

Deflated Krylov Acceleration of the Schwarz Domain Decomposition Method

Kees Vuik, A. Segal and F.J. Vermolen

`c.vuik@math.tudelft.nl`

`http://ta.twi.tudelft.nl/users/vuik/`

Delft University of Technology

GAMM Workshop Numerical Linear Algebra

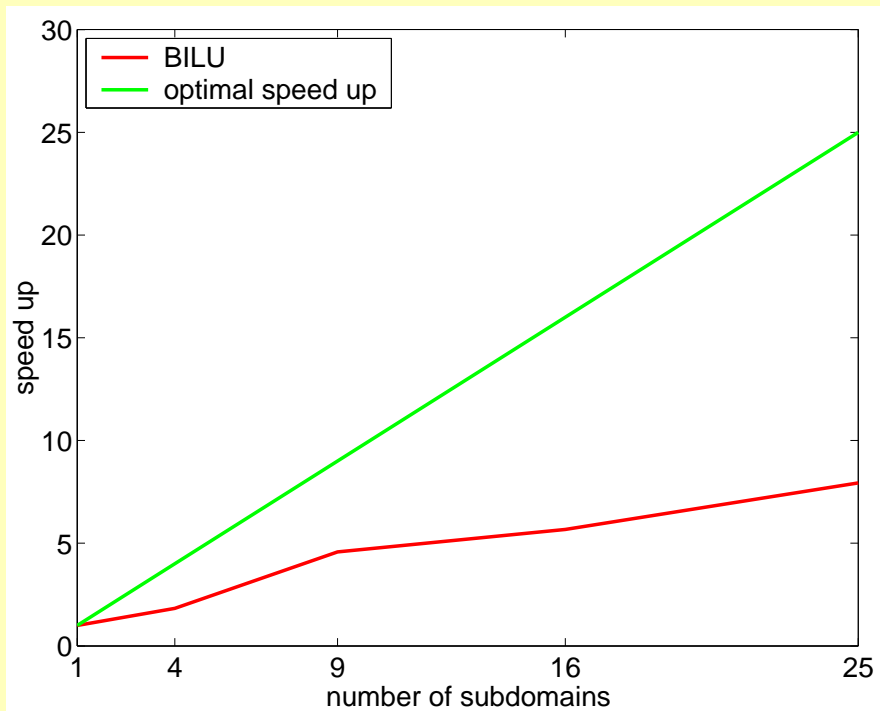
University of Bielefeld, Germany, September 13-14, 2002

Contents

1. Introduction
2. Deflated Krylov methods
3. Deflation compared with Coarse Grid Correction
4. Convergence of DICCG
5. Vertex centered approach
6. Conclusions

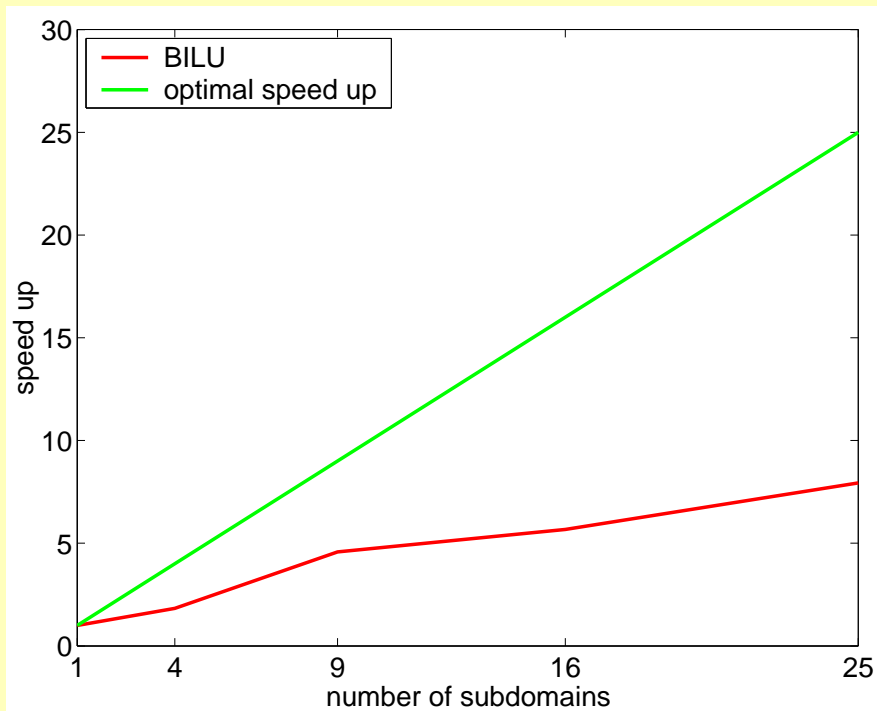
1. Introduction

Block ILU

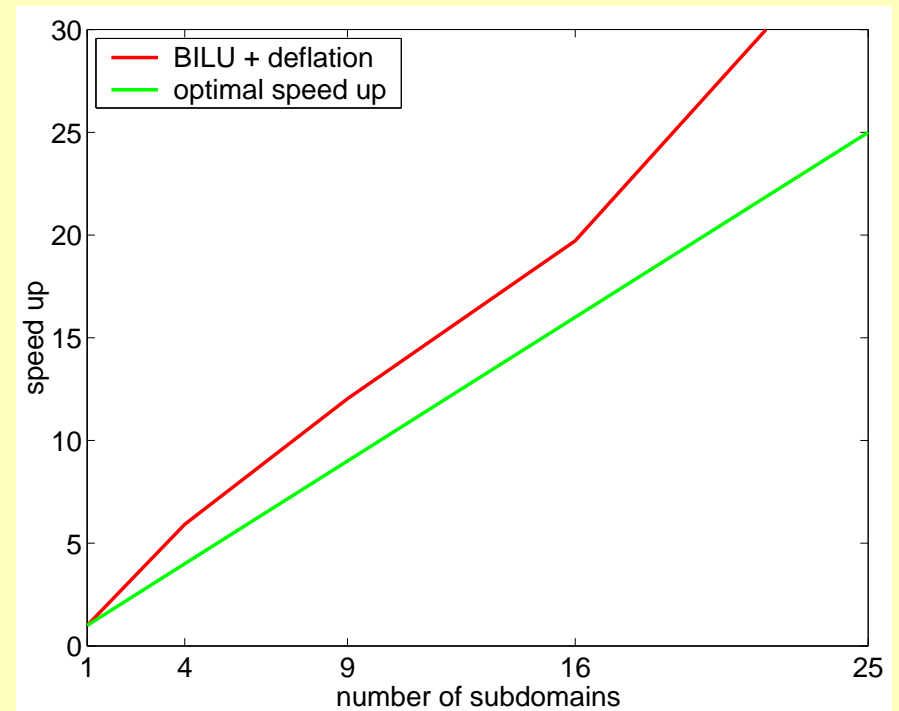


1. Introduction

Block ILU

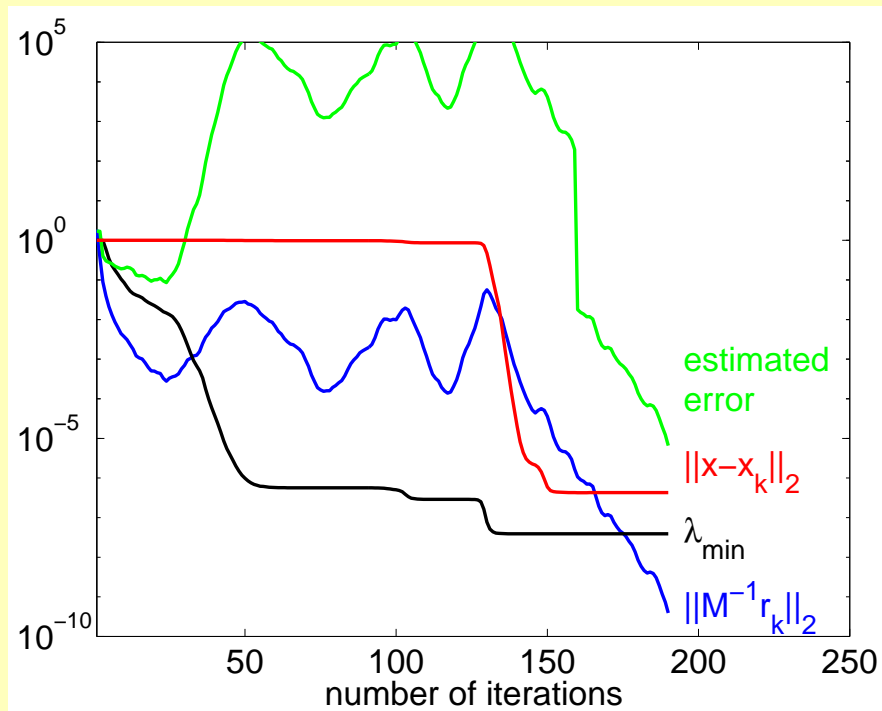


Block ILU with deflation



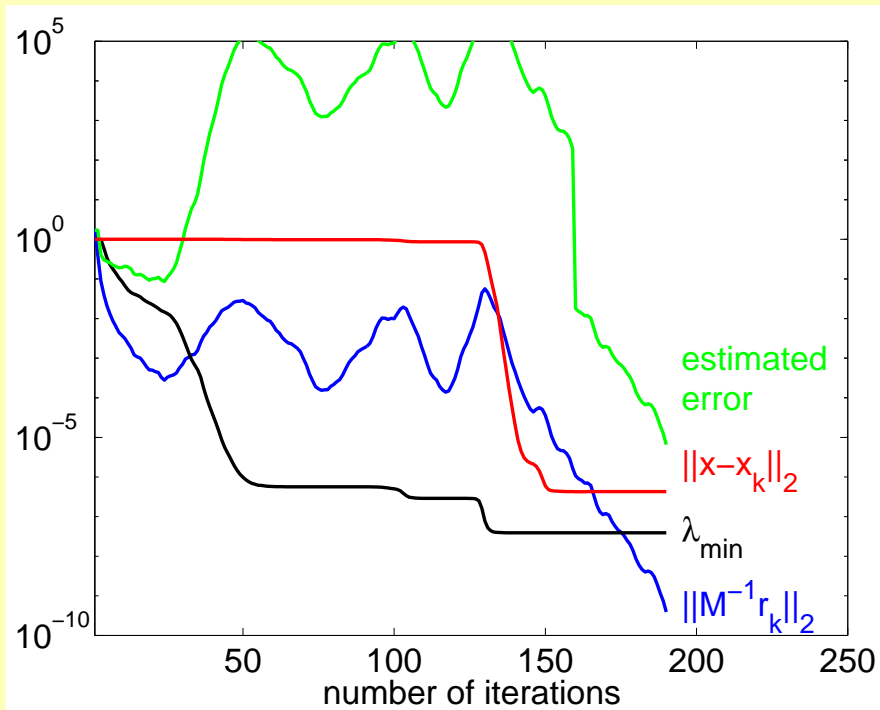
Results oil flow problem

ICCG

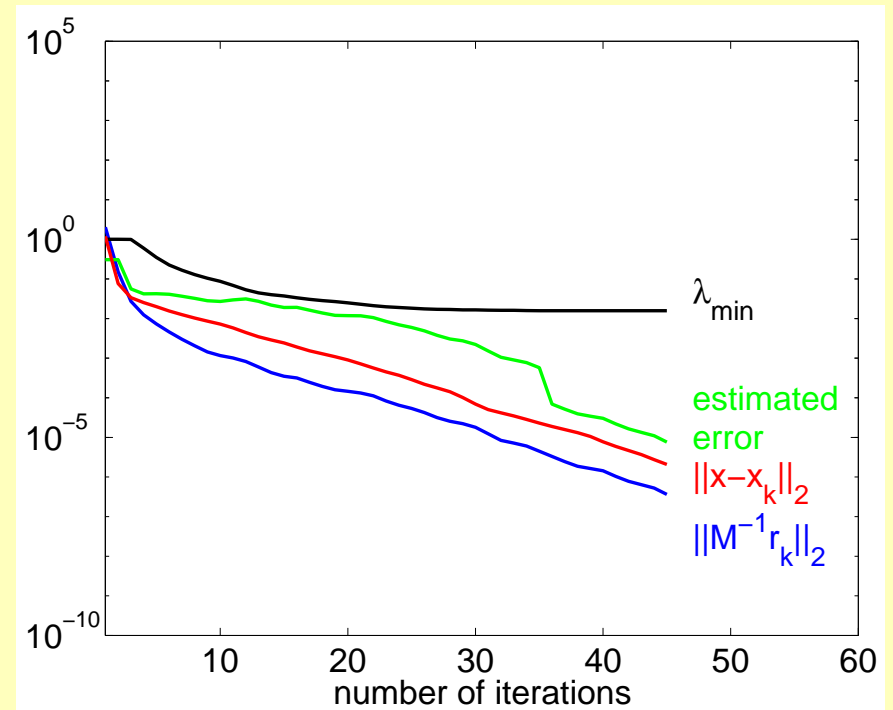


Results oil flow problem

ICCG



Deflated ICCG



Literature review

- Robust preconditioners
(M)ICCG vd Vorst, Meijering, Gustafsson
ILUT Saad, MRILU Ploeg, Wubs
Navier-Stokes Elman, Silvester, Wathen, Golub
RIF Benzi, Tuma

Literature review

- Robust preconditioners
(M)ICCG [vd Vorst](#), [Meijering](#), [Gustafsson](#)
ILUT [Saad](#), MRILU [Ploeg](#), [Wubs](#)
Navier-Stokes [Elman](#), [Silvester](#), [Wathen](#), [Golub](#)
RIF [Benzi](#), [Tuma](#)
- Parallel preconditioners
Block variants [see above](#)
ILU [Bastian](#), [Horton](#), [Vuik](#), [Nooyen](#), [Wesseling](#)
SPAI [Grote](#), [Huckle](#), [Benzi](#), [Tuma](#), [Chow](#), [Saad](#)

Literature review

- Robust preconditioners
(M)ICCG vd Vorst, Meijering, Gustafsson
ILUT Saad, MRILU Ploeg, Wubs
Navier-Stokes Elman, Silvester, Wathen, Golub
RIF Benzi, Tuma
- Parallel preconditioners
Block variants see above
ILU Bastian, Horton, Vuik, Nooyen, Wesseling
SPAI Grote, Huckle, Benzi, Tuma, Chow, Saad
- Acceleration of parallel preconditioners
CGC Notay, vd Velde, Benzi, Frommer, Nabben, Szyld, Chan,
Mathew, Dryja, Widlund, Padiy, Axelsson, Polman
Deflation Nicolaidis, Mansfield, Frank, Vuik
Morgan, Chapman, Saad, Burrage, Ehrel, Pohl
FETI Farhat, Roux, Mandel, Klawonn, Widlund

2. Deflated Krylov methods

A is SPD, Conjugate Gradients

$$P = I - AZE^{-1}Z^T \text{ with } E = Z^T AZ$$

and $Z = [z_1 \dots z_m]$, where z_1, \dots, z_m are independent deflation vectors.

Properties

1. $P^T Z = 0$ and $PAZ = 0$
2. $P^2 = P$
3. $AP^T = PA$

Deflated ICCG

$$x = (I - P^T)x + P^T x,$$

Deflated ICCG

$$x = (I - P^T)x + P^T x,$$

$$(I - P^T)x = ZE^{-1}Z^T Ax = ZE^{-1}Z^T b,$$

Deflated ICCG

$$x = (I - P^T)x + P^T x,$$

$$(I - P^T)x = ZE^{-1}Z^T Ax = ZE^{-1}Z^T b, \quad AP^T x = PAx = Pb.$$

Deflated ICCG

$$x = (I - P^T)x + P^T x,$$

$$(I - P^T)x = ZE^{-1}Z^T Ax = ZE^{-1}Z^T b,$$

$$AP^T x = PAx = Pb.$$

DICCG

$$k = 0, \hat{r}_0 = Pr_0, p_1 = z_1 = L^{-T}L^{-1}\hat{r}_0;$$

while $\|\hat{r}_k\|_2 > \varepsilon$ **do**

$$k = k + 1;$$

$$\alpha_k = \frac{(\hat{r}_{k-1}, z_{k-1})}{(p_k, PAp_k)};$$

$$x_k = x_{k-1} + \alpha_k p_k;$$

$$\hat{r}_k = \hat{r}_{k-1} - \alpha_k PAp_k;$$

$$z_k = L^{-T}L^{-1}\hat{r}_k;$$

$$\beta_k = \frac{(\hat{r}_k, z_k)}{(\hat{r}_{k-1}, z_{k-1})};$$

$$p_{k+1} = z_k + \beta_k p_k;$$

end while

Deflation for non-symmetric matrices

$$P = I - AZE^{-1}Y^T \text{ with } E = Y^T AZ$$
$$Q = I - ZE^{-1}Y^T A$$

and $Z = [z_1 \dots z_m]$, $Y = [y_1 \dots y_m]$ where z_1, \dots, z_m and y_1, \dots, y_m are independent sets of deflation vectors.

Properties

1. $PAZ = Y^T P = 0$ and $Y^T A Q = QZ = 0$
2. $P^2 = P$ and $Q^2 = Q$
3. $PA = A Q$

Deflated Krylov methods

$$x = (I - Q)x + Qx,$$

Deflated Krylov methods

$$x = (I - Q)x + Qx,$$

$$(I - Q)x = ZE^{-1}Y^T Ax = ZE^{-1}Y^T b,$$

Deflated Krylov methods

$$x = (I - Q)x + Qx,$$

$$(I - Q)x = ZE^{-1}Y^T Ax = ZE^{-1}Y^T b, \quad AQx = PAx = Pb.$$

Deflated Krylov methods

$$x = (I - Q)x + Qx,$$

$$(I - Q)x = ZE^{-1}Y^T Ax = ZE^{-1}Y^T b, \quad AQx = PAx = Pb.$$

Preconditioning

$$K^{-1}PA\tilde{x} = K^{-1}Pb, \quad Qx = Q\tilde{x}$$

$$PAK^{-1}\tilde{y} = Pb, \quad Qx = QK^{-1}\tilde{y}$$

Deflated Krylov methods

$$x = (I - Q)x + Qx,$$

$$(I - Q)x = ZE^{-1}Y^T Ax = ZE^{-1}Y^T b, \quad AQx = PAx = Pb.$$

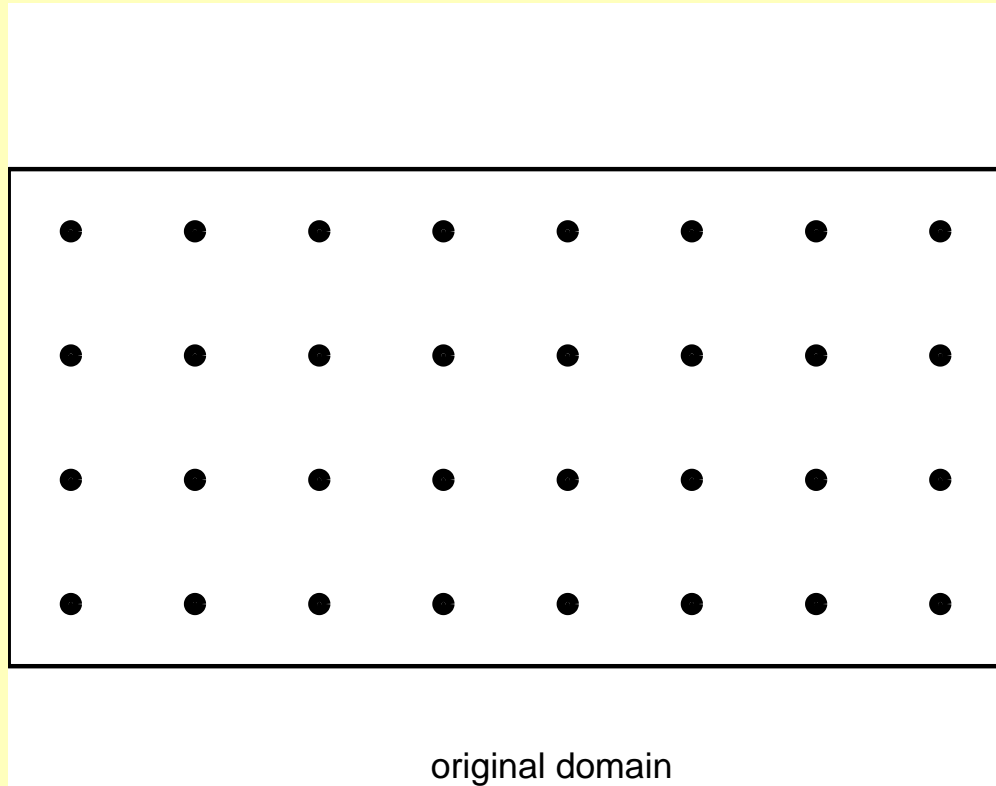
Preconditioning

$$K^{-1}PA\tilde{x} = K^{-1}Pb, \quad Qx = Q\tilde{x}$$

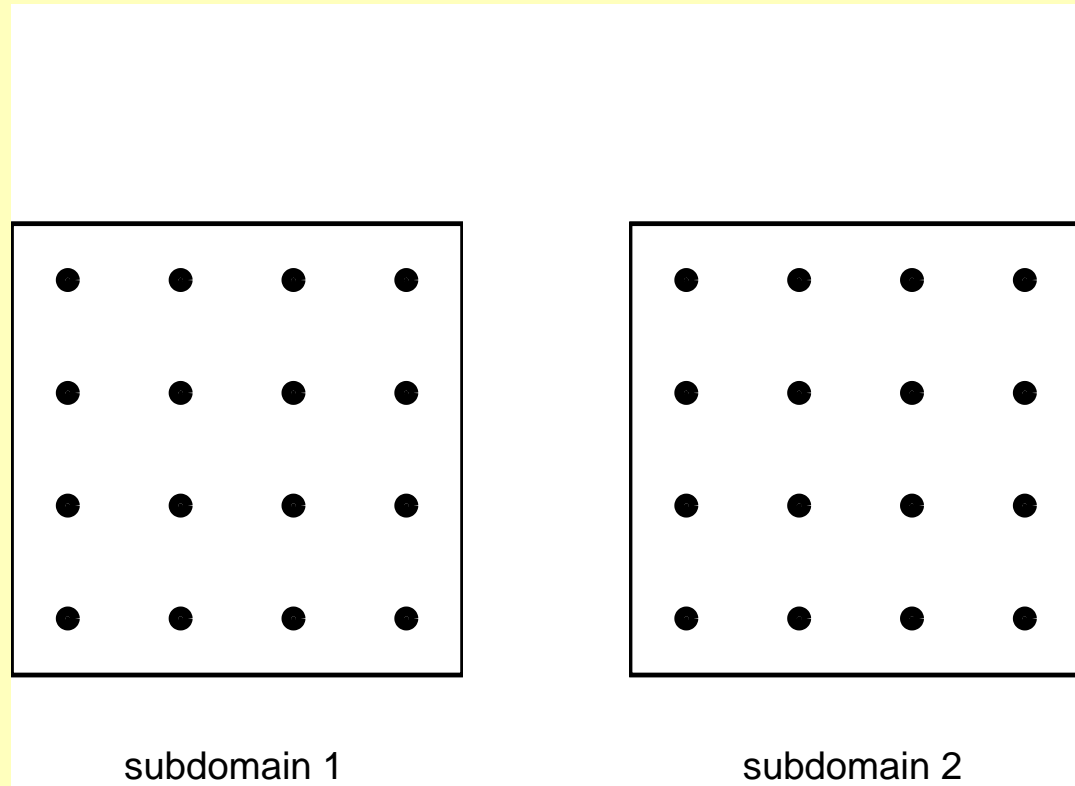
$$PAK^{-1}\tilde{y} = Pb, \quad Qx = QK^{-1}\tilde{y}$$

Systems can be solved by: **GMRES, GCR, Bi-CGSTAB,...**

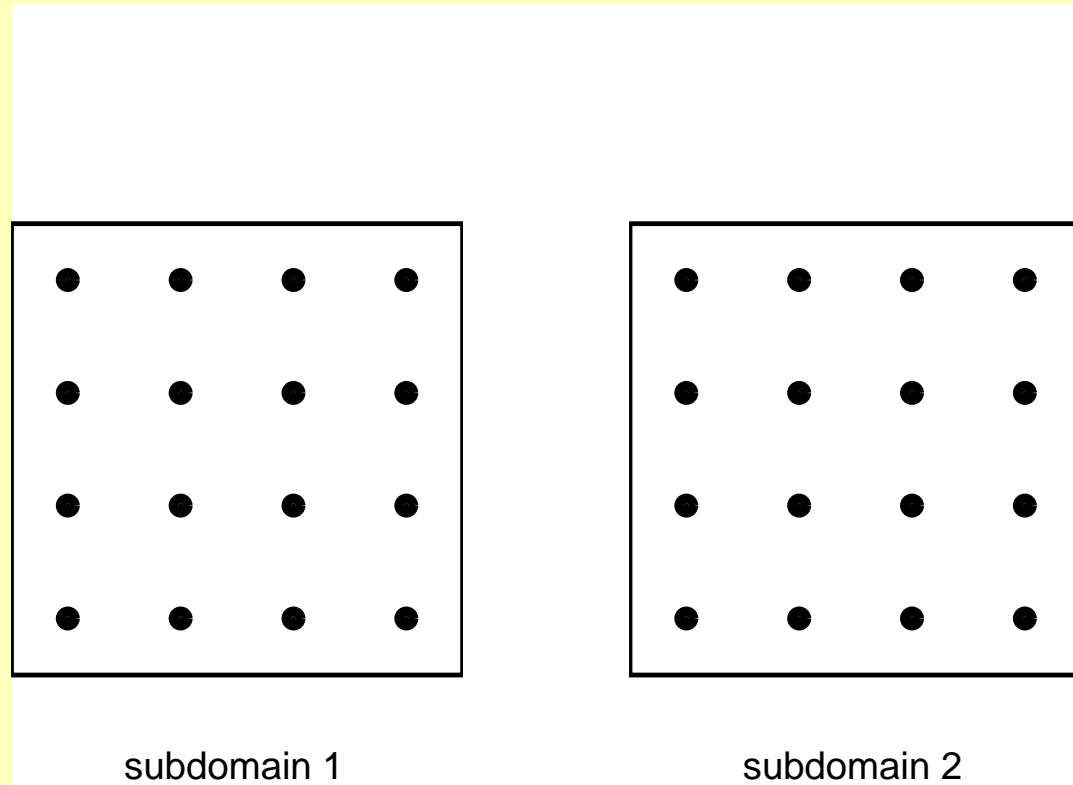
Decomposition of a cell centered domain



Decomposition of a cell centered domain



Decomposition of a cell centered domain



$$\bar{\Omega} = \bigcup_{i=1}^m \bar{\Omega}_i$$

Choice of the deflation vectors

m is number of subdomains z_1, \dots, z_m deflation vectors

- $z_i = 1$ on $\bar{\Omega}_i$
- $z_i = 0$ on $\Omega \setminus \bar{\Omega}_i$

Remarks

- The matrix E is sparse
- $K_{eff}(PA)$ decreases for increasing m
- Work to invert E increases for increasing m
- Optimal value of m ?

3. Comparison of Deflation with Coarse Grid Correction

Definition: $P_D = I - AZE^{-1}Z^T$.

$$x = (I - P_D^T)x + P_D^T x,$$

where $(I - P_D^T)x = ZE^{-1}Z^T b$ and $AP_D^T x = P_D Ax = P_D b$

DICCG

$$k = 0, \hat{r}_0 = P_D r_0, p_1 = z_1 = L^{-T} L^{-1} \hat{r}_0;$$

while $\|\hat{r}_k\|_2 > \varepsilon$ **do**

$$k = k + 1;$$

$$\alpha_k = \frac{(\hat{r}_{k-1}, z_{k-1})}{(p_k, P_D A p_k)};$$

$$x_k = x_{k-1} + \alpha_k p_k;$$

$$\hat{r}_k = \hat{r}_{k-1} - \alpha_k P_D A p_k;$$

$$z_k = L^{-T} L^{-1} \hat{r}_k;$$

$$\beta_k = \frac{(\hat{r}_k, z_k)}{(\hat{r}_{k-1}, z_{k-1})}; \quad p_{k+1} = z_k + \beta_k p_k;$$

end while

Coarse Grid Correction of ICCG

Definition

- $Z \in \mathbb{R}^{n \times m}$ with independent columns.
- $E = Z^T A Z \in \mathbb{R}^{m \times m}$, E is SPD.
- $P_C = L^{-T} L^{-1} + \gamma Z E^{-1} Z^T$.

CICCG

$k = 0$, $r_0 = b - Ax_0$, $p_1 = z_1 = L^{-T} L^{-1} r_0$;
while $\|r_k\|_2 > \varepsilon$ **do**
 $k = k + 1$;
 $\alpha_k = \frac{(r_{k-1}, z_{k-1})}{(p_k, Ap_k)}$;
 $x_k = x_{k-1} + \alpha_k p_k$;
 $r_k = r_{k-1} - \alpha_k A p_k$;
 $z_k = P_C r_k = L^{-T} L^{-1} r_k + \gamma Z E^{-1} Z^T r_k$;
 $\beta_k = \frac{(r_k, z_k)}{(r_{k-1}, z_{k-1})}$; $p_{k+1} = z_k + \beta_k p_k$;

end while

Properties of Deflation and CGC

$$P_D = I - AZE^{-1}Z^T$$

$$P_C = I + ZE^{-1}Z^T$$

Properties of Deflation and CGC

$$P_D = I - AZE^{-1}Z^T$$

$$P_C = I + ZE^{-1}Z^T$$

Properties of P_D

- $P_D A$ is symmetric and positive semidefinite
- P_D is a projection, $P_D A Z = 0$
- since $P_D A$ is singular, a good termination criterion is important

Properties of Deflation and CGC

$$P_D = I - AZE^{-1}Z^T$$

$$P_C = I + ZE^{-1}Z^T$$

Properties of P_D

- $P_D A$ is symmetric and positive semidefinite
- P_D is a projection, $P_D A Z = 0$
- since $P_D A$ is singular, a good termination criterion is important

Properties of P_C

- P_C is symmetric positive definite
- $A^{\frac{1}{2}}(P_C - I)A^{\frac{1}{2}}$ is a projection

Properties of Deflation and CGC

Definition

Eigenpair $\{\lambda_i, v_i\}$, so $Av_i = \lambda_i v_i$ with $0 < \lambda_1 \leq \dots \leq \lambda_n$.

Take $Z = [v_1 \dots v_m]$.

Properties of Deflation and CGC

Definition

Eigenpair $\{\lambda_i, v_i\}$, so $Av_i = \lambda_i v_i$ with $0 < \lambda_1 \leq \dots \leq \lambda_n$.

Take $Z = [v_1 \dots v_m]$.

Theorem

- the spectrum of $P_D A$ is $\{0, \dots, 0, \lambda_{m+1}, \dots, \lambda_n\}$

- the spectrum of $P_C A$ is $\{1 + \lambda_1, \dots, 1 + \lambda_m, \lambda_{m+1}, \dots, \lambda_n\}$

Properties of Deflation and CGC

Definition

Eigenpair $\{\lambda_i, v_i\}$, so $Av_i = \lambda_i v_i$ with $0 < \lambda_1 \leq \dots \leq \lambda_n$.

Take $Z = [v_1 \dots v_m]$.

Theorem

- the spectrum of $P_D A$ is $\{0, \dots, 0, \lambda_{m+1}, \dots, \lambda_n\}$

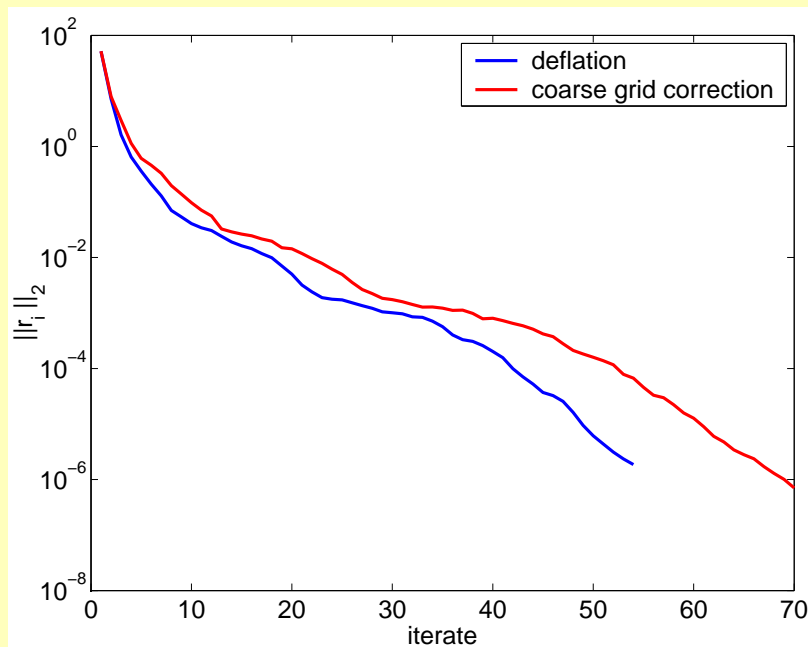
- the spectrum of $P_C A$ is $\{1 + \lambda_1, \dots, 1 + \lambda_m, \lambda_{m+1}, \dots, \lambda_n\}$

Corollary

DICCG converges **faster** than CICCG.

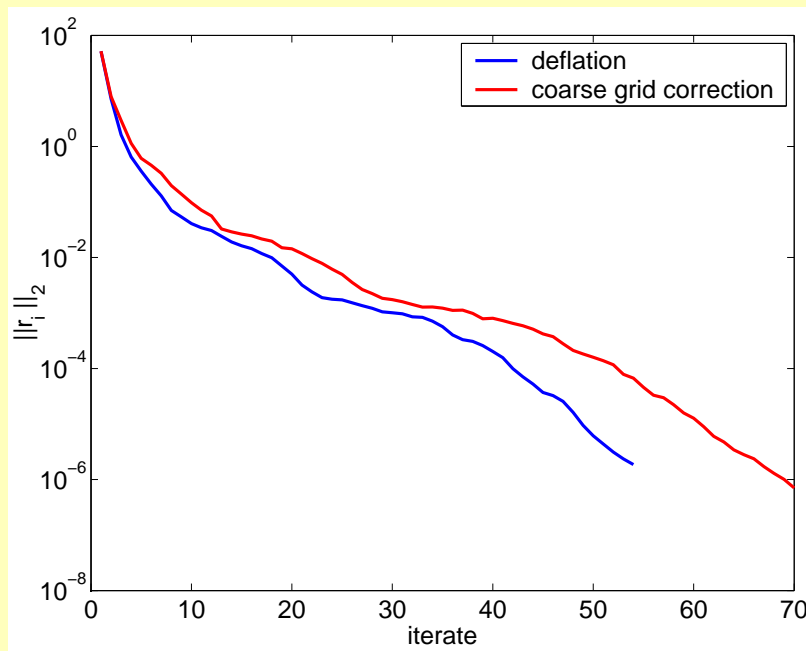
Experiments with Deflation and CGC

Residual with Block IC

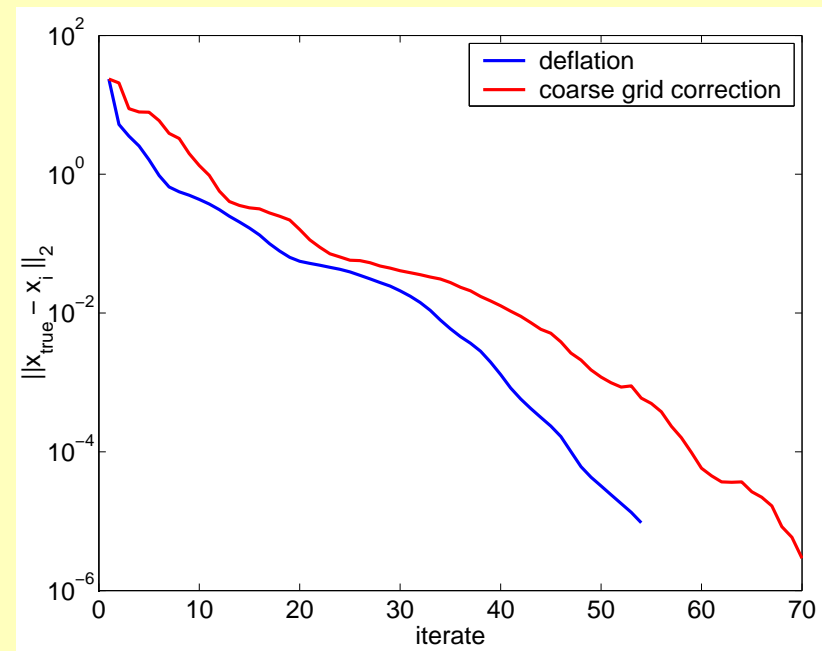


Experiments with Deflation and CGC

Residual with Block IC

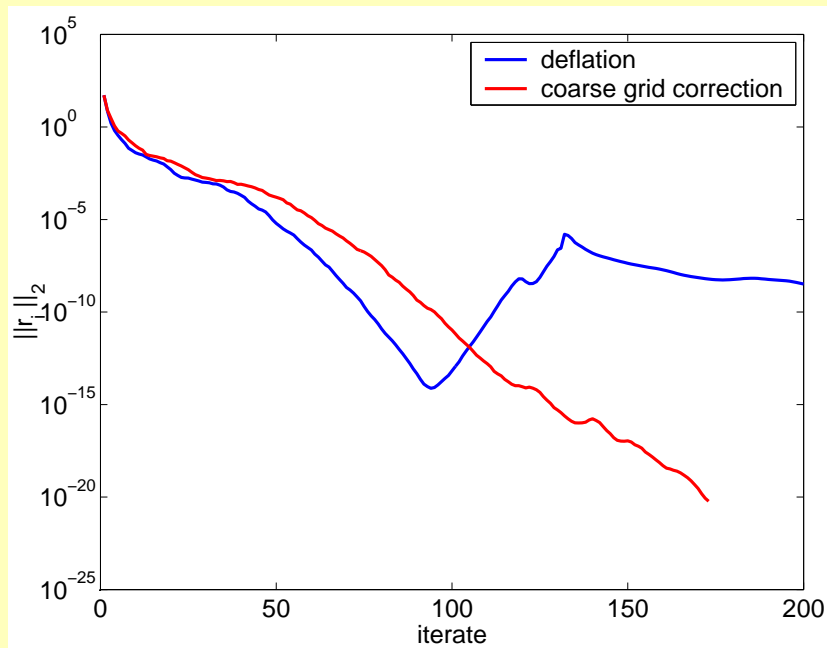


Error with Block IC



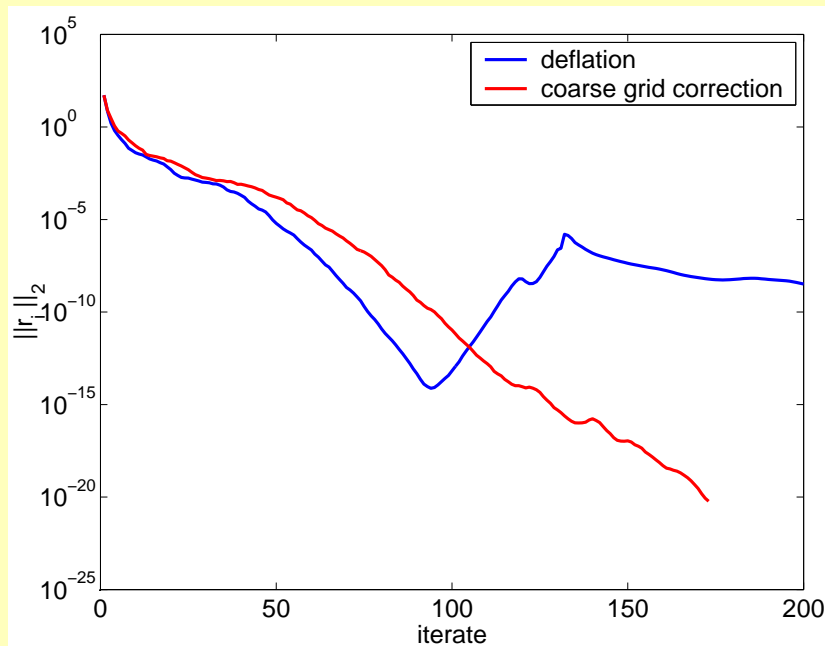
Too severe termination criterion

Residual with Block IC

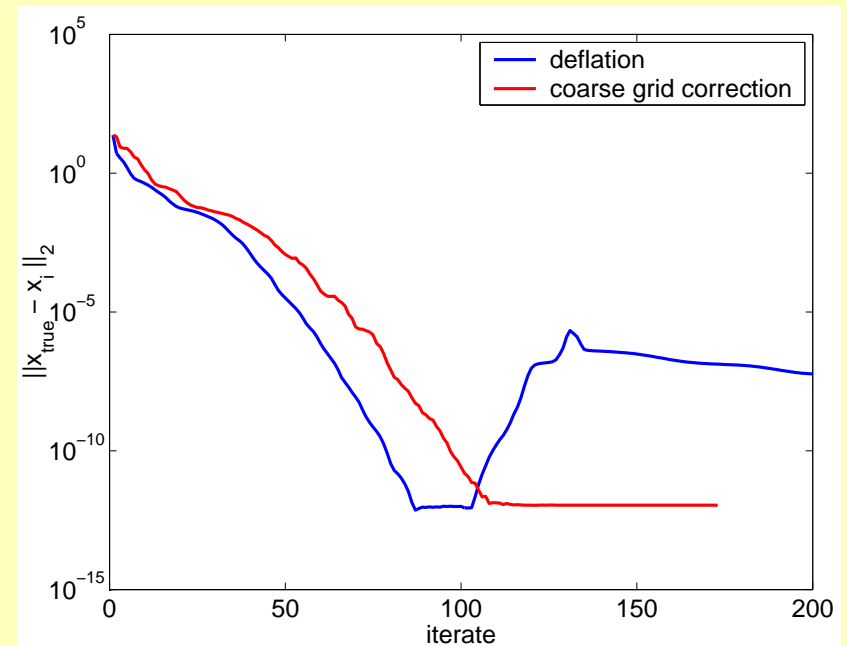


Too severe termination criterion

Residual with Block IC



Error with Block IC



4. Convergence of DICCG

Theorem

Let A be SPD, and A^* , $A - A^*$ be SPSD, and $\text{span}Z = \text{null}(A^*)$
Furthermore preconditioner K is SPD and $K = LL^T$ then

$$K_{eff}(L^{-1}PAL^{-T}) \leq \frac{\lambda_n(L^{-1}AL^{-T})}{\lambda_{m+1}(L^{-1}A^*L^{-T})}.$$

1. Removing the smallest eigenvalues from the spectrum leads to the greatest improvement for PDE problems.
2. A good preconditioner for A^* may be attractive (Kaasschieter).
3. A preconditioner for A^* may increase the largest eigenvalue of $L^{-1}AL^{-T}$.

Symmetric M-matrices (Stieltjes)

Block system:

$$\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \dots & A_{mm} \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Subdomain block Jacobi matrix $K(A) \in \mathbb{R}^{n \times n}$

$$K(A) = \begin{bmatrix} A_{11} & & \\ & \ddots & \\ & & A_{mm} \end{bmatrix}$$

Block matrices $K_{jj} = A_{jj}$ are Stieltjes matrices

Symmetric M-matrices (Stieltjes)

Definition

Define the matrix A^* by

$$A^* = K - \text{diag}(r_1, \dots, r_n),$$

where $r_i = \sum_{j=1}^n k_{ij}$ (i^{th} rowsum)

Block matrices A_{jj}^* have zero rowsums $\Rightarrow Z$ is a basis for $\text{null}(A^*)$.

Theorem

If A is an irreducibly diagonally dominant Stieltjes matrix and A^* has only irreducible blocks, then the hypotheses of the previous theorem are met.

Spectral properties

$$A = \left(\begin{array}{ccc|ccc} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \hline 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right)$$

$$A^* = \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right)$$

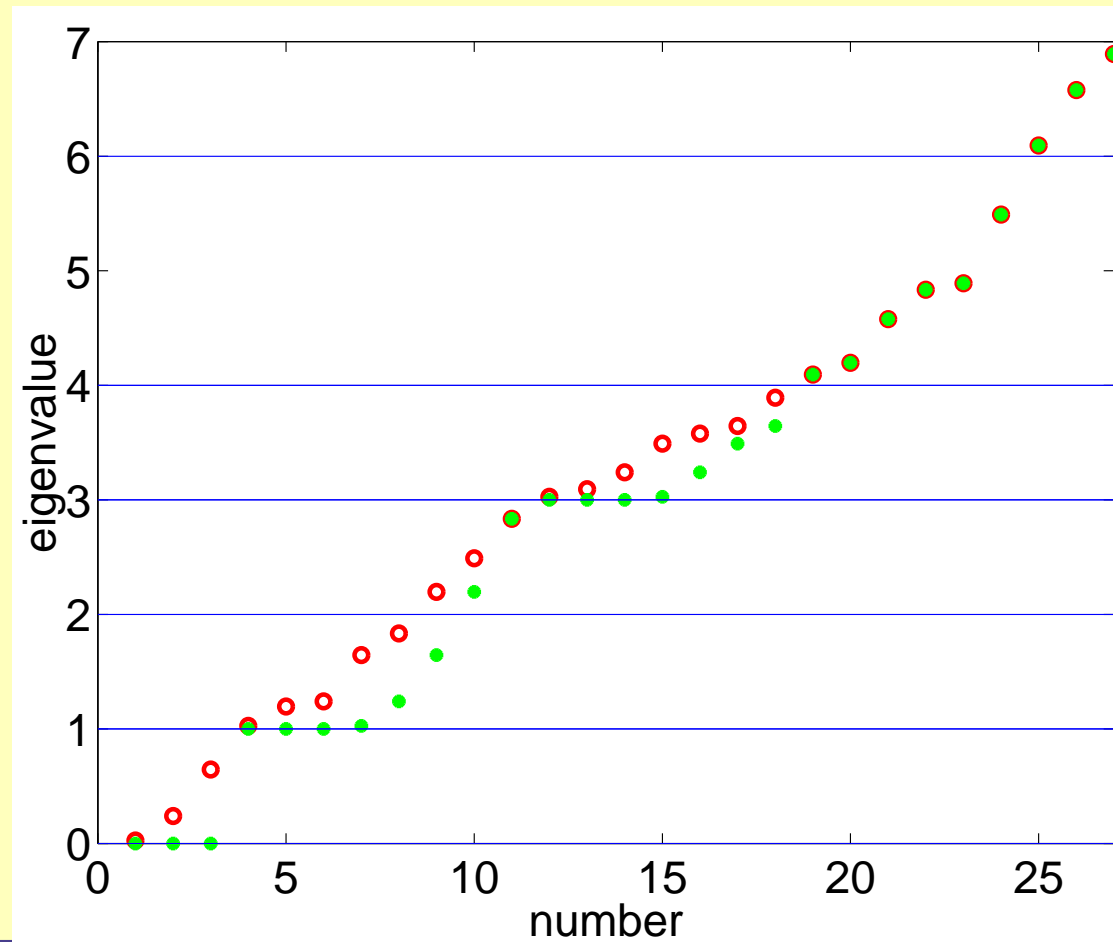
Spectral properties

2D problem, 3×9 points, 3 blocks

A

A^*

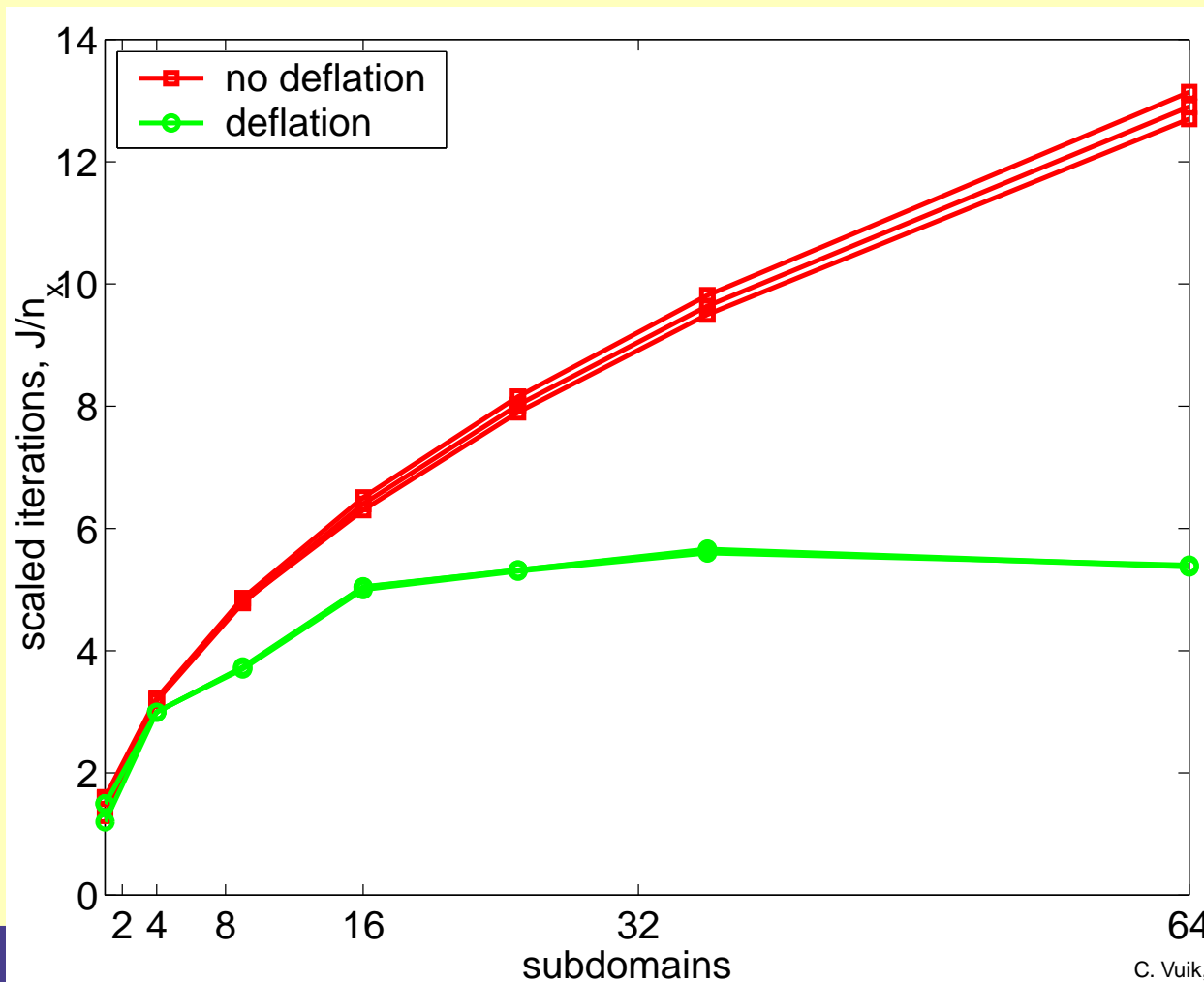
PA



Results (near grid independence)

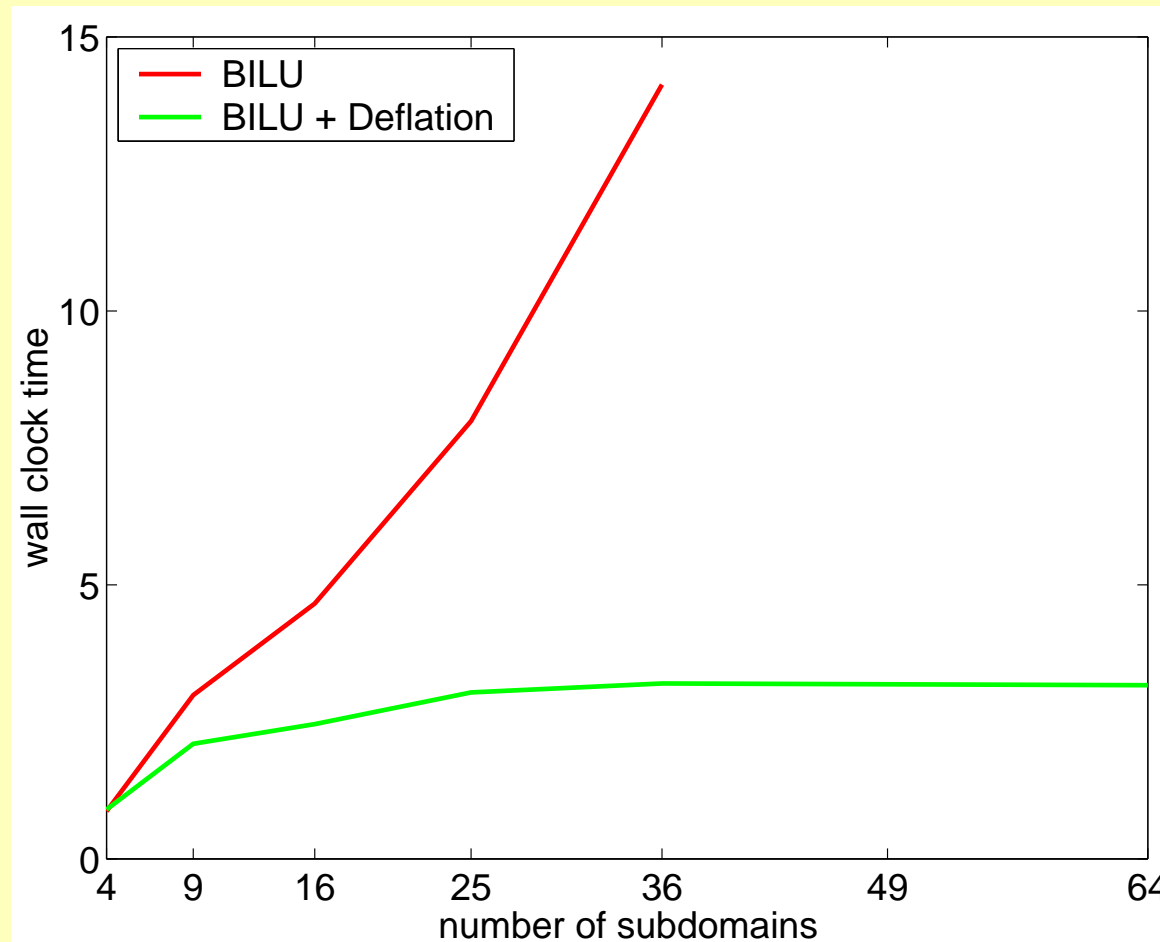
Poisson equation $-\Delta u(x, y) = f$.

Iterations divided by the subdomain resolution $n_x \equiv n_y \in \{10, 50, 200\}$



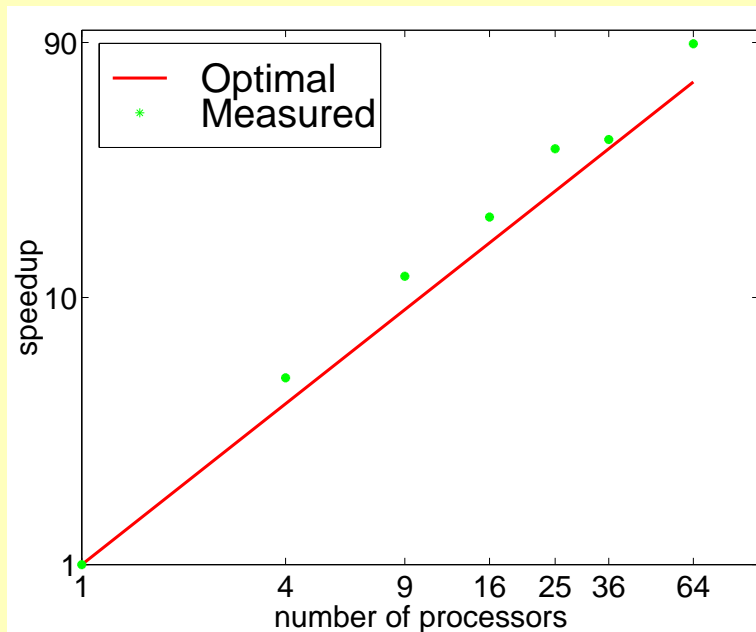
Results (parallel scalability)

subdomain grid size 50×50 , wall clock time, Cray T3E



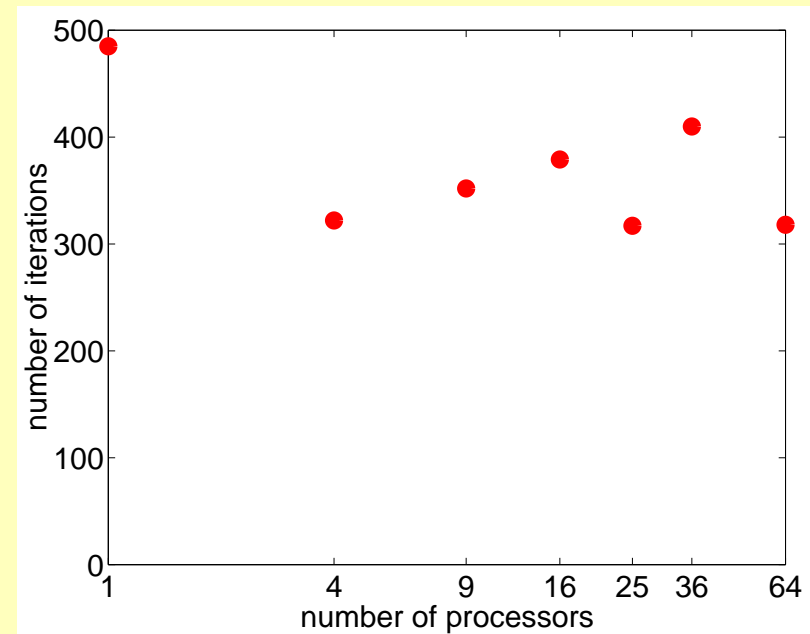
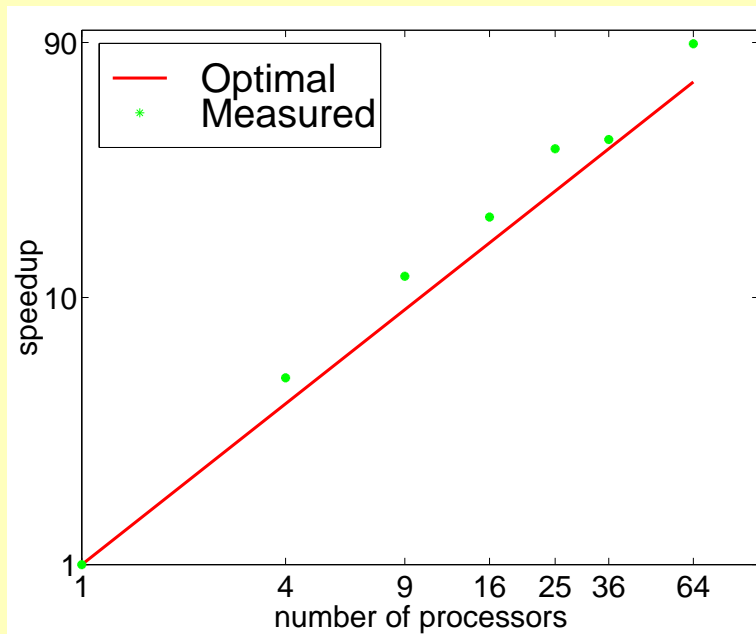
Results (parallel speedup)

480 × 480 grid, Cray T3E



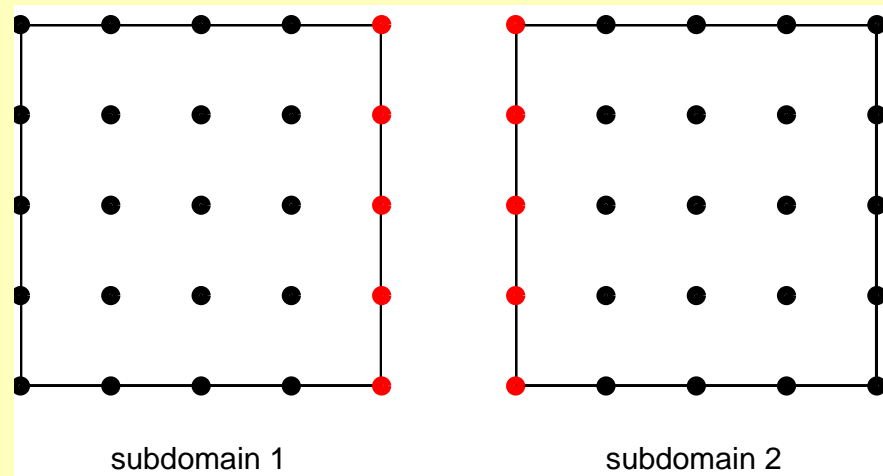
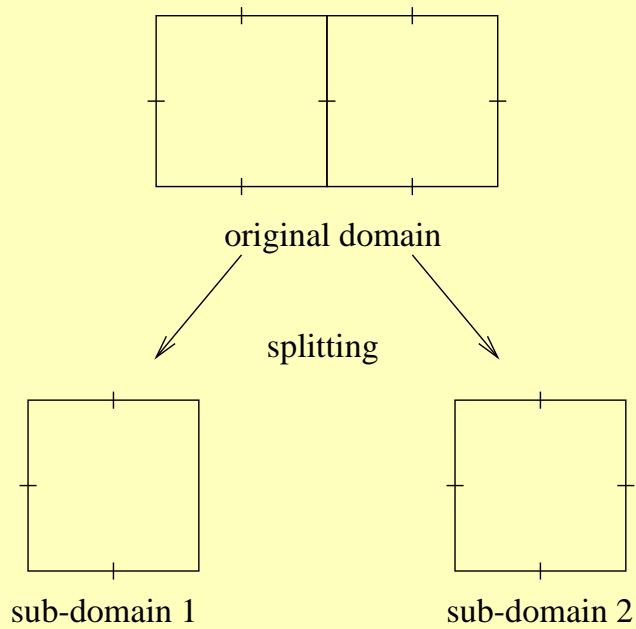
Results (parallel speedup)

480 × 480 grid, Cray T3E



5. Vertex centered approach

Data distribution



Variants for values at interfaces

$$z_i = 1 \text{ on } \Omega_i \text{ and } z_i = 0 \text{ on } \Omega \setminus \bar{\Omega}_i$$

1. no overlap

$z_i = 1$ at one subdomain

$z_i = 0$ at other subdomains

2. complete overlap

$z_i = 1$ at all subdomains

3. average overlap

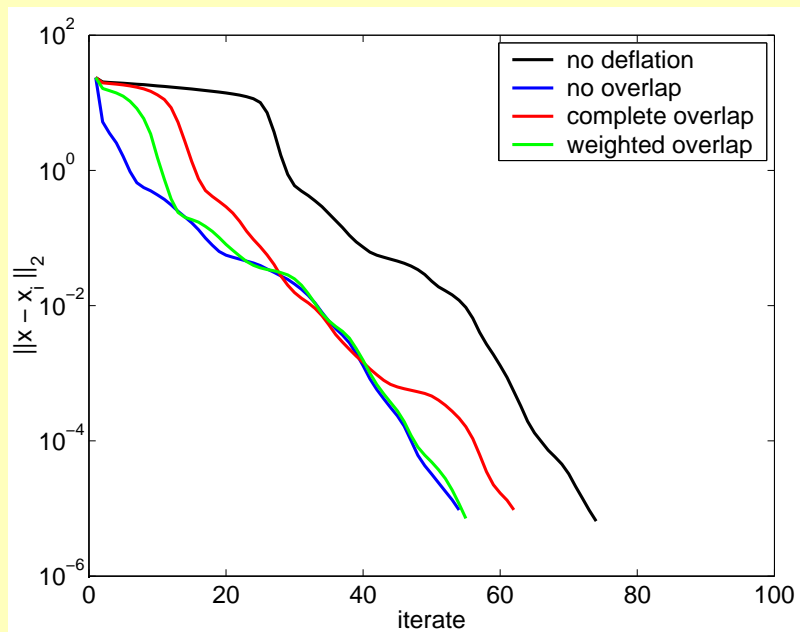
$z_i = \frac{1}{n_{neighbors}}$ at all subdomains

4. weighted overlap ($-\text{div}(\sigma \nabla u) = f$)

$$z_i = \frac{\sigma(i)}{\sum \sigma(neighbors)}$$

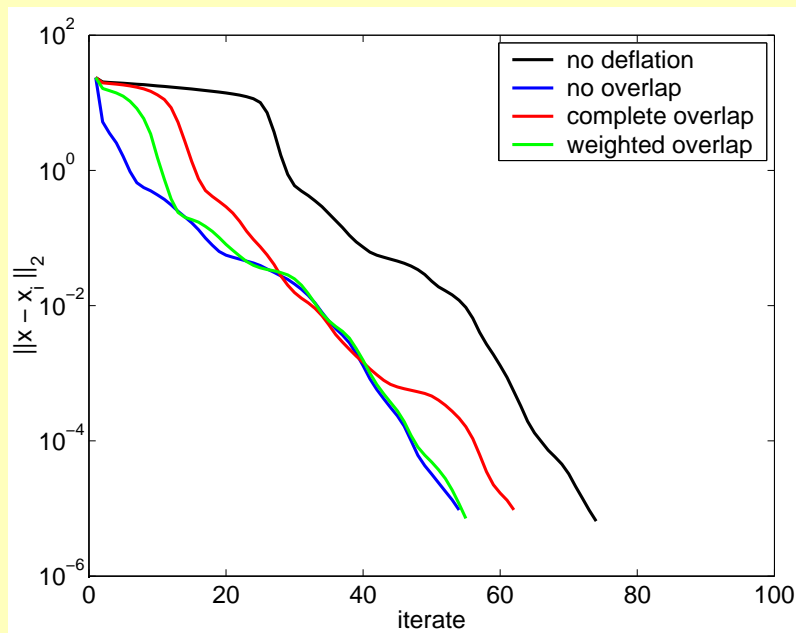
Results for constant coefficients

Error for Block IC and Deflation

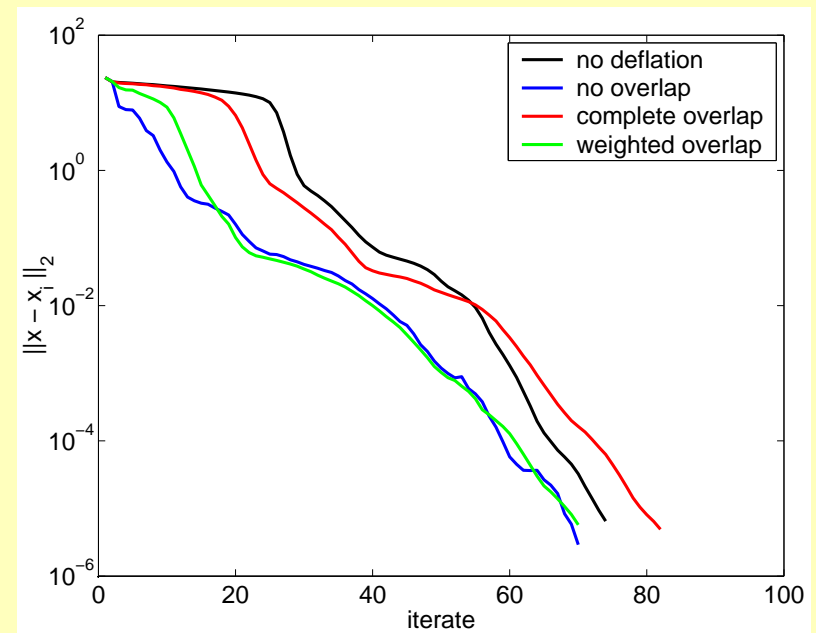


Results for constant coefficients

Error for Block IC and Deflation

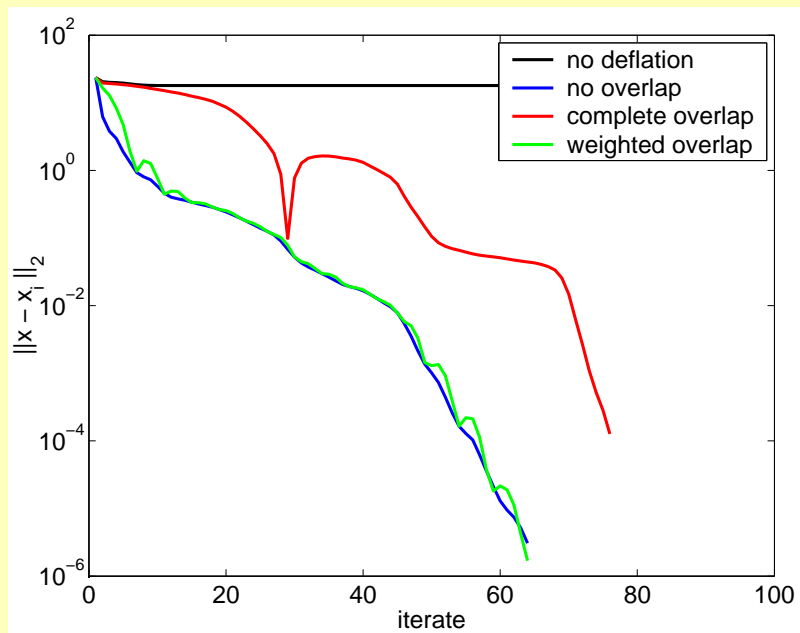


Error for Block IC and CGC



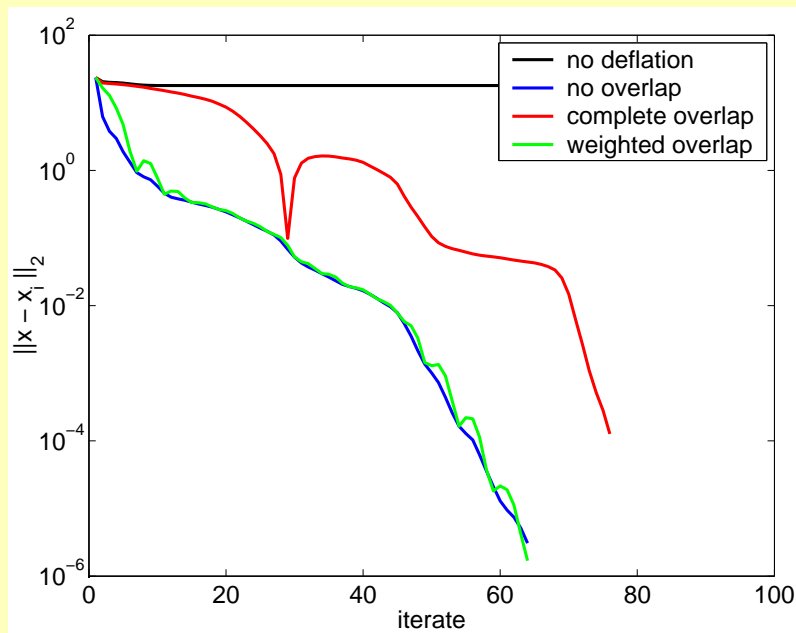
Results for discontinuous coefficients

Error for Block IC and Deflation

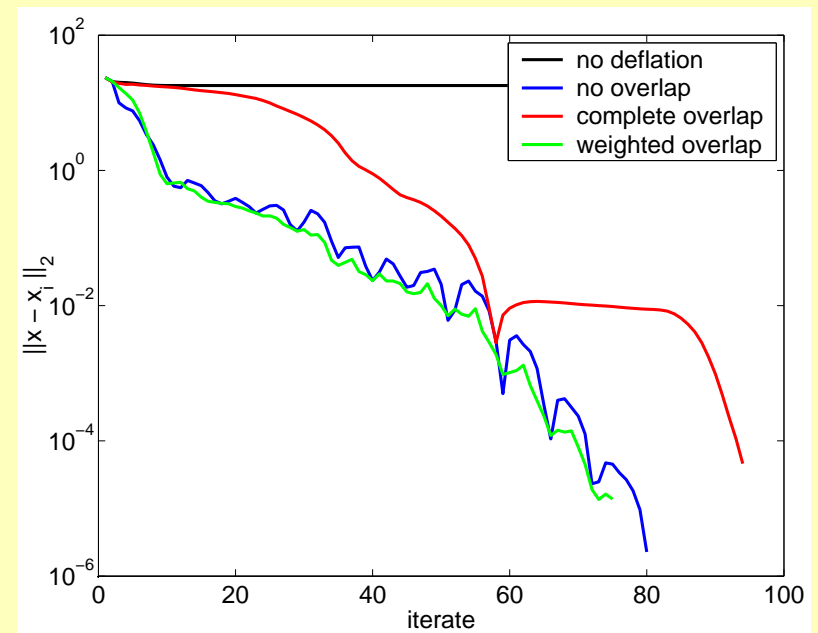


Results for discontinuous coefficients

Error for Block IC and Deflation

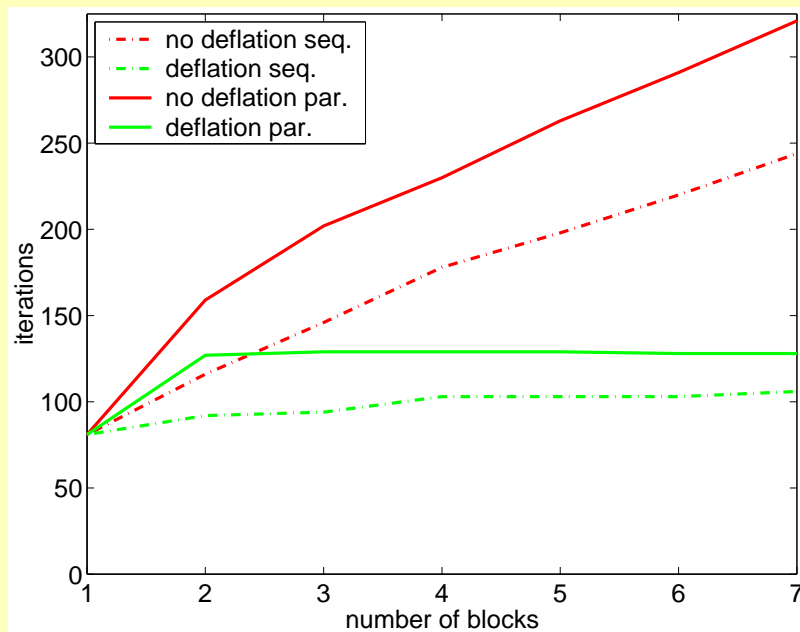


Error for Block IC and CGC



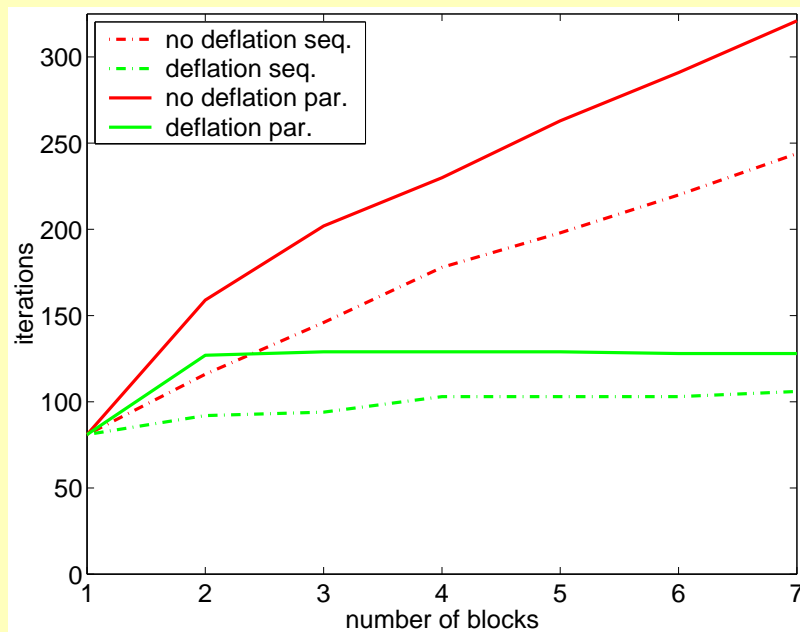
Results for Finite Element Method

Iterations

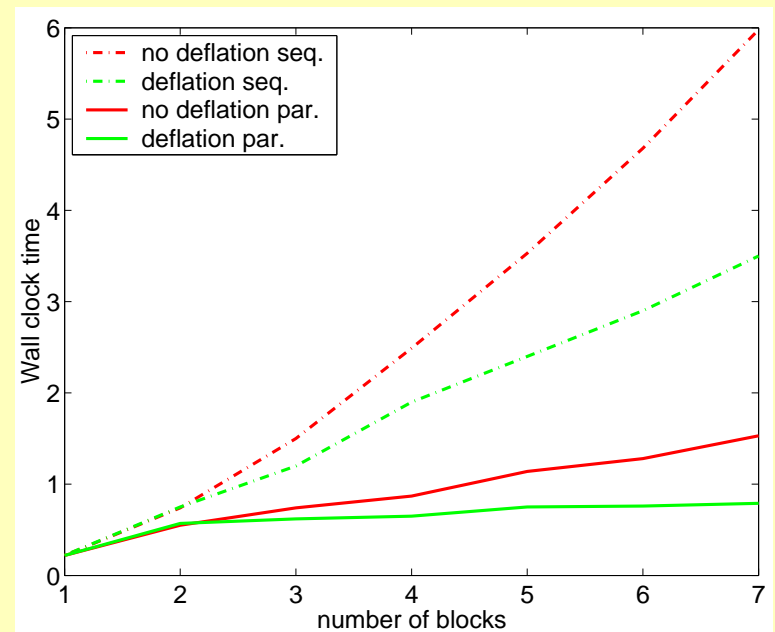


Results for Finite Element Method

Iterations



Wall clock time



6. Conclusions

- DICCG is more efficient than CICCG.
- Block preconditioned Krylov methods combined with Deflation or CGC are well parallelizable (scalable, good speed up).
- The theoretical results give insight into an optimal subdomain distribution and a suitable preconditioner.
- For the vertex centered case, the weighted overlap strategy is optimal
- Choices for the deflation vectors lead to comparable results in DICCG and CICCG.
- DICCG is a robust and efficient method to solve diffusion problems with discontinuous coefficients.

Further information

- http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_def.html
- C. Vuik A. Segal J.A. Meijerink
An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients
J. Comp. Phys., 152, pp. 385-403, 1999.
- J. Frank and C. Vuik
On the construction of deflation-based preconditioners
SIAM Journal on Scientific Computing, 23, pp. 442–462, 2001
- C. Vuik and A. Segal and L. El Yaakoubi and E. Dufour
A comparison of various deflation vectors applied to elliptic problems with discontinuous coefficients
Applied Numerical Mathematics, 41, pp. 219–233, 2002

Overview

Krylov

$$Ar$$

Preconditioned Krylov

$$L^{-T} L^{-1} Ar$$

Block Preconditioned Krylov

$$\sum_{i=1}^m (L_i^{-T} L_i^{-1}) Ar$$

Block Preconditioned Deflated Krylov

$$\sum_{i=1}^m (L_i^{-T} L_i^{-1}) P Ar$$