

A distance bound for pseudospectra of matrix polynomials

Panayiotis J. Psarrakos¹

June 13, 2006

Abstract

In this note, we obtain a lower bound for the distance between the pseudospectrum of a matrix polynomial and a given point that lies out of it, generalizing a known result on pseudospectra of matrices.

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

AMS Subject Classifications: 15A60; 65F30

1 Introduction and definitions

Let $\mathbb{C}^{n \times n}$ be the algebra of all $n \times n$ complex matrices, and consider the higher order linear system $A_m u^{(m)}(t) + A_{m-1} u^{(m-1)}(t) + \dots + A_1 u^{(1)}(t) + A_0 u(t) = f(t)$, where $A_j \in \mathbb{C}^{n \times n}$ ($j = 0, 1, \dots, m$) with $\det A_m \neq 0$, $u(t) \in \mathbb{C}^n$ is the unknown vector function and $f(t) \in \mathbb{C}^n$ is piecewise continuous (the indices on $u(t)$ denote derivatives with respect to the independent variable t). Applying the Laplace transformation yields the *matrix polynomial*

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

where λ is a complex variable. The study of matrix polynomials has a long history, especially with regard to their spectral analysis, which leads to the solutions of the corresponding systems of differential equations [1].

A scalar $\lambda_0 \in \mathbb{C}$ is said to be an *eigenvalue* of the matrix polynomial $P(\lambda)$ in (1) if the system $P(\lambda_0)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$. This solution x_0 is known as an *eigenvector* of $P(\lambda)$ corresponding to λ_0 . The set of all eigenvalues of $P(\lambda)$ is the *spectrum* of $P(\lambda)$, namely, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, and since $\det A_m \neq 0$, it contains no more than nm distinct (finite) elements.

¹Department of Mathematics, National Technical University, Zografou Campus, Athens 15780, Greece (E-mail: ppsarr@math.ntua.gr). Research supported by a grant of the EPEAEK project PYTHAGORAS II. The project is co-funded by the European Social Fund (75%) and Greek National Resources (25%).

We are interested in the spectra of perturbations of the matrix polynomial $P(\lambda)$ in (1) of the form

$$P_{\Delta}(\lambda) = (A_m + \Delta_m)\lambda^m + (A_{m-1} + \Delta_{m-1})\lambda^{m-1} + \cdots + (A_1 + \Delta_1)\lambda + A_0 + \Delta_0,$$

where the matrices $\Delta_0, \Delta_1, \dots, \Delta_m \in \mathbb{C}^{n \times n}$ are arbitrary. For a given $\varepsilon > 0$ and a given set of nonnegative weights $\mathbf{w} = \{w_0, w_1, \dots, w_m\}$ with at least one nonzero element, the (weighted) ε -pseudospectrum of $P(\lambda)$ is defined by

$$\sigma_{\varepsilon, \mathbf{w}}(P) = \{\lambda \in \mathbb{C} : \det P_{\Delta}(\lambda) = 0, \|\Delta_j\|_2 \leq \varepsilon w_j, j = 0, 1, \dots, m\},$$

where $\|\cdot\|_2$ denotes the *spectral norm*, i.e., the matrix norm subordinate to the Euclidean vector norm. The parameters $w_0, w_1, \dots, w_m \geq 0$ allow freedom in how perturbations are measured; for example, in an absolute sense when $w_0 = w_1 = \cdots = w_m = 1$, or in a relative sense when $w_j = \|A_j\|_2$ ($j = 0, 1, \dots, m$).

If $P(\lambda) = I\lambda - A$ for some $A \in \mathbb{C}^{n \times n}$, then $\sigma(P)$ coincides with the standard spectrum of A , $\sigma(A)$. If in addition, we set $\mathbf{w} = \{w_0, w_1\} = \{1, 0\}$, then $\sigma_{\varepsilon, \mathbf{w}}(P)$ coincides with the ε -pseudospectrum of the matrix A [2, 3, 4, 5], that is,

$$\sigma_{\varepsilon}(A) = \{\lambda \in \mathbb{C} : \lambda \in \sigma(A + E), \|E\|_2 \leq \varepsilon\}.$$

Denote by $s_{\min}(\cdot)$ and $s_{\max}(\cdot)$ the minimum and the maximum singular values of a complex matrix, respectively. If we consider the scalar polynomial

$$q_{\mathbf{w}}(\lambda) = w_m \lambda^m + w_{m-1} \lambda^{m-1} + \cdots + w_1 \lambda + w_0, \quad (2)$$

then by [6, Lemma 2.1],

$$\sigma_{\varepsilon, \mathbf{w}}(P) = \{\lambda \in \mathbb{C} : s_{\min}(P(\lambda)) \leq \varepsilon q_{\mathbf{w}}(|\lambda|)\}.$$

As the parameter $\varepsilon > 0$ increases, the ε -pseudospectrum of $P(\lambda)$ enlarges, and for ε large enough, $\sigma_{\varepsilon, \mathbf{w}}(P)$ is no longer bounded. On the other hand, since the leading coefficient A_m is nonsingular, for sufficiently small ε , $\sigma_{\varepsilon, \mathbf{w}}(P)$ consists of no more than nm bounded connected components, each one containing a single (possibly multiple) eigenvalue of $P(\lambda)$. Moreover, by Theorems 2.2 and 2.3 of [7], we know that the pseudospectrum $\sigma_{\varepsilon, \mathbf{w}}(P)$ is bounded if and only if $\varepsilon w_m < s_{\min}(A_m)$, and in this case, it has no more than nm connected components.

Pseudospectra provide important insights into the sensitivity of eigenvalues under perturbations and have several applications (see [2, 3, 4, 5, 6, 7] and the references therein). In this article, we continue the investigation of the ε -pseudospectrum of the matrix polynomial $P(\lambda)$ in (1), constructing a lower bound for the distance between $\sigma_{\varepsilon, \mathbf{w}}(P)$ and a given point $\lambda_0 \notin \sigma_{\varepsilon, \mathbf{w}}(P)$.

2 The distance lower bound

A simple inclusion-exclusion algorithm for the estimation of pseudospectra of complex matrices was recently proposed by Koutis and Gallopoulos [8] (this work can be downloaded from [4]). Their methodology is based on the following result (see also [2, 3]).

Theorem 1 [8, Theorem 2.4] *Let $A \in \mathbb{C}^{n \times n}$, $\varepsilon > 0$ and $\lambda_0 \notin \sigma_\varepsilon(A)$. Then the distance $\text{dist}(\lambda_0, \sigma_\varepsilon(A))$ from the point λ_0 to the ε -pseudospectrum of A satisfies*

$$\text{dist}(\lambda_0, \sigma_\varepsilon(A)) \geq s_{\min}(I\lambda_0 - A) - \varepsilon.$$

Consider now an $n \times n$ matrix polynomial $P(\lambda)$ as in (1), an $\varepsilon > 0$, some weights $w_0, w_1, \dots, w_m \geq 0$ and the corresponding polynomial $q_{\mathbf{w}}(\lambda)$ in (2). For a given $\lambda_0 \notin \sigma_{\varepsilon, \mathbf{w}}(P)$, we obtain a lower bound for the distance $\text{dist}(\lambda_0, \sigma_{\varepsilon, \mathbf{w}}(P))$, generalizing Theorem 1. The following two lemmas are necessary for our discussion. The first lemma can be found in [9], and the second one is a simple exercise in polynomials.

Lemma 2 *For any $A, B \in \mathbb{C}^{n \times n}$, $|s_{\min}(A + B) - s_{\min}(A)| \leq s_{\max}(B)$.*

Lemma 3 *Let $p(\lambda) = a_m \lambda^m + a_{m-1} \lambda^{m-1} + \dots + a_1 \lambda - a_0$ be a scalar polynomial with $a_0 > 0$, $a_1, a_2, \dots, a_m \geq 0$ and at least one of the coefficients a_1, a_2, \dots, a_m positive. Then $p(\lambda)$ has exactly one positive zero.*

Theorem 4 *For any $\lambda_0 \notin \sigma_{\varepsilon, \mathbf{w}}(P)$, we have the following two cases:*

- (i) *Suppose that at least one of the given weights w_1, w_2, \dots, w_m is positive, and r_1 is the positive root of*

$$\frac{q_{\mathbf{w}}^{(m)}(|\lambda_0|)}{m!} \lambda^m + \dots + \frac{q_{\mathbf{w}}^{(1)}(|\lambda_0|)}{1!} \lambda - \left(\frac{s_{\min}(P(\lambda_0))}{\varepsilon} - q_{\mathbf{w}}(|\lambda_0|) \right) = 0.$$

For any $\gamma \in (0, 1)$, let r_γ be the positive root of the equation

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} \lambda^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} \lambda - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0| + \gamma r_1)) = 0.$$

Then $\text{dist}(\lambda_0, \sigma_{\varepsilon, \mathbf{w}}(P)) \geq \min\{\gamma r_1, r_\gamma\}$.

- (ii) *If $w_1 = w_2 = \dots = w_m = 0$ and r_0 is the positive root of*

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} \lambda^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} \lambda - (s_{\min}(P(\lambda_0)) - \varepsilon w_0) = 0,$$

then $\text{dist}(\lambda_0, \sigma_{\varepsilon, \mathbf{w}}(P)) \geq r_0$.

Proof Suppose that $\lambda_0 \notin \sigma_{\varepsilon, \mathbf{w}}(P)$, or equivalently, $s_{\min}(P(\lambda_0)) > \varepsilon q_{\mathbf{w}}(|\lambda_0|)$. Then for any nonzero $\mu \in \mathbb{C}$, we have

$$P(\lambda_0 + \mu) = P(\lambda_0) + \frac{P^{(1)}(\lambda_0)}{1!} \mu + \dots + \frac{P^{(m)}(\lambda_0)}{m!} \mu^m,$$

where the matrix $P^{(m)}(\lambda_0)/(m!) = A_m$ is nonsingular. By Lemma 2 and norm properties, it follows

$$\begin{aligned} |s_{\min}(P(\lambda_0 + \mu)) - s_{\min}(P(\lambda_0))| &\leq s_{\max} \left(\sum_{j=1}^m \frac{P^{(j)}(\lambda_0)}{j!} \mu^j \right) \\ &\leq \sum_{j=1}^m \frac{\|P^{(j)}(\lambda_0)\|_2}{j!} |\mu|^j. \end{aligned}$$

Hence,

$$-\sum_{j=1}^m \frac{\|P^{(j)}(\lambda_0)\|_2}{j!} |\mu|^j \leq s_{\min}(P(\lambda_0 + \mu)) - s_{\min}(P(\lambda_0)),$$

or equivalently,

$$s_{\min}(P(\lambda_0)) - \sum_{j=1}^m \frac{\|P^{(j)}(\lambda_0)\|_2}{j!} |\mu|^j \leq s_{\min}(P(\lambda_0 + \mu)).$$

Thus, for

$$\varepsilon < \frac{1}{q_{\mathbf{w}}(|\lambda_0 + \mu|)} \left(s_{\min}(P(\lambda_0)) - \sum_{j=1}^m \frac{\|P^{(j)}(\lambda_0)\|_2}{j!} |\mu|^j \right),$$

or equivalently, for

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} |\mu|^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} |\mu| - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0 + \mu|)) < 0, \quad (3)$$

we have $s_{\min}(P(\lambda_0 + \mu)) > \varepsilon q_{\mathbf{w}}(|\lambda_0 + \mu|)$, i.e., $\lambda_0 + \mu \notin \sigma_{\varepsilon, \mathbf{w}}(P)$. Furthermore, observe that the difference $s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0 + \mu|)$ (in the constant coefficient of the scalar polynomial in the left-hand part of (3)) is positive when $s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0| + |\mu|) > 0$, or equivalently, when

$$\frac{q_{\mathbf{w}}^{(m)}(|\lambda_0|)}{m!} |\mu|^m + \dots + \frac{q_{\mathbf{w}}^{(1)}(|\lambda_0|)}{1!} |\mu| - \left(\frac{s_{\min}(P(\lambda_0))}{\varepsilon} - q_{\mathbf{w}}(|\lambda_0|) \right) < 0. \quad (4)$$

Next we consider the two cases of the theorem:

(i) Assume that at least one of the weights w_1, w_2, \dots, w_m is positive. Since $s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0|) > 0$, by Lemma 3, the polynomial

$$\frac{q_{\mathbf{w}}^{(m)}(|\lambda_0|)}{m!} \lambda^m + \dots + \frac{q_{\mathbf{w}}^{(1)}(|\lambda_0|)}{1!} \lambda - \left(\frac{s_{\min}(P(\lambda_0))}{\varepsilon} - q_{\mathbf{w}}(|\lambda_0|) \right)$$

has exactly one positive zero, r_1 . Then for every nonzero $\mu \in \mathbb{C}$ with $|\mu| < r_1$, (4) holds and $s_{\min}(P(\lambda_0)) > \varepsilon q_{\mathbf{w}}(|\lambda_0| + |\mu|)$. Hence, for any $\gamma \in (0, 1)$,

$$s_{\min}(P(\lambda_0)) > \varepsilon q_{\mathbf{w}}(|\lambda_0| + \gamma r_1),$$

and consequently, the scalar polynomial

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} \lambda^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} \lambda - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0| + \gamma r_1))$$

satisfies the conditions of Lemma 3 and has exactly one positive zero, r_γ . Furthermore, for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < \min\{\gamma r_1, r_\gamma\}$, we have

$$\begin{aligned} & \frac{\|P^{(m)}(\lambda_0)\|_2}{m!} |\mu|^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} |\mu| - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0 + \mu|)) \\ & \leq \frac{\|P^{(m)}(\lambda_0)\|_2}{m!} |\mu|^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} |\mu| - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0| + |\mu|)) \\ & < \frac{\|P^{(m)}(\lambda_0)\|_2}{m!} |\mu|^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} |\mu| - (s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0| + \gamma r_1)) \\ & < 0. \end{aligned}$$

Thus, for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < \min\{\gamma r_1, r_\gamma\}$, both (3) and (4) hold, and as a consequence, $\lambda_0 + \mu \notin \sigma_{\varepsilon, \mathbf{w}}(P)$.

(ii) Assume that $w_1 = w_2 = \dots = w_m = 0$ and $w_0 > 0$. Then $q_{\mathbf{w}}(\lambda) = w_0$ for every $\lambda \in \mathbb{C}$. Hence, for every $\mu \in \mathbb{C}$, the difference $s_{\min}(P(\lambda_0)) - \varepsilon q_{\mathbf{w}}(|\lambda_0 + \mu|) = s_{\min}(P(\lambda_0)) - \varepsilon w_0$ is positive. The scalar polynomial

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} \lambda^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} \lambda - (s_{\min}(P(\lambda_0)) - \varepsilon w_0)$$

satisfies the conditions of Lemma 3 and has exactly one positive zero, r_0 . As in case (i), for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < r_0$,

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!} |\mu|^m + \dots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!} |\mu| - (s_{\min}(P(\lambda_0)) - \varepsilon w_0) < 0,$$

i.e., (3) holds. Thus, $s_{\min}(P(\lambda_0 + \mu)) > \varepsilon w_0$, or equivalently, the point $\lambda_0 + \mu$ does not belong to $\sigma_{\varepsilon, \mathbf{w}}(P)$. \square

Note that r_γ in part (i) of this theorem is a continuous decreasing function of the variable $\gamma \in (0, 1)$ with $\lim_{\gamma \rightarrow 1^-} r_\gamma = 0$. As a consequence, the curve $\{(\gamma, r_\gamma) : \gamma \in (0, 1)\}$ has exactly one common point with the line segment $\{(\gamma, \gamma r_1) : \gamma \in (0, 1)\}$, which is the only maximum of the function $\min\{\gamma r_1, r_\gamma\}$ (see Figure 2 below). If this common point is $(\gamma_0, r_{\gamma_0}) = (\gamma_0, \gamma_0 r_1)$ (for a $\gamma_0 \in (0, 1)$), then $r_{\gamma_0} = \gamma_0 r_1$ is the best lower bound that Theorem 4 can give, as it is illustrated in the following example.

Example The spectrum of the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \lambda + \begin{bmatrix} -2 & 8 & 0 \\ 10 & 6 & 0 \\ 8 & -8 & 10 \end{bmatrix}$$

is $\sigma(P) = \{-3.9698, -1.9194, 1.6868, 4.6209, 0.2908 \pm i3.9250\}$. For $\varepsilon = 0.4$ and $\mathbf{w} = \{1, 1, 1\}$, the pseudospectrum $\sigma_{\varepsilon, \mathbf{w}}(P)$ is bounded and its boundary is drawn in Figure 1, where the eigenvalues of $P(\lambda)$ are plotted as '+' and the point $0 \notin \sigma_{\varepsilon, \mathbf{w}}(P)$ is marked with an asterisk. For the distance $\text{dist}(0, \sigma_{\varepsilon, \mathbf{w}}(P))$, we

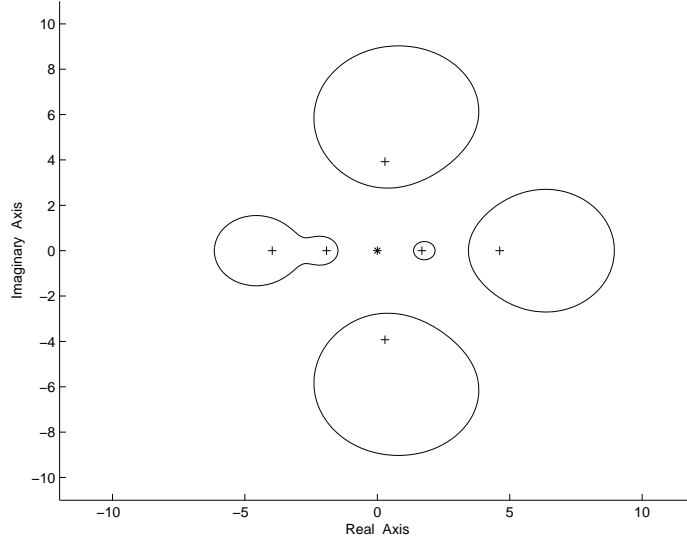


Figure 1: The pseudospectrum $\sigma_{0.4, \mathbf{w}}(P)$ with five connected components.

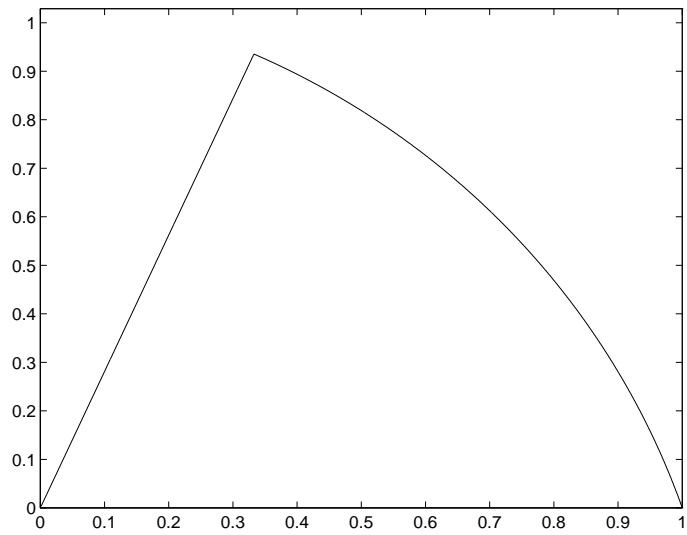


Figure 2: The function $\min\{\gamma r_1, r_\gamma\}$ for $\gamma \in (0, 1)$.

verify that $r_1 = 2.8113$ and, as one can see in Figure 2, the best lower bound that Theorem 4 (i) can imply is 0.9355 (which corresponds to $\gamma = 0.3328$). This bound is satisfactory, keeping in mind that the closest to the origin real boundary point of $\sigma_{\varepsilon, \mathbf{w}}(P)$ is 1.3686. \square

Theorem 4, the relative discussion and straightforward calculations yield the following result.

Corollary 5 *Let $Q(\lambda) = A_1\lambda + A_0$ be a linear pencil with $\det A_1 \neq 0$. Then for any $\lambda_0 \notin \sigma_{\varepsilon, \mathbf{w}}(Q)$, we have*

$$\text{dist}(\lambda_0, \sigma_{\varepsilon, \mathbf{w}}(Q)) \geq \frac{s_{\min}(A_1\lambda_0 + A_0) - \varepsilon(w_1|\lambda_0| + w_0)}{\|A_1\|_2 + \varepsilon w_1}.$$

We remark that for $A_1\lambda + A_0 = I\lambda - A$ and $\mathbf{w} = \{1, 0\}$, the above corollary implies directly Theorem 1.

Suppose now that $w_m > 0$. If the magnitude of λ_0 is sufficiently large, then the quantity $q_{\mathbf{w}}^{(j)}(|\lambda_0|)/(j!)$ can be approximated by $\binom{m}{j} w_m |\lambda_0|^{m-j}$ for every $j = 0, 1, \dots, m$. Furthermore, $s_{\min}(P(\lambda_0))$ can be estimated by $s_{\min}(A_m)|\lambda_0|^m$ (which is positive since $\det A_m \neq 0$). As a consequence, (4) is approximated by the inequality

$$\binom{m}{m} w_m |\mu|^m + \dots + \binom{m}{1} w_m |\lambda_0|^{m-1} |\mu| - \left(\frac{s_{\min}(A_m)}{\varepsilon} - w_m \right) |\lambda_0|^m < 0,$$

where $s_{\min}(A_m)/\varepsilon - w_m > 0$ if and only if $\sigma_{\varepsilon, \mathbf{w}}(P)$ is bounded [7, Theorem 2.2]. Dividing by $|\lambda_0|^m$, it follows

$$\binom{m}{m} \frac{w_m}{|\lambda_0|^m} |\mu|^m + \dots + \binom{m}{1} \frac{w_m}{|\lambda_0|} |\mu| - \left(\frac{s_{\min}(A_m)}{\varepsilon} - w_m \right) < 0,$$

where all the positive coefficients of the (positive) powers of $|\mu|$ are relatively small. Hence, we conclude that if one of the weights w_1, w_2, \dots, w_m is positive, $\sigma_{\varepsilon, \mathbf{w}}(P)$ is bounded and $|\lambda_0|$ is sufficiently large, then r_1 in Theorem 4 becomes relatively large. In particular, it becomes proportional to $|\lambda_0|$.

References

- [1] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York (1982).
- [2] M. Embree and L.N. Trefethen, Generalizing eigenvalue theorems to pseudospectra theorems, *SIAM J. Sci. Comput.* **23** (2001) 583-590.
- [3] M. Embree and L.N. Trefethen, *Spectra and Pseudospectra: The Behavior of Non-normal Matrices and Operators*, Princeton University Press (2005).

- [4] M. Embree and L.N. Trefethen, *Pseudospectra Gateway*, <http://www.comlab.ox.ac.uk/pseudospectra/>.
- [5] L.N. Trefethen, Pseudospectra of linear operators, *SIAM Rev.* **39** (1997) 383-406.
- [6] F. Tisseur and N. Higham, Structured pseudospectra for polynomial eigenvalue problems with applications, *SIAM J. Matrix Anal. Appl.* **23** (2001) 187-208.
- [7] P. Lancaster and P. Psarrakos, On the pseudospectra of matrix polynomials, *SIAM J. Matrix Anal. Appl.* **27** (2005) 115-129.
- [8] I. Koutis and E. Gallopoulos, Exclusion regions and fast estimation of pseudospectra, *Technical Report*, Department of Computer Engineering and Informatics, HP-CLAB, University of Patras, Patras, Greece (2000).
- [9] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge (1991).