On the distance from a matrix polynomial to matrix polynomials with some prescribed eigenvalues

E. Kokabifar[*], P.J. Psarrakos[†] and G.B. Loghmani[*]

Abstract

Consider an $n \times n$ matrix polynomial $P(\lambda)$ and a set Σ consisting of $k \leq n$ complex numbers. Recently, Kokabifar, Loghmani, Psarrakos and Karbassi studied a (weighted) spectral norm distance from $P(\lambda)$ to the $n \times n$ matrix polynomials whose spectra contain the specified set Σ, under the assumption that all the entries of Σ are distinct. In this paper, the case in which some or all of the desired eigenvalues can be multiple is discussed. Lower and upper bounds for the distance are computed, and a perturbation of $P(\lambda)$ associated to the upper bound is constructed. A detailed numerical example illustrates the efficiency and validity of the proposed computational method.

Keywords: Matrix polynomial, Eigenvalue, Perturbation, Singular value, Jordan chain.

AMS Classification: 15A18, 65F35.

1 Introduction

Assume that all the eigenvalues of a matrix $A \in \mathbb{C}^{n\times n}$ are simple. Computing the distance from A to the set of $n \times n$ (complex) matrices having multiple eigenvalues is known as Wilkinson's problem. Wilkinson introduced this problem in [22] and computed bounds for this distance, known as Wilkinson's distance, in [23–26]. Demmel [2] and Ruhe [19] also calculated alternative bounds for Wilkinson's distance. In 1999, Malyshev [14] obtained a singular value optimization characterization for the spectral norm distance from A to the set of all $n \times n$ complex matrices that have a fixed multiple eigenvalue; his work can be construed as a solution to Wilkinson's problem.

Expanding and improving the methodology used in [14], Gracia [4] and Lippert [13] studied a spectral norm distance from A to $n \times n$ complex matrices with two prescribed

[*]Department of Mathematics, Faculty of Science, Yazd University, Yazd, Iran (e.kokabifar@yahoo.com, loghmani@yazd.ac.ir). These authors were supported by grant no. 95002829 of Iran National Science Foundation (INSF).

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

eigenvalues. Moreover, a spectral norm distance from $A \in \mathbb{C}^{n\times n}$ to the set of $n\times n$ matrices with $k \leq n$ fixed eigenvalues is studied, geometrically by Lippert [12], and computationally by Kokabifar, Loghmani and Karbassi [8]. Papathanasiou and Psarrakos [17], Kokabifar, Loghmani, Nazari and Karbassi [7], and Psarrakos [18] studied a (weighted) spectral norm distance from an $n \times n$ matrix polynomial $P(\lambda)$ to the $n \times n$ matrix polynomials that have a prescribed multiple eigenvalue, two prescribed distinct eigenvalues, and a prescribed eigenvalue of specified algebraic multiplicity, respectively. The results achieved in [17] and [7] can be interpreted as generalizations to matrix polynomials of results obtained in [14] and [4,13], respectively.

Recently, Kokabifar, Loghmani, Psarrakos and Karbassi [9] extended the results of [7] to the case of $k \leq n$ distinct eigenvalues. However, a question that arises in a natural way is the following: *What one can say if some of the desired eigenvalues are multiple*? In this paper, we investigate this problem, and obtain an upper and a lower bound for the distance from an $n \times n$ matrix polynomial $P(\lambda)$ to matrix polynomials that have $k \leq n$ given eigenvalues which are not necessarily distinct, generalizing and unifying results of [7,17,18]. To achieve this, we need to modify definitions, lemmas and techniques presented in [7–9,17,18].

In the next section, we review some standard definitions on matrix polynomials, and also give some new definitions which are necessary for the remainder. In Section 3, we construct an admissible perturbation of $P(\lambda)$ that has the desired k eigenvalues, by extending and modifying the techniques presented in [7–9,17,18]. In Section 4, we apply the results of Section 3 to compute an upper and a lower bound for the distance. Finally, in Section 5, we give a comprehensive numerical example to illustrate the proposed method. Some partial results on the behavior of the Jordan structure of matrix polynomial under perturbations, which are necessary for the construction of the lower bound, are presented in an appendix.

2 Preliminaries

For $A_0, A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$, with $A_m \neq 0$, and a complex variable λ, we define the *matrix polynomial*

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0 = \sum_{j=0}^{m} A_j\lambda^j. \tag{1}$$

The spectral analysis and the Jordan structure of $P(\lambda)$ leads to the solutions of the higher order linear systems of differential equations $A_m\frac{d^m u(t)}{dt^m} + A_{m-1}\frac{d^{m-1}u(t)}{dt^{m-1}} + \cdots + A_0u(t) = f(t)$ (where $f(t)$ is a given $\mathbb{C}^n$-valued piecewise continuous function of the real variable t) and of difference equations $A_mu_{j+m} + A_{m-1}u_{j+m-1} + \cdots + A_0u_j = f_j$ (where $\{f_0, f_1, \ldots\}$ is a given sequence of vectors in $\mathbb{C}^n$ and $\{u_0, u_1, \dots\}$ is a sequence to be found) [3]. As a consequence, in the last decades, the study of matrix polynomials has received much attention of several researchers and has met many applications in diverse areas of applied mathematics such as boundary value problems, systems theory and control, vibrating and gyroscopic systems, wave theory, and stochastic models. Suggested references for the theory and applications of matrix polynomials are [3, 5, 10, 15, 16, 20, 21] and the references therein.

Suppose that for a scalar $\lambda_0 \in \mathbb{C}$ and a nonzero vector $v \in \mathbb{C}^n$, it holds that $P(\lambda_0)v = 0$. Then the scalar λ_0 is called an *eigenvalue* of $P(\lambda)$, and the vector v is known as a *(right) eigenvector* of $P(\lambda)$ corresponding to λ_0. Similarly, a nonzero vector $u \in \mathbb{C}^n$ is known as a *left eigenvector* of $P(\lambda)$ corresponding to λ_0 when $u^* P(\lambda_0) = 0$. The *spectrum* of $P(\lambda)$, denoted by $\sigma(P)$, is the set of its eigenvalues. The multiplicity of an eigenvalue $\lambda_0 \in \sigma(P)$ as a root of the scalar polynomial $\det P(\lambda)$ is called the *algebraic multiplicity* of λ_0, and the dimension of the null space of the (constant) matrix $P(\lambda_0)$ is known as the *geometric multiplicity* of λ_0. The algebraic multiplicity of an eigenvalue is always greater than or equal to its geometric multiplicity. An eigenvalue is called *semisimple* if its algebraic and geometric multiplicities are equal; otherwise, it is known as *defective*. Throughout this paper, it is assumed that:

(a) The coefficient matrix A_m is *nonsingular*; this implies that $P(\lambda)$ has exactly mn finite eigenvalues, counting algebraic multiplicities.

(b) The spectrum $\sigma(P)$ has exactly nm entries, where each eigenvalue appears as many times as its algebraic multiplicity.

The singular values of $P(\lambda)$ are the nonnegative roots of the eigenvalue functions of $P(\lambda)^* P(\lambda)$, ordered in non-increasing order, and they are denoted by

$$s_1(P(\lambda)) \geq s_2(P(\lambda)) \geq \cdots \geq s_n(P(\lambda)) \geq 0.$$

Let λ_0 be an eigenvalue of $P(\lambda)$, and let q be a positive integer less than or equal to the algebraic multiplicity of λ_0. If there exist q vectors $v_1, v_2, \dots, v_q$, with $v_1 \neq 0$, such that

$$\begin{aligned}
P(\lambda_0)v_1 &= 0, \\
\frac{1}{1!}P'(\lambda_0)v_1 + P(\lambda_0)v_2 &= 0, \\
\frac{1}{2!}P''(\lambda_0)v_1 + \frac{1}{1!}P'(\lambda_0)v_2 + P(\lambda_0)v_3 &= 0, \\
\vdots \qquad &\quad \vdots \\
\sum_{i=0}^{q-1} \frac{1}{i!}P^{(i)}(\lambda_0)v_{q-i} &= 0,
\end{aligned}$$

where $P^{(i)}(\lambda)$ denotes the i-th derivative of $P(\lambda)$ with respect to λ, then the set $\{v_1, v_2, \dots, v_q\}$ is called a *(right) Jordan chain of length* q of $P(\lambda)$ corresponding to λ_0. The vector v_1 ($\neq 0$) is clearly an eigenvector of $P(\lambda)$ associated to λ_0, and the vectors $v_2, v_3, \dots, v_q$ are known as *generalized eigenvectors* of λ_0 corresponding to the eigenvector v_1. When $m > 1$, the vectors in a Jordan chain need not be linearly independent [3, Subsection 1.4].

For convenience, for every $r = 1, 2, \dots, q$, we say that the matrix $\frac{1}{(r-1)!}P^{(r-1)}(\lambda_0)$ is the r*-th Jordan chain coefficient* of $P(\lambda)$ corresponding to λ_0. In addition, it is assumed that we are given a set of s distinct scalars $\mu_1, \mu_2, \dots, \mu_s \in \mathbb{C}$, with each μ_i having multiplicity equal to $q_i \geq 1$ $(i = 1, 2, \dots, s)$, where $q_1 + q_2 + \cdots + q_s = k \leq n$. This set is denoted by

$$\Sigma = \{ \underbrace{\mu_1, \dots, \mu_1}_{q_1\text{-times}}, \underbrace{\mu_2, \dots, \mu_2}_{q_2\text{-times}}, \dots, \underbrace{\mu_s, \dots, \mu_s}_{q_s\text{-times}} \}; \tag{2}$$

i.e., each μ_i appears exactly q_i times $(i = 1, 2, \dots, s)$. A class of additive perturbations of $P(\lambda)$, an associated spectral norm distance from $P(\lambda)$ to $n \times n$ matrix polynomials whose spectra contain the set Σ in (2), and an $nk \times nk$ matrix which is crucial for our discussion, are described in the next three definitions.

Definition 2.1. For a matrix polynomial $P(\lambda)$ as in (1), and for arbitrary matrices $\Delta_0, \Delta_1, \dots, \Delta_m \in \mathbb{C}^{n \times n}$, consider (additive) perturbations of $P(\lambda)$ of the form

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

Also, for $\varepsilon > 0$ and a set of given nonnegative weights $w = \{w_0, w_1, \dots, w_m\}$, with $w_0 > 0$, define the class of admissible perturbed matrix polynomials

$$\mathcal{B}(P, \varepsilon, w) = \big\{Q(\lambda) \text{ as in (3)} : \|\Delta_j\|_2 \leq \varepsilon w_j,\ j = 0, 1, \dots, m\big\}$$

and the scalar polynomial $w(\lambda) = w_m\lambda^m + w_{m-1}\lambda^{m-1} + \cdots + w_1\lambda + w_0$.

Definition 2.2. For a matrix polynomial $P(\lambda)$ as in (1) and a set of complex numbers Σ as in (2), the distance from $P(\lambda)$ to the set of matrix polynomials whose spectra include Σ is defined and denoted by

$$D_w(P, \Sigma) = \min\{\varepsilon \geq 0 : \exists\, Q(\lambda) \in \mathcal{B}(P, \varepsilon, w) \text{ such that } \Sigma \subseteq \sigma(Q)\}.$$

Definition 2.3. Let $P(\lambda)$ be a matrix polynomial as in (1), and let a set of complex numbers Σ as in (2) be given. For any nonzero scalar $\gamma \in \mathbb{C}$, define the $nk \times nk$ block lower triangular matrix

$$F_\gamma[P, \Sigma] = \begin{bmatrix} F_{1,1} & 0 & \cdots & 0 & 0 \\ F_{2,1} & F_{2,2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{k-1,1} & F_{k-1,2} & \cdots & F_{k-1,k-1} & 0 \\ F_{k,1} & F_{k,2} & \cdots & F_{k,k-1} & F_{k,k} \end{bmatrix},$$

in which:

(i) All the blocks $F_{i,j}$ $(1 \leq j \leq i \leq k)$ are $n \times n$.

(ii) The (main) diagonal blocks of $F_\gamma[P,\Sigma]$ are given by

$$F_{1,1} = \cdots = F_{q_1,q_1} = P(\mu_1), \quad F_{q_1+1,q_1+1} = \cdots = F_{q_1+q_2,q_1+q_2} = P(\mu_2),$$
$$\ldots, \quad F_{q_1+\cdots+q_{s-1}+1,q_1+\cdots+q_{s-1}+1} = \cdots = F_{q_1+\cdots+q_{s-1}+q_s,q_1+\cdots+q_{s-1}+q_s} = P(\mu_s).$$

(iii) For all $\mu_i \neq \mu_j$, set $\theta_{j,i} = \frac{\gamma}{\mu_j-\mu_i}$. The blocks below the diagonal are given by the recursive formula (starting from the diagonal blocks)

$$F_{i,j} = \begin{cases} \gamma^{i-j}(\text{the next Jordan chain coefficient of } F_{i-1,j}) & \text{if } F_{i-1,j} = F_{i,j+1}, \\ \theta_{j,i}\,(F_{i-1,j} - F_{i,j+1}) & \text{if } F_{i-1,j} \neq F_{i,j+1}. \end{cases}$$

(Here and elsewhere, we say that $F_{i-1,j} = F_{i,j+1}$ if and only if the i-th block-row and the j-th block-column of the matrix $F_\gamma[P,\Sigma]$ correspond to the same desired eigenvalue.)

For example, if $\Sigma = \{\mu_1, \mu_1, \mu_2, \mu_3, \mu_3, \mu_3\}$, then

$$F_\gamma[P,\Sigma] = \begin{bmatrix} F_{1,1} & 0 & 0 & 0 & 0 & 0 \\ F_{2,1} & F_{2,2} & 0 & 0 & 0 & 0 \\ F_{3,1} & F_{3,2} & F_{3,3} & 0 & 0 & 0 \\ F_{4,1} & F_{4,2} & F_{4,3} & F_{4,4} & 0 & 0 \\ F_{5,1} & F_{5,2} & F_{5,3} & F_{5,4} & F_{5,5} & 0 \\ F_{6,1} & F_{6,2} & F_{6,3} & F_{6,4} & F_{6,5} & F_{6,6} \end{bmatrix}$$

$$= \begin{bmatrix} P(\mu_1) & 0 & 0 & 0 & 0 & 0 \\ \gamma P'(\mu_1) & P(\mu_1) & 0 & 0 & 0 & 0 \\ \gamma^2 \frac{P'(\mu_1) - \frac{P(\mu_1)-P(\mu_2)}{\mu_1-\mu_2}}{\mu_1-\mu_2} & \gamma \frac{P(\mu_1)-P(\mu_2)}{\mu_1-\mu_2} & P(\mu_2) & 0 & 0 & 0 \\ \gamma\, \frac{F_{3,1}-F_{4,2}}{\mu_1-\mu_3} & \gamma^2 \frac{\frac{P(\mu_1)-P(\mu_2)}{\mu_1-\mu_2} - \frac{P(\mu_2)-P(\mu_3)}{\mu_2-\mu_3}}{\mu_1-\mu_3} & \gamma \frac{P(\mu_2)-P(\mu_3)}{\mu_2-\mu_3} & P(\mu_3) & 0 & 0 \\ \gamma\, \frac{F_{4,1}-F_{5,2}}{\mu_1-\mu_3} & \gamma\, \frac{F_{4,2}-F_{5,3}}{\mu_1-\mu_3} & \gamma^2 \frac{\frac{P(\mu_2)-P(\mu_3)}{\mu_2-\mu_3} - P'(\mu_3)}{\mu_2-\mu_3} & \gamma P'(\mu_3) & P(\mu_3) & 0 \\ \gamma\, \frac{F_{5,1}-F_{6,2}}{\mu_1-\mu_3} & \gamma\, \frac{F_{5,2}-F_{6,3}}{\mu_1-\mu_3} & \gamma^3 \frac{\frac{\frac{P(\mu_2)-P(\mu_3)}{\mu_2-\mu_3} - P'(\mu_3)}{\mu_2-\mu_3} - \frac{P''(\mu_3)}{2!}}{\mu_2-\mu_3} & \frac{\gamma^2}{2!} P''(\mu_3) & \gamma P'(\mu_3) & P(\mu_3) \end{bmatrix}. \quad (4)$$

3 Construction of a perturbation

In this section, an $n \times n$ matrix polynomial $\Delta_\gamma(\lambda)$ is constructed such that the spectrum of the perturbed matrix polynomial $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$ contains a given set Σ as in

(2). In the remainder, without loss of generality, it is assumed that the parameter γ is real positive [18]. Moreover, for convenience, set $\rho = nk - k + 1$.

For the construction of the desired perturbation of $P(\lambda)$, consider a pair

$$u(\gamma) = \begin{bmatrix} u_1(\gamma) \\ u_2(\gamma) \\ \vdots \\ u_k(\gamma) \end{bmatrix}, \; v(\gamma) = \begin{bmatrix} v_1(\gamma) \\ v_2(\gamma) \\ \vdots \\ v_k(\gamma) \end{bmatrix} \in \mathbb{C}^{nk} \quad (u_j(\gamma), v_j(\gamma) \in \mathbb{C}^n, \; j = 1, 2, \ldots, k)$$

of left and right singular vectors corresponding to the ρ-th singular value of matrix $F_\gamma[P,\Sigma]$, $s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)$. By the definition of $u(\gamma)$ and $v(\gamma)$, it follows

$$F_\gamma[P,\Sigma]v(\gamma) = s_\rho\left(F_\gamma[P,\Sigma]\right)u(\gamma), \tag{5}$$

or equivalently, the following matrix equations hold:

$$\begin{array}{rcl} F_{1,1}v_1(\gamma) & = & s_\rho\left(F_\gamma[P,\Sigma]\right)u_1(\gamma), \\ F_{2,1}v_1(\gamma) + F_{2,2}v_2(\gamma) & = & s_\rho\left(F_\gamma[P,\Sigma]\right)u_2(\gamma), \\ \vdots & & \vdots \\ F_{k-1,1}v_1(\gamma) + F_{k-1,2}v_2(\gamma) + \cdots + F_{k-1,k-1}v_{k-1}(\gamma) & = & s_\rho\left(F_\gamma[P,\Sigma]\right)u_{k-1}(\gamma), \\ F_{k,1}v_1(\gamma) + F_{k,2}v_2(\gamma) + \cdots + F_{k,k-1}v_{k-1}(\gamma) + F_{k,k}v_k(\gamma) & = & s_\rho\left(F_\gamma[P,\Sigma]\right)u_k(\gamma). \end{array} \tag{6}$$

Let $\hat{F}_\gamma[P,\Sigma]$ be the $nk \times nk$ block matrix (with blocks of order n) that has the (out of fractions) Jordan chain coefficients of $P(\lambda)$ exactly at the same positions as $F_\gamma[P,\Sigma]$ and zero blocks elsewhere, and denote by $I_{n\times n}$ the $n \times n$ identity matrix. By the definition of matrix $F_\gamma[P,\Sigma]$ (see Definition 2.3), it is apparent that each nonzero block $F_{i,j}$ $(i \geq j)$ of $F_\gamma[P,\Sigma]$ is either, a Jordan chain coefficient, or a linear combination of the Jordan chain coefficients lying in the i-th row and in the j-th column. As a consequence, by applying the Gauss elimination in an appropriate way, all nonzero blocks $F_{i,j}$ $(i > j)$ which are not Jordan chain coefficients can be vanished. In particular, there exist $nk \times nk$ elementary block matrices (with blocks of order n), i.e., matrices which are equal to $I_{nk\times nk}$ or differ from it by one single elementary block-row or block-column operation,

$$E_{2,1}, E_{3,1}, E_{3,2}, \ldots, E_{k-1,1}, E_{k-1,2}, \ldots, E_{k-1,k-1}, E_{k,1}, E_{k,2}, \ldots, E_{k,k-1} \text{ (block-row operations)}$$

and

$$\hat{E}_{2,1}, \hat{E}_{3,1}, \hat{E}_{3,2}, \ldots, \hat{E}_{k-1,1}, \hat{E}_{k-1,2}, \ldots, \hat{E}_{k-1,k-1}, \hat{E}_{k,1}, \hat{E}_{k,2}, \ldots, \hat{E}_{k,k-1} \text{ (block-column operations),}$$

such that:

the $(k,1)$-th block of $E_{k,1}F_\gamma[P,\Sigma]\hat{E}_{k,1}$ is either zero or a Jordan chain coefficient,

the $(k,1)$-th and $(k,2)$-th blocks of $E_{k,2}E_{k,1}F_\gamma[P,\Sigma]\hat{E}_{k,1}\hat{E}_{k,2}$ are either zero or Jordan chain coefficients,

$$\vdots$$

the $(k,1)$-th, $(k,2)$-th, $\ldots$, $(k,k-1)$-th and $(k-1,1)$-th blocks of $E_{k-1,1}E_{k,k-1}\cdots E_{k,2}E_{k,1}F_\gamma[P,\Sigma]\hat{E}_{k,1}\hat{E}_{k,2}\cdots\hat{E}_{k,k-1}\hat{E}_{k-1,1}$ are either zero or Jordan chain coefficients,

$$\vdots$$

the matrix $\left(\prod_{\substack{i=2,3,\ldots,k\\ j=i-1,i-2,\ldots,1}} E_{i,j}\right) F_\gamma[P,\Sigma]\left(\prod_{\substack{i=k,k-1,\ldots,2\\ j=1,2,\ldots,i-1}} \hat{E}_{i,j}\right)$ is equal to $\hat{F}_\gamma[P,\Sigma]$.

The symmetry of the divided differences in part (iii) of Definition 2.3 and the standard properties of elementary matrices imply that $\hat{E}_{i,j} = E_{i,j}^{-1}$, $1 \le j < i \le k$. Hence, if we define the $nk \times nk$ block upper triangular matrix

$$\begin{aligned} T &= E_{2,1}\,E_{3,2}\,E_{3,1}\,\cdots\,E_{k-1,k-2}\,\cdots\,E_{k-1,2}\,E_{k-1,1}\,E_{k,k-1}\,\cdots\,E_{k,2}\,E_{k,1} \\ &= \begin{bmatrix} I_n & 0 & 0 & \cdots & 0 & 0 \\ t_{2,1}I_n & I_n & 0 & \cdots & 0 & 0 \\ t_{3,1}I_n & t_{3,2}I_n & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ t_{k-1,1}I_n & t_{k-1,2}I_n & t_{k-1,3}I_n & \cdots & I_n & 0 \\ t_{k,1}I_n & t_{k,2}I_n & t_{k,3}I_n & \cdots & t_{k,k-1}I_n & I_n \end{bmatrix}, \end{aligned}$$

then it follows

$$TF_\gamma[P,\Sigma]T^{-1} = \hat{F}_\gamma[P,\Sigma] \quad\Leftrightarrow\quad F_\gamma[P,\Sigma] = T^{-1}\hat{F}_\gamma[P,\Sigma]T. \tag{7}$$

Remark 3.1. For every $r = 1,2,\ldots,s$, $i = q_1+q_2+\cdots+q_{r-1}+1, q_1+q_2+\cdots+q_{r-1}+2,\ldots,q_1+q_2+\cdots+q_r$ and $1 \le j \le i$ (where for $r=1$, we set $q_1+q_2+\cdots+q_{r-1}=0$), the entry $t_{i,j}$ is the (scalar) coefficient of $P(\mu_r)$ in the (i,j)-th block of matrix $F_\gamma[P,\Sigma]$. In particular, for every $r=1,2,\ldots,s$ and $q_1+q_2+\cdots+q_{r-1}+1 \le j < i \le q_1+q_2+\cdots+q_r$, it holds that $E_{i,j} = I_n$ and $t_{i,j} = 0$.

Remark 3.2. In the case where all the desired eigenvalues are assumed to be simple, i.e., when $s = k$, $q_1 = q_2 = \cdots = q_s = 1$ and $\Sigma = \{\mu_1, \mu_2, \mu_3, \ldots, \mu_k\}$, the entries $t_{i,j}$ of matrix T are given explicitly in [9]. Moreover, if there is only one desired eigenvalue of algebraic multiplicity at least k ($\leq n$), i.e., if $s = 1$, $q_1 = k$ and $\Sigma = \{\mu_1, \mu_1, \mu_1, \ldots, \mu_1\}$, then $T = I_{nk}$ (see also [18]).

For example, consider the matrix $F_\gamma[P, \Sigma]$ in (4). Then, it is straightforward to verify that for the matrix

$$\begin{aligned} T &= I_6\, E_{3,2}\, E_{3,1}\, E_{4,3}\, E_{4,2}\, E_{4,1}\, I_6\, E_{5,3}\, E_{5,2}\, E_{5,1}\, I_6\, I_6\, E_{6,3}\, E_{6,2}\, E_{6,1} \\ &= \begin{bmatrix} I_n & 0 & 0 & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 & 0 & 0 \\ \theta_{1,2}^2 I_n & -\theta_{1,2} I_n & I_n & 0 & 0 & 0 \\ -\theta_{1,3}^2\theta_{2,3} I_n & \theta_{1,3}\theta_{2,3} I_n & -\theta_{2,3} I_n & I_n & 0 & 0 \\ -(2\theta_{1,3}^3\theta_{2,3} + \theta_{1,3}^2\theta_{2,3}^2) I_n & (\theta_{1,3}^2\theta_{2,3} + \theta_{1,3}\theta_{2,3}^2) I_n & -\theta_{2,3}^2 I_n & 0 & I_n & 0 \\ -(3\theta_{1,3}^4\theta_{2,3} + 2\theta_{1,3}^3\theta_{2,3}^2 + \theta_{1,3}^2\theta_{2,3}^3) I_n & (\theta_{1,3}^3\theta_{2,3} + \theta_{1,3}^2\theta_{2,3}^2 + \theta_{1,3}\theta_{2,3}^3) I_n & -\theta_{2,3}^3 I_n & 0 & 0 & I_n \end{bmatrix}, \end{aligned} \tag{8}$$

it holds that

$$T^{-1} = \begin{bmatrix} I_n & 0 & 0 & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 & 0 & 0 \\ -\theta_{1,2}^2 I_n & \theta_{1,2} I_n & I_n & 0 & 0 & 0 \\ (\theta_{1,3}^2\theta_{2,3} - \theta_{1,2}^2\theta_{2,3}) I_n & (\theta_{1,2}\theta_{2,3} - \theta_{1,3}\theta_{2,3}) I_n & \theta_{2,3} I_n & I_n & 0 & 0 \\ (-\theta_{1,2}^2\theta_{2,3}^2 + 2\theta_{1,3}^3\theta_{2,3} + \theta_{1,3}^2\theta_{2,3}^2) I_n & (-\theta_{1,3}^2\theta_{2,3} - \theta_{1,3}\theta_{2,3}^2 + \theta_{1,2}\theta_{2,3}^2 I_n & \theta_{2,3}^2 I_n & 0 & I_n & 0 \\ (-\theta_{1,2}^2\theta_{2,3}^3 + 3\theta_{1,3}^4\theta_{2,3} + 2\theta_{1,3}^3\theta_{2,3}^2 + \theta_{1,3}^2\theta_{2,3}^3) I_n & (-\theta_{1,3}^3\theta_{2,3} - \theta_{1,3}^2\theta_{2,3}^2 - \theta_{1,3}\theta_{2,3}^2 + \theta_{1,2}\theta_{2,3}^3) I_n & \theta_{2,3}^3 I_n & 0 & 0 & I_n \end{bmatrix}$$

and

$$\hat{F}_\gamma[P, \Sigma] = T\, F_\gamma[P, \Sigma]\, T^{-1} = \begin{bmatrix} P(\mu_1) & 0 & 0 & 0 & 0 & 0 \\ \gamma P'(\mu_1) & P(\mu_1) & 0 & 0 & 0 & 0 \\ 0 & 0 & P(\mu_2) & 0 & 0 & 0 \\ 0 & 0 & 0 & P(\mu_3) & 0 & 0 \\ 0 & 0 & 0 & \gamma P'(\mu_3) & P(\mu_3) & 0 \\ 0 & 0 & 0 & \frac{\gamma^2}{2!} P''(\mu_3) & \gamma P'(\mu_3) & P(\mu_3) \end{bmatrix}.$$

Now, by (5) and (7), it follows

$$T^{-1} \hat{F}_\gamma[P, \Sigma]\, T\, v(\gamma) = s_\rho\, (F_\gamma[P, \Sigma])\, u(\gamma),$$

or equivelantly,

$$\hat{F}_\gamma[P, \Sigma]\, (T\, v(\gamma)) = s_\rho\, (F_\gamma[P, \Sigma])\, (T\, u(\gamma))\, .$$

Hence, for the vectors

$$\hat{u}(\gamma) = T\,u(\gamma) = \begin{bmatrix} \hat{u}_1(\gamma) \\ \hat{u}_2(\gamma) \\ \vdots \\ \hat{u}_k(\gamma) \end{bmatrix} \in \mathbb{C}^{nk} \quad (\hat{u}_j(\gamma) \in \mathbb{C}^n,\ j = 1, 2, \ldots, k) \tag{9}$$

and

$$\hat{v}(\gamma) = T\,v(\gamma) = \begin{bmatrix} \hat{v}_1(\gamma) \\ \hat{v}_2(\gamma) \\ \vdots \\ \hat{v}_k(\gamma) \end{bmatrix} \in \mathbb{C}^{nk} \quad (\hat{v}_j(\gamma) \in \mathbb{C}^n,\ j = 1, 2, \ldots, k)\,, \tag{10}$$

the system (6) is written

$$\left\{ \begin{aligned} P(\mu_1)\hat{v}_1(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_1(\gamma), \\ \gamma P'(\mu_1)\hat{v}_1(\gamma) + P(\mu_1)\hat{v}_2(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_2(\gamma), \\ &\vdots \\ \tfrac{\gamma^{q_1-1}}{(q_1-1)!}P^{(q_1-1)}(\mu_1)\hat{v}_1(\gamma) + \cdots + P(\mu_1)\hat{v}_{q_1}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1}(\gamma), \end{aligned} \right.$$

$$\left\{ \begin{aligned} P(\mu_2)\hat{v}_{q_1+1}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+1}(\gamma), \\ \gamma P'(\mu_2)\hat{v}_{q_1+1}(\gamma) + P(\mu_2)\hat{v}_{q_1+2}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+2}(\gamma), \\ &\vdots \\ \tfrac{\gamma^{q_2-1}}{(q_2-1)!}P^{(q_2-1)}(\mu_2)\hat{v}_{q_1+1}(\gamma) + \cdots + P(\mu_2)\hat{v}_{q_1+q_2}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+q_2}(\gamma), \end{aligned} \right.$$

$$\vdots \tag{11}$$

$$\left\{ \begin{aligned} P(\mu_s)\hat{v}_{q_1+\cdots+q_{s-1}+1}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+\cdots+q_{s-1}+1}(\gamma), \\ \gamma P'(\mu_s)\hat{v}_{q_1+\cdots+q_{s-1}+1}(\gamma) + P(\mu_s)\hat{v}_{q_1+\cdots+q_{s-1}+2}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+\cdots+q_{s-1}+2}(\gamma), \\ &\vdots \\ \tfrac{\gamma^{q_s-1}}{(q_s-1)!}P^{(q_s-1)}(\mu_s)\hat{v}_{q_1+\cdots+q_{s-1}+1}(\gamma) + \cdots + P(\mu_s)\hat{v}_{q_1+\cdots+q_{s-1}+q_s}(\gamma) &= s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)\hat{u}_{q_1+\cdots+q_{s-1}+q_s}(\gamma). \end{aligned} \right.$$

Assume that the k vectors $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma) \in \mathbb{C}^n$ are linearly independent. Then, for every $i = 1, 2, \ldots, k$, the vector $\hat{v}_i(\gamma)$ is nonzero, and it is a linear combination of the vectors $v_1(\gamma), v_2(\gamma), \ldots, v_i(\gamma)$, where the coefficient of $v_i(\gamma)$ in this combination is equal to 1. The k vectors $\hat{v}_1(\gamma), \hat{v}_2(\gamma), \ldots, \hat{v}_k(\gamma)$ play a leading role in computing the desired perturbation $\Delta_\gamma(\lambda)$.

Define the $n \times k$ matrices

$$\hat{U}(\gamma) = \left[\, \hat{u}_1(\gamma)\ \hat{u}_2(\gamma)\ \cdots\ \hat{u}_k(\gamma) \,\right] \quad \text{and} \quad \hat{V}(\gamma) = \left[\, \hat{v}_1(\gamma)\ \hat{v}_1(\gamma)\ \cdots\ \hat{v}_k(\gamma) \,\right],$$

and observe that the linear independence of the vectors $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma)$ implies also that $\operatorname{rank}(\hat{V}(\gamma)) = k$ and $\hat{V}(\gamma)^\dagger \hat{V}(\gamma) = I_k$, where $\hat{V}(\gamma)^\dagger$ denotes the *Moore-Penrose pseudoinverse* of $\hat{V}(\gamma)$.

We are constructing the desired perturbed matrix polynomial $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$ by assembling a perturbation $\Delta_\gamma(\lambda) = \sum_{j=0}^{m} \Delta_{\gamma,j} \lambda^j$, in which

$$\Delta_{\gamma,j} = \frac{1}{s} \sum_{i=1}^{s} \left(\frac{1}{w(|\mu_i|)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j w_j \right) \Delta_\gamma, \quad j = 0, 1, \ldots, m, \tag{12}$$

for some $n \times n$ matrix Δ_γ that has to be computed.

Denoting by $\Delta_\gamma^{(p)}(\lambda)$ the p-th derivative of $\Delta_\gamma(\lambda)$ with respect to λ, we have

$$\Delta_\gamma^{(p)}(\lambda) = \sum_{j=p}^{m} \Delta_{\gamma,j} \left(\prod_{\xi=0}^{p-1} (j - \xi) \right) \lambda^{j-p}.$$

Thus, substituting $\Delta_{\gamma,j}$ into $\Delta_\gamma^{(p)}(\lambda)$ and calculating the derivative for the scalar μ_r ($r = 1, 2, \ldots, s$) yield

$$\Delta_\gamma^{(p)}(\mu_r) = \underbrace{\frac{1}{s} \sum_{j=p}^{m} \left\{ \left(\sum_{i=1}^{s} \frac{1}{w(|\mu_i|)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j \mu_r^{j-p} \right) \left(\prod_{\xi=0}^{p-1} (j - \xi) \right) w_j \right\}}_{\beta_{r,p}} \Delta_\gamma.$$

Motivated by this relation, for $r = 1, 2, \ldots, s$ and $p = 0, 1, \ldots, q_r$, we define the quantities

$$\beta_{r,p} = \frac{1}{s} \sum_{j=p}^{m} \left\{ \left(\sum_{i=1}^{s} \frac{1}{w(|\mu_i|)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j \mu_r^{j-p} \right) \left(\prod_{\xi=0}^{p-1} (j - \xi) \right) w_j \right\},$$

where, for convention, we set $\frac{\overline{\mu}_i}{|\mu_i|} = 0$ whenever $\mu_i = 0$. In particular, for $p = 0$,

$$\beta_{r,0} = \frac{1}{s} \sum_{j=0}^{m} \left[\sum_{i=1}^{s} \left(\frac{1}{w(|\mu_i|)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \mu_r \right)^j \right) w_j \right], \quad r = 1, 2, \ldots, s. \tag{13}$$

Hence, we have

$$\Delta_\gamma^{(p)}(\mu_r) = \beta_{r,p} \Delta_\gamma, \quad r = 1, 2, \ldots, s, \; p = 0, 1, \ldots, q_r. \tag{14}$$

To construct the matrix Δ_γ assume that the quantities $\beta_{1,0}, \beta_{2,0}, \ldots, \beta_{s,0}$ are nonzero. Then, for any $r = 1, 2, \ldots, s$, define the $q_r \times q_r$ Toeplitz upper triangular matrix

$$M^{[r]} = \left[M_{i,j}^{[r]}\right] = \begin{bmatrix} \frac{1}{\beta_{r,0}} & -\gamma\frac{\beta_{r,1}}{\beta_{r,0}^2} & -\frac{\gamma^2}{2!}\frac{\beta_{r,2}}{\beta_{r,0}^2} + \gamma^2\frac{\beta_{r,1}^2}{\beta_{r,0}^3} & \cdots & * & * \\ 0 & \frac{1}{\beta_{r,0}} & -\gamma\frac{\beta_{r,1}}{\beta_{r,0}^2} & \cdots & * & * \\ 0 & 0 & \frac{1}{\beta_{r,0}} & \ddots & \vdots & \vdots \\ \vdots & \vdots & 0 & \ddots & -\gamma\frac{\beta_{r,1}}{\beta_{r,0}^2} & -\frac{\gamma^2}{2!}\frac{\beta_{r,2}}{\beta_{r,0}^2} + \gamma^2\frac{\beta_{r,1}^2}{\beta_{r,0}^3} \\ \vdots & \vdots & \vdots & & \frac{1}{\beta_{r,0}} & -\gamma\frac{\beta_{r,1}}{\beta_{r,0}^2} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{\beta_{r,0}} \end{bmatrix}$$

whose diagonal entries $M_{1,1}^{[r]}, M_{2,2}^{[r]}, \ldots, M_{q_r,q_r}^{[r]}$ are all equal to $\frac{1}{\beta_{r,0}}$, and the entries above the (main) diagonal are given by the following recursive formula:

$$M_{i,j}^{[r]} = -\frac{1}{\beta_{r,0}} \sum_{\xi=1}^{j-i} \frac{\gamma^\xi}{\xi!} \beta_{r,\xi} M_{i,j-\xi}^{[r]}, \quad 1 \le i < j \le q_r. \tag{15}$$

Define also the $k \times k$ block diagonal matrix $M = \operatorname{diag}\left\{M^{[1]}, M^{[2]}, \ldots, M^{[s]}\right\}$. Eventually, the $n \times n$ matrix Δ_γ that we are looking for, is of the form

$$\Delta_\gamma = -s_\rho\left(F_\gamma\left[P, \Sigma\right]\right) \hat{U}(\gamma) M \hat{V}(\gamma)^\dagger. \tag{16}$$

Consider the perturbation

$$\Delta_\gamma(\lambda) = \sum_{j=0}^{m} \Delta_{\gamma,j}\lambda^j = \sum_{j=0}^{m} \left[\frac{1}{s}\sum_{i=1}^{s}\left(\frac{1}{w\left(|\mu_i|\right)}\left(\frac{\bar{\mu}_i}{|\mu_i|}\right)^j w_j\right)\Delta_\gamma\right]\lambda^j.$$

In the remainder of this section, it will be obtained that the prescribed scalars $\mu_1, \mu_2, \ldots, \mu_s$ are eigenvalues of the perturbed matrix polynomial

$$Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda) = \sum_{j=0}^{m} \left(A_j + \Delta_{\gamma,j}\right)\lambda^j \tag{17}$$

with their multiplicities greater than or equal to $q_1, q_2, \ldots, q_s$, respectively. In particular, for $\gamma > 0$, it can be proved that the sets

$$\begin{gathered} \left\{\hat{v}_1(\gamma),\ \frac{1}{\gamma}\,\hat{v}_2(\gamma),\ \frac{1}{\gamma^2}\,\hat{v}_3(\gamma),\ \ldots,\ \frac{1}{\gamma^{q_1-1}}\,\hat{v}_{q_1}(\gamma)\right\}, \\ \left\{\hat{v}_{q_1+1}(\gamma),\ \frac{1}{\gamma}\,\hat{v}_{q_1+2}(\gamma),\ \frac{1}{\gamma^2}\,\hat{v}_{q_1+3}(\gamma),\ \ldots,\ \frac{1}{\gamma^{q_2-1}}\,\hat{v}_{q_1+q_2}(\gamma)\right\}, \\ \vdots \\ \left\{\hat{v}_{q_1+\cdots+q_{s-1}+1}(\gamma),\ \frac{1}{\gamma}\,\hat{v}_{q_1+\cdots+q_{s-1}+2}(\gamma),\ \frac{1}{\gamma^2}\,\hat{v}_{q_1+\cdots+q_{s-1}+3}(\gamma),\ \ldots,\ \frac{1}{\gamma^{q_s-1}}\,\hat{v}_{q_1+\cdots+q_s}(\gamma)\right\} \end{gathered} \tag{18}$$

form s Jordan chains of $Q_\gamma(\lambda)$, corresponding to its eigenvalues $\mu_1, \mu_2, \ldots, \mu_s$, respectively.

In order to avoid unnecessary prolix computations, we restrict ourselves in proving that the vectors $\hat{v}_1(\gamma), \frac{1}{\gamma}\hat{v}_2(\gamma), \ldots, \frac{1}{\gamma^{q_1-1}}\hat{v}_{q_1}(\gamma)$ form a Jordan chain of $Q_\gamma(\lambda)$ corresponding to μ_1 as one of its eigenvalues; the extension to the remaining scalars $\mu_2, \mu_3, \ldots, \mu_s$ is straightforward. For this purpose, in the system (11), we consider the p-th equation ($1 \leq p \leq q_1$) of the subsystem that corresponds to μ_1. Since the matrix $\hat{V}(\gamma)$ is assumed to be full column rank, we have

$$\Delta_\gamma \hat{V}(\gamma) = -s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \hat{U}(\gamma) M.$$

In addition, since we are dealing with the first q_1 equations of system (11), we only consider the first block of matrix M, $M^{[1]}$, which is of size $q_1 \times q_1$, and the vectors $\hat{u}_1(\gamma), \hat{u}_2(\gamma), \ldots, \hat{u}_{q_1}(\gamma)$ and $\hat{v}_1(\gamma), \hat{v}_2(\gamma), \ldots, \hat{v}_{q_1}(\gamma)$. For this case, we have

$$\Delta_\gamma \left[\hat{v}_1(\gamma)\ \hat{v}_2(\gamma)\ \cdots\ \hat{v}_{q_1}(\gamma)\right] = -s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \left[\hat{u}_1(\gamma)\ \hat{u}_2(\gamma)\ \cdots\ \hat{u}_{q_1}(\gamma)\right] M^{[1]}.$$

After doing these matrix multiplications, it is straightforward to see that the j-th column of the result in the left-hand side is $\Delta_\gamma \hat{v}_j(\gamma)$ ($j = 1, 2, \ldots, q_1$), while the j-th column of the result in the right-hand side is $-s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \sum_{i=1}^{j} M_{i,j}^{[1]} \hat{u}_i(\gamma)$. Replacing j with $p-j$ yields

$$\Delta_\gamma \hat{v}_{p-j}(\gamma) = -s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \sum_{i=1}^{p-j} M_{i,p-j}^{[1]} \hat{u}_i(\gamma), \quad j = p-1, p-2, \ldots, p-q_1. \tag{19}$$

By the p-th equation in (11), (14) and (19), it follows that the perturbed matrix polynomial $Q_\gamma(\lambda)$ in (17) satisfies

$$\begin{aligned}
\sum_{j=0}^{p-1} \frac{\gamma^j}{j!} Q_\gamma^{(j)}(\mu_1) \hat{v}_{p-j}(\gamma) &= \sum_{j=0}^{p-1} \frac{\gamma^j}{j!} P^{(j)}(\mu_1) \hat{v}_{p-j}(\gamma) + \sum_{j=0}^{p-1} \frac{\gamma^j}{j!} \Delta_\gamma^{(j)}(\mu_1) \hat{v}_{p-j}(\gamma) \\
&= s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \hat{u}_p(\gamma) + \sum_{j=0}^{p-1} \frac{\gamma^j}{j!} \beta_{1,j} \Delta_\gamma \hat{v}_{p-j}(\gamma) \\
&= s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \hat{u}_p(\gamma) + \sum_{j=0}^{p-1} \left(\frac{\gamma^j}{j!} \beta_{1,j} \left(-s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \sum_{i=1}^{p-j} M_{i,p-j}^{[1]} \hat{u}_i(\gamma) \right) \right) \\
&= s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \hat{u}_p(\gamma) - s_\rho \left(F_\gamma \left[P, \Sigma\right]\right) \sum_{j=0}^{p-1} \left(\frac{\gamma^j}{j!} \beta_{1,j} \sum_{i=1}^{p-j} M_{i,p-j}^{[1]} \hat{u}_i(\gamma) \right).
\end{aligned} \tag{20}$$

In (20), we observe that

$$\begin{aligned}\sum_{j=0}^{p-1}\left(\frac{\gamma^j}{j!}\beta_{1,j}\sum_{i=1}^{p-j}M^{[1]}_{i,p-j}\hat{u}_i(\gamma)\right) &= \beta_{1,0}\sum_{i=1}^{p}M^{[1]}_{i,p}\hat{u}_i(\gamma)+\gamma\beta_{1,1}\sum_{i=1}^{p-1}M^{[1]}_{i,p-1}\hat{u}_i(\gamma)\\ &\quad+\cdots+\beta_{1,p-2}\frac{\gamma^{p-2}}{(p-2)!}\sum_{i=1}^{2}M^{[1]}_{i,1}\hat{u}_i(\gamma)\\ &\quad+\beta_{1,p-1}\frac{\gamma^{p-1}}{(p-1)!}M^{[1]}_{1,1}\hat{u}_1(\gamma),\end{aligned}$$

Denoting the coefficients of $\hat{u}_1(\gamma),\hat{u}_2(\gamma),\ldots,\hat{u}_{p-1}(\gamma)$ in (20) by $\alpha_1,\alpha_2,\ldots,\alpha_{p-1}$, respectively, for any $i=1,2,\ldots,p-1$, we have

$$\begin{aligned}\alpha_i &= -s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\sum_{\xi=0}^{p-i}\frac{\gamma^\xi}{\xi!}\beta_{1,\xi}M^{[1]}_{i,p-\xi}\\ &= -s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\beta_{1,0}M^{[1]}_{i,p}-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\sum_{\xi=1}^{p-i}\frac{\gamma^\xi}{\xi!}\beta_{1,\xi}M^{[1]}_{i,p-\xi}\\ &= -s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\beta_{1,0}\left(-\frac{1}{\beta_{1,0}}\sum_{\xi=1}^{p-i}\frac{\gamma^\xi}{\xi!}\beta_{1,\xi}M^{[1]}_{i,p-\xi}\right)-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\sum_{\xi=1}^{p-i}\frac{\gamma^\xi}{\xi!}\beta_{1,\xi}M^{[1]}_{i,p-\xi}\\ &= 0.\end{aligned}$$

Moreover, it is apparent that

$$-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\beta_{1,0}M^{[1]}_{p,p}=-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\beta_{1,0}\frac{1}{\beta_{1,0}}=-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right).$$

As a consequence, for any $p\in\{1,2,\ldots,q_1\}$, it holds

$$\sum_{j=0}^{p-1}\frac{\gamma^j}{j!}Q^{(j)}_\gamma(\mu_1)\hat{v}_{p-j}(\gamma)=s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\hat{u}_p(\gamma)-s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\hat{u}_p(\gamma)=0.$$

Dividing this relation by γ^{1-p} yields

$$\sum_{j=0}^{p-1}\frac{1}{j!}Q^{(j)}_\gamma(\mu_1)\left(\gamma^{j-p+1}\hat{v}_{p-j}(\gamma)\right)=0,$$

which means that μ_1 is an eigenvalue of algebraic multiplicity at least q_1 of $Q_\gamma(\lambda)$, with

$$\left\{\hat{v}_1(\gamma),\ \frac{1}{\gamma}\hat{v}_2(\gamma),\ \frac{1}{\gamma^2}\hat{v}_3(\gamma),\ \ldots,\ \frac{1}{\gamma^{q_1-1}}\hat{v}_{q_1}(\gamma)\right\}$$

as a corresponding Jordan chain.

The next theorem summarizes the results obtained so far.

Theorem 3.3. *Let $P(\lambda)$ be a matrix polynomial as in* (1)*, and let Σ be a set of prescribed scalars as in* (2) *such that the quantities $\beta_{1,0}, \beta_{2,0}, \ldots, \beta_{s,0}$ defined by* (13) *are nonzero. Then, for any $\gamma > 0$ such that the vectors $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma)$ are linearly independent, the scalars $\mu_1, \mu_2, \ldots, \mu_s$ are eigenvalues of the perturbed matrix polynomial $Q_\gamma(\lambda)$ given by* (17)*, with algebraic multiplicities greater than or equal to $q_1, q_2, \ldots, q_s$, respectively. Moreover, the sets in* (18) *are Jordan chains of $Q_\gamma(\lambda)$ corresponding to $\mu_1, \mu_2, \ldots, \mu_s$, respectively.*

Remark 3.4. The discussion in this section and the construction of the perturbed matrix polynomial $Q_\gamma(\lambda)$ in (17) generalize main results of [9, 18]; in particular, they yield the results of [9, Section 3] when $s = k\,(\leq n)$, and the results of [18, Section 3] when $s = 1$.

Remark 3.5. As mentioned in [9, 18], it is not easy to obtain conditions ensuring that the quantities $\beta_{1,0}, \beta_{2,0}, \ldots, \beta_{s,0}$ are nonzero and/or the vectors $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma)$ are linearly independent. However, in all our experiments, these two required conditions hold generically.

4 Bounds for $D_w(P, \Sigma)$

In this section, we give an upper and a lower bound for the spectral norm distance $D_w(P, \Sigma)$ introduced in Definition 2.2. First, we see that an upper bound for $D_w(P, \Sigma)$ is directly obtained by the construction of the perturbed matrix polynomial $Q_\gamma(\lambda)$ in (17). In particular, by (12), it follows

$$\|\Delta_{\gamma,j}\|_2 \leq \frac{w_j}{s}\left(\sum_{i=1}^{s}\frac{1}{w(|\mu_i|)}\right)\|\Delta_\gamma\|_2, \quad j = 0, 1, \ldots, m.$$

Assume that $\beta_{i,0} \neq 0$, $i = 1, 2, \ldots, s$, and the vectors $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma)$ are linearly independent for some $\gamma > 0$. Recalling Definition 2.1, the distance $D_w(P, \Sigma)$ satisfies

$$D_w(P, \Sigma) \leq \frac{1}{s}\left(\sum_{i=1}^{s}\frac{1}{w(|\mu_i|)}\right)\|\Delta_\gamma\|_2\,. \tag{21}$$

In the remainder of this section, a lower bound for $D_w(P, \Sigma)$ is computed. At this point, it is necessary to recall that having an eigenvalue μ_i of the matrix polynomial $P(\lambda)$ with algebraic multiplicity q_i does not necessarily mean that μ_i has a Jordan chain of length q_i. Actually, it means that if the eigenvalue μ_i has geometric multiplicity g_i, then $P(\lambda)$ has g_i Jordan chains associated to g_i (not necessarily linearly independent) eigenvectors, with total number of vectors equal to the algebraic multiplicity q_i [3, 10, 15]. Thus, some concepts and discussions are needed to cope with this difficulty by considering what is presented in Appendix A. Moreover, to compute a lower bound, linear independence of the vectors $v_1(\gamma), v_2(\gamma), \dots, v_k(\gamma)$ is not required, but the weights $w_1, \dots, w_{m-1}$ are needed to be positive; recall that from the definition of perturbations of $P(\lambda)$, it is assumed that $w_0 > 0$.

Lemma 4.1. *Let $P(\lambda)$ be a matrix polynomial as in* (1)*, and let Σ be a set as described in* (2)*. Suppose that the spectrum of $P(\lambda)$ contains Σ and each eigenvalue $\mu_i \in \sigma(P)$ $(i = 1, 2, \dots, s)$ has a Jordan chain of length q_i. Then, $s_\rho\left(F_\gamma[P, \Sigma]\right) = 0$ for any $\gamma > 0$ (recall that $\rho = nk - k + 1$).*

Proof. Suppose that $\mu_1, \mu_2, \dots, \mu_s$ are eigenvalues of the matrix polynomial $P(\lambda)$ with algebraic multiplicities at least $q_1, q_2, \dots, q_s$, respectively, and $q_1 + q_2 + \cdots + q_s = k \le n$. Let also γ be a positive number. By hypothesis, there exist k (not necessarily linearly independent) vectors $y_1, y_2, \dots, y_k \in \mathbb{C}^n$ such that $y_1, y_{q_1+1}, y_{q_1+q_2+1}, \dots, y_{q_1+\cdots+q_{s-1}+1}$ are nonzero and the following s sets of matrix equations are satisfied:

$$
\begin{cases}
P(\mu_1)\, y_1 = 0, \\
\gamma P'(\mu_1)\, y_1 + P(\mu_1)\, y_2 = 0, \\
\quad \vdots \qquad\qquad \vdots \\
\frac{\gamma^{q_1-1}}{(q_1-1)!} P^{(q_1-1)}(\mu_1)\, y_1 + \cdots + P(\mu_1)\, y_{q_1} = 0,
\end{cases}
$$

$$
\begin{cases}
P(\mu_2)\, y_{q_1+1} = 0, \\
\gamma P'(\mu_2)\, y_{q_1+1} + P(\mu_2)\, y_{q_1+2} = 0, \\
\quad \vdots \qquad\qquad \vdots \\
\frac{\gamma^{q_2-1}}{(q_2-1)!} P^{(q_2-1)}(\mu_2)\, y_{q_1+1} + \cdots + P(\mu_2)\, y_{q_1+q_2} = 0,
\end{cases}
$$

$$
\vdots \tag{22}
$$

$$
\begin{cases}
P(\mu_s)\, y_{q_1+\cdots+q_{s-1}+1} = 0, \\
\gamma \hat{P}'(\mu_s)\, y_{q_1+\cdots+q_{s-1}+1} + P(\mu_s)\, y_{q_1+\cdots+q_{s-1}+2} = 0, \\
\quad \vdots \qquad\qquad \vdots \\
\frac{\gamma^{q_s-1}}{(q_s-1)!} P^{(q_s-1)}(\mu_s)\, y_{q_1+\cdots+q_{s-1}+1} + \cdots + P(\mu_s)\, y_{q_1+\cdots+q_{s-1}+q_s} = 0.
\end{cases}
$$

(In other words, the vectors $y_1, \frac{1}{\gamma} y_2, \dots, \frac{1}{\gamma^{q_1-1}} y_{q_1}$ form a Jordan chain of length q_1 corresponding to μ_1, the vectors $y_{q_1+1}, \frac{1}{\gamma} y_{q_1+2}, \dots, \frac{1}{\gamma^{q_2-1}} y_{q_1+q_2}$ form a Jordan chain of length q_2 corresponding to μ_2, and so on.)

Recall the $nk \times nk$ matrix $\hat{F}_\gamma[P, \Sigma]$ and the $nk \times nk$ nonsingular block upper triangular matrix T which are defined in the previous section and satisfy (7). Consider also the (nonzero) linearly independent vectors

$$\begin{bmatrix} 0_n \\ \vdots \\ 0_n \\ 0_n \\ y_1 \\ 0_{nq_2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \begin{bmatrix} 0_n \\ \vdots \\ 0_n \\ y_1 \\ y_2 \\ 0_{nq_2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \ldots, \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{q_1-1} \\ y_{q_1} \\ 0_{nq_2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \begin{bmatrix} 0_{nq_1} \\ 0_n \\ \vdots \\ 0_n \\ 0_n \\ y_{q_1+1} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \begin{bmatrix} 0_{nq_1} \\ 0_n \\ \vdots \\ 0_n \\ y_{q_1+1} \\ y_{q-1+2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \ldots, \begin{bmatrix} 0_{nq_1} \\ y_{q_1+1} \\ y_{q_1+2} \\ \vdots \\ y_{q_1+q_2-1} \\ y_{q_1+q_2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_s} \end{bmatrix}, \ldots, \begin{bmatrix} 0_{nq_1} \\ 0_{nq_2} \\ 0_{nq_3} \\ \vdots \\ 0_{nq_{s-1}} \\ y_{q_1+\cdots+q_{s-1}+1} \\ y_{q_1+\cdots+q_{s-1}+2} \\ \vdots \\ y_k \end{bmatrix},$$

where 0_n denotes the zero vector of order n. By (22), it follows readily that all these vectors are null vectors of matrix $\hat{F}_\gamma[P, \Sigma]$. By the similarity of $F_\gamma[P, \Sigma]$ and $\hat{F}_\gamma[P, \Sigma]$, the proof is complete. □

Consider a perturbation of matrix polynomial $P(\lambda)$, $Q(\lambda) = P(\lambda) + \Delta(\lambda)$. By following exactly the structure of $F_\gamma[P, \Sigma]$ in Definition 2.3, construct the $nk \times nk$ block lower triangular matrices $F_\gamma[Q, \Sigma]$ and $F_\gamma[\Delta, \Sigma]$. Clearly, it holds that $F_\gamma[Q, \Sigma] = F_\gamma[P, \Sigma] + F_\gamma[\Delta, \Sigma]$. As a consequence, Lemma 4.1 and the Weyl inequalities for singular values (e.g. see Corollary 5.1 of [1]) imply the following:

Corollary 4.2. *Let $P(\lambda)$ be a matrix polynomial as in* (1)*, and let Σ be a set of prescribed scalars as in* (2)*. Suppose that $\gamma > 0$, Σ is a subset of the spectrum of an $n \times n$ matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$, and each eigenvalue $\mu_i \in \sigma(Q)$ $(i = 1, 2, \ldots, s)$ has a Jordan chain of length q_i. Then, $s_\rho(F_\gamma[P, \Sigma]) \leq \|F_\gamma[\Delta, \Sigma]\|_2$.*

For convenience, denote

$$F_\gamma[\Delta, \Sigma] = \begin{bmatrix} \mathcal{F}_{1,1} & 0 & \cdots & 0 \\ \mathcal{F}_{2,1} & \mathcal{F}_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{F}_{k,1} & \mathcal{F}_{k,2} & \cdots & \mathcal{F}_{k,k} \end{bmatrix} \in \mathbb{C}^{nk \times nk},$$

with $\mathcal{F}_{i,j} \in \mathbb{C}^{n \times n}$, $1 \leq j \leq i \leq k$. Moreover, for the weight polynomial $w(\lambda) = w_m \lambda^m + w_{m-1}\lambda^{m-1} + \cdots + w_1 \lambda + w_0$, assuming that one can use the term "Jordan coefficient" for scalar polynomials as for matrix polynomials, define the $k \times k$ lower triangular matrix

$$\mathrm{F}_\gamma[w, \Sigma] = \begin{bmatrix} f_{1,1} & 0 & \cdots & 0 \\ f_{2,1} & f_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{k,1} & f_{k,2} & \cdots & f_{k,k} \end{bmatrix}$$

such that $f_{i,i} = w(|\mu_i|)$ $(i = 1, 2, \ldots, k)$, and analogous to (iii) of Definition 2.3, the entries below the diagonal are given by the recursive formula (starting from the diagonal entries)

$$f_{i,j} = \begin{cases} \gamma^{i-j}(\text{the next Jordan chain coefficient of } f_{i-1,j}) & \text{if } f_{i-1,j} = f_{i,j+1}, \\ |\theta_{j,i}|\,(f_{i-1,j} + f_{i,j+1}) & \text{if } f_{i-1,j} \neq f_{i,j+1}. \end{cases}$$

Next lemma yields a lower bound of the distance $D_w(P, \Sigma)$.

Lemma 4.3. *Let $P(\lambda)$ be a matrix polynomial as in* (1)*, Σ be a set of prescribed scalars as in* (2)*, and all the weights $w_0, w_1, \ldots, w_{m-1}$ be positive. Consider a matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$ in $\mathcal{B}(P, \varepsilon, w)$ having the set Σ in its spectrum. Then, for any $\gamma > 0$,*

$$\varepsilon \geq \frac{s_\rho\left(F_\gamma\left[P, \Sigma\right]\right)}{\|\mathrm{F}_\gamma[w, \Sigma]\|_2}.$$

Proof. The set $\mathcal{B}(P, \varepsilon, w)$ is closed and for any positive integer p, there is a matrix polynomial $Q_p(\lambda) \in \mathcal{B}(Q, 1/p, w)$ that lies in the interior of $\mathcal{B}(P, \varepsilon, w)$ and has a nonsingular leading coefficient. Moreover, by Corollary A.6 of Appendix, for each p and any positive integer q, there is a $Q_{p,q}(\lambda) \in \mathcal{B}(Q_p, 1/q, w)$ such that $\sigma(Q_{p,q}) = \sigma(Q_p)$ and all the eigenvalues of $Q_{p,q}(\lambda)$ have geometric multiplicity 1 (i.e., every eigenvalue of $Q_{p,q}(\lambda)$ has exactly one Jordan chain of length equal to the algebraic multiplicity of the eigenvalue). Hence, there is a sequence of matrix polynomials in the interior of $\mathcal{B}(P, \varepsilon, w)$ having all their eigenvalues of geometric multiplicity 1, which converges to the perturbed matrix polynomial $Q(\lambda)$. As a consequence, by the continuity of the Jordan structure and the singular value decomposition (with respect to matrix entries), without loss of generality, we may assume that each eigenvalue μ_i of $Q(\lambda)$ $(i = 1, 2, \ldots, s)$ has a Jordan chain of length q_i. Then, by Corollary 4.2, $s_\rho(F_\gamma[P, \Sigma]) \leq \|F_\gamma[\Delta, \Sigma]\|_2$.

The rest of the proof is devoted to obtain the inequality $\|F_\gamma[\Delta, \Sigma]\|_2 \leq \varepsilon \|\mathrm{F}_\gamma[w, \Sigma]\|_2$. To do this, for $i = 1, 2, \ldots, k$, observe that

$$\|\Delta(\mu_i)\|_2 \leq \sum_{j=0}^{m} \|\Delta_j\|_2 |\mu_i|^j \leq \varepsilon \sum_{j=0}^{m} w_j |\mu_i|^j = \varepsilon w(|\mu_i|). \tag{23}$$

Moreover, for every $i = 1, 2, \ldots, s$ and $p = 1, 2, \ldots, q_i$,

$$\begin{aligned} \left\|\Delta^{(p)}(\mu_i)\right\|_2 &\leq \sum_{j=p}^{m} j(j-1)\cdots(j-p+1)\,\|\Delta_j\|_2 |\mu_i|^{j-p} \\ &\leq \varepsilon \sum_{j=p}^{m} j(j-1)\cdots(j-p+1) w_j |\mu_i|^{j-p} = \varepsilon w^{(p)}(|\mu_i|). \end{aligned} \tag{24}$$

Consequently, keeping in mind the definition of $\mathrm{F}_\gamma[w,\Sigma]$, and using (23) and (24), one can verify that (see the proof of Lemma 4.2 of [9] and the discussion before it)

$$\|\mathcal{F}_{i,i}\|_2 \le \varepsilon f_{i,i},$$

and

$$\begin{aligned}
\|\mathcal{F}_{i,j}\|_2 &= \|\theta_{i,j}\left(\mathcal{F}_{i-1,j} - \mathcal{F}_{i,j+1}\right)\|_2 \\
&\le |\theta_{i,j}|\left(\|\mathcal{F}_{i-1,j}\|_2 + \|\mathcal{F}_{i,j+1}\|_2\right) \\
&\le \varepsilon\,|\theta_{i,j}|\left(f_{i-1,j} + f_{i,j+1}\right) \,=\, \varepsilon f_{i,j}.
\end{aligned}$$

Then, for any $\gamma \neq 0$, a unit vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{C}^{kn} \quad (x_i \in \mathbb{C}^n,\ i = 1,2,\ldots,k)$$

can be considered such that

$$\begin{aligned}
\|F_\gamma[\Delta,\Sigma]\|_2^2 &= \|F_\gamma[\Delta,\Sigma]\,x\|_2^2 \\
&= \|\mathcal{F}_{1,1}x_1\|_2^2 + \|\mathcal{F}_{2,1}x_1 + \mathcal{F}_{2,2}x_2\|_2^2 + \cdots + \left\|\sum_{i=1}^{k}\mathcal{F}_{k,i}x_i\right\|_2^2 \\
&\le (\varepsilon f_{1,1})^2\|x_1\|_2^2 + (\varepsilon f_{2,1})^2\|x_1\|_2^2 + (\varepsilon f_{2,2})^2\|x_2\|_2^2 \\
&\quad + (\varepsilon f_{2,1})(\varepsilon f_{2,2})\|x_1\|_2\|x_2\|_2 + \cdots + (\varepsilon f_{k,k})^2\|x_k\|_2^2 \\
&= \varepsilon^2\left\|\begin{bmatrix} f_{1,1} & 0 & \cdots & 0 \\ f_{2,1} & f_{2,2} & \cdots & 0 \\ f_{3,1} & f_{3,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{k,1} & f_{k,2} & \cdots & f_{k,k}\end{bmatrix}\begin{bmatrix}\|x_1\|_2 \\ \|x_2\|_2 \\ \vdots \\ \|x_k\|_2\end{bmatrix}\right\|_2^2 \\
&\le \varepsilon^2\|\mathrm{F}_\gamma[w,\Sigma]\|_2^2.
\end{aligned}$$

This completes the proof. □

By the above lemma (and according to Definition 2.2), it follows

$$D_w(P,\Sigma) \ge \frac{s_\rho\left(F_\gamma[P,\Sigma]\right)}{\|\mathrm{F}_\gamma[w,\Sigma]\|_2}. \tag{25}$$

It will be convenient to denote the lower bound in (25) by $\beta_{low,\delta}(P,\Sigma,\gamma)$ and the upper bound in (21) by $\beta_{up}(P,\Sigma,\gamma)$, i.e.,

$$\beta_{low,\delta}(P,\Sigma,\gamma) = \frac{s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)}{\left\|\mathrm{F}_\gamma\left[w,\Sigma\right]\right\|_2}$$

and

$$\beta_{up}(P,\Sigma,\gamma) = \frac{1}{s}\sum_{i=1}^{s}\left(\frac{1}{w\left(|\mu_i|\right)}\right)\|\Delta_\gamma\|_2.$$

The main results of this section are summarized in the next theorem.

Theorem 4.4. *Let $P(\lambda)$ be a matrix polynomial as in* (1)*, and let Σ be a set of prescribed scalars as in* (2)*.*

(a) *If all the weights $w_0, w_1, \ldots, w_m$ are positive, then for any $\gamma > 0$, $D_w(P,\Sigma) \geq \beta_{low,\delta}(P,\Sigma,\gamma)$.*

(b) *If the scalars $\beta_{1,0}, \beta_{2,0}, \ldots, \beta_{s,0}$ in* (13) *are nonzero, then for any $\gamma > 0$ such that $v_1(\gamma), v_2(\gamma), \ldots, v_k(\gamma)$ are linearly independent vectors, $D_w(P,\Sigma) \leq \beta_{up}(P,\Sigma,\gamma)$. Moreover, the matrix polynomial $Q_\gamma(\gamma)$ in* (17) *lies on the boundary of $\mathcal{B}(P,\beta_{up}(P,\Sigma,\gamma),w)$.*

5 A numerical example

In this section, we verify the validity and the effectiveness of our results and discussions in the previous sections, by presenting a comprehensive numerical experiment. Following the methodology given in Section 3, we construct an associated perturbation of a given matrix polynomial. Also, we compute lower and upper bounds for the distance $D_w(P,\Sigma)$, according to the results obtained in Section 4. All computations were performed in MATLAB with 16 significant digits; however, for simplicity, all numerical results are shown with 4 decimal places.

Example 5.1. Consider the 6×6 matrix polynomial

$$P(\lambda) = I_6\lambda^2 + \begin{bmatrix} -4 & 1 & 0 & -1 & 4 & 0 \\ 0 & 8 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & -5 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}\lambda + \begin{bmatrix} 3 & 0 & 0 & 0 & -3 & 2 \\ 0 & 16 & -8 & 1 & 0 & -1 \\ 0 & 0 & 4 & 0 & 11 & 0 \\ 0 & 0 & 0 & 4 & -5 & 0 \\ 0 & 0 & 0 & 0 & 9 & 2 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$$

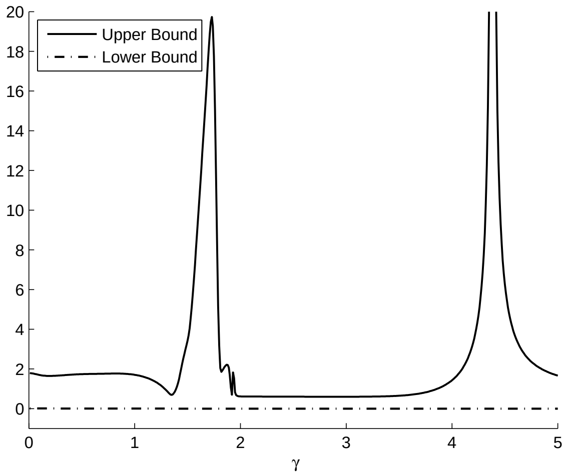

Figure 1: The graphs of the upper bound $\beta_{up}(P, \Sigma, \gamma)$ and the lower bound $\beta_{low}(P, \Sigma, \gamma)$.

and the set

$$\Sigma = \{\mu_1, \mu_1, \mu_2, \mu_3, \mu_3, \mu_3\} = \{1, 1, 2, 3, 3, 3\},$$

i.e., $s = 3$, $q_1 = 2$, $q_2 = 1$, $q_3 = 3$ and $k = q_1 + q_2 + q_3 = 6\ (= n)$. Consider also the set of weights $w = \{18.2014, 10.9003, 1\}$, which are the spectral norms of the corresponding coefficient matrices. Figure 1 illustrates the graphs of the upper bound $\beta_{up}(P, \Sigma, \gamma)$ and the lower bound $\beta_{low}(P, \Sigma, \gamma)$ for $\gamma \in (0, 5]$.

For the value $\gamma = 3$, we construct a perturbed matrix polynomial $Q_3(\lambda) = P(\lambda) + \Delta_3(\lambda)$, whose spectrum contains the set Σ. The matrix $F_3[P, \Sigma]$ is given by (4), the singular value $s_\rho(F_\gamma[P, \Sigma]) = s_{31}(F_3[P, \Sigma])$ is equal to 1.0984, and the matrix T in (8) is of the form

$$T = \begin{bmatrix} I_6 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_6 & 0 & 0 & 0 & 0 \\ 9I_6 & 3I_6 & I_6 & 0 & 0 & 0 \\ 6.75I_6 & 4.5I_6 & 3I_6 & I_6 & 0 & 0 \\ 43.8750I_6 & -20.25I_6 & -9I_6 & 0 & I_6 & 0 \\ 167.0625I_6 & 70.8750I_6 & 27I_6 & 0 & 0 & I_6 \end{bmatrix}.$$

By (9), (10) and (16), we compute the required vectors $\hat{v}_i(3)$ and $\hat{u}_i(3)$ $(i = 1, 2, \ldots, 6)$, and the matrix

$$\Delta_3 = -s_{31}\left(F_3\left[P,\Sigma\right]\right)\hat{U}(3)M\hat{V}(3)^{\dagger}$$
$$= \begin{bmatrix} 0.1471 & 0.4017 & -0.1227 & 0.5755 & -0.8616 & 0.5150 \\ -0.019 & -0.0632 & 0.0204 & -0.0236 & 0.0102 & 0.6517 \\ -0.0956 & -0.2241 & -0.1835 & 0.4688 & 1.4087 & 16.8831 \\ -1.2610 & -3.3113 & 1.1005 & -2.7597 & 6.0712 & -3.1099 \\ 0.9518 & 4.1610 & -0.5813 & -0.2851 & -10.2916 & -1.2288 \\ 0.1051 & 0.3406 & -0.2436 & -0.1013 & -0.4117 & -17.7001 \end{bmatrix},$$

where

$$M = \operatorname{diag}\left\{M^{[1]}, M^{[2]}, M^{[3]}\right\}$$
$$= \operatorname{diag}\left\{\begin{bmatrix} 1.3720 & -1.7639 \\ 0 & 1.3720 \end{bmatrix}, \begin{bmatrix} 0.9386 \end{bmatrix}, \begin{bmatrix} 0.6894 & -0.5835 & 0.3903 \\ 0 & 0.6894 & -0.5835 \\ 0 & 0 & 0.6894 \end{bmatrix}\right\}.$$

By (12), we obtain the perturbation

$$\Delta_3(\lambda) = \begin{bmatrix} 0.0036 & 0.0097 & -0.0030 & 0.0139 & -0.0209 & 0.0125 \\ -0.0005 & -0.0015 & 0.0005 & -0.0006 & 0.0002 & 0.0158 \\ -0.0023 & -0.0054 & -0.0044 & 0.0114 & 0.0341 & 0.4088 \\ -0.0305 & -0.0802 & 0.0266 & -0.0668 & 0.1470 & -0.0753 \\ 0.0230 & 0.1008 & -0.0141 & -0.0069 & -0.2492 & -0.0298 \\ 0.0025 & 0.0082 & -0.0059 & -0.0025 & -0.0100 & -0.4286 \end{bmatrix}\lambda^2$$
$$+ \begin{bmatrix} 0.0388 & 0.1060 & -0.0324 & 0.1519 & -0.2274 & 0.1359 \\ -0.0050 & -0.0167 & 0.0054 & -0.0062 & 0.0027 & 0.1720 \\ -0.0252 & -0.0592 & -0.0484 & 0.1237 & 0.3718 & 4.4561 \\ -0.3328 & -0.8740 & 0.2904 & -0.7284 & 1.6024 & -0.8208 \\ 0.2512 & 1.0982 & -0.1534 & -0.0753 & -2.7163 & -0.3243 \\ 0.0277 & 0.0899 & -0.0643 & -0.0267 & -0.1087 & -4.6717 \end{bmatrix}\lambda$$
$$+ \begin{bmatrix} 0.0648 & 0.1770 & -0.0541 & 0.2536 & -0.3797 & 0.2270 \\ -0.0084 & -0.0279 & 0.0090 & -0.0104 & 0.0045 & 0.2872 \\ -0.0422 & -0.0988 & -0.0809 & 0.2066 & 0.6209 & 7.4408 \\ -0.5558 & -1.4594 & 0.4850 & -1.2162 & 2.6757 & -1.3706 \\ 0.4195 & 1.8338 & -0.2562 & -0.1257 & -4.5357 & -0.5416 \\ 0.0463 & 0.1501 & -0.1074 & -0.0446 & -0.1815 & -7.8008 \end{bmatrix}.$$

The perturbed matrix polynomial $Q_3(\lambda) = P(\lambda) + \Delta_3(\lambda)$ lies on the boundary of the set $\mathcal{B}(P, \beta_{up}(P,\Sigma,3), w) = \mathcal{B}(P, 0.5991, w)$, and its spectrum

$$\sigma(Q_3) = \{0.5719, 0.9726, 1, 1, 2, 3, 3, 3, 0.1786 \pm 1.2680\,\mathrm{i}, -4.0589 \pm 0.7284\,\mathrm{i}\}$$

contains Σ. Moreover, for $\gamma = 1$, it is straightforward to compute $\beta_{low}(P, \Sigma, 1) = 0.0034$. As a consequence,

$$\beta_{low}(P, \Sigma, 1) \;=\; 0.0034 \;\leq\; D_w\,(P, \Sigma) \;\leq\; 0.5991 \;=\; \beta_{up}(P, \Sigma, 3).$$

Finally, we remark that for $\gamma = 1.7$, the vectors $v_1(1.7), v_2(1.7), \ldots, v_6(1.7)$ are close to be linearly dependent with $s_6(\hat{V}(1.7)) = 0.0006$, and for $\gamma = 4.4$, the vectors $v_1(4.4), v_2(4.4), \ldots, v_6(4.4)$ are linearly dependent with $s_6(\hat{V}(4.4)) = 0$. This explains the transient behaviour of the graph in Figure 1, around these two values of γ.

References

[1] J.W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

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

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

[4] J.M. Gracia, *Nearest matrix with two prescribed eigenvalues*, *Linear Algebra Appl.*, 401 (2005), 277–294.

[5] T. Kaczorek, *Polynomial and Rational Matrices: Applications in Dynamical Systems Theory*, Springer-Verlag, London, 2007.

[6] 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.

[7] E. Kokabifar, G.B. Loghmani, A.M. Nazari and S.M. Karbassi, On the distance from a matrix polynomial to matrix polynomials with two prescribed eigenvalues, *Wavelets and Linear Algebra*, 2 (2015), 25–38.

[8] E. Kokabifar, G.B. Loghmani, and S.M. Karbassi, Nearest matrix with prescribed eigenvalues and its applications, *J. Comput. Appl. Math.*, 298 (2016), 53-–63.

[9] E. Kokabifar, G.B. Loghmani, P.J. Psarrakos and S.M. Karbassi, On the distance from a matrix polynomial to matrix polynomials with k prescribed distinct eigenvalues, *Linear Multilinear Algebra*, 65 (2017), 658—676.

[10] P. Lancaster, *Lambda-Matrices and Vibrating Systems*, Dover Publications, 2002.

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

[12] R.A. Lippert, Fixing multiple eigenvalues by a minimal perturbation, *Linear Algebra Appl.*, 432 (2010), 1785–1817.

[13] R.A. Lippert, Fixing two eigenvalues by a minimal perturbation, *Linear Algebra Appl.*, 406 (2005) 177–200.

[14] 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.

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

[16] M.F. Neuts, *Matrix-Geometric Solutions in Stochastic Models*, Springer, Berlin, 1984.

[17] 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.

[18] P.J. Psarrakos, Distance bounds for prescribed multiple eigenvalues of matrix polynomials, *Linear Algebra Appl.*, 436 (2012), 4107–4119.

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

[20] F. Tisseur and K Meerbergen, The quadratic eigenvalue problem, *SIAM Rev.*, 43 (2001), 235–286.

[21] A.I.G. Vardulakis, *Linear Multivariable Control: Algebraic Analysis and Synthesis Methods*, Wiley, New York, 1991.

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

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

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

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

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

Appendix A On the sensitivity of Jordan structure

Let $A = R_A J_A R_A^{-1}$ be the Jordan canonical form of a matrix $A \in \mathbb{C}^{n\times n}$, where J_A is a Jordan matrix and R_A is a nonsingular matrix with columns the Jordan chains (eigenvectors and generalized eigenvectors) of A. Suppose also that the first h Jordan blocks of J_A are of the form

$$J(\lambda_0, s_i) = \begin{bmatrix} \lambda_0 & 1 & \cdots & 0 \\ 0 & \lambda_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_0 \end{bmatrix} \in \mathbb{C}^{s_i\times s_i}, \quad s_1 \geq s_2 \geq \cdots \geq s_h,$$

and correspond to an eigenvalue $\lambda_0 \in \sigma(A)$.

For a real $\varepsilon \in (0,1)$, we consider the $n \times n$ matrices

$$L_\varepsilon = I_{s_1} \oplus \varepsilon^{-s_1} I_{s_2} \oplus \varepsilon^{-(s_1+s_2)} I_{s_2} \oplus \cdots \oplus \varepsilon^{-(s_1+s_2+\cdots+s_h)} I_{s_h}$$

and

$$T_\varepsilon = [\tau_{i,j}] \quad \text{with} \quad \tau_{i,j} = \begin{cases} \varepsilon^{s_l} & \text{if } (i,j) = (s_1+s_2+\cdots+s_l, s_1+s_2+\cdots+s_l+1), \\ 0 & \text{otherwise,} \end{cases}$$

$l = 1, 2, \ldots, h-1$. Then it is straightforward to verify that $L_\varepsilon^{-1} J_A L_\varepsilon = J_A$, and that $L_\varepsilon^{-1} T_\varepsilon L_\varepsilon$ is the $n \times n$ matrix with ones at positions (s_1, s_1+1), $(s_1+s_2, s_1+s_2+1), \ldots,$ $(s_1+s_2+\cdots+s_{h-1}, s_1+s_2+\cdots+s_{h-1}+1)$ and zeros elsewhere. As a consequence, $L_\varepsilon^{-1}(J_A+T_\varepsilon)L_\varepsilon$ is the Jordan matrix that follows from J_A by replacing the $(s_1+s_2+\cdots+s_h) \times (s_1+s_2+\cdots+s_h)$ principal submatrix $J(\lambda_0, s_1) \oplus J(\lambda_0, s_2) \oplus \cdots \oplus J(\lambda_0, s_k)$ by the Jordan block $J(\lambda_0, s_1+s_2+\cdots+s_h) \in \mathbb{C}^{(s_1+s_2+\cdots+s_h)\times(s_1+s_2+\cdots+s_h)}$.

Defining $J_{A_\varepsilon} = L_\varepsilon^{-1}(J_A + T_\varepsilon)L_\varepsilon$ and $A_\varepsilon = (R_A L_\varepsilon)\left[L_\varepsilon^{-1}(J_A+T_\varepsilon)L_\varepsilon\right](R_A L_\varepsilon)^{-1}$, we observe that $A = A_\varepsilon - R_A T_\varepsilon R_A^{-1}$ with $\|A - A_\varepsilon\|_2 = \|R_A T_\varepsilon R_A^{-1}\|_2 \le \varepsilon^{s_k} \|R_A\|_2 \|R_A^{-1}\|_2$. Moreover, the matrices A and A_ε have the same characteristic polynomial; in other words, they have exactly the same eigenvalues with the same algebraic multiplicities.

Definition A.1. Let $A, B \in \mathbb{C}^{n\times n}$ be two matrices with the same characteristic polynomial. We say that the Jordan structure of A *assemblingly majorizes* the Jordan structure of B if there exist Jordan canonical forms of A and B, with the Jordan blocks of each eigenvalue of A not necessarily in an nonincreasing order of sizes,

$$A = R_A J_A R_A^{-1} \quad \text{and} \quad B = R_B J_B R_B^{-1},$$

such that for each eigenvalue $\lambda_0 \in \sigma(A) = \sigma(B)$, the associated Jordan blocks of A, $J_A(\lambda_0, s_i)$ $(i = 1, 2, \ldots, p)$, and of B, $J_B(\lambda_0, r_i)$ $(i = 1, 2, \ldots, q)$, satisfy $p \le q$ and

$$s_1 = r_1 + \cdots + r_{\xi_1}, \;\; s_2 = r_{\xi_1+1} + \cdots + r_{\xi_2}, \;\; \ldots, \;\; s_p = r_{\xi_{p-1}} + \cdots + r_q$$

for some $1 \le \xi_1 < \xi_2 < \cdots < \xi_{p-1} < q$.

For example, consider the matrices

$$A = \begin{bmatrix} \lambda_0 & 1 & 0 & 0 \\ 0 & \lambda_0 & 1 & 0 \\ 0 & 0 & \lambda_0 & 0 \\ 0 & 0 & 0 & \lambda_0 \end{bmatrix}, \quad B = \begin{bmatrix} \lambda_0 & 1 & 0 & 0 \\ 0 & \lambda_0 & 0 & 0 \\ 0 & 0 & \lambda_0 & 0 \\ 0 & 0 & 0 & \lambda_0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} \lambda_0 & 1 & 0 & 0 \\ 0 & \lambda_0 & 0 & 0 \\ 0 & 0 & \lambda_0 & 1 \\ 0 & 0 & 0 & \lambda_0 \end{bmatrix}.$$

Then the Jordan structure of A assemblingly majorizes the Jordan structure of B but not the Jordan structure of C.

By Definition A.1 and the above discussion, the next results follow readily.

Proposition A.2. *For any matrix $A \in \mathbb{C}^{n\times n}$ and any $\delta \in (0,1)$, there is a matrix $\hat{A} \in \mathbb{C}^{n\times n}$ such that the Jordan structure of $\hat{A}$ assemblingly majorizes the Jordan structure of A, and $\|A-\hat{A}\|_2 \le \delta$.*

Corollary A.3. *For any matrix $A \in \mathbb{C}^{n\times n}$ and any $\delta \in (0,1)$, there is an $\hat{A} \in \mathbb{C}^{n\times n}$ such that $\hat{A}$ and A have the same characteristic polynomial, all the eigenvalues of $\hat{A}$ have geometric multiplicity* 1 *(i.e., each eigenvalue of $\hat{A}$ corresponds to exactly one Jordan block of $J_{\hat{A}}$ and has a Jordan chain of length equal to the algebraic multiplicity of the eigenvalue), and $\|A-\hat{A}\|_2 \le \delta$.*

Consider now an $n\times n$ matrix polynomial $P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0$ as in (1), with nonsingular leading coefficient A_m. Any eigenvalue of $P(\lambda)$ of geometric multiplicity g has g maximal Jordan chains associated to g (nonzero) 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 an eigenvalue $\lambda_0 \in \sigma(P)$ is known as the *index of annihilation* of λ_0 [6]. An $n \times nm$ matrix X_P with columns maximal Jordan chains of $P(\lambda)$ and an $nm \times nm$ Jordan matrix J_P form a *Jordan pair* (X_P, J_P) of $P(\lambda)$ [3] if the matrix

$$S_P = \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix} \in \mathbb{C}^{nm\times nm}$$

is nonsingular and

$$A_m X_P J_P^m + A_{m-1} X_P J_P^{m-1} + \cdots + A_1 X_P J_P + A_0 X_P = 0.$$

The index of annihilation of an eigenvalue $\lambda_0 \in \sigma(P)$ coincides with the size of the largest Jordan blocks of J_P corresponding to λ_0; for details on the Jordan structure of matrix polynomials, see [3, 11].

By Theorem 2.4 in [3], we have the following proposition.

Proposition A.4. *Let (X_P, J_P) be a Jordan pair of the matrix polynomial $P(\lambda)$ in* (1)*, and let $S_P^{-1} = [\, V_1 \ V_2 \ \cdots \ V_m \,]$, $V_1, V_2, \ldots, V_m \in \mathbb{C}^{n\times n}$. Then $P(\lambda)$ admits the representation*

$$\begin{aligned} P(\lambda) &= A_m\lambda^m - A_m X_P J_P^m S_P^{-1} \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix} \\ &= A_m\lambda^m - A_m X_P J_P^m \left(V_1 + V_2\lambda + \cdots + V_m\lambda^{m-1} \right). \end{aligned}$$

By the previous discussion on matrices and Proposition A.2, there are appropriate matrices $T_\varepsilon, L_\varepsilon \in \mathbb{C}^{nm\times nm}$ such that the $nm \times nm$ Jordan matrix $\hat{J}_P = L_\varepsilon^{-1}(J_P + T_\varepsilon)L_\varepsilon$ assemblingly majorizes J_P and the distance $\|J_P - \hat{J}_P\|_2$ is arbitrarily small. Furthermore, the representation of $P(\lambda)$ in Proposition A.4 yields

$$\begin{aligned} P(\lambda) &= A_m\lambda^m - A_m X_P (L_\varepsilon \hat{J}_P L_\varepsilon^{-1} - T_\varepsilon)^m \begin{bmatrix} X_P \\ X_P(L_\varepsilon \hat{J}_P L_\varepsilon^{-1} - T_\varepsilon) \\ \vdots \\ X_P(L_\varepsilon \hat{J}_P L_\varepsilon^{-1} - T_\varepsilon)^{m-1} \end{bmatrix}^{-1} \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix} \\ &= A_m\lambda^m - A_m X_P L_\varepsilon \hat{J}_P^m \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix}^{-1} \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix} - \hat{E}(\lambda) \end{aligned}$$

for some matrix polynomial $\hat{E}(\lambda) = \hat{E}_{m-1}\lambda^{m-1} + \hat{E}_{m-2}\lambda^{m-2} + \cdots + \hat{E}_1\lambda + \hat{E}_0$. Without loss of generality, we assume that T_ε and L_ε are chosen such as the matrix

$$\hat{S}_P = \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix} \in \mathbb{C}^{nm\times nm}$$

is nonsingular; by continuity, this is true for sufficiently small $\varepsilon > 0$.

By Theorem 7.8 in [3], $(X_P L_\varepsilon, \hat{J}_P)$ is a Jordan pair of the perturbed matrix polynomial

$$\hat{Q}(\lambda) = P(\lambda) + \hat{E}(\lambda) = A_m\lambda^m - A_m X_P L_\varepsilon \hat{J}_P^m \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix}^{-1} \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix}.$$

The difference $\hat{E}(\lambda) = P(\lambda) - \hat{Q}(\lambda) = \hat{E}_{m-1}\lambda^{m-1} + \cdots + \hat{E}_1\lambda + \hat{E}_0$ is written

$$\hat{E}(\lambda) = A_m X_P \left(L_\varepsilon \hat{J}_P^m \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix}^{-1} - J_P^m \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix}^{-1} \right) \begin{bmatrix} I_n \\ \lambda I_n \\ \vdots \\ \lambda^{m-1} I_n \end{bmatrix},$$

which means that for every $j = 0, 1, \ldots, m-1$,

$$\hat{E}_j = A_m X_P \left(L_\varepsilon \hat{J}_P^m \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix}^{-1} - J_P^m \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix}^{-1} \right) \begin{bmatrix} 0 \\ \vdots \\ I_n \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow j\text{-th position} \\ \\ \\ \end{matrix} .$$

As a consequence,

$$\begin{aligned} \|\hat{E}_j\|_2 &\leq \|A_m\|_2 \|X_P\|_2 \left\| L_\varepsilon \hat{J}_P^m \begin{bmatrix} X_P L_\varepsilon \\ X_P L_\varepsilon \hat{J}_P \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} \end{bmatrix}^{-1} - J_P^m \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix}^{-1} \right\|_2 \\ &= \|A_m\|_2 \|X_P\|_2 \left\| L_\varepsilon \hat{J}_P^m L_\varepsilon^{-1} \begin{bmatrix} X_P \\ X_P L_\varepsilon \hat{J}_P L_\varepsilon^{-1} \\ \vdots \\ X_P L_\varepsilon \hat{J}_P^{m-1} L_\varepsilon^{-1} \end{bmatrix}^{-1} - J_P^m \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix}^{-1} \right\|_2 \\ &= \|A_m\|_2 \|X_P\|_2 \left\| (J_P + T_\varepsilon)^m \begin{bmatrix} X_P \\ X_P (J_P + T_\varepsilon) \\ \vdots \\ X_P (J_P + T_\varepsilon)^{m-1} \end{bmatrix}^{-1} - J_P^m \begin{bmatrix} X_P \\ X_P J_P \\ \vdots \\ X_P J_P^{m-1} \end{bmatrix}^{-1} \right\|_2 . \end{aligned} \quad (26)$$

For sufficient small $\varepsilon > 0$, the upper bound (26) can be arbitrarily small, and thus, Proposition A.2 and Corollary A.3 are generalized to the case of matrix polynomials.

Proposition A.5. *Let $P(\lambda)$ be an $n \times n$ matrix polynomial as in* (1)*, and let the weights $w_0, w_1, \ldots, w_{m-1}$ be positive. Then, for any $\delta \in (0, 1)$, there is an $n \times n$ matrix polynomial $\hat{P}(\lambda) \in \mathcal{B}(P, \delta, w)$, with leading coefficient A_m, such that the Jordan matrix $J_{\hat{P}}$ assemblingly majorizes the Jordan matrix J_P.*

Corollary A.6. *Let $P(\lambda)$ be an $n \times n$ matrix polynomial as in* (1)*, and let the weights $w_0, w_1, \ldots, w_{m-1}$ be positive. Then, for any $\delta \in (0, 1)$, there is an $n \times n$ matrix polynomial $\hat{P}(\lambda) \in \mathcal{B}(P, \delta, w)$, with leading coefficient A_m, such that* $\det P(\lambda) = \det \hat{P}(\lambda)$ *and all the eigenvalues of $\hat{P}(\lambda)$ have geometric multiplicity* 1 *(i.e., every eigenvalue of $\hat{P}(\lambda)$ corresponds to exactly one Jordan block of $J_{\hat{P}}$ and has a Jordan chain of length equal to the algebraic multiplicity of the eigenvalue).*

Example A.7. (See Example 14.4 of [11]) The matrix polynomial

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

has exactly one eigenvalue, $\lambda_0 = 0$, and a Jordan pair of the form

$$(X_P, J_P) = \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \end{bmatrix} \right).$$

Moreover,

$$S_P = \begin{bmatrix} X_P \\ X_P J_P \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad S_P^{-1} = \begin{bmatrix} X_P \\ X_P J_P \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix}.$$

For $\varepsilon = 10^{-3}$, we consider the set of positive weights $w = \{w_0, w_1, w_2\} = \{1, 1, 1\}$ and the matrices

$$L_\varepsilon = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 10^9 \end{bmatrix} \quad \text{and} \quad T_\varepsilon = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10^{-9} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and we define

$$\hat{J}_P = L_\varepsilon^{-1}(J_P + T_\varepsilon) L_\varepsilon = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is straightforward to verify that the pair

$$(X_P L_\varepsilon, \hat{J}_P) = \left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 10^9 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right)$$

is a Jordan pair of the perturbed matrix polynomial

$$\hat{P}(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & -1 \\ 10^{-9} & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 \\ 0 & -10^{-9} \end{bmatrix} \in \mathcal{B}(P, 10^{-9}, w).$$