Distance bounds for prescribed multiple eigenvalues of matrix polynomials

Panayiotis J. Psarrakos*

February 15, 2012

Abstract

In this paper, motivated by a problem posed by Wilkinson, we study the coefficient perturbations of a (square) matrix polynomial to a matrix polynomial that has a prescribed eigenvalue of specified algebraic multiplicity and index of annihilation. For an $n \times n$ matrix polynomial $P(\lambda)$ and a given scalar $\mu \in \mathbb{C}$, we introduce two weighted spectral norm distances, $\mathcal{E}_r(\mu)$ and $\mathcal{E}_{r,k}(\mu)$, from $P(\lambda)$ to the $n \times n$ matrix polynomials that have μ as an eigenvalue of algebraic multiplicity at least r and to those that have μ as an eigenvalue of algebraic multiplicity at least r and maximum Jordan chain length (exactly) k, respectively. Then we obtain a lower bound for $\mathcal{E}_{r,k}(\mu)$, and derive an upper bound for $\mathcal{E}_r(\mu)$ by constructing an associated perturbation of $P(\lambda)$.

Keywords: matrix polynomial, eigenvalue, algebraic multiplicity, index of annihilation, perturbation, singular value, singular vector.

AMS Subject Classifications: 15A18, 15A22, 65F15, 65F35.

1 Introduction and preliminaries

Wilkinson's problem [14, 15] concerns computing the spectral norm distance (known as *Wilkinson's distance*) from an $n \times n$ matrix with n distinct eigenvalues to the set of $n \times n$ matrices having multiple eigenvalues, and has a strong connection to ill-conditioning of eigenvalue problems. Malyshev [9] provided a solution to Wilkinson's problem by obtaining a singular value characterization of Wilkinson's distance. Recently, Mengi [11] (extending Malyshev's methodology) derived a singular value optimization characterization for the smallest perturbation to a matrix that has an eigenvalue of specified algebraic multiplicity. Papathanasiou and Psarrakos [13] studied the case of polynomial eigenvalue problems, and applied Malyshev's technique to derive lower and upper bounds for a weighted distance from a given $n \times n$ matrix polynomial to the $n \times n$ matrix polynomials that have a prescribed multiple eigenvalue.

Motivated by the above, we 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 $A_m \neq 0$. We also assume that $P(\lambda)$ is *regular*, that is, the determinant $\det P(\lambda)$ is not identically zero.

*Department of Mathematics, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece (ppsarr@math.ntua.gr).

The study of matrix polynomials, especially with regard to their spectral analysis, has a long history and important applications; see [4, 8, 10] and the references therein.

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. The set of all eigenvalues of $P(\lambda)$, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, is the *spectrum* of $P(\lambda)$, and since $P(\lambda)$ is regular, it contains no more than nm 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 matrix $P(\lambda_0)$. An eigenvalue of $P(\lambda)$ is called *semisimple* if its algebraic and geometric multiplicities are equal, and it is called *defective* otherwise.

Suppose that for an eigenvalue $\lambda_0 \in \sigma(P)$, there exist $x_0, x_1, \ldots, x_k \in \mathbb{C}^n$ with $x_0 \neq 0$, such that

$$\sum_{j=0}^{\xi} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{\xi-j} \;=\; 0\,; \quad \xi = 0, 1, \ldots, k, \tag{2}$$

where $P^{(j)}(\lambda)$ denotes the jth derivative of $P(\lambda)$ and $k+1$ cannot exceed the algebraic multiplicity of λ_0. Then the vector x_0 is clearly an eigenvector of λ_0, and the vectors $x_1, x_2, \ldots, x_k$ are known as *generalized eigenvectors*. The set $\{x_0, x_1, \ldots, x_k\}$ is called a *Jordan chain of length* $k+1$ of $P(\lambda)$ corresponding to the eigenvalue λ_0. Moreover, it is apparent that any set $\{x_0, x_1, \ldots, x_\xi\}$, $\xi = 0, 1, \ldots, k-1$, is also a Jordan chain of $P(\lambda)$ corresponding to λ_0. Any eigenvalue of $P(\lambda)$ of geometric multiplicity p has p maximal Jordan chains associated with p linearly independent eigenvectors, with total number of eigenvectors and generalized eigenvectors equal to the algebraic multiplicity of this eigenvalue. The largest length of Jordan chains of $P(\lambda)$ corresponding to $\lambda_0 \in \sigma(P)$ is known as the *index of annihilation* of λ_0 [7]. This index coincides with the size of the largest Jordan blocks of the Jordan canonical form of $P(\lambda)$ corresponding to λ_0, and it is equal to 1 if and only if the eigenvalue λ_0 is semisimple; for details on the Jordan structure of matrix polynomials, see [4, 8].

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

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

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

$$\mathcal{B}(P, \varepsilon, \mathrm{w}) \;=\; \{Q(\lambda) \text{ as in (3)} : \|\Delta_j\| \leq \varepsilon\, w_j,\ j = 0, 1, \ldots, m\}\,,$$

where $\|\cdot\|$ denotes the *spectral matrix norm* (i.e., the norm subordinate to the Euclidean vector norm), and the polynomial $w(\lambda) = w_m\lambda^m + \cdots + w_1\lambda + w_0$. The weights $w_0, w_1, \ldots, w_m$ allow freedom in how perturbations are measured. Moreover, $\mathcal{B}(P, \varepsilon, \mathrm{w})$ is convex and compact, with respect to the max norm $\|P(\lambda)\|_\infty = \max\limits_{0\leq j\leq m} \|A_j\|$ [1].

Now we can introduce weighted distances from $P(\lambda)$ to the matrix polynomials that have a prescribed eigenvalue of algebraic multiplicity at least r, and to those that have a prescribed eigenvalue of algebraic multiplicity at least r and index of annihilation (exactly) k.

Definition 1.1. 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 an eigenvalue of algebraic multiplicity at least r* by

$$\begin{aligned}\mathcal{E}_r(\mu) \;=\; \inf\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P,\varepsilon,\mathrm{w}) \text{ with } \mu \text{ as an eigenvalue of} \\ \text{algebraic multiplicity at least } r\}\,,\end{aligned}$$

and the *distance from $P(\lambda)$ to μ as an eigenvalue of algebraic multiplicity at least r and index of annihilation k* by

$$\begin{aligned}\mathcal{E}_{r,k}(\mu) \;=\; \inf\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P,\varepsilon,\mathrm{w}) \text{ with } \mu \text{ as an eigenvalue of algebraic} \\ \text{multiplicity at least } r \text{ and index of annihilation } k\}\,.\end{aligned}$$

Notice that we have the identity $\mathcal{E}_r(\mu) = \min\limits_{k=1,2,\ldots,mn} \mathcal{E}_{r,k}(\mu)$. Here, we allow values of k greater than r because the optimal perturbed matrix polynomial $Q(\lambda)$ may have μ as an eigenvalue of (both) algebraic multiplicity and index of annihilation greater than r.

The singular values of an $n \times n$ complex matrix A, i.e., the nonnegative square roots of the eigenvalues of A^*A, are denoted by $s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A) \geq 0$. It is apparent that for the matrix polynomial $P(\lambda)$, $\sigma(P) = \{\lambda \in \mathbb{C} : s_n(P(\lambda)) = 0\}$, and a scalar $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ of geometric multiplicity exactly p if and only if $s_1(P(\lambda_0)) \geq \cdots \geq s_{n-p}(P(\lambda_0)) > s_{n-p+1}(P(\lambda_0)) = \cdots = s_n(P(\lambda_0)) = 0$.

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, $\sigma(A)$, and if in addition $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, then $\mathcal{B}(P,\varepsilon,\mathrm{w}) = \{I\lambda - (A + E) : \|E\| \leq \varepsilon\}$. In this case, Malyshev [9] (inspired by [12]) has proved that

$$\mathcal{E}_2(\mu) \;=\; \sup_{\gamma>0} s_{2n-1}\left(\begin{bmatrix} I\mu - A & 0 \\ \gamma I & I\mu - A \end{bmatrix}\right).$$

Extending Malyshev's methodology, Ikramov and Nazari [6] (for $r = 3$), and Mengi [11] have obtained (always for $P(\lambda) = I\lambda - A$ and $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, i.e., for constant matrices)

$$\mathcal{E}_r(\mu) \;=\; \sup_{\gamma_{i,j}\,\in\,\mathbb{C}\setminus\{0\}} s_{rn-r+1}\left(\begin{bmatrix} I\mu - A & 0 & 0 & \cdots & 0 \\ \gamma_{2,1}I & I\mu - A & 0 & \cdots & 0 \\ \gamma_{3,1}I & \gamma_{3,2}I & I\mu - A & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{r,1}I & \gamma_{r,2}I & \gamma_{r,3}I & \cdots & I\mu - A \end{bmatrix}\right). \quad (4)$$

In this work, our goal is to derive a primer estimation of the distances $\mathcal{E}_{r,k}(\mu)$ and $\mathcal{E}_r(\mu)$ for matrix polynomials. Unfortunately, in the case of matrix polynomials, Malyshev's technique leads to lower and upper bounds for the distance $\mathcal{E}_2(\mu)$ and not to its exact value [13]. Hence, to be able to exploit the definition of Jordan chains of

matrix polynomials in (2) and avoid multivariable optimization problems, we consider block-Toeplitz matrices of higher order with only one real parameter $\gamma > 0$ and its powers instead of the $r(r-1)/2$ complex parameters $\gamma_{2,1}, \gamma_{3,1}, \gamma_{3,2}, \ldots, \gamma_{r,r-1}$ in (4) (see Definition 2.1 below and the discussion before Theorem 2.4). Generalizing results of [13], in Section 2, we obtain a lower bound for $\mathcal{E}_{r,k}(\mu)$, and in Section 3, we construct an upper bound for $\mathcal{E}_r(\mu)$ and an associated perturbation of $P(\lambda)$. Simple numerical examples are given in Section 4 to illustrate our results.

2 A lower bound for the distance $\mathcal{E}_{r,k}(\mu)$

By Theorem 4 of [13], we know that if both the required algebraic and geometric multiplicities of μ are equal to r (i.e., the index of annihilation is equal to 1), then $\mathcal{E}_{r,1}(\mu) \geq s_{n-r+1}(P(\mu))/w(|\mu|)$. Hence, it follows that

$$\mathcal{E}_1(\mu) \,=\, \frac{s_n(P(\mu))}{w(|\mu|)} \,\leq\, \mathcal{E}_r(\mu) \,\leq\, \frac{s_{n-r+1}(P(\mu))}{w(|\mu|)} \,\leq\, \mathcal{E}_{r,1}(\mu).$$

So, in the special case $s_n(P(\mu)) = s_{n-1}(P(\mu)) = \cdots = s_{n-r+1}(P(\mu))$, it is clear that $\mathcal{E}_r(\mu) = s_{n-r+1}(P(\mu))/w(|\mu|)$, and an optimal perturbation of $P(\lambda)$ is given by [13, formula (3)]. Thus, in what follows, we assume that $s_n(P(\mu)) \neq s_{n-r+1}(P(\mu))$ and consider perturbations of $P(\lambda)$ such that the perturbed matrix polynomial has μ as a defective eigenvalue (i.e., $k \geq 2$) of algebraic multiplicity at least r. The next definition is necessary for the remainder.

Definition 2.1. For the matrix polynomial $P(\lambda)$ in (1), a positive integer k and a scalar $\gamma \in \mathbb{C}$, we define the $kn \times kn$ matrix polynomial

$$F_k[P(\lambda);\gamma] \,=\, \begin{bmatrix} P(\lambda) & 0 & \cdots & 0 \\ \gamma\, P^{(1)}(\lambda) & P(\lambda) & \cdots & 0 \\ \frac{\gamma^2}{2!}\, P^{(2)}(\lambda) & \gamma\, P^{(1)}(\lambda) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma^{k-1}}{(k-1)!}\, P^{(k-1)}(\lambda) & \frac{\gamma^{k-2}}{(k-2)!}\, P^{(k-2)}(\lambda) & \cdots & P(\lambda) \end{bmatrix}.$$

We observe that a scalar $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ if and only if it is an eigenvalue of $F_k[P(\lambda);\gamma]$. Furthermore, if λ_0 is an eigenvalue of $P(\lambda)$ of algebraic multiplicity r and index of annihilation k, then for every $\gamma \neq 0$, the null space of matrix $F_k[P(\lambda_0);\gamma]$ has dimension at least r (for $\gamma = 1$, see [5, Lemma 2.5] and [7]).

Lemma 2.2. *Suppose $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $P(\lambda)$ of algebraic multiplicity at least r and index of annihilation k. Then for any nonzero $\gamma \in \mathbb{C}$, we have $s_{kn-r+1}(F_k[P(\lambda_0);\gamma]) = 0$.*

Proof. By hypothesis, we may assume that $P(\lambda)$ has p Jordan chains corresponding to λ_0, namely,

$$\{x_{1,0}, x_{1,1}, \ldots, x_{1,k_1}\},\ \{x_{2,0}, x_{2,1}, \ldots, x_{2,k_2}\},\ \ldots,\ \{x_{p,0}, x_{p,1}, \ldots, x_{p,k_p}\},$$

with $x_{1,0}, x_{2,0}, \ldots, x_{p,0}$ linearly independent eigenvectors, $k-1 = k_1 \geq k_2 \geq \cdots \geq k_p$ and $(k_1+1) + (k_2+1) + \cdots + (k_p+1) = r$. Notice that the first Jordan chain,

$\{x_{1,0}, x_{1,1}, \ldots, x_{1,k_1}\}$, is necessarily maximal. Then, recalling (2), we see that the r vectors

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{i,0} \\ \gamma x_{i,1} \\ \gamma^2 x_{i,2} \\ \vdots \\ \gamma^{k_i-1} x_{i,k_i-1} \\ \gamma^{k_i} x_{i,k_i} \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ x_{i,0} \\ \gamma x_{i,1} \\ \vdots \\ \gamma^{k_i-2} x_{i,k_i-2} \\ \gamma^{k_i-1} x_{i,k_i-1} \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ x_{i,0} \end{bmatrix} \in \mathbb{C}^{kn}; \quad i = 1, 2, \ldots, p \qquad (5)$$

satisfy

$$F_k[P(\lambda_0); \gamma] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{i,0} \\ \gamma x_{i,1} \\ \gamma^2 x_{i,2} \\ \vdots \\ \gamma^{\xi-1} x_{i,\xi-1} \\ \gamma^{\xi} x_{i,\xi} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\lambda_0) x_{i,0} \\ \gamma \sum_{j=0}^{1} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{i,1-j} \\ \gamma^2 \sum_{j=0}^{2} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{i,2-j} \\ \vdots \\ \gamma^{\xi} \sum_{j=0}^{\xi} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{i,\xi-j} \end{bmatrix} = 0$$

for all $i = 1, 2, \ldots, p$ and $\xi = 0, 1, \ldots, k_i$. Hence, they lie in the null space of matrix $F_k[P(\lambda_0); \gamma]$. Moreover, one can verify that the vectors in (5) are linearly independent, keeping in mind their block form and the linear independence of the eigenvectors $x_{1,0}, x_{2,0}, \ldots, x_{p,0} \in \mathbb{C}^n$. As a consequence, the dimension of the null space of $F_k[P(\lambda_0); \gamma]$ is greater than or equal to r (i.e., 0 is an eigenvalue of matrix $F_k[P(\lambda_0); \gamma]$ of geometric multiplicity at least r), and thus, $s_{kn-r+1}(F_k[P(\lambda_0); \gamma]) = 0$. □

By this lemma, if a scalar $\mu \in \mathbb{C}$ is a multiple eigenvalue of a perturbed matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$ (as in (3)) of algebraic multiplicity at least r and index of annihilation k, then for any nonzero $\gamma \in \mathbb{C}$, μ is an eigenvalue of the $kn \times kn$ matrix polynomial $F_k[Q(\lambda); \gamma]$ of geometric multiplicity at least r. This observation and the discussion in Section 3 of [13] yield readily the following lemma.

Lemma 2.3. *If $\mu \in \mathbb{C}$ is an eigenvalue of a matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$ of algebraic multiplicity at least r and index of annihilation k, then for every $\gamma \neq 0$,*

$$s_{kn-r+1}(F_k[P(\mu); \gamma]) \leq \|F_k[\Delta(\mu); \gamma]\|.$$

It is straightforward to verify that $\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} \in \mathbb{C}^{kn}$ (u_j, $v_j \in \mathbb{C}^n$, $j = 1, 2, \ldots, k$) is a pair of left and right singular vectors corresponding to a singular

value of $F_k[P(\mu);\gamma]$ $(\gamma \neq 0)$ if and only if $\begin{bmatrix} u_1 \\ (\overline{\gamma}/|\gamma|)u_2 \\ \vdots \\ (\overline{\gamma}/|\gamma|)^{k-1}u_k \end{bmatrix}$, $\begin{bmatrix} v_1 \\ (\overline{\gamma}/|\gamma|)v_2 \\ \vdots \\ (\overline{\gamma}/|\gamma|)^{k-1}v_k \end{bmatrix}$ is a pair of left and right singular vectors of $F_k[P(\mu);|\gamma|]$ 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 positive.

Theorem 2.4. *Suppose $\mu \in \mathbb{C}$ is an eigenvalue of a perturbed matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathrm{w})$ of algebraic multiplicity at least r and index of annihilation k. Then for every $\gamma > 0$,*

$$\varepsilon \geq \frac{\|F_k[\Delta(\mu);\gamma]\|}{\|F_k[w(|\mu|);|\gamma|]\|} \geq \frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);|\gamma|]\|}\,.$$

Proof. For the matrix polynomial $\Delta(\lambda)$ and its derivatives, we have

$$\|\Delta(\mu)\| \leq \sum_{j=0}^{m} \|\Delta_j\|\,|\mu|^j \leq \varepsilon\, w(|\mu|)$$

and

$$\left\|\Delta^{(i)}(\mu)\right\| \leq \sum_{j=i}^{m} j\,(j-1)\cdots(j-i+1)\,\|\Delta_j\|\,|\mu|^{j-i} \leq \varepsilon\, w^{(i)}(|\mu|)\,;\;\; i = 1, 2, \ldots, r-1.$$

Thus, for any $\gamma > 0$, there is a unit vector $\hat{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \in \mathbb{C}^{kn}$ with $y_1, y_2, \ldots, y_k \in \mathbb{C}^n$ such that

$$\begin{aligned}
\|F_k[\Delta(\mu);\gamma]\|^2 &= \|F_k[\Delta(\mu);\gamma]\,\hat{y}\|^2 = \left\| \begin{bmatrix} \Delta(\mu)y_1 \\ \gamma\Delta^{(1)}(\mu)y_1 + \Delta(\mu)y_2 \\ \vdots \\ \sum_{i=0}^{k-1} \frac{\gamma^i}{i!}\,\Delta^{(i)}(\mu)y_{k-i} \end{bmatrix} \right\|^2 \\
&= \|\Delta(\mu)y_1\|^2 + \left\|\gamma\Delta^{(1)}(\mu)y_1 + \Delta(\mu)y_2\right\|^2 + \cdots + \left\|\sum_{i=0}^{k-1} \frac{\gamma^i}{i!}\,\Delta^{(i)}(\mu)y_{k-i}\right\|^2 \\
&\leq \|\Delta(\mu)\|^2\,\|y_1\|^2 + \gamma^2 \left\|\Delta^{(1)}(\mu)\right\|^2 \|y_1\|^2 \\
&\quad + 2\,\gamma\,\|\Delta(\mu)\| \left\|\Delta^{(1)}(\mu)\right\| \|y_1\|\|y_2\| + \|\Delta(\mu)\|^2\,\|y_2\|^2 \\
&\quad + \cdots + \|\Delta(\mu)\|^2\,\|y_k\|^2
\end{aligned}$$

$$
\begin{aligned}
&\leq (\varepsilon\, w(|\mu|))^2\|y_1\|^2+\gamma^2(\varepsilon\, w^{(1)}(|\mu|))^2\|y_1\|^2\\
&\quad +2\,\gamma(\varepsilon\, w(|\mu|))(\varepsilon w^{(1)}(|\mu|))\|y_1\|\|y_2\|+(\varepsilon\, w(|\mu|))^2\|y_2\|^2\\
&\quad +\cdots+(\varepsilon\, w(|\mu|))^2\|y_k\|^2\\
&= \left\|\left[\begin{array}{c} \varepsilon\, w(|\mu|)\,\|y_1\| \\ \gamma\varepsilon\, w^{(1)}(|\mu|)\,\|y_1\|+\varepsilon\, w(|\mu|)\|y_2\| \\ \vdots \\ \sum_{i=0}^{k-1}\frac{\gamma^i}{i!}\,\varepsilon\, w^{(i)}(|\mu|)\|y_{k-i}\| \end{array}\right]\right\|^2\\
&\leq \varepsilon^2\,\|F_k[w(|\mu|);\gamma]\|^2\,.
\end{aligned}
$$

The proof is completed by Lemma 2.3. □

As mentioned before, for $k=1$, $\mathcal{E}_{r,1}(\mu)\geq s_{n-r+1}(P(\mu))/w(|\mu|)$ [13, Theorem 4].

Corollary 2.5. *For any $\gamma>0$, the inequalities*

$$
\mathcal{E}_{r,k}(\mu)\;\geq\;\frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);\gamma]\|}\,;\quad k=1,2,\ldots,r
$$

and

$$
\mathcal{E}_{r}(\mu)\;\geq\;\min_{k=1,2,\ldots,nm}\frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);\gamma]\|}
$$

hold.

Let us denote by $\eta(r,k)$ the smallest integer that is greater than or equal to $n-\frac{r-1}{k}$. One can see that

$$
\frac{s_{kn-r+1}(F_k[P(\mu);\gamma])}{\|F_k[w(|\mu|);\gamma]\|}\;\longrightarrow\;\frac{s_{\eta(r,k)}(P(\mu))}{w(|\mu|)}\;\leq\;\mathcal{E}_{n-\eta(r,k)+1,1}(\mu)
$$

as $\gamma\longrightarrow 0^+$.

3 An upper bound for the distance $\mathcal{E}_r(\mu)$

The technique applied in the proof of Theorem 11 of [13] can be extended for the derivation of an upper bound for the distance $\mathcal{E}_r(\mu)$ from $P(\lambda)$ in (1) to the $n\times n$ matrix polynomials that have $\mu\in\mathbb{C}$ as an eigenvalue of algebraic multiplicity at least r. The following definition is necessary.

Definition 3.1. For any $\gamma>0$ and $r\in\{2,3,\ldots,n\}$, let $\left[\begin{array}{c} u_1(\gamma)\\ u_2(\gamma)\\ \vdots\\ u_r(\gamma)\end{array}\right]$, $\left[\begin{array}{c} v_1(\gamma)\\ v_2(\gamma)\\ \vdots\\ v_r(\gamma)\end{array}\right]\in\mathbb{C}^{rn}$ ($u_j(\gamma)$, $v_j(\gamma)\in\mathbb{C}^n$, $j=1,2,\ldots,r$) be a pair of left and right singular vectors of $s_{rn-r+1}(F_r[P(\mu);\gamma])$, respectively. Then we define the $n\times r$ matrices

$$
U(\gamma)=[\,u_1(\gamma)\;\;u_2(\gamma)\;\;\cdots\;\;u_r(\gamma)\,]\quad\text{and}\quad V(\gamma)=[\,v_1(\gamma)\;\;v_2(\gamma)\;\;\cdots\;\;v_r(\gamma)\,]\,.
$$

For any $\gamma > 0$ with $\mathrm{rank}(V(\gamma)) = r \in \{2, 3 \ldots, n\}$, we will construct a perturbation $\Delta_\gamma(\lambda)$ such that the perturbed matrix polynomial $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$ has μ as a defective eigenvalue with an associated (not necessarily maximal) Jordan chain of length r. First we define the quantities

$$\phi_i \;=\; \frac{w^{(i)}(|\mu|)}{(i!)\, w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|}\right)^i \;;\;\; i = 1, 2, \ldots, r,$$

setting $\overline{\mu}/|\mu| = 0$ whenever $\mu = 0$, and recalling that $w_0 > 0$. We also consider the $r \times r$ upper triangular Toeplitz matrix

$$\Theta_\gamma \;=\; [\theta_{i,j}] \;=\; \begin{bmatrix} 1 & -\gamma\phi_1 & \gamma^2(\phi_1^2 - \phi_2) & \gamma^3(2\phi_1\phi_2 - \phi_3 - \phi_1^3) & \cdots \\ 0 & 1 & -\gamma\phi_1 & \gamma^2(\phi_1^2 - \phi_2) & \cdots \\ 0 & 0 & 1 & -\gamma\phi_1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

whose entries above the main diagonal are given by the recursive formulae

$$\theta_{i,j} \;=\; -\,\theta_{i,i}\gamma^{j-i}\phi_{j-i} - \theta_{i,i+1}\gamma^{j-(i+1)}\phi_{j-(i+1)} - \cdots - \theta_{i,j-1}\gamma\phi_1\,;\;\; 1 \le i < j \le r. \quad (6)$$

Denoting by $V(\gamma)^\dagger$ the Moore-Penrose pseudoinverse of $V(\gamma)$, we consider the $n \times n$ matrix

$$\Delta_\gamma \;=\; -s_{rn-r+1}(F_r[P(\mu);\gamma])\, U(\gamma)\, \Theta_\gamma V(\gamma)^\dagger,$$

and define the matrices

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

and the matrix polynomial

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

We observe that $\Delta_\gamma(\mu) = \Delta_\gamma,$ and for $i = 1, 2, \ldots, r,$

$$\begin{aligned} \Delta_\gamma^{(i)}(\mu) \;&=\; \sum_{j=i}^{m} j\,(j-1)\cdots(j-i+1)\,\frac{w_j}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|}\right)^j \Delta_\gamma\, \mu^{j-i} \\ &=\; \Delta_\gamma\, \frac{1}{w(|\mu|)} \sum_{j=i}^{m} j\,(j-1)\cdots(j-i+1) w_j \left(\frac{\overline{\mu}}{|\mu|}\right)^i \left(\frac{\overline{\mu}}{|\mu|}\right)^{j-i} \mu^{j-i} \\ &=\; \Delta_\gamma\, \frac{w^{(i)}(|\mu|)}{w(|\mu|)} \left(\frac{\overline{\mu}}{|\mu|}\right)^i \;=\; (i!)\,\phi_i \Delta_\gamma\,. \end{aligned}$$

Since $u(\gamma), v(\gamma) \in \mathbb{C}^{rn}$ are a left and a right singular vector of $s_{rn-r+1}(F_r[P(\mu);\gamma])$, respectively, it holds that

$$F_r[P(\mu);\gamma]\, v(\gamma) \;=\; s_{rn-r+1}(F_r[P(\mu);\gamma])\, u(\gamma). \quad (7)$$

Denote also by $e_1, e_2, \ldots$ the vectors of the standard basis, i.e., the columns of the identity matrix. If $\mathrm{rank}(V(\gamma)) = r$ ($\in \{2, 3, \ldots, n\}$), or equivalently, if $V(\gamma)^\dagger V(\gamma) = I_r$ (the $r \times r$ identity matrix), then the perturbed matrix polynomial

$$Q_\gamma(\lambda) \;=\; P(\lambda) + \Delta_\gamma(\lambda) \;=\; \sum_{j=0}^{m} (A_j + \Delta_{\gamma,j})\lambda^j \tag{8}$$

satisfies

$$\begin{aligned} Q_\gamma(\mu)\, v_1(\gamma) &= P(\mu)\, v_1(\gamma) + \Delta_\gamma\, v_1(\gamma) \\ &= s_{rn-r+1}(F_r[P(\mu);\gamma])\, u_1(\gamma) - s_{rn-r+1}(F_r[P(\mu);\gamma])\, U(\gamma)\, e_1 \\ &= s_{rn-r+1}(F_r[P(\mu);\gamma])\, u_1(\gamma) - s_{rn-r+1}(F_r[P(\mu);\gamma])\, u_1(\gamma) \\ &= 0. \end{aligned}$$

By straightforward computations, setting $\phi_0 = 1$ and recalling (6) and (7), we verify that for every $i = 1, 2, \ldots, r-1$,

$$\begin{aligned} &\sum_{j=0}^{i} \frac{1}{j!}\, \gamma^j\, Q_\gamma^{(j)}(\mu)\, v_{i-j+1}(\gamma) \\ =\;& \sum_{j=0}^{i} \frac{1}{j!}\, \gamma^j\, P^{(j)}(\mu)\, v_{i-j+1}(\gamma) + \sum_{j=0}^{i} \frac{1}{j!}\, \gamma^j\, \Delta_\gamma^{(j)}(\mu)\, v_{i-j+1}(\gamma) \\ =\;& s_{rn-r+1}(F_r[P(\mu);\gamma])\, u_{i+1}(\gamma) - \sum_{j=0}^{i} \gamma^j\, \phi_j\, \Delta_\gamma\, v_{i-j+1}(\gamma) \\ =\;& s_{rn-r+1}(F_r[P(\mu);\gamma]) \left(u_{i+1}(\gamma) - \sum_{j=0}^{i} \gamma^j\, \phi_j\, U(\gamma)\, \Theta_\gamma V(\gamma)^\dagger\, v_{i-j+1}(\gamma) \right) \\ =\;& s_{rn-r+1}(F_r[P(\mu);\gamma]) \left(u_{i+1}(\gamma) - \sum_{j=0}^{i} \gamma^j\, \phi_j\, U(\gamma)\, \Theta_\gamma\, e_{i-j+1} \right) \\ =\;& s_{rn-r+1}(F_r[P(\mu);\gamma]) \left(u_{i+1}(\gamma) - \sum_{j=0}^{i} \sum_{\xi=1}^{i-j+1} \gamma^j\, \phi_j\, \theta_{\xi,i-j+1}\, u_\xi(\gamma) \right) \\ =\;& s_{rn-r+1}(F_r[P(\mu);\gamma]) \left(u_{i+1}(\gamma) - \theta_{i+1,i+1}\, u_{i+1}(\gamma) - \sum_{\xi=1}^{i} \left(\sum_{j=\xi}^{i+1} \gamma^{i-j+1}\, \phi_{i-j+1}\, \theta_{\xi,j} \right) u_\xi(\gamma) \right) \\ =\;& -\, s_{rn-r+1}(F_r[P(\mu);\gamma]) \left(\sum_{\xi=1}^{i} \left(\sum_{j=\xi}^{i+1} \gamma^{i-j+1}\, \phi_{i-j+1}\, \theta_{\xi,j} \right) u_\xi(\gamma) \right) \\ =\;& 0. \end{aligned}$$

Dividing by $\gamma^i \neq 0$ yields

$$\sum_{j=0}^{i} \frac{1}{j!}\, Q_\gamma^{(j)}(\mu) \left(\gamma^{-(i-j)}\, v_{i-j+1}(\gamma) \right) = 0.$$

As a consequence, if $\operatorname{rank}(V(\gamma)) = r$ ($\in \{2, 3, \ldots, n\}$), then μ is a defective eigenvalue of $Q_\gamma(\lambda)$ with $\left\{v_1(\gamma), \gamma^{-1}v_2(\gamma), \gamma^{-2}v_3(\gamma), \ldots, \gamma^{-(r-1)}v_r(\gamma)\right\}$ as an associated Jordan chain of length r (recall the definition of Jordan chains in (2)).

Furthermore, we see that

$$\|\Delta_\gamma(\mu)\| \;=\; \|\Delta_\gamma\| \;=\; s_{rn-r+1}(F_r[P(\mu);\gamma]) \left\|U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger\right\|$$

and

$$\|\Delta_{\gamma,j}\| \;=\; w_j\,\frac{s_{rn-r+1}(F_r[P(\mu);\gamma])}{w(|\mu|)}\,\left\|U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger\right\|\;;\;\; j = 0, 1, \ldots, r.$$

Hence, we have the next result, which is a direct generalization of the second part of Theorem 11 in [13].

Theorem 3.2. *Let $P(\lambda)$ be a matrix polynomial as in (1), $\mu \in \mathbb{C}$, and $r \in \{2, 3, \ldots, n\}$. Then for every $\gamma > 0$ such that $\operatorname{rank}(V(\gamma)) = r$, it holds that*

$$\mathcal{E}_r(\mu) \;\leq\; \frac{s_{rn-r+1}(F_r[P(\mu);\gamma])}{w(|\mu|)}\,\left\|U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger\right\|,$$

and $Q_\gamma(\lambda)$ in (8) lies on the boundary of $\mathcal{B}\left(P, \frac{s_{rn-r+1}(F_r[P(\mu);\gamma])}{w(|\mu|)}\left\|U(\gamma)\,\Theta_\gamma V(\gamma)^\dagger\right\|, \mathrm{w}\right)$ and has μ as a defective eigenvalue with a (not necessarily maximal) Jordan chain of length r.

We remark that for $r \geq 3$, it is not easy to find values of γ for which the condition "$\operatorname{rank}(V(\gamma)) = r$" is ensured, as it was done in [13] for $r = 2$. On the other hand, in all our experiments, this rank condition appears to hold generically.

4 Numerical examples

To illustrate the proposed (lower and upper) bounds and their tightness, we begin with the special case of constant matrices.

Example 4.1. Consider the 6×6 smoke matrix that can be generated by the Matlab command `gallery('smoke',6)`, the corresponding linear pencil $L(\lambda) = I_6\lambda - S$, the weights $\mathrm{w} = \{w_0, w_1\} = \{1, 0\}$, and the scalar $\mu = 0.3841 + \mathrm{i}\,0.6767$. By [11], we know that

$$\mathcal{E}_3(0.3841 + \mathrm{i}\,0.6767) \;=\; \mathcal{E}_{3,3}(0.3841 + \mathrm{i}\,0.6767) \;=\; 0.3270.$$

The graphs of our lower bound for $\mathcal{E}_{3,3}(0.3841 + \mathrm{i}\,0.6767)$ (see Corollary 2.5) and our upper bound for $\mathcal{E}_3(0.3841 + \mathrm{i}\,0.6767)$ (see Theorem 3.2) are depicted in Figure 1 for $\gamma \in (0, 5]$. For $\gamma = 0.6748$, Corollary 2.5 implies the lower bound 0.3145. Moreover, for $\gamma = 0.5004$, Theorem 3.2 yields the upper bound 0.4694 and an associated perturbed matrix $S + \Delta$ with $\|\Delta\| = 0.4694$, which has $\mu = 0.3841 + \mathrm{i}\,0.6767$ as a defective eigenvalue of algebraic multiplicity 3 and geometric multiplicity 1.

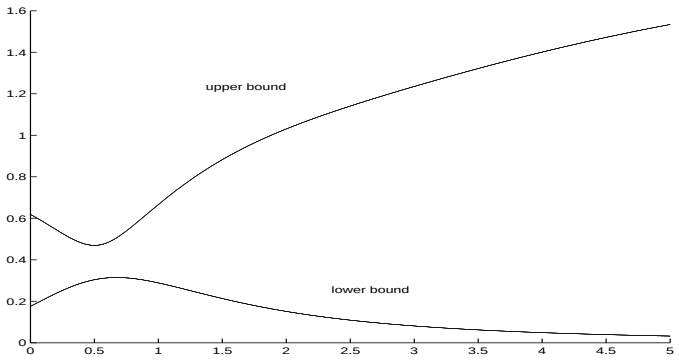

Figure 1: Lower and upper bounds for $\mathcal{E}_3(0.3841 + \mathrm{i}\,0.6767),\ 0 < \gamma \leq 5$.

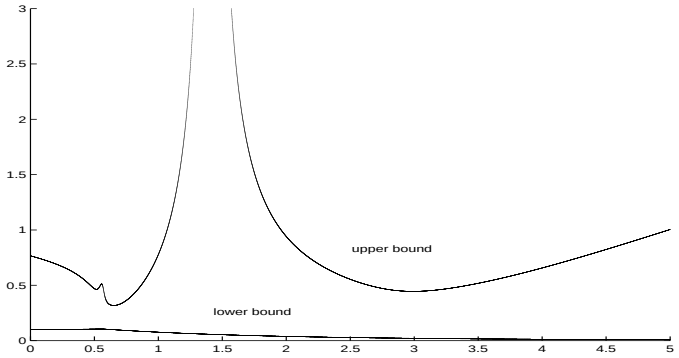

Figure 2: Lower bound for $\mathcal{E}_{3,3}(-1.1105)$ and upper bound for $\mathcal{E}_3(-1.1105)$, $0 < \gamma \leq 5$.

For our second example, we consider a real quadratic matrix polynomial.

Example 4.2. 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 $\mathrm{w} = \{w_0, w_1, w_2\} = \{10, 6.1108, 3\}$ (the norms of the coefficient matrices). For the scalar $\mu = -1.1105$, by Example 2 of [13] (see also Proposition 17 and Theorem 18 in [1]), we know that

$$\mathcal{E}_2(-1.1105) = \mathcal{E}_{2,2}(-1.1105) = \mathcal{E}_1(-1.1105) = 0.1002. \tag{9}$$

The graphs of our lower bound for $\mathcal{E}_{3,3}(-1.1105)$ and our upper bound for $\mathcal{E}_3(-1.1105)$ are illustrated in Figure 2 for $\gamma \in (0, 5]$. Setting $\gamma = 0.5530$ and $\gamma = 0.6518$, we get the lower bound 0.1048 and the upper bound 0.3177, respectively. It is worth noting that these bounds are clearly compatible with (9). Furthermore, Theorem 3.2 yields

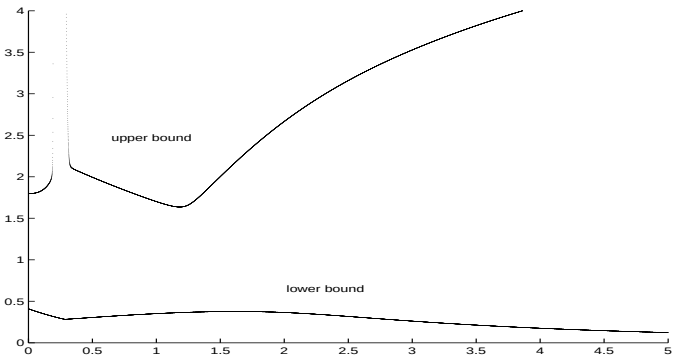

Figure 3: Lower bound for $\mathcal{E}_{3,2}(2)$ and upper bound for $\mathcal{E}_3(2)$, $0 < \gamma \leq 5$.

the perturbed matrix polynomial

$$Q(\lambda) = \begin{bmatrix} 0.5323 & -0.0453 & -0.6286 \\ 0.1664 & 1.6410 & 0.4567 \\ -0.0043 & -0.0977 & 2.7268 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0.9526 & 1.0922 & 1.2805 \\ -0.3389 & 3.7312 & 0.0696 \\ 0.0087 & -0.8009 & 6.5565 \end{bmatrix} \lambda + \begin{bmatrix} 0.4411 & 0.8490 & -2.0955 \\ -0.4454 & 1.8035 & 1.5225 \\ -0.0142 & -0.3258 & 9.0893 \end{bmatrix}$$

that lies on the boundary of $\mathcal{B}(P, 0.3177, \mathrm{w})$, and has $\mu = -1.1105$ as a defective eigenvalue of algebraic multiplicity 3 and geometric multiplicity 1, with an associated eigenvector $x = \begin{bmatrix} 0.9571 \\ 0.2890 \\ -0.0194 \end{bmatrix}$.

The matrix polynomial in our last example is triangular, and hence, we can directly compute a (non optimal) perturbation with the desired properties.

Example 4.3. Consider the 4×4 quadratic matrix polynomial

$$P(\lambda) = \begin{bmatrix} 2\lambda^2 - 5\lambda + 4 & 0 & 2\lambda + 1 & -\lambda + 6 \\ 0 & \lambda^2 + 2\lambda - 5 & -\lambda^2 & 0 \\ 0 & 0 & 2\lambda^2 - \mathrm{i}\,8 & \mathrm{i}\,\lambda \\ 0 & 0 & 0 & \lambda^2 - \lambda + 15 \end{bmatrix}$$

and the weights $\mathrm{w} = \{w_0, w_1, w_2\} = \{1, 1, 1\}$. The graphs of the proposed lower bound for $\mathcal{E}_{3,2}(2)$ and upper bound for $\mathcal{E}_3(2)$ are plotted in Figure 3 for $\gamma \in (0, 5]$. As $\gamma \longrightarrow 0^+$ and for $\gamma = 1.1827$, we get the lower bound 0.4093 for $\mathcal{E}_{3,2}(2)$ and the upper bound 1.6357 for $\mathcal{E}_3(2)$, respectively. The exact values of the distances $\mathcal{E}_{3,2}(2)$ and $\mathcal{E}_3(2)$ are not known, but these bounds are consistent with the fact that the matrix polynomial

$$R(\lambda) = \begin{bmatrix} \lambda^2 - 4\lambda + 4 & 0 & 2\lambda + 1 & -\lambda + 6 \\ 0 & \lambda^2 + \lambda - 6 & -\lambda^2 & 0 \\ 0 & 0 & \lambda^2 - \mathrm{i}\,9 & \mathrm{i}\,\lambda \\ 0 & 0 & 0 & \lambda^2 + 16 \end{bmatrix}$$

lies on the boundary of $\mathcal{B}(P, 1, \mathrm{w})$, and has $\mu = 2$ as a defective eigenvalue of algebraic multiplicity 3 and index of annihilation 2. Recalling the comment in the last paragraph of Section 2 (and applying the methodology in [13, Section 3]), we also observe that the lower bound 0.4093 coincides with the distance $\mathcal{E}_{2,1}(2)$.

Corollary 2.5 and Theorem 3.2 provide an infinite number of lower bounds of $\mathcal{E}_{r,k}(\mu)$ and upper bounds of $\mathcal{E}_r(\mu)$, respectively, in such a way that the best lower/upper bound is given by the maximum/minimum of these bounds over all $\gamma > 0$. Since these optimization problems are over only one real variable, in our examples above, we apply the standard grid search with respect to $\gamma \in (0, 5]$. The difficulty with this brute-force approach is that it usually requires too many bound evaluations.

Alternatively, we can use golden section search [3, pp. 656–659], or Brent's method [2, Chapter 5]. The latter algorithm is based on the combination of inverse quadratic interpolation and golden section search, and one of its implementations is the Matlab function `fminbnd`. As it is expected, these two derivative-free optimization methods compute at best a local extremum, and there is no guarantee that the global extremum will be found (so, the grid search with a few grid points can be used to initiate an interval of the global optimum). For example, we consider the upper bound of the distance $\mathcal{E}_3(-1.1105)$ in Example 2 (see Figure 2). Applying `fminbnd` on the interval $(0, 2]$, we get the global minimum 0.3177 at $\gamma = 0.6518$ after 8 iterations. Over the interval $(0, 5]$, `fminbnd` finds the local minimum 0.4449 at $\gamma = 2.9905$ after 10 iterations.

Acknowledgment. The author is grateful to the anonymous referees for their valuable comments and suggestions.

References

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

[2] R.P. Brent, *Algorithms for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

[3] G. Dahlquist and A. Björck, *Numerical Methods in Scientific Computing*, Society for Industrial and Applied Mathematics, Philadelphia, 2008.

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

[5] I. Gohberg and L. Rodman, Analytic matrix functions with prescribed local data, *J. Anal. Math.*, **40** (1981), 90–128.

[6] Kh.D. Ikramov and A.M. Nazari, On the distance to the closest matrix with triple zero eigenvalue, *Math. Notes*, **73** (2003), 511–520.

[7] G. Kalogeropoulos, P. Psarrakos and N. Karcanias, On the computation of the Jordan canonical form of regular matrix polynomials, *Linear Algebra Appl.*, **385** (2004), 117–130.

[8] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd edition, Academic Press, Orlando, 1985.

[9] 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.

[10] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.

[11] E. Mengi, Locating a nearest matrix with an eigenvalue of prescribed algebraic multiplicity, *Numer. Math.*, **118** (2011), 109–135.

[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, *Automatika*, **31** (1995), 879–890.

[13] N. Papathanasiou and P. Psarrakos, The distance from a matrix polynomial to matrix polynomials with a prescribed multiple eigenvalue, *Linear Algebra Appl.*, **429** (2008), 1453–1477.

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

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