Deflated Krylov Acceleration of the Schwarz Domain Decomposition Method

Kees Vuik, A. Segal, F.J. Vermolen and J. Frank
c.vuik@math.tudelft.nl

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

Delft University of Technology

Woudschoten Conference of the Dutch-Flemish Numerical Analysis Communities

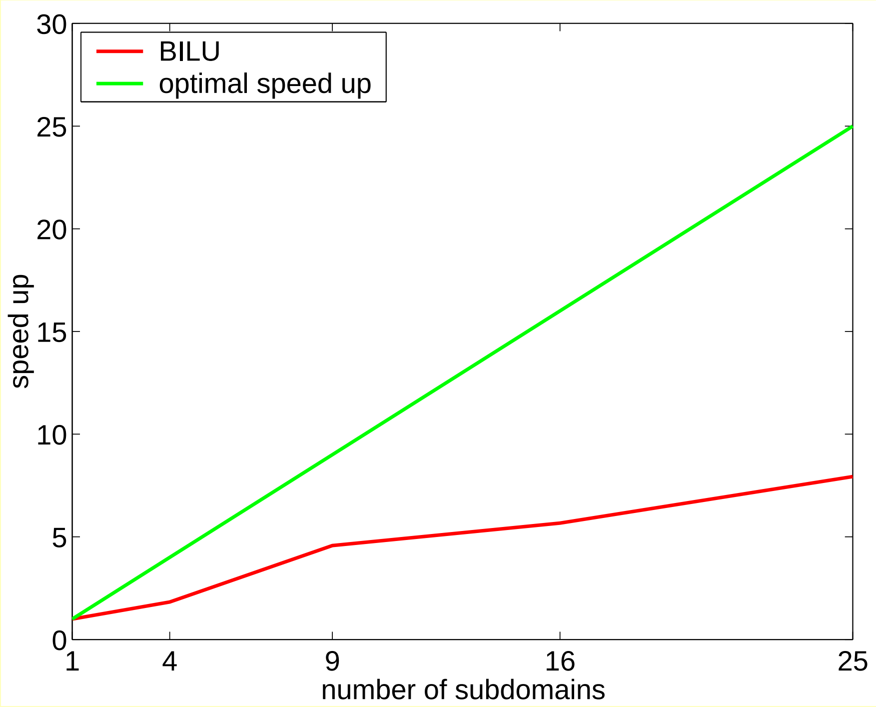
Woudschoten, 25 - 27 September, 2002

Contents

TU Delft

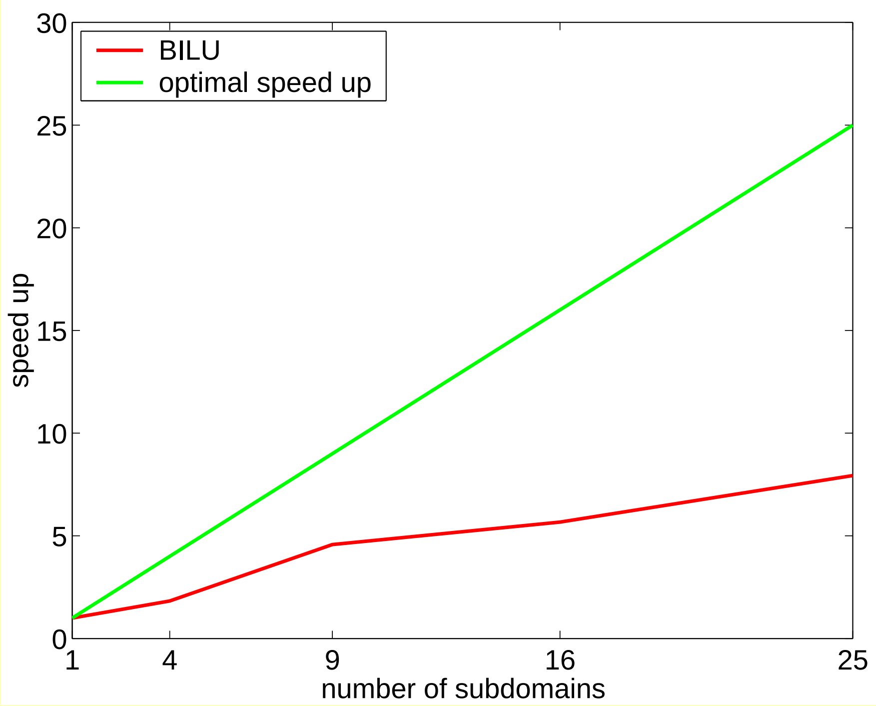
1. Introduction

Block ILU

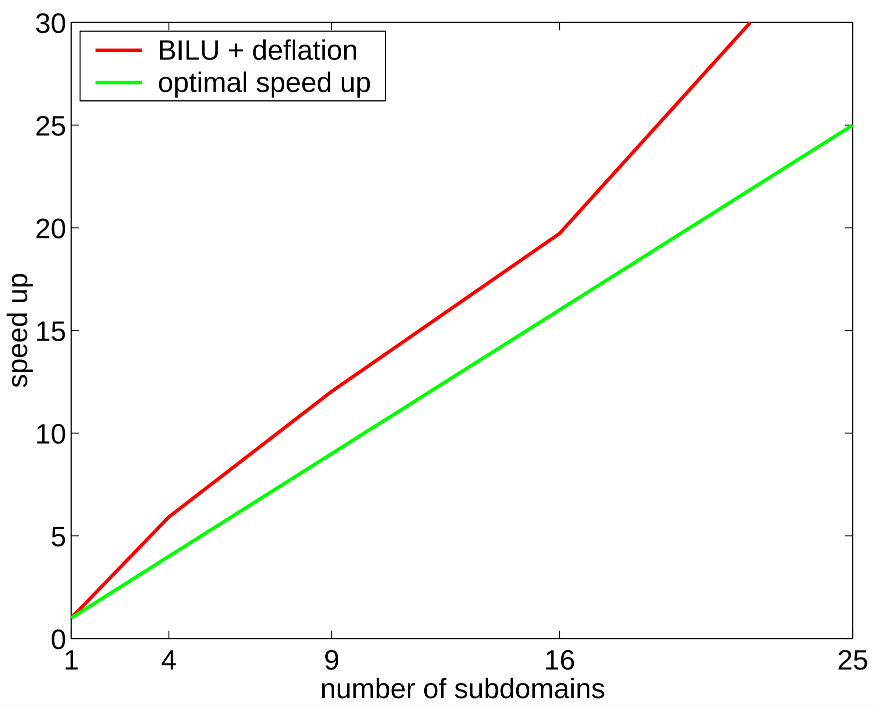


1. Introduction

Block ILU



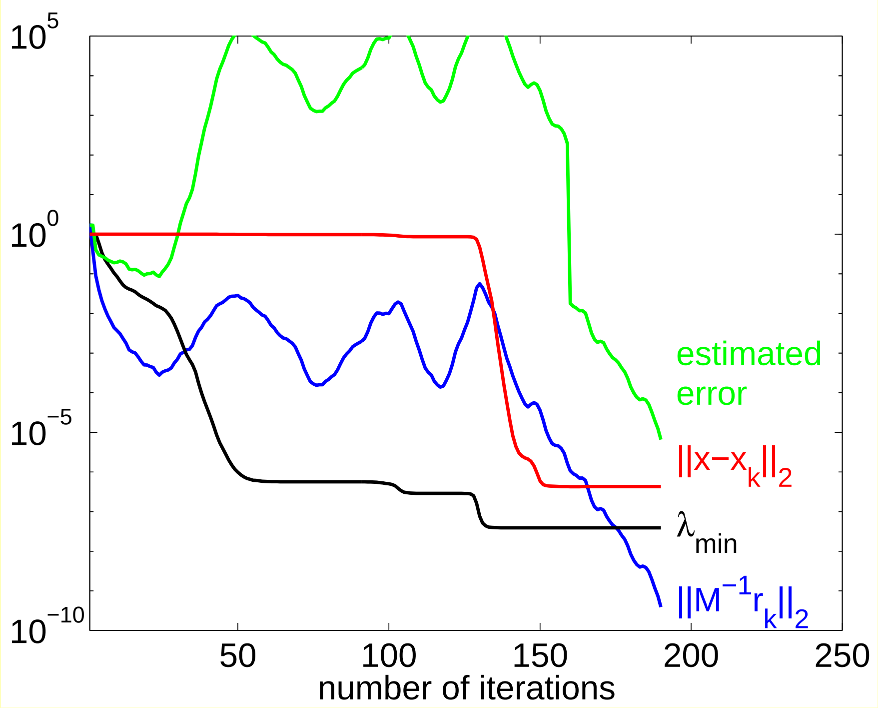
Block ILU with deflation



TU Delft

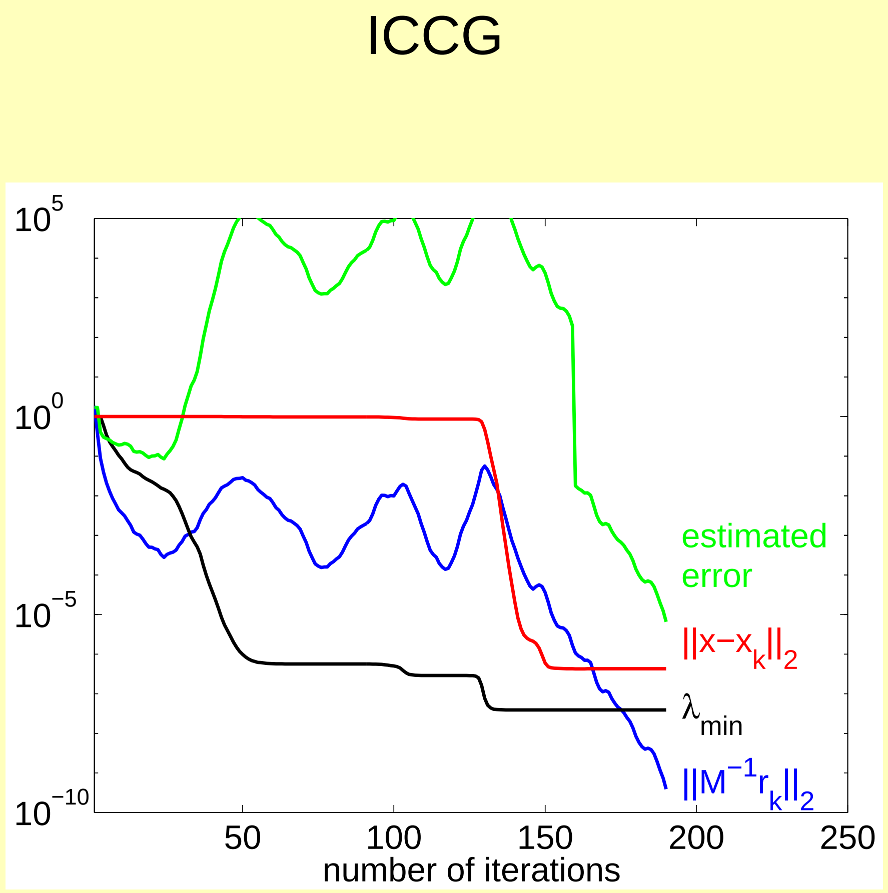
Results oil flow problem

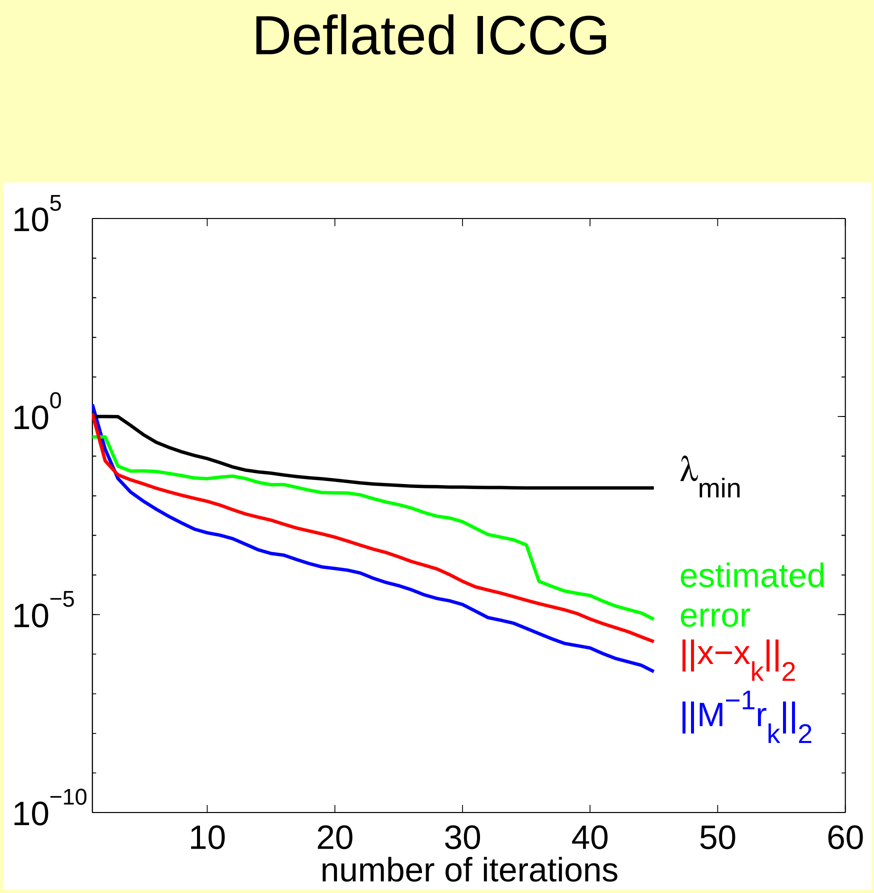
ICCG



TU Delft

Results oil flow problem





TU Delft

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 Nicolaides, Mansfield, Frank, Vuik
  Morgan, Chapman, Saad, Burrage, Ehrel, Pohl
  FETI Farhat, Roux, Mandel, Klawonn, Widlund

2. Deflated Krylov methods

A is SPD, Conjugate Gradients

$P = I - AZE^{-1}Z^T$ with $E = Z^TAZ$

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

Properties

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

Deflated ICCG

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

Deflated ICCG

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

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

TU Delft

Deflated ICCG

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

$$(I - P^T)x = ZE^{-1}Z^TAx = ZE^{-1}Z^Tb, \qquad AP^Tx = PAx = Pb.$$

TU Delft

Deflated ICCG

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

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

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**

$k = k + 1;$

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

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

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

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

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

end while

TU Delft

Deflation for non-symmetric matrices

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

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

Properties

1. $PAZ = Y^TP = 0$ and $Y^TAQ = QZ = 0$
2. $P^2 = P$ and $Q^2 = Q$
3. $PA = AQ$

Deflated Krylov methods

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

TU Delft

Deflated Krylov methods

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

$$(I - Q)x = ZE^{-1}Y^TAx = ZE^{-1}Y^Tb,$$

TU Delft

Deflated Krylov methods

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

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

TU Delft

Deflated Krylov methods

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

$$(I - Q)x = ZE^{-1}Y^TAx = ZE^{-1}Y^Tb, \qquad AQx = PAx = Pb.$$

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

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

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

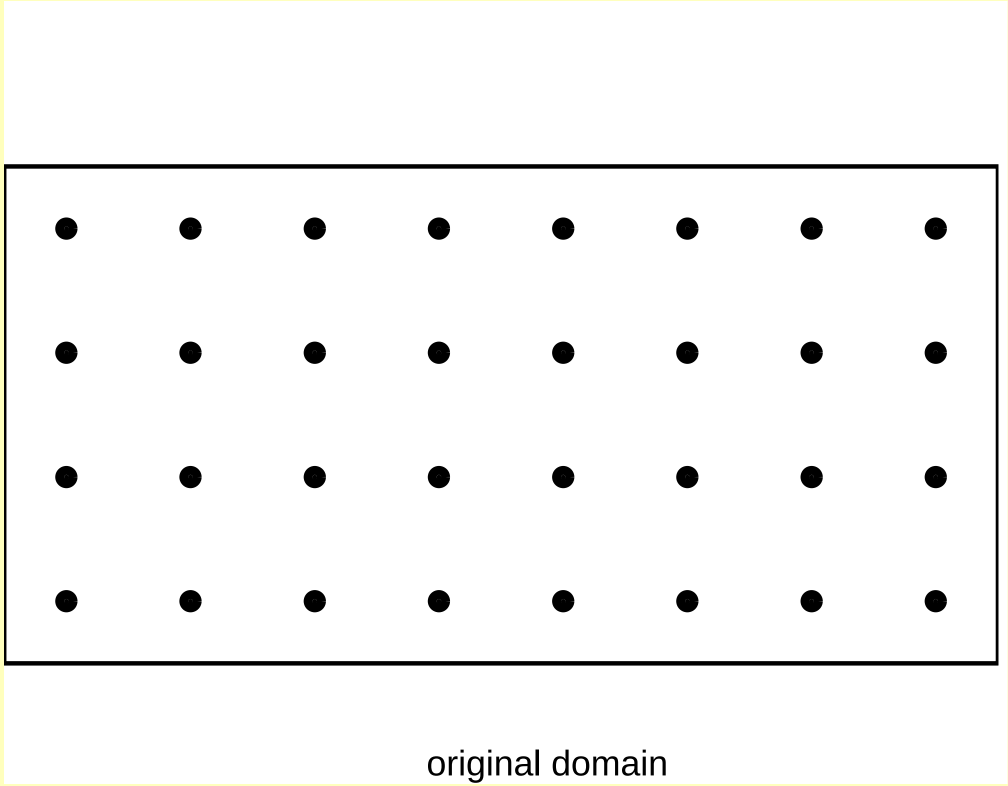
Preconditioning

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

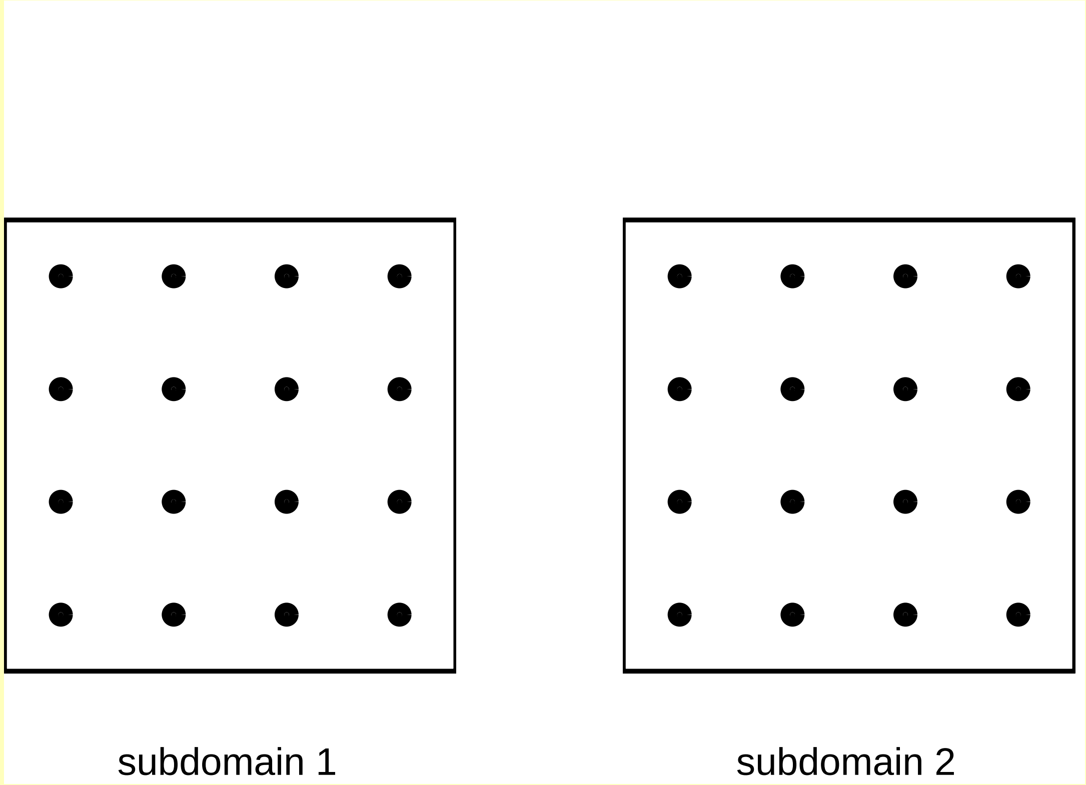
$$PAK^{-1}\tilde{y} = Pb, \quad Qx = QK^{-1}\tilde{y}$$

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

Decomposition of a cell centered domain



Decomposition of a cell centered domain



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

Choice of the deflation vectors

m is number of subdomains $z_1, ..., z_m$ 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 $AP_D^T x = P_D A x = P_D b$

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**

$k = k + 1;$

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

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

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

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

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

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

$P_D = I - AZE^{-1}Z^T$ $\qquad$ $P_C = I + ZE^{-1}Z^T$

TU Delft

Properties of Deflation and CGC

$P_D = I - AZE^{-1}Z^T$ $\qquad$ $P_C = I + ZE^{-1}Z^T$

Properties of P_D

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

Properties of Deflation and CGC

$$P_D = I - AZE^{-1}Z^T \qquad\qquad P_C = I + ZE^{-1}Z^T$$

Properties of P_D

- $P_D A$ is symmetric and positive semidefinite
- P_D is a projection, $P_D AZ = 0$
- 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 $0 < \lambda_1 \leq \ldots \leq \lambda_n$.

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

Properties of Deflation and CGC

Definition
Eigenpair $\{\lambda_i, v_i\}$, so $Av_i = \lambda_i v_i$ with $0 < \lambda_1 \leq \ldots \leq \lambda_n$.
Take $Z = [v_1 \ldots v_m]$.

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

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

Properties of Deflation and CGC

Definition

Eigenpair $\{\lambda_i, v_i\}$, so $Av_i = \lambda_i v_i$ with $0 < \lambda_1 \leq \ldots \leq \lambda_n$.
Take $Z = [v_1 \ldots v_m]$.

Theorem 1

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

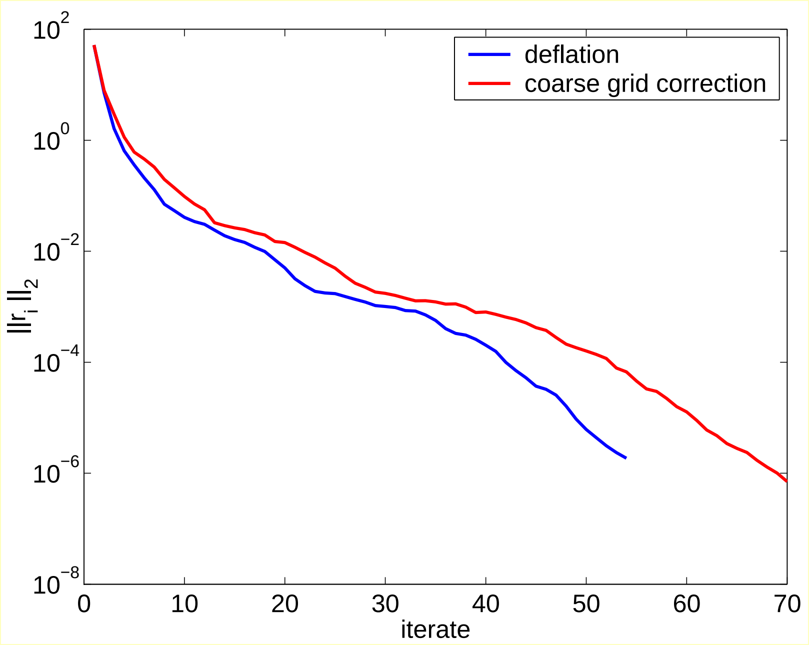
- the spectrum of $P_C A$ is $\{1 + \lambda_1, \ldots, 1 + \lambda_m, \lambda_{m+1}, \ldots, \lambda_n\}$

Corollary

DICCG converges faster than CICCG.

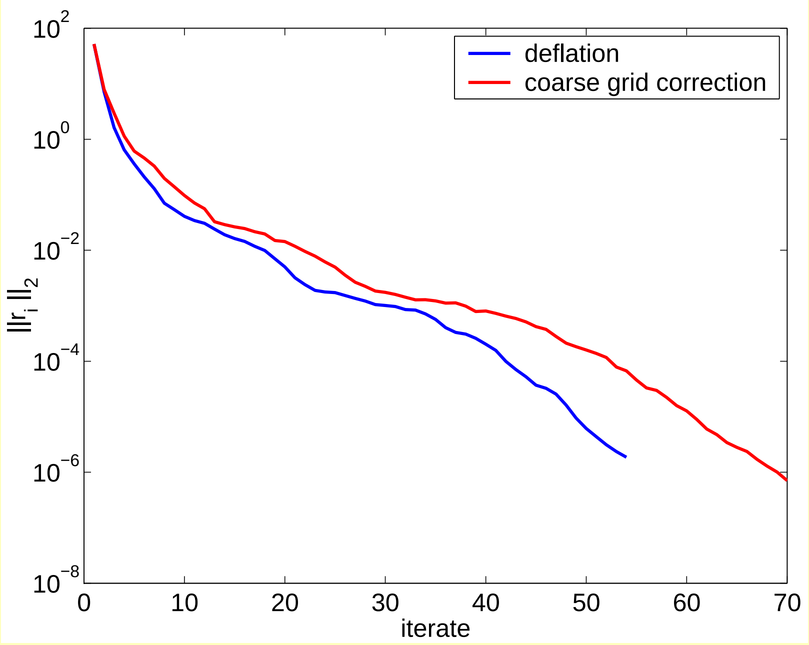
Experiments with Deflation and CGC

Residual with Block IC

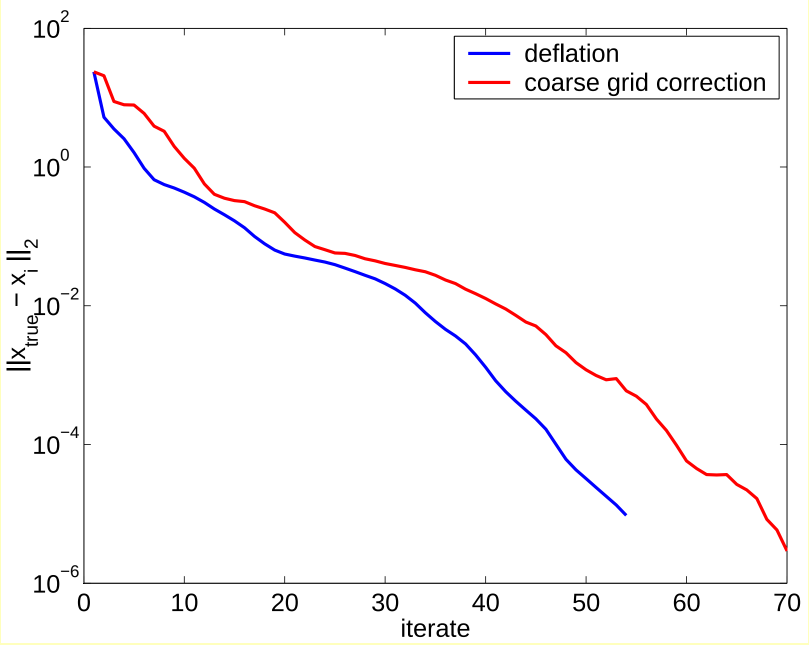


Experiments with Deflation and CGC

Residual with Block IC



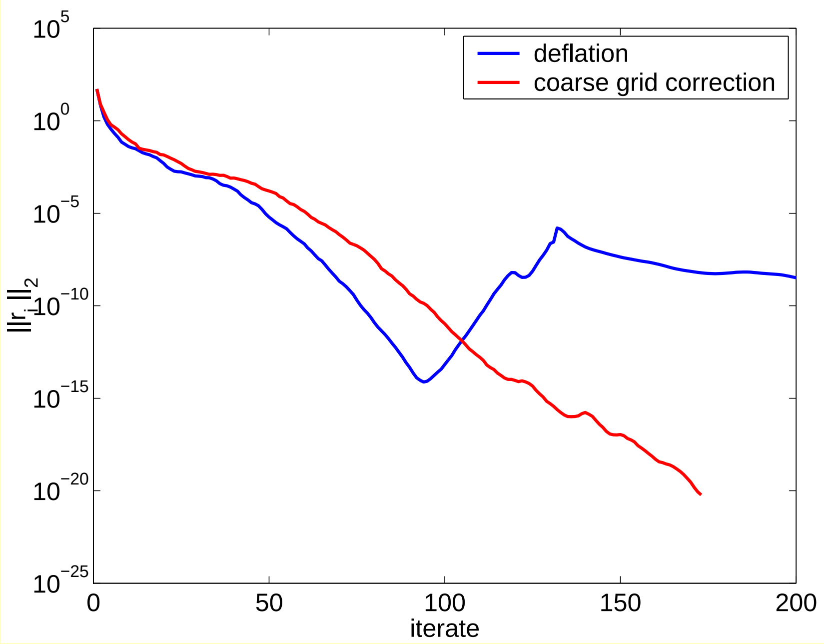
Error with Block IC



TU Delft

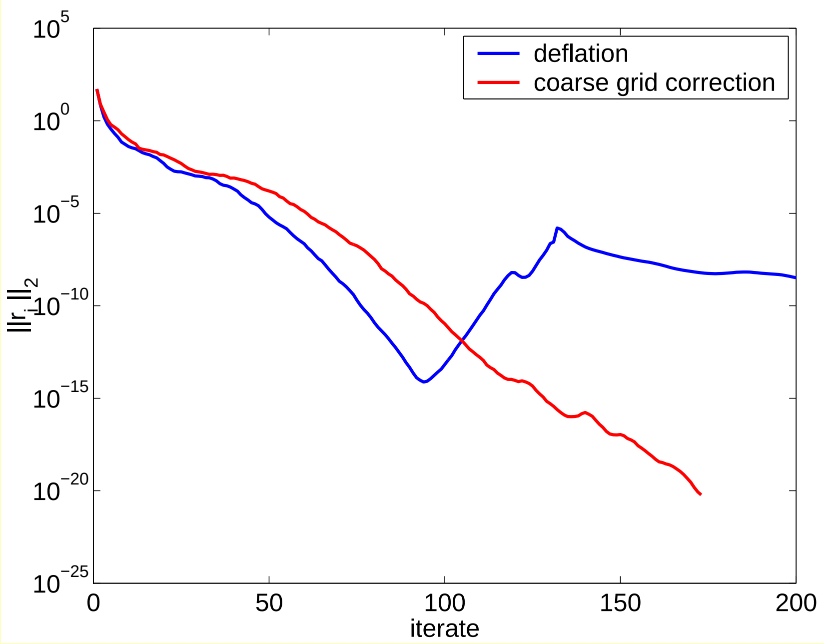
Too severe termination criterion

Residual with Block IC

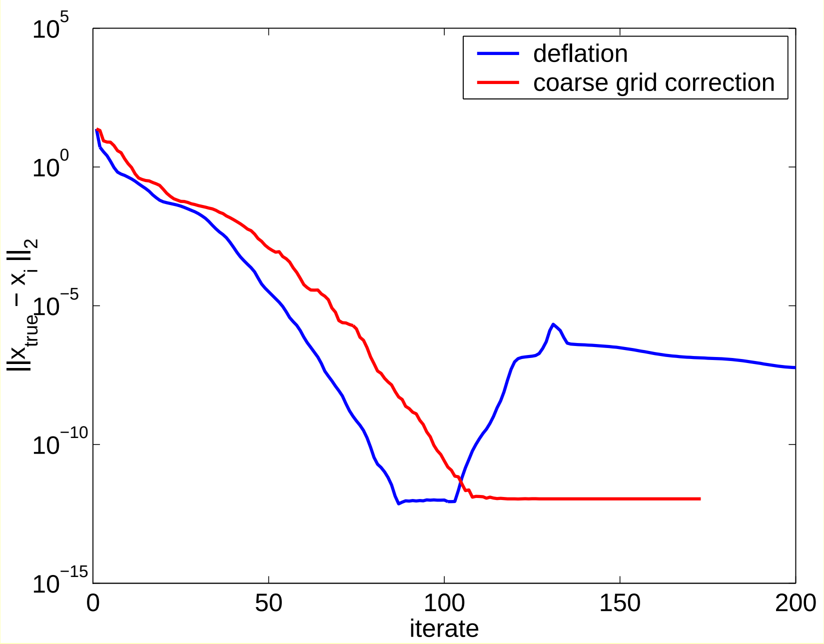


Too severe termination criterion

Residual with Block IC



Error with Block IC



TU Delft

4. Convergence of DICCG

Theorem 2

Let A be SPD, and A^*, $A - A^*$ be SPSD, and span$Z$ = 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

$$A^* = K - diag(r_1, ..., r_n),$$

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 null(A^*).

Theorem 3

If A is an irreducibly diagonally dominant Stieltjes matrix and A^* has only irreducible blocks, then the hypotheses of Theorem 2 are met.

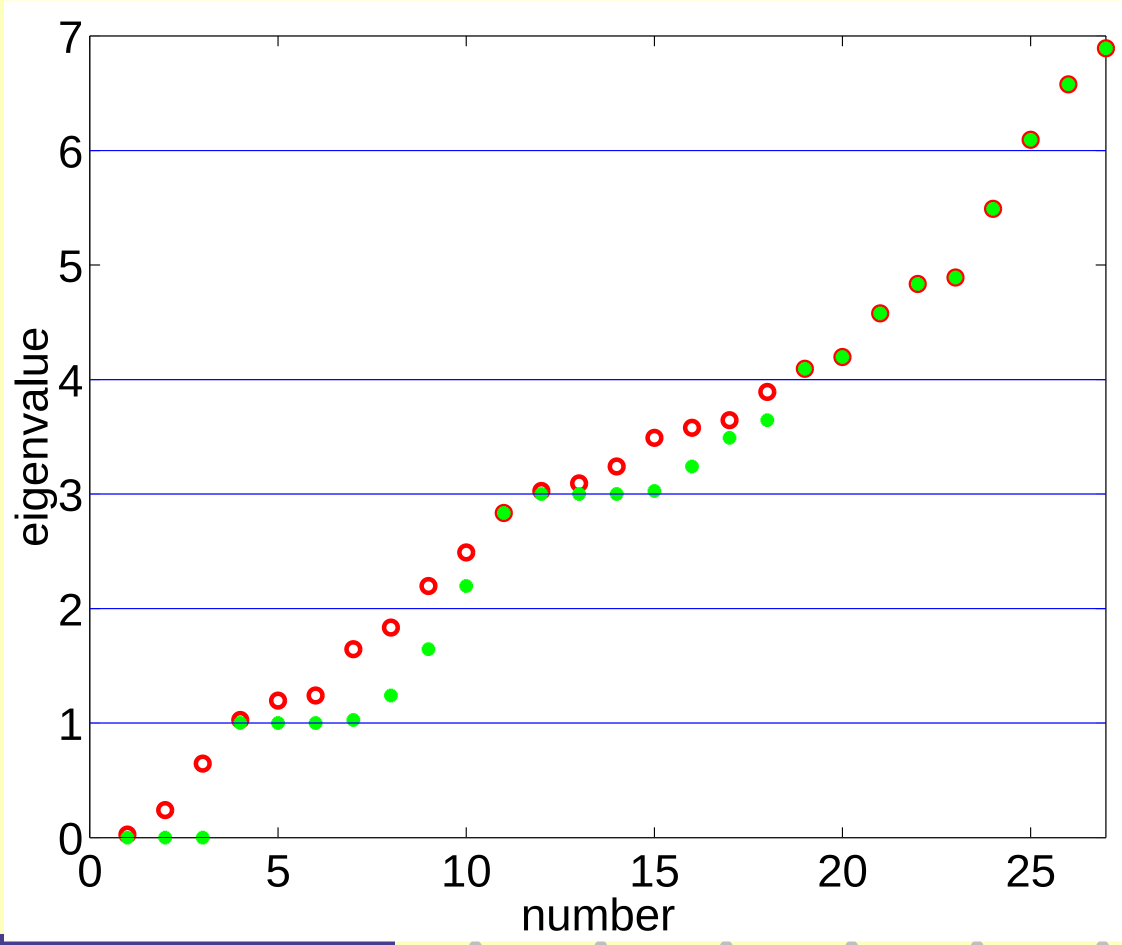
Spectral properties

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

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

Spectral properties

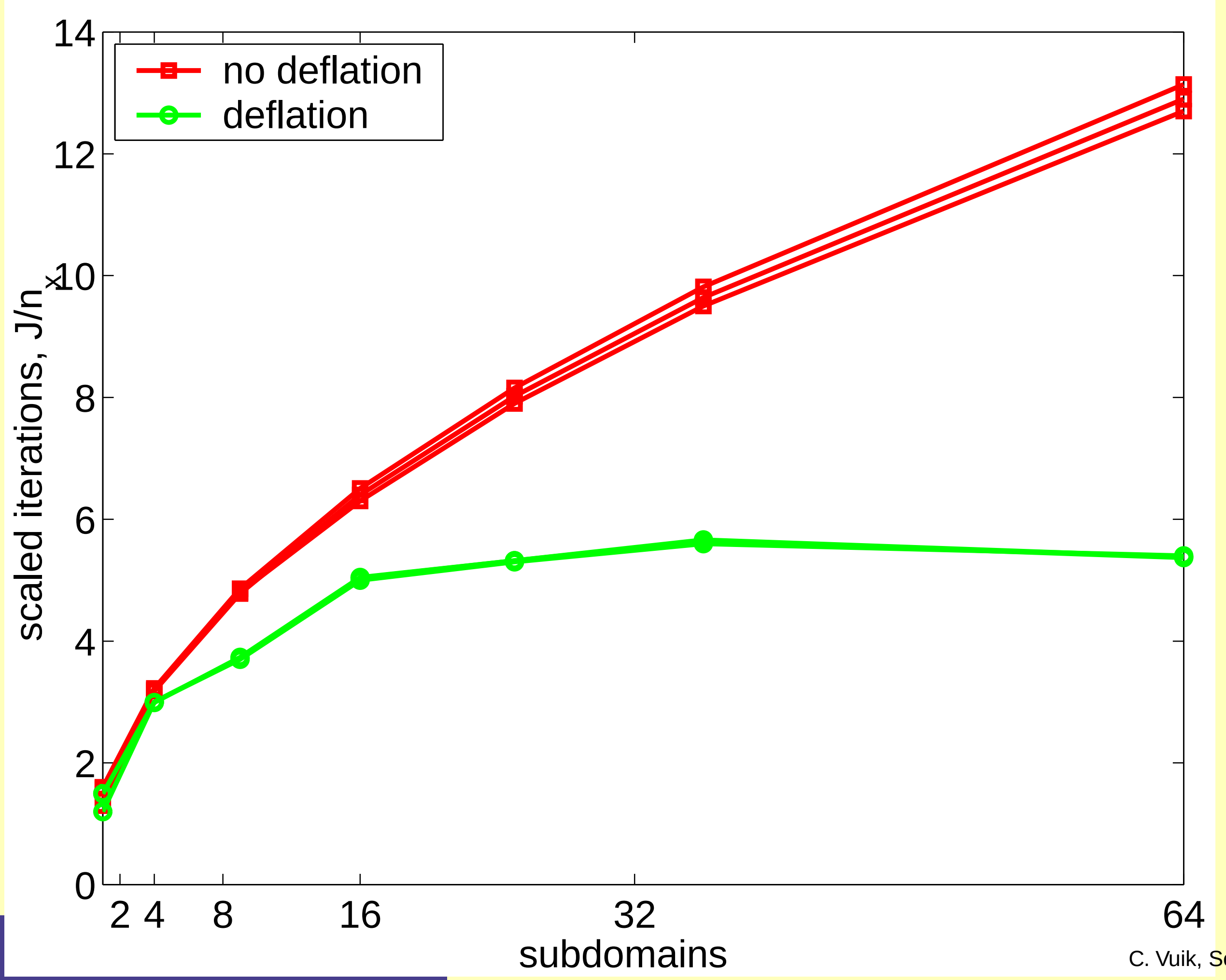
2D problem, 3×9 points, 3 blocks A A^* PA



Results (near grid independence)

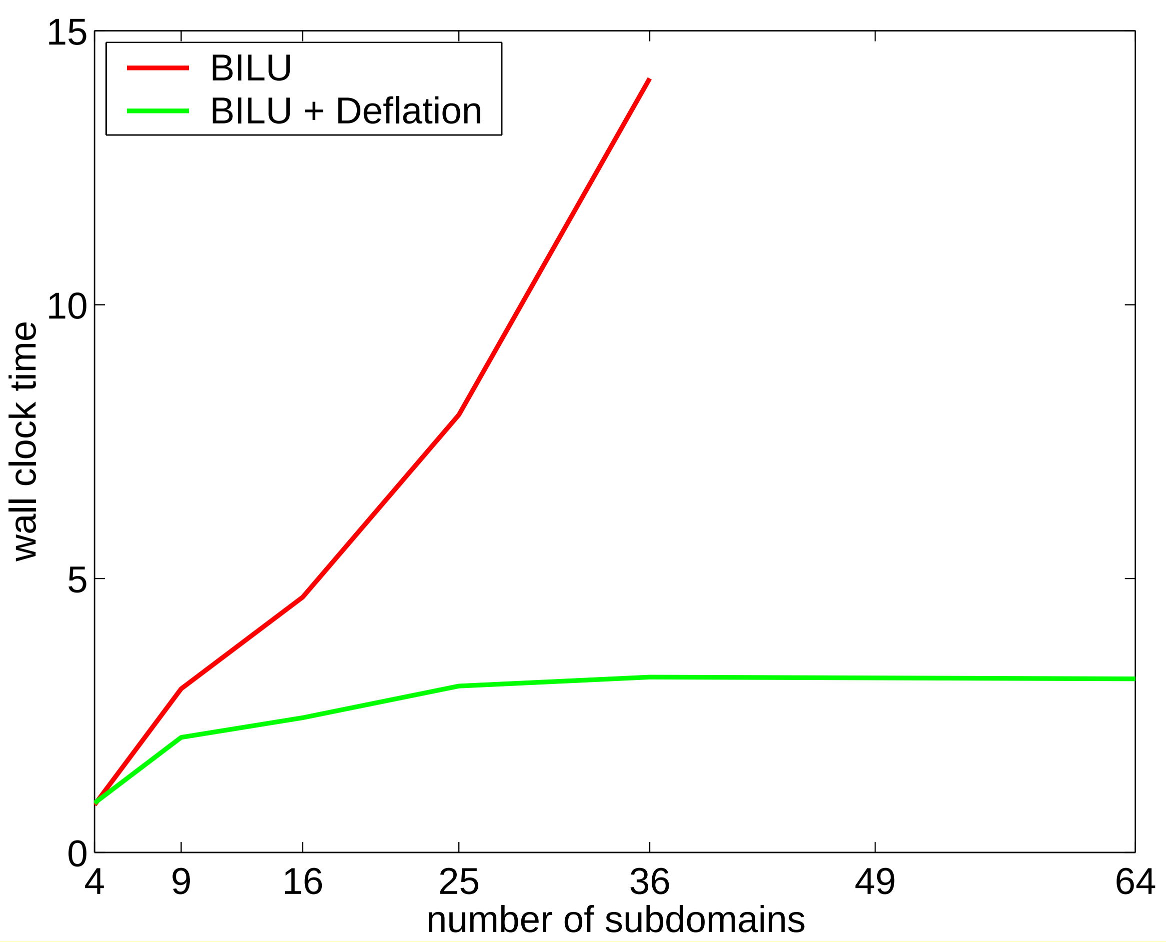
Poisson equation $-\Delta u(x,y) = f$.

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



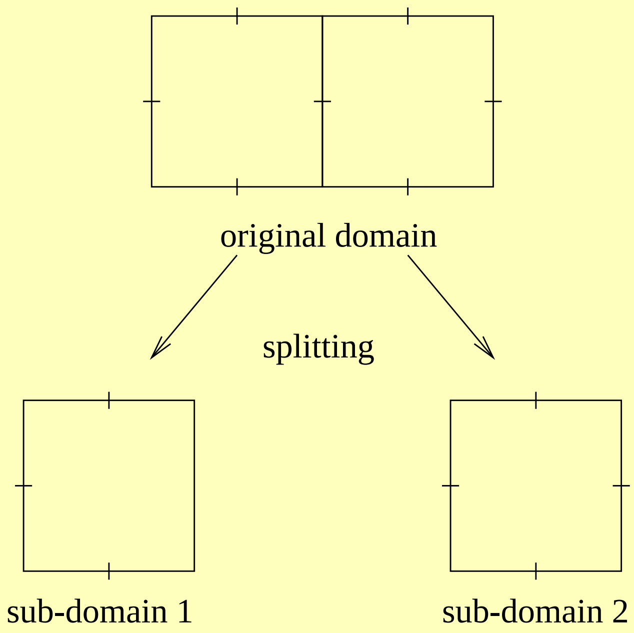
Results (parallel scalability)

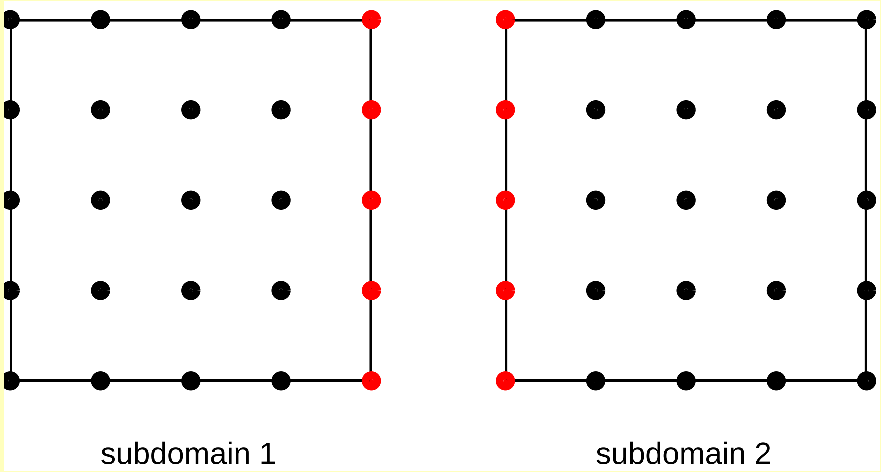
subdomain grid size 50×50, wall clock time, Cray T3E



TU Delft

Data distribution





Variants for values at interfaces

$z_i = 1$ on Ω_i and $z_i = 0$ 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**

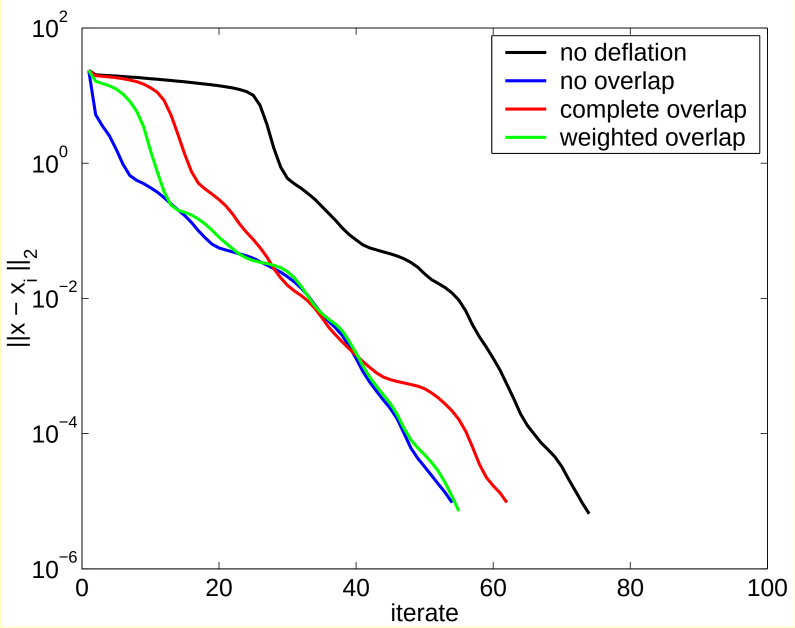
   $z_i = \frac{1}{n_{neighbors}}$ at all subdomains

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

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

Results for constant coefficients

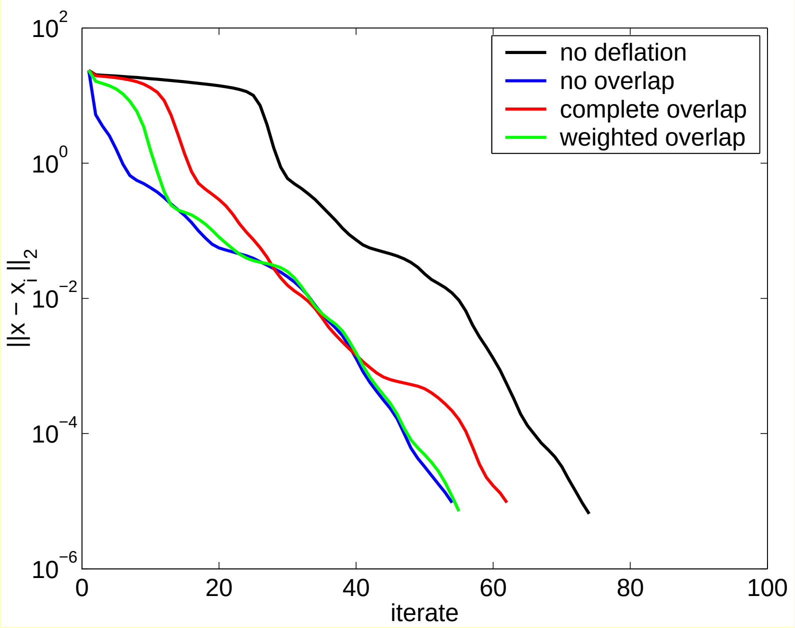
Error for Block IC and Deflation



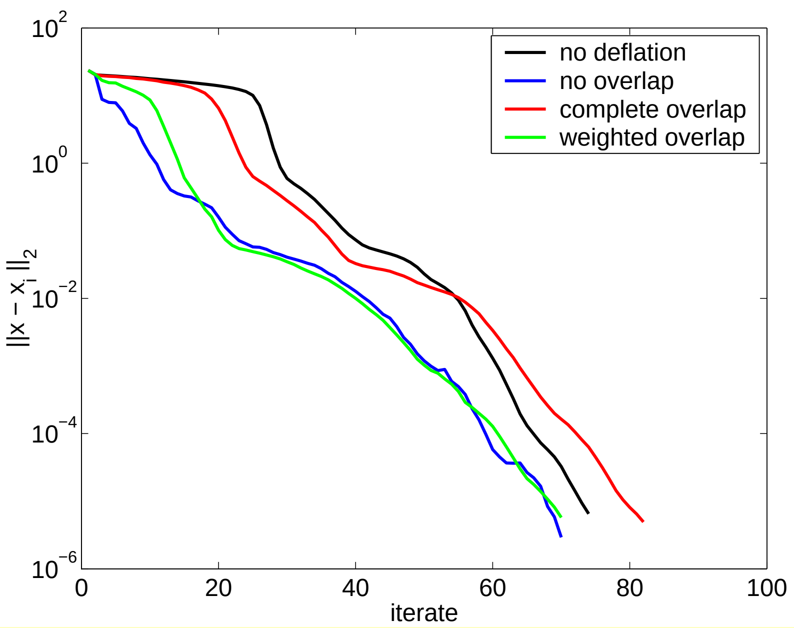
TU Delft

Results for constant coefficients

Error for Block IC and Deflation



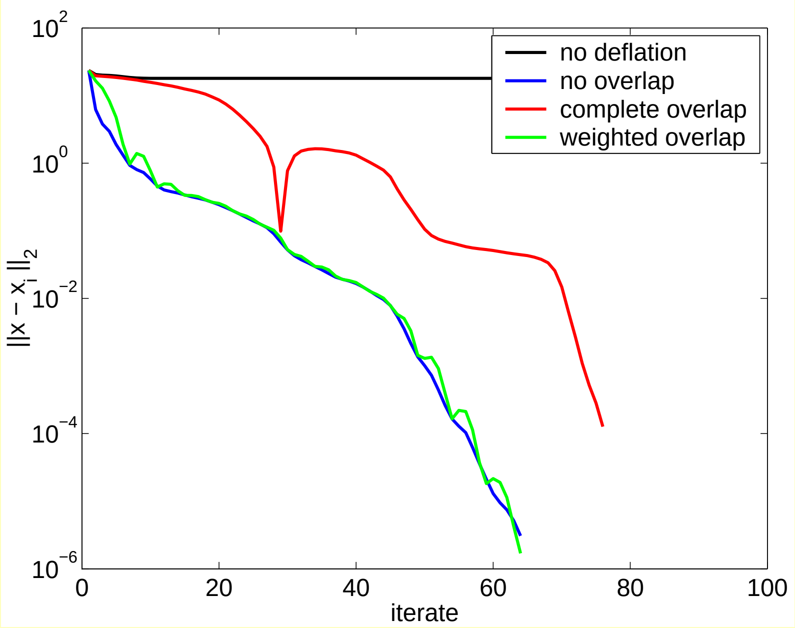
Error for Block IC and CGC



TU Delft

Results for discontinuous coefficients

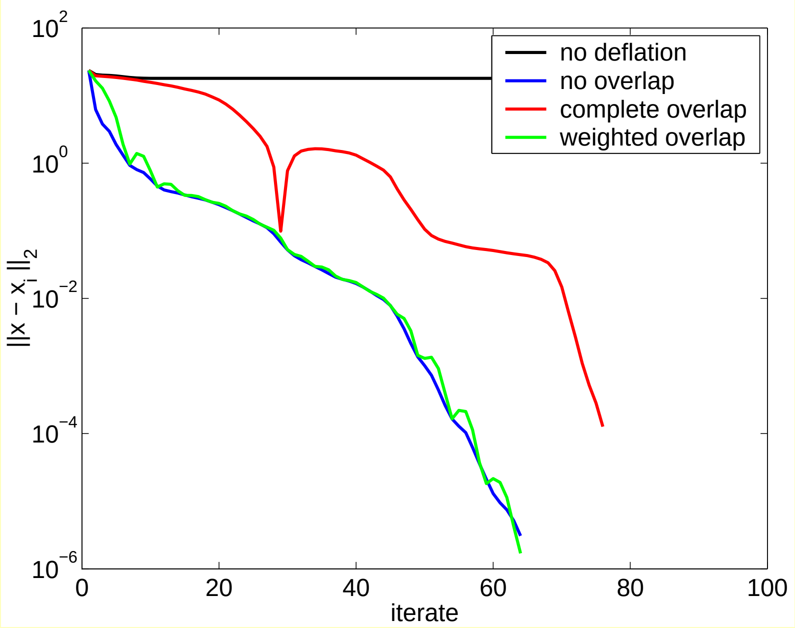
Error for Block IC and Deflation



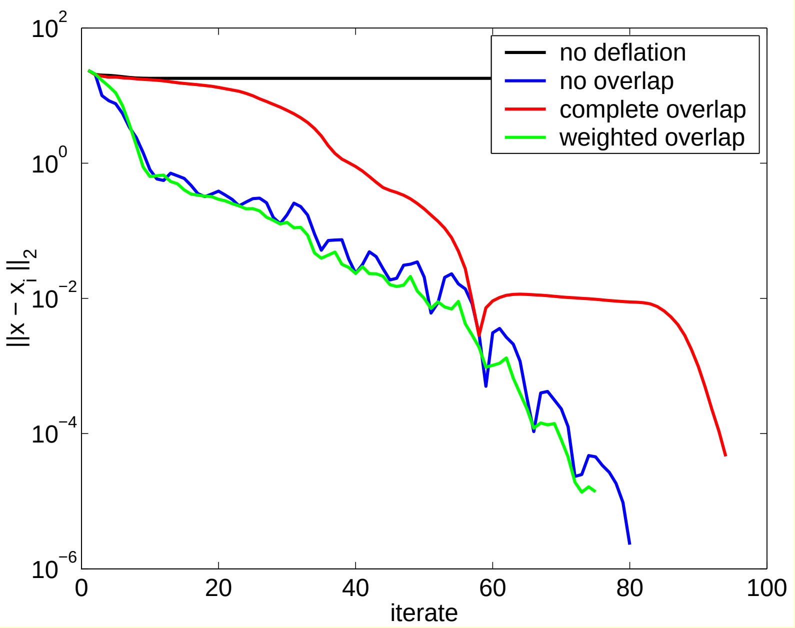
TU Delft

Results for discontinuous coefficients

Error for Block IC and Deflation



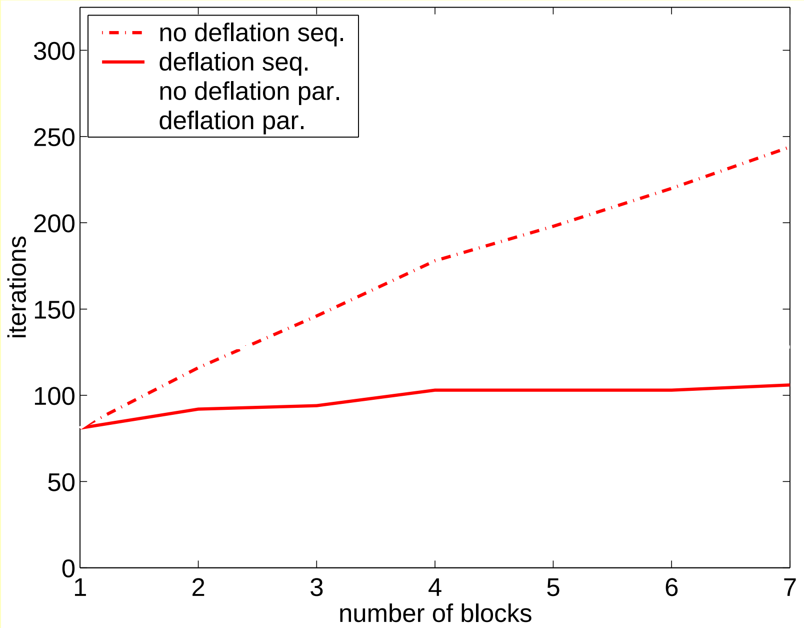
Error for Block IC and CGC



TU Delft

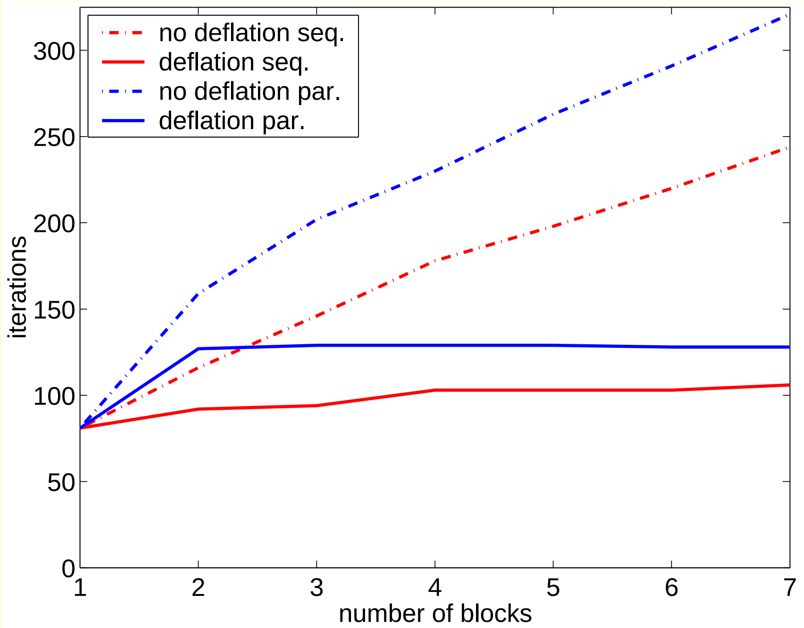
Results for Finite Element Method

Iterations



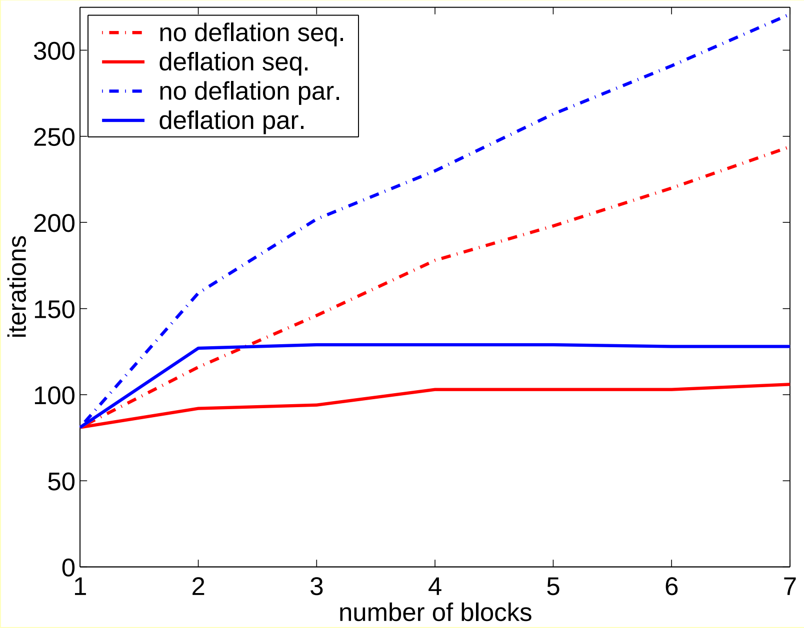
Results for Finite Element Method

Iterations

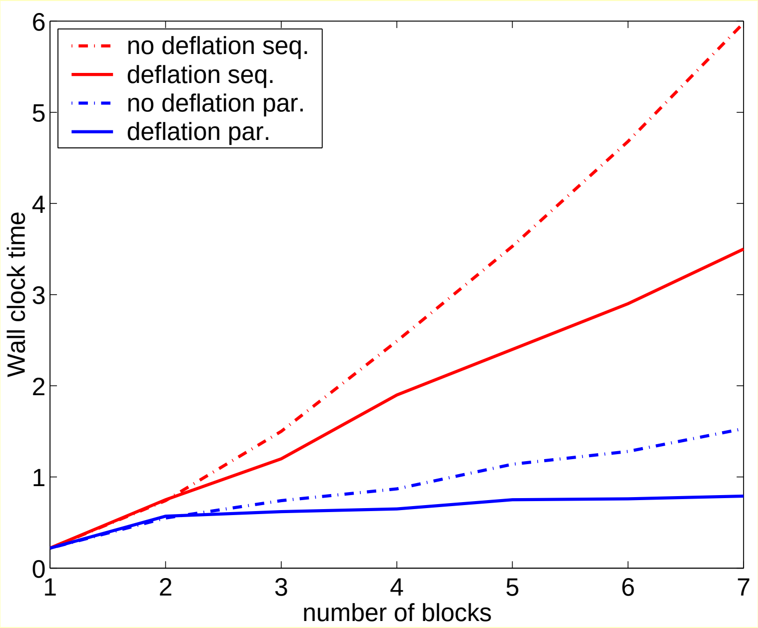


Results for Finite Element Method

Iterations



Wall clock time



TU Delft

6. Conclusions

- DICCG is more efficient than CICCG
- Block preconditioned Krylov methods combined with Deflation or CGC are well parallelizable (scalable, good speed up)
- 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	$\sum_{i=1}^{m}(L_i^{-T}L_i^{-1})PAr$

Results for a non-symmetric problem

$$\nabla \cdot (\mathbf{a}(x,y)u(x,y)) - \Delta u(x,y) = f \text{ on } (0,1) \times (0,1)$$

recirculating wind field $a_1(x,y) = -80xy(1-x)$, $a_2(x,y) = 80xy(1-y)$
boundary conditions $u(x,0) \equiv u(y,0) \equiv u(x,1) \equiv 0$, $u_x(1,y) = 0$

Subdomain grid 50×50, truncated GCR

m	no deflation	deflation
1	42	42
4	122	122
9	224	191
16	314	235
25	369	250
36	518	283
49	1007	377

TU Delft