From Direct to Iterative Substructuring: some Parallel Experiences in 2 and 3D

Luc Giraud

N7-IRIT, Toulouse

MUMPS Day

October 24, 2006, ENS-INRIA, Lyon, France

Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

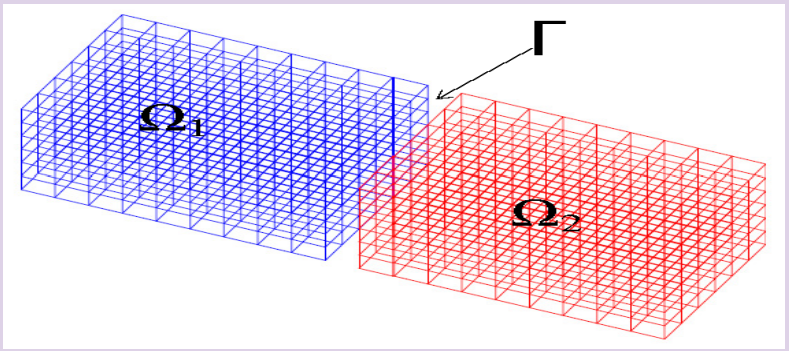
Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

Background in 3D

The model PDE

$$\begin{cases} -div(K.\nabla u) &= f \quad in \quad \Omega, \\ u &= 0 \quad on \quad \partial\Omega_{Dirichlet}, \\ (K\nabla u, n) &= 0 \quad on \quad \partial\Omega_{Neumann}. \end{cases}$$



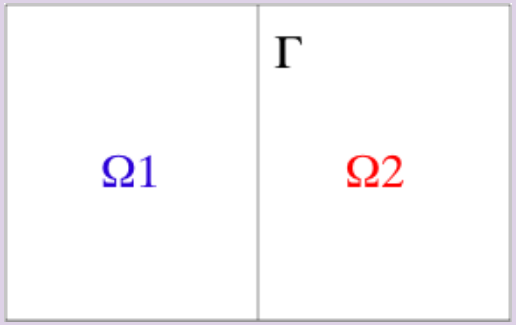
The associated linear system

$$Au = f \equiv \begin{pmatrix} A_{11} & A_{1\Gamma} & 0 \\ A_{1\Gamma}^T & A_\Gamma^{(1)} + A_\Gamma^{(2)} & A_{2\Gamma}^T \\ 0 & A_{2\Gamma} & A_{22} \end{pmatrix} \begin{pmatrix} u_1 \\ u_\Gamma \\ u_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_\Gamma \\ f_2 \end{pmatrix}$$

Background in 2D

The same model PDE but in 2D

$$\begin{cases} -\frac{\partial}{\partial x}(a(x,y)\frac{\partial v}{\partial x}) - \frac{\partial}{\partial y}(b(x,y)\frac{\partial v}{\partial y}) = F(x,y) & \text{in} \quad \Omega, \\ v = 0 & \text{on} \quad \partial\Omega. \end{cases}$$



The associated linear system

$$A_h x = b \equiv \begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ A_{1\Gamma}^T & A_{2\Gamma}^T & A_\Gamma^{(1)} + A_\Gamma^{(2)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_\Gamma \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_\Gamma \end{pmatrix}$$

Algebraic Gaussian elimination

Space dimension free exposure

Algebraic splitting and block Gaussian elimination

$$\underbrace{\begin{pmatrix} A_{11} & 0 & A_{1\Gamma} \\ 0 & A_{22} & A_{2\Gamma} \\ A_{1\Gamma}^T & A_{2\Gamma}^T & A_{\Gamma}^{(1)} + A_{\Gamma}^{(2)} \end{pmatrix}}_{A_h} \begin{pmatrix} x_1 \\ x_2 \\ x_\Gamma \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_\Gamma \end{pmatrix}$$

$$Sx_\Gamma = b_\Gamma - \sum_{i=1}^{2} A_{i\Gamma}^T A_{ii}^{-1} b_i \qquad S = \sum_{i=1}^{2} A_{\Gamma}^{(i)} - A_{i\Gamma}^T A_{ii}^{-1} A_{i\Gamma} = \sum_{i=1}^{2} S^{(i)}$$

Local unassembled Schur complement associated with Ω_i

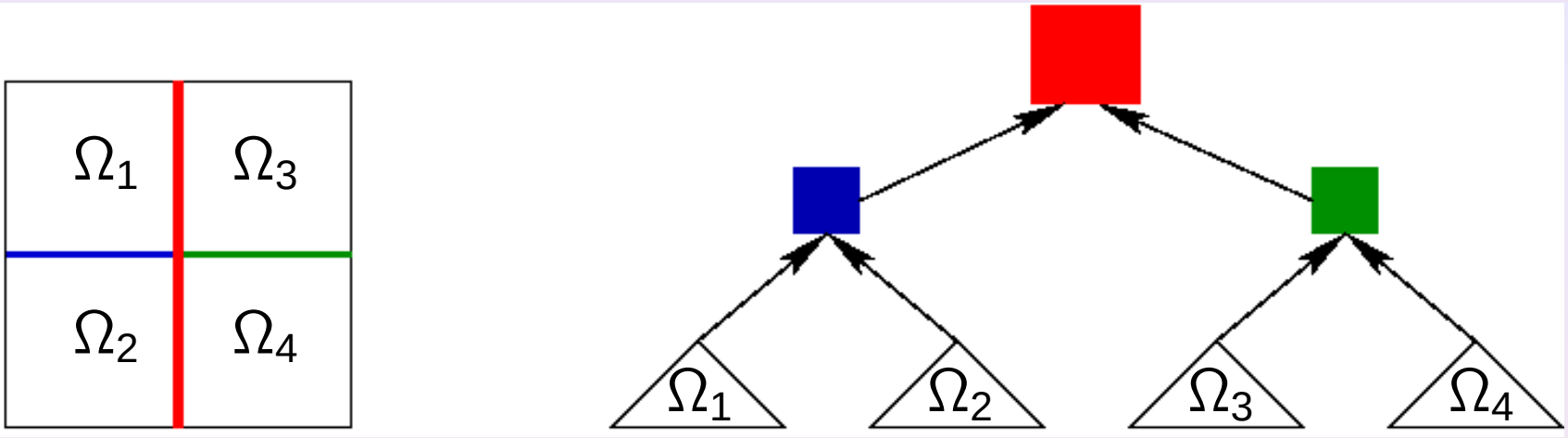
$$A^{(\Omega_i)} = \begin{pmatrix} A_{ii} & A_{i\Gamma_i} \\ A_{i\Gamma_i}^T & A_{\Gamma_i}^{(i)} \end{pmatrix} \Rightarrow S^{(i)} = A_{\Gamma_i}^{(i)} - \underbrace{A_{i\Gamma_i}^T A_{ii}^{-1} A_{i\Gamma_i}}_{\text{Contribution matrix}}$$

Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

The sparse direct multifrontal view point

[Amestoy, Duff, L'Excellent - 98] [Amestoy, Duff, Koster, L'Excellent - 01], [Amestoy, Guermouche, L'Excellent, Pralet - 06]



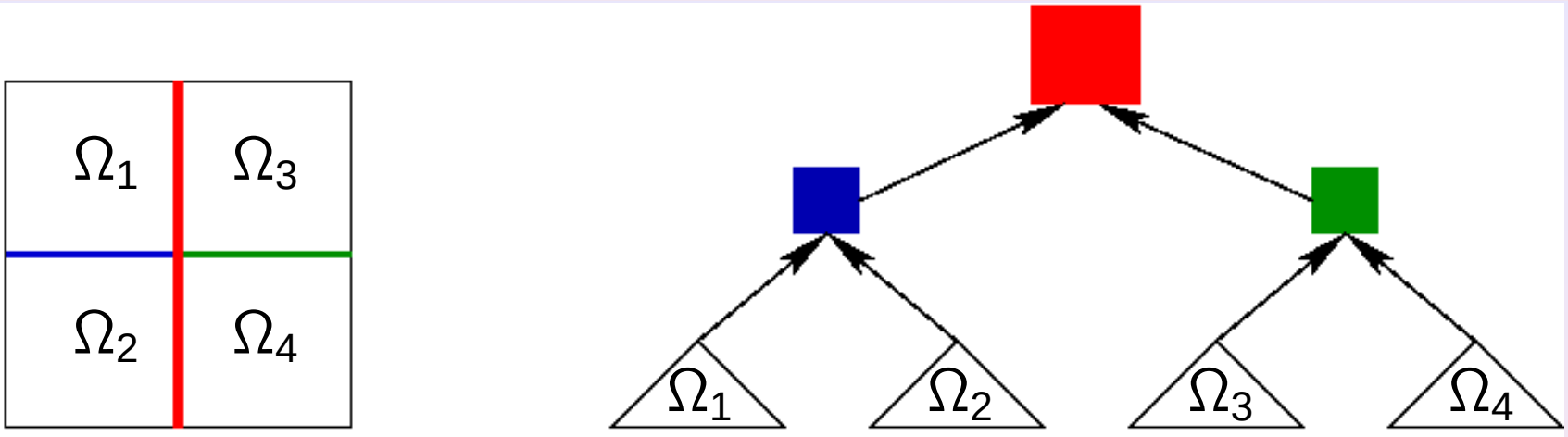
The direct substructuring approach on top of MUMPS

1: Call *#domains* sequential MUMPS in parallel on each $A^{(\Omega_i)}$ to form the local unassembled Schur complement (local ordering)
2: Convert the local ordering on the interface into a global ordering (simple pre-processing)
3: Call one parallel MUMPS on the distributed matrix
$\Rightarrow$ elemental distributed matrices

$$S = \underbrace{\sum_{i=1}^{\#domains} S^{(i)}}_{\text{feature of MUMPS for redundant coordinates}}$$

The sparse direct multifrontal view point

[Amestoy, Duff, L'Excellent - 98] [Amestoy, Duff, Koster, L'Excellent - 01], [Amestoy, Guermouche, L'Excellent, Pralet - 06]



The direct substructuring approach on top of MUMPS

1: Call *#domains* sequential MUMPS in parallel on each $A^{(\Omega_i)}$ to form the local unassembled Schur complement (local ordering)

2: Convert the local ordering on the interface into a global ordering (simple pre-processing)

3: Call one parallel MUMPS on the distributed matrix

$\Rightarrow$ elemental distributed matrices

$$S = \underbrace{\sum_{i=1}^{\#domains} S^{(i)}}_{\text{feature of MUMPS for redundant coordinates}}$$

Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

Elliptic PDEs properties

Algebraic splitting and block Gaussian elimination: # domains = N sub-domains case

$$\begin{pmatrix} A_{I_1 I_1} & \dots & 0 & A_{I_1 \Gamma_1} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & A_{I_N I_N} & A_{I_N \Gamma_N} \\ A_{\Gamma_1 I_1} & \dots & A_{\Gamma_N I_N} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_{I_1} \\ \vdots \\ u_{I_N} \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_{I_1} \\ \vdots \\ f_{I_N} \\ f_\Gamma \end{pmatrix}$$

$$Su_\Gamma = \left(\sum_{i=1}^{N} R_{\Gamma_i}^T S^{(i)} R_{\Gamma_i} \right) u_\Gamma = f_\Gamma - \sum_{i=1}^{N} R_{\Gamma_i}^T A_{\Gamma_i I_i} A_{I_i I_i}^{-1} f_{I_i}$$

where $\quad S^{(i)} = A_{\Gamma_i \Gamma_i}^{(i)} - A_{\Gamma_i I_i} A_{I_i I_i}^{-1} A_{I_i \Gamma_i}$

Spectral properties for elliptic PDE's

$$\kappa(A) = \mathcal{O}(h^{-2}) \qquad \kappa(S) = \mathcal{O}(h^{-1})$$

$$||e^{(k)}||_A \leq 2 \cdot \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k ||e^{(0)}||_A$$

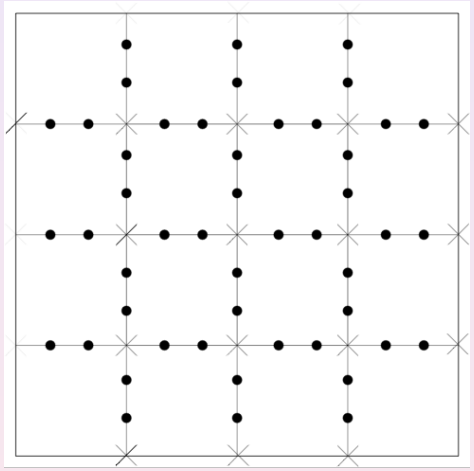
A simple mathematical framework

Block preconditioners

- U a algebraic space of vectors associated with unknowns on Γ
- U_ℓ subspaces of U such that $U = U_1 + ... + U_n$
- R_ℓ : the canonical pointwise restriction from $U \quad \mapsto \quad U_\ell$

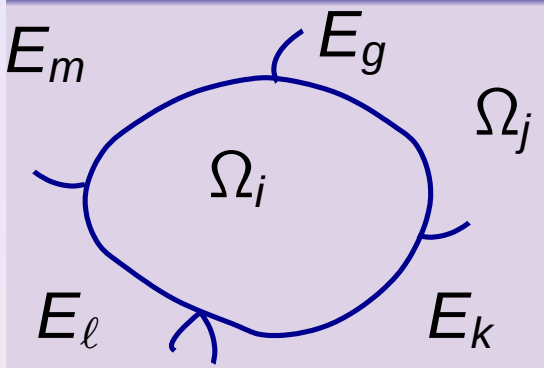
$$M = \sum_{\ell=1}^{n} R_\ell^T M_\ell^{-1} R_\ell \text{ where } M_\ell = R_\ell S R_\ell^T$$

- Examples :
  - U_ℓ associated with each edge/face: block Jacobi (geometric information needed)
  - U_ℓ associated with $\partial\Omega_\ell$: additive Schwarz



Structure of the Local Schur Complement

Non-Overlapping Domain Decomposition



$$\partial\Omega_i = \Gamma_i = E_\ell \cup E_k \cup E_m \cup E_g$$

Distributed Schur Complement

$$S^{(i)} = \begin{pmatrix} S_{mm}^{(i)} & S_{mg} & S_{mk} & S_{m\ell} \\ S_{gm} & S_{gg}^{(i)} & S_{gk} & S_{g\ell} \\ S_{km} & S_{kg} & S_{kk}^{(i)} & S_{k\ell} \\ S_{\ell m} & S_{\ell g} & S_{\ell k} & S_{\ell\ell}^{(i)} \end{pmatrix} \qquad S_{gg} = S_{gg}^{(i)} + S_{gg}^{(j)}$$

- If A is SPD then S is also SPD $\Rightarrow$ CG
- In a distributed memory environment: S is distributed non-assembled

Additive Schwarz preconditioner [Carvalho, Giraud, Meurant - 01]

Preconditionner properties

- U_i associated with the entire interface Γ_i of sub-domain
- $M_{AS} = \sum_{i=1}^{\#domains} R_i^T \left(\bar{S}^{(i)}\right)^{-1} R_i$

$$\bar{S}^{(i)} = \begin{pmatrix} S_{mm} & S_{mg} & S_{mk} & S_{m\ell} \\ S_{gm} & S_{gg} & S_{gk} & S_{g\ell} \\ S_{km} & S_{kg} & S_{kk} & S_{k\ell} \\ S_{\ell m} & S_{\ell g} & S_{\ell k} & S_{\ell\ell} \end{pmatrix} \qquad S^{(i)} = \begin{pmatrix} S^{(i)}_{mm} & S_{mg} & S_{mk} & S_{m\ell} \\ S_{gm} & S^{(i)}_{gg} & S_{gk} & S_{g\ell} \\ S_{km} & S_{kg} & S^{(i)}_{kk} & S_{k\ell} \\ S_{\ell m} & S_{\ell g} & S_{\ell k} & S^{(i)}_{\ell\ell} \end{pmatrix}$$

Assembled local Schur complement

local Schur complement

- $\bar{S}^{(i)}$ and $S^{(i)}$ are dense matrices ...

Remarks

M_{AS} is SPD if S is SPD

Cheaper Additive Shwarz preconditioner form

Main characteristics

- Cheaper in memory space
- Flops reduction
- Without any additional communication cost

Sparsification strategy for $\bar{S}^{(i)}$

$$\widehat{s}_{k\ell} = \begin{cases} \bar{s}_{k\ell} & \text{if} \quad \bar{s}_{k\ell} \geq \epsilon(|\bar{s}_{kk}| + |\bar{s}_{\ell\ell}|) \\ 0 & \text{else} \end{cases}$$

We end-up with a sparse variant of the preconditioner

$$M_{spAS} = \sum_{i=1}^{\#domains} R_i^T \left(\hat{S}^{(i)}\right)^{-1} R_i$$

M_{AS} v.s. Neumann-Neumann

Neumann-Neumann preconditioner

[J.F Bourgat, R. Glowinski, P. Le Tallec and M. Vidrascu - 89]

[Y.H. de Roek, P. Le Tallec and M. Vidrascu - 91]

$$S^{(1)} = S^{(2)} = \frac{S}{2} \Rightarrow S^{-1} = \frac{1}{2}\left(\left(S^{(1)}\right)^{-1} + \left(S^{(2)}\right)^{-1}\right)\frac{1}{2}$$

$$A^{(i)} = \begin{pmatrix} A_{ii} & A_{i\Gamma} \\ A_{i\Gamma} & A_{\Gamma}^{(i)} \end{pmatrix} = \begin{pmatrix} I & 0 \\ A_{i\Gamma}A_{ii}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{ii} & 0 \\ 0 & S^{(i)} \end{pmatrix}\begin{pmatrix} I & A_{ii}^{-1}A_{\Gamma i} \\ 0 & I \end{pmatrix}$$

$$\left(S^{(i)}\right)^{-1} = \begin{pmatrix} 0 & I \end{pmatrix}\left(A^{(i)}\right)^{-1}\begin{pmatrix} 0 \\ I \end{pmatrix}$$

$$M_{NN} = \sum_{i=1}^{\#domains} R_i^T \left(D_i \left(S^{(i)}\right)^{-1} D_i\right) R_i \quad \text{while} \quad M_{AS} = \sum_{i=1}^{\#domains} R_i^T \left(\bar{S}^{(i)}\right)^{-1} R_i$$

Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

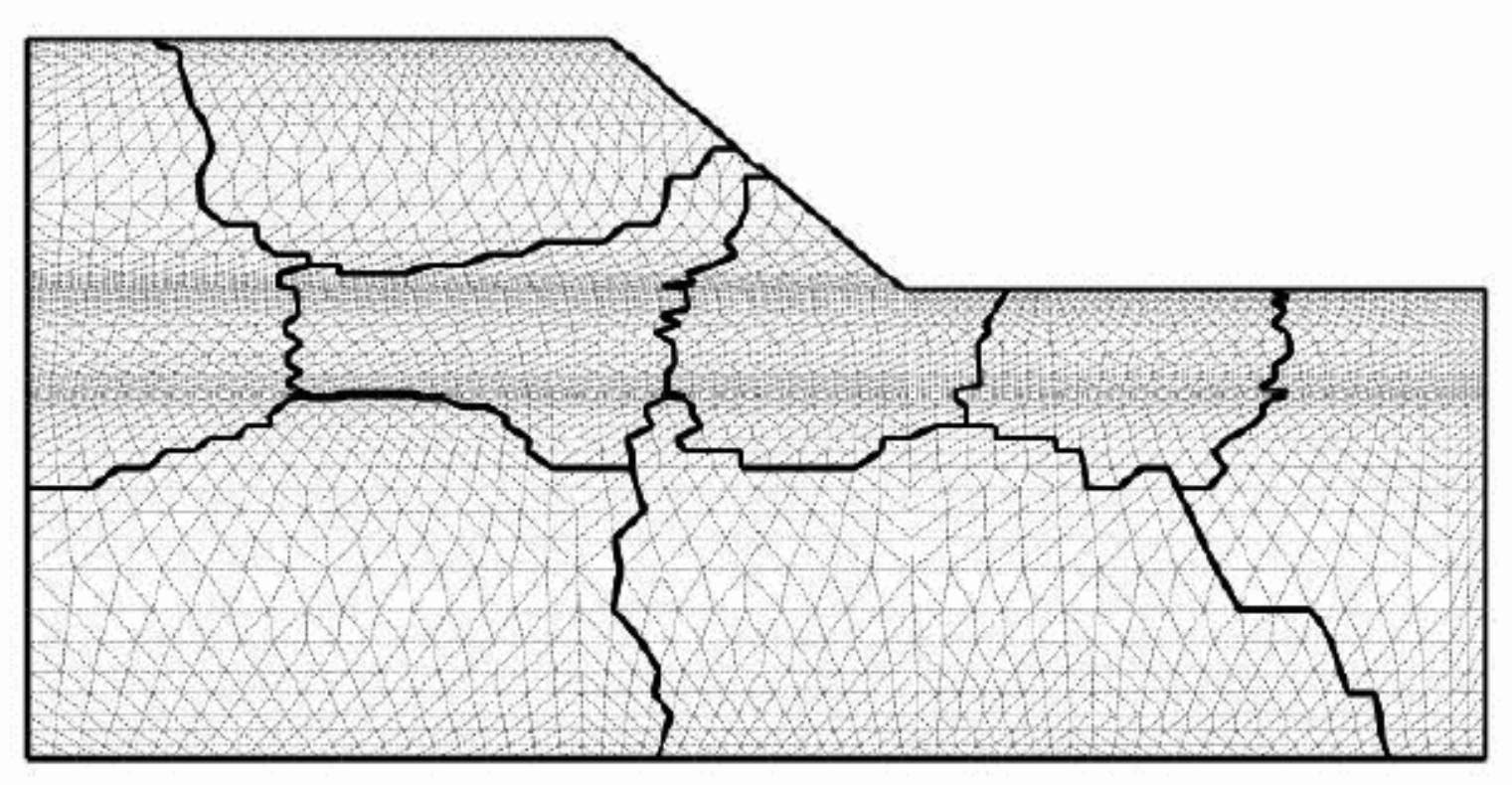
Application to semiconductor device modelling

[Giraud, Koster, Marroco, Rioual - 01], [Giraud, Marroco, Rioual - 05]

The drift diffusion equations

$$\begin{cases} \text{Find } (\phi, \phi_n, \phi_p) \text{ so that} \\ -div(\epsilon \nabla \phi) + q[N(\phi, \phi_n) - P(\phi, \phi_p) - Dop] = 0, \\ -div(q\mu_n N(\phi, \phi_n) \nabla \phi_n) + qGR(\phi, \phi_n, \phi_p) = 0, \\ -div(q\mu_p P(\phi, \phi_p) \nabla \phi_p) - qGR(\phi, \phi_n, \phi_p) = 0, \\ \text{with mixed Dirichlet-Neumann boundary conditions.} \end{cases}$$

Heterojunction device (mixed finite elements)



	Mesh size	# subdomains	Size Schur
Medium 16	365 701 edges	16	2273
Large 32	1 214 758 edges	32	5180

Explicit Schur complement calculation

Experiments on a SGI O2K

$$S = \sum_{i=1}^{N} \underbrace{A_{\Gamma}^{(i)} - A_{i\Gamma} A_{ii}^{-1} A_{\Gamma i a}}_{implicit} = \sum_{i=1}^{N} \underbrace{S^{(i)}}_{explicit}$$

	200×200			400×400		
	Facto	Matvec	20 Krylov	Facto	Matvec	20 Krylov
Implicit	1.94	0.37	9.34	70.36	7.61	222.6
Explicit	3.39	0.01	3.59	119.23	0.46	128.4

Krylov solvers

- CG on SPD systems (Poisson equations)
- Full-GMRES on unsymmetric systems (electrons/holes equations)

Iterative solvers

Embeded iterative schemes: numerical effects

	ε_{Krylov}					
	10^{-5}	10^{-7}	10^{-9}	10^{-11}	10^{-13}	10^{-15}
Newton Steps	$\times$	214	182	176	176	176

Performance of the preconditioners

	Medium 16		*Large 32*	
	M_{bJ}	M_{AS}	M_{bJ}	M_{AS}
Newton its	182	176	228	175
iter CG	38	24	68	32
iter GMRES	34	24	94	34
time (s)	788	809	2892	1654

Direct v.s. Hybrid iterative solve

	Medium 16		*Large 32*	
	M_{AS}	*Dss*	M_{AS}	*Dss*
Newton	176	173	175	166
time (s)	809	1140	1654	2527

3D experiments [Giraud, Haidar, Watson - ongoing]

3D Poisson problem:

- Number of CG iterations where either:
  - $\frac{H}{h}$ constant while # sub-domains is varied horizontal view $\rightarrow$
  - Increasing mesh size $\frac{H}{h}$ while # sub-domains kept constant vertical view $\downarrow$

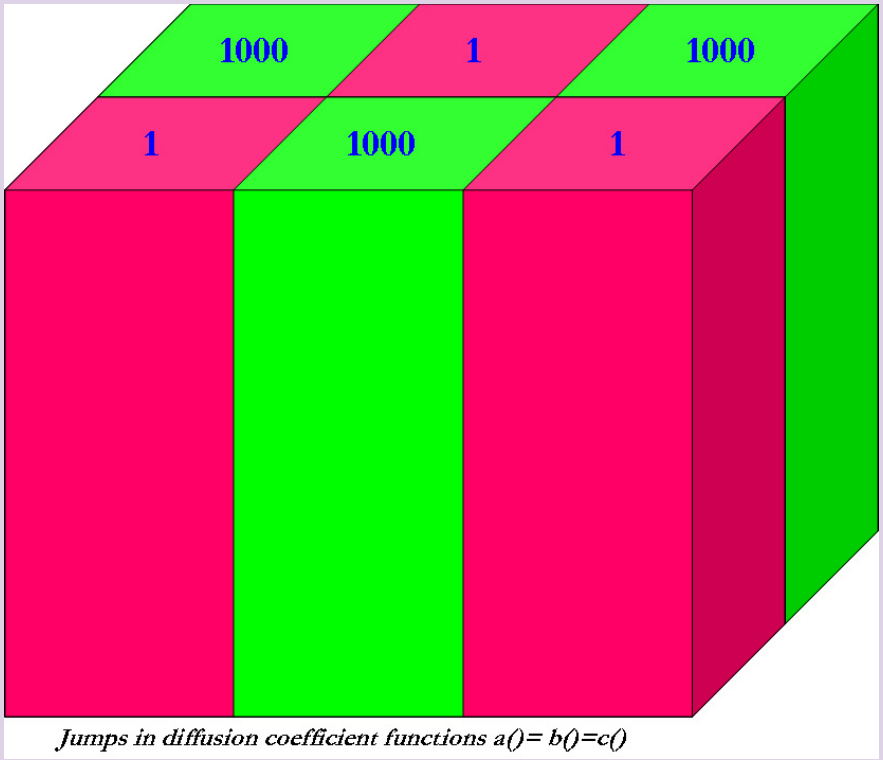
		# sub-domains $\equiv$ # processors							
sub-domains size		27	64	125	216	343	512	729	1000
$20 \times 20 \times 20$	M_{AS}	16	23	25	29	32	35	39	42
	M_{SpAS}	16	23	26	31	34	39	43	46
$25 \times 25 \times 25$	M_{AS}	17	24	26	31	33	37	40	43
	M_{SpAS}	17	25	28	34	37	42	45	49
$30 \times 30 \times 30$	M_{AS}	18	25	27	32	34	39	42	45
	M_{SpAS}	18	26	29	36	40	44	48	52
$35 \times 35 \times 35$	M_{AS}	19	26	30	33	35	43	44	47
	M_{SpAS}	19	28	30	38	46	46	50	56

- The solved problem size vary from 1.1 up to 42.8 Millions of unknowns
- The number of iterations increases slightly when increasing # sub-domains
- This increase is less significant when the local mesh size $\frac{H}{h}$ grows
- With a few hundred processors those runs cannot be performed using a direct substructuring approach with current version of MUMPS

Numerical scalability

3*D* Discontinuous problem :

Jumps in diffusion coefficient functions $a() = b() = c()$: $1 - 10^3$



Numerical scalability

3*D* Discontinuous problem :

Jumps in diffusion coefficient functions $a() = b() = c()$: $1 - 10^3$

- Number of CG iterations where either:
  - $\frac{H}{h}$ constant while # sub-domains is varied horizontal view$\longrightarrow$
  - Increasing mesh size $\frac{H}{h}$ while # sub-domains kept constant vertical view$\downarrow$

		# sub-domains $\equiv$ # processors							
sub-domains size		27	64	125	216	343	512	729	1000
$20 \times 20 \times 20$	M_{AS}	32	37	44	53	58	68	78	82
	M_{SpAS}	32	42	48	58	63	75	85	91
$25 \times 25 \times 25$	M_{AS}	29	41	46	52	60	71	80	85
	M_{SpAS}	34	45	51	63	66	82	89	99
$30 \times 30 \times 30$	M_{AS}	34	43	46	57	61	75	84	87
	M_{SpAS}	30	47	52	68	70	90	96	105
$35 \times 35 \times 35$	M_{AS}	31	43	49	62	63	80	87	92
	M_{SpAS}	29	51	58	71	84	92	105	116

Parallel performance

3*D* Discontinuous problem:

- Implementation details:
  - Setup Schur: MUMPS
  - Setup Precond: dense Schur (LAPACK)- sparse Schur (MUMPS)
  - Target computer : System Xserve MAC G5
- Parallel elapsed time: 10^3 processors $\frac{H}{h}$ *vary* $\epsilon = 10^{-4}$
  - Jumps in diffusion coefficient functions $a() = b() = c()$: $1 - 10^3$

Sub-domains size	$20 \times 20 \times 20$		$25 \times 25 \times 25$		$30 \times 30 \times 30$		$35 \times 35 \times 35$	
setup Schur	1.30	1.30	4.20	4.20	11.2	11.2	26.8	26.8
setup Precond	0.93	0.50	3.05	1.60	8.73	3.51	21.4	6.22
time per iter	0.08	0.05	0.23	0.13	0.50	0.28	0.77	0.37
total	8.79	6.17	26.8	18.6	63.0	44.1	119	75.9
# iter	82	91	85	99	87	105	92	116

dense local Schur Precond M_{AS} - sparse local Schur Precond M_{SpAS}

Local data storage

M_{AS} vs M_{SpAS} Memory behaviour

	M_{AS}	M_{SpAS}	
Subdomains size		$\epsilon = 10^{-5}$	$\epsilon = 10^{-4}$
$20 \times 20 \times 20$	35.85_{MB}	7.5_{MB} (10%)	1.8_{MB} (5%)
$25 \times 25 \times 25$	91.23_{MB}	12.7_{MB} (14%)	2.7_{MB} (3%)
$30 \times 30 \times 30$	194.4_{MB}	19.4_{MB} (10%)	3.8_{MB} (2%)
$35 \times 35 \times 35$	367.2_{MB}	28.6_{MB} (7%)	10.2_{MB} (2%)

Outline

1. General Framework
2. The direct substructuring approach
3. The iterative substructuring approach
   - Description of the preconditioners
   - Variant of Additive Shwarz preconditioner M_{AS}
   - M_{AS} v.s. Neumann-Neumann
4. Parallel numerical experiments
   - 2D experiments in semiconductor device modelling
   - 3D experiments on the academic cube
5. Prospectives

Prospectives

Objective

- Control the growth of iterations when increasing the # processors

Various possibilities (future work)

- Numerical remedy: two-level preconditioner
  - Coarse space correction, ie solve a closed problem on a coarse space
  - Various choices for the coarse component (eg one d.o.f. per sub-domain)
- Computer Science remedy : several processors per sub-domain
  - two-level of parallelism
  - 2D cyclic data storage

MUMPS is back !!!

Final comments

The advances in sparse directs techniques anables us to consider preconditioning techniques that were out of reach before

Prospectives

Objective

- Control the growth of iterations when increasing the # processors

Various possibilities (future work)

- Numerical remedy: two-level preconditioner
  - Coarse space correction, ie solve a closed problem on a coarse space
  - Various choices for the coarse component (eg one d.o.f. per sub-domain)
- Computer Science remedy : several processors per sub-domain
  - two-level of parallelism
  - 2D cyclic data storage

MUMPS is back !!!

Final comments

The advances in sparse directs techniques anables us to consider preconditioning techniques that were out of reach before

Prospectives

Objective

- Control the growth of iterations when increasing the # processors

Various possibilities (future work)

- Numerical remedy: two-level preconditioner
  - Coarse space correction, ie solve a closed problem on a coarse space
  - Various choices for the coarse component (eg one d.o.f. per sub-domain)
- Computer Science remedy : several processors per sub-domain
  - two-level of parallelism
  - 2D cyclic data storage

MUMPS is back !!!

Final comments

The advances in sparse directs techniques anables us to consider preconditioning techniques that were out of reach before

Credit to co-workers

Co-workers

- L. M. Carvalho (UFRJ Brasil, mainly during PhD at CERFACS),
- A. Haidar (CERFACS),
- J. Koster (Norway ministery of research, mainly when at Parallab),
- P. Le Tallec (Ecole Polytechnique & Paris Dauphine, mainly when at Paris Dauphine & INRIA),
- J. C. Rioual (NEC, mainly during PhD at CERFACS),
- A. Marrocco (INRIA),
- G. Meurant (CEA),
- L. Watson (Virginia Tech).

and of course to

the MUMPS team

More information

http://www.n7.fr/ ~giraud

Merci pour votre attention

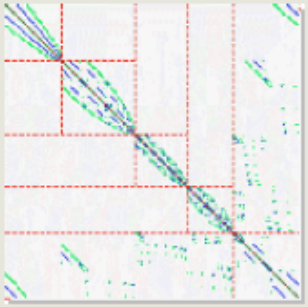
a final “personal” advert ...

More information

http://www.n7.fr/ ~giraud

Merci pour votre attention

a final “personal” advert ...



2007 International Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications

July 9-12, 2007
Météopole
Toulouse, France

conference info

- home page
- themes and motivations
- topics
- program committee
- invited presentations
- contributed papers
- important dates
- conference site

Program Chairs

Luc Giraud	ENSEEIHT-IRIT, France
Esmond G. Ng	Lawrence Berkeley National Laboratory, USA
Yousef Saad	The University of Minnesota, USA
Wei-Pai Tang	The Boeing Company, USA

Program Committee

Cleve Ashcraft	Livermore Software Technology Corp., USA
Michele Benzi	Emory University, USA
Matthias Bollhoefer	Technische Universität Berlin, Germany
Iain Duff	Rutherford Appleton Laboratory, UK and CERFACS, France
Stéphane Grihon	Airbus, France
Misha Kilmer	Tufts University, USA
Gérard Meurant	CEA, France
Arnold Reusken	University of Aachen, Germany
Jean Roman	LaBRI-INRIA Futurs, France
Valeria Simoncini	University of Bologna, Italy

Local Organization

Patrick Amestoy	ENSEEIHT-IRIT, France
Luc Giraud	ENSEEIHT-IRIT, France
Serge Gratton	CERFACS, France