Multiscale model reduction for flows in heterogeneous media

Y. Efendiev 1 and J. Galvis 1 and E. Gildin 1 and X.H. Wu^2

¹Texas A&M University, ²ExxonMobil

Introduction

Subsurface properties vary over many scales and there are many small features that can impact the overall flow.



 Typical approaches involve upscaling where effective (homogenized) equations are postulated and effective parameters are tabulated (e.g., Darcy law, ...).

Introduction. Multiscale porous media.

Flow in porous media: $div(k(x)\nabla p) = f$, k(x) is permeability.



- Small scales (no periodcity) and high contrast are present.
- ► Examples. Periodic small scales $k(x) = k(x, \frac{x}{\epsilon})$. High-contrast $\frac{\max k(x)}{\min k(x)} \propto \epsilon^{-\beta}$.

Local upscaling/homogenization.



Comparisons between fine and coarse models are done using average energies.

- ▶ Local flow problem $div(k\nabla\phi) = 0$, $\phi = \phi_{bndr}$ is solved. $F(k) = \int k\nabla\phi\nabla\phi$.
- Similarly, $div(k^*\nabla\phi^*) = 0$, $\phi^* = \phi_{bndr}$ is solved. $F^*(k^*) = \int k^*\nabla\phi^*\nabla\phi^*$.
- Minimization $||k k^*||_V = F(k) F^*(k^*)(=0)$. For $\phi_{bndr} = x$ gives $k^* = \frac{1}{|D|} \int k \nabla \phi \nabla \phi$ (e.g., Durlofsky et al., 1991, Bourgeat et al., 1987...).

Reduced multiscale models

We are interested in approximating the solution in V of div(k(x)∇p) = f.



- ▶ Instead: we look for a coarse-scale representation of $p = \sum_{i=1}^{\text{fine}} p_i \phi_i$, $p^* = \sum_{i=1}^{\text{coarse}} p_i^* \Phi_i$, such that $\|p p^*\|_V$ is small given a coarse grid.
- ▶ Q: how determine Φ_i; sparsity; coarse grid and control accuracy? Applications to preconditioners.

Multiscale FE¹.

• Approximation via basis functions defined on a coarse grid: $V_0 = \text{Span}\{\Phi_i\}.$

Then, mass conservation or Galerkin formulation can be written to compute the coarse degrees of freedoms.



- Basis: $div(k(x)\nabla\Phi_i) = 0$ in ω_i , $\Phi_i = \Phi_i^0$ on bdr.
- If $div(k(x)\nabla p) = f$ (or L(p) = f). Then, $p = \sum_i c_i \Phi_i$, where c_i are found by writing mass conservation or Galerkin formulation. E.g., $\langle L(\sum_i c_i \Phi_i), \tilde{\Phi}_j \rangle = \langle f, \tilde{\Phi}_j \rangle$

¹Hou and Wu, JCP 1997, Babuska et al. 1984,...

Illustration of basis functions



Upscaling vs. multiscale

Upscaling (homogenization)

- local effective parameters
- different source/bc
- local changes:
 - few parameters

Multiscale FEM:

- Iocal basis functions
- different source/bc
- local changes:
 - (local) basis functions





Multiscale basis functions

- Some advantages: non-uniform gridding; built-in fine-scale recovery; multiple basis per node; enforcing global information;
- Multiscale basis functions. Boundary conditions (important!).



Piecewise linear boundary conditions result to large discrepancies near the edges of coarse blocks (e.g., the solution is u ≈ u₀ + εu₁(x, ^x/_ε) along coarse edge while MsFE solution is linear). Error ∝ ε/H.



Subgrid information at the boundaries needed to get an accurate solution.

Convergence property of MsFEM

Consider $k_{\epsilon}(x) = k(x/\epsilon)$, where k(y) is periodic in y. H - computational coarse-mesh size. **Theorem** (Hou, Wu, Cai, 1998) Denote p_{ϵ}^{H} the numerical solution obtained by MsFEM, and p_{ϵ} the solution of the original problem. Then, If $H \gg \epsilon$,

$$\|p_{\epsilon} - p_{\epsilon}^{H}\|_{H^{1}} \le C(H + \sqrt{\frac{\epsilon}{H}})$$

- The ratio ϵ/H reflects two intrinsic scales. We call ϵ/H the resonance error
- ► The theorem shows that there is a scale resonance when H ≈ ε. Numerical experiments confirm the scale resonance.
- For problems with scale separation, we can choose $H \gg \epsilon$ in order to avoid the resonance, but for problems with continuous spectrum of scales, we cannot avoid this resonance.

Resonance errors

 \blacktriangleright To demonstrate the influence of the boundary condition: Multiscale expansion of ϕ^i

$$\phi^{i} = \phi_{0}(x) + \epsilon \phi_{1}(x, x/\epsilon) + \epsilon \theta + \dots,$$

where θ satisfies $div(k_{\epsilon}\nabla\theta) = 0$, $\theta = -\phi_1(x, x/\epsilon)$ on ∂K .



Subgrid approximation improvement

Improving subgrid capturing errors: oversampling² (used to avoid artificial boundary condition effects).



- The errors depend on localization assumptions (Hou et al. 1999, Efendiev et al., 1999, Babuska et al. 2006, Nolen et al. 2008,)
- Limited global information (Owhadi and Zhang, 2006, Efendiev, Hou,, 2006, 2009, ...). Local-global approaches...
- Main question: can we add basis functions so that MsFEM solution converges to the fine-scale solution?

²Hou and Wu, JCP 1997

Classical upscaling or numerical homogenization.

. . .

- Multiscale finite element methods (J. Aarnes, Z. Cai, Y. Efendiev, V. Ginting, T. Hou, X. Wu....)
- Mixed multiscale finite element methods (Z. Chen, J. Aarnes, T. Arbogast, K.A. Lie, S. Krogstad,...)
- MsFV (P. Jenny, H. Tchelepi, S.H. Lee, Iliev,)
- Mortar multiscale methods (T. Arbogast, M. Peszynska, M. Wheeler, I. Yotov,...)
- Subgrid modeling and stabilization (by T. Arbogast, I. Babuska, F. Brezzi, T. Hughes, ...)
- Heterogeneous multiscale methods (E, Engquist, Abdulle, M. Ohlberger, ...)
- Numerical homogenization (NH) using two-scale convergence (C. Schwab, V.H. Hoang, M. Ohlberger, ...)
- NH (Bourgeat, Allaire, Gloria, Blanc & Le Bris, Madureira, Sarkis, Versieux, Cao, ...)
- AMG coarsening (P Vassilevski)

►
$$div(k(x)\nabla p) = f$$
,
Fine $V = \text{span}\{\phi_i\}_{i=1}^{N_f}$

Coarse:
$$V_0 = \operatorname{span} \{ \Phi_i \}_{i=1}^{N_c}$$

• Au = b, $A_0u_0 = b_0$, where $a_{ij} = (k\nabla\phi_i, \nabla\phi_j)$ and $a_{ij}^0 = (k\nabla\Phi_i, \nabla\Phi_j)$.

- Multiscale/upscaling goal: reduce ||u − u₀|| that depends on small scales (e.g., ε/H, ε is small scale) and the contrast (k_{max}/k_{min}).
- DD preconditioners iterate on a residual until convergence (number of iterations!). The number of iterations depends on the contrast in the presence of small scales.

$$b \quad div(k(x)\nabla p) = f, \\ \hline Fine \ V = \operatorname{span}\{\phi_i\}_{i=1}^{N_f} \\ \hline Coarse: \ V_0 = \operatorname{span}\{\Phi_i\}_{i=1}^{N_c} \\ \hline \end{cases}$$

• Au = b, $A_0u_0 = b_0$, where $a_{ij} = (k\nabla\phi_i, \nabla\phi_j)$ and $a_{ij}^0 = (k\nabla\Phi_i, \nabla\Phi_j)$.

- Multiscale/upscaling goal: reduce ||u − u₀|| that depends on small scales (e.g., ε/H, ε is small scale) and the contrast (k_{max}/k_{min}).
- DD preconditioners iterate on a residual until convergence (number of iterations!). The number of iterations depends on the contrast in the presence of small scales.

$$\begin{aligned} & \mathsf{div}(k(x)\nabla p) = f, \\ & \overline{\mathsf{Fine} \ V = \mathsf{span}\{\phi_i\}_{i=1}^{N_f}} \\ & \mathsf{A}u = b, \quad A_0u_0 = b_0, \end{aligned}$$

where $a_{ij} = (k \nabla \phi_i, \nabla \phi_j)$ and $a_{ij}^0 = (k \nabla \Phi_i, \nabla \Phi_j)$.

- ► Multiscale/upscaling goal: reduce ||u u₀|| that depends on small scales (e.g., ϵ/H, ϵ is small scale) and the contrast (k_{max}/k_{min}).
- DD preconditioners iterate on a residual until convergence (number of iterations!). The number of iterations depends on the contrast in the presence of small scales.

$$div(k(x)\nabla p) = f,$$
Fine $V = \text{span}\{\phi_i\}_{i=1}^{N_f}$
Coarse: $V_0 = \text{span}\{\Phi_i\}_{i=1}^{N_c}$.

where $a_{ij} = (k \nabla \phi_i, \nabla \phi_j)$ and $a_{ij}^0 = (k \nabla \Phi_i, \nabla \Phi_j).$

- ► Multiscale/upscaling goal: reduce ||u u₀|| that depends on small scales (e.g., ϵ/H, ϵ is small scale) and the contrast (k_{max}/k_{min}).
- DD preconditioners iterate on a residual until convergence (number of iterations!). The number of iterations depends on the contrast in the presence of small scales.

DD preconditioners

- Additive Schwarz preconditioner of the form $B^{-1} = R_0^T A_0^{-1} R_0 + \sum_{i=1}^N R_i^T A_i^{-1} R_i,$
- ▶ Goal: to keep the number of iterations independent of physical parameters (cond(B⁻¹A) ≤ C independent of contrast).



- We are interested in the cases with highly heterogeneous coefficients within coarse regions.
- There are many works for cases with nearly homogeneous coefficients within coarse regions (Widlund, Wheeler, Drya,...)
- M. Sarkis (LNCSE 2001, Cont. Math. 2002) quasi-monotonic coefficients. Also, the use of extra basis functions

. . .

- Aarnes and Hou (Acta Sinica 2002) the use of multiscale basis in DD preconditioners
- Graham, Scheichl (Num. Math. 2007) cases with inclusions, energy minimizing basis
- Xu and Zhu (M3AS 2008) Analysis of the spectrum of preconditioner operator
- T. Arbogast, I. Yotov, M. Wheeler, ... (Comp. Geo., 2007,...) -Preconditioning of various discretization of multiscale PDEs

- Eigenproblem $div(k(x)\nabla\psi_i) = \lambda_i k(x)\psi_i$
- The weight is very important for dimension reduction.
- "gap" \rightarrow $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L < \lambda_{L+1} \leq \ldots \leq \lambda_N$,
- ▶ $\lambda_1, ..., \lambda_L$ are small, asymptotically vanishing eigenvalues $(\int k |\nabla \psi_i|^2 / \int k |\psi_i|^2)$
- "No-separation" is possible when inclusions approach to each other (L. Borcea and G. Papanicolaou, 1997,...)



20 / 46

- Eigenproblem $div(k(x)\nabla\psi_i) = \lambda_i k(x)\psi_i$
- The weight is very important for dimension reduction.
- "gap" $\rightarrow \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L < \lambda_{L+1} \leq \ldots \leq \lambda_N$,

▶ $\lambda_1, ..., \lambda_L$ are small, asymptotically vanishing eigenvalues $(\int k |\nabla \psi_i|^2 / \int k |\psi_i|^2)$

 "No-separation" is possible when inclusions approach to each other (L. Borcea and G. Papanicolaou, 1997,...)



21 / 46

- Eigenproblem $div(k(x)\nabla\psi_i) = \lambda_i k(x)\psi_i$
- The weight is very important for dimension reduction.
- "gap" $\rightarrow \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L < \lambda_{L+1} \leq \ldots \leq \lambda_N$,
- ► λ_1 , ..., λ_L are small, asymptotically vanishing eigenvalues $(\int k |\nabla \psi_i|^2 / \int k |\psi_i|^2)$
- "No-separation" is possible when inclusions approach to each other (L. Borcea and G. Papanicolaou, 1997,...)



- Eigenproblem $div(k(x)\nabla\psi_i) = \lambda_i k(x)\psi_i$
- The weight is very important for dimension reduction.
- "gap" $\rightarrow \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L < \lambda_{L+1} \leq \ldots \leq \lambda_N$,
- ▶ $\lambda_1, ..., \lambda_L$ are small, asymptotically vanishing eigenvalues $(\int k |\nabla \psi_i|^2 / \int k |\psi_i|^2)$
- "No-separation" is possible when inclusions approach to each other (L. Borcea and G. Papanicolaou, 1997,...)



Coarse spaces

In our work, we take the coarse space to be

$$V_0 = Span\{\chi_i \psi_j^{\omega_i}\} = Span\{\Phi_i\},$$

where ω_i is the union of coarse grids with common vertex i (ω_K is the union of coarse grids with common edge with K), and χ_i is a partition of unity function for the node i.





Coarse spaces in preconditioners

It was proved (Galvis and Efendiev, 2009) that cond(B⁻¹A) ≤ C/(λ^{*}_{L+1}) if the coarse space "spans" the eigenfunctions. A main difficulty is to prove the following fact.

$$\int_{K} k(v - I_0 v)^2 \leq \frac{1}{\lambda_{K,L+1}} \int_{\omega_K} k |\nabla v|^2$$
$$\int_{K} k |\nabla I_0 v|^2 \leq \max\{1, \frac{1}{H^2 \lambda_{K,L+1}}\} \int_{\omega_K} k |\nabla v|^2,$$

where $I_0 v$ is the coarse-scale projection of v.

From the above theorem, one can show that $cond(B^{-1}A) \leq C(1 + (H/\delta)^2)$ independent of contrast.

Numerical result



Numerical result

η	Linear	MS	EMF	LSM_1	LSM
10^{4}	81(3.8e+3)	68(2.1e+3)	52(3.8e+2)	53(54.29)	46(32.04
10^{5}	88(3.8e+4)	73(2.1e+4)	60(3.8e+3)	41(56.25)	40(32.73
10^{6}	111(3.8e+5)	91(2.1e+5)	68(3.8e+4)	40(56.53)	37(33.40
10^{7}	141(3.8e+6)	112(2.1e+6)	76(3.8e+5)	37(56.52)	36(41.35
10^{8}	156(3.8e+7)	129(2.1e+7)	86(3.8e+6)	37(56.42)	33(42.47
10^{9}	175(3.8e+9)	143(2.1e+8)	73(2.3e+7)	30(54.88)	30(42.59

Number of iterations until convergence of the PCG and condition number for different values of the contrast η with the coefficient. We set the tolerance to 1e-10. Here H = 1/10 with h = 1/100. The classical (one basis per node) coarse problems size is 81×81 . The new coarse problem is of size 321×321 .

- Multiscale basis functions are constructed by solving local problems with linear boundary conditions, i.e., for each node *i*, $div(k\nabla\phi_i) = 0$ in ω_i , and ϕ_i is linear on $\partial\omega_i$.
- Energy minimizing basis functions (e.g., Xu and Zikatanov 2004) are constructed by minimizing $\sum \int \int k |\nabla \Phi_i|^2$.

Dimension reduction

- There can be many eigenfunctions
 - How to reduce the dimension of the coarse space?
 - Inclusions vs. channels. Reduction "for inclusions"
 - Inclusions high-conductivity regions inside the coarse block.
 Channels high-conductivity regions connecting one boundary to the other of the coarse block.



Multiscale capturing

Reducing the dimension of coarse space

How to reduce the dimension of the coarse space (and handle tensor coefficients)?





Intuition: Isolated inclusions can be represented with a single basis function per node (localization), while channels need to be represented separately.

Reducing the dimension



► Then, we take the "important" eigenvalues/eigenvectors of div(k∇ψ_i) = λ_ikψ_i, and form coarse space as before V₀ = Span{χ_iψ_j}

Dimension reduction

- The main objective is to minimize the variations in k as much as possible. Note that "important" eigenvalues are determined by high-conductivity part of k.
- Using initial multiscale finite element basis functions, one can take into account the information that can be localized and then complement these basis functions.



• It can be shown that $cond(B^{-1}A) \leq C$.

Numerical experiment



Figure: Left: Coarse mesh. Center: Original coefficient. Right: Coefficient \tilde{k} computed using (linear) multiscale basis functions.

Numerical experiment

η	MS	LSM ($\tilde{\kappa} = \kappa$)	LSM($\tilde{\kappa}$)
10^{4}	98(2490.75)	27(6.19)	28(7.34)
10^{5}	123(24866.24)	28(6.19)	29(7.35)
10^{6}	144(248621.33)	29(6.19)	29(7.35)
10^{7}	174(2486172.35)	29(6.19)	30(7.35)
Dim	49	102	69

Table: Number of iterations until convergence and estimated condition number for the PCG and different values of the contrast η . Here H=1/8 with h=1/80. The notation MS stands for the (linear boundary condition) multiscale coarse space, and LSM is the local spectral multiscale coarse space

Numerical result



Coarse-scale approximation

▶ This is joint work with J. Galvis and X.H. Wu.



The choice of basis functions.

High-contrast "homogenization" problems

- Previous MsFEM approaches do not consider the contrast in the problem, k_{max}/k_{min}, (except, Chu, Graham, Hou, 2009 and Berlyand and Owhadi, 2009).
- Homogenization and network results for high-contrast problems (L. Borcea, G. Papanicolaou, Berlyand, Novikov,... 1997, 1998, ... Smyshlyaev et al., 2005-2008, ...).

Multiple basis per node

 Multiple basis per node is needed to represent distinct important features of the solution



Pressures



Errors

H	MS	EMF	LSM_1	LSM-RE	RLSM
1/10	0.98(0.98)	0.24(0.24)	0.11(0.11)	0.06(0.06)	0.10(0.10)
1/20	1.46(1.47)	0.85(0.87)	0.14(0.14)	0.09(0.09)	0.13(0.13)

Relative energy error. Here h = 1/100, $\eta = 10^4$ ($\eta = 10^6$).

H	MS	EMF	LSM_1	LSM-RE	RLSM
$\frac{1}{10}$	0.03(0.03)	0.001(0.001)	0.0003(0.0003)	0.0003(0.0003)	0.0002(0.0002)
$\frac{1}{20}$	0.06(0.07)	0.03(0.04)	0.00004(0.00006)	0.0001(0.0001)	0.0004(0.0003)

Relative weighted L^2 error. Here h = 1/100, $\eta = 10^4$ ($\eta = 10^6$).

Convergence

For sufficiently large $\Lambda_* = \min_{\omega_i}$ and under additional assumptions, we have

$$\int_D \kappa |\nabla (u - u_H)|^2 \preceq \left(\frac{H}{\Lambda_*}\right)^{\gamma},$$

for some $\gamma > 0$.

+	LSM, $H = 1/10$	\widetilde{LSM} , $H = 1/10$	LSM, $H = 1/20$	\widetilde{LSM} , $H = 1/20$
0	0.14 (λ=.8)	0.18 (λ =0.27)	0.18 (λ =0.25)	0.13 (λ=.4)
+2	0.048 (λ =3.34)	0.039 (λ =1.18)	0.021 (λ =.4.2)	0.019 (λ =1.11)
+3	0.043 (λ =5.3)	0.028 (λ =2.25)	0.012 (λ =5.1)	0.0189 (λ =1.89)
+4	0.03 (λ =7.1)	0.022 (λ =3.4)	0.007 (λ =8.2)	0.05 (λ =2.7)

- Our objective is to reach to higher values of Λ_* with smaller coarse spaces.
- \blacktriangleright The choice of initial multiscale basis functions is important and allows fast increase in Λ_*
- Convergence rate is $H^{1+\beta}/\Lambda_*$, where $\beta \ge 0$ smoothness.

Local-global multiscale model reduction

- Multiscale methods are typically designed to provide approximations for arbitrary coarse-level inputs
- Can we develop local-global multiscale models if the input set belongs to a smaller set?



Local-global multiscale model reduction



- The main idea is to use an appropriate number of basis functions locally and choose a certain number of global modes.
- ▶ p_t = Ap + Bw, q = Cp, where A is the fine-scale forward model. w is assumed to be coarse-scale input data.
- Our goal is to find q_0^r such that $||q q_0^r||$ is small. We will consider $q q_0^r = (q q_0) + (q_0 q_0^r)$.
- To obtain reduced model for a given input-output relation, Balanced Truncation method is used. Balanced truncation involves solving Lyapunov equations AP + PA^T + BB^T = 0, AQ + QA^T + CC^T = 0 to identify important basis for the global system.

Local-global multiscale model reduction

• Coarse system
$$(p_0)_t = A_0 p_0 + B_0 w$$
, $q_0 = C_0 p_0$.

- ▶ Balanced Truncation for an appropriate coarse system $A_0P_0 + P_0A_0^T + B_0B_0^T = 0$, $A_0Q_0 + Q_0A_0^T + C_0C_0^T = 0$
- ► It can be shown that $||q q_0^r|| \le ||C||_A H^{\gamma} / \Lambda_* + \sum_{i=l+1} \sigma_i^0$.

MS	MS Error	BT Error	Total
+0	0.12(0.12)	0.23(0.04)	0.29(0.12)
+1	0.079(0.079)	0.25(0.06)	0.29(0.109)
+2	0.062(0.062)	0.26(0.06)	0.29(0.099)

() - BT with 10 SV, else with 3 SV.

Eigenvector computation

The computations of eigenfunctions, though local, can be expensive. We show that one can still achieve an optimal result with hierarchical computation of eigenfunctions.



(YE, J. Galvis, P. Vassilevski, Spectral agglomerate AMGe method for high-contrast problems, DD19.).

Eigenvector computation

The computations of eigenfunctions, though local, can be expensive. We show that one can still achieve an optimal result with hierarchical computation of eigenfunctions.



(YE, J. Galvis, P. Vassilevski, Spectral agglomerate AMGe method for high-contrast problems, DD19.).

Conclusions

- Reduced model approximation of the solution of PDEs via coarse-scale basis functions.
- Eigenvalue problems for capturing important effects
- Dimension reduction inclusions vs. channels. The choice of initial multiscale (one per node) basis functions.
- Optimal (w.r.t. the contrast) preconditioners.
- Subgrid accuracy of multiscale finite element methods
- Model reduction
- Hierarchical computation of eigenvectors