Preconditioning techniques for mixed finite element equations with multiple scales

Jørg Espen Aarnes & Stein Krogstad SINTEF ICT, Dept. of Applied Mathematics

Multiscale modeling in fluid flow and material science October 18-20, Oslo, Norway

Outline

- Model problem and mixed FEM formulation.
- Preconditioning mixed FEM eqs. and eqs. with multiple scales.
- The construction of a multiscale DD preconditioner.
- A family of multiscale multigrid preconditioners for elliptic eqs.
- Two alternative iterative schemes for solving mixed FEM eqs.
- Numerical results and concluding remarks.



Model problem

We consider the following elliptic model problem:

$$\begin{aligned} -\nabla \cdot k(x) \nabla u + c(x) u &= f & \text{in } \Omega, \\ (-k(x) \nabla u) \cdot n &= 0 & \text{on } \partial \Omega. \end{aligned}$$

Here c is a non-negative function in $L^2(\Omega)$ and k is a symmetric positive definite tensor with uniform upper and lower bounds:

$$0 < \alpha \leq \frac{\xi^T k(x)\xi}{\xi^T \xi} \leq \beta < \infty \qquad \forall \xi \in \mathcal{R}^d \setminus \{0\}, \quad \forall x \in \Omega.$$



back

The mixed formulation

The mixed formulation of the model problem reads: Find $q\in H^{1,{
m div}}_0(\Omega)$ and $u\in L^2(\Omega)$ such that

$$\int_{\Omega} k^{-1} q \cdot p \, dx \quad - \quad \int_{\Omega} u \, \nabla \cdot p \, dx \quad = \quad 0$$
$$\int_{\Omega} v \, \nabla \cdot q \, dx \quad + \quad \int_{\Omega} cuv \, dx \quad = \quad \int_{\Omega} fv \, dx$$

for all $p \in H_0^{1,\operatorname{div}}(\Omega)$ and $v \in L^2(\Omega)$.



◄ back ►

The mixed FEM formulation

Replacing $H_0^{1,\text{div}}(\Omega)$ and $L^2(\Omega)$ with finite dimensional subspaces $Q = \text{span}\{\psi_i\}$ and $V = \text{span}\{\phi_m\}$ we obtain:

Find $q = \sum_{i} q_i \psi_i$ and $u = \sum_{m} u_m \phi_m$ such that

$$\int_{\Omega} k^{-1} q \cdot \psi_j \, dx \quad - \quad \int_{\Omega} u \nabla \cdot \psi_j \, dx \quad = 0$$
$$\int_{\Omega} \phi_n \nabla \cdot q \, dx \quad + \quad \int_{\Omega} c u \phi_n \, dx \quad = \int_{\Omega} f \phi_n \, dx$$

for all ψ_j and ϕ_n .



4 of 33

back

Thus, the mixed FEM formulation gives rise to the linear system

$$\begin{bmatrix} \mathbf{B} & -\mathbf{C}^{\mathrm{T}} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{f} \end{bmatrix},$$

where

$$q = \sum_{i} q_{i} \psi_{i}, \quad u = \sum_{m} u_{m} \phi_{m}, \quad f_{m} = \int_{\Omega} f \phi_{m} \, dx,$$

and

$$\mathbf{B} = \left[\int_{\Omega} k^{-1} \psi_i \cdot \psi_j dx\right], \mathbf{C} = \left[\int_{\Omega} \phi_m \operatorname{div}(\psi_j) dx\right], \mathbf{D} = \left[\int_{\Omega} c \phi_m \phi_n dx\right].$$





◀ back ►

• Mixed formulations give rise to saddle point problems:



◄ back ►

- Mixed formulations give rise to saddle point problems:
 - The mixed linear system is indefinite.



- Mixed formulations give rise to saddle point problems:
 - The mixed linear system is indefinite.
- **B** is SPD, and \mathbf{B}^{-1} is dense.



- Mixed formulations give rise to saddle point problems:
 - The mixed linear system is indefinite.
- **B** is SPD, and \mathbf{B}^{-1} is dense.
- **D** is non-negative and $\mathbf{D} + \mathbf{CB}^{-1}\mathbf{C}^{T}$ is SPD.



- Mixed formulations give rise to saddle point problems:
 - The mixed linear system is indefinite.
- **B** is SPD, and \mathbf{B}^{-1} is dense.
- **D** is non-negative and $\mathbf{D} + \mathbf{C}\mathbf{B}^{-1}\mathbf{C}^{T}$ is SPD.
- **B** and **D** (may) contain multiple scales.





• Details at all scales have a strong impact on the solution:



- Details at all scales have a strong impact on the solution:
 - we need to construct subspace correction operators that reflect "all" scales, and employ proper intergrid transfer operators.



- Details at all scales have a strong impact on the solution:
 - we need to construct subspace correction operators that reflect "all" scales, and employ proper intergrid transfer operators.
- Multiscale¹ finite element methods (MsFEMs) honor the subgrid scales and give rise to natural intergrid transfer operators that are adaptive to the local property of the differential operator.

 $^{1}Multiscale methods$: Methods that incorporate fine scale information into a set of coarse scale equations in a way which is consistent with the local property of the differential operator.





 Multigrid methods and domain decomposition methods are for the most part geared toward positive definite systems.



- Multigrid methods and domain decomposition methods are for the most part geared toward positive definite systems.
- To use MG or DD techniques to construct preconditioners for indefinite systems on the present form, we can



◄ back ►

- Multigrid methods and domain decomposition methods are for the most part geared toward positive definite systems.
- To use MG or DD techniques to construct preconditioners for indefinite systems on the present form, we can
 - employ an inexact Uzawa type algorithm and develop a MG or DD preconditioner for the resulting systems,



- Multigrid methods and domain decomposition methods are for the most part geared toward positive definite systems.
- To use MG or DD techniques to construct preconditioners for indefinite systems on the present form, we can
 - employ an inexact Uzawa type algorithm and develop a MG or DD preconditioner for the resulting systems,
 - develop a preconditioner for the full mixed system where some blocks are MG or DD preconditioners for a submatrix (e.g., **B**).



Preconditioning elliptic eqs. with multiple scales

The convergence rate of traditional MG methods and DD methods may deteriorate for elliptic problems with multiple scale coefficients.

Define c(x) = 0 and let k(x) be a scalar periodic function:





✓ back ▶

We now scale the coefficients so that

$$\max(k(x)) / \min(k(x)) = 2^p,$$

and investigate a DD method with an optimal rate of convergence.





Analysis:



Analysis:

• The coarse subspace correction operator does not reflect smaller scales, i.e., the scales that are not resolved by the coarse grid.



Analysis:

- The coarse subspace correction operator does not reflect smaller scales, i.e., the scales that are not resolved by the coarse grid.
 - the subspace correction has poor approximation properties at the "coarse grid nodal points".



Analysis:

- The coarse subspace correction operator does not reflect smaller scales, i.e., the scales that are not resolved by the coarse grid.
 - the subspace correction has poor approximation properties at the "coarse grid nodal points".
 - the coarse to fine grid interpolation operator (induced by the FEM approximation space) do not honor subgrid information.





the inter grid transfer operators are obtained from MsFEM approximation spaces.



◄ back ►

- the inter grid transfer operators are obtained from MsFEM approximation spaces.
- the coarse grid operator is based on an algebraic variant of the MsFEM construction.



- the inter grid transfer operators are obtained from MsFEM approximation spaces.
- the coarse grid operator is based on an algebraic variant of the MsFEM construction.
- The multigrid smoothers are replaced with a single level domain decomposition sweep.



The multiscale finite element method (MsFEM)

The MsFEM seeks a function u^{ms} in a multiscale approximation space V^{ms} and solves the variational formulation

$$a(u^{\mathsf{ms}}, v) = (f, v) \qquad \forall v \in V^{\mathsf{ms}}.$$

Here (\cdot, \cdot) is the inner product in L^2 and $a(\cdot, \cdot)$ is the bilinear form

$$a(u,v) = \int_{\Omega} \nabla u \cdot k \nabla v + cuv \ dx.$$

 V^{ms} is spanned by special multiscale base functions ϕ_i .



◄ back ►

The base functions ϕ_i are constructed by solving a homogeneous equation inside a family of coarse grid elements $\mathcal{K} = \{K\}$.

$$-\nabla \cdot k \nabla u + c u = 0 \quad \text{in } K \in \mathcal{K},$$

and prescribed boundary conditions on ∂K .





Assume that the model equation is discretized at the subgrid scale with a FEM so that the corresponding linear system is on the form

$$\mathbf{Ax} = \mathbf{b}$$
 where $a_{i,j} = a(\xi_i, \xi_j)$, and $b_i = (f, \xi_i)$.

Furthermore, express the MsFEM base functions ϕ_i as a linear combination of the FEM base functions:

$$\phi_i = \sum_j r_{j,i} \xi_j$$

and define the coarse to fine grid interpolation operator $\mathbf{R}_0 = [r_{i,j}]$.



The MsFEM correction, in algebraic form, now reads as follows.

$$\tilde{\mathbf{x}} = \mathbf{x}^{n} + \mathbf{P}^{ms}(\mathbf{x} - \mathbf{x}^{n}) = \mathbf{x}^{n} + \mathbf{Q}^{ms}(\mathbf{b} - \mathbf{A}\mathbf{x}^{n})$$

Here

- $\boldsymbol{\cdot} \ \mathbf{P}^{\text{ms}} = \mathbf{Q}^{\text{ms}} \mathbf{A} = \mathbf{R}_0 (\mathbf{R}_0^T \mathbf{A} \mathbf{R}_0)^{-1} \mathbf{R}_0^T \mathbf{A}$
- + $\mathbf{r}_{:,j} = (\mathbf{I} \sum_i \mathbf{P}_{\mathbf{K}_i})\mathbf{R}_{\mathbf{BC}}$
- $\cdot \mathbf{P}_{\mathbf{K}} = \mathbf{Q}_{\mathbf{K}}\mathbf{A} = \mathbf{R}_{\mathbf{K}}(\mathbf{R}_{\mathbf{K}}^{\mathbf{T}}\mathbf{A}\mathbf{R}_{\mathbf{K}})^{-1}\mathbf{R}_{\mathbf{K}}^{\mathbf{T}}\mathbf{A}$



◄ back ►



◄ back ►

• We shall use the algebraic form of the MsFEM to construct robust coarse grid solvers for "arbitrary" SPD matrices.



17 of 33

back

- We shall use the algebraic form of the MsFEM to construct robust coarse grid solvers for "arbitrary" SPD matrices.
- By combining these coarse solvers with distributed local subspace corrections, we obtain DD preconditioners that are less sensitive to the problem coefficients than traditional DD preconditioners.



- We shall use the algebraic form of the MsFEM to construct robust coarse grid solvers for "arbitrary" SPD matrices.
- By combining these coarse solvers with distributed local subspace corrections, we obtain DD preconditioners that are less sensitive to the problem coefficients than traditional DD preconditioners.
- These auxiliary DD preconditioners will then be used to design effective iterative schemes for mixed systems with multiple scales.



• Sufficient overlap: $\forall x \in \Omega \quad \exists i : \operatorname{distance}(x, \partial \Omega_i) > \delta H.$



◀ back ▶

• Sufficient overlap: $\forall x \in \Omega \quad \exists i : \text{distance}(x, \partial \Omega_i) > \delta H$.





◄ back ►

• Sufficient overlap: $\forall x \in \Omega \quad \exists i : distance(x, \partial \Omega_i) > \delta H$.



• Bounded overlap: at most C subdomains overlap.



◀ back ►

• Sufficient overlap: $\forall x \in \Omega \quad \exists i : \operatorname{distance}(x, \partial \Omega_i) > \delta H$.



• Bounded overlap: at most C subdomains overlap.





◄ back ▶

A super convergence result: In 1D, that is if $\Omega \subset \mathcal{R}$, then $H^1(\Omega) = V^{\mathsf{ms}} + H^1_0(\mathcal{K})$.



Since V^{ms} is orthogonal to $H_0^1(\mathcal{K})$ with respect to $a(\cdot, \cdot)$, this decomposition is a direct sum.



If u is the solution to the variational formulation

 $u \in H^1(\Omega)$ such that a(u,v) = (f,v) $\forall v \in H^1(\Omega)$,

then $u = u^{ms} + u^*$ where

 $\begin{array}{ll} u^* \in H^1_0(\mathcal{K}) & \text{such that} & a(u^*,v) = (f,v) & \forall v \in H^1_0(\mathcal{K}), \\ u^{ms} \in V^{ms} & \text{such that} & a(u^{ms},v) = (f,v) & \forall v \in V^{ms}. \end{array}$

The preconditioner $\Psi^{-1} = \mathbf{Q}^{ms} + \sum_{\mathbf{K} \in \mathcal{K}} \mathbf{Q}_{\mathbf{K}}$ is an ideal preconditioner for one dimensional problems: $\Psi^{-1}\mathbf{A} = \mathbf{I}$.



Additive and multiplicative DD algorithms

The DD preconditioner is now determined by the order in which we perform the subspace corrections. Two possible choices are



For the AS algorithm, the resulting DD preconditioner Ψ becomes

$$\Psi^{-1} = \mathbf{Q}^{\mathsf{ms}} + \sum_{\Omega_i \in \mathcal{F}_1} \mathbf{Q}_{\Omega_i} + \sum_{\Omega_i \in \mathcal{F}_2} \mathbf{Q}_{\Omega_i}, \quad \mathbf{Q} = \mathbf{R} (\mathbf{R}^{\mathsf{T}} \mathbf{A} \mathbf{R})^{-1} \mathbf{R}^{\mathsf{T}}.$$



ৰ back 🕨

Condition number estimates

The Schwarz analysis framework can be used to estimate the condition number of the preconditioned linear system $\Psi^{-1}A$.

For the two-level additive Schwarz algorithm we have

$$\kappa(\boldsymbol{\Psi}_{\mathbf{A}}^{-1}\mathbf{A}) \le C_0(1+C)\gamma(k,c).$$

Here C_0 depends on the subdomain overlap and $\gamma(k, c)$ depends on the regularity of the coefficients k and c. In particular, $\gamma(k, c)$ will depend strongly on coefficient aspect ratios if we do not use a multiscale coarse solver.



Multiscale DD preconditioner as part of multigrid

The multiscale DD preconditioner construction can also be incorporated into a multigrid framework.



Here the algebraic form of the MsFEM is used to construct intergrid transfer operators, and linear systems at the next coarser level.





◀ back ►

 The smoothing sweeps must reflect subscale information, e.g., a sweep of a single level DD preconditioner.



- The smoothing sweeps must reflect subscale information, e.g., a sweep of a single level DD preconditioner.
- By employing algebraic MsFEM to create the coarser level linear systems we automatically ensure that:



- The smoothing sweeps must reflect subscale information, e.g., a sweep of a single level DD preconditioner.
- By employing algebraic MsFEM to create the coarser level linear systems we automatically ensure that:
 - interpolation operators properly reflect subscale information,



- The smoothing sweeps must reflect subscale information, e.g., a sweep of a single level DD preconditioner.
- By employing algebraic MsFEM to create the coarser level linear systems we automatically ensure that:
 - interpolation operators properly reflect subscale information,
 - coarse systems incorporate subgrid information in a manner that is consistent with local properties of the elliptic operator.



Two alternative solution strategies

One can solve the mixed system with PCG by hybridization:

$$\mathbf{S}\mathbf{u} = \mathbf{f}, \qquad \mathbf{S} = \mathbf{D} + \mathbf{C}\mathbf{B}^{-1}\mathbf{C}^{\mathrm{T}},$$

for u with a suitable preconditioner, e.g., $\mathbf{M}_S = \mathbf{D} + \mathbf{C} \mathbf{B}_0^{-1} \mathbf{C}^T.$

 $\mathbf{M}_{\mathbf{S}}$ is the matrix that we obtain from a TPFA finite volume scheme with interface transmissibilities equal to

$$b_{i,i}^{-1} = \left(\int_{\Omega} \psi_i \cdot k^{-1} \psi_i \, dx\right)^{-1}.$$



..., or one can solve the full system using preconditioned GMRES:

$$\begin{bmatrix} \Psi_B^{-1} & \Psi_B^{-1} C^T \Psi_S^{-1} \\ 0 & \Psi_S^{-1} \end{bmatrix} \begin{bmatrix} B & -C^T \\ C & D \end{bmatrix} \begin{bmatrix} q \\ u \end{bmatrix} = \begin{bmatrix} \Psi_B^{-1} C^T \Psi_S^{-1} f \\ \Psi_S^{-1} f \end{bmatrix}.$$

Note that solving the mixed linear system is equivalent to solving

$$\begin{bmatrix} \mathbf{B} & -\mathbf{C}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{f} \end{bmatrix}$$

and that

$$\begin{bmatrix} \mathbf{B} & -\mathbf{C}^{\mathbf{T}} \\ \mathbf{0} & \mathbf{S} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{B}^{-1} & \mathbf{B}^{-1}\mathbf{C}^{\mathbf{T}}\mathbf{S}^{-1} \\ \mathbf{0} & \mathbf{S}^{-1} \end{bmatrix}.$$



back

Numerical test case (flow in porous media)

For our numerical examples we extract permeabilities $\lambda(x)$ and porosities $\phi(x)$ from an oil reservoir model that was used to test and validate upscaling techniques for reservoir simulation.





Map of the logarithm of the Permeability in the top layer.

Map of the logarithm of the Permeability in the bottom layer.



The full 3D reservoir model consist of $60 \times 220 \times 85$ cells.

The top 35 layers is a so-called Tarbert formation, while the bottom 50 layers is a fluvial Upper Ness formation.

For our numerical examples, we define $k(x) = \lambda(x)$ and $c(x) = \nu \phi(x)$ where ν is a constant positive parameter.

We thus test the proposed precondition iterative schemes on mixed FEM equations that arise from the following equation:

$$\begin{array}{rclcrcl} -\nabla\cdot\lambda(x)\nabla u+\nu\phi(x)u&=&f\quad \text{in}\quad\Omega,\\ (-\lambda(x)\nabla u)\cdot n&=&0\quad \text{on}\quad\partial\Omega. \end{array}$$



back



each iteration of the PCG algorithm involves solving two SPD systems: one for B (the action of S) and one for M_S.



29 of 33

- each iteration of the PCG algorithm involves solving two SPD systems: one for \mathbf{B} (the action of \mathbf{S}) and one for $\mathbf{M}_{\mathbf{S}}$.
- For the PCG algorithm we report $N_0(N_B/N_{M_S})$ where



- each iteration of the PCG algorithm involves solving two SPD systems: one for ${f B}$ (the action of ${f S}$) and one for ${f M}_{f S}$.
- For the PCG algorithm we report $N_0(N_B/N_{M_S})$ where
 - $N_0 = \#$ PCG iterations for $\mathbf{Su} = \mathbf{f}$ with preconditioner $\mathbf{M}_{\mathbf{S}}$.



- each iteration of the PCG algorithm involves solving two SPD systems: one for ${f B}$ (the action of ${f S}$) and one for ${f M}_{f S}$.
- For the PCG algorithm we report $N_0(N_B/N_{M_S})$ where
 - $N_0 = \#$ PCG iterations for $\mathbf{Su} = \mathbf{f}$ with preconditioner $\mathbf{M_S}$.
 - $N_B = \#$ PCG iterations for $\mathbf{Bp} = \mathbf{q}$ with preconditioner $\Psi_{\mathbf{B}}$.



- each iteration of the PCG algorithm involves solving two SPD systems: one for ${f B}$ (the action of ${f S}$) and one for ${f M}_{f S}$.
- For the PCG algorithm we report $N_0(N_B/N_{M_S})$ where
 - $N_0 = \#$ PCG iterations for $\mathbf{Su} = \mathbf{f}$ with preconditioner $\mathbf{M_S}$.
 - $N_B = \#$ PCG iterations for $\mathbf{Bp} = \mathbf{q}$ with preconditioner $\Psi_{\mathbf{B}}$.
 - $N_S = \#$ PCG iterations for $\mathbf{M}_{\mathbf{S}}\mathbf{u} = \mathbf{f}$ with preconditioner $\Psi_{\mathbf{S}}$.



Each iteration of the PGMRES involves one sweep of each DD preconditioner.

Indeed, the action of the full preconditioner on a vector $[\mathbf{p}, \mathbf{v}]$ is

$$\begin{bmatrix} \Psi_{\mathrm{B}}^{-1} & \Psi_{\mathrm{B}}^{-1} \mathbf{C}^{\mathrm{T}} \Psi_{\mathrm{S}}^{-1} \\ \mathbf{0} & \Psi_{\mathrm{S}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \Psi_{\mathrm{B}}^{-1} (\mathbf{p} + \mathbf{C}^{\mathrm{T}} \Psi_{\mathrm{S}}^{-1} \mathbf{v}) \\ \Psi_{\mathrm{S}}^{-1} \mathbf{v} \end{bmatrix}$$

Hence, we compute first $\mathbf{r} = \Psi_{\mathbf{S}}^{-1}\mathbf{v}$ and then $\Psi_{\mathbf{B}}^{-1}(\mathbf{p} + \mathbf{C}^{T}\mathbf{r})$. Thus, for the PGMRES algorithm we have $N_{S} = N_{B} = N_{0}$.



2D test-cases sampled from the reservoir: $\nu = 10^{-4}$.

Case 1: Top layer in Tarbert formation.

DD algorithm	PGMR	Ms-PGMR	PCG	Ms-PCG
Additive Swz.	199	53	15(161/448)	10(53/56)
Multipt. Swz.	55	20	13(28/206)	10(22/25)

Case 2: Bottom layer in Upper Ness formation.

DD algorithm	PGMR	Ms-PGMR	PCG	Ms-PCG
Additive Swz.	183	57	10(98/378)	10(44/88)
Multipt. Swz.	73	21	10(45/204)	10(22/33)



Full 3D test-cases: Tarbert formation and Upper Ness formation

Geomodel	Tarbert		Upper Ness	
Algorithm	Ms-PGMR	Ms-PCG	Ms-PGMR	Ms-PCG
$\nu = 1$	20	9(30/29)	8	5(13/5)
$\nu = 10^{-2}$	40	10(33/72)	45	9(28/98)
$\nu = 10^{-4}$	59	10(33/88)	150	10(33/255)





 Multiscale methods can be used to construct MG and DD preconditioners that accelerate PCG and PGRMRES significantly compared with traditional MG and DD preconditioners.



- Multiscale methods can be used to construct MG and DD preconditioners that accelerate PCG and PGRMRES significantly compared with traditional MG and DD preconditioners.
- The PGMRES algorithm performed better than the PCG algorithm for most problems, but it also requires more memory.



- Multiscale methods can be used to construct MG and DD preconditioners that accelerate PCG and PGRMRES significantly compared with traditional MG and DD preconditioners.
- The PGMRES algorithm performed better than the PCG algorithm for most problems, but it also requires more memory.
- Combined with a mixed multiscale FEM, this methodology can help bridge the gap between the geoscale and the simulation scale in reservoir simulation.

