Iterative solution of saddle point problems

Miroslav Rozložník, Pavel Jiránek

Institute of Computer Science,
Czech Academy of Sciences, Prague

and

Faculty of Mechatronics and Interdisciplinary Engineering Studies,
Technical University of Liberec

SNA 2007, Ostrava, 22. – 26. 1. 2007

Outline

Problem statement – the most frequent definition:

$$\mathcal{A}u \equiv \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \equiv b$$

- $A \in \mathbb{R}^{n \times n}$ is a square matrix of order n,
- $B \in \mathbb{R}^{n \times m}$ is an overdetermined matrix with $n \geq m$,
- $f \in \mathbb{R}^n$ is an n-dimensional right-hand side vector.

Problem statement – generalizations and generalized saddle point problems:

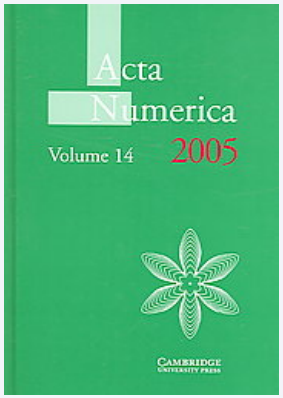
$$\begin{pmatrix} A & B \\ B^T & -C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

- $C \in \mathbb{R}^{m \times m}$ is a square matrix of order m,
- $g \in \mathbb{R}^m$ is an m-dimensional right-hand side vector.

$$\begin{pmatrix} A & B \\ D^T & -C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

- $D \in \mathbb{R}^{n \times m}$ is an overdetermined matrix with $n \geq m$.

Basic reference



M. Benzi, G. H. Golub, J. Liesen. *Numerical solution of saddle point problems.* Acta Numerica, 2005, pp. 1–137.

Applications leading to saddle point problems

- computational fluid dynamics (Glowinski, Quarteroni, Valli, Temam, Turek, Wesseling),
- constrained and weighted least squares problems (Björck, Golub, Van Loan),
- constrained optimization (Gill, Murray, Wright),
- economics (Arrow, Hurwitz, Uzawa, Duchin, Szyld, Leontief),
- electrical circuits and networks (Bergen, Chua, Desoer, Kuh, Strans, Tropper),
- electromagnetism (Bossavit, Perugia, Simoncini, Arioli),
- finance (Markowitz, Perold),
- image reconstruction and registration (Hall, Haber, Modersitzki),
- interpolation of scattered data (Lyche, Nilssen, Winther, Sibson, Stone),
- linear elasticity (Braess, Ciarlet),
- mesh generation (Liesen, de Sturler, Sheffer, Aydin, Siefert),
- mixed finite element method for elliptic PDEs (Brezzi, Fortin, Quarteroni, Valli),
- model order reduction (Freund, Heres, Schilders, Stykel),
- optimal control (Battermann, Heinkenschloss, Sachs, Betts, Biros, Ghattas, Nguyen),
- parameter identification problems (Burger, Mühlhuber, Haber, Asher, Oldenburg).

Saddle point problems in Czech republic:

- **L. Lukšan, J. Vlček.** *Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems.* Numerical Linear Algebra with Applications 5, 1998, pp. 219–247.
- **Z. Dostál, D. Lukáš**. *Multigrid preconditioned augmented Lagrangians for the Stokes problem.* Proceedings of SNA'06, ICS AS CR, 2006, pp. 25–28.
- **J. Kruis**. *Reinforcement-matrix interaction modelled by FETI method.* Proceedings of SNA'06, ICS AS CR, 2006, pp. 51–54.
- **R. Kučera, J. Haslinger, T. Kozubek**. *An algorithm for solving nonsymmetric saddle-point linear systems arising in FDM.* Proceedings of PANM 13, MI AS CR, 2006.
- **M. Hokr**. *Modelling of flow and transport problems in geological media.* Proceedings of SNA'07, UGN AS CR, 2007.

Uranium deposit Stráž – geographical location and hydrogeological situation:

- classical deep mining: 1966 – 1993,
- underground acidic leaching: 1968 – 1996,
- produced 14000 tons of uranium,
- $4 \cdot 10^6$ tons of H_2SO_4 injected in sandstone area,
- $190 \cdot 10^6$ m^3 of contaminated water in cretaceous collectors,
- hydrological barrier (injection of clean water),
- drainage channels (pumping out the solution), mine drainage.

Application fields solved in Diamo s.e.:

- modelling of the underground water flow and transport of contaminants,
- modelling of remediation scenarios,
- modelling of flooding of the deep uranium mines,
- modelling of chemical leakage from the waste pond,
- fractured rock flow, modelling of radioactive waste deposit.

Mathematical models used in Diamo s.e.:

- structural and situation models: describe the structure and state of objects, provide input data for other models,
- models of flow and transport: computational models based on the FEM method, space and time discretization,
- thermodynamical and kinetical models: modelling of chemical processes and reactions,
- economical and optimization models: making decisions support and tools.

Particular application – porous media flow:

- impermeable (generally nonparallel) bottom and top layers,
- modelled area covers approximately 120 km,
- vertical thickness up to 200 m,
- trilateral prismatic discretization,
- 3D meshes of the order from $2 \cdot 10^5$ to $8 \cdot 10^5$ elements,
- systems of order 10^6 at every time step,
- implementation and GWS software developed at Diamo s.e.

References:

- M. Rozložník, V. Simoncini. *Krylov subspace method for saddle point problems with indefinite preconditioning.* SIAM J. Mat. Anal. Appl. 24, 2002, pp. 368–391.
- J. Maryška, M. Rozložník, M. Tůma. *The potential fluid flow problem and the convergence rate of the minimal residual method.* Numer. Lin. Alg. Appl. 3, pp. 525–542.
- M. Arioli, J. Maryška, M. Rozložník, M. Tůma. *Dual variable methods for mixed-hybrid finite element approximation of the potential fluid flow problem in porous media.* ETNA 22, 2006, pp. 17–40.
- J. Maryška, M. Rozložník, M. Tůma. *Schur complement systems in the mixed-hybrid finite element approximation of the potential fluid flow problem.* SIAM J. Sci. Comp. 22, 2000, pp. 704–723.
- J. Maryška, M. Rozložník, M. Tůma. *Mixed hybrid finite element approximation of the potential fluid flow problem.* J. Comp. Appl. Math. 63, pp. 383–392.

References:

- J. Maryška, M. Rozložník, M. Tůma. *Primal vs. dual variable approach for mixed-hybrid finite element approximation of the potential fluid flow problem in porous media.* Proceedings of the 3rd International Conference on "Large-Scale Scientific Computations", Lecture Notes in Computer Science 2179, Sv. Margenov, J. Wasniewski, P. Yalamov (eds.), June 6-10, 2001, pp. 417–424.
- P. Jiránek, M. Rozložník. *Maximum attainable accuracy of inexact saddle point solvers.* submitted to SIMAX, 2006.
- P. Jiránek, M. Rozložník. *Limiting accuracy of segregated solution methods for nonsymmetric saddle point problems.* submitted to JCAM, 2006.

Saddle point systems – equivalent definitions and properties

A is nonsingular $\Rightarrow$

$$\mathcal{A} \text{ is nonsingular } \Leftrightarrow \; B^T A^{-1} B \text{ is nonsingular,}$$

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^T A^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & -B^T A^{-1} B \end{pmatrix}.$$

A symmetric positive definite and B of full column rank
$\Rightarrow B^T A^{-1} B$ is symmetric positive definite
$\Rightarrow \mathcal{A}$ is nonsingular.

A symmetric positive semidefinite and B of full column rank $\Rightarrow$

$$N(A) \cap N(B^T) = 0 \;\Leftrightarrow\; \mathcal{A} \text{ is nonsingular.}$$

A nonnegative real ($\frac{1}{2}(A + A^T)$ symmetric positive semidefinite) and B of full rank $\Rightarrow$

$$N(\tfrac{1}{2}(A + A)^T) \cap N(B^T) = 0 \;\Rightarrow \mathcal{A} \text{ is nonsingular.}$$

Equality constrained quadratic programming problem:

$$\text{minimize } \frac{1}{2}(Au, u) - (f, u) \text{ subject to } B^T u = g.$$

The solution is the saddle point problem of the Lagrangian

$$\mathcal{L}(u, v) \equiv \frac{1}{2}(Au, u) - (f, u) + (B^T u - g, v),$$

$$\mathcal{L}(x, v) \leq \mathcal{L}(x, y) \leq \mathcal{L}(u, y) \quad \forall u \in \mathbb{R}^N, \ v \in \mathbb{R}^m,$$

$$\mathcal{L}(x, y) = \min_{u \in \mathbb{R}^n} \max_{v \in \mathbb{R}^m} \mathcal{L}(u, v).$$

Weighted least squares problem:

A symmetric positive definite, B of full column rank,

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix} = \begin{pmatrix} A^{-1}(I - B(B^T A^{-1} B)^{-1} B^T A^{-1}) f \\ -(B^T A^{-1} B)^{-1} B^T A^{-1} f \end{pmatrix},$$

Vector y is the solution of weighted least squares problem

$$\|x\|_A = \|f - By\|_{A^{-1}} = \min_{v \in \mathbb{R}^m} \|f - Bv\|_{A^{-1}}.$$

Least squares problem – augmented formulation:

B is of full column rank $\Rightarrow$

$$\begin{pmatrix} I & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} I & B \\ B^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} f \\ 0 \end{pmatrix} = \begin{pmatrix} (I - \Pi) f \\ (B^T B)^{-1} B^T f \end{pmatrix},$$

$\Pi = B(B^T B)^{-1} B^T$ is the orthogonal projector onto $R(B)$.

Vector y is the solution of the least squares problem $By \approx f$,

$$y = B^\dagger f = (B^T B)^{-1} B^T f.$$

Vector x is the least squares residual

$$\|x\| = \|f - By\| = \min_{v \in \mathbb{R}^m} \|f - Bv\|.$$

A symmetric positive definite and B has full rank.

Block factorization of $\mathcal{A}$:

$$\mathcal{A} = \begin{pmatrix} I & 0 \\ B^T A^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -B^T A^{-1} B \end{pmatrix} \begin{pmatrix} I & A^{-1} B \\ 0 & I \end{pmatrix}$$

$\Rightarrow \mathcal{A}$ has n positive and m negative eigenvalues

$\Rightarrow A$ is symmetric, but indefinite.

Spectral bounds: [Rusten, Winther, 1992]

$$\lambda(A) \subset [\lambda_n, \lambda_1],\ \lambda_n > 0,\ \sigma(B) \subset [\sigma_m, \sigma_1] \Rightarrow$$

$$\lambda(\mathcal{A}) \subset I^- \cup I^+,$$

$$I^- \equiv \left[\tfrac{1}{2}(\lambda_n - \sqrt{\lambda_n^2 + 4\sigma_1^2}), \tfrac{1}{2}(\lambda_1 - \sqrt{\lambda_1^2 + 4\sigma_m^2}) \right],$$

$$I^+ \equiv \left[\lambda_n, \tfrac{1}{2}(\lambda_1 + \sqrt{\lambda_1^2 + 4\sigma_1^2}) \right].$$

Potential fluid flow problem:

The Darcy's law and the continuity equation in the space domain Ω:

$$Ku + \nabla p = 0, \quad \nabla \cdot u = q.$$

Dirichlet and Neumann boundary conditions on the boundary $\partial\Omega$:

$$p = p_D \text{ on } \partial\Omega_D, \quad u \cdot n = u_N \text{ on } \partial\Omega_N.$$

Mixed-hybrid finite element approximation:

- general prismatic discretization of the domain,
- low-degree Raviart-Thomas approximation of u,
- element-wise constant approximation of p,
- face-wise constant approximation of λ (pressure traces on faces of elements).

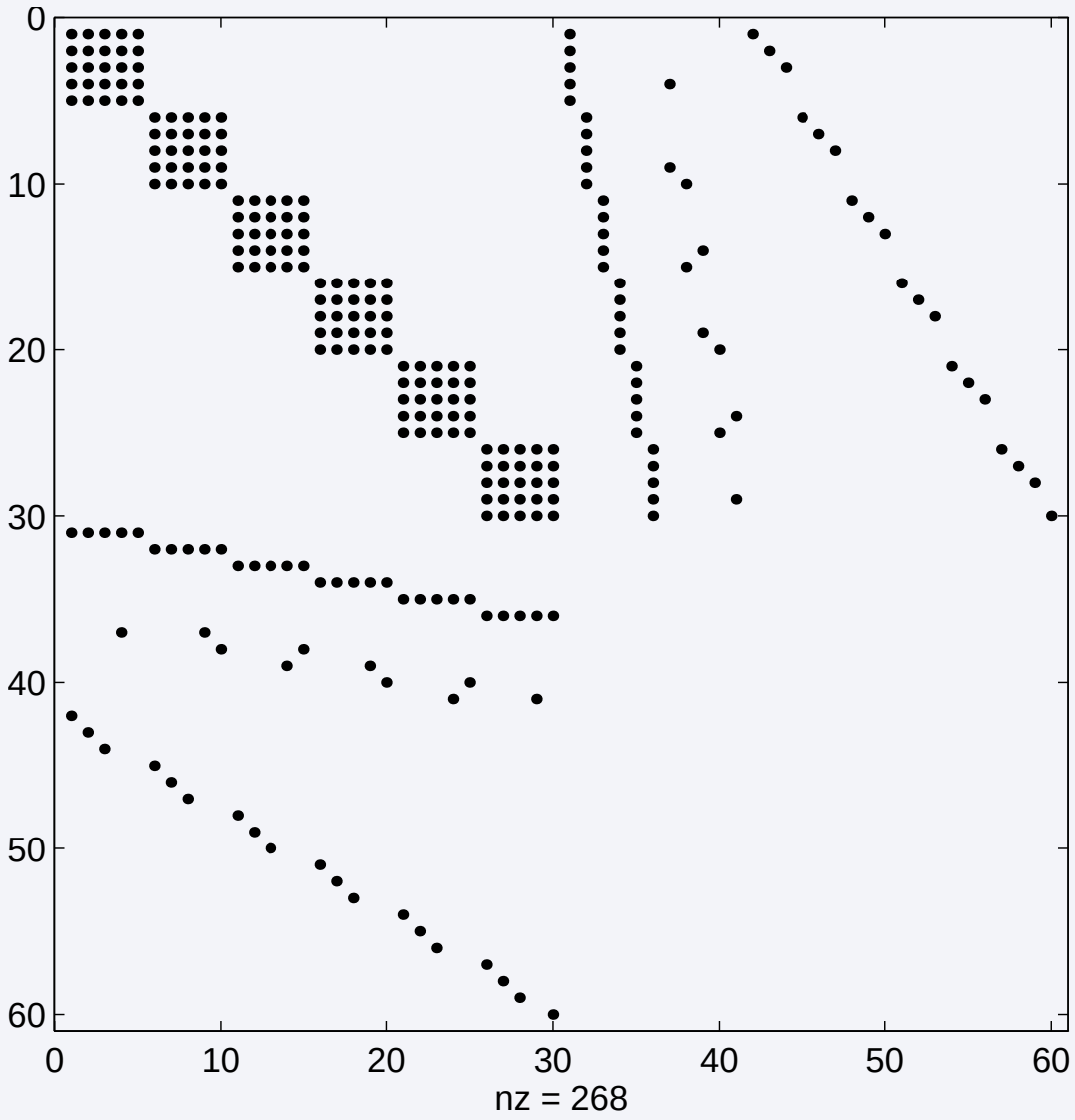
Particular saddle point problem:

$$\begin{pmatrix} A & B_1 & B_2 & B_3 \\ B_1^T & 0 & 0 & 0 \\ B_2^T & 0 & 0 & 0 \\ B_3^T & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix},$$

$$u \in \mathbb{R}^{5 \cdot NE},\ p \in \mathbb{R}^{NE}, \lambda_1 \in \mathbb{R}^{NIF},\ \lambda_2 \in \mathbb{R}^{NNC},$$

- NE – number of elements,
- NIF – number of interior inter-element faces,
- NNC – number of faces with Neumann conditions.

- A represents the hydraulic permeability,
- B_1^T stands for the continuity equation in the elements, $B_1^T B_1 = 5I$,
- B_2^T represents the continuity equation across the interior inter-element faces, $B_2^T B_2 = 2I$,
- B_3^T stands for the fulfillment of Neumann conditions on the boundary, $B_3^T B_3 = I$.


0
10
20
30
40
50
60
0
10
20
30
40
50
60
nz = 268

Model problem - potential fluid flow problem in a rectangular domain:

mesh size	Discretization parameters			N
	NE	NIF	NNC	
1/5	250	525	199	2224
1/10	2000	4600	796	17396
1/15	6750	15975	1791	58266
1/20	16000	38400	3184	137584
1/30	54000	131400	7164	462564
1/40	128000	313600	12736	1094336

Spectral properties of matrix blocks and of the saddle point matrix:

- spectrum of A lies in $[c_1 h^{-1}, c_2 h^{-1}]$,
- singular values of $B = (B_1, B_2, B_3)$ lie in $[c_3 h, c_4]$.

mesh h	spectrum of A	singular values of B
1/2	[0.11e-2, 0.66e-2]	[0.276, 2.577]
1/3	[0.16e-2, 0.10e-1]	[0.197, 2.611]
1/4	[0.22e-2, 0.13e-1]	[0.152, 2.624]
1/6	[0.33e-2, 0.19e-1]	[0.104, 2.635]
1/8	[0.44e-2, 0.26e-1]	[0.079, 2.639]
1/12	[0.66e-2, 0.40e-1]	——

Diagonal scaling of the saddle point matrix:

$$\tilde{\mathcal{A}} = \begin{pmatrix} \tilde{A} & B \\ B & 0 \end{pmatrix} = \begin{pmatrix} h^{1/2} I & 0 \\ 0 & h^{-1/2} I \end{pmatrix} \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} h^{1/2} I & 0 \\ 0 & h^{-1/2} I \end{pmatrix},$$

- spectrum of $\tilde{A}$ lies in $[c_1, c_2]$.

Spectrum of scaled system matrix:

$$\lambda(\tilde{\mathcal{A}}) \in I^- \cup I^+,$$

$$I^- = \left[\tfrac{1}{2}(c_1 - \sqrt{c_1^2 + 4c_4^2}), -c_3^2 c_2^{-1} h^2\right],$$

$$I^+ = \left[c_1, \tfrac{1}{2}(c_2 + \sqrt{c_2^2 + 4c_4^2})\right].$$

mesh size h	negative part	positive part
1/2	[-2.57, -0.27]	[0.13e-2, 2.57]
1/3	[-2.60, -0.19]	[0.18e-2, 2.61]
1/4	[-2.62, -0.15]	[0.23e-2, 2.62]
1/6	[-2.63, -0.13]	[0.34e-2, 2.63]
1/8	[-2.63, -0.10]	[0.44e-2, 2.64]

Basic solution approaches

Basic solution schemes for saddle point problems:

1. segregated methods:
   - reduce the whole problem to a smaller one, compute the component x or y as a solution of the reduced problem,
   - back-substitution into the original system to obtain the remaining component.

2. coupled methods:
   - do not explicitly use the block structure of the problem,
   - compute the components x and y at once.

Schur complement reduction:

Block LU factorization of the saddle point matrix:

$$\begin{pmatrix} I & 0 \\ B^T A^{-1} & -I \end{pmatrix} \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^T A^{-1} & -I \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}$$

$$\Downarrow$$

$$\begin{pmatrix} A & B \\ 0 & B^T A^{-1} B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ B^T A^{-1} f - g \end{pmatrix}$$

1. Solve the system with the Schur complement matrix
$$B^T A^{-1} B y = B^T A^{-1} f - g.$$
2. Back-substitution
$$Ax = f - By.$$

Iterative solution: $\|y - y_k\|_{B^T A^{-1} B} \to 0$, solve $Ax_k = f - By_k$.

Also called: static condensation, nodal analysis, displacement method, range-space method.

Null-space method:

$B^T x = 0 \Rightarrow x \in N(B^T)$,

$$\begin{pmatrix} I-\Pi \\ B^T \end{pmatrix} \begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (I-\Pi)f \\ f \end{pmatrix}$$

$$\Downarrow$$

$$\begin{pmatrix} (I-\Pi)A(I-\Pi) & 0 \\ B^T A & B^T B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (I-\Pi)f \\ f \end{pmatrix},$$

$\Pi = B(B^T B^{-1})B^T$ is the orthogonal projector onto $R(B)$.

1. Solve the projected system

$$(I-\Pi)A(I-\Pi)x = (I-\Pi)f.$$

2. Back-substitution

$$B^T By = B^T(f - Ax) \;\Leftrightarrow\; By \approx f - Ax.$$

Iterative solution: $\|x - x_k\|_{(I-\Pi)A(I-\Pi)} \to 0$, solve $By_k \approx f - Ax_k$.

Also called: reduced Hessian method, loop analysis, force method.

The Schur complement reduction:

$$\mathcal{A} = \begin{pmatrix} A & B_1 & B_2 & B_3 \\ B_1^T & 0 & 0 & 0 \\ B_2^T & 0 & 0 & 0 \\ B_3^T & 0 & 0 & 0 \end{pmatrix}$$

The Schur complement matrix:

$$-\mathcal{A}/A = \begin{pmatrix} B_1^T \\ B_2^T \\ B_3^T \end{pmatrix} A^{-1} \begin{pmatrix} B_1 & B_2 & B_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{12}^T & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33}. \end{pmatrix}$$

Second Schur complement:

$$(-\mathcal{A}/A)/A_{11} = \begin{pmatrix} A_{22} & A_{23} \\ A_{23}^T & A_{33} \end{pmatrix} - \begin{pmatrix} A_{12}^T \\ A_{13}^T \end{pmatrix} A_{11}^{-1} \begin{pmatrix} A_{12} & A_{13} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix}.$$

Third Schur complement:

$$((-\mathcal{A}/A)/A_{11})/B_{22} = B_{11} - B_{12} B_{22}^{-1} B_{12}^T.$$

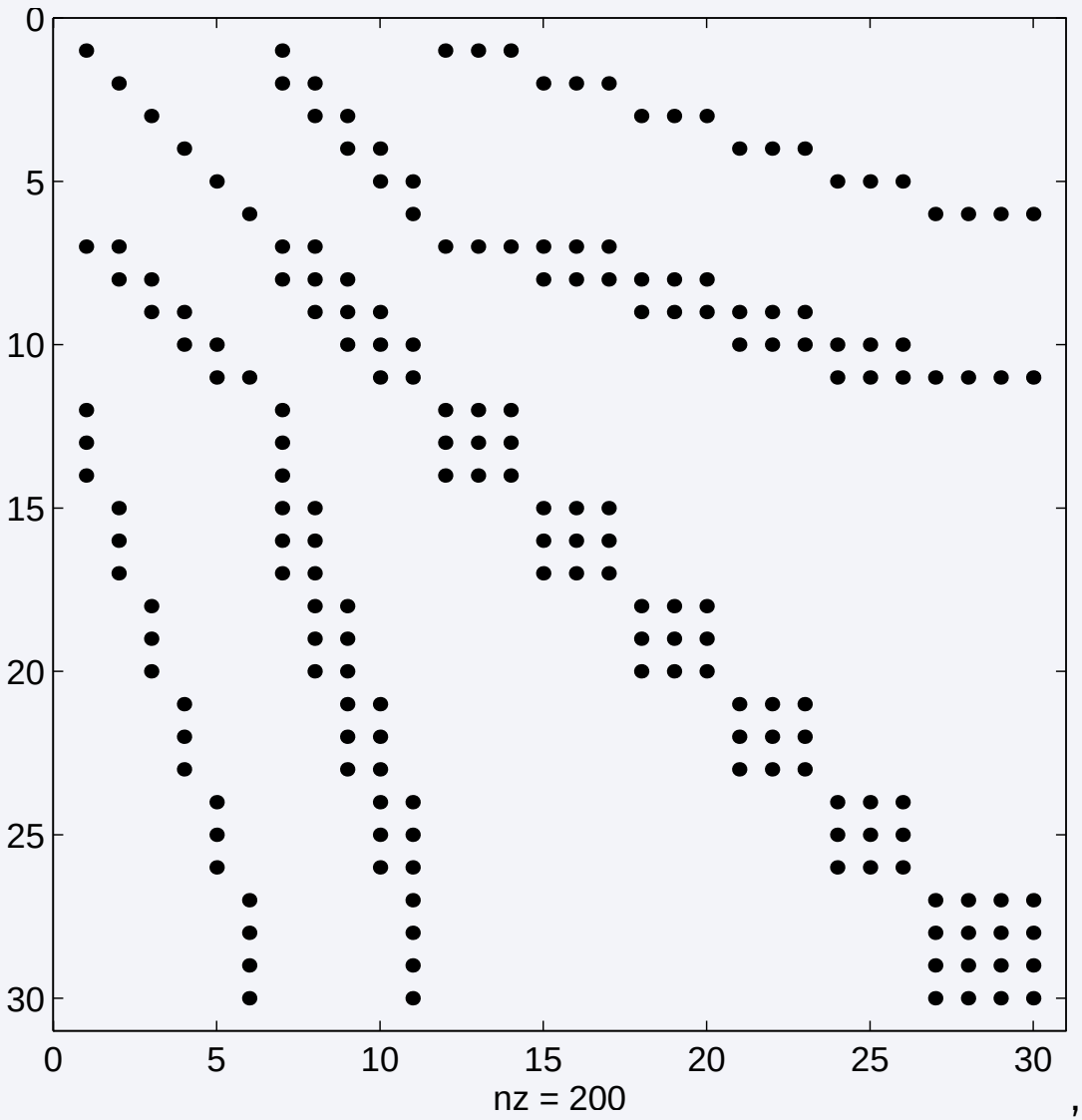
- Subsequent reduction to the Schur complement systems without additional fill-in:

$$\mathcal{A} \rightarrow -\mathcal{A}/A \rightarrow (\mathcal{A}/A)/A_{11} - ((\mathcal{A}/A)/A_{11})/B_{22}.$$

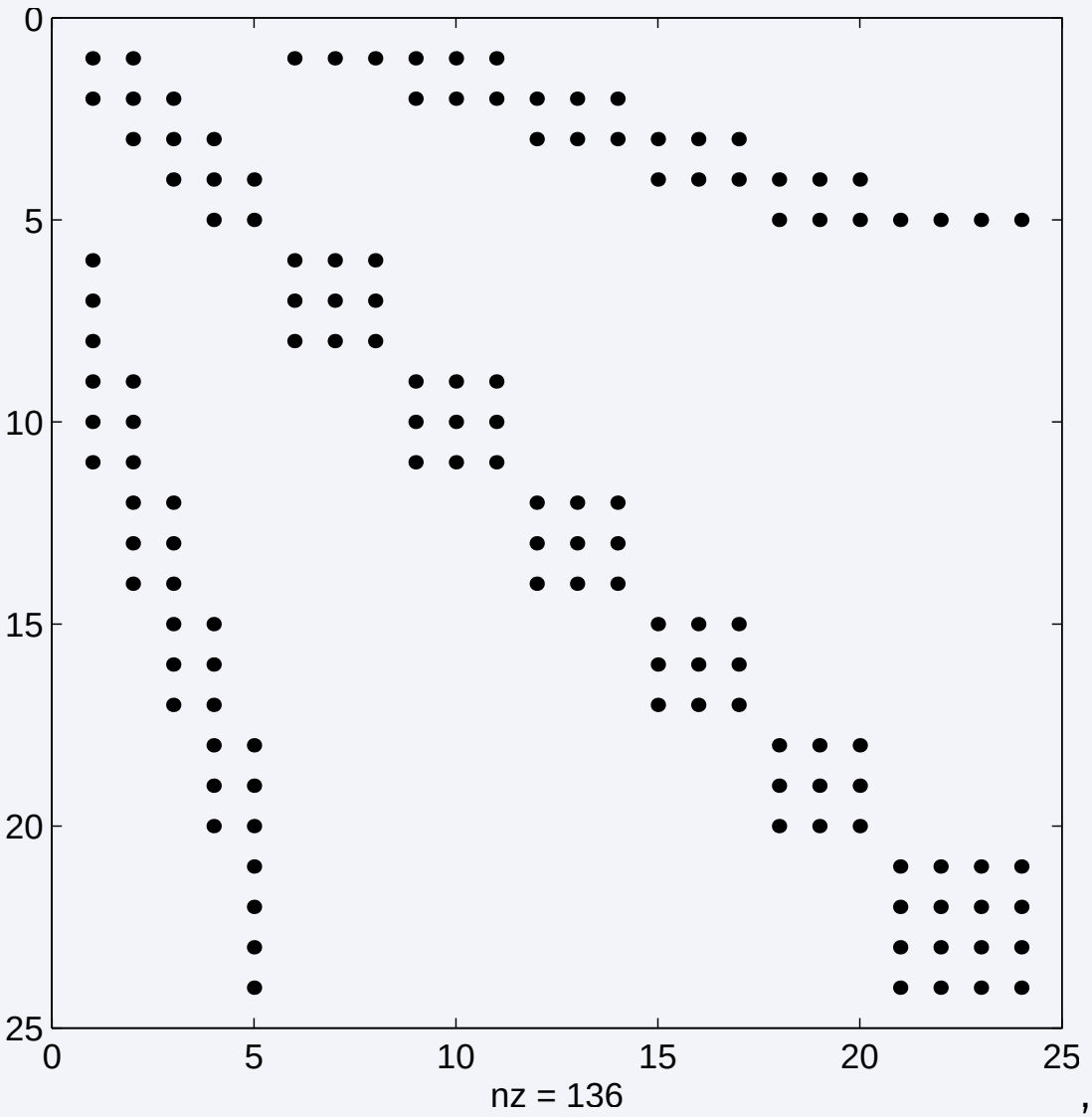
- Iterative solution of the final Schur complement system with the matrix $-((\mathcal{A}/A)/A_{11})/B_{22}$ for the unknown vector λ_1 by CG.
- Block back-substitution process for the unknown vectors λ_2, p and u using the factors from the Schur complement reduction.

Schur complement reduction:

elements NE (mesh size h)	Matrix dimensions			
	$\mathcal{A}$	$-\mathcal{A}/A$	SC_2	SC_3
250 (1/5)	2224	974	724	525
2000 (1/10)	17396	7396	5396	4600
6750 (1/15)	58266	24516	17766	15975
16000 (1/20)	137584	57584	41584	38400
54000 (1/30)	462564	192564	138564	131400
128000 (1/40)	1094336	454336	326336	313600


0
5
10
15
20
25
30
0
5
10
15
20
25
30
nz = 200

,


0
5
10
15
20
25
0
5
10
15
20
25
nz = 136

,

Spectral properties of Schur complement matrices do not deteriorate:

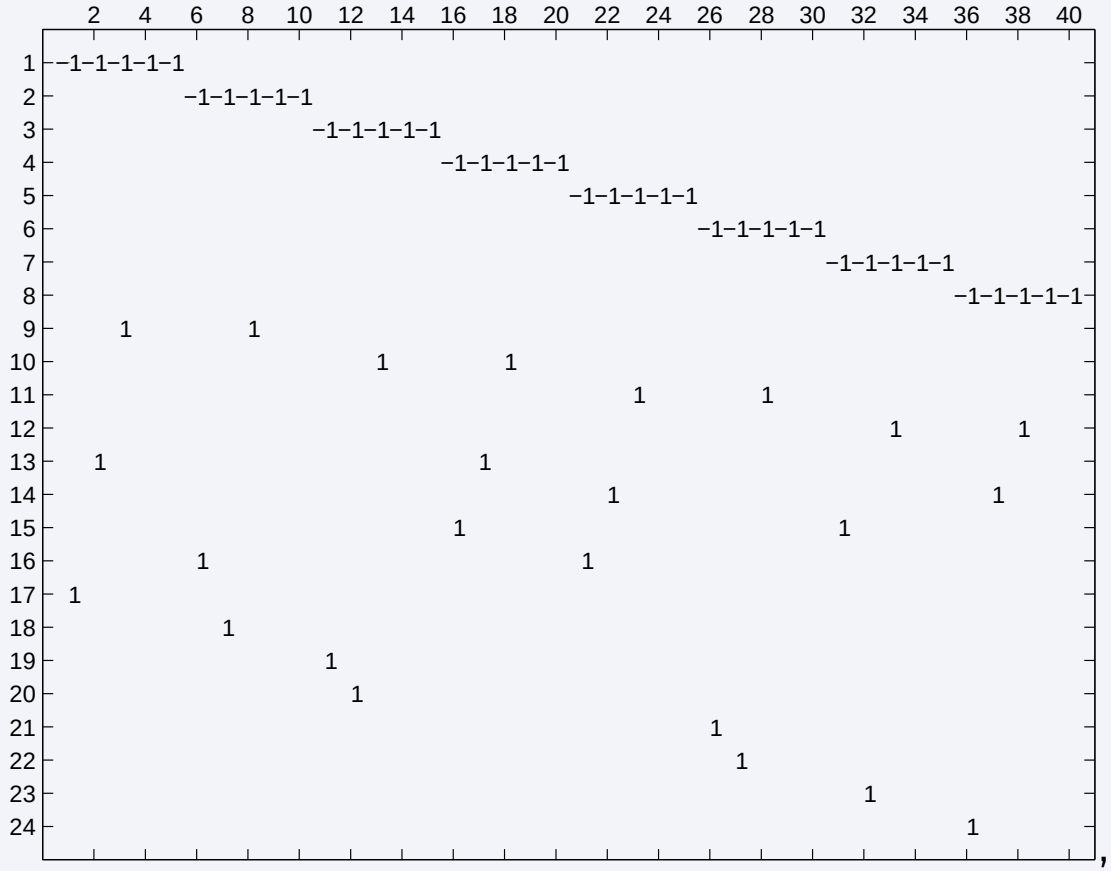
$$\lambda(-\mathcal{A}/A) \subset [c_2^{-1} c_3^2 h^2, c_1^{-1} c_4^2],$$
$$\lambda(((-\mathcal{A}/A)/A_{11})/B_{22}) \subset \lambda((-\mathcal{A}/A)/A_{11}) \subset \lambda(-\mathcal{A}/A).$$

NE	spectral properties of SC matrices		
	$-\mathcal{A}/A$	SC_2	SC_3
250	[0.10e0, 0.34e4]	[0.11e0, 0.12e4]	[0.2e0, 0.12e4]
2000	[0.16e-1, 0.17e4]	[0.22e-1, 0.60e3]	[0.26e-1, 0.60e3]
6750	[0.52e-2, 0.12e4]	[0.72e-2, 0.40e3]	[0.80e-2, 0.40e3]
16000	[0.23e-2, 0.87e3]	[0.32e-2, 0.30e3]	[0.34e-2, 0.30e3]
54000	[0.70e-3, 0.58e3]	[0.98e-3, 0.20e3]	[0.10e-2, 0.20e3]
128000	[0.30e-3, 0.43e3]	[0.42e-3, 0.15e3]	[0.43e-3, 0.15e3]

Null-space projection:

- B is an incomplete incidence matrix of certain graph,
- fixed geometry of the domain, iterative change of material (physical) properties (solving inverse problems or sequences of time-dependent or nonlinear problems),
- use divergence-free finite elements (null-space approach embedded in formulation) vs. fully algebraic mixed of mixed-hybrid approach.

$$B = (B_1,\ B_2,\ B_3) =$$


2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
1 −1−1−1−1−1
2 −1−1−1−1−1
3 −1−1−1−1−1
4 −1−1−1−1−1
5 −1−1−1−1−1
6 −1−1−1−1−1
7 −1−1−1−1−1
8 −1−1−1−1−1
9 1 1
10 1 1
11 1 1
12 1 1
13 1 1
14 1 1
15 1 1
16 1 1
17 1
18 1
19 1
20 1
21 1
22 1
23 1
24 1

,

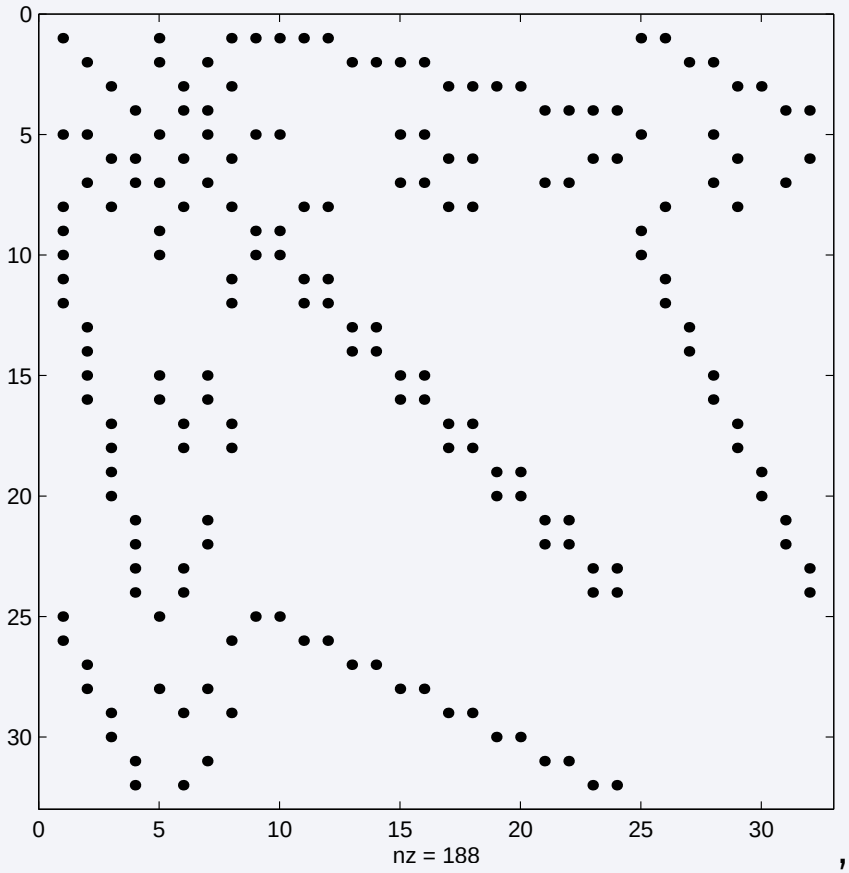
Approach based on a null-space basis of $(B_1,\ B_2,\ B_3)^T$:

- find a null-space basis Z of B^T,
- find some particular solution of $B^T u_1 = (q_2^T, q_3^T, q_4^T)^T$,
- solve iteratively the symmetric positive definite system

$$Z^T A Z u_2 = Z^T (q_1 - A u_1),$$

- find p and λ such that

$$B \begin{pmatrix} p \\ \lambda \end{pmatrix} = q_1 - Au,\ u = u_1 + Z u_2.$$


0
5
10
15
20
25
30
0
5
10
15
20
25
30
nz = 188

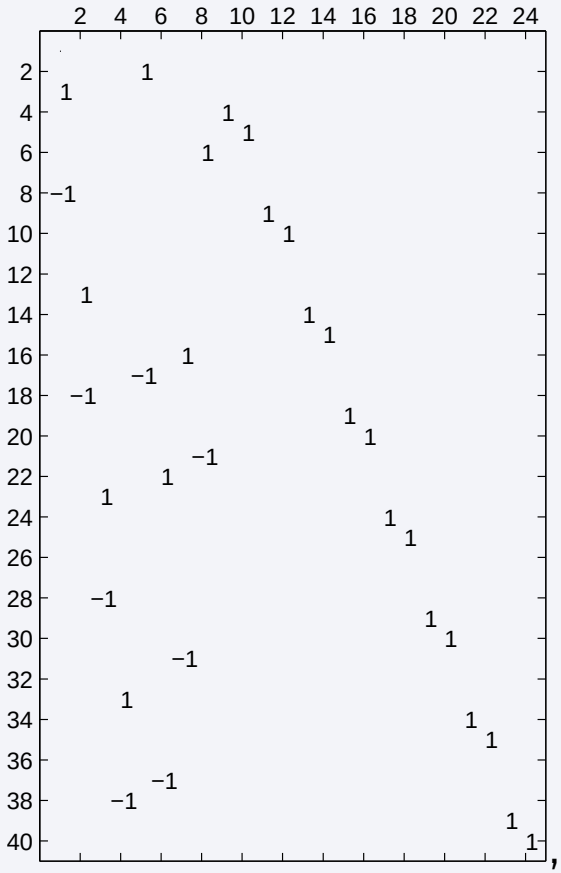
,

Approach based on a (partial) null-space basis of the block $(B_2,\ B_3)^T$: [Arioli, Maryška, Rozložník, Tůma, 2004]

- find an orthogonal null-space basis Z of the matrix block $(B_2\ B_3)^T$,
- find some particular solution of $(B_2\ B_3)^T u_1 = q_3$,
- solve iteratively the symmetric indefinite projected system

$$\begin{pmatrix} Z^T A Z & Z^T B \\ B^T Z & 0 \end{pmatrix} \begin{pmatrix} u_2 \\ p \end{pmatrix} = \begin{pmatrix} Z^T (q_1 - A u_1) \\ q_2 - B^T u_1 \end{pmatrix},$$

- set $u = u_1 + Z u_2$, find unknown vector λ such that $(B_2\ B_3)\lambda = q_1 - Au - Bp$.


2 4 6 8 10 12 14 16 18 20 22 24
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40

Discretization parameters					Dimension of null-spaces	
h	NE	NIF	NDC	NNC	$NZ1$	$NZ2$
$1/5$	250	525	100	100	375	625
$1/10$	2000	4600	400	400	3000	5000
$1/15$	6750	15975	900	900	10125	16875
$1/20$	16000	38400	1600	1600	24000	40000
$1/25$	31250	75625	2500	2500	46875	78125
$1/30$	54000	131400	3600	3600	81000	135000
$1/35$	87750	209475	4900	4900	138625	226375
$1/40$	128000	313600	6400	6400	192000	320000

Choice of the null-space basis:

- **Fundamental cycle null-space basis** based on incidence vectors of cycles in the mesh: find a shortest path spanning tree; form cycles using non-tree edges;

$$\sigma(Z) \subset [1, c_5 h^{-2}],\ \lambda(Z^T A Z) \subset [c_1, c_2 c_5^2 h^{-4}].$$

- **Orthogonal null-space basis** based on the QR decomposition of B (MA49 from HSL); projected system with $Z^T \tilde{A} Z$ does not depend on the mesh size h.

- **Partial null-space approach**: Z has orthogonal columns and can be explicitly constructed without any computation,

$$\sigma(Z^T B) \subset [c_7 h, c_8].$$

Iterative methods for linear systems

Continuous problem

$\downarrow$

Discretization

$\downarrow$

Linear algebraic system

- **Dense or sparse direct solver?**
  If we can solve the system directly, let's do it!
- If not, use **iterative method** with some **preconditioner**.
- Use as much **information from the problem** as possible!

Linear system, iterative methods:

$$\mathcal{A}u = b, \quad \mathcal{A} \in \mathbb{R}^{N\times N}, \; b \in \mathbb{R}^N.$$

$$u_0, r_0 = b - Au_0,$$

$$u_k = \ldots, \quad r_k = b - Au_k,$$

$$u_k \longrightarrow u, \quad r_k \longrightarrow 0.$$

Iterative methods:

- stationary methods:

  $\longrightarrow$ solvers,

  $\longrightarrow$ preconditioners,

  $\longrightarrow$ smoothers.

  Uzawa method, augmented Lagrangian methods, other splittings,...

- Krylov subspace methods:

  $\longrightarrow$ solvers,

  $\longleftarrow$ need preconditioners.

  CG, MINRES, GMRES,...

- Algebraic multigrid, aggregation and multilevel methods

  $\longleftarrow$ need smoothers,

  $\longleftarrow$ need solvers,

  $\longleftarrow$ need preconditioners.

References:

Overviews:

- J. Stoer, *Solution of large linear systems of conjugate gradient type methods*, In Mathematical Programming, Springer, Berlin, 1983, pp. 540–565.
- R. Freund, G. H. Golub, N. Nachtigal, *Iterative solution of linear systems*, Acta Numerica 1, 1992, pp. 1–44.

Theory:

- A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

Practical issues:

- Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Pub. Co., Boston, 1996 (2nd edition: SIAM, Philadelphia, 2003).

Uzawa iteration [Uzawa, 1958]:

1. choose y_0
2. for $k = 0, 1, 2, \ldots$ until convergence
3. given y_k compute x_{k+1} such that $\mathcal{L}(x_{k+1}, y_k)$ is minimized:
   $x_{k+1} = A^{-1}(f - By_k)$
4. perform the Richardson update:
   $y_{k+1} = y_k + \beta_k(B^T x_{k+1} - g)$
5. end

Uzawa method as the fixed point iteration:

$$\mathcal{M} u_{k+1} = \mathcal{N} u_k + b,$$

$$\mathcal{M} = \begin{pmatrix} A & 0 \\ B^T & -\beta^{-1} I \end{pmatrix}, \quad \mathcal{N} = \begin{pmatrix} 0 & -B \\ 0 & -\beta^{-1} I \end{pmatrix}.$$

The iteration matrix

$$\mathcal{T} = \mathcal{M}^{-1}\mathcal{N} = \begin{pmatrix} 0 & -A^{-1}B \\ 0 & I - \beta B^T A^{-1} B. \end{pmatrix}$$

Richardson method applied to the Schur complement system

$$Sy \equiv B^T A^{-1} B y = B^T A^{-1} f.$$

Inexact Uzawa method:

Preconditioned inexact Uzawa iteration [Bramble, Pasciak, Vassilev, 1997], [Elman, Golub, 1994]:

1. choose y_0
2. for $k = 0, 1, 2, \ldots$ until convergence
3. $\quad x_{k+1} = x_k + \alpha_k \hat{A}^{-1}(f - Ax_k - By_k)$
4. $\quad y_{k+1} = y_k + \beta_k \hat{S}^{-1}(B^T x_{k+1} - g)$
5. end

$\hat{A}$ and $\hat{S}$ are approximations to A and S, respectively.

Inexact Uzawa method:

When $\hat{A}$ and $\hat{S}$ is spectrally equivalent to A and S, respectively, the inexact Uzawa method is convergent [Bramble, Pasciak, Vassilev, 1997].

- $\hat{A} = I_n$, $\hat{S} = I_m$ – Arrow-Hurwitz method [Arrow, Hurwitz, 1958]; convergence analysis [Fortin, Glowinski, 1983], [Verfürth, 1984],
- multigrid methods [Bramble, Pasciak, Vassilev, 1997], [Elman, 1996].

Augmented Lagrangian method:

Augmented Lagrangian method [Hestenes, 1969], [Powell, 1969], [Fortin, Glowinski, 1983] – penalized Lagrangian:

$$\mathcal{L}_\gamma(u, v) = \mathcal{L}(u, v) + \frac{1}{2}\gamma\|B^T u - g\|^2, \ \gamma > 0.$$

Uzawa iteration:

1. choose y_0
2. for $k = 0, 1, 2, \ldots$ until convergence
3. given y_k compute x_{k+1} such that $\mathcal{L}_\gamma(x_{k+1}, y_k)$ is minimized: $x_{k+1} = (A + \gamma BB^T)^{-1}(f - By_k + \gamma Bg)$
4. perform Richardson update: $y_{k+1} = y_k + \gamma(B^T x_{k+1} - g)$
5. end

HSS iteration:

A is positive real, positive real formulation of $\mathcal{A}z = b$:

$$\mathcal{A}'u \equiv \begin{pmatrix} A & B \\ -B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

Splitting of the system matrix proposed in [Bai, Golub, Ng, 2003], [Benzi, Golub, 2004]:

$$\mathcal{A}' = (\alpha I_{n+m} + \mathcal{H}) - (\alpha I_{n+m} - \mathcal{S}) = (\alpha I_{n+m} + \mathcal{S}) - (\alpha I_{n+m} - \mathcal{H})$$

with $\alpha > 0$ and $\mathcal{H} = \frac{1}{2}(\mathcal{A}' + \mathcal{A}'^T)$, $\mathcal{S} = \frac{1}{2}(\mathcal{A}' - \mathcal{A}'^T)$.

Two-level iteration:

$$(\alpha I_{n+m} + \mathcal{H})u_{k+\frac{1}{2}} = (\alpha I_{n+m} - \mathcal{S})u_k + b,$$
$$(\alpha I_{n+m} + \mathcal{S})u_{k+1} = (\alpha I_{n+m} - \mathcal{H})u_{k+\frac{1}{2}} + b.$$

Other stationary methods for saddle point problems:

- (Block-)SOR methods [Strikwerda, 1984], [Barlow, Nichols, Plemmons, 1988], [Plemmons, 1986], [Chen, 1998], [Golub, Wu, Yuan, 2001];
- Alternating direction methods [Brown, 1982], [Douglas, Durán, Pietra, 1986];
- Methods with indefinite splitting (constraint preconditioners) [Dyn, Ferguson, 1983], [Bank, Welfert, Yserentant, 1990], [Golub, Wathen, 1998], [Braess, Sarazin, 1997].

  $A = M - N \rightarrow \mathcal{M} u_{k+1} = \mathcal{N} u_k + b$, where

$$\mathcal{M} = \begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix}, \mathcal{N} = \begin{pmatrix} N & 0 \\ 0 & 0 \end{pmatrix}.$$

Starting from an initial guess u_0, build the sequence of nested spaces $\mathcal{K}_k$ (approximation space) and $\mathcal{L}_k$ (constraint space) such that

$$u_k \in u_0 + \mathcal{K}_k, \quad r_k \perp \mathcal{L}_k.$$

Krylov subspace methods: $\mathcal{K}_k = \mathcal{K}_k(\mathcal{A}, r_0) = \text{span}(r_0, \mathcal{A}r_0, \ldots, \mathcal{A}^{k-1}r_0)$.

- Orthogonal residual method – $\mathcal{L}_k = \mathcal{K}_k$ (CG [Hestenes, Stiefel, 1952] for symmetric positive definite systems, FOM [Saad, 1981]).
- Minimal residual method – $\mathcal{L}_k = A\mathcal{K}_k$ (CR [Stiefel, 1955], MINRES [Paige, Saunders, 1975] for symmetric systems, GMRES [Saad, Schultz, 1986]).

Other methods: SYMMLQ [Paige, Saunders, 1985] for symmetric systems, Petrov-Galerkin condition $\mathcal{L}_k = \mathcal{K}_k(\mathcal{A}^T, r_0)$ (BiCG [Lanczos, 1950], [Fletcher, 1976]), BiCGStab [van der Vorst, 1992] (combination of BiCG + GMRES(1)), QMR [Freund, Nachtigal, 1991], TFQMR [Freund, 1993],...

Symmetric positive definite case (Schur complement system, projected system) – CR, CG.

Symmetric indefinite case (the whole saddle point system) – MINRES, SYMMLQ.

- CG applied to symmetric but indefinite system: CG iterate exists at least at every second step [Paige, Saunders, 1975].
- Peak/plateau behaviour [Cullum, Greenbaum, 1996]:
  CG converges fast $\Rightarrow$ MINRES is not much better than CG,
  CG norm increases $\Rightarrow$ MINRES stagnates
- Residual smoothing techniques [Walker, Zhou, 1994], [Weiss, 1990].

Orthogonal residual methods (CG, FOM):
$\mathcal{A}$ symmetric positive definite $\Rightarrow$

$$r_k \perp \mathcal{K}_k(\mathcal{A}, r_0) \Leftrightarrow \|u - u_k\|_{\mathcal{A}} = \|r_k\|_{\mathcal{A}^{-1}} = \min_{\tilde{u} \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)} \|b - \mathcal{A}\tilde{u}\|_{\mathcal{A}^{-1}}$$
$$\Leftrightarrow \|u - u_k\|_{\mathcal{A}} = \|r_k\|_{\mathcal{A}^{-1}} = \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \|p_k(\mathcal{A}) r_0\|_{\mathcal{A}^{-1}}.$$

Minimal residual methods (CR, MINRES, GMRES):

$$r_k \perp \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0) \Leftrightarrow \|r_k\| = \min_{\tilde{u} \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)} \|b - \mathcal{A}\tilde{u}\|$$
$$\Leftrightarrow \|r_k\| = \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \|p_k(\mathcal{A}) r_0\|.$$

Minimal residual methods – convergence analysis:

$$\|r_k\| = \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \|p_k(\mathcal{A}) r_0\|$$

$$\mathcal{A} = \mathcal{V}\mathcal{D}\mathcal{V}^{-1} \text{ is diagonalizable} \Rightarrow$$

$$\begin{aligned} \frac{\|r_k\|}{\|r_0\|} &\leq \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in \lambda(\mathcal{A})} \|\mathcal{V} p_k(\mathcal{D}) \mathcal{V}^{-1}\| \\ &\leq \kappa(\mathcal{V}) \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in \lambda(\mathcal{A})} |p_k(\lambda)|. \end{aligned}$$

Normal systems – eigenvalues play an important role:

$$\frac{\|r_k\|}{\|r_0\|} \leq \min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in \lambda(\mathcal{A})} |p_k(\lambda)|.$$

Diagonalizable systems – $\kappa(\mathcal{V})$ can be large;

General nonnormal (nondiagonalizable) systems – we can get any convergence behaviour independently on the spectrum [Greenbaum, Pták, Strakoš, 1996].

Another approaches: field of values and pseudospectra [Ernst, 2000], [Starke, 1997], [Nachtigal, Reddy, Trefethen, 1992], polynomial numerical hull [Greenbaum, 2002].

The initial residual should be included in the analysis [Liesen, Strakoš, 2005].

Symmetric case:

The convergence of symmetric iterative methods is essentially determined by the eigenvalue distribution.

Positive definite case:

$$\lambda(\mathcal{A}) \subset [a, b],\ 0 < b < a \Rightarrow$$

$$\min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in \lambda(\mathcal{A})} |p_k(\lambda)| \leq 2 \left(\frac{\sqrt{a} - \sqrt{b}}{\sqrt{a} + \sqrt{b}} \right)^k ,$$

Indefinite case:

$$\lambda(\mathcal{A}) \subset [-a,-b] \cup [c,d],\ 0 < b < a,\ 0 < c < d,\ a-b = d-c \Rightarrow$$

$$\min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in \lambda(\mathcal{A})} |p_k(\lambda)| \leq 2 \left(\frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \right)^{\left[\frac{k}{2}\right]},$$

The asymptotic convergence rate can be estimated

$$\lim_{k \to \infty} \left(\frac{\|r_k\|}{\|r_0\|} \right)^{\frac{1}{k}} \leq \lim_{k \to \infty} \left(\min_{\substack{p_k \in P_k \\ p_k(0)=1}} \max_{\lambda \in I^- \cup I^+} |p_k(\lambda)| \right)^{\frac{1}{k}}.$$

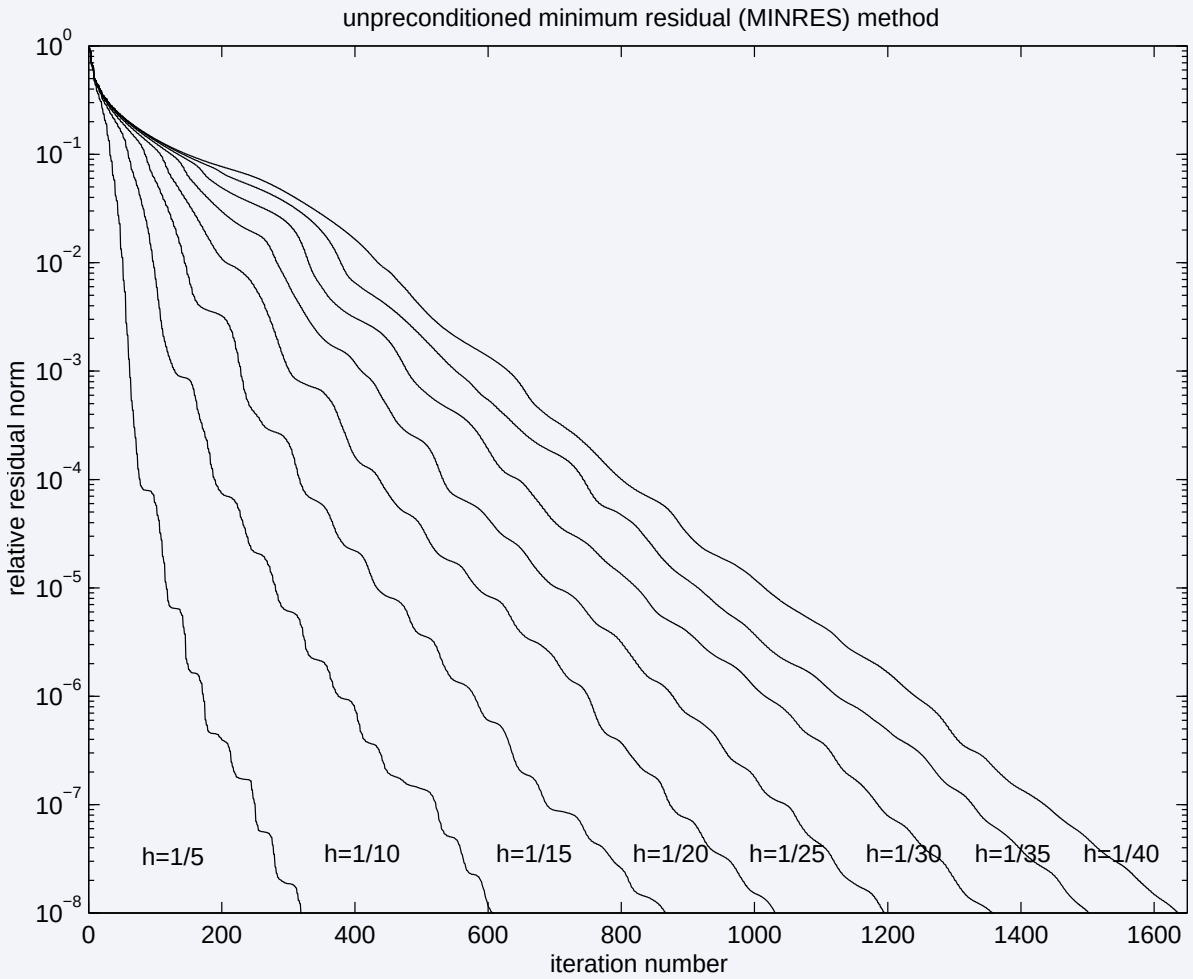
Unpreconditioned MINRES on the whole saddle point system:

$$\lim_{k\to\infty} \left(\frac{\|r_k\|}{\|r_0\|} \right)^{\frac{1}{k}} \leq 1 - c_1 h$$

iteration count / average contraction count:

h	$\frac{\|r_n\|}{\|r_0\|} = 10^{-4}$	$\frac{\|r_n\|}{\|r_0\|} = 10^{-8}$	$\frac{\|r_n\|}{\|r_0\|} = 10^{-12}$
1/2	58/0.853	78/0.789	98/0.754
1/3	164/0.945	245/0.927	365/0.927
1/4	229/0.960	389/0.953	520/0.948
1/6	350/0.974	669/0.973	1033/0.973
1/8	395/0.977	805/0.977	1168/0.977
1/12	529/0.982	1083/0.983	1682/0.983

Unpreconditioned MINRES on the whole saddle point system:

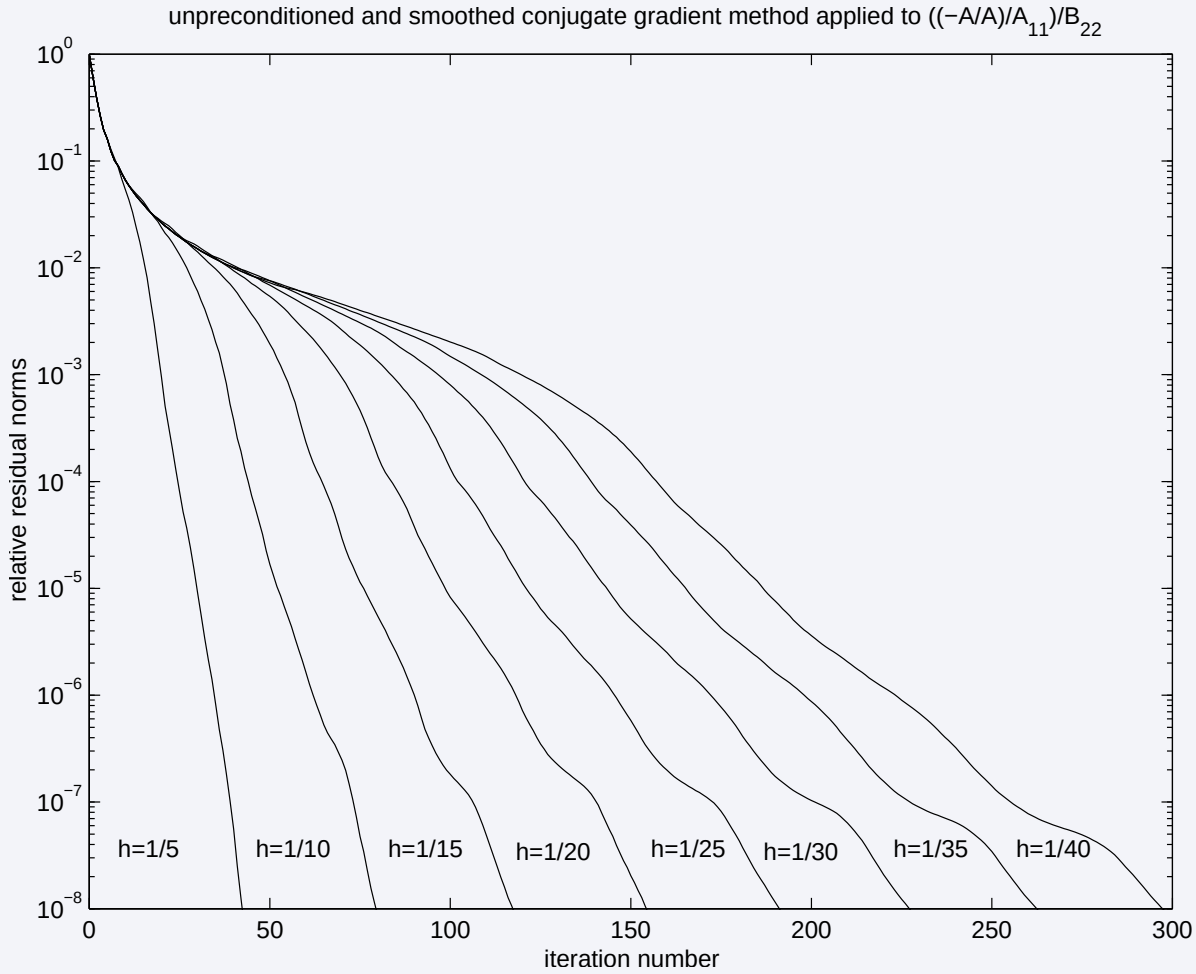


(Preconditioned) CR on the Schur complement systems:

$$\lim_{k\to\infty}\left(\frac{\|r_k\|}{\|r_0\|}\right)^{\frac{1}{k}} \leq 1 - c_6 h.$$

	unpreconditioned / preconditioned MR			
NE	$\mathcal{A}$	$-\mathcal{A}/A$	SC_2	SC_3
250	1247	247/29	139/20	110/20
2000	2182	529/53	291/39	262/39
6750	2951	763/75	416/58	394/58
16000	3198	1001/95	545/78	525/78
54000	4209	1471/137	800/117	776/117
128000	4545	1934/177	1027/151	1023/155

(Preconditioned) CR on the Schur complement systems:



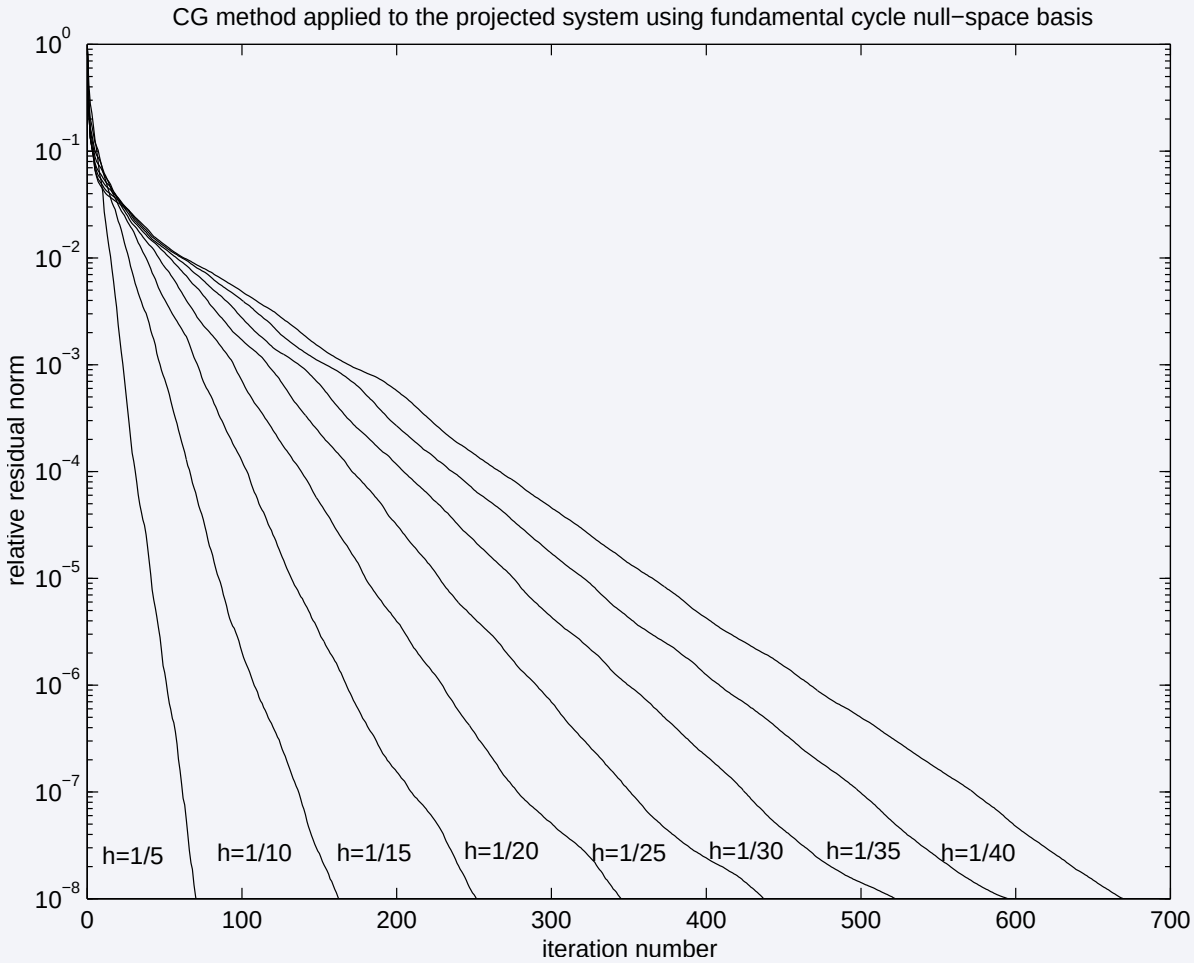
CG on the projected system (fundamental cycle null-space basis):

$$\lim_{k \to \infty} \left(\frac{\|r_k\|}{\|r_0\|} \right)^{\frac{1}{k}} \leq 1 - c_6 h^2$$

(in practice better than in theory).

h	memory requirements		iteration counts	
	QR $NNZ(QR)$	FC $NNZ(Z1)$	QR QR/SN	FC UN
1/5	28360 (3e-2)	3360 (7e-3)	22/20 (0.17/0.44)	71 (0.08)
1/10	410466 (0.97)	47120 (0.07)	22/21 (1.87/4.23)	163 (1.57)
1/15	1979203 (9.73)	226780 (0.30)	22/21 (8.48/17.1)	252 (19.9)
1/20	7120947 (59.6)	697840 (0.93)	22/21 (25.0/48.6)	346 (75.9)
1/25	18105131 (237)	1675800 (2.21)	22/21 (57.2/107)	438 (222)
1/30	40837823 (980)	3436160 (4.60)	21/21 (110/214)	523 (510)
1/35	—	6314420 (8.64)	—	596 (1009)
1/40	—	10706080 (14.8)	—	670 (1900)

CG on the projected system (fundamental cycle null-space basis):

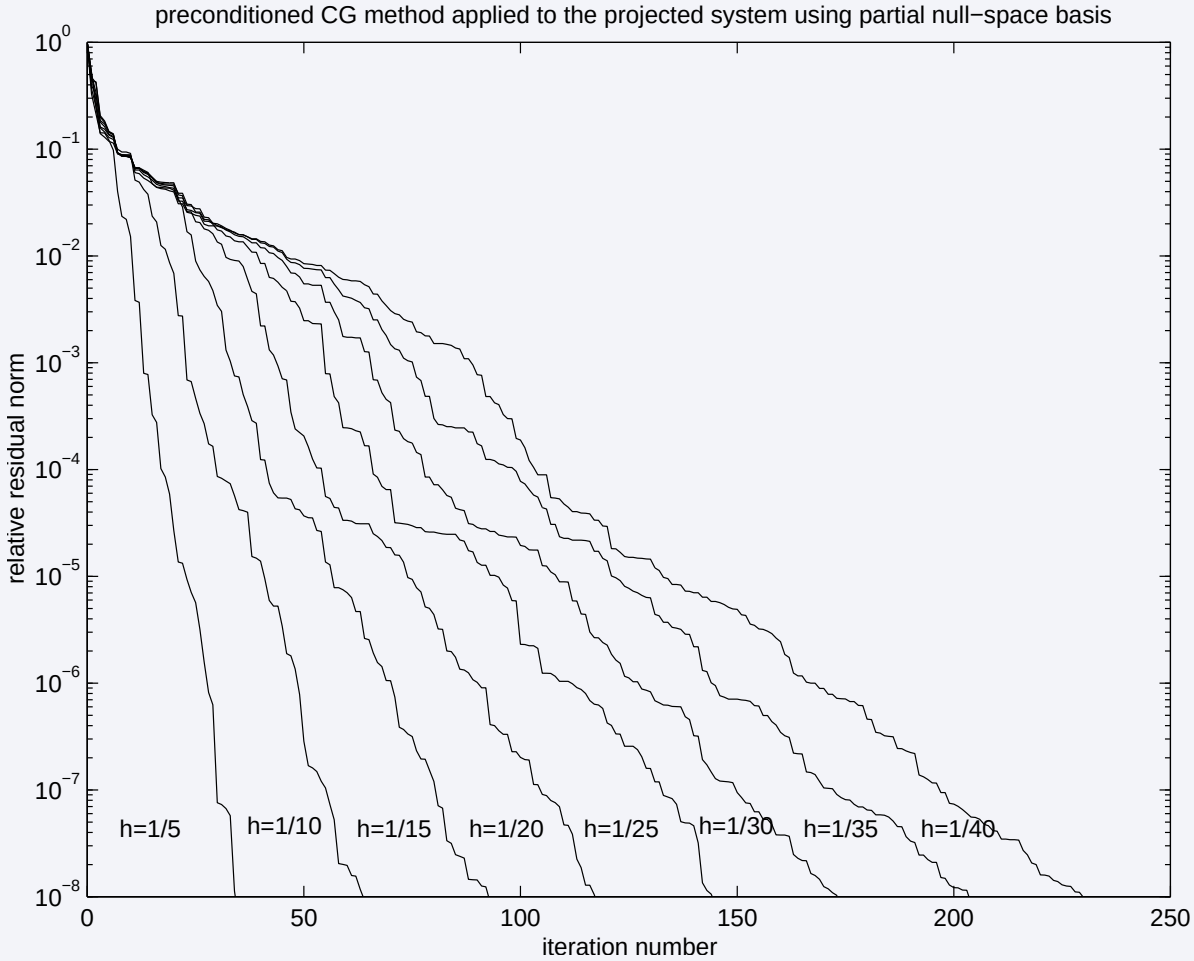


CG on the projected system (partial null-space basis):

$$\lim_{k\to\infty}\left(\frac{\|r_k\|}{\|r_0\|}\right)^{\frac{1}{k}} \leq 1 - c_6 h.$$

h	NNZ	implicit	sparse QR	
		IP/IQ	$NNZ(QR)$	QR/SN
1/5	14375	62/35 (0.05/0.03)	20834 (0.02)	18/14 (0.09/0.09)
1/10	123000	103/64 (0.68/0.48)	356267 (0.35)	19/16 (1.11/0.89)
1/15	424125	144/93 (5.17/3.79)	1840670 (3.14)	21/15 (6.09/4.63)
1/20	1016000	186/118 (20.2/14.2)	6322468 (17.97)	21/15 (18.3/14.94)
1/25	1996875	225/145 (50.8/37.4)	16661544 (86.6)	23/15 (47.0/27.8)
1/30	3465000	260/174 (111/84.2)	40669978 (584)	22/15 (96.7/85.5)
1/35	5518625	295/204 (224/173)	—	—
1/40	8256000	331/230 (383/295)	—	—

CG on the projected system (partial null-space basis):



Preconditioning of saddle point problems

Preconditioning = transformation of $\mathcal{A}u = b$ into another system.

- left: $\mathcal{M}^{-1}\mathcal{A}u = \mathcal{M}^{-1}b$,
- right: $\mathcal{A}\mathcal{M}^{-1}v = b$, $u = \mathcal{M}^{-1}v$,
- two-sided: $\mathcal{M} = \mathcal{M}_1\mathcal{M}_2$, $\mathcal{M}_1^{-1}\mathcal{A}\mathcal{M}_2^{-1}v = \mathcal{M}_1^{-1}b$, $u = \mathcal{M}_2^{-1}v$.

- Better convergence properties on the preconditioned system ($\mathcal{M}$ should be a "good" approximation to $\mathcal{A}$),
- $\mathcal{M}$ (or $\mathcal{M}^{-1}$) should be easily computed and the system with $\mathcal{M}$ should be easily solved.

For symmetric systems, the convergence of iterative methods depends on the distribution of eigenvalues of the system matrix $\longrightarrow$ the cluster of eigenvalues and/or reduced conditioning ensures fast convergence.

For nonsymmetric systems, the cluster of eigenvalues may be not enough (but it is in a practice) – reduction of minimal polynomial degree.

- Pure algebraic preconditioners – incomplete factorizations, sparse approximate inverses, algebraic multilevel methods [Saad, 2003], [Benzi, 2002].
- Application dependent preconditioners – the information about the underlying continuous problem is needed.

  The preconditioner quality depends on how much information from the original problem we use.

  The range of problems, that can be treated by a particual preconditioner, is limited.

Iterative solution – the method of choice:

- Symmetric positive definite case + positive preconditioner:
  $\rightarrow$ CG.
- Symmetric indefinite case + positive definite preconditioner:
  $\rightarrow$ CG, MINRES, SYMMLQ.
- Symmetric indefinite case + indefinite preconditioner:
  $\rightarrow$ GMRES; Simplified BiCG and QMR.

Symmetric indefinite system + symmetric positive definite preconditioner:

$$\mathcal{A}u = b$$

$\mathcal{A}$ symmetric indefinite, $\mathcal{M}$ symmetric positive definite.

$$\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}v = \mathcal{M}^{-\frac{1}{2}}b,\ u = \mathcal{M}^{-\frac{1}{2}}v,$$

$\tilde{\mathcal{A}}\tilde{u} = \tilde{b}$, $\tilde{\mathcal{A}}$ is symmetric, but indefinite!

Symmetric indefinite system + indefinite preconditioner:

$$\mathcal{A}u = b$$

$\mathcal{A}$ symmetric indefinite, $\mathcal{M} = \mathcal{M}_1\mathcal{M}_2$ symmetric indefinite.

$\mathcal{M}_1$ and $\mathcal{M}_2$ can be nonsymmetric.

$$\mathcal{M}_1^{-1}\mathcal{A}\mathcal{M}_2^{-1}v = \mathcal{M}_1^{-1}b,\ u = \mathcal{M}_2^{-1}v,$$

$$\tilde{\mathcal{A}}\tilde{u} = \tilde{b},\ \ \tilde{\mathcal{A}} \text{ is nonsymmetric!}$$

Iterative solution of indefinitely preconditioned nonsymmetric system:

$$\tilde{\mathcal{A}} = \mathcal{M}_1^{-1}\mathcal{A}\mathcal{M}_2^{-1},\ \mathcal{J} = \mathcal{M}_1^T\mathcal{M}_2$$
$$\Downarrow$$
$$\mathcal{A}^T\mathcal{J} = \mathcal{J}\mathcal{A}$$

Simplified $\mathcal{J}$-symmetric Lanczos process

$$\mathcal{A}V_k = V_{k+1}T_{k+1,k},\ \ \mathcal{A}^T W_k = W_{k+1}\tilde{T}_{k+1,k},$$
$$W_k^T V_k = I \ \Rightarrow\ W_n = \mathcal{J}V_n.$$

$\mathcal{J}$-symmetric variant of Bi-CG and QMR [Freund, Nachtigal, 1995].

Iterative solution of preconditioned system with simplified Lanczos process:

$\mathcal{J}$-symmetric Bi-CG algorithm **is nothing but** classical CG algorithm preconditioned with indefinite matrix $\mathcal{J}$!

Preconditioned conjugate gradients method applied to indefinite system with indefinite preconditioning **is in fact** conjugate gradients method applied to nonsymmetric (and often nonnormal) preconditioned system with $\mathcal{A}\mathcal{M}^{-1}$.

Nevertheless, it frequently works in practice [R, Simoncini, 2002].

$\mathcal{J}$-symmetric Bi-CG + QMR-smoothing $\Rightarrow$ $\mathcal{J}$-symmetric QMR [Freund, Nachtigal, 1995], [Walker, Zhou, 1994].

Preconditioners for saddle point problems:

Preconditioners for saddle point problems exploit their block structure and we need the information about the problem – (good) saddle point preconditioners are application dependent.

Basic preconditioning schemes for saddle point problems (overview [Zulehner, 2002], [Axelsson, Neytcheva, 2003]):

- block preconditioners,
- constraint preconditioners,
- incomplete factorizations for symmetric indefinite systems.

Block preconditioners rely on the availability of (approximate) solution of systems with A and $S = B^T A^{-1} B$.

Block factorization of $\mathcal{A}$:

$$\mathcal{A} = \underbrace{\begin{pmatrix} I & 0 \\ B^T A^{-1} & -I \end{pmatrix} \underbrace{\begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}}_{\text{block diagonal preconditioner}} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}}_{\text{block triangular preconditioner}}.$$

Block diagonal preconditioners:

$$\mathcal{M} = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}.$$

The preconditioned saddle point matrix is diagonalizable and

$$\lambda(\mathcal{M}^{-1}\mathcal{A}) = \{1, \tfrac{1}{2}(1 \pm \sqrt{5})\}$$

[Murphy, Golub, Wathen, 2000] – GMRES terminates in at most three steps (also true for $C \neq 0$ [Ipsen, 2001]).

Application of the exact preconditioner is expensive – inexact preconditioning

$$\hat{\mathcal{M}} = \begin{pmatrix} \hat{A} & 0 \\ 0 & \hat{S} \end{pmatrix}, \ \hat{A} \approx A, \ \hat{S} \approx S.$$

Block diagonal preconditioners:

- In [Silvester, Wathen, 1993, 1994], the case of $C \neq 0$ is considered. They give the spectral bounds for various choices of $\hat{A}$ and $\hat{S}$ assuming the spectral equivalence of A and $\hat{A}$ and $B^T A^{-1} B$ and $\hat{S}$.
- In [Fisher, Ramage, Silvester, Wathen, 1998] (A spd, $C = 0$), the preconditioner has the form $\hat{A} = \alpha A$, $\hat{S} = \beta \tilde{S}$ with $\alpha > 0$, $\beta \in \{-1, 1\}$ and $\tilde{S} \approx B^T A^{-1} B$. The spectrum of for $\beta = -1$ and $\beta = 1$ is different but the convergence of MINRES and GMRES is the same for both choices of β (for a fixed α) for the initial residual of the form $r_0 = (0, *)^T$.
- In [de Sturler, Liesen, 2005], the authors gave bounds for the spectrum of $\hat{\mathcal{M}}$ with respect to the spectrum of $\mathcal{M}$ with $\hat{A}$ and $\hat{S} = B^T \hat{A}^{-1} B$ (for the generalized saddle point problem).

Block diagonal preconditioners:

- Stokes problem with and without stabilization – using diagonal scaling [Wathen, Silvester, 1993]:

$$\lambda(\mathcal{M}^{-1}\mathcal{A}) \subset (-a, -bh) \cup (ch^2, d).$$

  Estimates for the asymptotic convergence rate [Wathen, Fisher, Silvester, 1995].

- Optimization [Battermann, Heinkenschloss, 1998], [Lukšan, Vlček, 1998], scattered data interpolation [Lyche, Nilssen, Winther, 2002], unsteady Stokes problem [Mardal, Winther, 2004], elasticity and Stokes problems [Pavarino, 1997, 1998], [Chizhonkov, 2001], [Klawonn, 1998], [Krzyżanowski, 2001], [Peters, Reichelt, Reusken, 2004], mixed finite approximation of elliptic PDEs [Kuznetsov, 1995, 2004], [Perugia, Simoncini, 1999], [Powel, Silvester, 2004], [Vassilevski, Lazarov, 1996]. etc.

Block triangular preconditioners:

$$\mathcal{M} = \begin{pmatrix} A & B \\ 0 & S \end{pmatrix}.$$

The preconditioning matrix is diagonalizable and its spectrum is

$$\lambda(\mathcal{M}^{-1}\mathcal{A}) = \{\pm 1\}$$

$$\mathcal{M} = \begin{pmatrix} A & B \\ 0 & -S \end{pmatrix}.$$

The preconditioned saddle point matrix has the spectrum

$$\lambda(\mathcal{M}^{-1}\mathcal{A}) = \{1\}$$

but is not diagonalizable.

In both cases, the minimal polynomial degree is equal to 2 [Murphy, Golub, Wathen, 2000], [Ipsen, 2001] – GMRES terminates in at most two steps.

Block triangular preconditioners:

Inexact preconditioning

$$\hat{\mathcal{M}} = \begin{pmatrix} \hat{A} & B \\ 0 & \hat{S} \end{pmatrix}, \ \hat{A} \approx A, \ \hat{S} \approx S.$$

The application of block triangular preconditioner:

$$\hat{\mathcal{M}}^{-1} = \begin{pmatrix} \hat{A}^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & -B \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & \hat{S}^{-1} \end{pmatrix}.$$

Block triangular preconditioners:

- $\hat{\mathcal{M}}^{-1}\mathcal{A}$ is nonnormal – field of values analysis [Klawonn, Starke, 1999], [Loghin, Wathen, 2004].
- For symmetric problems, the symmetry is destroyed (symmetrization [Bramble, Pasciak, 1988] seldom necessary) – compensated by the fast convergence of GMRES.
- Other analyses and applications: [Batterman, Heinkenschloss, 1998], [Cao, 2004], [Elman, Silvester, Wathen, 2002], [Kanschat, 2003], [Klawonn, 1998], [Krzyżanovski, 2001], [Pavarino, 1998]; inexact preconditioning: [Bramble, Pasciak, 1988], [Simoncini, 2004], [Zulehner, 2002].

Constraint preconditioners:

$$\mathcal{M} = \begin{pmatrix} M & B \\ B^T & 0 \end{pmatrix}.$$

Mixed finite element approximations of elliptic PDEs: [Axelsson, Neytcheva, 2003], [Bank, Welfert, Yserentant, 1990], [Ewing, Lazarov, Lu, Vassilevski, 1990], [Perugia, Simoncini, 2000], [R, Simoncini, 2002], [Tong, Sameh, 1998].

Optimization: [Lukšan, Vlček, 1998], [Dyn, Ferguson, 1983], [Gould, Hribar, Nocedal, 2001], [Keller, Gould, Wathen, 2000], [Bergamaschi, Gondzio, Zilli, 2004].

Properties of the preconditioned matrix:

$\mathcal{M}^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{M}^{-1}$ are not diagonalizable (and have at most 2×2 Jordan blocks).

The degree of the minimal polynomial of $\mathcal{M}^{-1}\mathcal{A}$ is equal to $n - m + 2$ [Lukšan, Vlček, 1998], [Keller, Gould, Wathen, 2000].

A and M symmetric, B of full rank ($m < n$), $\mathcal{A}$ and $\mathcal{M}$ nonsingular $\Rightarrow$

$$\lambda(\mathcal{M}^{-1}\mathcal{A}) = \underbrace{\{1\}}_{\text{multiplicity } 2m} \cup \underbrace{\lambda\left([(I-\Pi)M(I-\Pi)]^{\dagger}(I-\Pi)A(I-\Pi)\right) \setminus \{0\}}_{n-m \text{ eigenvalues}}$$

[Keller, Gould, Wathen, 2000].

The classical (null-space projection) constaint preconditioner:

$$\mathcal{A}\mathcal{M}^{-1} = \begin{pmatrix} A(I-\Pi)+\Pi & (A-I)B(B^TB)^{-1} \\ 0 & I \end{pmatrix}, \quad \Pi = B(B^TB)^{-1}B^T.$$

The spectrum of (nondiagonalizable) $\mathcal{A}\mathcal{M}^{-1}$ has the form

$$\lambda(\mathcal{A}\mathcal{M}^{-1}) \subset \{1\} \cup \lambda(A(I-\Pi)+\Pi) \subset \{1\} \cup \lambda((I-\Pi)A(I-\Pi)) \setminus \{0\}.$$

The solution of preconditioned system (with $r_0 = (r_0^{(x)}, 0)^T$) is equivalent to the null-space projection method, i.e. to the solution of

$$(I-\Pi)A(I-\Pi)x = (I-\Pi)f.$$

$\Rightarrow$ PCG can be applied (with an appropriate safeguarding strategy) [Lukšan, Vlček, 1998], [R, Simoncini, 2002], [Gould, Hribar, Nocedal, 2001].

Constraint preconditioner:

- [Golub, Wathen, 1998] – A nonsymmetric but positive real (Oseen equation), $M = \frac{1}{2}(A + A^T)$; effective for sufficiently large viscosities; inexact solves [Baggag, Sameh, 2004].
- [Botchev, Golub, 2004] – small viscosity flows with M incorporating the skew-symmetric part of A.
- [Axelsson, Neytcheva, 2003], [Bergamashi, Gondzio, Zilli, 2004], [Dollar, 2005], [Durazzi, Ruggiero, 2001], [Perugia, Simoncini, 2000], [Toh, Phoon, Chan, 2004], [Zulehner, 2002] – $A = A^T$, $C \neq 0$; systems with $C \neq 0$ are often easier to solve – iterative methods converge faster – regularized preconditioning (for problems with $C = 0$) [Axelsson, 1979] with

$$\mathcal{M} = \begin{pmatrix} A & B \\ B^T & -\varepsilon I \end{pmatrix}, \ \varepsilon > 0.$$

Implementation and numerical stability

Effects of rounding errors:

Delay of convergence:

- rounding errors slow down the real rate of convergence,
- rounding errors lead to loss of numerical rank of computed basis.

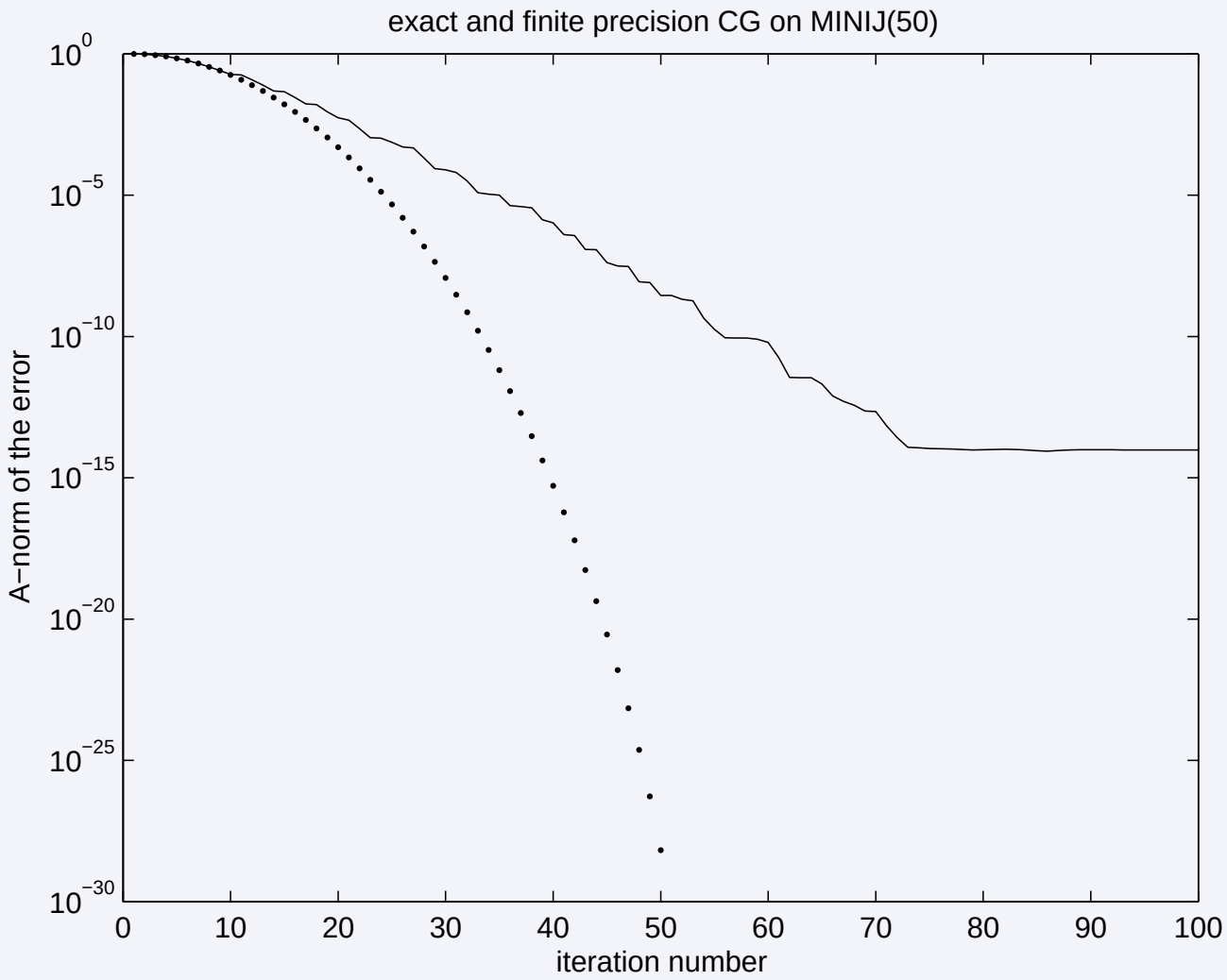
Limiting accuracy:

- there is a limit in the accuracy of computed iterates,

Look for better preconditioning or better methods which perhaps mitigate these effects.

Try more stable (but also more expensive) implementations.

Stopping criterion: level of termination cannot be arbitrarily small – it should be above the maximum attainable accuracy level, stopping criteria based on the backward error or related to the problem.


exact and finite precision CG on MINIJ(50)
A–norm of the error
iteration number
10^0
10^−5
10^−10
10^−15
10^−20
10^−25
10^−30
0
10
20
30
40
50
60
70
80
90
100

The Schur complement reduction:

$$\mathcal{A} = \begin{pmatrix} A & B_1 & B_2 & B_3 \\ B_1^T & 0 & 0 & 0 \\ B_2^T & 0 & 0 & 0 \\ B_3^T & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ p \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}$$

- Subsequent reduction to the Schur complement systems without additional fill-in: [Kaaschieter, Huijben, 1992], [Maryška, R, Tůma, 1996]

$$\mathcal{A} \to -\mathcal{A}/A \to (\mathcal{A}/A)/A_{11} - ((\mathcal{A}/A)/A_{11})/B_{22}.$$

  Numerical stability of the block LU decomposition [Demmel, Higham, 1995], [Higham, 1996], [Wilkinson, Reinsch, 1971], [Golub, Van Loan, 1989].
- Iterative solution of the final (symmetric positive definite) Schur complement system with the matrix $-((\mathcal{A}/A)/A_{11})/B_{22}$ for the unknown vector λ_1 by CG with the prescribed backward error tol.
- Block back-substitution process for the unknown vectors λ_2, p and u using the factors from the Schur complement reduction.

Maximum attainable accuracy of the computed approximate solutions:

$$\mathcal{A}\bar{u} = b + \Delta b,\ \mathcal{A} \in \mathbb{R}^{N\times N},$$

$$\begin{aligned}\|\Delta b\| \leq\ & O(\max\{\mathrm{tol}, N^{\frac{3}{2}}\varepsilon\}) \\ & \times \Big(\|A\| + \|B\| + (1 + \|B\|)\|A^{-1}\| \\ & \times \max\{\|A\|, \|B\|(1 + \|B\|)\|A^{-1}\|\}\sqrt{\kappa(A)\kappa(B)} \Big) \|\bar{u}\|.\end{aligned}$$

$$(\mathcal{A} + \mathcal{E})\bar{u} = b,$$

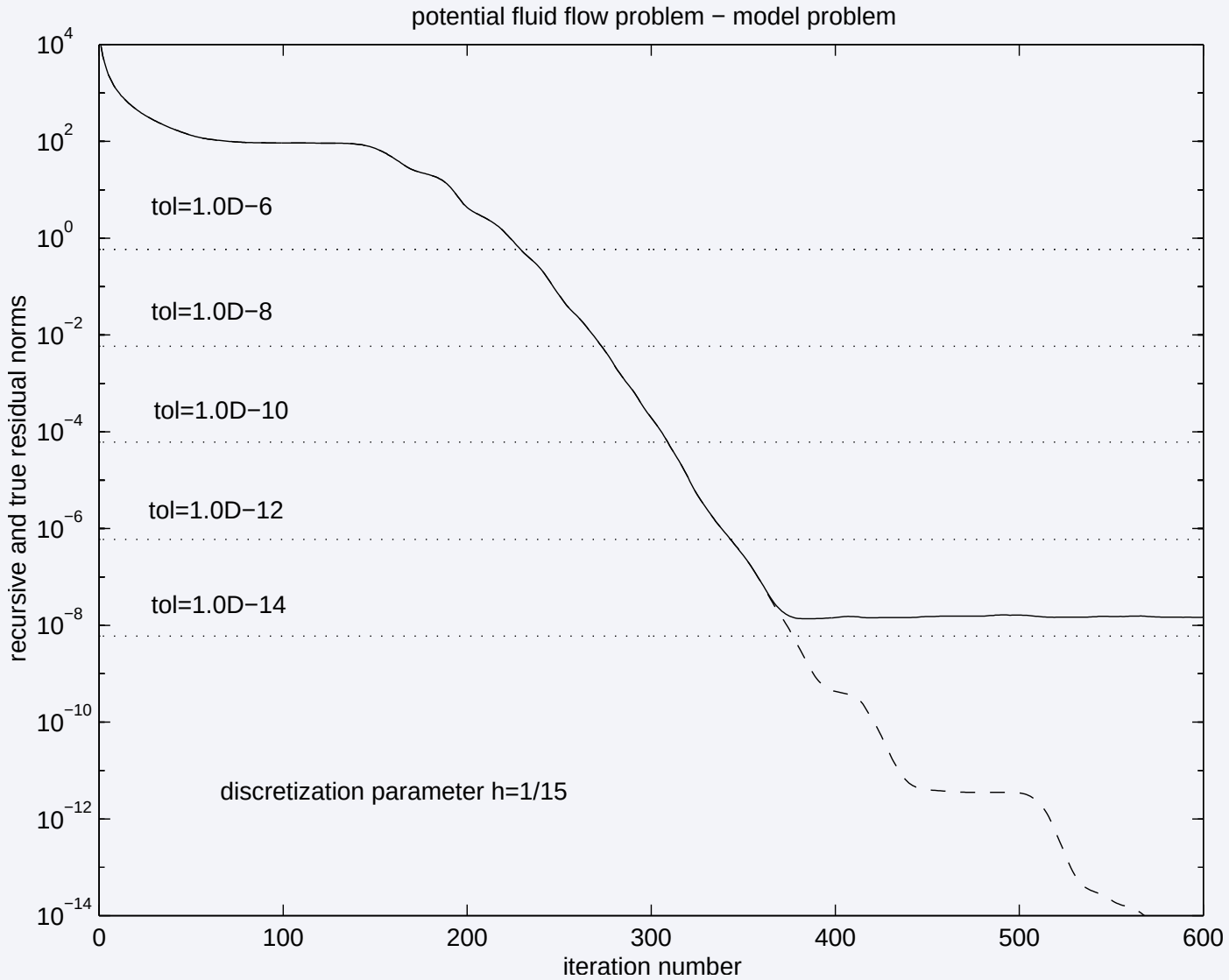
$$\|\mathcal{E}\| \leq O(\max\{\mathrm{tol}, N^{\frac{3}{2}}\varepsilon\} N^{-\frac{1}{3}}).$$

h	solution norm	true residual norm
1/5	12602.08	0.9455e-8
1/10	37302.81	0.1051e-7
1/15	69515.73	0.1481e-7
1/20	107777.54	0.2109e-7
1/30	199370.25	0.3202e-7
1/40	307999.99	0.4291e-7

True residual and solution norms for the tolerance tol$=10^{-14}$

tol	true residual	$\|\Delta b_1\|_\infty$	$\|\Delta b_2\|_\infty$	$\|\Delta b_3\|_\infty$
10^{-6}	0.5879e-0	0.568e-13	0.239e-9	0.213e-1
10^{-8}	0.5865e-2	0.568e-13	0.237e-9	0.269e-3
10^{-10}	0.6112e-4	0.567e-13	0.210e-9	0.312e-5
10^{-12}	0.5971e-6	0.491e-13	0.229e-9	0.224e-7
10^{-14}	0.1566e-7	0.380e-13	0.204e-9	0.659e-9

True residual norm by blocks for $h = 1/15$


potential fluid flow problem − model problem
recursive and true residual norms
iteration number
tol=1.0D−6
tol=1.0D−8
tol=1.0D−10
tol=1.0D−12
tol=1.0D−14
discretization parameter h=1/15
10^4
10^2
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
0
100
200
300
400
500
600

Maximum attainable accuracy of segregated methods:

- Segregated methods compute x or y as the solution of a reduced system, compute the remaining component by the back-substitution into the original system (this can be done using various back-substitution formulas).
- The limiting accuracy level of computed approximate solutions x_k and y_k measured by their (true) residuals $f - Ax_k - By_k$ and $-B^T x_k$ depends on:
  - the maximum over the norms of the iterates beginning from the initial step up to the current iteration step [Greenbaum, 1994, 1997],
  - the backward error associated with the approximate solutions of inner systems which are solved inexactly [Simoncini, Szyld, 2003], [van den Eshof, Sleijpen, 2004],
  - the back-substitution formula [J, R, 2006].

Schur complement reduction:

Solution of inner systems $Au = b$ with the backward error τ:

$$(A + \Delta A)\bar{u} = b + \Delta b, \ \frac{\|\Delta A\|}{\|A\|}, \ \frac{\|\Delta b\|}{\|b\|} \leq \tau, \ \tau\kappa(A) \ll 1.$$

- Choose y_0, $x_0 = A^{-1}(f - By_0)$, $r_0^{(y)} = -B^T x_0$.
- Update the approximation y_k and the residual $r_k^{(y)}$:

$$y_{k+1} = y_k + \alpha_k p_k^{(y)}, \ r_{k+1}^{(y)} = r_k^{(y)} + \alpha_k B^T A^{-1} B p_k^{(y)}.$$

- Compute the iterate x_{k+1}:

$$x_{k+1} = x_k - \alpha_k A^{-1} B p_k^{(y)} \qquad \text{(generic update)},$$
$$x_{k+1} = A^{-1}(f - By_{k+1}) \qquad \text{(direct substitution)},$$
$$x_{k+1} = x_k + A^{-1}(f - Ax_k - By_{k+1}) \qquad \text{(corrected direct substitution)}.$$

Null-space projection:

Solution of inner systems $Bv \approx c$ with the backward error τ:

$$(B + \Delta B)v \approx c + \Delta c, \ \frac{\|\Delta B\|}{\|B\|}, \ \frac{\|\Delta c\|}{\|c\|} \leq \tau, \ \tau\kappa(B) \ll 1.$$

- Choose $x_0 \in N(B^T)$, $y_0 = B^\dagger(f - Ax_0)$, $r_0^{(x)} = f - Ax_0 - By_0$.
- Update the approximation x_k and the residual $r_k^{(x)}$:

$$x_{k+1} = x_k + \alpha_k p_k^{(x)}, \ p_k^{(x)} \in N(B^T), \ r_{k+1}^{(x)} = r_k^{(x)} - \alpha_k(I - \Pi)A(I - \Pi)p_k^{(x)},$$

- Compute the iterate y_{k+1}:

$$y_{k+1} = y_k - \alpha_k B^\dagger A p_k^{(x)} \qquad \text{(generic update)},$$
$$y_{k+1} = B^\dagger(f - Ax_{k+1}) \qquad \text{(direct substitution)},$$
$$y_{k+1} = y_k + B^\dagger(f - Ax_{k+1} - By_k) \qquad \text{(corrected direct substitution)}.$$

Numerical experiment:

$$A = \text{tridiag}(1, 10^{-5}, -1) \in \mathbb{R}^{100\times 100},\ B = \text{rand}(100, 50) \in \mathbb{R}^{100\times 50}, f = (1, \ldots, 1)^T.$$

$$\kappa(A) = \|A\|\|A^{-1}\| \approx 2.00 \cdot 32.15 \approx 64.27,$$
$$\kappa(B) = \|B\|\|B^{\dagger}\| \approx 7.39 \cdot 0.75 \approx 5.55.$$

The Schur complement system and the projected system are solved with the GMRES, CGNE, BiCG and CGS method.

The Schur complement residual:

$$\| - B^T A^{-1} f + B^T A^{-1} B y_{k+1} - r_{k+1}^{(y)} \| \leq \frac{O(\tau)\kappa(A)}{1 - \tau\kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\| Y_{k+1}),$$

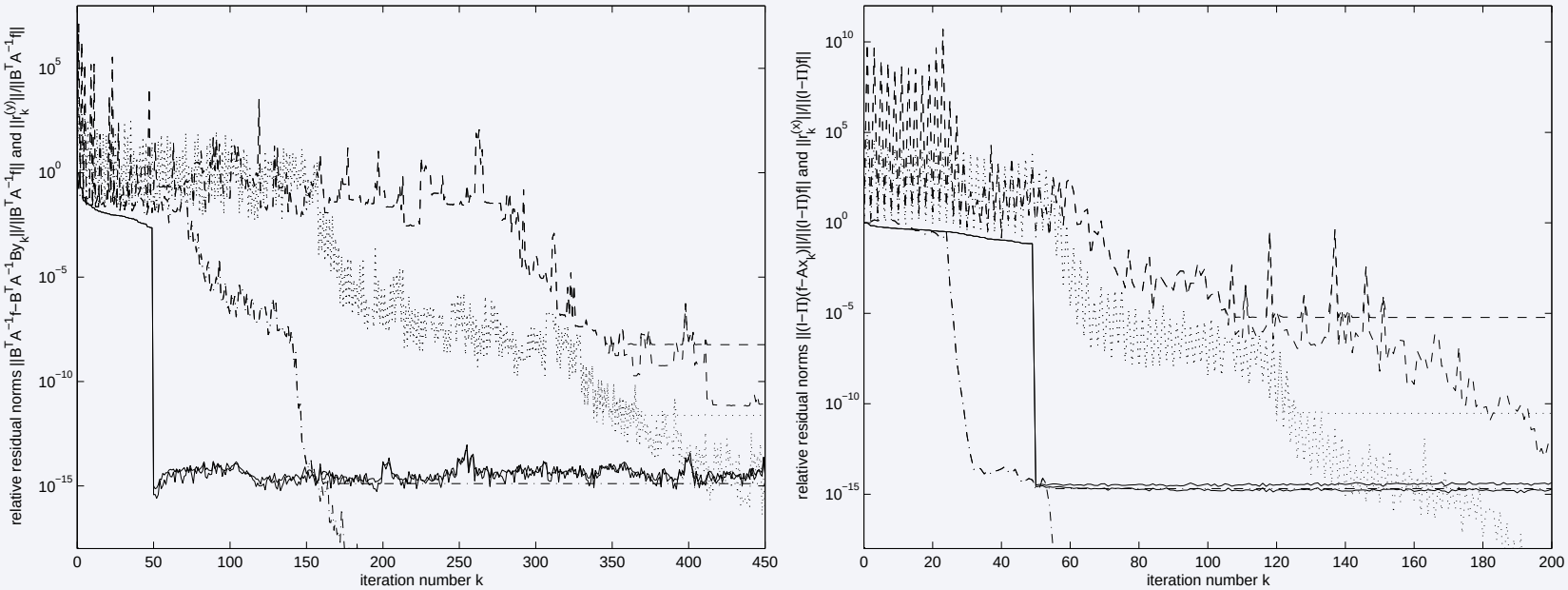
The projected residual:

$$\|(I - \Pi)(f - A x_{k+1} - r_{k+1}^{(x)})\| \leq \frac{O(u)\kappa(B)}{1 - \tau\kappa(B)} (\|f\| + \|A\| X_{k+1}),$$

$$X_{k+1} = \max_{i=0,1,\ldots,k+1} \|x_i\|, \quad Y_{k+1} = \max_{i=0,1,\ldots,k+1} \|y_i\|.$$

The quantities Y_{k+1} and X_{k+1} in the Schur complement reduction and in the null-space projection methods:

	Schur complement reduction		Null-space projection	
	Y_{k+1} (dir. sol.)	Y_{k+1} ($\tau = 10^{-12}$)	X_{k+1} (dir. sol.)	X_{k+1} ($\tau = 10^{-9}$)
GMRES	$1.6155 \cdot 10^{1}$	$1.6155 \cdot 10^{1}$	$3.9445 \cdot 10^{1}$	$3.9445 \cdot 10^{1}$
CGNE	$1.6157 \cdot 10^{1}$	$1.6156 \cdot 10^{1}$	$3.9445 \cdot 10^{1}$	$3.9445 \cdot 10^{1}$
BiCG	$9.8556 \cdot 10^{4}$	$1.5190 \cdot 10^{6}$	$6.5733 \cdot 10^{5}$	$6.5733 \cdot 10^{5}$
CGS	$3.3247 \cdot 10^{7}$	$7.7455 \cdot 10^{9}$	$5.2896 \cdot 10^{10}$	$5.2896 \cdot 10^{10}$



GMRES (solid), CGNE (dash-dotted), BiCG (dotted), CGS (dashed)

Schur complement reduction method:

Generic update:

$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(\tau)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|Y_{k+1}),$$

$$\| - B^T x_{k+1}\| \lesssim \frac{O(u)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|Y_{k+1}).$$

Direct substitution:

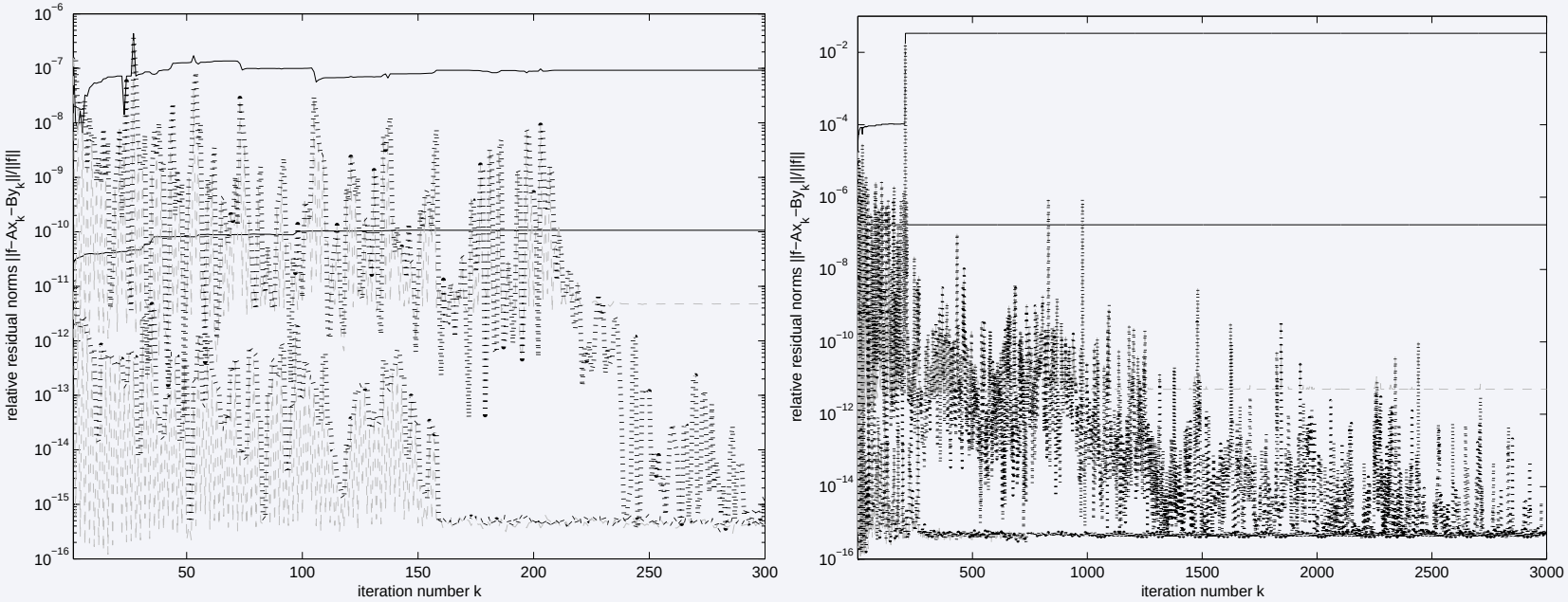
$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(\tau)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|\|y_{k+1}\|),$$

$$\| - B^T x_{k+1}\| \lesssim \frac{O(\tau)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|Y_{k+1}).$$

Corrected direct substitution:

$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(u)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|Y_{k+1}),$$

$$\| - B^T x_{k+1}\| \lesssim \frac{O(\tau)\kappa(A)}{1 - \tau\kappa(A)}(\|f\| + \|B\|Y_{k+1}).$$



BiCG (left), CGS (right)

Generic update (solid), direct substitution (dashed), corrected direct substitution (dotted)

Null-space projection method:

Generic update:

$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(u)\kappa(B)}{1 - \tau\kappa(B)}(\|f\| + \|A\|X_{k+1}).$$

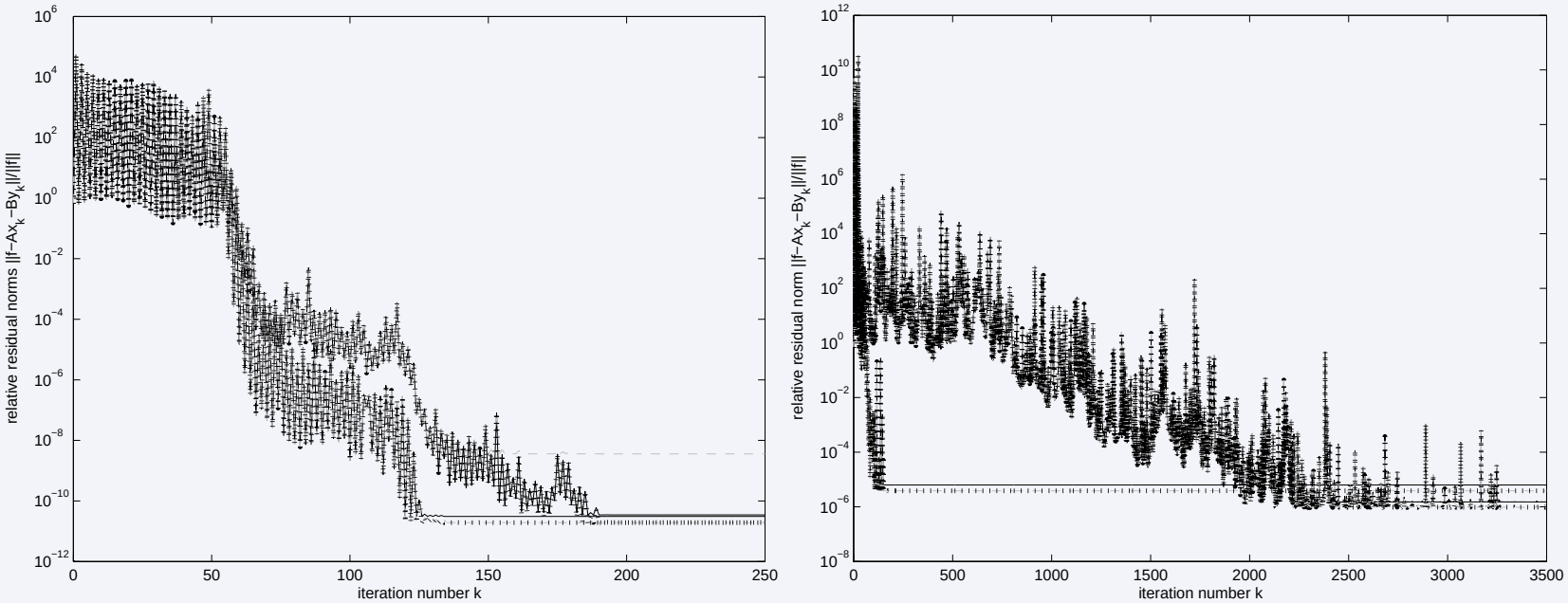
Direct substitution:

$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(\tau)\kappa(B)}{1 - \tau\kappa(B)}(\|f\| + \|A\|\|x_{k+1}\|) + \frac{O(u)\kappa(B)}{1 - \tau\kappa(B)}(\|f\| + \|A\|X_{k+1}).$$

Corrected direct substitution:

$$\|f - Ax_{k+1} - By_{k+1}\| \leq \frac{O(u)\kappa(B)}{1 - \tau\kappa(B)}(\|f\| + \|A\|X_{k+1}).$$

Level of $-B^T x_{k+1}$ depends on the departure of x_{k+1} from $N(B^T)$, $\| - B^T x_{k+1}\| \sim \kappa(B)\|B\|X_{k+1}$.



BiCG (left), CGS (right)

Generic update (solid), direct substitution (dashed), corrected direct substitution (dotted)

Saddle point system preconditioned with the constraint preconditioner:

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad \mathcal{M} = \begin{pmatrix} I & B \\ B^T & 0 \end{pmatrix},$$

PCG applied to $\mathcal{A}\mathcal{M}^{-1}$ with $r_0 = (r_0^{(x)}, 0)^T$ satisfies $r_k = (r_k^{(x)}, 0)^T$ for all k and is equivalent to CG applied to the projected system

$$(I - \Pi)A(I - \Pi)x = (I - \Pi)f$$

where $\Pi = B(B^T B)^{-1} B^T$ and

$$\|x - x_k\|_A = \min_{\tilde{x} \in x_0 + (I-\Pi)\,\mathrm{span}\{r_i^{(x)}\}_{i=0}^{k-1}} \|x - \tilde{x}\|_A$$

[Lukšan, Vlček, 1998], [Gould, Wathen, Keller, 1999]

The convergence of the residual norm:

In exact arithmetic, $x_k \longrightarrow x$, but in general $y_k \not\longrightarrow y$.

$$\Downarrow$$

$$r_{k+1} = \begin{pmatrix} r_k^{(x)} \\ 0 \end{pmatrix} \not\longrightarrow 0.$$

The approximations x_k are the iterates of CG applied to the projected system,

$$x - x_k = p_k((I - \Pi)A(I - \Pi))(x - x_0).$$

The first block component of the residual satisfies

$$r_k^{(x)} = p_k(A(I - \Pi) + \Pi)r_0^{(x)}.$$

Safeguarding techniques for $y_k \to y$:

- Scaling of the saddle point matrix [R, Simoncini, 2002]:

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \to \begin{pmatrix} DAD & B \\ B^T & 0 \end{pmatrix} \equiv \tilde{\mathcal{A}}$$

  such that

$$\{1\} \in \lambda((I - \Pi)DAD(I - \Pi)).$$

  Possible choices: $D = \alpha I = \tau^{-1} I$, where $\tau > 0$ lies in the field of values of $(I - \Pi)A(I - \Pi)$; scaling by the diagonal $D = \mathrm{diag}(A)$.

- Choose another direction vector [Braess, Deuflhard, Lipikov, 1999], [Hribar, Gould, Nocedal, 1999]:

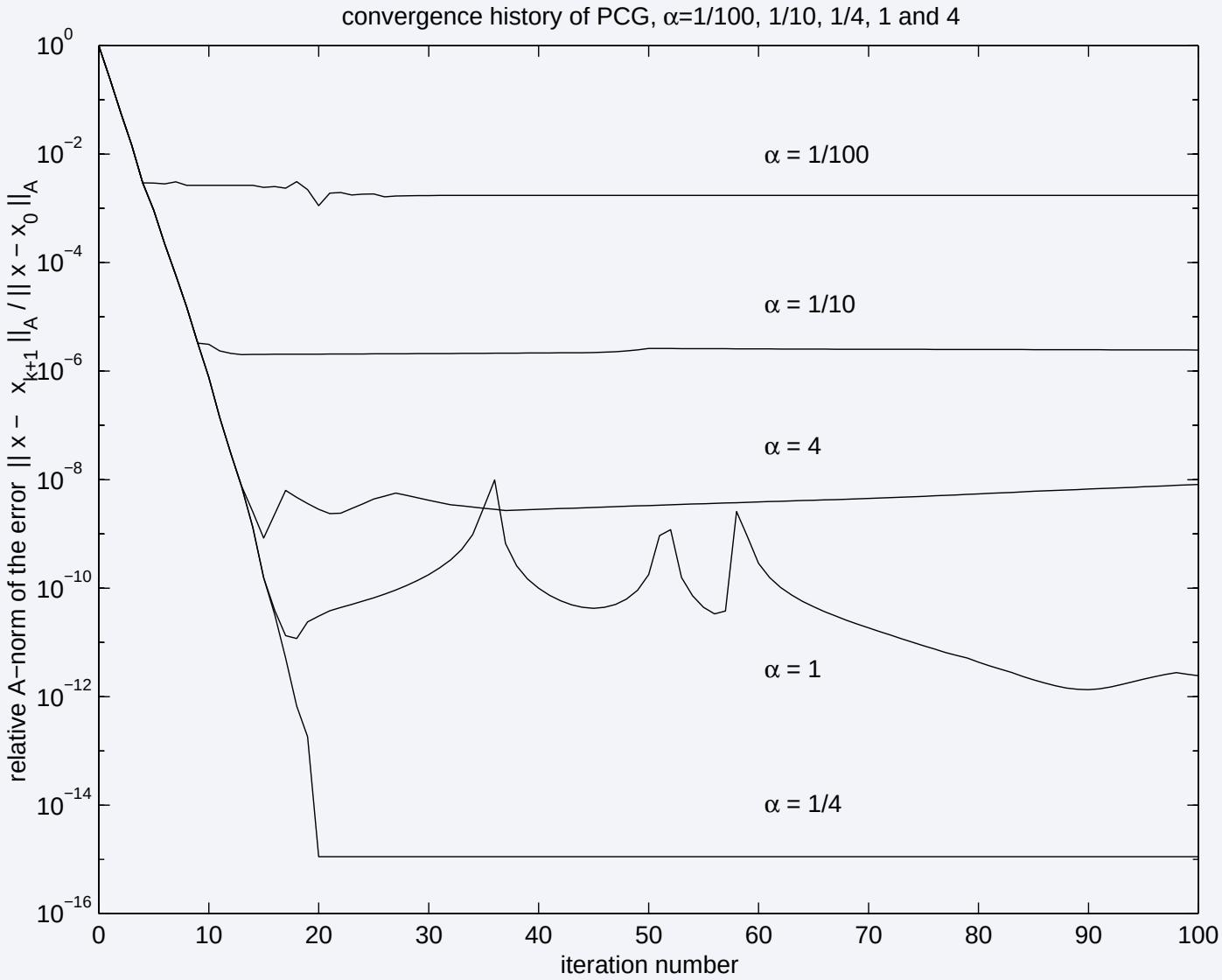
$$y_{k+1} = y_k + B^\dagger r_k^{(x)} \;\Leftrightarrow\; \|f - Ax_{k+1} - By_{k+1}\| = \min_{\tilde{y}} \|f - Ax_{k+1} - B\tilde{y}\|.$$

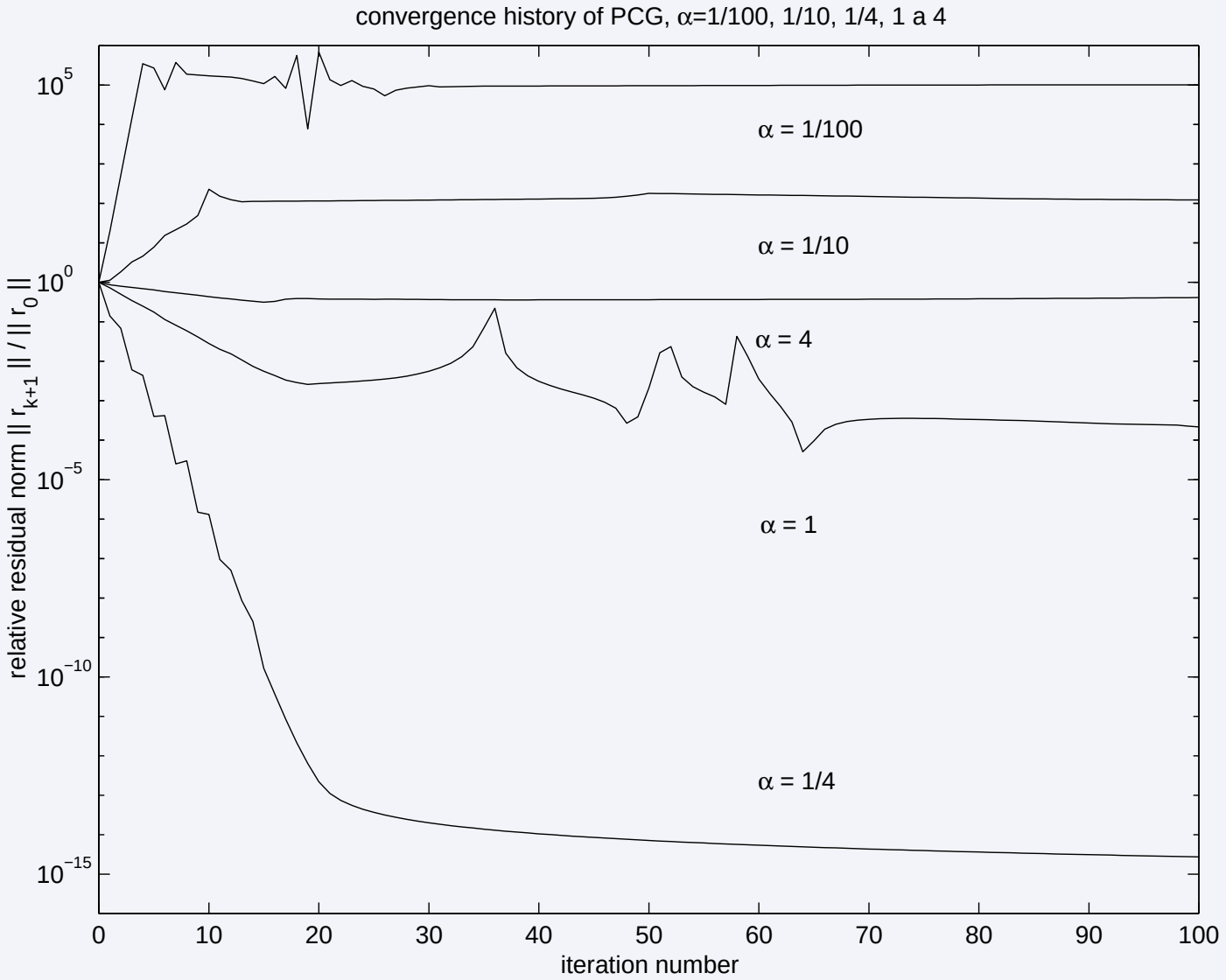
Numerical example:

$$A = \text{tridiag}(1, 4, 1) \in \mathbb{R}^{25\times 25},\ B = \text{rand}(25, 5) \in \mathbb{R}^{25\times 5},\ f = \text{rand}(25, 1) \in \mathbb{R}^{25}.$$

$$\lambda(A) \subset [2.0146, 5.9854]$$

α	spectrum of $\tilde{\mathcal{A}}\mathcal{M}^{-1}$
1	$[0.2067, 0.5856] \cup \{1\}$
1/4	$[0.5170, 1.4641]$
1/10	$\{1\} \cup [8.2712, 23.4252]$


convergence history of PCG, α=1/100, 1/10, 1/4, 1 and 4
relative A−norm of the error || x − x_{k+1} ||_A / || x − x_0 ||_A
iteration number
α = 1/100
α = 1/10
α = 4
α = 1
α = 1/4
10^0
10^−2
10^−4
10^−6
10^−8
10^−10
10^−12
10^−14
10^−16
0
10
20
30
40
50
60
70
80
90
100


convergence history of PCG, α=1/100, 1/10, 1/4, 1 a 4
relative residual norm || r_{k+1} || / || r_0 ||
iteration number
α = 1/100
α = 1/10
α = 4
α = 1
α = 1/4
10^5
10^0
10^−5
10^−10
10^−15
0
10
20
30
40
50
60
70
80
90
100

Behaviour in finite precision arithmetic:

The A-norm of the error $x - x_k$:

$$\begin{aligned}\|x - x_k\|_A &\leq \gamma_1 \|B^T(x - x_k)\| + \gamma_2 \|(I - \Pi)(f - Ax_k - By_k)\| \\ &\lesssim \gamma_1 \|r_k^{(y)}\| + \gamma_2 \|(I - \Pi) r_k^{(x)}\|\end{aligned}$$

(we can scale the preconditioner to minimize the departure [Björck, 1992] or use the residual update [Hribar, Gould, Nocedal, 1999]).

[Greenbaum, 1994, 1997], [Sleijpen et al. 1994]

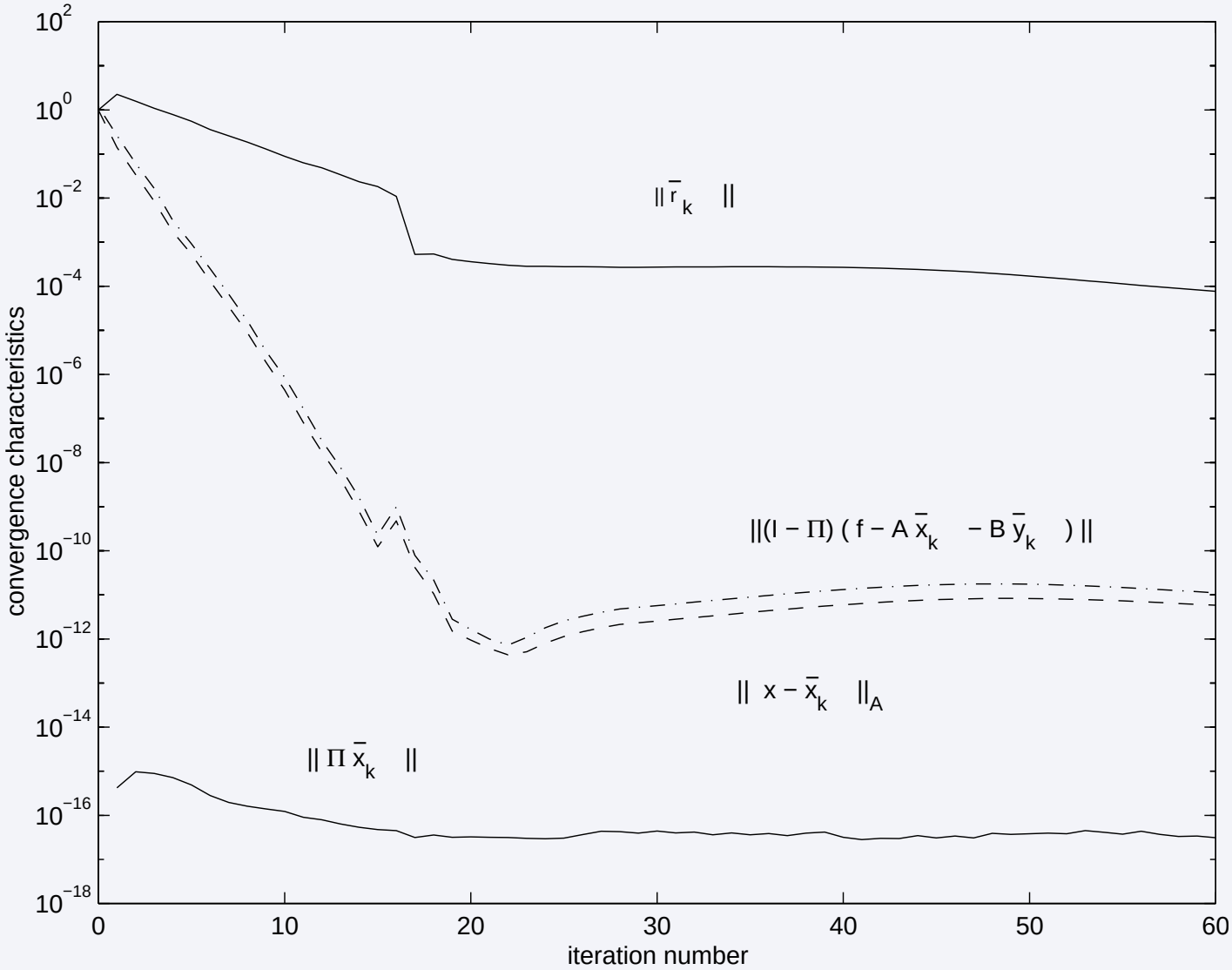
$$\|b - \mathcal{A}u_k\| \leq c\varepsilon\kappa(\mathcal{A}) \max_{i=0,1,\ldots,k} \|r_i\|$$

Good scaling $\Rightarrow \|r_k\| \to 0$ (nearly) monotonically $\Rightarrow$,

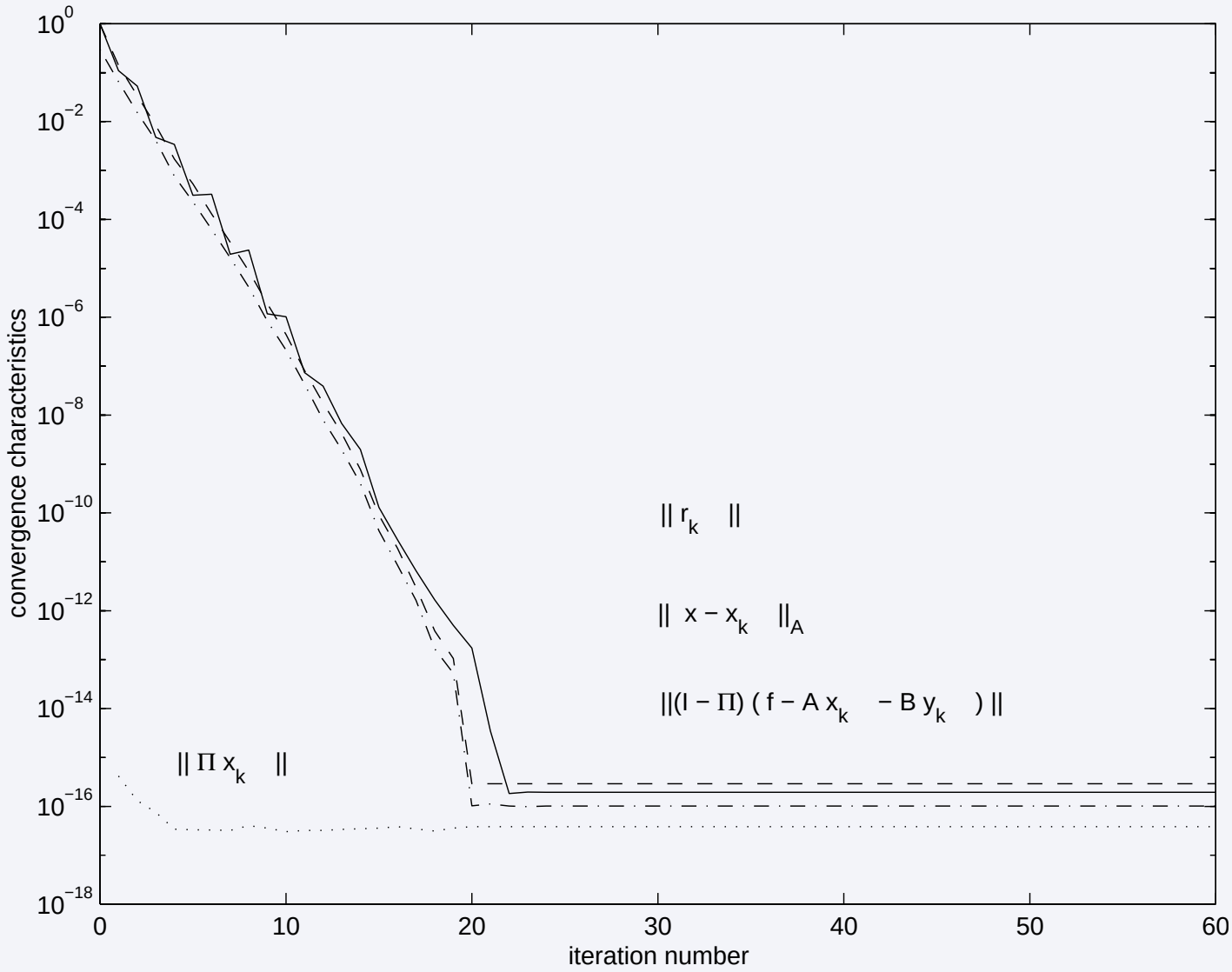
$$\|r_0\| \approx \max_{i=0,1,\ldots,k} \|r_i\|,$$

$$\|b - \mathcal{A}u_k\| \lesssim c\varepsilon\kappa(\mathcal{A})\|r_0\|.$$

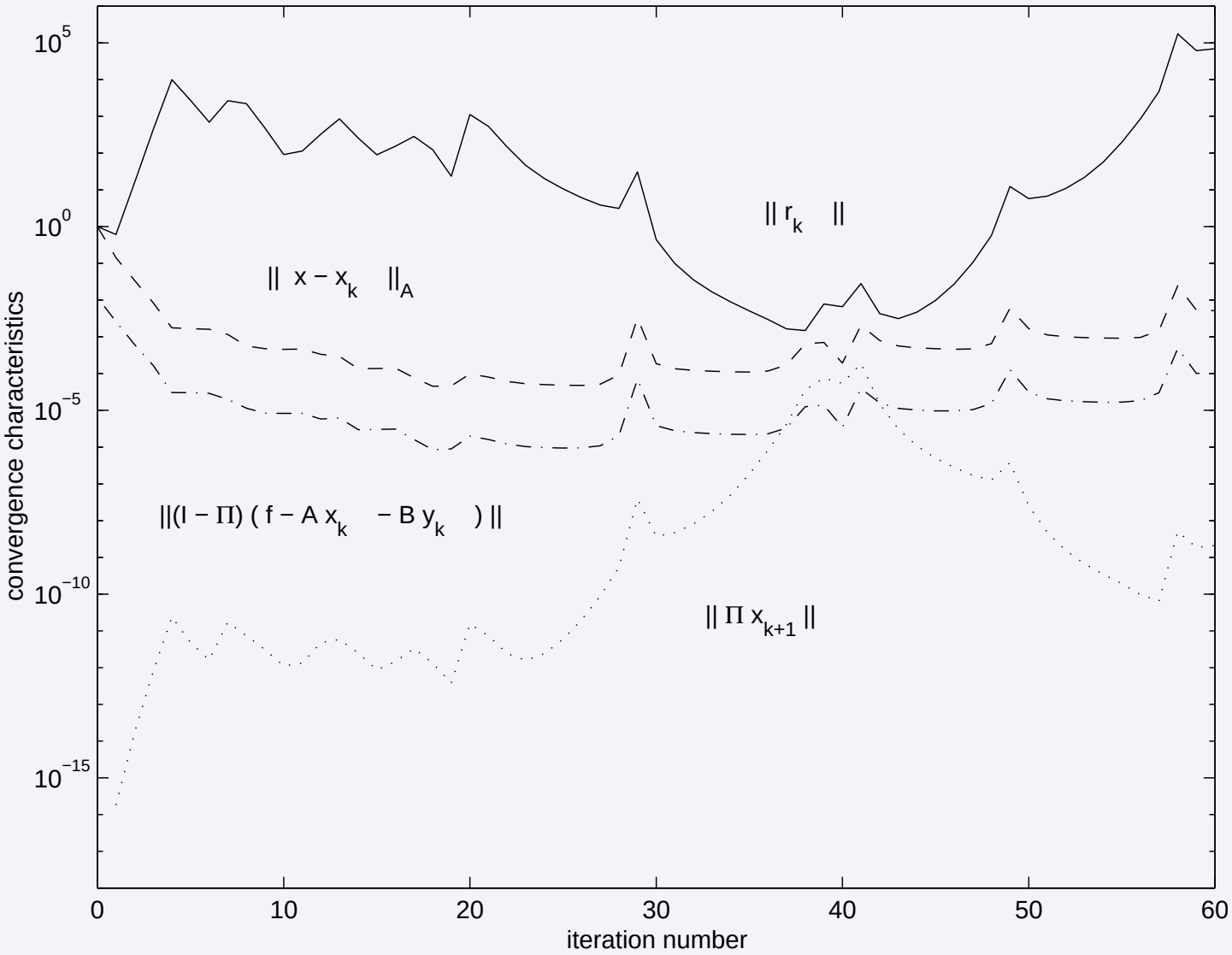
$\alpha = 1$


convergence characteristics
iteration number
$\| \bar{r}_k \|$
$\|(I - \Pi)(f - A\bar{x}_k - B\bar{y}_k)\|$
$\| x - \bar{x}_k \|_A$
$\| \Pi \bar{x}_k \|$
10^{2}
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
0
10
20
30
40
50
60

$$\alpha = 1/4$$


convergence characteristics
iteration number
|| r_k ||
|| x − x_k ||_A
||(I − Π) (f − A x_k − B y_k) ||
|| Π x_k ||

$\alpha = 1/100$


convergence characteristics
iteration number
10^{5}
10^{0}
10^{-5}
10^{-10}
10^{-15}
0
10
20
30
40
50
60
$\| r_k \|$
$\| x - x_k \|_A$
$\|(I - \Pi)(f - A x_k - B y_k)\|$
$\| \Pi x_{k+1} \|$