The distance from a matrix polynomial to matrix polynomials with a prescribed multiple eigenvalue[1]

Nikolaos Papathanasiou and Panayiotis Psarrakos[2]
Department of Mathematics, National Technical University of Athens
Zografou Campus, 15780 Athens, Greece

March 31, 2008

Abstract

For a matrix polynomial $P(\lambda)$ and a given complex number μ, we introduce a (spectral norm) distance from $P(\lambda)$ to the matrix polynomials that have μ as an eigenvalue of geometric multiplicity at least κ, and a distance from $P(\lambda)$ to the matrix polynomials that have μ as a multiple eigenvalue. Then we compute the first distance and obtain bounds for the second one, constructing associated perturbations of $P(\lambda)$.

Keywords: matrix polynomial; eigenvalue; multiplicity; perturbation; ε-pseudospectrum

AMS Subject Classifications: 65F15; 65F35

1 Introduction

The distance from a matrix $A \in \mathbb{C}^{n\times n}$ with simple eigenvalues to the set of matrices with multiple eigenvalues, and its relationship with the conditioning of the eigenproblem of A, were originally studied by Householder [8] and Wilkinson [16]. Several bounds for this distance have been obtained by Ruhe [13], Wilkinson [17, 18, 19, 20] and Demmel [2]. Nearness to matrices with (multiple) defective eigenvalues can also explain transient behaviors of the matrix exponential [3].

Using *Singular Value Decomposition (SVD)* and standard arguments of matrix analysis one can easily verify the following result, which was first published (in a slightly different form) by Golub, Klema and Stewart [5] (see also [6, Theorem 2.5.3]). Note that $\|\cdot\|$ denotes the spectral matrix norm, i.e., that norm subordinate to the Euclidean vector norm.

Theorem 1 (Golub, Klema and Stewart, 1976) *Let $A \in \mathbb{C}^{n\times n}$ and $\mu \in \mathbb{C}$. Suppose that the matrix $I\mu - A$ has an SVD of the form*

$$I\mu - A = U\,\Sigma\,V^* = U\,\mathrm{diag}\left\{s_1(I\mu - A), s_2(I\mu - A), \ldots, s_n(I\mu - A)\right\}\,V^*,$$

[1] Research supported by a grant of the EPEAEK project PYTHAGORAS II. The project is co-funded by the European Social Fund (75%) and Greek National Resources (25%).

[2] E-mail: ppsarr@math.ntua.gr. Corresponding author.

where the matrices $U, V \in \mathbb{C}^{n\times n}$ *are unitary and* $s_1(I\mu - A) \ge s_2(I\mu - A) \ge \cdots \ge s_n(I\mu - A) \ge 0$ *are the singular values of* $I\mu - A$*. Then the distance from* A *to the* $n\times n$ *matrices* X *that have* μ *as an eigenvalue of geometric multiplicity* $\ge \kappa$*,*

$$\min\left\{\|X - A\| : \mu \textit{ is an eigenvalue of } X \textit{ with geometric multiplicity } \ge \kappa\right\},$$

is equal to the singular value $s_{n-\kappa+1}(I\mu - A)$*, and an optimal perturbation of* A *is*

$$X_\mu = I\mu - U\,\mathrm{diag}\left\{s_1(I\mu - A), \ldots, s_{n-\kappa}(I\mu - A), 0, \ldots, 0\right\}\,V^*.$$

The next theorem was recently proved by Malyshev [11], and gives the (spectral norm) distance from A to the set of matrices with $\mu \in \mathbb{C}$ as a multiple eigenvalue. Here and elsewhere in the paper, when we consider a *pair of a left singular vector* $u \in \mathbb{C}^n$ *and a right singular vector* $v \in \mathbb{C}^n$ of a matrix $A \in \mathbb{C}^{n\times n}$ corresponding to the singular value $s_j(A)$, we *always* assume that there is an SVD of A,

$$A = U\,\Sigma\,V^* = U\,\mathrm{diag}\left\{s_1(A), s_2(A), \ldots, s_n(A)\right\}V^*$$

($s_1(A) \ge s_2(A) \ge \cdots \ge s_n(A) \ge 0$) with u and v as the j-th columns of the unitary matrices U and V, respectively. This means that these (unit) singular vectors are not arbitrarily chosen, and they satisfy $Av = s_j(A)u$ and $u^*A = s_j(A)v^*$.

Theorem 2 (Malyshev, 1999) *The distance from a matrix* $A \in \mathbb{C}^{n\times n}$ *to the set of* $n\times n$ *matrices* X *that have a given* $\mu \in \mathbb{C}$ *as a multiple eigenvalue,*

$$\min\left\{\|X - A\| : \mu \textit{ is a multiple eigenvalue of } X\right\},$$

is equal to the maximum (with respect to $\gamma \ge 0$*) singular value*

$$s_* = \max_{\gamma\ge 0} s_{2n-1}\left(\begin{bmatrix} \mu I - A & 0 \\ \gamma I & \mu I - A \end{bmatrix}\right).$$

Furthermore, if s_* *corresponds to the value* $\gamma_* > 0$*, then there is a pair* $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k,\, v_k \in \mathbb{C}^n,\ k = 1, 2)$ *of left and right singular vectors of* s_**, respectively, such that an optimal perturbation of* A *is* $X_\mu = A + s_*\,[u_1\ u_2]\,[v_1\ v_2]^\dagger$*, where* $[v_1\ v_2]^\dagger$ *is the Moore-Penrose pseudoinverse of* $[v_1\ v_2]$*. If* s_* *corresponds to the value* $\gamma_* = 0$ *and* $u, v \in \mathbb{C}^n$ *is a pair of left and right singular vectors of* $I\mu - A$ *for the singular value* s_**, respectively, then an optimal perturbation of* A *is* $X_\mu = A + s_* uv^*$*.*

In this article, we generalize the above theorems to the case of matrix polynomials. In Section 3, we estimate the distance from a matrix polynomial to the set of matrix polynomials that have a given complex number as an eigenvalue of geometric multiplicity at least κ, and construct an optimal perturbation (Theorem 4). Then, in Sections 4–6, we extend the methodology of Malyshev [11] (see also [12]), and obtain lower and upper bounds for the distance from a matrix polynomial to matrix polynomials that have a prescribed multiple eigenvalue (Theorems 11, 19 and 20). Perturbations that lead to our upper bounds are also given. Moreover, in Section 7, we confirm that our results in Sections 4–6 are direct generalizations of the results of [11]. Finally, in Section 8, we present three illustrative examples.

2 Definitions for matrix polynomials

Consider an $n \times n$ matrix polynomial

$$P(\lambda) = A_m \lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0, \tag{1}$$

where λ is a complex variable and $A_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ with $\det A_m \neq 0$. The study of matrix polynomials, especially with regard to their spectral analysis, has a long history and important applications [4].

A scalar $\lambda_0 \in \mathbb{C}$ is called an *eigenvalue* of $P(\lambda)$ if the system $P(\lambda_0)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$. This solution x_0 is known as a *(right) eigenvector* of $P(\lambda)$ corresponding to λ_0. A nonzero vector $y_0 \in \mathbb{C}^n$ that satisfies $y_0^* P(\lambda_0) = 0$ is called a *left eigenvector* of $P(\lambda)$ corresponding to λ_0. The set of all eigenvalues of $P(\lambda)$ is the *spectrum* of $P(\lambda)$, namely, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, and since $\det A_m \neq 0$, it contains no more than nm distinct (finite) elements. The *algebraic multiplicity* of a $\lambda_0 \in \sigma(P)$ is the multiplicity of λ_0 as a zero of the (scalar) polynomial $\det P(\lambda)$, and it is always greater than or equal to the *geometric multiplicity* of λ_0, that is, the dimension of the null space of the matrix $P(\lambda_0)$. A multiple eigenvalue of $P(\lambda)$ is called *defective* if its algebraic multiplicity is greater than its geometric one.

We are interested in perturbations of the matrix polynomial $P(\lambda)$ of the form

$$Q(\lambda) \;=\; P(\lambda) + \Delta(\lambda) \;=\; \sum_{j=0}^{m} (A_j + \Delta_j)\lambda^j, \tag{2}$$

where the matrices $\Delta_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ are arbitrary. For a given parameter $\varepsilon > 0$ and a given set of nonnegative weights $\mathrm{w} = \{w_0, w_1, \ldots, w_m\}$ with $w_0 > 0$, we define the class of admissible perturbed matrix polynomials

$$\mathcal{B}(P, \varepsilon, \mathrm{w}) = \{Q(\lambda) \text{ as in (2)} : \|\Delta_j\| \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m\}.$$

The weights w_j $(j = 0, 1, \ldots, m)$ allow freedom in how perturbations are measured; for example, in an absolute sense when $w_0 = w_1 = \cdots = w_m = 1$, or in a relative sense when $w_j = \|A_j\|$ $(j = 0, 1, \ldots, m)$. Moreover, $\mathcal{B}(P, \varepsilon, \mathrm{w})$ is convex and compact [1], with respect to the max norm $\|P(\lambda)\|_\infty = \max_{0\leq j\leq m}\|A_j\|$.

Next we introduce the distance from $P(\lambda)$ to the set of matrix polynomials that have a prescribed eigenvalue of algebraic multiplicity at least 2, or of a given geometric multiplicity.

Definition 3 For the matrix polynomial $P(\lambda)$ in (1) and a given $\mu \in \mathbb{C}$, we define the *distance from $P(\lambda)$ to μ as a multiple eigenvalue* by

$$\mathcal{E}_a(\mu) \;=\; \min\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w}) \text{ with } \mu \text{ as a multiple eigenvalue}\},$$

and the *distance from $P(\lambda)$ to μ as an eigenvalue with geometric multiplicity κ* by

$$\mathcal{E}_{g,\kappa}(\mu) \;=\; \min\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w}) \text{ with } \mu \text{ as an eigenvalue of geometric multiplicity at least } \kappa\}.$$

If $P(\lambda) = I\lambda - A$ for some $A \in \mathbb{C}^{n\times n}$, then $\sigma(P)$ coincides with the standard spectrum of A, $\sigma(A)$. If in addition, $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, then $\mathcal{B}(P, \varepsilon, \mathrm{w}) = \{I\lambda - (A + E) : \|E\| \leq \varepsilon\}$, and the distances $\mathcal{E}_{g,\kappa}(\mu)$ and $\mathcal{E}_a(\mu)$ are given by Theorems 1 and 2, respectively.

3 Computation of the distance $\mathcal{E}_{g,\kappa}(\mu)$

Consider the matrix polynomial $P(\lambda)$ in (1), a set of weights $\mathrm{w} = \{w_0, w_1, \ldots, w_m\}$ with $w_0 > 0$, and perturbations $Q(\lambda)$ of the form (2). For any $\lambda \in \mathbb{C}$, the singular values of $P(\lambda)$, i.e., the nonnegative roots of the eigenvalue functions of $P(\lambda)^* P(\lambda)$, are denoted by $s_1(P(\lambda)) \geq s_2(P(\lambda)) \geq \cdots \geq s_n(P(\lambda)) \geq 0$. Observe that $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ of geometric multiplicity κ if and only if the matrix $P(\lambda_0)$ is of rank $n - \kappa$, or equivalently, if and only if

$$s_1(P(\lambda_0)) \geq \cdots \geq s_{n-\kappa}(P(\lambda_0)) > s_{n-\kappa+1}(P(\lambda_0)) = \cdots = s_n(P(\lambda_0)) = 0.$$

Suppose that $\mu \in \mathbb{C}$ is not an eigenvalue of $P(\lambda)$ with geometric multiplicity $\geq \kappa$. In this section, we compute the distance $\mathcal{E}_{g,\kappa}(\mu)$ (i.e., the minimum $\varepsilon > 0$ such that μ is an eigenvalue of some $Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})$ of geometric multiplicity $\geq \kappa$) and an optimal perturbation of $P(\lambda)$. We consider an SVD of the matrix $P(\mu)$,

$$\begin{aligned} P(\mu) &= \hat{U}\,\Sigma_\mu\,\hat{V}^* \\ &= [\hat{u}_1\ \hat{u}_2\ \cdots\ \hat{u}_n]\ \mathrm{diag}\,\{s_1(P(\mu)), s_2(P(\mu)), \ldots, s_n(P(\mu))\}\ [\hat{v}_1\ \hat{v}_2\ \cdots\ \hat{v}_n]^*, \end{aligned}$$

and we define the matrix (see also [1, 15])

$$\begin{aligned} E &= -\hat{U}\,\mathrm{diag}\,\{0, \ldots, 0, s_{n-\kappa+1}(P(\mu)), \ldots, s_n(P(\mu))\}\ \hat{V}^* \\ &= -[\hat{u}_{n-\kappa+1}\ \cdots\ \hat{u}_n]\ \mathrm{diag}\,\{s_{n-\kappa+1}(P(\mu)), \ldots, s_n(P(\mu))\}\ [\hat{v}_{n-\kappa+1}\ \cdots\ \hat{v}_n]^*. \end{aligned}$$

By Theorem 1, the matrix $P(\mu) + E$ is a nearest matrix to $P(\mu)$ that has 0 as an eigenvalue of geometric multiplicity $\geq \kappa$. Then $E\hat{v}_j = -s_j(P(\mu))\hat{u}_j$ and $\hat{u}_j^* E = -s_j(P(\mu))\hat{v}_j^*$ for every $j = n-\kappa+1, \ldots, n$, and $\|E\| = s_{n-\kappa+1}(P(\mu))$. We also define the scalar polynomial $w(\lambda) = w_m\lambda^m + \cdots + w_1\lambda + w_0$ and the matrices

$$\hat{\Delta}_j = \frac{w_j}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|}\right)^j E\ ; \quad j = 0, 1, \ldots, m,$$

where we set $\overline{\mu}/|\mu| = 0$ whenever $\mu = 0$. The matrix polynomial

$$\hat{\Delta}(\lambda) = \sum_{j=0}^{m} \hat{\Delta}_j \lambda^j$$

satisfies

$$\hat{\Delta}(\mu) = \left(\sum_{j=0}^{m} w_j |\mu|^j\right) w(|\mu|)^{-1} E = E,$$

and for the perturbation

$$\hat{Q}(\lambda) \;=\; P(\lambda) + \hat{\Delta}(\lambda) \;=\; \sum_{j=0}^{m} (A_j + \hat{\Delta}_j)\lambda^j \tag{3}$$

of $P(\lambda)$ (introduced in [15]), it is clear that

$$\hat{Q}(\mu)\hat{v}_j = P(\mu)\hat{v}_j + \hat{\Delta}(\mu)\hat{v}_j = 0 \quad \text{and} \quad \hat{u}_j^* \hat{Q}(\mu) = \hat{u}_j^* P(\mu) + \hat{u}_j^* \hat{\Delta}(\mu) = 0$$

for every $j = n - \kappa + 1, \ldots, n$. As a consequence, $\mu \in \sigma(\hat{Q})$ with geometric multiplicity $\geq \kappa$, (right) eigenvectors $\hat{v}_{n-\kappa+1}, \ldots, \hat{v}_n$ and left eigenvectors $\hat{u}_{n-\kappa+1}, \ldots, \hat{u}_n$. Moreover, $\|\hat{\Delta}_j\| = w_j\, w(|\mu|)^{-1} s_{n-\kappa+1}(P(\mu))$ $(j = 0, 1, \ldots, m)$, and hence, $\hat{Q}(\lambda)$ lies on the boundary $\partial\mathcal{B}(P, s_{n-\kappa+1}(P(\mu))/w(|\mu|), \mathrm{w})$.

Assume now that for a positive $\varepsilon < s_{n-\kappa+1}(P(\mu))/w(|\mu|)$, there is a $Q(\lambda) = P(\lambda) + \Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})$ that has μ as an eigenvalue of geometric multiplicity $\geq \kappa$. Then the matrix polynomial $\Delta(\lambda)$ is of the form

$$\Delta(\lambda) = \sum_{j=0}^{m} \Delta_j \lambda^j,$$

where

$$\|\Delta_j\| \leq \varepsilon\, w_j \; ; \;\; j = 0, 1, \ldots, m,$$

and thus,

$$\|\Delta(\mu)\| \leq \sum_{j=0}^{m} \|\Delta_j\|\, |\mu|^j \leq \varepsilon \sum_{j=0}^{m} w_j\, |\mu|^j = \varepsilon\, w(|\mu|) < s_{n-\kappa+1}(P(\mu)).$$

This is a contradiction because the matrix $Q(\mu) = P(\mu) + \Delta(\mu)$ is a perturbation of the matrix $P(\mu)$ that has 0 as an eigenvalue of geometric multiplicity $\geq \kappa$, and by Theorem 1, $\|\Delta(\mu)\| \geq s_{n-\kappa+1}(P(\mu))$. Hence, we have the following result.

Theorem 4 *Consider the matrix polynomial $P(\lambda)$ in (1) and a scalar $\mu \in \mathbb{C}$. Then the distance from $P(\lambda)$ to μ as an eigenvalue of geometric multiplicity κ, is*

$$\mathcal{E}_{g,\kappa}(\mu) = \frac{s_{n-\kappa+1}(P(\mu))}{w(|\mu|)}.$$

Furthermore, the perturbation $\hat{Q}(\lambda)$ in (3) lies on $\partial\mathcal{B}(P, \mathcal{E}_{g,\kappa}(\mu), \mathrm{w})$ and has μ as an eigenvalue of geometric multiplicity $\geq \kappa$.

If we consider the linear pencil $P(\lambda) = I\lambda - A$ and $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, then it is apparent that the above theorem is a direct generalization of Theorem 1.

4 Bounds for the distance $\mathcal{E}_a(\mu)$

By the definition of the distances $\mathcal{E}_a(\mu)$ and $\mathcal{E}_{g,\kappa}(\mu)$ (recall Definition 3), and the results of the previous section, it is obvious that

$$\frac{s_n(P(\mu))}{w(|\mu|)} \;=\; \mathcal{E}_{g,1}(\mu) \;\leq\; \mathcal{E}_a(\mu) \;\leq\; \mathcal{E}_{g,2}(\mu) \;=\; \frac{s_{n-1}(P(\mu))}{w(|\mu|)}.$$

If $s_n(P(\mu)) = s_{n-1}(P(\mu))$, then (see also Proposition 14 of [1])

$$\mathcal{E}_a(\mu) \;=\; \mathcal{E}_{g,2}(\mu) \;=\; \mathcal{E}_{g,1}(\mu) \;=\; \frac{s_n(P(\mu))}{w(|\mu|)}$$

and an optimal perturbation of $P(\lambda)$ is the matrix polynomial $\hat{Q}(\lambda)$ in (3) (for $\kappa = 2$). Hence, for the distance $\mathcal{E}_a(\mu)$, we may assume that $s_n(P(\mu)) \neq s_{n-1}(P(\mu))$ and study perturbations of $P(\lambda)$ that have μ as a defective eigenvalue of algebraic multiplicity ≥ 2 and geometric multiplicity 1. The next definition will be needed in the sequel.

Definition 5 For the matrix polynomial $P(\lambda)$ in (1) and a scalar $\gamma \in \mathbb{C}$, we define the $2n \times 2n$ matrix polynomial

$$F[P(\lambda);\gamma] = \begin{bmatrix} P(\lambda) & 0 \\ \gamma P'(\lambda) & P(\lambda) \end{bmatrix},$$

where $P'(\lambda)$ denotes the derivative of $P(\lambda)$ with respect to λ.

Clearly, a $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ if and only if it is an eigenvalue of $F[P(\lambda);\gamma]$. Furthermore, when $\gamma \neq 0$, λ_0 is a multiple eigenvalue of $P(\lambda)$ if and only if the null space of the matrix $F[P(\lambda_0);\gamma]$ has dimension ≥ 2, as shown next.

Lemma 6 *A scalar $\lambda_0 \in \mathbb{C}$ is a multiple eigenvalue of $P(\lambda)$ if and only if for any nonzero $\gamma \in \mathbb{C}$,*

$$s_{2n-1}(F[P(\lambda_0);\gamma]) = s_{2n-1}\left(\begin{bmatrix} P(\lambda_0) & 0 \\ \gamma P'(\lambda_0) & P(\lambda_0) \end{bmatrix}\right) = 0.$$

Proof For any $\gamma \neq 0$, the singular value $s_{2n-1}(F[P(\lambda_0);\gamma])$ is equal to 0 if and only if the null space of the matrix $F[P(\lambda_0);\gamma]$ has dimension at least 2, i.e., if and only if there exist two (nonzero) linearly independent vectors $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \in \mathbb{C}^{2n}$ $(x_k, y_k \in \mathbb{C}^n,\ k = 1,2)$ such that

$$F[P(\lambda_0);\gamma] \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \ ; \ \ k = 1,2$$

or equivalently,

$$P(\lambda_0)x_k = 0 \ \text{ and } \ \gamma P'(\lambda_0)x_k + P(\lambda_0)y_k = 0 \ ; \ \ k = 1,2.$$

These equations hold if and only if λ_0 is a multiple eigenvalue of $P(\lambda)$. In particular, if $x_k \neq 0$ (for $k = 1$ or 2), then the vectors $x_k, y_k \in \mathbb{C}^n$ form a Jordan chain of length 2, corresponding to $\lambda_0 \in \sigma(P)$ (see [4] for the definition and properties of Jordan chains of matrix polynomials). If $x_1 = x_2 = 0$, then $y_1, y_2 \in \mathbb{C}^n$ are linearly independent eigenvectors corresponding to $\lambda_0 \in \sigma(P)$. □

Corollary 7 *For any $\lambda_0 \in \mathbb{C}$, we have that either, $s_{2n-1}(F[P(\lambda_0);\gamma]) \neq 0$ for every $\gamma \neq 0$, or $s_{2n-1}(F[P(\lambda_0);\gamma]) \equiv 0$.*

Proof Suppose that for a $\gamma_0 \neq 0$, $s_{2n-1}(F[P(\lambda_0);\gamma_0]) = 0$. Then λ_0 is a multiple eigenvalue of $P(\lambda)$, and thus, for every $\gamma \neq 0$, $s_{2n-1}(F[P(\lambda_0);\gamma]) = 0$. □

By Lemma 6, a scalar $\mu \in \mathbb{C}$ is a multiple eigenvalue of a perturbation $Q(\lambda) = P(\lambda) + \Delta(\lambda)$ if and only if μ is an eigenvalue of the $2n \times 2n$ matrix polynomial $F[Q(\lambda);\gamma] = \begin{bmatrix} Q(\lambda) & 0 \\ \gamma Q'(\lambda) & Q(\lambda) \end{bmatrix}$ (for some $\gamma \neq 0$) of geometric multiplicity ≥ 2. Moreover, the results of Section 3 yield the following lemma.

Lemma 8 *If $\mu \in \mathbb{C}$ is a multiple eigenvalue of a matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$, then for every $\gamma \neq 0$,*

$$s_{2n-1}(F[P(\mu);\gamma]) \le \left\| \begin{bmatrix} \Delta(\mu) & 0 \\ \gamma\Delta'(\mu) & \Delta(\mu) \end{bmatrix} \right\| \; (= \|F[\Delta(\mu);\gamma]\|\,).$$

The next result leads directly to a lower bound of the distance $\mathcal{E}_a(\mu)$.

Lemma 9 *If $\mu \in \mathbb{C}$ is a multiple eigenvalue of a perturbation $Q(\lambda) = P(\lambda) + \Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})$, then for every $\gamma \neq 0$,*

$$\varepsilon \ge \frac{\left\| \begin{bmatrix} \Delta(\mu) & 0 \\ \gamma\Delta'(\mu) & \Delta(\mu) \end{bmatrix} \right\|}{\left\| \begin{bmatrix} w(|\mu|) & 0 \\ \gamma w'(|\mu|) & w(|\mu|) \end{bmatrix} \right\|} \ge \frac{s_{2n-1}(F[P(\mu);\gamma])}{\|F[w(|\mu|);\gamma]\|}.$$

Proof For the matrix polynomials $\Delta(\mu)$ and $\Delta'(\mu)$, we know that

$$\|\Delta(\mu)\| \le \sum_{j=0}^{m} \|\Delta_j\|\,|\mu|^j \le \varepsilon\, w(|\mu|) \quad \text{and} \quad \|\Delta'(\mu)\| \le \sum_{j=1}^{m} j\,\|\Delta_j\|\,|\mu|^{j-1} \le \varepsilon\, w'(|\mu|).$$

Hence, for any $\gamma \neq 0$, there is a unit vector $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{C}^{2n}$ $(x, y \in \mathbb{C}^n)$ such that

$$\begin{aligned}
\left\| \begin{bmatrix} \Delta(\mu) & 0 \\ \gamma\Delta'(\mu) & \Delta(\mu) \end{bmatrix} \right\|^2 &= \left\| \begin{bmatrix} \Delta(\mu) & 0 \\ \gamma\Delta'(\mu) & \Delta(\mu) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} \Delta(\mu)x \\ \gamma\Delta'(\mu)x + \Delta(\mu)y \end{bmatrix} \right\|^2 \\
&= \|\Delta(\mu)x\|^2 + \left\|\gamma\Delta'(\mu)x + \Delta(\mu)y\right\|^2 \\
&\le \|\Delta(\mu)\|^2\,\|x\|^2 + |\gamma|^2 \left\|\Delta'(\mu)\right\|^2 \|x\|^2 \\
&\quad +2|\gamma|\,\|\Delta(\mu)\|\,\left\|\Delta'(\mu)\right\|\,\|x\|\|y\| + \|\Delta(\mu)\|^2\,\|y\|^2 \\
&\le (\varepsilon w(|\mu|))^2\|x\|^2 + |\gamma|^2(\varepsilon w'(|\mu|))^2\|x\|^2 \\
&\quad +2|\gamma|(\varepsilon w(|\mu|))(\varepsilon w'(|\mu|))\|x\|\|y\| + (\varepsilon w(|\mu|))^2\|y\|^2 \\
&= \left\| \begin{bmatrix} \varepsilon w(|\mu|)\,\|x\| \\ |\gamma|\varepsilon w'(|\mu|)\,\|x\| + \varepsilon w(|\mu|)\|y\| \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} \varepsilon w(|\mu|) & 0 \\ |\gamma|\varepsilon w'(|\mu|) & \varepsilon w(|\mu|) \end{bmatrix} \begin{bmatrix} \|x\| \\ \|y\| \end{bmatrix} \right\|^2 \\
&\le \left\| \begin{bmatrix} \varepsilon w(|\mu|) & 0 \\ \gamma\varepsilon w'(|\mu|) & \varepsilon w(|\mu|) \end{bmatrix} \right\|^2 .
\end{aligned}$$

The proof is completed by Lemma 8. □

By the above lemma, it is clear that

$$s_{2n-1}(F[P(\mu);\gamma]) \le \mathcal{E}_a(\mu) \left\| \begin{bmatrix} w(|\mu|) & 0 \\ \gamma w'(|\mu|) & w(|\mu|) \end{bmatrix} \right\| = \mathcal{E}_a(\mu) w(|\mu|) \left\| \begin{bmatrix} 1 & 0 \\ \gamma\frac{w'(|\mu|)}{w(|\mu|)} & 1 \end{bmatrix} \right\|.$$

Hence, the distance from $P(\lambda)$ to μ as a multiple eigenvalue satisfies

$$\mathcal{E}_a(\mu) \geq \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)} \left\| \begin{bmatrix} 1 & 0 \\ \gamma \frac{w'(|\mu|)}{w(|\mu|)} & 1 \end{bmatrix} \right\|^{-1}. \tag{4}$$

Now we turn our attention to the derivation of an upper bound of $\mathcal{E}_a(\mu)$. For our discussion, it is necessary to define two $n \times 2$ matrices related to the singular vectors of $F[P(\mu);\gamma]$ corresponding to $s_{2n-1}(F[P(\mu);\gamma])$.

Definition 10 Let $\begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \end{bmatrix}, \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k(\gamma),\, v_k(\gamma) \in \mathbb{C}^n,\ k = 1, 2)$ be a pair of left and right singular vectors of $s_{2n-1}(F[P(\mu);\gamma])$, respectively (for some γ). Then we define the $n \times 2$ matrices $U(\gamma) = [u_1(\gamma)\ u_2(\gamma)]$ and $V(\gamma) = [v_1(\gamma)\ v_2(\gamma)]$.

It is easy to see that $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k,\, v_k \in \mathbb{C}^n,\ k = 1, 2)$ is a pair of left and right singular vectors corresponding to a singular value of $F[P(\mu);\gamma]$ $(\gamma \neq 0)$ if and only if $\begin{bmatrix} u_1 \\ (\overline{\gamma}/|\gamma|)u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ (\overline{\gamma}/|\gamma|)v_2 \end{bmatrix}$ is a pair of left and right singular vectors of $F[P(\mu);|\gamma|]$ corresponding to the same singular value. Hence, for convenience (and without loss of generality), from this point and in the remainder of the paper, we assume that the parameter γ is real nonnegative.

For any $\gamma > 0$ with $rank(V(\gamma)) = 2$, we will construct a matrix polynomial $\Delta_\gamma(\lambda)$ such that the perturbation $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$ has μ as a multiple eigenvalue. First we consider the quantity

$$\phi = \frac{w'(|\mu|)}{w(|\mu|)} \frac{\overline{\mu}}{|\mu|}$$

(recall that $w_0 > 0$, and that, by convention, $\overline{\mu}/|\mu| = 0$ whenever $\mu = 0$) and the matrix

$$\Delta_\gamma = -s_{2n-1}(F[P(\mu);\gamma])U(\gamma) \begin{bmatrix} 1 & -\gamma\phi \\ 0 & 1 \end{bmatrix} V(\gamma)^\dagger,$$

where $V(\gamma)^\dagger$ is the Moore-Penrose pseudoinverse of $V(\gamma)$. Then we define the $n \times n$ matrix polynomial

$$\Delta_\gamma(\lambda) = \sum_{j=0}^{m} \Delta_{\gamma,j} \lambda^j$$

with

$$\Delta_{\gamma,j} = \frac{w_j}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|} \right)^j \Delta_\gamma \ ; \ \ j = 0, 1, \ldots, m,$$

and observe that $\Delta_\gamma(\mu) = \Delta_\gamma$ and $\Delta'_\gamma(\mu) = \phi\, \Delta_\gamma$.

Since $\begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \end{bmatrix}, \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \end{bmatrix} \in \mathbb{C}^{2n}$ is a pair of left and right singular vectors of $s_{2n-1}(F[P(\mu);\gamma])$, respectively, it follows

$$\begin{bmatrix} P(\mu) & 0 \\ \gamma P'(\mu) & P(\mu) \end{bmatrix} \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \end{bmatrix} = s_{2n-1}(F[P(\mu);\gamma]) \begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \end{bmatrix}.$$

As a consequence, for the matrix polynomial

$$Q_\gamma(\mu) \;=\; P(\mu) + \Delta_\gamma(\mu) \;=\; \sum_{j=0}^{m}(A_j + \Delta_{\gamma,j})\lambda^j \tag{5}$$

we have (keeping in mind that the condition $rank(V(\gamma)) = 2$ implies $V(\gamma)^\dagger V(\gamma) = I$)

$$\begin{aligned} Q_\gamma(\mu)v_1(\gamma) &= P(\mu)v_1(\gamma) + \Delta_\gamma v_1(\gamma) \\ &= s_{2n-1}(F[P(\mu);\gamma])u_1(\gamma) - s_{2n-1}(F[P(\mu);\gamma])u_1(\gamma) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} &\gamma Q'_\gamma(\mu)v_1(\gamma) + Q_\gamma(\mu)v_2(\gamma) \\ &\quad= \gamma P'(\mu)v_1(\gamma) + P(\mu)v_2(\gamma) + \gamma\Delta'_\gamma(\mu)v_1(\gamma) + \Delta_\gamma(\mu)v_2(\gamma) \\ &\quad= s_{2n-1}(F[P(\mu);\gamma])u_2(\gamma) + \gamma\,\phi\Delta_\gamma v_1(\gamma) + \Delta_\gamma v_2(\gamma) \\ &\quad= s_{2n-1}(F[P(\mu);\gamma])u_2(\gamma) - \gamma\,\phi\, s_{2n-1}(F[P(\mu);\gamma])u_1(\gamma) \\ &\qquad - s_{2n-1}(F[P(\mu);\gamma])U(\gamma)\begin{bmatrix} -\gamma\,\phi \\ 1 \end{bmatrix} \\ &\quad= s_{2n-1}(F[P(\mu);\gamma])u_2(\gamma) - \gamma\,\phi\, s_{2n-1}(F[P(\mu);\gamma])u_1(\gamma) \\ &\qquad + \gamma\,\phi\, s_{2n-1}(F[P(\mu);\gamma])u_1(\gamma) - s_{2n-1}(F[P(\mu);\gamma])u_2(\gamma) \\ &\quad= 0. \end{aligned}$$

This means that if $rank(V(\gamma)) = 2$, then μ is a defective eigenvalue of $Q_\gamma(\lambda)$ with $v_1(\gamma), v_2(\gamma) \in \mathbb{C}^n$ as an associated Jordan chain of length 2. Furthermore, it holds that

$$\|\Delta_{\gamma,j}\| = w_j\,\frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)}\left\|U(\gamma)\begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix}V(\gamma)^\dagger\right\| \;;\;\; j = 0, 1, \ldots, m.$$

Thus, for any $\gamma > 0$ with $rank(V(\gamma)) = 2$, the distance $\mathcal{E}_a(\mu)$ satisfies

$$\mathcal{E}_a(\mu) \;\le\; \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)}\left\|U(\gamma)\begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix}V(\gamma)^\dagger\right\|. \tag{6}$$

For $\gamma \ge 0$, we define

$$\begin{aligned} \beta_{low}(P,\mu,\gamma) &= \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)}\left\|\begin{bmatrix} 1 & 0 \\ \gamma\frac{w'(|\mu|)}{w(|\mu|)} & 1 \end{bmatrix}\right\|^{-1} \\ &= \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)}\left\|\begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix}\right\|^{-1} \end{aligned} \tag{7}$$

and

$$\beta_{up}(P,\mu,\gamma) \;=\; \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)}\left\|U(\gamma)\begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix}V(\gamma)^\dagger\right\|. \tag{8}$$

Then (4) and (6) imply that these quantities are a lower bound and an upper bound of $\mathcal{E}_a(\mu)$.

Theorem 11 *Suppose $P(\lambda)$ is a matrix polynomial as in (1) and $\mu \in \mathbb{C}$. Then for every $\gamma > 0$, $\mathcal{E}_a(\mu) \geq \beta_{low}(P,\mu,\gamma)$, and if $rank(V(\gamma)) = 2$, then $\mathcal{E}_a(\mu) \leq \beta_{up}(P,\mu,\gamma)$, where the bounds $\beta_{low}(P,\mu,\gamma)$ and $\beta_{up}(P,\mu,\gamma)$ are given by (7) and (8), respectively. Furthermore, if $rank(V(\gamma)) = 2$, then $Q_\gamma(\lambda)$ in (5) lies on the boundary of $\mathcal{B}(P,\beta_{up}(P,\mu,\gamma),\mathrm{w})$ and has μ as a defective eigenvalue.*

Note that if μ is not a multiple eigenvalue of $P(\lambda)$, then the upper bound $\beta_{up}(P,\mu,\gamma)$ and the lower bound $\beta_{low}(P,\mu,\gamma)$ can be strictly greater and less than the distance $\mathcal{E}_a(\mu)$, respectively. This is clear in Examples 1 and 3 below. On the other hand, if μ is a multiple eigenvalue of $P(\lambda)$, then $\beta_{up}(P,\mu,\gamma) = \beta_{low}(P,\mu,\gamma) = 0$ and $Q_\gamma(\lambda) = P(\lambda)$ for every $\gamma > 0$, and $\mathcal{E}_a(\mu) = 0$.

If we denote by $\|\cdot\|_F$ the Frobenius norm of a matrix, then we see that

$$\|U(\gamma)\|_F = \left\| \begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \end{bmatrix} \right\| = 1 \quad \text{and} \quad \|V(\gamma)\|_F = \left\| \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \end{bmatrix} \right\| = 1.$$

Since the $n \times 2$ matrices $U(\gamma)$ and $V(\gamma)$ are of rank 1 or 2, by [7, p. 315], it follows

$$\frac{\sqrt{2}}{2} \;\leq\; \|U(\gamma)\|, \|V(\gamma)\| \;\leq\; 1,$$

and thus, $\|V(\gamma)^\dagger\| \geq 1$. Moreover, the difference of the proposed bounds satisfies

$$\begin{aligned}
&\beta_{up}(P,\mu,\gamma) - \beta_{low}(P,\mu,\gamma) \\
&\quad = \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)} \left(\left\| U(\gamma) \begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix} V(\gamma)^\dagger \right\| - \left\| \begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix} \right\|^{-1} \right) \\
&\quad \leq \frac{s_{2n-1}(F[P(\mu);\gamma])}{w(|\mu|)} \left(\left\| \begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix} \right\| \left\| V(\gamma)^\dagger \right\| - \left\| \begin{bmatrix} 1 & -\gamma\,\phi \\ 0 & 1 \end{bmatrix} \right\|^{-1} \right),
\end{aligned}$$

and vanishes in a special case described next.

We observe that as $\gamma \longrightarrow 0^+$ or $\phi \longrightarrow 0^+$,

$$\beta_{up}(P,\mu,\gamma) \;\longrightarrow\; \frac{s_n(P(\mu))}{w(|\mu|)} \left\| U(\gamma)V(\gamma)^\dagger \right\| \;\leq\; \frac{s_n(P(\mu))}{w(|\mu|)} \left\| V(\gamma)^\dagger \right\|$$

and

$$\beta_{low}(P,\mu,\gamma) \;\longrightarrow\; \frac{s_n(P(\mu))}{w(|\mu|)} \;=\; \mathcal{E}_{g,1}(\mu).$$

If $\gamma \longrightarrow 0^+$ and $\left\| U(0)V(0)^\dagger \right\| = 1$, then both bounds $\beta_{up}(P,\mu,\gamma)$ and $\beta_{low}(P,\mu,\gamma)$ converge to $\mathcal{E}_{g,1}(\mu) = s_n(P(\mu))/w(|\mu|)$. This special case is illustrated in Example 2 below.

5 A value of γ that ensures $rank(V(\gamma)) = 2$

In this section, we define and study a special value of the parameter $\gamma > 0$ that implies $rank(V(\gamma)) = 2$.

Definition 12 Let $\gamma_* \geq 0$ be a point where the singular value $s_{2n-1}(F[P(\mu);\gamma])$ attains its maximum value (if any). For the sake of simplicity, we denote this maximum value by $s_* = s_{2n-1}(F[P(\mu);\gamma_*])$.

The case $\gamma_* > 0$ is considered below, and the case $\gamma_* = 0$ is treated in the next section. In particular, we obtain a simplification of the upper bound $\beta_{up}(P,\mu,\gamma)$ in (8), which allows the connection of our results with the results in [11].

First we derive a sufficient condition for the existence of γ_*.

Lemma 13 *Let B be an $n \times n$ matrix of rank ≥ 2. Then as $\gamma \longrightarrow \infty$ $(\gamma \geq 0)$,*

$$s_{2n-1}\left(\begin{bmatrix} P(\mu) & 0 \\ \gamma B & P(\mu) \end{bmatrix}\right) \longrightarrow 0.$$

Proof Suppose $\mu \notin \sigma(P)$, i.e., the matrix $P(\mu)$ is nonsingular. Then for every $\gamma \geq 0$,

$$\begin{bmatrix} P(\mu) & 0 \\ \gamma B & P(\mu) \end{bmatrix}^{-1} = \begin{bmatrix} P(\mu)^{-1} & 0 \\ -\gamma P(\mu)^{-1} B P(\mu)^{-1} & P(\mu)^{-1} \end{bmatrix}$$

and

$$\begin{aligned} s_{2n-1}\left(\begin{bmatrix} P(\mu) & 0 \\ \gamma B & P(\mu) \end{bmatrix}\right) &= \frac{1}{s_2\left(\begin{bmatrix} P(\mu)^{-1} & 0 \\ -\gamma P(\mu)^{-1} B P(\mu)^{-1} & P(\mu)^{-1} \end{bmatrix}\right)} \\ &= \frac{1}{s_2\left(\begin{bmatrix} P(\mu)^{-1} & 0 \\ \gamma P(\mu)^{-1} B P(\mu)^{-1} & P(\mu)^{-1} \end{bmatrix}\right)}. \end{aligned}$$

By Weyl's Theorem [7, Theorem 4.3.7] (see also [7, Exersice 7.3.16]), it follows that

$$\begin{aligned} & s_2\left(\begin{bmatrix} P(\mu)^{-1} & 0 \\ \gamma P(\mu)^{-1} B P(\mu)^{-1} & P(\mu)^{-1} \end{bmatrix}\right) \\ \geq\ & s_2\left(\begin{bmatrix} 0 & 0 \\ \gamma P(\mu)^{-1} B P(\mu)^{-1} & 0 \end{bmatrix}\right) - s_1\left(\begin{bmatrix} P(\mu)^{-1} & 0 \\ 0 & P(\mu)^{-1} \end{bmatrix}\right) \\ =\ & \gamma\, s_2(P(\mu)^{-1} B P(\mu)^{-1}) - \left\| P(\mu)^{-1} \right\|. \end{aligned}$$

Since $rank(B) \geq 2$, we have $s_2(P(\mu)^{-1} B P(\mu)^{-1}) > 0$. Thus, as $\gamma \longrightarrow \infty$,

$$s_{2n-1}\left(\begin{bmatrix} P(\mu) & 0 \\ \gamma B & P(\mu) \end{bmatrix}\right) \longrightarrow 0.$$

Suppose now that the matrix $P(\mu)$ is singular. For any $\delta > 0$, there is a $\mu_\delta \in \mathbb{C}$ sufficiently close to μ such that $\|P(\mu) - P(\mu_\delta)\| < \delta$ and $\det P(\mu_\delta) \neq 0$. By the first part of the proof, there is a real $\gamma_\delta > 0$ such that for every $\gamma \geq \gamma_\delta$,

$$s_{2n-1}\left(\begin{bmatrix} P(\mu_\delta) & 0 \\ \gamma B & P(\mu_\delta) \end{bmatrix}\right) < \delta.$$

As a consequence, Weyl's Theorem also yields

$$s_{2n-1}\left(\begin{bmatrix} P(\mu) & 0 \\ \gamma B & P(\mu) \end{bmatrix}\right) \le s_{2n-1}\left(\begin{bmatrix} P(\mu_\delta) & 0 \\ \gamma B & P(\mu_\delta) \end{bmatrix}\right) + \|P(\mu) - P(\mu_\delta)\| < 2\delta,$$

and the proof is complete. □

Corollary 14 *If $rank(P'(\mu)) \ge 2$, then as $\gamma \longrightarrow \infty$ $(\gamma \ge 0)$,*

$$s_{2n-1}(F[P(\mu);\gamma]) \longrightarrow 0 \quad \textit{and} \quad \beta_{low}(P,\mu,\gamma) \longrightarrow 0.$$

By this corollary, it is obvious that if $rank(P'(\mu)) \ge 2$, then there is a $\gamma_* \ge 0$ where the singular value $s_{2n-1}(F[P(\mu);\gamma])$ attains its maximum, $s_* = s_{2n-1}(F[P(\mu);\gamma_*])$. Moreover, since the leading coefficient mA_m of $P'(\lambda)$ is nonsingular, the spectrum $\sigma(P')$ has no more than $n(m-1)$ elements, and if $\mu \notin \sigma(P')$, then clearly $rank(P'(\mu)) = n \ge 2$.

The left and right singular vectors of $F[P(\mu);\gamma]$ corresponding to $s_{2n-1}(F[P(\mu);\gamma])$ possess a remarkable property, which will be useful in the sequel.

Lemma 15 *Let $\mu \in \mathbb{C}$ and $\gamma \ge 0$ such that $s_{2n-1}(F[P(\mu);\gamma]) > 0$, and let $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k,\, v_k \in \mathbb{C}^n,\ k = 1,2)$ be a pair of left and right singular vectors corresponding to $s_{2n-1}(F[P(\mu);\gamma])$, respectively. Then it holds that $u_2^* u_1 = v_2^* v_1$.*

Proof Let $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{C}^{2n}$ be a pair of left and right singular vectors of $s_{2n-1}(F[P(\mu);\gamma])$, respectively, i.e., they satisfy

$$\begin{bmatrix} P(\mu) & 0 \\ \gamma P'(\mu) & P(\mu) \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = s_{2n-1}(F[P(\mu);\gamma])\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{9}$$

and

$$[u_1^*\ u_2^*]\begin{bmatrix} P(\mu) & 0 \\ \gamma P'(\mu) & P(\mu) \end{bmatrix} = s_{2n-1}(F[P(\mu);\gamma])\,[v_1^*\ v_2^*]\,. \tag{10}$$

Multiplying (9) from the left by $[u_2^*\ 0]$ and (10) from the right by $\begin{bmatrix} 0 \\ v_1 \end{bmatrix}$, we get

$$u_2^* P(\mu) v_1 = s_{2n-1}(F[P(\mu);\gamma]) u_2^* u_1 \quad \text{and} \quad u_2^* P(\lambda) v_1 = s_{2n-1}(F[P(\mu);\gamma]) v_2^* v_1,$$

respectively. Since $s_{2n-1}(F[P(\mu);\gamma]) > 0$, it follows that $u_2^* u_1 = v_2^* v_1$. □

Next we obtain that for every $\mu \notin \sigma(P')$, the value $\gamma = \gamma_*$ ensures the condition $rank(V(\gamma)) = 2$. In the remainder of this paper, we need Lemma 5 of [11] (see also [14]).

Lemma 16 *Let $G(\zeta) \in \mathbb{C}^{n_1 \times n_2}$ be an analytic matrix function on an open set $\Gamma \subseteq \mathbb{R}$, and let $s_1(G(\zeta)) \ge s_2(G(\zeta)) \ge \cdots \ge s_{\min\{n_1,n_2\}}(G(\zeta)) \ge 0$ be its singular values. If $s_j(G(\zeta)) > 0$ at a local extremum $\zeta_* \in \Gamma$, then there is a pair of a left singular vector $u \in \mathbb{C}^{n_1}$ and a right singular vector $v \in \mathbb{C}^{n_2}$ of $G(\zeta_*)$ corresponding to $s_j(G(\zeta_*))$ such that $\mathrm{Re}(u^* G'(\zeta_*) v) = 0$.*

Applying this lemma to $F[P(\mu);\gamma]$ yields the following result.

Lemma 17 *Let $\mu \in \mathbb{C}$, $\gamma_* > 0$ be a point of local extremum of $s_{2n-1}(F[P(\mu);\gamma])$, and $s_* = s_{2n-1}(F[P(\mu);\gamma_*]) > 0$. Then there exists a pair $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k(\gamma_*),\ v_k(\gamma_*) \in \mathbb{C}^n,\ k=1,2)$ of left and right singular vectors of s_*, respectively, such that*

1. *$u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) = 0$, and*
2. *the $n \times 2$ matrices $U(\gamma_*) = [u_1(\gamma_*)\ u_2(\gamma_*)]$ and $V(\gamma_*) = [v_1(\gamma_*)\ v_2(\gamma_*)]$ satisfy $U(\gamma_*)^* U(\gamma_*) = V(\gamma_*)^* V(\gamma_*)$.*

Proof By Lemma 16, we know that there is a pair $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k(\gamma_*),\ v_k(\gamma_*) \in \mathbb{C}^n,\ k=1,2)$ of left and right singular vectors of s_*, respectively, such that

$$\begin{aligned} 0 &= \operatorname{Re}\left([u_1(\gamma_*)^*\ u_2(\gamma_*)^*] \frac{dF[P(\mu);\gamma_*]}{d\gamma} \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \right) \\ &= \operatorname{Re}\left([u_1(\gamma_*)^*\ u_2(\gamma_*)^*] \begin{bmatrix} 0 & 0 \\ P'(\mu) & 0 \end{bmatrix} \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \right) \\ &= \operatorname{Re}\left(u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) \right). \end{aligned}$$

Multiplying the relation (9) (for $\gamma = \gamma_*$) from the left by $[u_1(\gamma_*)^*\ \ -u_2(\gamma_*)^*]$ and the relation (10) from the right by $\begin{bmatrix} v_1(\gamma_*) \\ -v_2(\gamma_*) \end{bmatrix}$, we obtain

$$\begin{aligned} &u_1(\gamma_*)^* P(\mu) v_1(\gamma_*) - \gamma_* u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) - u_2(\gamma_*)^* P(\mu) v_2(\gamma_*) \\ &\qquad = s_*(u_1(\gamma_*)^* u_1(\gamma_*) - u_2(\gamma_*)^* u_2(\gamma_*)) \end{aligned}$$

and

$$\begin{aligned} &u_1(\gamma_*)^* P(\mu) v_1(\gamma_*) + \gamma_* u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) - u_2(\gamma_*)^* P(\mu) v_2(\gamma_*) \\ &\qquad = s_*(v_1(\gamma_*)^* v_1(\gamma_*) - v_2(\gamma_*)^* v_2(\gamma_*)), \end{aligned}$$

respectively. Then it follows

$$\begin{aligned} 2\gamma_* u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) &= s_*\,(v_1(\gamma_*)^* v_1(\gamma_*) - v_2(\gamma_*)^* v_2(\gamma_*) \\ &\quad - u_1(\gamma_*)^* u_1(\gamma_*) + u_2(\gamma_*)^* u_2(\gamma_*)), \end{aligned}$$

where the right hand side of the equation is a real number. Consequently, the number $u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*)$ is also real, and hence, $u_2(\gamma_*)^* P'(\mu) v_1(\gamma_*) = 0$. This means that

$$u_1(\gamma_*)^* u_1(\gamma_*) - u_2(\gamma_*)^* u_2(\gamma_*) = v_1(\gamma_*)^* v_1(\gamma_*) - v_2(\gamma_*)^* v_2(\gamma_*),$$

where $u_1(\gamma_*)^* u_1(\gamma_*) + u_2(\gamma_*)^* u_2(\gamma_*) = v_1(\gamma_*)^* v_1(\gamma_*) + v_2(\gamma_*)^* v_2(\gamma_*) = 1$. As a consequence,

$$u_1(\gamma_*)^* u_1(\gamma_*) = v_1(\gamma_*)^* v_1(\gamma_*) \quad \text{and} \quad u_2(\gamma_*)^* u_2(\gamma_*) = v_2(\gamma_*)^* v_2(\gamma_*). \tag{11}$$

By these equations and Lemma 15, it is straightforward to see that $U(\gamma_*)^* U(\gamma_*) = V(\gamma_*)^* V(\gamma_*)$. $\square$

Now we can prove that for every $\mu \notin \sigma(P')$, the matrices $U(\gamma_*)$ and $V(\gamma_*)$ can be chosen to be of full (column) rank.

Lemma 18 *If* $\mu \in \mathbb{C}\backslash\sigma(P')$, $\gamma_* > 0$, *and* $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix}$ *are the singular vectors of the previous lemma, then* $v_1(\gamma_*) \neq 0$ *and* $rank(U(\gamma_*)) = rank(V(\gamma_*)) = 2$.

Proof Both parts of the lemma will be proved by contradiction. For $\gamma = \gamma_* > 0$, (10) is written

$$[u_1(\gamma_*)^* \; u_2(\gamma_*)^*] \begin{bmatrix} P(\mu) & 0 \\ \gamma_* P'(\mu) & P(\mu) \end{bmatrix} = [s_* v_1(\gamma_*)^* \; s_* v_2(\gamma_*)^*] \, .$$

If we assume that $v_1(\gamma_*)^* = 0$, then the first equality in (11) implies $u_1(\gamma_*)^* = 0$. Thus,

$$\gamma_* u_2(\gamma_*)^* P'(\mu) \; = \; 0.$$

Since $\det P'(\mu) \neq 0$, it follows that $u_2(\gamma_*) = 0$. This is a contradiction because $u_1(\gamma_*)^* u_1(\gamma_*) + u_2(\gamma_*)^* u_2(\gamma_*) = 1$, and hence, $v_1(\gamma_*) \neq 0$.

Assume now that $rank(U(\gamma_*)) < 2$ or $rank(V(\gamma_*)) < 2$. Recall (11), and observe that $u_2(\gamma_*) = 0$ if and only if $v_2(\gamma_*) = 0$. In this case, (9) implies $\gamma_* P'(\mu) v_1(\gamma_*) = 0$. Since $P'(\mu)$ is invertible, it follows that $v_1(\gamma_*) = 0$; this is a contradiction. As a consequence, $u_2(\gamma_*)$ and $v_2(\gamma_*)$ are nonzero, and there is a scalar $c \neq 0$ such that $u_1(\gamma_*) = c\, u_2(\gamma_*)$ and $v_1(\gamma_*) = c\, v_2(\gamma_*)$. In this case, (9) yields

$$P(\mu) v_2(\gamma_*) = s_* u_2(\gamma_*) \quad \text{and} \quad \gamma_* P'(\mu) v_1(\gamma_*) + P(\mu) v_2(\gamma_*) = s_* u_2(\gamma_*),$$

and as a consequence, $P'(\mu) v_1(\gamma_*) = 0$. Since $\det P'(\mu) \neq 0$ and $v_1(\gamma_*) \neq 0$, we have a contradiction. $\square$

Before writing the upper bound of Theorem 11 for $\gamma = \gamma_*$, once again we explicitly construct a suitable perturbation of $P(\lambda)$ as in (5). In particular, for the pair of singular vectors $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \in \mathbb{C}^{2n}$ of Lemma 17, we define the matrix

$$\Delta_{\gamma_*} = -s_* U(\gamma_*) \begin{bmatrix} 1 & -\gamma_* \phi \\ 0 & 1 \end{bmatrix} V(\gamma_*)^\dagger$$

(recall that $s_* = s_{2n-1}(F[P(\mu); \gamma_*]) > 0$ and $rank(V(\gamma_*)) = 2$) and the associated perturbation

$$Q_{\gamma_*}(\lambda) = \sum_{j=0}^{m} (A_j + \Delta_{\gamma_*,j}) \lambda^j = \sum_{j=0}^{m} \left(A_j + \frac{w_j}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|} \right)^j \Delta_{\gamma_*} \right) \lambda^j. \qquad (12)$$

From the relation $U(\gamma_*)^* U(\gamma_*) = V(\gamma_*)^* V(\gamma_*)$ of Lemma 17, it follows that the $n \times 2$ matrices $U(\gamma_*)$ and $V(\gamma_*)$ have the same nonzero singular values and the same associated right singular vectors. Thus, there exists an $n \times n$ unitary matrix W such that $U(\gamma_*) = WV(\gamma_*)$. Consequently, the upper bound (8) for the distance $\mathcal{E}_a(\mu)$ is

$$\begin{aligned} \beta_{up}(P,\mu,\gamma_*) &= \frac{s_*}{w(|\mu|)} \left\| U(\gamma_*) \begin{bmatrix} 1 & -\gamma_* \phi \\ 0 & 1 \end{bmatrix} V(\gamma_*)^\dagger \right\| \\ &= \frac{s_*}{w(|\mu|)} \left\| V(\gamma_*) \begin{bmatrix} 1 & -\gamma_* \phi \\ 0 & 1 \end{bmatrix} V(\gamma_*)^\dagger \right\|, \end{aligned}$$

and, keeping in mind Lemma 18, we have the main result of this section.

Theorem 19 *Suppose that $\mu \in \mathbb{C}\backslash\sigma(P')$, $\gamma_* > 0$ is a point of maximum value of $s_{2n-1}(F[P(\mu);\gamma])$, and $s_* = s_{2n-1}(F[P(\mu);\gamma_*]) > 0$. Then there exists a pair $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix} \in \mathbb{C}^{2n}$ $(u_k(\gamma_*),\ v_k(\gamma_*) \in \mathbb{C}^n,\ k = 1, 2)$ of left and right singular vectors of s_*, respectively, such that*

$$\mathcal{E}_a(\mu) \leq \beta_{up}(P,\mu,\gamma_*) = \frac{s_*}{w(|\mu|)} \left\| V(\gamma_*) \begin{bmatrix} 1 & -\gamma_* \phi \\ 0 & 1 \end{bmatrix} V(\gamma_*)^\dagger \right\|,$$

and the matrix polynomial $Q_{\gamma_*}(\lambda)$ in (12) lies on $\partial\mathcal{B}(P, \beta_{up}(P,\mu,\gamma_*), \mathrm{w})$ and has μ as a defective eigenvalue.

If the singular value s_* of the matrix $F[P(\mu);\gamma_*]$ is simple, then the pair of singular vectors $\begin{bmatrix} u_1(\gamma_*) \\ u_2(\gamma_*) \end{bmatrix}, \begin{bmatrix} v_1(\gamma_*) \\ v_2(\gamma_*) \end{bmatrix}$ in the above theorem can be chosen arbitrarily (as far as they correspond to the same SVD of $F[P(\mu);\gamma_*]$). On the other hand, if s_* is a multiple singular value, then we can estimate these singular vectors by using the second part of the proof of [11, Lemma 5].

6 The non-generic case $\gamma_* = 0$

Suppose that the singular value $s_{2n-1}(F[P(\mu);\gamma])$ attains its maximum at $\gamma_* = 0$, and consider the matrix

$$F[P(\mu);\gamma_*] = F[P(\mu);0] = \begin{bmatrix} P(\mu) & 0 \\ 0 & P(\mu) \end{bmatrix}.$$

The condition $s_* = s_{2n-1}(F[P(\mu);0]) > 0$ implies that $s_* = s_n(P(\mu)) > 0$, i.e., $\mu \notin \sigma(P)$, and we have two cases (with respect to the singular values of $P(\mu)$), namely,

$$s_* = s_n(P(\mu)) = s_{n-1}(P(\mu)) \quad \text{and} \quad s_* = s_n(P(\mu)) < s_{n-1}(P(\mu)).$$

Case 1: $s_* = s_{n-1}(P(\mu)) = s_n(P(\mu))$.

This case was already discussed at the beginning of Section 4, and we have that

$$\mathcal{E}_a(\mu) = \mathcal{E}_{g,2}(\mu) = \mathcal{E}_{g,1}(\mu) = \frac{s_n(P(\mu))}{w(|\mu|)},$$

and an optimal perturbation of $P(\lambda)$ is the matrix polynomial $\hat{Q}(\lambda)$ in (3) (for $\kappa = 2$).

Case 2: $s_* = s_n(P(\mu)) < s_{n-1}(P(\mu))$.

Let $u, v \in \mathbb{C}^n$ be a pair of left and right singular vectors of $P(\mu)$ corresponding to $s_* = s_n(P(\mu))$, respectively. First we obtain that $u^* P'(\mu) v = 0$, following the steps of Malyshev's methodology [11]. Only here we use the fact that the local extremum $\gamma_* = 0$ is a maximum.

Consider the analytic matrix function (with respect to $\gamma \in \mathbb{R}$)

$$F[P(\mu);\gamma] = \begin{bmatrix} P(\mu) & 0 \\ \gamma P'(\mu) & P(\mu) \end{bmatrix}$$

for such small $|\gamma| > 0$, where $s_{2n-2}(F[P(\mu);\gamma]) > s_{2n-1}(F[P(\mu);\gamma])$. We denote this range of γ by Γ. (Note that the definition of Γ is always possible since $s_{2n-1}(F[P(\mu);0]) = s_n(P(\mu)) < s_{n-1}(P(\mu)) = s_{2n-2}(F[P(\mu);0])$.) By Theorem S6.3 of [4] (see also Theorem II-6.1 of [9]), $F[P(\mu);\gamma]$ has analytic unordered singular values $\tilde{s}_{2n-1}(\gamma)$ and $\tilde{s}_{2n}(\gamma)$ satisfying $\tilde{s}_{2n-1}(0) = \tilde{s}_{2n}(0) = s_*$. Without loss of generality, the neighborhood Γ can be chosen sufficiently small such that for every $\gamma \in \Gamma$, $\tilde{s}_{2n-1}(\gamma)$ and $\tilde{s}_{2n}(\gamma)$ are not greater than s_*. We also consider a pair $\tilde{u}_{2n-1}(\gamma), \tilde{v}_{2n-1}(\gamma)$ of left and right singular vectors of $\tilde{s}_{2n-1}(\gamma)$, and a pair $\tilde{u}_{2n}(\gamma), \tilde{v}_{2n}(\gamma)$ of left and right singular vectors of $\tilde{s}_{2n}(\gamma)$. All singular vectors are analytic and with respect to the same SVD of the matrix $F[P(\mu);\gamma]$.

Since $\tilde{s}_{2n-1}(0) = \tilde{s}_{2n}(0) = s_* = s_n(P(\mu)) < s_{n-1}(P(\mu))$, it follows that

$$\tilde{u}_{2n-1}(0) = \begin{pmatrix} w_{11}u \\ w_{21}u \end{pmatrix}, \quad \tilde{u}_{2n}(0) = \begin{pmatrix} w_{12}u \\ w_{22}u \end{pmatrix},$$

$$\tilde{v}_{2n-1}(0) = \begin{pmatrix} w_{11}v \\ w_{21}v \end{pmatrix} \quad \text{and} \quad \tilde{v}_{2n}(0) = \begin{pmatrix} w_{12}v \\ w_{22}v \end{pmatrix},$$

where the matrix $\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \in \mathbb{C}^{2\times 2}$ is unitary. Then there exists a unit vector $\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in \mathbb{C}^2$ such that

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{1 + |u^* P'(\mu) v|^2}} \begin{bmatrix} 1 \\ u^* P'(\mu) v \end{bmatrix}.$$

Consider the unit vectors

$$x(\gamma) = \alpha\, \tilde{u}_{2n-1}(\gamma) + \beta\, \tilde{u}_{2n}(\gamma) \quad \text{and} \quad y(\gamma) = \alpha\, \tilde{v}_{2n-1}(\gamma) + \beta\, \tilde{v}_{2n}(\gamma),$$

for which $x(0), y(0)$ is a pair of left and right singular vectors of $F[P(\mu);0]$ corresponding to s_*. Then for every $\gamma \in \Gamma$,

$$\begin{aligned} F[P(\mu);\gamma] y(\gamma) &= \alpha\, \tilde{s}_{2n-1}(\gamma) \tilde{u}_{2n-1}(\gamma) + \beta\, \tilde{s}_{2n}(\gamma) \tilde{u}_{2n}(\gamma) \\ &= [\tilde{u}_{2n-1}(\gamma) \;\; \tilde{u}_{2n}(\gamma)] \begin{bmatrix} \alpha\, \tilde{s}_{2n-1}(\gamma) \\ \beta\, \tilde{s}_{2n}(\gamma) \end{bmatrix}. \end{aligned}$$

Hence, since $\tilde{s}^2_{2n-1}(\gamma)$ and $\tilde{s}^2_{2n}(\gamma)$ are analytic with local maximum at $\gamma = 0$, it follows

$$\begin{aligned} 0 &= \frac{d}{d\gamma}\left(|\alpha|^2\tilde{s}^2_{2n-1}(\gamma) + |\beta|^2\tilde{s}^2_{2n}(\gamma)\right)\Big|_{\gamma=0} \\ &= \frac{d}{d\gamma}(y(\gamma)^* F[P(\mu);\gamma]^* F[P(\mu);\gamma] y(\gamma))\Big|_{\gamma=0}. \end{aligned}$$

The condition $y(\gamma)^*y(\gamma) = 1$ implies that $\frac{dy^*(0)}{d\gamma}y(0) + y^*(0)\frac{dy(0)}{d\gamma} = 0$, and differentiating $y(\gamma)^*F[P(\mu);\gamma]^*F[P(\mu);\gamma]y(\gamma)$ at $\gamma = 0$ yields

$$\begin{aligned} 0 &= \frac{d}{d\gamma}\left(y(\gamma)^* F[P(\mu);\gamma]^* F[P(\mu);\gamma] y(\gamma)\right)\Big|_{\gamma=0} \\ &= \Bigg(\frac{dy(\gamma)^*}{d\gamma}F[P(\mu);\gamma]^*F[P(\mu);\gamma]y(\gamma) + y(\gamma)^*\frac{dF[P(\mu);\gamma]^*}{d\gamma}F[P(\mu);\gamma]y(\gamma) \\ &\quad + y(\gamma)^*F[P(\mu);\gamma]^*\frac{dF[P(\mu);\gamma]}{d\gamma}y(\gamma) + y(\gamma)^*F[P(\mu);\gamma]^*F[P(\mu);\gamma]\frac{dy(\gamma)}{d\gamma}\Bigg)\Bigg|_{\gamma=0} \\ &= s_*^2\left(\frac{dy(0)^*}{d\gamma}y(0) + y(0)^*\frac{dy(0)}{d\gamma}\right) \\ &\quad + s_*\left(y(0)^*\begin{bmatrix} 0 & (P'(\mu))^* \\ 0 & 0 \end{bmatrix}x(0) + x(0)^*\begin{bmatrix} 0 & 0 \\ P'(\mu) & 0 \end{bmatrix}y(0)\right) \\ &= s_*\left(y(0)^*\begin{bmatrix} 0 & (P'(\mu))^* \\ 0 & 0 \end{bmatrix}x(0) + x(0)^*\begin{bmatrix} 0 & 0 \\ P'(\mu) & 0 \end{bmatrix}y(0)\right). \end{aligned}$$

Furthermore, we can see that

$$\begin{aligned} x(0) &= \alpha\begin{bmatrix} w_{11}u \\ w_{21}u \end{bmatrix} + \beta\begin{bmatrix} w_{12}u \\ w_{22}u \end{bmatrix} = u \otimes \left(\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix}\right) \\ &= u \otimes \left(\frac{1}{\sqrt{1+|u^*P'(\mu)v|^2}}\begin{bmatrix} 1 \\ u^*P'(\mu)v \end{bmatrix}\right) \\ &= \frac{1}{\sqrt{1+|u^*P'(\mu)v|^2}}\begin{bmatrix} u \\ (u^*P'(\mu)v)u \end{bmatrix}, \end{aligned}$$

and similarly,

$$y(0) = \frac{1}{\sqrt{1+|u^*P'(\mu)v|^2}}\begin{bmatrix} v \\ (u^*P'(\mu)v)v \end{bmatrix}.$$

As a consequence,

$$y(0)^*\begin{bmatrix} 0 & (P'(\mu))^* \\ 0 & 0 \end{bmatrix}x(0) = x(0)^*\begin{bmatrix} 0 & 0 \\ P'(\mu) & 0 \end{bmatrix}y(0) = \frac{|u^*P'(\mu)v|^2}{1+|u^*P'(\mu)v|^2},$$

and thus,

$$0 = \frac{d}{d\gamma}(y(\gamma)^*F[P(\mu);\gamma]^*F[P(\mu);\gamma]y(\gamma))\Big|_{\gamma=0} = 2s_*\frac{|u^*P'(\mu)v|^2}{1+|u^*P'(\mu)v|^2},$$

which implies $u^*P'(\mu)v = 0$.

We define now the constant matrix polynomial

$$\Delta_0(\lambda) = \Delta_{0,0} = -s_* uv^*$$

with $\|\Delta_{0,0}\| = s_* = w_0(s_*/w_0)$ (recall that $w_0 > 0$). Then the perturbation

$$Q_0(\lambda) \;=\; P(\lambda) + \Delta_{0,0} \;=\; A_m\lambda^m + \cdots + A_1\lambda + A_0 + \Delta_{0,0} \tag{13}$$

of $P(\lambda)$ lies on the boundary of $\mathcal{B}(P, s_*/w_0, \mathrm{w})$, and satisfies

$$Q_0(\mu)v = P(\mu)v - s_* uv^* v = s_* u - s_* u = 0,$$

$$u^* Q_0(\mu) = u^* P(\mu) - s_* u^* uv^* = s_* v^* - s_* v^* = 0$$

and

$$u^* Q_0'(\mu)v = u^* P'(\mu)v = 0.$$

Thus, by Proposition 16 of [1], μ is a multiple eigenvalue of $Q_0(\lambda)$.

The main results of this section can be summarized in the following.

Theorem 20 *If $\mu \in \mathbb{C}\backslash\sigma(P)$, $s_{2n-1}(F[P(\mu);\gamma])$ attains a maximum value at $\gamma_* = 0$ and $s_* = s_{2n-1}(F[P(\mu);0]) = s_n(P(\mu))$ (> 0), then $\mathcal{E}_a(\mu) \le s_*/w_0$. Furthermore, if $u, v \in \mathbb{C}^n$ is a pair of left and right singular vectors of $P(\mu)$ corresponding to s_*, respectively, then the matrix polynomial $Q_0(\lambda)$ in (13) lies on $\partial\mathcal{B}(P, s_*/w_0, \mathrm{w})$ and has μ as a multiple eigenvalue.*

Remark 21 If we allow perturbations only of the constant coefficient of $P(\lambda)$, i.e., if $w_0 > 0$ and $w_1 = w_2 = \cdots = w_m = 0$, then $w(|\mu|) = w_0$ and $\phi = w'(|\mu|) = 0$. As a consequence, if the singular value $s_{2n-1}(F[P(\mu);\gamma])$ attains its maximum s_* at $\gamma_* \ge 0$, then the definition of $\beta_{low}(P, \mu, \gamma)$ in (7), and Theorems 19 and 20 imply that

$$\mathcal{E}_a(\mu) \;=\; \beta_{low}(P, \mu, \gamma_*) \;=\; \beta_{up}(P, \mu, \gamma_*) \;=\; \frac{s_*}{w_0}.$$

Moreover, an optimal perturbation of $P(\lambda)$ that lies on $\partial\mathcal{B}(P, s_*/w_0, \mathrm{w})$ and has μ as a multiple eigenvalue is given by (12) when $\gamma_* > 0$, and by (13) when $\gamma_* = 0$.

7 Connection with Malyshev's results

Suppose that the matrix polynomial $P(\lambda)$ is of the form $P(\lambda) = I\lambda - A$ for some $A \in \mathbb{C}^{n\times n}$, and the set of weights is $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, i.e., we consider the standard eigenproblem associated to matrix A. Then obviously,

$$P'(\lambda) = I, \;\; w(|\mu|) = w_0 = 1 \;\text{ and }\; \phi = w'(|\mu|) = 0$$

(see also Remark 21). The existence of γ_* is ensured, and if $\gamma_* > 0$, then the upper bound of Theorem 19 is $\beta_{up}(P, \mu, \gamma_*) = s_* \left\| V(\gamma_*)V(\gamma_*)^\dagger \right\| = s_*$ and coincides with the lower bound $\beta_{low}(P, \mu, \gamma_*)$. Hence, the distance from $P(\lambda) = I\lambda - A$ (or equivalently, from matrix A) to $\mu \in \mathbb{C}$ as a multiple eigenvalue is $\mathcal{E}_a(\mu) = s_*$, and the perturbation

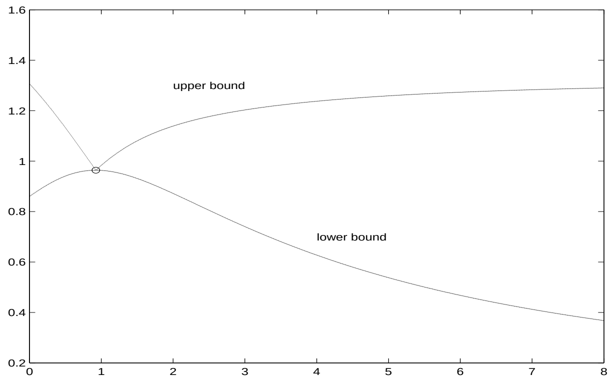

Figure 1: The graphs of the bounds $\beta_{up}(P, -2+i, \gamma)$ and $\beta_{low}(P, -2+i, \gamma)$.

$Q_{\gamma_*}(\lambda)$ in (12) is written $Q_{\gamma_*}(\lambda) = I\lambda - \big(A + s_* U(\gamma_*)V(\gamma_*)^\dagger\big)$. Moreover, if $\gamma_* = 0$, then the upper bound of Theorem 20 and the lower bound $\beta_{low}(P, \mu, 0)$ are equal to s_*. Thus, the distance from $P(\lambda) = I\lambda - A$ to $\mu \in \mathbb{C}$ as a multiple eigenvalue is $\mathcal{E}_a(\mu) = s_*$, and the perturbation $Q_0(\lambda)$ in (13) is written $Q_0(\lambda) = I\lambda - (A + s_* uv^*)$. This means that our results in the previous three sections are direct generalizations of Malyshev's results [11] to the case of matrix polynomials.

For example, we consider the matrix

$$A = \begin{bmatrix} -1 & -2 & -3 \\ 0 & -4 & -5 \\ 0 & 0 & -6+i \end{bmatrix},$$

the corresponding linear matrix polynomial $P(\lambda) = I\lambda - A$ and the scalar $\mu = -2+i$. The graphs of the upper bound $\beta_{up}(P, -2+i, \gamma)$ and the lower bound $\beta_{low}(P, -2+i, \gamma)$ ($\gamma \in [0, 8]$) are plotted in Figure 1, and they meet at the point $(0.9249, 0.9639)$, which is marked as "o". Consequently, $\gamma_* = 0.9249$ and the maximum value of $s_5(F[P(-2+i); \gamma])$ is $s_* = s_5(F[P(-2+i); 0.9249]) = 0.9639$. By Theorem 11, it follows that $\mathcal{E}_a(-2+i) = 0.9639$, the matrix polynomial

$$\begin{aligned} Q_{0.9249}(\lambda) &= I\lambda - (A + \Delta) \\ &= I\lambda - \begin{bmatrix} -1.3608 + i\,0.7100 & -1.7947 - i\,0.1331 & -3.0561 + i\,0.0407 \\ 0.3571 - i\,0.3035 & -3.7284 + i\,0.3732 & -5.1005 - i\,0.0209 \\ -0.1042 - i\,0.2500 & -0.6269 - i\,0.4520 & -5.7867 + i\,0.9972 \end{bmatrix} \end{aligned}$$

(or equivalently, the matrix $A + \Delta$) has $\mu = -2+i$ as a defective eigenvalue, and $\|\Delta\| = 0.9639$.

8 Numerical examples

We present three numerical examples to illustrate our results and verify the quality of the bounds $\beta_{up}(P, \mu, \gamma)$ and $\beta_{low}(P, \mu, \gamma)$. The matrix polynomials of the first two examples were borrowed from [1], and all the computations were performed in Matlab.

For our discussion, it is necessary to recall the definition of the *ε-pseudospectrum* of the matrix polynomial $P(\lambda)$,

$$\begin{aligned}\sigma_{\varepsilon,\mathrm{w}}(P) &= \{\lambda \in \mathbb{C} : \det Q(\lambda) = 0,\ Q(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})\} \\ &= \{\lambda \in \mathbb{C} : \det Q(\lambda) = 0,\ \|\Delta_j\| \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m\},\end{aligned}$$

i.e., the set of the eigenvalues of all perturbations of $P(\lambda)$ in $\mathcal{B}(P, \varepsilon, \mathrm{w})$. The pseudospectrum $\sigma_{\varepsilon,\mathrm{w}}(P)$ is a closed subset of the complex plane, has no more than nm connected components, and it is bounded if and only if $s_n(A_m) > \varepsilon\, w_m$. The suggested references on pseudospectra of matrix polynomials are [1, 10, 15]. The following result of Boulton, Lancaster and Psarrakos (see Lemma 8, Corollary 15 and Theorem 18 (i) of [1]) is also necessary.

Proposition 22 *Suppose that, as the parameter $\varepsilon > 0$ increases, two different connected components of $\sigma_{\varepsilon,\mathrm{w}}(P)$ ($\neq \mathbb{C}$) meet at $\mu \in \mathbb{C}$. If $\mu \neq 0$, then it is a multiple eigenvalue of a perturbation $Q(\lambda) \in \partial\mathcal{B}(P, \varepsilon, \mathrm{w})$ and*

$$\mathcal{E}_a(\mu) \;=\; \mathcal{E}_{g,1}(\mu) = \frac{s_n(P(\mu))}{w(|\mu|)} \quad (= \varepsilon).$$

The special case of self-intersection points of pseudospectra described in this proposition is the only known to the authors (non-trivial) case of scalars $\mu \neq 0$ where one can estimate the true value of the distance $\mathcal{E}_a(\mu)$, and it is illustrated in the first two examples. It is also worth noting that in this special case, we always have $\mathcal{E}_a(\mu) = \beta_{low}(P, \mu, 0)$, i.e., the maximum of our lower bound coincides with the exact value of the distance.

Example 1 Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} -2 & 1 \\ 0 & -4 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

and the set of weights $\mathrm{w} = \{1, 1, 1\}$. The boundaries of the ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.005, 0.0091, 0.02, 0.03$ are drawn in Figure 2. The eigenvalues of $P(\lambda)$, 1 and 2, are marked in the figure as "+", $\sigma_{0.005,\mathrm{w}}(P)$ has two connected components and $\sigma_{0.0091,\mathrm{w}}(P)$ is connected with a node point $\mu = 1.4145$ (marked as an asterisk). By Proposition 22 and the relative discussion, $\mu = 1.4145$ is a multiple eigenvalue of a matrix polynomial $Q(\lambda) \in \partial\mathcal{B}(P, 0.0091, \mathrm{w})$ and

$$\mathcal{E}_a(1.4145) \;=\; \mathcal{E}_{g,1}(1.4145) \;=\; 0.0091 \;=\; \frac{s_2(P(1.4145))}{w(1.4145)} \;=\; \beta_{low}(P, 1.4145, 0).$$

In Figure 3, the graphs of the upper bound $\beta_{up}(P, 1.4145, \gamma)$ and the lower bound $\beta_{low}(P, 1.4145, \gamma)$ are plotted for $\gamma \in [0, 8]$. The vertical line corresponds to the value $\gamma_* = 0.7738$, and the bounds $\beta_{up}(P, 1.4145, 0.7738)$ and $\beta_{low}(P, 1.4145, 0.7738)$ are marked as "o". The maximum value of $s_3(F[P(1.4145); \gamma])$ is

$$s_* = s_3(F[P(1.4145); 0.7738]) = 0.0471,$$

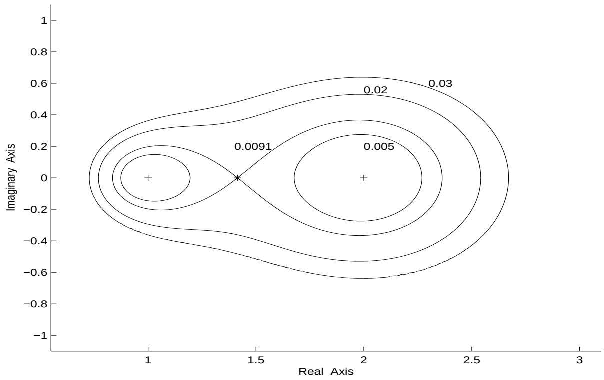

Figure 2: A single intersection point.

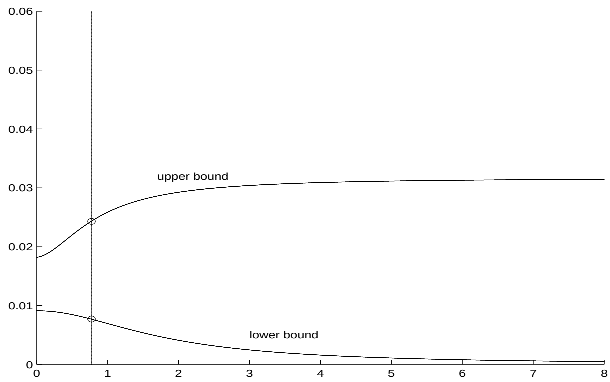

Figure 3: The graphs of the bounds $\beta_{up}(P, 1.4145, \gamma)$ and $\beta_{low}(P, 1.4145, \gamma)$.

and one can see that

$$\begin{aligned} 0.0077 \;=\; \beta_{low}(P, 1.4145, 0.7738) \;&\leq\; \mathcal{E}_a(1.4145) \;=\; 0.0091 \\ &\leq\; \beta_{up}(P, 1.4145, 0.7738) \;=\; 0.0243. \end{aligned}$$

A matrix polynomial that lies on the boundary of $\mathcal{B}(P, 0.0243, \mathrm{w})$ and has $\mu = 1.4145$ as a defective eigenvalue of algebraic multiplicity 2 and geometric multiplicity 1, with associated eigenvector $v_1(0.7738) = \begin{bmatrix} 0.8051 \\ -0.1061 \end{bmatrix}$, is

$$Q_{0.7738}(\lambda) = \begin{bmatrix} 1.0054 & 0.0156 \\ 0.0121 & 1.0140 \end{bmatrix} \lambda^2 + \begin{bmatrix} -1.9946 & 1.0156 \\ 0.0121 & -3.9860 \end{bmatrix} \lambda + \begin{bmatrix} 1.0054 & 0.0156 \\ 0.0121 & 4.0140 \end{bmatrix}.$$

This matrix polynomial is given by (5) and (12), and it is directly computable by the procedures described in Sections 4 and 5. Thus, Theorems 11 and 19 are confirmed. Furthermore,

$$|u_2(0.7738)^* P'(1.4145) v_1(0.7738)| \;=\; 1.4471 \cdot 10^{-10}$$

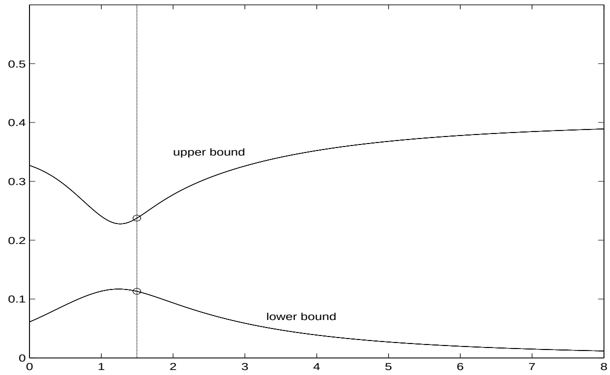

Figure 4: The graphs of the bounds $\beta_{up}(P,3,\gamma)$ and $\beta_{low}(P,3,\gamma)$.

and

$$\|U(0.7738)^*U(0.7738) - V(0.7738)^*V(0.7738)\| \;=\; 2.3778\cdot 10^{-9},$$

verifying Lemma 17.

Consider now the scalar $\mu = 3$. The graphs of the upper bound $\beta_{up}(P,3,\gamma)$ and the lower bound $\beta_{low}(P,3,\gamma)$ ($\gamma\in[0,8]$) are plotted in Figure 4. The vertical line corresponds to the value $\gamma_* = 1.4952$ and the bounds $\beta_{up}(P,3,1.4952)$ and $\beta_{low}(P,3,1.4952)$ are marked as "o". In this figure, we see that

$$0.1131 \;=\; \beta_{low}(P,3,1.4952) \;\le\; \mathcal{E}_a(3) \;\le\; \beta_{up}(P,3,1.4952) \;=\; 0.2375.$$

The perturbation of $P(\lambda)$ in (12) that lies on $\partial\mathcal{B}(P,0.2375,\mathrm{w})$ and has $\mu=3$ as a defective eigenvalue of algebraic multiplicity 2 and geometric multiplicity 1, with associated eigenvector $v_1(1.4952) = \begin{bmatrix} 0.5845 \\ -0.3104 \end{bmatrix}$, is

$$Q_{1.4952}(\lambda) = \begin{bmatrix} 0.7840 & -0.0581 \\ -0.0328 & 0.8614 \end{bmatrix}\lambda^2 + \begin{bmatrix} -2.2160 & 0.9419 \\ -0.0328 & -4.1386 \end{bmatrix}\lambda + \begin{bmatrix} 0.7840 & -0.0581 \\ -0.0328 & 3.8614 \end{bmatrix}.$$

By Figure 4, it is also clear that for every $\gamma\in(0,2.9)$, the lower bound $\beta_{low}(P,3,\gamma)$ is greater than $\mathcal{E}_{g,1}(3) = s_2(P(3))/w(3) = \beta_{low}(P,3,0) = 0.0611$. □

In the first part of the following example, we consider a self-intersection point of a pseudospectrum where both the minimum of our upper bound and the maximum of our lower bound coincide with the true value of the distance $\mathcal{E}_a(\mu)$.

Example 2 Let

$$P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\lambda^2 + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & 6 \end{bmatrix}\lambda + \begin{bmatrix} 2 & 1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 10 \end{bmatrix}$$

and let $\mathrm{w} = \{10, 6.1108, 3\}$ (the norms of the coefficient matrices). Figure 5 contains the boundaries of the ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.05, 0.1002, 0.16$. The eigen-

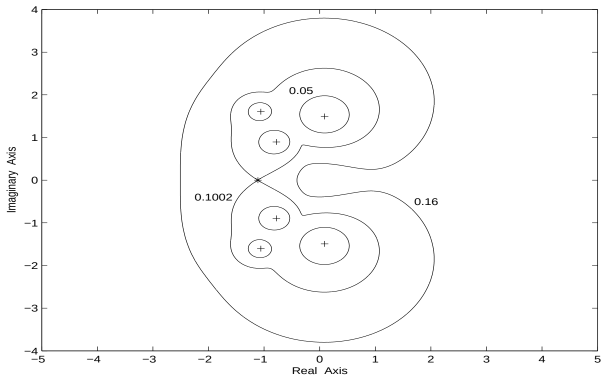

Figure 5: A single intersection point.

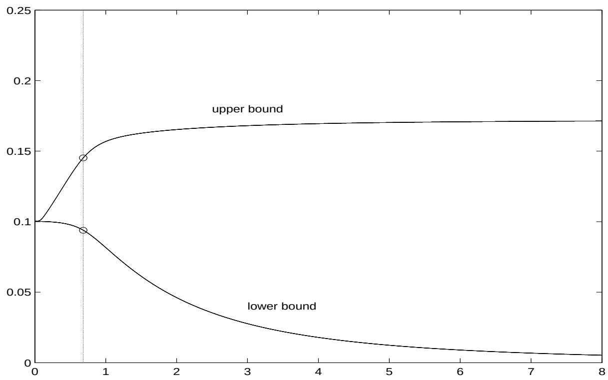

Figure 6: The graphs of the bounds $\beta_{up}(P, -1.1105, \gamma)$ and $\beta_{low}(P, -1.1105, \gamma)$.

values of $P(\lambda)$, $0.0877 \pm i\, 1.4940$, $-1.0590 \pm i\, 1.6051$ and $-0.7787 \pm i\, 0.8958$, are marked as "+", $\sigma_{0.05,\mathrm{w}}(P)$ has six connected components and $\sigma_{0.1002,\mathrm{w}}(P)$ is connected with a node point $\mu = -1.1105$ (marked as an asterisk). As in Example 1,

$$\mathcal{E}_a(-1.1105) \;=\; \mathcal{E}_{g,1}(-1.1105) \;=\; 0.1002 \;=\; \frac{s_3(P(-1.1105))}{w(1.1105)} \;=\; \beta_{low}(P, -1.1105, 0).$$

The graphs of the bounds $\beta_{up}(P, -1.1105, \gamma)$ and $\beta_{low}(P, -1.1105, \gamma)$ for $\gamma \in [0, 8]$ are drawn in Figure 6, where the vertical line corresponds to the value $\gamma_* = 0.6824$. The maximum value of $s_5(F[P(-1.1105); \gamma])$ is $s_* = s_5(F[P(-1.1105); 0.6824]) = 2.3769$, and

$$\begin{aligned} 0.0939 \;=\; \beta_{low}(P, -1.1105, 0.6824) \;&\leq\; \mathcal{E}_a(-1.1105) \;=\; 0.1002 \\ &\leq\; \beta_{up}(P, -1.1105, 0.6824) \;=\; 0.1452. \end{aligned}$$

It is worth mentioning that since the matrices

$$V(0) \;=\; \begin{bmatrix} -0.3393 & 0 \\ -0.9334 & 0 \\ 0.1173 & 0 \end{bmatrix} \quad \text{and} \quad U(0) \;=\; \begin{bmatrix} 0.4841 & 0 \\ 0.8689 & 0 \\ 0.1029 & 0 \end{bmatrix}$$

satisfy $\left\|U(0)V(0)^\dagger\right\| = 1$,

$$\beta_{up}(P, -1.1105, 0) \;=\; \beta_{low}(P, -1.1105, 0) \;=\; \mathcal{E}_a(-1.1105) \;=\; 0.1002$$

(recall the commentary at the end of Section 4). Furthermore, (5) yields a matrix polynomial that lies on the boundary of $\mathcal{B}(P, 0.1002, \mathrm{w})$ and has $\mu = -1.1105$ as a defective eigenvalue of algebraic multiplicity 2 and geometric multiplicity 1, with corresponding eigenvector $v_1(0) = \begin{bmatrix} -0.3393 \\ -0.9334 \\ 0.1173 \end{bmatrix}$. This matrix polynomial is

$$\begin{aligned} Q_0(\lambda) \;&=\; \begin{bmatrix} 0.9506 & -0.1358 & 0.0171 \\ -0.0886 & 1.7562 & 0.0306 \\ -0.0105 & -0.0289 & 3.0036 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0.1006 & 1.2767 & -0.0348 \\ 0.1805 & 3.4966 & 0.9376 \\ 0.0214 & -0.9412 & 5.9926 \end{bmatrix} \lambda \\ &\quad + \begin{bmatrix} 1.8354 & 0.5472 & 0.0569 \\ -1.2954 & 2.1874 & 0.1021 \\ -0.0350 & -0.0963 & 10.0121 \end{bmatrix}. \end{aligned}$$

If we set $\mu = 3 + i$, then by Theorem 4, the distance from $P(\lambda)$ to $3 + i$ as an eigenvalue of geometric multiplicity 2 is

$$\mathcal{E}_{g,2}(3+i) = \frac{s_2(P(3+i))}{w(|3+i|)} \;=\; \frac{32.1524}{59.3240} = 0.5420.$$

Using the methodology proposed in Section 3, we obtain the matrix polynomial

$$\begin{aligned} \hat{Q}(\lambda) \;&=\; \begin{bmatrix} 0.4134 + i\,0.0607 & -0.1918 + i\,0.0808 & -0.0090 - i\,0.0002 \\ 0.0403 - i\,0.0302 & 0.4100 + i\,0.2467 & -0.0611 - i\,0.0173 \\ -0.0037 + i\,0.0010 & 0.0599 + i\,0.0176 & 3.0019 + i\,0.0016 \end{bmatrix} \lambda^2 \\ &\quad + \begin{bmatrix} -1.1726 - i\,0.2606 & 0.5773 + i\,0.0325 & -0.0172 - i\,0.0062 \\ 0.0974 - i\,0.0325 & -0.2315 - i\,0.5476 & 0.8931 - i\,0.0727 \\ -0.0077 - i\,0.0004 & -0.8955 + i\,0.0727 & 6.0025 + i\,0.0043 \end{bmatrix} \lambda \\ &\quad + \begin{bmatrix} 0.3144 - i\,1.0114 & 0.3269 - i\,0.1683 & -0.0235 - i\,0.0185 \\ -0.8319 + i\,0.0000 & -1.7334 - i\,2.5223 & -0.1283 - i\,0.1682 \\ -0.0118 - i\,0.0046 & 0.1245 + i\,0.1669 & 10.0017 + i\,0.0080 \end{bmatrix} \end{aligned}$$

in (3), which lies on $\partial\mathcal{B}(P, 0.5420, \mathrm{w})$ and has $\mu = 3 + i$ as an eigenvalue of geometric multiplicity 2.

The graphs of the bounds $\beta_{up}(P, 3+i, \gamma)$ and $\beta_{low}(P, 3+i, \gamma)$ ($\gamma \in [0, 8]$) are plotted in Figure 7. The vertical line corresponds to the value $\gamma = 1.9$, which is different than $\gamma_* = 2.0680$, and the bounds $\beta_{up}(P, 3+i, 1.9)$ and $\beta_{low}(P, 3+i, 1.9)$ are marked as "o". It is straightforward to see that

$$0.2149 \;=\; \beta_{low}(P, 3+i, 1.9) \;\leq\; \mathcal{E}_a(3+i) \;\leq\; \beta_{up}(P, 3+i, 1.9) \;=\; 0.4901,$$

where our upper bound is smaller than $\mathcal{E}_{g,2}(3+i) = 0.5420$. The perturbation of $P(\lambda)$ in (5) that lies on $\partial\mathcal{B}(P, 0.4901, \mathrm{w})$ and has $\mu = 3 + i$ as a defective eigenvalue

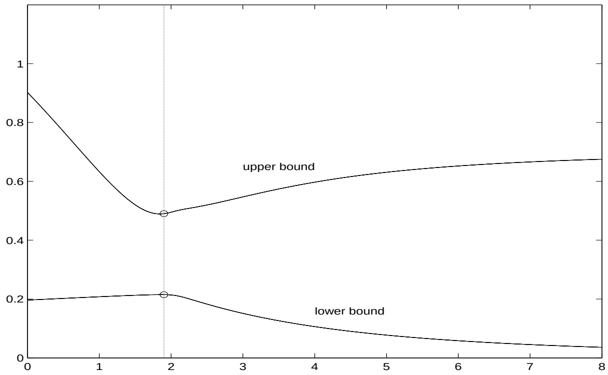

Figure 7: The graphs of the bounds $\beta_{up}(P, 3+i, \gamma)$ and $\beta_{low}(P, 3+i, \gamma)$.

of algebraic multiplicity 2 and geometric multiplicity 1, with associated eigenvector $v_1(0.6824) = \begin{bmatrix} 0.8076 \\ 0.1593 - i\,0.0078 \\ 0.0050 - i\,0.0028 \end{bmatrix}$, is

$$
\begin{aligned}
Q_{1.9}(\lambda) \;&=\; \begin{bmatrix} 0.3586 + i\,0.0490 & 0.0819 + i\,0.1531 & 0.0015 + i\,0.0067 \\ 0.0108 - i\,0.0215 & 0.5621 + i\,0.2087 & -0.0570 - i\,0.0132 \\ 0.0003 + i\,0.0010 & 0.0695 + i\,0.0095 & 3.0025 + i\,0.0014 \end{bmatrix} \lambda^2 \\
&+ \begin{bmatrix} -1.2710 - i\,0.3184 & 1.0597 + i\,0.3487 & -0.0014 + i\,0.0140 \\ 0.0347 - i\,0.0345 & 0.0870 - i\,0.5229 & 0.8983 - i\,0.0622 \\ -0.0001 + i\,0.0020 & -0.8718 + i\,0.0631 & 6.0039 + i\,0.0042 \end{bmatrix} \lambda \\
&+ \begin{bmatrix} 0.1915 - i\,1.1520 & 0.9122 + i\,0.5722 & -0.0094 + i\,0.0211 \\ -0.9282 - i\,0.0356 & -1.2517 - i\,2.3192 & -0.1257 - i\,0.1493 \\ -0.0012 + i\,0.0031 & 0.1664 + i\,0.1643 & 10.0038 + i\,0.0086 \end{bmatrix}. \;\square
\end{aligned}
$$

In our last example, the maximum value of the lower bound $\beta_{low}(P, \mu, \gamma)$ (with respect to $\gamma \geq 0$) significantly differs from the exact distance $\mathcal{E}_a(\mu)$; this was not clear in the previous two examples.

Example 3 Consider the matrix polynomial

$$
P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} -5 & 0.1 \\ 0 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix}
$$

and the set of weights $\mathrm{w} = \{1, 1, 1\}$. The boundaries of the ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.3, 0.5, 0.7$ are drawn in Figure 8. The eigenvalues of $P(\lambda)$, 0, 5 and $\pm i\,3$, are marked in the figure as "+", and each pseudospectrum is compact with four connected components. By the continuity of the eigenvalues with respect to the entries of the coefficient matrices, it follows that every matrix polynomial in $\mathcal{B}(P, 0.7, \mathrm{w})$ has four distinct eigenvalues (see Theorem 2.3 of [10]). Hence, for every $\mu \in \mathbb{C}$, $\mathcal{E}_a(\mu) > 0.7$.

For $\mu = 6$, we see that $\max_{\gamma \geq 0} \beta_{low}(P, 6, \gamma) = \beta_{low}(P, 6, 3.7670) = 0.3729$ and $\min_{\gamma \geq 0} \beta_{up}(P, 6, \gamma) = \beta_{up}(P, 6, 3.8115) = 1.1455$ (where both values of γ are different

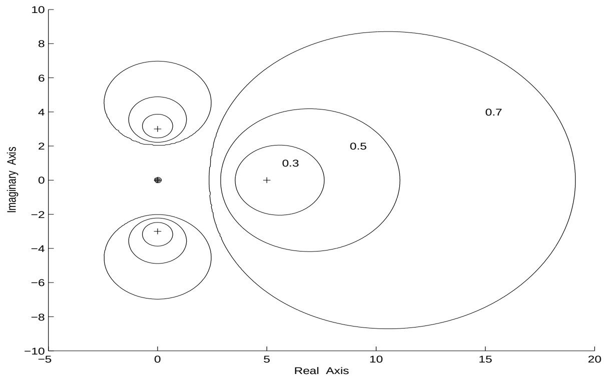

Figure 8: Pseudospectra with four connected components.

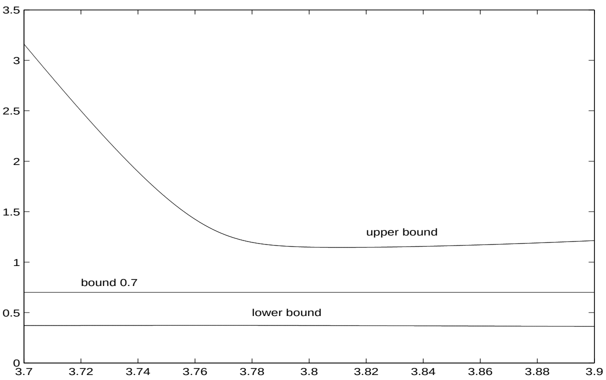

Figure 9: The graphs of the bounds $\beta_{up}(P,6,\gamma)$ and $\beta_{low}(P,6,\gamma)$.

than $\gamma_* = 3.7846$). Hence, it follows

$$\max_{\gamma\geq 0} \beta_{low}(P,6,\gamma) \;=\; 0.3729 \;<\; 0.7 \;<\; \mathcal{E}_a(6) \;\leq\; 1.1455 \;=\; \min_{\gamma\geq 0} \beta_{up}(P,6,\gamma).$$

In Figure 9, the graphs of the upper bound $\beta_{up}(P,6,\gamma)$ and the lower bound $\beta_{low}(P,6,\gamma)$ are plotted for $\gamma \in [3.7, 3.9]$. The horizontal line between these two graphs corresponds to the lower bound 0.7. The matrix polynomial

$$Q_{3.8115}(\lambda) = \begin{bmatrix} 0.6184 & 0.1258 \\ 0.1333 & -0.1235 \end{bmatrix} \lambda^2 + \begin{bmatrix} -5.3816 & 0.2258 \\ 0.1333 & -1.1235 \end{bmatrix} \lambda + \begin{bmatrix} -0.3816 & 0.1258 \\ 0.1333 & 7.8765 \end{bmatrix}$$

is given by (5), lies on the boundary of $\mathcal{B}(P, 1.1455, \mathrm{w})$, and has $\mu = 6$ as a defective eigenvalue of algebraic multiplicity 2 and geometric multiplicity 1 with corresponding eigenvector $v_1(3.8115) = \begin{bmatrix} 0.2880 \\ 0.4987 \end{bmatrix}$. □

References

[1] L. Boulton, P. Lancaster and P. Psarrakos, On pseudospecta of matrix polynomials and their boundaries, *Math. Comp.*, **77** (2008) 313–334.

[2] J.W. Demmel, On condition numbers and the distance to the nearest ill-posed problem, *Numer. Math.*, **51** (1987) 251–289.

[3] M. Embree and L.N. Trefethen, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.

[4] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

[5] G. Golub, V. Klema and G. Stewart, Rank degeneracy and least squares problems, *Stanford Univ. Technical Report STAN-CS-76-559*, 1976.

[6] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1996.

[7] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[8] A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdel, New York, 1964.

[9] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, New York, 1980.

[10] P. Lancaster and P. Psarrakos, On the pseudospectra of matrix polynomials, *SIAM J. Matrix Anal. Appl.*, **27** (2005) 115–129.

[11] A.N. Malyshev, A formula for the 2-norm distance from a matrix to the set of matrices with a multiple eigenvalue, *Numer. Math.*, **83** (1999) 443–454.

[12] L. Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young and J.C. Doyle, A formula for computation of the real stability radius, *Automatica*, **31** (1995) 879–890.

[13] A. Ruhe, Properties of a matrix with a very ill-conditioned eigenproblem, *Numer. Math.*, **15** (1970) 57–60.

[14] J.-G. Sun, A note on simple non-zero singular values, *J. Comput. Math.*, **6** (1988) 258–266.

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

[16] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Claredon Press, Oxford, 1965.

[17] J.H. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, *Numer. Math.*, **19** (1972) 175–178.

[18] J.H. Wilkinson, On neighbouring matrices with quadratic elementary divisors, *Numer. Math.*, **44** (1984) 1–21.

[19] J.H. Wilkinson, Sensitivity of eigenvalues, *Util. Math.*, **25** (1984) 5–76.

[20] J.H. Wilkinson, Sensitivity of eigenvalues II, *Util. Math.*, **30** (1986) 243–286.