Toward black-box adaptive domain decomposition methods

Frédéric Nataf

Laboratory J.L. Lions (LJLL), CNRS, Alpines Inria and Univ. Paris VI

joint work with
Victorita Dolean (Univ. Nice Sophia-Antipolis)
Patrice Hauret (Michelin, Clermont-Ferrand)
Frédéric Hecht (LJLL)
Pierre Jolivet (LJLL)
Clemens Pechstein (Johannes Kepler Univ., Linz)
Robert Scheichl (Univ. Bath)
Nicole Spillane (LJLL)

JLPC Lyon 2013

Outline

Outline

Motivation

Large Sparse discretized system with strongly heterogeneous coefficients (high contrast, nonlinear, multiscale)

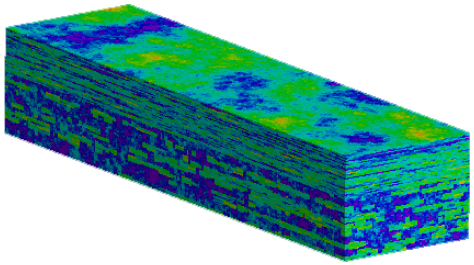
E.g. Darcy pressure equation, P^1-finite elements:

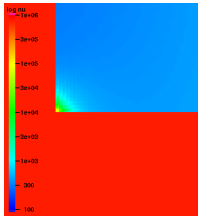
$$A\mathbf{U} = \mathbf{F}$$

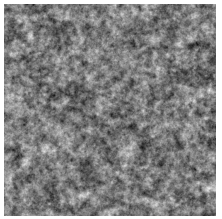
$$\text{cond}(A) \sim \frac{\alpha_{\max}}{\alpha_{\min}} h^{-2}$$

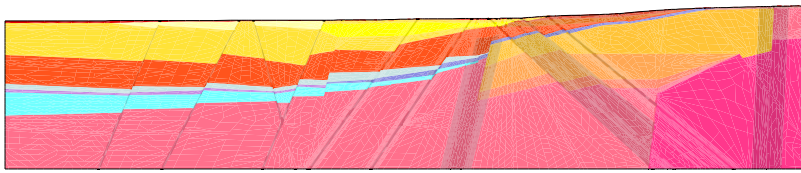
Goal:
iterative solvers
robust in size and heterogeneities

Applications:
flow in heterogeneous / stochastic / layered media
structural mechanics
time dependent waves
etc.

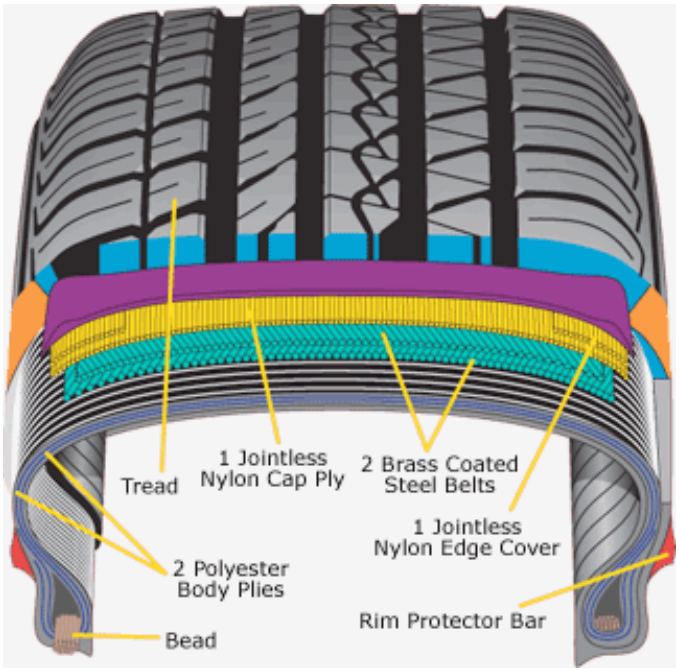








Problem



$$-\mathrm{div}(\sigma(\mathbf{u})) = \mathbf{f},$$

where

$$\sigma_{ij}(\mathbf{u}) = 2\mu\varepsilon_{ij}(\mathbf{u}) + \lambda\delta_{ij}\mathrm{div}(\mathbf{u}), \quad \varepsilon_{ij}(\mathbf{u}) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right),$$
$$\mathbf{f} = (0, g)^T = (0, 10)^T, \quad \mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}.$$

After a FEM (fintie element method) discretization:

$$\Rightarrow A\mathbf{U} = \mathbf{F}$$

where A is a large sparse highy heterogeneous matrix.

Black box solvers (solve(MAT,RHS,SOL))

	Direct Solvers	Iterative Solvers
Pros	Robustness	Naturally $\|\|$
Cons	Difficult to $\|\|$	Robustness

Domain Decomposition Methods (DDM): Hybrid solver $\rightarrow$ should be naturally parallel and robust

General form:

$$Au = f, \text{ solved with PCG for a preconditioner } M^{-1}.$$

What's classical: Robustness with respect to problem size (scalability)
What's New here: Provable Robustness in the SPD case with respect to:

- Coefficients jumps
- System of PDEs

Black box solvers (solve(MAT,RHS,SOL))

	Direct Solvers	Iterative Solvers
Pros	Robustness	Naturally $\|\|$
Cons	Difficult to $\|\|$	Robustness

Domain Decomposition Methods (DDM): Hybrid solver $\rightarrow$ should be naturally parallel and robust

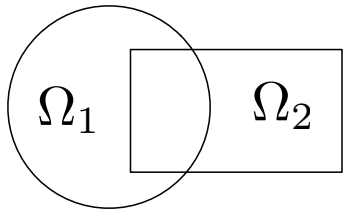
General form:

$$Au = f, \text{ solved with PCG for a preconditioner } M^{-1}.$$

What's classical: Robustness with respect to problem size (scalability)
What's New here: Provable Robustness in the SPD case with respect to:

- Coefficients jumps
- System of PDEs

The original Schwarz Method (H.A. Schwarz, 1870)



$$-\Delta(u) = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega.$$

Schwarz Method : $(u_1^n, u_2^n) \to (u_1^{n+1}, u_2^{n+1})$ with

$$-\Delta(u_1^{n+1}) = f \quad \text{in } \Omega_1$$
$$u_1^{n+1} = 0 \text{ on } \partial\Omega_1 \cap \partial\Omega$$
$$u_1^{n+1} = u_2^n \quad \text{on } \partial\Omega_1 \cap \overline{\Omega_2}.$$

$$-\Delta(u_2^{n+1}) = f \quad \text{in } \Omega_2$$
$$u_2^{n+1} = 0 \text{ on } \partial\Omega_2 \cap \partial\Omega$$
$$u_2^{n+1} = u_1^{n+1} \quad \text{on } \partial\Omega_2 \cap \overline{\Omega_1}.$$

Parallel algorithm, converges but very slowly, overlapping subdomains only.

The parallel version is called **Jacobi Schwarz method (JSM)**.

Jacobi and Schwarz (I): Algebraic point of view

The set of indices is partitioned into two sets $\mathcal{N}_1$ and $\mathcal{N}_2$:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

The block-Jacobi algorithm reads:

$$\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \end{pmatrix} = - \begin{pmatrix} 0 & A_{12} \\ A_{21} & 0 \end{pmatrix} \begin{pmatrix} x_1^{k} \\ x_2^{k} \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

It corresponds to solving a Dirichlet boundary value problem in each subdomain with Dirichlet data taken from the other one at the previous step $\Longleftrightarrow$ Schwarz method with minimal overlap

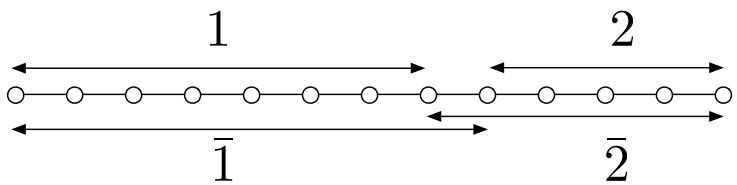


Figure : Domain decomposition with minimal overlap

Jacobi and Schwarz (II): Larger overlap

Let δ be a non negative integer

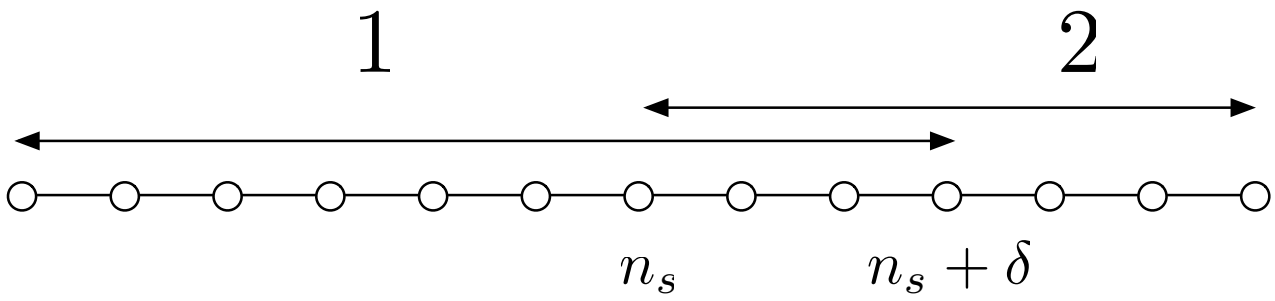


Figure : Domain decomposition with overlap

$$A = \begin{pmatrix} & n_s & & \\ & \vdots & & \\ & & & \\ & n_s + 1 & & \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} & n_s + \delta & & \\ & \vdots & & \\ & & & \\ & n_s & & \end{pmatrix}$$

Figure : Matrix decomposition Without or with overlap

Strong and Weak scalability

How to evaluate the efficiency of a domain decomposition?

Strong scalability (Amdahl)

"How the solution time varies with the number of processors for a fixed *total* problem size"

Weak scalability (Gustafson)

"How the solution time varies with the number of processors for a fixed problem size *per processor*."

Not achieved with the one level method

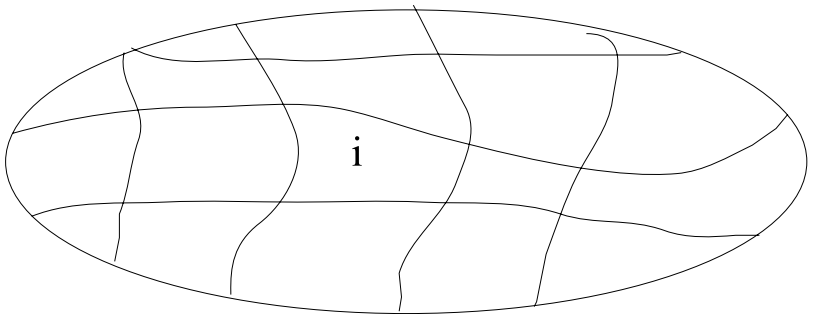
Number of subdomains	8	16	32	64
ASM	18	35	66	128

The iteration number increases linearly with the number of subdomains in one direction.

How to achieve scalability

Stagnation corresponds to a few very low eigenvalues in the spectrum of the preconditioned problem. They are due to the lack of a global exchange of information in the preconditioner.

$$-\Delta u = f \text{ in } \Omega$$
$$u = 0 \text{ on } \partial\Omega$$



The mean value of the solution in domain i depends on the value of f on all subdomains.
A classical remedy consists in the introduction of a coarse problem that couples all subdomains. This is closely related to deflation technique classical in linear algebra (see Nabben and Vuik's papers in SIAM J. Sci. Comp, 200X) and multigrid techniques.

Adding a coarse space

We add a coarse space correction (*aka* second level)
Let V_H be the coarse space and Z be a basis, $V_H = \operatorname{span} Z$,
writing $R_0 = Z^T$ we define the two level preconditioner as:

$$M^{-1}_{ASM,2} := R_0^T (R_0 A R_0^T)^{-1} R_0 + \sum_{i=1}^{N} R_i^T A_i^{-1} R_i.$$

The Nicolaides approach is to use the kernel of the operator as a coarse space, this is the constant vectors, in local form this writes:

$$Z := (R_i^T D_i R_i \mathbf{1})_{1 \leq i \leq N}$$

where D_i are chosen so that we have a partition of unity:

$$\sum_{i=1}^{N} R_i^T D_i R_i = Id.$$

Theoretical convergence result

Theorem (Widlund, Dryja)

Let $M^{-1}_{ASM,2}$ *be the two-level additive Schwarz method:*

$$\kappa(M^{-1}_{ASM,2}\, A) \leq C\left(1 + \frac{H}{\delta}\right)$$

where δ *is the size of the overlap between the subdomains and* H *the subdomain size.*

This does indeed work very well

Number of subdomains	8	16	32	64
ASM	18	35	66	128
ASM + Nicolaides	20	27	28	27

Failure for Darcy equation with heterogeneities

$$
\begin{aligned}
-\nabla \cdot (\alpha(x,y)\nabla u) &= 0 \quad \text{in} \quad \Omega \subset \mathbb{R}^2, \\
u &= 0 \quad \text{on} \quad \partial\Omega_D, \\
\frac{\partial u}{\partial n} &= 0 \quad \text{on} \quad \partial\Omega_N.
\end{aligned}
$$

Decomposition

$\alpha(x,y)$

Jump	1	10	10^2	10^3	10^4
ASM	39	45	60	72	73
ASM + Nicolaides	30	36	50	61	65

Our approach

Fix the problem by an optimal and proven choice of a coarse space Z.

Outline

Objectives

Strategy

Define an appropriate coarse space $V_{H2} = \text{span}(Z_2)$ and use the framework previously introduced, writing $R_0 = Z_2^T$ the two level preconditioner is:

$$P_{ASM2}^{-1} := R_0^T (R_0 A R_0^T)^{-1} R_0 + \sum_{i=1}^{N} R_i^T A_i^{-1} R_i.$$

The coarse space must be

- Local (calculated on each subdomain) $\rightarrow$ parallel
- Adaptive (calculated automatically)
- Easy and cheap to compute
- Robust (must lead to an algorithm whose convergence is proven not to depend on the partition nor the jumps in coefficients)

Abstract eigenvalue problem

Gen.EVP per subdomain:

Find $p_{j,k} \in V_{h|\Omega_j}$ and $\lambda_{j,k} \geq 0$:

$$a_{\Omega_j}(p_{j,k},\, v) \;=\; \lambda_{j,k}\, a_{\Omega_j^\circ}(\Xi_j p_{j,k},\, \Xi_j v) \qquad \forall v \in V_{h|\Omega_j}$$

$$A_j \mathbf{p}_{j,k} \;=\; \lambda_{j,k}\, \mathbf{X}_j A_j^\circ \mathbf{X}_j\, \mathbf{p}_{j,k} \qquad (\mathbf{X}_j \ldots \text{diagonal})$$

$a_D \ldots$ restriction of a to D

In the two-level ASM:

Choose first m_j eigenvectors per subdomain:

$$V_0 \;=\; \text{span}\left\{\Xi_j p_{j,k}\right\}_{k=1,\ldots,m_j}^{j=1,\ldots,N}$$

This automatically includes Zero Energy Modes.

Abstract eigenvalue problem

Gen.EVP per subdomain:

Find $p_{j,k} \in V_{h|\Omega_j}$ and $\lambda_{j,k} \geq 0$:

$$a_{\Omega_j}(p_{j,k},\, v) \;=\; \lambda_{j,k}\, a_{\Omega_j^\circ}(\Xi_j p_{j,k},\, \Xi_j v) \qquad \forall v \in V_{h|\Omega_j}$$

$$A_j \mathbf{p}_{j,k} \;=\; \lambda_{j,k}\, \mathbf{X}_j A_j^\circ \mathbf{X}_j\, \mathbf{p}_{j,k} \qquad (\mathbf{X}_j \ldots \text{diagonal})$$

$a_D \ldots$ restriction of a to D

In the two-level ASM:

Choose first m_j eigenvectors per subdomain:

$$V_0 \;=\; \text{span}\left\{\Xi_j p_{j,k}\right\}_{k=1,\ldots,m_j}^{j=1,\ldots,N}$$

This automatically includes Zero Energy Modes.

Comparison with existing works

Galvis & Efendiev (SIAM 2010):

$$\int_{\Omega_j} \kappa \nabla p_{j,k} \cdot \nabla v \, dx \;=\; \lambda_{j,k} \int_{\Omega_j} \kappa \, p_{j,k} \, v \, dx \qquad \forall v \in V_{h|\Omega_j}$$

Efendiev, Galvis, Lazarov & Willems (submitted):

$$a_{\Omega_j}(p_{j,k},\, v) \;=\; \lambda_{j,k} \sum_{i \in \text{neighb}(j)} a_{\Omega_j}(\xi_j \, \xi_i \, p_{j,k},\, \xi_j \, \xi_i \, v) \qquad \forall v \in V_{|\Omega_j}$$

$\xi_j \ldots$ partition of unity, calculated adaptively (MS)

Our gen.EVP:

$$a_{\Omega_j}(p_{j,k},\, v) \;=\; \lambda_{j,k} \, a_{\Omega_j^\circ}(\Xi_j p_{j,k},\, \Xi_j v) \qquad \forall v \in V_{h|\Omega_j}$$

both matrices typically singular $\Longrightarrow$ $\lambda_{j,k} \in [0,\, \infty]$

Two technical assumptions.

Theorem (Spillane, Dolean, Hauret, N., Pechstein, Scheichl)

If for all j: $0 < \lambda_{j,m_{j+1}} < \infty$*:*

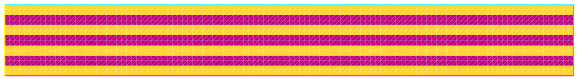
$$\kappa(M_{ASM,2}^{-1}A) \leq (1+k_0)\left[2 + k_0(2k_0+1)\max_{j=1}^{N}\left(1+\frac{1}{\lambda_{j,m_j+1}}\right)\right]$$

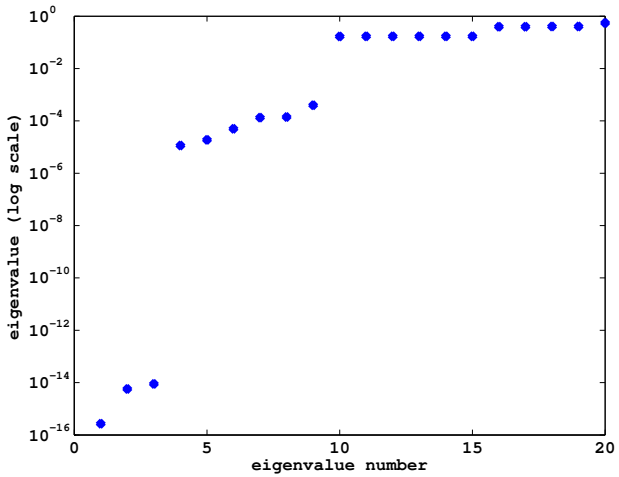
Possible criterion for picking m_j: (used in our Numerics)

$$\lambda_{j,m_j+1} < \frac{\delta_j}{H_j}$$

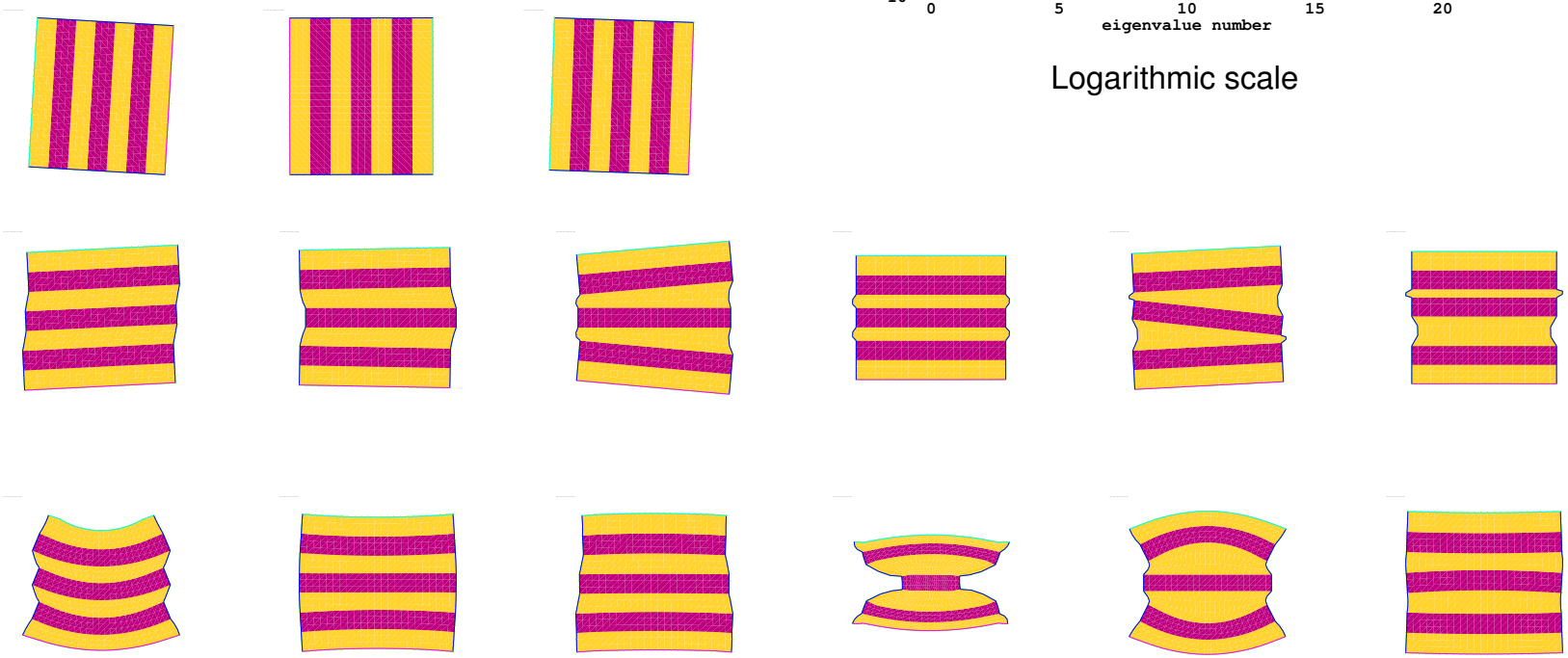
H_j . . . subdomain diameter, δ_j . . . overlap

Eigenvalues and eigenvectors

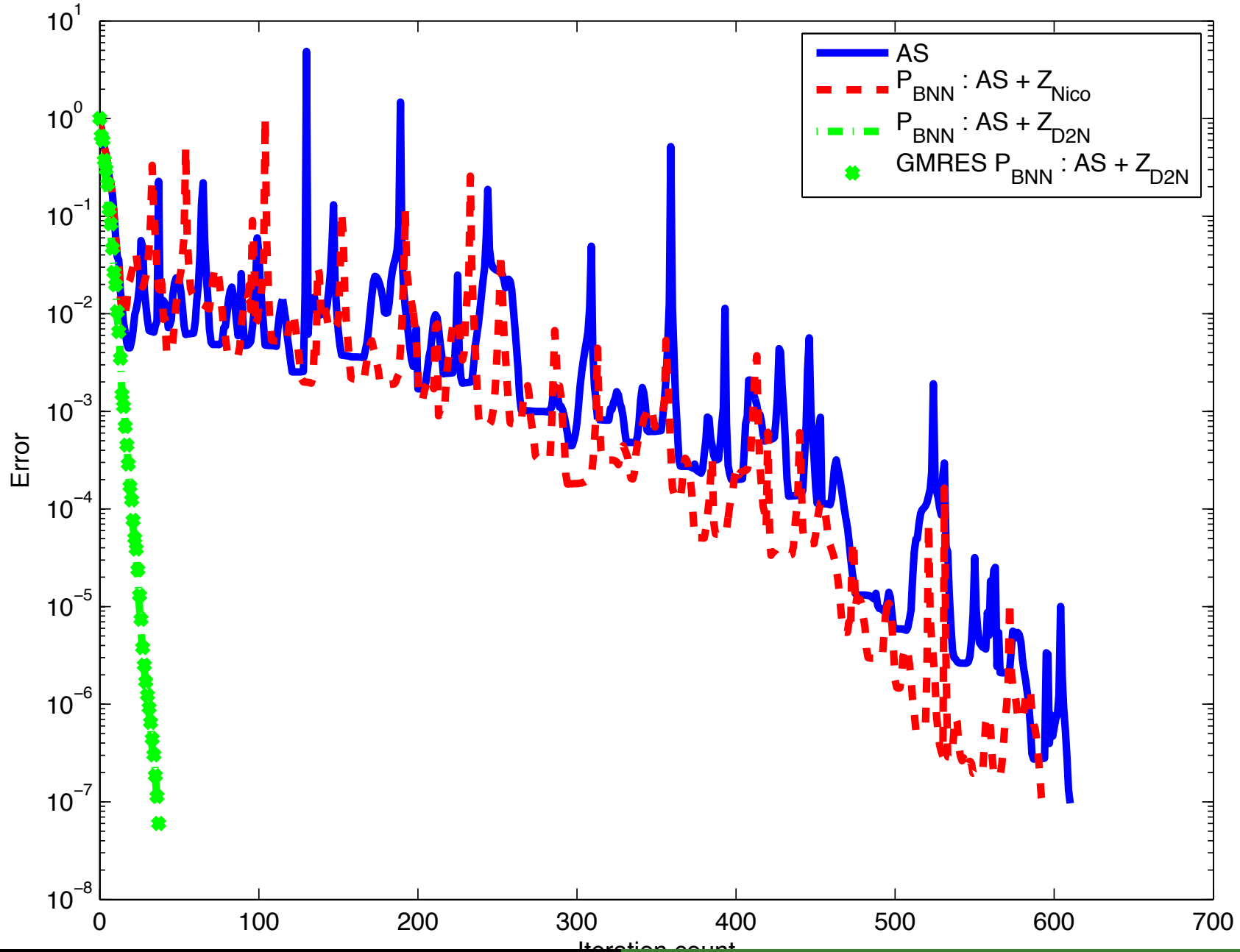




Logarithmic scale



Convergence



Numerical results via a Domain Specific Language

`FreeFem++` (`http://www.freefem.org/ff++`), with:

- Metis Karypis and Kumar 1998
- SCOTCH Chevalier and Pellegrini 2008
- UMFPACK Davis 2004
- ARPACK Lehoucq et al. 1998
- MPI Snir et al. 1995
- Intel MKL
- PARDISO Schenk et al. 2004
- MUMPS Amestoy et al. 1998
- PaStiX Hénon et al. 2005
- PETSC

Why use a DS(E)L instead of `C/C++/Fortran/..` ?

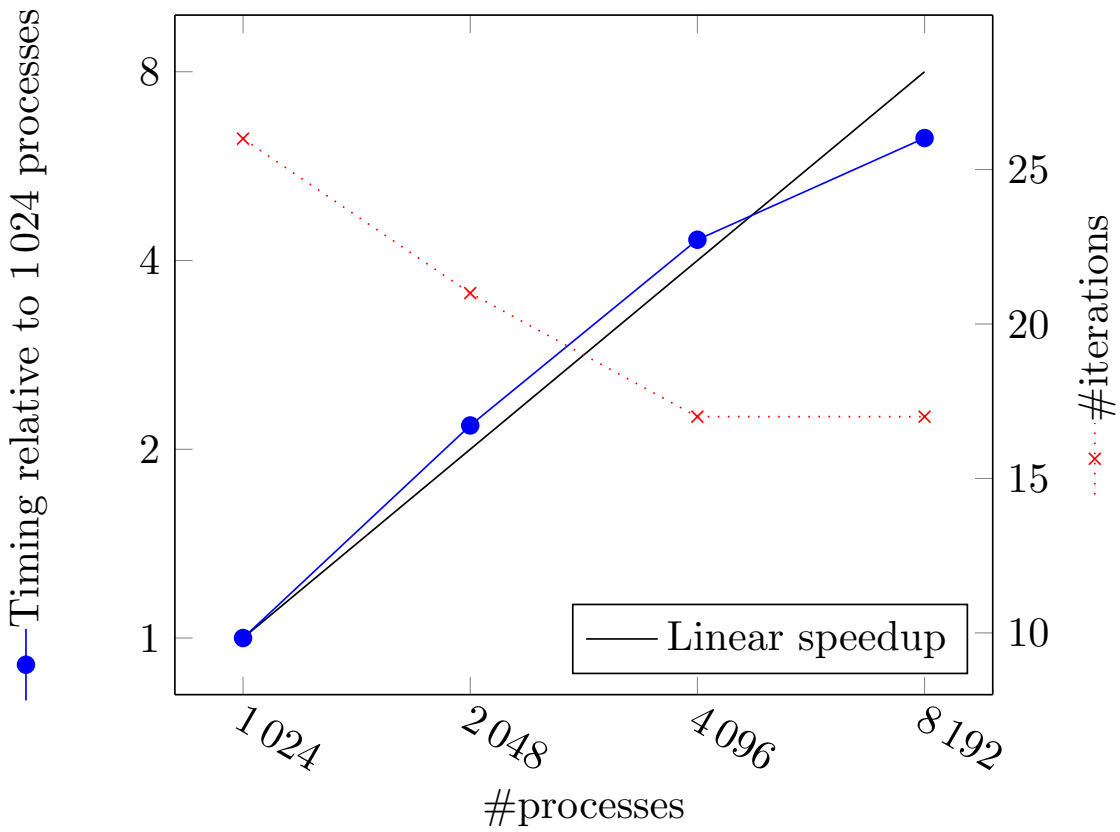
- performances close to low-level language implementation,
- hard to beat something as simple as:

$$\begin{aligned}\mathbf{varf}\ \mathrm{a}(\mathrm{u},\ \mathrm{v}) = \mathbf{int3d}(\mathrm{mesh})([\mathbf{dx}(\mathrm{u}),\ \mathbf{dy}(\mathrm{u}),\ \mathbf{dz}(\mathrm{u})]'\ *\ [\mathbf{dx}(\mathrm{v}),\ \mathbf{dy}(\mathrm{v}),\ \mathbf{dz}(\mathrm{v})]) \\ + \mathbf{int3d}(\mathrm{mesh})(\mathrm{f}\ *\ \mathrm{v}) + \mathbf{on}(\mathrm{boundary\_mesh})(\mathrm{u} = 0)\end{aligned}$$

- Much More in F. Hecht's talk tomorrow

Strong scalability in two dimensions heterogeneous elasticity (P. Jolivet with Frefeem ++)

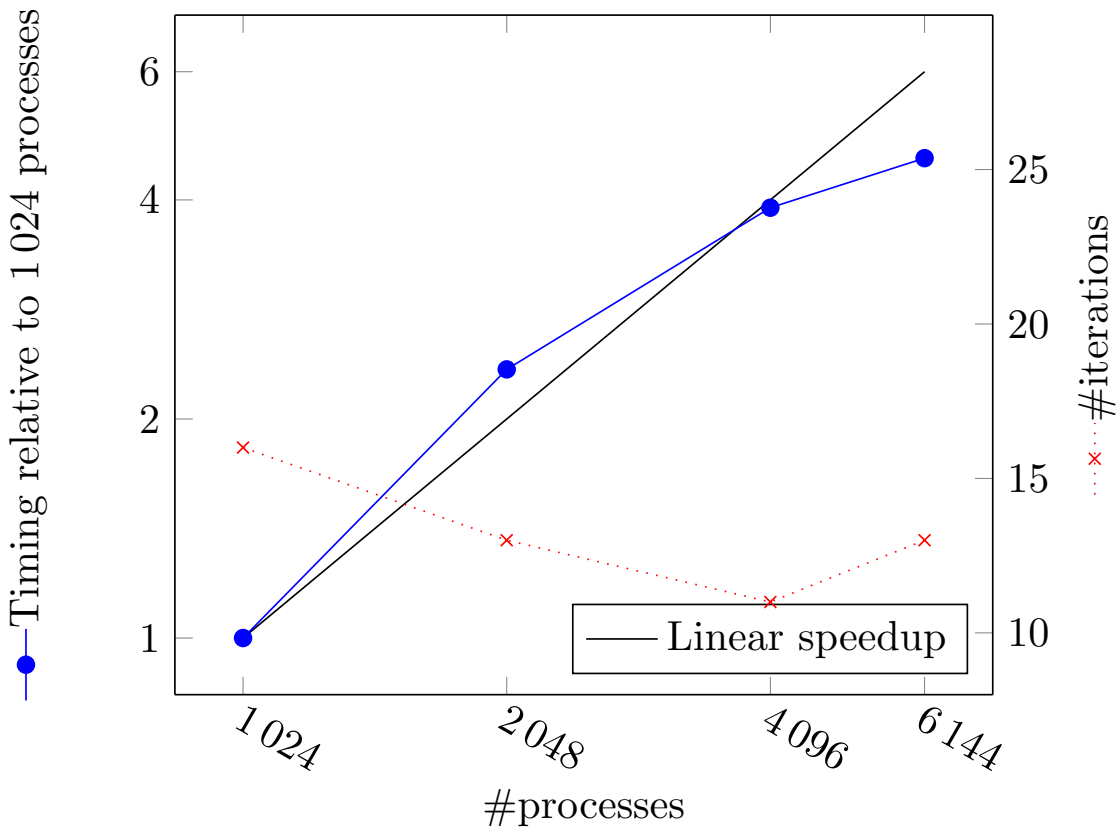
Elasticity problem with heterogeneous coefficients



Speed-up for a 1.2 billion unknowns 2D problem. Direct solvers in the subdomains. Peak performance wall-clock time: 26s.

Strong scalability in three dimensions heterogeneous elasticity

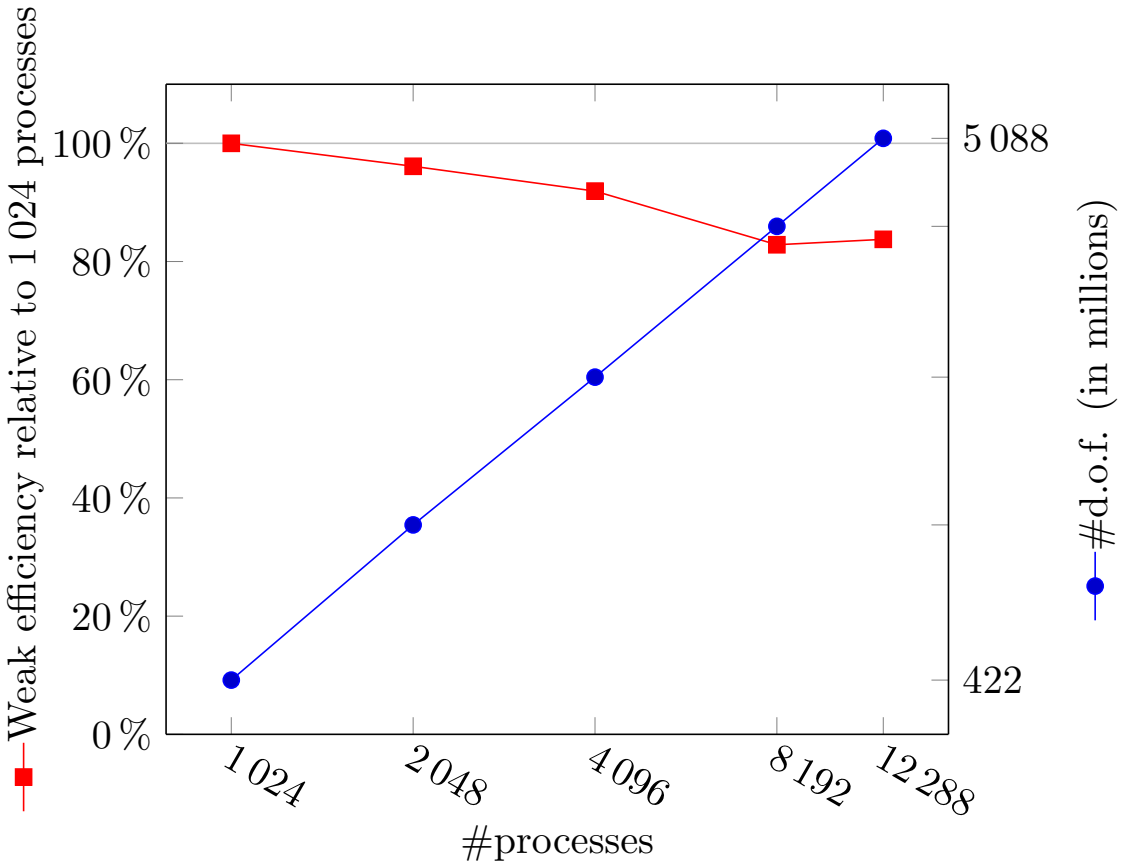
Elasticity problem with heterogeneous coefficients



Speed-up for a 160 million unknowns 3D problem. Direct solvers in subdomains. Peak performance wall-clock time: 36s.

Weak scalability in two dimensions

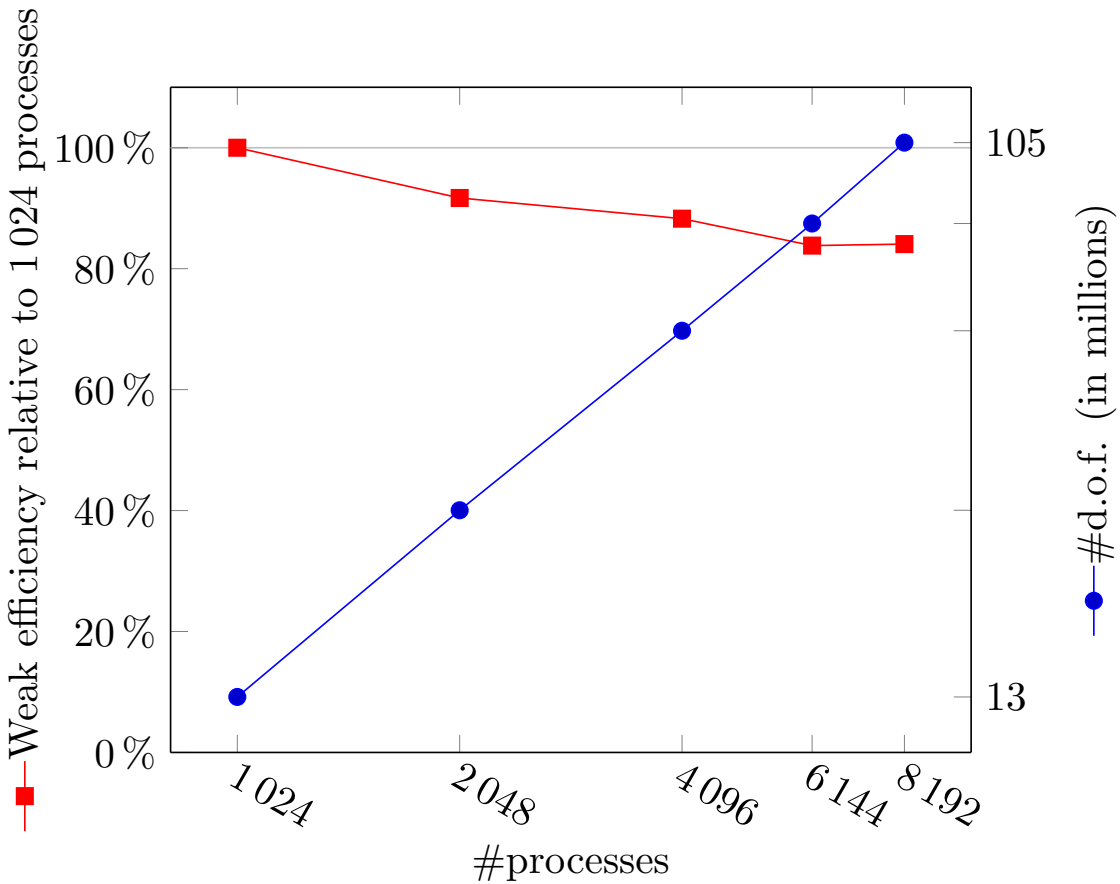
Darcy problems with heterogeneous coefficients



Efficiency for a 2D problem. Direct solvers in the subdomains.
Final size: 22 billion unknowns. Wall-clock time: $\simeq$ 200s.

Weak scalability in three dimensions

Darcy problems with heterogeneous coefficients



Efficiency for a 3D problem. Direct solvers in the subdomains.
Final size: 2 billion unknowns. Wall-clock time: $\simeq$ 200s.

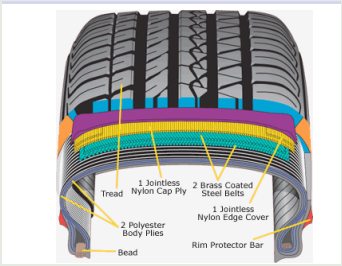
Outline

Conclusion

Summary

- Using generalized eigenvalue problems and projection preconditioning we are able to achieve a targeted convergence rate.
- Works for ASM, BNN and FETI methods
- This process can be implemented in a black box algorithm.

Future work

- Build the coarse space on the fly.
- 

- Multigrid like three (or more) level methods

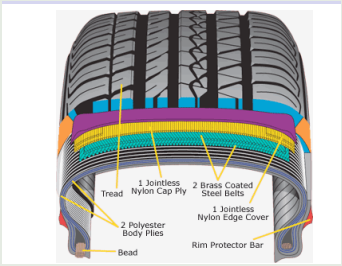
THANK YOU FOR YOUR ATTENTION!

Conclusion

Summary

- Using generalized eigenvalue problems and projection preconditioning we are able to achieve a targeted convergence rate.
- Works for ASM, BNN and FETI methods
- This process can be implemented in a black box algorithm.

Future work

- Build the coarse space on the fly.
- 

- Multigrid like three (or more) level methods

THANK YOU FOR YOUR ATTENTION!