# Subspace Methods for Approximating the Numerical Range and Associate Quantities

Daniel Kressner

Institute of Mathematics / EPF Lausanne
currently at TU Munich

daniel.kressner@epfl.ch
http://anchp.epfl.ch

Joint work with Ding Lu and Bart Vandereycken.

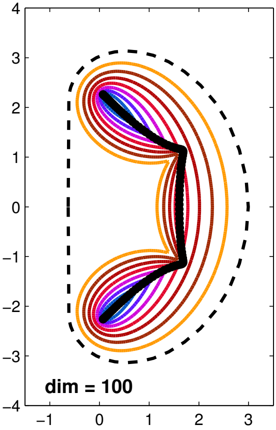
WONRA 2018

# Numerical range

For matrix $A \in \mathbb{C}^{n\times n}$, consider numerical range :

$$\mathcal{F}(A) = \left\{ \frac{x^*Ax}{x^*x} \;:\; x \neq 0 \right\} \subset \mathbb{C}$$

$A$ is large and data-sparse!




Associated quantities:

- Numerical radius:

$$\text{numrad}(A) = \max |\mathcal{F}(A)|$$

- Crawford number:

$$\text{dist}\big(\mathcal{F}(A), 0\big)$$

$=$ coercivity constant

# Example

Coercivity constants of boundary integral operators in acoustic scattering [Betcke/Spence'2011].

Consider bilinear form

$$a_k(u,v) = \int_\Gamma B_k u(y) \cdot \overline{v(y)}\, \mathrm{d}s(y), \quad \text{with} \quad B_k = I + K_k - \mathrm{i}kS_k, \quad (1)$$

with

- wave number $k > 0$
- boundary $\Gamma$ of sound-soft bounded obstacle in $\mathbb{R}^3$
- $u(x), v(x) \in L^2(\Gamma)$

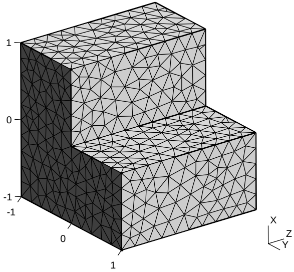
$I$ is the identity operator, and $K_k$ and $S_k$ are defined by

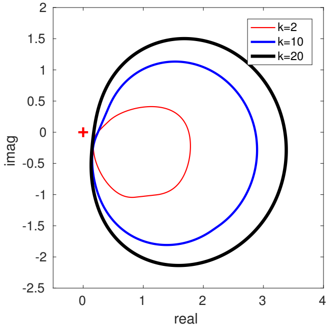
$$K_k u(x) = 2\int_\Gamma \frac{\partial \Phi(x,y)}{\partial n(x)} u(y)\, \mathrm{d}s(y), \qquad S_k u(x) = 2\int_\Gamma \Phi(x,y) u(y)\, \mathrm{d}s(y),$$

where $\Phi(x,y) = e^{\mathrm{i}k\|x-y\|_2}/(2\pi\|x-y\|_2)$ for $x, y \in \mathbb{R}^3$, $x \neq y$, $n(x)$ is the outpoint unit normal at $\Gamma$.

# Example

Standard boundary element discretization $\rightsquigarrow$ discretized boundary integral operator $A_k^h$




Left: L-shaped obstacle discretized with mesh size $h = 0.2$.
Right: Numerical range of the discretized boundary integral operator $A_k^h$ for wave numbers $k = 2, 10, 20$ and $h = 0.2$.

Coercivity constant $\inf_u |a_k(u,u)|/\|u\|^2$ approximated by Crawford number of $A_k^h$.

# Subspace methods

# Numerical range and parameter-dependent EVP

Consider Hermitian part $H(A) = (A + A^*)/2$:

$$\mathcal{F}(A) \cap \mathbb{R} = \left[\lambda_{\min}(H(A)), \lambda_{\max}(H(A))\right]$$

Existing methods (see [Psarrakos/Tsatsomeros'02]) based on constructing polygonal hull from

$$\left[\lambda_{\min}(H(e^{i\theta}A)), \lambda_{\max}(H(e^{i\theta}A))\right]$$

for $\theta \in [0, 2\pi]$.

- Note that $\lambda_{\max}(H(e^{i(\theta+\pi)}A)) = -\lambda_{\min}(H(e^{i\theta}A))$.
- Numerical range computation equivalent to solving parameter-dependent Hermitian eigenvalue problem

$$\varphi(\theta) = \lambda_{\min}(H(e^{i\theta}A)), \quad \theta \in [0, 2\pi].$$

# Subspace approach

Let $V \in \mathbb{C}^{n\times k}$ be orthonormal basis of subspace $\mathcal{V} \subset \mathbb{C}^n$. Then

$$\mathcal{F}(V^*AV) = \left\{ \frac{x^*Ax}{x^*x} : \ x \in \mathcal{V}, x \neq 0 \right\} \subset \mathcal{F}(A).$$

Goal: Choose $\mathcal{V}$ such that $\mathcal{F}(V^*AV)$ is a good approximation of $\mathcal{F}(A)$ (or at least for quantities of interest).

Suitable candidate:

- Sample eigenvectors $v(\theta_j)$, $\theta_j \in [0, 2\pi]$ belonging to $\lambda_{\min}(H(e^{i\theta}A))$.
- Set $\mathcal{V} = \mathrm{span}\{v(\theta_1), \ldots, v(\theta_N)\}$ and compute orthonormal basis of $\mathcal{V}$ via QR decomposition or via SVD (dimensionality reduction).

# Subspace approach: interpolation properties

Properties of $\varphi(\theta) = \lambda_{\min}(H(e^{i\theta}A))$ vs. $\varphi(\theta; V) = \lambda_{\min}(H(e^{i\theta}V^*AV))$.

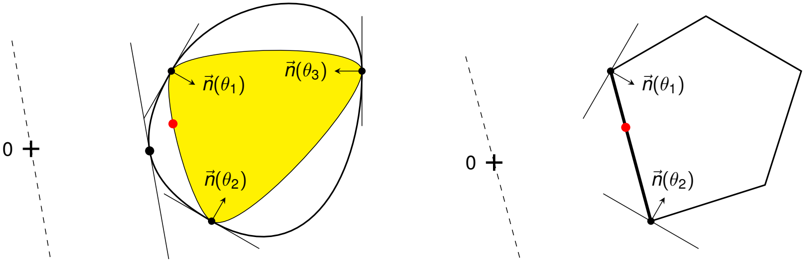
**Lemma.**

(a) Monotonicity: $\varphi(\theta; V) \geq \varphi(\theta)$.

(b) Interpolation: Suppose that $\mathcal{V}$ contains eigenvector belonging to eigenvalue $\lambda_{\min}(H(e^{i\theta}A))$ of $H(e^{i\theta}A)$. Then,

$$\varphi(\theta; V) = \varphi(\theta), \quad \varphi'(\theta; V) = \varphi'(\theta).$$

(latter requires $\lambda_{\min}(H(e^{i\theta}A))$ to be a simple eigenvalue)



# Subspace approach: existence of good subspace

- Let $A : [-1, 1] \to \mathbb{C}^{n\times n}$ be real analytic such that $A(\theta)$ is Hermitian for all $\theta \in [-1, 1]$.
- Let $v(\theta)$ be suitably normalized eigenvector belonging to $\lambda_{\min}(A(\theta))$. If $\lambda_{\min}(A(\theta))$ is simple then there exists analytic extension

$$v : E_R \to \mathbb{C}^n, \quad \text{Bernstein ellipse } E_R, \quad R > 1.$$

See [Baumgärtel'1985], [Kato'1980].

Choose Chebyshev nodes

$$\theta_j = \cos\left(\frac{2j-1}{2N}\pi\right), \quad j = 1, \dots, N,$$

and let

$$p_N(\theta) = \ell_1(\theta)v(\theta_1) + \ell_2(\theta)v(\theta_2) + \cdots + \ell_N(\theta)v(\theta_N),$$

with Lagrange polynomials $\ell_1, \dots, \ell_N : [-1, 1] \to \mathbb{R}$.

# Subspace approach: existence of good subspace

On the one hand, standard interpolation error estimates [Mason/Handcomb'2003] imply

$$\|v(\theta) - p_N(\theta)\|_2 \lesssim R^{-N}, \quad \forall \theta \in [-1, 1].$$

On the other hand, $p_N(\theta)$ is a linear combination of $v(\theta_1), \ldots, v(\theta_N)$ and thus

$$p_N(\theta) \in \mathcal{V} := \mathrm{span}\{v(\theta_1), \ldots, v(\theta_N)\}.$$

In summary: The distance between any eigenvector and $\mathcal{V}$,

$$\mathrm{dist}(\mathcal{V}, v(\theta)) := \min\{\|v(\theta) - v\|_2 : v \in \mathcal{V}\},$$

satisfies

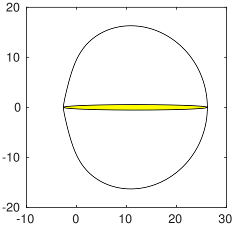
$$\mathrm{dist}(\mathcal{V}, v(\theta)) \lesssim R^{-N}.$$

In turn

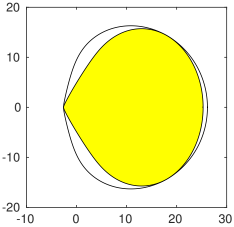
$$\lambda_{\min}(V^* A(\theta) V) = \lambda_{\min}(A(\theta)) + \mathcal{O}(R^{-N}).$$

# Subspace approach: example

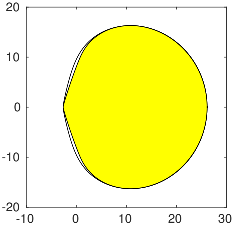
$A = 100 \times 100$ random upper Hessenberg matrix. White: $\mathcal{F}(A)$, yellow: $\mathcal{F}(V^*AV)$ with $\mathcal{V}$ containing $N$ uniformly sampled $v(\theta)$.



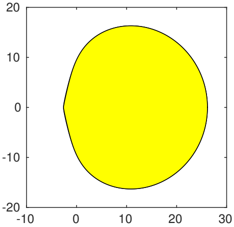

$N = 2$




$N = 3$




$N = 4$




$N = 10$

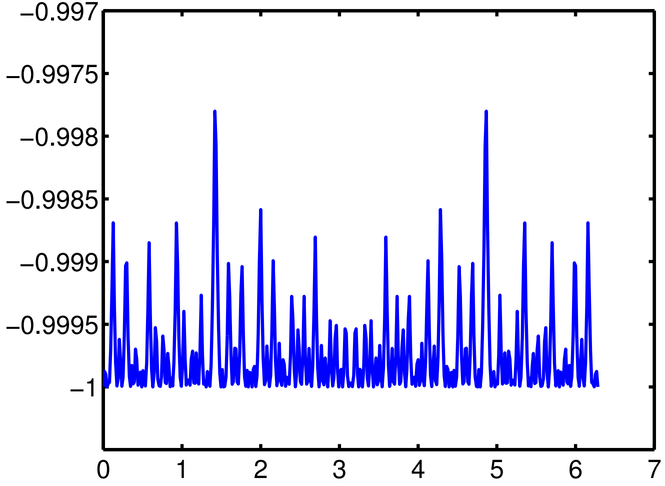
# Subspace approach: remarks

- Instead of uniform sampling: Greedy sampling strategy guided by lower and upper bounds [Sirković/Kressner'2016].
- Good convergence requires, at the minimum, that *n*-width of $\{v(\theta)\}$ decays sufficiently fast.

# Subspace approach: counterexample

$A = 100 \times 100$ random orthogonal matrix.

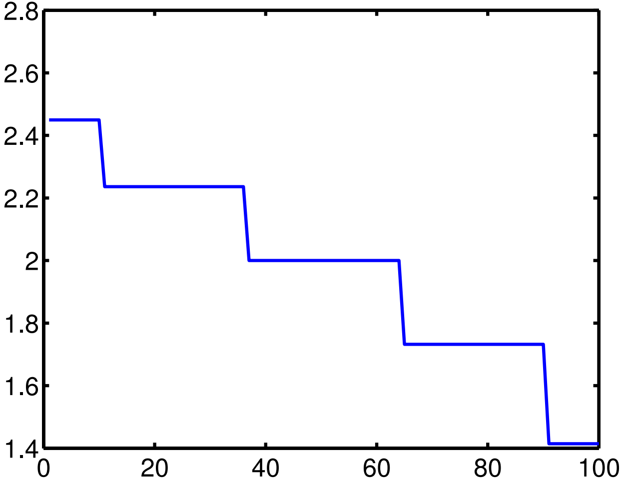
$\lambda_{\min}(H(e^{i\theta}A))$ for $\theta \in [0, 2\pi]$



# Subspace approach: counterexample

$A = 100 \times 100$ random orthogonal matrix.

Singular values of $[v(\theta_1), \ldots, v(\theta_N)]$



# Crawford number

# Crawford number

Crawford [1976] number of $A \in \mathbb{C}^{n \times n}$ is defined distance of $\mathcal{F}(A)$ to zero:

$$\gamma(A) = \min\{\, |z| : z \in \mathcal{F}(A)\,\}.$$

Equivalent characterization:

$$\gamma(A) = \max\left\{ \max_{\theta \in [0,2\pi]} \lambda_{\min}(H(e^{\mathrm{i}\theta} A)), \quad 0 \right\}.$$

In applications, it is usually of interest to verify $\gamma(A) > 0$.

Subspace methods for eigenvalue optimization developed in:

- [Kressner/Vandereycken'2014]: pseudospectral abscissa computation
- [Kangal et al.'2015]: general multiparameter optimization of Hermitian eigenvalue problems
- [Aliyev et al.'2016]: $\mathcal{H}_\infty$ norm computation
- [Kressner/Lu/Vandereycken'2018]: Crawford number computation
- ...

# Eigenvalue optimization

Subspace method for general eigenvalue optimization problem

$$\max_{\theta \in \Omega} \varphi(\theta) \quad \text{with} \quad \varphi(\theta) = \lambda_{\min}(A(\theta)),$$

```
Compute λ_0 = λ_min(A(θ_0)) and corresponding normalized eigenvector v_0.
Initialize V_0 = v_0.
for k = 0, 1, ..., N − 1 do
    Solve θ_{k+1} = arg max_{θ∈Ω} λ_min(V_k^* A(θ) V_k) and set
    φ_{k+1} = λ_min(V_k^* A(θ_{k+1}) V_k).
    Stopping criteria: if φ_{k+1} − λ_k < tol · |φ_{k+1}| then terminate.
    Compute the smallest eigenvalue λ_{k+1} with normalized eigenvector v_{k+1} of A(θ_{k+1}).
    Subspace update: V_{k+1} = orth([V_k, v_{k+1}]).
end for
```

# Convergence of subspace method

Let $V_k = \text{span}(v_1, \ldots, v_k)$. Optimality condition $\varphi'(\theta_{k+1}; V_k) = 0$ combined with Taylor expansion give

$$0 = \varphi'(\theta_{k+1}; V_k) = \varphi'(\theta_*;\ V_k) + \varphi''(\xi; V_k)(\theta_{k+1} - \theta_*),$$

for some $\xi$ between $\theta_*$ and $\theta_{k+1}$. Therefore,

$$|\theta_{k+1} - \theta_*| = |\varphi'(\theta_*;\ V_k)|/|\varphi''(\xi; V_k)|.$$

It remains to control $|\varphi'(\theta_*;\ V_k)|$.

Because of interpolation conditions, $\varphi'(\theta;\ V_k) - \varphi'(\theta)$ is zero at $2k + 1$ distinct points. Standard interpolation error estimates for analytic functions $\rightsquigarrow$

$$|\varphi'(\theta_*; V_k) - \varphi'(\theta_*)| \lesssim |\theta_* - \theta_k| \prod_{i=0}^{k-1} |\theta_* - \theta_i|^2.$$

Thus,

$$|\theta_{k+1} - \theta_*| \lesssim |\theta_* - \theta_k| \prod_{i=0}^{k-1} |\theta_* - \theta_i|^2$$

# Convergence of subspace method

**Theorem.** If algorithm converges to $\theta_*$ with $\lambda_{\min}(A(\theta_*))$ simple and $\varphi''(\theta_*) < 0$ then $|\theta_* - \theta_k| \leq e_k$ with a sequence $(e_k)$ that converges to zero with local *R*-convergence order

$$\sigma = 1 + \sqrt{2} \approx 2.4142\ldots$$

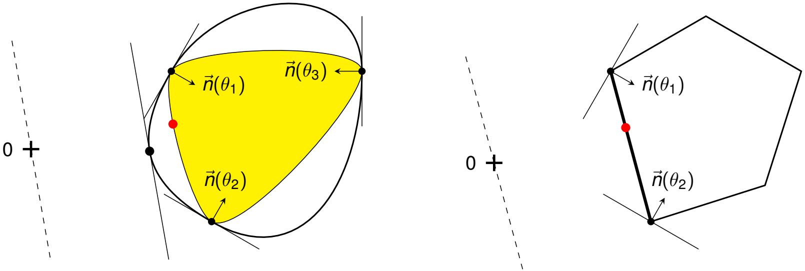
# Application to Crawford number

$$\varphi(\theta) = \lambda_{\min}(H(e^{i\theta} A))$$

Compute $\lambda_0 = \lambda_{\min}(H(e^{i\theta_0} A))$ and corresponding normalized eigenvector $v_0$.
Initialize $V_0 = v_0$.
**for** $k = 0, 1, \ldots, N-1$ **do**
    Solve $\theta_{k+1} = \arg\max_{\theta \in \Omega} \lambda_{\min}(H(e^{i\theta} V_k^* A V_k))$ and set $\varphi_{k+1} = \lambda_{\min}(H(e^{i\theta} V_k^* A V_k))$.
    *Stopping criteria*: **if** $\varphi_{k+1} - \lambda_k < \text{tol} \cdot |\varphi_{k+1}|$ or $\varphi_{k+1} \le 0$ **then** terminate.
    Compute the smallest eigenvalue $\lambda_{k+1}$ with normalized eigenvector $v_{k+1}$ of $H(e^{i\theta_{k+1}} A)$.
    *Subspace update*: $V_{k+1} = \text{orth}([V_k, v_{k+1}])$.
**end for**

Solution of reduced problem requires Crawford number of $k \times k$ matrix $V_k^* A V_k$. Use criss-cross algorithm [Mengi/Overton'2005].

# Application to Crawford number



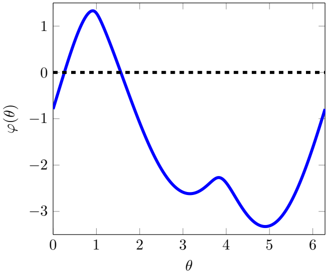
# Application to Crawford number

$$\varphi(\theta) = \lambda_{\min}(H(e^{i\theta}A))$$

**Theorem** [Guo/Higham/Tisseur'2009], [Lu/K./Vandereycken'2018]. Let $\gamma(A) > 0$. Then there exists $\theta_0 \in \mathbb{R}$ such that

$$\{\theta \colon \varphi(\theta) > 0\} \cap [\theta_0, \theta_0 + 2\pi]$$

is an open, nonempty interval $(\ell, u)$ of length at most $\pi$. Moreover, $\varphi(\theta)$ is strongly concave on any closed subinterval of $(\ell, u)$.



# Application to Crawford number

$$\varphi(\theta) = \lambda_{\min}(H(e^{\mathrm{i}\theta} A))$$

**Theorem** [Guo/Higham/Tisseur'2009], [Lu/K./Vandereycken'2018]. Let $\gamma(A) > 0$. Then there exists $\theta_0 \in \mathbb{R}$ such that

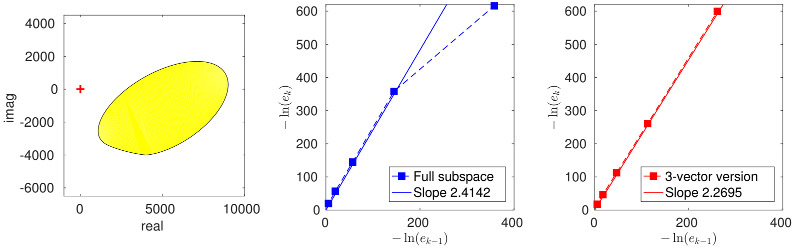
$$\{\theta \colon \varphi(\theta) > 0\} \cap [\theta_0, \theta_0 + 2\pi]$$

is an open, nonempty interval $(\ell, u)$ of length at most $\pi$. Moreover, $\varphi(\theta)$ is strongly concave on any closed subinterval of $(\ell, u)$.

Result has several important consequences:

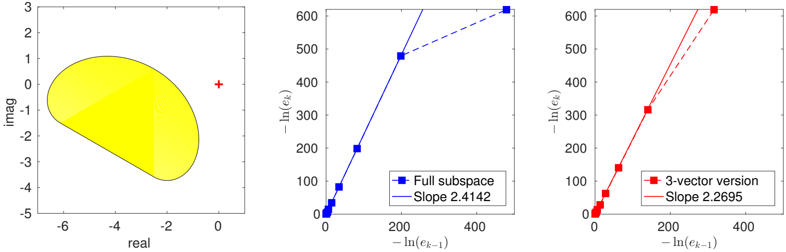
- Global maximizers of $\varphi(\theta)$ and $\varphi(\theta; V_k)$ unique.
- Condition $\varphi''(\theta_*) < 0$ in local convergence result always holds.
- Can develop a variant of subspace method that only maintains three vectors $v(\ell_k), v(\theta_k), v(u_k)$. The three-vector subspace method has local $R$-convergence order 2.27 . . ..

# Numerical results



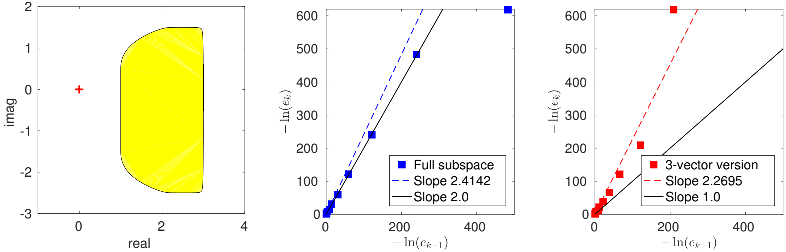

Numerical range and observed order of convergence for a shifted sum of the Fiedler and Moler matrices.

# Numerical results




Numerical range and observed order of convergence for a rotated Grcar matrix.

# Numerical results




Numerical range and observed order of convergence for the matrix

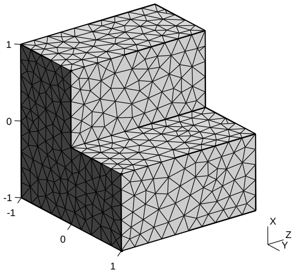
$$A = \text{tridiag}\begin{pmatrix} & \mathrm{i} & \mathrm{i} & \dots & \mathrm{i} \\ 1 & 1 & a_3 & \dots & a_n \\ & \mathrm{i} & \mathrm{i} & \dots & \mathrm{i} \end{pmatrix} + 0.5\mathrm{i}\cdot I_n \quad \text{with} \quad a_j = 2 + \frac{j}{n}, n = 120.$$

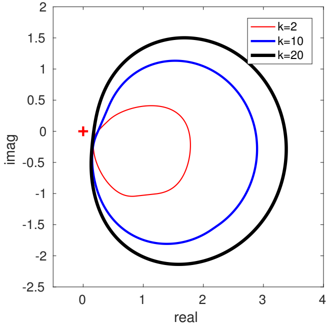
This matrix is constructed such that $\lambda_{\min}(H(e^{\mathrm{i}\theta_*}A))$ is *not* simple.

# Numerical results

Consider boundary integral operator from beginning:

- Discretization with $n = 396\,162$ degrees of freedom ($h = 0.02$) using BEM++ [Śmigaj et al.'2015].
- Discretized operator $A_k^h$ is dense but can be compressed in hierarchical matrix format. Increasingly difficult as wave number $k$ increases.
- Used tolerance $10^{-13}$ for terminating algorithms.



# Numerical results

| $k$ | memory | | Crawford number | its. | time (h) |
|---|---|---|---|---|---|
| 2 | 46 GB | Full | 1.556145884392413e-01 | 5 | 2.5 |
| | | 3-vec | 1.556145884392399e-01 | 5 | 2.5 |
| | | Uhlig | 1.556145884392416e-01 | 11 | 4.4 |
| 10 | 54 GB | Full | 1.880394259192281e-01 | 7 | 6.1 |
| | | 3-vec | 1.880394259192268e-01 | 7 | 6.1 |
| | | Uhlig | 1.880394259192323e-01 | 30 | 29.9 |
| 20 | 60 GB | full | 1.777716873410842e-01 | 8 | 8.1 |
| | | 3-vec | 1.777716873410808e-01 | 10 | 10.8 |
| | | Uhlig's | 1.777716873410810e-01 | 24 | 24.9 |

# Conclusions

- Subspaces of sampled eigenvectors speeds up numerical range computations.
- Rigorous local and global convergence analysis for Crawford number.
- Analogous results exist for pseudospectra.

Selected references:


- D. Kressner, D. Lu, B. Vandereycken. Subspace acceleration for the Crawford number and related eigenvalue optimization problems, SIAM J. Matrix Anal. Appl., vol. 39, no. 2, pp. 961–982.
- P. Sirković and D. Kressner. Subspace acceleration for large-scale parameter-dependent Hermitian eigenproblems. SIAM J. Matrix Anal. Appl., 37(2):695â718, 2016
- D. Kressner and B. Vandereycken. Subspace methods for computing the pseudospectral abscissa and the stability radius. SIAM J. Matrix Anal. Appl., 35(1):292–313, 2014


`https://anchp.epfl.ch`