Deflation acceleration of the PCG method applied to porous media flow

C. Vuik

Delft Institute of Applied Mathematics

c.vuik@tudelft.nl

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

Computation of flow and transport in heterogeneous media

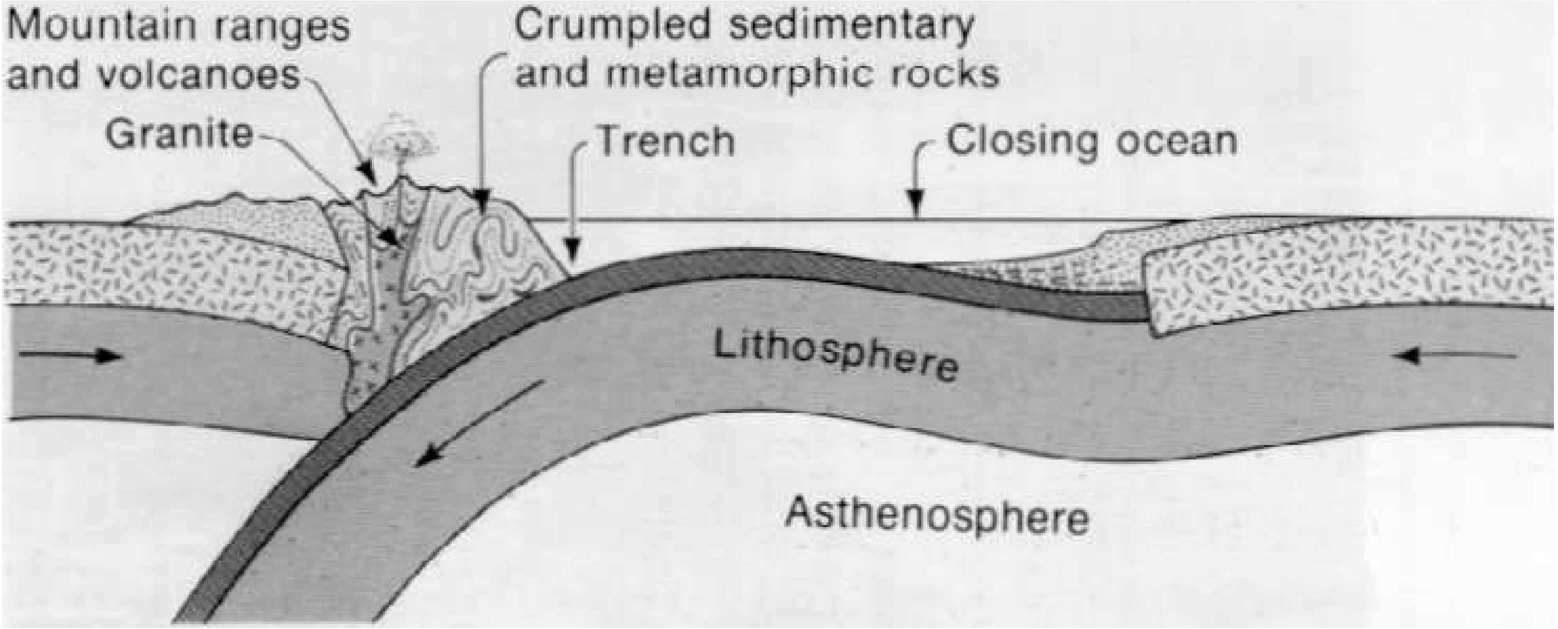
June 19-20, 2006, Bath, UK

Contents

1. Introduction

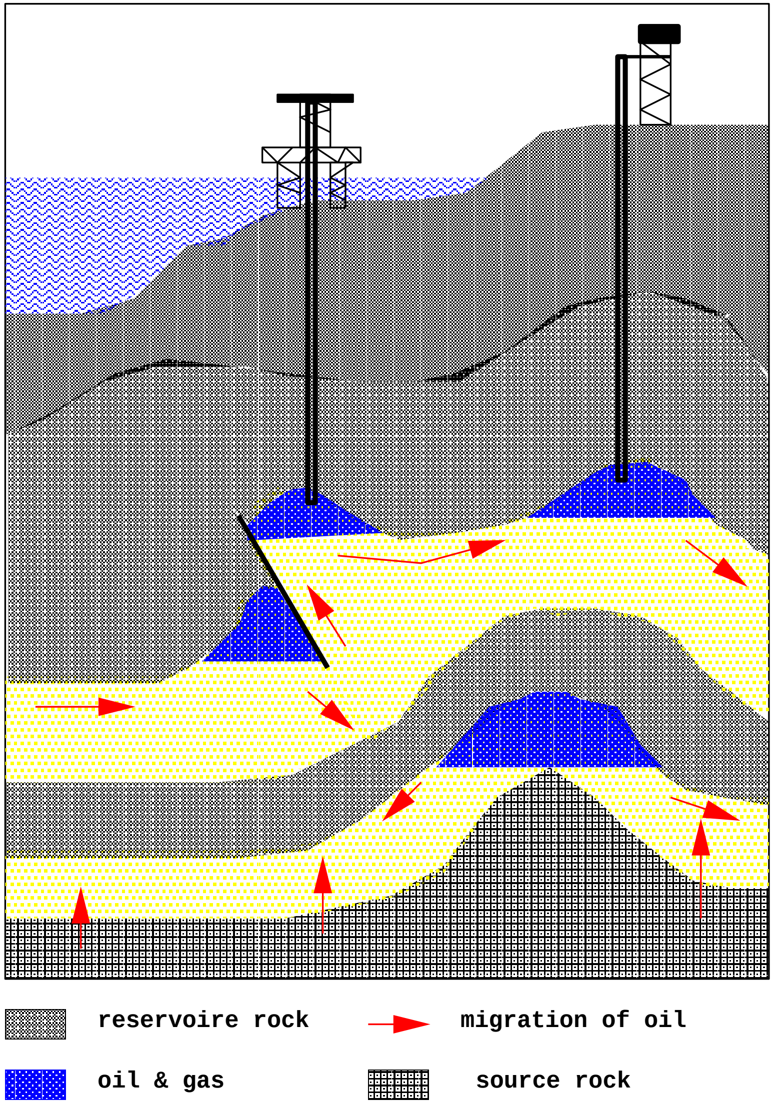
Motivation

Knowledge of the fluid pressure in rock layers is important for an oil company to predict the presence of oil and gas in reservoirs.



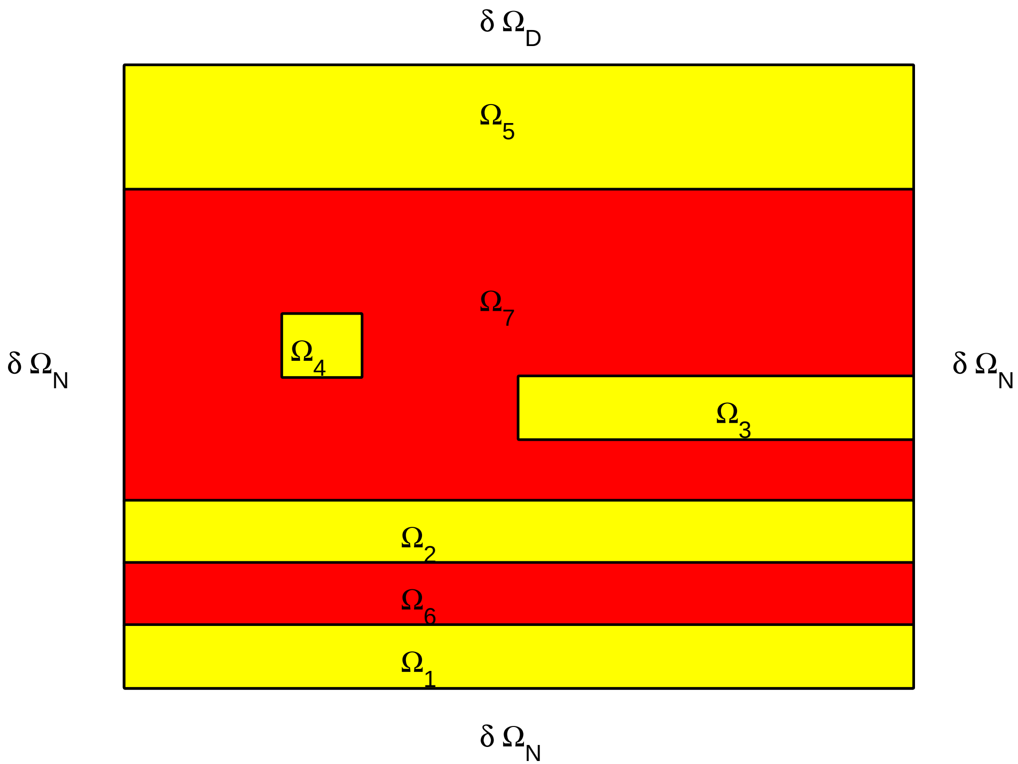
The earth's crust has a layered structure

Drilling for oil



Mathematical model

Computation of fluid pressure $-\text{div}(\sigma \nabla p(x)) = 0$ on Ω, p fluid pressure, σ permeability



$\sigma_h = 1$ (sand) $\sigma_l = \varepsilon = 10^{-7}$ (shale)

Properties and Applications

$$Ax = b$$

A is sparse and SPD

Condition number of A is $O(10^7)$, due to large contrast in permeability

Applications

- reservoir simulations
- porous media flow
- electrical power networks
- semiconductors
- magnetic field simulations

- fictitious domain methods

2. IC preconditioned CG

Error estimate

$$Ax = b$$

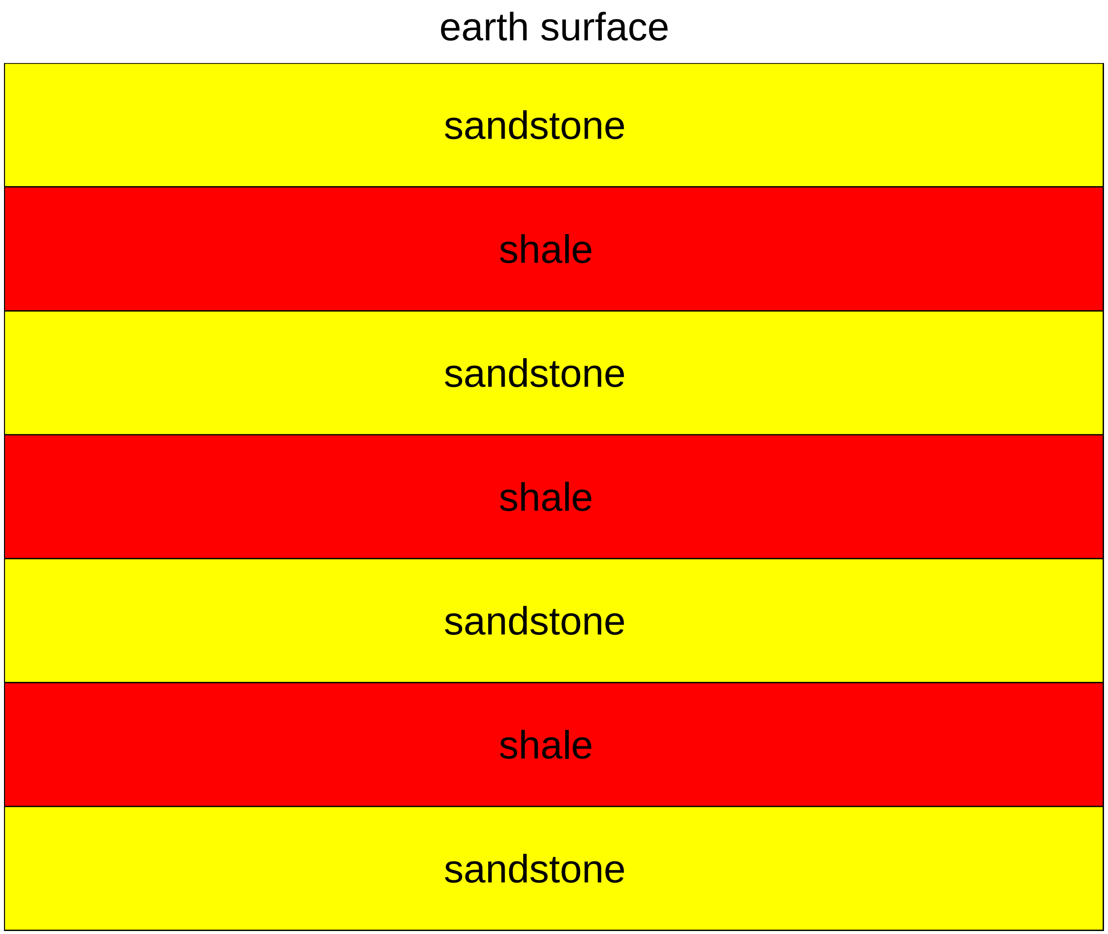
$$M^{-1}Ax = M^{-1}b$$

$$x - x_k = (M^{-1}A)^{-1}M^{-1}A(x - x_k)$$

$$\|x - x_k\|_2 \leq \frac{1}{\lambda_{min}}\|M^{-1}r_k\|_2$$

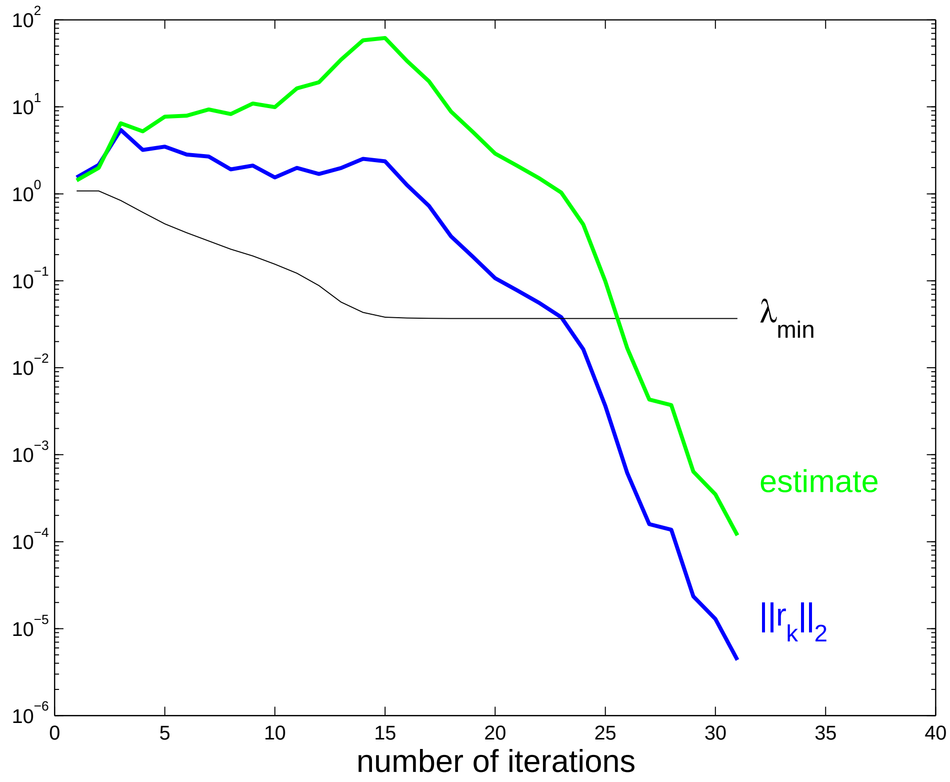
λ_{min}: smallest eigenvalue of $M^{-1}A$

Test problem



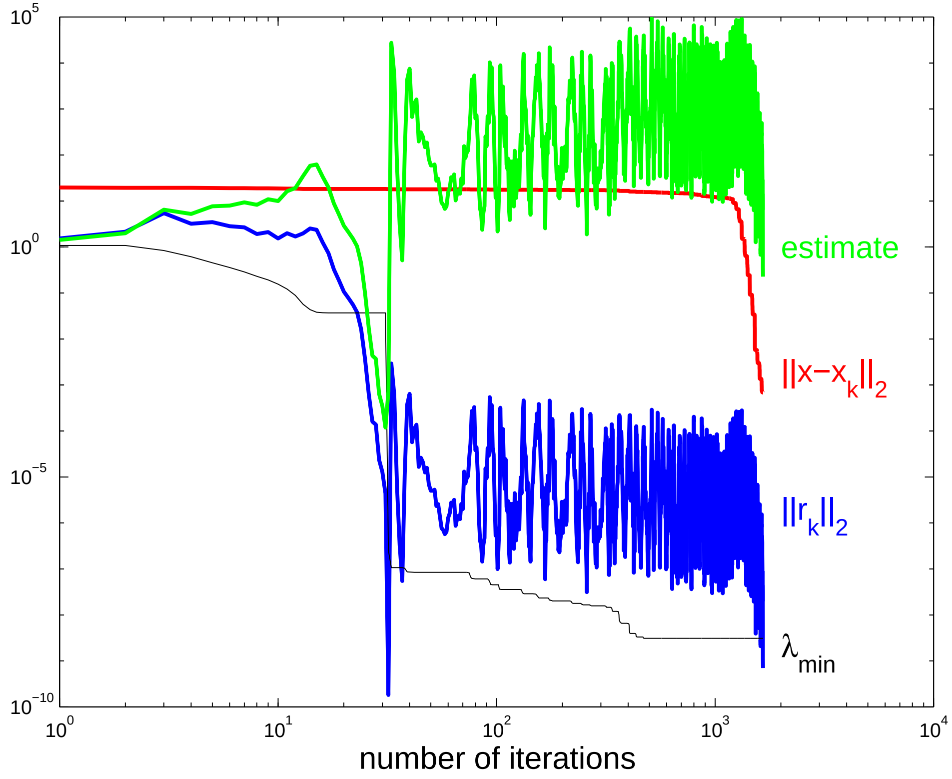
Configuration with 7 straight layers

Convergence CG



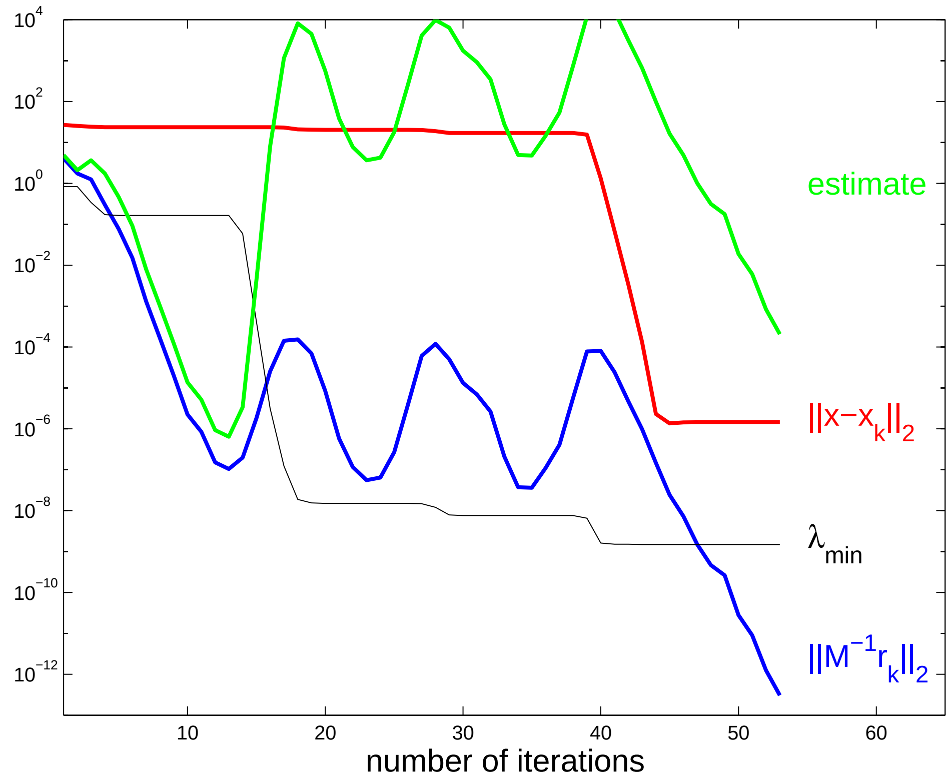
Convergence behavior of CG without preconditioning

Convergence CG



Convergence behavior of CG without preconditioning

Convergence ICCG



Convergence behavior of ICCG

Spectrum of IC preconditioned matrix

L is the Incomplete Cholesky factor of A

k^s is the number of high-permeability domains not connected to a Dirichlet boundary

D is a diagonal matrix $(d_{ii} > 0)$ and $\hat{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$

Theorem 1 (scaling invariance)
$L^{-1} A L^{-T}$ and $\hat{L}^{-1} \hat{A} \hat{L}^{-T}$ are identical.

Proof:

$\hat{L} = D^{-\frac{1}{2}} L$ and $\hat{L}^{-1} \hat{A} \hat{L}^{-T} = L^{-1} D^{\frac{1}{2}} (D^{-\frac{1}{2}} A D^{-\frac{1}{2}}) D^{\frac{1}{2}} L^{-T} = L^{-1} A L^{-T}$.

Spectrum of IC preconditioned matrix

Take $D =$diag(A)

Theorem 2

$\hat{A}$ has k^s eigenvalues of $O(\varepsilon)$, where ε is the ratio between high and low permeability.

Theorem 3

The IC preconditioned matrix $L^{-1}AL^{-T}$ has k^s eigenvalues of $O(\varepsilon)$.

Proof: Scaling invariance (Theorem 1) implies

$$\text{spectrum}(L^{-1}AL^{-T}) = \text{spectrum}(\hat{L}^{-1}\hat{A}\hat{L}^{-T})$$

In [Vuik, Segal, Meijerink, Wijma, 2001] we have shown that the number and size of small eigenvalues of $\hat{A}$ and $\hat{L}^{-1}\hat{A}\hat{L}^{-T}$ are the same. The theorem is proven by using Theorem 2. ☒

3. Deflated ICCG

Idea: remove the bad eigenvectors from the error/residual.

Krylov	Ar
Preconditioned Krylov	$M^{-1}Ar$
Block Preconditioned Krylov	$\sum_{i=1}^{m}(M_i^{-1})Ar$
Block Preconditioned Deflated Krylov	$\sum_{i=1}^{m}(M_i^{-1})PAr$

3. Deflated ICCG

Idea: remove the bad eigenvectors from the error/residual.

Various choices are possible:

- Projection vectors
  Physical vectors, eigenvectors, coarse grid projection vectors (constant, linear, ...)
- Projection method
  Deflation, coarse grid projection, balancing, augmented, FETI
- Implementation
  sparseness, with(out) using projection properties, optimized, ...

Literature

Deflated CG and coarse grid projection vectors

Nicolaides 1987, Mansfield 1990, Graham and Hagger 1997, 1999, Kolotilina 1998, Vuik, Segal and Meijerink 1999, Saad, Yeung, Erhel and Guyomarc'h 2000, Frank and Vuik 2001, Rodriguez, Klie and Wheeler 2006, Nabben and Vuik 2004, 2006

Literature

Deflated CG and coarse grid projection vectors

Nicolaides 1987, Mansfield 1990, Graham and Hagger 1997, 1999, Kolotilina 1998, Vuik, Segal and Meijerink 1999, Saad, Yeung, Erhel and Guyomarc'h 2000, Frank and Vuik 2001, Rodriguez, Klie and Wheeler 2006, Nabben and Vuik 2004, 2006

Additive Coarse Grid Correction

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

Literature

Deflated CG and coarse grid projection vectors

Nicolaides 1987, Mansfield 1990, Graham and Hagger 1997, 1999, Kolotilina 1998, Vuik, Segal and Meijerink 1999, Saad, Yeung, Erhel and Guyomarc'h 2000, Frank and Vuik 2001, Rodriguez, Klie and Wheeler 2006, Nabben and Vuik 2004, 2006

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

Deflated ICCG

A is SPD, Conjugate Gradients

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

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

Properties

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

Deflated ICCG

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

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

Deflated ICCG

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

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

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

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

Convergence and termination criterion

Choose z_1 , z_2 , z_3 eigenvectors of $L^{-T}L^{-1}A$

Convergence

$$\|P^T x - P^T x_k\|_2 \leq 2\sqrt{K}\|P^T x - P^T x_0\|_2 \left(\frac{\sqrt{K}-1}{\sqrt{K}+1}\right)^k$$

where $K = \frac{\lambda_n}{\lambda_4}$

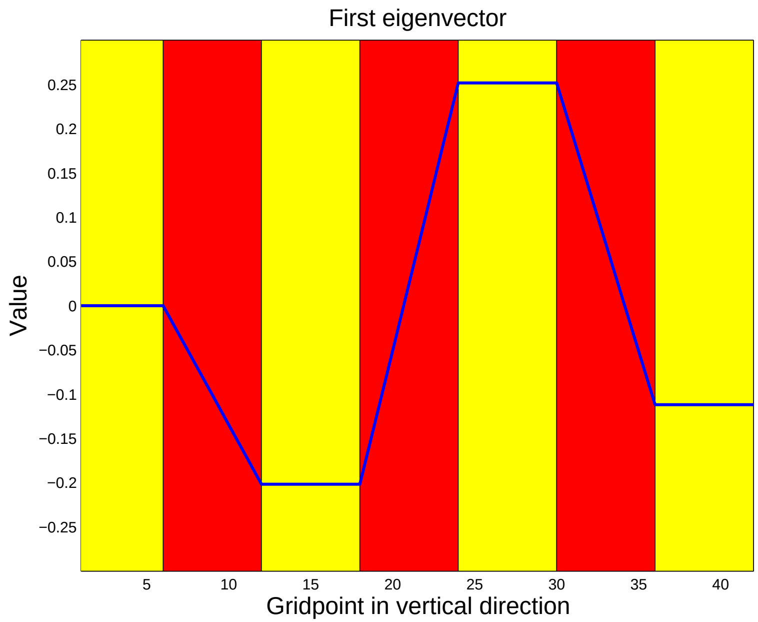
Termination criterion

$$\|L^{-T}L^{-1}Pb - L^{-T}L^{-1}PAx_k\|_2 \leq \frac{\delta}{\lambda_4} \text{ implies } \|P^T x - P^T x_k\|_2 \leq \delta$$

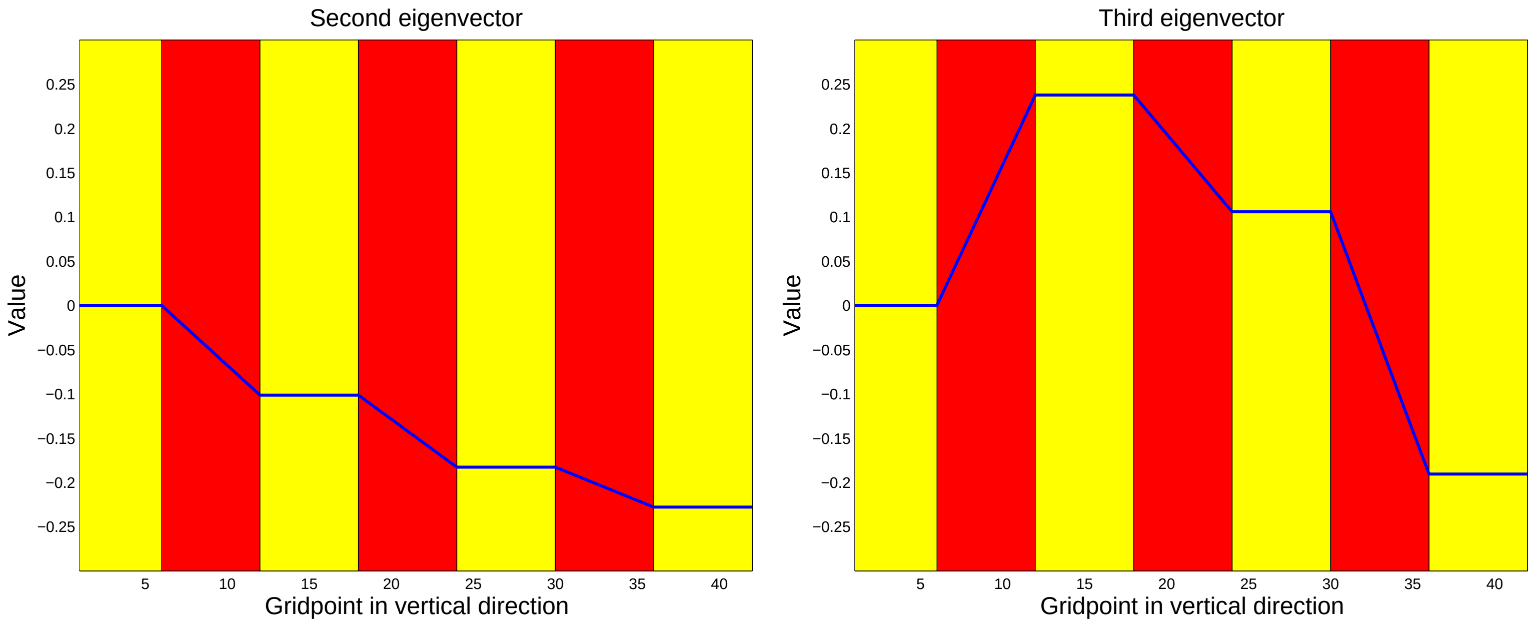
Deflation vectors

Choose eigenvectors of $L^{-T}L^{-1}A$. Properties of cross sections:

- a constant value in sandstone layers
- in shale layers their graph is linear



Eigenvectors of $L^{-T}L^{-1}A$



TU Delft

4. Physical deflation vectors

k is number of subdomains

$\Omega_i, i = 1, ..., k^s$ high-permeability subdomains without a Dirichlet b.c.;
$i = k^s + 1, ..., k^h$ remaining high-permeability subdomains

- define z_i for $i \in \{1, ..., k^s\}$
- $z_i = 1$ on $\bar{\Omega}_i$ and $z_i = 0$ on $\bar{\Omega}_j, j \neq i, j \in \{1, ..., k^h\}$
- z_i satisfies equation:

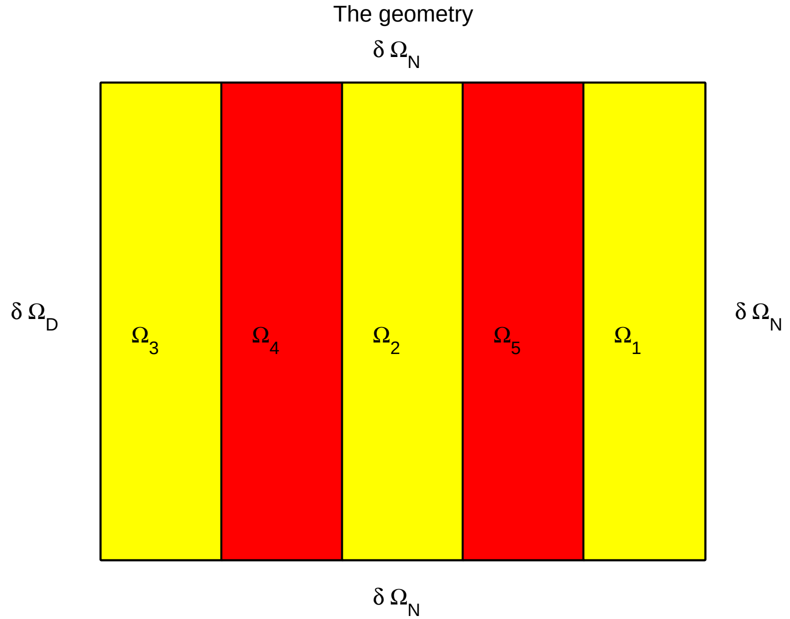
$$-\text{div}(\sigma_j \nabla z_i) = 0 \text{ on } \Omega_j, j \in \{k^h + 1, ..., k\},$$

with appropriate boundary conditions

Sparse vectors, subproblems are cheap to solve

Physical deflation vectors

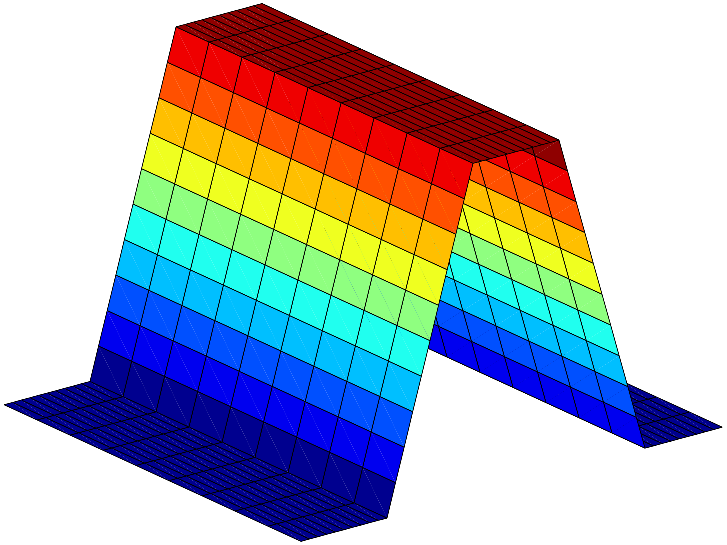
Example with $k_s = 2, k_h = 3$, and $k = 5$



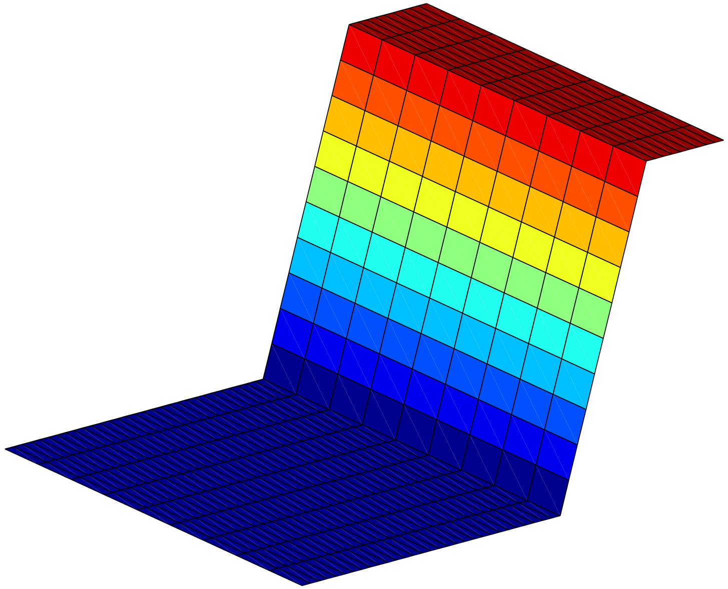
Physical deflation vectors

Example with $k_s = 2, k_h = 3$, and $k = 5$

The first projection vector



The second projection vector



Properties

Theorem 4

The deflation vectors are such that for $D = \text{diag}(A)$

- $\|D^{-1}Az_i\|_\infty = O(\varepsilon)$
- $\|L^{-T}L^{-1}Az_i\|_2 = O(\varepsilon)$

Define $Z = [z_1 ... z_{k^s}]$ and $U = [u_1 ... u_{k^s}]$, where u_i are 'small' eigenvectors.

Theorem 5

There is a matrix X such that $Z = UX + E$, with $\|E\|_2 = O(\sqrt{\varepsilon})$

Sensitivity of deflation vectors

- Random vector added in shale layers (amplitude $\alpha/2$)

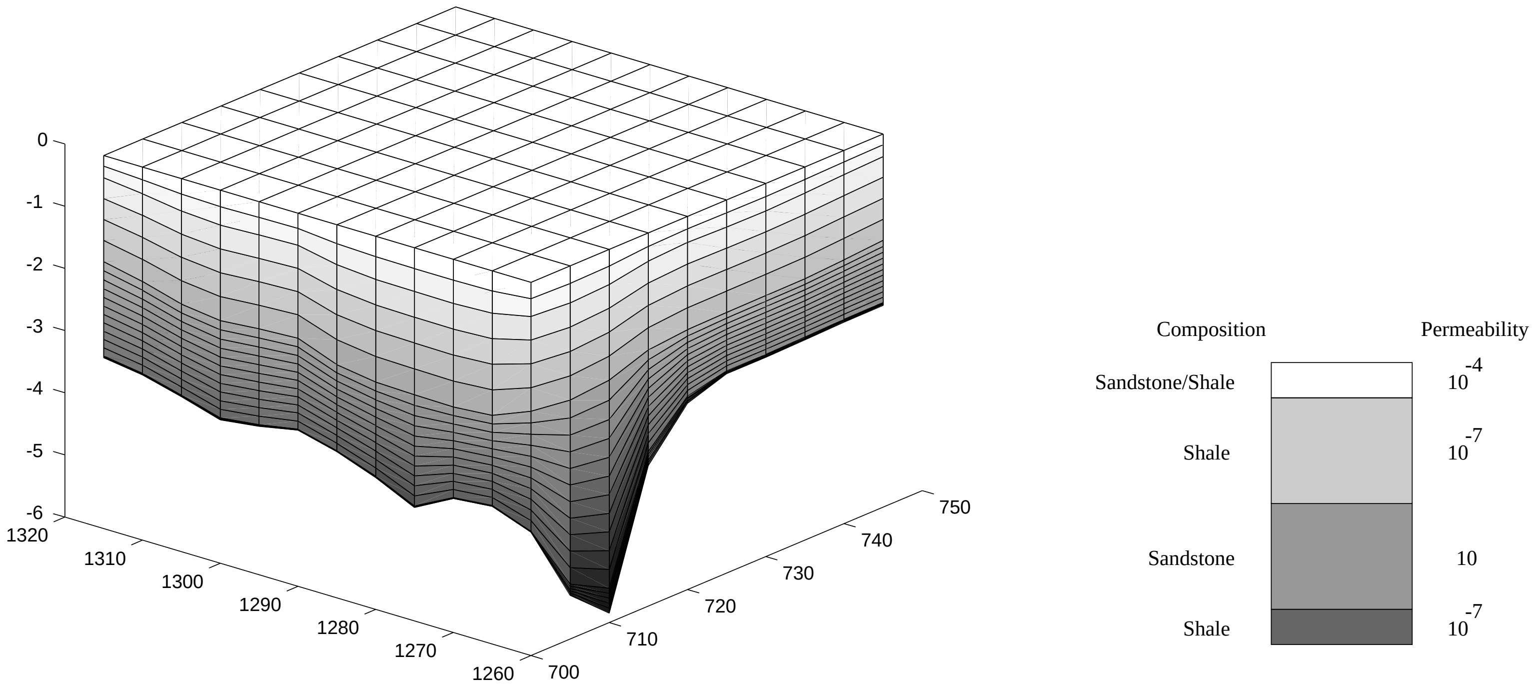
α	0	10^{-1}	1	ICCG
λ_{per}	0.164	0.164	$8.2 \cdot 10^{-3}$	$1.6 \cdot 10^{-9}$
iter	14	15	24	54

- Random vector added to the nonzero parts

α	0	10^{-3}	10^{-1}	ICCG
λ_{per}	0.164	$9 \cdot 10^{-4}$	$9 \cdot 10^{-8}$	$1.6 \cdot 10^{-9}$
iter	14	27	56	54

After perturbation the smallest eigenvalues remain exactly zero, however, the smallest non-zero eigenvalue can change considerably.

Geometry oil flow problem



Results oil flow problem

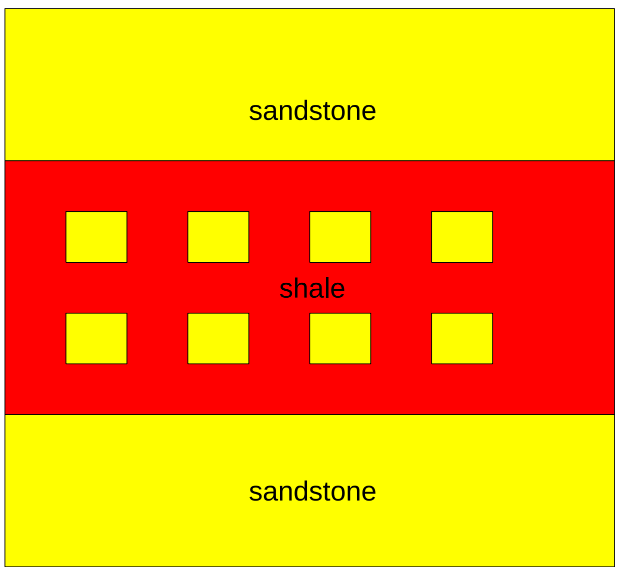
Varying σ_{shale}

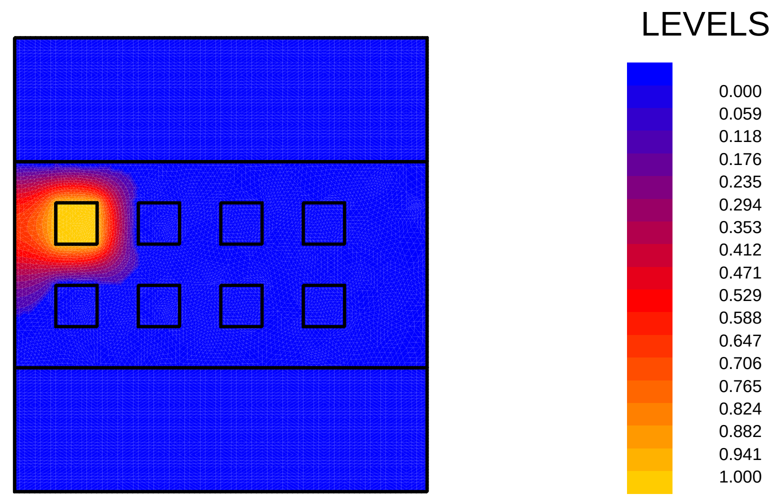
σ	ICCG		DICCG	
	λ_{min}	iter	λ_{min}	iter
10^{-3}	$1.5 \cdot 10^{-2}$	26	$6.9 \cdot 10^{-2}$	20
10^{-5}	$2.2 \cdot 10^{-4}$	59	$7.7 \cdot 10^{-2}$	20
10^{-7}	$2.3 \cdot 10^{-6}$	82	$7.7 \cdot 10^{-2}$	20

Varying accuracy

accuracy	ICCG		DICCG	
	iter	CPU	iter	CPU
10^{-5}	82	18.9	20	6.3
10^{-3}	78	18.0	12	4.1
10^{-1}	75	17.2	2	1.2

Inclusions





A

problem with sandstone inclusions

Inclusions

Varying *droptol*

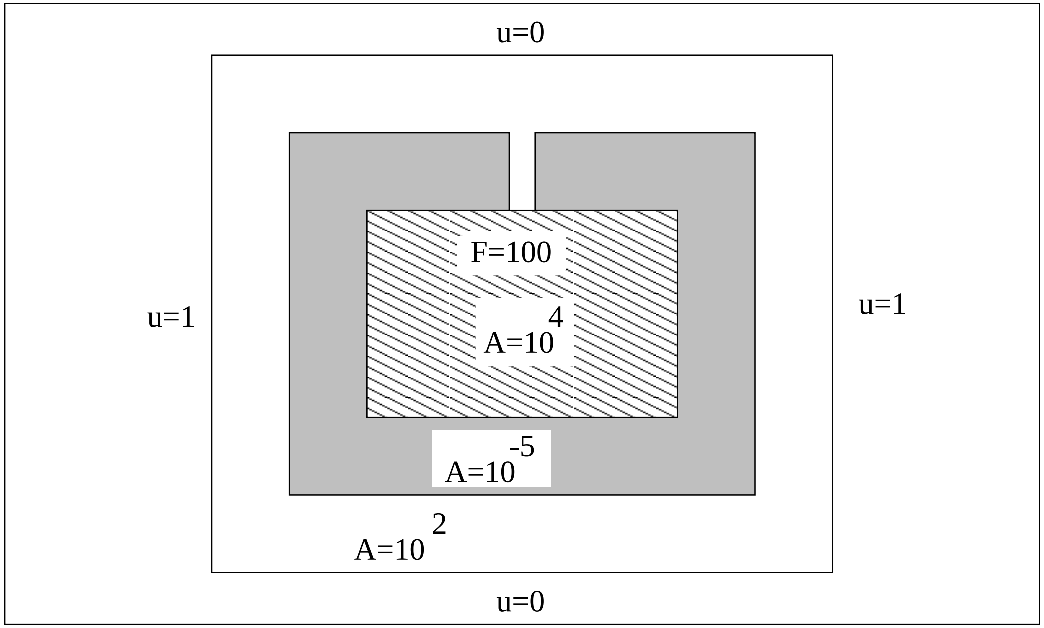
	ICCG	DICCG		
droptol		0	10^{-2}	10^{-1}
CPU	44	12	8.3	8.1
iterations	616	76	76	76
$\lambda_{\min}$	10^{-9}	10^{-3}	10^{-3}	10^{-3}
memory	0	$3.9n$	$1.6n$	$1.2n$

A groundwater flow problem

The pressure in groundwater satisfies the equation:

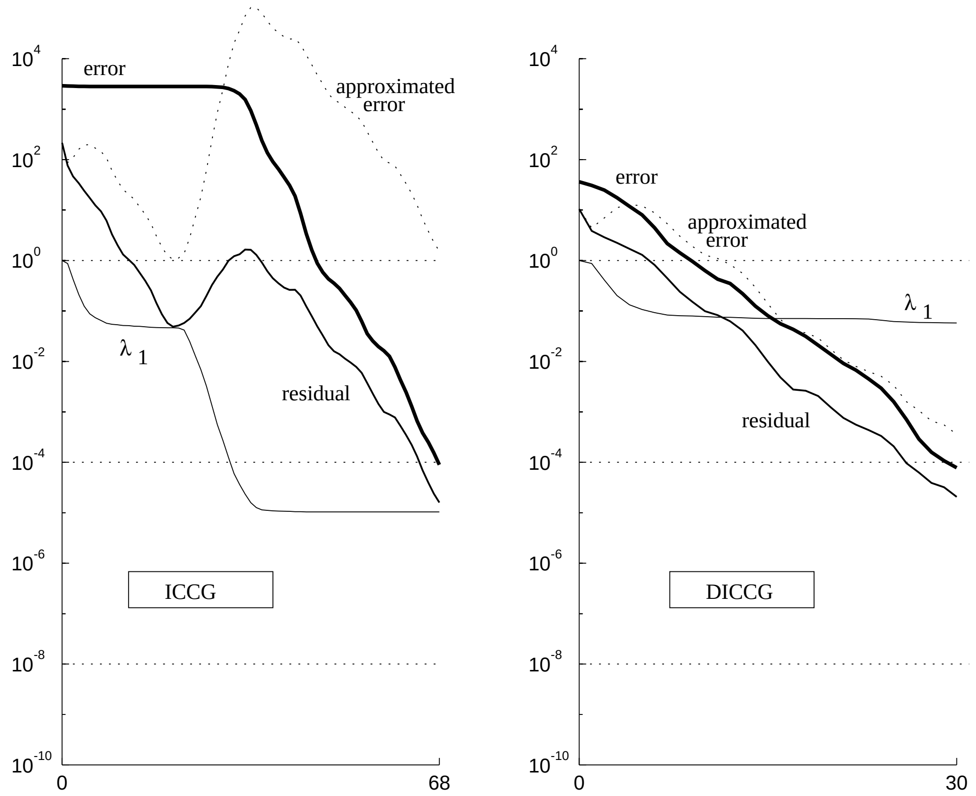
$$-\nabla \cdot (A\nabla u) = F, \tag{1}$$

where the coefficients and geometry of the problem are:



A groundwater flow problem

The low permeable layer ($A = 10^{-5}$) and the jump in permeabilities between the two sand sections lead to a 'small' eigenvalue.



TU Delft

5. Comparison of Deflation and Additive Coarse Grid Correction

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

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

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 6

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

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

Comparison of Deflation and Additive Coarse Grid Correction

Corollary

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

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

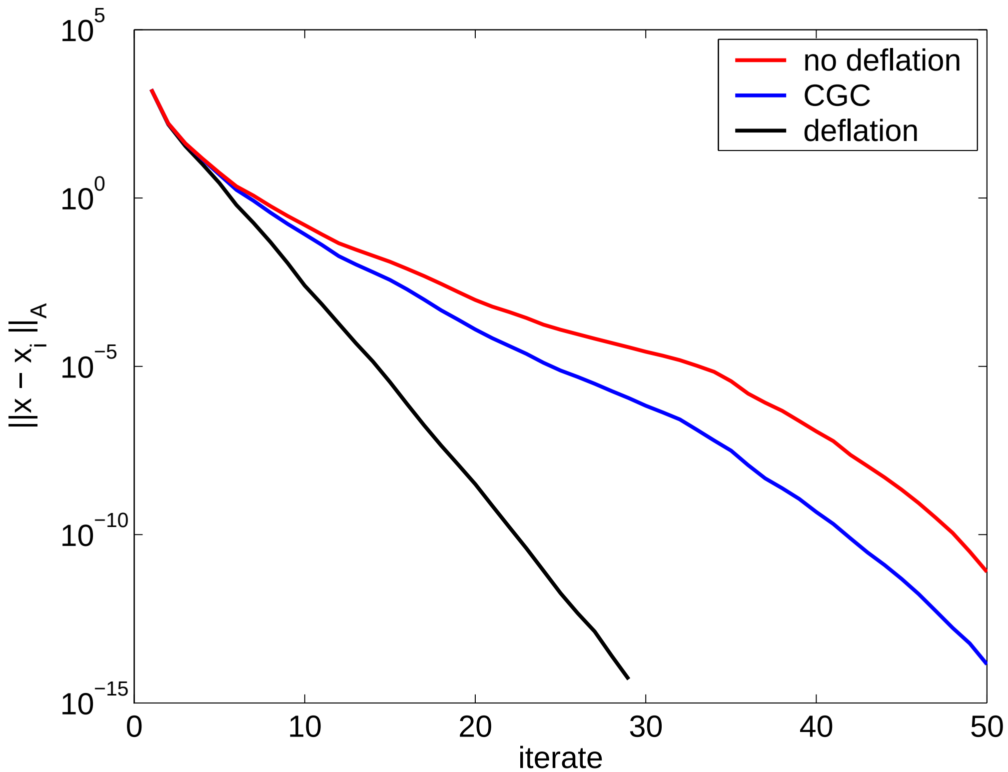
Conclusion

Deflation is asymptotically better than additive coarse grid correction!

Results for eigenvectors

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

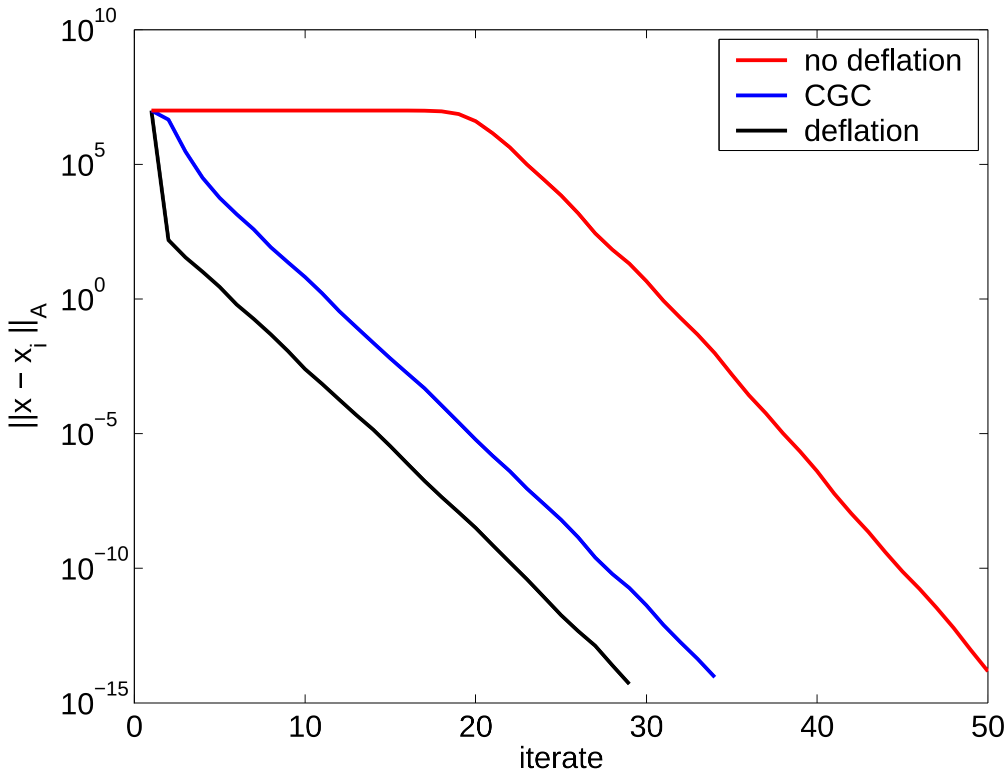
The eigenvectors $v_1, \ldots, v_{10}$ are used as projection vectors.



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.



6. 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 $Z = [v_1 \ldots v_r]$ and $M = I$.

Theorem 7

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

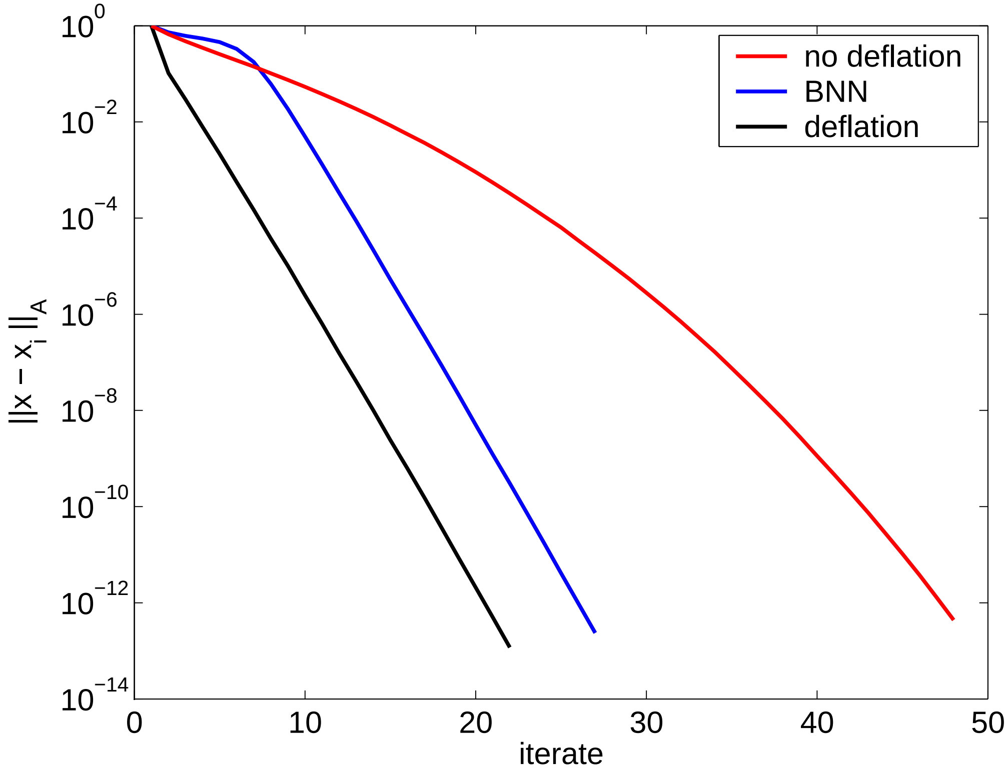
- the spectrum of $P_B A$ is $\{1, \ldots, 1, \lambda_{r+1}, \ldots, \lambda_n\}$

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

Deflation is asymptotically better than Balancing!

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

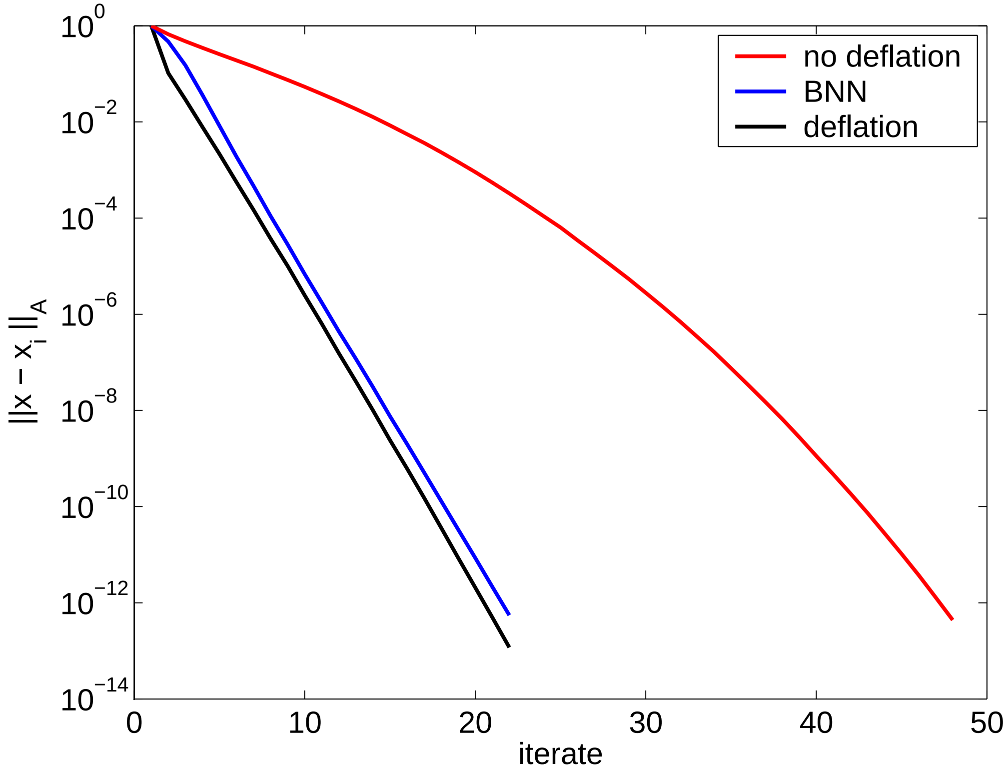
The eigenvalues of A are $1, 2, 3, \ldots, 99, 100$.



TUDelft

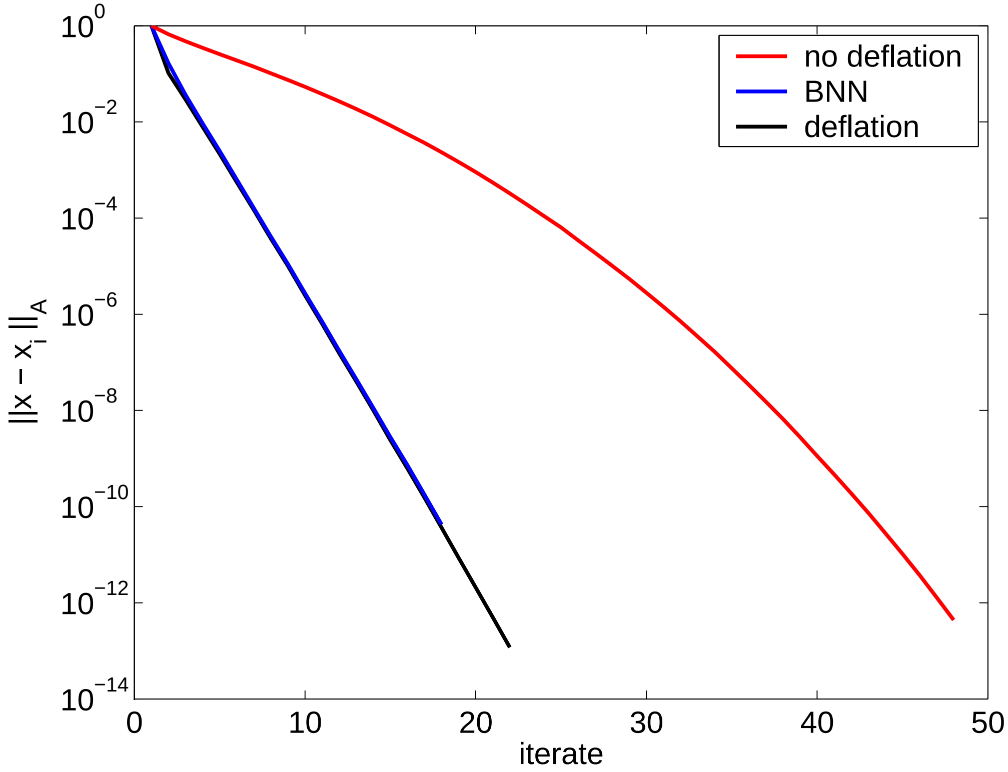
Results for eigenvectors $v_1, \ldots, v_{10}$

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 $0.01, 0.02, 0.03, \ldots, 0.99, 1$.



7. Conclusions

- DICCG is a robust and efficient method to solve diffusion problems with discontinuous coefficients.
- The choice of the projection vectors is important for the success of a projection method.
- 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 Balancing, but uses less work per iteration.

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, 27, pp. 1742-1759, 2006

Projection type methods

Krylov	Ar
Preconditioned Krylov	$M^{-1}Ar$
Block Preconditioned Krylov	$\sum_{i=1}^{m}(M_i^{-1})Ar$
Block Preconditioned Deflated Krylov	$\sum_{i=1}^{m}(M_i^{-1})PAr$