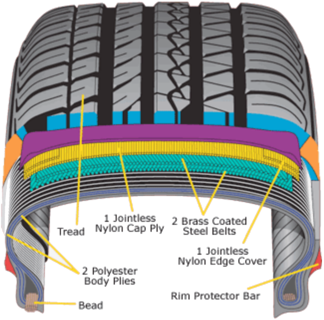
Domain Decomposition Methods with Adaptive Multipreconditioning

Nicole Spillane
(CNRS, CMAP, École Polytechnique)

June 18th, 2018
SIAM UKIE
Student Chapter Meeting
at the University of Bath

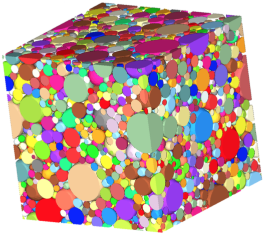


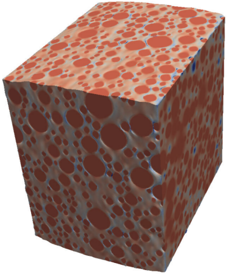
Find $\mathbf{u}_* \in \mathbb{R}^n$ s.t.

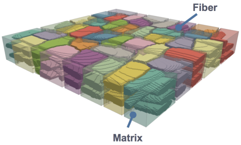
$$\mathbf{K}\mathbf{u}_* = \mathbf{f}$$

where

- $\mathbf{K}$ is symmetric positive definite (spd)
- $\mathbf{K} = \sum_{\{\tau \in \mathcal{T}_h\}} \mathbf{K}_\tau$
- All $\mathbf{K}_\tau$ are symmetric positive semi-definite (spsd).





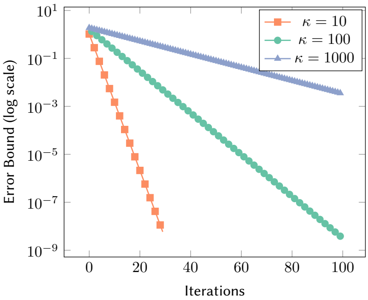


Two observations about parallel linear solvers

1. Direct solvers are robust (but require a lot of memory).
   e.g., Cholesky factorization w/ forward backward substitutions

$$\begin{pmatrix} \mathbf{K} \end{pmatrix} = \begin{pmatrix} \mathbf{L} \end{pmatrix} \begin{pmatrix} \mathbf{L}^\top \end{pmatrix} \qquad \begin{pmatrix} \mathbf{u}_* \end{pmatrix} = \begin{pmatrix} \mathbf{L}^\top \end{pmatrix} \backslash \begin{pmatrix} \mathbf{L} \end{pmatrix} \backslash \begin{pmatrix} \mathbf{f} \end{pmatrix}$$

2. Iterative solvers are naturally parallel (but not robust).
   e.g., Preconditioned Conjugate Gradient



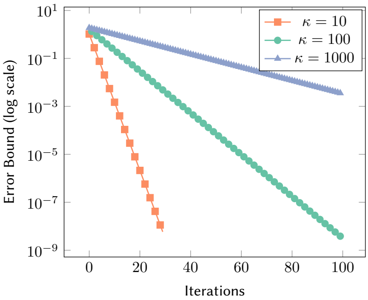
$$\frac{\|\mathbf{u}_* - \mathbf{u}_m\|_\mathbf{A}}{\|\mathbf{u}_* - \mathbf{u}_0\|_\mathbf{A}} \leqslant 2\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^m, \quad \kappa = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

Two observations about parallel linear solvers

1. Direct solvers are robust (but require a lot of memory).
   e.g., Cholesky factorization w/ forward backward substitutions

$$\left(\mathbf{K}\right) = \left(\mathbf{L}\right)\left(\mathbf{L}^\top\right) \qquad \left(\mathbf{u}_*\right) = \left(\mathbf{L}^\top\right) \backslash \left(\mathbf{L}\right) \backslash \left(\mathbf{f}\right)$$

2. Iterative solvers are naturally parallel (but not robust).
   e.g., Preconditioned Conjugate Gradient



$$\frac{\|\mathbf{u}_* - \mathbf{u}_m\|_\mathbf{A}}{\|\mathbf{u}_* - \mathbf{u}_0\|_\mathbf{A}} \leqslant 2\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^m, \quad \kappa = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

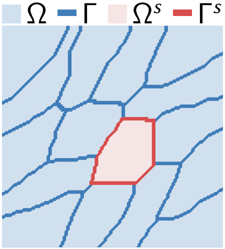
Hybrid Direct/Iterative solvers should be **both naturally parallel and robust**

e.g., Domain Decomposition

Balancing Domain Decomposition

Balancing Domain Decomposition (BDD): $\mathbf{K}\mathbf{u}_* = \mathbf{f}$ ($\mathbf{K}$ spd)

Global Formulation



J. Mandel. Balancing domain decomposition. *Comm. Numer. Methods Engrg.*, 9(3):233–241, 1993.

$$\mathbf{K}\mathbf{u}_* = \mathbf{f}$$

$$\Leftrightarrow$$

$$\begin{pmatrix} \mathbf{K}_{\Gamma\Gamma} & \mathbf{K}_{\Gamma I} \\ \mathbf{K}_{I\Gamma} & \mathbf{K}_{II} \end{pmatrix} \begin{pmatrix} \mathbf{u}_{*,\Gamma} \\ \mathbf{u}_{*,I} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_\Gamma \\ \mathbf{f}_I \end{pmatrix} \begin{matrix} \textit{(boundaries)} \\ \textit{(interiors)} \end{matrix}$$

$$\Leftrightarrow$$

$$\mathbf{u}_{*,I} = \mathbf{K}_{II}^{-1}(\mathbf{f}_I - \mathbf{K}_{I\Gamma}\mathbf{u}_{*,\Gamma})$$

and

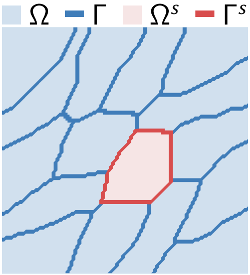
$$\underbrace{(\mathbf{K}_{\Gamma\Gamma} - \mathbf{K}_{\Gamma I}\mathbf{K}_{II}^{-1}\mathbf{K}_{I\Gamma})}_{:=\mathbf{A}}\mathbf{u}_{*,\Gamma} = \underbrace{\mathbf{f}_\Gamma - \mathbf{K}_{\Gamma I}\mathbf{K}_{II}^{-1}\mathbf{f}_I}_{:=\mathbf{b}}.$$

BDD is a hybrid solver:

- It solves $\mathbf{A}\mathbf{u}_{*,\Gamma} = \mathbf{b}$ with PCG.
- Within each iteration, a direct solve is performed: $\mathbf{K}_{II}^{-1}$.

Balancing Domain Decomposition (BDD): $\mathbf{K}\mathbf{u}_* = \mathbf{f}$ ($\mathbf{K}$ spd)

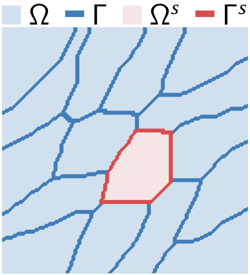
Parallel Formulation



$$\mathbf{A} = \mathbf{K}_{\Gamma\Gamma} - \mathbf{K}_{\Gamma I}\mathbf{K}_{II}^{-1}\mathbf{K}_{I\Gamma}; \; \mathbf{K}_{II} = \begin{pmatrix} \mathbf{K}^1_{I_1 I_1} & & \\ & \ddots & \\ & & \mathbf{K}^N_{I_N I_N} \end{pmatrix}.$$

Balancing Domain Decomposition (BDD): $\mathbf{K}\mathbf{u}_* = \mathbf{f}$ ($\mathbf{K}$ spd)

Parallel Formulation



$$\mathbf{A} = \mathbf{K}_{\Gamma\Gamma} - \mathbf{K}_{\Gamma I}\mathbf{K}_{II}^{-1}\mathbf{K}_{I\Gamma}; \quad \mathbf{K}_{II} = \begin{pmatrix} \mathbf{K}^1_{I_1 I_1} & & \\ & \ddots & \\ & & \mathbf{K}^N_{I_N I_N} \end{pmatrix}.$$

The operator $\mathbf{A}$ is a sum of local contributions :

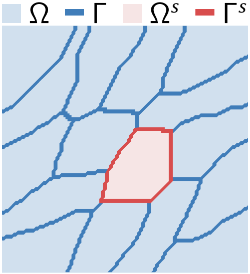
$$\mathbf{A} = \sum_{s=1}^{N} \mathbf{R}^{s\top}\mathbf{S}^s\mathbf{R}^s$$

$$\mathbf{S}^s := \mathbf{K}^s_{\Gamma^s\Gamma^s} - \mathbf{K}^s_{\Gamma^s I^s}(\mathbf{K}^s_{I^s I^s})^{-1}\mathbf{K}^s_{I^s\Gamma^s}$$

$$\Gamma_s \underset{\mathbf{R}^s}{\overset{\mathbf{R}^{s\top}}{\rightleftarrows}} \Gamma$$

Balancing Domain Decomposition (BDD): $\mathbf{K}\mathbf{u}_* = \mathbf{f}$ ($\mathbf{K}$ spd)

Parallel Formulation



Good preconditioner:

- *good* approximation of $\mathbf{A}^{-1}$,
- *cheap* to compute.

$$\mathbf{A} = \mathbf{K}_{\Gamma\Gamma} - \mathbf{K}_{\Gamma I}\mathbf{K}_{II}^{-1}\mathbf{K}_{I\Gamma};\ \mathbf{K}_{II} = \begin{pmatrix} \mathbf{K}^1_{I_1 I_1} & & \\ & \ddots & \\ & & \mathbf{K}^N_{I_N I_N} \end{pmatrix}.$$

The operator $\mathbf{A}$ is a sum of local contributions :

$$\mathbf{A} = \sum_{s=1}^{N} \mathbf{R}^{s\top}\mathbf{S}^s\mathbf{R}^s \quad \Big| \quad \mathbf{S}^s := \mathbf{K}^s_{\Gamma^s\Gamma^s} - \mathbf{K}^s_{\Gamma^s I^s}(\mathbf{K}^s_{I^s I^s})^{-1}\mathbf{K}^s_{I^s\Gamma^s}$$

$$\Gamma_s \underset{\mathbf{R}^s}{\overset{\mathbf{R}^{s\top}}{\rightleftarrows}} \Gamma$$

The preconditioner $\mathbf{H}$ also:

$$\mathbf{H} := \sum_{s=1}^{N} \mathbf{R}^{s\top}\mathbf{D}^s\mathbf{S}^{s-1}\mathbf{D}^s\mathbf{R}^s, \text{ with } \sum_{s=1}^{N} \mathbf{R}^{s\top}\mathbf{D}^s\mathbf{R}^s = \mathbf{I}.$$

Some remarks

- The kernel of $\mathbf{S}^s$ is handled by deflation and

$$\mathbf{H} := \sum_{s=1}^{N} \mathbf{R}^{s\top} \mathbf{D}^s \mathbf{S}^{s\dagger} \mathbf{D}^s \mathbf{R}^s \quad (\cdot^{\dagger}\text{: pseudoinverse}).$$

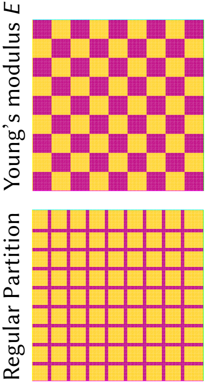
- The choice of the scaling matrices $\mathbf{D}^s$ is important (problems with heterogeneous coefficients or subdomains).
- Eigenvalues of the preconditioned operator satisfy

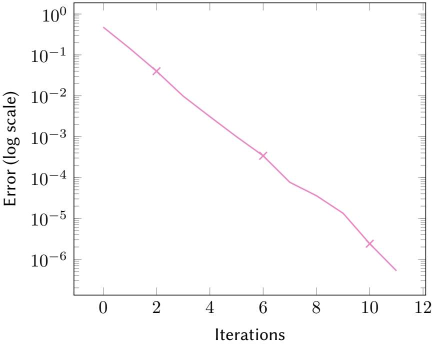
$$\lambda_{\min}(\mathbf{HA}) \geqslant 1$$

as a consequence of: $\sum_{s=1}^{N} \mathbf{R}^{s\top} \mathbf{D}^s \mathbf{R}^s = \mathbf{I}$.

Elasticity with homogeneous subdomains

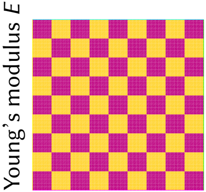
$N = 81$ subdomains, $\nu = 0.4$, $E_1 = 10^7$ and $E_2 = 10^{12}$.

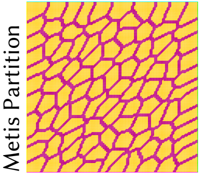


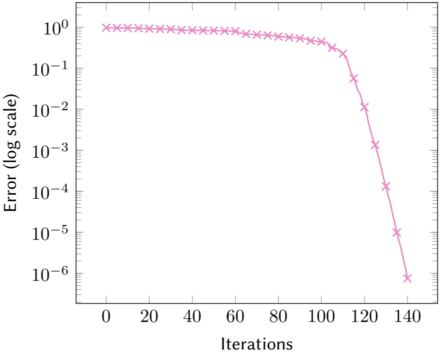


Lack of Robustness: Heterogeneous Elasticity

$N = 81$ subdomains, $\nu = 0.4$, $E_1 = 10^7$ and $E_2 = 10^{12}$.







Three objectives for new DD methods:

- **Reliability**: robustness and scalability.
- **Efficiency**: adapt automatically to difficulty.
- **Simplicity**: non invasive implementation.

GenEO Coarse Space

Deflation, Coarse space: Ideal choice ($(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$)

General convergence result for PCG:

$$\frac{\|\mathbf{x}_* - \mathbf{x}_m\|_{\mathbf{A}}}{\|\mathbf{x}_* - \mathbf{x}_0\|_{\mathbf{A}}} \leqslant 2\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^m, \text{ where } \kappa = \frac{\lambda_{\max}(\mathbf{HA})}{\lambda_{\min}(\mathbf{HA})}.$$

Deflation, Coarse space: Ideal choice ($(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$)

Convergence result for BDD [Mandel, Tezaur, 1996]:

$$\frac{\|\mathbf{x}_* - \mathbf{x}_m\|_{\mathbf{A}}}{\|\mathbf{x}_* - \mathbf{x}_0\|_{\mathbf{A}}} \leqslant 2 \left[\frac{\sqrt{\lambda_{\max}} - 1}{\sqrt{\lambda_{\max}} + 1} \right]^m, \text{ because } \lambda_{\min} \geqslant 1.$$

Deflation, Coarse space: Ideal choice ($(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$)

Convergence result for BDD [Mandel, Tezaur, 1996]:

$$\frac{\|\mathbf{x}_* - \mathbf{x}_m\|_{\mathbf{A}}}{\|\mathbf{x}_* - \mathbf{x}_0\|_{\mathbf{A}}} \leqslant 2 \left[\frac{\sqrt{\lambda_{\max}} - 1}{\sqrt{\lambda_{\max}} + 1} \right]^m, \text{ because } \lambda_{\min} \geqslant 1.$$

$\rightarrow$ **Deflation strategy [Nicolaides, 1987] [Dostál, 1988]**

(i) Compute $(\lambda_k, \mathbf{x}_k)$ such that $\mathbf{H}\mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k$.

(ii) Define $\mathbf{U} := \{\mathbf{x}_k; \lambda_k \geqslant \tau\}$ (span($\mathbf{U}$) is the **coarse space**)
$(\mathbf{I} - \mathbf{\Pi})$: $\mathbf{A}$-orthogonal projection onto span($\mathbf{U}$).

(iii) Solve $\underbrace{\mathbf{\Pi}^\top \mathbf{A}\mathbf{x}_* = \mathbf{\Pi}^\top b}_{\text{DD solver}}$ and $\underbrace{(\mathbf{I} - \mathbf{\Pi}^\top)\mathbf{A}\mathbf{x}_* = (\mathbf{I} - \mathbf{\Pi})^\top \mathbf{b}}_{\text{direct solver}}$.

Guaranteed Convergence: $$\frac{\|\mathbf{x}_* - \mathbf{x}_m\|_{\mathbf{A}}}{\|\mathbf{x}_* - \mathbf{x}_0\|_{\mathbf{A}}} \leqslant 2 \left[\frac{\sqrt{\tau} - 1}{\sqrt{\tau} + 1} \right]^m$$

BUT too expensive to compute.

GenEO coarse space for BDD

(Generalized Eigenvalues in the Overlaps)

1. In each subdomain we want $\langle \mathbf{x}^s, \mathbf{S}^s \mathbf{x}^s \rangle \geqslant \tau \langle \mathbf{x}^s, \mathbf{R}^s \mathbf{A} \mathbf{R}^{s\top} \mathbf{x}^s \rangle$ for all $\mathbf{x}^s$ so we solve:

$$\mathbf{S}^s \mathbf{x}_k^s = \lambda_k^s \mathbf{R}^s \mathbf{A} \mathbf{R}^{s\top} \mathbf{x}_k^s, \quad \left(\mathbf{A} = \sum_{s=1}^{N} \mathbf{R}^{s\top} \mathbf{S}^s \mathbf{R}^s \right).$$

2. Let the coarse space be

$$\mathbf{U} := \left\{ \mathbf{R}^{s\top} \mathbf{x}_k^s ; s = 1, \dots, N \quad \text{and} \quad \lambda_k^s \leqslant \tau \right\} (\text{ w/ } \textit{e.g. } \tau = 0.1).$$

3. Let $(\mathbf{I} - \mathbf{\Pi})$ be the $\mathbf{A}$-orthogonal projection onto span$(\mathbf{U})$

$$\mathbf{A}\mathbf{x}_* = \mathbf{b} \Leftrightarrow \underbrace{\mathbf{\Pi}^\top \mathbf{A}\mathbf{x}_* = \mathbf{\Pi}^\top \mathbf{b}}_{\text{DD solver}} \quad \text{and} \quad \underbrace{(\mathbf{I} - \mathbf{\Pi}^\top)\mathbf{A}\mathbf{x}_* = (\mathbf{I} - \mathbf{\Pi})^\top \mathbf{b}}_{\text{direct solver}}.$$

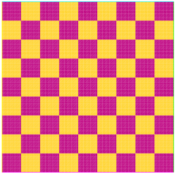
$\rightarrow$ **Guaranteed convergence:**

$$\frac{\|\mathbf{x}_* - \mathbf{x}_m\|_{\mathbf{A}}}{\|\mathbf{x}_* - \mathbf{x}_0\|_{\mathbf{A}}} \leqslant 2 \left[\frac{\sqrt{\mathcal{N}/\tau} - 1}{\sqrt{\mathcal{N}/\tau} + 1} \right]^m, \mathcal{N}: \text{ number of neighbours.}$$

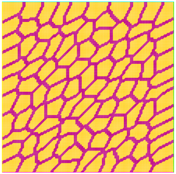
Numerical Illustration : Heterogeneous Elasticity

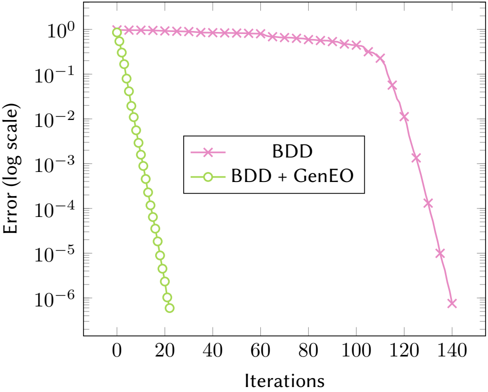
$N = 81$ subdomains, $\nu = 0.4$, $E_1 = 10^7$ and $E_2 = 10^{12}$, $\tau = 0.1$

Young's modulus E



Metis Partition





Size of the coarse space: $n_0 = 349$ including 212 rigid body modes.

N. S. and D. J. Rixen.
Int. J. Numer. Meth. Engng., 2013.

N. S., V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl.
Numer. Math., 2014.

Multipreconditioned BDD

What is multipreconditioning ? $(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$

PCG : Minimize $\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_{\mathbf{A}}$ over $\mathbf{x}_i + \widetilde{\text{span}}(\mathbf{H}\mathbf{r}_i)$.

An enlarged minimization space implies better convergence:

- deflation (*e.g.* with a spectral coarse space),
- block methods,

What is multipreconditioning ? $(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$

PCG : Minimize $\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_{\mathbf{A}}$ over $\mathbf{x}_i + \widetilde{\text{span}}(\mathbf{H}\mathbf{r}_i)$.

An enlarged minimization space implies better convergence:

- deflation (*e.g.* with a spectral coarse space),
- block methods,
- or **multipreconditioning !**

MPCG: Minimize $\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_{\mathbf{A}}$ over $\mathbf{x}_i + \widetilde{\text{span}}(\mathbf{H}^1\mathbf{r}_i, \dots, \mathbf{H}^N\mathbf{r}_i)$ where $\mathbf{H}^1, \dots, \mathbf{H}^N$ are N preconditioners.

R. Bridson and C. Greif.
A multipreconditioned conjugate gradient algorithm. *SIAM J. Matrix Anal. Appl.*, 2006.

Very natural for DD: one preconditioner per subdomain

$$\mathbf{H} = \sum_{s=1}^{N} \underbrace{\mathbf{R}^{s\top}\mathbf{D}^s\mathbf{S}^{s-1}\mathbf{D}^s\mathbf{R}^s}_{:=\mathbf{H}^s}.$$

What is multipreconditioning ? $(\mathbf{H})\mathbf{A}\mathbf{x}_* = (\mathbf{H})\mathbf{b}$

PCG : Minimize $\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_{\mathbf{A}}$ over $\mathbf{x}_i + \widetilde{\text{span}}(\mathbf{H}\mathbf{r}_i)$.

An enlarged minimization space implies better convergence:

- deflation (*e.g.* with a spectral coarse space),
- block methods,
- or **multipreconditioning !**

MPCG: Minimize $\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_{\mathbf{A}}$ over $\mathbf{x}_i + \widetilde{\text{span}}(\mathbf{H}^1\mathbf{r}_i, \dots, \mathbf{H}^N\mathbf{r}_i)$ where $\mathbf{H}^1, \dots, \mathbf{H}^N$ are N preconditioners.

D. J. Rixen.
Substructuring and Dual Methods in Structural Analysis. PhD thesis, 1997.

R. Bridson and C. Greif.
A multipreconditioned conjugate gradient algorithm. *SIAM J. Matrix Anal. Appl.*, 2006.

C. Greif, T. Rees, and D. Szyld.
MPGMRES: a generalized minimum residual method with multiple preconditioners. *Technical report*, 2011 (now published in SeMA Journal).

C. Greif, T. Rees, and D. Szyld.
Additive Schwarz with variable weights. *DD21 proceedings*, 2014.

P. Gosselet, D. J. Rixen, F.-X. Roux, and N. S.
Simultaneous FETI and block FETI… *International Journal for Numerical Methods in Engineering*, 2015.

Multipreconditioned CG for $\mathbf{A}\mathbf{x}_* = \mathbf{b}$ prec. by $\{\mathbf{H}^s\}_{s=1,\dots,N}$

- $\mathbf{A}$, $\mathbf{H} \in \mathbb{R}^{n\times n}$ spd,
- $\mathbf{H} = \sum_{s=1}^N \mathbf{H}^s$, where $\mathbf{H}^s$ spsd.

MPCG		PCG
$\mathbf{x}_0$ given ;	← *Initial Guess*	$\mathbf{x}_0$ given ;
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;		$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{Z}_0 = [\mathbf{H}^1\mathbf{r}_0 \mid \dots \mid \mathbf{H}^N\mathbf{r}_0]$;		$\mathbf{z}_0 = \mathbf{H}\mathbf{r}_0$;
$\mathbf{P}_0 = \mathbf{Z}_0$;	← *Initial search directions*	$\mathbf{p}_0 = \mathbf{z}_0$;
for $i = 0, 1, \dots,$ *convergence* **do**		**for** $i = 0, 1, \dots,$ *conv.* **do**
$\quad\mathbf{Q}_i = \mathbf{A}\mathbf{P}_i$;		$\quad\mathbf{q}_i = \mathbf{A}\mathbf{p}_i$;
$\quad\boldsymbol{\alpha}_i = (\mathbf{Q}_i^\top \mathbf{P}_i)^\dagger (\mathbf{P}_i{}^\top \mathbf{r}_i)$;		$\quad\alpha_i = (\mathbf{q}_i^\top \mathbf{p}_i)^{-1}(\mathbf{p}_i^\top \mathbf{r}_i)$;
$\quad\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{P}_i\boldsymbol{\alpha}_i$;	← *Update approximate solution*	$\quad\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i\mathbf{p}_i$;
$\quad\mathbf{r}_{i+1} = \mathbf{r}_i - \mathbf{Q}_i\boldsymbol{\alpha}_i$;	← *Update residual*	$\quad\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i\mathbf{q}_i$;
$\quad\mathbf{Z}_{i+1} = [\mathbf{H}^1\mathbf{r}_{i+1} \mid \dots \mid \mathbf{H}^N\mathbf{r}_{i+1}]$;	← *Multiprecondition*	$\quad\mathbf{z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$;
$\quad\boldsymbol{\beta}_{i,j} = (\mathbf{Q}_j^\top \mathbf{P}_j)^\dagger(\mathbf{Q}_j^\top \mathbf{Z}_{i+1}), \quad j = 0, \dots, i$;		$\quad\beta_i = (\mathbf{q}_i^\top \mathbf{p}_i)^{-1}(\mathbf{q}_i^\top \mathbf{z}_{i+1})$;
$\quad\mathbf{P}_{i+1} = \mathbf{Z}_{i+1} - \sum_{j=0}^i \mathbf{P}_j\boldsymbol{\beta}_{i,j}$;	← *Project and orthog.*	$\quad\mathbf{p}_{i+1} = \mathbf{z}_{i+1} - \beta_i\mathbf{p}_i$;
end		**end**
Return $\mathbf{x}_{i+1}$;		Return $\mathbf{x}_{i+1}$;

Multipreconditioned CG for $\mathbf{A}\mathbf{x}_* = \mathbf{b}$ prec. by $\{\mathbf{H}^s\}_{s=1,\dots,N}$

- $\mathbf{A}, \mathbf{H} \in \mathbb{R}^{n\times n}$ spd,
- $\mathbf{H} = \sum_{s=1}^{N} \mathbf{H}^s$, where $\mathbf{H}^s$ spsd.

MPCG

$\mathbf{x}_0$ given ; ⇐ *Initial Guess*
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$;
$\mathbf{Z}_0 = \left[\mathbf{H}^1\mathbf{r}_0 \mid \dots \mid \mathbf{H}^N\mathbf{r}_{i+1}\right]$; ⇐ *Multiprecondition*
$\mathbf{P}_0 = \mathbf{Z}_0$; ⇐ *Initial search directions*
for $i = 0, 1, \dots,$ *convergence* **do**
    $\mathbf{Q}_i = \mathbf{A}\mathbf{P}_i$;
    $\boldsymbol{\alpha}_i = (\mathbf{Q}_i^\top \mathbf{P}_i)^\dagger\, (\mathbf{P}_i{}^\top \mathbf{r}_i)$;
    $\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{P}_i\boldsymbol{\alpha}_i$; ⇐ *Update approximate solution*
    $\mathbf{r}_{i+1} = \mathbf{r}_i - \mathbf{Q}_i\boldsymbol{\alpha}_i$; ⇐ *Update residual*
    $\mathbf{Z}_{i+1} = \left[\mathbf{H}^1\mathbf{r}_{i+1} \mid \dots \mid \mathbf{H}^N\mathbf{r}_{i+1}\right]$; ⇐ *Multiprecondition*
    $\boldsymbol{\beta}_{i,j} = (\mathbf{Q}_j^\top \mathbf{P}_j)^\dagger (\mathbf{Q}_j^\top \mathbf{Z}_{i+1})$, $\quad j = 0, \dots, i$;
    $\mathbf{P}_{i+1} = \mathbf{Z}_{i+1} - \sum_{j=0}^{i} \mathbf{P}_j \boldsymbol{\beta}_{i,j}$; ⇐ *Orthogonalize*
end
Return $\mathbf{x}_{i+1}$;

Remark

$$\mathbf{x}_i, \mathbf{r}_i \in \mathbb{R}^n.$$

$$\mathbf{Z}_i,\ \mathbf{P}_i, \in \mathbb{R}^{n\times N}$$

- n: size of problem,
- N: nb of precs.

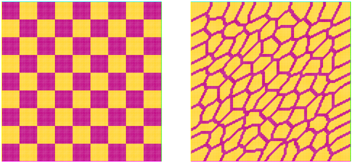
Properties

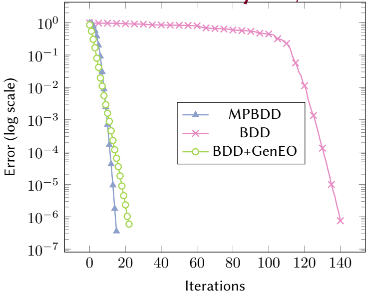
1. Global minimization over

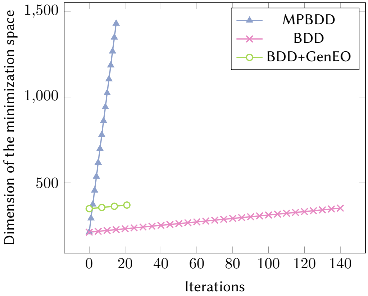
$$\mathbf{x}_0 + \bigoplus_{j=0}^{i-1} \operatorname{range}(\mathbf{P}_j)$$

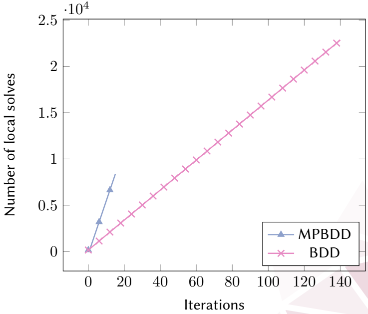
   of dimension: $i \times N$.
2. $\mathbf{P}_i^\top \mathbf{A}\mathbf{P}_j = 0$ $(i \neq j)$, but **no** short recurrence.

Numerical Illustration (Heterogeneous Elasticity $E_2/E_1 = 10^5$)



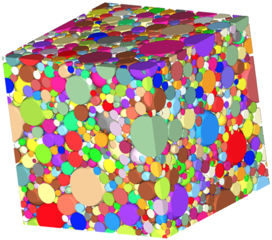


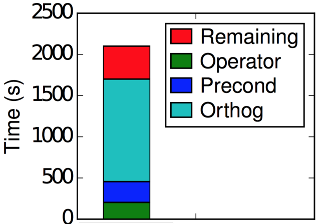


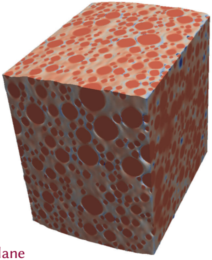


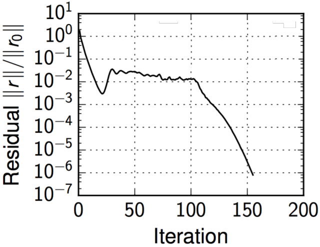
Solid propellant (linear elasticity with FETI) – in collaboration with C. Bovet (ONERA)

6721 stiff inclusions, $E_2/E_1 \leqslant 10^6$, 57 Mdofs, 360 subdomains on 1440 cores









Good convergence but a possible limitation

- ✓ Local contributions $\mathbf{H}^s \mathbf{r}_i$ form a good minimization space.
- ✗ Not adaptive: invert $\mathbf{P}_i^\top \mathbf{A} \mathbf{P}_i \in \mathbb{R}^{N \times N}$ at **each** iteration in $\boldsymbol{\alpha}_i = (\mathbf{P}_i^\top \mathbf{A} \mathbf{P}_i)^\dagger (\mathbf{P}_i{}^\top \mathbf{r}_i)$.

Introduce adaptativity into multipreconditioned CG.

- a few local vectors should suffice to accelerate convergence (previous work on GenEO).

Adaptive Multipreconditioned BDD

Adaptive Multipreconditioned CG for $\mathbf{A}\mathbf{x}_* = \mathbf{b}$ preconditioned by $\sum_{s=1}^{N} \mathbf{H}^s$.

($\tau \in \mathbb{R}^+$ is chosen by the user – *e.g.*, $\tau = 0.1$)

$\mathbf{x}_0$ chosen;
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$; $\mathbf{Z}_0 = \mathbf{H}\mathbf{r}_0$; $\mathbf{P}_0 = \mathbf{Z}_0$;
for $i = 0, 1, \dots,$ *convergence* **do**
  $\mathbf{Q}_i = \mathbf{A}\mathbf{P}_i$;
  $\boldsymbol{\alpha}_i = (\mathbf{Q}_i^\top \mathbf{P}_i)^\dagger (\mathbf{P}_i^\top \mathbf{r}_i)$;
  $\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{P}_i \boldsymbol{\alpha}_i$;
  $\mathbf{r}_{i+1} = \mathbf{r}_i - \mathbf{Q}_i \boldsymbol{\alpha}_i$;
  $t_i = \dfrac{(\mathbf{P}_i\boldsymbol{\alpha}_i)^\top \mathbf{A} (\mathbf{P}_i\boldsymbol{\alpha}_i)}{\mathbf{r}_{i+1}^\top \mathbf{H} \mathbf{r}_{i+1}}$;
  if $t_i < \tau$ **then** ← *τ-test*
    $\mathbf{Z}_{i+1} = \left[\mathbf{H}^1 \mathbf{r}_{i+1} \mid \dots \mid \mathbf{H}^N \mathbf{r}_{i+1}\right]$; ← *Multiprec.*
  else
    $\mathbf{Z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$; ← *Precondition*
  end
  $\boldsymbol{\beta}_{i,j} = (\mathbf{Q}_j^\top \mathbf{P}_j)^\dagger (\mathbf{Q}_j^\top \mathbf{Z}_{i+1})$, $\quad j = 0, \dots, i$;
  $\mathbf{P}_{i+1} = \mathbf{Z}_{i+1} - \sum_{j=0}^{i} \mathbf{P}_j \boldsymbol{\beta}_{i,j}$;
end
Return $\mathbf{x}_{i+1}$;

Remark

$\mathbf{x}_i, \mathbf{r}_i \in \mathbb{R}^n$.

$\mathbf{Z}_i, \mathbf{P}_i \in \mathbb{R}^{n \times N}$ or $\mathbb{R}^n$,

- n: size of problem,
- N: nb of precs.

Properties

1. Global minimization over
$$\mathbf{x}_0 + \bigoplus_{j=0}^{i-1} \text{range}(\mathbf{P}_j)$$
of dim $i \leqslant dim \leqslant i \times N$.
2. $\mathbf{P}_i^\top \mathbf{A} \mathbf{P}_j = 0$ $(i \neq j)$, but **no** short recurrence.

Theoretical Results (1/2): Choice of the τ-test ?

Theorem

If the τ−test returns $t_i \geqslant \tau$ then

$$\frac{\|\mathbf{x}_* - \mathbf{x}_{i+1}\|_\mathbf{A}}{\|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}} \leqslant \left(\frac{1}{1+\tau}\right)^{1/2}.$$

$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{P}_i\boldsymbol{\alpha}_i$;

$t_i = \dfrac{(\mathbf{P}_i\boldsymbol{\alpha}_i)^\top \mathbf{A}(\mathbf{P}_i\boldsymbol{\alpha}_i)}{\mathbf{r}_{i+1}^\top \mathbf{H}\mathbf{r}_{i+1}}$;

if $t_i < \tau$ **then**

$\mathbf{Z}_{i+1} = \left[\mathbf{H}^1\mathbf{r}_{i+1} \mid \dots \mid \mathbf{H}^N\mathbf{r}_{i+1}\right]$

else

$\mathbf{Z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$;

end

Proof (inspired by [Axelsson, Kaporin, '01]):

$$\mathbf{x}_* = \mathbf{x}_0 + \sum_{i=0}^{n} \mathbf{P}_i\boldsymbol{\alpha}_i = \mathbf{x}_i + \sum_{j=i}^{n} \mathbf{P}_j\boldsymbol{\alpha}_j$$

$$\Rightarrow \|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}^2 = \sum_{j=i}^{n} \|\mathbf{P}_j\boldsymbol{\alpha}_j\|_\mathbf{A}^2 \Rightarrow \|\mathbf{x}_* - \mathbf{x}_{i-1}\|_\mathbf{A}^2 = \|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}^2 + \|\mathbf{P}_{i-1}\boldsymbol{\alpha}_{i-1}\|_\mathbf{A}^2.$$

$$\Rightarrow \frac{\|\mathbf{x}_* - \mathbf{x}_{i-1}\|_\mathbf{A}^2}{\|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}^2} = 1 + \underbrace{\frac{\|\mathbf{r}_i\|_\mathbf{H}^2}{\|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}^2}}_{= \|\mathbf{x}_* - \mathbf{x}_i\|_{\mathbf{AHA}}^2 / \|\mathbf{x}_* - \mathbf{x}_i\|_\mathbf{A}^2 \geqslant \lambda_{min} \geqslant 1} \frac{\|\mathbf{P}_{i-1}\boldsymbol{\alpha}_{i-1}\|_\mathbf{A}^2}{\|\mathbf{r}_i\|_\mathbf{H}^2} \geqslant 1 + \underbrace{\frac{\|\mathbf{P}_{i-1}\boldsymbol{\alpha}_{i-1}\|_\mathbf{A}^2}{\|\mathbf{r}_i\|_\mathbf{H}^2}}_{= t_{i-1}} \geqslant 1 + \tau.$$

Theoretical Results (2/2): No extra work for well conditioned problems

Theorem

$t_i \geqslant 1/\lambda_{max}$ so if $\tau < 1/\lambda_{max}$ then AMPCG is PCG.

Local Adaptive MPCG for $\sum_{s=1}^N \mathbf{A}^s \mathbf{x}_* = \mathbf{b}$ preconditioned by $\sum_{s=1}^N \mathbf{H}^s$.

($\tau \in \mathbb{R}^+$ is chosen by the user)

$\mathbf{x}_0$ chosen; $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$; $\mathbf{Z}_0 = \mathbf{H}\mathbf{r}_0$; $\mathbf{P}_0 = \mathbf{Z}_0$;
for $i = 0, 1, \dots,$ *convergence* **do**
  $\mathbf{Q}_i = \mathbf{A}\mathbf{P}_i$;
  $\boldsymbol{\alpha}_i = (\mathbf{Q}_i^\top \mathbf{P}_i)^\dagger (\mathbf{P}_i{}^\top \mathbf{r}_i)$;
  $\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{P}_i \boldsymbol{\alpha}_i$;
  $\mathbf{r}_{i+1} = \mathbf{r}_i - \mathbf{Q}_i \boldsymbol{\alpha}_i$;
  $\mathbf{Z}_{i+1} = \mathbf{H}\mathbf{r}_{i+1}$; ⇽ *initialize* $\mathbf{Z}_{i+1}$
  for $s = 1, \dots, N$ **do**
    $t_i^s = \dfrac{\langle \mathbf{P}_i \boldsymbol{\alpha}_i, \mathbf{A}^s \mathbf{P}_i \boldsymbol{\alpha}_i \rangle}{\mathbf{r}_{i+1}^\top \mathbf{H}^s \mathbf{r}_{i+1}}$;
    if $t_i^s < \tau$ **then** ⇽ *local* τ*-test*
      $\mathbf{Z}_{i+1} = \left[\mathbf{Z}_{i+1} \,|\, \mathbf{H}^s \mathbf{r}_{i+1}\right]$;
    end
  end
  $\boldsymbol{\beta}_{i,j} = (\mathbf{Q}_j^\top \mathbf{P}_j)^\dagger (\mathbf{Q}_j^\top \mathbf{Z}_{i+1})$, $\quad j = 0, \dots, i$;
  $\mathbf{P}_{i+1} = \mathbf{Z}_{i+1} - \sum_{j=0}^{i} \mathbf{P}_j \boldsymbol{\beta}_{i,j}$;
end
Return $\mathbf{x}_{i+1}$;

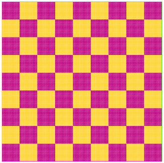
An additional assumption

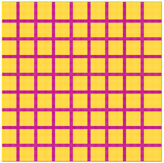
$$\mathbf{A} = \sum_{s=1}^{N} \underbrace{\mathbf{R}^{s\top} \mathbf{S}^s \mathbf{R}^s}_{:=\mathbf{A}^s}.$$

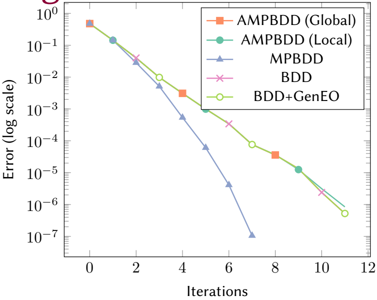
Benefits

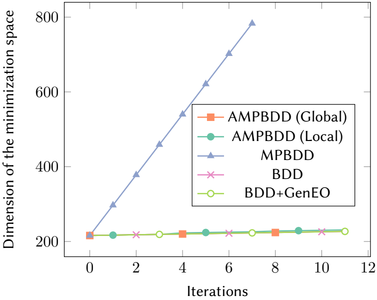
- Between 1 and N search directions per iteration.
- Reduces the cost of storage and of inverting $\mathbf{P}_i^\top \mathbf{A}\mathbf{P}_i$.

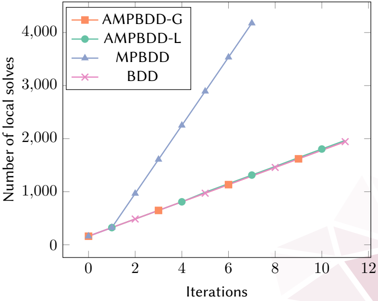
Numerical Illustration (0/6): Homogenous subdomains



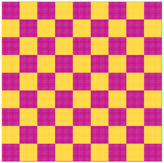


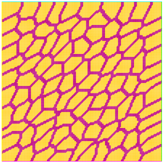


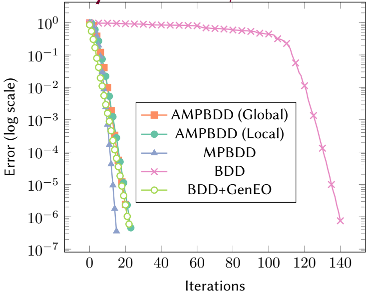


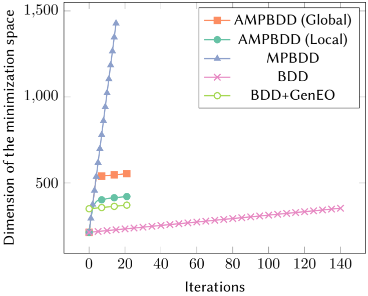


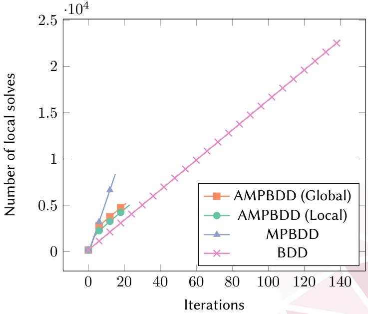
Numerical Illustration (1/6) – Elasticity with $E_2/E_1 = 10^5$



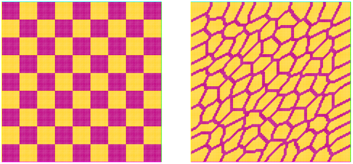


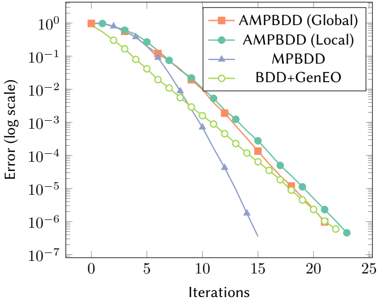


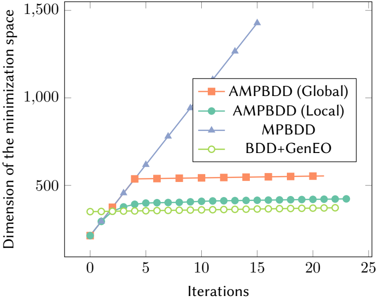


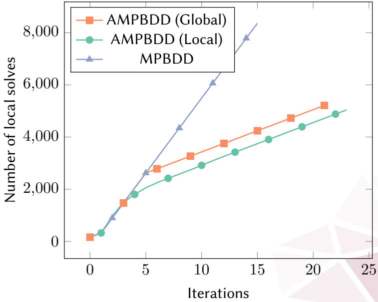


Numerical Illustration (2/6): Zoom on new methods

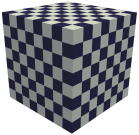


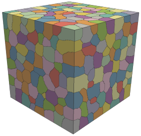




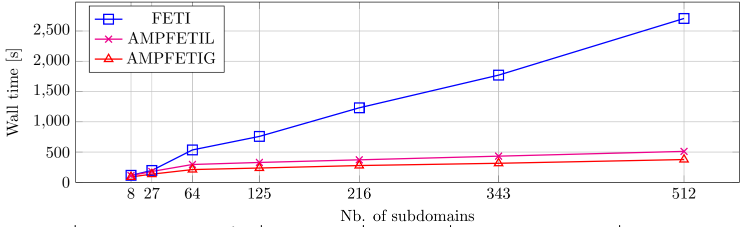


Numerical Illustration (3/6): Weak Scalability – FETI





- Software: Z-Set
- Cluster: Cobalt at CCRT/TGCC
- 1422 computational nodes with Intel Broadwell Processors: 2.4 GHz, 28 cores
- 128 Go SDRAM, infiniband Mellanox network
- 7 cores per subdomain and local factorizations are performed with mumps.

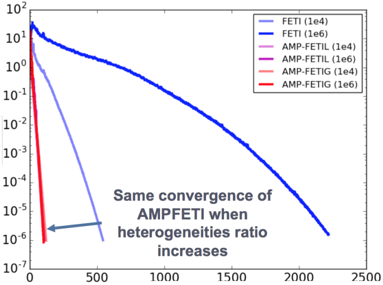


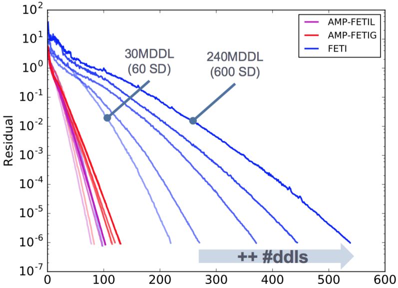
N	#DOFs ($\times 10^6$)	#cores	#iter.	# min. space	time (s)
8	1.6	56	69	132	89.58
64	12.5	448	112	742	210.80
216	42.0	1512	107	2257	277.30
512	99.2	3584	108	5729	376.30

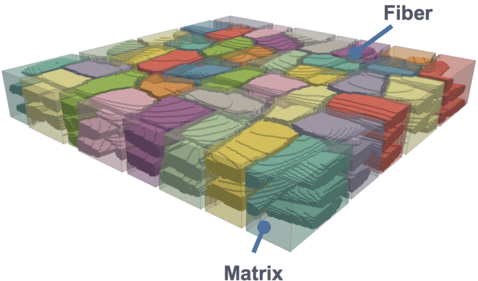
C. Bovet, P. Gosselet, A. Parret Freaud, and N. S.
Adaptive multi preconditioned FETI: scalability results and robustness assessment.
Computers and Structures, 2017.

Numerical Illustration (4/6): Composite Weave Pattern

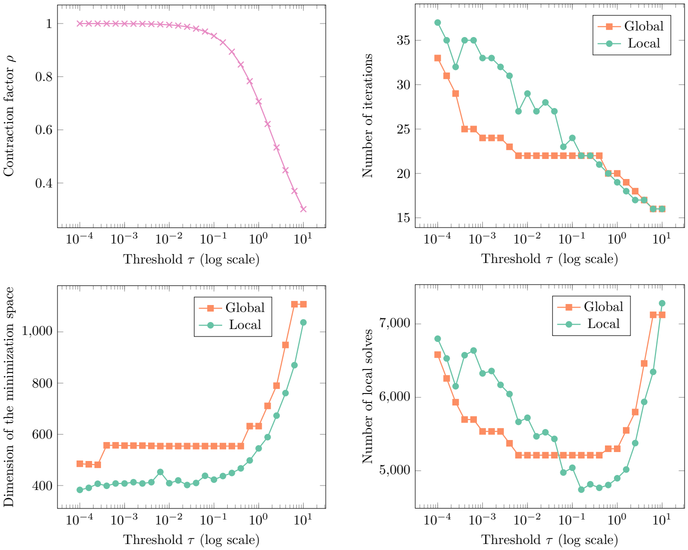
FETI – in collaboration with A. Parret Freaud (Safran Tech)







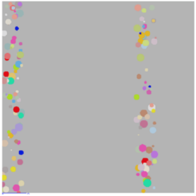
Numerical Illustration (5/6): Choice of τ

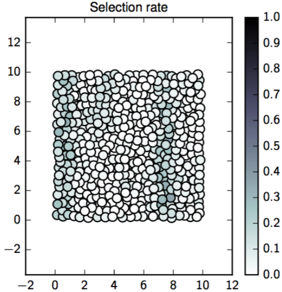


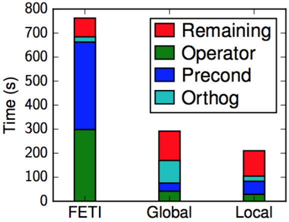
Numerical Illustration (6/6): 400 stiff inclusions – FETI

in collaboration with C. Bovet (ONERA)

$\tau = 0.01$, $E_{max}/E_{min} = 10^5$, 480 subdomains, 32 Mdofs







Conclusion

AMPDD → Robustness and Efficiency.

Perspectives

- Best adaptation process in non symmetric cases ? Best adaptation process in a fully algebraic context ?
- PETSc4py demonstration code (in collaboration with Loïc Gouarin: `github.com/gouarin/GenEO`
- Restart ! Recycle ! within AMPDD Krylov subspace solvers.

N. S.
An Adaptive Multipreconditioned Conjugate Gradient
SISC, 2016.

N. S.
Algebraic Adaptive Multi Preconditioning applied to Restricted Additive Schwarz.
DD23 Proceedings, 2016.

C. Bovet, P. Gosselet, A. Parret Freaud, and N. S.
Adaptive multi preconditioned FETI: scalability results and robustness assessment.
Computers and Structures, 2017.

C. Bovet, P. Gosselet, and N. S.
Multipreconditioning for nonsymmetric problems: the case of orthomin and biCG.
Note au CRAS, 2017.