



Robust Aggregation–Based Coarsening for Multiscale PDEs





Robert Scheichl

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

in collaboration with **Eero Vainikko** (Tartu, Estonia)

DD17, Strobl, Austria, July 2006

Model Problem:

$$\begin{vmatrix} -\nabla \cdot \left(\mathcal{A}(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \right) &= f(\boldsymbol{x}) & \text{ in } \Omega \subset \mathbb{R}^d \quad (d = 1, 2, 3) \\ u &= g & \text{ on } \Gamma_D \subset \partial \Omega \\ \mathcal{A} \nabla u \cdot \boldsymbol{n} &= 0 & \text{ on } \Gamma_N = \partial \Omega \backslash \Gamma_D \end{cases}$$

- Modelling **flow in heterogeneous porous media** (*u* fluid pressure)
- Emergent properties of materials with microstructures (thermal, electrical, mechanical)
- In today's talk we will only consider $\Gamma_N = \emptyset$, $g \equiv 0$ and $A(x) = \alpha(x)I$
- Main difficulty: highly variable (discontinuous) coefficient function $\alpha(x)$

Model Problem:

$$\begin{array}{rcl} -\nabla \cdot \left(\mathcal{A}(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \right) &=& f(\boldsymbol{x}) & \text{ in } \Omega \subset \mathbb{R}^d & (d = 1, 2, 3) \\ \\ u &=& g & \text{ on } \Gamma_D \subset \partial \Omega \\ \\ \mathcal{A} \nabla u \cdot \boldsymbol{n} &=& 0 & \text{ on } \Gamma_N = \partial \Omega \backslash \Gamma_D \end{array}$$

- Modelling **flow in heterogeneous porous media** (*u* fluid pressure)
- Emergent properties of materials with microstructures (thermal, electrical, mechanical)
- In today's talk we will only consider $\Gamma_N = \emptyset$, $g \equiv 0$ and $\mathcal{A}(\boldsymbol{x}) = \alpha(\boldsymbol{x})I$
- Main difficulty: highly variable (discontinuous) coefficient function $\alpha(\boldsymbol{x})$

Provides **important insight** for more complicated problems (e.g. oil reservoir simulation)!

Typical multiscale coefficient functions

Stochastic Model: $\log \alpha(x) =$ (homogeneous, isotropic) Gaussian random field with (i.e. log-normal $\alpha(x)$) mean 0, variance σ^2 , correlation length scale λ

Typical Realisation of log-normal $\alpha({m x})$

 $(n=512^2$, $\sigma^2=8$ and $\lambda=rac{1}{64})$



"Clipped" Realisation ("two-phase" media) $(n = 512^2 \text{ and } \lambda = \frac{1}{64})$



Typical multiscale coefficient functions

Stochastic Model: $\log \alpha(\mathbf{x}) =$ (homogeneous, isotropic) Gaussian random field with (i.e. log-normal $\alpha(\mathbf{x})$) mean 0, variance σ^2 , correlation length scale λ

Typical Realisation of log-normal a(x) "Clipped" Realisation ("two-phase" media) $(n=512^2, \sigma^2=8 \text{ and } \lambda=\frac{1}{64})$ $(n = 512^2 \text{ and } \lambda = \frac{1}{64})$ $\sup_{\mathbf{x},\mathbf{y}\in\Omega}\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})} = O(10^{10}) \nearrow$ $\sup_{\mathbf{x},\mathbf{y}\in\Omega}\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}=O(10^5)$ **Variance** σ^2 determines "contrast" !

Robust Aggregation-Based Coarsening for Multiscale PDEs

R. Scheichl



Correlation length scale λ determines **"roughness"** !



Robust Aggregation–Based Coarsening for Multiscale PDEs

R. Scheichl

Finite Element Method:

$$\begin{split} a(u,v) &:= \int_{\Omega} \alpha \nabla u \cdot \nabla v \quad \dots \quad \text{bilinear form on } (H_0^1(\Omega))^2 \text{ corresponding to model problem} \\ \mathcal{V}^h \subset H^1(\Omega) \qquad \dots \quad \text{FE space of continuous, piecewise linear functions on } \tau \in \mathcal{T}^h \\ \{\varphi_p\} \subset \mathcal{V}^h \cap H_0^1(\Omega) \qquad \dots \quad \text{"hat" functions corresponding to the nodes } \boldsymbol{x}_p^h \text{ of } \mathcal{T}^h \\ \text{Then the FE approximation } u_h = \sum_{p=1}^n U_p \varphi_p \in \mathcal{V}^h \cap H_0^1(\Omega) \text{ where } \mathbf{U} = (U_p)_{p=1,\dots,n} \text{ satisfies} \\ \hline A \mathbf{U} = \mathbf{b} \\ \text{with } A_{p,q} = \sum_{\tau} \alpha_\tau \int_{\tau} \nabla \varphi_p \cdot \nabla \varphi_q \text{ and } \alpha_\tau := \frac{1}{|\tau|} \int_{\tau} \alpha \ . \\ \hline \kappa(A) \lesssim \max_{\tau, \tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}}\right) h^{-2} = O(10^{16}) \text{ for } h = \frac{1}{512}, \ \sigma^2 = 8 \ !! \end{split}$$

Assume w.l.o.g. that α is piecewise constant and $\min_{\tau \in \mathcal{T}^h} \alpha_{\tau} = 1$ (otherwise rescale!).

Finite Element Method:

$$\begin{split} a(u,v) &:= \int_{\Omega} \alpha \nabla u \cdot \nabla v \quad \dots \quad \text{bilinear form on } (H_0^1(\Omega))^2 \text{ corresponding to model problem} \\ \mathcal{V}^h \subset H^1(\Omega) \qquad \dots \quad \text{FE space of continuous, piecewise linear functions on } \tau \in \mathcal{T}^h \\ \{\varphi_p\} \subset \mathcal{V}^h \cap H_0^1(\Omega) \qquad \dots \quad \text{"hat" functions corresponding to the nodes } x_p^h \text{ of } \mathcal{T}^h \\ \text{Then the FE approximation } u_h = \sum_{p=1}^n U_p \varphi_p \in \mathcal{V}^h \cap H_0^1(\Omega) \text{ where } \mathbf{U} = (U_p)_{p=1,\dots,n} \text{ satisfies} \\ \hline A \mathbf{U} = \mathbf{b} \\ \text{with } A_{p,q} = \sum_{\tau} \alpha_\tau \int_{\tau} \nabla \varphi_p \cdot \nabla \varphi_q \text{ and } \alpha_\tau := \frac{1}{|\tau|} \int_{\tau} \alpha \text{ .} \\ \hline \kappa(A) \lesssim \max_{\tau, \tau' \in \mathcal{T}^h} \left(\frac{\alpha_\tau}{\alpha_{\tau'}}\right) h^{-2} \\ \hline \mathbf{Meaning of } \lesssim \mathbf{I} \end{split}$$

Assume w.l.o.g. that α is piecewise constant and $\min_{\tau \in \mathcal{T}^h} \alpha_{\tau} = 1$ (otherwise rescale!).





IMPORTANT: choice of covering $\{\Omega_i\}$, coarse space \mathcal{V}_0 and operator R_0 !

Coarsening via Aggregation

Robust Aggregation-Based Coarsening for Multiscale PDEs

R. Scheichl

Linear Coarse Space (classical):

Dryja, Widlund, Bramble, Pasciak, Schatz, ..., Sarkis [Toselli & Widlund, 2005]

Assume: coarse mesh \mathcal{T}^H , for $K \in \mathcal{T}^H$: $\Omega_K \sim K$, $H_K = \operatorname{diam}(K)$ and $\delta_K = \operatorname{overlap} \operatorname{of} \Omega_K$. Then:

$$\kappa(\mathcal{P}_{AS}^{-1}A) \lesssim \max_{K \in \mathcal{T}^{H}} \max_{\tau, \tau' \subset \omega_{K}} \left(\frac{\alpha_{\tau}}{\alpha_{\tau'}}\right) \left(1 + \frac{H_{K}}{\delta_{K}}\right)$$
$$\kappa(\mathcal{P}_{AS}^{-1}A) \lesssim \mathbf{C}(\mathbf{H/h}) \max_{K \in \mathcal{T}^{H}} \max_{\tau, \tau' \subset K} \left(\frac{\alpha_{\tau}}{\alpha_{\tau'}}\right) \left(1 + \frac{H_{K}}{\delta_{K}}\right)$$

where $\omega_K := \bigcup_{\{K': K \cap K' \neq \emptyset\}} K'$ and $\mathbf{C}(\mathbf{H}/\mathbf{h}) = \log(\mathbf{H}/\mathbf{h})$ in **2D** and \mathbf{H}/\mathbf{h} in **3D**.

Linear Coarse Space (classical):

Dryja, Widlund, Bramble, Pasciak, Schatz, ..., Sarkis [Toselli & Widlund, 2005] Assume: coarse mesh \mathcal{T}^H , for $K \in \mathcal{T}^H$: $\Omega_K \sim K$, $H_K = \operatorname{diam}(K)$ and $\delta_K = \operatorname{overlap}$ of Ω_K . Then:

$$\kappa(\mathcal{P}_{AS}^{-1}A) \lesssim \max_{K \in \mathcal{T}^{H}} \max_{\tau, \tau' \subset \boldsymbol{\omega}_{\mathbf{K}}} \left(\frac{\alpha_{\tau}}{\alpha_{\tau'}}\right) \left(1 + \frac{H_{K}}{\delta_{K}}\right)$$
$$\kappa(\mathcal{P}_{AS}^{-1}A) \lesssim \mathbf{C}(\mathbf{H}/\mathbf{h}) \max_{K \in \mathcal{T}^{H}} \max_{\tau, \tau' \subset \mathbf{K}} \left(\frac{\alpha_{\tau}}{\alpha_{\tau'}}\right) \left(1 + \frac{H_{K}}{\delta_{K}}\right)$$

where $\omega_K := \bigcup_{\{K': K \cap K' \neq \emptyset\}} K'$ and $C(H/h) = \log(H/h)$ in **2D** and H/h in **3D**.

- i.e. "jump-independence" by resolving discontinuities with the coarse mesh!
 - Results extend to non-standard (partition of unity) coarse spaces [Sarkis, 1993-]
 - Unresolved layers \Rightarrow spectral clustering [Graham & Hagger, 1999]

BUT what about case of large variation of α inside coarse elements/subdomains?

Abstract Coarse Space:

 $\mathcal{V}_0 = \operatorname{span}\{\Phi_j : j = 1, \dots, N\}$

where $\{\Phi_j : j = 1, ..., N_H\} \subset \mathcal{V}^h$ are linearly independent coarse space basis functions such that $\Phi_j \in H_0^1(\Omega)$ for all $j \leq N$ and

(C1)
$$\sum_{j=1}^{N_H} \Phi_j(x) = 1$$
 for all $x \in \overline{\Omega}$

(C2) $\|\Phi_j\|_{L_{\infty}(\Omega)} \lesssim 1$ for all $j = 1, \dots, N_H$

(C3) $\forall j = 1, \ldots, N_H \quad \exists i \in \{1, \ldots, s\}$ such that $\omega_j := \operatorname{supp}\{\Phi_j\} \subset \Omega_i$

Important Parameters: $H_j := \operatorname{diam}\{\omega_j\}$, $H := \max_j H_j$, $\delta_j := \operatorname{overlap} \operatorname{for} \omega_j$, $\delta := \min_j \delta_j$.

Definition. (Coarse Space Robustness Indicator)

$$\gamma(\alpha) := \max_{j=1}^{N_H} \delta_j^2 \left\| \alpha |\nabla \Phi_j|^2 \right\|_{L_{\infty}(\Omega)}$$

Note. For the theory we need also shape regularity, uniform overlap, finite covering of the $\{\omega_j\}$.

Theorem (S., 2006).
$$\kappa(P_{AS}^{-1}A) \lesssim \gamma(\alpha) \left(1 + \max_{j=1}^{N_H} \frac{H_j}{\delta_j}\right)$$

Remarks:

• **Previous results** had extra assumption $\|\nabla \Phi_j\|_{L_{\infty}(\Omega)}^2 \lesssim \delta_j^{-2}$ (or a weaker L_2 -version):

Here, **interplay** between coarse space basis functions and α is made **explicit** in $\gamma(\alpha)$!

- No dependency on the subdomain sizes (even for $H^{sub} \gg H$) !
- The dependency on the "mesh" parameters is **sharp**, i.e. linear in H/δ .
- Previously best theoretical result for smoothed aggregation coarse spaces (for $\alpha \equiv 1$):

$$\kappa(P_{
m AS}^{-1}A) \lesssim 1 + rac{H}{\delta} + rac{H^{sub}}{\delta}$$
 [Sala, Shadid & Tuminaro, SIMAX 2006]

Numerical Evidence (Dependency on H and H^{sub})

Example (Laplacian): $\alpha \equiv 1 \implies \gamma(\alpha) \sim 1$



 $n = 1024 \times 1024$, minimal overlap (i.e. $\delta = 3h$), $1 \times$ smoothed aggregation

Note. All CPU times were obtained on a 3GHz INTEL Pentium 4 processor.

Example (two media) – Linear Coarsening

- \mathcal{T}^h and \mathcal{T}^H uniform.
- Square subdomains Ω_i (consisting of 8 coarse elements $K \in \mathcal{T}^H$), i.e. overlap $\delta = H$.



For
$$h = \frac{1}{256}$$
 and $H = 8h$:

\hat{lpha}	$\kappa(P_{AS}^{-1}A)$	$\gamma(lpha)$
10^{0}	5.2	4
10^{1}	9.1	40
10^{2}	58.1	400
10^{3}	471	4000
10^{4}	1821	4.0(+4)
10^{5}	2561*	4.0(+5)

* same as 1-level method (i.e. no coarse grid)

Idea of Proof. ($\gamma(\alpha)$ crucial ! Then use classical Schwarz theory introducing "weight" α !)

- To bound $\lambda_{\max}(P_{AS}^{-1}A)$ use a colouring argument.
- To bound $\lambda_{\min}(P_{AS}^{-1}A)$ find stable splitting for each $u \in \mathcal{V}^h$: $v_0 \in \mathcal{V}_0$ and $v_j \in \mathcal{V}^h(\omega_j)$ s.t. $u = \sum_{j=0}^{N_H} v_j$ and $\sum_{j=0}^{N_H} a(v_j, v_j) \lesssim \gamma(\alpha) \left(1 + \max_j \frac{H_j}{\delta_j}\right) a(u, u)$ i.e. $v_0 := \sum_{j=0}^N \bar{u}_j \Phi_j$ with $\bar{u}_j := |\omega_j|^{-1} \int_{\omega_j} u$ and $v_j := \begin{cases} I_h(\Phi_j(u - \bar{u}_j)), & j \leq N, \\ I_h(\Phi_j u), & j > N. \end{cases}$ Now, $a(v_0, v_0) \lesssim \gamma(\alpha) \left(\max_j \frac{H_j}{\delta_j}\right) a(u, u)$ and $a(v_j, v_j) \lesssim \|\alpha|\nabla\Phi_j|^2\|_{L_\infty(\omega_j)} \|u - \bar{u}_j\|_{L_2(\omega_{j,\delta_j})}^2 + |u - \bar{u}_j|_{H^1(\omega_j),\alpha}^2 \lesssim \gamma(\alpha) \left(1 + \frac{H_j}{\delta_j}\right) a(u, u)$
- To find stable splitting in $\mathcal{V}^h(\Omega_i)$ set $u_i := \sum_{j \in \mathcal{I}_i} v_j$ where $\mathcal{I}_i := \{j : \omega_j \subset \Omega_i\}$ and use again a **colouring argument**.

Geometric Multigrid Idea: Relaxation schemes (like ω -Jacobi or SOR) **smooth** the error

 \implies restrict to a (geometrically) **coarser** grid. Breaks down for large variation in $\alpha(\boldsymbol{x})$!

Geometric Multigrid Idea: Relaxation schemes (like ω -Jacobi or SOR) **smooth** the error

 \implies restrict to a (geometrically) **coarser** grid. Breaks down for large variation in $\alpha(x)$!

Algebraic Multigrid Idea: Relaxation schemes smooth error along strong connections in A: Use graph \mathcal{G} associated with A: node x_q^h is strongly connected to x_p^h if $|A_{p,q}|$ is large,

 $|\tilde{A}_{p,q}| \geq \varepsilon \max_{k \neq p} |\tilde{A}_{p,k}|$ where $\tilde{A} := D^{-1/2} A D^{-1/2}$ [Bastian, 1996] e.g.

⇒ { (a) Select "well-connected" coarse nodes, and interpolate to strongly connected neighbours based on heuristic Ae = r ≈ 0 [Ruge & Stueben, 1985].
 (b) Aggregate strongly connected neighbours, and use p.w. constant prolongation, smoothed along the strong connections [Vanek, Mandel & Brezina, 1995].

Main Idea:

Extend the notion of **strong connections** to **any** pair of nodes by considering **paths** in \mathcal{G} !

Define the strongly-connected graph-r neighbourhood of x_p^h by the set of all nodes x_q^h s.t. there exists a path $\gamma_{pq} := [x_p^h = x_{p_0}, x_{p_1}, \dots, x_{p_k} = x_q^h]$ of length $k \leq r$ from x_p^h to x_q^h , and x_{p_i} is strongly connected to $x_{p_{i-1}}$ for all $i = 1, \dots, k$.

• This is readily available from compressed row storage of so-called "filtered" matrix

$$A_{p,q}^{\varepsilon} := \begin{cases} A_{p,p} + \sum_{\substack{x_k^h \text{ not str. con.} \\ x_k^h}} A_{p,k} & \text{if } x_q^h = x_p^h \\ A_{p,q} & \text{if } x_q^h \text{ strongly connected to } x_p^h \\ 0 & \text{otherwise} \end{cases}$$

(i.e. graph \mathcal{G} not needed !)

- To find good seed nodes x_p^h for the aggregates we use an advancing front [Raw, 1996]
- Related to "aggressive" coarsening in algebraic multigrid, e.g. [Papadopoulos, 2004]



5. Subdomains are aggregated using same algorithm on A_0 (Overlap comes from smoothing!)



5. Subdomains are aggregated using same algorithm on A_0 (Overlap comes from smoothing!)

Novel Aspects (apart from new theory)

- Making use of strong connections in a domain decomposition method
- Second aggregation for subdomains $\longrightarrow H^{sub} \gg H$

¹see also [Vanek & Brezina, 1999], [Jenkins et al., 2001], [Lasser & Toselli, 2002]



Obtained with r = 2 and $\varepsilon = 0.67$ for typical realisation $(n = 32^2, \lambda = \frac{1}{8}, \max_{\tau, \tau'} \frac{\alpha_{\tau}}{\alpha_{\tau'}} \approx 10^3)$

Example (clipped random fields):



with $n=256\times 256~$ and $~\lambda=1/64$

CG–Iterations	(b =	1,	$tol = 10^{-6}$	³)
---------------	--------------	----	-----------------	----------------

CPU	I-Time	(in secs)	
		(<u> </u>	

σ^2	$\max_{\tau,\tau'} \frac{\alpha_{\tau}}{\alpha_{\tau'}}$	New	AMG	DOUG	σ^2	New	AMG
2	$1.5*10^1$	24			2	2.12	
4	$2.2*\mathbf{10^2}$	27			4	2.14	
6	${\bf 3.3}*{f 10^3}$	29			6	2.34	
8	$4.9*\mathbf{10^4}$	26			8	2.41	

All numerical results with r = 2 and $\varepsilon = 0.67$ and no smoothing!

UMFPACK

Example (clipped random fields):



AMG	 Aggregation–type Algebraic Multigrid [Bastian] (no smoothing – piecewise constant prolongation)
DOUG	 Classical Additive Schwarz with linear coarsening (parallel run on slow network – CPU-times pessimistic)
UMFPACK	 Sparse direct solver [Davies & Duff]

with $n=256\times 256~$ and $~\lambda=1/64$

CG –Iterations	(b = 1,	$tol = 10^{-6}$)
-----------------------	---------	-------------------

σ^2	$\max_{\tau,\tau'} \frac{\alpha_{\tau}}{\alpha_{\tau'}}$	New	AMG	DOUG
2	$1.5 * 10^{1}$	24	14	32
4	$2.2 * 10^2$	27	27	89
6	$3.3 * 10^3$	29	40	296
8	$4.9 * 10^4$	26	77	498

CPU–Time (in secs)

σ^2	New	AMG	UMFPACK
2	2.12	1.35	1.85
4	2.14	2.27	1.70
6	2.34	3.31	1.33
8	2.41	6.23	4.88

All numerical results with r = 2 and $\varepsilon = 0.67$ and no smoothing!

CG-Iterations (b = 1, tol = 10^{-6})				CPU-Tin	ne (in sec	s)	
λ	New	AMG	DOUG	λ	New	AMG	UMFPACK
1/16	26			1/16	2.20		
1/32	27			1/32	2.24		
1/64	26			1/64	2.41		
1/128	33			1/128	2.71		
1/256	48			1/256	3.84		

Clipped random fields with $n = 256 \times 256$ and $\sigma^2 = 8$.

CG–Iterations ($\mathbf{b} = 1$, $\mathrm{tol} = 10^{-6}$)					
n	New	AMG	DOUG		
128^{2}	25				
256^{2}	26				
512^{2}	34				
1024^2	74				

CPU–Time (in secs)					
n	New	AMG	UMFPACK		
128^{2}	0.46				
256^{2}	2.41				
512^{2}	16.8				
1024^{2}	105.9				

Clipped random fields with $\sigma^2 = 8$ and $\lambda = 4h$.

CG-Iterations (b = 1, tol = 10^{-6})					
λ	New	AMG	DOUG		
1/16	26	18	355		
1/32	27	64	430		
1/64	26	77	498		
1/128	33	70	655		
1/256	48	166	858		

CPU-Time (in secs) AMG λ New **UMFPACK** 1/162.20 1.67 4.52 1/32 2.24 5.14 4.77 1/64 2.41 6.23 4.88 2.71 1/128 5.77 7.48 1/2563.84 13.5 10.2

Clipped random fields with $n = 256 \times 256$ and $\sigma^2 = 8$.

CG–Iterations (b = 1, tol = 10^{-6})							
n	New	AMG	DOUG				
128^{2}	25	35	136				
256^{2}	26	77	498				
512^{2}	34	100	1111				
1024^{2}	74	422	***				

		•	,
n	New	AMG	UMFPACK
128^{2}	0.46	0.68	0.52
256^{2}	2.41	6.23	4.88
512^{2}	16.8	33.8	88.8
1024^{2}	105.9	540	***

CPU-Time (in secs)

Clipped random fields with $\sigma^2 = 8$ and $\lambda = 4h$.





<u>Idea of Proof.</u> In the case $R_0^T = P$ assumptions (C1)–(C3) are satisfied by construction, we have uniform overlap, and finite covering. Also $|\nabla \Phi_j(x)| \leq \delta_j^{-1} \quad \forall x \in \Omega$.



 $\implies \text{ (for } r \text{ sufficiently large) } \alpha(x) = 1 \text{ wherever } \nabla \Phi_j(x) \neq 0 \implies \gamma(\alpha) \leq 1$ Moreover $\delta_j = O(h)$ and since diam(islands) $\lesssim h$ we can choose r s.t. $H_j \lesssim h \implies \kappa(P_{AS}^{-1}A) \leq C$ where C is independent of h, r and α but may depend on the shape of the islands.

Example (log-normal $\alpha(\mathbf{x})$):



	New	AMG	DOUG	UMFPACK
Iterations	19	38	62	
CPU-time	8.3s	13.1s	29.7s	10.3s

Note. Simpler than clipped fields !!

with n=512 imes 512, $\sigma^2=8$ and $\lambda=1/64$

Current/Future Work:

- Parallelisation
- Multiplicative Schwarz
- Extension of theory for discontinuous coefficients to Algebraic Multigrid
- Combination with multiscale FE interpolation

Two preprints available at

http://www.bath.ac.uk/math-sci/BICS