Parallel Computing and Iterative Solution of Linear Systems

R. Blaheta
Institute of Geonics AS CR, Ostrava, Czech Republic
&
VŠB-TU Ostrava
www.ugn.cas.cz

SNA '05
Ostrava, February 7-11, 2005

Outline of the talk

Aim: describe parallel iterative solvers for PDE-FEM linear algebraic systems, discuss their mathematical/computer implementation aspects

- Paralelization of CG and other iterative methods
- Block diagonal and full block preconditioners
- Overlapping Schwarz DD methods
- Convergence analysis
- Two-level Schwarz methods
- Algebraic constructions of coarse grid problem
- Nonoverlapping DD methods
- What is not considered

Literature

- Chan, T.F., Mathew, T.P., Domain Decomposition Algorithms. Acta Numerica 1994, 61-143
- Smith, B.F., Bjørstad, P.E., Gropp, W.D., Domain Decomposition Parallel Multilevel Methods for Elliptic Partial Differential Equations. Cambridge University Press, Cambridge 1996
- A. Toselli, O. Widlund, Domain Decomposition Methods - Algorithms and Theory, Springer 2004
- R. Blaheta, Space decomposition Preconditioners and Parallel Solvers. In: M. Feistauer et al. eds. Numerical Mathematics and Advanced Appl., Springer 2004, pp. 20-38 (ENUMATH 2003 invited lecture)
- R. Blaheta, O. Jakl, J. Starý, Linear sytem solvers based on space decompositions and parallel computations, Inženýrská Mechanika, 10(2003), pp. 439-454

The Problem

For modelling of different phenomena as diffusion etc. , it is necessary to solve a a symmetric elliptic boundary value problem in Ω like:

$$\begin{aligned} -\sum_{ij} \frac{\partial}{\partial x_i}\left(k_{ij}\frac{\partial u}{\partial x_j}\right) &= f \qquad \text{in } \Omega \\ u &= \hat{u} \qquad \text{on } \Gamma_0 \subset \partial\Omega \\ \sum_{ij} k_{ij}\frac{\partial u}{\partial x_j} n_i &= \hat{f} \qquad \text{on } \Gamma_1 \subset \partial\Omega \end{aligned}$$

Discretization by FEM leads to an algebraic problem

$$Au = b$$

with a matrix A, which is SPD, large scale (dimension $n \sim 10^5 - 10^7$) and ill conditioned.

Linear System Solvers

- direct methods (Gauss elimination type)
- iterative methods - history
  - relaxation (Gauss, Jacobi, Southwell)
  - SOR (Young 1954, Varga 1962)
  - CG (Hestenes, Stiefel, Lanczos, 1952)
  - practical CG (Reid 1971)
  - ILU preconditioning (Axelsson 1972, van der Vorst 1977,...)
  - multigrid (Fedorenko 1964, Hackbusch 1977, ...), AMG
  - GMRES (Saad, Schulz 1986)
  - DD, //iterative methods (Dryja, Widlund, ...1st DD conf. 1989)
- iterative methods - today: Numerical scalability (optimal linear complexity), Parallel scalability, Robustness, Nonsymmetric Systems..

Sequential CG

given u^0

compute $r^0 = b - Au^0$,

$g^0 = B^{-1} r^0,\ r^0 = g^0,\ \ \sigma = \langle r^0, g^0 \rangle$

for $i = 0, 1, \ldots$ **until** convergence **do**

$w^i = Av^i$

$\alpha_i = \sigma / \langle v^i, w^i \rangle$

$u^{i+1} = u^i + \alpha_i v^i$

$r^{i+1} = r^i - \alpha_i w^i$

$g^{i+1} = B^{-1} r^{i+1}$

$\sigma_0 = \sigma, \quad \sigma = \langle r^{i+1}, g^{i+1} \rangle$

$\beta_i = \sigma / \sigma_0$

$v^{i+1} = g^{i+1} + \beta_i v^i$

end

CG iteration needs

- matrix–by–vector multiplication
- vector updates
- two inner products

Preconditioning

$g = G(r) = B^{-1} r$

standard B - SPD

#it$\sim$

$\sqrt{\mathrm{cond}(B^{-1}A)}\,\ln(1/\varepsilon)$

Convergence test

$\| r^i \| \leq \varepsilon \| b \|$

Parallelization by Splitting to Blocks

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ & A_{22} & A_{23} & A_{24} \\ & & A_{33} & A_{34} \\ sym & & & A_{44} \end{bmatrix} \begin{matrix} \}\, \underline{A}_1 \\ \\ \}\, \underline{A}_2 \end{matrix} \qquad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \begin{matrix} \}\, \underline{v}_1 \\ \\ \}\, \underline{v}_2 \end{matrix} \qquad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \begin{matrix} \}\, \underline{w}_1 \\ \\ \}\, \underline{w}_2 \end{matrix}$$

v, w , A are full (**accumulated**) vectors and matrix ,

$\underline{v} = (\underline{v}_1,\ \underline{v}_2)^T,\ w = (\underline{w}_1,\ \underline{w}_2)^T,\ \underline{A} = (A_1,\ A_2)^T$ are **distributed** data.

Distribution on two processors $\mathbb{P}_1$ and $\mathbb{P}_2$, $\mathbb{P}_k \Leftrightarrow \{A_k,\ \underline{v}_k,\ \underline{w}_k,\ v,\ w\}$.

- Matrix-by-vector multiplication $\underline{w}_k = A_k \cdot v$
- inner product $\sigma = \langle v,\, w\rangle$ or $\sigma_k = \langle v,\, \underline{w}_k\rangle \rightarrow \sigma = \sum \sigma_k$
- vector updates $v = v + \alpha \cdot w$ or $\underline{v}_k = \underline{v}_k + \alpha \cdot \underline{w}_k$

Preconditioning by Splitting to Blocks

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ & A_{22} & A_{23} & A_{24} \\ & & A_{33} & A_{34} \\ sym & & & A_{44} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \Rightarrow \quad B = \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix}$$

B additive (block diagonal, block Jacobi) preconditioner

$$A = \begin{bmatrix} I & \\ B_{21}B_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} B_{11} & \\ & S \end{bmatrix} \begin{bmatrix} I & B_{11}^{-1}B_{12} \\ 0 & I \end{bmatrix}$$

$$\Rightarrow \quad B_M = \begin{bmatrix} I & \\ B_{21}B_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix} \begin{bmatrix} I & B_{11}^{-1}B_{12} \\ & I \end{bmatrix}$$

B_M symmetric multiplicative (full block, Gauss-Seidel) preconditioner

Parallel CG

given u^0

compute $\mathbf{r}^0 = \mathbf{b} - \mathbf{A}u^0$,

$\mathbf{g}^0 = B^{-1}\mathbf{r}^0,\ v^0 = \mathbf{g}^0,\ \ \sigma = \langle \mathbf{r}^0, \mathbf{g}^0 \rangle$

for $i = 0, 1, \ldots$ **until** convergence **do**

$\mathbf{w}^i = \mathbf{A}v^i$

$\alpha_i = \sigma / \langle v^i, \mathbf{w}^i \rangle$ **(C)**

$u^{i+1} = u^i + \alpha_i v^i$

$\mathbf{r}^{i+1} = \mathbf{r}^i - \alpha_i \mathbf{w}^i$

$\mathbf{g}^{i+1} = B^{-1}\mathbf{r}^{i+1}$

$\sigma_0 = \sigma,\ \ \sigma = \langle \mathbf{r}^{i+1}, \mathbf{g}^{i+1} \rangle$ **(C)**

$\beta_i = \sigma / \sigma_0$

$v^{i+1} = \mathbf{g}^{i+1} + \beta_i v^i$ **(C)**

end

Parallel computation:

multiple arithmetic units & better memory access

v accumulated

$\mathbf{v} = \underline{v}$ distributed

(C) communication points

Block diagonal preconditioner

Properties of the //CG & Preconditioning

- parallel implementation of the block diagonal preconditioning, no communication, parallel scalability depends on the cost of the necessary communications,
- numerical scalability (efficiency): number of iterations vs. complexity of one iteration.
- for a general splitting into m blocks and m processors:
  1. for m small, the solution of systems with B_k is expensive,
  2. for m big, the preconditioner is not efficient.

  Note: $m \to n$: $B =$diag$(A) \Rightarrow$ cond$(B^{-1}A) = O(h^{-2})$. **Remedies:**

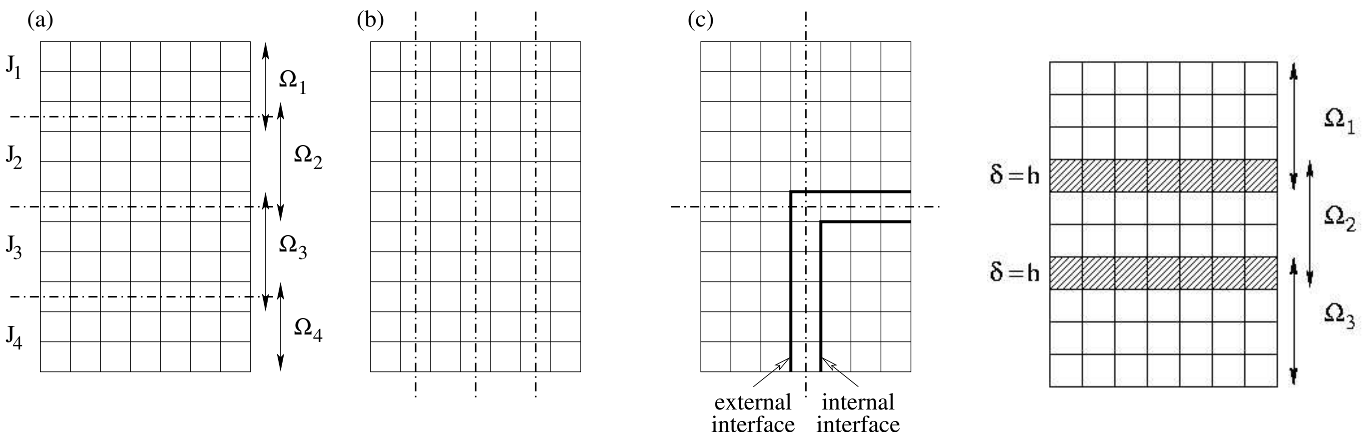
1. for m small, use an approximation to B_k, e.g. incomplete factorization,
2. for m big, use more efficient efficient methods, e.g. Schwarz overlapping DD methods.

Block diagonal - overlapping DD preconditioners

$$u = \begin{bmatrix} \underline{u}_1 \\ \vdots \\ \underline{u}_m \end{bmatrix} \quad A = \begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mm} \end{bmatrix}$$

- additive blok diagonal preconditioner B,
- R_k Boolean … $n_k \times n$,
- $\underline{u}_k = R_k u$,
- $B^{-1} r = \sum R_k^T A_{kk}^{-1} R_k r$

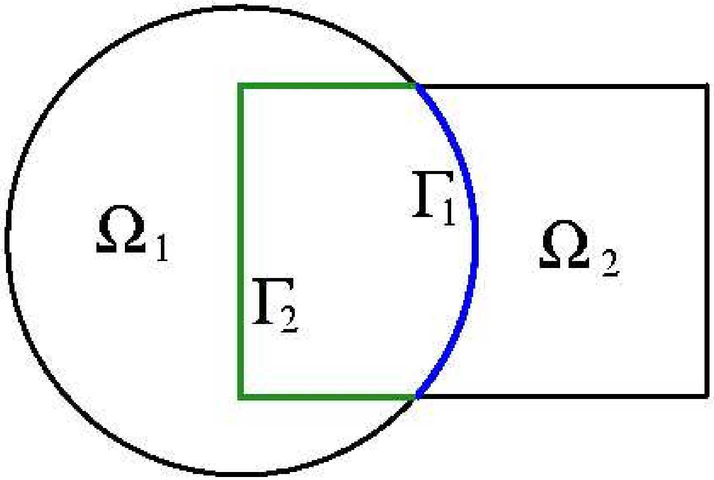
Block splitting - DD:



increase of δ: overlapping DD Schwarz methods: $B^{-1} r = \sum R_k^T A_{kk}^{-1} R_k r$

space decomposition - subspace coorrection methods

Schwarz method/preconditioner



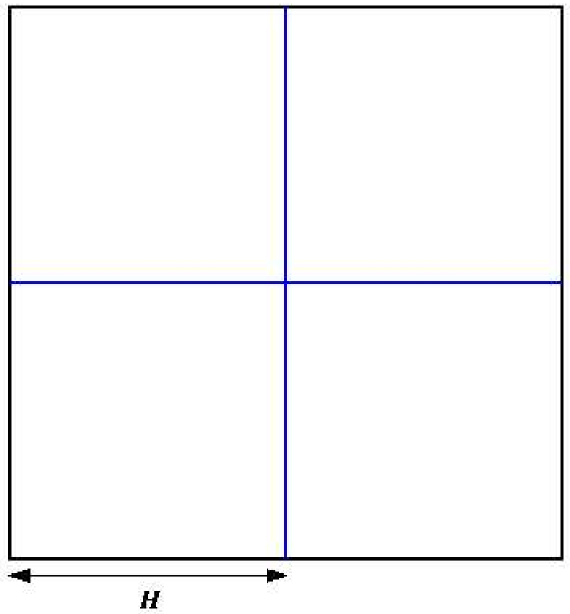
Alternating Schwarz method - existence of the solution of BVPs in more general domains, H.A. Schwarz 1870

Preconditioner

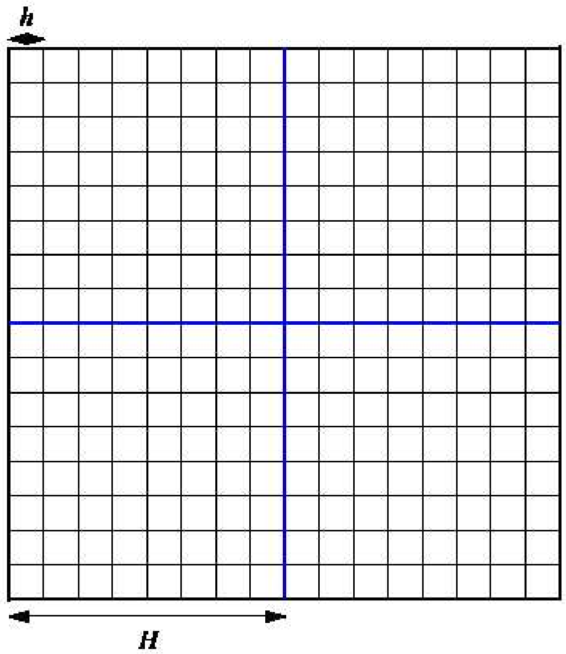
$$B^{-1}r = \sum R_k^T A_{kk}^{-1} R_k r$$

- A.M. Matsokin, S.V. Nepomnyaschikh (1985)
- Lions, P.L. (1987): On the Schwarz alternating methods *I.* In: *1st Internat. Symposium on Domain Decomposition Methods for PDE*, SIAM, pp. 1-42
- M. Dryja, O.B. Widlund (1989), Towards a unified theory of domain decomposition algorithms for elliptic problems, 3rd *Internat. Symposium on Domain Decomposition Methods for PDE*, SIAM, pp. 3-21

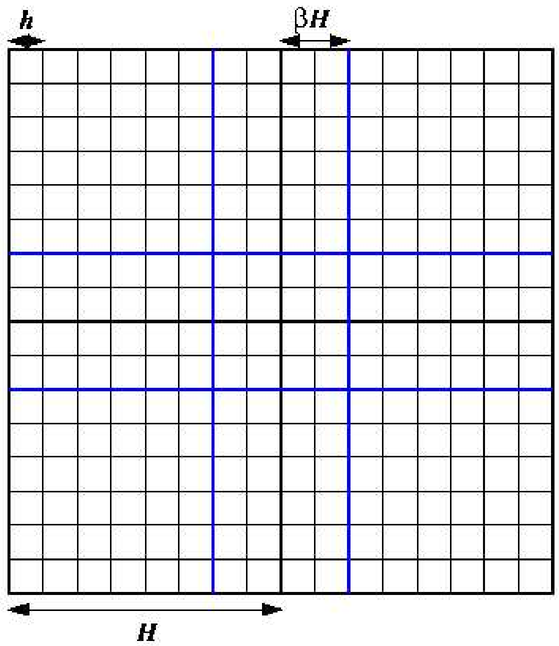
Overlapping Subdomains



Coarse partition of Ω into Ω_k^H with auxiliary grid of size H.

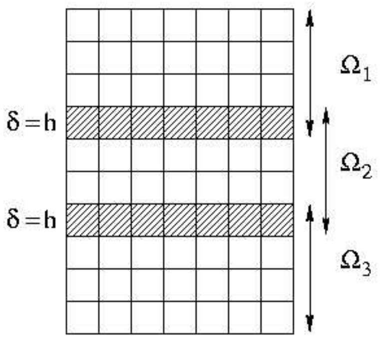


Refinement of the domain partition with grid size h.



Extension of Ω_k^H with size βH into overlapping subdomains Ω_k.

Parallelization of the overlapping DD



Subdomain matrix $\underline{A}_k$ e.g. FEM matrix corresponding to the problem in Ω_k with the Dirichlet BC on inner boundary Γ_k realized by penalty scaling of diagonal elements. $A_k = R_k^T \, A \, R_k$ is involved in $\underline{A}_k$.

Subdomain vector $\underline{v}_k$ includes all DOFs in the subdomain $\Omega_k \setminus \partial\Omega_{Dirichlet}$ including the DOFs on Γ_k.

v, w , A are full (**global**) vectors and matrix ,

Distribution on processors $\mathbb{P}_k \Leftrightarrow \{\underline{A}_k, \ \underline{v}_k, \ \underline{w}_k\}$ (only **local** data).

- Matrix-by-vector multiplication $\underline{w}_k = \underline{A}_k \cdot \underline{v}_k$ and transfer of data to DOFs on Γ_k.
- inner product $\sigma_k = \langle \underline{v}_k, \underline{w}_k \rangle - \sum_{j\, in\, overlap} \frac{1-K_j}{K_j} \underline{v}_{k,j} \cdot \underline{w}_{k,j} \rightarrow \sigma = \sum \sigma_k$
- vector updates $\underline{v}_k = \underline{v}_k + \alpha \cdot \underline{w}_k$

Parallel CG with overlapping DD preconditioner

given u^0

compute $\mathbf{r}^0 = \mathbf{b} - \mathbf{A}u^0$,

$\mathbf{g}^0 = B^{-1}\mathbf{r}^0,\ \mathbf{v}^0 = \mathbf{g}^0,\ \ \sigma = \langle \mathbf{r}^0, \mathbf{g}^0 \rangle$

for $i = 0, 1, \ldots$ **until** convergence **do**

$\mathbf{w}^i = \mathbf{A}\mathbf{v}^i$ **(C)**

$\alpha_i = \sigma / \langle \mathbf{v}^i, \mathbf{w}^i \rangle$ **(C)**

$\mathbf{u}^{i+1} = \mathbf{u}^i + \alpha_i \mathbf{v}^i$

$\mathbf{r}^{i+1} = \mathbf{r}^i - \alpha_i \mathbf{w}^i$

$\mathbf{g}^{i+1} = B^{-1}\mathbf{r}^{i+1}$ **(C)**

$\sigma_0 = \sigma,\ \ \sigma = \langle \mathbf{r}^{i+1}, \mathbf{g}^{i+1} \rangle$ **(C)**

$\beta_i = \sigma / \sigma_0$

$\mathbf{v}^{i+1} = \mathbf{g}^{i+1} + \beta_i \mathbf{v}^i$

end

CG with additive Schwarz preconditioner

Only local vectors $\mathbf{v} = \underline{v}$ and local matrices $\mathbf{A}$

(C) communication points

Effect of the overlap

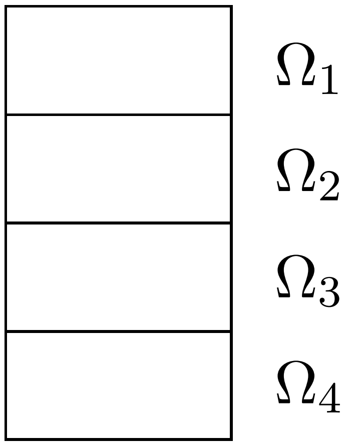
Model problem:

$\Omega = \langle 0, 2 \rangle \times \langle 0, 3 \rangle$

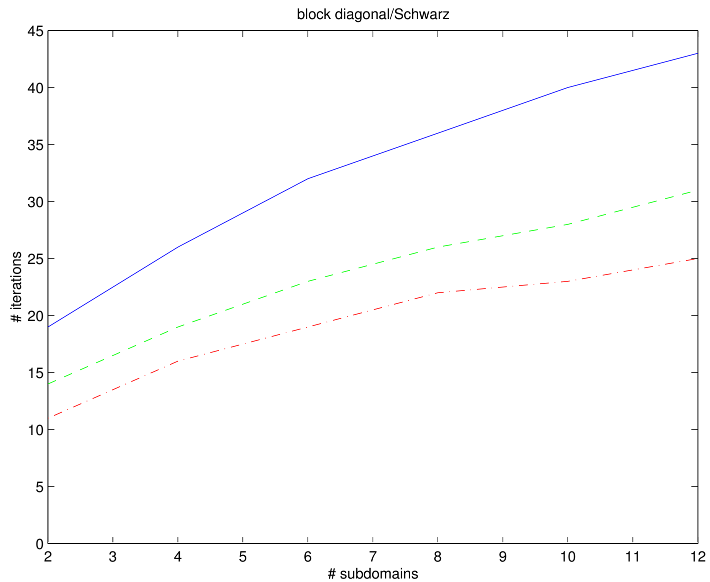
$-\triangle u = f$ in Ω

$u = 0$ on $\partial\Omega$

h=1/30, n=5100



	δ:	h	$2h$	$3h$
m	H:	–	–	–
2		19	14	11
4		26	19	16
6		32	23	19
8		36	26	22
10		40	28	23
12		43	31	25



Numbers of iterations for $\varepsilon = 10^{-4}$ with additive precond., overlap $\delta = h$ - Block Jacobi, $\delta > h$ Schwarz

Space decompositions

$$R^n,\;\; R_k : R^n \to R^{n_k}\;(k=1,\ldots,m),\;\; I_k : R^{n_k} \to R^n,\;\; I_k = R_k^T$$

$$R^n = \textstyle\sum_k \,\text{range}(I_k)$$

$$A_k = R_k^T\, A\, R_k$$

Abstract setting - finite dimensional Hilbert space V

$$V = V_1 + \ldots + V_m$$

$$V,\;\; R_k : V \to V_k\;(k=1,\ldots,m),\;\; I_k : V_k \to V,\;\; I_k = R_k^T$$

$$Au \;=\; b\,,\;\; u \in V,\;\; b \in V',\;\; A \in L(V,V')$$

$$A \text{ is SPD } \;\leftrightarrow\; \langle Au, v\rangle = \langle Av, u\rangle,\; \langle Av, v\rangle > 0 \;\Rightarrow\; \langle u, v\rangle_A$$

$$A_k = R_k^T\, A\, R_k$$

- analysis in the FE spaces,
- investigation of DD and other decomp. (DiD) in a unique framework.

Space decomposition preconditioners

$$G : r \mapsto g, \;\; g \sim A^{-1} r \;(pseudores.)$$

$g = 0$
for $k = 1, \ldots, m$
$\quad g \leftarrow g + I_k \tilde{A}_k^{-1} R_k\, z_k$
end

additive preconditioner G_A:
$z_k = r, \;\; k = 1, \ldots, m$
multiplicative preconditioner G_M:
$z_k = r - Ag$
hybrid preconditioner G_H, e.g.:
$z_k = r \;\; k = 1, \ldots, m-1$
$z_m = r - Ag$

$\tilde{A}_k \sim R_k A I_k = A_k$
$S_k(z) \sim A_k^{-1} z$
exact
inexact – linear
– nonlinear

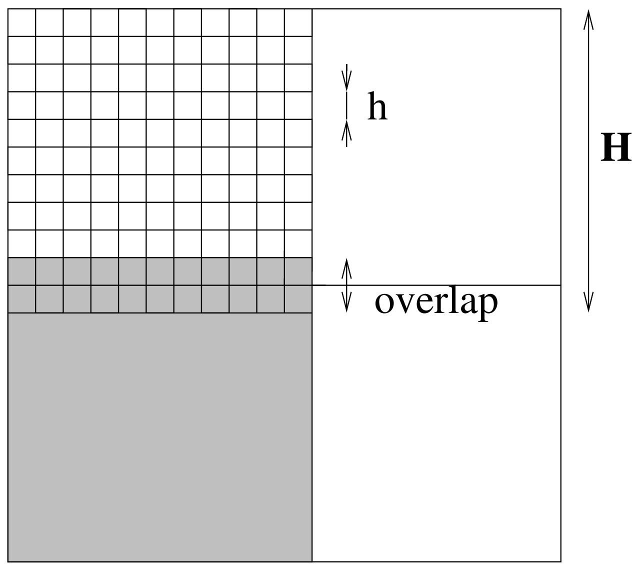
symmetric SD preconditioner
nonsymmetric
symmetrization
$k = 1, \ldots, m-1, m, m-1, \ldots, 1$

SD analysis, $V = V_1 + ... + V_m$

- Assumption **A1** (stability) $\forall v \in V \ \exists v_k \in V_k : \ v = v_1 + \ldots + v_m$
  $$\sum_k \| v_k \|_A^2 \leq K_0 \| v \|_A^2.$$
- Assumption **A2** $\forall v \in V \ \forall v_k \in V_k : \ v = v_1 + \ldots + v_m$
  $$\| v \|_A^2 \leq K_1 \sum_k \| v_k \|_A^2.$$
- **Note:** Trivial bound $K_1 = m$. Potentially m-independent bounds:
  If $\mathcal{E} = (\varepsilon_{kl})$, $\varepsilon_{kl} = \cos(V_k, V_l)_A$, then $K_1 \leq \rho(\mathcal{E}) \leq \max \sum_l \varepsilon_{kl}$.
  If $\mathcal{E}_0 = (\varepsilon_{kl} : \ k, l \neq j)$, then $K_1 \leq 2(1 + \rho(\mathcal{E}_0))$.
- **Theorem:** Let A1, A2 hold. Then
  $\lambda_{\min}(G_A A) \geq 1/K_0$, $\lambda_{\max}(G_A A) \leq K_1$, cond $(G_A A) \leq K_0 K_1$
  $$\| I - G_S A \|_A = \| I - G_M A \|_A^2 \leq \left(1 - \frac{1}{K_0(1+K_1)^2}\right)^2.$$

Matsokin, Nepomnyaschikh 1985, Lions 1988, Dryja, Widlund 1987,1989, Bramble, Pasciak, Wang, Xu 1991, Bjørstad, Mandel 1991

Domain decomposition (DD)



$\Omega = \Omega_1 \cup \ldots \cup \Omega_m$ aligned with $\mathcal{T}_h$

overlap $\delta = \text{dist}\left(\partial\Omega_k \cap \Omega_l, \partial\Omega_l \cap \Omega_k\right)$

subdomain size $\mathbf{H} = \max \ \text{diam}\left(\Omega_k\right)$

$\delta = \beta\mathbf{H}, \ \ \mathbf{H} \to 0 \Rightarrow \delta \to 0$

$\Omega_k \to V_k = \{v \in V_h : \ v = 0 \quad \text{on } \Omega \backslash \Omega_k\}$

- ex. partition of unity $\theta_1, \ldots, \theta_m$

  $\sum \theta_k = 1$ on Ω

  $\theta_k \in C^\infty\left(R^d\right), \theta_k = 0$ on $R^d \backslash \Omega_k$ $\Rightarrow$

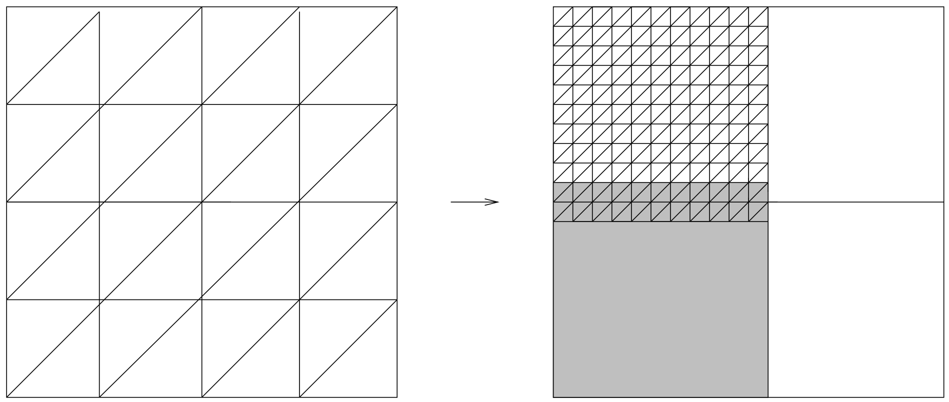
  $\| \nabla\theta_k \|_{L^\infty(\Omega)} \leq c/\delta$

- ex. interpolation $\Pi_h : C\left(\Omega\right) \to V_h$

- $V_h = V_1 + \ldots + V_m$, $v \in V_h \Rightarrow$
- $v = \sum_k v_k, \ \ v_k = \Pi_h\left(\theta_k v\right)$
- $K_0 = C\left(1 + \delta^{-2}\right)$
- $K_1 \leq \rho\left(\mathcal{E}\right)$ independent on m

$\text{cond}\left(G_A A\right) \leq C\left(1 + \delta^{-2}\right),$ etc.

Two-level domain decomposition



$\Omega = \Omega_1 \cup \ldots \cup \Omega_m$

$V_h = V_0 + V_1 + \ldots + V_m$

$V_0 = \{v \in V : \; v|_T \in P_1 \, \forall T \in \mathcal{T}_H\}$

$H \leq \mathbf{H}$, qualitative analysis $H = \mathbf{H}$

$v \in V_h \;\Rightarrow$ ex. $\quad v_0 = Qv \in V_0$

(1) $|\, v_0 \,|_{H^1(\Omega)} \leq c_1 \,|\, v \,|_{H^1(\Omega)}$

(2) $\|\, v - v_0 \,\|_{L_2(\Omega)} \leq c_2 H \,|\, v \,|_{H^1(\Omega)}$

Properties (1), (2) are valid if Q is L_2- orthogonal projection onto V_0 (Bramble, Xu 1991)

- $v \in V \;\Rightarrow\; v = v_0 + v_1 + \ldots + v_m$
- $v_0 = Qv, \quad v_k = \Pi_h\left(\theta_k(v - v_0)\right)$
- $K_0 = C\left(1 + \delta^{-2} H^2\right)$
- $K_1 \leq 2(1 + \rho(\mathcal{E}))$

$\mathrm{cond}\,(G_A A) \leq C\left(1 + \delta^{-2} H^2\right)$, etc.

Overlapping DD preconditioners: A model problem

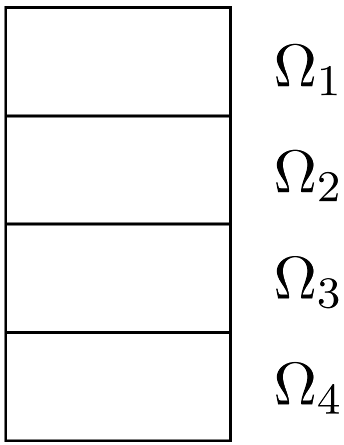
Model problem:

$\Omega = \langle 0,2 \rangle \times \langle 0,3 \rangle$

$-\triangle u = f$ in Ω

$u = 0$ on $\partial\Omega$

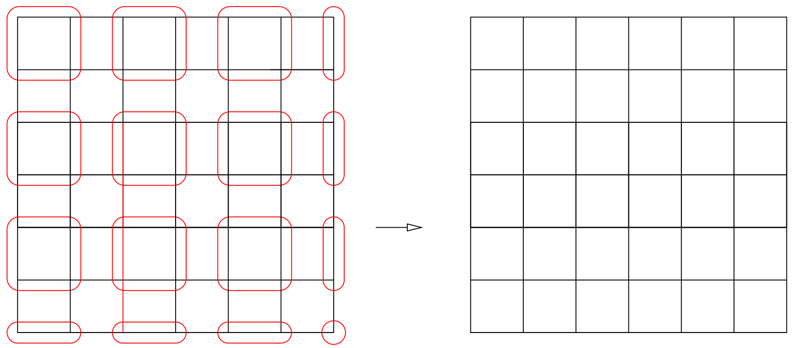
h=1/30, n=5100



	δ:	h	$2h$	$3h$	h	$2h$	$3h$	$3h$	$3h$	$3h$
m	H:	–	–	–	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{4}$
2		19	14	11	9	8	8	8	9	9
4		26	19	16	9	8	8	8	10	10
6		32	23	19	9	9	9	9	10	10
8		36	26	22	9	9	9	9	10	10
10		40	28	23	10	9	9	9	10	10
12		43	31	25	9	9	9	9	10	10

Numbers of iterations for $\varepsilon = 10^{-4}$ with additive precond., overlap $\delta = h$ - Block Jacobi, $\delta > h$ Schwarz

Two-level domain decomposition with aggregations



$\Omega = \Omega_1 \cup \ldots \cup \Omega_m$

$V_h = \text{span } \left\{\phi_i^h\right\}_{i=1}^n$

$\{1, \ldots, n\} = J_1 \cup \ldots \cup J_N$, disjoint

$V_0 = \text{span}\left\{\psi_i\right\}, \ \ \psi_i = \sum\limits_{j \in J_i} \phi_j^h$

$c_1 H^d \leq |\text{supp}\psi_i| \leq c_2 H^d$, $H \leq \{\mathbf{H}, kh\}$

$Q : V_h \to V_0, \quad Qv = \sum \alpha_k\left(v\right)\psi_k$

$\alpha_k = \frac{1}{\mu(\text{supp}\psi_k)} \int\limits_{\text{supp}\psi_k} u\left(x\right)dx$

Then:

- $|\ Qv\ |^2_{H^1(\Omega)} \leq c\frac{H}{h}\ |\ v\ |^2_{H^1(\Omega)}$
- $\|\ v - Qv\ \|_{L_2(\Omega)} \leq\ CH\ |\ v\ |_{H^1(\Omega)}$

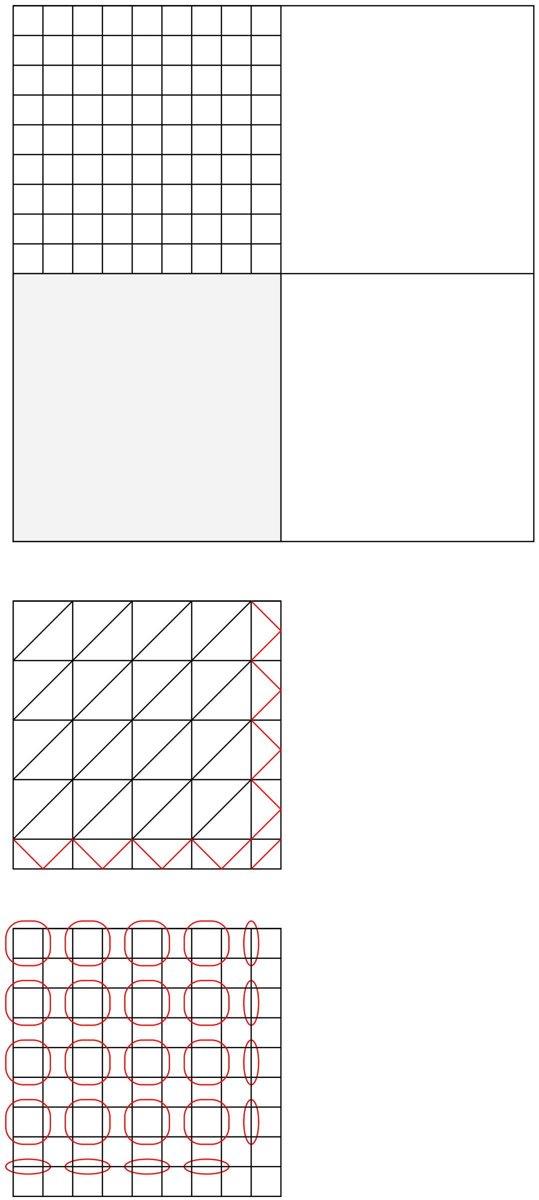
  (weak approximation property)

- $V_h = V_0 + V_1 + \ldots + V_m$
- $v \in V \ \Rightarrow\ v = v_0 + v_1 + \ldots + v_m$
- $v_0 = Qv, \quad v_k = \Pi_h\left(\theta_k(v - v_0)\right)$
- $K_0 = C\left(1 + h^{-1}H + \delta^{-2}H^2\right)$
- $K_1 \leq 2(1 + \rho(\mathcal{E}))$

$\text{cond}\left(G_A A\right) \leq K_0 K_1$, etc.

Blaheta 1986,1989, Braess 1994, Vaněk, Mandel, Brezina 1996, Brezina 1997, Jenkins et al. 2001

Two-level DD with interfaces on coarse grid



$$\bar{\Omega} = \bar{\Omega}_1 \cup \ldots \cup \bar{\Omega}_m, \quad \Omega_k \cap \Omega_l = \emptyset \text{ for } k \neq l$$

$$\Omega_k \to V_k : \varepsilon_{kl} = \cos\left(v_k, v_l\right)_A = 0 \text{ for } k \neq l$$

$$V_1 + \ldots + V_m = W \neq V_h$$

$V_0 \ldots$ by coarse triangulation or by aggregations with missing DOF $\Rightarrow$ $V_h = V_0 + W$,

$R_0 : V_h \to V_0$ interpolation to nodes of $\mathcal{T}_H$ or selected nodes from J_k

$R_k : V_h \to V_k,\ R_k v(x) = v(x)$ for nodes from Ω_k

$$v \in V_h \to v = v_0 + v_1 + \ldots + v_m$$

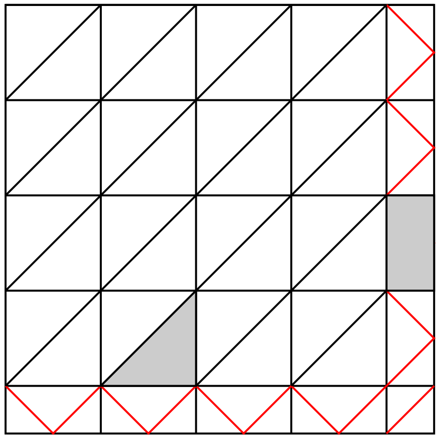
$$v_0 = R_0 v,\ v_k = R_k\left(v - v_0\right)$$

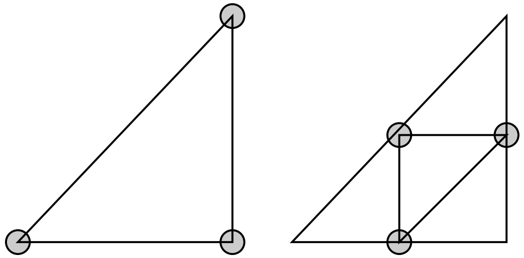
$$V_h = V_0 \oplus W_0,\ W_0 \subset W,\ W_0 = \sum R_k (I - R_0) V_h$$

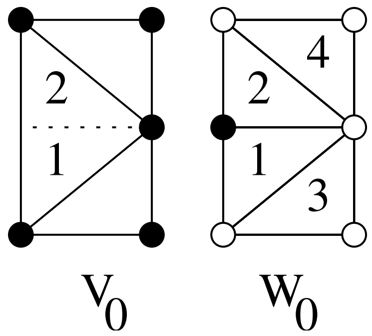
$$\gamma = \cos\left(V_0, W_0\right)_A$$

$$K_0 = 1/\left(1 - \gamma\right),\ K_1 = 1 + \gamma$$

CBS constant - boundary macroelements







$$\gamma = \sup\left\{ \frac{a(v,w)}{\sqrt{a(v,v)}\sqrt{a(w,w)}} : v \in V_0, w \in W_0,\ v, w \neq 0 \right\}$$

$$a\,(v,w) = \sum_E a_E\,(v,w) = \sum_E \int_E \langle D\,\nabla v,\ \nabla w\rangle\, dx$$

$\gamma = \max \gamma_E$, E - inner/interface macroelement

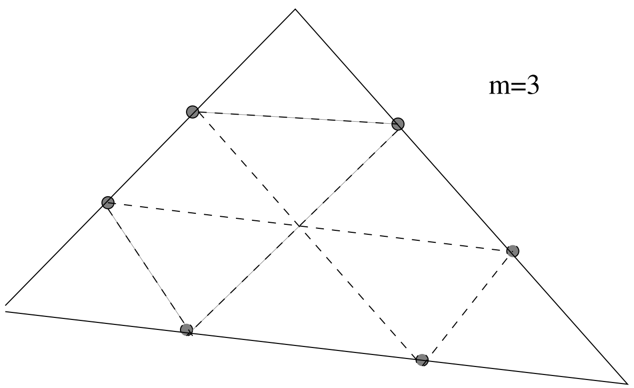
Interface: $v \in V_0(E) \rightarrow \nabla v = (\delta_x,\ \delta_y)$ in $T_1,\ T_2$.
$w \in W_0(E) \rightarrow \nabla w = (d_x,\ d_y)$ in T_1,
$\nabla w = (d_x,\ -d_y)$ in T_2, $\nabla w = 0$ in T_3, T_4. If $D = \mathrm{diag}[k_x,\ k_y]$, then

$$\left.\begin{aligned} a_E\,(v,w) &= 2k_x\delta_x\, d_x\,\Delta \\ a_E\,(v) &\leq 2k_x\delta_x^2\,\Delta \\ a_E\,(w) &\leq 2(k_x d_x^2 + k_y d_x^2)\Delta \end{aligned}\right\} \Rightarrow \gamma_E \leq \sqrt{\frac{k_x}{k_x+k_y}}.$$

CBS constant - inner macroelements

Maitre, Musy 1981:

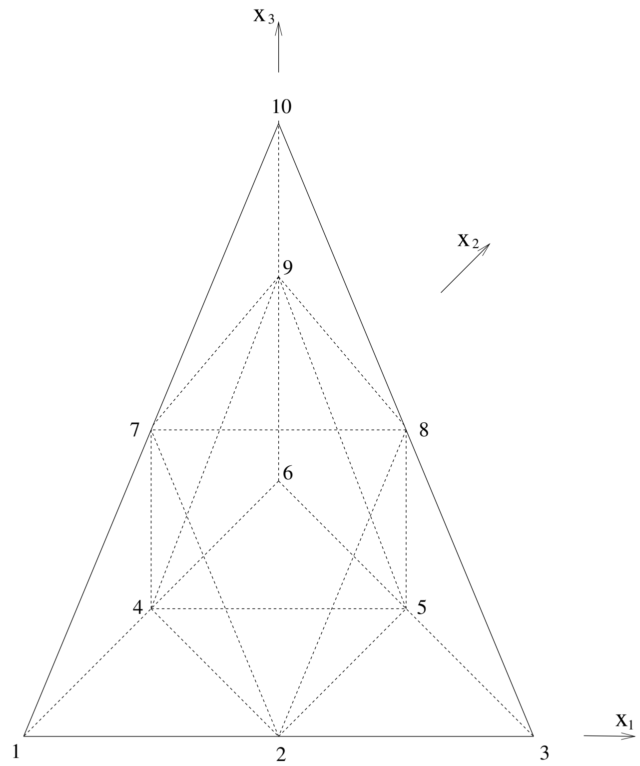
$\gamma = \sqrt{1/2}$ for isotropic Laplacian and rectangular elements, $m_p = 2$.



- m_p fold refinement!
- anisotropic Laplacian
- general elasticity $c_{i,j,k,l}$!
- arbitrary element shape

$\gamma \leq \sqrt{\frac{m_p^2-1}{m_p^2}}$ Axelsson, RB 2001

m_p^3 tetrahedra



$\gamma \leq \sqrt{1 - \frac{2}{m_p^4+m_p^2}}$

RB NLAA 2003

Nonlinear and Nonsymmetric Preconditioners

- For A_k corresponding to subdomain, we successfuly replace A_k by incomplete factorization.
- For coarse grid subproblem A_0, the same strategy is inefficient. More accurate approximation to A_0 is necessary, e.g. by inner PCG iterations. Then the pseudoresidual is $g = G(r)$ but G is *not linear*.
- It can be difficult to construct A_0 and solve this subproblem in a time comparable with the solution of subdomain problems A_k. Then it may be advantageous to use hybrid aditive-multiplicative algorithm:
  - coarse grid correction is computed individually,
  - residual is updated and subdomain problems are solved in parallel.

The hybrid algorithm gives $g = G(r)$ with G linear (for linear subproblem solvers) but not symmetric. The symmetrization is relative expensive and not necessary, an alternative is *nonsymmetric preconditioner* with GPCG.

Inner iterations GPCG[s] method

given $u^0 \rightarrow r^0 = b - Au^0,\ g^0 = G(r^0),\ v^0 = g^0$

for $i = 0, 1, \ldots$ **until** $\| r^i \| \leq \varepsilon \| b \|$ **do**

$$w^i = Av^i$$

$$\alpha_i = \sigma_i / \langle w^i, v^i \rangle$$

$$u^{i+1} = u^i + \alpha_i v^i$$

$$r^{i+1} = r^i - \alpha_i w^i$$

$$v^{i+1} = g^{i+1} = G(r^{i+1})$$

for $k = 1, \ldots, \min\{i+1, s\}$ **do**

$$\beta_{i+1}^{(k)} = (\langle g^{i+1}, r^{i+2-k} \rangle - \langle g^{i+1}, r^{i+1-k} \rangle)/\sigma_{i+1-k}$$

$$v^{i+1} = v^{i+1} + \beta_{i+1}^{(k)} v^{i+1-k}$$

end $\Rightarrow \sigma_{i+1}$

end

Extra:

$s-1\times$ vector storage

$r^i,\ r^{i-1}, \ldots, r^{i+2-s}(r^0)$

$s\times$ inner products

$\langle g^{i+1}, r^{i+1-k} \rangle$,

$k = 1, \ldots, \min\{i+1, s\}$

For s=1: one extra inner product

DD preconditioners: A model problem

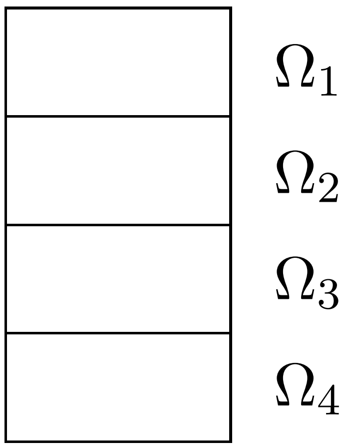
Model problem:

$\Omega = \langle 0, 2 \rangle \times \langle 0, 3 \rangle$

$- \triangle u = f$ in Ω

$u = 0$ on $\partial\Omega$

h=1/30, n=5100



Overlap 2h, #subdomains:	2	4	8	12	16	24
no coarse grid	12	16	22	23	31	37
c-grid H=3h, AP	7	7	8	8	7	8
c-grid H=3h, HP	6	6	6	6	6	6
aggreg. 2h, AP	12	13	15	16	17	17
aggreg. 2h, HP	9	10	11	11	11	11
interface & aggreg. 2h, AP	13	14	14	14	14	14
interface & aggreg. 2h, HP	7	8	7	8	8	8

Numbers of iterations for $\varepsilon = 10^{-3}$. AP=additive preconditioner, HP=hybrid preconditioner + GPCG[1]

DD preconditioners: Another model problem

$-\triangle u = 2\pi^2 \sin(\pi x)\sin(\pi y)$ in $\Omega = \langle 0, 1\rangle^2$ and $u = 0$ on $\partial\Omega$

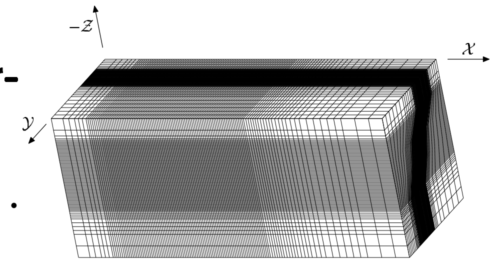
Fine grid	p=4		p=16		p=64	
	$\mathbf{B}^{as}$	$\mathbf{B}^{ms}$	$\mathbf{B}^{as}$	$\mathbf{B}^{ms}$	$\mathbf{B}^{as}$	$\mathbf{B}^{ms}$
$h = 1/48$	37	20	110	58	377	193
$h = 1/96$	41	22	124	65	423	221
$h = 1/192$	45	24	138	72	478	248

Table 1: overlap $\beta = 1/6$, accuracy $\varepsilon = 10^{-4}$, (see Knut-Andreas Lie, Uni. Oslo, 2001)

Fine grid	p=4		p=16		p=64	
	$\mathbf{B}_H^{as}$	$\mathbf{B}_H^{ms}$	$\mathbf{B}_H^{as}$	$\mathbf{B}_H^{ms}$	$\mathbf{B}_H^{as}$	$\mathbf{B}_H^{ms}$
$h = 1/48$	9	6	12	5	7	4
$h = 1/96$	10	6	13	5	9	5
$h = 1/192$	11	7	15	6	10	5

Table 2:
two-level Schwarz preconditioners

AS preconditioners - a geotechnical problem.



Overlap 2h. The sub-solvers: subdomain = incomplete factorization, the aggregated problem = inner PCG with $\varepsilon_0 = 10^{-1}$.

Left: #iterations. Right: times [s] on THEA.

# subd.	one–level method	two–level method 3x3x3	two–level method 6x6x6
2	92	45	56
3	102	47	60
4	110	51	64
6	121	55	70
7	125	57	72
8	128	–	–

one-level method	two–level method 3x3x3	two–level method 6x6x6
386	267	242
289	241	175
242	240	145
190	244	115
170	265	111
161	–	–

Robustness of DD preconditioners: anisotropy

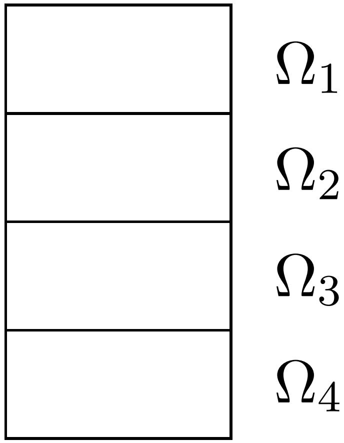
Model problem:

$\Omega = \langle 0,2\rangle \times \langle 0,3\rangle$

$-k_x \frac{\partial^2 u}{\partial x^2} - k_y \frac{\partial^2 u}{\partial y^2} = f$ in Ω

$u = 0$ on $\partial\Omega$

h=1/30, n=5100



δ 2h, #subd's:	4			8			16		
no coarse grid	16	$\frac{8}{27}$	$\frac{4}{26}$	22	$\frac{10}{37}$	$\frac{4}{38}$	31	$\frac{14}{52}$	$\frac{5}{61}$
c-grid H=3h, A	7	$\frac{7}{10}$	$\frac{7}{16}$	8	$\frac{7}{10}$	$\frac{7}{20}$	7	$\frac{7}{11}$	$\frac{7}{26}$
c-grid H=3h, H	6	$\frac{6}{8}$	$\frac{6}{13}$	6	$\frac{6}{8}$	$\frac{7}{16}$	6	$\frac{6}{10}$	$\frac{7}{22}$
aggreg. 2h, A	13	$\frac{10}{16}$	$\frac{8}{19}$	15	$\frac{12}{16}$	$\frac{8}{23}$	17	$\frac{13}{19}$	$\frac{9}{30}$
aggreg. 2h, H	10	$\frac{8}{11}$	$\frac{7}{15}$	11	$\frac{9}{12}$	$\frac{7}{19}$	11	$\frac{9}{13}$	$\frac{8}{25}$
if. & agg. 2h, A	14	$\frac{11}{14}$	$\frac{8}{22}$	14	$\frac{12}{16}$	$\frac{8}{30}$	14	$\frac{12}{19}$	$\frac{9}{40}$
if. & agg. 2h, H	8	$\frac{6}{8}$	$\frac{4}{14}$	7	$\frac{6}{9}$	$\frac{5}{16}$	8	$\frac{7}{10}$	$\frac{5}{22}$

Numbers of iterations for $\varepsilon = 10^{-3}$. A=additive preconditioner, H=hybrid preconditioner + GPCG[1].

Columns: (1)= isotropy, (2)=$\frac{k_x/k_y=10}{k_x/k_y=0.1}$, (3)=$\frac{k_x/k_y=100}{k_x/k_y=0.01}$.

DD preconditioners: A model parabolic problem

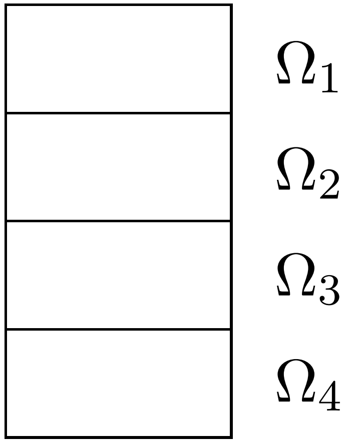
Model problem:

$\Omega = \langle 0, 2\rangle \times \langle 0, 3\rangle$

$\frac{\partial u}{\partial t} - \triangle u = f$

$u = 0$ on $\partial\Omega$

h=1/30, n=5100



Overlap 2h, #subd's:	4			8			16		
no coarse grid	16	8	3	22	9	3	31	13	3
c-grid H=3h, AP	7	8	4	8	8	4	7	10	5
c-grid H=3h, HP	6	8	3	6	8		6	10	4
aggreg. 2h, AP	13	8	5	15	8	5	17	10	5
aggreg. 2h, HP	10	6		11	8		11	8	
interf. & aggr. 2h, AP	14	9	5	14	10	5	14	10	5
interf. & aggr. 2h, HP	8	5	3	7	5	3	8	5	3

Numbers of iterations for $\varepsilon = 10^{-3}$. AP=additive preconditioner, HP=hybrid preconditioner + GPCG[1]. Column 1: matrix K, columns 2,3: matrix $M + \xi K$ with $\xi = h$, h^2, respectively.

Overlapping DD methods

:-) overlapping DD is an efficient tool for data decomposition, building preconditioners and construction of parallel algorithms with a small amount of communications

:-) the efficiency can be substantially increased by adding a rough global problem, which can be defined by coarse grid or aggregations

:-) there is a variety of DD methods: overlap, nested/non-nested coarse grid, simple/smoothed aggregations, RAS and RASHO, etc.

:-) new CGR with interface: robustness w.r.t. coefficient jumps between macroelements, no communication between subdomain problems in preconditioning, efficient hybrid version, clear quantitative analysis

:-) inexact sub-solvers, automatic partition/aggreg. (not PDE systems)

:-(increase of subproblems due to overlap, decrease of efficiency due to anisotropy

Final remarks

- overlapping DD - can be applied to other classes of problems (nonsymmetric, parabolic, saddle point etc), are easy to implement, potentialy fully algebraical (black box),
- there are also other classes of nonoverlapping DD methods handling the interface through Schur complement or Lagrangian multipliers (Neumann-Neumann, FETI),
- there are many possible decompositions, which can be used for efficient parallel solvers: beside DD, also composite grid FEM, HB decomposition, AMLI, DiD etc.
- it is possible to combine different decompositions and different (additive/multiplicative) algorithms,
- the two-level concept can be developed to multi-level one for better balance, better efficiency and development of optimal solvers.

Thank you for your attention.