



Matrix functions and large-scale matrix equations: Applications

V. Simoncini

Dipartimento di Matematica, Università di Bologna

`valeria@dm.unibo.it`

Appendix of talk I: Krylov vs. rational approximation

$f \rightarrow \mathcal{R}_\nu$

CG approximation of linear systems:

$$\begin{aligned}\mathcal{R}_\nu(A)v &= \omega_0 v + \sum_{j=1}^{\nu} \omega_j (A - \xi_j I)^{-1} v \\ &\approx \omega_0 v + \sum_{j=1}^{\nu} \omega_j V_m (H_m - \xi_j I)^{-1} e_1 \\ &= V_m \left(\omega_0 e_1 + \sum_{j=1}^{\nu} \omega_j (H_m - \xi_j I)^{-1} e_1 \right) \equiv V_m \mathcal{R}_\nu(H_m) e_1\end{aligned}$$

Appendix of talk I: Krylov vs. rational approximation

$f \rightarrow \mathcal{R}_\nu$

CG approximation of linear systems:

$$\begin{aligned}\mathcal{R}_\nu(A)v &= \omega_0 v + \sum_{j=1}^{\nu} \omega_j (A - \xi_j I)^{-1} v \\ &\approx \omega_0 v + \sum_{j=1}^{\nu} \omega_j V_m (H_m - \xi_j I)^{-1} e_1 \\ &= V_m \left(\omega_0 e_1 + \sum_{j=1}^{\nu} \omega_j (H_m - \xi_j I)^{-1} e_1 \right) \equiv V_m \mathcal{R}_\nu(H_m) e_1\end{aligned}$$

Krylov approximation: $x \approx V_m f(H_m) e_1$

Projection methods for matrix function - related application problems

- The Lyapunov matrix equation
- The matrix Transfer function
- An ill-posed PDE problem

Solving the Lyapunov equation. The problem

Approximate X in:

$$AX + XA^\top + BB^\top = 0$$

$$A \in \mathbb{R}^{n \times n} \text{ pos.real} \quad B \in \mathbb{R}^{n \times s}, \quad s \geq 1$$

Solving the Lyapunov equation. The problem

Approximate X in:

$$AX + XA^\top + BB^\top = 0$$

$$A \in \mathbb{R}^{n \times n} \text{ pos.real} \quad B \in \mathbb{R}^{n \times s}, \quad s \geq 1$$

Applications: signal processing, system and control theory

Time-invariant linear system:

$$\mathbf{x}'(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \quad \mathbf{x}(0) = x_0$$

Analytic solution:

$$X = \int_0^\infty e^{-tA} B B^\top e^{-tA^\top} dt = \int_0^\infty x x^\top dt \quad \text{with } x = \exp(-tA)B$$

see, e.g., Antoulas '05, Benner '06

Standard Krylov subspace projection

$$X \approx X_m \quad X_m \in \mathcal{K}$$

Galerkin condition: $R := AX_m + X_m A^\top + bb^\top \perp \mathcal{K}$

$$V_m^\top R V_m = 0 \quad \mathcal{K} = \text{range}(V_m)$$

Standard Krylov subspace projection

$$X \approx X_m \quad X_m \in \mathcal{K}$$

Galerkin condition: $R := AX_m + X_m A^\top + bb^\top \perp \mathcal{K}$

$$V_m^\top R V_m = 0 \quad \mathcal{K} = \text{range}(V_m)$$

Assume $V_m^\top V_m = I_m$ and let $X_m := V_m Y_m V_m^\top$.

Projected Lyapunov equation:

$$(V_m^\top A V_m) Y_m + Y_m (V_m^\top A^\top V_m) + V_m^\top b b^\top V_m = 0$$

$$\Updownarrow$$

$$T_m Y_m + Y_m T_m^\top + e_1 e_1^\top = 0$$

with $b = V_m e_1$ (Saad '90, for $\mathcal{K} = \mathcal{K}_m(A, b)$; Jaimoukha & Kasenally '94)

Standard Krylov projection. In quest of error bounds

$$AX + XA^\top + BB^\top = 0, \quad X \approx X_m \in K_m(A, B)$$

$$\|X - X_m\| \leq ??$$

Standard Krylov projection. In quest of error bounds

$$AX + XA^\top + BB^\top = 0, \quad X \approx X_m \in K_m(A, B)$$

$$\|X - X_m\| \leq ??$$

(Simoncini & Druskin '09.) Analytic solution:

$$X = \int_0^\infty e^{-tA} BB^\top e^{-tA^\top} dt = \int_0^\infty xx^\top dt$$

with $x = \exp(-tA)B$, $B = b$, $\|b\| = 1$

Let $\alpha_{\min} = \lambda_{\min}((A + A^\top)/2) > 0$. Then

$$\|x\| \leq \exp(-t\alpha_{\min})\|B\|$$

First (key) step

Krylov subspace proj.: $X_m = V_m Y_m V_m^\top$, $\text{range}(V_m) = K_m(A, b)$

$$T_m Y_m + Y_m T_m^\top + e_1 e_1^\top = 0$$

Clearly,

$$X_m = V_m \left(\int_0^\infty e^{-tT_m} e_1 e_1^\top e^{-tT_m^\top} dt \right) V_m^\top = \int_0^\infty x_m x_m^\top dt$$

with $x_m = V_m \exp(-tT_m) e_1$, and $\|x_m\| \leq \exp(-t\alpha_{\min})$.

First (key) step

Krylov subspace proj.: $X_m = V_m Y_m V_m^\top$, $\text{range}(V_m) = K_m(A, b)$

$$T_m Y_m + Y_m T_m^\top + e_1 e_1^\top = 0$$

Clearly,

$$X_m = V_m \left(\int_0^\infty e^{-tT_m} e_1 e_1^\top e^{-tT_m^\top} dt \right) V_m^\top = \int_0^\infty x_m x_m^\top dt$$

with $x_m = V_m \exp(-tT_m) e_1$, and $\|x_m\| \leq \exp(-t\alpha_{\min})$.

II step: $\|X - X_m\| = \left\| \int_0^\infty (xx^\top - x_m x_m^\top) dt \right\|$, so that

$$\|X - X_m\| \leq \int_0^\infty \|xx^\top - x_m x_m^\top\| dt \leq 2 \int_0^\infty e^{-t\alpha_{\min}} \|x - x_m\| dt$$

The case of A symmetric

A symmetric $\Rightarrow \alpha_{\min} = \lambda_{\min}(A)$

Let $0 < \hat{\lambda}_{\min} \leq \dots \leq \hat{\lambda}_{\max}$ eigs of $A + \lambda_{\min}I$, $\hat{\kappa} := \frac{\hat{\lambda}_{\max}}{\hat{\lambda}_{\min}}$

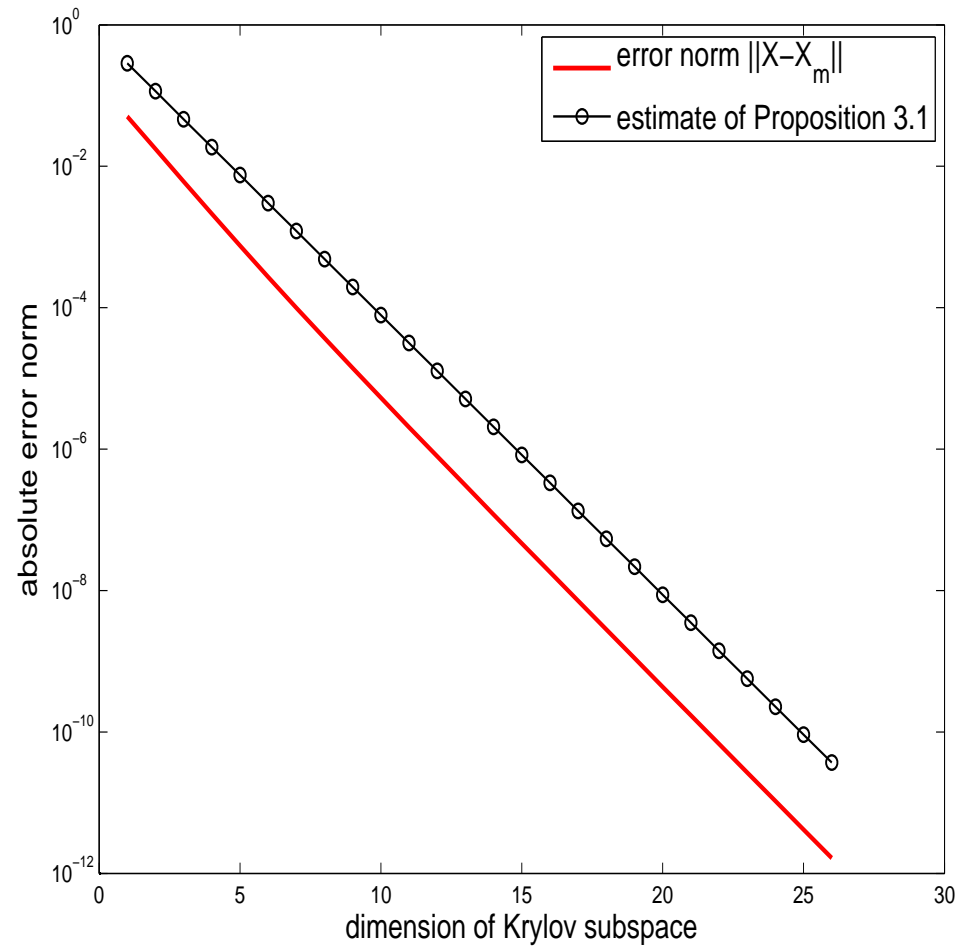
Then

$$\|X - X_m\| \leq \frac{\sqrt{\hat{\kappa}} + 1}{\hat{\lambda}_{\min} \sqrt{\hat{\kappa}}} \left(\frac{\sqrt{\hat{\kappa}} - 1}{\sqrt{\hat{\kappa}} + 1} \right)^m$$

Note: same rate as CG for $(A + \lambda_{\min}I)z = b$

(cf. works of Knizhnerman)

The case of A symmetric. An example



A : 400×400 diagonal with uniformly distributed eigenvalues in $[1, 10]$

$(\alpha_{\min} = \lambda_{\min} = 1)$

The case of $W(A)$ in an ellipse

Assume $W(A) \subseteq E \subset \mathbb{C}^+$

(E ellipse of center $(c, 0)$, foci $(c \pm d, 0)$ and major semi-axis a)

Then

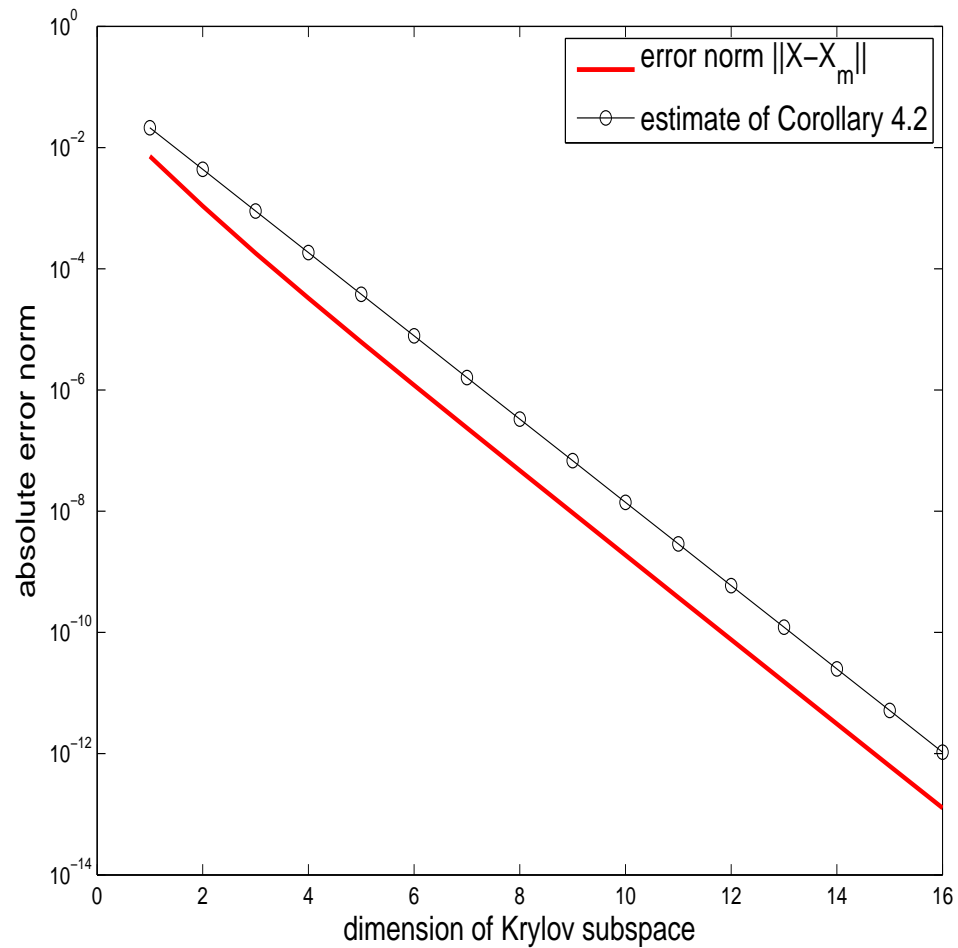
$$\|X - X_m\| \leq \frac{4}{\alpha_{\min}} \frac{r_2}{r_2 - r} \left(\frac{r}{r_2} \right)^m$$

where

$$r = \frac{a}{d} + \sqrt{\left(\frac{a}{d}\right)^2 - 1}, \quad r_2 = \frac{c + \alpha_{\min}}{d} + \sqrt{\left(\frac{c + \alpha_{\min}}{d}\right)^2 - 1}$$

Note: same rate as FOM for $(A + \alpha_{\min}I)z = b$

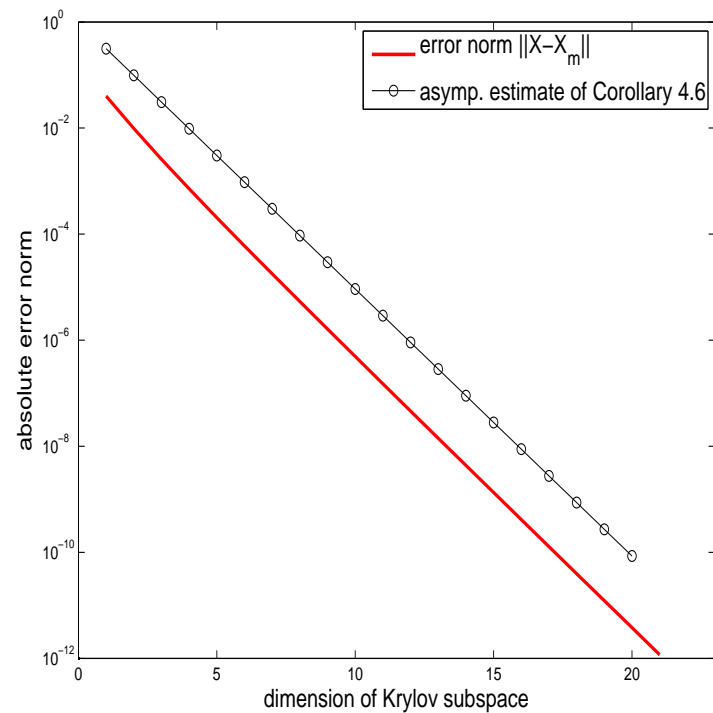
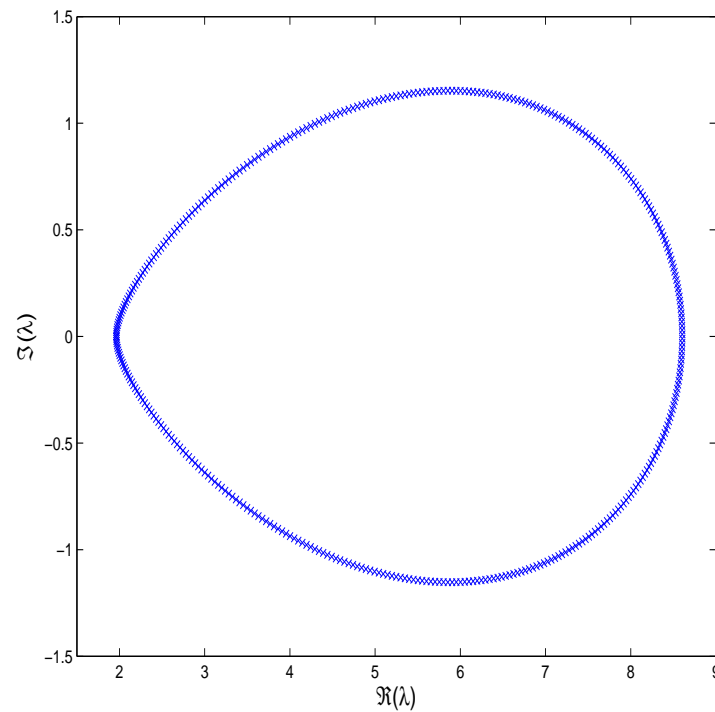
The case of $W(A)$ in an ellipse. First example



A normal with eigenvalues on an elliptic curve

The case of $W(A)$ in a wedge-shaped set. An example

Generalization to a wedge-shaped convex set of \mathbb{C}^+ .



A : diagonal (normal) matrix on the wedge-shaped curve.

(Inclusion set from Hochbruck & Lubich, 1997)

Cyclic low rank Smith method

(ADI made efficient)

(see, e.g., Li 2000, Penzl 2000)

$$X_0 = 0, X_j = -2p_j(A + p_j I)^{-1} B B^\top (A + p_j I)^{-\top} \quad j = 1, \dots, \ell \\ + (A + p_j I)^{-1} (A - p_j I) X_{j-1} (A - p_j I)^\top (A + p_j I)^{-\top}$$

with

$$r_\ell(t) = \prod_{j=1}^{\ell} (t - p_j), \quad \{p_1, \dots, p_\ell\} = \operatorname{argmin} \max_{t \in \Lambda(A)} \left| \frac{r_\ell(t)}{r_\ell(-t)} \right|$$

Cyclic low rank Smith method

(ADI made efficient)

(see, e.g., Li 2000, Penzl 2000)

$$X_0 = 0, X_j = -2p_j(A + p_j I)^{-1} B B^\top (A + p_j I)^{-\top} \quad j = 1, \dots, \ell \\ + (A + p_j I)^{-1} (A - p_j I) X_{j-1} (A - p_j I)^\top (A + p_j I)^{-\top}$$

with

$$r_\ell(t) = \prod_{j=1}^{\ell} (t - p_j), \quad \{p_1, \dots, p_\ell\} = \operatorname{argmin}_{t \in \Lambda(A)} \max \left| \frac{r_\ell(t)}{r_\ell(-t)} \right|$$

Convergence considerations:

Convergence depends on choice of $\{p_j\}$. For A spd:

$$\|X - X_\ell\| \approx \left(\frac{\sqrt{\kappa_{adi}} - 2}{\sqrt{\kappa_{adi}} + 2} \right)^\ell, \quad \kappa_{adi} = \frac{\lambda_{\max}}{\lambda_{\min}}$$

Extended Krylov subspace method

Galerkin condition: $X_m \in \mathcal{K}$ s.t.

$$R := AX_m + X_m A^\top + bb^\top \perp \mathcal{K}$$

$$\mathcal{K} = \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B), \quad \text{range}(\mathcal{V}_m) = \mathcal{K}$$

(Druskin & Knizhnerman '98, Simoncini '07) $X_m = \mathcal{V}_m Y_m \mathcal{V}_m^\top$

Projected Lyapunov equation:

$$(\mathcal{V}_m^\top A \mathcal{V}_m) Y_m + Y_m (\mathcal{V}_m^\top A^\top \mathcal{V}_m) + \mathcal{V}_m^\top b b^\top \mathcal{V}_m = 0$$

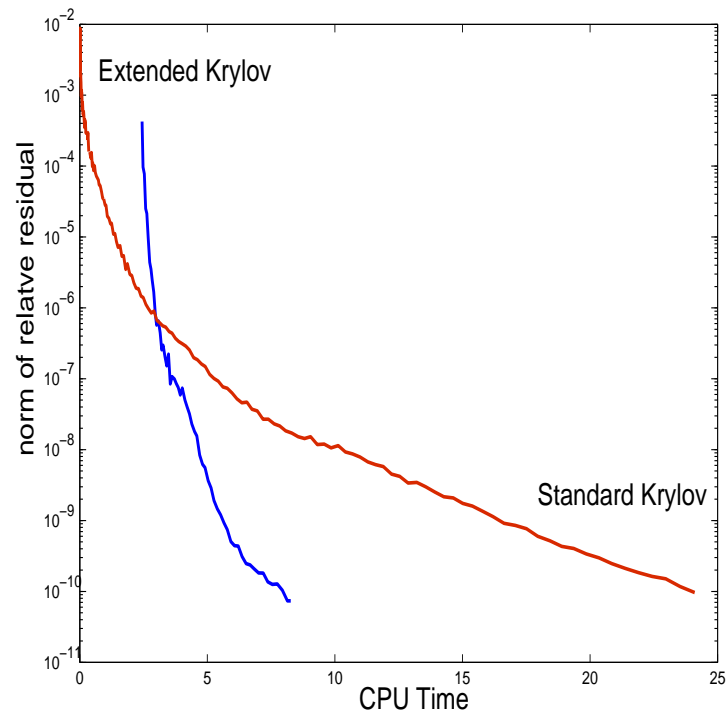
\Updownarrow

$$\mathcal{T}_m Y_m + Y_m \mathcal{T}_m^\top + e_1 e_1^\top = 0$$

Performance evaluation. I

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10z\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t)$$

A matrix $18^3 \times 18^3$



approximation space dim.: 146 (Standard Krylov) 112 (Extended Krylov)

Performance evaluation. II

Stopping criterion: norm of difference in solution

	s	EKSM		CF-ADI	
		time(#its)	dim.space	time (#its)	dim.space
Example	1	5.95 (12)	24	31.66 (6)	120
rail_5177	2	8.08 (10)	40	30.83 (5)	200
tol=10 ⁻⁵	4	11.11 (7)	56	40.20 (5)	400
	7	18.12 (6)	84	54.22 (5)	700
Example (*)	1	38.95 (34)	68	588.68 (5)	150
tol=10 ⁻⁸	2	50.50 (33)	132	633.41 (5)	300
	4	90.69 (33)	264	722.92 (5)	600
	7	204.91 (32)	448	857.57 (5)	1050

$$\mathbf{x}' = \mathbf{x}_{xx} + \mathbf{x}_{yy} + \mathbf{x}_{zz} - 10x\mathbf{x}_x - 1000y\mathbf{x}_y - 10z\mathbf{x}_z + \mathbf{b}(x, y)\mathbf{u}(t) \quad (*)$$

Convergence analysis of Extended Krylov

General considerations

$$AX + XA^\top + BB^\top = 0$$

$$A^{-1}X + XA^{-\top} + A^{-1}BB^\top A^{-\top} = 0$$

Summing up for any $\gamma \in \mathbb{R}$, we obtain yet a Lyapunov equation:

$$(A + \gamma A^{-1})X + X(A^\top + \gamma A^{-\top}) + [B, \sqrt{\gamma}A^{-1}B][B^\top; \sqrt{\gamma}B^\top A^{-\top}] = 0$$

Convergence analysis of Extended Krylov

General considerations

$$AX + XA^\top + BB^\top = 0$$

$$A^{-1}X + XA^{-\top} + A^{-1}BB^\top A^{-\top} = 0$$

Summing up for any $\gamma \in \mathbb{R}$, we obtain yet a Lyapunov equation:

$$(A + \gamma A^{-1})X + X(A^\top + \gamma A^{-\top}) + [B, \sqrt{\gamma}A^{-1}B][B^\top; \sqrt{\gamma}B^\top A^{-\top}] = 0$$

with $\mathcal{K}_m(A + \gamma A^{-1}, [B, \sqrt{\gamma}A^{-1}B]) \subsetneq \mathcal{K}_m(A, B) + \mathcal{K}_m(A^{-1}, A^{-1}B)$

$$\mathcal{A}X + X\mathcal{A}^\top + \mathcal{B}\mathcal{B}^\top = 0$$

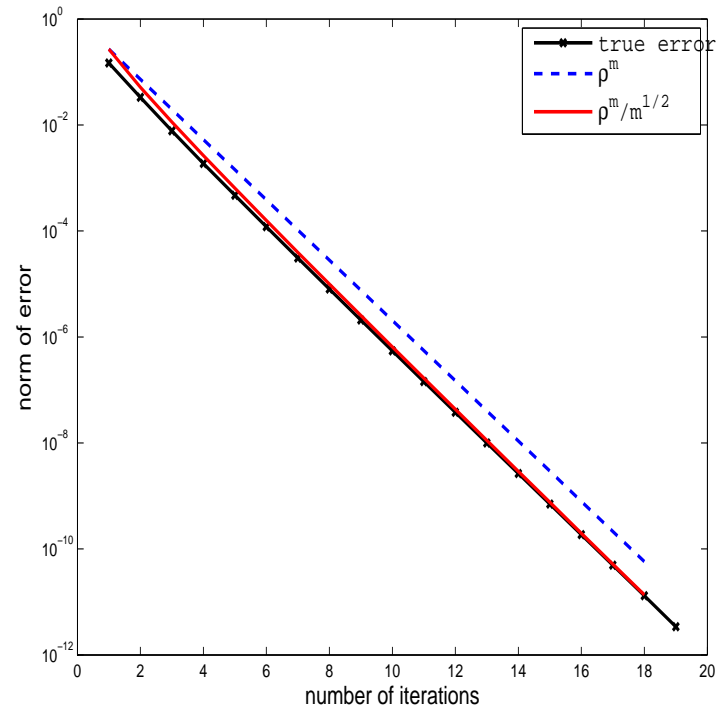
Convergence analysis of Extended Krylov: A symmetric pos.def.

Kressner & Tobler tr'09:

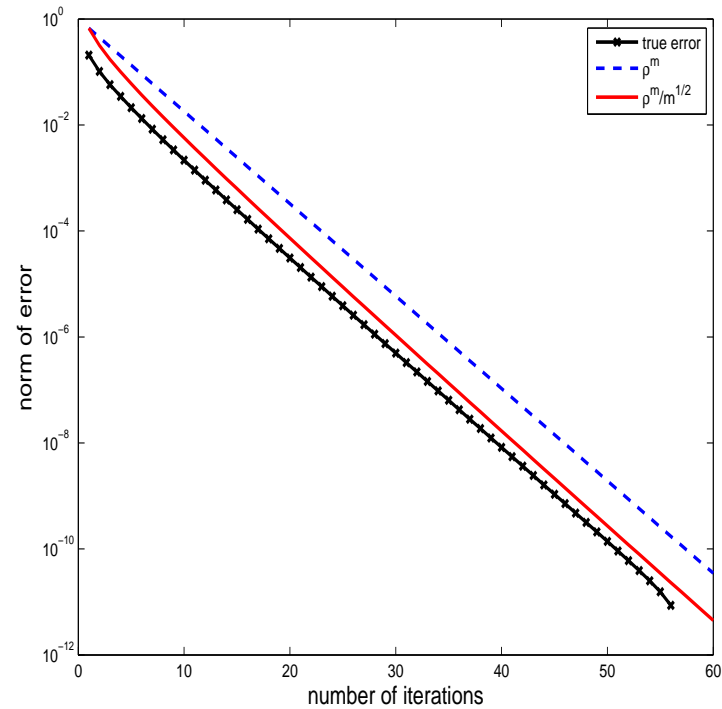
$$\|X - X_m\| \lesssim \left(\underbrace{\left(\frac{\sqrt[4]{\kappa} - 1}{\sqrt[4]{\kappa} + 1} \right)^{\frac{1}{2}}}_{\rho} \right)^m \quad \kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$$

A symmetric. An example

True error norm and asymptotic estimates for A symmetric.



Left: Spectrum in $[0.1, 10]$.

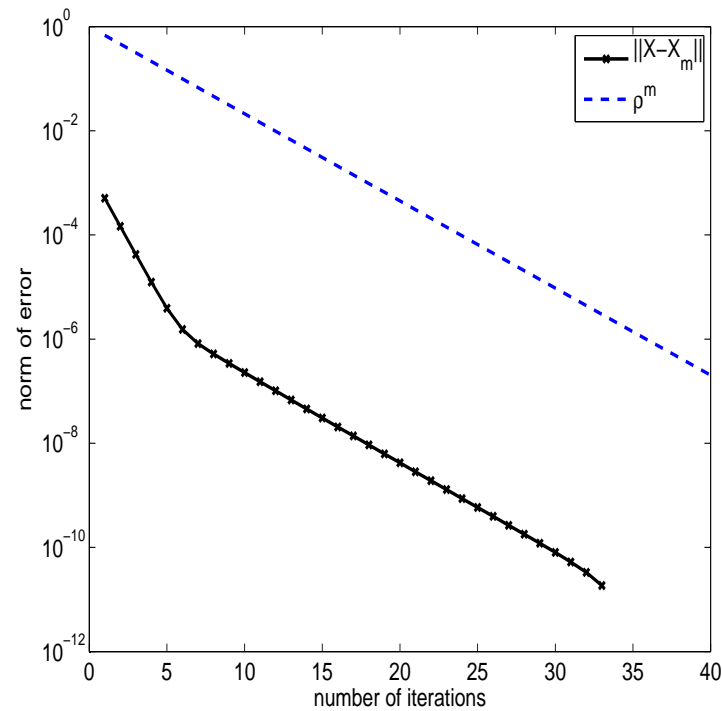


Right: Spectrum in $[0.01, 100]$

Convergence analysis of Extended Krylov: A nonsymmetric

$W(A) \subset \mathbb{C}^+$ a disk of center c and radius r .

$$\|X - X_m\| \lesssim \left(\frac{r^2}{4c^2 - 3r^2} \right)^m$$



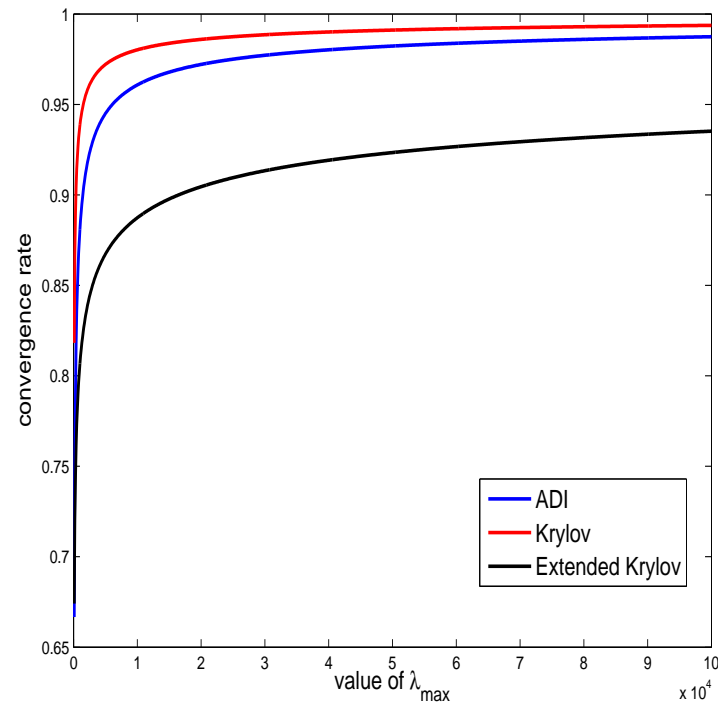
Knizhnerman & Simoncini (in prep.)

Comparison of convergence rates: A symmetric

ADI iteration: $\varepsilon_{adi,j} \approx \left(\frac{\sqrt{\kappa_{adi}} - 2}{\sqrt{\kappa_{adi}} + 2} \right)^j$

Standard Krylov: $\varepsilon_{kr,j} \approx \left(\frac{\sqrt{\hat{\kappa}} - 1}{\sqrt{\hat{\kappa}} + 1} \right)^j$

Extended Krylov: $\varepsilon_{ek,l} \approx \left(\left(\frac{4\sqrt{\kappa} - 1}{4\sqrt{\kappa} + 1} \right)^{1/2} \right)^l$



$$\lambda_{\min} = 1, \lambda_{\max} \in [10^2, 10^5]$$

Transfer function approximation (cf. MOR)

$$h(\sigma) = c^*(A - i\sigma I)^{-1}b, \quad \sigma \in [\alpha, \beta]$$

Given space \mathcal{K} and V s.t. $\mathcal{K} = \text{range}(V)$,

$$h(\sigma) \approx (V^*c)^*(V^*AV - \sigma I)^{-1}(V^*b)$$

For $\mathcal{K} = K_m(A, b)$ (standard Krylov):

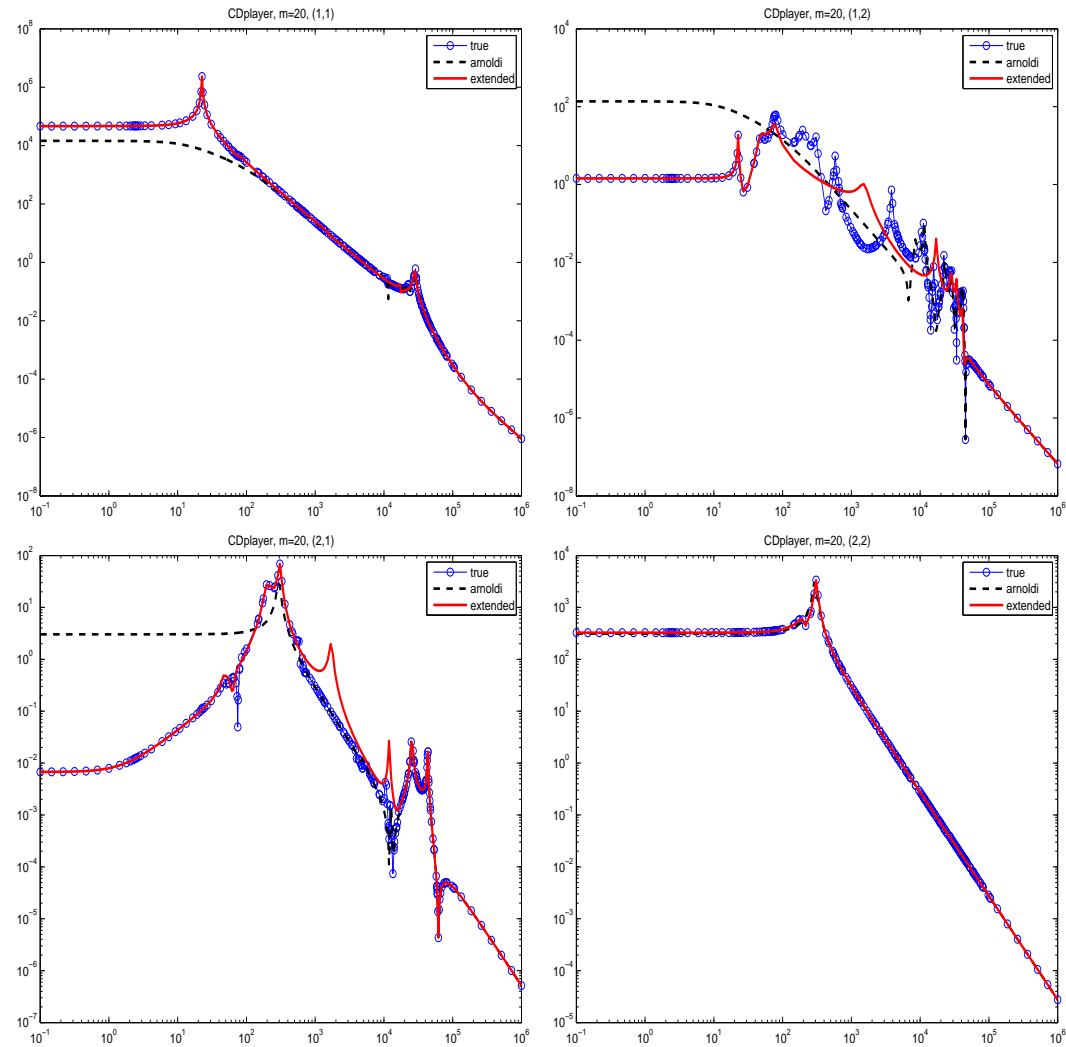
$$h_m(\sigma) = (V_m^*c)^*(H_m - \sigma I)^{-1}e_1 \|b\|$$

For $\mathcal{K} = K_m(A, b) + K_m(A^{-1}, A^{-1}b)$ (EKSM):

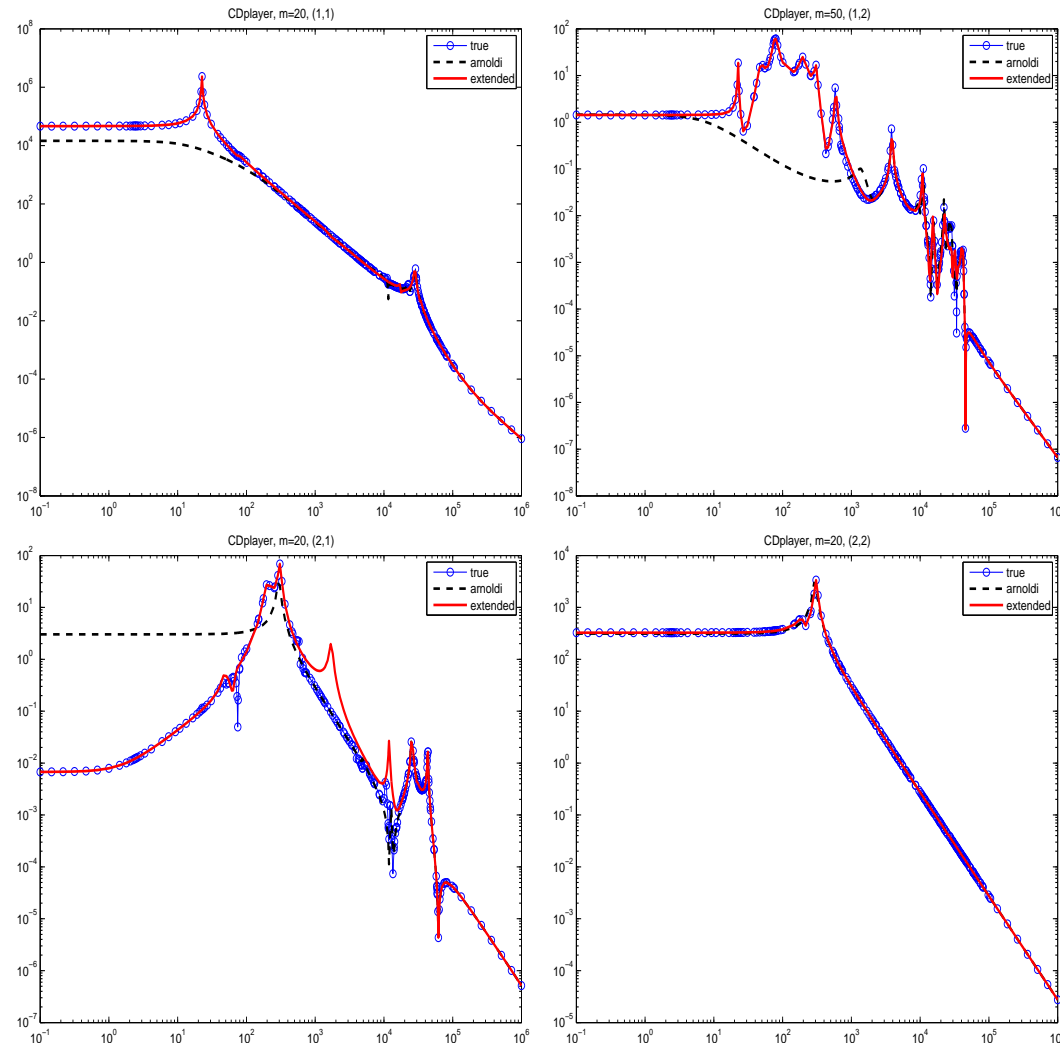
$$h_m(\sigma) = (U_m^*c)^*(T_m - \sigma I)^{-1}e_1 \|b\|$$

Alternative: Rational Krylov (Grimme-Gallivan-VanDooren etc.)
choosing the poles unresolved issue

An example: CD Player, $|h(\sigma)| = |C_{:,i}^*(A - i\sigma I)^{-1}B_{:,j}|$



An example: CD Player, $|h(\sigma)| = |C_{:,i}^*(A - i\sigma I)^{-1}B_{:,j}|$



Other related problems

- Projected generalized Lyapunov equation

$$EXA^\top + AXE^\top = -P_l BB^\top P_l^\top, \quad X = P_r X P_r^\top$$

(Stykel & Simoncini (in prep.))

- Sylvester equation: $AX + XB + C = 0$ (Heyouni '08)

- Riccati equation: $AX + XA^\top - XGX + C = 0$

(Heyouni & Jbilou '08)

- Shifted systems: $(A - \sigma I)x = b$ with many σ 's

(..., Simoncini tr'09)

- Special Sylvester equation: $AX + X\Sigma = [b(\sigma_1), \dots, b(\sigma_s)]$

(Simoncini tr'09)

Approximation space: Extended Krylov subspace

Applications. Ill-posed problem. I

$$\left\{ \begin{array}{l} u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1] \\ u(x, y, z) = 0, \quad (x, y, z) \in \partial\Omega \times [0, z_1] \\ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega \\ u_z(x, y, 0) = \mathbf{0}, \quad (x, y) \in \Omega \end{array} \right.$$

L elliptic oper., linear, self-adjoint, positive def.

Pb: determine u for $z = z_1$: $f(x, y) = u(x, y, z_1), (x, y) \in \Omega.$

Applications. Ill-posed problem. I

$$\left\{ \begin{array}{l} u_{zz} - Lu = 0, \quad (x, y, z) \in \Omega \times [0, z_1] \\ u(x, y, z) = 0, \quad (x, y, z) \in \partial\Omega \times [0, z_1] \\ u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega \\ u_z(x, y, 0) = \mathbf{0}, \quad (x, y) \in \Omega \end{array} \right.$$

L elliptic oper., linear, self-adjoint, positive def.

Pb: determine u for $z = z_1$: $f(x, y) = u(x, y, z_1)$, $(x, y) \in \Omega$.

Separation of variables: $u(x, y, z) = \cosh(z\sqrt{L})g$

★: L unbounded $\Rightarrow \cosh(z\sqrt{L})g$ unstable (wrto data perturbations)

L. Eldén & Simoncini '09

Applications. Ill-posed problem. II

Regularization: \tilde{g} perturbed data

$$u(x, y, z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y)$$

$$\Rightarrow v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, \tilde{g} \rangle s_k(x, y)$$

(λ_k^2, s_k) eigenpairs of L

Applications. Ill-posed problem. II

Regularization: \tilde{g} perturbed data

$$u(x, y, z) = \sum_{k=1}^{\infty} \cosh(\lambda_k z) \langle s_k, g \rangle s_k(x, y)$$

$$\Rightarrow v(x, y, z) = \sum_{\lambda_k \leq \lambda_c} \cosh(\lambda_k z) \langle s_k, \tilde{g} \rangle s_k(x, y)$$

(λ_k^2, s_k) eigenpairs of L

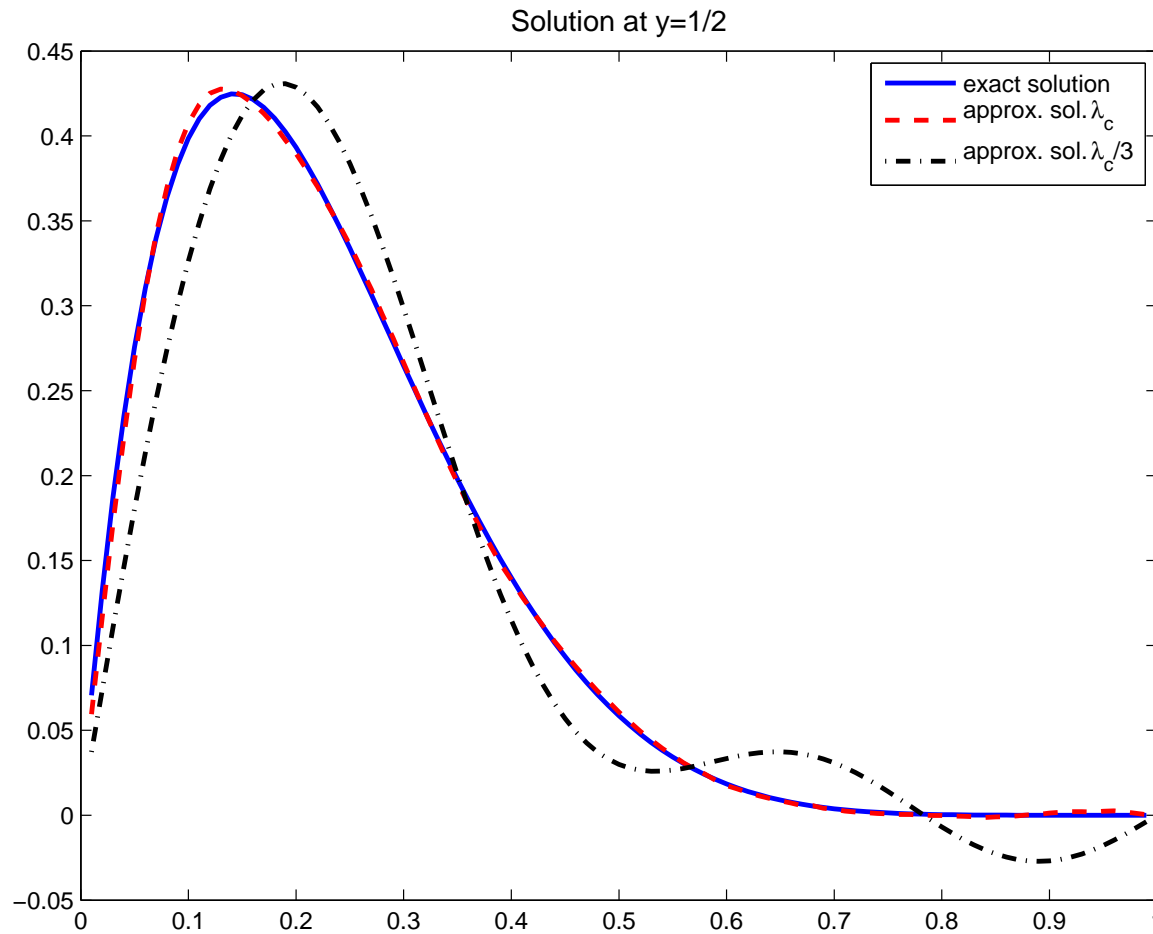
Approx using (e.g.) inverted Krylov subspace $K_m(L^{-1}, \tilde{g})$:

$$u^{(m)}(z) = V_m \cosh(z \sqrt{H_m^{-1}}) e_1 \|g\| \Rightarrow$$

$$\Rightarrow v^{(m)}(z) = V_m \sum_{1/\theta_j^{(m)} \leq \lambda_c} y_j^{(m)} \cosh(z/\theta_j^{(m)}) (y_j^{(m)})^\top e_1 \|\tilde{g}\|$$

$((\theta_j^{(m)})^2, y_j^{(m)})$ eigenpairs of $H_m = V_m^\top L^{-1} V_m$

Approximate regularized solution.



Exact solution (solid line); Approximate solution computed with $\lambda_c = 40$ and $k = 18$ (dashed) and for $\lambda_c = 40/3$ and $k = 30$ (dashed-dotted)

Conclusions

- Great potential of using “ $f(A)$ ”-machinery in application problems
- Exploit low cost of using A instead of $f(A)$
- Projection combined with matrix function evaluations

`valeria@dm.unibo.it`

`http://www.dm.unibo.it/~simoncin`

References

1. Leonid Knizhnerman and V. Simoncini, *Convergence analysis of the Extended Krylov Subspace Method for the Lyapunov equation*, In preparation.
2. V. Simoncini, *The Extended Krylov subspace for parameter dependent systems*, August 2009, pp.1-13.
3. Lars Eldèn and V. Simoncini, *Solving Ill-Posed Cauchy Problems by a Krylov Subspace Method*, *Inverse Problems*, v.25, n.6 (June 2009).
4. V. Simoncini and Vladimir Druskin, *Convergence analysis of projection methods for the numerical solution of large Lyapunov equations*, *SIAM J. Numerical Analysis*. Volume 47, Issue 2, pp. 828-843 (2009).
5. V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, *SIAM J. Scient. Computing*, v.29, n.3 (2007), pp. 1268-1288.