Krylov subspace solvers and indefinite preconditioning of saddle point algebraic linear systems

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria.simoncini@unibo.it

Joint work with W. Zulehner and W. Krendl

The problem. The setting

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

- Iterative solution by means of Krylov subspace methods

- Structural properties. Focus for this talk:
  - ⋆ A symmetric positive (semi)definite or indefinite
  - ⋆ B^T tall or square nonsing
  - ⋆ 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: $y_d(x,t) = y_d(x)e^{i\omega t}, \; \omega = \frac{2\pi k}{T}$

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

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

Discrete version:

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

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, 2012

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:

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

- Robustness. Convergence of MINRES bounded independently of the mesh, frequency 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:

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

- Performance. Accurate estimates for the spectral intervals:

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

- Robustness. Convergence of MINRES bounded independently of the mesh, frequency and regularization parameters (h, ω, ν)

An example for the time-periodic Stokes constraint

$\omega \backslash \nu$	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}	10^{2}	10^{4}	10^{6}
10^{-2}	58	58	61	48	32	22	20	20
10^{0}	58	58	61	48	36	32	32	32
10^{2}	58	57	66	62	62	62	62	62
10^{4}	48	56	60	60	60	60	60	60
10^{6}	30	30	30	30	30	30	30	30
10^{8}	16	16	16	16	16	16	16	16

(Taylor-Hood pair of FE spaces (P2-P1))

final tolerance: $\texttt{tol}=10^{-12}$

Practical block diagonal preconditioning

Ideal real Block diagonal preconditioner:

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

Practical case:

$$S^{-1} \approx (1 + \omega\sqrt{\nu}) M_p^{-1} + \omega\sqrt{\nu} K_p^{-1}$$

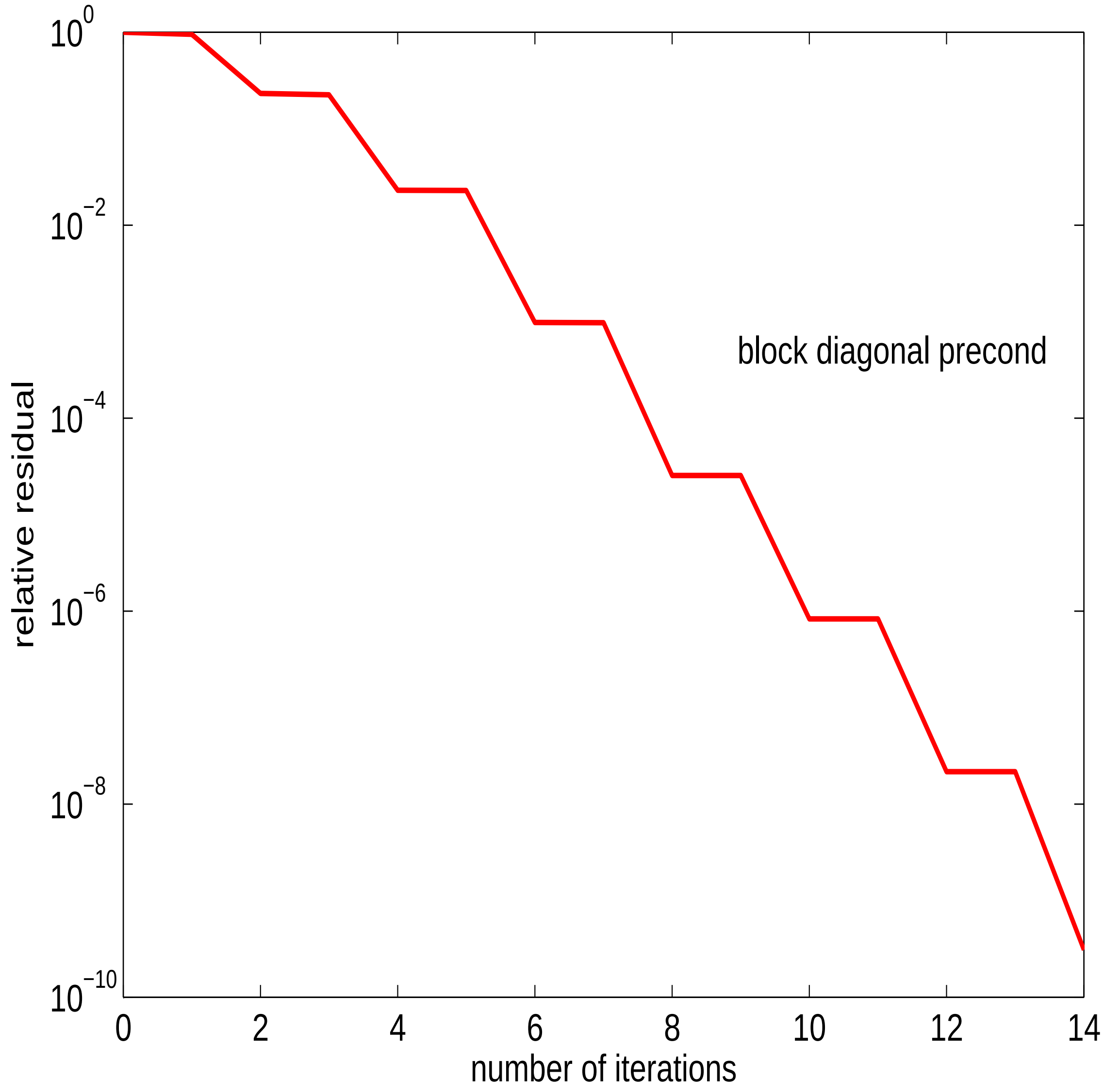
(Cahout-Charbard preconditioner)

with M_p, K_p the mass matrix and the discretized negative Laplacian in the finite element space for the pressure

$\Rightarrow$ M_p, K_p then replaced by, e.g., Multigrid versions

(Mardal, Winther, Bramble, Pasciak, Olshanskii, Peters, Reusken, ...)

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)

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 $\text{spec}(\mathcal{A}) \subseteq [-b, -a] \cup [a, b]$

$\Rightarrow$ MINRES roughly makes progress only at even iterations

Attempts to bypass quasi-stagnation. The time-periodic parabolic case

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

An alternative (indefinite) preconditioner :

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

Attempts to bypass quasi-stagnation. The time-periodic parabolic case

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

An alternative (indefinite) preconditioner :

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

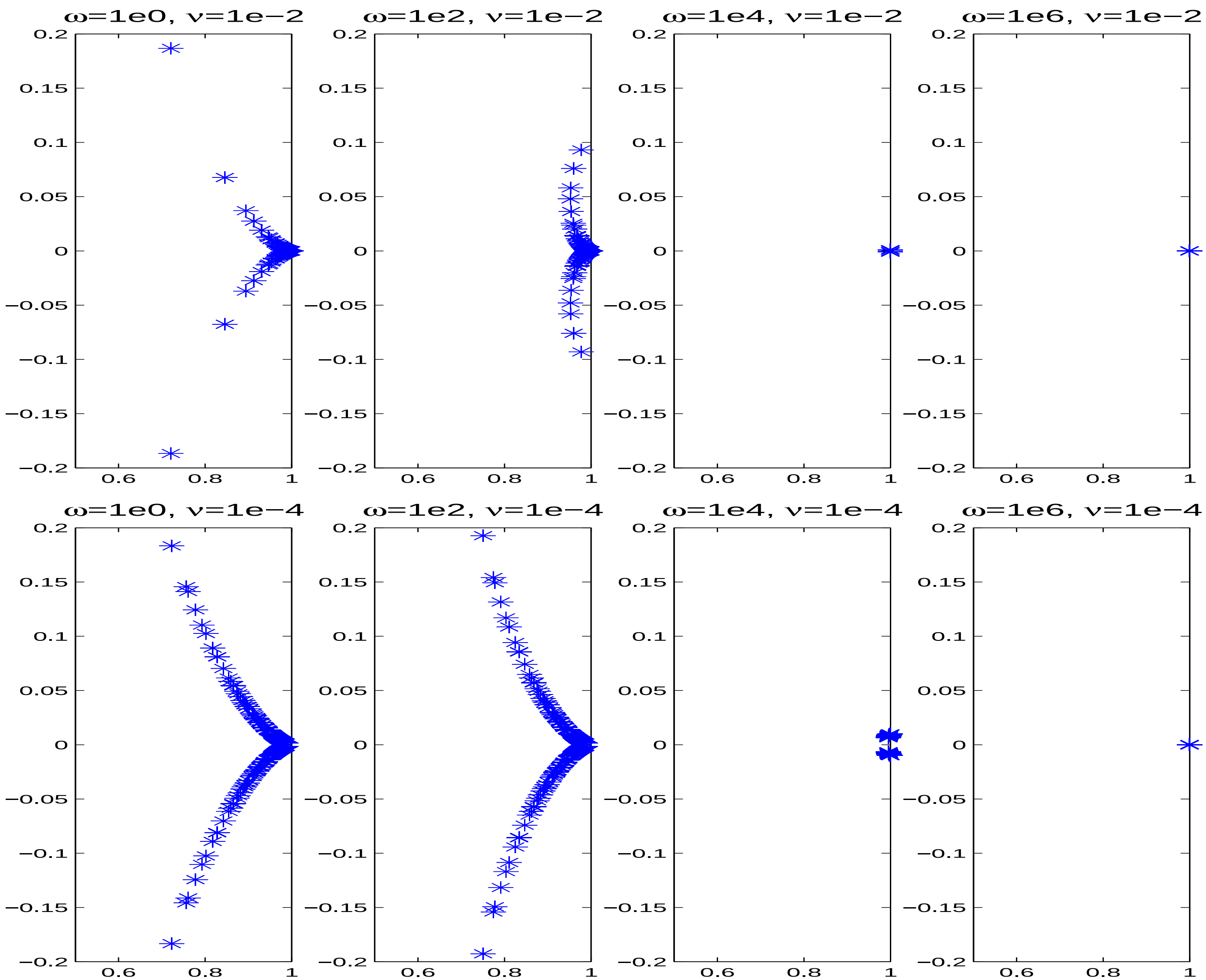
Spectral independence wrto parameters: It holds that

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

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

* Mesh independence in the spectral pattern

Spectral pattern



Spectral properties

An alternative (indefinite) preconditioner :

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

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

Eigenvectors:

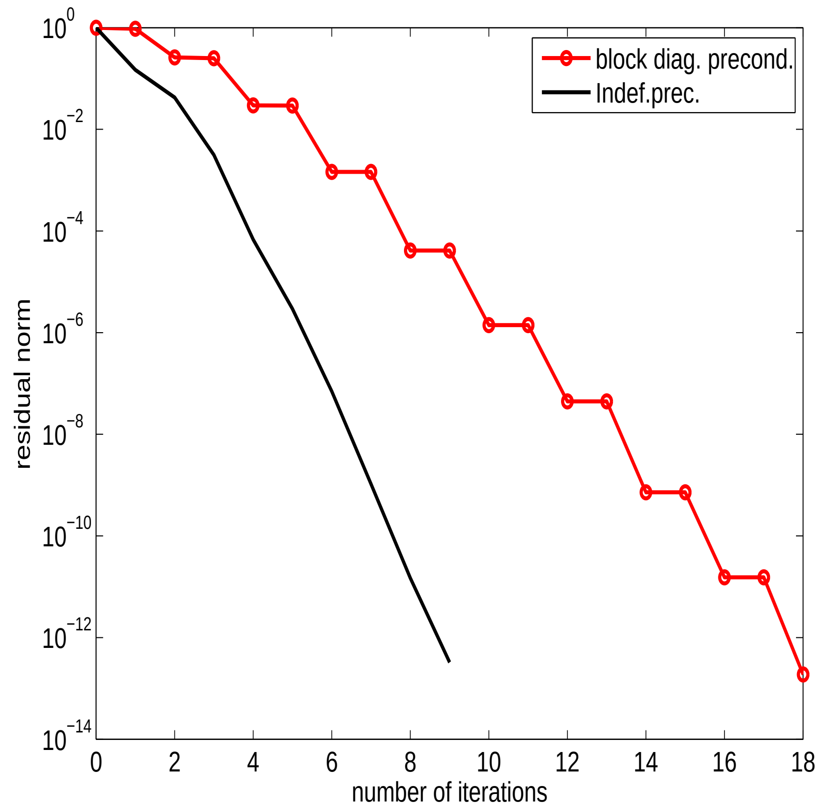
Let $M^{-1}K = X_0 \Lambda X_0^{-1}$. An eigenvector basis of $\mathcal{P}^{-1}\mathcal{A}$ is given by

$$X = \begin{bmatrix} X_0 & 0 \\ 0 & X_0 \end{bmatrix} \begin{bmatrix} I & I \\ \frac{1}{2}I + i\Gamma_+ & \frac{1}{2}I + i\Gamma_- \end{bmatrix}$$

where $\Gamma_\pm = \mathrm{diag}(\omega\sqrt{\nu} \pm \sqrt{\Re(\lambda) + \frac{3}{4} + \omega^2\nu})$

A numerical example. Time-periodic parabolic pb.

$\omega = 1, \nu = 10^{-2}$, $n = 3482(=\text{size}(K))$



MINRES vs GMRES

Number of iterations. Time-periodic parabolic pb.

Block diagonal preconditioner: MINRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	15	27	29	30	28	18	12
1	15	27	29	30	28	18	14
10^2	15	27	29	32	36	30	28
10^6	13	15	16	16	16	16	16
10^8	8	8	8	8	8	8	8

Number of iterations. Time-periodic parabolic pb.

Block diagonal preconditioner: MINRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	15	27	29	30	28	18	12
1	15	27	29	30	28	18	14
10^2	15	27	29	32	36	30	28
10^6	13	15	16	16	16	16	16
10^8	8	8	8	8	8	8	8

Block indefinite preconditioner: GMRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	16	32	42	32	17	9	6
1	16	32	42	32	17	9	6
10^2	16	32	42	32	16	8	5
10^6	12	7	4	3	3	2	2
10^8	3	3	2	2	2	2	1

Similar results with CGSTAB(ℓ)

Application to optimal control for the time-periodic Stokes equations

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

Application to optimal control for the time-periodic Stokes equations

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

Some reordering

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

Matrix of the type

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \qquad \text{with} \quad A \quad \text{Hermitian indefinite}$$

Indefinite preconditioning for the time-periodic Stokes equations

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \qquad \text{with} \quad A \quad \text{Hermitian indefinite}$$

Typically,

$$\mathcal{P}_{\text{indef}} = \begin{bmatrix} \widetilde{A} & B^T \\ B & 0 \end{bmatrix}, \qquad \widetilde{A} \approx A$$

(Bramble, Dollar, Ewing, Golub, Gould, Keller, Krzyzanowski, Lazarov, Lu, Murphy, Luksan, Pasciak, Perugia, Pestana, Rozloznik, Schilders, Schoeberl, Vassilevski, Vlcek, Wathen, Zulehner,...)

We can take:

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

Number of iterations. Time-periodic Stokes pb.

Block diagonal preconditioner: MINRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	31	52	62	60	62	48	32
1	31	52	62	60	62	48	36
10^2	31	52	62	60	68	64	62
10^6	22	32	34	34	34	34	34
10^8	16	16	16	16	16	16	16

Number of iterations. Time-periodic Stokes pb.

Block diagonal preconditioner: MINRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	31	52	62	60	62	48	32
1	31	52	62	60	62	48	36
10^2	31	52	62	60	68	64	62
10^6	22	32	34	34	34	34	34
10^8	16	16	16	16	16	16	16

Block indefinite preconditioner: GMRES # its

$\omega \backslash \nu$	10^{-12}	10^{-10}	10^{-8}	10^{-6}	10^{-4}	10^{-2}	10^{0}
0	16	31	38	28	13	6	4
1	16	31	38	28	13	6	4
10^2	16	31	38	28	13	6	4
10^6	13	7	4	3	3	2	2
10^8	4	3	2	2	2	2	2

Similar results with CGSTAB(ℓ)

A simpler indefinite strategy

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

Preconditioner

$$P_{\text{indef}} = \begin{bmatrix} \widetilde{A} & B^T \\ B & 0 \end{bmatrix}$$

with

$$\widetilde{A} \equiv \mathcal{P}_b = \begin{bmatrix} & \mathbf{M} + \sqrt{\nu}(\mathbf{K} - i\omega\mathbf{M}) \\ \mathbf{M} + \sqrt{\nu}(\mathbf{K} + i\omega\mathbf{M}) & 0 \end{bmatrix}$$

$\Rightarrow$ Clustered eigenvalues, parameters independent (proof)

$\Rightarrow$ well-conditioned eigenvectors - robust (proof)

Current issues

- Implement and analyze *inexact* preconditioners: use M_p, K_p in

$$S^{-1} \approx (1 + \omega\sqrt{\nu})M_p^{-1} + \omega\sqrt{\nu}K_p^{-1}$$

- Explore real version of the whole problem

Reference:

W. Krendl, V. Simoncini and W. Zulehner,
Stability Estimates and Structural Spectral Properties of Saddle Point Problems,
Numerische Mathematik, v. 124 (2013).