Structural Spectral properties of symmetric saddle point problems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

Joint work with W. Krendl and W. Zulehner

The problem

$$\begin{bmatrix} A & B^* \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

- Large scale

  $\Rightarrow$ Iterative solution with preconditioned Krylov subspace methods

- Structural properties. Focus for this talk:
  - $\star$ A symmetric positive (semi)definite
  - $\star$ B square, nonsingular
  - $\star$ C symmetric positive (semi)definite

Distributed optimal control for time-periodic parabolic equations

Problem: *Find the state* $y(x,t)$ *and the control* $u(x,t)$ *that minimize the cost functional*

$$J(y,u) = \frac{1}{2}\int_0^T \int_\Omega |y(x,t) - y_d(x,t)|^2 \; dx \; dt + \frac{\nu}{2}\int_0^T \int_\Omega |u(x,t)|^2 \; dx \; dt$$

subject to the time-periodic parabolic problem

$$\begin{aligned} \frac{\partial}{\partial t} y(x,t) - \Delta y(x,t) &= u(x,t) && \text{in } \Omega \times (0,T), \\ y(x,t) &= 0 && \text{on } \partial\Omega \times (0,T), \\ y(x,0) &= y(x,T) && \text{in } \Omega, \\ u(x,0) &= u(x,T) && \text{in } \Omega. \end{aligned}$$

Here $y_d(x,t)$ is a given target (or desired) state and $\nu > 0$ is a cost or regularization parameter.

Time-harmonic solution

Assume that y_d is time-harmonic: $\quad y_d(x,t) = y_d(x)e^{i\omega t}, \ \omega = \frac{2\pi k}{T}$

Then there exists a time-periodic solution
$y(x,t) = y(x)e^{i\omega t}, u(x,t) = u(x)e^{i\omega t}$, where $y(x), u(x)$ solve:

Minimize

$$\frac{1}{2}\int_\Omega |y(x) - y_d(x)|^2 \ dx + \frac{\nu}{2}\int_\Omega |u(x)|^2 \ dx$$

subject to

$$\begin{aligned} i\omega\, y(x) - \Delta y(x) &= u(x) \quad \textit{in } \Omega, \\ y(x) &= 0 \qquad\ \ \textit{on } \partial\Omega \end{aligned}$$

Discrete version:

$$\frac{1}{2}(y - y_d)^* M(y - y_d) + \frac{\nu}{2}u^* M u, \quad \text{subject to} \quad i\omega My + Ky = Mu$$

M, K real mass and stiffness matrices.

Solution of the discrete problem

Solution using Lagrange multipliers gives

$$\begin{bmatrix} M & 0 & K - i\omega M \\ 0 & \nu M & -M \\ K + i\omega M & -M & 0 \end{bmatrix} \begin{bmatrix} y \\ u \\ p \end{bmatrix} = \begin{bmatrix} M y_d \\ 0 \\ 0 \end{bmatrix}$$

Elimination of the control ($\nu M u = M p$) yields:

$$\begin{bmatrix} M & K - i\omega\, M \\ K + i\omega\, M & -\frac{1}{\nu} M \end{bmatrix} \begin{bmatrix} y \\ p \end{bmatrix} = \begin{bmatrix} M y_d \\ 0 \end{bmatrix}$$

Zulehner, 2011 (for $\omega = 0$); Kolmbauer and Kollmann, tr2011

Solving the saddle point linear system

After simple scaling,

$$\begin{bmatrix} M & \sqrt{\nu}\,(K - i\omega\, M) \\ \sqrt{\nu}\,(K + i\omega\, M) & -M \end{bmatrix} \begin{bmatrix} y \\ \frac{1}{\sqrt{\nu}}\, p \end{bmatrix} = \begin{bmatrix} M y_d \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \mathcal{A}x = b$$

Ideal (Real) Block diagonal Preconditioner:

$$\mathcal{P} = \begin{bmatrix} M + \sqrt{\nu}\,(K + \omega\, M) & 0 \\ 0 & M + \sqrt{\nu}\,(K + \omega\, M) \end{bmatrix}$$

- Performance. Accurate estimates for the spectral intervals:

$$\operatorname{spec}(\mathcal{P}^{-1}\mathcal{A}) \in \left[-1, -\frac{1}{\sqrt{3}}\right] \cup \left[\frac{1}{\sqrt{3}}, 1\right]$$

- Robustness. Convergence of MINRES independent of the mesh, periodicity and regularization parameters (h, ω, ν)

Distributed optimal control for the time-periodic Stokes equations. I

The problem.

Find the velocity $u(x,t)$, the pressure $p(x,t)$, and the force $f(x,t)$ that minimize the cost functional

$$J(u,f) = \frac{1}{2}\int_0^T \int_\Omega |u(x,t) - u_d(x,t)|^2 \, dx \, dt + \frac{\nu}{2}\int_0^T \int_\Omega |f(x,t)|^2 \, dx \, dt$$

subject to the time-periodic Stokes problem

$$\begin{aligned}
\frac{\partial}{\partial t}\mathbf{u}(x,t) - \Delta\mathbf{u}(x,t) + \nabla p(x,t) &= \mathbf{f}(x,t) && \text{in } \Omega \times (0,T), \\
\nabla \cdot \mathbf{u}(x,t) &= 0 && \text{in } \Omega \times (0,T), \\
\mathbf{u}(x,t) &= 0 && \text{on } \partial\Omega \times (0,T), \\
\mathbf{u}(x,0) &= \mathbf{u}(x,T) && \text{in } \Omega, \\
p(x,0) &= p(x,T) && \text{in } \Omega, \\
\mathbf{f}(x,0) &= \mathbf{f}(x,T) && \text{in } \Omega.
\end{aligned}$$

Distributed optimal control for the time-periodic Stokes equations. II

Similar solution strategy (time-harmonic solution, Lagrange multipliers, scaling) leads to a familiar structure:

$$\left[\begin{array}{cc|cc} \mathbf{M} & 0 & \sqrt{\nu}(\mathbf{K} - i\omega\,\mathbf{M}) & -\sqrt{\nu}\mathbf{D}^T \\ 0 & 0 & -\sqrt{\nu}\mathbf{D} & 0 \\ \hline \sqrt{\nu}(\mathbf{K} + i\omega\,\mathbf{M}) & -\sqrt{\nu}\mathbf{D}^T & -\mathbf{M} & 0 \\ -\sqrt{\nu}\mathbf{D} & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \mathbf{u} \\ p \\ \frac{1}{\sqrt{\nu}}\mathbf{w} \\ \frac{1}{\sqrt{\nu}}r \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{u}_d \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(new setting for $\omega \neq 0$)

Optimal preconditioning technique

$$\left[\begin{array}{cc|cc} \mathbf{M} & 0 & \sqrt{\nu}(\mathbf{K}-i\omega\,\mathbf{M}) & -\sqrt{\nu}\mathbf{D}^T \\ 0 & 0 & -\sqrt{\nu}\mathbf{D} & 0 \\ \hline \sqrt{\nu}(\mathbf{K}+i\omega\,\mathbf{M}) & -\sqrt{\nu}\mathbf{D}^T & -\mathbf{M} & 0 \\ -\sqrt{\nu}\mathbf{D} & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \underline{\mathbf{u}} \\ \underline{p} \\ \frac{1}{\sqrt{\nu}}\underline{\mathbf{w}} \\ \frac{1}{\sqrt{\nu}}\underline{r} \end{bmatrix} = \begin{bmatrix} \mathbf{M}\underline{\mathbf{u}}_d \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Ideal real Block diagonal preconditioner:

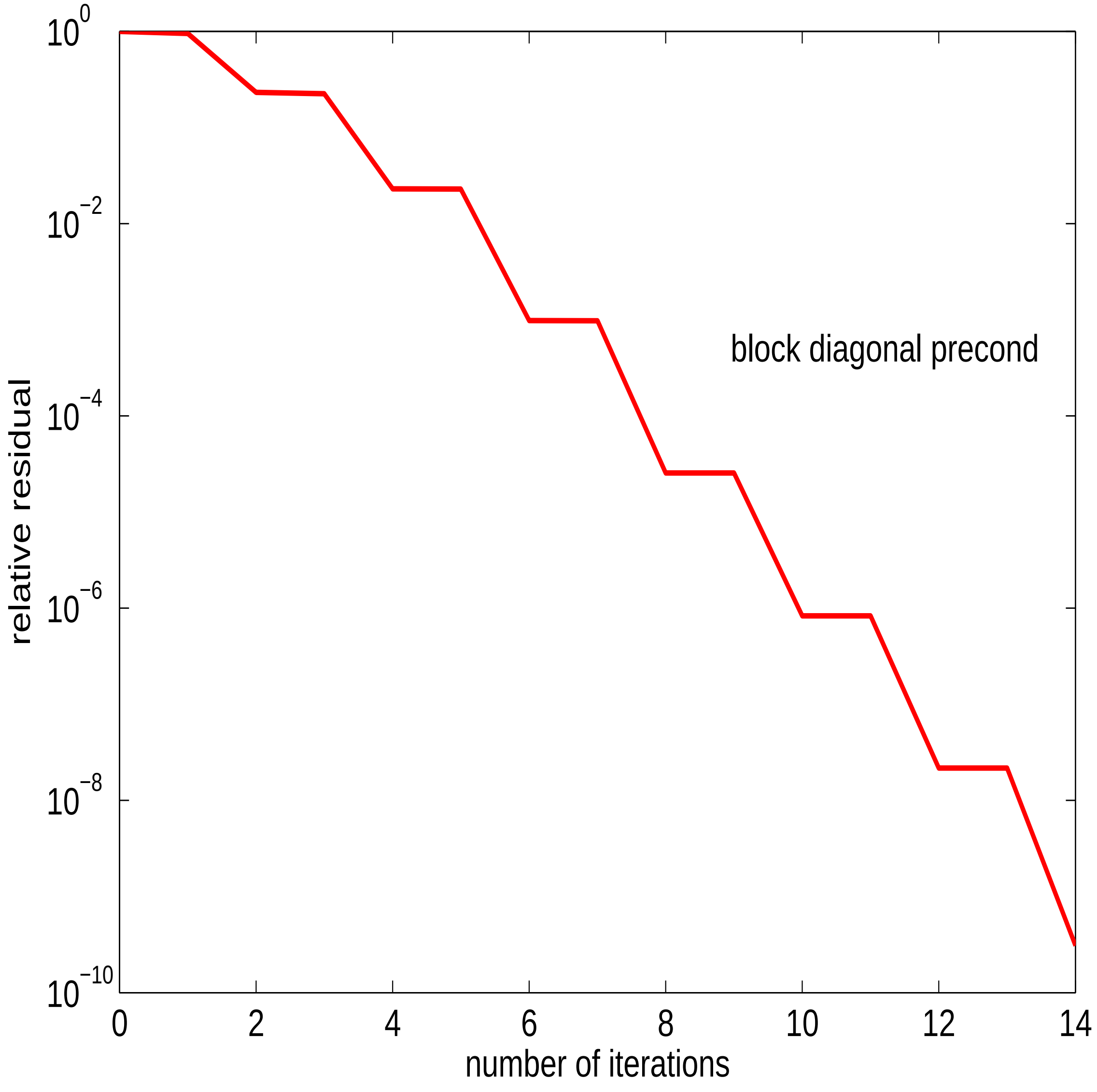
$$\mathcal{P} = \begin{bmatrix} P & & & \\ & S & & \\ & & P & \\ & & & S \end{bmatrix}, \qquad \begin{aligned} P &= M + \sqrt{\nu}(K+\omega M), \\ S &= D(M+\sqrt{\nu}(K+\omega M))^{-1}D^T \end{aligned}$$

- Performance. Accurate estimates for the spectral intervals:

$$\mathrm{spec}(\mathcal{P}^{-1}\mathcal{A}) \in \left[-\frac{1}{2}(1+\sqrt{5}), -\phi\right] \cup \left[\phi, \frac{1}{2}(1+\sqrt{5})\right], \quad \phi = 0.306...$$

- Robustness. Convergence of MINRES independent of the mesh, periodicity and regularization parameters (h, ω, ν)

Convergence history. Staircase behavior



Explanation of the Staircase behavior

Both matrices have the form:

$$\mathcal{A} = \begin{bmatrix} A & B^* \\ B & -A \end{bmatrix} \in \mathbb{C}^{2n \times 2n},$$

with:

$A \in \mathbb{R}^{n \times n}$ symmetric and semidefinite

$B \in \mathbb{C}^{n \times n}$ complex symmetric (i.e., $B = B^T$)

THEOREM: Assume that B is nonsingular. Then the eigenvalues μ of $\mathcal{A}$ come in pairs, $(\mu, -\mu)$, with $\mu \in \mathbb{R}$.

(cf. Hamiltonian matrices)

Consequence: spec($\mathcal{A}$) is symmetric with respect to the origin, and $\operatorname{spec}(\mathcal{A}) \subseteq [-b, -a] \cup [a, b]$

Symmetric spectrum. Consequences.

A classical result (e.g., Greenbaum 1997): Consider the linear system $\mathcal{A}x = r_0$

Let $\mathcal{A}$ be a Hermitian matrix, with spectrum in $[-a, -b] \cup [c, d]$, $a, b, c, d > 0$.
Assume that $|b - a| = |d - c|$.

Then after m iterations, the MINRES residual r_m satisfies

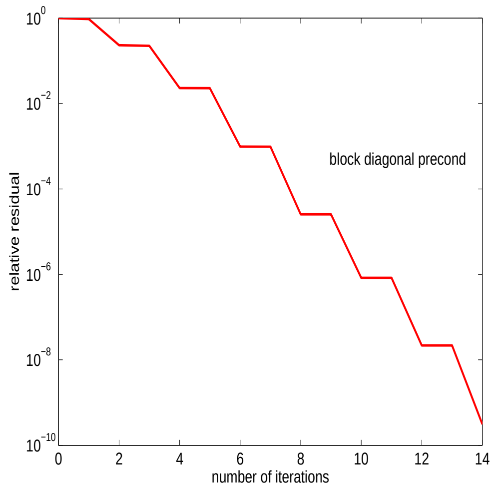
$$\frac{\|r_m\|}{\|r_0\|} \leq 2 \left(\frac{\sqrt{|ad|} - \sqrt{|bc|}}{\sqrt{|ad|} + \sqrt{|bc|}} \right)^{[m/2]}$$

For equal intervals (our case):

$$\frac{\|r_m\|}{\|r_0\|} \leq 2 \left(\frac{d/c - 1}{d/c + 1} \right)^{[m/2]}$$

$\Rightarrow$ MINRES roughly behaves like CG on a matrix having only the squared (!) positive eigenvalues

Convergence history. Quasi-stagnation...



Numerically, we do not usually see complete stagnation at even iterations...

More dramatic behavior (complete stagnation at every other iteration) for certain different settings:

Fischer, Ramage, Silvester and Wathen (BIT, 1998)

Fischer and Peherstorfer (ETNA, 2001)

Attempts to bypass quasi-stagnation

$$\mathcal{A} = \begin{bmatrix} M & \sqrt{\nu}(K - i\omega M) \\ \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}$$

An alternative (indefinite) preconditioner - work in progress:

$$\mathcal{P} = \begin{bmatrix} & M + \sqrt{\nu}(K - i\omega M) \\ M + \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}.$$

Attempts to bypass quasi-stagnation

$$\mathcal{A} = \begin{bmatrix} M & \sqrt{\nu}(K - i\omega M) \\ \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}$$

An alternative (indefinite) preconditioner - work in progress:

$$\mathcal{P} = \begin{bmatrix} & M + \sqrt{\nu}(K - i\omega M) \\ M + \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}.$$

Spectral independence wrto parameters: It holds that

$$\text{spec}(\mathcal{A}\mathcal{P}^{-1}) \subset \ [\tfrac{1}{2},1) \times [-1, 1] \in \mathbb{C}^{+}$$

The actual rectangle may be much smaller, depending on ν, ω, h

Attempts to bypass quasi-stagnation

$$\mathcal{A} = \begin{bmatrix} M & \sqrt{\nu}(K - i\omega M) \\ \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}$$

An alternative (indefinite) preconditioner - work in progress:

$$\mathcal{P} = \begin{bmatrix} & M + \sqrt{\nu}(K - i\omega M) \\ M + \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix}.$$

Spectral independence wrto parameters: It holds that

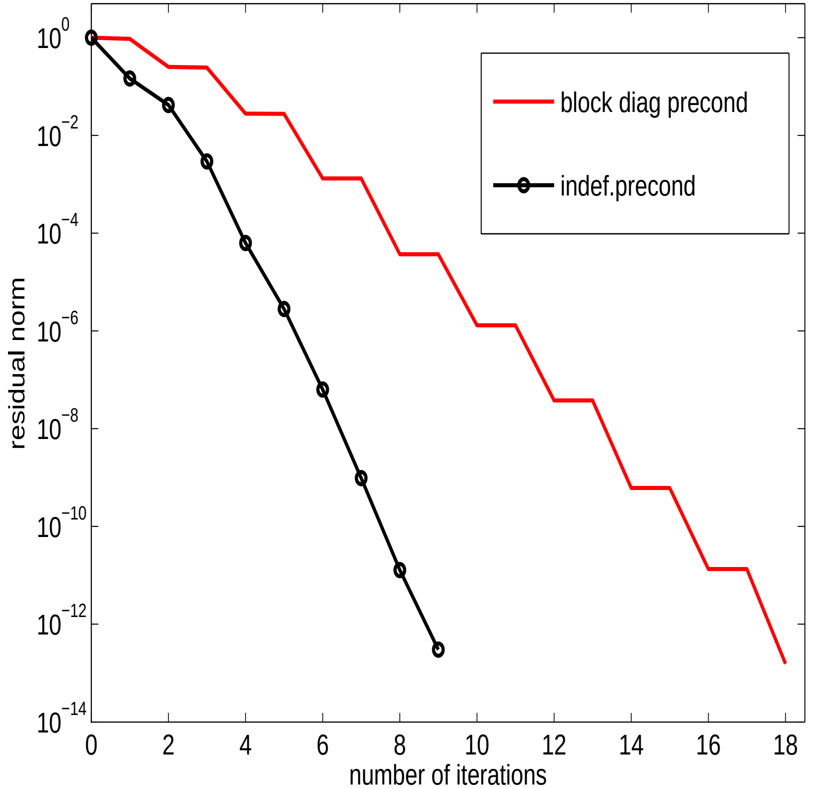
$$\mathrm{spec}(\mathcal{A}\mathcal{P}^{-1}) \subset \ [\tfrac{1}{2},1) \times [-1, 1] \in \mathbb{C}^{+}$$

The actual rectangle may be much smaller, depending on ν, ω, h

- Preconditioner not sensitive to $K \pm i\omega M$
- No results on eigenvectors

(Very) Preliminary numerical evidence. Time-periodic parabolic pb.

$\omega = 1, \nu = 10^{-2}$, $n = 1741$



MINRES vs GMRES

(Very) Preliminary numerical evidence. Time-periodic parabolic pb.

Block diagonal preconditioner: MINRES # its

$\omega \backslash \nu$	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}	10^{2}	10^{4}	10^{6}
10^{-8}	29	30	26	16	10	8	6	6
10^{-6}	29	30	26	16	10	8	6	6
10^{-4}	29	30	26	16	10	8	6	6
10^{-2}	29	30	26	16	10	8	8	8
10^{-0}	29	30	26	18	14	14	14	14
10^{2}	29	38	34	30	30	30	30	30
10^{6}	26	30	30	30	30	30	30	30
10^{8}	10	10	10	10	10	10	10	10

(Very) Preliminary numerical evidence. Time-periodic parabolic pb.

Block diagonal preconditioner: MINRES # its

$\omega\backslash\nu$	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}	10^{2}	10^{4}	10^{6}
10^{-8}	29	30	26	16	10	8	6	6
10^{-6}	29	30	26	16	10	8	6	6
10^{-4}	29	30	26	16	10	8	6	6
10^{-2}	29	30	26	16	10	8	8	8
10^{-0}	29	30	26	18	14	14	14	14
10^{2}	29	38	34	30	30	30	30	30
10^{6}	26	30	30	30	30	30	30	30
10^{8}	10	10	10	10	10	10	10	10

Block indefinite preconditioner: GMRES # its

$\omega\backslash\nu$	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}	10^{2}	10^{4}	10^{6}
10^{-8}	42	32	15	8	5	4	3	3
10^{-6}	42	32	15	8	5	4	3	3
10^{-4}	42	32	15	8	5	4	3	3
10^{-2}	42	32	15	8	5	4	3	3
10^{0}	42	32	15	8	5	4	3	3
10^{2}	42	29	11	6	4	3	3	2
10^{4}	11	5	4	3	2	2	2	2
10^{6}	3	3	2	2	2	1	1	1

Similar results with CGSTAB(ℓ)

Similar results for the Distributed optimal control for the time-periodic Stokes eqn

A side consideration

Is the complex matrix formulation needed?

$$
\begin{aligned}
\mathcal{A} &= \begin{bmatrix} M & \sqrt{\nu}(K - i\omega M) \\ \sqrt{\nu}(K + i\omega M) & -M \end{bmatrix} \\
&= \begin{bmatrix} I & 0 \\ i\omega\sqrt{\nu}I & I \end{bmatrix} \begin{bmatrix} M & \sqrt{\nu}K \\ \sqrt{\nu}K & -(1+\nu\omega^2)M \end{bmatrix} \begin{bmatrix} I & -i\omega\sqrt{\nu}I \\ 0 & I \end{bmatrix} \equiv R\mathcal{A}_r R^*
\end{aligned}
$$

(similar transformation in the Stokes case)

$$\mathcal{A}x = b \quad \Leftrightarrow \quad \mathcal{A}_r\hat{x} = \hat{b}$$

$\Rightarrow$ Convergence estimates (and expected performance) for real matrices

Final remarks

- Optimal block diagonal preconditioning emphasizes redundant information
- (Spectrally) Optimal indefinite preconditioners possible
- Other alternatives?

Reference for this talk

W. Krendl, V. Simoncini and W. Zulehner, *Stability Estimates and Structural Spectral Properties of Saddle Point Problems*, submitted, 2012.