

Algebraic multilevel preconditioning and aggregation

Yvan Notay*

Université Libre de Bruxelles

Service de Métrologie Nucléaire



September 10, 2004, Bulgaria

* Supported by the “Fonds National de la Recherche Scientifique”, Maître de recherches

<http://homepages.ulb.ac.be/~ynotay>

Introduction

Algebraic multilevel preconditioning:

- **multilevel ILU**: combines ideas from ILU and (algebraic) multigrid;

Introduction

Algebraic multilevel preconditioning:

- **multilevel ILU**: combines ideas from ILU and (algebraic) multigrid;
- well established for regular grids & geometric coarsening;

Introduction

Algebraic multilevel preconditioning:

- **multilevel ILU**: combines ideas from ILU and (algebraic) multigrid;
- well established for regular grids & geometric coarsening;
- **here: towards a truly algebraic method;**

Introduction

Algebraic multilevel preconditioning:

- **multilevel ILU**: combines ideas from ILU and (algebraic) multigrid;
- well established for regular grids & geometric coarsening;
- **here: towards a truly algebraic method**;
- **using coarsening by aggregation** (control of setup cost & memory requirements).

Basic principle (two-level scheme) (1)

System to solve:

$$A \mathbf{u} = \mathbf{b}$$

(1) Select a F/C partitioning of the unknowns;
partition A accordingly:

$$A = \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix}$$

Basic principle (two-level scheme) (2)

(2) The exact block factorization:

$$A = \begin{pmatrix} A_{FF} & \\ A_{CF} & S_A \end{pmatrix} \begin{pmatrix} I & A_{FF}^{-1} A_{FC} \\ & I \end{pmatrix}$$

where

$$S_A = A_{CC} - A_{CF} A_{FF}^{-1} A_{FC} \quad (\text{Schur complement})$$

is approximated by

$$B = \begin{pmatrix} P_{FF} & \\ A_{CF} & S \end{pmatrix} \begin{pmatrix} I & P_{FF}^{-1} A_{FC} \\ & I \end{pmatrix} .$$

Basic principle (two-level scheme) (2)

(2) The exact block factorization:

$$A = \begin{pmatrix} A_{FF} & \\ A_{CF} & S_A \end{pmatrix} \begin{pmatrix} I & A_{FF}^{-1} A_{FC} \\ & I \end{pmatrix}$$

where

$$S_A = A_{CC} - A_{CF} A_{FF}^{-1} A_{FC} \quad (\text{Schur complement})$$

is approximated by

$$B = \begin{pmatrix} P_{FF} & \\ A_{CF} & S \end{pmatrix} \begin{pmatrix} I & P_{FF}^{-1} A_{FC} \\ & I \end{pmatrix} .$$

Needed ingredients: $P_{FF} \approx A_{FF}$, $S \approx S_A$

Coarsening by aggregation (1)

Main idea:

- group together nodes having strong negative connections between them (partitioning of $[1, n]$);
- in each group select one node i to be a C node; the other ones are F nodes;
- noting G_i the group of nodes from which the C node i has been extracted (set of nodes aggregated with i), the coarse grid matrix (**aggregated matrix**) is obtained by summing entries:

$$S_{ij} = \sum_{i \in G_i} \sum_{j \in G_j} a_{ij} \ .$$

Tends to reproduce the stencil from the fine grid.

Coarsening by aggregation (2)

$$S = \begin{pmatrix} J_{FC}^T & I \end{pmatrix} \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} J_{FC} \\ I \end{pmatrix}$$

where

$$\forall i \in F, j \in C : (J_{FC})_{ij} = \begin{cases} 1 & \text{if } i \in G_j \\ 0 & \text{otherwise} \end{cases} .$$

Coarsening by aggregation (2)

$$S = \begin{pmatrix} J_{FC}^T & I \end{pmatrix} \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} J_{FC} \\ I \end{pmatrix}$$

where

$$\forall i \in F, j \in C : (J_{FC})_{ij} = \begin{cases} 1 & \text{if } i \in G_j \\ 0 & \text{otherwise} \end{cases} .$$

J_{FC} corresponds to a cheap but crude interpolation: each F node takes its value from exactly one C node.

Coarsening by aggregation (2)

$$S = \begin{pmatrix} J_{FC}^T & I \end{pmatrix} \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} J_{FC} \\ I \end{pmatrix}$$

where

$$\forall i \in F, j \in C : (J_{FC})_{ij} = \begin{cases} 1 & \text{if } i \in G_j \\ 0 & \text{otherwise} \end{cases} .$$

J_{FC} corresponds to a cheap but crude interpolation: each F node takes its value from exactly one C node.

Bad interpolation: does not work well with multigrid.

Coarsening by aggregation (2)

$$S = \begin{pmatrix} J_{FC}^T & I \end{pmatrix} \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} J_{FC} \\ I \end{pmatrix}$$

where

$$\forall i \in F, j \in C : (J_{FC})_{ij} = \begin{cases} 1 & \text{if } i \in G_j \\ 0 & \text{otherwise} \end{cases} .$$

J_{FC} corresponds to a cheap but crude interpolation: each F node takes its value from exactly one C node.

Bad interpolation: does not work well with multigrid.

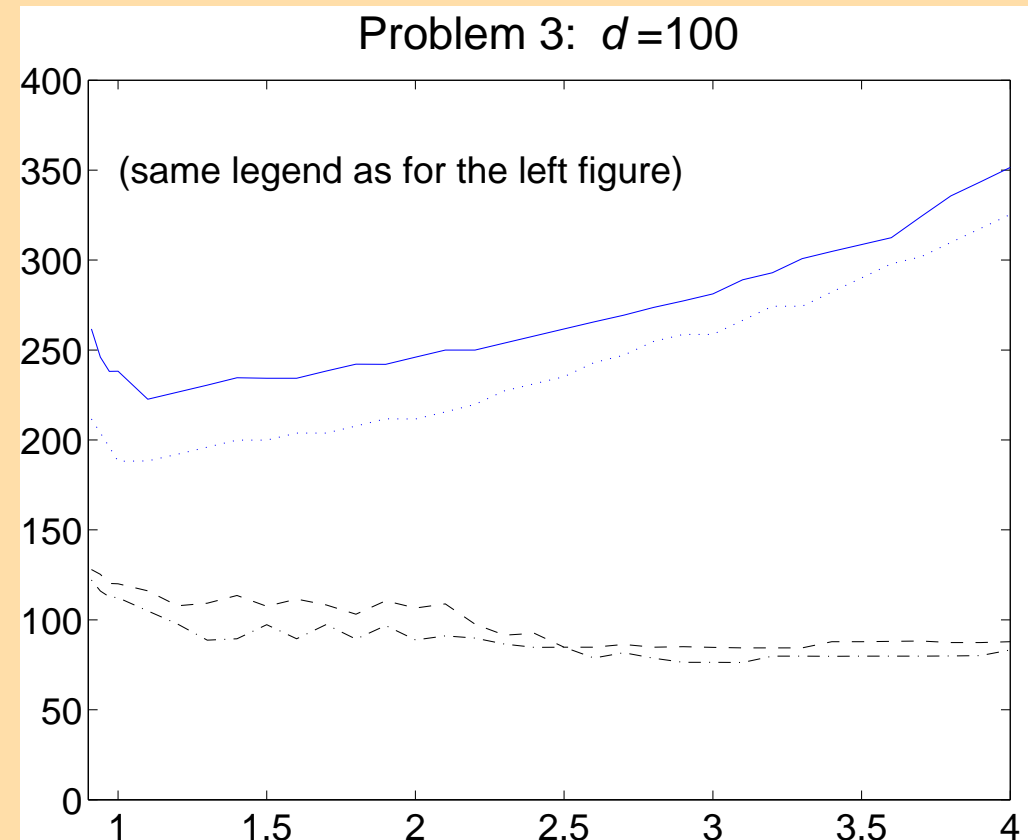
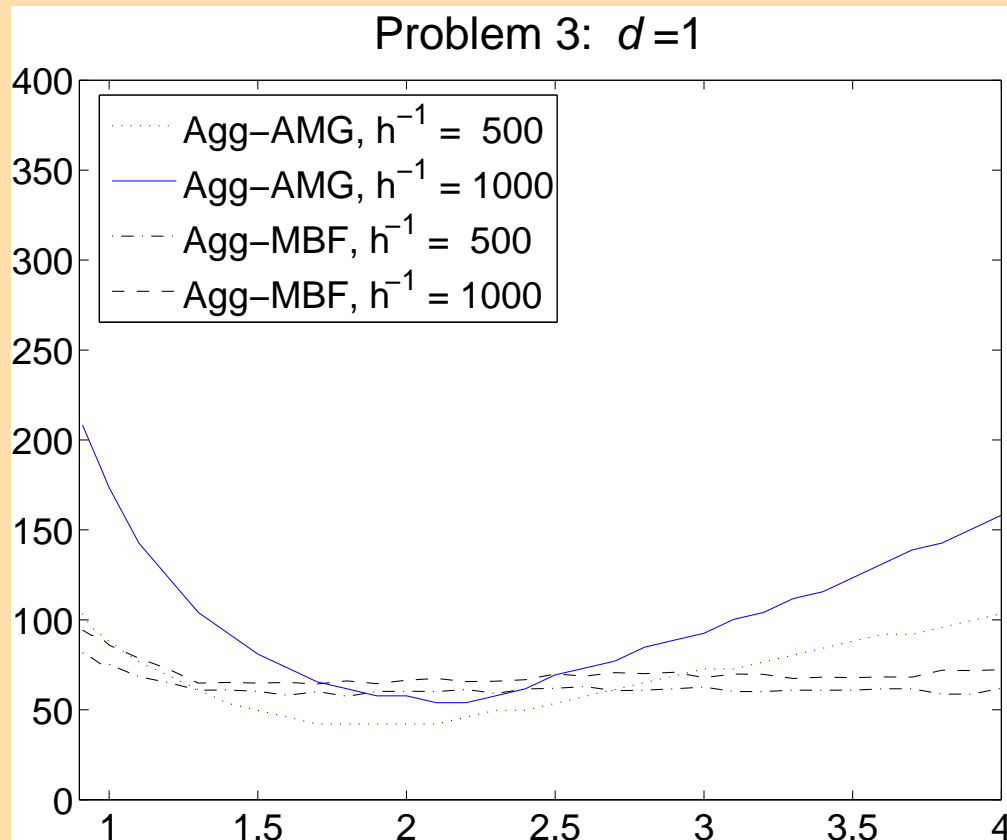
But produces reasonable coarse grid matrices S
(up to a scaling factor).

Coarsening by aggregation (3)

Relative solution cost – vs – Scaling factor

Model problem

Non model problem



Double pairwise aggregation (1)

Algorithm outline:

- group the nodes pairwise following the strongest (negative) coupling – some singletons left;
- form the corresponding aggregated matrix;
- repeat the process: group pairwise the groups (pairs and singletons) obtained from the first pass, following the strongest (negative) coupling in the aggregated matrix;
- in each resulting group (mostly quadruplets), select one node to be a C node; the other ones are F nodes.

Double pairwise aggregation (2)

Features:

- $\frac{n_C}{n} \approx \frac{1}{4}$ independently of the problem

→ control of setup cost & memory requirements

(if semi-coarsening: mesh size multiplied by 4
in the direction of coarsening);

Double pairwise aggregation (2)

Features:

- $\frac{n_C}{n} \approx \frac{1}{4}$ independently of the problem

→ control of setup cost & memory requirements

(if semi-coarsening: mesh size multiplied by 4
in the direction of coarsening);

- not very sensitive to Strong/Weak coupling threshold;

Double pairwise aggregation (2)

Features:

- $\frac{n_C}{n} \approx \frac{1}{4}$ independently of the problem

→ control of setup cost & memory requirements

(if semi-coarsening: mesh size multiplied by 4 in the direction of coarsening);

- not very sensitive to Strong/Weak coupling threshold;
- easy to parallelize.

Approximation P_{FF} to A_{FF} (1)

$$\begin{aligned} B^{-1} &= \begin{pmatrix} I & -P_{FF}^{-1}A_{FC} \\ & I \end{pmatrix} \begin{pmatrix} P_{FF}^{-1} & \\ & S \end{pmatrix} \begin{pmatrix} I & \\ -A_{CF}P_{FF}^{-1} & I \end{pmatrix} \\ &= \begin{pmatrix} I \\ 0 \end{pmatrix} P_{FF}^{-1} \begin{pmatrix} I & 0 \end{pmatrix} + \begin{pmatrix} -P_{FF}^{-1}A_{FC} \\ I \end{pmatrix} S^{-1} \begin{pmatrix} -A_{CF}P_{FF}^{-1} & I \end{pmatrix} \end{aligned}$$

Additive two-level scheme with prolongation

$$\tilde{p} = \begin{pmatrix} -P_{FF}^{-1}A_{FC} \\ I \end{pmatrix}$$

→ P_{FF} such that $-P_{FF}^{-1}A_{FC}$ is a “correct” interpolation.

Approximation P_{FF} to A_{FF} (2)

MILU factorization

Cheap approximation that satisfies

$$P_{FF} \mathbf{e}_F = A_{FF} \mathbf{e}_F ,$$

where $\mathbf{e} = \left(1 \ \dots \ 1 \right)^T$.

Hence, $A_{FF} \mathbf{e}_F + A_{FC} \mathbf{e}_C \approx 0$ entails

$$-P_{FF}^{-1} A_{FC} \mathbf{e}_C \approx \mathbf{e}_F .$$

For scalar PDEs, this is enough to ensure that the interpolation is “correct”.

Approximation P_{FF} to A_{FF} (3)

How to ensure that

- the MILU factorization is well defined (no break down);
- it is an accurate approximation to A_{FF} ?

Approximation P_{FF} to A_{FF} (3)

How to ensure that

- the MILU factorization is well defined (no break down);
- it is an accurate approximation to A_{FF} ?

In many cases, no problem arises because A_{FF} is well conditioned (diagonally dominant).

Approximation P_{FF} to A_{FF} (3)

How to ensure that

- the MILU factorization is well defined (no break down);
- it is an accurate approximation to A_{FF} ?

In many cases, no problem arises because A_{FF} is well conditioned (diagonally dominant).

For better robustness: **dynamic** MILU factorization:
 F nodes for which the pivot would not be at least $\frac{3}{5}$ of the corresponding diagonal element are rejected to the C set.

Guarantees existence **and** good conditioning properties of the MILU factorization.

From two- to multi-level (1)

Preconditioner:

$$B = \begin{pmatrix} P_{FF} & \\ A_{CF} & S \end{pmatrix} \begin{pmatrix} I & P_{FF}^{-1} A_{FC} \\ & I \end{pmatrix} .$$

Requires solving a system with S .

Recursivity: S has a structure similar to that of A

→ the same technique is applied to precondition S .

The system with S is then solved by preconditioned iterations (**flexible** conjugate gradient of **FGMRES**).

From two- to multi-level (2)

Stability: consider

$$B = \begin{pmatrix} P_{FF} & \\ A_{21} & S \end{pmatrix} \begin{pmatrix} I & P_{FF}^{-1}A_{12} \\ & I \end{pmatrix} .$$

The error in the solution to

$$B \mathbf{w} = \mathbf{r} \quad (2)$$

comes from the error made in solving

$$S \mathbf{x}_C = \mathbf{y}_C \quad (\mathbf{y}_C = \mathbf{r}_C - A_{CF}P_{FF}^{-1}\mathbf{r}_F) . \quad (3)$$

From two- to multi-level (2)

Stability: consider

$$B = \begin{pmatrix} P_{FF} & \\ A_{21} & S \end{pmatrix} \begin{pmatrix} I & P_{FF}^{-1}A_{12} \\ & I \end{pmatrix} .$$

The error in the solution to

$$B \mathbf{w} = \mathbf{r} \quad (1)$$

comes from the error made in solving

$$S \mathbf{x}_C = \mathbf{y}_C \quad (\mathbf{y}_C = \mathbf{r}_C - A_{CF}P_{FF}^{-1}\mathbf{r}_F) . \quad (2)$$

Good news:

- the relative error in energy norm in (1) = that in (2);
- the residual in (1) is the one in (2) padded with zeros.

From two- to multi-level (3)

Some technical details:

- stopping criterion for inner iterations: relative residual error ≤ 0.35 ;
- at most a few iterations allowed (to bound the cost of the application of the preconditioner);
- the coarse grid matrix S is factorized exactly as soon as the cost of this factorization represents less than a fraction of an unpreconditioned CG iteration for the whole system.

Numerical results: Problem 1

Model 2D anisotropic problem, 5 point FD

$$-a_x \frac{\partial^2 u}{\partial x^2} - a_y \frac{\partial^2 u}{\partial y^2} = 1 \quad \text{in } \Omega = (0, 1) \times (0, 1)$$

$$u = 0 \quad \text{on } x = 1, 0 \leq y \leq 1$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{elsewhere on } \partial\Omega$$

Numerical results: Problem 1

Model 2D anisotropic problem, 5 point FD

$$-a_x \frac{\partial^2 u}{\partial x^2} - a_y \frac{\partial^2 u}{\partial y^2} = 1 \quad \text{in } \Omega = (0, 1) \times (0, 1)$$

$$u = 0 \quad \text{on } x = 1, 0 \leq y \leq 1$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{elsewhere on } \partial\Omega$$

Geometric multigrid:

V cycle, damped Jacobi smoothing ($\omega = 0.5$),

1 pre- and 1 post-smoothing step, preconditioner for CG.

Numerical results: Problem 1

Model 2D anisotropic problem, 5 point FD

$$-a_x \frac{\partial^2 u}{\partial x^2} - a_y \frac{\partial^2 u}{\partial y^2} = 1 \quad \text{in } \Omega = (0, 1) \times (0, 1)$$

$$u = 0 \quad \text{on } x = 1, 0 \leq y \leq 1$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{elsewhere on } \partial\Omega$$

Geometric multigrid:

V cycle, damped Jacobi smoothing ($\omega = 0.5$),
1 pre- and 1 post-smoothing step, preconditioner for CG.

sol. : $\frac{\text{Solution cost}}{\text{Cost of 1 unprec. CG iter.}}$

Numerical results: Problem 1

$h^{-1} = 600$

$h^{-1} = 1200$

a_x	a_y	$\frac{n}{n_c}$	inner	iter.	sol.	$\frac{n}{n_c}$	inner	iter.	sol.
Geometric multigrid									
1	1	3.99	1.00	9	28.3	4.00	1.00	9	28.5
Aggregation-based Algebraic multilevel									
1	1	3.99	1.94	18	63.3	4.00	2.00	19	70.3
1	2	3.98	1.95	19	63.0	3.99	1.95	20	68.4
1	4	3.97	2.10	20	70.6	3.98	2.10	21	76.2
1	10	3.95	2.05	22	72.1	3.98	2.10	21	70.7
1	10^2	3.95	2.00	22	69.0	3.98	2.06	18	58.9
1	10^4	3.95	1.94	18	54.9	3.98	1.89	18	55.4

Numerical results: Problem 2

Model 2D anisotropic problem, bilinear finite elements

$$-a_x \frac{\partial^2 u}{\partial x^2} - a_y \frac{\partial^2 u}{\partial y^2} = 1 \quad \text{in } \Omega = (0, 1) \times (0, 1)$$

$$u = 0 \quad \text{everywhere on } \partial\Omega$$

Stencil:

$a_x = a_y = 1$	$a_x = 1, a_y = 2$	$a_x = 1, a_y = 100$
-2 -2 -2	-3 -6 -3	-101 -398 -101
-2 16 -2	24	+196 808 +196
-2 -2 -2	-3 -6 -3	-101 -398 -101

Numerical results: Problem 2

$h^{-1} = 600$

$h^{-1} = 1200$

a_x	a_y	$\frac{n}{n_c}$	inner	iter.	sol.	$\frac{n}{n_c}$	inner	iter.	sol.
1	1	3.99	1.94	17	55.2	3.99	1.94	17	56.4
1	2	3.99	2.20	15	54.0	4.00	2.20	15	55.6
1	4	3.99	2.11	19	68.9	4.00	2.05	20	72.2
1	10	3.99	2.00	20	68.5	4.00	2.05	20	74.2
1	10^2	3.99	1.95	20	69.1	4.00	1.75	24	80.0
1	10^4	3.99	1.75	20	65.5	4.00	1.83	18	60.3

Numerical results: Problem 5

3D problem with discontinuity, 7 point FD

$$-a_x \frac{\partial^2 u}{\partial x^2} - a_y \frac{\partial^2 u}{\partial y^2} - a_z \frac{\partial^2 u}{\partial z^2} = f \quad \text{in } \Omega = (0, 1) \times (0, 1) \times (0, 1)$$

$$\begin{aligned} u &= 0 && \text{on } z = 1, 0 \leq x, y \leq 1 \\ \frac{\partial u}{\partial n} &= 0 && \text{elsewhere on } \partial\Omega \end{aligned}$$

$$a_x = a_y = a_z = \begin{cases} d & \text{in } \left(\frac{1}{4}, \frac{3}{4}\right) \times \left(\frac{1}{4}, \frac{3}{4}\right) \times \left(\frac{1}{4}, \frac{3}{4}\right) \\ 1 & \text{elsewhere} \end{cases}$$

$$d : 1 \rightarrow 10^6$$

Numerical results: Problem 5

d	101 × 101 × 101 grid				201 × 201 × 201 grid			
	$\frac{n}{n_c}$	inner	iter.	sol.	$\frac{n}{n_c}$	inner	iter.	sol.
Uniform mesh								
1	4.00	2.05	19	85.0	4.00	2.10	20	94.4
10	4.00	2.05	21	97.1	4.00	2.05	21	104.5
10 ²	4.00	2.05	22	104.8	4.00	2.05	22	108.3
10 ⁶	4.00	2.09	22	106.5	4.00	2.04	23	113.7
Non uniform mesh								
1	3.93	2.05	22	100.3	3.97	2.04	24	105.2
10	3.89	2.05	22	103.8	3.95	2.04	24	111.7
10 ²	3.89	2.04	23	106.0	3.95	2.04	24	110.4
10 ⁶	3.89	2.04	23	108.3	3.95	2.13	24	112.9

Numerical results: Problem 7

3D convection diffusion problem, 7 point FD (upwind)

$$-\nu \Delta u + \bar{v} \bar{\nabla} u = 0 \quad \text{in } \Omega = (0, 1) \times (0, 1) \times (0, 1)$$

$$\begin{cases} u = 1 & \text{on } z = 1, 0 \leq x, y \leq 1 \\ u = 0 & \text{elsewhere on } \partial\Omega \end{cases}$$

$$\bar{v}(x, y, z) = \begin{pmatrix} 2x(1-x)(2y-1)z \\ -(2x-1)y(1-y) \\ -(2x-1)(2y-1)z(1-z) \end{pmatrix} ;$$

$\nu = \infty \rightarrow$ Laplace equation (no convection).

Numerical results: Problem 7

	101 × 101 × 101 grid				201 × 201 × 201 grid			
ν	$\frac{n}{n_c}$	inner	iter.	sol.	$\frac{n}{n_c}$	inner	iter.	sol.
Uniform mesh								
∞	4.00	2.00	15	79.6	4.00	2.00	15	82.0
10^{-1}	3.80	2.00	16	99.4	3.81	2.00	17	111.2
10^{-2}	3.75	2.00	18	119.7	3.84	1.94	18	111.7
10^{-4}	3.93	2.00	21	136.9	3.93	2.00	21	136.8
Stretched mesh								
∞	3.91	1.94	16	79.3	3.94	1.88	17	85.8
10^{-1}	3.91	2.00	16	79.2	3.94	1.88	17	85.7
10^{-2}	3.92	1.65	20	95.5	3.95	1.94	17	85.3
10^{-4}	3.46	1.81	21	117.1	3.64	1.87	23	125.8

Conclusions

- The resulting method offers a good compromise between robustness and efficiency.

Conclusions

- The resulting method offers a good compromise between robustness and efficiency.
- **Grey box method:**
applicable to any sparse matrix without user handling;
but
not advisable to use it as in all cases,
e.g. use specialized P_{FF} for systems of PDEs
(exact for all rigid body modes).

Conclusions

- The resulting method offers a good compromise between robustness and efficiency.
- **Grey box method:**
applicable to any sparse matrix without user handling;
but
not advisable to use it as in all cases,
e.g. use specialized P_{FF} for systems of PDEs
(exact for all rigid body modes).
- Further testing on unstructured grid problems is welcome.

Conclusions

- The resulting method offers a good compromise between robustness and efficiency.
- **Grey box method:**
applicable to any sparse matrix without user handling;
but
not advisable to use it as in all cases,
e.g. use specialized P_{FF} for systems of PDEs
(exact for all rigid body modes).
- Further testing on unstructured grid problems is welcome.
- The method seems not too difficult to parallelize.

Conclusions

- The resulting method offers a good compromise between robustness and efficiency.
- **Grey box method:**
applicable to any sparse matrix without user handling;
but
not advisable to use it as in all cases,
e.g. use specialized P_{FF} for systems of PDEs
(exact for all rigid body modes).
- Further testing on unstructured grid problems is welcome.
- The method seems not too difficult to parallelize.

Thank you for your attention !