Backward error, condition numbers, and pseudospectra for the multiparameter eigenvalue problem

Michiel E. Hochstenbach [a], Bor Plestenjak [b,*,1]

[a] *Department of Mathematics, Utrecht University, PO Box 80 010, 3508 TA Utrecht, The Netherlands.*

[b] *Department of Mathematics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia.*

Abstract

We define and evaluate the normwise backward error and condition numbers for the multiparameter eigenvalue problem (MEP). The pseudospectrum for the MEP is defined and characterized. We show that the distance from a right definite MEP to the closest non right definite MEP is related to the smallest unbounded pseudospectrum. Some numerical results are given.

AMS classification: 65F15, 15A18, 15A69

Key words: Multiparameter eigenvalue problem, Right definiteness, Backward error, Condition number, Pseudospectrum, Nearness problem

---

1 Introduction

We study the backward error, condition numbers and pseudospectra for the multiparameter eigenvalue problem (MEP)

$$W_i(\boldsymbol{\lambda})x_i = 0, \quad 0 \neq x_i \in \mathbb{C}^{n_i}, \quad i = 1, \dots, k, \tag{1}$$

* Corresponding author.
Email addresses: `hochstenbach@math.uu.nl` (Michiel E. Hochstenbach), `bor.plestenjak@fmf.uni-lj.si` (Bor Plestenjak).

1 Supported in part by the Ministry of Education, Science, and Sport of Slovenia.

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \mathbb{C}^k$,

$$W_i(\boldsymbol{\lambda}) = V_{i0} - \sum_{j=1}^{k} \lambda_j V_{ij},$$

and V_{ij} are $n_i \times n_i$ matrices over $\mathbb{C}$. We will abbreviate the MEP (1) by $\boldsymbol{W}$. For $k = 1$, a MEP is a generalized eigenvalue problem $V_{10}x_1 = \lambda_1 V_{11} x_1$.

A k-tuple $\boldsymbol{\lambda}$ that satisfies (1) is called an *eigenvalue* and the tensor product $\boldsymbol{x} = x_1 \otimes \cdots \otimes x_k$ is the corresponding *right eigenvector*. A *left eigenvector* corresponding to the eigenvalue $\boldsymbol{\lambda}$ is $\boldsymbol{y} = y_1 \otimes \cdots \otimes y_k$, where $0 \neq y_i \in \mathbb{C}^{n_i}$ and $y_i^* W_i(\boldsymbol{\lambda}) = 0$ for $i = 1, \ldots, k$.

The backward error and condition numbers are important tools in numerical linear algebra that reveal the quality and sensitivity of numerical solutions. The theory of backward error and conditioning for eigenproblems is well developed for the generalized eigenvalue problem (see, e.g., [8]) and the polynomial eigenvalue problem (see, e.g., [5,11]).

Multiparameter eigenvalue problems arise in a variety of applications [1], particularly in mathematical physics when the method of separation of variables is used to solve boundary value problems [15]. The result of the separation is a multiparameter system of ordinary differential equations.

To a MEP (1) which satisfies a certain regularity condition, a k-tuple of commuting linear transformations on a tensor product space is associated, as follows. The tensor product space $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k}$ is isomorphic to $\mathbb{C}^N$, where $N = n_1 \cdots n_k$. The linear transformations V_{ij} induce linear transformations $V_{ij}^\dagger$ on $\mathbb{C}^N$. For a decomposable tensor,

$$V_{ij}^\dagger(x_1 \otimes \cdots \otimes x_k) = x_1 \otimes \cdots \otimes V_{ij}x_i \otimes \cdots \otimes x_k.$$

$V_{ij}^\dagger$ is then extended to all of $\mathbb{C}^N$ by linearity. On $\mathbb{C}^N$ we define operator determinants

$$\Delta_0 = \begin{vmatrix} V_{11}^\dagger & V_{12}^\dagger & \cdots & V_{1k}^\dagger \\ V_{21}^\dagger & V_{22}^\dagger & \cdots & V_{2k}^\dagger \\ \vdots & \vdots & & \vdots \\ V_{k1}^\dagger & V_{k2}^\dagger & \cdots & V_{kk}^\dagger \end{vmatrix}$$

and

$$\Delta_i = \begin{vmatrix} V_{11}^\dagger & \cdots & V_{1,i-1}^\dagger & V_{10}^\dagger & V_{1,i+1}^\dagger & \cdots & V_{1k}^\dagger \\ V_{21}^\dagger & \cdots & V_{2,i-1}^\dagger & V_{20}^\dagger & V_{2,i+1}^\dagger & \cdots & V_{2k}^\dagger \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ V_{k1}^\dagger & \cdots & V_{k,i-1}^\dagger & V_{k0}^\dagger & V_{k,i+1}^\dagger & \cdots & V_{kk}^\dagger \end{vmatrix}$$

for $i = 1, \ldots, k$.

A MEP is called *nonsingular* if the corresponding operator determinant Δ_0 is invertible. A nonsingular MEP is equivalent to the associated problem

$$\Delta_i \boldsymbol{x} = \lambda_i \Delta_0 \boldsymbol{x}, \quad i = 1, \ldots, k, \tag{2}$$

for decomposable tensors $\boldsymbol{x} = x_1 \otimes \cdots \otimes x_k \in \mathbb{C}^N$, where the matrices $\Gamma_i := \Delta_0^{-1}\Delta_i$ commute for $i = 1, \ldots, k$ (see [2]).

If $\boldsymbol{\lambda}$ is an eigenvalue of $\boldsymbol{W}$ then

$$d_a := \dim \left(\bigcap_{\substack{j_1+\cdots+j_k=N \\ j_1,\ldots,j_k \geq 0}} \operatorname{Ker} \left[(\Gamma_1 - \lambda_1 I)^{j_1} \cdots (\Gamma_k - \lambda_k I)^{j_k} \right] \right)$$

is the *algebraic multiplicity* and

$$d_g := \dim \left(\bigcap_{i=1}^{k} \operatorname{Ker} (\Gamma_i - \lambda_i I) \right) = \prod_{i=1}^{k} \dim \left(\operatorname{Ker} W_i(\boldsymbol{\lambda}) \right)$$

is the *geometric multiplicity* of the eigenvalue (see [2]). We say that an eigenvalue $\boldsymbol{\lambda}$ is *geometrically* or *algebraically simple* when $d_g = 1$ or $d_a = 1$, respectively. It can be seen that $d_a \geq d_g$, so an eigenvalue that is algebraically simple is also geometrically simple.

Let $\boldsymbol{\lambda}$ be an eigenvalue of $\boldsymbol{W}$ with the corresponding left and right eigenvectors $\boldsymbol{x}$ and $\boldsymbol{y}$. We form a $k \times k$ matrix

$$B_0 = \begin{bmatrix} y_1^* V_{11} x_1 & y_1^* V_{12} x_1 & \cdots & y_1^* V_{1k} x_1 \\ y_2^* V_{21} x_2 & y_2^* V_{22} x_2 & \cdots & y_2^* V_{2k} x_2 \\ \vdots & \vdots & & \vdots \\ y_k^* V_{k1} x_k & y_k^* V_{k2} x_k & \cdots & y_k^* V_{kk} x_k \end{bmatrix}. \tag{3}$$

The following lemma is a consequence of [9, Lemma 3].

Lemma 1 *If $\boldsymbol{\lambda}$ is an algebraically simple eigenvalue of the multiparameter eigenvalue problem $\boldsymbol{W}$ then B_0 is nonsingular.*

A MEP is called *Hermitian* when all matrices V_{ij} are Hermitian. Furthermore, a Hermitian MEP is called *right definite* if

$$\begin{vmatrix} x_1^* V_{11} x_1 & x_1^* V_{12} x_1 & \cdots & x_1^* V_{1k} x_1 \\ x_2^* V_{21} x_2 & x_2^* V_{22} x_2 & \cdots & x_2^* V_{2k} x_2 \\ \vdots & \vdots & & \vdots \\ x_k^* V_{k1} x_k & x_k^* V_{k2} x_k & \cdots & x_k^* V_{kk} x_k \end{vmatrix} \geq \delta \tag{4}$$

for all vectors $x_i \in \mathbb{C}^{n_i}$, $\|x_i\| = 1$, $i = 1, \ldots, k$, and some $\delta > 0$. Condition (4) is equivalent to the positive definiteness of Δ_0 [2, Theorem 7.8.2]. This implies that if $\boldsymbol{W}$ is right definite then there exist N linearly independent eigenvectors. If $\boldsymbol{\lambda}$ is an eigenvalue of a right definite problem $\boldsymbol{W}$ then $\boldsymbol{\lambda} \in \mathbb{R}^k$. Furthermore, if all matrices V_{ij} of a right definite problem $\boldsymbol{W}$ are real, then the eigenvectors can be chosen real. For a real geometrically simple eigenvalue of a Hermitian MEP the corresponding left and right eigenvectors coincide.

After preliminaries in Section 2, we study the backward error in Section 3. The condition numbers for eigenvalues and eigenvectors are discussed in Section 4. The pseudospectra, examined in Section 5, are another valuable tool for the study of the sensitivity of eigenvalues to perturbations of matrices. In Section 6, we give some numerical experiments for right definite two-parameter eigenvalue problems, where pseudospectra can be visualized in $\mathbb{R}^2$.

2 Preliminaries

Throughout the paper we assume that the MEP $\boldsymbol{W}$ is nonsingular. The matrices E_{ij} for $i = 1, \ldots, k$; $j = 0, \ldots, k$ represent tolerances for the perturbations ΔV_{ij} of V_{ij}, defined by $\|\Delta V_{ij}\| \leq \varepsilon \|E_{ij}\|$ for some $\varepsilon > 0$. Usually we take either $E_{ij} = V_{ij}$ considering *normwise relative perturbations*, or $E_{ij} = I$ considering *normwise absolute perturbations. Elementwise perturbations* $|\Delta V_{ij}| \leq \varepsilon |E_{ij}|$ can also be considered, see Remark 5. We define

$$\Delta W_i(\boldsymbol{\lambda}) := \Delta V_{i0} - \sum_{j=1}^{k} \lambda_j \Delta V_{ij}.$$

We will denote the perturbed MEP with matrices $V_{ij} + \Delta V_{ij}$ by $\boldsymbol{W} + \Delta \boldsymbol{W}$. For a complex λ the *sign of* λ is defined as (cf. [8, p. 495])

$$\operatorname{sign}(\lambda) := \begin{cases} \overline{\lambda}/|\lambda|, & \lambda \neq 0 \\ 0, & \lambda = 0. \end{cases}$$

Suppose that we are looking for the maximum Euclidean norm of Az where $A \in \mathbb{C}^{k\times k}$ and $z \in \mathbb{C}^k$ is such that $|z_i| \le \theta_i$ for $i = 1,\ldots,k$, where $\theta_1,\ldots,\theta_k$ are given positive constants. According to Bauer's maximum principle (both the function $\|\cdot\|$ and its domain are convex), the maximum is attained by z for which $|z_i| = \theta_i$ for $i = 1,\ldots,k$. For $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_n]^T$ we define the $\boldsymbol{\theta}$*-weighted norm of* A as

$$\|A\|_{\boldsymbol{\theta}} := \max\{\, \|Az\|_2 \ : \ z \in \mathbb{C}^k,\ |z_i| = \theta_i \text{ for } i = 1,\ldots,k \,\}. \tag{5}$$

Clearly,

$$\|A\|_{\boldsymbol{\theta}} \le \|A\|_2 \|\boldsymbol{\theta}\|_2. \tag{6}$$

One may verify that $\|\cdot\|_{\boldsymbol{\theta}}$ is indeed a matrix norm. One may also see that $\|\cdot\|_{\boldsymbol{\theta}}$ is not a consistent norm as it does not necessarily satisfy $\|AB\|_{\boldsymbol{\theta}} \le \|A\|_{\boldsymbol{\theta}}\|B\|_{\boldsymbol{\theta}}$ (for a counterexample, take $A = B = I$ and $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta}\|_2 < 1$).

From now on, $\|\cdot\|$ stands for $\|\cdot\|_2$. We say that a decomposable tensor $\boldsymbol{z} = z_1 \otimes \cdots \otimes z_k$ is *normalized* if $\|z_i\| = 1$ for $i = 1,\ldots,k$. From $\|\boldsymbol{z}\| = \|z_1\| \cdots \|z_k\|$ it follows that $\|\boldsymbol{z}\| = 1$. In this paper we will assume that the eigenvectors are normalized.

3 Backward error

Let $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}})$ be an approximate eigenpair of $\boldsymbol{W}$ and let $\widetilde{\boldsymbol{x}}$ be normalized. We define the *normwise backward error of* $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}})$ by

$$\begin{aligned} \eta(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}}) := \min\{\ & \varepsilon : (W_i(\widetilde{\boldsymbol{\lambda}}) + \Delta W_i(\widetilde{\boldsymbol{\lambda}}))\widetilde{x}_i = 0, \\ & \|\Delta V_{ij}\| \le \varepsilon \|E_{ij}\|,\ i = 1,\ldots,k;\ j = 0,\ldots,k\ \}. \end{aligned} \tag{7}$$

The following theorem is a generalization of the backward errors for the case $k = 1$ given in [7, Lemma 2.1] and [8, Theorem 2.1].

Theorem 2 *For the normwise backward error* $\eta(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}})$ *we have*

$$\eta(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}}) = \max_{i=1,\ldots,k} \frac{\|r_i\|}{\widetilde{\theta}_i}, \tag{8}$$

where $r_i := W_i(\widetilde{\boldsymbol{\lambda}})\widetilde{x}_i$ *is the residual and*

$$\widetilde{\theta}_i := \|E_{i0}\| + \sum_{j=1}^{k} |\widetilde{\lambda}_j| \|E_{ij}\|$$

for $i = 1, \ldots, k$*.*

Proof. From $r_i = -\Delta W_i(\widetilde{\boldsymbol{\lambda}})\widetilde{x}_i$ it follows that $\|r_i\| \le \widetilde{\theta}_i\varepsilon$ for $i = 1, \ldots, k$. Therefore, the right-hand side of (8) is a lower bound for $\eta(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}})$. The lower bound is attained for the perturbations

$$\Delta V_{i0} = -\frac{1}{\widetilde{\theta}_i}\|E_{i0}\| r_i \widetilde{x}_i^*, \quad \Delta V_{ij} = \frac{\operatorname{sign}(\widetilde{\lambda}_j)}{\widetilde{\theta}_i}\|E_{ij}\| r_i \widetilde{x}_i^*$$

for $i, j = 1, \ldots, k$. $\square$

If $\boldsymbol{W}$ is Hermitian then it is of interest to consider a backward error in which the perturbations ΔV_{ij} are Hermitian. The *backward error for a Hermitian MEP* can be defined as

$$\begin{aligned}\eta_{\mathrm{H}}(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}}) := \min\{\ \varepsilon : (W_i(\widetilde{\boldsymbol{\lambda}}) + \Delta W_i(\widetilde{\boldsymbol{\lambda}}))\widetilde{x}_i = 0,\ \Delta V_{ij}^* = \Delta V_{ij},\\ \|\Delta V_{ij}\| \le \varepsilon \|E_{ij}\|,\ i = 1, \ldots, k;\ j = 0, \ldots, k\}.\end{aligned} \tag{9}$$

It is clear that $\eta_{\mathrm{H}}(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}}) \ge \eta(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}})$ and that the optimal perturbations in (7) are not Hermitian in general. The next lemma, which is is a generalization of [8, Lemma 2.6], shows that in the case when $\widetilde{\boldsymbol{\lambda}}$ is real requiring the perturbations to be Hermitian has no effect on the backward error.

Theorem 3 *If* $\boldsymbol{W}$ *is Hermitian and* $\widetilde{\boldsymbol{\lambda}}$ *is real then*

$$\eta_{\mathrm{H}}(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}}) = \eta(\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{\lambda}}). \tag{10}$$

Proof. Let $r_i = W_i(\widetilde{\boldsymbol{\lambda}})\widetilde{x}_i$. It follows from $\widetilde{\boldsymbol{\lambda}}$ being real that $\widetilde{x}_i^* r_i$ is real. We are looking for a Hermitian matrix S_i such that $S_i\widetilde{x}_i = -r_i$. We take $S_i = \|r_i\| I$

if r_i is a negative multiple of $\widetilde{x}_i$; otherwise we take $S_i = \|r_i\| H_i$ where H_i is a Householder matrix that maps $\widetilde{x}_i$ to $-r_i/\|r_i\|$. Such an H_i exists because $\widetilde{x}_i^* r_i$ is real and is equal to $I - 2(w_i^* w_i)^{-1} w_i w_i^*$, where $w_i = \widetilde{x}_i + r_i/\|r_i\|$.

Let ΔV_{ij} be Hermitian matrices defined by

$$\Delta V_{i0} = \frac{1}{\widetilde{\theta}_i}\|E_{i0}\| H_i, \quad \Delta V_{ij} = -\frac{1}{\widetilde{\theta}_i}\,\mathrm{sign}(\widetilde{\lambda}_j)\|E_{ij}\| H_i \tag{11}$$

for $i, j = 1, \ldots, k$. It follows that $\Delta W_i(\widetilde{\boldsymbol{\lambda}}) = S_i$ and the first constraint in (9) is satisfied. Using (8), we get

$$\|S_i\| = \|r_i\| \le \eta(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}})\widetilde{\theta}_i$$

for $i = 1, \ldots, k$. From (11) we deduce $\eta_{\mathrm{H}}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}}) \le \eta(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}})$. Since $\eta_{\mathrm{H}}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}}) \ge \eta(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}})$ by definition, equality (10) must hold. $\square$

We remark that one can see from $\widetilde{x}_i^* S_i \widetilde{x}_i = -\widetilde{x}_i r_i$ that a Hermitian matrix S_i such that $S_i \widetilde{x}_i = -\widetilde{x}_i r_i$ exists only when $\widetilde{x}_i^* r_i$ is real. This is the reason why Lemma 3 cannot be generalized for nonreal approximations $\widetilde{\boldsymbol{\lambda}}$. As it is reasonable to assume that $\widetilde{\boldsymbol{\lambda}}$ is real if $\boldsymbol{\lambda}$ is real, Lemma 3 can also be applied for a right definite MEP.

If we are interested only in the approximate eigenvalue $\widetilde{\boldsymbol{\lambda}}$, then a more appropriate measure of the backward error may be

$$\eta(\widetilde{\boldsymbol{\lambda}}) := \min\{\, \eta(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{\lambda}}) \ : \ \widetilde{\boldsymbol{x}} \text{ normalized} \,\}.$$

Proposition 4

$$\eta(\widetilde{\boldsymbol{\lambda}}) = \max_{i=1,\ldots,k} \frac{1}{\widetilde{\theta}_i}\, \sigma_{\min}(W_i(\widetilde{\boldsymbol{\lambda}})).$$

Proof. The result follows from Theorem 2 by using the equality

$$\min_{\|x\|=1} \|Ax\| = \sigma_{\min}(A). \qquad \square$$

Remark 5 Although in this paper we do not consider componentwise backward errors, componentwise results from [8] can be generalized as well.

4 Condition numbers

In this section, we assume that $\boldsymbol{\lambda}$ is a nonzero algebraically simple eigenvalue of a nonsingular MEP $\boldsymbol{W}$ with corresponding normalized right eigenvector $\boldsymbol{x}$ and left eigenvector $\boldsymbol{y}$.

4.1 Eigenvalue condition number

A *normwise condition number of* $\boldsymbol{\lambda}$ can be defined by

$$\begin{aligned}\kappa(\boldsymbol{\lambda}, \boldsymbol{W}) := \limsup_{\varepsilon \to 0} \Bigg\{ & \frac{\|\Delta \boldsymbol{\lambda}\|}{\varepsilon} : \\ & \Big(V_{i0} + \Delta V_{i0} - \sum_{j=1}^{k} (\lambda_j + \Delta\lambda_j)(V_{ij} + \Delta V_{ij}) \Big) (x_i + \Delta x_i) = 0, \\ & \|\Delta V_{ij}\| \leq \varepsilon \|E_{ij}\|, \ i = 1, \ldots, k; \ j = 0, \ldots, k \Bigg\}. \end{aligned} \tag{12}$$

The following results can be considered as generalizations of the theory in [8, Section 2.2].

Theorem 6 *The condition number* $\kappa(\boldsymbol{\lambda}, \boldsymbol{W})$ *is given by*

$$\kappa(\boldsymbol{\lambda}, \boldsymbol{W}) = \|B_0^{-1}\|_{\boldsymbol{\theta}}, \tag{13}$$

where

$$\theta_i := \|E_{i0}\| + \sum_{j=1}^{k} |\lambda_j| \|E_{ij}\|$$

for $i = 1, \ldots, k$, *and* $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_k]^T$.

Proof. If we expand the equality constraints in (12) and keep only the first order terms then we get

$$\Delta W_i(\boldsymbol{\lambda}) x_i + \sum_{j=1}^{k} \Delta\lambda_j V_{ij} x_i + W_i(\boldsymbol{\lambda}) \Delta x_i = \mathcal{O}(\varepsilon^2). \tag{14}$$

Premultiplying by y_i^* yields

$$y_i^* \Delta W_i(\boldsymbol{\lambda}) x_i + y_i^* \sum_{j=1}^{k} \Delta\lambda_j V_{ij} x_i = \mathcal{O}(\varepsilon^2)$$

for $i = 1, \ldots, k$. By rearranging the equations we obtain the linear system

$$\begin{bmatrix} y_1^* V_{11} x_1 & \cdots & y_1^* V_{1k} x_1 \\ \vdots & & \vdots \\ y_k^* V_{k1} x_k & \cdots & y_k^* V_{kk} x_k \end{bmatrix} \begin{bmatrix} \Delta\lambda_1 \\ \vdots \\ \Delta\lambda_k \end{bmatrix} = \begin{bmatrix} y_1^* \Delta W_1(\boldsymbol{\lambda}) x_1 \\ \vdots \\ y_k^* \Delta W_k(\boldsymbol{\lambda}) x_k \end{bmatrix} + \mathcal{O}(\varepsilon^2),$$

or in shorter form

$$B_0 \Delta\boldsymbol{\lambda} = \begin{bmatrix} y_1^* \Delta W_1(\boldsymbol{\lambda}) x_1 \\ \vdots \\ y_k^* \Delta W_k(\boldsymbol{\lambda}) x_k \end{bmatrix} + \mathcal{O}(\varepsilon^2).$$

Since $\boldsymbol{\lambda}$ is an algebraically simple eigenvalue, it follows from Lemma 1 that B_0 is nonsingular. Thus,

$$\Delta\boldsymbol{\lambda} = B_0^{-1} \begin{bmatrix} y_1^* \Delta W_1(\boldsymbol{\lambda}) x_1 \\ \vdots \\ y_k^* \Delta W_k(\boldsymbol{\lambda}) x_k \end{bmatrix} + \mathcal{O}(\varepsilon^2)$$

and we conclude

$$\|\Delta\boldsymbol{\lambda}\| \le \|B_0^{-1}\|_{\varepsilon\boldsymbol{\theta}} + \mathcal{O}(\varepsilon^2) = \varepsilon \|B_0^{-1}\|_{\boldsymbol{\theta}} + \mathcal{O}(\varepsilon^2).$$

Hence, the expression in (13) is an upper bound for the condition number. To show that this bound can be attained we consider the matrices

$$\Delta V_{i0} = \varepsilon \|E_{i0}\| y_i x_i^*, \quad \Delta V_{ij} = -\mathrm{sign}(\widetilde{\lambda}_j) \varepsilon \|E_{ij}\| y_i x_i^*$$

for $i, j = 1, \ldots, k$. □

As for the backward error, if the MEP $\boldsymbol{W}$ is Hermitian then it is natural to restrict the perturbations ΔV_{ij} in (12) to be Hermitian. We denote

$$\kappa_{\mathrm{H}}(\boldsymbol{\lambda}, \boldsymbol{W}) := \limsup_{\varepsilon \to 0} \left\{ \frac{\|\Delta\boldsymbol{\lambda}\|}{\varepsilon} : \right.$$

$$\left.\left(V_{i0}+\Delta V_{i0}-\sum_{j=1}^{n}(\lambda_j+\Delta\lambda_j)(V_{ij}+\Delta V_{ij})\right)(x_i+\Delta x_i)=0,\quad \Delta V_{ij}^*=\Delta V_{ij},\ \|\Delta V_{ij}\|\le\varepsilon\|E_{ij}\|,\ i=1,\ldots,k;\ j=0,\ldots,k\right\}.$$

Lemma 7 *If $\boldsymbol{\lambda}$ is a real algebraically simple eigenvalue of a Hermitian multiparameter eigenvalue problem $\boldsymbol{W}$ then*

$$\kappa_{\rm H}(\boldsymbol{\lambda},\boldsymbol{W})=\kappa(\boldsymbol{\lambda},\boldsymbol{W}).$$

Proof. For a Hermitian MEP and algebraically simple eigenvalue $\boldsymbol{\lambda}$ we can take $\boldsymbol{y}=\boldsymbol{x}$ and then the matrices H_i in the proof of Theorem 6 are Hermitian. It follows that the perturbations for which the bound is attained are also Hermitian. □

As in Section 3 let us remark that Lemma 7 can also be applied to a right definite MEP.

4.2 Eigenvector condition number

In order to study the condition number of the eigenvector of an algebraically simple eigenvalue we introduce the following approach. If an eigenvector $\boldsymbol{x}=x_1\otimes\cdots\otimes x_k$ is perturbed to $\widetilde{\boldsymbol{x}}=(x_1+\Delta x_1)\otimes\cdots\otimes(x_k+\Delta x_k)$, then we are interested in $\|\operatorname{vec}(\Delta\boldsymbol{x})\|$, where

$$\operatorname{vec}(\Delta\boldsymbol{x})=[\,\Delta x_1^T\quad\cdots\quad\Delta x_k^T\,]^T$$

is a vector in $\mathbb{C}^{n_1+\cdots+n_k}$. Therefore we define a *normwise condition number of* $\boldsymbol{x}$ by

$$\kappa(\boldsymbol{x},\boldsymbol{W}):=\limsup_{\varepsilon\to0}\left\{\frac{\|\operatorname{vec}(\Delta\boldsymbol{x})\|}{\varepsilon}:\ \left(V_{i0}+\Delta V_{i0}-\sum_{j=1}^{k}(\lambda_j+\Delta\lambda_j)(V_{ij}+\Delta V_{ij})\right)(x_i+\Delta x_i)=0,\ g_i^*x_i=g_i^*(x_i+\Delta x_i)=1,\ \|\Delta V_{ij}\|\le\varepsilon\|E_{ij}\|,\ i=1,\ldots,k;\ j=0,\ldots,k\right\},\qquad(15)$$

where the vectors g_i that are used for the normalization of $\widetilde{\boldsymbol{x}}$ are such that $g_i^* x_i \neq 0$ for $i = 1, \ldots, k$ and that the matrix

$$\begin{bmatrix} g_1^* V_{11} x_1 & \cdots & g_1^* V_{1k} x_1 \\ \vdots & & \vdots \\ g_k^* V_{k1} x_k & \cdots & g_k^* V_{kk} x_k \end{bmatrix} \tag{16}$$

is nonsingular. We can for instance take $g_i = y_i$, since in this case the matrix (16) is equal to B_0, which is nonsingular for algebraically simple eigenvalues by Lemma 1.

Let $m = n_1 + \cdots + n_k$. We can combine all the equations (14) into one equation in $\mathbb{C}^m$ as

$$D \operatorname{vec}(\Delta \boldsymbol{x}) = -\operatorname{diag}(\Delta W_i(\boldsymbol{\lambda})) \operatorname{vec}(\boldsymbol{x}) - V \Delta \boldsymbol{\lambda} + \mathcal{O}(\varepsilon^2), \tag{17}$$

where

$$V = \begin{bmatrix} V_{11} x_1 & \cdots & V_{1k} x_1 \\ \vdots & & \vdots \\ V_{k1} x_k & \cdots & V_{kk} x_k \end{bmatrix}, \quad D = \begin{bmatrix} W_1(\boldsymbol{\lambda}) & & \\ & \ddots & \\ & & W_k(\boldsymbol{\lambda}) \end{bmatrix},$$

$$\operatorname{diag}(\Delta W_i(\boldsymbol{\lambda})) = \begin{bmatrix} \Delta W_1(\boldsymbol{\lambda}) & & \\ & \ddots & \\ & & \Delta W_k(\boldsymbol{\lambda}) \end{bmatrix},$$

$$\Delta \boldsymbol{\lambda} = [\Delta\lambda_1 \ \cdots \ \Delta\lambda_k]^T, \quad \text{and} \quad \operatorname{vec}(\boldsymbol{x}) = [\, x_1^T \ \cdots \ x_k^T \,]^T.$$

If we define the $m \times k$ matrix

$$G = \begin{bmatrix} g_1 & 0 & \cdots & 0 \\ 0 & g_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & g_k \end{bmatrix}$$

then $G^* V$ is equal to (16). As a result $G^* V$ is nonsingular and we can define an oblique projection

$$P = I - V (G^* V)^{-1} G^*$$

onto $\operatorname{range}(G)^\perp$ along $\operatorname{range}(V)$. It follows that $PV = 0$ and when we multiply (17) by P we obtain

$$PD \operatorname{vec}(\Delta \boldsymbol{x}) = -P \operatorname{diag}(\Delta W_i(\boldsymbol{\lambda})) \operatorname{vec}(\boldsymbol{x}) + \mathcal{O}(\varepsilon^2). \tag{18}$$

From $g_i^* \Delta x_i = 0$ for $i = 1, \ldots, k$ it follows that $G^* \operatorname{vec}(\Delta \boldsymbol{x}) = 0$ and thus $P \operatorname{vec}(\Delta \boldsymbol{x}) = \operatorname{vec}(\Delta \boldsymbol{x})$. Now we can rewrite (18) as

$$PDP \operatorname{vec}(\Delta \boldsymbol{x}) = -P \operatorname{diag}(\Delta W_i(\boldsymbol{\lambda})) \operatorname{vec}(\boldsymbol{x}) + \mathcal{O}(\varepsilon^2). \tag{19}$$

Lemma 8 *The operator T defined by $T := PDP$ is a bijection as an operator from $\mathcal{G}^\perp$ onto $\mathcal{G}^\perp$, where $\mathcal{G}^\perp := \operatorname{range}(G)^\perp$*

Proof. Since T clearly maps to $\mathcal{G}^\perp$, it is enough to show that T is injective. Suppose that there exists a $\boldsymbol{z} \in \mathcal{G}^\perp$ such that $T\boldsymbol{z} = 0$. Since $P\boldsymbol{z} = \boldsymbol{z}$, there exists an $h \in \mathbb{C}^k$ such that

$$D\boldsymbol{z} = Vh. \tag{20}$$

If we left-multiply (20) by Y^*, where Y is the $m \times k$ matrix

$$Y = \begin{bmatrix} y_1 & 0 & \cdots & 0 \\ 0 & y_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & y_k \end{bmatrix},$$

we obtain $Y^*Vh = 0$ and since $Y^*V = B_0$ is nonsingular it follows that $h = 0$. As a result we have $W_i(\boldsymbol{\lambda}) z_i = 0$ for $i = 1, \ldots, k$ where $\boldsymbol{z}$ is partitioned conformally with $\operatorname{vec}(\boldsymbol{x})$. Since $\boldsymbol{\lambda}$ is algebraically simple by assumption it follows that $\dim \operatorname{Ker} W_i(\boldsymbol{\lambda}) = 1$ and therefore $z_i = \gamma_i x_i$ for certain $\gamma_i \in \mathbb{C}$. Now we know that $G^*\boldsymbol{z} = 0$ on the one hand and on the other hand $G^*\boldsymbol{z} = [\gamma_1 \ \cdots \ \gamma_k]^T$ so $\gamma_i = 0$ for $i = 1, \ldots, k$ from which we conclude that $\boldsymbol{z} = 0$. $\square$

It follows from Lemma 8 and (19) that

$$\operatorname{vec}(\Delta \boldsymbol{x}) = \left(PDP|_{\mathcal{G}^\perp}\right)^{-1} P \operatorname{diag}(\Delta W_i(\boldsymbol{\lambda})) \operatorname{vec}(\boldsymbol{x}),$$

where $PDP|_{\mathcal{G}^\perp}$ is a restriction of PDP to $\mathcal{G}^\perp$. This gives

$$\| \operatorname{vec}(\Delta \boldsymbol{x})\| \le \varepsilon \left\| \left(PDP|_{\mathcal{G}^\perp}\right)^{-1} P \right\|_{\boldsymbol{\theta},\boldsymbol{n}} + \mathcal{O}(\varepsilon^2), \tag{21}$$

where

$$\|A\|_{\boldsymbol{\theta},\boldsymbol{n}} := \max \Big\{ \|Az\| : \boldsymbol{z} = [\, z_1^T \quad \cdots \quad z_k^T \,]^T,$$

$$z_i \in \mathbb{C}^{n_i},\ \|z_i\| \le \theta_i,\ i = 1, \ldots, k \Big\}$$

and $\boldsymbol{n} = [n_1 \ \cdots \ n_k]^T$. One can view this $\boldsymbol{\theta}, \boldsymbol{n}$-norm as a block version of (5). This leads to the next theorem.

Theorem 9

$$\kappa(\boldsymbol{x}, \boldsymbol{W}) = \left\| \left(PDP|_{\mathcal{G}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta}, \boldsymbol{n}}. \tag{22}$$

Proof. In the discussion preceding the theorem we showed in (21) that

$$\kappa(\boldsymbol{x}, \boldsymbol{W}) \le \left\| \left(PDP|_{\mathcal{G}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta}, \boldsymbol{n}}.$$

What remains is to construct a perturbation for which equality is attained.

Suppose that for $\boldsymbol{z} = [\, z_1^T \quad \cdots \quad z_k^T \,]^T$ such that $\|z_i\| \le \theta_i$ for $i = 1, \ldots, k$ we have

$$\left\| \left(PDP|_{\mathcal{G}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta}, \boldsymbol{n}} = \left\| \left(PDP|_{\mathcal{G}^\perp} \right)^{-1} P\boldsymbol{z} \right\|. \tag{23}$$

Equality in (21) is then attained if we take

$$\Delta V_{i0} = -\frac{\varepsilon \|E_{i0}\|}{\alpha_i} z_i x_i^*, \quad \Delta V_{ij} = \operatorname{sign}(\lambda_j) \frac{\varepsilon \|E_{ij}\|}{\alpha_i} z_i x_i^*$$

for $i, j = 1, \ldots, k$. $\square$

Remark 10 If we take $G = Y$ then D is a bijection as an operator from $\mathcal{Y}^\perp$ to $\mathcal{Y}^\perp$, where $\mathcal{Y} := \operatorname{range}(Y)$, and $\left\| \left(PDP|_{\mathcal{Y}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta}, \boldsymbol{n}} = \left\| P \left(D|_{\mathcal{Y}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta}, \boldsymbol{n}}$.

From (21) we can produce upper bounds for the norm of $\widetilde{\boldsymbol{x}} - \boldsymbol{x}$. If we consider only first order terms then we have

$$\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}\| \le \|\Delta x_1\| + \cdots + \|\Delta x_k\| + \mathcal{O}(\varepsilon^2)$$

and it follows that

$$\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}\| \le \sqrt{k}\, \|\operatorname{vec}(\Delta \boldsymbol{x})\| + \mathcal{O}(\varepsilon^2).$$

As we insert (21) we obtain the bound

$$\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}\| \le \sqrt{k} \left\| \left(PDP|_{\mathcal{G}^\perp} \right)^{-1} P \right\|_{\boldsymbol{\theta},\boldsymbol{n}} \varepsilon + \mathcal{O}(\varepsilon^2).$$

5 Pseudospectra

Another tool for the study of the sensitivity of the eigenvalues to perturbations are pseudospectra. They have been studied for the standard (see, e.g., [13,14]) and generalized eigenproblem [6] and for the polynomial eigenvalue problem (see, e.g., [12]). We extend the definition of pseudospectrum to the multiparameter eigenvalue problem.

We define the *ε-pseudospectrum of $\boldsymbol{W}$* by

$$\Lambda_\varepsilon(\boldsymbol{W}) = \Big\{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ W_i(\boldsymbol{\lambda}) + \Delta W_i(\boldsymbol{\lambda}) \text{ singular}, \\ \|\Delta V_{ij}\| \le \varepsilon \|E_{ij}\|, \ i = 1, \ldots, k; \ j = 0, \ldots, k \Big\}. \tag{24}$$

If we define the *ε-pseudospectrum of W_i* by

$$\Lambda_\varepsilon(W_i) = \Big\{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ W_i(\boldsymbol{\lambda}) + \Delta W_i(\boldsymbol{\lambda}) \text{ singular}, \\ \|\Delta V_{ij}\| \le \varepsilon \|E_{ij}\|, \ j = 0, \ldots, k \Big\},$$

then it is easy to see that

$$\Lambda_\varepsilon(\boldsymbol{W}) = \Lambda_\varepsilon(W_1) \cap \Lambda_\varepsilon(W_2) \cap \cdots \cap \Lambda_\varepsilon(W_k). \tag{25}$$

Theorem 11

$$\begin{aligned} \Lambda_\varepsilon(\boldsymbol{W}) &= \{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ \eta(\boldsymbol{\lambda}) \le \varepsilon \text{ for } i = 1, \ldots, k \} \\ &= \{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ \sigma_{\min}(W_i(\boldsymbol{\lambda})) \le \varepsilon \widetilde{\theta}_i \text{ for } i = 1, \ldots, k \} \\ &= \{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ \|W_i(\boldsymbol{\lambda})^{-1}\| \ge 1/(\varepsilon \widetilde{\theta}_i) \text{ for } i = 1, \ldots, k \} \\ &= \{ \boldsymbol{\lambda} \in \mathbb{C}^k \ : \ \exists u_i, \ \|u_i\| = 1, \text{ such that} \\ &\qquad \|W_i(\boldsymbol{\lambda}) u_i\| \le \varepsilon \widetilde{\theta}_i \text{ for } i = 1, \ldots, k \}. \end{aligned}$$

Proof. The first equality follows readily from the definition (24). For the second equality Proposition 4 can be applied. The last two equalities follow from the identity $\min_{x\neq 0}\|Ax\|/\|x\| = \|A^{-1}\|^{-1} = \sigma_{\min}(A)$ with the convention that $\|A^{-1}\| = \infty$ if A is singular. $\square$

Pseudospectra for the MEP have a property that is different from pseudospectra for the standard eigenvalue problem $Ax = \lambda x$: if ε is large enough then $\Lambda_\varepsilon(\boldsymbol{W})$ will be unbounded. This is the subject of the rest of this section.

If $\boldsymbol{W}$ is a right definite MEP, then we may be interested in the smallest perturbation that makes $\boldsymbol{W} + \Delta\boldsymbol{W}$ not right definite. Again, here we restrict the perturbations ΔV_{ij} to be Hermitian. We can define the distance to the closest non right definite MEP as

$$\begin{aligned}\xi(\boldsymbol{W}) := \min\{\,\varepsilon : \boldsymbol{W} + \Delta\boldsymbol{W} \text{ is not right definite}, \Delta V_{ij}^* = \Delta V_{ij},\\ \|\Delta V_{ij}\| \le \varepsilon\|E_{ij}\|,\ i = 1,\dots,k;\ j = 0,\dots,k\,\}.\end{aligned}$$

In the next theorem we show that $\xi(\boldsymbol{W})$ is bounded by the minimal ε for which the pseudospectra is unbounded.

Theorem 12

$$\xi(\boldsymbol{W}) \le \min\{\,\varepsilon : \Lambda_\varepsilon(\boldsymbol{W}) \text{ is unbounded}\,\}. \tag{26}$$

Proof. If $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_k)$ is an eigenvalue of a right definite $\boldsymbol{W}$ with corresponding normalized eigenvector $\boldsymbol{x} = x_1 \otimes \cdots \otimes x_k$ then it follows that λ_i is equal to the tensor Rayleigh quotient [10]

$$\lambda_i = \frac{\boldsymbol{x}^*\Delta_i\boldsymbol{x}}{\boldsymbol{x}^*\Delta_0\boldsymbol{x}} \tag{27}$$

for $i = 1,\dots,k$.

Suppose now that ε is so small that $\boldsymbol{W} + \Delta\boldsymbol{W}$ is right definite for $\|\Delta V_{ij}\| \le \varepsilon\|E_{ij}\|,\ i = 1,\dots,k;\ j = 0,\dots,k$. Since the eigenvalues of $\boldsymbol{W} + \Delta\boldsymbol{W}$ can be expressed as Rayleigh quotients (27) it follows from right definiteness that the pseudospectrum $\Lambda_\varepsilon(\boldsymbol{W})$ is bounded. This yields the bound (26). $\square$

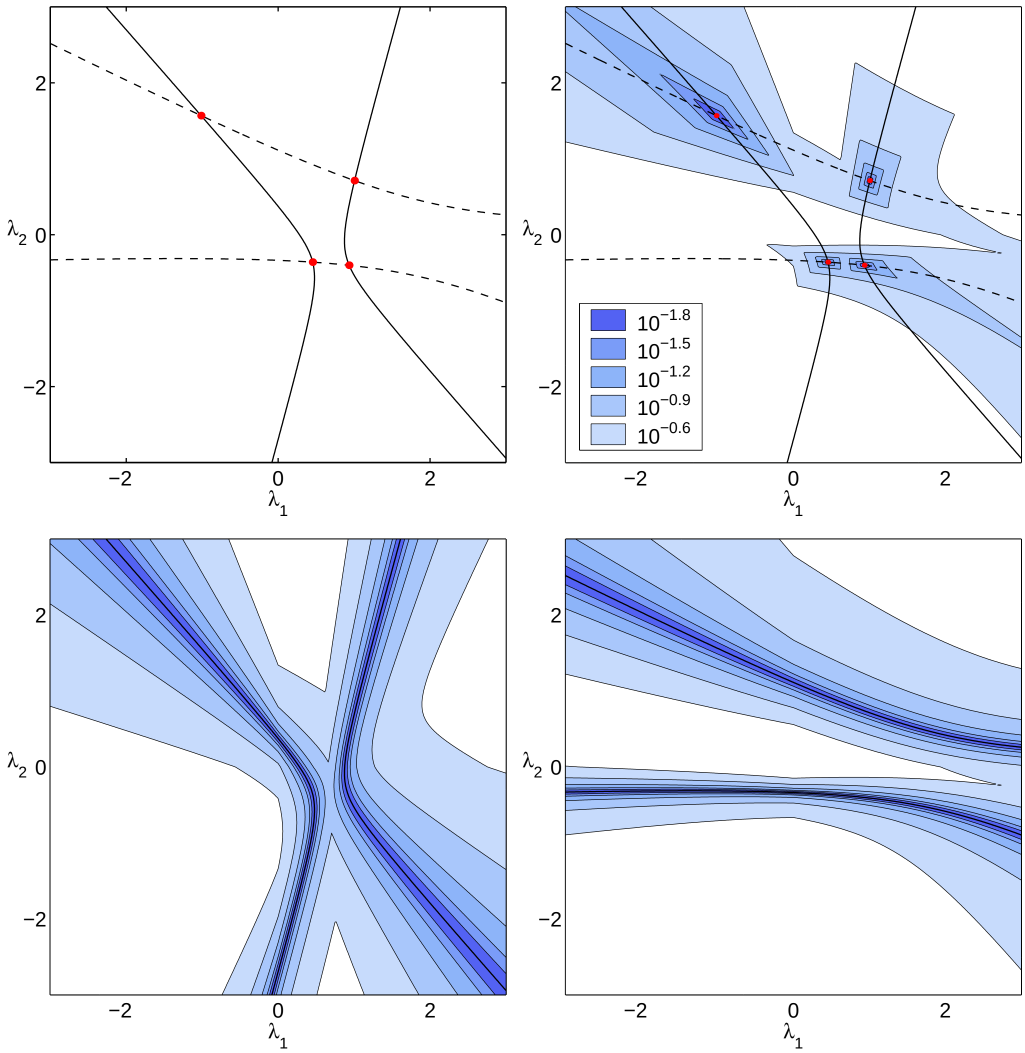
6 Numerical examples

We present some numerical examples obtained with Matlab 5.3. For all examples we take $E_{ij} = V_{ij}$ for all i, j. We draw all pseudospectra by computing $\sigma_{\min}(W_i(\boldsymbol{\lambda}))$ in all grid points by Matlab's `svd`. For more efficiency one could try to use similar ideas as mentioned in [13], but we will pay no further attention to this. The size of the grid used in the examples is 400×400.

Example 1. For the first numerical example we take the right definite two-parameter eigenvalue problem

$$W_1(\boldsymbol{\lambda}) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2.2 & 1 \\ 1 & 2.3 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0.1 & -1 \\ -1 & 0.1 \end{bmatrix},$$
$$W_2(\boldsymbol{\lambda}) = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 & -0.2 \\ -0.2 & -0.1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 2 & -0.1 \\ -0.1 & 4 \end{bmatrix}.$$

Fig. 1. Pseudospectra for Example 1. Top left: The eigenvalues are intersections of the eigencurves $\det W_1(\boldsymbol{\lambda}) = 0$ (solid line) and $\det W_2(\boldsymbol{\lambda}) = 0$ (dashed line). Top right: pseudospectra for $\varepsilon = 10^{-1.8}, 10^{-1.5}, 10^{-1.2}, 10^{-0.9}, 10^{-0.6}$. Bottom: pseudospectra for W_1 (left) and W_2 (right).



The eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ are intersection points of the eigenvalue curves $\det(W_1(\boldsymbol{\lambda})) = 0$ and $\det(W_2(\boldsymbol{\lambda})) = 0$ as depicted in the top left picture in Figure 1. The pseudospectra for $\varepsilon = 10^{-0.6}, 10^{-0.3}, 10^{0}, 10^{0.3}$ are shown in the top right picture in Figure 1. One can see that the boundaries of the pseudospectra are not differentiable. The reason is that pseudospectra are intersections of pseudospectra for W_1 and W_2, which are shown on the bottom left and bottom right picture in Figure 1, respectively.

Table 1
Eigenvalues and their condition numbers for the right definite two-parameter problem in Example 1.

λ_1	λ_2	$\kappa(\boldsymbol{\lambda}, \boldsymbol{W})$
-1.0142	1.5688	4.66
0.4556	-0.3613	2.42
0.9360	-0.4025	3.34
1.0069	0.7125	3.37

The eigenvalues together with the corresponding condition numbers are presented in Table 1. To obtain the condition number of an eigenvalue we have to compute $\|B_0^{-1}\|_{\boldsymbol{\theta}}$. Since the problem is right definite and all matrices V_{ij} are real we have to consider only real vectors in definition (5) of $\|B_0^{-1}\|_{\boldsymbol{\theta}}$. This fact makes it easy to compute the $\boldsymbol{\theta}$-norm as we only have to compute a finite number of norms. In particular, for a right definite two-parameter case we have

$$\|B_0^{-1}\|_{\boldsymbol{\theta}} = \max\{\, \|B_0^{-1}z\| \;:\; z \in \mathbb{R}^2,\ |z_i| = \theta_i \text{ for } i = 1, 2\,\}.$$

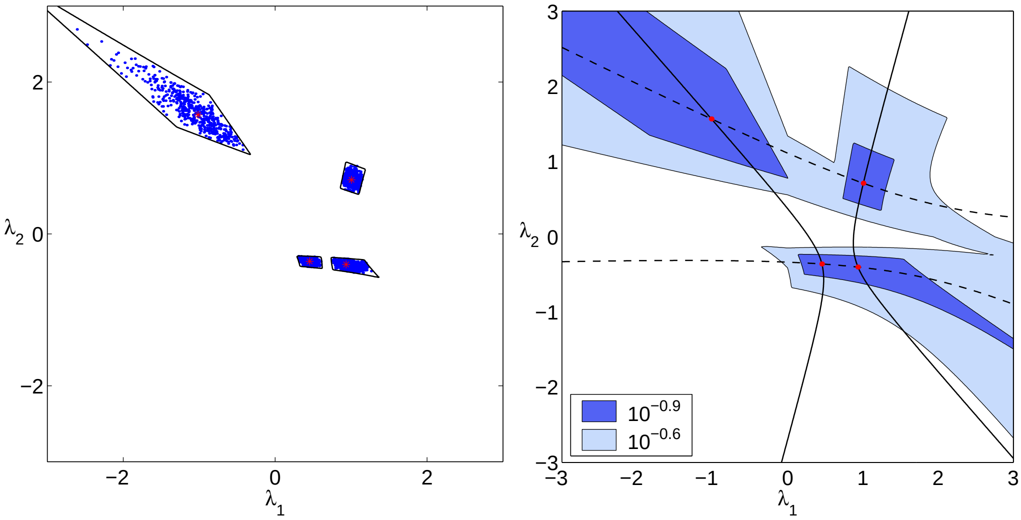
By comparing the results of Table 1 and Figure 1 one can see that the eigenvalue with the largest condition number has the largest pseudospectrum as may be expected.

The left figure in Figure 2 shows eigenvalues of 500 randomly perturbed problems, where each ΔV_{ij} is a random symmetric matrix such that $\|\Delta V_{ij}\| = 10^{-1.2}\|V_{ij}\|$. One can see that all dots in Figure 2 lie in the interior of the pseudospectrum for $\varepsilon = 10^{-1.2}$.

The right figure in Figure 2 presents pseudospectra for $\varepsilon = 10^{-0.9}$ and $\varepsilon = 10^{-0.6}$ on a larger area. One may suspect that here, in contrast to the eigenvalue problem $Ax = \lambda x$, a pseudospectrum may be unbounded.

Figures 1 and 2 suggest that the sensitivity of the eigenvalue is related to the angle of the intersection between the curves $\det(W_1(\boldsymbol{\lambda})) = 0$ and $\det(W_2(\boldsymbol{\lambda})) = 0$. We observe that the pseudospectrum is large when the angle of the intersection is small. The following proposition (which can be easily generalized to

Fig. 2. Left: Eigenvalues of 500 randomly perturbed two-parameter eigenvalue problems of Example 1, where each ΔV_{ij} is a symmetric matrix such that $\|\Delta V_{ij}\| = 10^{-1.2}\|V_{ij}\|$, and pseudospectrum for $\varepsilon = 10^{-1.2}$. Right: Pseudospectra for Example 1 for $\varepsilon = 10^{-0.9}$ and $\varepsilon = 10^{-0.6}$.



MEPs with more than two parameters) justifies this observation.

Proposition 13 *Let $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{R}^2$ be an algebraically simple eigenvalue of a real right definite two-parameter eigenvalue problem $\mathbf{W}$ and let $\boldsymbol{x} = x_1 \otimes x_2$ and $\boldsymbol{y} = y_1 \otimes y_2$ be the corresponding normalized right and left eigenvector, respectively. Then*

$$B_0 = \begin{bmatrix} \pm \prod_{j=1}^{n_1-1} \sigma_j^{(1)}(\boldsymbol{\mu}) & 0 \\ 0 & \pm \prod_{j=1}^{n_2-1} \sigma_j^{(2)}(\boldsymbol{\mu}) \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f_1}{\partial \lambda_1}(\boldsymbol{\mu}) & \frac{\partial f_1}{\partial \lambda_2}(\boldsymbol{\mu}) \\ \frac{\partial f_2}{\partial \lambda_1}(\boldsymbol{\mu}) & \frac{\partial f_2}{\partial \lambda_2}(\boldsymbol{\mu}) \end{bmatrix},$$

where $f_i(\boldsymbol{\lambda}) = \det W_i(\boldsymbol{\lambda})$ and where $\sigma_1^{(i)}(\boldsymbol{\mu}) \geq \cdots \geq \sigma_{n_i-1}^{(i)}(\boldsymbol{\mu}) > 0$ are nonzero singular values of $W_i(\boldsymbol{\mu})$ for $i = 1, 2$.

Proof. We define $Z(t) = V_{10} - tV_{11} - \mu_2 V_{12}$ and $g(t) = \det(Z(t))$. Since $Z(t)$ is a real analytic function of t, there exists an analytic singular value decomposition (see [4])

$$Z(t) = U(t)\Sigma(t)V(t)^T \tag{28}$$

such that

(1) $U(t)$ and $V(t)$ are orthogonal matrices,

(2) $\Sigma(t) = \text{diag}(\sigma_1(t), \ldots, \sigma_{n_1}(t))$ is a diagonal matrix,
(3) the elements of $U(t)$, $\Sigma(t)$, and $V(t)$ are analytic functions of t in a small neighborhood of μ_1, and
(4) $Z(\mu_1) = U(\mu_1)\Sigma(\mu_1)V(\mu_1)^T$ is a singular value decomposition of $W_i(\boldsymbol{\mu})$.

We may consider (28) as a singular value decomposition of $Z(t)$ where the singular values are not necessarily nonnegative and ordered. Let $u_{n_i}(t)$ and $v_{n_i}(t)$ denote the n_ith column of $U(t)$ and $V(t)$, respectively. Since $\boldsymbol{\mu}$ is an algebraically simple eigenvalue, $\sigma_{n_i}(\mu_1) = 0$, $\sigma_{n_i-1}(\mu_1) \neq 0$, $v_{n_i}(\mu_1) = x_i$, and $u_{n_i}(\mu_1) = y_i$.

If we differentiate $\sigma_{n_1}(t) = u_{n_1}(t)^T Z(t) v_{n_1}(t)$ then we obtain

$$\frac{d\sigma_{n_1}}{dt}(\mu_1) = -y_1^T V_{11} x_1 = -(B_0)_{11}. \tag{29}$$

From $g(t) = \mp\, \sigma_1(t)\sigma_2(t)\cdots\sigma_n(t)$ and (29) it follows that

$$\frac{\partial f_1}{\partial \lambda_1}(\boldsymbol{\mu}) = \frac{dg}{dt}(\mu_1) = \pm\, \sigma_1^{(1)}(\boldsymbol{\mu})\cdots\sigma_{n_1-1}^{(1)}(\boldsymbol{\mu})(B_0)_{11}.$$

In order to complete the proof one has to repeat the above procedure for all partial derivatives $\frac{\partial f_i}{\partial \lambda_j}(\boldsymbol{\mu})$ for $i, j = 1, 2$. $\square$

From Theorem 6 and (6) we can conclude that $\|B_0^{-1}\|$ has a great impact on the sensitivity of the eigenvalue $\boldsymbol{\lambda}$. As follows from Proposition 13, $\|B_0^{-1}\|$ may be large when the angle of the intersection between the curves $\det(W_1(\boldsymbol{\lambda})) = 0$ and $\det(W_2(\boldsymbol{\lambda})) = 0$ is small.

Example 2. For the second example we take the two-parameter Sturm–Liouville problem

$$\begin{aligned} W_1(\boldsymbol{\lambda})x_1(t_1) &= -x_1''(t_1) - (\lambda_1 + \lambda_2 \cos 2t_1)x_1(t_1), \\ W_2(\boldsymbol{\lambda})x_2(t_2) &= -x_2''(t_2) - \lambda_2 x_2(t_2) \end{aligned} \tag{30}$$

with boundary conditions $x_i(0) = x_i(\pi) = 0$ for $i = 1, 2$, studied in [3]. The second equation of (30) yields that $\lambda_2 = 1^2, 2^2, 3^2, \ldots$ and then it follows from the first equation of (30) that λ_1 is an eigenvalue of the Mathieu equation with parameter λ_2.

If we take $h = \pi/n$ and apply the finite-difference method to the two-parameter boundary-value problem (30) using symmetric differences $y_i' \approx (y_{i+1} - y_{i-1})/(2h)$

and $y_i'' \approx (y_{i+1} - 2y_i + y_{i-1})/h^2$ for the derivatives y' and y'', then we obtain an algebraic two–parameter problem where

$$V_{10} = V_{20} = \frac{1}{h^2}\,\mathrm{tridiag}(1, -2, 1),$$

$$V_{11} = I, \quad V_{21} = 0, \tag{31}$$

$$V_{12} = \mathrm{diag}\left(\cos\frac{2\pi}{n+1}, \cos\frac{4\pi}{n+1}, \ldots, \cos\frac{2n\pi}{n+1}\right), \quad V_{22} = I_n.$$

The eigenvalues of the above algebraic two-parameter problem are approximations to the eigenvalues of (30) with order of approximation $\mathcal{O}(h^2)$.

Fig. 3. Pseudospectra for the algebraic two-parameter approximation of Example 2, where $n = 10$ and $\varepsilon = 10^{-1.8}, 10^{-1.5}, 10^{-1.2}, 10^{-0.9}, 10^{-0.6}$.

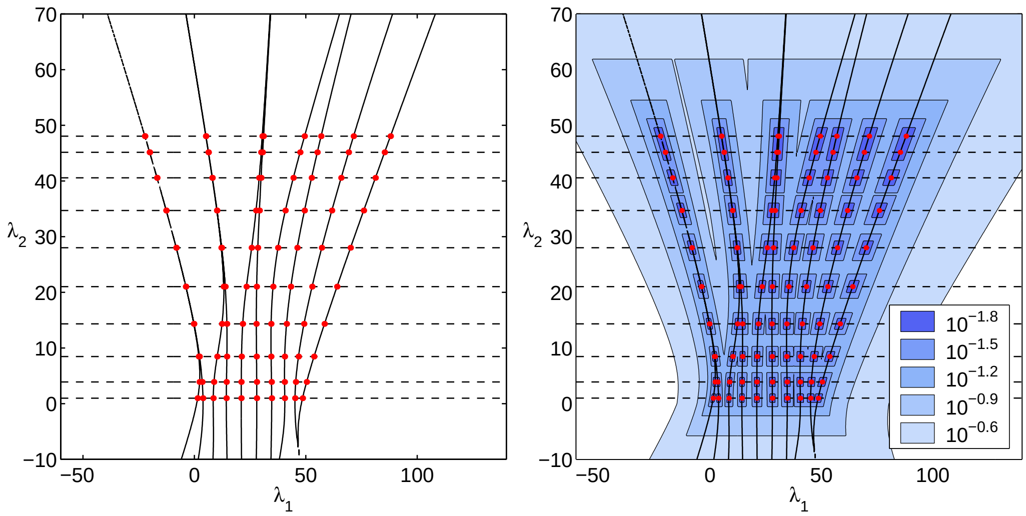


Figure 3 shows eigenvalues and pseudospectra for the algebraic two-parameter approximation (31) of (30) for $n = 10$. The left figure shows eigenvalues as the points where eigencurves $\det(W_1(\boldsymbol{\lambda})) = 0$ (solid line) and $\det(W_2(\boldsymbol{\lambda})) = 0$ (dashed line) intersect. One should note that the lines $\det(W_2(\boldsymbol{\lambda})) = 0$ do not agree with the known result $\lambda_2 = 1^2, 2^2, 3^2, \ldots$. The reason is that the eigenvalues in Figure 3 are the eigenvalues of the algebraic approximation (31) and not of the original problem (30). The eigenvalues occur in groups of two for a fixed λ_2. In some of these pairs the eigenvalues are so close together that they look like a single eigenvalue on Figure 3, an example of such pair is $(-12.6225, 34.7056)$ and $(-12.6215, 34.7056)$. The right figure with the pseudospectra for $\varepsilon = 10^{-1.8}, 10^{-1.5}, \ldots, 10^{-0.6}$ indicates that the fact that some of the eigenvalues are close together does not seem to influence their pseudospectra and the eigenvalues are well conditioned.

7 Conclusions

We have studied backward error, condition numbers, and pseudospectra for the MEP. The results can be viewed as a generalization of the theory for the generalized eigenvalue problem and are similar to the results for the polynomial eigenvalue problem. We also studied the nearness of a right definite MEP to a non right definite MEP and established that it is connected with unbounded pseudospectra.

References

[1] F. V. Atkinson, Multiparameter spectral theory, *Bull. Amer. Math. Soc.* **74** (1968) 1–27.

[2] F. V. Atkinson, *Multiparameter eigenvalue problems*, Academic Press, New York, 1972.

[3] P. A. Binding, P. J. Browne, X. Ji, A numerical method using the Prüfer transformation for the calculation of eigenpairs of two–parameter Sturm–Liouville problems, *IMA J. Num. Anal.* **13** (1993) 559-569.

[4] A. Bunse-Gerstner, R. Byers, V. Mehrmann, N. K. Nichols, Numerical computation of an analytic singular value decomposition of a matrix valued function, *Numer. Math.* **60** (1991) 1–39.

[5] J.-P. Dedieu, F. Tisseur, Perturbation theory for polynomial eigenvalue problems in homogeneous form, *Numerical Analysis Report 363*, Manchester Centre for Computational Mathematics, Manchester, 2001.

[6] V. Frayssé, M. Gueury, F. Nicoud, V. Toumazou, Spectral portraits for matrix pencils, *Tech. Report TR/PA/96/19*, CERFACS, Toulouse, 1996.

[7] V. Frayssé, V. Toumazou, A note on the normwise perturbation theory for the regular generalized eigenproblem, *Numer. Linear Algebra Appl.* **5** (1998) 1–10.

[8] D. J. Higham, N. J. Higham, Structured backward error and condition of generalized eigenvalue problems, *SIAM J. Matrix Anal. Appl.* **20** (1998) 493–512.

[9] T. Košir, Finite dimensional multiparameter spectral theory: the nonderogatory case, *Linear Algebra Appl.* **212/213** (1994) 45–70.

[10] B. Plestenjak, A continuation method for a right definite two-parameter eigenvalue problem, *SIAM J. Matrix Anal. Appl.* **21** (2000) 1163–1184.

[11] F. Tisseur, Backward error and condition of polynomial eigenvalue problems, *Linear Algebra Appl.* **309** (2000) 339–361.

[12] F. Tisseur, N. J. Higham, Structured pseudospectra for polynomial eigenvalue problems with applications, *SIAM J. Matrix Anal. Appl.* **23** (2001) 187–208.

[13] L. N. Trefethen, Computation of pseudospectra, *Acta Numer.* **8** (1999) 247–295.

[14] L. N. Trefethen, Spectra and pseudospectra, in: M. Ainsworth, J. Levesley, and M. Marletta, eds., *The Graduate Student's Guide to Numerical Analysis '98*, Springer, Berlin, 1999, 217–250.

[15] H. Volkmer, *Multiparameter Problems and Expansion Theorems*, vol. 1356 of Lecture Notes in Mathematics, Springer, Berlin, New York, 1988.