Deflated Krylov Acceleration of the Schwarz Domain Decomposition Method

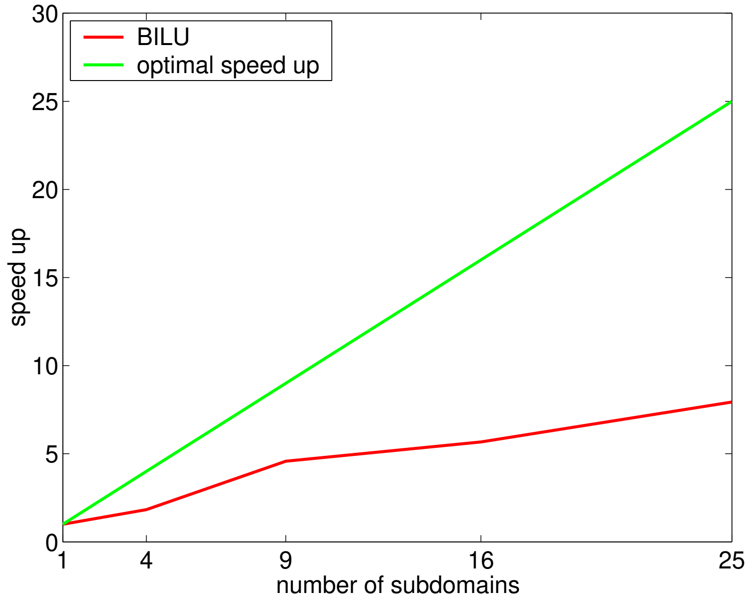
C. Vuik, A. Segal and F. Vermolen
Delft University of Technology

Householder Symposium XV, 2002
Peebles, Scotland
June 17th - 21st, 2002

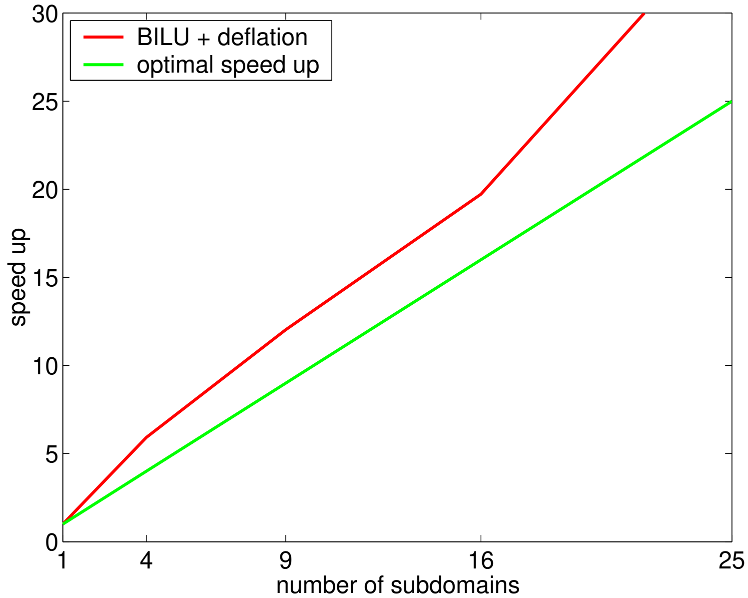
1. Introduction
2. Coarse Grid Correction and Deflation
3. Cell centered approach
4. Vertex centered approach
5. Conclusions

1. Introduction

Speed up for Block ILU

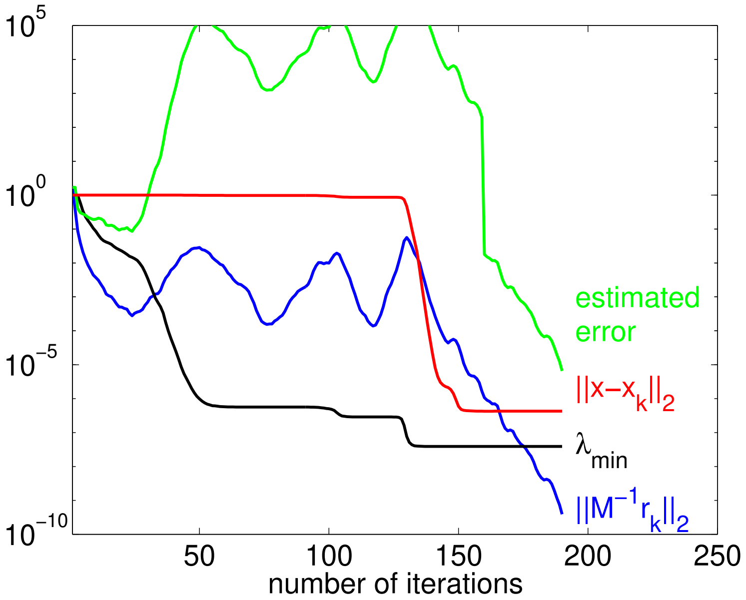


Speed up for Block ILU with deflation

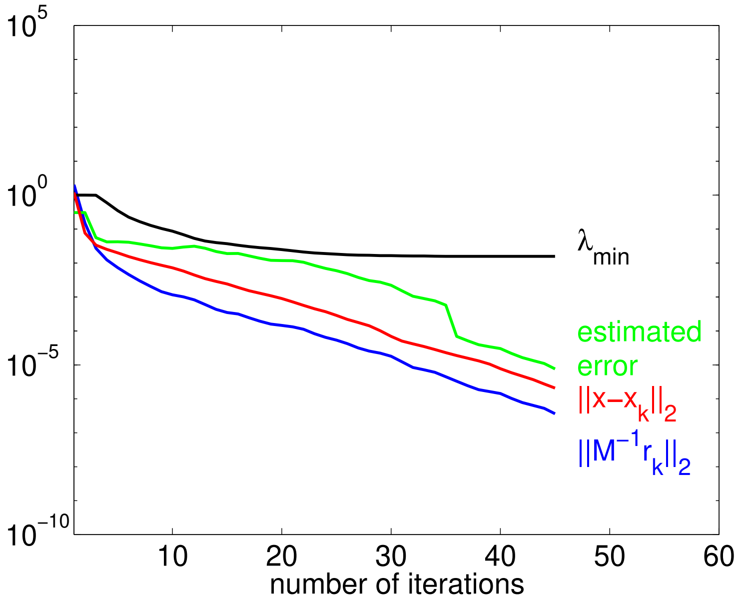


Results oil flow problem

The convergence behavior for ICCG



The convergence behavior for DICCG



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 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

2. Coarse Grid Correction and Deflation

Solve $Ax = b$, $A \in \mathbb{R}^{n \times n}$ and A is SPD

Definition

- $Z \in \mathbb{R}^{n \times m}$ with independent columns.
- $A_0 = Z^T A Z \in \mathbb{R}^{m \times m}$, A_0 is SPD.
- $P_C = L^{-T} L^{-1} + \gamma Z A_0^{-1} Z^T$.

CICCG

$k = 0,\ r_0 = b - Ax_0,\ p_1 = z_1 = L^{-T} L^{-1} r_0;$

while $\|r_k\|_2 > \varepsilon$ **do**

$$k = k + 1;$$
$$\alpha_k = \frac{(r_{k-1}, z_{k-1})}{(p_k, Ap_k)};$$
$$x_k = x_{k-1} + \alpha_k p_k;$$
$$r_k = r_{k-1} - \alpha_k A p_k;$$
$$z_k = P_C r_k = L^{-T} L^{-1} r_k + \gamma Z A_0^{-1} Z^T r_k;$$
$$\beta_k = \frac{(r_k, z_k)}{(r_{k-1}, z_{k-1})}; \qquad p_{k+1} = z_k + \beta_k p_k;$$

end while

Deflated ICCG

Definition: $P_D = I - AZA_0^{-1}Z^T$.

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

$$(I - P_D^T)x = ZA_0^{-1}Z^T Ax = ZA_0^{-1}Z^T b$$

$$AP_D^T x = P_D Ax = P_D b$$

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**

$k = k + 1;$

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

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

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

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

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

end while

Properties of CGC and Deflation

$P_C = I + ZA_0^{-1}Z^T$ $P_D = I - AZA_0^{-1}Z^T$

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 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 CGC and Deflation

Definition

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

Theorem

- the spectrum of $P_C A$ is

$$\{1 + \lambda_1, \ldots, 1 + \lambda_m, \lambda_{m+1}, \ldots, \lambda_n\}$$

- the spectrum of $P_D A$ is

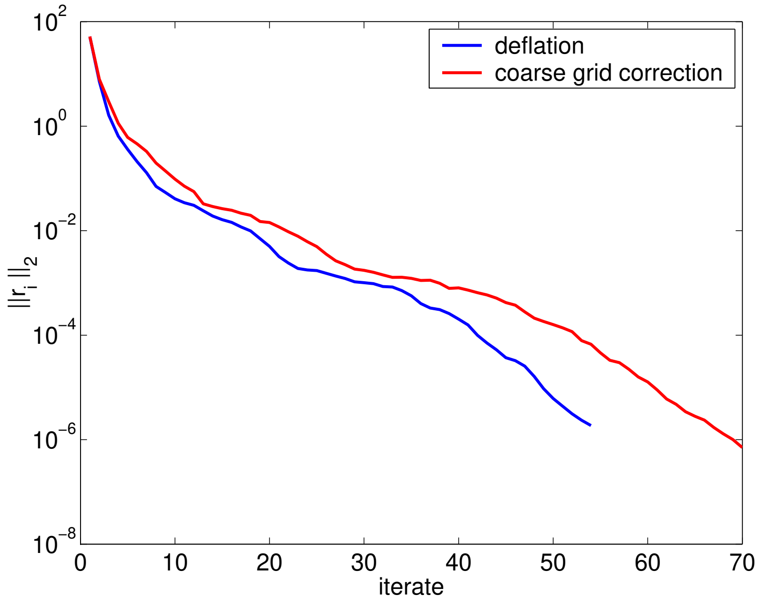
$$\{0, \ldots, 0, \lambda_{m+1}, \ldots, \lambda_n\}$$

Corollary

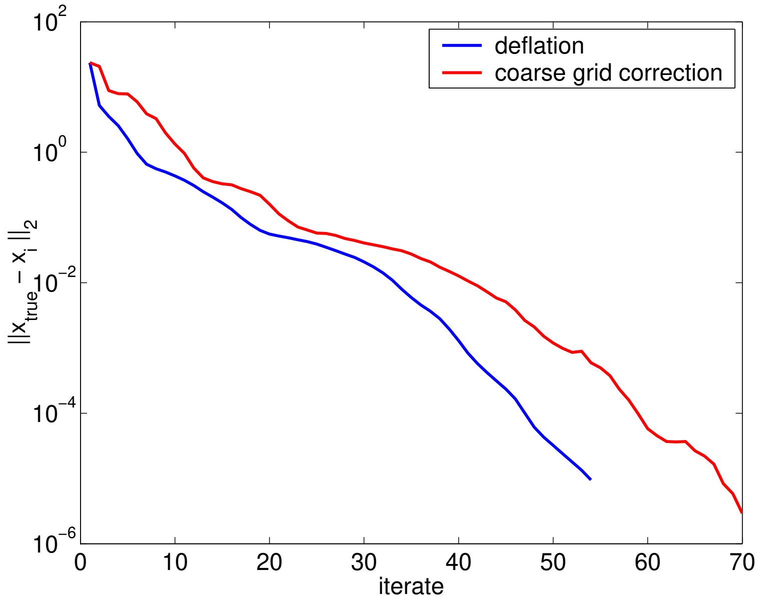
DICCG converges faster than CICCG.

Experiments with CGC and Deflation

Norm of the residual with Block IC

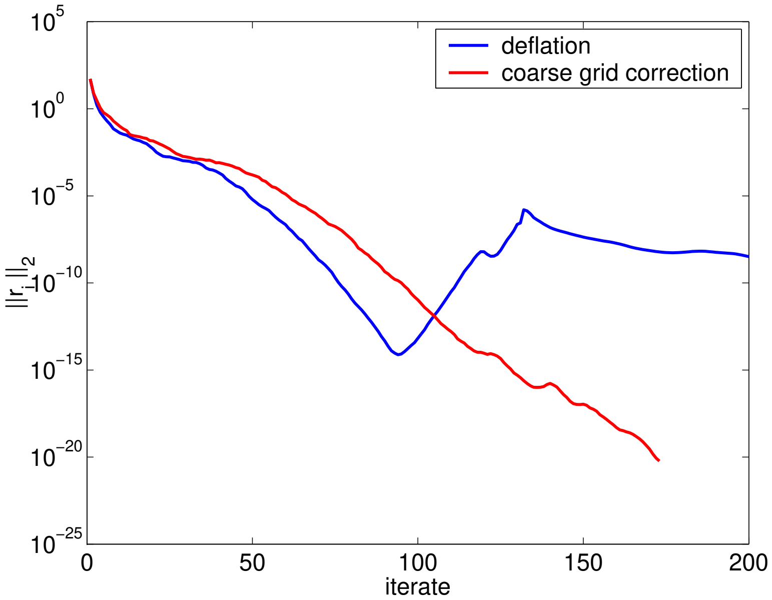


Norm of the error with Block IC

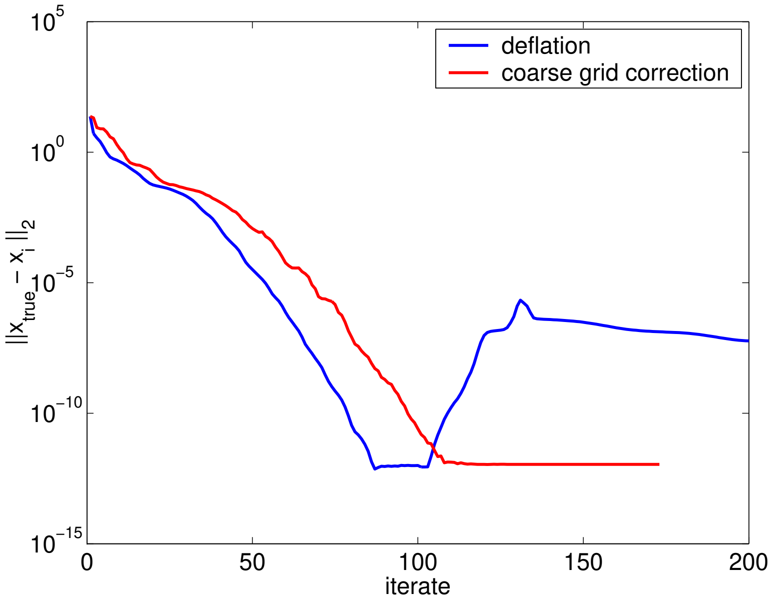


Too severe termination criterion

Norm of the residual with Block IC



Norm of the error with Block IC



3. Cell centered approach

Theorem

Let $A = C + R$ be SPD, and C, R be SPSD, and span$Z$ = null(C) furthermore preconditioner K is SPD and $K = LL^T$ then

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

Choice of deflation vectors

k is number of subdomains

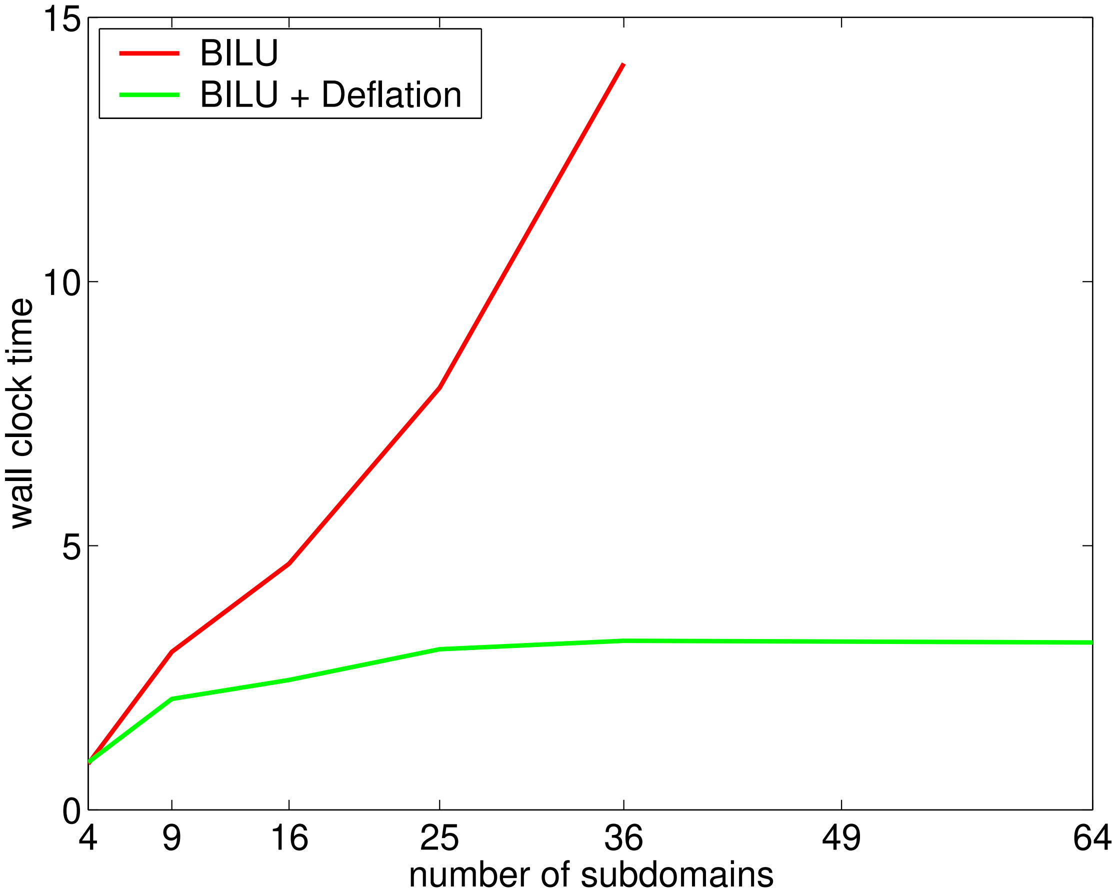
$z_1, \ldots, z_k$ deflation vectors

- $z_i = 1$ on Ω_i
- $z_i = 0$ on $\Omega \setminus \Omega_i$

Parallel scalability

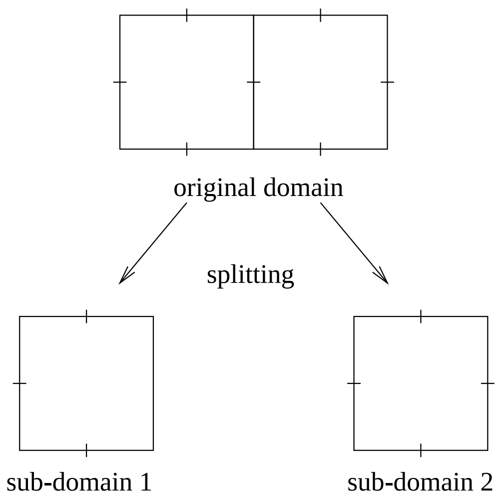
subdomain grid size 50×50

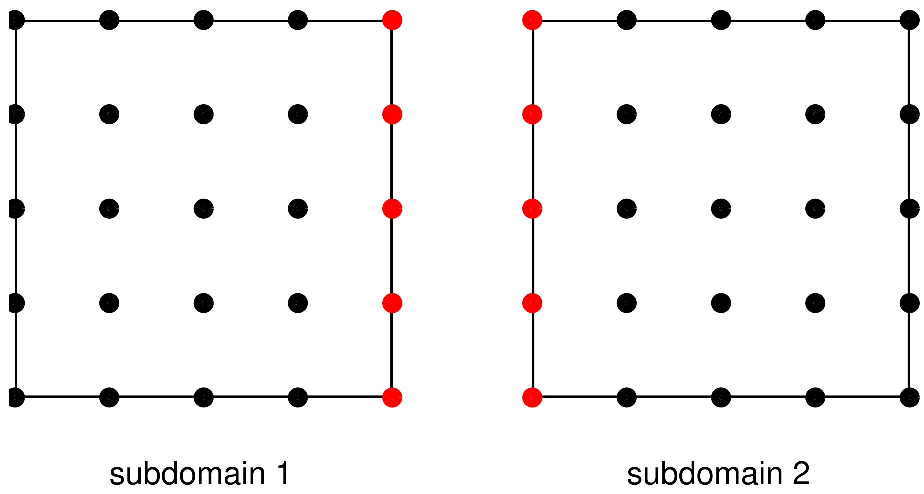
wall clock time, Cray T3E



4. Vertex centered approach

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

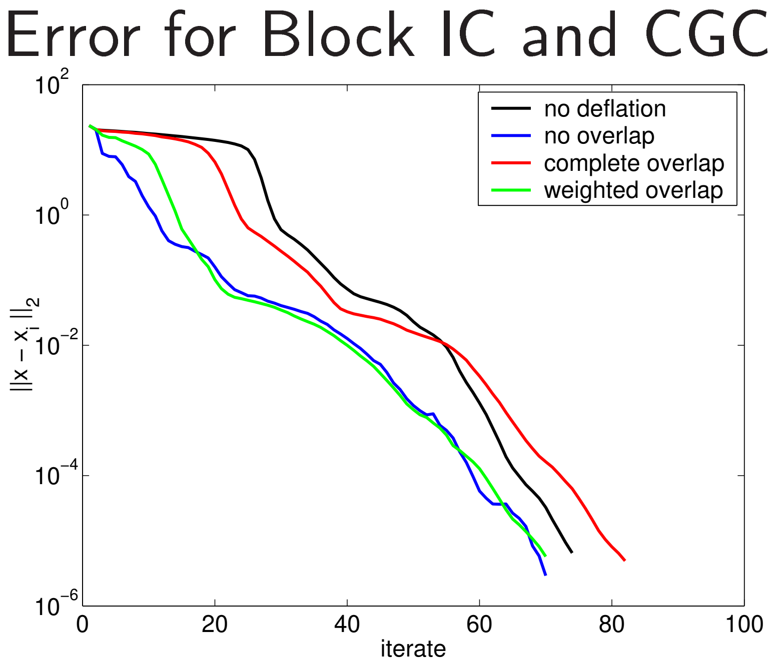
3. **average overlap**

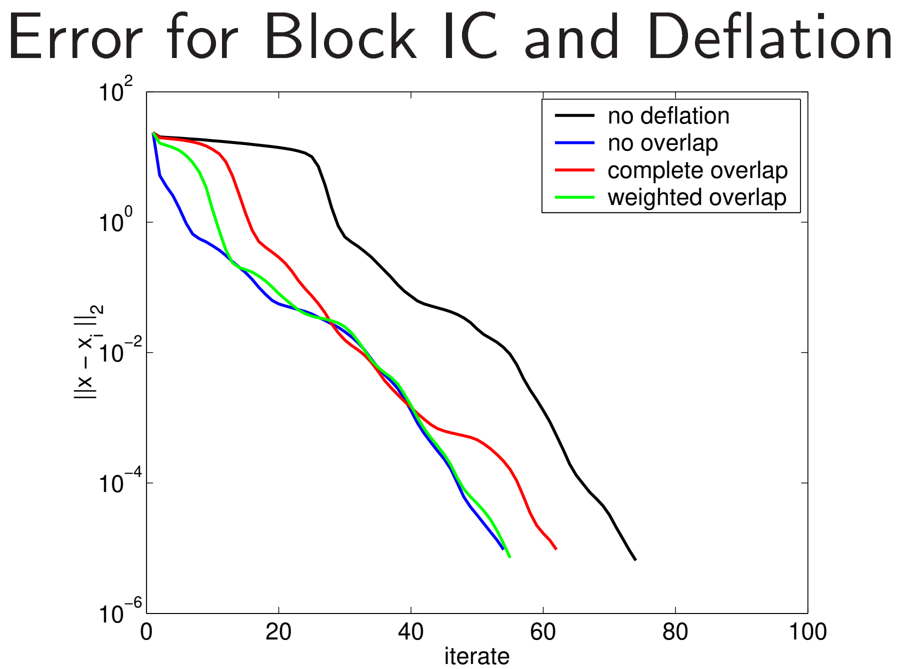
   $z_i = \frac{1}{n_{neighbors}}$ at all subdomains

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

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

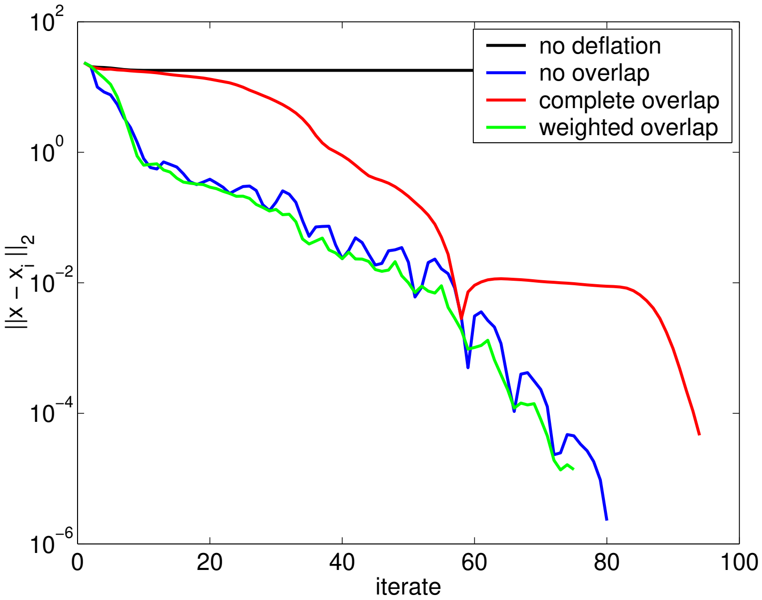
Results for constant coefficients



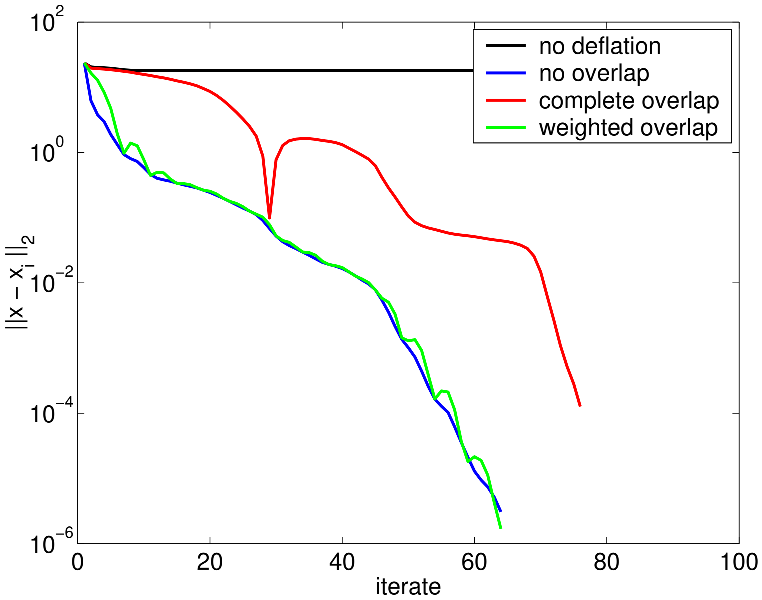


Results for disontinuous coefficients

Error for Block IC and CGC



Error for Block IC and Deflation



5 Conclusions

- Block preconditioned Krylov methods combined with CGC or Deflation are well parallelizable (scalable, good speed up).
- DICCG is more efficient than CICCG.
- Choices for the deflation vectors lead to comparable results in DICCG and CICCG.

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.

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

Overview

Krylov

$$Ar$$

Preconditioned Krylov

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

Block Preconditioned Krylov

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

Deflated Block Preconditioned Krylov

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