Deflation acceleration for a domain decomposition preconditioner

Kees Vuik and Reinhard Nabben

Delft Institute of Applied Mathematics

c.vuik@math.tudelft.nl

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

8th European Multigrid Conference

Scheveningen The Hague, The Netherlands

September 27-30, 2005

Contents

1. Introduction

Efficient solution of a linear system, where A is SPD,

$$Ax = b.$$

Conjugate Gradient, Preconditioner, Coarse Grid Acceleration.

1. Introduction

Efficient solution of a linear system, where A is SPD,

$$Ax = b.$$

Conjugate Gradient, Preconditioner, Coarse Grid Acceleration.

The convergence of CG depends on the effective condition number.

Coarse Grid Acceleration to eliminate the effect of ’bad’ eigenvalues.

1. Introduction

Efficient solution of a linear system, where A is SPD,

$$Ax = b.$$

Conjugate Gradient, Preconditioner, Coarse Grid Acceleration.

The convergence of CG depends on the effective condition number.

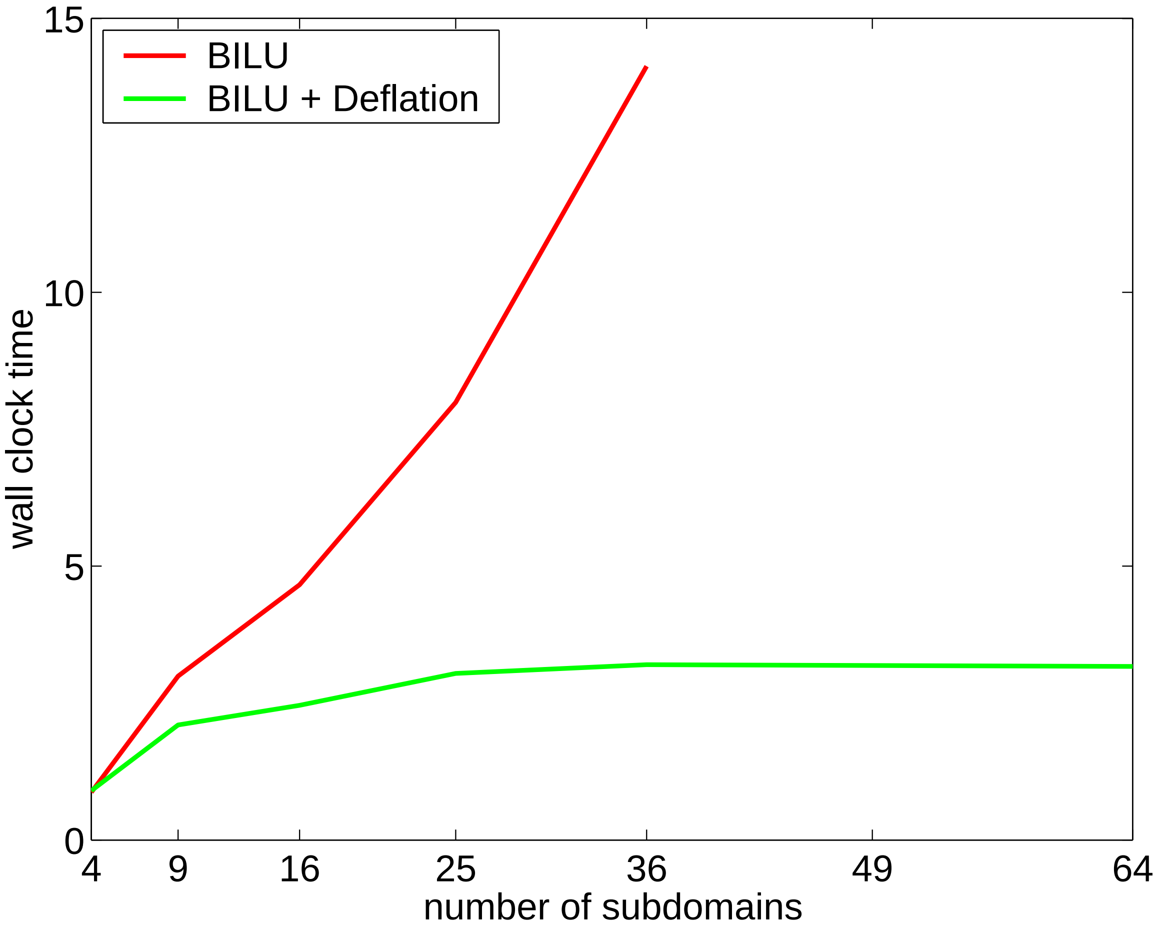
Coarse Grid Acceleration to eliminate the effect of 'bad' eigenvalues.

Motivation

- large jumps in the coefficients
- domain decomposition/block preconditioners (parallel)
- IC preconditioners (serial)

Parallel scalability

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



References

Deflated CG and coarse grid projection vectors

Nicolaides 1987, Mansfield 1990, Kolotilina 1998, Vuik, Segal and Meijerink 1999, Saad, Yeung, Erhel and Guyomarc'h 2000, Frank and Vuik 2001, Nabben and Vuik 2004

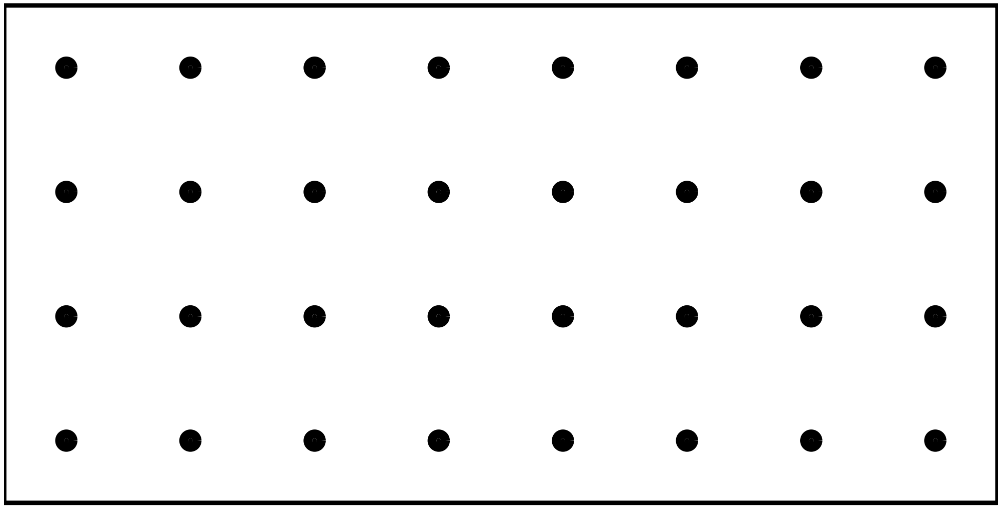
Additive Coarse Grid Correction

Bramble, Pasciak and Schatz 1986, Dryja and Widlund 1991, Smith, Bjorstad and Gropp 1996, Benzi, Frommer, Nabben and Szyld 2001

Balancing (Neumann-Neumann) preconditioner

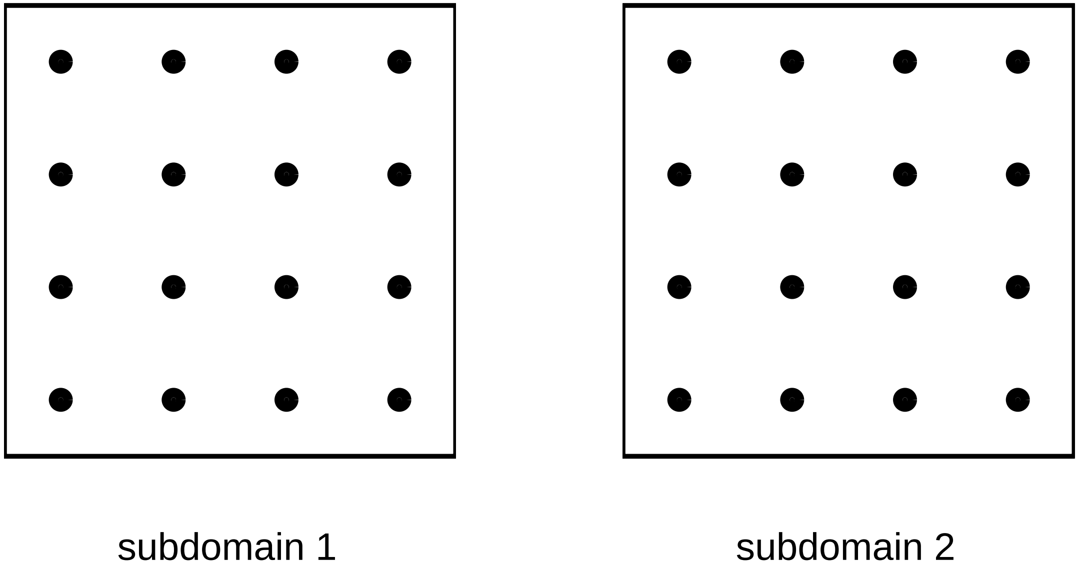
Mandel 1993, Dryja and Widlund 1995, Mandel and Brezina 1996, Pavarino and Widlund 2002

Decomposition of a cell centered domain (FDM and FVM)



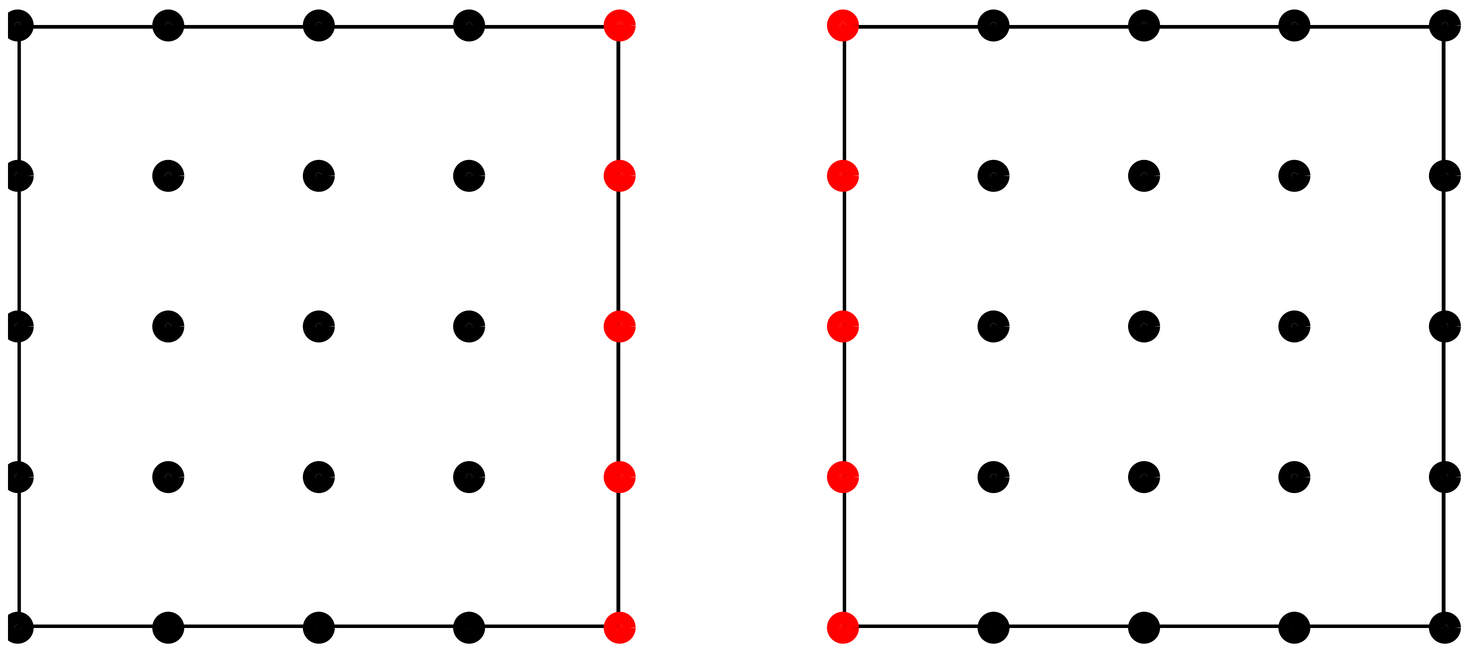
original domain

Decomposition of a cell centered domain (FDM and FVM)



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

Decomposition of a vertex centered domain (FEM)



subdomain 1

subdomain 2

Matrix for cell centered domain

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 $K(A) \in \mathbb{R}^{n \times n}$

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

Preconditioner for cell centered domain

Preconditioner $M \in \mathbb{R}^{n \times n}$

$$M = \begin{bmatrix} M_{11} & & \\ & \ddots & \\ & & M_{mm} \end{bmatrix}$$

where M_{ii} is a rough estimate of A_{ii} (e.g. IC decomposition).

2. Deflation

$$Z \in \mathbb{R}^{n \times r}$$

$$Ax = b, \qquad P_D = I - AZ(Z^T AZ)^{-1} Z^T$$

Note that $P_D A$ is a symmetric, positive semi definite singular matrix.

2. Deflation

$$Z \in \mathbb{R}^{n \times r}$$

$$Ax = b, \qquad P_D = I - AZ(Z^TAZ)^{-1}Z^T$$

Note that P_DA is a symmetric, positive semi definite singular matrix.

We use $x = (I - P_D^T)x + P_D^Tx$

Compute both terms:

1. $(I - P_D^T)x = Z(Z^TAZ)^{-1}Z^TAx = Z(Z^TAZ)^{-1}Z^Tb$,
2. Solve $P_DA\tilde{x} = P_Db$,
3. Form $P_D^T\tilde{x}$ (Theorem: $P_D^Tx = P_D^T\tilde{x}$).

Identity $P_D^T \tilde{x} = P_D^T x$

$$P_D = I - AZ(Z^T AZ)^{-1} Z^T$$

We have

$$AP_D^T = P_D A, \qquad P_D^T A^{-1} = A^{-1} P_D$$

$$AP_D^T \tilde{x} = P_D A \tilde{x} = P_D b$$

$$P_D^T \tilde{x} = A^{-1} A P_D^T \tilde{x} = A^{-1}(P_D A \tilde{x}) = A^{-1}(P_D b) = P_D^T A^{-1} b = P_D^T x$$

$\tilde{x}$ is not unique, but $P_D^T \tilde{x}$ is unique.

Choice of projection vectors

Choose projection vectors equal to coarse grid vectors

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

Enlargement of the projection space helps

Theorem (Nabben and Vuik 2004)
Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite. Let $Z_1 \in \mathbb{R}^{n \times r}$ and $Z_2 \in \mathbb{R}^{n \times s}$ with $rank Z_1 = r$ and $rank Z_2 = s$. Let $E_1 := Z_1^T A Z_1$ and $E_2 := Z_2^T A Z_2$. If $Im Z_1 \subseteq Im Z_2$, then

$$\begin{aligned} \lambda_n((I - AZ_1E_1^{-1}Z_1^T)A) &\geq \lambda_n((I - AZ_2E_2^{-1}Z_2^T)A) \\ \lambda_{r+1}((I - AZ_1E_1^{-1}Z_1^T)A) &\leq \lambda_{s+1}((I - AZ_2E_2^{-1}Z_2^T)A) \end{aligned}$$

Enlargement of the projection space helps

Theorem (Nabben and Vuik 2004)
Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite. Let $Z_1 \in \mathbb{R}^{n \times r}$ and $Z_2 \in \mathbb{R}^{n \times s}$ with $rank Z_1 = r$ and $rank Z_2 = s$. Let $E_1 := Z_1^T A Z_1$ and $E_2 := Z_2^T A Z_2$. If $Im Z_1 \subseteq Im Z_2$, then

$$\begin{aligned} \lambda_n((I - AZ_1E_1^{-1}Z_1^T)A) &\geq \lambda_n((I - AZ_2E_2^{-1}Z_2^T)A) \\ \lambda_{r+1}((I - AZ_1E_1^{-1}Z_1^T)A) &\leq \lambda_{s+1}((I - AZ_2E_2^{-1}Z_2^T)A) \end{aligned}$$

Theorem (Tang and Vuik 2005)

$$\begin{aligned} \lambda_n(A) &\geq \lambda_n((I - AZ_1E_1^{-1}Z_1^T)A) \\ \lambda_1(A) &\leq \lambda_{r+1}((I - AZ_1E_1^{-1}Z_1^T)A) \end{aligned}$$

3. Comparison of Deflation and Additive Coarse Grid Correction

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

$$M^{-1}P_D = M^{-1} - M^{-1}AZE^{-1}Z^T \qquad P_{CM^{-1}} = M^{-1} + \sigma ZE^{-1}Z^T$$

where $E = Z^T AZ$.

Work per iteration:

- 1 matrix vector product
- 1 preconditioner vector product
- 1 coarse grid operator

Comparison of Deflation and Additive Coarse Grid Correction

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

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

- the spectrum of $P_C A$ is $\{\sigma + \lambda_1, \ldots, \sigma + \lambda_r, \lambda_{r+1}, \ldots, \lambda_n\}$

Comparison of Deflation and Additive Coarse Grid Correction

Proof
Note that $P_D A = (I - ZZ^T)A$, so

$$
\begin{aligned}
P_D A v_i &= (I - ZZ^T)\lambda_i v_i = 0, \quad \text{for } 1 \leq i \leq r, \\
P_D A v_i &= (I - ZZ^T)\lambda_i v_i = \lambda_i v_i, \text{ for } r+1 \leq i \leq n.
\end{aligned}
$$

Since $P_C A = (A + \sigma ZZ^T)$ we obtain:

$$
\begin{aligned}
P_C A v_i &= (A + \sigma ZZ^T)v_i = (\lambda_i + \sigma)v_i, \text{ for } 1 \leq i \leq r, \\
P_C A v_i &= (A + \sigma ZZ^T)v_i = \lambda_i v_i, \quad \text{for } r+1 \leq i \leq n.
\end{aligned}
$$

Comparison of Deflation and Additive Coarse Grid Correction

Corollary

$$cond_{eff}(P_D A) = \frac{\lambda_n}{\lambda_{r+1}} \leq \frac{\max\{\lambda_n, \sigma + \lambda_r\}}{\min\{\lambda_{r+1}, \sigma + \lambda_1\}} = cond(P_C A)$$

- The eigenvalues of $P_C A$ has a worse distribution than the eigenvalues of $P_D A$

Conclusion

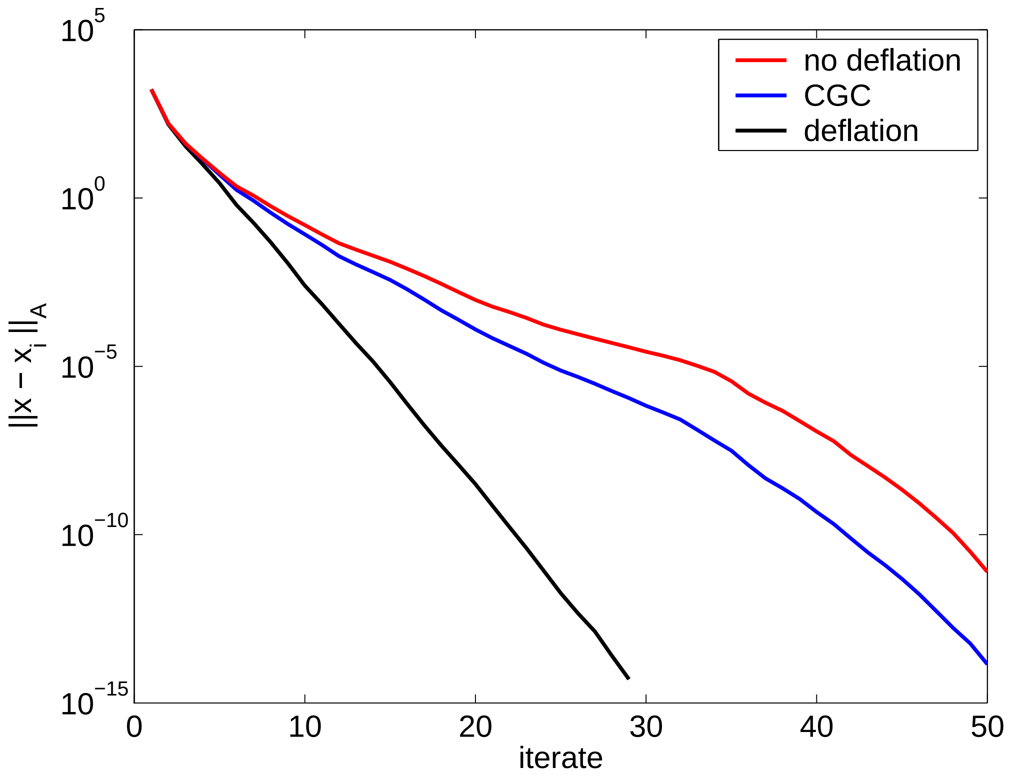
Deflation is asymptotically better than additive coarse grid correction!

This also holds for the preconditioning and general projection vectors

Results for eigenvectors

The eigenvalues of A are $1, 2, 3, \ldots, 99, 100$.

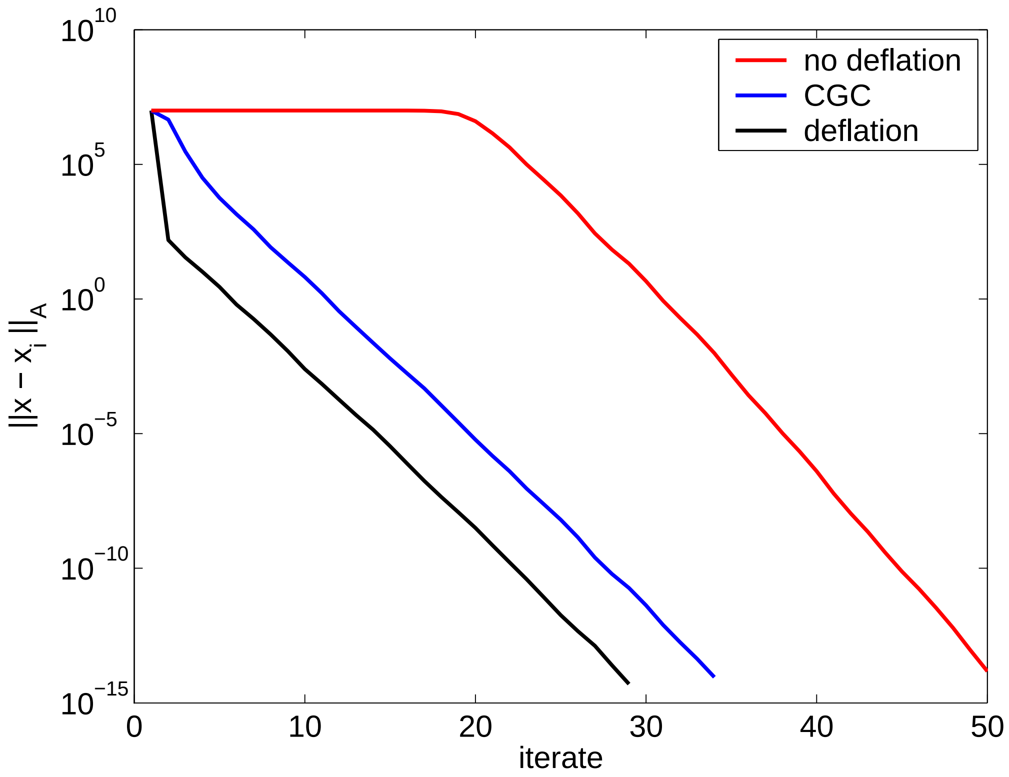
The eigenvectors $v_1, \ldots, v_{10}$ are used as projection vectors.



Results for eigenvectors

The eigenvalues of A are $10^{-6}, \ldots 10^{-6}, 11, 12, 13, \ldots, 99, 100$.

The eigenvectors $v_1, \ldots, v_{10}$ are used as projection vectors.



4. Comparison of Deflation and the Balancing preconditioner

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

$$P_B = (I - ZE^{-1}Z^TA)M^{-1}(I - AZE^{-1}Z^T) + ZE^{-1}Z^T$$

$$P_B = P_D^TM^{-1}P_D + ZE^{-1}Z^T$$

Work per iteration:

	Deflation	Balancing (depends on implementation)
matrix vector product	1	3
preconditioner vector product	1	1
coarse grid operator	1	2

Comparison of Deflation and the Balancing preconditioner

Take $Z = [v_1 \dots v_r]$ and $M = I$.

Theorem

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

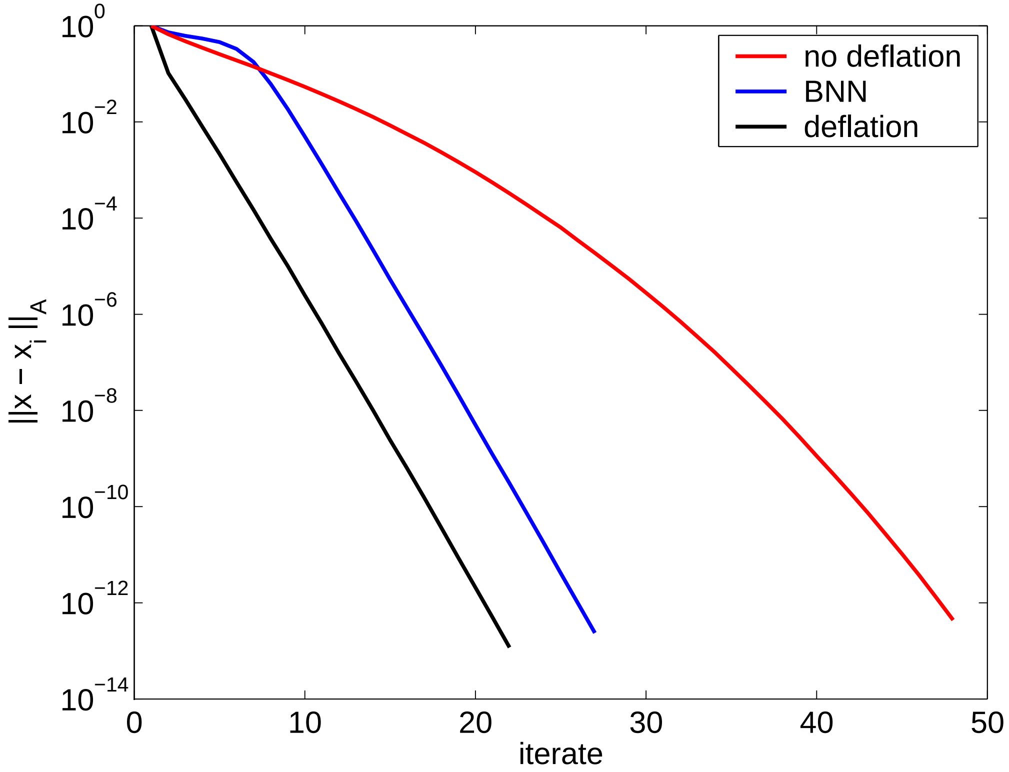
- the spectrum of $P_B A$ is $\{1, \dots, 1, \lambda_{r+1}, \dots, \lambda_n\}$

$$cond_{eff}(P_D A) = \frac{\lambda_n}{\lambda_{r+1}} \leq \frac{\max\{\lambda_n, 1\}}{\min\{\lambda_{r+1}, 1\}} = cond(P_B A)$$

Deflation is asymptotically better than the Balancing preconditioner!

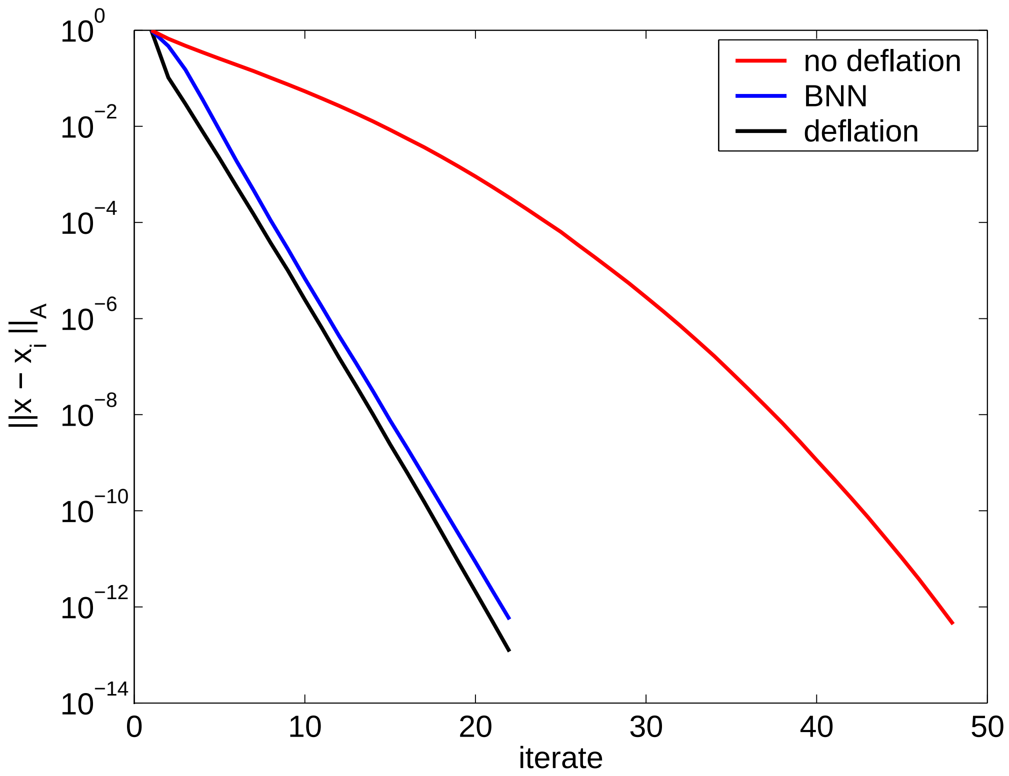
Results for eigenvectors $v_1, \ldots, v_{10}$

The eigenvalues of A are $1, 2, 3, \ldots, 99, 100$.



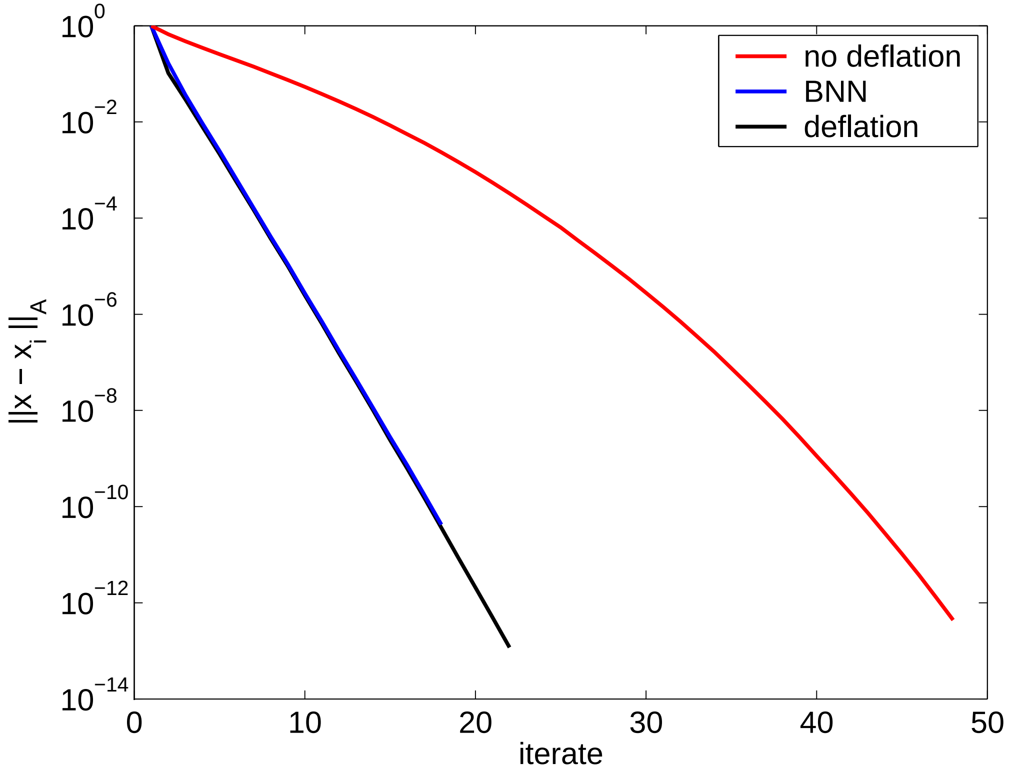
Results for eigenvectors $v_1, \ldots, v_{10}$

The eigenvalues of A are $0.1, 0.2, 0.3, \ldots, 9.9, 10.$



Results for eigenvectors $v_1, \ldots, v_{10}$

The eigenvalues of A are $0.01, 0.02, 0.03, \ldots, 0.99, 1$.



5. Deflation and multi grid?

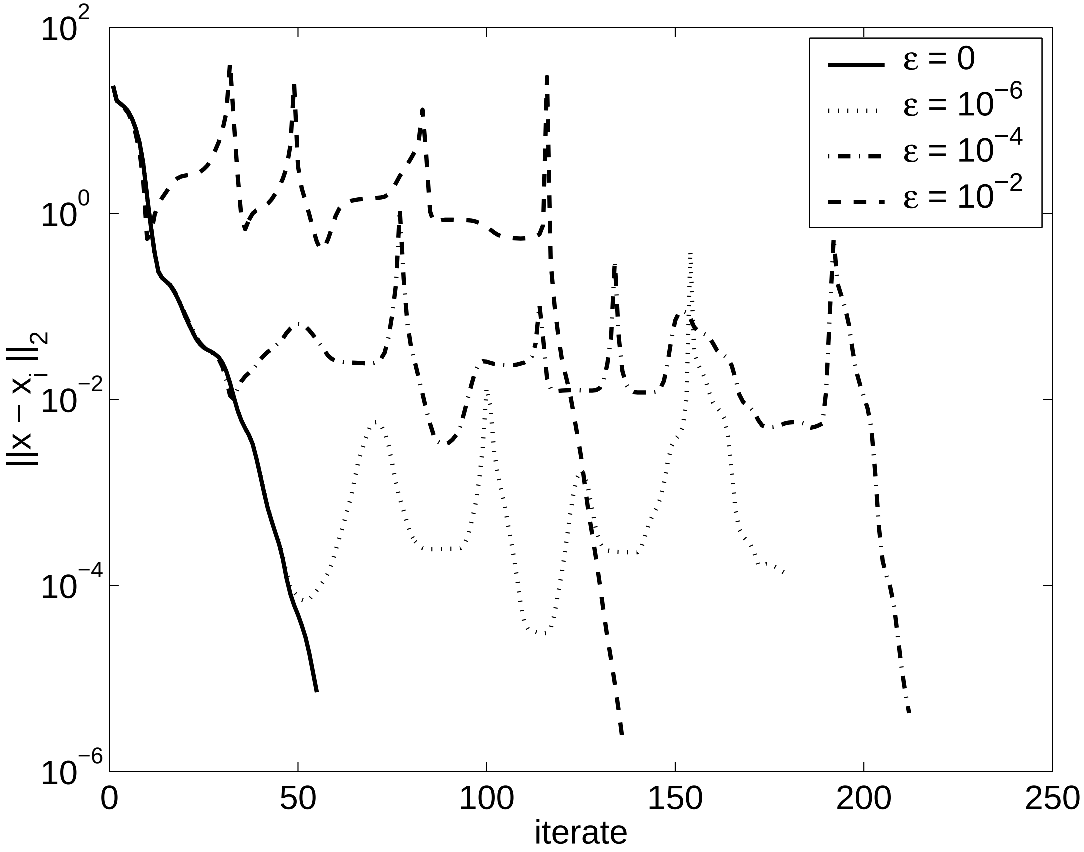
Is it possible to generalize the two grid Deflation approach to a multi grid Deflation?

Idea repeat the procedure on the small matrix E. Required accuracy of the inner iteration?

We replace E^{-1} by $\tilde{E}^{-1} = (I + \epsilon R) E^{-1} (I + \epsilon R)$, where R is a symmetric $r \times r$ matrix with random elements chosen from the interval $[-\frac{1}{2}, \frac{1}{2}]$.

Deflation and multi grid?

Poisson problem with 7 projection vectors



6. Conclusions

- Block preconditioned Krylov methods combined with Deflation, CGC, or BNN are well parallelizable (scalable, good speed up)
- Deflation needs less iterations than additive coarse grid correction, and uses the same amount of work per iteration
- Deflation uses less (approximately the same) iterations as the Balancing preconditioner, but uses less work per iteration.
- Generalization of two grid Deflation to multi grid Deflation is not straightforward.

Further information

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

Further information

- http://ta.twi.tudelft.nl/nw/users/vuik/pub_it_def.html
- C. Vuik, A. Segal and 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
- R. Nabben and C. Vuik
  A comparison of Deflation and Coarse Grid Correction applied to porous media flow
  SIAM J. on Numerical Analysis, 42, pp. 1631-1647, 2004
- R. Nabben and C. Vuik
  A comparison of Deflation and the Balancing preconditioner
  SIAM Journal on Scientific Computing, to appear