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

E. Kokabifar*, G.B. Loghmani*, P.J. Psarrakos† and S.M. Karbassi‡

July 26, 2016

Abstract

Consider an $n \times n$ matrix polynomial $P(\lambda)$ and a set Σ consisting of $k \leq n$ distinct complex numbers. In this paper, a (weighted) spectral norm distance from $P(\lambda)$ to the matrix polynomials whose spectra include the specified set Σ, is defined and studied. An upper and a lower bound for this distance are obtained, and an optimal perturbation of $P(\lambda)$ associated to the upper bound is constructed. Numerical examples are given to illustrate the efficiency of the proposed bounds.

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

AMS Classification: 15A18, 65F35.

1 Introduction

Let A be an $n \times n$ complex matrix, and let $\mathcal{M}$ be the set of all $n \times n$ complex matrices that have $\mu \in \mathbb{C}$ as a multiple eigenvalue. Malyshev [14] obtained the following singular value optimization characterization for the spectral norm distance from A to $\mathcal{M}$:

$$\min_{B \in \mathcal{M}} \|A - B\|_2 = \max_{\gamma \geq 0} s_{2n-1} \left(\begin{bmatrix} A - \mu I & \gamma I_n \\ 0 & A - \mu I \end{bmatrix} \right),$$

where $\| \cdot \|_2$ denotes the spectral matrix norm subordinate to the Euclidean vector norm, and s_i is the i-th singular value of the corresponding matrix sorted in a nonincreasing order. Malyshev's work can be considered as a solution to Wilkinson's problem, that is, the computation of the distance from a matrix $A \in \mathbb{C}^{n \times n}$ with (only) simple eigenvalues to the set of $n \times n$ matrices with multiple eigenvalues. This distance was introduced by Wilkinson in [25], and some bounds for it were computed by Ruhe [19], Wilkinson [21–24] and

*Department of Mathematics, Faculty of Science, Yazd University, Yazd, Iran (e.kokabifar@stu.yazd.ac.ir, loghmani@yazd.ac.ir).

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

‡Department of Mathematics, Yazd Branch, Islamic Azad University, Yazd, Iran (mehdikarbassi@gmail.com).

Demmel [2]. Spectral norm distances from A to matrices that have a prescribed eigenvalue of algebraic multiplicity 3, or any prescribed algebraic multiplicity, were considered by Ikramov and Nazri [6] and Mengi [16], respectively. Gracia [5] and Lippert [13] studied a spectral norm distance from A to matrices with two prescribed eigenvalues, and obtained a matrix closest to A having these two eigenvalues. Moreover, Kokabifar, Loghmani and Karbassi [9] and Lippert [12] investigated (computationally and geometrically) a spectral norm distance from A to matrices having k $(\leq n)$ prescribed eigenvalues.

In 2008, Papathanasiou and Psarrakos [17] generalized Malyshev's results for the case of matrix polynomials, in terms of a (weighted) spectral norm distance from an $n \times n$ matrix polynomial $P(\lambda)$ to the matrix polynomials that have a prescribed $\mu \in \mathbb{C}$ as a multiple eigenvalue, obtaining an upper and a lower bound for this distance. Lately, motivated by Mengi's results in [16], Psarrakos [18] introduced the matrix polynomials

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

where $P^{(i)}(\lambda)$ denotes the i-th derivative of $P(\lambda)$ with respect to λ. Then, he derived lower and upper bounds for a distance from $P(\lambda)$ to the matrix polynomials with a prescribed eigenvalue of a desired algebraic multiplicity by generalizing the methodology used in [17]. Recently, Kokabifar, Loghmani, Nazari and Karbassi [10] extended the results of [17] to the case of two distinct eigenvalues, by replacing the first order derivative of $P(\lambda)$ in $F_2\left[P(\lambda);\gamma\right]$ by a divided difference. Also, Karow and Mengi [8] studied systematically an alternative distance from a given $n \times n$ matrix polynomial to matrix polynomials with a specified number of eigenvalues at specified locations in the complex plane, deriving singular value optimization characterizations based on a Sylvester equation characterization.

In this paper, motivated by the literature above, we introduce and study a (weighted) spectral norm distance from an $n \times n$ matrix polynomial $P(\lambda)$ to the set of all matrix polynomials with $k \leq n$ prescribed distinct eigenvalues. In particular, we obtain an upper and a lower bound for this distance, and construct an optimal perturbation associated to the upper bound. Replacing the derivatives of $P(\lambda)$ in $F_k\left[P(\lambda);\gamma\right]$ by divided differences formulas, extending necessary definitions and lemmas of [10, 12, 17, 18], and constructing an appropriate perturbation of $P(\lambda)$ are the main ideas used herein. (Hence, this article can be considered as a generalization of the results obtained in [12] to the case of matrix polynomials, and also as an extension of [10, 17, 18] to the case of k arbitrary distinct eigenvalues). In the next section, we review standard definitions concerning matrix polynomials, and we also introduce some definitions which are necessary for the remainder. In Section 3, we construct an admissible perturbation of $P(\lambda)$ by extending the methods described in [10,17,18]. In Section 4, we obtain our bounds, and in Section 5, we give two numerical examples to illustrate the effectiveness of the proposed technique.

2 Preliminaries

In the last decades, the study of matrix polynomials, especially with regard to their spectral analysis, has received much attention of several researchers and has met many applications. Some basic references for the theory and applications of matrix polynomials are [4,7,11,15,20] and references therein.

For given $A_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ and a complex variable λ, we define the *matrix polynomial*

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

If, for a scalar $\mu \in \mathbb{C}$ and some nonzero vector $\upsilon \in \mathbb{C}^n$, it holds that $P(\mu)\upsilon = 0$, then the scalar μ is called an *eigenvalue* of $P(\lambda)$ and the vector υ is known as a *(right) eigenvector* of $P(\lambda)$ corresponding to μ. Similarly, a nonzero vector $\nu \in \mathbb{C}^n$ is known as a *(left) eigenvector* of $P(\lambda)$ corresponding to μ when $\nu^* P(\mu) = 0$. The *spectrum* of $P(\lambda)$, denoted by $\sigma(P)$, is the set of its eigenvalues. Throughout this paper, it is assumed that the coefficient matrix A_m is *nonsingular*, which implies that the spectrum of $P(\lambda)$ contains no more than mn distinct elements.

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*. The singular values of $P(\lambda)$ are the nonnegative roots of the eigenvalue functions of $P(\lambda)^* P(\lambda)$, and we denote them by $s_1(P(\lambda)) \geq s_2(P(\lambda)) \geq \cdots \geq s_n(P(\lambda))$.

Definition 2.1. Let $P(\lambda)$ be a matrix polynomial as in (1) and let $\Delta_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ be arbitrary matrices. Consider (additive) perturbations of the matrix polynomial $P(\lambda)$ of the form

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

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

$$\mathcal{B}(P, \varepsilon, w) = \left\{Q(\lambda) \text{ as in (2)} : \|\Delta_j\|_2 \leq \varepsilon w_j,\ j = 0, 1, \ldots, m\right\},$$

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. Let $P(\lambda)$ be a matrix polynomial as in (1), and let a set of distinct complex numbers $\Sigma = \{\mu_1, \mu_2, \ldots, \mu_k\}$ $(k \leq n)$ be given. 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. Consider a complex function f and k distinct scalars $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{C}$. The *divided difference at nodes* $\mu_i, \mu_{i+1}, \ldots, \mu_{i+t}$ $(1 \leq i \leq k-1,\ 1 \leq t \leq k-i)$ is denoted by $f[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t}]$ and defined by the following recursive formula [3]:

$$f[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t}] = \frac{f[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t-1}] - f[\mu_{i+1}, \mu_{i+2}, \ldots, \mu_{i+t}]}{\mu_i - \mu_{i+t}},$$

where $f[\mu_i] = f(\mu_i)$ $(i = 1, 2, \ldots, k)$.

Definition 2.4. Suppose that $P(\lambda)$ is a matrix polynomial as in (1) and a set of distinct complex numbers $\Sigma = \{\mu_1, \mu_2, \ldots, \mu_k\}$ $(k \leq n)$ is given. For any scalar $\gamma \in \mathbb{C}$, define the $nk \times nk$ matrix

$$F_\gamma[P, \Sigma] = \begin{bmatrix} P(\mu_1) & 0 & \cdots & 0 \\ \gamma P[\mu_1, \mu_2] & P(\mu_2) & \cdots & 0 \\ \gamma^2 P[\mu_1, \mu_2, \mu_3] & \gamma P[\mu_2, \mu_3] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{k-1} P[\mu_1, \ldots, \mu_k] & \gamma^{k-2} P[\mu_2, \ldots, \mu_k] & \cdots & P(\mu_k) \end{bmatrix}.$$

3 Construction of a perturbation

In this section, we construct an $n \times n$ matrix polynomial $\Delta_\gamma(\lambda)$ such that the given set of distinct scalars $\Sigma = \{\mu_1, \mu_2, \ldots, \mu_k\}$ $(k \leq n)$ is included in the spectrum of the perturbed matrix polynomial $Q_\gamma(\lambda) = P(\lambda) + \Delta_\gamma(\lambda)$. Without loss of generality, hereafter we can assume that the parameter γ is real nonnegative [18]. Moreover, for convenience, we set $\rho = nk - k + 1$.

Suppose now that $\gamma > 0$, and consider the nonzero quantities

$$\theta_{i,j} = \frac{\gamma}{\mu_i - \mu_j} \,(= -\theta_{j,i}), \quad i, j \in \{1, 2, \ldots, k\},\ i \neq j. \tag{3}$$

Definition 3.1. Let

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

be a pair of consistent left and right singular vectors of $s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)$, respectively. Define the vectors

$$\hat{v}_1(\gamma)=v_1(\gamma),\quad \hat{v}_p(\gamma)=v_p(\gamma)+\sum_{i=1}^{p-1}\left[(-1)^i\left(\prod_{j=p-i}^{p-1}\theta_{j,p}\right)v_{p-i}(\gamma)\right]\quad (p=2,3,\ldots,k) \tag{4}$$

and

$$\hat{u}_1(\gamma)=u_1(\gamma),\quad \hat{u}_p(\gamma)=u_p(\gamma)+\sum_{i=1}^{p-1}\left[(-1)^i\left(\prod_{j=p-i}^{p-1}\theta_{j,p}\right)u_{p-i}(\gamma)\right]\quad (p=2,3,\ldots,k). \tag{5}$$

Define also the $n\times k$ matrices

$$\begin{aligned}&\hat{U}(\gamma)=\left[\,\hat{u}_1(\gamma)\;\hat{u}_2(\gamma)\;\cdots\;\hat{u}_k(\gamma)\,\right],\quad \hat{V}(\gamma)=\left[\,\hat{v}_1(\gamma)\;\hat{v}_2(\gamma)\;\cdots\;\hat{v}_k(\gamma)\,\right] &(6)\\ &\text{and}\quad V(\gamma)=\left[\,v_1(\gamma)\;v_2(\gamma)\;\cdots\;v_k(\gamma)\,\right].\end{aligned}$$

In the remainder of the paper, we assume that $\operatorname{rank}(V(\gamma))=k$. It is also necessary to observe that by the definition of the quantities $\theta_{i,j}$ $(i,j\in\{1,2,\ldots,k\},\ i\neq j)$ in (3), for all distinct i, j and q in $\{1,2,\ldots,k\}$, it follows

$$\theta_{i,j}(\theta_{q,i}-\theta_{q,j})=\theta_{q,i}\,\theta_{q,j},\quad \theta_{i,j}(\theta_{i,q}-\theta_{j,q})=\theta_{i,q}\,\theta_{q,j}=\theta_{q,i}\,\theta_{j,q} \tag{7}$$

and

$$\theta_{i,j}(\theta_{i,q}+\theta_{q,j})=\theta_{i,q}\,\theta_{q,j},\quad \theta_{i,j}(\theta_{q,i}+\theta_{j,q})=\theta_{i,q}\,\theta_{j,q}=\theta_{q,i}\,\theta_{q,j}. \tag{8}$$

Using (7) and (8), one can verify that for every $p=2,3,\ldots,k$, the vectors $\hat{v}_p(\gamma)$ in (4) and $\hat{u}_p(\gamma)$ in (5) satisfy

$$\begin{aligned}\hat{v}_p(\gamma)&=v_p(\gamma)-\theta_{p-1,p}\,\hat{v}_{p-1}(\gamma)-(\theta_{p-2,p}\,\theta_{p-2,p-1})\,\hat{v}_{p-2}(\gamma)-\cdots-\left(\prod_{j=2}^{p}\theta_{1,j}\right)\hat{v}_1(\gamma)\\&=v_p\left(\gamma\right)-\sum_{i=1}^{p-1}\left(\prod_{j=i+1}^{p}\theta_{i,j}\right)\hat{v}_i\left(\gamma\right)\end{aligned} \tag{9}$$

and

$$\begin{aligned}\hat{u}_p(\gamma)&=u_p(\gamma)-\theta_{p-1,p}\,\hat{u}_{p-1}(\gamma)-(\theta_{p-2,p}\,\theta_{p-2,p-1})\,\hat{u}_{p-2}(\gamma)-\cdots-\left(\prod_{j=2}^{p}\theta_{1,j}\right)\hat{u}_1(\gamma)\\&=u_p(\gamma)-\sum_{i=1}^{p-1}\left(\prod_{j=i+1}^{p}\theta_{i,j}\right)\hat{u}_i(\gamma).\end{aligned} \tag{10}$$

Next we will define the desired perturbation of $P(\lambda)$ in terms of matrices $\hat{U}(\gamma)$ and $\hat{V}(\gamma)$ in (6). We consider the quantities

$$\alpha_{i,s} = \frac{1}{w\left(|\mu_i|\right)} \sum_{j=0}^{m} \left(\left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j \mu_s^j w_j \right) \quad \text{and} \quad \beta_s = \frac{1}{k} \sum_{i=1}^{k} \alpha_{i,s}, \quad i, s = 1, 2, \ldots, k, \tag{11}$$

where $w_0 > 0$ and, by convention, we set $\frac{\bar{\mu}_i}{|\mu_i|} = 0$ and $\alpha_{i,s} = 1$ whenever $\mu_i = 0$. If $\beta_1, \beta_2, \ldots, \beta_k$ are nonzero, then we define the $n \times n$ matrix

$$\Delta_\gamma = -s_\rho(F_\gamma[P, \Sigma])\hat{U}(\gamma)\,\mathrm{diag}\left\{ \frac{1}{\beta_1}, \frac{1}{\beta_2}, \ldots, \frac{1}{\beta_k} \right\} \hat{V}(\gamma)^\dagger, \tag{12}$$

where $\hat{V}(\gamma)^\dagger$ denotes the *Moore-Penrose pseudoinverse* of $\hat{V}(\gamma)$. Furthermore, we define the matrices

$$\Delta_{\gamma,j} = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{1}{w\left(|\mu_i|\right)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j w_j \right) \Delta_\gamma, \quad j = 0, 1, \ldots, m, \tag{13}$$

and the $n \times n$ matrix polynomial

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

By straightforward computations, we see that the matrix polynomial $\Delta_\gamma(\lambda)$ satisfies

$$\Delta_\gamma\left(\mu_s\right) \;=\; \sum_{j=0}^{m} \left[\frac{1}{k} \sum_{i=1}^{k} \left(\frac{1}{w\left(|\mu_i|\right)} \left(\frac{\bar{\mu}_i}{|\mu_i|} \right)^j \right) w_j \mu_s^j \right] \Delta_\gamma \;=\; \beta_s \Delta_\gamma, \quad s = 1, 2, \ldots, k. \tag{14}$$

We also remark that the condition $\mathrm{rank}(V(\gamma)) = k$ implies $\mathrm{rank}(\hat{V}(\gamma)) = k$. As a consequence, $\hat{V}(\gamma)^\dagger \hat{V}(\gamma) = I_k$, where I_k denotes the $k \times k$ identity matrix.

Since $u(\gamma), v(\gamma)$ is a pair of left and right singular vectors of $s_\rho\left(F_\gamma[P, \Sigma]\right)$, by definition,

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

or equivalently,

$$\begin{array}{rcl}
s_\rho\left(F_\gamma[P, \Sigma]\right) u_1(\gamma) &=& P(\mu_1) v_1(\gamma), \\
s_\rho\left(F_\gamma[P, \Sigma]\right) u_2(\gamma) &=& \gamma P[\mu_1, \mu_2] v_1(\gamma) + P(\mu_2) v_2(\gamma), \\
s_\rho\left(F_\gamma[P, \Sigma]\right) u_3(\gamma) &=& \gamma^2 P[\mu_1, \mu_2, \mu_3] v_1(\gamma) + \gamma P[\mu_2, \mu_3] v_2(\gamma) + P(\mu_3) v_3(\gamma), \\
\vdots & & \qquad \vdots \qquad\qquad\qquad \vdots \\
s_\rho\left(F_\gamma[P, \Sigma]\right) u_k(\gamma) &=& \gamma^{k-1} P[\mu_1, \ldots, \mu_k] v_1(\gamma) + \gamma^{k-2} P[\mu_2, \ldots, \mu_k] v_2(\gamma) + \cdots + P(\mu_k) v_k(\gamma).
\end{array} \tag{15}$$

Recall that the vectors $\hat{v}_1(\gamma), \hat{v}_2(\gamma), \ldots, \hat{v}_k(\gamma)$ are defined as in (4) and (9) and the vectors $\hat{u}_1(\gamma), \hat{u}_2(\gamma), \ldots, \hat{u}_k(\gamma)$ are defined as in (5) and (10). By expressing the divided differences in system (15) in terms of $P(\mu_1), P(\mu_2), \ldots, P(\mu_k)$, one can verify that on the right-hand side of the p-th equation of system (15), the sum of all vectors multiplied by $P(\mu_p)$ is equal to $\hat{v}_p(\gamma)$. Moreover, moving all the remaining vectors to the left-hand side of the p-th equation yields $s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_p(\gamma)$. In particular, the following hold:

For $p = 1$, $\hat{v}_1(\gamma) = v_1(\gamma)$ and $\hat{u}_1(\gamma) = u_1(\gamma)$. Thus, the first equation of system (15) implies

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_1(\gamma) = P(\mu_1)\hat{v}_1(\gamma).$$

For $p = 2$, the second equation of system (15) yields

$$\begin{aligned} s_\rho\left(F_\gamma[P,\Sigma]\right)u_2(\gamma) &= \gamma P[\mu_1,\mu_2]v_1(\gamma) + P(\mu_2)v_2(\gamma) \\ &= \theta_{1,2}\left(P(\mu_1) - P\left(\mu_2\right)\right)v_1(\gamma) + P(\mu_2)v_2(\gamma) \\ &= P(\mu_2)\left(v_2(\gamma) - \theta_{1,2}v_1(\gamma)\right) + \theta_{1,2}P(\mu_1)v_1(\gamma), \end{aligned}$$

or equivalently,

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\left(u_2(\gamma) - \theta_{1,2}u_1(\gamma)\right) = P(\mu_2)\left(v_2(\gamma) - \theta_{1,2}v_1(\gamma)\right).$$

By (4) and (5) (also, by (9) and (10)), we have $\hat{v}_2(\gamma) = v_2(\gamma) - \theta_{1,2}v_1(\gamma)$ and $\hat{u}_2(\gamma) = u_2(\gamma) - \theta_{1,2}u_1(\gamma)$, respectively. As a consequence,

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_2(\gamma) = P(\mu_2)\hat{v}_2(\gamma).$$

For the sake of induction, assume that for a given $p \in \{2, 3, \ldots, k-1\}$, it holds

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_i(\gamma) = P(\mu_i)\hat{v}_i(\gamma), \quad i = 1, 2, \ldots, p. \tag{16}$$

By expressing the divided differences in the $(p+1)$-th equation of system (15) in terms of $P(\mu_1), P(\mu_2), \ldots, P(\mu_{p+1})$, and applying (7) and (8) (to construct appropriate products of $\theta_{i,j}$'s), and (4) and (5) (to construct $\hat{v}_i(\gamma)$'s), straightforward computations yield

$$\begin{aligned} & s_\rho\left(F_\gamma[P,\Sigma]\right)u_{p+1}(\gamma) \\ = \;& \gamma^p P[\mu_1,\ldots,\mu_{p+1}]v_1(\gamma) + \gamma^{p-1}P[\mu_2,\ldots,\mu_{p+1}]v_2(\gamma) \\ & + \cdots + \gamma^2 P[\mu_{p-1},\mu_p,\mu_{p+1}]v_{p-1}(\gamma) + \gamma P[\mu_p,\mu_{p+1}]v_p(\gamma) + P(\mu_{p+1})v_{p+1}(\gamma) \\ = \;& \gamma^{p-1}\theta_{1,p+1}\left(P[\mu_1,\ldots,\mu_p] - P[\mu_2,\ldots,\mu_{p+1}]\right)v_1(\gamma) + \gamma^{p-2}\theta_{2,p+1}\left(P[\mu_2,\ldots,\mu_p] - P[\mu_3,\ldots,\mu_{p+1}]\right)v_2(\gamma) \\ & + \cdots + \left[(\theta_{p-1,p+1}\theta_{p-1,p})(P(\mu_{p-1}) - P(\mu_p)) - (\theta_{p-1,p+1}\theta_{p,p+1})(P(\mu_p) - P(\mu_{p+1}))\right]v_{p-1}(\gamma) \\ & + \theta_{p,p+1}\left(P(\mu_p) - P(\mu_{p+1})\right)v_p(\gamma) + P(\mu_{p+1})v_{p+1}(\gamma) \\ = \;& P(\mu_{p+1})\left[v_{p+1}(\gamma) - \theta_{p,p+1}\hat{v}_p(\gamma) - (\theta_{p-1,p+1}\,\theta_{p-1,p})\hat{v}_{p-1}(\gamma) - \cdots - \left(\prod_{j=2}^{p+1}\theta_{1,j}\right)\hat{v}_1(\gamma)\right] \\ & + \left[\theta_{p,p+1}P(\mu_p)\hat{v}_p(\gamma) + (\theta_{p-1,p+1}\,\theta_{p-1,p})P(\mu_{p-1})\hat{v}_{p-1}(\gamma) + \cdots + \left(\prod_{j=2}^{p+1}\theta_{1,j}\right)P(\mu_1)\hat{v}_1(\gamma)\right]. \end{aligned}$$

By (16), it follows

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\left[u_{p+1}(\gamma)-\theta_{p,p+1}\hat{u}_p(\gamma)-(\theta_{p-1,p+1}\,\theta_{p-1,p})\hat{u}_{p-1}(\gamma)-\cdots-\left(\prod_{j=2}^{p+1}\theta_{1,j}\right)\hat{u}_1(\gamma)\right]$$
$$=P(\mu_{p+1})\left[v_{p+1}(\gamma)-\theta_{p,p+1}\hat{v}_p(\gamma)-(\theta_{p-1,p+1}\,\theta_{p-1,p})\hat{v}_{p-1}(\gamma)-\cdots-\left(\prod_{j=2}^{p+1}\theta_{1,j}\right)\hat{v}_1(\gamma)\right],$$

and by (9) and (10), it is apparent that

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_{p+1}(\gamma)=P(\mu_{p+1})\hat{v}_{p+1}(\gamma).$$

Hence, it is obtained that

$$s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_i(\gamma)=P\left(\mu_i\right)\hat{v}_i(\gamma),\quad i=1,2,\ldots,k.$$

Recalling the definition of $n\times n$ matrices Δ_γ in (12) and $\Delta_{\gamma,0},\Delta_{\gamma,1},\ldots,\Delta_{\gamma,m}$ in (13), we define the matrix polynomial

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

Then, by (14), we have that for every $i=1,2,\ldots,k$,

$$\begin{aligned}
Q_\gamma\left(\mu_i\right)\hat{v}_i(\gamma) &= P\left(\mu_i\right)\hat{v}_i(\gamma)+\Delta_\gamma\left(\mu_i\right)\hat{v}_i(\gamma)\\
&= s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_i(\gamma)+\beta_i\Delta_\gamma\hat{v}_i(\gamma)\\
&= s_\rho\left(F_\gamma[P,\Sigma]\right)\hat{u}_i(\gamma)+\beta_i\left(-s_\rho\left(F_\gamma[P,\Sigma]\right)\frac{1}{\beta_i}\right)\hat{u}_i(\gamma)\\
&= 0.
\end{aligned}$$

As a consequence, if $\operatorname{rank}(V(\gamma))=k$ (recall that all $\beta_1,\beta_2,\ldots,\beta_k$ in (11) are nonzero), then $\mu_1,\mu_2,\ldots,\mu_k$ are eigenvalues of the matrix polynomial $Q_\gamma(\lambda)$ in (17) with $\hat{v}_1(\gamma),\hat{v}_2(\gamma),\ldots,\hat{v}_k(\gamma)$ as their associated eigenvectors, respectively.

The next result follows immediately.

Theorem 3.2. *Consider an $n\times n$ matrix polynomial $P(\lambda)$ as in (1) and a given set of $k\le n$ distinct complex numbers $\Sigma=\{\mu_1,\mu_2,\ldots,\mu_k\}$, and suppose that the quantities $\beta_1,\beta_2,\ldots,\beta_k$ in (11) are nonzero. For every $\gamma>0$ such that* $\operatorname{rank}(V(\gamma))=k$, *the scalars $\mu_1,\mu_2,\ldots,\mu_k$ are eigenvalues of the matrix polynomial $Q_\gamma(\lambda)$ in (17), with corresponding eigenvectors $\hat{v}_1(\gamma),\hat{v}_2(\gamma),\ldots,\hat{v}_k(\gamma)$, respectively.*

Remark 3.3. For the case $k = 2$, by [10, Section 2] (see also [17, Section 5]), if the matrix $P[\mu_1, \mu_2]$ is nonsingular and $\gamma_* > 0$ is a point where the singular value $s_{2n-1}(F_\gamma[P, \{\mu_1, \mu_2\}])$ attains its maximum value, then $\mathrm{rank}(V(\gamma_*)) = 2\,(= k)$. But for the case $k > 2$, as mentioned in [18], it is not easy to obtain conditions ensuring $\mathrm{rank}(V(\gamma)) = k$. However, in our experiments, the condition $\mathrm{rank}(V(\gamma)) = k$ holds typically.

Remark 3.4. At this time, we have no clear understanding of the case where $\beta_s = 0$ for some $s \in \{1, 2, \ldots, k\}$ (even for $k = 2$). Describing fully the set $\Sigma = \{\mu_1, \mu_2, \ldots, \mu_k\}$ ($k \le n$) of desired eigenvalues and the set of weights $w = \{w_0, w_1, \ldots, w_m\}$ which lead to $\beta_s = 0$ is still an open problem. However, the condition $\beta_1, \beta_2, \ldots, \beta_k \ne 0$ appears to hold generically in our experiments.

We conclude this section by providing a matrix $\Delta_0 \in \mathbb{C}^{n\times n}$ so that the prescribed scalars are eigenvalues of $P(\lambda) + \Delta_0$. For $i = 1, 2, \ldots, k$, let $\tilde{u}_i, \tilde{v}_i \in \mathbb{C}^n$ be a pair of left and right singular vectors of $P(\mu_i)$ corresponding to $\sigma_i = s_n(P(\mu_i))$, respectively. If the vectors $\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_k$ are linearly independent, then we define the constant matrix

$$\Delta_0 = -\left[\, \tilde{u}_1 \ \tilde{u}_2 \ \cdots \ \tilde{u}_k \,\right] \mathrm{diag}\left\{\sigma_1, \sigma_2, \ldots, \sigma_k\right\} \left[\, \tilde{v}_1 \ \tilde{v}_2 \ \cdots \ \tilde{v}_k \,\right]^{\dagger}$$

and observe that $\left[\, \tilde{v}_1 \ \tilde{v}_2 \ \cdots \ \tilde{v}_k \,\right]^{\dagger} \left[\, \tilde{v}_1 \ \tilde{v}_2 \ \cdots \ \tilde{v}_k \,\right] = I_k$. Therefore, the matrix polynomial

$$Q_0(\lambda) = P(\lambda) + \Delta_0(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + (A_0 + \Delta_0)\,, \tag{18}$$

lies on the boundary of $\mathcal{B}\left(P, \frac{\|\Delta_0\|_2}{w_0}, w\right)$ and satisfies

$$Q_0(\mu_i)\tilde{v}_i = P(\mu_i)\tilde{v}_i + \Delta_0(\mu_i)\tilde{v}_i = \sigma_i\tilde{u}_i - \sigma_i\tilde{u}_i = 0, \quad i = 1, 2, \ldots, k.$$

Hence, the scalars $\mu_1, \mu_2, \ldots, \mu_k$ are eigenvalues of the matrix polynomial $Q_0(\lambda)$ in (18) with corresponding eigenvectors $\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_k$, respectively.

Proposition 3.5. *Let $\tilde{u}_i, \tilde{v}_i \in \mathbb{C}^n$ be a pair of left and right singular vectors of $P(\mu_i)$ corresponding to $\sigma_i = s_n(P(\mu_i))$, respectively, for every $i = 1, 2, \ldots, k$. If the vectors $\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_k$ are linearly independent, then the matrix polynomial $Q_0(\lambda)$ in (18) lies on the boundary of $\mathcal{B}\left(P, \frac{\|\Delta_0\|_2}{w_0}, w\right)$ and has $\mu_1, \mu_2, \ldots, \mu_k$ as eigenvalues.*

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

The construction of the perturbed matrix polynomial $Q_\gamma(\lambda)$ in (17) yields immediately an upper bound for the distance $D_w(P, \Sigma)$. In particular, from (13) we have

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

where the $n \times n$ matrix Δ_γ is defined as in (12). Consequently, if all scalars $\beta_1, \beta_2, \ldots, \beta_k$ in (11) are nonzero, then for any $\gamma > 0$ such that $\operatorname{rank}(V(\gamma)) = k$, it follows

$$D_w(P, \Sigma) \leq \frac{1}{k} \sum_{i=1}^{k} \left(\frac{1}{w\left(|\mu_i|\right)} \right) \|\Delta_\gamma\|_2 . \tag{19}$$

Next, we compute a lower bound for $D_w(P, \Sigma)$. It is worth mentioning that for calculating this lower bound, the condition $\operatorname{rank}(V(\gamma)) = k$ is not necessary.

Lemma 4.1. *Suppose that $P(\lambda)$ is an $n \times n$ matrix polynomial as in (1), and $\mu_1, \mu_2, \ldots, \mu_k$ are $k \leq n$ distinct eigenvalues of $P(\lambda)$. Then, for every $\gamma > 0$, $s_\rho\left(F_\gamma[P, \Sigma]\right) = 0$ (recall that $\rho = nk - k + 1$).*

Proof. Since $\mu_1, \mu_2, \ldots, \mu_k$ are distinct eigenvalues of $P(\lambda)$, there exist k nonzero (but not necessarily linearly independent) vectors $\nu_1, \nu_2, \ldots, \nu_k$ satisfying $P(\mu_i)\nu_i = 0$, $i = 1, 2, \ldots, k$. Recalling Definition 2.4 and the quantities $\theta_{i,j}$ $(i, j \in \{1, 2, \ldots, k\},\ i \neq j)$ defined by (3), the $nk \times nk$ matrix $F_\gamma\left[P, \Sigma\right]$ can be written in the form

$$F_\gamma\left[P, \Sigma\right] = \begin{bmatrix} P(\mu_1) & 0 & 0 & \cdots & 0 \\ \theta_{1,2}(P(\mu_1) - P(\mu_2)) & P(\mu_2) & 0 & \cdots & 0 \\ \theta_{1,3}[\theta_{1,2}P(\mu_1) - (\theta_{1,2} + \theta_{2,3})P(\mu_2) + \theta_{2,3}P(\mu_3)] & \theta_{2,3}(P(\mu_2) - P(\mu_3)) & P(\mu_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & P(\mu_k) \end{bmatrix}.$$

As a consequence, denoting the (i, j)-th $n \times n$ block of this matrix by $F_{i,j}$, it follows that

$$F_{i,j} = \theta_{j,i}(F_{i-1,j} - F_{i,j+1}), \quad 1 \leq j < i \leq k. \tag{20}$$

Now recall (7) and (8), and consider the k (nonzero) linearly independent vectors

$$\begin{bmatrix} \nu_1 \\ \theta_{1,2}\nu_1 \\ \theta_{1,2}\theta_{1,3}\nu_1 \\ \vdots \\ \left(\prod_{j=2}^{k-1} \theta_{1,j} \right) \nu_1 \\ \left(\prod_{j=2}^{k} \theta_{1,j} \right) \nu_1 \end{bmatrix}, \begin{bmatrix} 0 \\ \nu_2 \\ \theta_{2,3}\nu_2 \\ \vdots \\ \left(\prod_{j=3}^{k-1} \theta_{2,j} \right) \nu_2 \\ \left(\prod_{j=3}^{k} \theta_{2,j} \right) \nu_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \nu_{k-1} \\ \theta_{k-1,k}\nu_{k-1} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \nu_k \end{bmatrix}. \tag{21}$$

It is apparent that

$$\begin{bmatrix} P(\mu_1) & 0 & 0 & \cdots & 0 \\ F_{2,1} & P(\mu_2) & 0 & \cdots & 0 \\ F_{3,1} & F_{3,2} & P(\mu_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{k,1} & F_{k,2} & F_{k,3} & \cdots & P(\mu_k) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \nu_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ P(\mu_k)\nu_k \end{bmatrix} = 0$$

and (applying (20) for $i = k$ and $j = k-1$)

$$\begin{bmatrix} P(\mu_1) & 0 & 0 & \cdots & 0 \\ F_{2,1} & P(\mu_2) & 0 & \cdots & 0 \\ F_{3,1} & F_{3,2} & P(\mu_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{k,1} & F_{k,2} & F_{k,3} & \cdots & P(\mu_k) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \nu_{k-1} \\ \theta_{k-1,k}\nu_{k-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\mu_{k-1})\nu_{k-1} \\ \theta_{k-1,k}P(\mu_{k-1})\nu_{k-1} \end{bmatrix} = 0.$$

For any $s = k-2, k-3, \ldots, 1$, consider the vector

$$\begin{aligned} \psi_s &= \begin{bmatrix} P(\mu_1) & 0 & 0 & \cdots & 0 \\ F_{2,1} & P(\mu_2) & 0 & \cdots & 0 \\ F_{3,1} & F_{3,2} & P(\mu_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{k,1} & F_{k,2} & F_{k,3} & \cdots & P(\mu_k) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \nu_s \\ \theta_{s,s+1}\nu_s \\ (\theta_{s,s+1}\theta_{s,s+2})\,\nu_s \\ \vdots \\ \left(\prod_{j=s+1}^{k} \theta_{s,j}\right)\nu_s \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\mu_s)\nu_s \\ \left[F_{s+1,s} + \theta_{s,s+1}P(\mu_{s+1})\right]\nu_s \\ \left[F_{s+2,s} + \theta_{s,s+1}F_{s+2,s+1} + (\theta_{s,s+1}\theta_{s,s+2})\,P(\mu_{s+2})\right]\nu_s \\ \vdots \\ \left[F_{k,s} + \theta_{s,s+1}F_{k,s+1} + \cdots + \left(\prod_{j=s+1}^{k-1}\theta_{s,j}\right)F_{k,k-1} + \left(\prod_{j=s+1}^{k}\theta_{s,j}\right)P(\mu_k)\right]\nu_s \end{bmatrix}. \end{aligned}$$

By applying (20) (to express all the entries of ψ_s in terms of $P(\mu_s), P(\mu_{s+1}), \ldots, P(\mu_k)$), and (7) and (8) (to construct appropriate products of $\theta_{s,j}$, $j = s+1, s+2, \ldots, k$), straightforward computations yield

$$
\begin{aligned}
\psi_s &= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\mu_s)\nu_s \\ \left[\theta_{s,s+1}(P(\mu_s) - P(\mu_{s+1})) + \theta_{s,s+1}P(\mu_{s+1})\right]\nu_s \\ \left[\theta_{s,s+2}(F_{s+1,s} - F_{s+2,s+1}) + (\theta_{s,s+1}\theta_{s+1,s+2})\left(P(\mu_{s+1}) - P(\mu_{s+2})\right) + (\theta_{s,s+1}\theta_{s,s+2})\, P(\mu_{s+2})\right]\nu_s \\ \vdots \\ \left[\theta_{s,k}(F_{k-1,s} - F_{k,s+1}) + \cdots + \left(\prod_{j=s+1}^{k-1}\theta_{s,j}\right)\theta_{k-1,k}(P(\mu_{k-1}) - P(\mu_k)) + \left(\prod_{j=s+1}^{k}\theta_{s,j}\right)P(\mu_k)\right]\nu_s \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\mu_s)\nu_s \\ \theta_{s,s+1}P(\mu_s)\nu_s \\ \left[(\theta_{s,s+1}\theta_{s,s+2})\, P(\mu_s) + (\theta_{s,s+1}\theta_{s+1,s+2} - \theta_{s,s+2}(\theta_{s,s+1} + \theta_{s+1,s+2}))P(\mu_{s+1}) + (\theta_{s,s+2}(\theta_{s,s+1} + \theta_{s+1,s+2}) - \theta_{s,s+1}\theta_{s+1,s+2})P(\mu_{s+2})\right]\nu_s \\ \vdots \\ \left[(\theta_{s,k}\theta_{s,k-1})(F_{k-2,s} - F_{k-1,s+1}) - (\theta_{s,k}\theta_{s+1,k})(F_{k-1,s+1} - F_{k,s+2}) + \cdots + \left(\prod_{j=s+1}^{k-1}\theta_{s,j}\right)\theta_{k-1,k}(P(\mu_{k-1}) - P(\mu_k)) + \left(\prod_{j=s+1}^{k}\theta_{s,j}\right)P(\mu_k)\right]\nu_s \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ P(\mu_s)\nu_s \\ \theta_{s,s+1}P(\mu_s)\nu_s \\ (\theta_{s,s+1}\theta_{s,s+2})\, P(\mu_s)\nu_s \\ \vdots \\ \left(\prod_{j=s+1}^{k}\theta_{s,j}\right)P(\mu_s)\nu_s \end{bmatrix} = 0.
\end{aligned}
$$

Thus, the k linearly independent vectors in (21) lie in the null space of matrix $F_\gamma\,[P,\Sigma]$. This means that the rank of $F_\gamma\,[P,\Sigma]$ is less than or equal to $kn - k = \rho - 1$, and the proof is complete. $\square$

The next lemma yields a lower bound for $D_w(P,\Sigma)$. We define the nonnegative quantities

$$
\begin{aligned}
&\varpi\,[\mu_i] = w\left(|\mu_i|\right), \quad i = 1, 2, \ldots, k, \\
&\varpi\,[\mu_i, \mu_{i+1}] = \sum_{j=0}^{m} w_j \frac{\left|\mu_i^j - \mu_{i+1}^j\right|}{|\mu_i - \mu_{i+1}|}, \quad i = 1, 2, \ldots, k-1,
\end{aligned}
$$

and (recursively)

$$\varpi[\mu_i,\mu_{i+1},\ldots,\mu_{i+t}] = \frac{\varpi[\mu_i,\ldots,\mu_{i+t-1}] + \varpi[\mu_{i+1},\ldots,\mu_{i+t}]}{|\mu_i-\mu_{i+t}|}, \quad i=1,2,\ldots,k-2,\ t=2,3,\ldots,k-i,$$

and the $k\times k$ matrix

$$F_\gamma[\varpi,\Sigma] = \begin{bmatrix} \varpi[\mu_1] & 0 & \cdots & 0 \\ \gamma\varpi[\mu_1,\mu_2] & \varpi[\mu_2] & \cdots & 0 \\ \gamma^2\varpi[\mu_1,\mu_2,\mu_3] & \gamma\varpi[\mu_2,\mu_3] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{k-1}\varpi[\mu_1,\mu_2,\ldots,\mu_k] & \gamma^{k-2}\varpi[\mu_2,\mu_3,\ldots,\mu_k] & \cdots & \varpi[\mu_k] \end{bmatrix}.$$

Lemma 4.2. *Suppose that the matrix polynomial $Q(\lambda)=P(\lambda)+\Delta(\lambda)$ belongs to $\mathcal{B}(P,\varepsilon,w)$. If $k\le n$ distinct scalars $\mu_1,\mu_2,\ldots,\mu_k\in\mathbb{C}$ are eigenvalues of $Q(\lambda)$, then for any $\gamma>0$,*

$$\varepsilon \ge \frac{s_\rho(F_\gamma[P,\Sigma])}{\|F_\gamma[\varpi,\Sigma]\|_2}.$$

Proof. Since the k distinct scalars $\mu_1,\mu_2,\ldots,\mu_k$ are eigenvalues of $Q(\lambda)=P(\lambda)+\Delta(\lambda)$, Lemma 4.1 implies that $s_\rho(F_\gamma[Q,\Sigma])=0$. As a consequence, the Weyl inequalities for singular values (e.g., see Corollary 5.1 of [1]), applied to $F_\gamma[Q,\Sigma]=F_\gamma[P,\Sigma]+F_\gamma[\Delta,\Sigma]$, yield $s_\rho(F_\gamma[P,\Sigma])\le\|F_\gamma[\Delta,\Sigma]\|_2$ for any $\gamma>0$. Keeping in mind that $\Delta(\lambda)$ is of the form $\Delta(\lambda)=\sum_{j=0}^{m}\Delta_j\lambda^j$ with $\Delta_j\in\mathbb{C}^{n\times n}$ satisfying $\|\Delta_j\|_2\le\varepsilon w_j$ $(j=0,1,\ldots,m)$, the rest of the proof is devoted to establish the inequality

$$\|F_\gamma[\Delta,\Sigma]\|_2 \le \varepsilon\,\|F_\gamma[\varpi,\Sigma]\|_2\,.$$

It is easy to see that

$$\|\Delta(\mu_i)\|_2 \le \sum_{j=0}^{m}\|\Delta_j\|_2\,|\mu_i|^j \le \varepsilon\sum_{j=0}^{m}w_j\,|\mu_i|^j = \varepsilon w(|\mu_i|) = \varepsilon\varpi[\mu_i]\,, \quad i=1,2,\ldots,k,$$

and

$$\|\Delta[\mu_i,\mu_{i+1}]\|_2 \le \sum_{j=0}^{m}\|\Delta_j\|_2\left|\frac{\mu_i^j-\mu_{i+1}^j}{\mu_i-\mu_{i+1}}\right| \le \varepsilon\varpi[\mu_i,\mu_{i+1}]\,, \quad i=1,2,\ldots,k-1.$$

For the sake of induction, we assume that for a given $t\in\{1,2,\ldots,k-2\}$, it holds

$$\|\Delta[\mu_i,\mu_{i+1},\ldots,\mu_{i+t}]\|_2 \le \varepsilon\varpi[\mu_i,\mu_{i+1},\ldots,\mu_{i+t}]\,, \quad i=1,2,\ldots,k-t.$$

Then it follows

$$\begin{aligned}
\left\|\Delta\left[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t+1}\right]\right\|_2 &\leq \frac{\left\|\Delta\left[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t}\right]\right\|_2+\left\|\Delta\left[\mu_{i+1}, \mu_{i+2}, \ldots, \mu_{i+t+1}\right]\right\|_2}{\left|\mu_i-\mu_{i+t+1}\right|} \\
&\leq \frac{\varepsilon \varpi\left[\mu_i, \mu_{i+1}, \ldots, \mu_{i+t}\right]+\varepsilon \varpi\left[\mu_{i+1}, \mu_{i+2}, \ldots, \mu_{i+t+1}\right]}{\left|\mu_i-\mu_{i+t+1}\right|} \\
&= \varepsilon \varpi\left[\mu_i, \mu_{i+1},, \ldots, \mu_{i+t+1}\right]
\end{aligned}$$

for every $i=1,2, \ldots, k-t-1$. Hence, we obtain

$$\left\|\Delta\left[\mu_i, \ldots, \mu_{i+t}\right]\right\|_2 \leq \varepsilon \varpi\left[\mu_i, \ldots, \mu_{i+t}\right], \quad t=0,1, \ldots, k-1, \; i=1,2, \ldots, k-t.$$

As in the proof of Theorem 2.4 of [18], we can consider a unit vector

$$x=\left[\begin{array}{c}
x_1 \\
x_2 \\
\vdots \\
x_k
\end{array}\right] \in \mathbb{C}^{kn} \quad\left(x_i \in \mathbb{C}^n, \; i=1,2, \ldots, k\right)$$

such that

$$\begin{aligned}
\left\|F_\gamma[\Delta, \Sigma]\right\|_2^2 &= \left\|F_\gamma[\Delta, \Sigma] x\right\|_2^2 \\
&= \left\|\Delta\left(\mu_1\right) x_1\right\|_2^2+\left\|\gamma \Delta\left[\mu_1, \mu_2\right] x_1+\Delta\left(\mu_2\right) x_2\right\|_2^2 \\
&\quad +\cdots+\left\|\sum_{i=1}^k \gamma^{k-i} \Delta\left[\mu_i, \ldots, \mu_k\right] x_i\right\|_2^2 \\
&\leq \left(\varepsilon \varpi\left[\mu_1\right]\right)^2\left\|x_1\right\|_2^2+\left(\gamma \varepsilon \varpi\left[\mu_1, \mu_2\right]\right)^2\left\|x_1\right\|_2^2+\left(\varepsilon \varpi\left[\mu_2\right]\right)^2\left\|x_2\right\|_2^2 \\
&\quad +2 \gamma\left(\varepsilon \varpi\left[\mu_1, \mu_2\right]\right)\left(\varepsilon \varpi\left[\mu_2\right]\right)\left\|x_1\right\|_2\left\|x_2\right\|_2+\cdots+\left(\varepsilon \varpi\left[\mu_k\right]\right)^2\left\|x_k\right\|_2^2 \\
&= \varepsilon^2\left\|\left[\begin{array}{cccc}
\varpi\left[\mu_1\right] & 0 & \cdots & 0 \\
\gamma \varpi\left[\mu_1, \mu_2\right] & \varpi\left[\mu_2\right] & \cdots & 0 \\
\gamma^2 \varpi\left[\mu_1, \mu_2, \mu_3\right] & \gamma \varpi\left[\mu_2, \mu_3\right] & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\gamma^{k-1} \varpi\left[\mu_1, \mu_2, \ldots, \mu_k\right] & \gamma^{k-2} \varpi\left[\mu_2, \mu_3, \ldots, \mu_k\right] & \cdots & \varpi\left[\mu_k\right]
\end{array}\right]\left[\begin{array}{c}
\left\|x_1\right\|_2 \\
\left\|x_2\right\|_2 \\
\vdots \\
\left\|x_k\right\|_2
\end{array}\right]\right\|_2^2 \\
&\leq \varepsilon^2\left\|F_\gamma[\varpi, \Sigma]\right\|_2^2 .
\end{aligned}$$

This completes the proof. □

Keeping in mind Definition 2.2, the above lemma yields a lower bound for $D_w(P, \Sigma)$, namely,

$$D_w(P, \Sigma) \geq \frac{s_\rho\left(F_\gamma[P, \Sigma]\right)}{\left\|F_\gamma[\varpi, \Sigma]\right\|_2}. \tag{22}$$

It will be convenient to denote the lower bound in (22) by $\beta_{low}(P,\Sigma,\gamma)$ and the upper bound in (19) by $\beta_{up}(P,\Sigma,\gamma)$, i.e.,

$$\beta_{low}(P,\Sigma,\gamma)=\frac{s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)}{\left\|F_\gamma\left[\varpi,\Sigma\right]\right\|_2} \tag{23}$$

and (recalling the $n\times n$ matrix Δ_γ defined by (12))

$$\beta_{up}(P,\Sigma,\gamma)=\frac{1}{k}\sum_{i=1}^{k}\left(\frac{1}{w\left(|\mu_i|\right)}\right)\left\|\Delta_\gamma\right\|_2. \tag{24}$$

Our main results are summarized in the following theorem.

Theorem 4.3. *Consider an $n\times n$ matrix polynomial $P(\lambda)$ as in (1) and a given set of $k\le n$ distinct complex numbers $\Sigma=\{\mu_1,\mu_2,\ldots,\mu_k\}$.*

(a) *For any $\gamma>0$, $D_w(P,\Sigma)\ge\beta_{low}(P,\Sigma,\gamma)$.*

(b) *If the quantities $\beta_1,\beta_2,\ldots,\beta_k$ in (11) are nonzero, then for any $\gamma>0$ such that $\mathrm{rank}(V(\gamma))=k$, $D_w(P,\Sigma)\le\beta_{up}(P,\Sigma,\gamma)$ and the matrix polynomial $Q_\gamma(\gamma)$ in (17) lies on the boundary of $\mathcal{B}(P,\beta_{up}(P,\Sigma,\gamma),w)$.*

In the next remark, we give an upper and a lower bound for a spectral norm distance from an $n\times n$ matrix A to the set of all matrices with k prescribed eigenvalues. This issue is explained in [9,12] in detail.

Remark 4.4. Consider the standard eigenproblem of a matrix $A\in\mathbb{C}^{n\times n}$. In this special case, we set $P(\lambda)=I\lambda-A$ and $w=\{w_0,w_1\}=\{1,0\}$. Thus, for every $i=1,2,\ldots,k$, $\varpi\left[\mu_i\right]=w\left(|\mu_i|\right)=w_0$ and $\varpi\left[\mu_i,\ldots,\mu_j\right]=0$ for every $j=\{i+1,i+2,\ldots,k\}$. Consequently, the matrix $F_\gamma\left[\varpi,\Sigma\right]$ becomes the identity matrix I_k and the lower bound in (23) turns into $\beta_{low}(P,\Sigma,\gamma)=s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)$. Furthermore, it is easy to see that $\alpha_{i,s}=1$ and $\beta_s=1$ for every $i,s=1,2,\ldots,k$. Therefore, the upper bound in (24) becomes

$$\beta_{up}(P,\Sigma,\gamma)=\left\|\Delta_\gamma\right\|_2=s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\left\|\hat{U}\left(\gamma\right)\hat{V}(\gamma)^\dagger\right\|_2.$$

Moreover, the associated perturbed matrix polynomial $Q_\gamma(\lambda)$ in (17) is now given by

$$Q_\gamma(\lambda)=P(\lambda)+\Delta_\gamma(\lambda)=P(\lambda)+\Delta_\gamma=I\lambda-\left(A+s_\rho\left(F_\gamma\left[P,\Sigma\right]\right)\hat{U}\left(\gamma\right)\hat{V}(\gamma)^\dagger\right). \tag{25}$$

5 Numerical examples

In this section, the validity of the results in the previous sections is verified by two numerical examples. The lower and upper bounds for the distance $D_w(P,\Sigma)$ are computed by applying the procedures described in Section 4, and by using the MATLAB function `fminbnd` which finds a minimum of a function of one variable within a fixed interval. As it was mentioned in Remark 3.3, the condition $\mathrm{rank}(V(\gamma)) = k$ appears to be generic when $\gamma > 0$. 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 3×3 matrix polynomial

$$P(\lambda) = \begin{bmatrix} 7 & 9 & -2 \\ 0 & -2 & 0 \\ 6 & -3 & -1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 9 & -3 & 3 \\ -5 & 8 & 10 \\ 4 & -3 & 0 \end{bmatrix} \lambda + \begin{bmatrix} -5 & 0 & 5 \\ -2 & -2 & 10 \\ 1 & 9 & 2 \end{bmatrix},$$

whose spectrum is $\sigma(P) = \{76.9807, 0.9284, 0.3034, -1.0283, -0.9421 \pm 0.9281\,\mathrm{i}\}$. Let $w = \{w_0, w_1, w_2\} = \{12.0731, 14.8523, 11.7991\}$ be the set of weights which are the norms of the coefficient matrices, and suppose that the set of desired eigenvalues is $\Sigma = \{1+\mathrm{i}, -2, 3\}$. By applying the MATLAB function `fminbnd`, it appears that the function $\beta_{up}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$ $(\gamma > 0)$ attains its minimum at $\gamma = 1.9656$, that is,

$$\beta_{up}(P, \{1+\mathrm{i}, -2, 3\}, 1.9656) = 1.0090,$$

and the function $\beta_{low}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$ $(\gamma > 0)$ attains its maximum at $\gamma = 5.9606 \cdot 10^{-5}$, that is,

$$\beta_{low}(P, \{1+\mathrm{i}, -2, 3\}, 5.9606 \cdot 10^{-5}) = 0.1320.$$

In Figure 1, the graphs of the upper bound $\beta_{up}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$ and the lower bound $\beta_{low}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$ are plotted for $\gamma \in (0, 10]$. Also, for the perturbation

$$\begin{aligned} \Delta_{1.9656}(\lambda) \; = \; & \begin{bmatrix} -1.5506 + 0.5852\,\mathrm{i} & -3.6805 - 3.7560\,\mathrm{i} & 3.2843 - 2.4550\,\mathrm{i} \\ -1.3951 + 1.1287\,\mathrm{i} & 0.8130 - 3.6071\,\mathrm{i} & 1.4666 + 0.2551\,\mathrm{i} \\ -4.9524 + 1.3272\,\mathrm{i} & -0.1817 - 0.1712\,\mathrm{i} & -0.1517 - 2.5523\,\mathrm{i} \end{bmatrix} \lambda^2 \\ & + \begin{bmatrix} -1.0045 + 0.6941\,\mathrm{i} & -3.2991 - 2.0307\,\mathrm{i} & 1.9114 - 2.3391\,\mathrm{i} \\ -0.7966 + 1.0550\,\mathrm{i} & -0.0602 - 2.7233\,\mathrm{i} & 1.0938 - 0.0784\,\mathrm{i} \\ -3.3045 + 1.8295\,\mathrm{i} & -0.1603 - 0.0901\,\mathrm{i} & -0.5623 - 1.7977\,\mathrm{i} \end{bmatrix} \lambda \\ & + \begin{bmatrix} -2.1779 - 1.0042\,\mathrm{i} & 0.1345 - 7.6081\,\mathrm{i} & 5.8658 + 0.8927\,\mathrm{i} \\ -2.5802 - 0.2920\,\mathrm{i} & 4.5439 - 2.8248\,\mathrm{i} & 1.2263 + 1.7709\,\mathrm{i} \\ -6.3971 - 3.7574\,\mathrm{i} & -0.0080 - 0.3612\,\mathrm{i} & 2.4770 - 2.7481\,\mathrm{i} \end{bmatrix} \end{aligned}$$

the perturbed matrix polynomial $Q_{1.9656}(\lambda) = P(\lambda) + \Delta_{1.9656}(\lambda)$ lies on the boundary of the set $\mathcal{B}(P, \beta_{up}(P, \{1+\mathrm{i}, -2, 3\}, 1.9656), w) = \mathcal{B}(P, 1.0090, w)$ and has Σ in its spectrum.

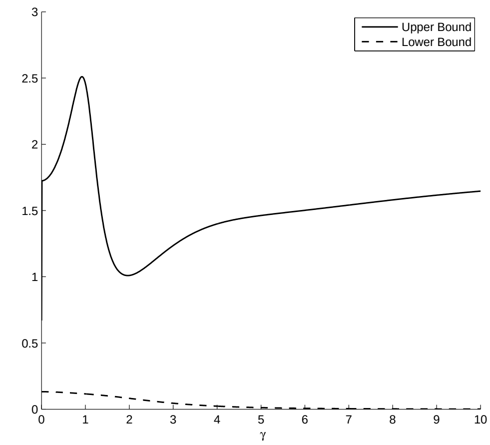

Fig 1: The graphs of $\beta_{low}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$ and $\beta_{up}(P, \{1+\mathrm{i}, -2, 3\}, \gamma)$.

It is worth mentioning that the discussion at the end of Section 3 yields the perturbation

$$\Delta_0(\lambda) = \Delta_0 = \begin{bmatrix} 0.0673 + 0.0158\,\mathrm{i} & 0.0656 - 0.0194\,\mathrm{i} & 0.0060 - 0.0079\,\mathrm{i} \\ 1.2669 - 0.1878\,\mathrm{i} & 0.0412 + 0.2304\,\mathrm{i} & -0.6315 + 0.0940\,\mathrm{i} \\ 0.3092 - 0.1368\,\mathrm{i} & -0.1210 + 0.1678\,\mathrm{i} & -0.2397 + 0.0684\,\mathrm{i} \end{bmatrix} \cdot 10^2.$$

The perturbed matrix polynomial $Q_0(\lambda) = P(\lambda) + \Delta_0$ lies on the boundary of $\mathcal{B}(P, 12.5337, w)$ and has Σ in its spectrum.

Our second example illustrates the applicability of Remark 4.4.

Example 5.2. Consider the Frank matrix of order 12,

$$F_{12} = \begin{bmatrix} 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 11 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 10 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 9 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 8 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 7 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 6 & 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

which has some small ill-conditioned eigenvalues. Suppose that the set of the desired

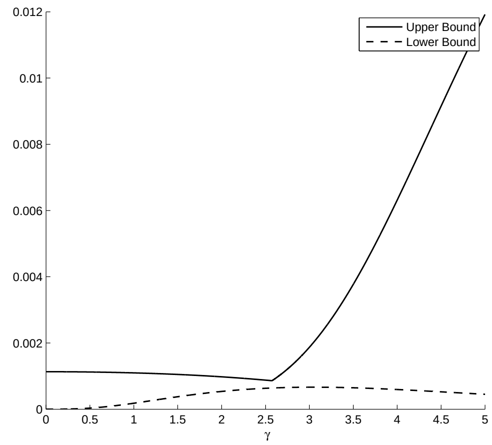

Fig 2: The graphs of $\beta_{low}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma)$ and $\beta_{up}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma)$.

eigenvalues is $\Sigma = \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}$. The optimal (spectral norm) distance from F_{12} to the set of matrices that have Σ in their spectrum is $6.9 \cdot 10^{-4}$ [12]. We consider the linear matrix polynomial $P(\lambda) = \lambda I_{12} - F_{12}$, and the weights $w_0 = 1$ and $w_1 = 0$ (i.e., we consider perturbations of the standard eigenproblem of matrix F_{12}). The MATLAB function `fminbnd` applied to the difference

$$\beta_{up}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma) - \beta_{low}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma)$$

yields $\gamma = 2.5730$. Then, according to the discussion in Remark 4.4, we have

$$\begin{aligned} \beta_{low}(P, \Sigma, 2.5730) = 6.4007 \cdot 10^{-4} &\leq 6.9 \cdot 10^{-4} = D_w(P, \Sigma) \\ &\leq 8.6167 \cdot 10^{-4} = \beta_{up}(P, \Sigma, 2.5730). \end{aligned}$$

Also, it is easy to see that the spectrum of the perturbed linear matrix polynomial $Q_\gamma(\lambda)$ in (25) includes the given set Σ. In Figure 2, the graphs of the upper bound $\beta_{up}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma)$ and the lower bound $\beta_{low}(P, \{0.1, -0.1, 0.1\,\mathrm{i}, -0.1\,\mathrm{i}\}, \gamma)$ are plotted for $\gamma \in (0, 5]$.

6 Concluding remarks

In this article, a spectral norm distance from an $n \times n$ matrix polynomial $P(\lambda)$ to the $n \times n$ matrix polynomials that have $k \leq n$ distinct complex numbers as eigenvalues is introduced and studied. An upper and a lower bound for this distance are obtained. Furthermore, a perturbation of $P(\lambda)$ with the given scalars as eigenvalues and associated to the upper bound is constructed under two conditions (namely, $\text{rank}(V(\gamma)) = k$ and $\beta_1, \beta_2, \ldots, \beta_k \neq 0$) which appear in our experiments to hold generically. The tightness of these bounds is illustrated by two numerical examples. Overall, the proposed methodology can be considered as an extension of the results in [5, 10, 12, 13, 17, 18].

A question that arises in a natural way, is what one can say about the case where some of the desired eigenvalues are multiple. In this case, it seems that it is necessary to replace some of the divided differences in the $nk \times nk$ matrix $F_\gamma[P, \Sigma]$ (see Definition 2.4) by derivatives of the matrix polynomial $P(\lambda)$. The mixture of divided differences and derivatives in the definition of $F_\gamma[P, \Sigma]$ yields several computational difficulties. Moreover, the new perturbations and bounds will be of different type than the perturbation constructed in Section 3 and the bounds obtained in Section 4. As a consequence, this problem requires the development of a modified technique based on the combination of the methodology given herein and the methods established in [9, 12, 17, 18]; this will be the subject of a future work.

Acknowledgments. We acknowledge with thanks an anonymous referee for useful comments and suggestions.

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] J.D. Faires and R.L. Burden, *Numerical Methods*, third edition, Brooks Cole, 2002.

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

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

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

[8] M. Karow and E. Mengi, Matrix polynomials with specified eigenvalues, *Linear Algbera Appl.*, 466 (2015), 457–482.

[9] 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.

[10] 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.

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

[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, Providence, RI, Translations of Mathematical Monographs, Vol. 71, 1988.

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

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

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

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

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

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