



Solution of structured algebraic linear systems in PDE-constrained optimization problem

V. Simoncini

Dipartimento di Matematica, Università di Bologna
and CIRSA, Ravenna, Italy
valeria@dm.unibo.it

The problem

$$\begin{bmatrix} A_1 & 0 & B_1^\top \\ 0 & A_2 & B_2^\top \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

Hypotheses:

- ★ $A_1 \in \mathbb{R}^{n_1 \times n_1}$ spd, $A_2 \in \mathbb{R}^{n_2 \times n_2}$ sym semidef.
- ★ B_2 square and nonsingular

A_1 cheap to solve with (e.g., diagonal)

The problem

$$\begin{bmatrix} A_1 & 0 & B_1^\top \\ 0 & A_2 & B_2^\top \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

Hypotheses:

- ★ $A_1 \in \mathbb{R}^{n_1 \times n_1}$ spd, $A_2 \in \mathbb{R}^{n_2 \times n_2}$ sym semidef.
- ★ B_2 square and nonsingular
- * A Simplified Optimal Control Problem
- * An Optimal transport problem

Reduced order problem (Schur complement)

$$\begin{bmatrix} A_1 & 0 & B_1^\top \\ 0 & A_2 & B_2^\top \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

into

$$\begin{bmatrix} A_2 & B_2^\top \\ B_2 & -B_1 A_1^{-1} B_1^\top \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_2 \\ f_3 - B_1^\top A_1^{-1} f_1 \end{bmatrix} \quad \mathcal{A}x = f$$

Both A_2 , $B_1 A_1^{-1} B_1^\top$ pos. (semi-)definite, B_2 square nonsing.

Structured Preconditioning techniques

$$\mathcal{A} = \begin{bmatrix} A_2 & B_2^\top \\ B_2 & -B_1 A_1^{-1} B_1^\top \end{bmatrix} \quad A_2 \geq 0, \quad B_2 \text{ square nonsing.}$$

- If $B_1 A_1^{-1} B_1^\top$ nonsingular:

$$\mathcal{P}_d = \begin{bmatrix} \tilde{B}_2 C^{-1} \tilde{B}_2^\top & 0 \\ 0 & C \end{bmatrix} \quad \begin{array}{l} C \approx B_1 A_1^{-1} B_1^\top \\ \tilde{B}_2 \approx B_2 \end{array}$$

Structured Preconditioning techniques

$$\mathcal{A} = \begin{bmatrix} A_2 & B_2^\top \\ B_2 & -B_1 A_1^{-1} B_1^\top \end{bmatrix} \quad A_2 \geq 0, \quad B_2 \text{ square nonsing.}$$

- If $B_1 A_1^{-1} B_1^\top$ nonsingular:

$$\mathcal{P}_d = \begin{bmatrix} \tilde{B}_2 C^{-1} \tilde{B}_2^\top & 0 \\ 0 & C \end{bmatrix} \quad \begin{array}{l} C \approx B_1 A_1^{-1} B_1^\top \\ \tilde{B}_2 \approx B_2 \end{array}$$

- If $B_1 A_1^{-1} B_1^\top$ **singular**:

$$\mathcal{P}_{ad} = \begin{bmatrix} D & 0 \\ 0 & C(D) \end{bmatrix}, \quad \begin{array}{l} D \approx A_2, \quad \mathbf{D} > \mathbf{0} \\ C(D) \approx B_1 A_1^{-1} B_1^\top + B_2 D^{-1} B_2^\top \end{array}$$

Constraint (indefinite) Preconditioning

Regardless of singularity of $B_1 A_1^{-1} B_1$:

$$\mathcal{P}_{indef} = \begin{bmatrix} 0 & \tilde{B}_2^\top \\ \tilde{B}_2 & -C \end{bmatrix}, \quad \tilde{B}_2 \approx B_2$$

for properly chosen C

Note: \tilde{B}_2 square and nonsing.

\Rightarrow solve for the two nondiag. blocks in sequence

Spectral properties

Under appropriate hypotheses on $B_i, A_i, i = 1, 2$ and properly choosing \tilde{B}_2 :

- (real indefinite) $\text{spec}(\mathcal{A}P_d^{-1})$ is bounded independently of the problem dimension
- For $\tilde{B}_2 = B_2$: (Positive real) $\text{spec}(\mathcal{A}P_{indef}^{-1})$ is bounded independently of the problem dimension
- For $\tilde{B}_2 \neq B_2$ approp. chosen: $\text{spec}(\mathcal{A}P_{indef}^{-1})$ is bounded independently of the problem dimension

Constrained Optimal Control Problem

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$.

Given \hat{u} (*desired state*) defined in $\hat{\Omega} \subseteq \Omega$, find u :

$$\min_{u, f} \frac{1}{2} \|u - \hat{u}\|_{L_2(\hat{\Omega})}^2 + \beta \|f\|_{L_2(\Omega)}^2$$

$$\text{s.t.} \quad -\nabla^2 u = f \quad \text{in } \Omega$$

$$u = \hat{u} \quad \text{on } \partial\Omega$$

- β : regularization parameter
- Typical extra condition: box constraints for f
- Alternative PDE constraints
- Alternative PDE boundary conditions

Thanks to Sue H. Thorne...

Constrained Optimal Control Problem. Discretization.

Discrete cost functional

$$\min_{u_h, f_h} \frac{1}{2} \|u_h - \hat{u}\|_2^2 + \beta \|f_h\|_2^2 = \min_{\mathbf{u}, \mathbf{f}} \frac{1}{2} \mathbf{u}^\top \bar{M} \mathbf{u} - \mathbf{u}^\top \mathbf{u} + \alpha + \beta \mathbf{f}^\top M \mathbf{f}$$

with $\alpha = \|\hat{u}\|_2^2$, M mass matrix, \bar{M} portion of mass matrix

Constraint:

$$-\nabla^2 u = f \quad \text{in } \Omega \Rightarrow \mathbf{K} \mathbf{u} = M \mathbf{f} + \mathbf{d}$$

K stiffness matrix (discrete Laplacian)

Solution via Lagrangian:

$$\mathcal{L}(\mathbf{f}, \mathbf{u}, \lambda) = \frac{1}{2} \mathbf{u}^\top \bar{M} \mathbf{u} - \mathbf{u}^\top \mathbf{b} + \alpha + \beta \mathbf{f}^\top M \mathbf{f} + \lambda^\top (\mathbf{K} \mathbf{u} - M \mathbf{f} - \mathbf{d})$$

The resulting saddle point problem

$$\mathcal{L}(\mathbf{f}, \mathbf{u}, \lambda) = \frac{1}{2} \mathbf{u}^\top \bar{M} \mathbf{u} - \mathbf{u}^\top \mathbf{b} + \alpha + \beta \mathbf{f}^\top M \mathbf{f} + \lambda^\top (K \mathbf{u} - M \mathbf{f} - \mathbf{d})$$

First order condition on Lagrangian yields:

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^\top \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \\ \mathbf{d} \end{bmatrix} \Rightarrow \begin{bmatrix} \bar{M} & K^\top \\ K & -\frac{1}{2\beta} M \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

★ $2\beta \mathbf{f} = \lambda$

The resulting saddle point problem

$$\mathcal{L}(\mathbf{f}, \mathbf{u}, \lambda) = \frac{1}{2} \mathbf{u}^\top \bar{M} \mathbf{u} - \mathbf{u}^\top \mathbf{b} + \alpha + \beta \mathbf{f}^\top M \mathbf{f} + \lambda^\top (K \mathbf{u} - M \mathbf{f} - \mathbf{d})$$

First order condition on Lagrangian yields:

$$\begin{bmatrix} 2\beta M & 0 & -M \\ 0 & \bar{M} & K^\top \\ -M & K & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \\ \mathbf{d} \end{bmatrix} \Rightarrow \begin{bmatrix} \bar{M} & K^\top \\ K & -\frac{1}{2\beta} M \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

- ★ $2\beta \mathbf{f} = \lambda$
- ★ \bar{M} could be singular (it depends on where \hat{u} is defined)
- ★ M nonsingular

Solution strategies

$$\mathcal{A} = \begin{bmatrix} \bar{M} & K^\top \\ K & -\frac{1}{2\beta}M \end{bmatrix}$$

★ Ideal case:

$$P_d = \begin{bmatrix} KD^{-1}K^\top & 0 \\ 0 & D \end{bmatrix}, \quad D = \frac{1}{2\beta}M$$

Solution strategies

$$\mathcal{A} = \begin{bmatrix} \bar{M} & K^\top \\ K & -\frac{1}{2\beta}M \end{bmatrix}$$

★ Ideal case:

$$P_d = \begin{bmatrix} KD^{-1}K^\top & 0 \\ 0 & D \end{bmatrix}, \quad D = \frac{1}{2\beta}M$$

Practical case:

$\tilde{K} \approx K$, e.g., (algebraic) multigrid method

$\tilde{D} \approx D$, e.g., $\tilde{D} = \text{diag}(\frac{1}{2\beta}M)$

Solution strategies

$$\mathcal{A} = \begin{bmatrix} \bar{M} & K^\top \\ K & -\frac{1}{2\beta}M \end{bmatrix}$$

★ Ideal case:

$$P_d = \begin{bmatrix} KD^{-1}K^\top & 0 \\ 0 & D \end{bmatrix}, \quad D = \frac{1}{2\beta}M$$

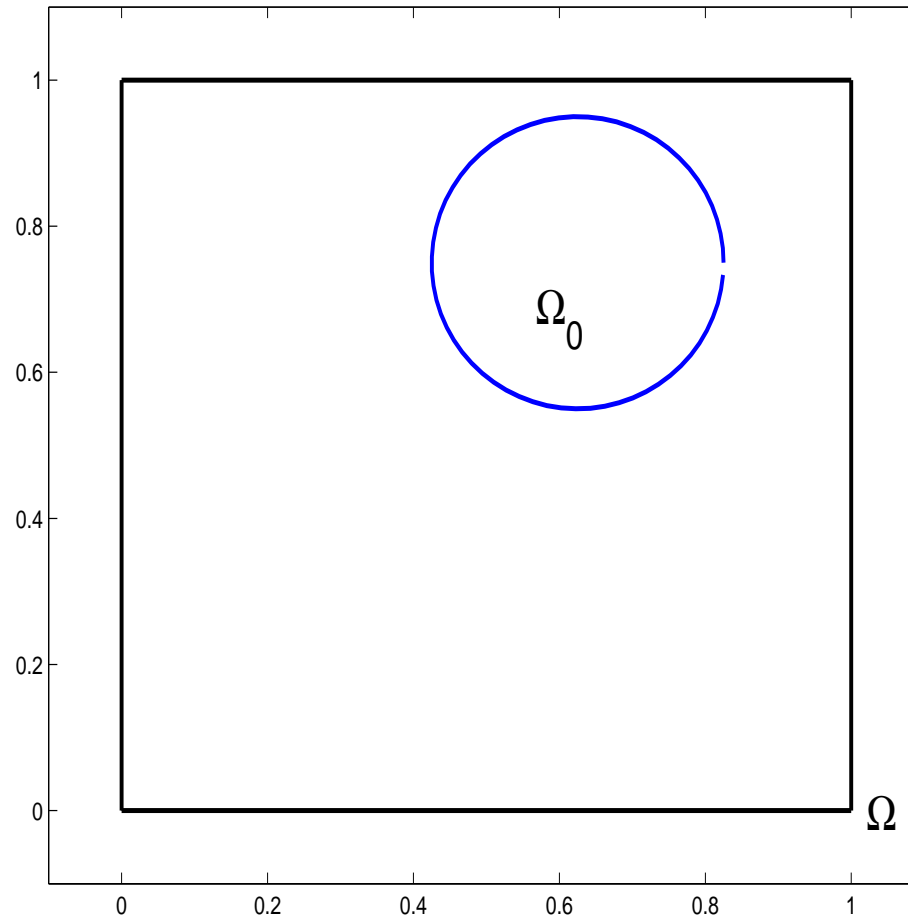
Practical case:

$\tilde{K} \approx K$, e.g., (algebraic) multigrid method

$\tilde{D} \approx D$, e.g., $\tilde{D} = \text{diag}(\frac{1}{2\beta}M)$

★ Analogously: $P_{indef} = \begin{bmatrix} 0 & \tilde{K} \\ \tilde{K} & \frac{1}{2\beta}M \end{bmatrix}$

Numerical results: 2D and 3D



2D: $\hat{u}(x, y) = 2$ in Ω_0 and $\hat{u}(x, y) = 0$ on $\partial\Omega$ (undefined elsewhere)

Data thanks to Sue H. Thorne, RAL, UK

Numerical results. Minimal Residual method. No of iterations.

$\text{tol}=10^{-8}$

\bar{M} singular, $K \in \mathbb{R}^{n \times n}$ sym

MINRES

2D:

β	n	GMRES w/ P_{indef}		MINRES w/ P_d	
		its	CPU Time	its	CPU Time
10^{-2}	961	3	0.15	31	0.18
	3969	3	0.17	28	0.37
	16129	3	0.29	25	1.22
	65025	3	0.88	25	4.28
	261121	4	4.54	27	21.31
10^{-5}	961	10	0.19	35	0.20
	3969	10	0.25	32	0.41
	16129	10	0.65	29	1.19
	65025	10	2.21	27	4.52
	261121	10	9.80	27	19.74

Numerical results. Minimal Residual method. No of iterations.

$$\text{tol}=10^{-8}$$

\bar{M} singular, $K \in \mathbb{R}^{n \times n}$ sym

MINRES

3D:

β	n	GMRES w/ P_{indef}		MINRES w/ P_d	
		its	CPU Time	its	CPU Time
10^{-2}	343	3	0.16	59	0.20
	3375	3	0.20	63	1.09
	29791	3	0.94	63	11.20
	250047	3	9.96	63	129.48
10^{-5}	343	8	0.13	57	0.20
	3375	9	0.30	61	1.03
	29791	9	2.04	64	11.28
	250047	9	23.02	65	132.36

Monge-Kantorovich mass transfer problem

Pb: Given two density functions u_0 and u_T on the set Ω , find an “optimal” mapping from u_0 to u_T

Formulation (time in $[0, T]$):

$$\min_{u, m} \frac{1}{2} \|u(T, \mathbf{x}) - u_T(\mathbf{x})\|^2 + \frac{1}{2} \alpha T \int_{\Omega} \int_0^T u \|m\|^2 dt d\mathbf{x}$$
$$s.t. \quad u_t + \nabla \cdot (um) = 0, \quad u(0, \mathbf{x}) = u_0$$

$u(t, \mathbf{x})$: density field $m(t, \mathbf{x})$: velocity field

A preconditioning technique for a class of PDE-constrained optimization problems

M. BENZI, E. HABER AND L. TARALLI, Adv. in Comput. Math. '10

Monge-Kantorovich mass transfer problem

Time and space discretization + Gauss-Newton approximation on the Lagrangian

Sequence of “Newton step-depending” linear systems:

$$\begin{bmatrix} Q^\top Q & 0 & B_1^\top \\ 0 & L & B_2^\top \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} \tilde{u}_k \\ \tilde{m}_k \\ \tilde{p}_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

with $Q^\top Q$ diagonal and highly singular, $L > 0$ diagonal

B_2 rank deficient

Reduced order problem

$$\begin{bmatrix} Q^\top Q & 0 & B_1^\top \\ 0 & L & B_2^\top \\ B_1 & B_2 & 0 \end{bmatrix} \begin{bmatrix} \tilde{u}_k \\ \tilde{m}_k \\ \tilde{p}_k \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

into

$$\begin{bmatrix} Q^\top Q & B_1^\top \\ B_1 & -B_2 L^{-1} B_2^\top \end{bmatrix} \begin{bmatrix} \tilde{u}_k \\ \tilde{p}_k \end{bmatrix} = \begin{bmatrix} b_2 \\ \tilde{b}_3 \end{bmatrix}$$

Both $Q^\top Q$, $B_2 L^{-1} B_2^\top$ pos. semi-definite, B_1 square nonsing.

Augmented Block Diagonal Preconditioning

$$\mathcal{A} = \begin{bmatrix} Q^\top Q & B_1^\top \\ B_1 & -B_2 L^{-1} B_2^\top \end{bmatrix}$$

Both diagonal blocks are singular. Augmentation:

$$\mathcal{P}_{ad} = \begin{bmatrix} D & 0 \\ 0 & C(D) \end{bmatrix}, \quad \begin{aligned} D &> 0 \\ C(D) &\approx B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top, \end{aligned}$$

Exact preconditioner: Spectral bounds

$$\mathcal{A} = \begin{bmatrix} Q^\top Q & B_1^\top \\ B_1 & -B_2 L^{-1} B_2^\top \end{bmatrix} = \left[\begin{array}{cc|c} O & & B_{11}^\top \\ & \Omega & B_{12}^\top \\ \hline B_{11} & B_{12} & -B_2 L^{-1} B_2^\top \end{array} \right]$$

Exact preconditioner: Spectral bounds

$$\mathcal{A} = \begin{bmatrix} Q^\top Q & B_1^\top \\ B_1 & -B_2 L^{-1} B_2^\top \end{bmatrix} = \left[\begin{array}{cc|c} O & & B_{11}^\top \\ & \Omega & B_{12}^\top \\ \hline B_{11} & B_{12} & -B_2 L^{-1} B_2^\top \end{array} \right]$$

$$\mathcal{P}_{ad} = \begin{bmatrix} D & O \\ O & C \end{bmatrix} \quad D = \begin{bmatrix} \gamma I & \\ & \Omega \end{bmatrix} \quad \gamma = \|L\|$$
$$C = B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top$$

Exact preconditioner: Spectral bounds

$$\mathcal{A} = \begin{bmatrix} Q^\top Q & B_1^\top \\ B_1 & -B_2 L^{-1} B_2^\top \end{bmatrix} = \left[\begin{array}{cc|c} O & & B_{11}^\top \\ & \Omega & B_{12}^\top \\ \hline B_{11} & B_{12} & -B_2 L^{-1} B_2^\top \end{array} \right]$$

$$\mathcal{P}_{ad} = \begin{bmatrix} D & O \\ O & C \end{bmatrix} \quad D = \begin{bmatrix} \gamma I & \\ & \Omega \end{bmatrix} \quad \gamma = \|L\|$$

$$C = B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top$$

$$\text{spec}(\mathcal{A}\mathcal{P}_{ad}^{-1}) \subset \left[-1, \frac{1}{2}(1 - \sqrt{5}) \right] \cup \left[\theta, \frac{1}{2}(1 + \sqrt{5}) \right]$$

$$\text{where } \theta = \frac{-1 + \sqrt{1 + 4s}}{2}, \quad s = \lambda_{\min} \left(\frac{1}{\gamma} B_{12}^\top C^{-1} B_{12} \right) \leq 1$$

The bound is sharp

$$\text{spec}(\mathcal{AP}_{ad}^{-1}) \subset \left[-1, \frac{1}{2}(1 - \sqrt{5})\right] \cup \left[\theta, \frac{1}{2}(1 + \sqrt{5})\right] \quad \theta = \frac{-1 + \sqrt{1 + 4s}}{2}$$

$$\mathcal{A} = \left[\begin{array}{cc|cc} 0 & 0 & -1 & \delta \\ 0 & 2 & 10 & 1 \\ \hline -1 & 10 & -20 & 0 \\ \delta & 1 & 0 & 0 \end{array} \right] \quad \mathcal{P} = \left[\begin{array}{cc|cc} 10 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 200 + \delta^2 & 0 \\ 0 & 0 & 0 & \frac{25}{2} \end{array} \right]$$

The bound is sharp

$$\text{spec}(\mathcal{A}\mathcal{P}_{ad}^{-1}) \subset \left[-1, \frac{1}{2}(1 - \sqrt{5})\right] \cup \left[\theta, \frac{1}{2}(1 + \sqrt{5})\right] \quad \theta = \frac{-1 + \sqrt{1 + 4s}}{2}$$

$$\mathcal{A} = \left[\begin{array}{cc|cc} 0 & 0 & -1 & \delta \\ 0 & 2 & 10 & 1 \\ \hline -1 & 10 & -20 & 0 \\ \delta & 1 & 0 & 0 \end{array} \right] \quad \mathcal{P} = \left[\begin{array}{cc|cc} 10 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 200 + \delta^2 & 0 \\ 0 & 0 & 0 & \frac{25}{2} \end{array} \right]$$

δ		spectral values			
10^{-2}	$\text{spec}(SP_{ad}^{-1})$	-1.0000	-0.61805	0.00603	1.6180
	$\mathcal{I}^-, \mathcal{I}^+$	-1.0000	-0.61803	0.00599	1.6180
10^0	$\text{spec}(SP_{ad}^{-1})$	-1.0000	-0.66646	0.42423	1.5774
	$\mathcal{I}^-, \mathcal{I}^+$	-1.0000	-0.61803	0.33426	1.6180
10^2	$\text{spec}(SP_{ad}^{-1})$	-1.0000	-0.70458	0.99980	1.4194
	$\mathcal{I}^-, \mathcal{I}^+$	-1.0000	-0.61803	0.61797	1.6180

Exact preconditioner. Numerical results

$$\mathcal{P}_{ad} = \begin{bmatrix} D & O \\ O & C \end{bmatrix} \quad D = \begin{bmatrix} \gamma I & \\ & \Omega \end{bmatrix} \quad \gamma = \|L\|$$

$$C = B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top$$

n_x	n_t	n	#	CPU time	n_x	n_t	n	#	CPU time
10	10	1000	12	0.19	10	10	1000	12	0.19
15	15	3375	13	1.28	20	10	4000	16	3.44
20	20	8000	14	8.17	30	10	9000	21	34.16
25	25	15625	15	33.31	40	20	32000	21	376.54
30	30	27000	16	109.25	50	20	50000	-	-
35	35	42875	16	309.97					
40	40	64000	-	-					

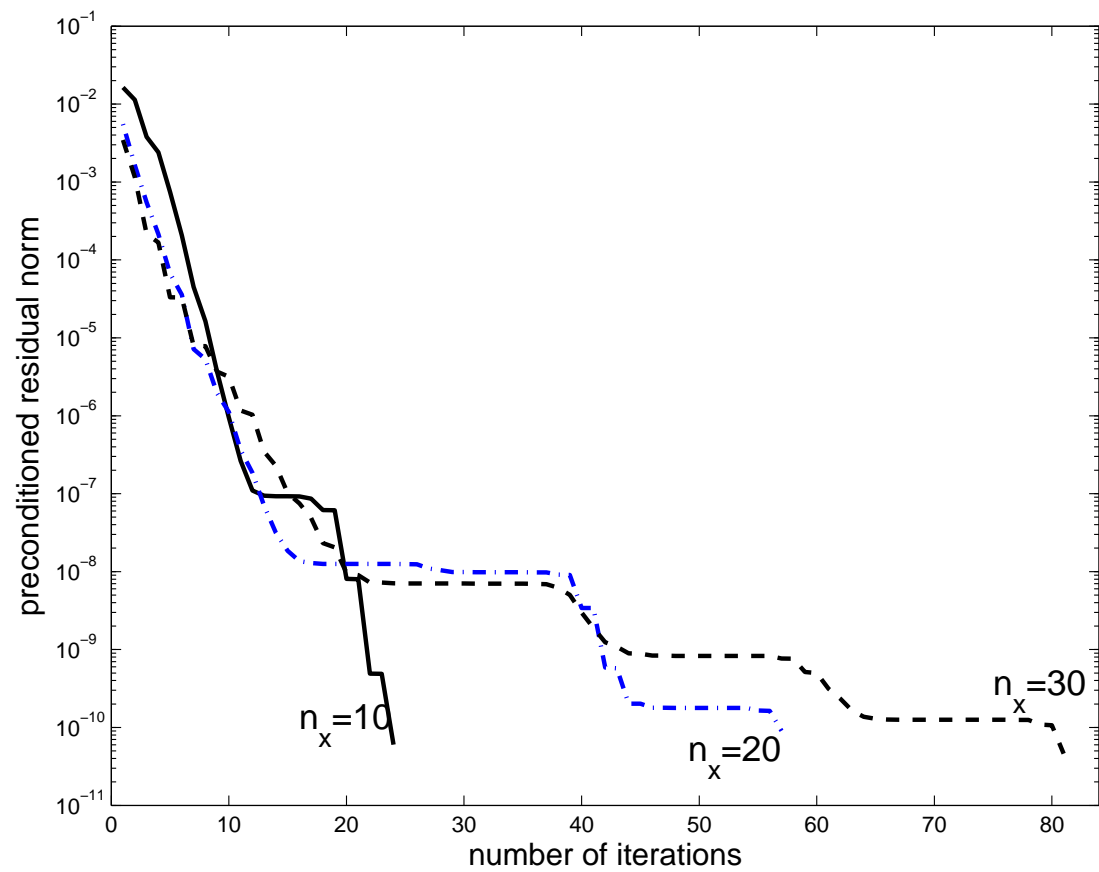
Practical preconditioner. Numerical results

$$\mathcal{P}_{ad} = \begin{bmatrix} D & O \\ O & C \end{bmatrix} \quad D = \begin{bmatrix} \gamma I & \\ & \Omega \end{bmatrix} \quad \gamma = \|L\|$$

$$C = \text{hsl_mi20}(B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top)$$

n_x	n_t	n	\mathcal{P}_{ad}		\mathcal{P}_{ad} w/mi20	
			# its	time	# its	time
10	10	1000	12	0.19	24	0.29
15	15	3375	13	1.28	52	1.43
20	20	8000	14	8.17	57	4.64
25	25	15625	15	33.31	85	14.64
30	30	27000	16	109.25	81	28.37
35	35	42875	16	309.97	122	65.82
40	40	64000	-		90	88.63

Numerical results with practical \mathcal{P}_{ad}



Approximate smallest eigenvalues

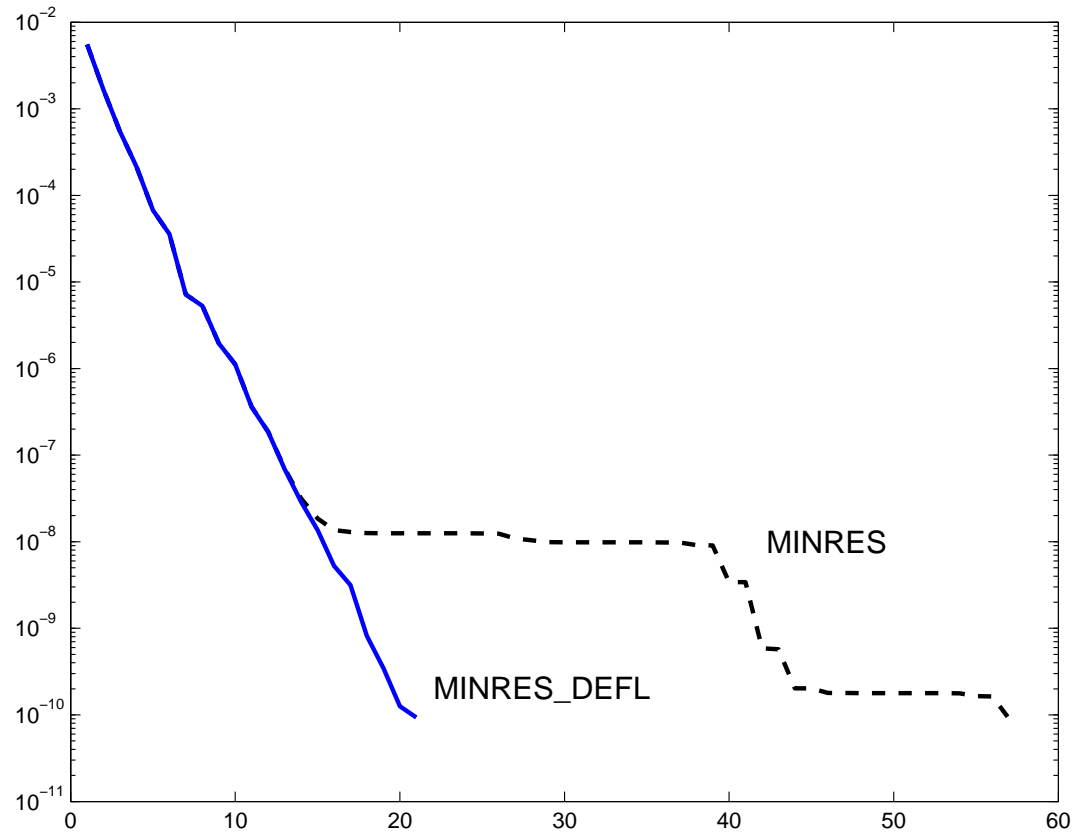
$$C = \text{hsl_mi20}(B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top)$$

Eigs of “preconditioned augmented Schur complement”

$$\text{spec}((B_2 L^{-1} B_2^\top + B_1 D^{-1} B_1^\top) C^{-1})$$

approx λ_i	$n_x = 10$	$n_x = 20$	$n_x = 30$
i=1	9.4632e-04	1.7676e-05	1.3247e-05
i=2	9.4999e-04	1.8274e-05	1.3937e-05
i=3	8.5302e-01	1.5449e-04	4.6116e-05
i=4	8.5400e-01	1.5484e-04	4.6200e-05
i=5	8.5722e-01	6.2973e-01	3.7707e-01
i=6	8.5752e-01	6.3079e-01	3.7791e-01
i=7	8.6410e-01	6.6089e-01	3.8408e-01
i=8	8.6825e-01	6.6091e-01	3.8409e-01
i=9	8.7875e-01	6.6557e-01	4.0425e-01
i=10	8.8097e-01	6.6674e-01	4.0514e-01

Augmented/Deflated MINRES



Current approximation:

50 its of Arnoldi method on precon'd Schur complement
(at first Newton step)

Complete timings

n	\mathcal{P}_{ad}		\mathcal{P}_{ad} w/mi20		\mathcal{P}_{ad} w/mi20+DEFL		
	# its	time	# its	time	# its	time	
1000	12	0.19	24	0.29	17	0.26	+ 0.52, 2 eigs
3375	13	1.28	52	1.43	25	0.80	+ 1.53, 5 eigs
8000	14	8.17	57	4.64	21	1.95	+ 4.51, 4 eigs
15625	15	33.31	85	14.64	36	6.68	+ 9.57, 5 eigs
27000	16	109.25	81	28.37	28	10.61	+19.03, 4 eigs
42875	16	309.97	122	65.82	47	27.33	+29.82, 5 eigs
64000	-		90	88.63	37	37.82	+52.23, 4 eigs

Conclusions

For the analyzed problems:

- First step: order reduction
- Standard or Augmented block diagonal preconditioning
- Fixes in case of stagnation
- Ad-hoc Algebraic Multigrid preconditioners?