MANCHESTER
1824

The University of Manchester

Tropical roots as approximations to eigenvalues of matrix polynomials

Vanni Noferini, Meisam Sharify and Francoise Tisseur

March 2014

MIMS EPrint: **2014.16**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

ISSN 1749-9097

TROPICAL ROOTS AS APPROXIMATIONS TO EIGENVALUES OF MATRIX POLYNOMIALS*

VANNI NOFERINI[†], MEISAM SHARIFY[‡], AND FRANÇOISE TISSEUR[‡]

Abstract. The tropical roots of $\mathsf{t}_{\times}p(x) = \max_{0\le j\le \ell} \|A_j\|x^j$ are points at which the maximum is attained at least twice. These roots, which can be computed in only $O(\ell)$ operations, can be good approximations to the moduli of the eigenvalues of the matrix polynomial $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$, in particular when the norms of the matrices A_j vary widely. Our aim is to investigate this observation and its applications. We start by providing annuli defined in terms of the tropical roots of $\mathsf{t}_{\times}p(x)$ that contain the eigenvalues of $P(\lambda)$. Our localization results yield conditions under which tropical roots offer order of magnitude approximations to the moduli of the eigenvalues of $P(\lambda)$. Our tropical localization of eigenvalues are less tight than eigenvalue localization results derived from a generalized matrix version of Pellet's theorem but they are easier to interpret. Tropical roots are already used to determine the starting points for matrix polynomial eigensolvers based on scalar polynomial root solvers such as the Ehrlich-Aberth method and our results further justify this choice. Our results provide the basis for analyzing the effect of Gaubert and Sharify's tropical scalings for $P(\lambda)$ on (a) the conditioning of linearizations of tropically scaled $P(\lambda)$ and (b) the backward stability of eigensolvers based on linearizations of tropically scaled $P(\lambda)$. We anticipate that the tropical roots of $\mathsf{t}_{\times}p(x)$, on which the tropical scalings are based, will help designing polynomial eigensolvers with better numerical properties than standard algorithms for polynomial eigenvalue problems such as that implemented in the MATLAB function `polyeig`.

Key words. Polynomial eigenvalue problem, matrix polynomial, tropical algebra, localization of eigenvalues, Rouché's theorem, Pellet's theorem, Newton's polygon, tropical scaling.

AMS subject classifications. 65F15, 15A22, 15A80, 15A18, 47J10.

1. Introduction.

Being able to cheaply locate the eigenvalues of a real or complex $n \times n$ matrix polynomial

$$P(\lambda) = \sum_{i=0}^{\ell} \lambda^i A_i, \quad A_\ell \neq 0, \tag{1.1}$$

is useful in a number of situations, such as, for example, when selecting the starting points in the Ehrlich-Aberth method for the numerical solution of polynomial eigenvalue problems [6], [7], or in choosing the contour in contour integral methods for polynomial eigenvalue problems of large dimensions [3]. Betcke's diagonal scaling [4, Sec. 5], whose aim is to improve the conditioning of P's eigenvalues near a target eigenvalue ω, requires a priori knowledge of the magnitude of ω.

The tropical roots of the tropical (or max-times) polynomial $f(x) = \max_{0\le i\le \ell} a_i x^i$ with $a_i, x \ge 0$ are points (i.e., nonnegative real numbers) at which the maximum is attained at least twice. They are easy and cheap to compute (see Section 2.1). Our aim is to investigate the order of magnitude approximation of the eigenvalues of $P(\lambda)$ in terms of the tropical roots of $\mathsf{t}_{\times}p(x) = \max_{0\le i\le \ell}\left(\|A_i\|x^i\right)$ for some matrix norm $\|\cdot\|$ subordinate to a vector norm.

*Version of April 17, 2014.

† School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK. (`Vanni.Noferini@manchester.ac.uk`). Supported by European Research Council Advanced Grant MATFUN (267526).

‡School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK. (`Meisam.Sharify@manchester.ac.uk`, `Francoise.Tisseur@manchester.ac.uk`). Both authors were supported by Engineering and Physical Sciences Research Council grant EP/I005293.

Gaubert and Sharify [8, Thm. 2] were the first to notice the tropical splitting of the eigenvalues of matrix polynomials. Indeed, for $n \times n$ heavily damped quadratics, i.e., quadratic matrix polynomials $Q(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ with $\|A_1\|^2 \geq \|A_0\|\|A_2\|$, they showed that

$$\operatorname{gap}\big(\Lambda(Q), \Lambda(L)\big) \leq g(\kappa(A_2))\alpha_{\max}\Big(\frac{\alpha_{\min}}{\alpha_{\max}}\Big)^{1/(2n)}, \tag{1.2}$$

$$\alpha_{\max}\kappa(A_1)^{-1} \leq |\lambda| \leq \alpha_{\max}\kappa(A_2) \quad \forall\lambda \in \Lambda(L), \tag{1.3}$$

where $\Lambda(P)$ denotes the spectrum of $P(\lambda)$, $\operatorname{gap}\big(\Lambda(Q), \Lambda(L)\big)$ is a measure of the distance between the n largest eigenvalues of $Q(\lambda)$ in modulus and the n eigenvalues of $L(\lambda) = A_2\lambda + A_1$, $g(\kappa(A_2))$ is more or less a constant times the matrix condition number $\kappa(A_2) = \|A_2\|\|A_2^{-1}\|$, and $\alpha_{\max}$ and $\alpha_{\min}$ are the largest and smallest tropical roots of $\mathsf{t}_\times q(x) := \max\big(\|A_0\|, \|A_1\|x, \|A_2\|x^2\big)$. The bounds (1.2)–(1.3) show that when the ratio $\alpha_{\min}/\alpha_{\max}$ is small enough and A_2, A_1 are well conditioned then there are precisely n eigenvalues of $Q(\lambda)$ with moduli of the order of $\alpha_{\max}$. Similarly, when A_1 and A_0 are both well conditioned, the moduli of the n smallest eigenvalues of $Q(\lambda)$ are close to the smallest tropical root $\alpha_{\min}$ of $\mathsf{t}_\times q(x)$.

For the particular case of matrix polynomials $P(\lambda)$ with coefficients matrices of the form $A_i = \sigma_i Q_i$ with $\sigma_i \geq 0$ and $Q_i^* Q_i = I$, and the 2-norm $\|\cdot\|_2$, Bini, Noferini and Sharify [7, Thm. 2.7] have identified annuli of small width defined in terms of the tropical roots of $\mathsf{t}_\times p(x)$ that contain the eigenvalues of $P(\lambda)$. We extend their results to arbitrary matrix polynomials and any subordinate matrix norm in Section 2.2. We obtain conditions under which tropical roots offer order of magnitude approximations to the moduli of the eigenvalues of $P(\lambda)$. As shown in Section 3 our tropical localization results are less tight than those from two generalized matrix version of Pellet's theorem [7, Thm. 2.1] and [16, Thm. 3.3] but they are easier to interpret. We illustrate our localization results with numerical examples in Section 4 and show experimentally how tropical roots can help in the design of a numerically stable polynomial eigensolver.

We note that a different approach, also involving tropical roots, is pursued in [2], where Akian, Gaubert and Sharify derive bounds for products of eigenvalues of $n \times n$ matrix polynomials $P(\lambda)$ of degree ℓ. Their results generalize to matrix polynomials bounds by Ostrowski and Pólya [17, 18] for products of roots of scalar polynomials.

2. Tropical bounds.

The *max-plus semiring* $\mathbb{R}_{\max}$ is the set $\mathbb{R}\cup\{-\infty\}$ equipped with the max operation denoted by $\oplus$ as addition and the usual addition denoted by $\otimes$ as multiplication. The zero and unit elements of this semiring are $-\infty$ and 0, respectively.

A variant of $\mathbb{R}_{\max}$ is the *max-times semiring* $\mathbb{R}_{\max,\times}$, which is the set of nonnegative real numbers $\mathbb{R}^+$ equipped with the max operation as addition and the usual multiplication as multiplication. This semiring is isomorphic to $\mathbb{R}_{\max}$ by the map $x \mapsto \log x$. So, every notion defined over $\mathbb{R}_{\max}$ has an $\mathbb{R}_{\max,\times}$ analogue. By the word "*tropical*", we refer to any of these algebraic structures.

2.1. Tropical polynomial and Newton polygon.

A max-plus *tropical polynomial* $\mathsf{t}p$ is a function of a variable $x \in \mathbb{R}_{\max}$ of the form

$$\mathsf{t}p(x) := \bigoplus_{i=0}^{\ell} a_i \otimes x^{\otimes i} = \max_{0\leq i\leq \ell}(a_i + ix), \tag{2.1}$$

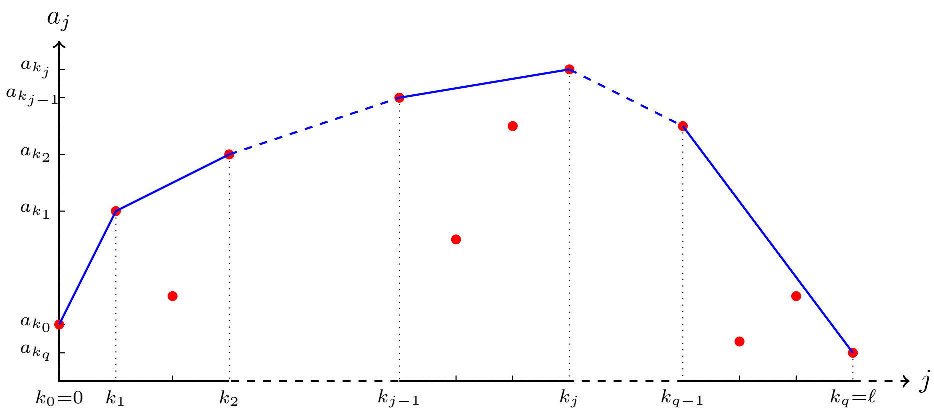

Fig. 2.1: Newton polygon (in blue) corresponding to a max-plus tropical polynomial $\mathsf{t}p(x) = \bigoplus_{j=0}^{\ell} a_j \otimes x^{\otimes j}$. Points (j, a_j) are denoted by red dots.

where ℓ is a nonnegative integer and $a_0, \ldots, a_\ell \in \mathbb{R}_{\max}$. The tropical polynomial $\mathsf{t}p$ is of *degree* ℓ if $a_\ell \neq -\infty$. If we assume that at least one of the coefficients $a_0, \ldots, a_\ell$ is finite then $\mathsf{t}p$ is a real valued convex function, piecewise affine, with integer slopes. The finite *tropical roots* of $\mathsf{t}p(x)$ are the points at which the maximum in the expression (2.1) is attained at least twice. If $a_0 = -\infty$ then $-\infty$ is a tropical root. A tropical polynomial of degree ℓ has ℓ tropical roots counting multiplicities. The multiplicity of a finite root α coincides with the variation of the derivative of the map $\mathsf{t}p$ at α, $\lim_{\epsilon\to 0} \frac{d\,\mathsf{t}p}{dx}|_{x+\epsilon} - \frac{d\,\mathsf{t}p}{dx}|_{x-\epsilon}$. The multiplicity of $-\infty$ as a root of $\mathsf{t}p$ is given by $\lim_{\epsilon\to 0} \frac{d\,\mathsf{t}p}{dx}|_{x+\epsilon}$ or equivalently by $\inf\{j \mid a_j \neq -\infty\}$.

The tropical roots can be obtained via *Newton polygons*. Define the Newton polygon of $\mathsf{t}p$ to be the upper boundary of the convex hull of the set of points (j, a_j), $j = 0, \ldots, \ell$ (see Fig. 2.1). This boundary consists of a number of linear segments. The opposites of the slopes of these segments are precisely the tropical roots and the multiplicity of a root coincides with the width of the corresponding segment measured by the difference of the abscissae of its endpoints (see [1, Prop. 2.10] or [15, Lem. 2.3]). Hence, if we denote by $k_0 = 0 < \cdots < k_q = \ell$ the abscissae of the vertices of the Newton polygon then $\mathsf{t}p(x)$ has q distinct roots given by

$$\alpha_j = -\frac{a_{k_j} - a_{k_{j-1}}}{k_j - k_{j-1}}, \quad j = 1, \ldots, q \tag{2.2}$$

with multiplicities $m_j = k_j - k_{j-1}$, $j = 1, \ldots, q$, respectively. Since the points (j, a_j) are already sorted by abscissae, the Graham scan algorithm [11] computes the convex hull of these points in $O(\ell)$ operations. As a result the tropical roots, counted with multiplicities, can be computed in $O(\ell)$ operations [8, Prop. 1].

In the "max-times" semiring $\mathbb{R}_{\max,\times}$, a tropical polynomial has the form $\mathsf{t}_\times p(x) = \max_{0\leq i\leq \ell} a_i x^i$, where $a_0, \ldots, a_\ell$ are nonnegative numbers, and x takes nonnegative values. The *tropical roots* of $\mathsf{t}_\times p(x)$ are, by definition, the exponentials of the tropical roots of the max-plus polynomial $\mathsf{t}p(x) = \max_{0\leq i\leq \ell}(\log a_i + ix)$. So on using (2.2), the q distinct tropical roots of $\mathsf{t}_\times p(x)$ and their multiplicities are given by

$$(a_{k_{j-1}}/a_{k_j})^{1/(k_j - k_{j-1})}, \quad m_j = k_j - k_{j-1}, \quad j = 1, \ldots, q,$$

respectively.

2.2. Eigenvalue location: tropical approach. Throughout the rest of this paper, any matrix polynomial denoted by $P(\lambda)$ is regular, i.e., $\det P(\lambda)$ is not identically zero. Then the finite eigenvalues of an $n \times n$ $P(\lambda)$ of degree ℓ are the roots of $\det P(\lambda) = 0$, and if $\det P(\lambda) = 0$ has degree $d \le \ell n$, then $P(\lambda)$ has $\ell n - d$ eigenvalues at infinity. To $P(\lambda) = \sum_{i=0}^{\ell} A_i \lambda^i$ we associate the max-times tropical scalar polynomial

$$\mathsf{t}_\times p(x) = \max_{0 \le i \le \ell} \|A_i\| x^i, \tag{2.3}$$

where $\|\cdot\|$ is any matrix norm subordinate to a vector norm. Our aim is to show that, under specific conditions, the tropical roots of $\mathsf{t}_\times p(x)$ are good approximations to the moduli of the eigenvalues of $P(\lambda)$. The key tool for this is a generalization of Rouché's theorem for matrix valued functions [9], [16].

Theorem 2.1 (Generalized Rouché theorem for matrix valued functions). *Let $P, Q : \Omega \to \mathbb{C}^{n\times n}$ be analytic matrix-valued functions, where Ω is an open connected subset of $\mathbb{C}$ and assume that $P(\lambda)$ is nonsingular for all λ on the simple closed curve $\Gamma \subseteq \Omega$. Let $\|\cdot\|$ be any matrix norm on $\mathbb{C}^{n\times n}$ induced by a vector norm on $\mathbb{C}^n$. If $\|P(\lambda)^{-1}Q(\lambda)\| < 1$ for all $\lambda \in \Gamma$, then* $\det(P+Q)$ *and* $\det(P)$ *have the same number of zeros inside Γ, counting multiplicities.*

Before deriving our new results, we set up the notation used throughout the reminder of this paper.

2.2.1. Notation. The variable i will usually be an index varying between 0 and the degree ℓ of $\mathsf{t}_\times p(x)$ in (2.3), whereas j will be an index with value between 1 and q, where q is the number of distinct tropical roots of $\mathsf{t}_\times p(x)$. These tropical roots will be denoted by α_j, $j = 1, \dots, q$ with

$$\alpha_j := (\|A_{k_{j-1}}\| / \|A_{k_j}\|)^{1/(k_j - k_{j-1})} \tag{2.4}$$

of multiplicity $m_j = k_j - k_{j-1}$, where

$$k_0 = 0 < \ldots < k_q = \ell$$

denote the abscissae of the Newton polygon associated with the max-plus polynomial $\mathsf{t}p(x) = \max_{0\le i\le \ell}(\log\|A_i\| + ix)$ (see Fig. 2.1). We write

$$\mathcal{K} := \{k_0, \dots, k_q\}. \tag{2.5}$$

Note that the tropical roots (2.4) have the property that $\alpha_1 < \cdots < \alpha_q$.

As in [7] and [8], the tropical roots (2.4) will be used to define an eigenvalue parameter scaling, $\lambda = \alpha_j \mu$, and a scaled matrix polynomial $\widetilde{P}(\mu)$ via

$$\big(\mathsf{t}_\times p(\alpha_j)\big)^{-1} P(\lambda) = \big(\|A_{k_{j-1}}\| \alpha_j^{k_{j-1}}\big)^{-1} P(\alpha_j \mu) = \sum_{i=0}^{\ell} \widetilde{A}_i \mu^i =: \widetilde{P}(\mu), \tag{2.6}$$

where

$$\widetilde{A}_i = \big(\|A_{k_{j-1}}\| \alpha_j^{k_{j-1}}\big)^{-1} A_i \alpha_j^i. \tag{2.7}$$

Clearly, μ is an eigenvalue of $\widetilde{P}(\mu)$ if and only if $\alpha_j\mu$ is an eigenvalue of $P(\lambda)$. Note that this scaling does not affect the condition number with respect to inversion,

$$\kappa(A_i) := \|A_i\|\|A_i^{-1}\| = \kappa(\widetilde{A}_i),$$

of any coefficient matrix A_i. By convention, $\kappa(A) = \infty$ when A is singular.

We will use disks and annuli to localize the eigenvalues of $P(\lambda)$. The closed (open) disk in the complex plane centered at 0 with radius r is denoted by $\mathcal{D}(r)$ $\big(\mathring{\mathcal{D}}(r)\big)$ and $\mathcal{A}(a,b) := \{\lambda \in \mathbb{C},\ \ a \le |\lambda| \le b\}$ denotes a closed annulus centered at 0 with a, b such that $0 < a < b$. We write $\mathring{\mathcal{A}}(a,b)$ for the interior of $\mathcal{A}(a,b)$. Finally, we denote by

$$\delta_j = \frac{\alpha_j}{\alpha_{j+1}} < 1, \quad 1 \le j < q, \tag{2.8}$$

the ratio between two consecutive roots.

2.2.2. Preliminary results. We now present preliminary lemmas, which will be needed in Section 2.2.3 to prove our localization results. We refer to Section 2.2.1 for the notation.

The norms of the coefficient matrices of the scaled matrix polynomial $\widetilde{P}(\mu)$ in (2.6) are at most 1 as shown by this first lemma.

LEMMA 2.2. *The norm of $\widetilde{A}_i$ in* (2.7) *satisfies*

$$\|\widetilde{A}_i\| \le \begin{cases} {\delta_{j-1}}^{k_{j-1}-i} & \text{if} \quad 0 \le i < k_{j-1}, \\ 1 & \text{if} \quad k_{j-1} < i < k_j, \\ {\delta_j}^{i-k_j} & \text{if} \quad k_j < i \le \ell, \end{cases} \qquad \|\widetilde{A}_{k_{j-1}}\| = \|\widetilde{A}_{k_j}\| = 1.$$

Proof. This is a corollary of [19, Lem. 3.3.2]. See also [7, Lem. 3.4]. ∎

The next lemma provides upper and lower bounds on the moduli of all the eigenvalues of $P(\lambda)$ in terms of the smallest and largest tropical roots α_1 and α_q of $\mathsf{t}_\times p(x)$, and the conditioning of A_0 and A_ℓ.

LEMMA 2.3. *Every eigenvalue λ of a matrix polynomial $P(\lambda)$ satisfies*

$$\big(1+\kappa(A_0)\big)^{-1}\alpha_1 \le |\lambda| \le \big(1+\kappa(A_\ell)\big)\alpha_q.$$

Furthermore, if both A_0 and A_ℓ are invertible, both inequalities are strict.

Proof. For the upper bound, we consider the scaled matrix polynomial $\widetilde{P}(\mu)$ in (2.6) with $j = q$. Observe first that if A_ℓ is singular then the right hand side is ∞, so there is nothing to prove and, if $P(\lambda)$ is regular, then the bound is attained in the sense that necessarily $P(\lambda)$ has an eigenvalue at infinity. Hence, we may assume that A_ℓ is invertible. Let $\theta = \max_{0\le i\le \ell-1} \|\widetilde{A}_i\|^{1/(\ell-i)}$. We now recall an argument from the proof of [14, Lem. 4.1]. For any eigenpair (μ, x) such that $|\mu| > \theta$ and $\|x\| = 1$,

$$\begin{aligned} 0 = \|\widetilde{P}(\mu)x\| &\ge |\mu^\ell|\bigg(\|\widetilde{A}_\ell^{-1}\|^{-1} - \sum_{i=0}^{\ell-1}\frac{\|\widetilde{A}_i\|}{|\mu^{\ell-i}|}\bigg) \\ &\ge |\mu^\ell|\bigg(\|\widetilde{A}_\ell^{-1}\|^{-1} - \sum_{i=1}^{\ell}\frac{\theta^i}{|\mu^i|}\bigg) \\ &\ge |\mu^\ell|\bigg(\|\widetilde{A}_\ell^{-1}\|^{-1} - \sum_{i=1}^{\infty}\frac{\theta^i}{|\mu^i|}\bigg) = |\mu^\ell|\bigg(\|\widetilde{A}_\ell^{-1}\|^{-1} - \frac{\theta}{|\mu|-\theta}\bigg). \end{aligned}$$

Hence, any eigenvalue must satisfy $|\mu| \le \beta := (1 + \|\widetilde{A}_\ell^{-1}\|)\theta$ (observe that $\theta < \beta$, so the initial assumption is no loss of generality). If we additionally assume that A_0 is also invertible, then $\theta > 0$. Thus, the argument above can be tightened as the inequality $\sum_{i=1}^{\ell} \frac{\theta^i}{|\lambda^i|} > \sum_{i=1}^{\infty} \frac{\theta^i}{|\lambda^i|}$ is in this case strict, yielding $|\mu| < \beta$.

To conclude the proof, we observe that from Lemma 2.2, $\theta = 1$ and $\|\widetilde{A}_\ell\| = 1$ so that $\|\widetilde{A}_\ell^{-1}\| = \kappa(\widetilde{A}_\ell) = \kappa(A_\ell)$. It follows that $|\mu| \le 1 + \kappa(A_\ell)$, that is, $|\lambda| \le (1+\kappa(A_\ell))\alpha_q$ since $\lambda = \alpha_q\mu$. Moreover, if $\theta > 0$, then the last inequality is also strict.

The lower bound is proved similarly, using $\widetilde{P}(\mu)$ in (2.6) with $j = 1$ and applying the above argument to $\text{rev}\widetilde{P}(\mu)$. We invite the reader to fill in the details. ∎

With the aim of invoking Theorem 2.1, we decompose $\widetilde{P}(\mu) = \big(\mathsf{t}_\times p(\alpha_j)\big)^{-1} P(\lambda)$ as the sum of two matrix polynomials,

$$S(\mu) = \sum_{i=k_{j-1}}^{k_j} \widetilde{A}_i \mu^i, \qquad Q(\mu) = \sum_{i=0}^{k_{j-1}-1} \widetilde{A}_i \mu^i + \sum_{i=k_j+1}^{\ell} \widetilde{A}_i \mu^i. \tag{2.9}$$

We will need the following localization result for the nonzero eigenvalues of $S(\mu)$.

LEMMA 2.4. *If $A_{k_{j-1}}$ and A_{k_j} are nonsingular then the nm_j nonzero eigenvalues of $S(\mu)$ in* (2.9) *are located in the open annulus $\mathring{\mathcal{A}}\big(\big(1 + \kappa(A_{k_{j-1}})\big)^{-1}, 1 + \kappa(A_{k_j})\big)$.*

Proof. Note that $S(\mu)$ is regular, of degree k_j, with nk_{j-1} zero eigenvalues. Hence $S(\mu)$ has $nk_j - nk_{j-1} = nm_j$ nonzero eigenvalues, which are eigenvalues of $\mu^{-k_{j-1}} S(\mu)$. The tropical polynomial associated with $\mu^{-k_{j-1}} S(\mu)$ has only one root, which is equal to 1. The lemma is then a direct consequence of Lemma 2.3. ∎

Bounds on the norms of $Q(\mu)$ and $S(\mu)^{-1}$ will also be needed.

LEMMA 2.5. *The following hold for $Q(\mu)$ and the inverse of $S(\mu)$ in* (2.9),

$$\|Q(\mu)\| \le \frac{\delta_{j-1}|\mu|^{k_{j-1}}}{|\mu| - \delta_{j-1}} + \frac{\delta_j |\mu|^{k_j+1}}{1 - \delta_j|\mu|} \qquad \text{if } \delta_{j-1} < |\mu| < \frac{1}{\delta_j},$$

$$\|S(\mu)^{-1}\| \le \begin{cases} \dfrac{\kappa(A_{k_j-1})|\mu|^{-k_{j-1}}(1 - |\mu|)}{1 - |\mu|\big(1 + \kappa(A_{k_j-1})(1 - |\mu|^{m_j})\big)} & \text{if } 0 < |\mu| \le \big(1 + \kappa(A_{k_j-1})\big)^{-1}, \\[2ex] \dfrac{\kappa(A_{k_j})|\mu|^{-k_j}(|\mu| - 1)}{|\mu| - 1 - \kappa(A_{k_j})(1 - |\mu|^{-m_j})} & \text{if } |\mu| \ge 1 + \kappa(A_{k_j}). \end{cases}$$

Proof. Assume that $\delta_{j-1} < |\mu| < \frac{1}{\delta_j}$. Using (2.9) and Lemma 2.2 we have that

$$\begin{aligned} \|Q(\mu)\| &\le \sum_{i=0}^{k_{j-1}-1} {\delta_{j-1}}^{k_{j-1}-i}|\mu|^i + \sum_{i=k_j+1}^{\ell} {\delta_j}^{i-k_j}|\mu|^i \\ &= \frac{\delta_{j-1}(|\mu|^{k_{j-1}} - {\delta_{j-1}}^{k_{j-1}})}{|\mu| - \delta_{j-1}} + \frac{\delta_j|\mu|^{k_j+1}(1 - (\delta_j|\mu|)^{\ell-k_j})}{1 - \delta_j|\mu|} \end{aligned}$$

and the bound in the lemma follows since $\delta_{j-1} < |\mu| < \frac{1}{\delta_j}$.

Assume that $\mathcal{S} := \{\mu \in \mathbb{C} : 0 < |\mu| \le \big(1 + \kappa(A_{k_j-1})\big)^{-1}\}$ is non empty, that is, that $A_{k_{j-1}}$ is nonsingular. By Lemma 2.4, the matrix $S(\mu)$ is nonsingular for all

$\mu \in \mathcal{S}$. We rewrite $S(\mu)$ as

$$S(\mu) = A(\mu)\big(I - (-A(\mu)^{-1}B(\mu))\big), \tag{2.10}$$

where $A(\mu) = \widetilde{A}_{k_{j-1}}\mu^{k_{j-1}}$ is nonsingular and $B(\mu) = \sum_{i=k_{j-1}+1}^{k_j} \widetilde{A}_i\mu^i$. Note that if we can show that $\|A(\mu)^{-1}B(\mu)\| < 1$, then the matrix $I-(-A(\mu)^{-1}B(\mu))$ is nonsingular and

$$\|S(\mu)^{-1}\| \le \frac{\|A(\mu)^{-1}\|}{1-\|A(\mu)^{-1}B(\mu)\|} \tag{2.11}$$

(see for example [10, Lem. 2.3.3]). Using Lemma 2.2 we have that

$$\|A(\mu)^{-1}B(\mu)\| \le \kappa(A_{k_{j-1}})|\mu|^{-k_{j-1}}\Big(\sum_{i=k_{j-1}+1}^{k_j} |\mu|^i\Big) = \kappa(A_{k_{j-1}})\frac{|\mu|(1-|\mu|^{m_j})}{1-|\mu|}. \tag{2.12}$$

Now $\mu \in \mathcal{S}$ implies $|\mu| < 1$ so that $1-|\mu|^{m_j} < 1$. Also, since $|\mu| \le (1+\kappa(A_{k_j-1}))^{-1}$, we have that $\frac{|\mu|}{1-|\mu|} \le \kappa(A_{k_j-1})^{-1}$ so that, using the upper bound in (2.12), we find that $\|A(\mu)^{-1}B(\mu)\| < 1$. The upper bound for $\|S(\mu)^{-1}\|$ when $\mu \in \mathcal{S}$ follows by combining (2.11) and (2.12), and by noting that $\|A(\mu)^{-1}\| = |\mu|^{-k_{j-1}}\kappa(A_{k_{j-1}})$.

We now consider $\mu \in \mathbb{C}$ such that $|\mu| \ge 1+\kappa(A_{k_j})$. Note that such a μ exists only if A_{k_j} is nonsingular. Lemma 2.4 implies that for such a μ, the matrix $S(\mu)$ is invertible. We rewrite $S(\mu)$ as in (2.10) with $A(\mu) = \widetilde{A}_{k_j}\mu^{k_j}$, $B(\mu) = \sum_{i=k_{j-1}}^{k_j-1} \widetilde{A}_i\mu^i$. The rest of the proof is then analogous to the case where $\mu \in \mathcal{S}$ so we omit it. ∎

Finally, this last technical lemma will be needed in the proof of our tropical localization results in Section 2.2.3, and in Section 3 when comparing the Pellet bounds to the tropical bounds.

LEMMA 2.6. *For given $c, \delta > 0$ such that $\delta \le (1+2c)^{-2}$, the quadratic polynomial*

$$p(r) = r^2 - \left(2 + \frac{1-\delta}{\delta(1+c)}\right)r + \frac{1}{\delta}$$

has two real roots

$$f := f(\delta, c) = \frac{(1+2c)\delta + 1 - \sqrt{(1-\delta)(1-(1+2c)^2\delta)}}{2\delta(1+c)}, \qquad g = (\delta f)^{-1},$$

with the properties that

(i) $1 < 1+c \le f \le g$,

(ii) $\dfrac{1}{f-1} + \dfrac{1}{g-1} = \dfrac{1}{c}$.

Proof. (i) The discriminant of $p(r)$ is not negative since $\delta \le (1+2c)^{-2}$ so p has two real roots, f, g such that $f \le g$. That $f \ge 1+c$ is easy to check.

(ii) Clearly,

$$\left(\frac{1}{f-1} + \frac{1}{g-1}\right)^{-1} = \frac{1-f-g+fg}{f+g-2}.$$

Since f and g are the roots of p, $f+g = 2+\dfrac{1-\delta}{\delta(1+c)}$. Hence, recalling that $fg=\delta^{-1}$, we get $\left(\dfrac{1}{f-1}+\dfrac{1}{g-1}\right)^{-1} = \dfrac{c-\delta c}{1-\delta} = c.$ $\square$

2.2.3. Main results. When $\|A_i\| \le \|A_0\|^{(\ell-i)/\ell}\|A_\ell\|^{i/\ell}$ for all $0\le i\le \ell$ then $P(\lambda)$ has only one tropical root given by $\alpha = (\|A_0\|/\|A_\ell\|)^{1/\ell}$ and we know from Lemma 2.3 that all the eigenvalues of $P(\lambda)$ lie in the annulus

$$\mathcal{A}\big((1+\kappa(A_0))^{-1}\alpha, (1+\kappa(A_\ell))\alpha\big).$$

So in this case, if A_0 and A_ℓ are well-conditioned then $P(\lambda)$ has $n\ell$ eigenvalues of modulus close to α. We now extend this type of result to the case where $\mathsf{t}_\times p(x)$ has more than one tropical root.

THEOREM 2.7. *Let $P(\lambda)=\sum_{i=0}^{\ell} A_i\lambda^i \in \mathbb{C}[\lambda]^{n\times n}$ be regular. For $1\le j\le q-1$, let $f_j = f(\delta_j,\kappa(A_{k_j}))$, where $f(\delta,c)$ is defined as in Lemma 2.6, and $g_j=(\delta_j f_j)^{-1}$. Then, in the notation of Section 2.2.1, the following statements hold.*

- (i) *If $\delta_j \le (1+2\,\kappa(A_{k_j}))^{-2}$ with $1\le j\le q-1$ then $P(\lambda)$ has exactly nk_j eigenvalues inside the disk $\mathcal{D}(f_j\alpha_j)$ and it does not have any eigenvalue in the open annulus $\mathring{\mathcal{A}}(f_j\alpha_j, g_j\alpha_j)$.*
- (ii) *If $\delta_j \le (1+2\,\kappa(A_{k_j}))^{-2}$ and $\delta_s \le (1+2\,\kappa(A_{k_s}))^{-2}$ with $1<j<s<q$ then $P(\lambda)$ has exactly $n(k_s-k_j)$ eigenvalues inside the annulus $\mathcal{A}(g_j\alpha_j, f_s\alpha_s)$.*
- (iii) *If $\delta_1 \le (1+2\,\kappa(A_{k_1}))^{-2}$ then $P(\lambda)$ has exactly nk_1 eigenvalues inside the annulus $\mathcal{A}\big((1+\kappa(A_0))^{-1}\alpha_1, f_1\alpha_1\big)$.*
- (iv) *If $\delta_{q-1} \le (1+2\,\kappa(A_{k_{q-1}}))^{-2}$ then $P(\lambda)$ has exactly nm_q eigenvalues inside the annulus $\mathcal{A}\big(g_{q-1}\alpha_{q-1}, (1+\kappa(A_\ell))\alpha_q\big)$.*

Proof. (i) We assume that $\delta_j \le (1+2\,\kappa(A_{k_j}))^{-2}$ and we partition $\widetilde{P}(\mu)$ as in (2.9). Let r be such that

$$1+\kappa(A_{k_j}) < r < 1/\delta_j. \tag{2.13}$$

Note that such r exists since $\delta_j \le (1+2\,\kappa(A_{k_j}))^{-2} < (1+\kappa(A_{k_j}))^{-1}$. By Lemma 2.4, $S(\mu)$ is nonsingular on the circle $\Gamma_r=\{\mu\in\mathbb{C} : |\mu|=r\}$. To apply Theorem 2.1 with $\widetilde{P}(\mu)=S(\mu)+Q(\mu)$ and Γ_r, we must check that $\|S(\mu)^{-1}Q(\mu)\|<1$ for all $\mu\in\Gamma_r$. Since $|\mu|=r$ with r such that

$$\delta_{j-1} < 1 < 1+\kappa(A_{k_j}) < r < 1/\delta_j, \tag{2.14}$$

we can apply the bounds in Lemma 2.5. These yield

$$\begin{aligned}\|S(\mu)^{-1}Q(\mu)\| &\le \|S(\mu)^{-1}\|\|Q(\mu)\| \\ &\le \frac{r^{-k_j}(r-1)\,\kappa(A_{k_j})}{\big(r-1-\kappa(A_{k_j})(1-r^{-m_j})\big)}\left(\frac{\delta_{j-1}r^{k_{j-1}}}{r-\delta_{j-1}}+\frac{\delta_j r^{k_j+1}}{1-\delta_j r}\right).\end{aligned}$$

The latter bound is less than 1 if

$$\frac{\delta_{j-1}r^{-m_j}}{r-\delta_{j-1}}+\frac{\delta_j r}{1-\delta_j r} < \frac{r-1-\kappa(A_{k_j})(1-r^{-m_j})}{(r-1)\,\kappa(A_{k_j})},$$

or equivalently, if

$$\frac{\delta_j r}{1-\delta_j r} < \frac{r-1-\kappa(A_{k_j})}{(r-1)\,\kappa(A_{k_j})} + r^{-m_j}\left(\frac{1}{r-1}-\frac{\delta_{j-1}}{r-\delta_{j-1}}\right).$$

Since $\frac{1}{r-1} > \frac{\delta_{j-1}}{r-\delta_{j-1}}$, the last inequality holds when $\frac{\delta_j r}{1-\delta_j r} < \frac{r-1-\kappa(A_{k_j})}{(r-1)\,\kappa(A_{k_j})}$, or equivalently when $p(r) < 0$, where p is as in Lemma 2.6 with $\delta = \delta_j$ and $c = \kappa(A_{k_j})$. It follows from Lemma 2.6 that $p(r)$ is negative for the values of r such that

$$f_j < r < g_j \tag{2.15}$$

recalling that by Lemma 2.6 f_j and g_j are the two roots of p. Note that, by the same lemma, $f_j \geq 1 + \kappa(A_{k_j})$ and $g_j \leq (\delta_j)^{-1}$ so (2.15) is sharper than (2.14). So, for any r satisfying (2.15), $\|S(\mu)^{-1}Q(\mu)\| < 1$ for all $\mu \in \Gamma_r$. By Theorem 2.1, $S(\mu)$ and $\widetilde{P}(\mu)$ have the same number of eigenvalues inside the disks $\mathcal{D}(r)$ for all r satisfying (2.15). Since $S(\mu)$ has nk_j eigenvalues inside any of these disks, $\widetilde{P}(\mu)$ does not have any eigenvalue in $\mathring{\mathcal{A}}(f_j, g_j)$ and has exactly nk_j eigenvalues inside the disk $\mathcal{D}(f_j)$. This completes the proof of (i) since $\lambda = \mu\alpha_j$.

(ii) If $\delta_s \leq \big(1 + 2\,\kappa(A_{k_s})\big)^{-2}$ then by (i), $P(\lambda)$ has nk_s eigenvalues inside the disk $\mathcal{D}(f_s\alpha_s)$ and no eigenvalues inside the annulus $\mathring{\mathcal{A}}(f_s\alpha_s, g_s\alpha_s)$. An analogous statement holds if we assume that $\delta_j < (1 + 2\,\kappa(A_{k_j}))^{-2}$. This implies that $P(\lambda)$ has exactly $n(k_s - k_j)$ eigenvalues which lie in the annulus $\mathcal{A}(g_j\alpha_j, f_s\alpha_s)$.

(iii) If $\delta_1 \leq (1 + 2\kappa(A_{k_1}))^{-2}$ then by (i) $P(\lambda)$ has nk_1 eigenvalues inside the disk $\mathcal{D}(f_1\alpha_1)$. Also by Lemma 2.3 $P(\lambda)$ does not have any eigenvalue inside the disk $\mathring{\mathcal{D}}((1 + \kappa(A_0))^{-1}\alpha_1)$.

(iv) If $\delta_{q-1} \leq (1 + 2\kappa(A_{q-1}))^{-2}$ then by (i) $P(\lambda)$ has nk_{q-1} eigenvalues inside the disk $\mathcal{D}\big(g_{q-1}\alpha_{q-1}\big)$. Also by Lemma 2.3 all the eigenvalues of $P(\lambda)$ lie inside the annulus $\mathcal{A}\big((1 + \kappa(A_0))^{-1}\alpha_1, (1 + \kappa(A_\ell))\alpha_q\big)$, which completes the proof. ∎

Note that for a fixed value of $c \geq 1$ and $\delta \leq (1 + 2c)^{-2}$, $f(\delta, c)$ in Lemma 2.6 is an increasing function of δ and its maximum value, which is $1 + 2c$, is achieved at $\delta = (1+2c)^{-2}$. This implies that $f\big(\delta_j, \kappa(A_{k_j})\big) \leq 1+2\,\kappa(A_{k_j})$ for any $\delta_j \leq (1+2\,\kappa(A_{k_j}))^{-2}$ and we have the following corollary.

Corollary 2.8. *In the notation of Section 2.2.1 and under the assumptions of Theorem 2.7, the following statements hold.*

- (i) *If $\delta_j \leq (1 + 2\,\kappa(A_{k_j}))^{-2}$ for some j such that $1 \leq j \leq q-1$, then $P(\lambda)$ has exactly nk_j eigenvalues inside the disk $\mathcal{D}((1+2\,\kappa(A_{k_j}))\alpha_j)$ and no eigenvalue in the open annulus $\mathring{\mathcal{A}}\big((1 + 2\,\kappa(A_{k_j}))\alpha_j, (1 + 2\,\kappa(A_{k_j}))^{-1}\alpha_{j+1}\big)$.*
- (ii) *If $\delta_j \leq (1 + 2\,\kappa(A_{k_j}))^{-2}$ and $\delta_s \leq (1 + 2\,\kappa(A_{k_s}))^{-2}$ for some j and s such that $1 < j < s < q$ then $P(\lambda)$ has exactly $n(k_s - k_j)$ eigenvalues inside the annulus $\mathcal{A}\big((1 + 2\,\kappa(A_{k_j}))^{-1}\alpha_{j+1}, (1 + 2\,\kappa(A_{k_s}))\alpha_s\big)$.*
- (iii) *If $\delta_1 \leq (1 + 2\,\kappa(A_{k_1}))^{-2}$ then $P(\lambda)$ has exactly nk_1 eigenvalues inside the annulus $\mathcal{A}\big((1 + \kappa(A_0))^{-1}\alpha_1, (1 + 2\,\kappa(A_{k_1}))\alpha_1\big)$.*
- (iv) *If $\delta_{q-1} < (1 + 2\,\kappa(A_{k_{q-1}}))^{-2}$ then $P(\lambda)$ has exactly nm_q eigenvalues inside the annulus $\mathcal{A}\big((1 + 2\,\kappa(A_{k_{q-1}}))^{-1}\alpha_q, (1 + \kappa(A_\ell))\alpha_q\big)$.*

It follows from Corollary 2.8 that if $A_{k_{j-1}}$ and A_{k_j} are well conditioned and the ratios $\delta_{j-1} = \alpha_{j-1}/\alpha_j$, $\delta_j = \alpha_j/\alpha_{j+1}$ are small enough then $P(\lambda)$ has nm_j eigenvalues of modulus close to α_j. In particular, if $\kappa(A_{k_{j-1}}) = \kappa(A_{k_j}) = 1$ and $\delta_{j-1}, \delta_j < \frac{1}{9}$ then $P(\lambda)$ has exactly nm_j eigenvalues in the annulus $\mathcal{A}(1/3\alpha_j, 3\alpha_j)$. This is an improvement over [7, Thm. 2.7]. When $n = 1$, the bounds in Corollary 2.8 are the same as the ones that appeared in [19, Thm. 3.3.3] for scalar polynomials.

Define

$$\mathcal{J} := \{j_0, j_1, \ldots, j_m\} \subseteq \{0, \ldots, q\} \tag{2.16}$$

to be such that

(a) $0 = j_0 < j_1 < \cdots < j_m = q$,

(b) for $1 \le i \le m-1$, $\delta_{j_i} \le \big(1 + 2\,\kappa(A_{k_{j_i}})\big)^{-2}$ if and only if $j_i \in \mathcal{J}$.

(Here m is implicitly defined by $|\mathcal{J}| = m+1$.) In other words, the j_i are all the indices such that δ_{j_i} satisfies the condition in Theorem 2.7 or Corollary 2.8. We also define

$$\widetilde{\mathcal{K}} = \{k_{j_i} \in \mathcal{K} : j_i \in \mathcal{J}\}. \tag{2.17}$$

If we let, in the notation of Theorem 2.7,

$$\begin{aligned} b_{j_0} &= \big(1 + \kappa(A_0)\big)^{-1}\alpha_1, & a_{j_m} &= \big(1 + \kappa(A_\ell)\big)\alpha_q, \\ a_{j_i} &= f_{j_i}\alpha_{j_i}, & b_{j_i} &= g_{j_i}\alpha_{j_i}, \quad 1 \le i < m, \end{aligned}$$

and

$$\begin{aligned} \widetilde{b}_{j_0} &= b_{j_0}, & \widetilde{a}_{j_m} &= a_{j_m}, \\ \widetilde{a}_{j_i} &= (1 + 2\,\kappa(A_{k_{j_i}}))\alpha_{j_i}, & \widetilde{b}_{j_i} &= \big(1 + 2\,\kappa(A_{k_{j_{i-1}}}))\big)^{-1}\alpha_{j_{i-1}+1}, \quad 1 \le i < m, \end{aligned}$$

then it follows from Lemma 2.3, Theorem 2.7 and Corollary 2.8 that

$$\Lambda(P) \subset \bigcup_{i=0}^{m-1} \mathcal{A}(b_{j_i}, a_{j_{i+1}}) \subseteq \bigcup_{i=0}^{m-1} \mathcal{A}(\widetilde{b}_{j_i}, \widetilde{a}_{j_{i+1}}), \tag{2.18}$$

where $\Lambda(P)$ denotes the spectrum of $P(\lambda)$.

3. Comparisons with Pellet's bounds. The generalized matrix versions of Pellet's theorem by Bini et al. [7, Thm. 2.1] and Melman [16, Thm. 3.3] also provide annuli containing the eigenvalues of $P(\lambda)$. Although [7, Thm. 2.1] is stated for the 2-norm, as mentioned in [16] the result can be extended to any subordinate norm by using Theorem 2.1. Throughout this section, roots of polynomials are counted with multiplicity: in particular, a double positive root is thought of as two coincident positive roots.

THEOREM 3.1 (Generalized Pellet theorem). *Let $P(\lambda) = \sum_{i=0}^{\ell} A_i\lambda^i$ with $\ell > 1$ and $A_0 \ne 0$, and let $\|\cdot\|$ denotes any subordinate matrix norm. For A_k nonsingular define*

$$q_k(x) := \sum_{i=0,\, i\ne k}^{\ell} \|A_k^{-1}A_i\|x^i - x^k. \tag{3.1}$$

The following statements hold.

(i) *If A_k is nonsingular for some k such that $1 \le k \le \ell - 1$ then $q_k(x)$ has either no real positive root or two real positive roots $s_k \le t_k$. In the latter case, $P(\lambda)$ has kn eigenvalues in the disk $\mathcal{D}(0, s_k)$ and no eigenvalue in the open annulus $\mathring{\mathcal{A}}(s_k, t_k)$.*

(ii) *If A_0 is nonsingular then $q_0(x)$ has only one real positive root t_0 and $P(\lambda)$ has no eigenvalue in the open disk $\mathring{\mathcal{D}}(0, t_0)$.*

(iii) *If A_ℓ is nonsingular then $q_\ell(x)$ has only one real positive root s_ℓ and all the eigenvalues of $P(\lambda)$ are located in the disk $\mathcal{D}(0, s_\ell)$.*

Melman's version of Pellet's theorem [16, Thm. 3.3] is analogous to Theorem 3.1 but with $q_k(x)$ replaced by

$$\widetilde{q}_k(x) := \sum_{i=0, i\neq k}^{\ell} \|A_i\| x^i - \|A_k^{-1}\|^{-1} x^k. \tag{3.2}$$

Although the coefficients of $\widetilde{q}_k(x)$ are less expensive to compute than those of $q_k(x)$, Theorem 3.1 provides tighter localization results than those in [16, Thm. 3.3]. This fact has appeared, in the form of a remark, in [7, p. 1713], [16, p. 1555]. The next proposition provides a detailed statement.

PROPOSITION 3.2. *Let $P(\lambda) = \sum_{i=0}^{\ell} A_i \lambda^i$ with $\ell > 1$ and $A_0 \neq 0$.*

(i) *If A_k is nonsingular for some k such that $1 \leq k \leq \ell - 1$, and $\widetilde{q}_k(x)$ has two real positive roots $\widetilde{s}_k, \widetilde{t}_k$ with $\widetilde{s}_k \leq \widetilde{t}_k$ then $q_k(x)$ has also two real positive roots, s_k, t_k such that $s_k \leq \widetilde{s}_k \leq \widetilde{t}_k \leq t_k$.*

(ii) *If A_0 is nonsingular and $\widetilde{q}_0(x)$ has a real positive root $\widetilde{t}_0$ then $q_0(x)$ has also a real positive root t_0 such that $\widetilde{t}_0 \leq t_0$.*

(iii) *If A_ℓ is nonsingular and $\widetilde{q}_\ell(x)$ has a real positive root $\widetilde{s}_n$ then $q_k(x)$ has also a positive root s_n such that $s_n \leq \widetilde{s}_n$.*

Proof. (i) Since for any subordinate matrix norm $\|AB\| \leq \|A\|\|B\|$, and $\widetilde{q}_k(\widetilde{s}_k) = \widetilde{q}_k(\widetilde{t}_k) = 0$, we have

$$\widetilde{s}_k^k = \sum_{i=0, i\neq k}^{\ell} \|A_k^{-1}\|\|A_i\|\widetilde{s}_k^i \geq \sum_{i=0, i\neq k}^{\ell} \|A_k^{-1} A_i\|\widetilde{s}_k^i \ ,$$

which implies that $q_k(\widetilde{s}_k) \leq 0$. Similarly we can show $q_k(\widetilde{t}_k) \leq 0$. Since $q_k(0) > 0$ and the leading coefficient of $q_k(x)$ is positive, this implies that $q_k(x)x$ has two positive roots $s_k \leq t_k$. But $q_k(\widetilde{s}_k), q_k(\widetilde{t}_k) \leq 0$, so we must have $s_k \leq \widetilde{s}_k \leq \widetilde{t}_k \leq t_k$. The statements (ii) and (iii) are proved in a similar way. ∎

The next result, when combined with Proposition 3.2, shows that the eigenvalue localization results from either version of the generalized Pellet theorem are always better than those presented in Theorem 2.7.

PROPOSITION 3.3. *Using the notation of Theorem 2.7, assume that $\delta_j \leq (1 + 2\,\kappa(A_{k_j}))^{-2}$ for some $1 \leq j \leq q-1$. Then the k_jth Pellet polynomial $\widetilde{q}_{k_j}(x)$ in (3.2) has two positive roots $\widetilde{s}_{k_j} < \widetilde{t}_{k_j}$ such that $\widetilde{s}_{k_j} < f_j \alpha_j$ and $\widetilde{t}_{k_j} > g_j \alpha_j$.*

Proof. Note that $\widetilde{q}_{k_j}(0) = \|A_0\| \geq 0$ and $\lim_{x\to\infty} \widetilde{q}_{k_j}(x) \sim x^\ell \|A_\ell\| > 0$. According to Pellet's theorem, $\widetilde{q}_{k_j}$ has either zero or two positive roots. Hence, if there exists two positive numbers $y_1 < y_2$ such that $\widetilde{q}_{k_j}(y_1) < 0$ and $\widetilde{q}_{k_j}(y_2) < 0$, then $\widetilde{q}_{k_j}$ has two positive roots $\widetilde{s}_{k_j}, \widetilde{t}_{k_j}$ such that $\widetilde{s}_{k_j} < y_1 < y_2 < \widetilde{t}_{k_j}$. Define $y_1 := f_j \alpha_j$ and $y_2 := g_j \alpha_j$. Next we show that $\widetilde{q}_{k_j}(y_1), \widetilde{q}_{k_j}(y_2) < 0$. Note that

$$\begin{aligned}\widetilde{q}_{k_j}(y_1) &= \sum_{i=0}^{k_j-1} \|A_i\|\alpha_j^i f_j^i - \|A_{k_j}^{-1}\|^{-1} f_j^{k_j}\alpha_j^{k_j} + \sum_{i=k_j+1}^{\ell} \|A_i\|\alpha_j^i f_j^i \\ &\leq \|A_{k_j}\|\alpha_j^{k_j}\left(\sum_{i=0}^{k_j-1} f_j^i - \frac{f_j^{k_j}}{\kappa(A_{k_j})} + \sum_{i=k_j+1}^{\infty} \delta_j^{i-k_j} f_j^i\right),\end{aligned}$$

since by Lemma 2.2 $\|A_i\|\alpha_j^i \leq \|A_{k_j}\|\alpha_j^{k_j}$ for $i < k_j$ and $\|A_i\|\alpha_j^i \leq \delta_j^{i-k_j}\|A_{k_j}\|\alpha_j^{k_j}$ for

$i > k_j$. Also $\delta_j f_j < 1$, by Lemma 2.6, which implies that

$$\sum_{i=k_j+1}^{\infty} \delta_j^{i-k_j} f_j^i = f_j^{k_j} \sum_{i=1}^{\infty} (\delta_j f_j)^i = f_j^{k_j} \frac{\delta_j f_j}{1-\delta_j f_j} = \frac{f_j^{k_j}}{g_j - 1},$$

recalling that $g_j = (\delta_j f_j)^{-1}$. This yields

$$\begin{aligned}\widetilde{q}_{k_j}(y_1) &\le \|A_{k_j}\| \alpha_j^{k_j} \left(\frac{f_j^{k_j} - 1}{f_j - 1} - \frac{f_j^{k_j}}{\kappa(A_{k_j})} + \frac{f_j^{k_j}}{g_j - 1} \right) \\ &< \|A_{k_j}\| \alpha_j^{k_j} f_j^{k_j} \left(\frac{1}{f_j - 1} - \frac{1}{\kappa(A_{k_j})} + \frac{1}{g_j - 1} \right).\end{aligned}$$

By Lemma 2.6, $\frac{1}{f_j-1} + \frac{1}{g_j-1} = \frac{1}{\kappa(A_{k_j})}$, which implies that $\widetilde{q}_{k_j}(y_1) < 0$. The proof for y_2 is similar to the one given above so we skip it. ∎

Let

$$\mathcal{H} = \{h_0, h_1, \ldots, h_p\} \tag{3.3}$$

be the set of all indices in $\{0, 1, \ldots, \ell\}$ such that
(a) $0 = h_0 < h_1 < \cdots < h_p = \ell$,
(b) A_{h_k} is nonsingular for $0 < k < p$,
(c) $q_{h_k}(x)$, $0 < k < p$ has two positive real roots $s_{h_k} \le t_{h_k}$.
We define

$$\widetilde{\mathcal{H}} = \{\widetilde{h}_0, \widetilde{h}_1, \ldots, \widetilde{h}_r\} \subseteq \{0, 1, \ldots, \ell\} \tag{3.4}$$

similarly to $\mathcal{H}$ but with $\widetilde{q}_k(x)$ in place of $q_k(x)$.

Theorem 3.4. *Let $\mathcal{K}$, $\widetilde{\mathcal{K}}$, $\mathcal{H}$ and $\widetilde{\mathcal{H}}$ be the set of indices defined by* (2.5), (2.17), (3.3) *and* (3.4). *Then,*

$$\widetilde{\mathcal{K}} \subseteq \widetilde{\mathcal{H}} \subseteq \mathcal{H} \subseteq \mathcal{K}.$$

Proof. The first two inclusions follow directly from Propositions 3.3 and 3.2. For the last inclusion, suppose that $k \notin \mathcal{K}$. Then there are $k_a < k < k_b$ such that k_a and k_b are two consecutive indices in $\mathcal{K}$. Modulo the appropriate scaling we may assume $\|A_{k_a}\| = \|A_{k_b}\| = 1$ and $\|A_k\| < 1$. Hence, if A_k is nonsingular then for all $x \ge 0$,

$$q_k(x) = \sum_{i \ne k} \|A_k^{-1} A_i\| x^i - x^k \ge \sum_{i \ne k} \frac{\|A_i\|}{\|A_k\|} x^i - x^k \ge \frac{x^{k_a} + x^{k_b}}{\|A_k\|} - x^k,$$

where we used $\|AB\| \ge \frac{\|B\|}{\|A^{-1}\|}$, which holds for any invertible A and any subordinate matrix norm. Therefore, if $0 < x < 1$, $\frac{q_k(x)}{x^{k_a}} \ge \frac{1}{\|A_k\|} - x^{k-k_a} > 1 - 1 = 0$. If $x = 1$, $q_k(1) > 1 + 1 - 1 > 0$. If $x > 1$, $\frac{q_k(x)}{x^k} \ge \frac{x^{b-k}}{\|A_k\|} - 1 > 1 - 1 = 0$. We conclude that $q_k(x) > 0$ for all $x > 0$. So $q_k(x)$ does not have a real positive root and hence $k \notin \mathcal{H}$. ∎

Theorem 3.4 shows that to compute the Pellet bounds we only need to construct the polynomials $q_k(x)$ and $\widetilde{q}_k(x)$ for $k \in \mathcal{K}$.

Let $s_0 = 0$ and let t_0 be the real positive root of $q_0(x)$ if A_0 is nonsingular and $t_0 = 0$ otherwise. Also, let $t_\ell = +\infty$ and let s_ℓ to be the real positive root of $q_\ell(x)$ if A_ℓ is nonsingular and set $s_\ell = +\infty$ otherwise. We define $\widetilde{s}_0, \widetilde{t}_0, \widetilde{s}_\ell$ and $\widetilde{t}_\ell$ similarly with respect to $\widetilde{q}_k(x)$, $k = 0, \ell$. It is shown in [7, Cor. 2.2] that $t_{h_j} \le s_{h_{j+1}}$ and it follows from Theorem 3.2 that $\widetilde{t}_{h_j} \le \widetilde{s}_{h_{j+1}}$. Then from Theorems 3.1–3.3 and (2.18) it follows that

$$\Lambda(P) \subseteq \bigcup_{j=0}^{p-1} \mathcal{A}(t_{h_j}, s_{h_{j+1}}) \subseteq \bigcup_{j=0}^{r-1} \mathcal{A}(\widetilde{t}_{\widetilde{h}_j}, \widetilde{s}_{\widetilde{h}_{j+1}}) \subset \bigcup_{i=0}^{m-1} \mathcal{A}(b_{j_i}, a_{j_{i+1}}) \subseteq \bigcup_{i=0}^{m-1} \mathcal{A}(\widetilde{b}_{j_i}, \widetilde{a}_{j_{i+1}}). \tag{3.5}$$

with $p \ge r \ge m$. In other words, (3.5) means that Bini et al.'s generalized Pellet theorem provides better eigenvalue localization results than Melman's generalized Pellet theorem, which in turn provides better localization results than our localization theorem based on tropical roots (see Theorem 2.7 and Corollary 2.8). However, the tropical roots and the results of Theorem 2.7 and its corollary remain interesting since these results can be easily interpreted and can be used in the numerical computation of the eigenvalues as we explain below. Importantly, the amount of information that Theorem 2.7 provides does not depend on the condition numbers of all the coefficients, but only on a selected number of them (A_k such that $k \in \mathcal{K}$). This fact can be used to give bounds on the sensitivity of the moduli of the eigenvalues when one coefficient $A_i, i \notin \mathcal{K}$, is perturbed. Even when the conditions of Theorem 2.7 are not satisfied, it can still happen that the tropical roots provide good approximations to the moduli of the eigenvalues. Indeed, they always lie inside the inclusion annuli defined by the generalized Pellet theorem, as we now show.

THEOREM 3.5. *Let $P(\lambda) = \sum_{i=0}^{\ell} A_i \lambda^i$ be a regular matrix polynomial. Also, for some $0 \le j \le p$ and some $0 \le i_1 < i_2 \le q$, let $h_j = k_{i_1}$ and $h_{j+1} = k_{i_2}$ be two consecutive indices in $\mathcal{H} \subseteq \mathcal{K}$, defined as in* (3.3) *and* (2.5)*. Then*

$$\{\alpha_{i_1+1}, \dots, \alpha_{i_2}\} \subset [t_{h_j}, s_{h_{j+1}}],$$

where $s_{h_j} \le t_{h_j}$ are the two positive real roots of $q_{h_j}(x)$ in (3.1)*.*

Proof. By (3.1) it holds $q_{h_j}(t_{h_j}) = 0$, implying that, for any index $c \ne h_j$, $(t_{h_j})^{h_j} = \sum_{i \ne h_j} \|A_{h_j}^{-1} A_i\| (t_{h_j})^i \ge \frac{\|A_c\|}{\|A_{h_j}\|} (t_{h_j})^c$. Therefore $(t_{h_j})^{h_j - c} \ge \frac{\|A_c\|}{\|A_{h_j}\|}$. If in particular $c > h_j$, then $t_{h_j} \le \left(\frac{\|A_{h_j}\|}{\|A_c\|}\right)^{\frac{1}{c-h_j}}$. Since $h_j = k_{i_1} \in \mathcal{K}$, taking $c = k_{i_1+1}$ and recalling (2.4) we obtain $t_{h_j} \le \alpha_{i_1+1}$. A similar argument shows that $s_{h_{j+1}} \ge \alpha_{i_2}$, and hence, $t_{h_j} \le \alpha_{i_1+1} < \cdots < \alpha_{i_2} \le s_{h_{j+1}}$. ∎

4. Numerical experiments and applications. We start with some experiments that illustrate the bounds of Sections 2.2 and 3 and show how well the tropical roots of $\mathsf{t}_\times(x)$ approximate the moduli of the eigenvalues of $P(\lambda)$. Our experiments were performed in MATLAB 7, for which the unit roundoff is $u = 2^{-53} \approx 1.1 \times 10^{-16}$.

Experiment 1. Our first example is a 20×20 quartic matrix polynomial $P(\lambda) = \sum_{i=0}^{4} \lambda^i A_i$ generated with the MATLAB commands

```
randn('state',48); n = 20;
A0 = 1e-5*randn(n); A1 = 1e2*randn(n); A2 = 1e2*randn(n);
A3 = 1e8*randn(n); A4 = 1e7*randn(n);
```

so as to have large variation in the norms of its coefficient matrices, the latter being fairly well conditioned (see Table 4.1). It follows from this table that the set of

Table 4.1: Norm and condition number of the coefficient matrices of $P(\lambda)$ in Experiment 1.

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$
$\|A_i\|_2$	7.5e-5	8.9e+2	8.6e+2	8.8e+8	7.7e+7
$\kappa_2(A_i)$	3.4e+1	3.0e+2	9.1e+1	1.3e+2	1.0e+2

Table 4.2: Pellet and tropical localizations of the spectrum $\Lambda(P)$ of $P(\lambda)$ in Experiment 1.

$\Lambda(P)$	20 eigenvalues in $\mathcal{A}(3.1\text{e-}8, 1.4\text{e-}6) =: \mathcal{A}_1$
	40 eigenvalues in $\mathcal{A}(1.3\text{e-}4, 2.5\text{e-}3) =: \mathcal{A}_2$
	20 eigenvalues in $\mathcal{A}(6.3\text{e-}1, 5.7\text{e+}1) =: \mathcal{A}_3$
Pellet 1	20 eigenvalues in $\mathcal{A}(4.3\text{e-}9, 1.6\text{e-}5) \supset \mathcal{A}_1$
	40 eigenvalues in $\mathcal{A}(6.2\text{e-}5, 8.4\text{e-}3) \supset \mathcal{A}_2$
	20 eigenvalues in $\mathcal{A}(2.1\text{e-}1, 6.1\text{e+}2) \supset \mathcal{A}_3$
Pellet 2	60 eigenvalues in $\mathcal{A}(2.4\text{e-}9, 1.2\text{e-}2) \supset \mathcal{A}_1 \cup \mathcal{A}_2$
	20 eigenvalues in $\mathcal{A}(8.8\text{e-}2, 1.2\text{e+}3) \supset \mathcal{A}_3$
Tropical	80 eigenvalues in $\mathcal{A}(2.4\text{e-}9, 1.2\text{e+}3) \supset \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$

abscissae of the Newton polygon associated with $\mathsf{t}_\times p(x)$ is $\mathcal{K} = \{0, 1, 3, 4\}$. Thus $\mathsf{t}_\times p(x)$ has three tropical roots,

$$\alpha_1 = \frac{\|A_0\|_2}{\|A_1\|_2} = 8.4\text{e-}8, \quad \alpha_2 = \left(\frac{\|A_1\|_2}{\|A_3\|_2}\right)^{1/2} = 1.0\text{e-}3, \quad \alpha_3 = \frac{\|A_3\|_2}{\|A_4\|_2} = 1.1\text{e+}1$$

of multiplicity one, two and one, respectively. The eigenvalues of $P(\lambda)$, which we computed with the MATLAB function `polyeig` are located in three separate annuli $\mathcal{A}_i$, $i = 1, 2, 3$ given in the first rows of Table 4.2. These are to be compared to the annuli from the Bini et al. generalized Pellet's theorem (see Theorem 3.1), Melman's version of the generalized Pellet's theorem (see [16, Thm. 3.3] or Theorem 3.1 with $\widetilde{q}_k(x)$ in place of $q_k(x)$), and that of Theorem 2.7 referred to as Pellet 1, Pellet 2, and Tropical, respectively, in Table 4.2. The generalized Pellet theorem identifies more annuli with $q_k(x)$ in (3.1) than with $\widetilde{q}_k(x)$ in (3.2), and the bounds provided by the former are tighter as expected from Proposition 3.2. Theorem 2.7 provides only a lower and upper bound for this particular example. It can be seen from Table 4.2 that the sets $\mathcal{H}, \widetilde{\mathcal{H}}$ and $\widetilde{\mathcal{K}}$, which are defined in (3.3),(3.4) and (2.17), are $\mathcal{H} = \{0, 1, 3, 4\}, \widetilde{\mathcal{H}} = \{0, 3, 4\}$ and $\widetilde{\mathcal{K}} = \{0, 4\}$ so that $\widetilde{\mathcal{K}} \subset \widetilde{\mathcal{H}} \subset \mathcal{H} \subseteq \mathcal{K}$ coherently with Theorem 3.4. Note that

$$0.4 \le \frac{|\lambda|}{\alpha_1} \le 17 \;\forall\; \lambda \in \mathcal{A}_1, \quad 0.1 \le \frac{|\lambda|}{\alpha_2} \le 2.5 \;\forall\; \lambda \in \mathcal{A}_2, \quad 0.06 \le \frac{|\lambda|}{\alpha_3} \le 5.0 \;\forall\; \lambda \in \mathcal{A}_3$$

so for this example, the tropical roots offer an order of magnitude approximation to the eigenvalues of $P(\lambda)$.

Experiment 2. Our next example is a class of matrix polynomials generated via

```
A0 = randn(n); A1 = 1e-3*randn(n); A2 = 1e3*randn(n);
A3 = 1e7*randn(n); A4 = 1e-3*randn(n);
```

for a given size $n > 3$ (but not too large). For this class of matrix polynomials, $\mathsf{t}_\times p(x)$ has only two tropical roots, a small root α_1 of multiplicity three and a large root α_2

of multiplicity one (for $n = 5$, $\alpha_1 = O(10^{-3})$ and $\alpha_2 = O(10^{10})$). Theorem 2.7 and Theorem 3.1 detect two annuli, one associated with α_1 containing $3n$ eigenvalues and one associated with α_2 and containing n eigenvalues. MATLAB's function `polyeig` does not find any eigenvalue in the largest annuli. Instead, it tends to return n eigenvalues at infinity. The leading coefficient A_4 is however generically nonsingular, so that there should be no eigenvalue at infinity.

The function `polyeig`, which solves the polynomial eigenvalue problem via linearization, is not numerically stable [12], [21]. The linearization process used in eigensolvers such as `polyeig` can also affect the sensitivity of the eigenvalues: a well conditioned eigenvalue for $P(\lambda)$ may be badly conditioned for the linearization [13]. As a result, `polyeig` can return eigenvalues with no digits of accuracy. We note that there is currently no eigensolver for dense matrix polynomials of degree $\ell > 2$ with guaranteed backward stability. With the aim of addressing this issue, Gaubert and Sharify [8] propose to solve q tropically scaled polynomial eigenvalue problems with the matrix polynomials $\widetilde{P}(\mu)$ in (2.6) scaled with α_i, $i = 1, \dots, q$. We recall below a version of [8, Alg. 1].

ALGORITHM 4.1. *Given* $P(\lambda) = \sum_{i=0}^{\ell} \lambda^i A_i \in \mathbb{C}[\lambda]^{n\times n}$, *the tropical roots* α_j *of* $\mathsf{t}_\times p(x) = \max_{0\le j\le \ell} \|A_j\| x^j$ *and their multiplicities* m_j, $j = 1, \dots, q$, *this algorithm computes the eigenvalues (and eigenvectors) of* P.

1 $k = 1$
2 for $j = 1 : q$
3   Scale $P(\lambda)$ into $\widetilde{P}(\mu)$ as in (2.6).
4   Solve $\widetilde{P}(\mu)x = 0$ with `polyeig` and scale back the eigenvalues, $\lambda_i = \alpha_j \mu_i$.
5   Sort the eigenvalues in modulus from small to large.
6   Keep $\lambda_k, \dots, \lambda_{k+nm_j-1}$ and the corresponding eigenvectors.
7     $k = k + nm_j$.
8 end

Gaubert and Sharify show experimentally that their algorithm tends to compute eigenpairs with smaller backward errors (see (4.1)) than those computed with the classical approach (i.e. without tropical scaling). Sharify and Tisseur [20] show that amongst the eigenpairs returned by Algorithm 4.1, those with eigenvalues of modulus within order one of α_i are computed with small backward errors and their condition numbers is not affected by the linearization process.

We note that Algorithm 4.1 is q times more expensive that `polyeig`, where $q \le \ell$ is the number of distinct tropical roots of $\mathsf{t}_\times q(x)$ but it has better numerical stability properties than `polyeig` and that it delivers more accurate eigenpairs. It is outside the scope of this paper to develop an efficient eigensolver and also to justify the selection of the computed eigenpairs (see lines 5–6 of Algorithm 4.1).

To illustrate the behaviour of Algorithm 4.1, we measure the backward error $\eta_p(\lambda, x)$ for a computed eigenpair (λ, x) of $P(\lambda)$ with λ finite and nonzero, with the scaled residual [21]

$$\eta_P(\lambda, x) = \frac{\|P(\lambda)x\|_2}{\left(\sum_{i=0}^{\ell} |\lambda|^i \|A_i\|_2\right)\|x\|_2}. \tag{4.1}$$

We consider the backward error to be small if $\eta(\lambda, x) \le (\ell n)u$. To measure the sensitivity of a simple, finite and nonzero eigenvalue λ of $P(\lambda) = \sum_{i=0}^{\ell} \lambda^i A_i \in \mathbb{C}[\lambda]^{n\times n}$ and of a linearization $L(\lambda) = A + \lambda B \in \mathbb{C}[\lambda]^{n\ell\times n\ell}$ of P we use the condition numbers

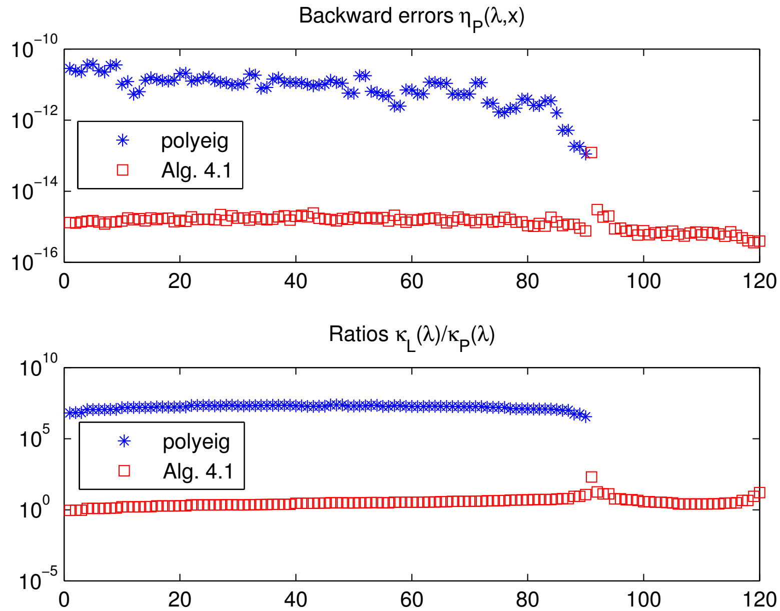

Fig. 4.1: Problem from Experiment 2. Backward errors of computed eigenpairs and ratios between the eigenvalue condition numbers.

[21]

$$\kappa_P(\lambda) = \frac{\left(\sum_{i=0}^{\ell} |\lambda|^i \|A_i\|_2\right)\|x\|_2\|y\|_2}{|\lambda||y^* P'(\lambda)x|}, \quad \kappa_L(\lambda) = \frac{\left(\|A\|_2 + |\lambda|\|B\|_2\right)\|z\|_2\|w\|_2}{|\lambda||w^* Bz|}, \quad (4.2)$$

where x, y are right and left eigenvectors of P with eigenvalue λ and z, w are right and left eigenvectors of L with eigenvalue λ. Ideally, we would like the linearization L of P to be such that $\kappa_L(\lambda) \approx \kappa_P(\lambda)$.

Experiment 3. The top plot in Figure 4.1 shows the backward errors for the computed eigenpairs via `polyeig` and Algorithm 4.1 for $P(\lambda)$ generated as in Experiment 2 by setting $n = 30$ and `randn('state',0)`. The bottom plot displays the ratios between the condition number $\kappa_L(\lambda)$ of λ as an eigenvalue of L and the condition number $\kappa_P(\lambda)$ of λ as an eigenvalue of $P(\lambda)$. In our figure, the x-axis is the eigenvalue index and the eigenvalues are sorted in increasing order of absolute value. Since `polyeig` wrongly returns 30 eigenvalues at infinity and the backward error in (4.1) and condition numbers in (4.2) are not defined at infinity, $\eta_P(\lambda, x)$ and $\kappa_L(\lambda)/\kappa_P(\lambda)$ are not plotted for these eigenvalues. The top plot shows that none of the eigenpairs returned by `polyeig` have a small backward error whereas Algorithm 4.1 returns all eigenpairs with a backward error close to the unit roundoff except for the $3n + 1 = 91$st eigenvalue $|\lambda_{91}| = 7.5 \times 10^8$ for which $\eta_P(\lambda_{91}, x_{91}) = 1.2 \times 10^{-13}$. The matrix polynomial $P(\lambda)$ has two tropical roots $\alpha_1 = 4.5 \times 10^{-3}$ with multiplicity 3 and $\alpha_2 = 10^{10}$ with multiplicity 1. The largest tropical root α_2 does not quite provide an order of magnitude approximation to λ_{91} and tropical scaling with α_2 yields a slightly too large backward error for the eigenpair (λ_{91}, x_{91}).

The linearization used by `polyeig` is the reversal of the first companion linearization of the reversal of $P(\lambda)$ defined by $\text{rev}P(\lambda) = \lambda^\ell(P(1/\lambda)$. The bottom plot

Table 4.3: Size n, degree ℓ and norm of the coefficient matrices for the `butterfly` and `orr_sommerfeld` problems.

Problem	n		$i=0$	$i=1$	$i=2$	$i=3$	$i=4$
`butterfly`	9	$\|A_i\|_2$	1.0e+0	5.8e+3	1.7e+6	2.4e+7	2.0e+12
		$\kappa_2(A_i)$	2.1e+0	∞	5.8e+0	∞	5.8e+0
`orr_sommerfeld`	64	$\|A_i\|_2$	1.8e+2	2.8e-2	6.8e+2	2.8e+0	6.8e-3
		$\kappa_2(A_i)$	1.0e+0	4.3e+2	1.1e+3	∞	5.0e+8

shows that, for this example, when no scaling is applied to $P(\lambda)$, the linearization process increases the eigenvalue condition numbers by a factor 10^7 but when scaling is used such as in Algorithm 4.1, then $\kappa_L(\lambda) \approx \kappa_P(\lambda)$.

Experiment 4. We consider two quartics from the NLEVP collection of nonlinear eigenvalue problems [5] namely the `butterfly` problem and the `orr_sommerfeld` problem. The coefficient matrices of the `butterfly` problem are generated as follows:

```
c = kron([1e2 1e-2 1e2 1 1e-3],[1 1]);
coeffs = nlevp('butterfly',9,c);
A0 = coeffs{1}; A1 = coeffs{2}; A2 = coeffs{3};
A3 = coeffs{3}; A4 = coeffs{4};
```

Both problems have variations in the norms of their coefficient matrices as shown in Table 4.3. The moduli of the eigenvalues of these matrix polynomials, the tropical roots of $\mathsf{t}_\times p(x)$ as well as the intervals from the generalized Pellet theorem (Theorem 3.1) which contain the moduli of the eigenvalues of P are all plotted in Figure 4.2. The backward errors for eigenpairs computed with `polyeig` and Algorithm 4.1 are plotted in Figure 4.3, and the ratios between the condition number $\kappa_L(\lambda)$ of λ as an eigenvalue of the linearization L used by the eigensolvers and the condition number $\kappa_P(\lambda)$ of λ as an eigenvalue of $P(\lambda)$ are shown in Figure 4.4.

For the `butterfly` problem, $\mathcal{K} = \{0, 2, 3, 4\}$ (see (2.5)) and since A_3 is singular, Theorem 2.7 and the generalized Pellet's theorems identify two annuli, one containing $2n = 18$ eigenvalues with magnitude around $\alpha_1 = \left(\|A_0\|_2/\|A_2\|_2\right)^{1/2} \approx 5.1 \times 10^{-1}$, and the second and wider annulus containing the remaining $2n$ eigenvalues and the two tropical roots $\alpha_2 = \|A_2\|_2/\|A_3\|_2 \approx 2.4 \times 10^2$ and $\alpha_3 = \|A_3\|_2/\|A_4\|_2 \approx 4.1 \times 10^2$. The top plot in Figure 4.2 shows that the three tropical roots associated with the `butterfly` problem are good approximations to the magnitude of the eigenvalues. As a consequence of this and the analysis in [20], the eigenpairs computed by Algorithm 4.1 have small backward errors (see top plot in Figure 4.3) and the linearization process used by the eigensolver does not increase the eigenvalue condition numbers (see top of Figure 4.4). We note that `polyeig` returns eigenpairs with backward errors as large at 10^{-10} and the linearization process increases the eigenvalue condition numbers by a factor 10^{10} for the $2n$ largest eigenvalues.

For the `orr_sommerfeld` problem, Theorem 2.7 and the generalized Pellet's theorems identify only one annulus. The tropical roots do not offer order of magnitude approximations to all the eigenvalues, in particular for the largest ones (see bottom of Figure 4.2). Nevertheless, Algorithm 4.1 returns eigenpairs with backward errors all less than $10^{-13} \approx 2(\ell n)u$, whereas those returned by `quadeig` can be as large at 10^{-4} (see bottom plot in Figure 4.3). The linearization process used by the eigensolvers increases the eigenvalue condition numbers by a factor at most 10^5 for Algorithm 4.1 and up to 10^{15} for `polyeig` (see bottom plot in Figure 4.4).

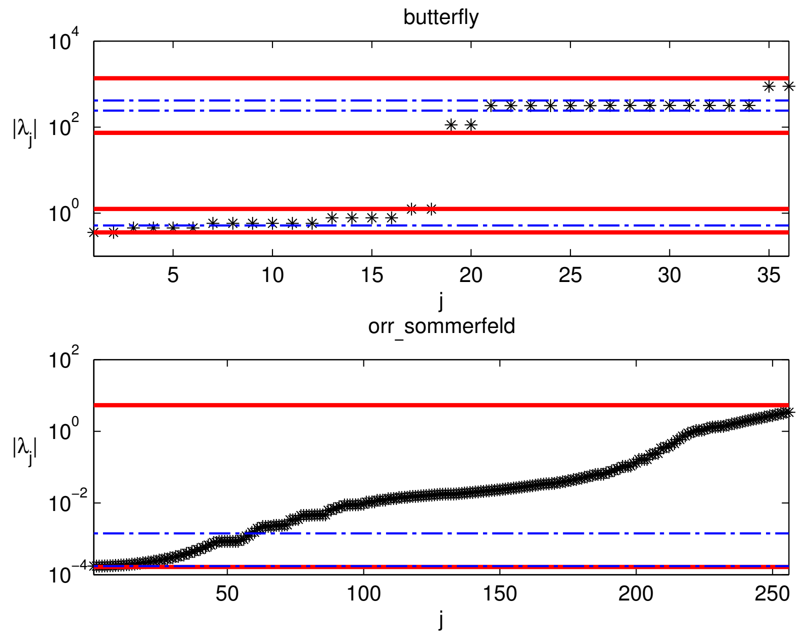

Fig. 4.2: Butterfly and Orr-Sommerfeld problems. The moduli of the eigenvalues are plotted as '$*$'. The thick red lines define intervals $[t_{h_j}, s_{h_j}]$ from the generalized Pellet theorem containing the $|\lambda_j|$. The blue dash-dotted lines indicate the value of the tropical roots.

5. Concluding remarks. We have identified sufficient conditions under which the tropical roots of $\mathsf{t}_\times p(x) = \max_{0\le i\le \ell}\left(\|A_i\| x^i\right)$ are good order of magnitude approximations to the eigenvalues of $P(\lambda) = \sum_{i=0}^{\ell} \lambda^i A_i$. These tropical roots are interesting from the numerical point of view since they are cheap to compute and can be used to define a family of eigenvalue parameter scalings for matrix polynomials that can both improve the backward stability of polynomial eigensolvers based on linearizations and help not to increase the eigenvalue condition numbers of the linearized problem (see Section 4). This is confirmed by the analysis in [20]. We anticipate that these tropical roots will help designing a more numerically stable version of `polyeig`.

REFERENCES

[1] M. Akian, R. Bapat, and S. Gaubert, *Min-plus methods in eigenvalue perturbation theory and generalised Lidskii-Vishik-Ljusternik theorem*, Technical Report arXiv:math/0402090v3, 2006.

[2] M. Akian, S. Gaubert, and M. Sharify, *Log-majorization of the moduli of the eigenvalues of a matrix polynomial by tropical roots*, Technical Report arXiv:1304.2967v2, 2014.

[3] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, and K. Kimura, *A numerical method for polynomial eigenvalue problems using contour integral*, Japan J. Indust. Appl. Math., 27 (2010), pp. 73–90.

[4] T. Betcke, *Optimal scaling of generalized and polynomial eigenvalue problems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1320–1338.

[5] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), pp. 7:1–7:28.

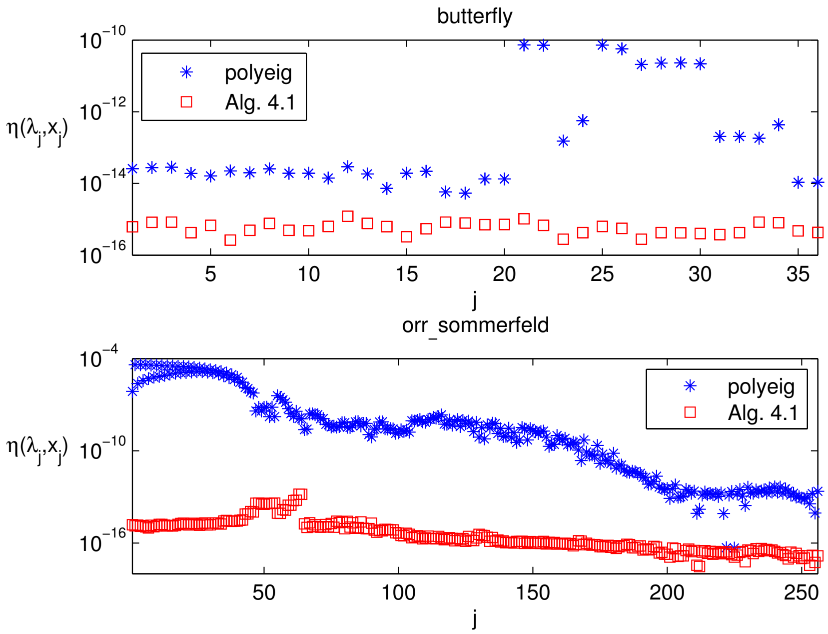

Fig. 4.3: Butterfly and Orr-Sommerfeld problems. Backward errors for computed eigenpairs.

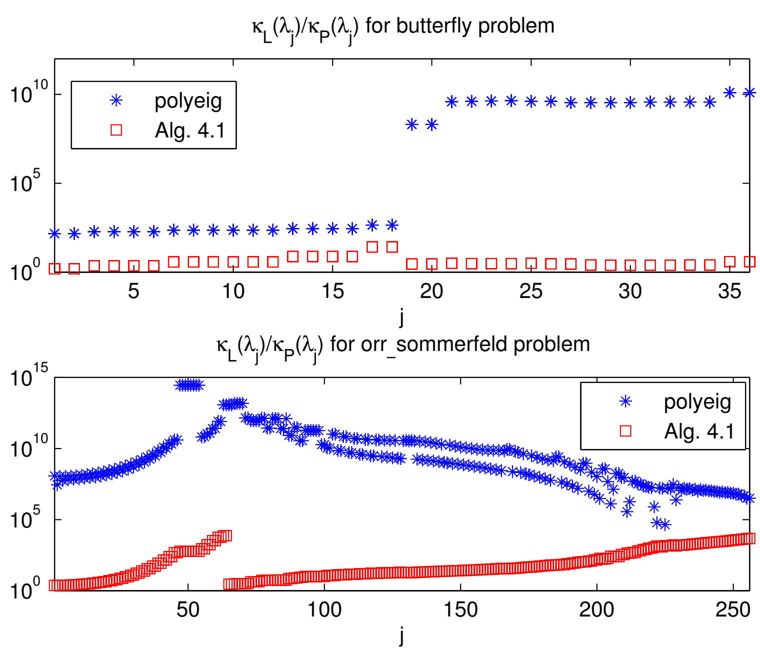

Fig. 4.4: Butterfly and Orr-Sommerfeld problems. Ratios between eigenvalue condition numbers for linearization L and eigenvalue condition numbers for P.

[6] D. A. Bini and V. Noferini, *Solving polynomial eigenvalue problems by means of the Ehrlich-Aberth method*, Linear Algebra Appl., 439 (2013), pp. 1130–1149.

[7] D. A. Bini, V. Noferini, and M. Sharify, *Locating the eigenvalues of matrix polynomials*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1708–1727.

[8] S. Gaubert and M. Sharify, *Tropical scaling of polynomial matrices*, in Positive systems, vol. 389 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 2009, pp. 291–303.

[9] I. Gohberg and E. Sigal, *An operator generalization of the logarithmic residue theorem and Rouché's theorem*, Math. USSR Sb., 13 (1971), pp. 603–625.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, USA, fourth ed., 2013.

[11] R. L. Graham, *An efficient algorithm for determining the convex hull of a finite planar set*, Inf. Proc. Lett., 1 (1972), pp. 132–133.

[12] N. J. Higham, R.-C. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1218–1241.

[13] N. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1005–1028.

[14] N. J. Higham and F. Tisseur, *Bounds for eigenvalues of matrix polynomials*, Linear Algebra Appl., 358 (2003), pp. 5–22.

[15] J. Hook and F. Tisseur, *Tropical eigenvalues and singular values*, MIMS EPrint 2014.6, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Feb. 2014.

[16] A. Melman, *Generalization and variations of Pellet's theorem for matrix polynomials*, Linear Algebra Appl., 439 (2013), pp. 1550–1567.

[17] A. Ostrowski, *Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent*, Acta Math., 72 (1940), pp. 99–155.

[18] ———, *Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent. Chapitres III et IV*, Acta Math., 72 (1940), pp. 157–257.

[19] M. Sharify, *Scaling Algorithms and Tropical Methods in Numerical Matrix Analysis: Application to the Optimal Assignment Problem and to the Accurate Computation of Eigenvalues*, PhD thesis, Ecole Polytechnique, Palaiseau, France, Sept. 2011.

[20] M. Sharify and F. Tisseur, *Effect of tropical scaling on linearizations of matrix polynomials: backward error and conditioning*, MIMS EPrint, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2014. In preparation.

[21] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.