An improved grid method for the computation of the pseudospectra of matrix polynomials*

Stavros Fatouros† and Panayiotis Psarrakos‡
Department of Mathematics, National Technical University of Athens
Zografou Campus, 15780 Athens, Greece

May 27, 2008

Abstract

Pseudospectra of matrix polynomials have been systematically investigated in the last years, since they provide important insights into the sensitivity of polynomial eigenvalue problems. An accurate approximation of the pseudospectrum of a matrix polynomial $P(\lambda)$ by means of the standard grid method is computationally high demanding. In this paper, we propose an improvement of the grid method, which reduces the computational cost and retains the robustness and the parallelism of the method. In particular, after giving two lower bounds for the distance from a point to the boundary of the pseudospectrum of $P(\lambda)$, we present two algorithms for the estimation of the pseudospectrum, using exclusion discs. Furthermore, two illustrative examples and an application of pseudospectra on elliptic (quadratic) eigenvalue problems are given.

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

AMS Subject Classifications: 65F15; 65F35

1 Introduction and preliminaries

Consider the *matrix polynomial*

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

where λ is a complex variable and $A_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ with $\det A_m \neq 0$. The study of matrix polynomials has a long history, especially with regard to their spectral analysis, which leads to the solutions of higher order linear systems of differential equations [9, 19].

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

*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%).

†E-mail: s.fatouros@city.ac.uk.

‡E-mail: ppsarr@math.ntua.gr. Corresponding author.

as an *eigenvector* of $P(\lambda)$ corresponding to λ_0. The set of all eigenvalues of $P(\lambda)$ is the *spectrum* of $P(\lambda)$, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, and since $\det A_m \neq 0$, it contains no more than nm distinct (finite) elements.

We are interested in the spectra of perturbations of $P(\lambda)$ 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 $\Delta_0, \Delta_1, \ldots, \Delta_m \in \mathbb{C}^{n\times n}$ are arbitrary. For a given $\varepsilon > 0$ and a given set of nonnegative weights $\mathbf{w} = \{w_0, w_1, \ldots, w_m\}$ with at least one nonzero element, we define the set of perturbed matrix polynomials

$$\mathcal{B}(P, \varepsilon, \mathbf{w}) \;=\; \{P_\Delta(\lambda) : \|\Delta_j\|_2 \leq \varepsilon\, w_j,\; j = 0, 1, \ldots, m\}\,.$$

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

The (weighted) *ε-pseudospectrum* of $P(\lambda)$ (introduced by Tisseur and Higham [18]) is then defined by

$$\begin{aligned}\sigma_{\varepsilon,\mathbf{w}}(P) &= \{\lambda \in \mathbb{C} : \det P_\Delta(\lambda) = 0,\; P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})\}\\ &= \{\lambda \in \mathbb{C} : \det P_\Delta(\lambda) = 0,\; \|\Delta_j\|_2 \leq \varepsilon\, w_j,\; j = 0, 1, \ldots, m\}\,.\end{aligned}$$

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, $\mathbf{w} = \{w_0, w_1\} = \{1, 0\}$, then $\sigma_{\varepsilon,\mathbf{w}}(P)$ coincides with the well understood *ε-pseudospectrum* of the matrix A [5, 6, 7, 20], that is, $\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : \lambda \in \sigma(A + E),\, \|E\|_2 \leq \varepsilon\}$.

For any matrix $A \in \mathbb{C}^{n\times n}$, we denote by $s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A) \geq 0$ the singular values of A. If we consider the scalar polynomial

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

then by [18, Lemma 2.1],

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

Thus, by the continuity of the eigenvalues of $P_\Delta(\lambda)$ with respect to the entries of the coefficient matrices, it follows that the boundary of $\sigma_{\varepsilon,\mathbf{w}}(P)$, $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$, satisfies

$$\partial\sigma_{\varepsilon,\mathbf{w}}(P) \;\subseteq\; \{\lambda \in \mathbb{C} : s_n(P(\lambda)) = \varepsilon\, w(|\lambda|)\}\,. \tag{3}$$

As the parameter $\varepsilon > 0$ increases, the ε-pseudospectrum of $P(\lambda)$ enlarges, and for ε large enough, $\sigma_{\varepsilon,\mathbf{w}}(P)$ may become unbounded. Moreover, if there is a $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$ with identically zero determinant, then $\sigma_{\varepsilon,\mathbf{w}}(P)$ coincides with the complex plane. 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, $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded if and only if $\varepsilon\, w_m < s_n(A_m)$, and in this case, it has no more than nm connected components [15].

Pseudospectra provide important insights into the sensitivity of eigenvalues under perturbations and have several applications (see [3, 5, 6, 7, 15, 18, 20] and the references therein). Hence, the problem of their computation is of special interest. Tisseur and Higham [18] use the standard grid method (`GRID`) and formula (3) for the estimation of the ε-pseudospectrum of a matrix polynomial. In particular, they obtain a graphical representation of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ by evaluating $s_n(P(\lambda))$ on a predefined grid of points in the complex plane and then plotting the ε-contours. This simple method is robust and gain large parallelism, but it is also too costly since it is based on a two dimensional grid.

In [15], Lancaster and Psarrakos propose a path-following algorithm for the numerical determination of the boundary of $\sigma_{\varepsilon,\mathbf{w}}(P)$, extending the work of Brühl (for matrices) [4]. This alternative methodology has been shown to be effective, and since it is essentially based on an one dimensional grid, it is much less demanding in terms of the number of points λ at which it is necessary to compute $s_n(P(\lambda))$. As might be expected, difficulties may appear near singular points of the boundary $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ and near points where the distance between distinct connected components of $\sigma_{\varepsilon,\mathbf{w}}(P)$ becomes relatively small. Furthermore, `GRID` handles several ε's at once, while one has to repeat the path-tracing algorithm of [15] for each connected component and each ε. Thus, the standard grid method still appears to offer greater robustness and more parallelism.

Several techniques for the estimation of pseudospectra of matrices can be found in the literature [1, 2, 4, 6, 7, 8, 10, 14, 21]. In this article, motivated by a recent paper of Koutis and Gallopoulos [14] (which can be downloaded from [7]), we present a modified grid method that reduces drastically the number of points λ where the evaluation of $s_n(P(\lambda))$ is required, reducing the computational cost, and also retains the robustness and the parallelism of the method. Thus, despite its apparent simplicity, it can be considered as an effective tool for computing pseudospectra. In Section 2, we give two lower bounds for the distance from a given point $\lambda_0 \notin \partial\sigma_{\varepsilon,\mathbf{w}}(P)$ to the boundary $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$ (one bound for exterior points and one for interior points). In Section 3, we propose two algorithms for the estimation of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$, using our theoretical results and appropriate exclusion discs. Moreover, illustrative examples with several comparisons and an application of pseudospectra on elliptic eigenvalue problems are presented in Sections 4 and 5, respectively.

2 Two distance lower bounds

In [14], Koutis and Gallopoulos propose a simple inclusion-exclusion algorithm for the estimation of pseudospectra of complex matrices, which is based on the following result (see also [5, 6]).

Theorem 1 [14, Theorem 2.4] *Let $A \in \mathbb{C}^{n\times n}$, $\varepsilon > 0$ and $\lambda_0 \notin \sigma_\varepsilon(A)$. Then the distance* $\operatorname{dist}(\lambda_0, \sigma_\varepsilon(A))$ *from the point λ_0 to the ε-pseudospectrum of A satisfies* $\operatorname{dist}(\lambda_0, \sigma_\varepsilon(A)) \geq s_n(I\lambda_0 - A) - \varepsilon$.

Consider now an $n \times n$ matrix polynomial $P(\lambda)$ as in (1), an $\varepsilon > 0$, some weights $w_0, w_1, \ldots, w_m \geq 0$ (with at least one of them nonzero) and the associated polynomial

$w(\lambda)$ in (2). Theorem 1 has been generalized by Psarrakos [17]. Here (and elsewhere), the indices denote the derivatives of the polynomials.

Theorem 2 [17, Theorem 4] *Let* $\lambda_0 \notin \sigma_{\varepsilon,\mathbf{w}}(P)$, *and let* r_1 *be the positive root of the equation*

$$\frac{w^{(m)}(|\lambda_0|)}{m!}\,\lambda^m + \cdots + \frac{w^{(1)}(|\lambda_0|)}{1!}\,\lambda - \left(\frac{s_n(P(\lambda_0))}{\varepsilon} - w(|\lambda_0|)\right) \,=\, 0, \tag{4}$$

assuming that $r_1 = +\infty$ *whenever this equation has no positive roots. For any* $\gamma \in (0,1)$, *let* ρ_γ *be the positive root of*

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}\,\lambda^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}\,\lambda - (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| + \gamma r_1)) \,=\, 0. \tag{5}$$

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

We remark that $r_1 = +\infty$ if and only if $w_1 = w_2 = \cdots = w_m = 0$ and $w_0 > 0$. In this case, (5) is written

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

and does not depend on γ.

Suppose that at least one of the weights $w_1, w_2, \ldots, w_m$ is positive (or equivalently, that $r_1 < +\infty$). As mentioned in [17], ρ_γ is a continuous decreasing function of the variable $\gamma \in (0,1)$, with $\lim_{\gamma\to 1^-} \rho_\gamma = 0$. Consequently, the curve $\{(\gamma, \rho_\gamma) : \gamma \in (0,1)\}$ 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 $\min\{\gamma r_1, \rho_\gamma\}$. If this common point is $(\gamma_0, \rho_{\gamma_0}) = (\gamma_0, \gamma_0 r_1)$ (for a $\gamma_0 \in (0,1)$), then $\rho_{\gamma_0} = \gamma_0 r_1$ is the best lower bound Theorem 2 can imply. Furthermore, if $w_m > 0$, $\sigma_{\varepsilon,\mathbf{w}}(P)$ is bounded and $|\lambda_0|$ is sufficiently large, then r_1 becomes relatively large. In particular, it becomes proportional to $|\lambda_0|$ [17].

By Theorem 2 and straightforward calculations, we see that for the linear pencil $P(\lambda) = A_1\lambda + A_0$ with $\det A_1 \neq 0$, and for any $\lambda_0 \notin \sigma_{\varepsilon,\mathbf{w}}(P)$,

$$\operatorname{dist}(\lambda_0, \sigma_{\varepsilon,\mathbf{w}}(P)) \,\geq\, \frac{s_n(A_1\lambda_0 + A_0) - \varepsilon\,(w_1|\lambda_0| + w_0)}{\|A_1\|_2 + \varepsilon\, w_1}\,.$$

Moreover, for $A_1\lambda + A_0 = I\lambda - A$ and $\mathbf{w} = \{1, 0\}$, Theorem 1 follows immediately.

It is known that for any point λ_0 of the ε-pseudospectrum of a matrix $A \in \mathbb{C}^{n\times n}$, there is an $E \in \mathbb{C}^{n\times n}$ such that $\lambda_0 \in \sigma(A+E)$ and $\|E\|_2 = s_n(I\lambda_0 - A) \leq \varepsilon$ [6, 7]. Moreover, for any matrix $\Delta \in \mathbb{C}^{n\times n}$ with $\|\Delta\|_2 \leq \varepsilon - s_n(I\lambda_0 - A)$, it is clear that $\|E + \Delta\|_2 \leq \|E\|_2 + \|\Delta\|_2 \leq \varepsilon$, and thus, $\sigma(A + E + \Delta) \subset \sigma_\varepsilon(A)$. Choosing $\Delta = zI$ ($z \in \mathbb{C}$, $0 \leq |z| \leq \varepsilon - s_n(I\lambda_0 - A)$), one can easily verify the next proposition, which is, in some sense, complementary to Theorem 1.

Proposition 3 *Let* $A \in \mathbb{C}^{n\times n}$, $\varepsilon > 0$ *and* $\lambda_0 \in \operatorname{Int}[\sigma_\varepsilon(A)]$, *the interior of the pseudospectrum. Then the distance from the point* λ_0 *to the boundary* $\partial\sigma_\varepsilon(A)$ *satisfies* $\operatorname{dist}(\lambda_0, \partial\sigma_\varepsilon(A)) \,\geq\, \varepsilon - s_n(I\lambda_0 - A)$.

Next we generalize this proposition to the case of matrix polynomials, borrowing arguments from the proof of [17, Theorem 4]. As in [17], the following two lemmas are necessary for our discussion. The first lemma can be found in [13], and the second one is a simple exercise in scalar polynomials.

Lemma 4 *For any* $A, B \in \mathbb{C}^{n\times n}$, $|s_n(A+B) - s_n(A)| \leq s_1(B)$.

Lemma 5 *Let* $p(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda - a_0$ *be a scalar polynomial with* $a_0 > 0$, $a_1, a_2, \ldots, a_m \geq 0$ *and at least one of the coefficients* $a_1, a_2, \ldots, a_m$ *positive. Then* $p(\lambda)$ *has exactly one positive zero.*

Theorem 6 *For any* $\lambda_0 \in \mathrm{Int}[\sigma_{\varepsilon,\mathbf{w}}(P)]$ *with* $s_n(P(\lambda_0)) < \varepsilon\, w(|\lambda_0|)$, *let* r *be the smallest positive root of the equation*

$$\frac{w^{(m)}(|\lambda_0|)}{m!}(-1)^m\,\lambda^m + \cdots + \frac{w^{(1)}(|\lambda_0|)}{1!}(-1)\,\lambda - \left(\frac{s_n(P(\lambda_0))}{\varepsilon} - w(|\lambda_0|)\right) = 0 \quad (6)$$

(assuming that $r = +\infty$ *whenever this equation has no positive roots), and let* $r_2 = \min\{r, |\lambda_0|\}$. *For any* $\delta \in (0,1)$, *let* ρ_δ *be the positive root of*

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}\,\lambda^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}\,\lambda + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| - \delta r_2)) = 0. \quad (7)$$

Then $\mathrm{dist}(\lambda_0, \partial\sigma_{\varepsilon,\mathbf{w}}(P)) \geq \min\{\delta r_2, \rho_\delta\}$.

Proof Suppose that $\lambda_0 \in \mathrm{Int}[\sigma_{\varepsilon,\mathbf{w}}(P)]$ with $s_n(P(\lambda_0)) < \varepsilon\, w(|\lambda_0|)$. Then for any nonzero $\mu \in \mathbb{C}$, we have

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

where the leading coefficient $P^{(m)}(\lambda_0)/(m!) = A_m$ is nonsingular. By Lemma 4 and norm properties, it follows

$$\begin{aligned} |s_n(P(\lambda_0+\mu)) - s_n(P(\lambda_0))| &\leq s_1\left(\sum_{j=1}^{m}\frac{P^{(j)}(\lambda_0)}{j!}\,\mu^j\right) \\ &\leq \sum_{j=1}^{m}\frac{\|P^{(j)}(\lambda_0)\|_2}{j!}\,|\mu|^j. \end{aligned}$$

As a consequence,

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

or equivalently,

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

For a moment, assume that $|\mu| \leq |\lambda_0|$. Then we observe that for

$$\varepsilon \;>\; \frac{1}{w(|\lambda_0| - |\mu|)} \left(\sum_{j=1}^{m} \frac{\|P^{(j)}(\lambda_0)\|_2}{j!} \, |\mu|^j + s_n(P(\lambda_0)) \right),$$

or equivalently, for

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}|\mu|^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}|\mu| + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| - |\mu|)) \;<\; 0, \quad (8)$$

we have $s_n(P(\lambda_0+\mu)) < \varepsilon\, w(|\lambda_0|-|\mu|) \leq \varepsilon\, w(|\lambda_0+\mu|)$, i.e., $\lambda_0+\mu$ lies in the interior of $\sigma_{\varepsilon,\mathbf{w}}(P)$. Furthermore, observe that the difference $s_n(P(\lambda_0))-\varepsilon\, w(|\lambda_0|-|\mu|)$ (in the constant coefficient of the scalar polynomial in the left-hand part of (8)) is negative if and only if

$$\frac{w^{(m)}(|\lambda_0|)}{m!}(-1)^m|\mu|^m + \cdots + \frac{w^{(1)}(|\lambda_0|)}{1!}(-1)|\mu| - \left(\frac{s_n(P(\lambda_0))}{\varepsilon} - w(|\lambda_0|)\right) \;>\; 0. \quad (9)$$

Next we consider two cases:

(i) Suppose that at least one of the weights $w_1, w_2, \ldots, w_m$ is positive. Since $s_n(P(\lambda_0)) < \varepsilon\, w(|\lambda_0|)$, the polynomial

$$q(\lambda) \;=\; \frac{w^{(m)}(|\lambda_0|)}{m!}\,(-1)^m\lambda^m + \cdots + \frac{w^{(1)}(|\lambda_0|)}{1!}(-1)\,\lambda - \left(\frac{s_n(P(\lambda_0))}{\varepsilon} - w(|\lambda_0|)\right)$$

satisfies $q(0) > 0$. Let r be the smallest positive zero of this polynomial, assuming that $r = +\infty$ whenever the polynomial has no positive zeros. Then for every nonzero $\mu \in \mathbb{C}$ with $|\mu| < r_2$ $(= \min\{r, |\lambda_0|\})$, (9) holds and

$$s_n(P(\lambda_0)) \;<\; \varepsilon\, w(|\lambda_0| - |\mu|) \;\leq\; \varepsilon\, w(|\lambda_0 + \mu|).$$

Hence, for any $\delta \in (0,1)$,

$$s_n(P(\lambda_0)) \;<\; \varepsilon\, w(|\lambda_0| - \delta r_2),$$

and consequently, the scalar polynomial

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}\lambda^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}\lambda + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| - \delta r_2))$$

satisfies the conditions of Lemma 5 and has exactly one positive zero, ρ_δ. Furthermore, for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < \min\{\delta r_2, \rho_\delta\}$, we have

$$\begin{aligned}
&\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}|\mu|^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}|\mu| + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0 + \mu|)) \\
&\leq\; \frac{\|P^{(m)}(\lambda_0)\|_2}{m!}|\mu|^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}|\mu| + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| - |\mu|)) \\
&<\; \frac{\|P^{(m)}(\lambda_0)\|_2}{m!}|\mu|^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}|\mu| + (s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0| - \delta r_2)) \\
&<\; 0.
\end{aligned}$$

Thus, for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < \min\{\delta r_2, \rho_\delta\}$, both (8) and (9) hold, and as a consequence, $\lambda_0 + \mu \in \mathrm{Int}[\sigma_{\varepsilon,\mathbf{w}}(P)]$.

(ii) Suppose that $w_1 = w_2 = \cdots = w_m = 0$ and $w_0 > 0$. Then $w(\lambda) = w_0$ for every $\lambda \in \mathbb{C}$. Hence, for every $\mu \in \mathbb{C}$, the difference $s_n(P(\lambda_0)) - \varepsilon\, w(|\lambda_0 + \mu|) = s_n(P(\lambda_0)) - \varepsilon\, w_0$ is negative, i.e., we may assume that $r = +\infty$. The scalar polynomial in the left-hand side of (7) is of the form

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}\,\lambda^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}\,\lambda + (s_n(P(\lambda_0)) - \varepsilon\, w_0)\,,$$

satisfies the conditions of Lemma 5, and has exactly one positive zero, ρ. As in case (i), for every nonzero $\mu \in \mathbb{C}$ such that $|\mu| < \rho$,

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}|\mu|^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}|\mu| + (s_n(P(\lambda_0)) - \varepsilon\, w_0) \;<\; 0,$$

i.e., (8) holds. Thus, $s_n(P(\lambda_0 + \mu)) < \varepsilon\, w_0$, or equivalently, the point $\lambda_0 + \mu$ belongs to the interior of $\sigma_{\varepsilon,\mathbf{w}}(P)$. □

We remark that in part (ii) of the above proof, i.e., when $w_1 = w_2 = \cdots = w_m = 0$ and $w_0 > 0$, we have $r = +\infty$ and $|\lambda_0|$ is removed from the construction of the bound. Moreover, (7) is written

$$\frac{\|P^{(m)}(\lambda_0)\|_2}{m!}\,\lambda^m + \cdots + \frac{\|P^{(1)}(\lambda_0)\|_2}{1!}\,\lambda + (s_n(P(\lambda_0)) - \varepsilon\, w_0) \;=\; 0,$$

and does not depend on δ and r_2. It is also clear that for $P(\lambda) = I\lambda - A$ and $\mathbf{w} = \{1, 0\}$, this part of the proof yields directly Proposition 3.

If $r < +\infty$, then as in Theorem 2, ρ_δ is a continuous decreasing function of $\delta \in (0, 1)$ with $\lim_{\gamma \to 1^-} \rho_\delta = 0$. Hence, the curve $\{(\delta, \rho_\delta) : \delta \in (0, 1)\}$ has exactly one common point with the line segment $\{(\delta, \delta r_2) : \delta \in (0, 1)\}$, say $(\delta_0, \rho_{\delta_0}) = (\delta_0, \delta_0 r_2)$, which is the only maximum of $\min\{\delta r_2, \rho_\delta\}$, and $\rho_{\delta_0} = \delta_0 r_2$ is the best lower bound Theorem 6 can give.

Example 1 The boundaries of the ε-pseudospectra of the matrix polynomial

$$H(\lambda) \;=\; \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \lambda^2 + \begin{bmatrix} 1.75 & 0 & 0 \\ 0 & 7.5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \lambda + \begin{bmatrix} 3.5 & 1 & 0 \\ 1 & 8 & 1 \\ 0 & 1 & 4 \end{bmatrix},$$

for $\varepsilon = 0.15,\ 0.3,\ 0.5,\ 0.6,\ 0.8$ and $\mathbf{w} = \{1, 1, 1\}$, are drawn in Figure 1. The eigenvalues of $H(\lambda)$, $-0.4710 \pm \mathrm{i}\,1.2448, -1.1794 \pm \mathrm{i}\,1.0335, -0.6621 \pm \mathrm{i}\,0.7491$, are plotted in the figure as '+'. Clearly, the origin (marked with an asterisk) does not belong to $\sigma_{0.8,\mathbf{w}}(H)$. For the distance $\mathrm{dist}(0, \sigma_{0.8,\mathbf{w}}(H))$, we verify that $r_1 = 1.3080$. Furthermore, Theorem 2 and the bisection method (after 12 iterations) yield the lower bound 0.3229, which corresponds to $\gamma_0 = 0.2468$.

Next we consider the point $\lambda_0 = -1.5 + \mathrm{i}$ (also marked with an asterisk), which is an interior point of $\sigma_{0.8,\mathbf{w}}(H)$ and satisfies $|\lambda_0| = 1.8028$. For the distance $\mathrm{dist}(-1.5 + \mathrm{i}, \partial\sigma_{0.8,\mathbf{w}}(H))$, we see that $r_2 = r = 1.1110$. Theorem 6 and the bisection method (after 12 iterations) imply the lower bound 0.2515, which corresponds to $\delta_0 = 0.2263$. □

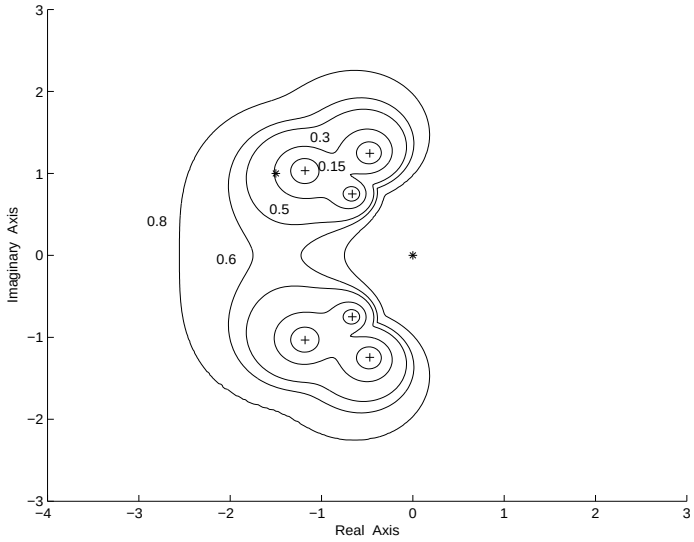

Figure 1: The pseudospectra $\sigma_{\varepsilon,\mathbf{w}}(H)$ for $\varepsilon = 0.15,\ 0.3,\ 0.5,\ 0.6,\ 0.8$.

3 Numerical algorithms

For any $\lambda_0 \in \mathbb{C}$ such that $s_n(P(\lambda_0)) \neq \varepsilon\, w(|\lambda_0|)$, Theorems 2 and 6 yield an open disc that does not intersect the boundary of the pseudospectrum $\sigma_{\varepsilon,\mathbf{w}}(P)$. As a consequence, for the estimation of $\partial\sigma_{\varepsilon,\mathbf{w}}(P)$, we can use a simple modification of `GRID` applied to a predefined grid of points in a region Ω that contains $\sigma_{\varepsilon,\mathbf{w}}(P)$ (or a part of it). In particular, when evaluating $s_n(P(z)) \neq \varepsilon\, w(|z|)$ at a grid point $z \in \Omega$, we set $s_n(P(z'))/w(|z'|) := s_n(P(z))/w(|z|)$ for every grid point z' in the associated exclusion disc

$$\mathcal{S}(z, R) \,=\, \{\lambda \in \mathbb{C} : |\lambda - z| < R\}\,,$$

where R is an appropriate lower bound from Theorem 2 or Theorem 6.

Our methodology is illustrated by the following algorithm, where we consider several values of the parameter ε, namely,

$$0 \,<\, \varepsilon_1 \,<\, \varepsilon_2 \,<\, \cdots \,<\, \varepsilon_k,$$

and combine Theorems 2 and 6. In particular, we use exterior exclusion discs for the pseudospectrum $\sigma_{\varepsilon_k,\mathbf{w}}(P)$, and interior exclusion discs for $\sigma_{\varepsilon_1,\mathbf{w}}(P)$. For the closure of the set $\sigma_{\varepsilon_k,\mathbf{w}}(P)\backslash\sigma_{\varepsilon_1,\mathbf{w}}(P)$, we apply the standard `GRID`.

Algorithm IGRID1

Input: The coefficients $A_0, A_1, \ldots, A_m$ of the matrix polynomial $P(\lambda)$,
the parameters $0 < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_k$,
the weights $w_0, w_1, \ldots, w_m \geq 0$,
the initial region $\Omega = [x_{\min}, x_{\max}] \times [\mathrm{i}\, y_{\min}, \mathrm{i}\, y_{\max}]$,
the lengths $h_x, h_y > 0$ of the grids of $[x_{\min}, x_{\max}]$ and $[\mathrm{i}\, y_{\min}, \mathrm{i}\, y_{\max}]$, respectively.

Step I Construct the grid of the rectangle $\Omega = [x_{\min}, x_{\max}] \times [\mathrm{i}\, y_{\min}, \mathrm{i}\, y_{\max}]$ that corresponds to the lengths h_x and h_y, and for every grid point $z = x + \mathrm{i}\, y$, set the initial value $s_n(P(z))/w(|z|) := 0$.

Step II For every grid point $z = x + \mathrm{i}\, y$ with $s_n(P(z))/w(|z|) = 0$, repeat the following:

(a) Compute the matrix $P(z)$, and the values $w(|z|)$ and $s_n(P(z))/w(|z|)$.

(b) While $s_n(P(z))/w(|z|) > \varepsilon_k$, repeat:

1. Compute the positive root r_1 of the equation (4).

2. Using the bisection method and computing the positive root r_γ of (5) for the chosen values of γ, estimate a $\gamma_0 \in (0,1)$ such that $r_{\gamma_0} \cong \gamma_0 r_1$. Then set $R := r_{\gamma_0}$.

(c) While $s_n(P(z))/w(|z|) < \varepsilon_1$, repeat:

1. Compute the minimum positive root r of the equation (6), and set $r_2 = \min\{r, |z|\}$.

2. Using the bisection method and computing the positive root r_δ of (7) for the chosen values of δ, estimate a $\delta_0 \in (0,1)$ such that $r_{\delta_0} \cong \delta_0 r_2$. Then set $R := r_{\delta_0}$.

(d) If $s_n(P(z))/w(|z|) > \varepsilon_k$ or $s_n(P(z))/w(|z|) < \varepsilon_1$, then for every grid point $z' = x' + \mathrm{i}\, y'$ in the open disc $\mathcal{S}(z, R)$ with $s_n(P(z'))/w(|z'|) = 0$, set $s_n(P(z'))/w(|z'|) := s_n(P(z))/w(|z|)$.

STEP III Plot the curves $\{\lambda \in \mathbb{C} : s_n(P(\lambda))/w(|\lambda|) = \varepsilon_j\} \cap \Omega,\ j = 1, 2, \ldots, k$.

In many cases, the radii of the exclusion discs centered at interior points of $\sigma_{\varepsilon_1,\mathbf{w}}(P)$ are relatively small. As a consequence, the construction of these discs may increase unnecessarily the cost. An alternative algorithm to IGRID1 follows by removing (c) in Step II, and replacing (d) of the same step by the following:

(d') If $s_n(P(z))/w(|z|) > \varepsilon_k$, then for every grid point $z' = x' + \mathrm{i}\, y'$ in the open disc $\mathcal{S}(z, R)$ with $s_n(P(z'))/w(|z'|) = 0$, set $s_n(P(z'))/w(|z'|) := s_n(P(z))/w(|z|)$.

Here, we evaluate $s_n(P(z))/w(|z|)$ at every grid point $z \in \sigma_{\varepsilon_k,\mathbf{w}}(P)$ and do not compute exclusion discs in $\mathrm{Int}[\sigma_{\varepsilon_1,\mathbf{w}}(P)]$. This algorithm, called IGRID2, results almost the same figures with GRID and IGRID1, is much cheaper than GRID when the area of $\Omega \backslash \sigma_{\varepsilon_k,\mathbf{w}}(P)$ is sufficiently larger than the area of $\sigma_{\varepsilon_k,\mathbf{w}}(P)$, and is cheaper than IGRID1 when the components of $\sigma_{\varepsilon_1,\mathbf{w}}(P)$ are sufficiently small. All these are clearly confirmed by the examples of the next section.

A question of interest is how one can choose the initial region Ω. It is easy to see that the spectrum of the matrix polynomial $P(\lambda)$ coincides with the spectrum of the $nm \times nm$ *block companion matrix* of $P(\lambda)$, that is,

$$C_P = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & -A_m^{-1}A_2 & \cdots & -A_m^{-1}A_{m-1} \end{bmatrix}.$$

Suppose now that the leading coefficient A_m of $P(\lambda)$ is unperturbed, i.e., that $w_m = 0$, and consider a matrix polynomial $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$. Then the difference of the two associated block companion matrices satisfies

$$\begin{aligned} \|C_P - C_{P_\Delta}\|_2^2 &= \left\| A_m^{-1} \left(\sum_{j=0}^{m-1} \Delta_j \Delta_j^* \right) (A_m^{-1})^* \right\|_2 \\ &\leq \|A_m^{-1}\|_2^2 \sum_{j=0}^{m-1} \|\Delta_j\|_2^2 \leq \frac{\varepsilon^2}{s_n(A_m)^2} \sum_{j=0}^{m-1} w_j^2. \end{aligned}$$

As a consequence, $\sigma_{\varepsilon,\mathbf{w}}(P) \subseteq \sigma_{\varepsilon\,\zeta}(C_P)$, where $\zeta = \sqrt{w_0^2 + \cdots + w_{m-1}^2} \,/\, s_n(A_m)$. If we denote by $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ the minimum and the maximum eigenvalues of a hermitian matrix, respectively, then by Theorem $13\,\varepsilon$ of [5], a region of interest that always contains the (bounded) pseudospectrum $\sigma_{\varepsilon,\mathbf{w}}(P)$ is the rectangle $\Omega_C = [x_{\min}, x_{\max}] \times [\mathrm{i}\, y_{\min}, \mathrm{i}\, y_{\max}]$, where

$$x_{\min} = \lambda_{\min}\left(\frac{C_P + C_P^*}{2} \right) - \varepsilon\,\zeta, \quad x_{\max} = \lambda_{\max}\left(\frac{C_P + C_P^*}{2} \right) + \varepsilon\,\zeta,$$

$$y_{\min} = \lambda_{\min}\left(\frac{C_P - C_P^*}{2\,\mathrm{i}} \right) - \varepsilon\,\zeta \quad \text{and} \quad y_{\max} = \lambda_{\max}\left(\frac{C_P - C_P^*}{2\,\mathrm{i}} \right) + \varepsilon\,\zeta.$$

Keeping in mind that this initial region Ω_C can be relatively large (see the results of [16] on numerical ranges), we can use it even in the case $w_m > 0$, as in Example 3 below.

4 Two numerical examples

The question of comparison of the methods GRID, IGRID1 and IGRID2 applied to matrix polynomials, especially with respect to their costs, arises in a natural way. We present two numerical examples performed in MATLAB 6.5[1] to illustrate our results, compare the algorithms IGRID1 and IGRID2, and verify that they are computationally much less demanding than GRID (as far as the radii of the exclusion discs are greater than the lengths h_x and h_y).

Example 2 The 20×20 matrix polynomial

$$\begin{aligned} P(\lambda) &= A_2\lambda^2 + A_1\lambda + A_0 \\ &= I\lambda^2 + \mathrm{i} \begin{bmatrix} I_{10} & 0 \\ 0 & 5I_{10} \end{bmatrix} \lambda + \begin{bmatrix} 1 & -1 & -1 & \cdots & -1 \\ -1 & 1 & -1 & \cdots & -1 \\ -1 & -1 & 1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \cdots & 1 \end{bmatrix} \end{aligned}$$

[1] Implementations of the algorithms IGRID1 and IGRID2 in Matlab are available at `http://www.math.ntua.gr/~ppsarr/pseinout.m` and `http://www.math.ntua.gr/~ppsarr/pseout.m`, respectively. These Matlab codes have been used in our examples. We remark that, for practical reasons, each exclusion disc has been replaced by a square (with sides parallel to the axes) that is almost inscribed in the disc.

corresponds to a gyroscopic system (see [19] for definitions and properties). We consider perturbations of $P(\lambda)$ measured in the absolute sense, i.e., $\mathbf{w} = \{1, 1, 1\}$. The boundaries of the ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.2, 0.4, 0.6, 0.7, 0.8,$ are drawn in the left part of Figure 2, and the boundary $\partial\sigma_{0.6,\mathbf{w}}(P)$ is sketched in the right part of the figure. In both parts, we have used IGRID1 on a 400×400 grid of the initial region $\Omega = [-25, 25] \times [-\mathrm{i}\,30, \mathrm{i}\,10]$, and the eigenvalues of $P(\lambda)$ (some of them multiple) are marked as "+".

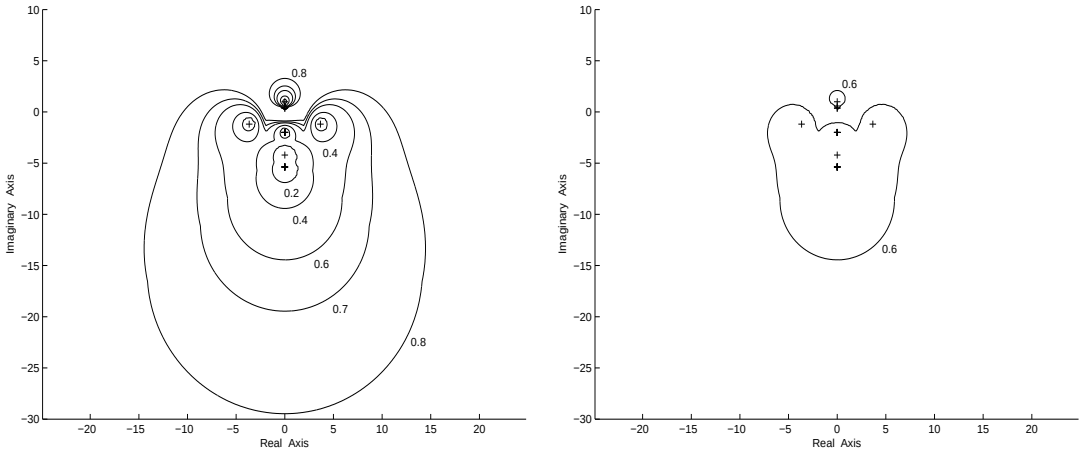

Figure 2: Pseudospectra of a gyroscopic system.

In Table 1, we record the execution time and the number of points λ where the evaluation of $s_{20}(P(\lambda))$ is required for the methods GRID, IGRID1 and IGRID2, and for two grids. For IGRID1 and IGRID2, we also give the number of the exterior points of $\sigma_{0.8,\mathbf{w}}(P)$ (left part) and $\sigma_{0.6,\mathbf{w}}(P)$ (right part) where the minimum singular value is computed. Note that this number is always the same for the two algorithms. The first two rows of the table correspond to the left part of Figure 2 (i.e., for $\varepsilon = 0.2, 0.4, 0.6, 0.7, 0.8$), and the last two rows correspond to the right part of the figure (i.e., for $\varepsilon = 0.6$). The results in the table clearly demonstrate that the new algorithms return (almost) the same approximation of the boundaries of pseudospectra as GRID at a much lower cost, although in the left part of the figure, the area of $\Omega\backslash\sigma_{0.8,\mathbf{w}}(P)$ is not much greater than the area of $\sigma_{0.8,\mathbf{w}}(P)$. Note also that both IGRID1 and IGRID2 behave well when the number of predefined grid points is increased since the number of grid points in each exclusion disc is also increased. As a consequence, the number of grid points λ, where the calculation of $s_{20}(P(\lambda))$ is needed, increases more slowly than the total number of grid points (compare for example the third and the fourth rows of Table 1).

Moreover, for the boundaries in the left part of Figure 2, we observe that IGRID1 and IGRID2 have almost the same computational cost since the area of $\sigma_{0.2,\mathbf{w}}(P)$ is relatively small. In particular, IGRID1 constructs only a few (small) exclusion discs in the interior of $\sigma_{0.2,\mathbf{w}}(P)$, which are not very helpful and require some extra computations. On the other hand, the pseudospectrum $\sigma_{0.6,\mathbf{w}}(P)$ is not so small, and as a consequence, IGRID1 is apparently preferable when drawing $\partial\sigma_{0.6,\mathbf{w}}(P)$ in

Table 1: Cost comparisons for $\sigma_{\varepsilon,\mathbf{w}}(P)$.

	Method GRID	Method IGRID1	Method IGRID2
$\varepsilon = 0.2, 0.4, 0.6, 0.7, 0.8$ 200×200 grid (left part)	27.57 sec 40000 points	11.84 sec 20494 points 5806 ext. points	11.88 sec 20630 points 5806 ext. points
$\varepsilon = 0.2, 0.4, 0.6, 0.7, 0.8$ 400×400 grid (left part)	112.93 sec 160000 points	37.55 sec 71081 points 12201 ext. points	37.75 sec 71749 points 12201 ext. points
$\varepsilon = 0.6$ 200×200 grid (right part)	27.57 sec 40000 points	4.07 sec 3812 points 2826 ext. points	4.42 sec 6107 points 2826 ext. points
$\varepsilon = 0.6$ 400×400 grid (right part)	112.93 sec 160000 points	8.53 sec 7144 points 4960 ext. points	11.48 sec 18151 points 4960 ext. points

the right part of the figure. IGRID1 is faster than IGRID2 since it requires the computation of the minimum singular value at less grid points than IGRID2, due to the exclusion discs in $\mathrm{Int}[\sigma_{0.6,\mathbf{w}}(P)]$. Note also that the area of $\Omega \backslash \sigma_{0.6,\mathbf{w}}(P)$ is quite larger than the one of $\sigma_{0.6,\mathbf{w}}(P)$, and hence, IGRID1 and IGRID2 are much cheaper than GRID. This becomes clear from the last two rows of Table 1. □

Example 3 We consider the 50×50 matrix polynomial

$$\begin{aligned} P(\lambda) &= A_2\lambda^2 + A_1\lambda + A_0 \\ &= I\lambda^2 + tridiag\{-3, 9, -3\}\lambda + tridiag\{-5, 15, -5\}, \end{aligned}$$

which corresponds to a damped mass-spring system described in [18, 19], and we set $\mathbf{w} = \{1, 1, 1\}$. The boundaries of the ε-pseudospectra of $P(\lambda)$ for $\varepsilon = 0.01, 0.05, 0.1,$

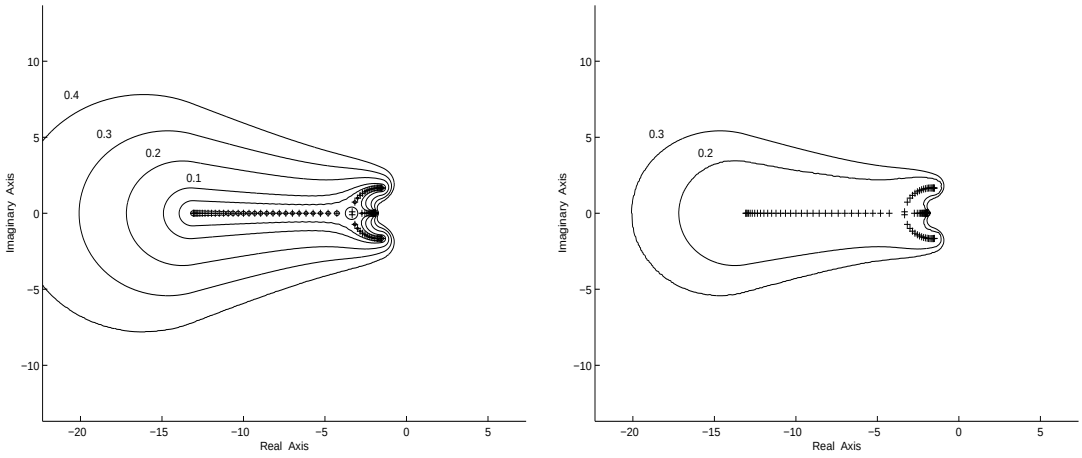

Figure 3: Pseudospectra of a damped mass-spring system.

Table 2: Cost comparisons for $\sigma_{\varepsilon,\mathbf{w}}(P)$.

	Method GRID	Method IGRID1	Method IGRID2
$\varepsilon = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4$ 200×200 grid (left part)	146.61 sec 40000 points	37.45 sec 14927 points 2182 ext. points	35.88 sec 14939 points 2182 ext. points
$\varepsilon = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4$ 400×400 grid (left part)	576.69 sec 160000 points	118.24 sec 54657 points 3878 ext. points	116.80 sec 54832 points 3878 ext. points
$\varepsilon = 0.2, 0.3$ 200×200 grid (right part)	146.61 sec 40000 points	19.52 sec 6293 points 2038 ext. points	26.20 sec 9671 points 2038 ext. points
$\varepsilon = 0.2, 0.3$ 400×400 grid (right part)	576.69 sec 160000 points	51.38 sec 15357 points 3449 ext. points	81.58 sec 33962 points 3449 ext. points

$0.2, 0.3, 0.4,$ are drawn in the left part of Figure 3, and for $\varepsilon = 0.2, 0.3,$ are drawn in the right part of the figure. In both parts, we have used IGRID1 on a 400×400 grid of the initial region

$$\begin{aligned}\Omega_C &= [x_{\min}, x_{\max}] \times [\mathrm{i}\, y_{\min}, \mathrm{i}\, y_{\max}] \\ &= [-21.6342 - 0.4\sqrt{2}, 6.6456 + 0.4\sqrt{2}] \times [-\mathrm{i}\,(12.9905 + 0.4\sqrt{2}), \mathrm{i}\,(12.9905 + 0.4\sqrt{2})] \\ &= [-22.1999, 7.2113] \times [-\mathrm{i}\, 13.5562, \mathrm{i}\, 13.5562]\end{aligned}$$

defined in the previous section. Furthermore, the eigenvalues of $P(\lambda)$ are marked as "+".

In Table 2, we give the execution time and the number of points λ where the evaluation of $s_{50}(P(\lambda))$ is needed for GRID, IGRID1 and IGRID2, and for two grids. For IGRID1 and IGRID2, we also give the number of the (common) exterior points where the minimum singular value is computed. The first two rows of the table correspond to the left part of Figure 2 (i.e., for $\varepsilon = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4$), and the rest two rows correspond to the right part of the figure (i.e., for $\varepsilon = 0.2, 0.3$). As in Example 2, the new algorithms return (almost) the same approximation of the boundaries of pseudospectra as GRID at a much lower cost, and once again, they behave well when the number of predefined grid points is increased.

It is worth noting that for the boundaries in the left part of Figure 2, IGRID2 is a little cheaper than IGRID1 since $\sigma_{0.01,\mathbf{w}}(P)$ consists of several tiny connected components whose total area is quite small. Hence, the exclusion discs constructed by IGRID1, in the interior of $\sigma_{0.01,\mathbf{w}}(P)$, are not beneficial and increase the cost. On the other hand, the pseudospectrum $\sigma_{0.2,\mathbf{w}}(P)$ is not so small, and as a consequence, IGRID1 is preferable when sketching the boundaries of $\sigma_{0.2,\mathbf{w}}(P)$ and $\sigma_{0.3,\mathbf{w}}(P)$ in the right part of the figure. The reason is the exclusion discs in the interior of $\sigma_{0.2,\mathbf{w}}(P)$, as one can conclude from the last two rows of Table 2. □

5 Distance to non-ellipticity

Quadratic eigenvalue problems appear in many applications (for a recent survey on this topic, see [19]). An important class of $n \times n$ quadratic matrix polynomials

$$H(\lambda) \;=\; H_2\lambda^2 + H_1\lambda + H_0 \tag{10}$$

are those in which the coefficient matrices H_0, H_1 and H_2 are hermitian, H_2 is positive definite, and $(x^*H_1x)^2 < 4(x^*H_2x)(x^*H_0x)$ for all nonzero $x \in \mathbb{C}^n$. A matrix polynomial that satisfies these properties is called *elliptic*, and has only non-real eigenvalues (in conjugate pairs). Moreover, H_0 is necessarily positive definite. We remark that by [9, 12], a quadratic matrix polynomial $H(\lambda)$ with hermitian coefficient matrices is elliptic if and only if $\sigma(H) \cap \mathbb{R} = \emptyset$.

For a given elliptic matrix polynomial as in (10), it is natural to ask how much the coefficient matrices must be perturbed, always under hermitian perturbations, for the property of ellipticity to be lost. This distance problem was systematically studied in [11, 12], but without any freedom of independent size perturbations for different coefficient matrices. Here, after defining a new (weighted) distance based on the class

$$\mathcal{B}_h(H,\varepsilon,\mathbf{w}) \;=\; \left\{ H_\Delta(\lambda) = \sum_{j=0}^{2} (H_j + \Delta_j)\lambda^j \,:\, \|\Delta_j\|_2 \le \varepsilon\, w_j,\; \Delta_j^* = \Delta_j,\; j = 0,1,2 \right\},$$

we give a solution of the problem by using pseudospectra and the smallest singular value.

For the elliptic matrix polynomial $H(\lambda)$ in (10), the *distance to non-ellipticity* is denoted and defined by

$$\mathcal{D}(H) \;=\; \min\left\{\varepsilon > 0 : \exists\; H_\Delta(\lambda) \in \mathcal{B}_h(H,\varepsilon,\mathbf{w}) \text{ that is non-elliptic}\right\}.$$

By our discussion on pseudospectra and Proposition 14 of [3], we verify that

$$\begin{aligned}
\mathcal{D}(H) &= \min\left\{\varepsilon > 0 : \exists\; H_\Delta(\lambda) \in \partial\mathcal{B}_h(H,\varepsilon,\mathbf{w}) \text{ that is non-elliptic}\right\} \\
&= \min\left\{\varepsilon > 0 : \exists\; H_\Delta(\lambda) \in \partial\mathcal{B}_h(H,\varepsilon,\mathbf{w}) \text{ with a real eigenvalue}\right\} \\
&\ge \min\left\{\varepsilon > 0 : \sigma_{\varepsilon,\mathbf{w}}(H) \cap \mathbb{R} \ne \emptyset\right\} \\
&= \min\left\{\varepsilon > 0 : \partial\sigma_{\varepsilon,\mathbf{w}}(H) \cap \mathbb{R} \ne \emptyset\right\}.
\end{aligned}$$

Furthermore, if $t_e \in \partial\sigma_{\varepsilon,\mathbf{w}}(H) \cap \mathbb{R}$, and $u_e, v_e \in \mathbb{C}^n$ is a pair of a left and a right singular vectors of the matrix $H(t_e)$ corresponding to $s_n(H(t_e))$ (with respect to the same singular value decomposition), respectively, then the matrix polynomial

$$\begin{aligned}
H_e(\lambda) \;=\; & \left(H_2 - \frac{w_2}{w(|t_e|)}\, s_n(H(t_e))u_ev_e^*\right)\lambda^2 \\
& + \left(H_1 - \frac{w_1}{w(|t_e|)}\,\frac{t_e}{|t_e|}\, s_n(H(t_e))u_ev_e^*\right)\lambda \\
& + H_0 - \frac{w_0}{w(|t_e|)}\, s_n(H(t_e))u_ev_e^*.
\end{aligned} \tag{11}$$

has all its coefficients hermitian, is non-elliptic and lies on $\partial\mathcal{B}_h(H,\varepsilon,\mathbf{w})$. Thus, the next result follows readily.

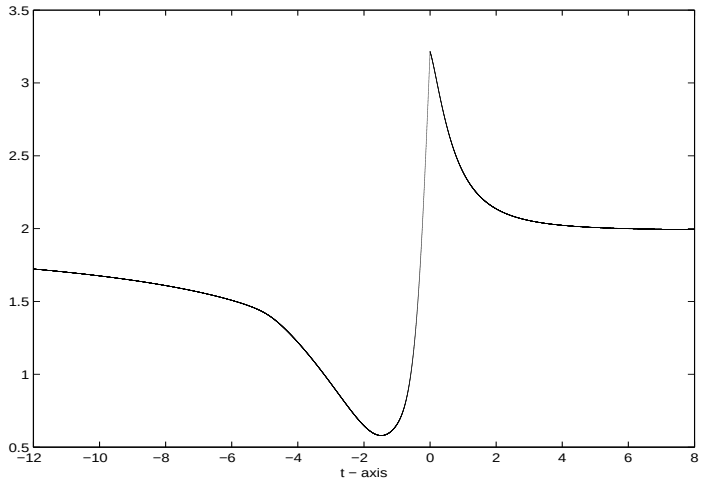

Figure 4: The graph of $s_3(H(t))/w(|t|),\ -12 \le t \le 8$.

Proposition 7 *The distance from an elliptic matrix polynomial $H(\lambda)$ to non-ellipticity is*

$$\mathcal{D}(H) \;=\; \min\left\{\varepsilon > 0 : \sigma_{\varepsilon,\mathbf{w}}(H) \cap \mathbb{R} \neq \emptyset\right\} \;=\; \min_{t\in\mathbb{R}} \frac{s_n(H(t))}{w(|t|)}\,.$$

Moreover, if this minimum is attained at $t_e \in \mathbb{R}$, then a non-elliptic matrix polynomial on $\partial\mathcal{B}_h(H, \mathcal{D}(H), \mathbf{w})$ is given by (11).

The computation of the distance $\mathcal{D}(H)$ is simple and direct. Note also that if $w_m > 0$, then for $t \longrightarrow \pm\infty$,

$$\frac{s_n(H(t))}{w(|t|)} \;\longrightarrow\; \frac{s_n(H_m)}{w_m}\,.$$

The quadratic matrix polynomial $H(\lambda)$ in Example 1 is elliptic (see Example 5 of [12]). By its pseudospectra in Figure 1, it is clear that the distance $\mathcal{D}(H)$ satisfies $0.5 < \mathcal{D}(H) < 0.6$. In Figure 4, we see the graph of the function $s_3(H(t))/w(|t|)$ and confirm that $\lim_{t\to+\infty} s_3(H(t))/w(|t|) = \lim_{t\to-\infty} s_3(H(t))/w(|t|) = s_3(H_2)/w_2 = 2$. Proposition 7 and straightforward calculations imply that $\mathcal{D}(H) = 0.5796$ and the closest to $H(\lambda)$ non-elliptic matrix polynomial given by (11) is

$$\begin{aligned} H_e(\lambda) \;=\; & \begin{bmatrix} 1.9327 & 0.1737 & -0.0656 \\ 0.1737 & 2.5516 & 0.1694 \\ -0.0656 & 0.1694 & 3.9360 \end{bmatrix} \lambda^2 \\ & + \begin{bmatrix} 1.8173 & -0.1737 & 0.0656 \\ -0.1737 & 7.9484 & -0.1694 \\ 0.0656 & -0.1694 & 5.0640 \end{bmatrix} \lambda + \begin{bmatrix} 3.4327 & 1.1737 & -0.0656 \\ 1.1737 & 7.5516 & 1.1694 \\ -0.0656 & 1.1694 & 3.9360 \end{bmatrix} \end{aligned}$$

with real (multiple) eigenvalue $t_e = -1.4785$.

References

[1] C. Becas and E. Gallopoulos, Cobra: parallel path following for computing the matrix pseudospectrum, *Parallel Comput.* **27** (2001) 1879–1896.

[2] C. Becas and E. Gallopoulos, Parallel computation of pseudospectra by fast descent, *Parallel Comput.* **28** (2002) 223–242.

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

[4] M. Brühl, A curve tracing algorithm for computing the pseudospectrum, *BIT* **36** (1996) 441–454.

[5] M. Embree and L.N. Trefethen, Generalizing eigenvalue theorems to pseudospectra theorems, *SIAM J. Sci. Comput.* **23** (2001) 583–590.

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

[7] M. Embree and L.N. Trefethen, *Pseudospectra Gateway*, `http://www.comlab.ox.ac.uk/pseudospectra/`.

[8] E. Gallestey, Computing sperctral value sets using the subharmonicity of the norm of rational matrices, *BIT* **38** (1998) 22–33.

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

[10] S.K. Godunov and M. Sadkane, Computation of pseudospectra via spectral projectors, *Linear Algebra Appl.* **279** (1998) 163–175.

[11] Y. Hachez and P. van Dooren, Elliptic and hyperbolic quadratic eigenvalue problems and associated distance problems, *Linear Algebra Appl.* **371** (2003) 31–44.

[12] N. Higham, F. Tisseur and P. van Dooren, Detecting a definite hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems, *Linear Algebra Appl.* **351/352** (2002) 455–474.

[13] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge (1991).

[14] I. Koutis and E. Gallopoulos, Exclusion and fast estimation of pseudospectra, *Technical Report*, Department of Computer Engineering and Informatics, HPCLAB, University of Patras, Patras, Greece (2000).

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

[16] J. Maroulas and P. Psarrakos, Geometrical properties of numerical range of matrix polynomials, *Computers Math. Applic.* **31** (1996) 42–47.

[17] P. Psarrakos, A distance bound for pseudospectra of matrix polynomials, *Appl. Math. Lett.* **20** (2007) 499–504.

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

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

[20] L.N. Trefethen, Pseudospectra of linear operators, *SIAM Rev.* **39** (1997) 383–406.

[21] L.N. Trefethen, Computation of pseudospectra, *Acta Numerica* **8** (1999) 247–295.