Approximated structured pseudospectra

S. Noschese L. Reichel

"Numerical Linear Algebra and Applications (NL2A)"
CIRM, October 24–28, 2016

Pseudospectra

The eigenvalue sensitivity can be studied by computing pseudospectra.

Let $\Lambda(A)$ denote the spectrum of the matrix $A \in \mathbb{C}^{n\times n}$. For $\varepsilon > 0$, the ε-pseudospectrum of the matrix A is defined as

$$\Lambda_\varepsilon(A) := \left\{\lambda \in \mathbb{C}\colon \lambda \in \Lambda(A+E), \quad E \in \mathbb{C}^{n\times n}, \quad \|E\| \leq \varepsilon\right\}.$$

The set $\Lambda_\varepsilon(A)$ depends on the choice of matrix norm, however, this dependence often is not important in applications when one is interested in determining which eigenvalues of the matrix A are most sensitive to perturbations.

An insightful discussion of the ε-pseudospectrum and many applications are presented by Trefethen and Embree.

Computational task

When one considers the spectral norm, approximations of the ε-pseudospectrum are determined by computing the smallest singular value of many matrices of the form $A - zI_n$, where $z \in \mathbb{C}$.

$$\boxed{\text{If } \sigma_n(A - zI_n) \leq \varepsilon, \text{ then } z \in \Lambda_\varepsilon(A).}$$

These computations are very demanding unless A is small.

To reduce the cost, one first computes the Schur factorization $A = URU^H$ and then determines the smallest singular value of the upper triangular matrix $R - zI_n$ for many z-values in $\mathbb{C}$ [Lui 1997].

$$\boxed{\text{If } \sigma_n(R - zI_n) \leq \varepsilon, \text{ then } z \in \Lambda_\varepsilon(A).}$$

Nevertheless, the computational task remains substantial.

The computation of the ε-pseudospectrum for a moderately sized matrix $A \in \mathbb{C}^{n \times n}$ for a fixed $\varepsilon > 0$ can be very time-consuming.

Structured approach

A structure-respecting eigenvalue algorithm may yield higher accuracy and require less computing time than a structure-ignoring method.

To assess its numerical properties (e.g., it is strongly backward stable if the computed eigenvalues are exact eigenvalues of a slightly perturbed matrix with the same structure [Bunch 1987]) suitable measures of the sensitivity of the eigenvalues should be used in order not to overestimate the worst-case effect of perturbations.

These measures include structured condition numbers [Karow, Kressner and Tisseur 2006], as well as the structured ε-pseudospectrum.

The structured ε-pseudospectrum can be applied to measure the sensitivity of the eigenvalues of a structured matrix to similarly structured perturbations.

$\mathcal{S}$-structured pseudospectra

Let $\mathcal{S}$ denote the subset of matrices in $\mathbb{C}^{n\times n}$ with a particular structure, such as bandedness, Toeplitz or Hankel. Then, for $\varepsilon > 0$, the $\mathcal{S}$-structured ε-pseudospectrum of a matrix $A \in \mathcal{S}$ is given by

$$\Lambda_\varepsilon^{\mathcal{S}}(A) := \{\lambda \in \mathbb{C}\colon \lambda \in \Lambda(A+E), \quad E \in \mathcal{S}, \quad \|E\| \leq \varepsilon\}$$

Note that when one considers the spectral norm the Schur factorization of A cannot be applied for the computation of the $\mathcal{S}$-structured ε-pseudospectrum.

In fact, the computation of structured ε-pseudospectra of a matrix has become an established tool in gaining insight into behavior of associated dynamical systems under structured perturbations.

However, besides plotting the spectra of matrices $A \in \mathbb{C}^{n\times n}$ with structured random perturbations, few methods are available for computing $\Lambda_\varepsilon^{\mathcal{S}}(A)$ [Graillat 2006; Rump 2006; Karow 2010].

Thus, computing the structured ε-pseudospectral abscissa, which provides a measure of robust structured stability, remains a significant computational challenge to date.

Approximating pseudospectra

The high computational cost of computing standard (unstructured) pseudospectra has given rise to algorithms that can be executed efficiently on a parallel computer [Bekas and Gallopoulos 2002; Mezher and Philippe 2002].

We propose a different approach to speed up the computations, that also can be applied to the determination of structured pseudospectra.

Fairly few rank-one matrices known to yield large eigenvalue perturbations according to Wilkinson's analysis provide insight into the ε-pseudospectrum and when different components of the pseudospectrum coalesce.

Our method can be used to inexpensively compute approximated (standard and structured) pseudospectra when software tools, such as Eigtool [Wright 2002] and Seigtool [Karow, Kokiopoulou and Kressner 2010], are too expensive to use.

Wilkinson perturbations

Consider λ simple eigenvalue of $A \in \mathbb{C}^{n\times n}$ with unit right and left eigenvectors x and y. Let $E \in \mathbb{C}^{n\times n}$, $\|E\|_{2,F} = 1$, and $\varepsilon > 0$ so that $\lambda_E(t)$ of $A + tE$ exists and is unique for all $0 \leq t < \varepsilon$. Then

$$\lambda_E(t) = \lambda + \frac{y^H E x}{y^H x} t + \mathcal{O}(t^2)$$

and one has $|y^H E x / y^H x| \leq 1/|y^H x|$ with equality for

$$\boxed{E = \eta\, y x^H}$$

for any unimodular $\eta \in \mathbb{C}$ [Wilkinson 1965]. We refer to such matrices as Wilkinson perturbations.

The condition number of the eigenvalue λ is defined as $\kappa(\lambda) = 1/|y^H x|$.

Structured maximal perturbations

We turn to $\mathcal{S} \subsetneq \mathbb{C}^{n\times n}$. The points in a $\mathcal{S}$-structured ε-pseudospectrum are exact eigenvalues of a nearby matrix in $\mathcal{S}$. This suggests that we may use standard results from the literature on structured eigenvalue sensitivity to machine perturbations.

To determine approximations of structured pseudospectra, we project suitable Wilkinson perturbations onto $\mathcal{S}$.

We endow $\mathcal{S}$ with the Frobenius norm. Let $M|_{\mathcal{S}}$ denote the matrix in $\mathcal{S}$ closest to $M \in \mathbb{C}^{n\times n}$. Also, consider the normalized projection

$$M|_{\widehat{\mathcal{S}}} := \frac{M|_{\mathcal{S}}}{\|M|_{\mathcal{S}}\|_F} \ .$$

When M is a Wilkinson perturbation, $M|_{\mathcal{S}}$ can be used in the definition of the structured condition number and $M|_{\widehat{\mathcal{S}}}$ in the definition of the structured maximal perturbations [S.N. and Pasquini 2006].

Toeplitz structure

Let $\mathcal{S} := \mathcal{T}$ be the considered subset of Toeplitz matrices. The structure is determined by the location of the nonzero diagonals of the Toeplitz matrix in $\mathcal{T}$.

The matrix in $\mathcal{T}$ closest to $M \in \mathbb{C}^{n \times n}$ with respect to the Frobenius norm, i.e. $M|_{\mathcal{T}}$, is obtained by replacing all elements in a nonzero diagonal of M by their arithmetic mean.

This construction can be generalized to Hankel matrices by considering anti-diagonals in place of diagonals. Every symmetry pattern, such as persymmetry and skew-persymmetry, can be handled similarly [S.N. and Pasquini 2007].

Maximal $\mathcal{T}$-structured perturbations

Let λ be a simple eigenvalue of a Toeplitz matrix $A \in \mathcal{T} \subset \mathbb{C}^{n\times n}$ with right and left eigenvectors normalized so that $\|x\|_F = \|y\|_F = 1$ and $y^H x > 0$.
In [S.N. and Pasquini 2007] it has been shown that the Toeplitz structured condition number is given by

$$\kappa^{\mathcal{T}}(\lambda) = \frac{\|yx^H|_{\mathcal{T}}\|_F}{y^H x} = \kappa(\lambda)\,\|yx^H|_{\mathcal{T}}\|_F$$

and that the maximal Toeplitz structured perturbation is given by

$$\boxed{E = yx^H|_{\widehat{\mathcal{T}}}}$$

In [Buttà, Guglielmi and S.N. 2012] the structured pseudospectral abscissa of a Toeplitz matrix has been computed. The key idea in the extension of the unstructured algorithm in [Guglielmi and Overton 2011] has been considering maximal Toeplitz structured perturbations.

The most Λ_ε-sensitive pair of eigenvalues

When $\mathcal{S} = \mathbb{C}^{n\times n}$, assume that machine epsilon ε_M satisfies $0 < \varepsilon_M \ll \varepsilon_*$, where ε_* is the distance from defectivity of the matrix A,

$$\varepsilon_* = \inf\{\|A - B\|_F \colon B \in \mathbb{C}^{n\times n} \text{ is defective}\} .$$

Then the component of $\Lambda_{\varepsilon_M}(A)$ that contains λ is approximately the disk $\mathcal{D}(\lambda, \varepsilon_M)$ of radius $\kappa(\lambda)\varepsilon_M = \varepsilon_M/|y^H x|$ centered at λ.
An estimate of ε_* is given by

$$\varepsilon := \min_{\substack{1\le i\le n\\ 1\le j\le n\\ j\neq i}} \frac{|\lambda_i - \lambda_j|}{\kappa(\lambda_i) + \kappa(\lambda_j)} , \tag{1}$$

Let the index pair $\{\hat{\imath}, \hat{\jmath}\}$ minimize this ratio over all distinct eigenvalue pairs. Then $\mathcal{D}(\lambda_{\hat{\imath}}, t)$ and $\mathcal{D}(\lambda_{\hat{\jmath}}, t)$ will coalesce first when increasing t.

We note that usually the most Λ_ε-sensitive pair of eigenvalues $\{\lambda_{\hat{\imath}}, \lambda_{\hat{\jmath}}\}$ are not the two worst conditioned ones.

The most $\Lambda_\varepsilon^{\mathcal{S}}$-sensitive pair of eigenvalues

We turn to the situation when $\mathcal{S} \subsetneq \mathbb{C}^{n\times n}$. Then the role of $\kappa(\lambda)$ is played by the structured condition number $\kappa^{\mathcal{S}}(\lambda)$. For $\varepsilon_M \ll \varepsilon_*^{\mathcal{S}}$, where $\varepsilon_*^{\mathcal{S}}$ denotes the structured distance from defectivity $\varepsilon_*^{\mathcal{S}}$ of A,

$$\varepsilon_*^{\mathcal{S}} = \inf\{\|A - B\|_F \colon B \in \mathcal{S} \text{ is defective}\}\,,$$

the component of $\Lambda_{\varepsilon_M}^{\mathcal{S}}(A)$ that contains λ is approximately the disk $\mathcal{D}^{\mathcal{S}}(\lambda, \varepsilon_M)$ of radius $\kappa^{\mathcal{S}}(\lambda)\varepsilon_M = \kappa(\lambda)\|yx^H|_{\mathcal{S}}\|_F\,\varepsilon_M$ centered at λ, and an estimate of $\varepsilon_*^{\mathcal{S}}$ is given by

$$\varepsilon^{\mathcal{S}} := \min_{\substack{1\le i\le n\\ 1\le j\le n\\ j\neq i}} \frac{|\lambda_i - \lambda_j|}{\kappa^{\mathcal{S}}(\lambda_i) + \kappa^{\mathcal{S}}(\lambda_j)} \ge \varepsilon\,. \tag{2}$$

Let the index pair $\{\hat{\imath}, \hat{\jmath}\}$ minimize this ratio. Then $\mathcal{D}^{\mathcal{S}}(\lambda_{\hat{\imath}}, t)$ and $\mathcal{D}^{\mathcal{S}}(\lambda_{\hat{\jmath}}, t)$ will coalesce first as t increases.

The most $\Lambda_\varepsilon^{\mathcal{S}}$-sensitive pair of eigenvalues $\{\lambda_{\hat{\imath}}, \lambda_{\hat{\jmath}}\}$ are not necessarily the worst conditioned eigenvalues with respect to structured perturbations.

The new approach

Standard (unstructured) ε-pseudospectra are well approximated by using Wilkinson perturbations associated with the most Λ_ε-sensitive pair of eigenvalues, that is to say by using "worst case" rank-one perturbations associated with two eigenvalues whose pseudospectral components are likely to first coalesce, as determined by (1).

For matrices in $\mathcal{S}$, the structured ε-pseudospectra can be well approximated by using normalized projections of Wilkinson perturbations associated with the most $\Lambda_\varepsilon^{\mathcal{S}}$-sensitive pair of eigenvalues, that is to say with two eigenvalues whose components in the structured pseudospectra are likely to first coalesce, as determined by (2).

Algorithm for computing an approximated pseudospectrum

Data: matrix A, eigensystem $\{\lambda_i, x_i, y_i,\ \forall i = 1 : n\}$

Result: approximated $\Lambda_\varepsilon(A)$

- compute ε, $\{\hat{\imath}, \hat{\jmath}\}$ by (1)
- compute $W_{\hat{\imath}} := \varepsilon\, y_{\hat{\imath}} x_{\hat{\imath}}^H$, $W_{\hat{\jmath}} := \varepsilon\, y_{\hat{\jmath}} x_{\hat{\jmath}}^H$
- display the spectrum of $A + \eta W_{\hat{\imath}}$ for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$, where $\theta_k = 2\pi(k-1)/10^3$, $k = 1 : 10^3$
- display the spectrum of $A + \eta W_{\hat{\jmath}}$ for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$, where $\theta_k = 2\pi(k-1)/10^3$, $k = 1 : 10^3$

Algorithm for computing an approximated structured pseudospectrum

Data: matrix A, eigensystem $\{\lambda_i, x_i, y_i,\ \forall i = 1:n\}$

Result: approximated $\Lambda^{\mathcal{S}}_{\varepsilon^{\mathcal{S}}}(A)$

- compute $\varepsilon^{\mathcal{S}}$, $\{\hat{\imath}, \hat{\jmath}\}$ by (2)
- compute $W^{\mathcal{S}}_{\hat{\imath}} := \varepsilon^{\mathcal{S}}\, y_{\hat{\imath}} x^H_{\hat{\imath}}|_{\widehat{\mathcal{S}}}$, $W^{\mathcal{S}}_{\hat{\jmath}} := \varepsilon^{\mathcal{S}}\, y_{\hat{\jmath}} x^H_{\hat{\jmath}}|_{\widehat{\mathcal{S}}}$
- display the spectrum of $A + \eta W^{\mathcal{S}}_{\hat{\imath}}$ for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$, where $\theta_k = 2\pi(k-1)/10^3$, $k = 1:10^3$
- display the spectrum of $A + \eta W^{\mathcal{S}}_{\hat{\jmath}}$ for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$, where $\theta_k = 2\pi(k-1)/10^3$, $k = 1:10^3$

Example 1

Consider a real tridiagonal Toeplitz matrix of order $n = 5$ with random diagonal and superdiagonal entries in the interval $[0, 1]$, and random subdiagonal entries in the interval $[0, 5]$.

This gives a matrix with fairly ill-conditioned eigenvalues. Indeed, the sensitivity of the eigenvalues grows exponentially with the ratio of the absolute values of the sub- and super-diagonal matrix entries [S.N., Pasquini and Reichel 2013].

The eigenvalues for a tridiagonal Toeplitz matrix of this kind and their standard and structured condition numbers are shown in Table 1.

Table 1

i	λ_i	$\kappa(\lambda_i)$	$\kappa^{\mathcal{T}}(\lambda_i)$
1	-0.4988	$1.153 \cdot 10^2$	$2.625 \cdot 10^0$
2	0.0564	$3.269 \cdot 10^2$	$1.559 \cdot 10^0$
3	0.8147	$4.243 \cdot 10^2$	$4.472 \cdot 10^{-1}$
4	1.5731	$3.269 \cdot 10^2$	$1.559 \cdot 10^0$
5	2.1283	$1.153 \cdot 10^2$	$2.625 \cdot 10^0$

Table: Example 1: Eigenvalue condition numbers. While the eigenvalues in the middle of the spectrum are the worst conditioned with respect to unstructured perturbations, the extremal eigenvalues are most sensitive to structured perturbations [S.N., Pasquini and Reichel 2013].

Approximated $\Lambda_{\varepsilon_1}(A)$. Figure 1

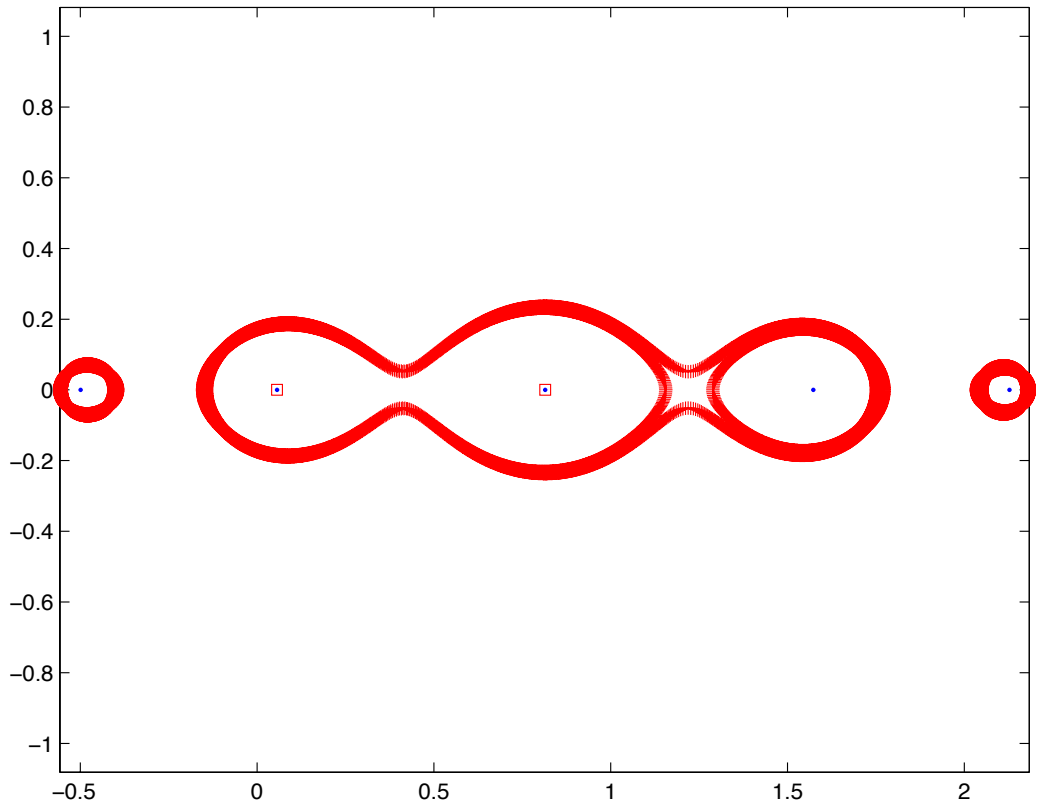


Figure: Eigenvalues of matrices of the form $A + \varepsilon_1 \eta W_2$ and $A + \varepsilon_1 \eta W_3$, where the W_j are Wilkinson perturbations associated with the eigenvalues λ_j, $j = 2, 3$ (marked by red squares), for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$, $\theta_k := 2\pi(k-1)/10^3$, $k = 1 : 10^3$, and $\varepsilon_1 = 10^{-3.2}$.

Approximated $\Lambda_{\varepsilon_1}(A)$. Figure 2

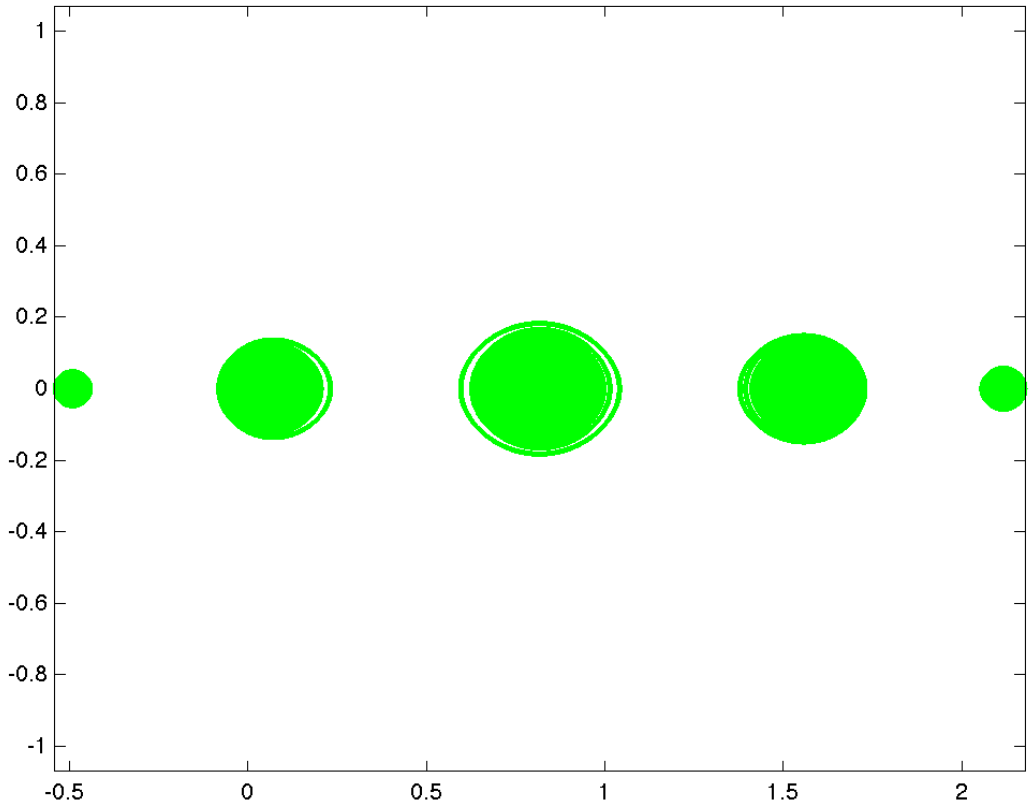


Figure: Eigenvalues of matrices of the form $A + \varepsilon_1 \mathrm{e}^{\mathrm{i}\theta_k} E_i$, with $\varepsilon_1 = 10^{-3.2}$, $\theta_k := 2\pi(k-1)/10^3$, $i, k = 1 : 10^3$. The E_i are unit-norm rank-one random perturbations.

ε-pseudospectra by Eigtool

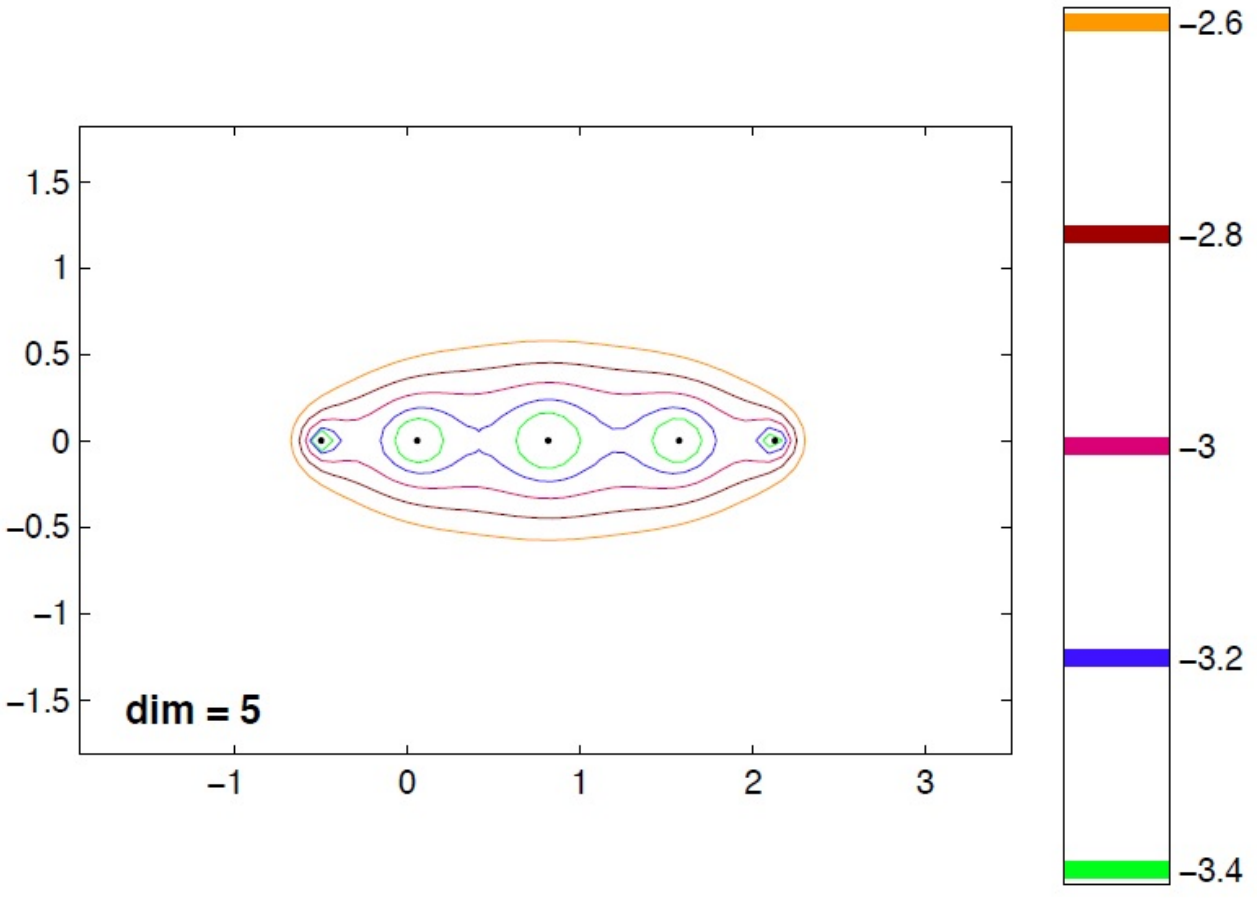


Figure: $\varepsilon = 10^k$, $k = -3.4 : 0.2 : -2.6$.

Comparing the figures

The estimate (1) of the (unstructured) distance from defectivity ε_* is $\varepsilon_1 = 10^{-3.2}$. It is achieved for the indices 2 and 3 of the most Λ_ε-sensitive pair of eigenvalues.

Comparing the ε_1-pseudospectrum given by Eigtool [Wright 2002] with Figure 1 illustrates the effectiveness of our simple approach.

In particular, the approximated ε_1-pseudospectrum of Figure 1 provides a much better approximation of the ε_1-pseudospectrum than the approximated ε_1-pseudospectrum of Figure 2 and requires the computation of many fewer spectra (10^3 versus 10^6). This makes our approach considerably faster.

We remark that Eigtool uses the spectral norm $\|\cdot\|$, while we apply the Frobenius norm for the matrices E_i. Since the E_i are of rank one, they have the same spectral and Frobenius norms.

Approximated $\Lambda_{\varepsilon_2}^{\mathcal{T}}(A)$. Figure 3

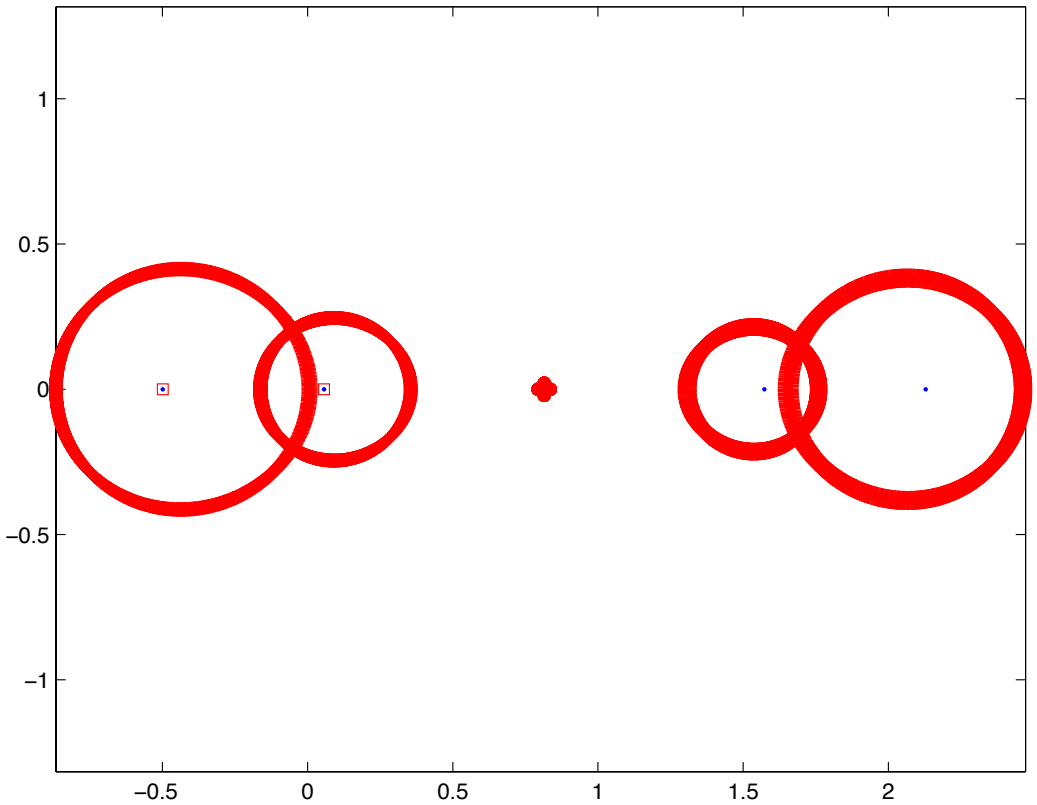


Figure: Eigenvalues of matrices of the form $A + \varepsilon_2 W_1^{\mathcal{T}}$ and $A + \varepsilon_2 W_2^{\mathcal{T}}$, for $\eta := \mathrm{e}^{\mathrm{i}\theta_k}$ and $\theta_k := 2\pi(k-1)/10^3$, $k = 1 : 10^3$, and $\varepsilon_2 = 10^{-0.8}$. The eigenvalues λ_j, $j = 1, 2$ are marked by red squares.

Approximated $\Lambda^{\mathcal{T}}_{\varepsilon_2}(A)$. Figure 4

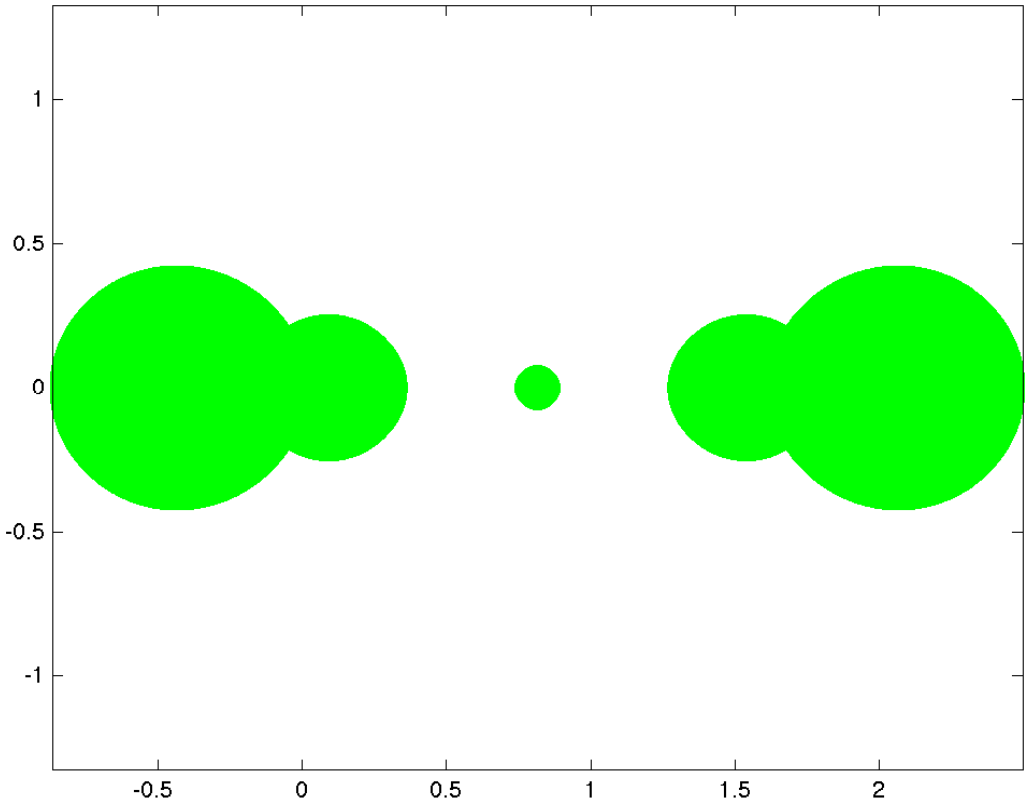


Figure: Eigenvalues of matrices of the form $A + \varepsilon_2 \mathrm{e}^{\mathrm{i}\theta_k} E_i^{\mathcal{T}}$, with $\varepsilon_2 = 10^{-0.8}$, $\theta_k := 2\pi(k-1)/10^3$. The $E_i^{\mathcal{T}}$ are unit-norm projected random perturbations in $\mathcal{T}$.

Comparing the structured figures

We follow Algorithm 2 with $\mathcal{S} = \mathcal{T}$, where $\mathcal{T}$ is the space of tridiagonal Toeplitz matrices of order 5. One obtains from (2) the estimate $\varepsilon_2 = 10^{-0.8}$ of the structured distance from defectivity $\varepsilon_*^{\mathcal{T}}$. It is achieved for the eigenvalues λ_1 and λ_2.

Figure 3 displays the spectra of matrices of the form $A + \varepsilon_2 W_1^{\mathcal{T}}$ and $A + \varepsilon_2 W_2^{\mathcal{T}}$, where $W_1^{\mathcal{T}} = W_1|_{\widehat{\mathcal{T}}}$ and $W_2^{\mathcal{T}} = W_2|_{\widehat{\mathcal{T}}}$ are the normalized projected Wilkinson perturbations onto $\mathcal{T}$.

Figure 4 displays the spectra of the matrices $A + \varepsilon_2 \mathrm{e}^{\mathrm{i}\theta_k} E_i^{\mathcal{T}}$, $i, k = 1 : 10^3$, where the $E_i^{\mathcal{T}}$ are random tridiagonal Toeplitz matrices scaled so that $\|E_i^{\mathcal{T}}\|_F = 1$.

Eigtool cannot be applied to determine structured pseudospectra.

Example 2

Consider a pentadiagonal Toeplitz matrix of order $n = 10$. Traditional and structured eigenvalue condition numbers are shown below. The estimate (1) of the distance to defectivity is $\varepsilon_1 = 10^{-2}$; it is achieved for the indices 9 and 10.

i	λ_i	$\kappa(\lambda_i)$	$\kappa^{\mathcal{T}}(\lambda_i)$
1	$5.4616 + 6.5356\,\mathrm{i}$	$1.039 \cdot 10^{1}$	$1.169 \cdot 10^{-1}$
2	$3.8552 + 5.1268\,\mathrm{i}$	$2.999 \cdot 10^{1}$	$8.646 \cdot 10^{-1}$
3	$1.7072 + 3.1264\,\mathrm{i}$	$5.643 \cdot 10^{1}$	$5.665 \cdot 10^{-1}$
4	$-3.9451 - 0.1224\,\mathrm{i}$	$1.534 \cdot 10^{1}$	$1.250 \cdot 10^{0}$
5	$-0.7339 - 3.2688\,\mathrm{i}$	$2.528 \cdot 10^{0}$	$8.553 \cdot 10^{-1}$
6	$0.3809 - 2.2234\,\mathrm{i}$	$4.908 \cdot 10^{0}$	$6.596 \cdot 10^{-1}$
7	$2.4409 - 0.7300\,\mathrm{i}$	$2.373 \cdot 10^{0}$	$8.623 \cdot 10^{-1}$
8	$1.5110 - 1.0247\,\mathrm{i}$	$8.071 \cdot 10^{0}$	$7.491 \cdot 10^{-1}$
9	$-2.2354 + 0.4417\,\mathrm{i}$	$5.207 \cdot 10^{1}$	$9.748 \cdot 10^{-1}$
10	$-0.2952 + 1.1966\,\mathrm{i}$	$7.775 \cdot 10^{1}$	$3.750 \cdot 10^{-1}$

Table: Example 2: Eigenvalue condition numbers.

Approximated $\Lambda_{\varepsilon_1}(A)$.

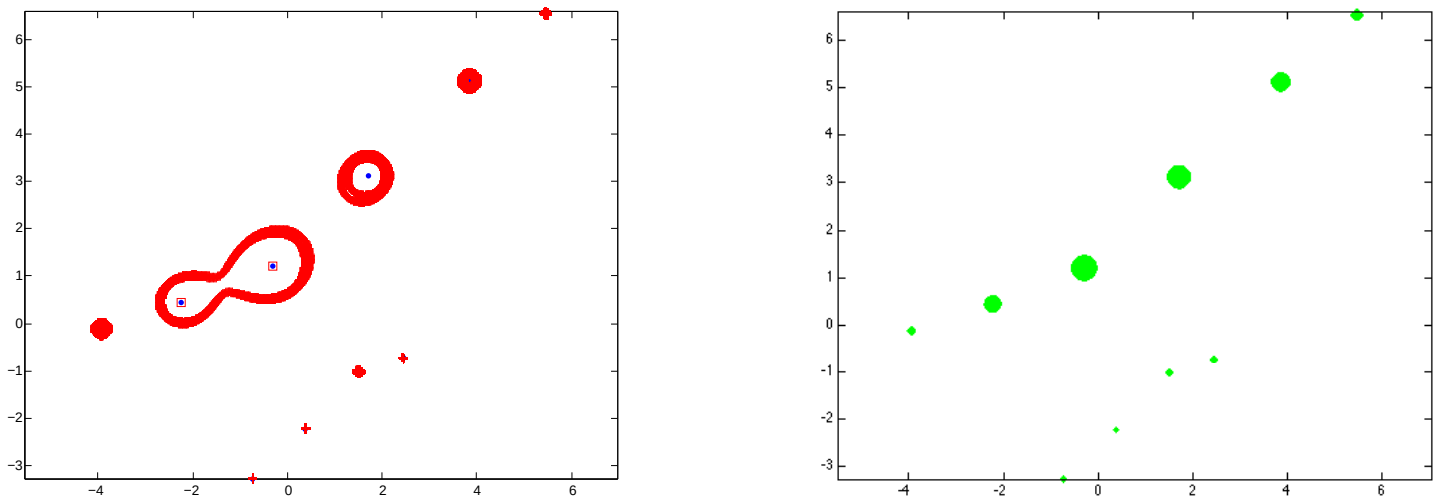


Figure: Our approach versus the random approach.

ε-pseudospectra by Eigtool

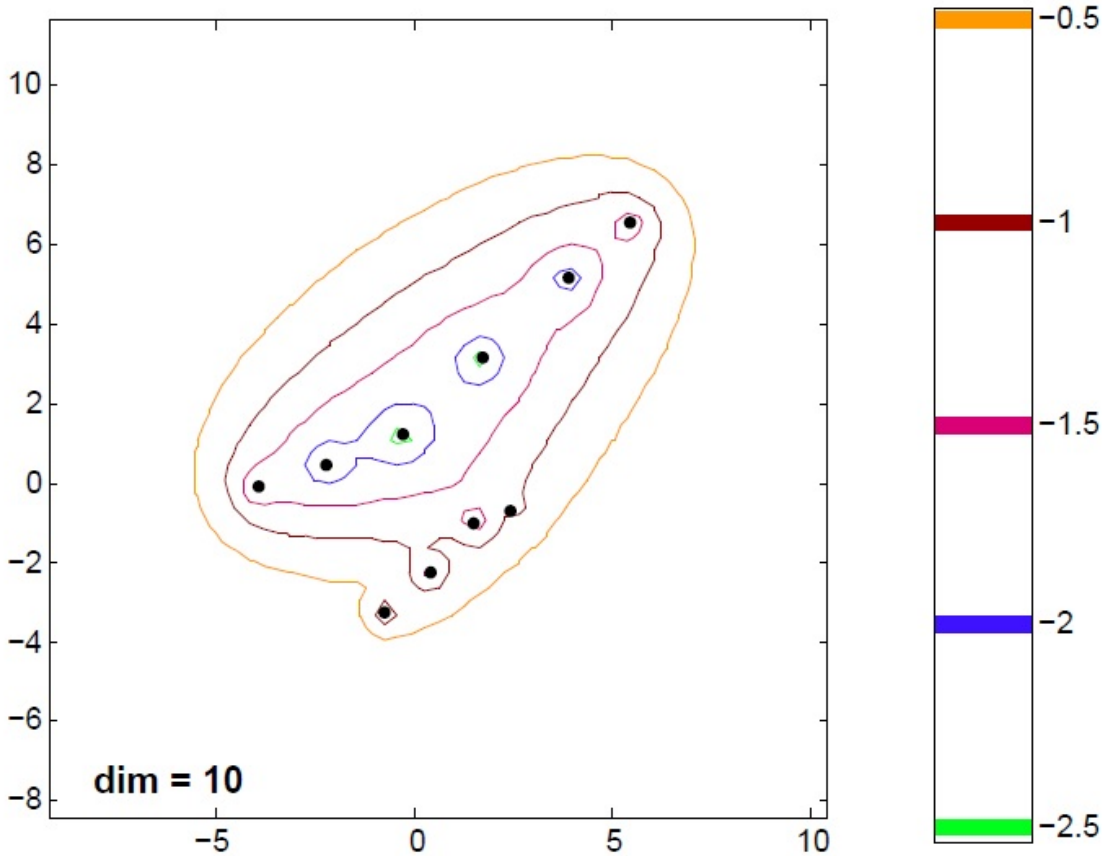


Figure: $\varepsilon = 10^k$, $k = -2.5 : 0.5 : -0.5$.

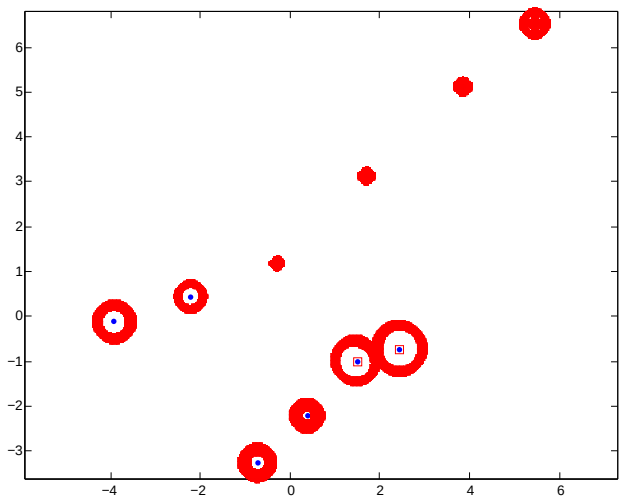
Exploiting the structure

Comparing with the ε_1-pseudospectrum determined by Eigtool shows that our simple computations can give more accurate approximations of pseudospectra and require less computational effort.

We now consider structured perturbations and pseudospectra. Let $\mathcal{T}$ be the space of pentadiagonal Toeplitz matrices of order 10. We obtain from (2) the estimate $\varepsilon_2 = 10^{-0.2}$ of the structured distance from defectivity $\varepsilon_*^{\mathcal{T}}$. It is achieved for the eigenvalues λ_7 and λ_8.

According to Table 2, the most $\Lambda_{\varepsilon_2}^{\mathcal{T}}$-sensitive pair of eigenvalues do not have the largest structured condition numbers.

Approximated $\Lambda^{\mathcal{T}}_{\varepsilon_2}(A)$.



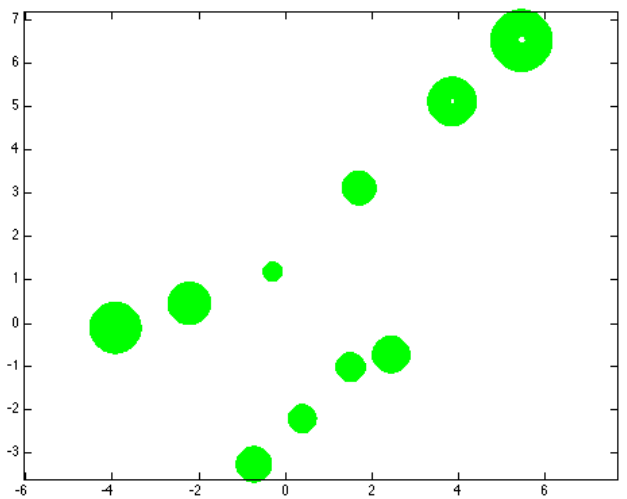


Figure: Our structured approach versus the random structured approach.

Thank you for your attention!