elliptic problems
- **Applications:** porous media flow, material science, etc...

Summary

- Solvers for **heterogeneous multiscale** elliptic problems
- **Applications:** porous media flow, material science, etc...
- **Coefficient-explicit analysis** of two-level Schwarz
  (when the coarse grid does **not resolve** the coefficients)
- Requirement on coarse space: **bounded/low energy** of basis

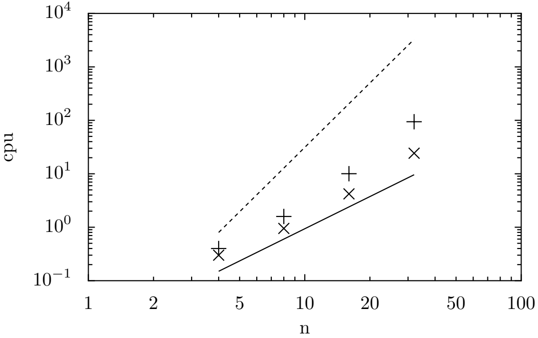
Summary

- Solvers for **heterogeneous multiscale** elliptic problems
- **Applications:** porous media flow, material science, etc...
- **Coefficient-explicit analysis** of two-level Schwarz
  (when the coarse grid does **not resolve** the coefficients)
- Requirement on coarse space: **bounded/low energy** of basis
- **Explicit energy minimisation** $\longrightarrow$ global system $B\boldsymbol{\lambda} = \mathbf{1}$
- **1-Level Schwarz** Preconditioner sufficient for $B\boldsymbol{\lambda} = \mathbf{1}$
- Efficient implementation via **Sherman-Morrison-Woodbury**
- **Choice of supports** crucial for random media

Thank you!

- **Mandel, Brezina, Vaněk**, Energy optimization of algebraic multigrid bases, **Computing**, 1999.
- **Wan, Chan, Smith**, An energy-minimizing interpolation for robust multigrid methods, **SISC**, 2000.
- **Xu, Zikatanov**, On an energy minimizing basis for algebraic multigrid methods, **Comput. Visual. Sci.**, 2004.
- **Brannick, Brezina, MacLachlan et al**, An energy-based AMG coarsening strategy, **Num. Lin. Alg. Appl.**, 2005.
- **Graham, Lechner, RS**, Domain decomposition for multiscale PDEs, **Numer. Math.**, 2007.
- **RS, Vainikko**, Additive Schwarz with aggregation-based coarsening for elliptic problems with highly variable coefficients, **Computing**, 2007.
- **Graham, RS**, Robust domain decomposition algorithms for multiscale PDEs, **Numer. Meth. PDEs**, 2007.
- **Aksoylu, Graham, Klie, RS**, Towards a rigorously justified algebraic preconditioner for high-contrast diffusion problems, **Comput. Visual. Sci.**, 2008.
- **Pechstein, RS**, Analysis of FETI methods for multiscale PDEs, submitted, 2008.
- **Van lent, RS, Graham**, Energy minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs, submitted, 2008.

Coarse Space Construction – Timings



$+$ energy minimising
$\times$ multiscale FE
— linear growth (with n^2)
- - quadr. growth (with n^2)
$n = H/h$ ratio coarse/fine
`cpu` in seconds

$\sim$2-3 times more expensive than multiscale FE (or linear) coarsening.