

Tropical methods for the localization of eigenvalues and application to their numerical computation

Andrea Marchesini

▶ **To cite this version:**

Andrea Marchesini. Tropical methods for the localization of eigenvalues and application to their numerical computation. Spectral Theory [math.SP]. Ecole polytechnique X, 2015. English. NNT : . tel-01285110

HAL Id: tel-01285110
https://pastel.hal.science/tel-01285110

Submitted on 8 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License







Thèse présentée pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Spécialité : Mathématiques Appliquées

par

Andrea Marchesini

Tropical methods for the localization of eigenvalues

and application to their numerical computation

soutenue le 15 décembre 2015 devant le jury composé de :

Marianne Akian	INRIA - CMAP École polytechnique	Directeur de thèse
Peter Butkovič	University of Birmingham	Rapporteur
Bernadette Charron-Bost	École polytechnique	Examinateur
Stéphane Gaubert	INRIA - CMAP École polytechnique	Co-directeur de thèse
Laurence Grammont	Université de Saint-Étienne	Examinateur
Françoise Tisseur	University of Manchester	Examinateur
Paul Van Dooren	Université catholique de Louvain	Rapporteur

Abstract

In this thesis we use tropical mathematics to locate and numerically compute eigenvalues of matrices and matrix polynomials. The first part of the work focuses on eigenvalues of matrices, while the second part focuses on matrix polynomials and adds a numerical experimental side along the theoretical one.

By "locating" an eigenvalue we mean being able to identify some bounds within which it must lie. This can be useful in situations where one only needs approximate eigenvalues; moreover, they make good starting values for iterative eigenvalue-finding algorithms. Rather than full location, our result for matrices is in the form of majorization bounds to control the absolute value of the eigenvalues. These bounds are to some extent a generalization to matrices of a result proved by Ostrowski for polynomials: he showed (albeit with different terminology) that the product of the k largest absolute values of the roots of a polynomial can be bounded from above and below by the product of its k largest tropical (max-times) roots, up to multiplicative factors which are independent of the coefficients of the polynomial. We prove an analogous result for matrices: the product of the k largest absolute values of eigenvalues is bounded, up to a multiplicative factor, by the product of the k largest tropical eigenvalues. It should be noted that tropical eigenvalues can be computed by using the solution to a parametric optimal assignment problem, in a way that is robust with respect to small perturbations in the data. Another thing worth mentioning is that the multiplicative factor in the bound is of combinatorial nature and it is reminiscent of a work by Friedland, who essentially proved a specialization of our result to the particular case $k = 1$ (i.e. for the largest eigenvalue only). We can interpret the absolute value as an archimedean valuation; in this light, there is a correspondence between the present result and previous work by Akian, Bapat and Gaubert, who dealt with the same problem for matrices over fields with non-archimedean valuation (specifically Puiseux series, with the leading exponent as valuation) and showed in that case more stringent bounds, with no multiplicative factor, and with generic equality rather than upper and lower bounds.

The second part of the thesis revolves around the computation of eigenvalues of matrix polynomials. For linear matrix polynomials, stable algorithms such as the QZ method have been known for a long time. Eigenproblems for matrix polynomials of higher degree are usually reduced to the linear case, using a linearization such as the companion linearization. This

however can worsen the condition number and backward error of the computed eigenvalue with respect to perturbations in the coefficients of the original polynomial (even if they remain stable in the coefficients of the linearized). To mitigate this inconvenience it is common to perform a scaling of the matrix polynomial before linearizing. Various scaling methods have been proposed. In our work, we introduce a two-sided diagonal scaling strategy based on the tropical eigenvalues of the matrix polynomial obtained by taking entrywise valuation of the original one (and we will consider both the archimedean and non-archimedean case). We study the effect of this scaling on the conditioning and backward error, with both analytic formulas and numerical examples, showing that it can increase the accuracy of the computed eigenvalues by several orders of magnitude.

Résumé

Dans cette thèse nous utilisons des outils d'algèbre tropicale pour localiser et calculer de façon numérique les valeurs propres de matrices et de polynômes matriciels. La première partie porte sur les valeurs propres de matrices, la deuxième se concentre sur les polynômes matriciels tout en rajoutant à l'étude théorique un coté numérique.

"Localiser" une valeur propre veut dire pouvoir identifier des bornes entre lesquelles elle se trouve. Cela peut être utile dans des situations où l'on n'a besoin que de valeurs propres approximées; ces approximations permettent aussi d'obtenir de bons points d'initialisation d'algorithmes itératifs de calcul des valeurs propres. Notre résultat pour les matrices prend la forme d'inégalités de type majorisation qui contrôlent le module des valeurs propres. Ces bornes peuvent être vues comme une généralisation au cas matriciel d'un résultat prouvé par Ostrowski pour les polynômes : il a montré (en utilisant une terminologie différente) que le produit des k plus grands modules des racines d'un polynôme sont bornées inférieurement et supérieurement par le produit de ses k plus grandes valeurs propres tropicales, à un facteur multiplicatif près qui est indépendent des coefficients du polynôme. Nous prouvons un résultat analogue pour le cas d'une matrice : le produit des k plus grands modules des valeurs propres est borné, à un facteur multiplicatif près, par le produit des k plus grandes valeurs propres tropicales. On notera que les valeurs propres tropicales peuvent être calculées au moyen de la solution d'un problème d'affectation optimale paramètrique, et ceci de façon stable par rapport à des perturbations des données. On notera aussi que le facteur multiplicatif est de nature combinatoire, et qu'il est inspiré d'un résultat de Friedland, lequel a démontré notre inégalité dans le cas particulier $k = 1$. On peut interpréter le module comme une valuation archimédienne; ainsi, il y a une correspondance entre le résultat présenté ici et un travail précédent d'Akian, Bapat et Gaubert, qui ont traité le même problème pour des matrices à coefficients dans un corps avec une valuation non-archimédienne (notamment le corps des séries de Puiseux, équipé de la valuation donnée par l'exposant dominant) et qui ont montré des bornes supérieures plus sérées pour ce cas, sans facteur multiplicatif, et avec égalité générique à la place de bornes supérieures et inférieures.

La deuxième partie de la thèse traite du calcul des valeurs propres de polynômes matriciels. Pour des polynôme matriciels linéaires, des algorithmes stables tels que la méthode QZ sont connus dépuis longtemps. L'approche pour les polynômes de degré supérieur con-

siste souvent à se ramener au cas linéaire, à l'aide de linéarisations telles que la linéarisation compagnon. Cela peut néanmoins dégrader le conditionnement et l'erreur inverse des valeurs propres calculées, par rapport aux coefficients du polynôme originel (même s'ils restent stables par rapport au linéarisé). Pour faire face à cet inconvénient il est commun de procéder à un changement d'échelle avant de linéariser le polynôme. Plusieurs techniques de changement d'échelle ont été proposées. Dans notre travail, nous introduisons un changement d'échelle par multiplication diagonale à gauche et à droite basée sur les valeurs propres tropicales du polynôme matriciel obtenu en prenant la valuation (archimédienne ou non, selon le cas) de chaque coefficient du polynôme originel. Nous étudions l'effet de ce scaling sur le conditionnement et sur l'erreur inverse, en obtenant des formules analytiques ainsi qu'en donnant des exemples numériques, et nous montrons que la précision des valeurs propres calculées peut être améliorée de plusieurs ordres de grandeur.

Acknowledgements

If I am defending this thesis, most of the credit goes to my advisors Marianne and Stéphane. Not only they provided invaluable mathematical help, but they also managed to put up with my laziness and motivated me all along. It has been a pleasure working with them.

I am indebted to Peter Butkovič and Paul Van Dooren for reviewing this manuscript, providing insightful remarks and suggestions to improve it. I not-so-proudly claim full paternity of all remaining errors.

Bernadette Charron-Bost, Laurence Grammont and Françoise Tisseur have honored me by accepting to sit in my thesis committee, and to them goes my gratitude. In particular I would like to thank Françoise, as well as James Hook, for the warm hospitality during my stays in Manchester and for the fruitful discussions that they sparked.

I must also thank the CMAP laboratory and Inria, both for the financial support and for their amazing administrative staff: without Nasséra, Alexandra, Jessica and all the others who helped me whenever I was lost in paperwork, this thesis would not have been possible.

I owe a lot to all my fellow PhD students, for the conversations, the tips, the laughs, the football matches, and in general for making CMAP a pleasant place. They are the last of a long list of friends who have been by my side throughout this long journey, from back in Verona, to Pisa, to Paris. It would be impossible to name all of them, but the ridiculously huge amount of happy memories that I have thanks to them is something that gives me shivers.

Finally I want to thank my parents, my sisters, my brother and all my family, for their love and support. This thesis is dedicated to them.

Contents

1 Introduction **1**

2 Preliminaries **5**
2.1 Basic elements of tropical algebra 5
2.1.1 Tropical polynomials 6
2.1.2 Tropical matrices and eigenvalues 9
2.2 Eigenvalues of matrix polynomials 11
2.2.1 Desirable properties of algorithms 12
2.2.2 Conditioning and backward error for the polynomial eigenvalue problem 13
2.2.3 Solving polynomial eigenvalue problem via linearization 13
2.2.4 Scaling 14
2.3 Tropical algebra and valuation theory 16
2.3.1 Valuation and roots 17
2.3.2 Valuations and eigenvalues 18

3 Majorization bounds for eigenvalues **19**
3.1 Upper bound 19
3.1.1 Friedland's Theorem 20
3.1.2 Spectral radius of exterior powers 20
3.1.3 Proof of Theorem 3.1 23
3.2 Lower bound 23
3.3 Optimality of upper bound and comparison with bounds for polynomial roots 26
3.3.1 Monomial matrices. 26
3.3.2 Full matrices. 26
3.3.3 Comparison with the Hadamard-Pólya's bounds for polynomial roots. 27

4 Hungarian scaling of asymptotic matrix polynomials **31**
4.1 Asymptotics of eigenvalues and parametric optimal assignment problems . . . 32
4.2 Preliminaries on Hungarian pairs 35

4.3 Asymptotics of eigenvectors 38
4.4 Asymptotics of eigenvalue condition numbers 45
4.5 Some general bounds on eigenvalue condition numbers 51
4.6 Eigenvalue Condition numbers of linearizations 57
4.7 Numerical examples 64

CHAPTER 1

Introduction

In this thesis we use tropical algebra to locate and numerically compute eigenvalues of matrices and matrix polynomials. Tropical algebra (named this way in honor of Brazilian mathematician Imre Simon, who pioneered the field in the '70s) was initially developed to tackle problems in theoretical computer science, such as automata theory, and it is essentially mathematics done on an idempotent semiring where max (or min) and + play the role of addition and multiplication, respectively, whence the alternative names of *max-plus* or *min-plus* algebra.

The first part of the present work focuses on the interplay between tropical algebra and eigenvalues of matrices. By "locating" an eigenvalue we mean being able to identify some bounds within which it must lie. This can be useful in situations where one only needs approximate eigenvalues; moreover, they make good starting values for iterative eigenvalue-finding algorithms, such as the Ehrlich-Aberth method [BNS13]. For the case of roots of complex polynomials, numerous classical bounds are known. Rather than full location, they are usually in the form of bounds on the absolute value of the roots. One notable instance of such bounds, from which we drew inspiration, is a result by Ostrowski [Ost40]. The tool that he used to prove his bounds is known as the *Newton polygon*. Given a complex polynomial $p = a_n z^n + \cdots + a_1 z + a_0$, its Newton polygon is the polygonal line obtained as the convex hull of the set $\{(0, -\log|a_0|), \ldots, (n, -\log|a_n|)\}$ in the two-dimensional plane. Ostrowski (building on previous work by Hadamard, and with the collaboration of Pólya) showed that the exponentials of the opposite of the slopes of the Newton polygon can be seen as approximations of the moduli of the roots of p. In modern terminology, the opposites of the slopes are called the *tropical roots* of the polynomial represented by the Newton polygon, and Ostrowski's result is best enounced as a majorization inequality: if $\zeta_1, \ldots, \zeta_n$ are the

roots of p, ordered by decreasing magnitude, and if $\alpha_1 \geq \cdots \geq \alpha_n$ are the exponentials of the tropical roots, then

$$\frac{1}{\binom{n}{k}}\alpha_1 \cdots \alpha_k \leq |\zeta_1 \cdots \zeta_k| \leq \sqrt{e(k+1)}\alpha_1 \cdots \alpha_k \qquad \forall k \in \{1, \ldots, n\} .$$

In [AGM14] we generalized Ostrowski's result to the matrix case, giving majorization bounds to control the absolute value of the eigenvalues of a matrix by means of its *tropical eigenvalues*. Tropical eigenvalues are defined in an algebraic fashion, as the tropical roots of a polynomial which is constructed to be the tropical transposition of the usual characteristic polynomial of a matrix; as it turns out, tropical eigenvalues can be thought of as solutions to an optimal assignment problem, and as such they can be computed in polynomial time ($O(n^3)$, see [GK10]) and in a way that is robust with respect to small perturbations in the data. The multiplicative factors in the bounds are of combinatorial nature, based on exterior algebra and on maximal mean cycles in a graph, being reminiscent in this respect of work by Friedland [Fri86], who essentially proved a specialization of our result to the particular case k=1 (i.e. a relation between the spectral radius of a matrix, and a quantity that will prove to be its largest tropical eigenvalue).

All bounds mentioned so far involve the absolute value, which we can interpret as an archimedean valuation; in this light, there is a correspondence between the present result and previous work by Akian, Bapat and Gaubert [ABG04], who dealt with the same problem for matrices over fields with non-archimedean valuation. More specifically, their model of non-archimedean field was the field of Puiseux series, with the leading exponent as valuation. In the non-archimedean case results are more exact: rather than having upper and lower bounds, they showed that if $\mathcal{A}$ is a matrix and A is the matrix obtained by taking the valuation entrywise, then the tropical eigenvalues of A coincide with the valuations of the eigenvalues of $\mathcal{A}$.

The second part of the thesis revolves around the computation of eigenvalues of matrix polynomials. For linear matrix polynomials, stable algorithms such as the QZ method [MS73] have been known for a long time. Eigenproblems for matrix polynomials of higher degree are usually reduced to the linear case; linearizations work by transforming a matrix polynomial of size n and degree d into a linear matrix polynomial of size nd with the same spectrum. While being theoretically equivalent, the two problems can have a very different numerical behavior. In fact, the condition number and backward error of a computed eigenvalue both deal with perturbations in the coefficients of the input data, not of the results: different inputs can thus lead to different values for these two measures of stability, often making them worse, especially in case of matrix polynomials with entries of different orders of magnitude. To mitigate this inconvenience it is common to perform a scaling of the matrix polynomial before linearizing. Various scaling methods have been proposed. The first kind of methods involves a scaling in the variable of the matrix polynomial: to compute the eigenvalues of $P(\lambda)$ one can compute the eigenvalues of $Q(\mu) = P(\alpha\mu)$ and then multiply them by α to obtain eigenvalues for the original polynomial; the choice of the parameter α influences the effectiveness of this method. Fan, Lin and Van Dooren [FLVD04] have proposed such a method for a quadratic matrix polynomial, choosing α to bring the norm of the matrices close to a constant (and then further dividing by this constant to get close to 1). Their method was generalized by Gaubert and Sharify [GS09] to allow application to higher degree polynomials via some sort of "patchwork scaling", whereby multiple scalings are performed successively, each one of them used to retrieve a portion of the eigenvalues. Another kind

of scaling is what is known as *balancing*, inspired from the practice of balancing a matrix, i.e. finding a diagonal change of coordinates which makes the norm of each i-th row equal to the norm of the corresponding i-th column. This technique has been adapted to matrix polynomial in several ways [War81, LD06, Bet08].

In a joint work with Akian, Gaubert, Hook and Tisseur [AGH+15], we combine these two kinds of scaling and we introduce a two-sided diagonal scaling strategy based on the tropical eigenvalues of the matrix polynomial obtained by taking entrywise valuation of the original one (and we will consider both the archimedean and non-archimedean case). The ideas behind this scaling come from the aforementioned work by Akian, Bapat and Gaubert, and they are based on solving a parametric optimal assignment problem by means of dual variables, using what is known as the Hungarian method. For this reason we called this the Hungarian scaling. We study its effect on the conditioning and backward error both from an analytic and a numerical point of view. We derive formulas for the before- to after-linearization ratios of condition number and backward error, and finally we show the results of numerical simulations performed in MATLAB showing that the accuracy of the computed eigenvalues can be increased by several orders of magnitude.

CHAPTER 2

Preliminaries

2.1 Basic elements of tropical algebra

The *max-plus semiring* $\mathbb{R}_{\max}$ is the set $\mathbb{R} \cup \{-\infty\}$ endowed with an addition $a \oplus b = \max(a, b)$ and a multiplication $a \otimes b = a + b$. The identity elements are $\mathbb{0} = -\infty$ for the addition and $\mathbb{1} = 0$ for the multiplication. Addition as we just defined it is not invertible (hence the "semi" in semiring); as a consequence, there is no such thing as max-plus subtraction.

The *max-times semiring* $\mathbb{T}$ is the set $\mathbb{R}_{\geq 0}$ endowed with the maximum as addition $a \oplus b = \max(a, b)$ and the usual multiplication $a \otimes b = ab$.

Note that the max-times semiring is the isomorphic image of the max-plus semiring under the logarithmic map (extended so that $\log(-\infty) = 0$). The isomorphism allows us to identify $\mathbb{R}_{\max}$ and $\mathbb{T}$ as the same algebraic object, and to refer to either of them as the *tropical semiring*.

Other forms of the tropical semiring notably include the min-plus semiring. We refer the reader to [CG79, But10, HOW06, Pin98] for more information on various tropical semirings.

Tropical semirings were initially introduced to solve applied problems arising from scheduling (see [CG94]), discrete optimization (see for instance [GM84]), or discrete event systems [BCOQ92], or to solve decision problems in automata and language theory [Sim78]. Since 2000, they have been extensively studied in pure mathematics. In particular, tropical structures have been used to develop combinatorial methods in algebraic geometry [GKZ94, Vir01, Mik04, EKL06, SS04, IMS07].

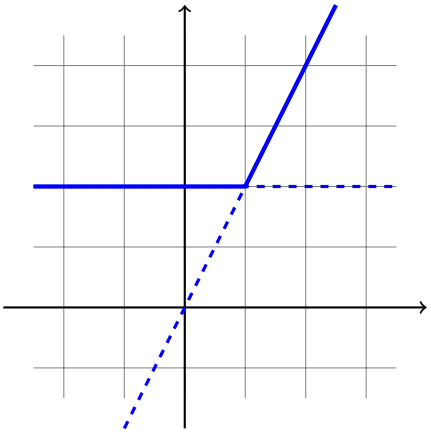
2.1.1 Tropical polynomials

The set of formal polynomials over $\mathbb{R}_{\max}$ (or $\mathbb{T}$) can be defined naturally and it is itself a semiring once we provide it with the usual formal addition and multiplication between polynomials. The notions of degree and valuation still make sense over the tropical semiring, and we shall use the standard constructions and notations with obvious modifications: if $p = a_0 + \cdots + a_d X^d$, $\deg p$ (resp. $\operatorname{val} p$) denotes the largest (resp. smallest) integer k such that $a_k \neq \mathbb{0}$.

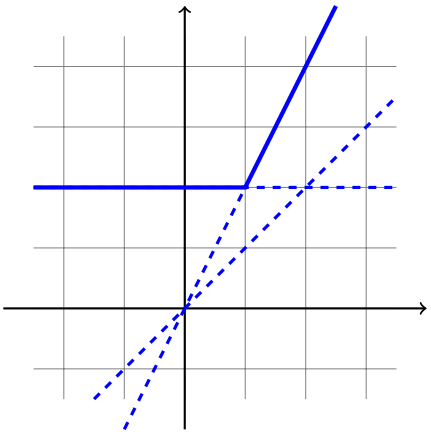
One crucial difference between tropical polynomials and polynomials over the real or complex field is that tropical polynomials are not in one-to-one correspondence with their associated polynomial functions: in other words, we can find two polynomials p and q with different coefficients such that $p(x) = q(x)$ for all $x \in \mathbb{R}_{\max}$ (or $\mathbb{T}$).

An example will help to get a more intuitive feel for tropical polynomials.

Example 2.1. Consider the max-plus polynomial $p = X^2 \oplus 2$. We draw the graph of its associated polynomial function in the $\mathbb{R}^2_{\max}$ plane. It is immediate to see that in general a monomial aX^k is represented in the max-plus plane by the line $a + kx$ (when $a = \mathbb{1} = 0$ we drop it and just write X^k). The graph of the whole polynomial function is then the max-plus sum (i.e. the pointwise maximum, i.e. the upper envelope) of all its monomials.



Now consider $q = X^2 \oplus X \oplus 2$. Clearly $q \neq p$, but the upper envelope does not change, so the associated functions coincide. It is also easy to see that every polynomial of the form $X^2 \oplus aX \oplus 2$ with $a \leq 1$ will yield the same function.



Definition 2.1. Given a max-plus polynomial p, the points of non-differentiability of the associated tropical polynomial function are called the *tropical roots* of p. Thinking of the function as an upper envelope, the tropical roots can be characterized equivalently as the points at which the maximum is attained at least twice (i.e. by at least two different monomials). The *multiplicity* of a tropical root α is the change of slope of the graph of the function at α, or equivalently the difference between the largest and the smallest exponent of the monomials of p which attain the maximum at α. If p has no constant term, then p is said to have a root at $-\infty$, whose multiplicity is the exponent of the first non-vanishing monomial term.

The choice of the term "roots" is justified by the following result, which we might consider a tropical analogue of the fundamental theorem of algebra.

Proposition 2.1 (Cuninghame-Green and Meijer, [CGM80])**.** *Every tropical polynomial $p = \bigoplus_{k=0}^{n} a_k X^k$ of degree n has exactly n tropical roots $\alpha_1 \geq \ldots \geq \alpha_n$, and the associated polynomial function can be uniquely factored as*

$$p(x) = a_n(x \oplus \alpha_1) \ldots (x \oplus \alpha_n) \ .$$

□

Convex analysis offers an alternative way of visualizing tropical polynomials and roots. Let $f\colon \mathbb{R} \to \overline{\mathbb{R}}$ be an extended real-valued function. The *lower semi-continuous convex envelope* (or *convexification*) and *upper semi-continuous concave envelope* (or *concavification*) of f are defined, respectively, as the largest convex l.s.c. minorant of f, and its smallest concave u.s.c. majorant. They can be characterized as

$$\begin{aligned} \operatorname{cvx} f &= \sup\{g \mid g \text{ affine},\ g \leqslant f\} \ , \\ \operatorname{cav} f &= \inf\{g \mid g \text{ affine},\ g \geqslant f\} \ . \end{aligned}$$

For any (formal) max-plus polynomial we can define an extended real-valued function on $\mathbb{R}$ that represents its coefficients: more precisely, to the max-plus polynomial $p = \bigoplus_{k\in\mathbb{N}} a_k X^k$ we associate the function

$$\operatorname{coef} p\colon \mathbb{R} \to \overline{\mathbb{R}}, \qquad (\operatorname{coef} p)(x) = \begin{cases} a_k & \text{if } x = k \in \mathbb{N} \\ -\infty & \text{otherwise.} \end{cases}$$

Then the max-plus polynomial function $p(x)$ corresponds to the Legendre-Fenchel transform of the map $-\operatorname{coef} p$ (notice the minus sign). This allows us to give dual characterizations of tropical polynomials and roots.

Definition 2.2. Let $p = \bigoplus_{k=0}^{n} a_k X^k$ be a max-plus polynomial. The *Newton polygon* $\Delta(p)$ of p is the graph of the function $\operatorname{cav}(\operatorname{coef} p)$ restricted to the interval where it takes finite values. In other terms, the Newton polygon of p is the upper boundary of the two-dimensional convex hull of the set of points $\{(k, a_k) \mid 0 \leqslant k \leqslant n,\ a_k \neq -\infty\}$.

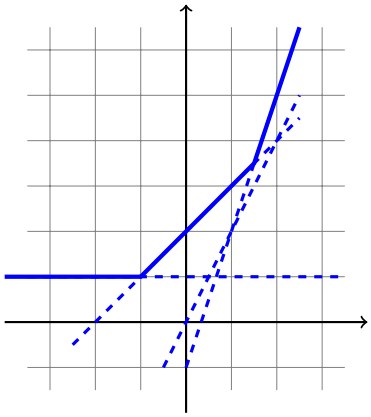
The values $\operatorname{cav}(\operatorname{coef} p)(0), \ldots, \operatorname{cav}(\operatorname{coef} p)(n)$ are called the *concavified coefficients* of p, and they are denoted by $\overline{a}_0, \ldots, \overline{a}_n$. An index k such that $a_k = \overline{a}_k$ (so that the point (k, a_k) lies on $\Delta(p)$) will be called a *saturated index*. With a slight abuse, we will extend the cav notation from functions to formal polynomials, defining the *concavified polynomial* of p as $\overline{p} = \operatorname{cav} p := \bigoplus_{k=0}^{n} \overline{a}_k X^k$.

We use similar notions and notation for max-times polynomials. Consider the isomorphism of semirings which sends a max-times polynomial $p = \bigoplus_{k=0}^{n} a_k X^k$ to the max-plus polynomial $\operatorname{Log} p := \bigoplus_{k=0}^{n} (\log a_k) X^k$. We define the *log-concavified polynomial* of p as $\widehat{p} = \operatorname{lcav} p := \operatorname{Log}^{-1} \circ \operatorname{cav} \circ \operatorname{Log} p$. Its coefficients are called the *log-concavified coefficients* of p, and they are denoted by $\widehat{a}_0, \dots, \widehat{a}_n$.

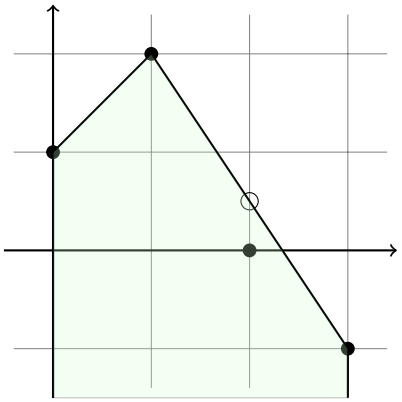
Proposition 2.2 (see e.g. [Roc70, Chap 12, p. 104])**.** *Two max-plus polynomials have the same associated polynomial function if and only if they have the same concavified coefficients, or equivalently the same Newton polygons.* □

Proposition 2.3 (see e.g. [ABG15, Proposition 2.4])**.** *Let $p \in \mathbb{R}_{\max}[X]$ be a max-plus polynomial. The roots of p coincide with the opposite of the slopes of the Newton polygon of p. The multiplicity of a root α of p coincides with the length of the interval where the Newton polygon has slope $-\alpha$.* □

The last proposition is illustrated in the following figure:



Graph of $p = -1X^3 \oplus X^2 \oplus 2X \oplus 1$. Note that the line with slope 2 does not appear in the convex hull.



Newton polygon of p. Black circles represent the coefficients. The coefficient of degree 2 is not saturated, and the white circle represents the concavified coefficient.

Figure 2.1: Graph and Newton polygon representations of a polynomial. The tropical roots are -1 (with multiplicity 1) and $3/2$ (with multiplicity 2), and one can check that the slopes of the polygonal line on the right are the opposite of the roots.

Corollary 2.4. *Let $p = \bigoplus_{k=0}^{n} a_k X^k$ be a max-plus polynomial, and let $\alpha_1 \geqslant \ldots \geqslant \alpha_n$ be its roots, counted with multiplicities. Then the following relation for the concavified coefficients of p holds:*

$$\overline{a}_{n-k} = a_n + \alpha_1 + \cdots + \alpha_k \qquad \forall k \in \{1, \dots, n\} .$$

Analogously, if $p = \bigoplus_{k=0}^{n} a_k X^k$ is a max-times polynomial with roots $\alpha_1 \geqslant \cdots \geqslant \alpha_n$, then the following relation for its log-concavified coefficients holds:

$$\widehat{a}_{n-k} = a_n \alpha_1 \cdots \alpha_k \qquad \forall k \in \{1, \dots, n\} .$$

As a consequence of the Newton polygon characterization, computing tropical roots has at most the complexity of computing the convex hull of n points in the plane: using the Graham scan algorithm [Gra72] this can be done in $O(n)$ time (assuming the points are already sorted by abscissa).

2.1.2 Tropical matrices and eigenvalues

For a matrix $A \in \mathbb{C}^{n\times n}$, eigenvalues can be defined either as the roots of the characteristic polynomial $\det(A - \lambda I)$ or as those λ that satisfy $Ax = \lambda x$ for some $x \neq 0$.

We can transpose both definitions in the tropical setting, but as we will see they will no longer be equivalent.

Definition 2.3. Let S be a semiring, with addition and multiplication denoted by $\oplus$ and $\otimes$, respectively. The *permanent* of a matrix $A \in S^{n\times n}$ is defined as $\operatorname{per}_S A = \bigoplus_{\sigma\in S_n} \bigotimes_{k=1}^{n} a_{k,\sigma(k)}$, where σ ranges over all the permutations of $\{1, \dots, n\}$. We will sometimes drop the subscript if the semiring is clear from the context.

Remark: when $S = \mathbb{R}_{\max}$, the permanent amounts to

$$\max_{\sigma\in S_n} \sum_{k=1}^{n} a_{k,\sigma(k)} \ .$$

A maximization problem of this kind is known in the literature as an *optimal assignment problem*. Many algorithms exist to solve it, the *Hungarian algorithm* being one of the most well-known; we will make extensive use of the Hungarian algorithm in Chapter 4.

Given a matrix $A \in \mathbb{R}_{\max}^{n\times n}$, the tropical characteristic polynomial of A is $\operatorname{per}_{\mathbb{R}_{\max}[\lambda]}(A \oplus \lambda\mathbb{I})$ (the doublestroke I is to emphasize that we are taking the max-plus identity matrix, with $\mathbb{1} = 0$ on the main diagonal and $\mathbb{0} = -\infty$ off of it). The *algebraic tropical eigenvalues* of A are the tropical roots of the tropical characteristic polynomial. The *geometric tropical eigenvalues* of A, on the other hand, are defined as the λ for which $A \otimes x = \lambda \otimes x$ holds for some nonzero vector x, called a *tropical eigenvector*. The exact same definitions apply for matrices in $\mathbb{T}^{n\times n}$.

Here is a simple example to show that the two types of eigenvalues do not coincide.

Example 2.2. Consider the max-plus matrix

$$A = \begin{pmatrix} 0 & -\infty \\ 0 & 1 \end{pmatrix} \ .$$

Its tropical characteristic polynomial is

$$\operatorname{per}_{\mathbb{R}_{\max}[\lambda]}(A \oplus \lambda\mathbb{I}) = \operatorname{per}_{\mathbb{R}_{\max}[\lambda]} \begin{pmatrix} \lambda \oplus 0 & -\infty \\ 0 & \lambda \oplus 1 \end{pmatrix} = (\lambda \oplus 0)(\lambda \oplus 1) \ ,$$

so its algebraic tropical eigenvalues are 0 and 1. On the other hand, the equation $A\otimes x = \lambda\otimes x$ leads to the system

$$\begin{cases} \max(x_1, -\infty) = \lambda + x_1 \\ \max(x_1, x_2 + 1) = \lambda + x_2 \end{cases}$$

whose only nontrivial solutions are of the form $(\lambda = 1,\ x_1 = -\infty,\ x_2)$ for any $x_2 \neq -\infty$, which means the only tropical geometric eigenvalue is 1.

It is no coincidence that in our example 1 is both a geometric and algebraic eigenvalue. Indeed, one can prove that the largest algebraic eigenvalue is guaranteed to be a geometric eigenvalue as well. We give here a more precise statement of this result in max-times terms. By analogy with the spectral radius, we denote by $\rho_{\mathbb{T}}(A)$ the largest algebraic eigenvalue of a max-times matrix A.

Definition 2.4. Consider a matrix $A \in \mathbb{T}^{n\times n}$, and let I be a subset of $\{1,\dots,n\}$ of cardinality ℓ. Denote by Ω_I the set of all cyclic permutations of I (i.e. permutations whose cycle decomposition consists of a single cycle of length ℓ). For any $\sigma \in \Omega_I$ we define the *weight* of σ with respect to A as $w_A(\sigma) = \prod_{i\in I} a_{i,\sigma(i)}$, and we define the *mean weight* of σ with respect to A as

$$\mu_A(\sigma) = w_A(\sigma)^{1/\ell} = \Big(\prod_{i\in I} a_{i,\sigma(i)}\Big)^{1/\ell} .$$

Finally, we define the *maximal cycle mean* of A as

$$\rho_{\max}(A) = \max_{I\subseteq\{1,\dots,n\}} \max_{\sigma\in\Omega_I} \mu_A(\sigma) .$$

If we interpret A as the adjacency matrix of a directed graph with weighted edges, then $\rho_{\max}(A)$ represents the maximum geometric-mean weight of a cycle over the graph.

Remark 2.3. Since any permutation can be factored into a product of cycles, we can equivalently define the maximal cycle mean in terms of general permutations instead of cyclical ones:

$$\rho_{\max}(A) = \max_{I\subseteq\{1,\dots,n\}} \max_{\sigma\in S_I} \mu_A(\sigma) .$$

Proposition 2.5 ([CG79], [CG83])**.** *Let $A \in \mathbb{T}^{n\times n}$ be an irreducible matrix. The following hold:*

- *there exists a unique geometric eigenvalue λ*
- *λ is equal to the maximal cycle mean $\rho_{\max}$*
- *$\rho_{\max}$ is equal to the largest algebraic tropical eigenvalue of A.*

□

We should remark that while in our previous toy example computations were easy, finding the coefficients of a tropical characteristic polynomial is in general a challenging problem, for which no polynomial time algorithm is known [BL07]. The tropical eigenvalues, however, only depend on the associated polynomial function, and can be computed by solving at most n optimal assignment problems, leading to the complexity bound of $O(n^4)$ of Burkard and Butkovič [BB03]. Gassner and Klinz [GK10] showed that this can be reduced to $O(n^3)$ using parametric optimal assignment techniques.

In any case, we can give explicit expressions for the coefficients of the tropical characteristic polynomial: if $A \in \mathbb{T}^{n\times n}$ and we write

$$\mathrm{per}_{\mathbb{T}[X]}(A \oplus \lambda\mathbb{I}) = X^n \oplus c_{n-1}X^{n-1} \oplus \cdots \oplus c_0 , \tag{2.1}$$

then it is not difficult to see that

$$c_{n-k} = \bigoplus_{\substack{I\subset[n]\\ \#I=k}} \bigoplus_{\sigma\in S_I} a_{1,\sigma(1)}\dots a_{n,\sigma(n)} = \bigoplus_{\substack{I\subset[n]\\ \#I=k}} \operatorname{per}_{\mathbb{T}} A[I,I] \quad \forall k\in\{1,\dots,n\}\,,$$

where $[n]=\{1,\dots,n\}$, S_I is the group of permutations of the set I, and $A[I,J]$ is the $k\times k$ submatrix obtained by selecting from A the rows $i\in I$ and the columns $j\in J$. It will be convenient to write the coefficients of the tropical characteristic polynomial in terms of the exterior powers of A.

Definition 2.5. The *k-th exterior power* of a matrix $A\in\mathbb{C}^{n\times n}$ is the matrix $\bigwedge^k A\in\mathbb{C}^{\binom{n}{k}\times\binom{n}{k}}$ whose rows and columns are indexed by the subsets of cardinality k of $\{1,\dots,n\}$, and whose entries are defined as

$$\left(\bigwedge^k A\right)_{I,J} = \det A[I,J]\,. \tag{2.2}$$

The k-th trace of A is then defined as

$$\operatorname{tr}^k A = \operatorname{tr}\left(\bigwedge^k A\right) = \sum_{\substack{I\subset[n]\\ \#I=k}} \det A[I,I]$$

for all $k\in\{1,\dots,n\}$. If we replace the determinant with the permanent in Equation (2.2), we get the *k-th permanental exterior power* of A, denoted by $\bigwedge^k_{\mathrm{per}} A$.

Analogously, for a matrix $A\in\mathbb{T}^{n\times n}$, we define the *$k$-th tropical exterior power* of A to be the matrix $\bigwedge^k_{\mathbb{T}} A\in\mathbb{T}^{\binom{n}{k}\times\binom{n}{k}}$ whose entries are

$$\left(\bigwedge^k_{\mathbb{T}} A\right)_{I,J} = \operatorname{per}_{\mathbb{T}} A[I,J]$$

for all subsets $I,J\subset\{1,\dots,n\}$ of cardinality k. The *k-th tropical trace* of A is defined as

$$\operatorname{tr}^k_{\mathbb{T}} A = \operatorname{tr}_{\mathbb{T}}\left(\bigwedge^k_{\mathbb{T}} A\right) = \max_{\substack{I\subset[n]\\ \#I=k}} \operatorname{per}_{\mathbb{T}} A[I,I]\,. \tag{2.3}$$

One readily checks that the coefficients of the tropical characteristic polynomial in (2.1) are given by $c_{n-k}=\operatorname{tr}^k_{\mathbb{T}} A$.

2.2 Eigenvalues of matrix polynomials

A *matrix polynomial* (also called a *matrix pencil*) over some ring R is a formal polynomial whose coefficients are matrices over R. We will generally denote the degree of the polynomial by d and consider square matrices of size n over the complex field, so a typical matrix polynomial is of the form

$$P = \sum_{k=0}^{d} A_k X^k, \qquad A_k\in\mathbb{C}^{n\times n}\,. \tag{2.4}$$

One can also see matrix polynomials as matrices whose entries are polynomials (there is an obvious isomorphism between $R^{n\times n}[X]$ and $R[X]^{n\times n}$). This interpretation grants us the right to speak of the *determinant* of a matrix polynomial, which will be a polynomial of degree at most nd. A matrix polynomial is *singular* if its determinant is identically zero, otherwise it is *regular*.

Definition 2.6. The *eigenvalues* of a regular complex matrix polynomial P are the roots of its determinant. If the degree of the determinant is $k < nd$ (which happens when the leading coefficient A_d is singular), the matrix polynomial is said to have an infinite eigenvalue of multiplicity $nd - k$. Vectors x and $y \in \mathbb{C}^{n\times n}$ such that $P(\lambda)x = y^*P(\lambda) = 0$ are the *right* and *left eigenvectors* associated to λ, respectively (here y^* denotes the conjugate transpose of y). We refer to (λ, x) and (λ, y) as right and left *eigenpairs*, and to (λ, x, y) as an *eigentriple*.

Remark 2.4. The multiplicity of the infinite eigenvalue is equal to the multiplicity of 0 as an eigenvalue of the reversed polynomial $X^dP(1/X)$.

The problem of finding eigenvalues and eigenvectors of a matrix polynomial goes under the name of *polynomial eigenvalue problem*.

2.2.1 Desirable properties of algorithms

Numerical computations involving real and complex numbers are usually carried out using floating point numbers. Since these are a discrete set of rational numbers, most of the time our data will not be representable exactly. Whenever this occurs, a rounding to the nearest representable number is automatically performed. Double-precision floating point numbers are laid out in such a way that the worst-case relative error between two consecutive representable numbers is $\sim 10^{-16}$. Elementary arithmetic operations on floating point numbers are also guaranteed to yield as a result the closest representable number to the true result. When the computations involve multiple operations, however, the small relative errors can combine and become relevant.

Consider an algorithm to compute some function $f\colon X \to Y$ at a point x, and assume X and Y are normed spaces (we denote both norms by $\|\cdot\|$ for convenience). Call y the exact value $f(x)$, and $\hat{y}$ the value returned by the algorithm. We would like the relative error $\frac{\|\hat{y}-y\|}{\|y\|}$ to be as small as possible.

The *backward error* is defined as the distance (in absolute or relative terms) from x to the set of pre-images of $\hat{y}$. In other words, the backward error is the answer to the question "how much does one have to perturb the original data in order for $\hat{y}$ to be an exact solution?", or "how close is the problem we actually solved to the problem we wanted to solve?"

An algorithm is *backward stable* if its backward error is small enough; the quantification of "small enough" is problem-dependent. For problems where inputs are not exact by nature (e.g. because they come from physical measurements with inherent uncertainty), the best one can hope for is to have a backward error of the same order as the uncertainty in the data: in this case the two errors may well cancel each other, and the solution would be exact.

Nonetheless, knowing that the problem we actually solved is close to the one we wanted to solve is no guarantee that the solutions of the two problems are close as well. There are problems where tiny differences in initial conditions lead to results which are far apart. The sensitivity of a problem to initial data is usually measured by means of the *condition number*. The condition number has different definitions for different problems, but in general one can think of it as a worst-case amplification factor between the error in the input and the resulting error in the output. It often takes the form of the operator norm of the first order approximation of the function to compute, which implies it becomes less reliable when the error in the input is not small to begin with.

Unlike the backward error, the condition number is an intrinsic characteristic of the problem at hand and it is independent of the algorithm used to solve it. Problems with small

condition numbers are called *well-conditioned*, while problems with large condition numbers are said to be *ill-conditioned*.

Knowing both the condition number and the backward error associated to a problem (and to the specific algorithm used), one can then take their product to obtain a first-order bound for the error in the solution.

2.2.2 Conditioning and backward error for the polynomial eigenvalue problem

Let P be a regular matrix polynomial of the form (2.4). Following definitions in [Tis00] we introduce condition number and backward error as follows.

Definition 2.7. Let λ be a nonzero simple eigenvalue of P with corresponding right eigenvector x and left eigenvector y. Let $\Delta P = \Delta A_0 + \cdots + \lambda^d \Delta A_d$ be a perturbation of P. The *normwise relative condition number* of λ is defined as

$$\kappa(\lambda, P) = \limsup_{\epsilon \to 0} \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} \;\middle|\; (P(\lambda+\Delta\lambda) + \Delta P(\lambda+\Delta\lambda))(x+\Delta x) = 0, \right.$$
$$\left. \|\Delta A_k\|_2 \le \epsilon \|A_k\|_2 \text{ for } k = 0, \ldots, d \right\}.$$

The backward error of an approximate eigentriple $(\tilde{\lambda}.\tilde{x}, \tilde{y})$ is defined as

$$\eta(\tilde{\lambda}, \tilde{x}, \tilde{y}) = \min\{\epsilon \mid P(\tilde{\lambda})\tilde{x} + \Delta P(\tilde{\lambda})\tilde{x} = 0, \tilde{y}^* P(\tilde{\lambda}) + \tilde{y}^* \Delta P(\tilde{\lambda}) = 0,$$
$$\|\Delta A_k\|_2 \le \epsilon \|A_k\|_2 \text{ for } k = 0, \ldots, d\}.$$

One can derive formulas for the condition number and backward error.

Proposition 2.6 ([Tis00]). *Let $\alpha = \sum_{k=0}^{d} |\lambda|^k \|A_k\|_2$, and let $\tilde{\alpha}$ be the analogous quantity for the approximate computed eigenvalue $\tilde{\lambda}$. Then,*

$$\kappa(\lambda, P) = \frac{\alpha \|x\|_2 \|y\|_2}{|\lambda| |y^* P'(\lambda) x|}$$

and

$$\eta(\tilde{\lambda}, \tilde{x}, \tilde{y}) = \frac{1}{\tilde{\alpha}} \max\left\{ \frac{\|P(\tilde{\lambda})\tilde{x}\|_2}{\|\tilde{x}\|_2}, \frac{\|\tilde{y}^* P(\tilde{\lambda})\|_2}{\|\tilde{y}\|_2} \right\}.$$

□

2.2.3 Solving polynomial eigenvalue problem via linearization

The eigenvalue problem for linear matrix polynomials ($d = 1$) has been studied extensively; it is referred to as the *generalized eigenvalue problem* and it is usually presented in the form $Ax = \lambda Bx$.

The generalized eigenvalue problem can be solved via the *QZ algorithm* by Moler and Stewart [MS73], which is backward stable. The QZ algorithm is not the most suitable one for very large sparse matrices, or when only a small number of eigenvalues are required. In this case one can resort to some form of power iteration.

The most common approach to compute the eigenvalues of a matrix polynomial P of degree $d > 1$ is to reduce to the linear case, that is to find $nd \times nd$ matrices A and B such

that the matrix pencil $A + \lambda B$ has the same spectrum as P, and then solve the generalized eigenvalue problem using QZ or some other algorithm. This process goes under the name of *linearization*. Note that while the eigenvalues of the linearized pencil are the same as those of P, the eigenvector are certainly not (the dimensions do not match). The convenience of the formula to recover the eigenvectors of P is one of the factors to weigh when choosing a linearization.

Among the most widely used linearizations are the so called *first* and *second companion linearization*, which have block structure respectively

$$C_1 = \lambda \begin{pmatrix} A_n & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{pmatrix} + \begin{pmatrix} A_{n-1} & A_{n-2} & \dots & A_0 \\ -I & 0 & & \\ & \ddots & \ddots & \\ & & -I & 0 \end{pmatrix}$$

and

$$C_2 = \lambda \begin{pmatrix} A_n & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{pmatrix} + \begin{pmatrix} A_{n-1} & -I & & \\ A_{n-2} & 0 & \ddots & \\ \vdots & & \ddots & -I \\ A_0 & & & 0 \end{pmatrix}.$$

Linearizations C_1 and C_2 can be seen as special cases of two larger classes of linearizations: Mackey et al. [MMMM06] showed that for each matrix polynomial P as in (2.4) there exist two vector subspaces of the space of linear pencils of size nd, denoted by $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ respectively, which are almost entirely composed of linearizations for P (the exceptions form a closed nowhere dense set of measure zero). One interesting characteristic is that $z \in \mathbb{C}^{nd}$ is a right eigenvector corresponding to an eigenvalue λ of a linearization in $\mathbb{L}_1(P)$ if and only if it is of the form $z = [\lambda^{d-1}x, \lambda^{d-2}x, \dots, \lambda x, x]^\top$ where $x \in \mathbb{C}^n$ is an eigenvector corresponding to the same eigenvalue λ for the original polynomial P. The corresponding property holds for left eigenvectors of linearizations in $\mathbb{L}_2(P)$.

Linearizations belonging to $\mathbb{DL}(P) := \mathbb{L}_1(P) \cap \mathbb{L}_2(P)$ are particularly convenient to study. $\mathbb{DL}(P)$ is a d-dimensional vector space, parametrized by a vector $v \in \mathbb{C}^d$. Higham, Mackey and Tisseur [HMT06] solved the problem of finding the parameter v for which the condition number of an eigenvalue of the linearization is minimized. They also compared the condition number $\kappa(\lambda, L)$ for any linearization $L \in \mathbb{DL}(P)$ to $\kappa(\lambda, C_1)$ and to $\kappa(\lambda, P)$, obtaining upper and lower bounds for the ratios of each pair of condition numbers considered. Similar study and comparisons have been done by Higham, Li and Tisseur [HLT07] for the backward error of a polynomial matrix P and its companion and $\mathbb{DL}$ linearizations.

As a final note it should be remarked that some methods that do not resort to linearization have been developed as well, such as iterative methods [BN13], reduction to lower-degree nonlinear matrix polynomials (quadratization and more generally ℓ-ification, see e.g. [DTDM14]), or one-sided factorization (which can also be used to solve nonhomogeneous parametric linear systems, see [GHT11]).

2.2.4 Scaling

When solving a polynomial eigenvalue problem via linearization we are essentially dealing with a different problem which happens to have the same solution, at least in theory; anyway,

due to the finite precision of floating point numbers, the computed solutions can be very different. This is where scaling comes into play; there is no unique definition of scaling, but rather many different approaches which have in common the goal of transforming one matrix or matrix polynomial into another with the same eigenvalues but with better numerical properties. Here we present a few of them.

Eigenvalue scaling Consider the eigenproblem for a matrix polynomial $P(\lambda) = A_0 + \lambda A_1 + \cdots + \lambda^d A_d$, and build the scaled polynomial $Q(\mu) = \beta P(\alpha\mu) = \tilde{A}_0 + \mu\tilde{A}_1 + \cdots + \mu^d \tilde{A}_d$, where $\tilde{A}_k = \beta\alpha^k A_k$. It is immediate to see that if (λ, x, y) is an eigentriple for P, then $(\mu = \lambda/\alpha, x, y)$ is an eigentriple for Q. We can thus find eigenvalues for P by solving the eigenproblem for Q and multiplying the obtained eigenvalues by α. Using the formulas of Proposition 2.6 one can check that the condition number and backward error do not change when we recover the eigenvalues of P from those of Q; however, by choosing wisely the scaling parameters α and β we can hope to find a polynomial Q for which the linearization algorithm is more stable.

Fan, Lin and Van Dooren [FLVD04] proposed such a scaling for the particular case of a quadratic polynomial, with the aim of bringing the norms of the scaled matrices close to one. More precisely, α and β were chosen as the minimizers of the min-max problem

$$\min_{\alpha,\beta} \max_{k=0,1,2} \left\{ \left| \|\tilde{A}_k\| - 1 \right| \right\} ,$$

which results in

$$\alpha = \sqrt{\frac{\|A_0\|}{\|A_2\|}} , \qquad \beta = \frac{2}{\|A_0\| + \alpha\|A_1\|} .$$

This leads to an improvement of the backward error by several orders of magnitude.

Gaubert and Sharify [GS09] interpreted the previous result in terms of tropical algebra. They noted that if the max-times polynomial $R(x) = \|A_0\| \oplus \|A_1\| x \oplus \|A_2\| x^2$ has one double tropical root, then this root must be the α found by Fan, Lin and Van Dooren. On the other hand, for cases where $R(x)$ has two distinct tropical roots $\alpha_1 \leq \alpha_2$ they proposed a differentiated scaling whereby they performed eigenvalue scaling twice, first with parameter α_1, then α_2. The idea is that the first scaling improves the accuracy of the n smallest eigenvalues, while the second improves the n largest ones. This can be easily generalized to matrix polynomials of higher degree: it is sufficient to perform one eigenvalue scaling for each tropical root of the polynomial $\bigoplus_{i=0}^{d} \|A_i\| x^i$, then by keeping the appropriate group of n eigenvalues from each scaling one can improve accuracy over the whole spectrum, at the cost of repeating the scaling step multiple times.

Balancing For the classical eigenvalue problem, the notion of *balancing* a matrix has been introduced by Osborne [Osb60] and refined by Parlett and Reinsch [PR69]. A matrix is said to be balanced (with respect to a given norm) if its i-th row and i-th column have the same norm for all i. A balanced matrix has minimal Frobenius norm among all matrices in its diagonal similarity orbit. This property, together with the observation that errors in the computation of eigenvalues are in general roughly proportional to the Frobenius norm, lead to the development of Osborne's iterative algorithm to balance a matrix before proceeding to compute its eigenvalues.

Diagonal scaling Ideas similar to balancing have been applied to the generalized eigenvalue problem. Given the problem $Ax = \lambda Bx$, Ward [War81] proposed a method to find diagonal nonsingular matrices D_1 and D_2 such that the elements of D_1AD_2 and D_1BD_2 would be close to 1 in magnitude. Later, Lemonnier and Van Dooren [LD06] gave a different interpretation of the original balancing algorithm, showing that the matrix which minimizes the Frobenius norm is also, in some precise sense, the closest to a normal matrix among those in the diagonal similarity orbit. Starting from this observation, they proposed to apply the same principle to linear pencils, after giving a notion of normal pencil and a quantification of the deviation from normality. The resulting scaling is a two-sided diagonal scaling that leads to a pencil $\lambda\tilde{B} - \tilde{A}$ with the property $\|Ae_i\|_2^2 + \|Be_i\|_2^2 = \|e_i^*A\|_2^2 + \|e_i^*B\|_2^2 = \text{constant}$ (here e_i denotes the i-th canonical basis vector).

A further generalization of the balancing scaling to matrix polynomials of degree larger than one was proposed by Betcke [Bet08]. Given P as in (2.4), with right and left eigenvectors x and y corresponding to an eigenvalue λ, he showed that if $|x_i| = |y_i| = 1\ \forall i \in \{1, \dots, n\}$, then $\kappa(\lambda, P)$ is almost optimal (more precisely, $\kappa(\lambda, P) \leq n^{3/2}\kappa(\lambda, D_1PD_2)$ for any choice of nonsingular diagonal matrices D_1 and D_2). He thus proposed a scaling aimed at getting eigenvectors with components as homogeneous as possible, imposing the balancing condition $\|D_1P(\lambda)D_2e_i\| \approx \|e_j^*D_1P(\lambda)D_2\| \approx 1\ \forall i, j \in \{1, \dots, n\}$. Since the value of λ is not known a priori, he introduced a weighting parameter ω meant as an initial guess of the magnitude of the wanted eigenvalue. We will show in the next chapters that tropical eigenvalues can play this role.

Betcke also proposed a generalization of Fan, Lin and Van Dooren scaling (which was defined for the quadratic case) to all matrix polynomials of degree $d \geq 2$. The target to minimize in this case is the ratio

$$\frac{\max_i \alpha^i \|A_i\|}{\min(\|A_0\|, \alpha^d \|A_d\|)},$$

which gives as a result the scaling parameter

$$\alpha = \left(\frac{\|A_0\|}{\|A_d\|}\right)^{\frac{1}{d}}.$$

It is not always immediate to see which scaling could bring more benefit to the computation. For an algorithm that implements (selectively) all the scalings discussed above, in the quadratic case, see [HMT13].

2.3 Tropical algebra and valuation theory

A *valuation* on a field $\mathbb{F}$ is a map $\nu\colon \mathbb{F} \to \mathbb{R} \cup \{-\infty\}$ satisfying

- $\nu(a) = -\infty \iff a = 0$
- $\nu(ab) = \nu(a) + \nu(b)$
- $\nu(a + b) \leq \max(\nu(a), \nu(b))$

for all $a, b \in \mathbb{F}$. In the last property, equality holds when $\nu(a) \neq \nu(b)$.

Valuations are closely related to non-archimedean absolute values, i.e. absolute values for which the triangular inequality is replaced by the stronger $|a+b| \leq \max(|a|,|b|)$: for these absolute values, $\nu(a) = \log|a|$ is a valuation, provided that we extend the logarithm so that $\log 0 = -\infty$.

For this reason we will also refer to valuations as *non-archimedean valuations*, and by analogy we will call *archimedean valuation* any map $\nu\colon a \mapsto \log|a|$ where $|\cdot|$ is an archimedean absolute value (as are for instance the usual absolute values on $\mathbb{R}$ and $\mathbb{C}$).

The max-plus semiring arises naturally when studying non-archimedean valuations, since the valuation axioms can be rewritten as

- $\nu(ab) = \nu(a) \otimes \nu(b)$
- $\nu(a+b) \leq \nu(a) \oplus \nu(b)$

which makes ν almost a homomorphism from $\mathbb{F}$ to $\mathbb{R}_{\max}$.

When dealing with vectors or matrices we will write, with a slight abuse of notation, $\nu(v)$ or $\nu(A)$, to mean entrywise valuation.

We need to introduce the notion of *majorization*, which will be used in the next paragraphs.

Definition 2.8. Let $u, v \in \mathbb{R}^n_{\max}$. Let $u_{[1]} \geqslant \cdots \geqslant u_{[n]}$ (resp. $v_{[1]} \geqslant \cdots \geqslant v_{[n]}$) denote the components of u (resp. v) in decreasing order. We say that u is *weakly sub-majorized* by v, and we write $u \prec_{\mathrm{w}} v$, if the following conditions hold:

$$u_{[1]} \cdots u_{[k]} \leqslant v_{[1]} \cdots v_{[k]} \quad \forall k = 1, \ldots, n \ .$$

2.3.1 Valuation and roots

Given a polynomial $p \in \mathbb{F}[x_1, \ldots, x_n]$, $p(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}^d} a_{\mathbf{k}} \mathbf{x}^{\mathbf{k}}$ and a valuation $\nu\colon \mathbb{F} \to \mathbb{R} \cup \{-\infty\}$, the *tropicalization* of p is defined as $p^{\tau} = \max_{\mathbf{k} \in \mathbb{N}^d} \{\nu(a_{\mathbf{k}}) + u_1 k_1 + \cdots + u_d k_d\}$, where $u_1, \ldots, u_d$ are the coordinates in $\mathbb{R}^d_{\max}$.

Einsiedler, Kapranov and Lind [EKL06] showed that the image of the algebraic hypersurface of p under a non-archimedean valuation map (applied coordinate-wise) coincides with the locus of non-differentiability points of its tropicalization. Since the non-differentiability points are the tropical roots, this morally means that the operations of applying a non-archimedean valuation and taking the roots commute.

The same result does not hold for archimedean valuations. In its place, majorization inequalities arise between the image of the roots of a polynomial and the tropical roots of its tropicalization. One result of this kind dates back to the work of Ostrowski [Ost40]: rephrased in a more modern terminology, he showed that if $\zeta_1, \ldots, \zeta_n$ are the roots of a univariate complex polynomial, ordered by decreasing absolute value, and $\log \alpha_1 \geq \cdots \geq \log \alpha_n$ are the points of non-differentiability of the (archimedean) tropicalization obtained via the valuation associated to the usual absolute value, i.e. $\nu(x) = \log|x|$, then one has

$$\frac{1}{\binom{n}{k}} \alpha_1 \cdots \alpha_k \leq |\zeta_1 \cdots \zeta_k| \leq \sqrt{e(k+1)}\, \alpha_1 \cdots \alpha_k \qquad \forall k \in \{1, \ldots, n\} \ . \tag{2.5}$$

2.3.2 Valuations and eigenvalues

The study of relations between algebraic objects and their valuation can be extended to the case of matrices. One question which arises naturally is whether one can find an equivalent of the Einsiedler-Kapranov-Lind theorem, i.e. whether the tropical eigenvalues of the valuated counterpart of a matrix correspond to the valuations of its classical eigenvalues.

While this is not true in general, Akian, Bapat and Gaubert [ABG04, ABG15] showed that if A is a square matrix over an algebraically closed field equipped with a non-archimedean valuation ν, then the sequence of valuations of the eigenvalues of A is weakly sub-majorized by the sequence of the (algebraic) tropical eigenvalues of $\nu(A)$. Moreover, they proved that this result remains valid for *quasi-valuations* (i.e. valuations ν where the requirement $\nu(ab) = \nu(a) + \nu(b)$ is weakened to $\nu(ab) \leq \nu(a) + \nu(b)$, as is the case for first order asymptotics), and they extended it to matrix polynomials. To do so, one first needs to define eigenvalues of a tropical matrix polynomial. Just as for the matrix case, this is done by replacing the determinant with the permanent in the definition of the characteristic polynomial. The tropical eigenvalues of a tropical matrix polynomial P are then the tropical roots of $\mathrm{per}_{\mathbb{R}_{\max}} P$.

CHAPTER 3

Majorization bounds for eigenvalues

In this chapter we provide majorization-type upper and lower bounds for the absolute value of the eigenvalues of a matrix A. These bounds are given in terms of the tropical eigenvalues of a max-times matrix associated to A. The content of this chapter has appeared in [AGM14].

3.1 Upper bound

In order to link classical and tropical eigenvalues, we must first define what are tropical eigenvalues for a complex matrix.

Definition 3.1. Given a complex polynomial $p = \sum_{k=0}^{n} a_k z^k$, we define its *max-times relative* $p^\times \in \mathbb{T}[X]$ as

$$p^\times = \bigoplus_{k=0}^{n} |a_k| X^k \, .$$

The tropical roots of p are defined as the the tropical roots of its max-times relative $p^\times$.

Definition 3.2. Given a complex matrix $A = (a_{i,j}) \in \mathbb{C}^{n\times n}$, the tropical eigenvalues of A are defined as the tropical eigenvalues of the associated max-times matrix $|A| = (|a_{i,j}|)$.

Theorem 3.1. *Let $A \in \mathbb{C}^{n\times n}$ be a complex matrix, and let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues, ordered by nonincreasing absolute value (i.e., $|\lambda_1| \geqslant \ldots \geqslant |\lambda_n|$). Moreover, let $\gamma_1 \geqslant \ldots \geqslant \gamma_n$ be the tropical eigenvalues of A. Then for all $k \in \{1, \ldots, n\}$, we have*

$$|\lambda_1 \cdots \lambda_k| \leqslant U_k \gamma_1 \cdots \gamma_k$$

where

$$U_k = \rho(\textstyle\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)) \ .$$

In the last formula, pat A denotes the *pattern matrix* of A, a $(0,1)$-matrix depending only on the position of nonzero entries of A; more precisely, it is defined as

$$(\mathrm{pat}\, A)_{i,j} = \begin{cases} 0 & \text{if } a_{i,j} = 0 \\ 1 & \text{otherwise} \end{cases}$$

In the following sections we present some auxiliary results needed to prove this theorem.

3.1.1 Friedland's Theorem

Let $A = (a_{i,j})$ and $B = (b_{i,j})$ be nonnegative matrices. We denote by $A \circ B$ the Hadamard (entrywise) product of A and B, and by $A^{[r]}$ the entrywise r-th power of A. That is:

$$(A \circ B)_{i,j} = a_{i,j} b_{i,j}, \qquad (A^{[r]})_{i,j} = a^r_{i,j} \ .$$

Theorem 3.2 (Friedland, [Fri86])**.** *Let A be a nonnegative matrix. Define the* limit eigenvalue *of A as*

$$\rho_\infty(A) = \lim_{r \to +\infty} \rho(A^{[r]})^{1/r}.$$

Then we have

$$\rho_\infty(A) = \rho_{\max}(A), \tag{3.1}$$

and also

$$\rho(A) \leqslant \rho(\mathrm{pat}\, A)\rho_{\max}(A) \ .$$

Friedland's result is related to the following log-convexity property of the spectral radius.

Theorem 3.3 (Kingman [Kin61], Elsner, Johnson and Da Silva, [EJD88])**.** *If A and B are nonnegative matrices, and α, β are two positive real numbers such that $\alpha + \beta = 1$, then $\rho(A^{[\alpha]} \circ B^{[\beta]}) \leqslant \rho(A)^\alpha \rho(B)^\beta$.*

Corollary 3.4. *If A and B are nonnegative matrices, then $\rho(A \circ B) \leqslant \rho(A)\rho_{\max}(B)$.*

Proof. Let p, q be two positive real numbers such that $\frac{1}{p} + \frac{1}{q} = 1$. By applying Theorem 3.3 to the nonnegative matrices $A^{[p]}$ and $B^{[q]}$, and $\alpha = \frac{1}{p}$, we get $\rho(A \circ B) \leqslant \rho(A^{[p]})^{\frac{1}{p}} \rho(B^{[q]})^{\frac{1}{q}}$. Then by taking the limit for $q \to \infty$ and using the identities of Theorem 3.2 we obtain $\rho(A \circ B) \leqslant \rho(A)\rho_{\max}(B)$. □

3.1.2 Spectral radius of exterior powers

The next two propositions are well known.

Proposition 3.5 (See e.g. [HJ90, Theorem 8.1.18])**.** *The following statements about the spectral radius hold:*

(a) For any complex matrix A we have $\rho(A) \leqslant \rho(|A|)$;

(b) If A and B are nonnegative matrices and $A \leqslant B$, then $\rho(A) \leqslant \rho(B)$.

Proposition 3.6 (See e.g. [MM92, 2.15.12]). *If $A \in \mathbb{C}^{n\times n}$ has eigenvalues $\lambda_1, \dots, \lambda_n$, then the eigenvalues of $\bigwedge^k A$ are the products $\prod_{i\in I} \lambda_i$ for all subsets $I \subset \{1, \dots, n\}$ of cardinality k.*

An immediate corollary of Proposition 3.6 is that if $|\lambda_1| \geqslant \dots \geqslant |\lambda_n|$, then the spectral radius of $\bigwedge^k A$ is

$$\rho(\textstyle\bigwedge^k A) = |\lambda_1 \cdots \lambda_k| \,.$$

In the tropical setting we can prove the following combinatorial result, which will be one of the key ingredients of the proof of Theorem 3.1.

Theorem 3.7. *Let $A \in \mathbb{C}^{n\times n}$ be a complex matrix, and let $\gamma_1 \geqslant \cdots \geqslant \gamma_n$ be its tropical eigenvalues. Then for any $k \in \{1, \dots, n\}$ we have*

$$\rho_{\mathbb{T}}(\textstyle\bigwedge^k_{\mathbb{T}} |A|) \leqslant \gamma_1 \cdots \gamma_k.$$

The proof of this theorem relies on the following result, which is a variation on classical theorems of Hall and Birkhoff on doubly stochastic matrices. Recall that a *circulation matrix* of size $n \times n$ is a nonnegative matrix $B = (b_{i,j})$ such that for all $i \in \{1, \dots, n\}$, $\sum_{j\in[n]} b_{i,j} = \sum_{j\in[n]} b_{j,i}$. The *weight* of this matrix is the maximum value of the latter sums as $i \in \{1, \dots, n\}$. We call *partial permutation matrix* a matrix having a permutation matrix as a principal submatrix, all the other entries being zero. The *support* of a partial permutation matrix consists of the row (or column) indices of this principal submatrix.

Lemma 3.8. *Every circulation matrix $B = (b_{i,j})$ with integer entries, of weight ℓ, can be written as the sum of at most ℓ partial permutation matrices.*

Proof. We set $s_i = \sum_{j\in[n]} b_{i,j} = \sum_{j\in[n]} b_{j,i}$, so that $s_i \leqslant \ell \quad \forall i \in \{1, \dots, n\}$. If we add to B the diagonal matrix $D = \mathrm{Diag}(\ell - s_1, \dots, \ell - s_n)$, we obtain a matrix with nonnegative integer entries in which the sum of each row and each column is ℓ. A well known theorem (see e.g. Hall, [Hal98, Theorem 5.1.9]), allows us to write

$$B + D = P^{(1)} + \cdots + P^{(\ell)}$$

where the $P^{(i)}$'s are permutation matrices. Furthermore we can write D as a sum of diagonal matrices $D^{(1)}, \dots, D^{(\ell)}$ such that $D^{(i)} \leqslant P^{(i)}$ $\forall i \in \{1, \dots, \ell\}$. In this way we have

$$B = (P^{(1)} - D^{(1)}) + \cdots + (P^{(\ell)} - D^{(\ell)}) = B^{(1)} + \cdots + B^{(\ell)}$$

where every $B^{(m)} = (b^{(m)}_{i,j})$ is a partial permutation matrix (possibly zero). □

Proof of Theorem 3.7. Let $A \in \mathbb{C}^{n\times n}$ be a complex matrix. By definition, the tropical eigenvalues $\gamma_1 \geqslant \cdots \geqslant \gamma_n$ of A are the roots of the tropical characteristic polynomial $q_{|A|} = \mathrm{per}_{\mathbb{T}}(|A| \oplus XI)$. Recall that $\mathrm{tr}^k_{\mathbb{T}} |A|$ is the $(n-k)$-th coefficient of $q_{|A|}$, with the convention $\mathrm{tr}^0_{\mathbb{T}} |A| = 1$. We shall denote by $\widehat{\mathrm{tr}}^k_{\mathbb{T}} |A|$ the $(n-k)$-th log-concavified coefficient of $q_{|A|}$.

In the following formulas we will denote by $S_{I,J}$ the set of bijections from I to J. By Proposition 2.5, we have

$$\begin{aligned}\rho_{\mathbb{T}}(\textstyle\bigwedge^k_{\mathbb{T}}|A|) &= \max_{\ell\in[n]} \max_{\substack{\#I_1=k\\ \vdots \\ \#I_\ell=k}} \left(\textstyle\bigwedge^k_{\mathbb{T}}|A|_{I_1I_2}\cdots\bigwedge^k_{\mathbb{T}}|A|_{I_lI_1}\right)^{1/\ell}\\ &= \max_{\ell\in[n]} \max_{\substack{\#I_1=k\\ \vdots \\ \#I_\ell=k}} \max_{\substack{\sigma_1\in S_{I_1,I_2}\\ \vdots \\ \sigma_\ell\in S_{I_\ell,I_1}}} \Big(\prod_{i_1\in I_1}|a_{i_1\sigma_1(i_1)}|\cdots\prod_{i_\ell\in I_\ell}|a_{i_\ell\sigma_\ell(i_\ell)}|\Big)^{1/\ell}. \end{aligned} \tag{3.2}$$

The product in parentheses is a monomial in the entries of $|A|$ of degree $k\cdot\ell$. We rewrite it as

$$\prod_{\substack{i\in[n]\\ j\in[n]}}|a_{i,j}|^{b_{i,j}},$$

where $b_{i,j}$ is the total number of times the element $|a_{i,j}|$ appears in the product. We can arrange the $b_{i,j}$ into a matrix $B=(b_{i,j})$, and observe that $\sum_{j\in[n]}b_{i,j}=\sum_{j\in[n]}b_{j,i}\ \forall i\in\{1,\dots,n\}$, so that B is a circulation matrix. In fact, for every $m\in\{1,\dots,\ell\}$, every index $i\in I_m$ contributes for 1 to the i-th row of B (because of the presence of $|a_{i,\sigma_m(i)}|$ in the product), and also for 1 to the i-th column of B (because of the presence of $|a_{\sigma^{-1}_{m-1}(i),i}|$ in the product). By Lemma 3.8, we can write $B=B^{(1)}+\cdots+B^{(r)}$ with $r\leqslant\ell$, where $B^{(1)},\dots,B^{(r)}$ are partial permutation matrices, with respective supports $I^{(1)},\dots,I^{(r)}$. We set $B^{(r+1)}=\cdots=B^{(\ell)}=0$ and $I^{(r+1)}=\cdots=I^{(\ell)}=\varnothing$.

The product in the definition of $\rho_{\mathbb{T}}(\bigwedge^k_{\mathbb{T}}|A|)$ (inside the parentheses in (3.2)) can thus be rewritten as

$$\begin{aligned}\prod_{\substack{i\in[n]\\ j\in[n]}}|a_{i,j}|^{b_{i,j}} &= \prod_{m=1}^{\ell}\Big(\prod_{\substack{i\in[n]\\ j\in[n]}}|a_{i,j}|^{b^{(m)}_{i,j}}\Big)\\ &\leqslant \prod_{m=1}^{\ell}\operatorname{tr}^{\#I^{(m)}}_{\mathbb{T}}|A|\\ &\leqslant \prod_{m=1}^{\ell}\widehat{\operatorname{tr}}^{\#I^{(m)}}_{\mathbb{T}}|A|\\ &\leqslant (\widehat{\operatorname{tr}}^k_{\mathbb{T}}|A|)^{\ell},\end{aligned}$$

where the last inequality follows from the log-concavity of $k\mapsto\widehat{\operatorname{tr}}^k_{\mathbb{T}}|A|$ and from the fact that $\frac{1}{\ell}\sum_{m=1}^{\ell}\#I^{(m)}=k$. So, using (3.2), we conclude that $\rho_{\mathbb{T}}(\bigwedge^k_{\mathbb{T}}|A|)\leqslant\widehat{\operatorname{tr}}^k_{\mathbb{T}}|A|$.

Now, $\widehat{\operatorname{tr}}^k_{\mathbb{T}}|A|$ is the $(n-k)$-th concavified coefficient of the tropical polynomial $q_{|A|}$, whose roots are $\gamma_1\geqslant\ldots\geqslant\gamma_n$. Applying Corollary 2.4, and recalling that $\operatorname{tr}^0_{\mathbb{T}}|A|=1$, we obtain

$$\widehat{\operatorname{tr}}^k_{\mathbb{T}}|A|=\gamma_1\cdots\gamma_k,$$

so we conclude that

$$\rho_{\mathbb{T}}(\textstyle\bigwedge^k_{\mathbb{T}}|A|)\leqslant\gamma_1\cdots\gamma_k.$$

□

3.1.3 Proof of Theorem 3.1

For all subsets I, J of $\{1, \dots, n\}$, we have

$$\begin{aligned} |\textstyle\bigwedge^k A|_{I,J} &= |\det A[I,J]| \leqslant \operatorname{per} |A[I,J]| \\ &\leqslant \#\{\sigma \in S_{I,J} | w_A(\sigma) \neq 0\} \cdot \max_{\sigma \in S_{I,J}} |w_A(\sigma)| \\ &= \Big(\textstyle\bigwedge^k_{\operatorname{per}}(\operatorname{pat} A)\Big)_{I,J} \Big(\textstyle\bigwedge^k_{\mathbb{T}} |A|\Big)_{I,J}. \end{aligned}$$

Since this holds for all I and J, we can write, in terms of matrices,

$$|\textstyle\bigwedge^k A| \leqslant \Big(\bigwedge^k_{\operatorname{per}}(\operatorname{pat} A)\Big) \circ \Big(\bigwedge^k_{\mathbb{T}} |A|\Big) \ . \tag{3.3}$$

We have

$$\begin{aligned} |\lambda_1 \dots \lambda_k| &= \rho(\textstyle\bigwedge^k A) \qquad \text{(by Proposition 3.6)} \\ &\leqslant \rho\big((\textstyle\bigwedge^k_{\operatorname{per}}(\operatorname{pat} A)) \circ (\bigwedge^k_{\mathbb{T}} |A|)\big) \qquad \text{(by (3.3) and Proposition 3.5)}, \\ &\leqslant \rho(\textstyle\bigwedge^k_{\operatorname{per}}(\operatorname{pat} A)) \rho_{\mathbb{T}}(\bigwedge^k_{\mathbb{T}} |A|) \qquad \text{(by Corollary 3.4 and Proposition 2.5)} \\ &\leqslant \rho(\textstyle\bigwedge^k_{\operatorname{per}}(\operatorname{pat} A)) \gamma_1 \cdots \gamma_k \qquad \text{(by Theorem 3.7)} \end{aligned}$$

and the proof of the theorem is complete. □

3.2 Lower bound

We next show that the product of the k largest absolute values of eigenvalues can be bounded from below in terms of the k largest tropical eigenvalues, under some quite restrictive non-degeneracy conditions.

Lemma 3.9. *Let $A = (a_{i,j}) \in \mathbb{C}^{n \times n}$ be a complex matrix, and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues, ordered by nonincreasing absolute value (i.e. $|\lambda_1| \geqslant \dots \geqslant |\lambda_n|$). Moreover, let $\gamma_1 \geqslant \dots \geqslant \gamma_n$ be the tropical eigenvalues of A. Let $k \in \{1, \dots, n\}$ be a saturated index for the tropical characteristic polynomial $q_{|A|}$. Suppose $\operatorname{tr}^k A \neq 0$, and let C_k be any positive constant such that*

$$C_k \operatorname{tr}^k_{\mathbb{T}} |A| \leqslant |\operatorname{tr}^k A| \ . \tag{3.4}$$

Then the following bound holds:

$$\frac{C_k}{\binom{n}{k}} \gamma_1 \cdots \gamma_k \leqslant |\lambda_1 \cdots \lambda_k| \ .$$

Proof. Thanks to Ostrowski's lower bound in (2.5), we already have

$$\alpha_1 \cdots \alpha_k \leqslant \binom{n}{k} |\lambda_1 \cdots \lambda_k|,$$

where $\alpha_1 \geqslant \dots \geqslant \alpha_n$ are the tropical roots of the ordinary characteristic polynomial

$$p_A(x) = \det(xI - A) = x^n - (\operatorname{tr} A) x^{n-1} + (\operatorname{tr}^2 A) x^{n-2} + \cdots + (-1)^n \operatorname{tr}^n A \ .$$

Moreover, by Corollary 2.4 we have

$$\begin{aligned} \alpha_1 \cdots \alpha_k &= \operatorname{coef}_k(\operatorname{lcav} p_A^\times) \geq \operatorname{coef}_k(p_A^\times) = |\operatorname{tr}^k A| \\ \gamma_1 \cdots \gamma_k &= \operatorname{coef}_k(\operatorname{lcav} q_{|A|}) \geq \operatorname{coef}_k(q_{|A|}) = \operatorname{tr}_{\mathbb{T}}^k |A| \,, \end{aligned} \tag{3.5}$$

where we denoted by $\operatorname{coef}_k(p)$ the coefficient of degree k of the polynomial p. Since k is a saturated index for $q_{|A|}$, the second line in (3.5) is actually a chain of equalities. Now we can use Equation (3.4) and write

$$\gamma_1 \cdots \gamma_k = \operatorname{tr}_{\mathbb{T}}^k |A| \leqslant \frac{1}{C_k} |\operatorname{tr}^k A| \leqslant \frac{1}{C_k} \operatorname{lcav}(|\operatorname{tr}^k A|) = \frac{1}{C_k} \alpha_1 \cdots \alpha_k \leqslant \frac{\binom{n}{k}}{C_k} |\lambda_1 \cdots \lambda_k| \,.$$

□

Theorem 3.10. *Let A, $\lambda_1, \ldots, \lambda_n$, $\gamma_1, \ldots, \gamma_n$ be as in Lemma 3.9, and let k be a saturated index for the tropical characteristic polynomial $q_{|A|}$. Suppose that among the subsets of cardinality k of $\{1, \ldots, n\}$ there is a unique subset $\overline{I}_k$ for which there exists a (possibly not unique) permutation $\overline{\sigma} \in S_{\overline{I}_k}$ that realizes the maximum*

$$\max_{\substack{I \subset [n] \\ \# I = k}} \max_{\sigma \in S_I} \prod_{i \in I} |a_{i,\sigma(i)}| \tag{3.6}$$

(that is, $w_{|A|}(\overline{\sigma}) = \operatorname{tr}_{\mathbb{T}}^k |A|$). Suppose $\det A[\overline{I}_k, \overline{I}_k] \neq 0$. Finally suppose that, for any permutation σ of any subset of cardinality k except $\overline{I}_k$, $w_{|A|}(\sigma) \leqslant \delta_k \cdot w_{|A|}(\bar{\sigma}) = \delta_k \operatorname{tr}_{\mathbb{T}}^k |A|$, with

$$\delta_k < \frac{|\det A[\overline{I}_k, \overline{I}_k]|}{\operatorname{tr}_{\mathbb{T}}^k |A| \left(\binom{n}{k} - 1\right) k!}.$$

Then the inequality

$$L_k \gamma_1 \cdots \gamma_k \leqslant |\lambda_1 \cdots \lambda_k|$$

holds with

$$L_k = \frac{1}{\binom{n}{k}} \left(\frac{|\det A[\overline{I}_k, \overline{I}_k]|}{\operatorname{tr}_{\mathbb{T}}^k |A|} - \delta_k \left(\binom{n}{k} - 1\right) k! \right).$$

Proof. To prove the theorem it is sufficient to show that (3.4) holds with

$$C_k = \left(\frac{|\det A[\overline{I}_k, \overline{I}_k]|}{\operatorname{tr}_{\mathbb{T}}^k |A|} - \delta_k \left(\binom{n}{k} - 1 \right) k! \right) .$$

We have

$$\begin{aligned}
|\operatorname{tr}^k A| &= \Big|\sum_{\substack{I\in[n]\\ \#I=k}} \det A[I,I]\Big| \\
&\geqslant \Big|\det A[\overline{I}_k,\overline{I}_k]\Big| - \Big|\sum_{\substack{\#I=k\\ I\neq\overline{I}_k}} \det A[I,I]\Big| \\
&\geqslant \Big|\det A[\overline{I}_k,\overline{I}_k]\Big| - \sum_{\substack{\#I=k\\ I\neq\overline{I}_k}} \operatorname{per}\big|A[I,I]\big| \\
&\geqslant \Big|\det A[\overline{I}_k,\overline{I}_k]\Big| - \left(\binom{n}{k}-1\right) k!\delta_k \operatorname{tr}_{\mathbb{T}}^k |A| \,, \\
&\geqslant \left(\frac{|\det A[\overline{I}_k,\overline{I}_k]|}{\operatorname{tr}_{\mathbb{T}}^k |A|} - \delta_k\left(\binom{n}{k}-1\right)k!\right)\operatorname{tr}_{\mathbb{T}}^k |A| \\
&= C_k \operatorname{tr}_{\mathbb{T}}^k |A| \,,
\end{aligned}$$

and the hypothesis on δ_k guarantees that $C_k > 0$. □

If the maximum in (3.6) is attained by exactly one permutation, then the statement of Theorem 3.10 can be slightly modified as follows.

Theorem 3.11. *Let A, $\lambda_1,\dots,\lambda_n$, $\gamma_1,\dots,\gamma_n$ and k be as in Theorem 3.10. Suppose that the maximum in (3.6) is attained for a unique permutation $\bar\sigma$, and that for any other permutation σ of any k-subset of $\{1,\dots,n\}$ the inequality $\frac{w_{|A|}(\sigma)}{w_{|A|}(\bar\sigma)} \leqslant \eta_k$ holds for some*

$$\eta_k < \frac{1}{\left(\binom{n}{k}k! - 1\right)} \,.$$

Then the inequality

$$L_k\gamma_1\cdots\gamma_k \leqslant |\lambda_1\cdots\lambda_k|$$

holds with

$$L_k = \frac{1}{\binom{n}{k}}\left(1-\eta_k\left(\binom{n}{k}k!-1\right)\right) \,.$$

Proof. The arguments of the proof are the same as for Theorem 3.10. In the present case, we have

$$\begin{aligned}
|\operatorname{tr}^k A| &= \Big|\sum_{\substack{I\in[n]\\ \#I=k}} \det A[I,I]\Big| \\
&\geqslant \Big|w_{|A|}(\bar\sigma)\Big| - \Big|\sum_{\sigma\neq\bar\sigma} w_{|A|}(\sigma)\Big| \\
&\geqslant \operatorname{tr}_{\mathbb{T}}^k |A| - \left(\binom{n}{k}k!-1\right)\eta_k \operatorname{tr}_{\mathbb{T}}^k |A| \,,
\end{aligned}$$

and we conclude applying Lemma 3.9. □

3.3 Optimality of upper bound and comparison with bounds for polynomial roots

We now discuss briefly the optimality of the upper bound for some special classes of matrices. Throughout this paragraph, if A is a complex $n \times n$ matrix, then $\lambda_1, \dots, \lambda_n$ will be its eigenvalues (ordered by nonincreasing absolute value), and $\gamma_1 \geqslant \cdots \geqslant \gamma_n$ will be its tropical eigenvalues.

3.3.1 Monomial matrices.

Recall that a *monomial matrix* is the product of a diagonal matrix (with non-zero diagonal entries) and of a permutation matrix. We next show that the upper bound is tight for monomial matrices.

Proposition 3.12. *If A is a monomial matrix, then, for all $k \in \{1, \dots, n\}$, the inequality in Theorem 3.1 is tight,*

$$|\lambda_1 \cdots \lambda_k| = \rho(\textstyle\bigwedge^k_{\mathrm{per}}(\operatorname{pat} A))\gamma_1 \cdots \gamma_k \qquad \forall k \in \{1, \dots, n\} \tag{3.7}$$

Proof. We claim that if A is a monomial matrix, then, the absolute values of the eigenvalues of A coincide with the tropical eigenvalues of $|A|$, counted with multiplicities.

To see this, assume that $A = DC$ where D is diagonal and C is a matrix representing a permutation σ. If σ consists of several cycles, then, DC has a block diagonal structure, and so, the characteristic polynomial of A is the product of the characteristic polynomials of the diagonal blocks of A. The same is true for the tropical characteristic polynomial of $|A|$. Hence, it suffices to show the claim when σ consists of a unique cycle. Then, denoting by $d_1, \dots, d_n$ the diagonal terms of D, expanding the determinant of $xI - A$ or the permanent of $xI \oplus A$, one readily checks that the characteristic polynomial of A is $x^n - d_1 \dots d_n$, whereas the tropical characteristic polynomial of $|A|$ is $x^n \oplus |d_1 \dots d_n|$. It follows that the eigenvalues of A are the nth roots of $d_1 \dots d_n$, whereas the tropical eigenvalues of $|A|$ are all equal to $|d_1 \dots d_n|^{1/n}$. So, the claim is proved.

It remains to show that $\rho(\bigwedge^k_{\mathrm{per}}(\operatorname{pat} A)) = 1$. Note that $\operatorname{pat} A = C$. We claim that $\bigwedge^k_{\mathrm{per}} C$ is a permutation matrix. In fact, for any fixed $k \in \{1, \dots, n\}$, let I be a subset of cardinality k of $\{1, \dots, n\}$. Since C is a permutation matrix, there is one and only one subset $J \subset \{1, \dots, n\}$ such that $\operatorname{per} C[I, J] \neq 0$: precisely, if C represents the permutation $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$, then $\operatorname{per} C[I, \sigma(I)] = 1$ and $\operatorname{per} C[I, J] = 0\ \forall J \neq \sigma(I)$. This means that each row of $\bigwedge^k_{\mathrm{per}} C$ contains exactly one 1, and the other entries are zeroes. Since the same reasoning is also valid for columns, we can conclude that $\bigwedge^k_{\mathrm{per}} C$ is a permutation matrix, and as such its spectral radius is 1. □

3.3.2 Full matrices.

Monomial matrices are among the sparsest matrices we can think of. One may wonder what happens in the opposite case, when all the matrix entries are nonzero. We next discuss a class of instances of this kind, in which the upper bound is not tight. We only consider the case $k = n$ for brevity, although it is not the only case for which the equality fails to hold.

Proposition 3.13. *Let $A = (a_{i,j})$ be a $n \times n$ complex matrix, $n \geqslant 3$, and suppose $|a_{i,j}| = 1$ for all $i, j \in \{1, \dots, n\}$. Then the inequality in Theorem 3.1 can not be tight for $k = n$.*

Proof. For any couple (I, J) of k-subsets of $\{1, \dots, n\}$, the (I, J) element of the matrix $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$ is given by the permanent of a $k \times k$ matrix of ones, that is $k!$; so $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$ is a $\binom{n}{k} \times \binom{n}{k}$ matrix with all entries equal to $k!$. Its spectral radius is therefore $\binom{n}{k}k!$ (and $(1, \dots, 1)^\top$ is an eigenvector for the maximum eigenvalue). For $k = n$, $\rho(\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A))$ reduces to $n!$, so our upper bound would be $|\lambda_1, \cdots, \lambda_n| \leqslant n! \cdot \gamma_1 \cdots \gamma_n$. Now, the left-hand side can be thought of as $|\det A|$, and on the other hand $\gamma_1 = \cdots = \gamma_n = 1$ (the tropical characteristic polynomial is $q_A(x) = x^n \oplus x^{n-1} \oplus \cdots \oplus x \oplus 1 = x^n \oplus 1 = (x \oplus 1)^n \ \forall x \geqslant 0$). So the inequality in Theorem 3.1 is equivalent to $|\det A| \leqslant n!$. But the well-known Hadamard bound for the determinant yields in this case $|\det A| \leqslant (\sqrt{n})^n = n^{n/2}$, and since $n^{n/2} < n! \ \forall n \geqslant 3$ the inequality of Theorem 3.1 can not be tight. □

3.3.3 Comparison with the Hadamard-Pólya's bounds for polynomial roots.

Finally, we discuss the behavior of the upper bound of Theorem 3.1 for the case of a companion matrix. Since the eigenvalues of a companion matrix are exactly the roots of its associated polynomial, this will allow a comparison between the present matrix bounds and the upper bound of Hadamard and Pólya discussed in the introduction. We start by showing that the usual property of companion matrices remains true in the tropical setting.

Lemma 3.14. *Consider the polynomial $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, and let A be its companion matrix. Then the tropical eigenvalues of A are exactly the tropical roots of p.*

Proof. The matrix is

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix} .$$

By definition, its tropical eigenvalues are the tropical roots of the polynomial $q_A(x) = \bigoplus_{k=0}^n \mathrm{tr}^k_{\mathbb{T}}|A|\, x^{n-k}$, so to verify the claim it is sufficient to show that $\mathrm{tr}^k_{\mathbb{T}}|A| = |a_{n-k}|$ for all $k \in \{0, \dots, n\}$. Recall that $\mathrm{tr}^k_{\mathbb{T}}|A|$ is the maximal tropical permanent of a $k \times k$ principal submatrix of $|A|$ (see Equation (2.3)). It is easy to check that the only principal submatrices with a non-zero contribution are those of the form $|A|[I_k, I_k]$ with $I_k = \{n - k + 1, \dots, n\}$, and in this case $\mathrm{per}_{\mathbb{T}}\, |A|[I_k, I_k] = |a_{n-k}|$. □

Lemma 3.15. *If A is the companion matrix of a polynomial of degree n, then,*

$$\rho(\textstyle\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)) \leqslant \min(k + 1, n - k + 1) \ .$$

Proof. First, we note that nonzero entries of $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$ can only be 1's, because $\mathrm{pat}\, A$ is a $(0, 1)$-matrix, and the tropical permanent of any of its square submatrices has at most one non-zero term. By computing explicitly the form of $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$, for example following the method used by Moussa in [Mou97], we see that each column of $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$ has either one or $k + 1$ nonzero entries, and each row has either one or $n - k + 1$ nonzero entries. In terms of matrix norms, we have $\|\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)\|_1 = k + 1$, and $\|\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)\|_\infty = n - k + 1$. Since both these norms are upper bounds for the spectral radius, we can conclude that $\rho(\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)) \leqslant \min(k + 1, n - k + 1)$. □

Thus, by specializing Theorem 3.1 to companion matrices, we recover the version of the upper bound (2.5) originally derived by Hadamard, with the multiplicative constant $k+1$. By comparison, the multiplicative constant in Lemma 3.15 is smaller due to its symmetric nature. However, it was observed by Ostrowski that the upper bound in (2.5) can be strengthened by exploiting symmetry. We give a formal argument for the convenience of the reader.

Lemma 3.16. *Let $P = \{p \in \mathbb{C}[z] \mid \deg p = n\}$ be the set of complex polynomials of degree n. Denote the roots and the tropical roots as above. Suppose that the inequality $|\zeta_1 \cdots \zeta_k| \leqslant f(k) \cdot \alpha_1 \cdots \alpha_k$ holds for some function f, for all $k \in \{1, \dots, n\}$ and for all polynomials $p \in P$. Then the inequality $|\zeta_1 \cdots \zeta_k| \leqslant f(n-k) \cdot \alpha_1 \cdots \alpha_k$ also holds for all $k \in \{1, \dots, n\}$ and for all polynomials $p \in P$.*

Proof. Consider a polynomial $p \in P$, $p(z) = a_n z^n + \cdots + a_0$ with roots $\zeta_1, \dots, \zeta_n$ (ordered by nonincreasing absolute value) and tropical roots $\alpha_1 \geqslant \ldots \geqslant \alpha_n$. Arguing by density, we may assume that $a_0 \neq 0$. Then, we build its reciprocal polynomial $p^*(z) = z^n p(1/z) = a_0 z^n + \cdots + a_n$. It is clear that the roots of p^* are $\zeta_1^{-1}, \dots, \zeta_n^{-1}$. Moreover, its tropical roots are $\alpha_1^{-1} \leqslant \ldots \leqslant \alpha_n^{-1}$: this can be easily proved by observing that the Newton polygon of $p*$ is obtained from the Newton polygon of p by symmetry with respect to a vertical axis, and thus it has opposite slopes.

Since $p^* \in P$, by hypothesis we can bound its $n-k$ largest roots:

$$\left| \frac{1}{\zeta_n} \cdots \frac{1}{\zeta_{k+1}} \right| \leqslant f(n-k) \cdot \frac{1}{\alpha_n} \cdots \frac{1}{\alpha_{k+1}} .$$

By applying Corollary 2.4 (and observing that 0 is a saturated index for the max-times relative $p^\times$) we also have

$$|a_0| = |a_n| \alpha_1 \cdots \alpha_n ,$$

so we can write

$$\begin{aligned} |\zeta_1 \cdots \zeta_k| &= |\zeta_1 \cdots \zeta_n| \left| \frac{1}{\zeta_n} \cdots \frac{1}{\zeta_{k+1}} \right| \\ &= \frac{|a_0|}{|a_n|} \left| \frac{1}{\zeta_n} \cdots \frac{1}{\zeta_{k+1}} \right| \\ &\leqslant \frac{|a_0|}{|a_n|} \cdot f(n-k) \cdot \frac{1}{\alpha_n} \cdots \frac{1}{\alpha_{k+1}} \\ &= \alpha_1 \cdots \alpha_n \cdot f(n-k) \cdot \frac{1}{\alpha_n} \cdots \frac{1}{\alpha_{k+1}} = f(n-k) \cdot \alpha_1 \cdots \alpha_k \end{aligned}$$

□

Therefore, it follows from the Pólya's upper bound (2.5) that

$$|\zeta_1 \cdots \zeta_k| \leqslant \min \left(\sqrt{\frac{(k+1)^{k+1}}{k^k}}, \sqrt{\frac{(n-k+1)^{n-k+1}}{(n-k)^{n-k}}} \right) \alpha_1 \cdots \alpha_k,$$

for all $k \in \{1, \dots, n\}$. This is tighter than the bound derived from Theorem 3.1 and Lemma 3.15. In the latter lemma, we used a coarse estimation of the spectral radius, via norms. A finer bound can be obtained by computing the true spectral radius of $\bigwedge^k_{\mathrm{per}}(\mathrm{pat}\, A)$

for the companion matrix A, but numerical experiments indicate this still does not improve Pólya's bound. This is perhaps not surprising as the latter is derived by analytic functions techniques (Jensen inequality and Parseval identity), which do not naturally carry over to the more general matrix case considered here.

CHAPTER 4

Hungarian scaling of asymptotic matrix polynomials

In this chapter we propose a new scaling technique to be applied to matrix polynomials before computing its eigenvalues via linearization, aimed at improving the numerical stability of the computed eigenvalues.

The chapter is organized as follows. In Section 4.1, we recall the main results of [ABG04], including the definition and properties of the tropical eigenvalues, in relation with the optimal assignment problem. In Section 4.2, some auxiliary properties on the optimal dual variables of this problem are established or recalled from [ABG15]. Then, the two following sections concern matrix polynomials depending on a small parameter ϵ. Since the condition number of an eigenvalue depends on the corresponding left and right eigenvectors, we first study in Section 4.3 the asymptotics of these eigenvectors. In particular, we show in Corollary 4.16 that, under irreducibility and genericity assumptions, the valuations of the eigenvectors are equal to the optimal dual variables of the assignment problem associated to the valuation of the eigenvalue. We then study, in Section 4.4, the asymptotics of eigenvalue condition numbers. We show in Theorem 4.19 that for a simple nonzero eigenvalue of the matrix polynomial, the condition number of the polynomial eigenvalue problem for the scaled matrix polynomial remains bounded as $\epsilon \to 0$. In contrast, under a genericity assumption, and if the scaling is not trivial, the same eigenvalue condition number tends to infinity for the unscaled matrix polynomial (Corollary 4.23).

Then, we consider matrix polynomials with fixed coefficients (independent of ϵ). We give in Section 4.5 quantitative bounds comparing the eigenvalue condition numbers in the scaled

and unscaled cases. These indicate that the eigenvalue condition number is reduced after scaling, under some structural assumptions on the orders of magnitudes of the entries of the matrices. We also consider in Section 4.6 the effect of the diagonal scaling on the condition number of the eigenvalue problem for the companion linearization of the matrix polynomial, and demonstrate a similar improvement, by showing that, after the diagonal scaling, the eigenvalue condition numbers for the matrix polynomial and for its companion linearization are close as soon the tropical eigenvalue is close to the modulus of the complex eigenvalue. We present in Section 4.7 numerical results, showing the efficiency of the scaling for families of deformed matrix polynomials.

The present chapter is part of a manuscript in preparation as a joint work with Akian, Gaubert, Hook and Tisseur [AGH+15].

4.1 Asymptotics of eigenvalues and parametric optimal assignment problems

We recall here the result of [ABG04], which determines the images by the nonarchimedean valuation of the eigenvalues of a matrix polynomial over the field of converging Puiseux series. More generally, the result there applies to entries in the ring of functions with a polynomial growth at point 0, that we next define.

We denote by $\mathcal{C}$ the set of germs at point 0 of complex valued continuous functions f of the nonnegative parameter ϵ, having an asymptotic expansion around zero of the form

$$f(\epsilon) = a\epsilon^{-A} + o(\epsilon^{-A})$$

for some $a \in \mathbb{C}$ and $A \in \mathbb{R} \cup \{-\infty\}$, with the convention that $\epsilon^{+\infty} = 0$. By germ, we mean that we identify two functions that coincide in a sufficiently small neighborhood of zero. The set $\mathcal{C}$ is a ring, which includes in particular the field $\mathbb{K}$ of Puiseux series in the parameter ϵ that are absolutely converging when ϵ is of small enough modulus. It also includes fields of generalized power series with real exponents [vdDS98].

When $a \neq 0$ or $A = -\infty$, we have the asymptotic equivalent $f(\epsilon) \sim a\epsilon^{-A}$ as ϵ tends to zero. Then, we shall say that f has a *first order asymptotics*, and that $a\epsilon^{-A}$ is the *leading monomial* of f and A the *leading exponent*.

We consider a matrix polynomial of degree d in the indeterminate Y, whose coefficients are $n \times n$ matrices $\mathcal{A}_0, \dots, \mathcal{A}_d$ with entries in $\mathcal{C}$:

$$\mathcal{A} = \mathcal{A}_0 + \mathsf{Y}\mathcal{A}_1 + \cdots + \mathsf{Y}^d\mathcal{A}_d \ . \tag{4.1}$$

Equivalently, $\mathcal{A}$ is a $n \times n$ matrix, the coefficients of which are polynomials of degree at most d with coefficients in $\mathcal{C}$ ($\mathcal{A}_{ij} \in \mathcal{C}[\mathsf{Y}]$). The *characteristic polynomial* of $\mathcal{A}$ is the polynomial $P_{\mathcal{A}} := \det(\mathcal{A}) \in \mathcal{C}[\mathsf{Y}]$. We say that $\mathcal{A}$ is *regular* if $P_{\mathcal{A}}$ is nonzero, and in that case, the *eigenvalues* of $\mathcal{A}$ are the roots of the polynomial $P_{\mathcal{A}}$. Also if the degree of $P_{\mathcal{A}}$ is smaller than nd, we say that ∞ is an eigenvalue of $\mathcal{A}$ with multiplicity $nd - \deg P_{\mathcal{A}}$, so that $P_{\mathcal{A}}$ has always nd finite or infinite eigenvalues, counting multiplicities.

We shall occasionally write $\mathcal{A}_\epsilon$ instead of $\mathcal{A}$ to emphasize the dependence in ϵ. We denote by $\mathcal{A}_{ij}$ the (i,j) entry of any matrix or matrix polynomial $\mathcal{A}$, so that for instance $(\mathcal{A}_k)_{ij}$ denotes the (i,j) entry of $\mathcal{A}_k$. We assume that, for all $k \in \{0, \dots, d\}$, there are matrices $a_k \in \mathbb{C}^{n\times n}$ and $A_k \in (\mathbb{R} \cup \{-\infty\})^{n\times n}$, such that

$$(\mathcal{A}_k)_{ij} = (a_k)_{ij}\epsilon^{-(A_k)_{ij}} + o(\epsilon^{-(A_k)_{ij}}) \ , \qquad \text{for all } i,j \in \{1, \dots, n\} \ . \tag{4.2}$$

Our goal is to determine the leading monomials of the eigenvalues of $\mathcal{A}$ knowing only the coefficient matrices a_k and of the exponent matrices A_k. To this end we associate to the matrix polynomial $\mathcal{A} \in \mathcal{C}[\mathsf{Y}]^{n\times n}$ the $n \times n$ tropical matrix polynomial of degree d

$$A = A_0 \oplus \mathsf{Y} A_1 \oplus \cdots \oplus \mathsf{Y}^d A_d \ , \tag{4.3}$$

where the matrices A_k are $n \times n$ matrices over $\mathbb{R}_{\max}$, that is, as elements of $\mathbb{R}_{\max}^{n\times n}$, so that $A \in \mathbb{R}_{\max}[\mathsf{Y}]^{n\times n}$. The result of [ABG04] that we recall below shows that under some nondegeneracy assumptions the leading exponents of the eigenvalues of $\mathcal{A}$ are characterized by the *tropical eigenvalues* of A, i.e. the tropical roots the tropical characteristic polynomial of A, defined as

$$P_A := \mathrm{per}_{\mathbb{R}_{\max}[\mathsf{Y}]} A \in \mathbb{R}_{\max}[\mathsf{Y}] \ . \tag{4.4}$$

In particular, if P_A has a positive valuation $\mathrm{val}\, P_A$, $\mathbb{0}$ is a tropical eigenvalue of A, with multiplicity $\mathrm{val}\, P_A$. Moreover, when the degree $\deg P_A$ is less than nd, we shall say, by convention, that $+\infty$ is a tropical eigenvalue of A with multiplicity $nd - \deg P_A$.

The tropical eigenvalues of A can be computed in polynomial time. To see this, observe that evaluating the tropical characteristic polynomial function P_A at a given point y reduces to compute $\mathrm{per}\, \widehat{A}(y)$, where $\widehat{A}(y) = A_0 \oplus A_1 y \oplus \cdots \oplus A_d y^d \in \mathbb{R}_{\max}^{n\times n}$ is the evaluation of A at y. For any matrix $B \in \mathbb{R}_{\max}^{n\times n}$, we have

$$\mathrm{per}\, B = \max_{\sigma\in S_n} \sum_{1\leqslant i\leqslant n} B_{i\sigma(i)} \ .$$

As we anticipated in Chapter 2, this latter expression is the value of the optimal assignment problem with weight matrix B. It can be computed in strongly polynomial time, by the Hungarian algorithm which uses $O(n^3)$ arithmetic operations. Alternatively, it can be computed in weakly polynomial time, by an algorithm of Gabow and Tarjan [GT88], which uses $O(n^{2.5}\log(nW))$ operations, where W denotes the greatest absolute value of the finite entries of B, assuming that these entries are integers. We refer the reader to [BR97, § 2.4] or [Sch03, § 17] for more background on the optimal assignment problem.

Computing the tropical characteristic polynomial function P_A amounts to solving a *parametric optimal assignment* problem, in which which the weight matrix depends of the parameter y. Burkard and Butkovič [BB03] considered the situation in which $d = 1$ and $A_1 = I$ is the identity matrix, and showed that the eigenvalues of the matrix polynomial $A = A_0 + I\mathsf{Y}$ can be computed in $O(n^4)$ time, by calling $O(n)$ times an oracle solving an optimal assignment problem. Gassner and Klinz [GK10] showed that this can be reduced in $O(n^3)$ using parametric optimal assignment techniques. The same ideas apply to tropical matrix polynomial functions of any degree. In particular, an adaptation of the method of [BB03] readily leads to an algorithm running in time $O(n^4 d)$, whereas Hook [Hoo13] adapted the method of [GK10] to get an algorithm running in time $O(n^3 d^2)$. Therefore, computing the tropical eigenvalues is not more expensive than computing the classical eigenvalues, and it is actually cheaper as combinatorial algorithms are available. The interest of the tropical eigenvalues is also that they can be computed in a numerically robust way, as the optimal assignment algorithms are insensitive to numerical errors.

The following result motivates the introduction of tropical eigenvalues. We shall say that a property depending on N complex parameters holds *generically* if the set of parameters for which it fails is included in an algebraic hypersurface of $\mathbb{C}^N$.

Theorem 4.1 ([ABG04, ABG15]). *Let $\mathcal{A} = \mathcal{A}_0 + \mathsf{Y}\mathcal{A}_1 + \cdots + \mathsf{Y}^d\mathcal{A}_d \in \mathcal{C}[\mathsf{Y}]^{n\times n}$, and assume that the entries of the matrices $\mathcal{A}_0, \ldots, \mathcal{A}_d$ have asymptotic expansions of the form (4.2). Then, for generic values of the parameters $(a_k)_{ij}$, $i, j \in [n]$, $k \in [d]$, the leading exponents of the eigenvalues of $\mathcal{A}$ are precisely the tropical eigenvalues of A, and the multiplicities are the same.*

This theorem can be made more precise, by characterizing the leading coefficients (not only the leading exponents) of the eigenvalues of $\mathcal{A}$. These coefficients will be seen to depend on the leading coefficients of $\mathcal{A}$ that participate to optimal permutations, in a sense that we next explain. For this purpose, we need to recall the characterization of optimal assignments in terms of the optimal dual variables.

To any matrix $B \in \mathbb{R}_{\max}^{n\times n}$, we associate the directed graph $G(B)$ with set of nodes $[n]$ and an arc (i, j) if $B_{ij} \neq -\infty$. In the sequel, we shall omit the word "directed" as all graphs will be directed.

The optimal assignment problem for a weight matrix B can be cast as the following linear programming problem

$$\max_X \sum_{\substack{i,j\in[n]\\ B_{ij}\neq-\infty}} B_{ij}X_{ij} \ ,$$

where the maximum is taken over the set of bistochastic matrices (nonnegative matrices of column and row sum one) whose support is included in the set of arcs of B. This linear programming problem has a dual, which reads

$$\min_{U,V\in\mathbb{R}^n} \sum_{i\in[n]} U_i + \sum_{j\in[n]} V_j, \qquad U_i + V_j \geqslant B_{ij},\ i, j \in [n] \ .$$

If $\operatorname{per} B \neq -\infty$, meaning the primal problem is feasible, we define the graph $\operatorname{Opt}(B)$ as the set of arcs belonging to optimal assignments: the set of nodes of $\operatorname{Opt}(B)$ is $[n]$ and there is an arc from i to $j \in [n]$ if there exists $\sigma \in S_n$ such that $j = \sigma(i)$ and $\operatorname{per} B = B_{1\sigma(1)} + \cdots + B_{n\sigma(n)}$.

We shall call *Hungarian pair* with respect to B any optimal solution (U, V) of the dual problem. The strong duality theorem in linear programming implies that $U_1 + \cdots + U_n + V_1 + \cdots + V_n = \operatorname{per} B$.

For any Hungarian pair (U, V), we now define the *saturation graph*, $\operatorname{Sat}(B, U, V)$, which has set of nodes $[n]$ and an arc from i to $j \in [n]$ if $B_{ij} = U_i + V_j$.

By complementary slackness, if σ is an optimal permutation, the support of σ is included in $\operatorname{Sat}(B, U, V)$, and conversely, if σ is a permutation whose support is included in $\operatorname{Sat}(B, U, V)$, then it is optimal. Therefore $\operatorname{Opt}(B)$ is included in $\operatorname{Sat}(B, U, V)$, moreover, an arc is in $\operatorname{Opt}(B)$ if and only if it belongs to a cover of $\operatorname{Sat}(B, U, V)$ by disjoint oriented cycles, so $\operatorname{Opt}(B)$ can be recovered from $\operatorname{Sat}(B, U, V)$. The interest of $\operatorname{Sat}(U, V, B)$ is that it can be readily computed.

For any tropical matrix polynomial A, any scalar $\gamma \in \mathbb{R}_{\max}$, and any $k \in \{0, \ldots, d\}$, we denote by $\mathcal{G}_k(A, \gamma)$ the graph with set of nodes $[n]$, and an arc from i to $j \in [n]$ if $\gamma^k(A_k)_{ij} = \widehat{A}_{ij}(\gamma)$ and $\widehat{A}_{ij}(\gamma) \neq \mathbb{0}$. This is a subgraph of the graph of $\widehat{A}(\gamma)$. For any graphs G and G', the intersection $G \cap G'$ is the graph the set of nodes (resp. arcs) of which is the intersection of the sets of nodes (resp. arcs) of G and G'. If G is any graph with set of nodes $S \subset [n]$, and if $b \in \mathbb{C}^{n\times n}$, we define the matrix $b^G \in \mathbb{C}^{S\times S}$ by $(b^G)_{ij} = b_{ij}$ if $(i, j) \in G$, and $(b^G)_{ij} = 0$ if $(i, j) \in S \times S \setminus G$.

The following result was stated in [ABG04] (with the notation of the min-plus semifield), the proof can be found in [ABG15].

Theorem 4.2 ([ABG04, Thm 1.1] and [ABG15, Thm 8.2 and 8.3]). *Let $\mathcal{A}$ be a regular matrix polynomial over $\mathcal{C}$ as in (4.1) satisfying (4.2) and denote by A the tropical matrix polynomial (4.3) with coefficients A_k as in (4.2). Then A is a regular matrix polynomial, and we have:* $\operatorname{val} P_{\mathcal{A}_\epsilon} \geqslant \operatorname{val} P_A$, $\deg P_{\mathcal{A}_\epsilon} \leqslant \deg P_A$.

Let γ denote any finite ($\neq \pm\infty$) algebraic eigenvalue of A, and denote by $m_{\gamma,A}$ its multiplicity. Let us also denote by $m'_{\gamma,A}$ the sum of the multiplicities of all the algebraic eigenvalues of A smaller than γ ($-\infty$ included), putting $m'_{\gamma,A} = 0$ if no such eigenvalues exist. Let G be equal either to $\operatorname{Opt}(\widehat{A}(\gamma))$ or $\operatorname{Sat}(\widehat{A}(\gamma), U, V)$, for any choice of the Hungarian pair (U, V) with respect to $\widehat{A}(\gamma)$, and let $G_k = \mathcal{G}_k(A, \gamma) \cap G$ for all $0 \leqslant k \leqslant d$. Consider the matrix polynomial

$$a^{(\gamma,G)} := a_0^{G_0} + \mathsf{Y} a_1^{G_1} + \cdots + \mathsf{Y}^d a_d^{G_d} \in \mathbb{C}[\mathsf{Y}]^{n\times n} \quad , \tag{4.5}$$

and assume that it is regular. If $a^{(\gamma,G)}$ has m_γ nonzero eigenvalues, $\lambda_1, \ldots, \lambda_{m_\gamma}$, then $m_\gamma \leqslant m_{\gamma,A}$ and the matrix polynomial $\mathcal{A}_\epsilon$ has m_γ eigenvalues $\zeta_{\epsilon,1}, \ldots, \zeta_{\epsilon,m_\gamma}$ with first order asymptotics of the form $\zeta_{\epsilon,i} \sim \lambda_i \epsilon^{-\gamma}$. If 0 is an eigenvalue of the matrix polynomial $a^{(\gamma)}$ with multiplicity m'_γ, then $m'_\gamma \geqslant m'_{\gamma,A}$, and the matrix polynomial $\mathcal{A}_\epsilon$ has precisely m'_γ eigenvalues ζ_ϵ such that $\zeta_\epsilon = o(\epsilon^{-\gamma})$. All the other eigenvalues ζ_ϵ of $\mathcal{A}_\epsilon$ are such that the modulus of $\epsilon^\gamma \zeta_\epsilon$ converges to infinity.

Moreover, for generic values of the parameters $(a_k)_{ij}$, the matrix polynomial $a^{(\gamma,G)}$ is regular, and all the above inequalities are equalities: $\operatorname{val} P_{\mathcal{A}_\epsilon} = \operatorname{val} P_A$, $\deg P_{\mathcal{A}_\epsilon} = \deg P_A$, $m_\gamma = m_{\gamma,A}$ *and* $m'_\gamma = m'_{\gamma,A}$.

4.2 Preliminaries on Hungarian pairs

In the present section, we recall some of the results on Hungarian pairs that are proved in [ABG15, §9] as preliminaries for the proof of Theorem 4.2, and that we shall also need for the proof of our main results on asymptotics of eigenvectors and on condition numbers of eigenvalues. We also state and prove some additional necessary results on Hungarian pairs.

We shall adopt the following notation. For any $U \in \mathbb{R}^n$, we denote by $\mathsf{d}_{\mathrm{m}}(U)$ the diagonal $n \times n$ matrix over $\mathbb{R}_{\max}$ such that $(\mathsf{d}_{\mathrm{m}}(U))_{ii} = U_i$, so that $(\mathsf{d}_{\mathrm{m}}(U))_{ij} = -\infty$ for $i \neq j$. For any permutation σ of $[n]$, we denote by P^σ_{m} its associated tropical matrix: $(P^\sigma_{\mathrm{m}})_{ij} = \mathbb{1}$ if $j = \sigma(i)$ and $(P^\sigma_{\mathrm{m}})_{ij} = \mathbb{0}$ otherwise. We also denote by P^σ its associated complex matrix: $(P^\sigma)_{ij} = 1$ if $j = \sigma(i)$ and $(P^\sigma)_{ij} = 0$ otherwise. Moreover, if $U \in \mathbb{R}^n$, we use the notation $U_\sigma := P^\sigma_{\mathrm{m}} U$ which is simply $(U_{\sigma(i)})_{i=1,\ldots,n}$. We shall say that a matrix $M \in \mathbb{R}^{n\times n}_{\max}$ is *monomial* if it can be written as $M = \mathsf{d}_{\mathrm{m}}(U) P^\sigma_{\mathrm{m}}$ for some $U \in \mathbb{R}^n$ and $\sigma \in S_n$. We have equivalently $M = P^\sigma_{\mathrm{m}} \mathsf{d}_{\mathrm{m}}(V)$, by taking $V = U_{\sigma^{-1}}$. Recall that the monomial matrices are the only invertible matrices over $\mathbb{R}_{\max}$, and that $M^{-1} = P^{\sigma^{-1}}_{\mathrm{m}} \mathsf{d}_{\mathrm{m}}(-U)$. Now if G is a graph with set of nodes $[n]$, and $\sigma, \tau \in S_n$, we denote by $G_{\sigma,\tau}$ the graph with same set of nodes $[n]$, and an arc (i, j) if and only if $(\sigma(i), \tau^{-1}(j))$ is an arc of G.

We denote by $\mathbb{1}$ the vector of $\mathbb{R}^n$ with all its entries equal to $\mathbb{1} = 0$. The following results and their corollaries are proved in [ABG15, §9]. The first ones allow one to normalize matrices in a suitable way, the last ones show that one can take either $G = \operatorname{Sat}(B, U, V)$ or $G = \operatorname{Opt}(B)$ in Theorem 4.2.

Lemma 4.3 ([ABG15, §9]). *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, and let $M = \mathsf{d}_{\mathrm{m}}(W)P_{\mathrm{m}}^{\sigma}$ and $N = \mathsf{d}_{\mathrm{m}}(X)P_{\mathrm{m}}^{\tau}$ be monomial matrices, with $W, X \in \mathbb{R}^n$ and $\sigma, \tau \in S_n$. Then*

$$\operatorname{per}(MBN) = W_1 \cdots W_n X_1 \cdots X_n \operatorname{per} B = (\operatorname{per} M)(\operatorname{per} B)(\operatorname{per} N) \ ,$$

ν is an optimal permutation for MBN if and only if $\tau^{-1} \circ \nu \circ \sigma^{-1}$ is an optimal permutation for B, and $\operatorname{Opt}(MBN) = \operatorname{Opt}(B)_{\sigma,\tau}$.

Let (U, V) be a Hungarian pair with respect to B. Then $(MU, N^T V)$ is a Hungarian pair with respect to MBN, and we have $\operatorname{Sat}(MBN, MU, N^T V) = \operatorname{Sat}(B, U, V)_{\sigma,\tau}$.

Corollary 4.4 ([ABG15, §9]). *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, and (U, V) be a Hungarian pair with respect to B. Then $(\mathbb{1}, \mathbb{1})$ is a Hungarian pair with respect to the matrix $C := \mathsf{d}_{\mathrm{m}}(U)^{-1} B\, \mathsf{d}_{\mathrm{m}}(V)^{-1} \in \mathbb{R}_{\max}^{n\times n}$, $\operatorname{Sat}(B, U, V) = \operatorname{Sat}(C, \mathbb{1}, \mathbb{1})$ and $\operatorname{Opt}(B) = \operatorname{Opt}(C)$. In particular, $\operatorname{per} C = \mathbb{1}$, $C_{ij} \leqslant \mathbb{1} = 0$ for all $i, j \in [n]$ and (i, j) is an arc of $\operatorname{Sat}(B, U, V)$ if and only if $C_{ij} = \mathbb{1}$.*

Corollary 4.5 ([ABG15, §9]). *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, let σ be an optimal permutation for B, and let (U, V) be a Hungarian pair with respect to B. Then, the identity map is an optimal permutation for $P_{\mathrm{m}}^{\sigma^{-1}} B$ and $(U_{\sigma^{-1}}, V)$ is a Hungarian pair with respect to $P_{\mathrm{m}}^{\sigma^{-1}} B$. Moreover, we have $\operatorname{per}(P_{\mathrm{m}}^{\sigma^{-1}} B) = \operatorname{per}(B)$, $\operatorname{Opt}(P_{\mathrm{m}}^{\sigma^{-1}} B) = \operatorname{Opt}(B)_{\sigma^{-1},\mathrm{id}}$, and $\operatorname{Sat}(P_{\mathrm{m}}^{\sigma^{-1}} B, U_{\sigma^{-1}}, V) = \operatorname{Sat}(B, U, V)_{\sigma^{-1},\mathrm{id}}$.*

Corollary 4.6 ([ABG15, §9]). *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, and (U, V) be a Hungarian pair with respect to B. Then $\operatorname{Opt}(B) \subset \operatorname{Sat}(B, U, V)$.*

Proposition 4.7 ([ABG15, §9]). *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, (U, V) be a Hungarian pair with respect to B, and the identity map is an optimal permutation of B. Then $\operatorname{Opt}(B)$ is the disjoint union of the strongly connected components of $\operatorname{Sat}(B, U, V)$.*

The previous result was the main ingredient of the proof of the following result.

Lemma 4.8 ([ABG15, Lemma 10.1]). *Let b be a matrix polynomial with coefficients in $\mathbb{C}^{n\times n}$ and degree d, let $B \in \mathbb{R}_{\max}^{n\times n}$ be a tropical matrix such that $\operatorname{per} B \neq \mathbb{0}$, and let (U, V) be a Hungarian pair with respect to B. Then the matrix polynomials $b^G := b_0^G + \cdots + \mathsf{Y}^d b_d^G$ defined respectively with $G = \operatorname{Sat}(B, U, V)$ and with $G = \operatorname{Opt}(B)$ have same eigenvalues.*

We now prove some additional results, which will help us to find the best eigenvalue condition number.

Proposition 4.9. *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B = \mathbb{1}$, $(\mathbb{1}, \mathbb{1})$ is a Hungarian pair with respect to B and the identity map is an optimal permutation of B. Then (U, V) is a Hungarian pair with respect to B if and only if $U, V \in \mathbb{R}^n$, with $BU = U$, and $V_i = U_i^{-1}$ for all $i \in [n]$.*

Proof. Let (U, V) be a Hungarian pair with respect to B. By definition, $B_{ij} \leqslant U_i V_j$, and $U_1 \cdots U_n V_1 \cdots V_n = \mathbb{1} = \operatorname{per} B$. By Assumption, $(\mathbb{1}, \mathbb{1})$ is a Hungarian pair with respect to B, and since also (U, V) is a Hungarian pair with respect to B, we deduce, by Corollary 4.6, that $\operatorname{Opt}(B) \subset \operatorname{Sat}(B, U, V)$ and $\operatorname{Opt}(B) \subset \operatorname{Sat}(B, \mathbb{1}, \mathbb{1})$. Hence $B_{ij} = U_i V_j = \mathbb{1}$ for all arcs (i, j) in $\operatorname{Opt}(B)$. Since the identity map is an optimal permutation of B, we get that, for all $i \in [n]$, (i, i) is an arc of $\operatorname{Opt}(B)$, hence $B_{ii} = U_i V_i = \mathbb{1}$, so $V_i = U_i^{-1}$. Then $B_{ij} U_j \leqslant U_i$ for all $i, j \in [n]$ and since $B_{ii} U_i = U_i$, we deduce that $BU = U$. Conversely, if $U \in \mathbb{R}^n$, $BU = U$, and $V = (U_i^{-1})_{i\in[n]}$, then $B_{ij} \leqslant U_i V_j$ and $U_1 \cdots U_n V_1 \cdots V_n = \mathbb{1} = \operatorname{per} B$, so that (U, V) is a Hungarian pair with respect to B. □

For a matrix $A \in \mathbb{R}_{\max}^{n\times n}$, we denote by $A^\star$ its Kleene star, defined as $A^\star = \oplus_{n\geqslant 0} A^n$, with the convention that $A^0 = I$, and by $\rho_{\max}(A)$ the maximal eigenvalue of A, which is also the maximal circuit mean weight of the graph of A. Moreover, for $i \in [n]$, $A_{\cdot i}$ denotes the ith column of A.

Proposition 4.10. *Let B be as in Proposition 4.9. Let $c_1, \dots, c_m$ be the sets of nodes of the strongly connected components of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$. Consider the matrix $A \in \mathbb{R}_{\max}^{m\times m}$ such that $A_{kk} = \mathbb{0}$ and $A_{k\ell} = \max\{B_{ij} \mid i \in c_k,\ j \in c_\ell\}$, for all $k \neq \ell \in [m]$. Then $\alpha := \rho_{\max}(A) < \mathbb{1}$, and $W = \oplus_{i\in[m]} (\alpha^{-1}A)^\star_{\cdot i}$ belongs to $\mathbb{R}^m$ and is an eigenvector of A. Consider the vectors $U, V \in \mathbb{R}^n$ such that $U_i = V_i^{-1} = W_k$ for all $i \in c_k$ and $k \in [m]$. Then (U, V) is a Hungarian pair with respect to B, such that $B_{ij} \leqslant \alpha U_i V_j$ for all $i, j \in [n]$ belonging to different strongly connected components of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$, and it satisfies $\mathrm{Sat}(B, U, V) = \mathrm{Opt}(B)$.*

Proof. By Proposition 4.7, $c_1, \dots, c_m$ coincide with the strongly connected components of $\mathrm{Opt}(B)$. Let A be as in the proposition. Then $A_{k\ell} \leqslant \mathbb{1}$ for all $k, \ell \in [m]$, and all circuits of A have a weight strictly less than $\mathbb{1}$. Indeed, all circuits with length 1 have a weight $\mathbb{0}$ since $A_{kk} = \mathbb{0}$ for all $k \in [m]$. Now assume by contradiction that there exists a circuit with length $p > 1$ and weight $\mathbb{1}$ for A, $(k_1, \dots k_p, k_{p+1} = k_1)$. Since all entries of A are less or equal to $\mathbb{1}$, this implies that $A_{k_q, k_{q+1}} = \mathbb{1}$ for all $q \in [p]$. Moreover, by definition of A, for all $q \in [p]$, there exists $i_q \in c_{k_q}$ and $j_{q+1} \in c_{k_{q+1}}$, such that $B_{i_q, j_{q+1}} = A_{k_q, k_{q+1}} = \mathbb{1}$. Denote $j_1 = j_{p+1}$. Since c_{k_q} is the set of nodes of a strongly connected component of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$, there exists, for all $q \in [p]$, a path from j_q to i_q in c_{k_q} with weights $\mathbb{1}$ for B. Concatenating these paths with the arcs (i_q, j_{q+1}), we obtain a circuit all the weights of witch are equal to $\mathbb{1}$ for B, hence a circuit of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$. This implies that i_q and j_{q+1} belong to the same connected component of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$, a contradiction. This shows that $\alpha = \rho_{\max}(A) < \mathbb{1}$. Let W be as in the proposition. Since all columns Z of $(\alpha^{-1}A)^\star$ are in $\mathbb{R}_{\max}^m$ and satisfy $\alpha^{-1}AZ \leqslant (\alpha^{-1}A)^\star Z = Z$, we deduce that $W \in \mathbb{R}_{\max}^n$ and $AW \leqslant \alpha W$. Moreover since $(\alpha^{-1}A)^\star \geqslant I$, we get that $W = \oplus_{i\in[m]} (\alpha^{-1}A)^\star_{\cdot i} \geqslant \mathbb{1}$, hence $W \in \mathbb{R}^m$. Consider now $U, V \in \mathbb{R}^n$ as in the proposition. We have that $B_{ij}U_j \leqslant \alpha U_i$ for all $i, j \in [n]$ belonging to different strongly connected components of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$, and $B_{ij}U_j \leqslant U_i$ for all $i, j \in [n]$. Then U is a fixed point of B, so that (U, V) is a Hungarian pair with respect to B. Since $\alpha < \mathbb{1}$, $\mathrm{Sat}(B, U, V)$ is exactly the union of the strongly connected subgraphs of $\mathrm{Sat}(B, \mathbb{1}, \mathbb{1})$, hence it is equal to $\mathrm{Opt}(B)$ by Proposition 4.7. □

Corollary 4.11. *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$. Then there exists a Hungarian pair (U, V) with respect to B such that $\mathrm{Sat}(B, U, V) = \mathrm{Opt}(B)$.*

Proof. Let (U, V) be a Hungarian pair with respect to B, C be as in Corollary 4.4, σ be an optimal permutation for B, or equivalently for C and consider $C' = P_{\mathrm{m}}^{\sigma^{-1}} C$. Then, by Corollaries 4.4 and 4.5, and Lemma 4.3, $(\mathbb{1}, \mathbb{1})$ is a Hungarian pair with respect to C', $\operatorname{per} C' = \operatorname{per} C = \mathbb{1}$, and the identity map is an optimal permutation for C'. Then, by Proposition 4.10, there exists a Hungarian pair U', V' with respect to C' such that $\mathrm{Sat}(C', U', V') = \mathrm{Opt}(C')$. By Lemma 4.3, we deduce the same assertion for B. □

Lemma 4.12. *Let $B \in \mathbb{R}_{\max}^{n\times n}$ such that $\operatorname{per} B \neq \mathbb{0}$, the identity map is an optimal permutation for B and $\mathrm{Opt}(B)$ is strongly connected. Then, for any permutation σ, the graph $\mathrm{Opt}(P_{\mathrm{m}}^{\sigma^{-1}} B) = \mathrm{Opt}(B)_{\sigma^{-1},\mathrm{id}}$ is strongly connected. If in addition σ is an optimal permutation for B, then the identity map is an optimal permutation for $P_{\mathrm{m}}^{\sigma^{-1}} B$ and so $\mathrm{Opt}(B)_{\sigma^{-1},\mathrm{id}}$ contains all loops (i, i) with $i \in [n]$.*

Proof. Let $G = \mathrm{Opt}(B)$. Since the identity map is an optimal permutation for B, G contains all loops (i,i) with $i \in [n]$. Let $a = (a_{ij})$ be the matrix with 0 or 1 entries, such that $a_{ij} = 1$ if (i,j) is an arc of G and $a_{ij} = 0$ otherwise. Then G is the graph of a, and G is strongly connected if and only if a is irreducible. Since G contains all loops, $a_{ii} = 1$ for all $i \in [n]$. Then, by [BR91, Thm 4.2.3], a is a fully indecomposable matrix, which means that there does not exist permutations σ and τ of $[n]$, such that $P^\sigma a P^\tau$ is block triangular (with blocks of size $\leqslant n-1$). This implies in particular that $P^{\sigma^{-1}} a$ is irreducible for all permutations σ. Since the graph of $P^{\sigma^{-1}} a$ is equal to $G_{\sigma^{-1},\mathrm{id}}$, we deduce that $G_{\sigma^{-1},\mathrm{id}}$ is strongly connected, which shows the first assertion of the lemma. The last assertion follows from Corollary 4.5. □

4.3 Asymptotics of eigenvectors

Let us add some notations. For any $U \in \mathbb{R}^n$, we denote by $\mathsf{d}_\epsilon(U)$ the diagonal $n \times n$ matrix over $\mathcal{C}$ such that $(\mathsf{d}_\epsilon(U))_{ii} = \epsilon^{-U_i}$ (and $(\mathsf{d}_\epsilon(U))_{ij} = 0$ for $i \neq j$). Then $\mathsf{d}_\epsilon(U)$ is invertible (in $\mathcal{C}^{n\times n}$) and $\mathsf{d}_\epsilon(U)^{-1} = \mathsf{d}_\epsilon(-U)$. Denote by $(\mathbb{C}^n)^*$ the set of n dimensional row vectors. Recall that if $\mathcal{A} \in \mathbb{C}[\mathsf{Y}]^{n\times n}$ is a complex matrix polynomial, we say that a vector $z \in \mathbb{C}^n$ (resp. $w \in (\mathbb{C}^n)^*$) is a right (resp. left) eigenvector of $\mathcal{A}$ with respect to the eigenvalue $\ell \in \mathbb{C}$ of $\mathcal{A}$ if $\mathcal{A}(\ell)z = 0$ (resp. $w\mathcal{A}(\ell) = 0$). We shall adopt an analogue definition for eigenvectors of matrix polynomials $\mathcal{A}$ over the ring $\mathcal{C}$, as in (4.1): eigenvectors are elements of $\mathcal{C}^n$ or $(\mathcal{C}^n)^*$. Finally, for a $n \times n$ matrix B (with coefficients in any ring), and any subsets M, N of $[n]$, we shall denote by B_{MN} the $M \times N$ submatrix of B. Also $B_{(i)(j)}$ will be the $([n]\setminus\{i\}) \times ([n]\setminus\{j\})$ submatrix of B.

Theorem 4.13. *Let $\mathcal{A}$ be a regular matrix polynomial over $\mathcal{C}$ as in (4.1) satisfying (4.2) and denote by A the tropical matrix polynomial (4.3) with coefficients A_k as in (4.2). Let γ be a finite algebraic eigenvalue of A, and let (U,V) be a Hungarian pair with respect to $\widehat{A}(\gamma)$. Consider the scaled matrix polynomial*

$$\mathcal{B}_\epsilon = \mathsf{d}_\epsilon(-U)\mathcal{A}(\epsilon^{-\gamma}\mathsf{Y})\,\mathsf{d}_\epsilon(-V) \ .$$

Then, for $G = \mathrm{Sat}(\widehat{A}(\gamma), U, V)$, $G_k = \mathcal{G}_k(A,\gamma) \cap G$ for all $0 \leqslant k \leqslant d$, and $a^{(\gamma,G)}$ as in (4.5), we have

$$\lim_{\epsilon\to 0} \mathcal{B}_\epsilon = a^{(\gamma,G)} \ .$$

Assume that $a^{(\gamma,G)}$ is a regular matrix polynomial, that λ is a simple nonzero eigenvalue of $a^{(\gamma,G)}$, and let z and w be respectively right and left eigenvectors of $a^{(\gamma,G)}$ for the eigenvalue λ. Let ℓ_ϵ be an eigenvalue of $\mathcal{A}_\epsilon$ with first order asymptotics $\ell_\epsilon \sim \lambda\epsilon^{-\gamma}$, as in Theorem 4.2. Then, for $\epsilon > 0$ small enough, ℓ_ϵ is a simple nonzero eigenvalue of $\mathcal{A}_\epsilon$, $m_\epsilon = \ell_\epsilon\epsilon^\gamma$ is a simple eigenvalue of $\mathcal{B}_\epsilon$ and $\lim_{\epsilon\to 0} m_\epsilon = \lambda$. Moreover, z_ϵ and w_ϵ are respectively right and left eigenvectors of $\mathcal{A}_\epsilon$ for the eigenvalue ℓ_ϵ if and only if $x_\epsilon = \mathsf{d}_\epsilon(V)z_\epsilon$ and $y_\epsilon = w_\epsilon\,\mathsf{d}_\epsilon(U)$ are respectively right and left eigenvectors of $\mathcal{B}_\epsilon$ for the eigenvalue m_ϵ. Also, one can choose z_ϵ and w_ϵ, so that $z \in \mathcal{C}^n$, $w \in (\mathcal{C}^n)^*$, and

$$z = \lim_{\epsilon\to 0} x_\epsilon, \quad w = \lim_{\epsilon\to 0} y_\epsilon \ .$$

In particular if all the entries of z and w are nonzero, then the entries of z_ϵ and w_ϵ have first order asymptotics:

$$(z_\epsilon)_i \sim z_i\epsilon^{V_i}, \quad (w_\epsilon)_i \sim w_i\epsilon^{U_i}, \quad i \in [n] \ .$$

Proof. Let us consider the matrix polynomial:

$$b := a_0^{\mathcal{G}_0(A,\gamma)} + \Upsilon a_1^{\mathcal{G}_1(A,\gamma)} + \cdots + \Upsilon^d a_d^{\mathcal{G}_d(A,\gamma)}.$$

It is easy to see that for all $\gamma \in \mathbb{R}$, all $y \in \mathbb{C}$, and all $i, j \in [n]$, we have

$$(\mathcal{A}_\epsilon(\epsilon^{-\gamma} y))_{ij} = b_{ij}(y)\epsilon^{-(\widehat{A}(\gamma))_{ij}} + o(\epsilon^{-(\widehat{A}(\gamma))_{ij}}) \ .$$

Then

$$(\mathcal{B}_\epsilon(y))_{ij} = b_{ij}(y)\epsilon^{-(\widehat{B}(\gamma))_{ij}} + o(\epsilon^{-(\widehat{B}(\gamma))_{ij}}) \ ,$$

where $\widehat{B}(\gamma) = \mathsf{d}_{\mathrm{m}}(-U)\widehat{A}(\gamma)\,\mathsf{d}_{\mathrm{m}}(-V) = \mathsf{d}_{\mathrm{m}}(U)^{-1}\widehat{A}(\gamma)\,\mathsf{d}_{\mathrm{m}}(V)^{-1}$. By Corollary 4.4, $\widehat{B}(\gamma)$ has its entries less or equal to $\mathbb{1}$ and $\mathrm{Sat}(\widehat{B}(\gamma), \mathbb{1}, \mathbb{1}) = \mathrm{Sat}(\widehat{A}(\gamma), U, V) = G$, which implies that $\lim_{\epsilon\to 0} \mathcal{B}_\epsilon(y) = (b(y))^G = a^{(\gamma,G)}(y)$.

The following assertions of the theorem up to the last one are immediate. Indeed, $m_\epsilon = \ell_\epsilon \epsilon^\gamma$ is an eigenvalue of $\mathcal{B}_\epsilon$, it satisfies $\lim_{\epsilon\to 0} m_\epsilon = \lambda$, and since λ is simple, then for $\epsilon > 0$ small enough, m_ϵ is a simple eigenvalue of $\mathcal{B}_\epsilon$, so is ℓ_ϵ for $\mathcal{A}_\epsilon$. Moreover, z_ϵ and w_ϵ are respectively right and left eigenvectors of $\mathcal{A}_\epsilon$ for the eigenvalue ℓ_ϵ if and only if $\chi_\epsilon = \mathsf{d}_\epsilon(V) z_\epsilon$ and $y_\epsilon = w_\epsilon\, \mathsf{d}_\epsilon(U)$ are respectively right and left eigenvectors of $\mathcal{B}_\epsilon$ for the eigenvalue m_ϵ. Since λ and m_ϵ are simple eigenvalues of $a^{(\gamma,G)}$ and $\mathcal{B}_\epsilon$ respectively, there are unique eigenvectors z and w of $a^{(\gamma,G)}$ for the eigenvalue λ up to a multiplicative constant, and there are unique eigenvectors χ_ϵ and y_ϵ of $\mathcal{B}_\epsilon$ up to a multiplicative constant. Moreover, if we denote by $\mathcal{C}_\epsilon \in \mathcal{C}^{n\times n}$ the adjugate matrix of $\mathcal{B}_\epsilon(m_\epsilon)$ (the transpose of the comatrix), and by $c \in \mathbb{C}^{n\times n}$ the adjugate matrix of $a^{(\gamma,G)}(\lambda)$, we have that $\lim_{\epsilon\to 0} \mathcal{C}_\epsilon = c$ and since λ is a simple eigenvalue of $a^{(\gamma,G)}$, there exists $i, j \in [n]$ such that $c_{ij} \neq 0$, and so for ϵ small enough $(\mathcal{C}_\epsilon)_{ij} \neq 0$. Then z (resp. χ_ϵ) is proportional to the column j of c (resp. $\mathcal{C}_\epsilon$), w (resp. y_ϵ) is proportional to the row i of c (resp. $\mathcal{C}_\epsilon$), and we have $z_i \neq 0$, $w_j \neq 0$, $(\chi_\epsilon)_i \neq 0$, and $(y_\epsilon)_j \neq 0$. Moreover, we can choose the eigenvectors χ_ϵ and y_ϵ such that they are continuous functions of ϵ satisfying $(\chi_\epsilon)_i = z_i$ and $(y_\epsilon)_j = w_j$. Then χ_ϵ and y_ϵ converge towards z and w respectively, and z_ϵ, w_ϵ, χ_ϵ, and y_ϵ have all their entries in $\mathcal{C}$. Then, the last assertion is immediate. □

Theorem 4.14. *With the notations and assumptions of Theorem 4.13, we have that, for generic values of the parameters $(a_k)_{ij}$, the matrix polynomial $a^{(\gamma,G)}$ is regular, it has $m_{\gamma,A}$ nonzero eigenvalues, where $m_{\gamma,A}$ is the multiplicity of the finite eigenvalue γ of A, and all these nonzero eigenvalues are simple.*

Proof. By Theorem 4.2, we have that for generic values of the parameters $(a_k)_{ij}$, the matrix polynomial $a^{(\gamma,G)}$ is regular, that it has $m_{\gamma,A}$ nonzero eigenvalues, where $m_{\gamma,A}$ is the multiplicity of the finite eigenvalue γ of A, and that its valuation is equal to the sum $m'_{\gamma,A}$ of the multiplicities of all the algebraic eigenvalues of A smaller than γ ($-\infty$ included). Then, for generic values of the parameters $(a_k)_{ij}$, $P := \det(a^{(\gamma,G)})/\Upsilon^{m'_{\gamma,A}}$ is a polynomial with valuation zero and degree $m_{\gamma,A}$. Then P has only simple nonzero roots when its discriminant is nonzero. Since the discriminant of a polynomial P is a polynomial expression in its coefficients P_k, $k = 0, \ldots, m_{\gamma,A}$, we get that P has only simple nonzero roots, for generic values of its coefficients P_k. Since the coefficients P_k of P, $k = 0, \ldots, m_{\gamma,A}$, are polynomial functions of the parameters $(a_k)_{ij}$, we get that P has only simple nonzero roots for generic values of the parameters $(a_k)_{ij}$, so has the polynomial $\det(a^{(\gamma,G)})$. □

Remark 4.1. Assume that the asymptotic expansions in (4.2) are true first order asymptotics, meaning that

$$\text{for all } 0 \leqslant k \leqslant d \text{ and } i,j \in [n], \text{ we have either } (a_k)_{ij} \neq 0 \text{ or } (A_k)_{ij} = -\infty \ . \tag{4.6}$$

Then, we can prove that the matrix polynomial $a^{(\gamma,G)}$ in (4.5) is nonzero. Indeed, G is nonempty by definition (it exists since $\operatorname{per}(\widehat{A}(\gamma)) \neq \mathbb{0}$). Let (i,j) be an arc of G, then $\widehat{A}(\gamma)_{ij} = U_i V_j \neq \mathbb{0}$, and there exists some $k = 0,\dots,d$, such that $(A_k)_{ij}\gamma^k = \widehat{A}(\gamma)_{ij}$, hence $(i,j) \in \mathcal{G}_k(A,\gamma)$, and $(A_k)_{ij} \neq -\infty$. By (4.6), this implies that $(a_k)_{ij} = (a_k^{(\gamma,G)})_{ij} \neq 0$, hence the polynomial $a_{ij}^{(\gamma,G)}$ is nonzero, and so is the matrix polynomial $a^{(\gamma,G)}$. However, even if (4.6) holds, some cancellations may occur in the computation of the asymptotics of the characteristic polynomial of $\mathcal{A}$, so that the characteristic polynomial of $a^{(\gamma,G)}$ is zero, hence the matrix polynomial $a^{(\gamma,G)}$ is singular. For instance consider the matrix polynomial over the field $\mathbb{K}$ of Puiseux series with degree 2 and dimension $n = 4$:

$$\mathcal{A} = \begin{bmatrix} \mathsf{Y}^2 & \mathsf{Y} & \mathsf{Y} & \epsilon \\ \mathsf{Y} & 1 & 1 & 0 \\ 0 & 1 & \mathsf{Y} & 0 \\ 0 & 0 & \epsilon & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \epsilon \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \epsilon & 1 \end{bmatrix} + \mathsf{Y} \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \mathsf{Y}^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \ .$$

We have $\det \mathcal{A} = -\epsilon^2 \mathsf{Y}$ so that $\mathcal{A}$ is regular and 0 as its unique finite eigenvalue. The corresponding matrix polynomial A is given by:

$$A = \begin{bmatrix} \mathsf{Y}^2 & \mathsf{Y} & \mathsf{Y} & -1 \\ \mathsf{Y} & 0 & 0 & -\infty \\ -\infty & 0 & \mathsf{Y} & -\infty \\ -\infty & -\infty & -1 & 0 \end{bmatrix} \ .$$

The characteristic polynomial of A is given by $\operatorname{per} A = \mathsf{Y}^3 \oplus \mathsf{Y}^2 \oplus (-2)\mathsf{Y}$, the tropical roots of which are $\mathbb{0}$, $\mathbb{1} = 0$ and -2, with multiplicities 1. Considering the finite tropical eigenvalue $\gamma = 0$ of A, we see that $(\mathbb{1},\mathbb{1})$ is Hungarian pair of

$$\widehat{A}(0) = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -\infty \\ -\infty & 0 & 0 & -\infty \\ -\infty & -\infty & -1 & 0 \end{bmatrix} \ .$$

Taking $G = \operatorname{Sat}(\widehat{A}(0),\mathbb{1},\mathbb{1})$, we obtain a matrix polynomial $a^{(\gamma,G)}$ equal to $\mathcal{A}_\epsilon$ for $\epsilon = 0$. Hence $\det(a^{(\gamma,G)}) = 0$, and the matrix polynomial $a^{(\gamma,G)}$ is singular.

Theorem 4.13 shows that, if the right and left eigenvectors z and w of $a^{(\gamma,G)}$ have only nonzero entries, then the eigenvectors of $\mathcal{A}_\epsilon$ have first order asymptotics. In the following result, we give sufficient conditions under which this holds generically.

Theorem 4.15. *Let $a = a_0 + \mathsf{Y}a_1 + \cdots + \mathsf{Y}^d a_d \in \mathbb{C}[\mathsf{Y}]^{n\times n}$ be a complex matrix polynomial. Let $G_0,\dots,G_d$ be some graphs with set of nodes $[n]$, and consider the matrix polynomial $b = a_0^{G_0} + \mathsf{Y}a_1^{G_1} + \cdots + \mathsf{Y}^d a_d^{G_d}$. Assume that $G := \cup_{k=0}^d G_k$ contains all the loops (i,i) with $i \in [n]$ (or equivalently that the identity map is a permutation of G) and that G is strongly connected. Then, for generic values of the parameters $(a_k)_{ij}$, all the nonzero eigenvalues of b (if such eigenvalues exist) are simple, and the entries of any right and left eigenvectors z and w of b associated to a nonzero simple eigenvalue are nonzero.*

Proof. Note that if all the coefficients $(a_k)_{ij}$ are nonzero (which holds generically), then G is the graph associated to b: (i,j) is an arc of G if and only if the polynomial b_{ij} is nonzero.

Let us consider the tropical matrix polynomial $A = A_0 \oplus \cdots \cdots \mathsf{Y}^d A_d$, which coefficients are given by: $(A_k)_{ij} = \mathbb{1}$ if $(i,j) \in G_k$ and $(A_k)_{ij} = \mathbb{0}$ otherwise. Then A corresponds to the tropical matrix polynomial (4.3) defined for the matrix polynomial $\mathcal{A}$ over $\mathcal{C}$ as in (4.1) satisfying (4.2) and such that $\mathcal{A}_\epsilon = b$ for all $\epsilon > 0$. We have $(\widehat{A}(\mathbb{1}))_{ij} = \mathbb{1}$ if and only if $(i,j) \in G$, and since G contains all loops, we get that $\operatorname{per} \widehat{A}(\mathbb{1}) = \mathbb{1}$, so that A is regular. Moreover, $(\mathbb{1}, \mathbb{1})$ is a Hungarian pair for $\widehat{A}(\mathbb{1})$, $G = \operatorname{Sat}(\widehat{A}(\mathbb{1}), \mathbb{1}, \mathbb{1})$, and since G is strongly connected and contains the identity permutation, we have also $G = \operatorname{Opt}(\widehat{A}(\mathbb{1}))$. Since all the entries of the A_k are $\mathbb{0}$ or $\mathbb{1}$, we get that the characteristic polynomial P_A of A is of the form (either formally or not) $P_A = \oplus_{k \in K} \mathsf{Y}^k$ for some subset K of $\{0, \dots, nd\}$, with $\min K = \operatorname{val} P_A$ and $\max K = \deg P_A$. So, either P_A has no nonzero roots (if K is reduced to a point), or $\mathbb{1}$ is the only nonzero root of P_A and its multiplicity is equal to $m := \deg P_A - \operatorname{val} P_A$. In the first case, b has no nonzero eigenvalue (since P_b is also equal to a single monomial), and there is nothing to prove. In the second case, b has at most m nonzero eigenvalues, and for generic values of the parameters $(a_k)_{ij}$, b has exactly m nonzero eigenvalues and they are all simple (use for instance Theorem 4.13 with $\mathcal{A}_\epsilon = b$). Also, the valuation of P_b is greater or equal to $\operatorname{val} P_A$, and for generic values of the parameters $(a_k)_{ij}$, it is equal to $\operatorname{val} P_A$.

Let λ be a nonzero simple eigenvalue of b. Then the right and left eigenvectors z and w of b are unique up to a multiplicative constant, and they are respectively proportional to the nonzero columns and rows of the adjugate of $b(\lambda)$, so that their entries are proportional to minors of size $n-1$ of $b(\lambda)$. Then all the entries of z and w are nonzero if and only if all the minors of size $n-1$ of $b(\lambda)$ are nonzero, that is, if $\det(b_{(s)(t)}(\lambda)) \neq 0$ for all $s, t \in [n]$.

Let $s, t \in [n]$ and $Q = P_{b_{(s)(t)}}$. Then Q is a (complex) polynomial, the coefficients of which are polynomials in the parameters $(a_k)_{ij}$. If $Q(\lambda) = 0$, then λ is a common nonzero root of Q and P_b. We know that $\operatorname{val} P_b \geqslant v_1 := \operatorname{val} P_B$ and $\deg P_b \leqslant d_1 := \deg P_B$. We also have that $\operatorname{val} Q \geqslant v_2 := \operatorname{val} P_{B_{(s)(t)}}$ and $\deg Q \leqslant d_2 := \deg P_{B_{(s)(t)}}$. Then λ is a common root of P_b/Y^{v_1} and Q/Y^{v_2}. Given two complex polynomials Q_1 and Q_2, satisfying the following bounds on the valuation and degree, $v_i \leqslant \operatorname{val} Q_i$ and $\deg Q_i \leqslant d_i$, $i = 1, 2$, we shall denote by $\mathrm{R}(Q_1, v_1, d_1, Q_2, v_2, d_2)$ the resultant of the polynomials Q_1/Y^{v_1} and Q_2/Y^{v_2} with respective degrees $d_1 - v_1$ and $d_2 - v_2$. Since P_b and Q are polynomials, the coefficients of which are polynomials in the parameters $(a_k)_{ij}$, their resultant $\mathrm{R}(P_b, v_1, d_1, Q, v_2, d_2) = \mathrm{R}(P_b, \operatorname{val} P_B, \deg P_B, P_{b_{(s)(t)}}, \operatorname{val} P_{B_{(s)(t)}}, \deg P_{B_{(s)(t)}})$ is a polynomial expression in the parameters $(a_k)_{ij}$, that we shall denote by $R_{b,s,t}(a)$. From the above remarks, $Q(\lambda) = 0$ implies that $R_{b,s,t}(a) = 0$. Therefore, the condition $R_{b,s,t}(a) \neq 0$ for all $s, t \in [n]$ would imply that all the entries of the right and left eigenvectors z and w of b are nonzero. If all the polynomials $R_{b,s,t}$, with $s, t \in [n]$, are nonzero, then the former condition would define a generic condition, and so, we would obtain that for generic values of the parameters $(a_k)_{ij}$, the entries of any right and left eigenvectors z and w of b associated to a nonzero simple eigenvalue are nonzero, hence the conclusion of the theorem.

To show that the polynomial $R_{b,s,t}$, with $s, t \in [n]$, is nonzero, we need to prove that there exists some parameters $(a_k)_{ij}$ such that $R_{b,s,t}(a) \neq 0$. Moreover, $R_{b,s,t}(a) \neq 0$ in particular

when

$$\operatorname{val} P_{b_{(s)(t)}} = \operatorname{val} P_{B_{(s)(t)}}, \tag{4.7a}$$

$$\deg P_{b_{(s)(t)}} = \deg P_{B_{(s)(t)}}, \quad \text{and} \tag{4.7b}$$

$$P_{b_{(s)(t)}} \text{ and } P_b \text{ have no common nonzero roots.} \tag{4.7c}$$

In the sequel, we fix the degree d, and we shall show, by induction on the size n, that for all $n \times n$ matrix polynomials b of degree d as in the theorem (so all graphs $G_0, \dots, G_d$ satisfying the conditions of the theorem), and for all $s, t \in [n]$, there exist some parameters $(a_k)_{ij}$ (or a $n \times n$ matrix polynomials a of degree d), such that (4.7) holds. This property holds trivially for $n = 1$, with the convention that an empty matrix has a determinant equal to 1, since then $P_{b_{(s)(t)}}$ has no root.

To show the induction, we now assume that the latter property holds for all sizes $\leqslant n-1$. Let us show under this assumption that the latter property is true for size n. For this, we consider a $n \times n$ matrix polynomials b of degree d as in the theorem, and indices $s, t \in [n]$, and we shall exhibit parameters $(a_k)_{ij}$ satisfying (4.7). Applying the same permutation τ on rows and columns of b and a changes the graph G into $G_{\tau,\tau^{-1}}$ which has the same properties as G. Exchanging t and n, we are reduced to prove the above result for $t = n$.

Assume first that either $s = n$ or $(s, t = n)$ is an arc of G. In the second case, there exists a permutation of G such that $\sigma(s) = t$. Applying the permutation σ^{-1} on rows of b and a changes the graph G into the graph $G_{\sigma^{-1},\mathrm{id}}$ which satisfies the same properties as G, by Lemma 4.12. Since the properties (4.7) for b and (s, t) are equivalent to the same properties for the new b and (n, n), we are reduced to prove the above result for $t = s = n$. Let G' be the restriction of G to $[n-1] \times [n-1]$. The graph G' contains all loops, but is not in general strongly connected. Let $\mathcal{S}$ denote the set of strongly connected components of G', $\mathcal{I}$ denote the set of initial strongly connected components of G', that is the set of $c \in \mathcal{S}$ such that there is no arc from $c' \in \mathcal{S}$, $c' \neq c$ (or from $[n-1] \setminus c$) to c, and $\mathcal{F}$ denote the set of the final ones, that is the set of $c \in \mathcal{S}$ such that there is no arc from c to $c' \in \mathcal{S}$, $c' \neq c$. We have:

$$P_{b_{(n)(n)}}(\mathsf{Y}) = \prod_{c \in \mathcal{S}} P_{b_{cc}}(\mathsf{Y}) \ , \tag{4.8}$$

so that the set of roots of $P_{b_{(n)(n)}}$ is the union of the sets of roots of the polynomials $P_{b_{cc}}$.

For any $c \in \mathcal{I}$, there exists an arc from n to some node j_c of c. Indeed, since G is strongly connected, for any $j \in \mathcal{I}$, there is a path from n to j, and since c is an initial strongly connected components of G', this path does not pass through $[n-1] \setminus c$, hence the first arc of this path is of the form (n, j_c) with $j_c \in c$. Moreover, since $G = \cup_{k=0}^{d} G_k$, there exists $k_c \in \{0, \dots, d\}$ such that $(n, j_c) \in G_{k_c}$. Similarly, for any $c \in \mathcal{F}$, there exists an arc from some node i_c of c to n and $k_c \in \{0, \dots, d\}$ such that $(i_c, n) \in G_{k_c}$. Let us fix some of the parameters $(a_k)_{ij}$, by taking $(a_k)_{ij} = 0$ for $k = 0, \dots, d$, and either $i = n$ and $j \notin \{j_c \mid c \in \mathcal{I}\}$, or $j = n$ and $i \notin \{i_c \mid c \in \mathcal{F}\}$, for $i = n$, $j = j_c$ and $k \neq k_c$ with $c \in \mathcal{I}$, and for $i = i_c$, $j = n$ and $k \neq k_c$ with $c \in \mathcal{F}$. Then

$$P_b(\mathsf{Y}) = \sum_{c \in \mathcal{I},\, c' \in \mathcal{F}} (a_{k_c})_{nj_c} (a_{k_{c'}})_{i_{c'}n} \mathsf{Y}^{k_c + k_{c'}} (-1)^{j_c + i_c - 1} \det((b_{(n)(j_c)})_{(i_{c'})(n)}(\mathsf{Y})), \tag{4.9}$$

and $\det((b_{(n)(j_c)})_{(i_{c'})(n)}(\mathsf{Y})) = \det((b_{(n)(n)})_{(i_{c'})(j_c)}(\mathsf{Y}))$.

The latter expression is the sum of the (signed) weights of the bijections from $[n-1]\setminus\{i_{c'}\}$ to $[n-1]\setminus\{j_c\}$. So it is nonzero if and only if there exists a permutation σ in $G'\cup\{(i_{c'},j_c)\}$ such that $\sigma(i_{c'})=j_c$. Since G' contains all loops, this is equivalent to the condition that there is a path in G' from j_c to $i_{c'}$. For any $c\in\mathcal{I}$ and $c'\in\mathcal{F}$, let us denote by $S_{c,c'}$ the set of nodes $i\in[n-1]$ such that there is a path from c to i and a path from i to c' in G', an let $S'_{c,c'}=[n-1]\setminus S_{c,c'}$. If $S_{c.c'}$ is empty, we get that there is no path from c to c', and so $\det((b_{(n)(n)})_{(i_{c'})(j_c)}(\mathsf{Y}))=0$.

Let c,c' be such that $S_{c.c'}$ is non empty. One can show that

$$\begin{aligned}\det((b_{(n)(j_c)})_{(i_{c'})(n)}(\mathsf{Y})) &= \det((b_{(n)(n)})_{(i_{c'})(j_c)}(\mathsf{Y}))\\ &= \det(b_{S'_{c,c'},S'_{c,c'}}(\mathsf{Y}))\det((b_{S_{c,c'},S'_{c,c'}})_{(i_{c'})(j_c)}(\mathsf{Y}))\\ &= \prod_{c''\in\mathcal{S},\, c''\subset S'_{c,c'}} P_{b_{c''c''}}(\mathsf{Y}))\det((b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}(\mathsf{Y}))\ . \end{aligned}\tag{4.10}$$

Using (4.8), we get that $P_{b_{(n)(n)}}(\mathsf{Y})=\prod_{c''\in\mathcal{S},\, c''\subset S'_{c,c'}}\det(b_{c''c''}(\mathsf{Y}))\det(b_{S_{c,c'},S_{c,c'}}(\mathsf{Y}))$. Then $\lambda\neq 0$ is a common root of $P_{b_{(n)(n)}}$ and $\det((b_{(n)(j_c)})_{(i_{c'})(n)}(\mathsf{Y}))$ if and only if either it is a root of some $P_{b_{c''c''}}$ with $c''\in\mathcal{S}$, $c''\subset S'_{c,c'}$, or it is a common root of $\det(b_{S_{c,c'},S_{c,c'}}(\mathsf{Y}))$ and $\det((b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}(\mathsf{Y}))$.

Since c is initial and c' is final, there is no arc $(i_{c'},j_c)$ in G'. We shall denote by $\tilde{b}_{cc'}$ the matrix polynomial $b_{S_{c,c'},S_{c,c'}}$ in which we add a parameter α (a polynomial of degree 0) in position $(i_{c'},j_c)$. Then $(b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}=(\tilde{b}_{cc'})_{(i_{c'})(j_c)}$ and $\det(\tilde{b}_{cc'}(\mathsf{Y}))=\det(b_{S_{c,c'},S_{c,c'}}(\mathsf{Y}))\pm\alpha\det((\tilde{b}_{cc'})_{(i_{c'})(j_c)}(\mathsf{Y}))$. So when $\alpha\neq 0$, λ is common eigenvalue of $b_{S_{c,c'},S_{c,c'}}$ and $(b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}$ if and only if λ is common eigenvalue of $\tilde{b}_{cc'}$ and $(\tilde{b}_{cc'})_{(i_{c'})(j_c)}$. Let $G_k^{cc'}$ be the restriction of the graph G_k to the set of nodes $S_{c,c'}$, and let us add the arc $(i_{c'},j_c)$ to $G_0^{cc'}$. Define also $G^{cc'}=\cup_{k=0}^d G_k^{cc'}$. From the definition of $S_{c,c'}$, $G^{cc'}$ is strongly connected. Moreover, since G' contains all loops, so does $G^{cc'}$. Denoting by $\tilde{a}$ the matrix polynomial $a_{S_{c,c'},S_{c,c'}}$, and denoting $\alpha=(\tilde{a}_0)_{i_{c'}j_c}$, we get that the matrix $\tilde{b}_{cc'}$ satisfies the conditions of the theorem with the graphs $G_k^{cc'}$ and the matrix $\tilde{a}$, and has a size $\leqslant n-1$. Applying the induction assumption, there exists some parameters some parameters $(\tilde{a}_k)_{ij}$ such that (4.7) holds for $\tilde{b}_{cc'}$, $s=i_{c'}$, and $t=j_c$. Therefore the polynomial $R_{\tilde{b}_{cc'},i_{c'},j_c}(\tilde{a})\neq 0$. Since $\tilde{a}$ depends only on the restriction of a to $[n-1]\times[n-1]$, and since the constraint $\alpha\neq 0$ does not change the genericity, we get that for generic values of the parameters $(a_k)_{ij}$, $i,j\in[n-1]$, there is no common eigenvalue of $b_{S_{c,c'},S_{c,c'}}$ and $(b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}$.

We know that for generic values of the parameters $(a_k)_{ij}$, $i,j\in[n-1]$, the properties (4.7a) and (4.7b) hold for $s=t=n$, and the eigenvalues of $b_{(n)(n)}$ are all simple. So applying the latter property to all $c\in\mathcal{I}$ and $c'\in\mathcal{F}$ such that $S_{c.c'}$ is non empty, we can choose the parameters $(a_k)_{ij}$, $i,j\in[n-1]$ so that (4.7a) and (4.7b) hold for $s=t=n$, the nonzero eigenvalues $\lambda_1,\dots,\lambda_m$ (with $m=\deg P_{b_{(n)(n)}}=\operatorname{val}P_{b_{(n)(n)}}$) of $b_{(n)(n)}$ are all simple, and there is no common eigenvalue of $b_{S_{c,c'},S_{c,c'}}$ and $(b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}$ for all $c\in\mathcal{I}$ and $c'\in\mathcal{F}$ such that $S_{c.c'}$ is non empty. Let these parameters be fixed and $\lambda=\lambda_k$ for some $k\in[m]$. From (4.9) and (4.10), we have

$$P_b(\lambda)=\sum_{c\in\mathcal{I},\, c'\in\mathcal{F}, S_{c.c'}\neq\emptyset}(a_{k_c})_{nj_c}(a_{k_{c'}})_{i_{c'}n}\xi_{cc'}\ ,$$

with

$$\xi_{cc'} = \lambda^{k_c+k_{c'}}(-1)^{j_c+i_c-1} \prod_{c''\in\mathcal{S},\, c''\subset S'_{c,c'}} P_{b_{c''c''}}(\lambda)\det((b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}(\lambda)) \ .$$

Since λ is a simple root of $b_{(n)(n)}$, there exists one and only one $c'' \in \mathcal{S}$ such that $P_{b_{c''c''}}(\lambda) = 0$. Moreover, by definition of initial and final strongly connected components, there exists $c \in \mathcal{I}$ and $c' \in \mathcal{F}$ such that $c'' \subset S_{c.c'}$. In that case, $S_{c.c'}$ is nonempty, and λ is an eigenvalue of $b_{S_{c,c'},S_{c,c'}}$ and since there is no common eigenvalue of $b_{S_{c,c'},S_{c,c'}}$ and $(b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}$, we get that $\det((b_{S_{c,c'},S_{c,c'}})_{(i_{c'})(j_c)}(\lambda))$. Moreover, $c'' \notin S'_{c,c'}$, so $\prod_{c''\in\mathcal{S},\, c''\subset S'_{c,c'}} P_{b_{c''c''}}(\lambda)) \neq 0$. Since also $\lambda \neq 0$, we get that $\xi_{cc'} \neq 0$. Then, for each λ, $P_b(\lambda)$ is a nonzero polynomial expression in the parameters $(a_k)_{nj}$ and $(a_k)_{in}$, with $i, j \in [n]$. We thus can choose the parameters $(a_k)_{nj}$ and $(a_k)_{in}$, with $i, j \in [n]$, in such a way that all these polynomial expressions are nonzero. This implies that the polynomials $P_{b_{(n)(n)}}$ and P_b have no common nonzero roots, that is (4.7c). This shows that for these parameters (4.7) holds for $s = t = n$, and proves the induction in the case where $(s, t = n)$ is an arc of G.

Assume now that $(s, t = n)$ is not an arc of G. Since G is strongly connected, there exists a path $i_0 = n, \dots, i_k = s$ in G from n to s. Let σ be the cyclic permutation $(i_0, \dots, i_k, i_0)$. Although, this is not a permutation of G, the graph $G_{\sigma^{-1},\mathrm{id}}$ is strongly connected, by Lemma 4.12. Moreover, since G contains all loops and the arcs of the path $i_0 = n, \dots, i_k = s$, we get that $G_{\sigma^{-1},\mathrm{id}}$ contains the loops (i, i) with $i \in [n-1]$. Defining G' as before, that is G' is the restriction of G to $[n-1] \times [n-1]$, we get that G' contains all loops. Since the above proof uses only the properties that G' contains all loops, and that G is strongly connected, it applies also in the present case, and so the induction is proved in all cases. □

Using Theorems 4.13 and 4.15, we deduce:

Corollary 4.16. *Let $\mathcal{A}, A, a, \gamma, (U, V), G$ be as in Theorem 4.13 and denote $b := a^{(\gamma,G)}$. Assume that the identity map is an optimal permutation of $\widehat{A}(\gamma)$, and that G is strongly connected. Let λ be a nonzero simple eigenvalue of b, let ℓ be an eigenvalue of $\mathcal{A}$ with first order asymptotics $\ell_\epsilon \sim \lambda\epsilon^{-\gamma}$, as in Theorem 4.2, and let z and w be associated right and left eigenvectors of $\mathcal{A}$ as in Theorem 4.13. Then, for generic values of the parameters $(a_k)_{ij}$, the entries of z and w have first order asymptotics:*

$$(z_\epsilon)_i \sim z_i\epsilon^{V_i}, \quad (y_\epsilon)_i \sim w_i\epsilon^{U_i}, \quad i \in [n] \ ,$$

where z and w are right and left eigenvectors of b with nonzero entries.

Corollary 4.17. *Let $\mathcal{A}, A, a, \gamma, (U, V), G$ be as in Theorem 4.13 and denote $b := a^{(\gamma,G)}$. Assume that the identity map is an optimal permutation of $\widehat{A}(\gamma)$. Let c be a strongly connected component of G such that γ is an eigenvalue of the restricted matrix polynomial A_{cc} with multiplicity m_c. Then, for generic values of the parameters $(a_k)_{ij}$, the matrix polynomial b_{cc} has exactly m_c nonzero simple eigenvalues. Any such an eigenvalue λ of b_{cc} is necessarily an eigenvalue of b. Moreover, for generic values of the parameters $(a_k)_{ij}$, any right and left eigenvectors z and w of b for the eigenvalue λ are such that their restrictions to c, z_c and w_c, are right and left eigenvectors of b_{cc} with nonzero entries.*

Proof. Let $\mathcal{A}_\epsilon$, A, γ, (U, V), G, b, c and m_c be as in the corollary. Let also m_γ be the multiplicity of γ for the matrix polynomial A. From Theorems 4.2 and 4.13, for generic values of the

parameters $(a_k)_{ij}$, the matrix polynomial b is regular, it has exactly m_γ nonzero eigenvalues $\lambda_1, \dots, \lambda_{m_\gamma}$, and all these nonzero eigenvalues are simple. Also, the matrix polynomial $\mathcal{A}_\epsilon$ has m_γ simple nonzero eigenvalues $\ell_{\epsilon,1}, \dots, \ell_{\epsilon,m_\gamma}$ with first order asymptotics $\ell_{\epsilon,i} \sim \lambda_i \epsilon^{-\gamma}$. Moreover, applying also Theorem 4.2 to the restricted matrix polynomial $(\mathcal{A}_\epsilon)_{cc}$, we obtain that, for generic values of the parameters $(a_k)_{ij}$, b_{cc} has exactly m_c nonzero eigenvalues (counted with multiplicities), and that 0 is an eigenvalue of b_{cc} if and only if $m'_c > 0$ and then its multiplicity is m'_c, where m'_c is the sum of the multiplicities of the tropical eigenvalues of A_{cc} smaller than γ ($-\infty$ included), putting $m'_c = 0$ if no such eigenvalues exist. The nonzero eigenvalues of b_{cc} belong necessarily to the set of eigenvalues of b, and then are simple.

When λ is a nonzero simple eigenvalue of b_{cc} and also a simple eigenvalue of b as above, and z and w are right and left eigenvectors b for b, we have, that their restriction to c, z_c and w_c are right and left eigenvectors of b_{cc}. Indeed, we can find a permutation τ such that applying the same permutation τ on rows and columns of the matrix polynomials $\mathcal{A}$, A and a, we obtain that the nodes of the strongly connected components $c_1, \dots, c_q$ of G are ordered from initial to final components: if (i, j) is an arc of G, $i \in c_k$ and $j \in c_\ell$ then $k \leqslant \ell$. Then b is block upper triangular with blocks equal to $b_{c_1c_1}, \dots, b_{c_qc_q}$, and there exists $1 \leqslant k \leqslant q$ such that $c = c_k$. Since λ is a simple eigenvalue of b, it is the eigenvalue of the $c \times c$ block only. Therefore, $z_{c_\ell} = 0$ for all $\ell > k$ and z_c is a right eigenvector of b_{cc} for the eigenvalue λ. Similarly, $w_{c_\ell} = 0$ for all $\ell < k$ and w_c is a left eigenvector of b_{cc} for the eigenvalue λ.

Since z_c and w_c only depend on the restriction of the matrix polynomials a and A to $c \times c$, and c is a strongly connected component of G, we can reduce the proof of the last assertion to the case where $c = [n]$ and G is strongly connected, containing the identity permutation, hence the loops. The matrix polynomial b depends only on the parameters $(a_k)_{ij}$ and on the graphs $G_k = \mathcal{G}_k(A, \gamma) \cap G$, $0 \leqslant k \leqslant d$, since $b = a_0^{G_0} + \mathsf{Y} a_1^{G_1} + \cdots + \mathsf{Y}^d a_d^{G_d}$. Moreover $\cup_{k=0}^d G_k = G$, so that a, b and the graphs G_k satisfy the assumptions of Theorem 4.15, and the result of the corollary follows from this theorem. □

Remark 4.2. Note that the same type of result holds without the assumption that the identity map is an optimal permutation of $\widehat{A}(\gamma)$. For this, consider a strongly connected component c of $G_{\sigma^{-1},\mathrm{id}}$, the restriction of matrix polynomials to the arcs in $\sigma^{-1}(c) \times c$ instead of $c \times c$, and $w_{\sigma^{-1}(c)}$ instead of w_c.

4.4 Asymptotics of eigenvalue condition numbers

We recall that the condition number of a complex matrix polynomial $\mathcal{A} = \sum_{k=0}^d \mathcal{A}_k \mathsf{Y}^k \in \mathbb{C}[\mathsf{Y}]^{n\times n}$ for the simple nonzero eigenvalue ℓ, with respect to relative errors is equal to [Tis00, Theorem 5]:

$$\kappa(\mathcal{A}, \ell) = \frac{\|w\|_2 \|z\|_2 (\sum_{k=0}^d \|\mathcal{A}_k\| |\ell|^k)}{|\ell| |w\mathcal{A}'(\ell)z|} \quad , \tag{4.11}$$

where $z \in \mathbb{C}^n$ and $w \in (\mathbb{C}^n)^*$ are respectively right and left eigenvectors of $\mathcal{A}$, and $\mathcal{A}'$ denotes the derivative of the matrix polynomial with respect to Y. Here $\|\cdot\|_2$ is the Euclidean norm, and for any matrix $A \in \mathbb{C}^{n\times m}$, we denote by $\|A\|$ the norm relative to the vectorial Euclidean norms on $\mathbb{C}^n$ and $\mathbb{C}^m$: $\|A\| = \max\{\|Av\|_2 \mid v \in \mathbb{C}^m,\ \|v\|_2 = 1\}$. Then, for all $\epsilon \geqslant 0$ small enough, the condition number formula can be applied to the matrix polynomial $\mathcal{A}_\epsilon$ and an eigenvalue ℓ_ϵ.

Theorem 4.18. *Let us use the notations and assumptions of Theorem 4.13. In particular, let $a^{(\gamma,G)}$ be regular, λ be a simple nonzero eigenvalue of $a^{(\gamma,G)}$, and μ_ϵ be a simple eigenvalue of $\mathcal{B}_\epsilon$ such that $\lim_{\epsilon\to 0}\mu_\epsilon = \lambda$. Then, we have*

$$\lim_{\epsilon\to 0}\kappa(\mathcal{B}_\epsilon,\mu_\epsilon) = \kappa(a^{(\gamma,G)},\lambda)\ .$$

Proof. This is immediate using Theorem 4.13. Indeed $\lim_{\epsilon\to 0}\mathcal{B}_\epsilon = a^{(\gamma,G)}$ means that $\mathcal{B}_\epsilon = \sum_{k=0}^d \mathcal{B}_{\epsilon,k}\mathsf{Y}^k$ and that $\lim_{\epsilon\to 0}\mathcal{B}_{\epsilon,k} = a_k^{(\gamma,G)} = a_k^{G_k}$. Then, if $\lambda = \lim_{\epsilon\to 0}\mu_\epsilon$, we have $\lim_{\epsilon\to 0}\mathcal{B}'_\epsilon(\mu_\epsilon) = (a^{(\gamma,G)})'(\lambda)$. Choosing right and left eigenvectors x_ϵ and y_ϵ of $\mathcal{B}_\epsilon$ as in Theorem 4.13, we get that all terms in the expression $\kappa(\mathcal{B}_\epsilon,\mu_\epsilon)$ have a limit which correspond to the same term in the expression of $\kappa(a^{(\gamma,G)},\lambda)$ defined using the limits z,w of x_ϵ and y_ϵ. This shows the convergence of the eigenvalue condition number. □

For any matrix $A\in\mathbb{C}^{n\times n}$ and any subsets I,J of $[n]$, we denote by A_{IJ} the $I\times J$ submatrix of A, $A_{IJ} = (A_{ij})_{i\in I,j\in J}$. Similarly, for any vector $v\in\mathbb{C}^n$, v_I denotes the vector $v_I = (v_i)_{i\in I}\in\mathbb{C}^I$.

Theorem 4.19. *Let us use the notations and assumptions of Theorem 4.13. In particular let $G = \mathrm{Sat}(\widehat{A}(\gamma),U,V)$ for some Hungarian pair (U,V) with respect to $\widehat{A}(\gamma)$, let $a^{(\gamma,G)}$ be regular, and let λ be a simple nonzero eigenvalue of $a^{(\gamma,G)}$. Then*

$$\kappa(a^{(\gamma,G)},\lambda) \geqslant \kappa(a^{(\gamma,\mathrm{Opt}(\widehat{A}(\gamma)))},\lambda)\ ,$$

and the inequality is an equality for any Hungarian pair satisfying $\mathrm{Sat}(\widehat{A}(\gamma),U,V) = \mathrm{Opt}(\widehat{A}(\gamma))$ as in Corollary 4.11.

Proof. Let $\sigma\in S_n$ be an optimal permutation of $\widehat{A}(\gamma)$. Then, by Corollary 4.5, the identity is an optimal permutation of $P_{\mathrm{m}}^{\sigma^{-1}}\widehat{A}(\gamma)$. By multiplying the matrix polynomial $\mathcal{A}_\epsilon$ by $P^{\sigma^{-1}}$ on the left, we do the same on $a^{(\gamma,G)}$ and we multiply the tropical matrix polynomial A by $P_{\mathrm{m}}^{\sigma^{-1}}$, and left eigenvectors are multiplied on the right by P^σ. Since this does not change the 2-norm, it is sufficient to prove the result in the case where the identity is an optimal permutation of $\widehat{A}(\gamma)$. Then, by Proposition 4.7, $\mathrm{Opt}(\widehat{A}(\gamma))$ is the disjoint union of the strongly connected components of G. Also, as in the proof of Corollary 4.17, we can find another permutation τ such that applying the same permutation τ on rows and columns of the matrix polynomials $\mathcal{A}$, A and a, we obtain that the nodes of the strongly connected components $c_1,\dots,c_m$ of G are ordered from initial to final components: if (i,j) is an arc of G, $i\in c_k$ and $j\in c_\ell$ then $k\leqslant\ell$. Then $a^{(\gamma,G)}$ is block upper triangular with blocks of sizes $|c_1|,\dots,|c_m|$, where $|c|$ is the size of the set c. Moreover, $a^{(\gamma,\mathrm{Opt}(\widehat{A}(\gamma)))}$ corresponds to the block diagonal part of $a^{(\gamma,G)}$.

So it is sufficient to prove that for any matrix polynomials a and b such that a is block upper triangular and b corresponds to the block diagonal part of a, and any simple nonzero eigenvalue λ of a or equivalently of b, we have $\kappa(a,\lambda)\geqslant\kappa(b,\lambda)$, which is an easy consequence of (4.11) formula. Indeed, let us denote by a^{ij} and b^{ij} the blocks of a and b, with $i,j\in[m]$, so that $a^{ij} = 0$ for $i>j$, $b^{ij} = 0$ for $i\neq j$ and $a^{ii} = b^{ii}$. Since λ is simple, there exists a unique $i\in[m]$ such that λ is an eigenvalue of the block $a^{ii} = b^{ii}$. Let z and w be right and left eigenvectors of a with respect to λ, and for all $j\in[m]$, let us denote by z^j and w^j the restriction of z and w to the block j. Then necessarily $z^\ell = 0$ for $\ell>i$, and $w^\ell = 0$ for $\ell<i$. Moreover z^i and w^i are right and left eigenvectors of a^{ii}. This implies that $|wa'(\lambda)z| = |w^i(a^{ii})'(\lambda)z^i|$. If now $\tilde{z}$ and $\tilde{w}$ are right and left eigenvectors of b, we get that

$\tilde{z}^\ell = 0$ and $\tilde{w}^\ell = 0$ for $\ell \neq i$, and that $\tilde{z}^i$ and $\tilde{w}^i$ are right and left eigenvectors of $b^{ii} = a^{ii}$. Since λ is a simple eigenvalue of a^{ii}, we have $\tilde{z}^i = z^i$ and $\tilde{w}^i = w^i$, up to a multiplicative constant. This implies that $|wa'(\lambda)z| = |\tilde{w}b'(\lambda)\tilde{z}|$. Since b_k is equal to the block diagonal part of a_k, we get that $\|b_k\| \leqslant \|a_k\|$. Indeed, it is easy to see that $\|b_k\| = \max_{i\in[m]} \|b_k^{ii}\|$, and that $\|a_k\| \geqslant \max_{i,j\in[m]} \|a_k^{ij}\| \geqslant \max_{i\in[m]} \|a_k^{ii}\|$. Moreover, $\tilde{z}^i = z^i$ and $\tilde{z}^\ell = 0$ for $\ell \neq i$, hence $\|\tilde{z}\|_2 \leqslant \|z\|_2$. The same holds for $\tilde{w}$ and w, so that we get that

$$\kappa(a,\lambda) = \frac{\|w\|_2\|z\|_2(\sum_{k=0}^d \|a_k\||\lambda|^k)}{|\lambda||wb'(\lambda)z|} \geqslant \frac{\|\tilde{w}\|_2\|\tilde{z}\|_2(\sum_{k=0}^d \|b_k\||\lambda|^k)}{|\lambda||\tilde{w}b'(\lambda)\tilde{z}|} = \kappa(b,\lambda) \ .$$

□

Theorem 4.20. *Let us use the notation and assumptions of Theorem 4.13. In particular let γ be a finite eigenvalue of A, (U,V) be any Hungarian pair with respect to $\widehat{A}(\gamma)$, $G = \mathrm{Sat}(\widehat{A}(\gamma),U,V)$, assume that $b := a^{(\gamma,G)}$ is regular, and let λ be a simple nonzero eigenvalue of b, z and w be respectively right and left eigenvectors of $a^{(\gamma,G)}$ for the eigenvalue λ, and ℓ_ϵ be an eigenvalue of $\mathcal{A}_\epsilon$ such that $\ell_\epsilon \sim \lambda\epsilon^{-\gamma}$.*

Denote

$$\delta := \min_{i\in[n],w_i\neq 0} U_i + \min_{j\in[n],z_j\neq 0} V_j - \max_{i,j\in[n],k=0,\dots,d,(b_k)_{ij}\neq 0} (U_i+V_j) \ . \tag{4.12}$$

Then $\delta \leqslant 0$ and

$$\liminf_{\epsilon\to 0} (\kappa(\mathcal{A}_\epsilon,\ell_\epsilon)\epsilon^{-\delta}) > 0 \ .$$

Proof. From Theorem 4.13, there exist right and left eigenvectors $z \in \mathcal{C}^n$ and $w \in (\mathcal{C}^n)^*$ of $\mathcal{A}$ for the eigenvalue ℓ, such that $x = \mathsf{d}_\epsilon(V)z$ and $y = w\,\mathsf{d}_\epsilon(U)$ are right and left eigenvectors of $\mathcal{B}$ for the simple eigenvalue $m = \ell\epsilon^\gamma$, and

$$z = \lim_{\epsilon\to 0} x_\epsilon, \quad w = \lim_{\epsilon\to 0} y_\epsilon \ . \tag{4.13}$$

We have:

$$\begin{aligned}
\kappa(\mathcal{A}_\epsilon,\ell_\epsilon) &= \frac{\|w_\epsilon\|_2\|z_\epsilon\|_2(\sum_{k=0}^d \|\mathcal{A}_{k,\epsilon}\||\ell_\epsilon|^k)}{|\ell_\epsilon||w_\epsilon\mathcal{A}'_\epsilon(\ell_\epsilon)z_\epsilon|} \\
&= \frac{\|y_\epsilon\,\mathsf{d}_\epsilon(-U)\|_2\|\,\mathsf{d}_\epsilon(-V)x_\epsilon\|_2(\sum_{k=0}^d \|\,\mathsf{d}_\epsilon(U)\mathcal{B}_{k,\epsilon}\,\mathsf{d}_\epsilon(V)\||m_\epsilon|^k)}{|m_\epsilon||y_\epsilon\mathcal{B}'_\epsilon(m_\epsilon)x_\epsilon|} \\
&= \kappa(\mathcal{B}_\epsilon,m_\epsilon)\frac{K_\epsilon}{Q_\epsilon} \ ,
\end{aligned}$$

with

$$\begin{aligned}
K_\epsilon &:= \|y_\epsilon\,\mathsf{d}_\epsilon(-U)\|_2\|\,\mathsf{d}_\epsilon(-V)x_\epsilon\|_2\left(\sum_{k=0}^d \|\,\mathsf{d}_\epsilon(U)\mathcal{B}_{k,\epsilon}\,\mathsf{d}_\epsilon(V)\||m_\epsilon|^k\right) \ , \\
Q_\epsilon &:= \|y_\epsilon\|_2\|x_\epsilon\|_2(\sum_{k=0}^d \|\mathcal{B}_{k,\epsilon}\||m_\epsilon|^k) \ .
\end{aligned}$$

From Theorem 4.18, $\kappa(\mathcal{B}_\epsilon,m_\epsilon)$ has a limit equal to $\kappa(b,\lambda)$ and from Theorem 4.13 and (4.13), Q_ϵ has a limit equal to $\|w\|_2\|z\|_2(\sum_{k=0}^d \|b_k\||\lambda|^k)$. These two limits are finite and nonzero

since b is regular and λ is a simple nonzero root of b. Let us lower bound K_ϵ. Since $\|z\|_2 \geqslant |z_i|$ for all $z \in \mathbb{C}^n$, and $i \in [n]$, and similarly $\|a\| \geqslant |a_{ij}|$ for all $a \in \mathbb{C}^{n\times n}$ and $i, j \in [n]$, we deduce that for all $i, j, i', j' \in [n]$, and $k = 0, \dots, d$, we have

$$\begin{aligned} K_\epsilon &\geqslant |(y_\epsilon)_i| \epsilon^{U_i} \epsilon^{V_j} |(\chi_\epsilon)_j| \epsilon^{-U_{i'}} |(\mathcal{B}_{k,\epsilon})_{i'j'}| \epsilon^{-V_{j'}} |m_\epsilon|^k \\ &= \epsilon^{U_i + V_j - U_{i'} - V_{j'}} |(y_\epsilon)_i| |(\chi_\epsilon)_j| |(\mathcal{B}_{k,\epsilon})_{i'j'}| |m_\epsilon|^k \ . \end{aligned}$$

Let δ be as in (4.12). Since z and w are different from zero (they are eigenvectors), the minimum in the expression of δ is nonempty so is finite. Moreover, since b is not identically zero, then there exists $k = 0, \dots, d$ such that b_k is non zero, hence there exists $i, j \in [n]$ such that $(b_k)_{ij} \neq 0$. This shows that the maximum in the expression of δ is nonempty so is finite, and δ is finite. Let us show that $\delta \leqslant 0$. Indeed, since λ is a simple eigenvalue of b, we get that $wb'(\lambda)z \neq 0$. Since $wb'(\lambda)z = \sum_{i,j\in[n]} \sum_{k\in[d]} (w_i (b_k)_{ij} k \lambda^{k-1} z_j)$, there exists $i, j \in [n]$ and $k \in [d]$ such that $w_i \neq 0$, $z_j \neq 0$ and $(b_k)_{ij} \neq 0$. Then, from the expression of δ in (4.12), we get that $\delta \leqslant U_i + V_j - (U_i + V_j) = 0$, which gives the result.

Let us now take i, j realizing the minimum in the expression of δ in (4.12) and i', j', k realizing the maximum. We get that

$$K_\epsilon \epsilon^{-\delta} \geqslant |(y_\epsilon)_{ij}| |(\chi_\epsilon)_j| |(\mathcal{B}_{k,\epsilon})_{i'j'}| |m_\epsilon|^k \ ,$$

and passing to the limit, we obtain:

$$\liminf_{\epsilon\to 0} K_\epsilon \epsilon^{-\delta} \geqslant |w_i| |z_j| |(b_k)_{i'j'}| |\lambda|^k \ ,$$

which is different from 0 by the definition of i, j and i', j', k. This implies the assertion of the theorem: $\liminf_{\epsilon\to 0} (\kappa(\mathcal{A}_\epsilon, \ell_\epsilon)\epsilon^{-\delta}) > 0$. □

Corollary 4.21. *Let us use the notations and assumptions of Theorem 4.20. Assume also that (4.6) holds. Then, we have:*

$$\delta = \min_{i\in[n], w_i\neq 0} U_i + \min_{j\in[n], z_j\neq 0} V_j - \max_{(i,j)\in G} (U_i + V_j) \ . \tag{4.14}$$

Moreover, assume that the identity is an optimal permutation of $\widehat{A}(\gamma)$ and that c is the set of nodes of the strongly connected component of G such that λ is an eigenvalue of b_{cc}. Then z_c and w_c are respectively right and left eigenvectors of b_{cc} for the eigenvalue λ. If all coordinates of z_c and w_c are non zero, and if either U_c or V_c is a non constant vector, then $\delta < 0$ and so

$$\liminf_{\epsilon\to 0} \kappa(\mathcal{A}_\epsilon, \ell_\epsilon) = +\infty \ .$$

Proof. Assume that (4.6) holds. By the arguments of Remark 4.1, any arc (i, j) of G is such that b_{ij} is not identically equal to 0, or equivalently that there exists $k = 0, \dots, d$ with $(b_k)_{ij} \neq 0$. Since $b_{ij} = 0$ if (i, j) is not an arc of G, we get that the maximum in the expression of δ in (4.12) is equal to the maximum over all arcs (i, j) of G, hence δ is equal to the expression in (4.14).

Assume in addition that the identity is an optimal permutation of $\widehat{A}(\gamma)$ and that c is the set of nodes of the strongly connected component of G such that λ is an eigenvalue of b_{cc}. Then by Corollary 4.17, z_c and w_c are respectively right and left eigenvectors of b_{cc} for the eigenvalue λ.

Now, restricting the max and min in (4.14) to nodes in c, and denoting by G_c the restriction of G to the nodes in c, we always get the following upper bound for δ:

$$\delta \leqslant \min_{i\in c, w_i\neq 0} U_i + \min_{j\in c, z_j\neq 0} V_j - \max_{(i,j)\in G_c} (U_i + V_j) \ .$$

If z_c and w_c have all their coordinates different from 0, the previous expression reduces to:

$$\delta \leqslant \min_{i\in c} U_i + \min_{j\in c} V_j - \max_{(i,j)\in G_c} (U_i + V_j) \ .$$

Assume also that U_c or V_c is not constant, and by contradiction that $\delta = 0$. Then $U_{i'} + V_{j'} \leqslant U_i + V_j$, for all $(i', j') \in G_c$ and $i, j \in c$. Since the identity is an optimal permutation of $\widehat{A}(\lambda)$, we get that $(i, i) \in G_c$ for all $i \in c$. Hence, Taking $j' = i' = j$, we get that $U_{i'} \leqslant U_i$, for all $i, i' \in c$, hence U_c is a constant vector. Similarly V_c is a constant vector, which leads to a contradiction. This concludes the proof of $\delta < 0$. Then the last assertion follows from Theorem 4.20. □

The following result gives restrictions on the connected component c in Corollary 4.21.

Lemma 4.22. *Let us use the notations and assumptions of Theorem 4.13, and assume that the identity is an optimal permutation of $\widehat{A}(\gamma)$. For all strongly connected components c of G, let m_c be the multiplicity of the eigenvalue γ of the restricted tropical matrix polynomial A_{cc}, where by convention $m_c = 0$ when γ is not an eigenvalue. Then there exists c such that $m_c \neq 0$ and the multiplicity of the eigenvalue γ of A is equal to the sum of the multiplicities m_c, for all the strongly connected components c of G. Moreover, for all strongly connected components c of G, such that b_{cc} has a nonzero eigenvalue λ, we have $m_c > 0$.*

Proof. By continuity, for γ' close to γ, the set of optimal permutations of $\widehat{A}(\gamma')$ is included in the set of optimal permutations of $\widehat{A}(\gamma)$, hence $\mathrm{Opt}(\widehat{A}(\gamma')) \subset \mathrm{Opt}(\widehat{A}(\gamma))$. Since, by Proposition 4.7, $\mathrm{Opt}(\widehat{A}(\gamma))$ is the disjoint union of the strongly connected components of G, we get that for γ' close to γ, $\mathrm{per}(\widehat{A}(\gamma')) = \prod_c \mathrm{per}(\widehat{A_{cc}}(\gamma'))$, where the product is over all strongly connected components c of G. Then γ is an eigenvalue of A if and only if it is an eigenvalue of A_{cc} for some c and then its multiplicity is equal to the sum of the multiplicities m_c of γ for A_{cc}.

If now, $m_c = 0$ for some strongly connected components of G, that is if γ is not an eigenvalue of A_{cc}, then the maximum in the tropical polynomial expression of $\mathrm{per}(\widehat{A_{cc}}(\gamma))$ is attained in an unique monomial. If k is the degree of this monomial, we also get that the matrix polynomial b_{cc} is such that $\det(b_{cc})$ is a monomial of degree k. This implies that b_{cc} has no nonzero eigenvalues, which shows the last assertion of the lemma. □

Corollary 4.23. *Let $\mathcal{A}$ be a regular matrix polynomial over $\mathcal{C}$ as in (4.1) satisfying (4.2) and denote by A the tropical matrix polynomial (4.3) with coefficients A_k as in (4.2). Let γ be a finite algebraic eigenvalue of A, let (U, V) be a Hungarian pair with respect to $\widehat{A}(\gamma)$. and denote $G = \mathrm{Sat}(\widehat{A}(\gamma), U, V)$, and $b = a^{(\gamma,G)}$. Assume that the identity map is an optimal permutation of $\widehat{A}(\gamma)$. Let c be a strongly connected component of G such that γ is an eigenvalue of the restricted matrix polynomial A_{cc} with multiplicity m_c, and either U_c or V_c is a non constant vector. Then, for generic values of the parameters $(a_k)_{ij}$, the matrix polynomial b_{cc} has exactly m_c nonzero simple eigenvalues. Any such an eigenvalue λ of b_{cc} is necessarily*

an eigenvalue of b, and for ϵ small enough, the matrix polynomial $\mathcal{A}_\epsilon$ has a simple nonzero eigenvalue $\ell_\epsilon \sim \lambda\epsilon^{-\gamma}$, and, for generic values of the parameters $(a_k)_{ij}$, it satisfies

$$\liminf_{\epsilon\to 0} \kappa(\mathcal{A}_\epsilon, \ell_\epsilon) = +\infty \ .$$

Proof. The first part of the corollary follows from the first part of Corollary 4.17. The last assertion of Corollary 4.17 shows that, for generic values of the parameters $(a_k)_{ij}$, any right and left eigenvectors z and w of b for a nonzero simple eigenvalue λ of b_{cc} are such that all the entries of z_c and w_c are nonzero. Then, by Corollary 4.21 and the assumption that either U_c or V_c is a non constant vector, we get that $\liminf_{\epsilon\to 0} \kappa(\mathcal{A}_\epsilon, \ell_\epsilon) = +\infty$. □

Remark 4.3. Note that the same type of result holds without the assumption that the identity map is an optimal permutation of $\widehat{A}(\gamma)$. For this, consider a strongly connected component c of $G_{\sigma^{-1},\mathrm{id}}$, the restriction of matrix polynomials to the arcs in $\sigma^{-1}(c) \times c$ instead of $c \times c$, and the restriction of U to $\sigma^{-1}(c)$ instead of c.

Example 4.4. We give now an example where the above genericity condition does not hold and showing that the assumption in Corollary 4.21 that all coordinates of z_c and w_c are non zero is necessary to get the conclusion. Let us consider the following linear matrix polynomial with dimension $n = 3$, depending on the parameter ϵ and also on another parameter $\delta > 0$ which will be chosen close to 0:

$$\mathcal{A} = \begin{bmatrix} (\mathsf{Y}-\delta)\epsilon^2 & \epsilon & \epsilon \\ \epsilon & \mathsf{Y}-\delta & 0 \\ \epsilon & 0 & \mathsf{Y}-\delta \end{bmatrix} = \begin{bmatrix} -\delta\epsilon^2 & \epsilon & \epsilon \\ \epsilon & -\delta & 0 \\ \epsilon & 0 & -\delta \end{bmatrix} + \mathsf{Y}\begin{bmatrix} \epsilon^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \ .$$

We have $P_{\mathcal{A}}(\mathsf{Y}) = \epsilon^2(\mathsf{Y}-\delta)((\mathsf{Y}-\delta)^2 - 2)$, so the eigenvalues of $\mathcal{A}$ are independent of ϵ, they are equal to δ and $\delta \pm \sqrt{2}$ and they are simple. Let us study the condition number of the eigenvalue δ. We have

$$\mathcal{A}(\delta) = \begin{bmatrix} 0 & \epsilon & \epsilon \\ \epsilon & 0 & 0 \\ \epsilon & 0 & 0 \end{bmatrix} \ ,$$

so that the vectors z and w such that $z^T = w = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$ are right and left eigenvectors of $\mathcal{A}$ for the eigenvalue δ. We have

$$\kappa(\mathcal{A}_\epsilon, \delta) = \frac{\|(\mathcal{A}_\epsilon)_0\| + \delta\|(\mathcal{A}_\epsilon)_1\|}{\delta} \ ,$$

and

$$\lim_{\epsilon\to 0} \kappa(\mathcal{A}_\epsilon, \delta) = \frac{\left\|\begin{bmatrix} 0 & 0 & 0 \\ 0 & -\delta & 0 \\ 0 & 0 & -\delta \end{bmatrix}\right\| + \delta \left\|\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right\|}{\delta} = 2 \ .$$

Let us construct the scaling of Theorem 4.13. The tropical matrix polynomial associated to $\mathcal{A}$ is given by:

$$A = \begin{bmatrix} -2 & -1 & -1 \\ -1 & 0 & -\infty \\ -1 & -\infty & 0 \end{bmatrix} \oplus \mathsf{Y}\begin{bmatrix} -2 & -\infty & -\infty \\ -\infty & 0 & -\infty \\ -\infty & -\infty & 0 \end{bmatrix} \ .$$

It has a unique eigenvalue $\mathbb{1} = 0$ with multiplicity 3, and:

$$\widehat{A}(\mathbb{1}) = \begin{bmatrix} -2 & -1 & -1 \\ -1 & 0 & -\infty \\ -1 & -\infty & 0 \end{bmatrix} .$$

A Hungarian pair (U, V) for $\widehat{A}(\mathbb{1})$ is given by $U = V = \begin{bmatrix} -1 & 0 & 0 \end{bmatrix}$, and the graph $G = \mathrm{Sat}(\widehat{A}(\mathbb{1}), U, V)$ of Theorem 4.13 contains the identity permutation and is strongly connected. The scaling of Theorem 4.13 leads to the scaled matrix polynomial:

$$\mathcal{B} = \begin{bmatrix} (\mathsf{Y} - \delta) & 1 & 1 \\ 1 & \mathsf{Y} - \delta & 0 \\ 1 & 0 & \mathsf{Y} - \delta \end{bmatrix} = \begin{bmatrix} -\delta & 1 & 1 \\ 1 & -\delta & 0 \\ 1 & 0 & -\delta \end{bmatrix} + \mathsf{Y} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

This matrix polynomial does not depend on ϵ, and the above vectors z and w are still right and left eigenvectors of $\mathcal{B}$ for the eigenvalue δ. Computing the condition number, we get:

$$\kappa(\mathcal{B}_\epsilon, \delta) = \frac{\|(\mathcal{B}_\epsilon)_0\| + \delta \|(\mathcal{B}_\epsilon)_1\|}{\delta} \geqslant \frac{\sqrt{2}}{\delta} .$$

Then, for $\ell_\epsilon = \delta = m_\epsilon$ and δ small, not only the property $\lim_{\epsilon \to 0} \kappa(\mathcal{A}_\epsilon, \ell_\epsilon) = \infty$ fails, but also $\kappa(\mathcal{B}_\epsilon, m_\epsilon)$ is much larger than $\kappa(\mathcal{A}_\epsilon, \ell_\epsilon)$, which means that in this case the scaling will diminish the efficiency of the computations.

4.5 Some general bounds on eigenvalue condition numbers

The previous results show the behavior of the eigenvalue condition numbers of matrix polynomials depending on the parameter ϵ, when ϵ goes to 0. Here we shall describe the behavior of these condition numbers when ϵ is in the order of 1. In particular, we shall compare the condition numbers before and after a diagonal scaling similar to the one considered in previous section and obtain bounds similar to the ones used in Theorem 4.20. We shall see that these scalings and bounds can also be used in the context of matrix polynomials depending on the parameter ϵ, leading to some improvements of the previous results.

We consider now a matrix polynomial

$$\mathcal{A} = \mathcal{A}_0 + \mathsf{Y}\mathcal{A}_1 + \cdots + \mathsf{Y}^d \mathcal{A}_d , \tag{4.15}$$

where Y is an indeterminate, and for every $0 \leqslant k \leqslant d$, $\mathcal{A}_k \in \mathbb{C}^{n \times n}$. For any scalar, vector, matrix, polynomial or matrix polynomial x with complex coefficients, we shall denote by $\mathsf{v}(x)$ the tropical scalar, vector, matrix, polynomial or matrix polynomial obtained by applying the logarithm of the absolute value entrywise or coordinatewise. In particular $\mathsf{v}(X) = (\log |X_{ij}|)_{i,j=1,\ldots,n} \in \mathbb{R}_{\max}^{n \times n}$ when $X = (X_{ij})_{i,j=1,\ldots,n} \in \mathbb{C}^{n \times n}$, and

$$\mathsf{v}(\mathcal{A}) = \mathsf{v}(\mathcal{A}_0) \oplus \mathsf{Y}\,\mathsf{v}(\mathcal{A}_1) \oplus \cdots \oplus \mathsf{Y}^d\,\mathsf{v}(\mathcal{A}_d) .$$

In the sequel, we shall use the following notations. For any $U \in \mathbb{R}^n$, we denote by $\mathsf{d}(U)$ the diagonal $n \times n$ complex matrix such that $(\mathsf{d}(U))_{ii} = \exp(U_i)$ (and $(\mathsf{d}(U))_{ij} = 0$ for $i \neq j$). Then $\mathsf{d}(U)$ is invertible (in $\mathbb{C}^{n \times n}$) and $\mathsf{d}(U)^{-1} = \mathsf{d}(-U)$. Note that formally, we have

$\mathsf{d}(U) = \mathsf{d}_\epsilon(U)$ for $\epsilon = e^{-1}$, but this time $\mathsf{d}(U)$ is a constant matrix, whereas $\mathsf{d}_\epsilon(U)$ was a matrix valued continuous function of the parameter ϵ.

For any vectors $x \in \mathbb{R}^p$, $p \geqslant 1$, we denote by:

$$\|x\|_\infty = \max(|x_1|, \dots, |x_p|) \ , \tag{4.16a}$$

$$|x|_{\mathrm{H}} = \max(x_1, \dots, x_p) - \min(x_1, \dots, x_p) \ , \tag{4.16b}$$

respectively for the sup-norm and for the Hilbert's seminorm (also called Hopf's oscillation) of x. We shall also use the notations:

$$\mathcal{T}(x) := \max(\max(|x_1|, \dots, |x_p|), \frac{1}{\min(|x_1|, \dots, |x_p|)}) = \exp(\|\,\mathsf{v}(x)\|_\infty) \ , \tag{4.17a}$$

$$\mathcal{H}(x) := \frac{\max(|x_1|, \dots, |x_p|)}{\min(|x_1|, \dots, |x_p|)} = \exp(|\,\mathsf{v}(x)|_{\mathrm{H}}) \ , \tag{4.17b}$$

with the convention that $\mathcal{T}(x) = \mathcal{H}(x) = +\infty$ if some entries of x are zero. The latter scalars are respectively the exponentials of the Thompson and Hilbert distances of the vector $(|x_i|)_{i\in[p]}$ to the vector $\mathbb{1}$ of $\mathbb{R}^p$. Then, if $U \in \mathbb{R}^p$, and e^U denotes the vector of $\mathbb{R}^p$ with i-th entry e^{U_i}, we have $\mathcal{T}(e^U) = \exp(\|U\|_\infty)$ and $\mathcal{H}(e^U) = \exp(|U|_{\mathrm{H}})$. When $x \in \mathbb{R}^p$ has zero entries, and $S \subset [p]$ is the support of x, that is the set of $i \in [p]$ such that $x_i \neq 0$, we shall also use the notation $\mathcal{H}_{\mathrm{su}}(x)$ for $\mathcal{H}(x_S)$. In the sequel, we shall use $\mathcal{T}$ for scalars only: $\mathcal{T}(a) = \max(|a|, 1/|a|)$. The Hilbert's seminorm $|\cdot|_{\mathrm{H}}$ will be used for additive objects, such as the Hungarian pairs, whereas $\mathcal{H}$ and $\mathcal{H}_{\mathrm{su}}$ will be used on complex vectors such as the eigenvectors of a complex matrix polynomial, or the matrix polynomial itself.

In the sequel, we shall consider the following *diagonal scaling* of the matrix polynomial $\mathcal{A}$.

Lemma 4.24. *Let $\mathcal{A}$ be a regular complex matrix polynomial as in (4.15), and let us consider the tropical matrix polynomial $A := \mathsf{v}(\mathcal{A})$. Then A is regular. Let γ be any finite eigenvalue of A, (U, V) be any Hungarian pair with respect to $\widehat{A}(\gamma)$, denote $G = \mathrm{Sat}(\widehat{A}(\gamma), U, V)$, and consider the scaled matrix polynomial*

$$\mathcal{B} = \mathsf{d}(-U)\mathcal{A}(e^\gamma \mathsf{Y})\,\mathsf{d}(-V) \ .$$

Then $|(\mathcal{B}_k)_{ij}| \leqslant 1$ for all $i, j = 1, \dots, n$ and $k = 0, \dots, d$. Moreover, for $0 \leqslant k \leqslant d$, we have

$$|(\mathcal{B}_k)_{ij}| = 1 \ \forall (i,j) \in G_k = \mathcal{G}_k(A, \gamma) \cap G \ .$$

Let $\mathfrak{m}$ be a simple nonzero eigenvalue of $\mathcal{B}$. Then $\mathfrak{x}$ and $\mathfrak{y}$ are respectively right and left eigenvectors of $\mathcal{B}$ for the eigenvalue $\mathfrak{m}$ if and only if $\mathfrak{z} = \mathsf{d}(-V)\mathfrak{x}$ and $\mathfrak{w} = \mathfrak{y}\,\mathsf{d}(-U)$ are respectively right and left eigenvectors of $\mathcal{A}$ for the eigenvalue $\mathfrak{l} = e^\gamma \mathfrak{m}$.

Proof. Since $\widehat{\mathsf{v}(\mathcal{B})}(\mathbb{1}) = \mathsf{d}_{\mathrm{m}}(-U)\widehat{A}(\gamma)\,\mathsf{d}_{\mathrm{m}}(-V)$ over $\mathbb{R}_{\max}$, we get that

$$\max_{k=0,\dots,d}(\log|(\mathcal{B}_k)_{ij}|) = [\widehat{\mathsf{v}(\mathcal{B})}(\mathbb{1})]_{ij} = (\widehat{A}(\gamma))_{ij} - U_i - V_j, \quad \text{for all } i, j = 1, \dots, n \ .$$

Moreover, (U, V) is a Hungarian pair with respect to $\widehat{A}(\gamma)$, hence $(\widehat{A}(\gamma))_{ij} - U_i - V_j \leqslant 0$ for all $i, j = 1, \dots, n$, and $(\widehat{A}(\gamma))_{ij} - U_i - V_j = 0$ for all $(i, j) \in G = \mathrm{Sat}(\widehat{A}(\gamma), U, V)$. This implies that $|(\mathcal{B}_k)_{ij}| \leqslant 1$ for all $i, j = 1, \dots, n$ and $k = 0, \dots, d$, and that $\max_{k=0,\dots,d}(|(\mathcal{B}_k)_{ij}|) = 1$ for $(i, j) \in G$. Since the former maximum is attained for $(i, j) \in \mathcal{G}_k(A, \gamma)$, we deduce that $|(\mathcal{B}_k)_{ij}| = 1$ for $(i, j) \in G_k = \mathcal{G}_k(A, \gamma) \cap G$. The last assertion of the lemma is immediate. □

Recall that the condition number $\kappa(\mathcal{A}, \ell)$ of the matrix polynomial $\mathcal{A}$ of (4.15) for the simple nonzero eigenvalue ℓ, with respect to relative errors satisfies (4.11).

Remark 4.5. The diagonal scaling defined in Lemma 4.24 differs in general from the tropical scaling proposed in [GS09] and further studied in [HMT13, ST14], that we next recall. Let $p := \oplus_{i=0}^{d} p_k \mathsf{Y}^k \in \mathbb{R}_{\max}[\mathsf{Y}]$ be the tropical polynomial with coefficients $p_k = \log \|\mathcal{A}_k\|$, $k = 0, \dots, d$, and let γ be a finite tropical root of p. Then, the tropical scaling relatively to γ transforms the matrix polynomial $\mathcal{A}$ into the matrix polynomial $\mathcal{B}$ such that $\mathcal{B}(\mathsf{Y}) = e^{-\widehat{p}(\gamma)} \mathcal{A}(e^{\gamma} \mathsf{Y})$. Since in general γ is not an eigenvalue of $A = \mathsf{v}(\mathcal{A})$ and the vectors U and V of a Hungarian pair as in Lemma 4.24 are not constant, the tropical scaling differs from the above diagonal scaling. As for the diagonal scaling, ℓ is an eigenvalue of $\mathcal{A}$ if and only if $m = e^{-\gamma} \ell$ is an eigenvalue of $\mathcal{B}$. However, the eigenvectors do not change, so as the condition number of the eigenvalue with respect to the matrix polynomial given in (4.11): $\kappa(\mathcal{A}, \ell) = \kappa(\mathcal{B}, m)$.

Proposition 4.25. *Let us use the notations of Lemma 4.24. Then*

$$\exp(-|U|_{\mathrm{H}}) \exp(-|V|_{\mathrm{H}}) \leqslant \frac{\kappa(\mathcal{A}, \ell)}{\kappa(\mathcal{B}, m)} \leqslant \exp(|U|_{\mathrm{H}}) \exp(|V|_{\mathrm{H}}) \ . \tag{4.18}$$

Proof. Let m, χ, y, ℓ, z, w be as in Lemma 4.24. We have

$$\begin{aligned} \kappa(\mathcal{A}, \ell) &= \frac{\|w\|_2 \|z\|_2 (\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k)}{|\ell| \ |w \mathcal{A}'(\ell) z|} \\ &= \frac{\|y \, \mathsf{d}(-U)\|_2 \| \, \mathsf{d}(-V) \chi\|_2 (\sum_{k=0}^{d} \| \, \mathsf{d}(U) \mathcal{B}_k \, \mathsf{d}(V)\| \ |m|^k)}{|m| \ |y \mathcal{B}'(m) \chi|} \ . \end{aligned}$$

Hence

$$\frac{\kappa(\mathcal{A}, \ell)}{\kappa(\mathcal{B}, m)} = \frac{\|y \, \mathsf{d}(-U)\|_2 \| \, \mathsf{d}(-V) \chi\|_2 (\sum_{k=0}^{d} \| \, \mathsf{d}(U) \mathcal{B}_k \, \mathsf{d}(V)\| \ |m|^k)}{\|y\|_2 \|\chi\|_2 (\sum_{k=0}^{d} \|\mathcal{B}_k\| \ |m|^k)} \ . \tag{4.19}$$

It is easy to see that $\| \, \mathsf{d}(V) \chi\|_2 \leqslant \exp(\max(V_1, \dots, V_n)) \|\chi\|_2$, for all $V \in \mathbb{R}^n$ and $\chi \in \mathbb{C}^n$, from which we deduce that

$$\| \, \mathsf{d}(U) C \, \mathsf{d}(V)\| \leqslant \exp(\max(U_1, \dots, U_n)) \exp(\max(V_1, \dots, V_n)) \|C\|$$

for all $U, V \in \mathbb{R}^n$ and $C \in \mathbb{C}^{n \times n}$. Applying these inequalities to (4.19), we obtain

$$\begin{aligned} \frac{\kappa(\mathcal{A}, \ell)}{\kappa(\mathcal{B}, m)} \leqslant & \exp(\max(-U_1, \dots, -U_n)) \exp(\max(-V_1, \dots, -V_n)) \\ & \exp(\max(U_1, \dots, U_n)) \exp(\max(V_1, \dots, V_n)) \\ = & \exp(|U|_{\mathrm{H}}) \exp(|V|_{\mathrm{H}}) \ . \end{aligned}$$

Symmetrically, we have

$$\begin{aligned} \frac{\kappa(\mathcal{B}, m)}{\kappa(\mathcal{A}, \ell)} &= \frac{\|w \, \mathsf{d}(U)\|_2 \| \, \mathsf{d}(V) z\|_2 (\sum_{k=0}^{d} \| \, \mathsf{d}(-U) \mathcal{A}_k \, \mathsf{d}(-V)\| \ |\ell|^k)}{\|w\|_2 \|z\|_2 (\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k)} \\ &\leqslant \exp(|U|_{\mathrm{H}}) \exp(|V|_{\mathrm{H}}) \ , \end{aligned}$$

which finishes the proof of the proposition. □

The previous bounds do not depend on any estimations of the eigenvalues. The aim of the following lower bounds is to show that for a given eigenvalue and the scaling adapted to this eigenvalue, a lower bound can be obtained similar to the upper bound in Proposition 4.25. These lower bounds depend in particular on the distance between $|m|$ and 1, so on the distance between $|\ell|$ and e^γ. Thus, they have to be associated to any general bounds on this distance, that we shall not discuss here. Such bounds are for instance given in [AGM14] in the case of matrices: we gave there some upper and lower bounds on the ratio between the moduli of the eigenvalues of a matrix $\mathcal{A}$ and the exponential of the tropical eigenvalues of the matrix $\mathsf{v}(\mathcal{A})$. Moreover, if $\mathcal{A}$ is a matrix polynomial as in (4.15) with either $\mathcal{A}_0 = I$ or $\mathcal{A}_d = I$, then a linearization of $\mathcal{A}$ transforms the eigenvalues of $\mathcal{A}$ into the eigenvalues of a matrix $\mathcal{C}$, and the tropical eigenvalues of A into the tropical eigenvalues of $\mathsf{v}(\mathcal{C})$, so that the bounds of [AGM14] can also be applied in that case.

Theorem 4.26. *Let us use the notations of Lemma 4.24. Let S, S' be the supports of the vectors x and y respectively. Let $\Delta = \max_{(i,j)\in G}(U_i + V_j)$. The following lower bounds hold:*

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\Delta - \max_{i\in[n]} U_i - \max_{j\in[n]} V_j)) \frac{1}{nd\,\mathcal{T}(m)^d} \tag{4.20}$$

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\Delta - \max_{i\in[n]} U_i - \min_{j\in S} V_j) \frac{1}{n^{3/2}d\,\mathcal{T}(m)^d\,\mathcal{H}_{\mathrm{su}}(x)} \tag{4.21}$$

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\Delta - \min_{i\in S'} U_i - \max_{j\in[n]} V_j) \frac{1}{n^{3/2}d\,\mathcal{T}(m)^d\,\mathcal{H}_{\mathrm{su}}(y)} \tag{4.22}$$

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\Delta - \min_{i\in S'} U_i - \min_{j\in S} V_j) \frac{1}{n^{2}d\,\mathcal{T}(m)^d\,\mathcal{H}_{\mathrm{su}}(x)\,\mathcal{H}_{\mathrm{su}}(y)} \ . \tag{4.23}$$

Proof. Let m, x, y, ℓ, z, w be as in Lemma 4.24. From the first assertion of this lemma, we deduce that $\|\mathcal{B}_k\| \leqslant n$, for all $k = 0, \dots, d$. Moreover, from the second one, we have $|(\mathcal{B}_k)_{ij}| = 1$, for $(i,j) \in G_k$, hence

$$\|\,\mathsf{d}(U)\mathcal{B}_k\,\mathsf{d}(V)\| \geqslant e^{U_i}|(\mathcal{B}_k)_{ij}|e^{V_j} = e^{U_i}e^{V_j} \quad \forall (i,j) \in G_k \ .$$

In particular, since $|m|^k \geqslant \min(1,|m|)^k \geqslant \min(1,|m|)^d$, we deduce that

$$\begin{aligned}\frac{\sum_{k=0}^d \|\,\mathsf{d}(U)\mathcal{B}_k\,\mathsf{d}(V)\|\ |m|^k}{\sum_{k=0}^d \|\mathcal{B}_k\|\ |m|^k} &\geqslant \frac{\max_{(i,j)\in G}(e^{U_i}e^{V_j})\min(1,|m|)^d}{n\sum_{k=0}^d |m|^k} \\ &\geqslant \max_{(i,j)\in G}(e^{U_i}e^{V_j})\frac{\min(1,|m|)^d}{nd\max(1,|m|)^d} \\ &\geqslant \exp(\Delta)\frac{1}{nd\,\mathcal{T}(m)^d}\end{aligned}$$

Combining this inequality with (4.19), we deduce

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \frac{\|y\,\mathsf{d}(-U)\|_2\|\,\mathsf{d}(-V)x\|_2}{\|y\|_2\|x\|_2}\exp(\Delta)\frac{1}{nd\,\mathcal{T}(m)^d} \ . \tag{4.24}$$

Since $\|x\|_2 \leqslant \exp(\max(V_1,\dots,V_n))\|\,\mathsf{d}(-V)x\|_2$, for all $V \in \mathbb{R}^n$ and $x \in \mathbb{C}^n$, and $\|y\|_2 \leqslant \exp(\max(U_1,\dots,U_n))\|y\,\mathsf{d}(-U)\|_2$, for all $U \in \mathbb{R}^n$ and $y \in (\mathbb{C}^n)^*$, (4.20) follows from (4.24).

Now, we also have $\|\mathsf{d}(-V)\chi\|_2 \geqslant e^{-V_i}|\chi_i|$ for all $i \in [n]$ such that $\chi_i \neq 0$. Hence, if $S \subset [n]$ is the support of χ, we have $\|\mathsf{d}(-V)\chi\|_2 \geqslant \max_{j\in S} e^{-V_j} \min_{j\in S} |\chi_j|$ and $\|\chi\|_2 \leqslant \max_{j\in S} |\chi_j|\sqrt{n}$. This implies

$$\frac{\|\mathsf{d}(-V)\chi\|_2}{\|\chi\|_2} \geqslant \max_{j\in S} e^{-V_j} \frac{1}{\mathcal{H}_{\mathrm{su}}(\chi)\sqrt{n}} \ .$$

Gathering this inequality with the previous lower bound of $\|y\,\mathsf{d}(-U)\|_2/\|y\|_2$, and with (4.24), we obtain (4.21). Inequality (4.22) is obtained similarly by exchanging the roles of χ and y. Moreover, (4.23) is obtained by applying the same second type of inequality for χ and y in (4.24). □

The lower bound (4.23) of Theorem 4.26 is similar to the lower bound used in the proof of Theorem 4.20. It implies easily the following one, which has to be compared with the result of Corollary 4.21.

Corollary 4.27. *Let us use the notations of Lemma 4.24. Assume that the supports of the vectors χ and y are full. Then*

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\max(|U|_{\mathrm{H}},|V|_{\mathrm{H}}))\frac{1}{n^2 d\,\mathcal{T}(m)^d\,\mathcal{H}(\chi)\,\mathcal{H}(y)} \ .$$

By applying the same arguments as in the proof of Theorem 4.26 to χ_C and y_C instead of χ and y, one can obtain the inequalities of Theorem 4.26, with G, S and S' restricted to the nodes of C, and up to the multiplication of the right hand sides by the scalar $\frac{\|\chi_C\|_2}{\|\chi\|_2}\frac{\|y_C\|_2}{\|y\|_2}$. In particular, one deduces the following result.

Corollary 4.28. *Let us use the notations of Lemma 4.24. Assume that C is a strongly connected component of G, and that the supports of the vectors χ and y both contain C. Then*

$$\frac{\kappa(\mathcal{A},\ell)}{\kappa(\mathcal{B},m)} \geqslant \exp(\max(|U_C|_{\mathrm{H}},|V_C|_{\mathrm{H}}))\frac{1}{n^2 d\,\mathcal{T}(m)^d\,\mathcal{H}(\chi_C)\,\mathcal{H}(y_C)}\frac{\|\chi_C\|_2}{\|\chi\|_2}\frac{\|y_C\|_2}{\|y\|_2} \ .$$

If one of the above lower bounds is greater than 1, then the diagonal scaling of Lemma 4.24 improves the condition number of the eigenvalue ℓ of $\mathcal{A}$. This holds in particular as soon as $|m|$ is close to 1, that the scaled eigenvectors χ, y have full supports and small oscillations, and that one of the vectors U and V is not constant and has large oscillations. The first property means that e^{γ} is a good approximation of the modulus of the complex eigenvalue ℓ of $\mathcal{A}$. The second one means that the vectors e^{-V} and $(e^{-U})^T$ are good approximations of the right and left eigenvectors z and w of $\mathcal{A}$ for the eigenvalue ℓ. To obtain the last one, one can apply the following result which is bounding from below the oscillations of the Hungarian pair, using the oscillations of some of the matrix polynomial entries.

Proposition 4.29. *Let us use the notations of Lemma 4.24. Then*

$$\max_{k=0,\dots,d} \mathcal{H}_{\mathrm{su}}(\mathcal{A}_k^{G_k}) \leqslant \exp(|U|_{\mathrm{H}}+|V|_{\mathrm{H}}) \leqslant \exp(2\max(|U|_{\mathrm{H}},|V|_{\mathrm{H}})) \ .$$

Proof. Since $|(\mathcal{A}_k)_{ij}|e^{k\gamma-U_i-V_j} = |(\mathcal{B}_k)_{ij}| = 1$ for all $(i,j) \in G_k = \mathcal{G}_k(A,\gamma)\cap G$, we get that for all $k = 0,\dots,d$ and $(i,j),(i',j') \in G_k$, we have $\frac{|(\mathcal{A}_k)_{ij}|}{|(\mathcal{A}_k)_{i'j'}|} = e^{U_i+V_j-U_{i'}-v_{j'}} \leqslant \exp(|U|_{\mathrm{H}}+|V|_{\mathrm{H}})$. Taking the maximum over all $(i,j),(i',j') \in G_k$, and then over all $k = 0,\dots,d$, we obtain the left inequality. The right inequality is immediate. □

The difficulty in the above lower bound is that it depends on the knowledge of the tropical eigenvalues and on the associated optimal permutations. However, gathered with Corollary 4.27 and the above observations, it shows that under some structural assumptions on the matrix polynomial, the eigenvalue condition number is reduced after scaling. Note that one can also obtain an upper bound on the oscillations of the Hungarian pair, by using the oscillations of all the matrix polynomial entries.

Example 4.6. Let us return to the matrix polynomial of Example 4.4. For any fixed $\epsilon > 0$, one can construct the scaling of Lemma 4.24 on the matrix polynomial $\mathcal{A}_\epsilon$. We have

$$A_\epsilon := \mathsf{v}(\mathcal{A}_\epsilon) = \begin{bmatrix} \log\delta + (\log\epsilon^2) & \log\epsilon & \log\epsilon \\ \log\epsilon & \log\delta & -\infty \\ \log\epsilon & -\infty & \log\delta \end{bmatrix} \oplus \mathsf{Y} \begin{bmatrix} \log\epsilon^2 & -\infty & -\infty \\ -\infty & 0 & -\infty \\ -\infty & -\infty & 0 \end{bmatrix}$$
$$= \begin{bmatrix} (\log\epsilon^2) \otimes (\mathsf{Y} \oplus \log\delta) & \log\epsilon & \log\epsilon \\ \log\epsilon & \mathsf{Y} \oplus \log\delta & -\infty \\ \log\epsilon & -\infty & \mathsf{Y} \oplus \log\delta \end{bmatrix} .$$

Since $\delta < 1$, we get $\mathrm{per}(A_\epsilon) = (\log\epsilon^2) \otimes (\mathsf{Y} \oplus \log\delta) \otimes (\mathsf{Y}^2 \oplus (\log\delta) \otimes \mathsf{Y} \oplus 0)$, the tropical roots of which are $\log\delta$ with multiplicity 1 and 0 with multiplicity 2. For $\gamma = \log\delta$, we have

$$\widehat{A_\epsilon}(\gamma) = \begin{bmatrix} 2\log\epsilon + \log\delta & \log\epsilon & \log\epsilon \\ \log\epsilon & \log\delta & -\infty \\ \log\epsilon & -\infty & \log\delta \end{bmatrix} ,$$

a Hungarian pair (U, V) of which is given by

$$U = \begin{bmatrix} \log\epsilon - \log\delta \\ 0 \\ 0 \end{bmatrix} , \quad V = \begin{bmatrix} \log\epsilon \\ \log\delta \\ \log\delta \end{bmatrix} .$$

The graph G contains the arcs $(1, i)$, $(i, 1)$ and (i, i) with $i = 2, 3$. The diagonal scaling of $\mathcal{A}_\epsilon$ proposed in Lemma 4.24 gives

$$\mathcal{B}_\epsilon = \mathsf{d}(-U)\mathcal{A}_\epsilon(e^\gamma \mathsf{Y})\,\mathsf{d}(-V) = \begin{bmatrix} (\mathsf{Y}-1)\delta^2 & 1 & 1 \\ 1 & \mathsf{Y}-1 & 0 \\ 1 & 0 & \mathsf{Y}-1 \end{bmatrix}$$
$$= \begin{bmatrix} -\delta^2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} + \mathsf{Y} \begin{bmatrix} \delta^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

The eigenvalue δ of $\mathcal{A}_\epsilon$ is transformed into the eigenvalue 1 of $\mathcal{B}_\epsilon$, and for this eigenvalue, we have:

$$\mathcal{B}_\epsilon(1) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} ,$$

so the vectors z and w such that $z^T = w = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$ are again the right and left eigenvectors of $\mathcal{B}_\epsilon$, then $\kappa(\mathcal{B}_\epsilon, 1) = \|(\mathcal{B}_\epsilon)_0\| + \|(\mathcal{B}_\epsilon)_1\|$, and we found

$$\sqrt{2} + 1 \leqslant \kappa(\mathcal{B}_\epsilon, 1) \leqslant \sqrt{7} + 1 .$$

Note that $\kappa(\mathcal{B}_\epsilon, 1) > \kappa(\mathcal{A}_\epsilon, \delta)$, so that the above diagonal scaling is still increasing the eigenvalue condition number, however the above upper bound shows that this time the increasing is not too large. This example shows that even in the case of a matrix polynomial depending on the parameter ϵ, it is more efficient to apply the diagonal scaling constructed from the archimedean valuation rather than the one constructed from the non-archimedean valuation. This is because the diagonal scaling constructed from the archimedean valuation has entries with general first order asymptotics, whereas the one constructed from the non-archimedean valuation has entries of the form ϵ^U only.

4.6 Eigenvalue Condition numbers of linearizations

In general, the numerical computation of eigenvalues is done by first linearizing the matrix polynomial, and then applying a QZ algorithm to the linearization. If we apply this method to compute the eigenvalues of a matrix polynomial, then the error on the eigenvalues can be bounded by the product of the eigenvalue condition number and of the backward error, either by considering these numbers with respect to the matrix polynomial under study, or by considering these numbers with respect to the linearized matrix polynomial used in the numerical computation. Then, given the property that the QZ algorithm is backward stable, the efficiency of the diagonal scaling of the previous section can be measured by its effect on the condition number of the eigenvalues with respect to the linearized matrix polynomial. The latter parameter can be studied by studying separately the condition number of the eigenvalues with respect to the matrix polynomial, and the ratio between the condition number of the eigenvalues with respect to the linearized matrix polynomial and the condition number of the same eigenvalues with respect to the initial matrix polynomial. The first study has already been done in the previous section. In the present section, we consider the second study. Note that another way to measure the efficiency would have been to combine the first study (the one of the previous section), with a study of the ratio between the backward error of the computed eigenvalues with respect to the initial matrix polynomial and the backward error with respect to the linearized matrix polynomial.

Let us consider the matrix polynomial of (4.15), and let $\mathcal{C}_\mathcal{A} = \mathcal{C}_0 + \mathsf{Y}\mathcal{C}_1$ be the companion linearization of $\mathcal{A}$ given in block form by:

$$\mathcal{C}_0 := \begin{pmatrix} \mathcal{A}_0 & 0 & \cdots & 0 \\ 0 & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I \end{pmatrix}, \quad \mathcal{C}_1 := \begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \cdots & \mathcal{A}_d \\ -I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I & 0 \end{pmatrix}. \tag{4.25}$$

Then any eigenvalue ℓ of $\mathcal{A}$ is an eigenvalue of $\mathcal{C}_\mathcal{A}$ with same multiplicity and vice versa. Moreover, any right and left eigenvectors x and y of $\mathcal{C}_\mathcal{A}$ associated to a simple nonzero eigenvalue of ℓ can be written in block form:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \cdots y_d \end{pmatrix},$$

with

$$\mathcal{X}_i = \ell^{i-1} z, \quad i = 1, \dots d, \tag{4.26a}$$
$$y_1 = w, \tag{4.26b}$$
$$y_i = \ell^{-(i-1)} w(\mathcal{A}_0 + \ell \mathcal{A}_1 + \cdots + \ell^{i-1}\mathcal{A}_{i-1}) \tag{4.26c}$$
$$= -\ell^{-(i-1)} w(\mathcal{A}_i \ell^i + \cdots + \ell^d \mathcal{A}_d), \quad i = 2, \dots d \ , \tag{4.26d}$$

where z and w are right and left eigenvectors of $\mathcal{A}$ associated to ℓ.

The following result appeared already in [HMT06], for the companion linearization exchanging the roles of $\mathcal{A}_k$ and $\mathcal{A}_{d-k}$. We give a proof for completeness.

Proposition 4.30 (Compare with [HMT06, Thm 7.3 and after])**.** *Let $\mathcal{A}$ be a matrix polynomial, and $\mathcal{C}_\mathcal{A}$ be the companion linearization of (4.25). Let ℓ be a simple nonzero eigenvalue of $\mathcal{A}$, and let $\mathcal{X}$ and y be right and left eigenvectors of $\mathcal{C}_\mathcal{A}$ associated to ℓ with block form (4.26). Denote $\lambda = \sqrt{1 + \cdots + |\ell|^{2(d-1)}}$. We have*

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} = \frac{\|y\|_2}{\|w\|_2} \lambda \frac{\|\mathcal{C}_0\| + \|\mathcal{C}_1\| \, |\ell|}{\sum_{k=0}^d \|\mathcal{A}_k\| \, |\ell|^k} \ .$$

Proof. We have

$$y\mathcal{C}'_\mathcal{A}(\ell)\mathcal{X} = y\mathcal{C}_1\mathcal{X} = w\mathcal{A}'(\ell)z,$$

hence

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} = \frac{\|y\|_2 \|\mathcal{X}\|_2 (\|\mathcal{C}_0\| + \|\mathcal{C}_1\| \, |\ell|)}{\|w\|_2 \|z\|_2 (\sum_{k=0}^d \|\mathcal{A}_k\| \, |\ell|^k)} \ .$$

In view of Equations (4.26), we get that

$$\|\mathcal{X}\|_2 = \|z\|_2 \sqrt{1 + \cdots + |\ell|^{2(d-1)}} = \|z\|_2 \lambda$$

which implies the result of the proposition. □

We deduce the following estimates (the lower bound appeared in [HMT06, §7] in the case of quadratic matrix polynomials):

Theorem 4.31. *Let $\mathcal{A}$ be a matrix polynomial, and $\mathcal{C}_\mathcal{A}$ be the companion linearization of (4.25). For any simple nonzero eigenvalue ℓ of $\mathcal{A}$, we have:*

$$\frac{1}{\sqrt{d}} \max(\frac{1}{a}, 1) \leqslant \frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant C_d \frac{\max(1, a)^2 \max(1, |\ell|)^d}{\sum_{k=0}^d \|\mathcal{A}_k\| \, |\ell|^k} \ ,$$

with $a = \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_d\|)$ and $C_d = d\sqrt{\frac{(d+1)(2d+1)}{6}}(\sqrt{d+1} + 1)$.

Proof. Let a be as in the theorem. In view of Equations (4.26), we obtain

$$\begin{aligned}
\|w\|_2 \leqslant \|y\|_2 &\leqslant \|w\|_2 \sqrt{1 + \max(\|\mathcal{A}_2\|, \dots, \|\mathcal{A}_d\|)^2 \sum_{i=2}^d (|\ell| + \cdots + |\ell|^{d-i+1})^2} \\
&\leqslant \|w\|_2 \sqrt{1 + a^2 \max(1, |\ell|)^{2(d-1)} \sum_{i=2}^d (d - i + 1)^2} \\
&\leqslant \|w\|_2 \max(1, a) \max(1, |\ell|)^{d-1} \sqrt{1 + \frac{d(d-1)(2d-1)}{6}}
\end{aligned}$$

and symmetrically:

$$\begin{aligned}\|y\|_2 &\leqslant \|w\|_2\sqrt{1+\max(\|\mathcal{A}_0\|,\dots,\|\mathcal{A}_{d-1}\|)^2\sum_{i=2}^{d}(1+\cdots+|\ell|^{-i+1})^2}\\ &\leqslant \|w\|_2\sqrt{1+a^2\max(1,|\ell|^{-1})^{2(d-1)}\sum_{i=2}^{d}i^2}\\ &\leqslant \|w\|_2\max(1,a)\max(1,|\ell|^{-1})^{d-1}\sqrt{\frac{d(d+1)(2d+1)}{6}}\end{aligned}$$

Taking the minimum of the two inequalities, and using that either $|\ell|\leqslant 1$ or $|\ell|^{-1}\leqslant 1$, and $d\geqslant 1$, we obtain:

$$\|w\|_2 \leqslant \|y\|_2 \leqslant \|w\|_2\max(1,a)C'_d \ . \tag{4.27}$$

with $C'_d := \sqrt{\frac{d(d+1)(2d+1)}{6}}$. Equation (4.25) implies

$$\|\mathcal{C}_0\| = \max(1,\|\mathcal{A}_0\|) \ . \tag{4.28}$$

Moreover

$$\|\mathcal{C}_1\|^2 = \max\{\|\mathcal{A}_1x_1+\cdots+\mathcal{A}_dx_d\|_2^2+\|x_1\|_2^2+\cdots\|x_{d-1}\|_2^2 \mid x,\ \|x_1\|_2^2+\cdots\|x_d\|_2^2=1\} \ ,$$

hence

$$\max(1,\|\mathcal{A}_1\|,\cdots,\|\mathcal{A}_d\|) \leqslant \|\mathcal{C}_1\| \leqslant \sqrt{1+\|\mathcal{A}_1\|^2+\cdots\|\mathcal{A}_d\|^2} \tag{4.29}$$

$$\leqslant \sqrt{d+1}\max(1,\|\mathcal{A}_1\|,\cdots,\|\mathcal{A}_d\|) \ . \tag{4.30}$$

Let us show the left inequality of the theorem. Using Proposition 4.30, (4.27),(4.28), and (4.29), we obtain in particular

$$\frac{\kappa(\mathcal{C}_\mathcal{A},\ell)}{\kappa(\mathcal{A},\ell)} \geqslant \lambda\frac{\max(1,\|\mathcal{A}_0\|)+\max(1,\|\mathcal{A}_1\|,\dots,\|\mathcal{A}_d\|)|\ell|}{\sum_{k=0}^{d}\|\mathcal{A}_k\|\ |\ell|^k} \ . \tag{4.31}$$

Now

$$\begin{aligned}\sum_{k=1}^{d}\|\mathcal{A}_k\|\ |\ell|^k &\leqslant |\ell|\max(\|\mathcal{A}_1\|,\cdots,\|\mathcal{A}_d\|)(\sum_{k=0}^{d-1}|\ell|^k)\\ &\leqslant |\ell|\min(a,1)\max(1,\|\mathcal{A}_1\|,\cdots,\|\mathcal{A}_d\|)\lambda\sqrt{d}\end{aligned}$$

and

$$\|\mathcal{A}_0\| \leqslant \min(a,1)\max(1,\|\mathcal{A}_0\|)\lambda\sqrt{d} \ .$$

So, we deduce that

$$\sum_{k=0}^{d}\|\mathcal{A}_k\|\ |\ell|^k \leqslant \min(a,1)\Big(\max(1,\|\mathcal{A}_0\|)+\max(1,\|\mathcal{A}_1\|,\cdots,\|\mathcal{A}_d\|)\ |\ell|\Big)\lambda\sqrt{d}$$

which, with (4.31), implies

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \frac{1}{\sqrt{d}} \frac{1}{\min(a, 1)} = \frac{1}{\sqrt{d}} \max(\frac{1}{a}, 1) \ .$$

Let us show now the right inequality of the theorem. From (4.27), (4.28), and (4.30), we get

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant \max(1, a) C'_d \lambda \frac{\max(1, \|\mathcal{A}_0\|) + |\ell| \sqrt{d+1} \max(1, \|\mathcal{A}_1\|, \ldots, \|\mathcal{A}_d\|)}{\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k} \ .$$

Using $\lambda \leqslant \sqrt{d} \max(1, |\ell|)^{d-1}$, and the previous inequality, we deduce:

$$\begin{aligned} \frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} &\leqslant C'_d \sqrt{d} \max(1, a)^2 \max(1, |\ell|)^{d-1} \frac{1 + |\ell| \sqrt{d+1}}{\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k} \\ &\leqslant d \sqrt{\frac{(d+1)(2d+1)}{6}} (\sqrt{d+1} + 1) \max(1, a)^2 \frac{\max(1, |\ell|)^d}{\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k} \ , \end{aligned}$$

which is the upper bound of the theorem. □

We then derive from the upper bound of Theorem 4.31 several coarser upper bounds which will allows us to determine sufficient conditions under which the eigenvalue condition numbers with respect to the linearized polynomial and the initial polynomial have same order of magnitude. The following one is similar to the one obtained in [HMT06, §7] for quadratic matrix polynomials.

Corollary 4.32. *Let $\mathcal{A}$ be a matrix polynomial, and $\mathcal{C}_\mathcal{A}$ be the companion linearization of (4.25). For any simple nonzero eigenvalue ℓ of $\mathcal{A}$, we have:*

$$\frac{1}{\sqrt{d}} \max(\frac{1}{a}, 1) \leqslant \frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant C_d \frac{\max(1, a)^2}{\min(\|\mathcal{A}_0\|, \|\mathcal{A}_d\|)} \ ,$$

with a and C_d as in Theorem 4.31.

Proof. Using

$$\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k \geqslant \min(\|\mathcal{A}_0\|, \|\mathcal{A}_d\|) \max(1, |\ell|)^d$$

together with the upper bound of Theorem 4.31, we get the upper bound in the corollary. □

The following one is new and is based on the Thompson distance between the eigenvalue modulus and 1 and the one between a and 1.

Corollary 4.33. *Let $\mathcal{A}$ be a matrix polynomial, and $\mathcal{C}_\mathcal{A}$ be the companion linearization of (4.25). For any simple nonzero eigenvalue ℓ of $\mathcal{A}$, we have:*

$$\frac{1}{\sqrt{d}} \leqslant \frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant C_d \mathcal{T}(a) \mathcal{T}(\ell)^d \ ,$$

with a and C_d as in Theorem 4.31. More precisely if $a = \|\mathcal{A}_k\|$ for $0 \leqslant k \leqslant d$, we have

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant C_d \mathcal{T}(a) \max(1/|\ell|^k, |\ell|^{d-k}) \ .$$

If in addition $a = \|\mathcal{A}_{k'}\|$ *for* $k < k' \leqslant d$*, then*

$$\frac{\kappa(\mathcal{C}_{\mathcal{A}}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant C_d \mathcal{T}(a) \max(1/|\ell|^k, |\ell|^{d-k'}) \ .$$

Proof. Using

$$\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k \geqslant \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_d\|) \min(1, |\ell|)^d = a \min(1, |\ell|)^d \ ,$$

and $\max(1, |\ell|)/\min(1, |\ell|) = \max(|\ell|, 1/|\ell|) = \mathcal{T}(\ell)$ in the upper bound of Theorem 4.31, we deduce the first upper bound of the corollary. Now if $a = \|\mathcal{A}_k\|$, we deduce that $\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k \geqslant a \ |\ell|^k$, and applying this inequality in the upper bound of Theorem 4.31, we get:

$$\begin{aligned} \frac{\kappa(\mathcal{C}_{\mathcal{A}}, \ell)}{\kappa(\mathcal{A}, \ell)} &\leqslant C_d \mathcal{T}(a) \frac{\max(1, |\ell|)^d}{|\ell|^k} \\ &= C_d \mathcal{T}(a) \max(1/|\ell|^k, |\ell|^{d-k}) \ . \end{aligned}$$

If in addition $a = \|\mathcal{A}_{k'}\|$ for $k < k' \leqslant d$, then applying the previous inequality, we get

$$\begin{aligned} \frac{\kappa(\mathcal{C}_{\mathcal{A}}, \ell)}{\kappa(\mathcal{A}, \ell)} &\leqslant C_d \mathcal{T}(a) \min(\max(1/|\ell|^k, |\ell|^{d-k}), \max(1/|\ell|^{k'}, |\ell|^{d-k'})) \\ &= C_d \mathcal{T}(a) \max(1/|\ell|^k, |\ell|^{d-k'}) \ . \end{aligned}$$

□

Remark 4.7. The previous result can be used to justify the use of the tropical scaling. Indeed, let p, γ and $\mathcal{B}$ be as in Remark 4.5, so as $\mathcal{B}$ is the matrix polynomial obtained after tropical scaling. Then, as explained there ℓ is an eigenvalue of $\mathcal{A}$ if and only if $m = e^{-\gamma}\ell$ is an eigenvalue of $\mathcal{B}$, and we have $\kappa(\mathcal{A}, \ell) = \kappa(\mathcal{B}, m)$. However, the tropical scaling is changing the eigenvalue condition number with respect to the linearized matrix polynomial. Denote by $\mathcal{C}_{\mathcal{B}}$ the companion linearization of $\mathcal{B}$. Let m be the multiplicity of the root γ of p. There exists $0 \leqslant k < k' = k + m \leqslant d$ such that $\widehat{p}(\gamma) = p_k\gamma^k = p_{k'}\gamma^{k'} \geqslant p_\ell\gamma^\ell$ (in $\mathbb{R}_{\max}$), for all $0 \leqslant \ell \leqslant d$. After scaling, we have $\|\mathcal{B}_\ell\| = e^{\ell\gamma - \widehat{p}(\gamma)}\|\mathcal{A}_\ell\|$ for all $\ell = 0, \ldots, d$, so that $a := \max(\|\mathcal{B}_\ell\|, \ \ell = 0, \ldots, d) = 1 = \|\mathcal{B}_k\| = \|\mathcal{B}_{k'}\|$. Applying Corollary 4.33 to $\mathcal{B}$, we get

$$\frac{1}{\sqrt{d}} \leqslant \frac{\kappa(\mathcal{C}_{\mathcal{B}}, m)}{\kappa(\mathcal{B}, m)} \leqslant C_d \max(1/|m|^k, |m|^{d-k'}) \ . \tag{4.32}$$

Also a similar bound has been obtained by Sharify and Tisseur [ST14], in the case of more general linearizations. This bound shows that, if e^γ is a good approximation of $|\ell|$, so that $|m|$ is close to 1, then $\kappa(\mathcal{C}_{\mathcal{B}}, m)$ is in the same order as $\kappa(\mathcal{B}, m) = \kappa(\mathcal{A}, \ell)$. In that case, the numerical computation of ℓ with QZ algorithm combined with the companion linearization will be accurate if and only if $\kappa(\mathcal{A}, \ell)$ is not too large. Moreover, a sufficient condition for the property that e^γ is a good approximation of some of the eigenvalues $|\ell|$ of $\mathcal{A}$ to hold is that the matrices $\mathcal{A}_k$ are well conditioned and that the tropical eigenvalues are well separated, see [GS09, BNS13, AGS13, NST14].

The following result shows that the above behavior of the tropical scaling holds similarly for the diagonal scaling.

Corollary 4.34. *Let us use the notations of Lemma 4.24. Denote by $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$ the companion linearizations of $\mathcal{A}$ and $\mathcal{B}$, respectively. Let $\mathfrak{m}$ be a simple eigenvalue of $\mathcal{B}$ and $\mathfrak{l} = e^\gamma \mathfrak{m}$ be the corresponding eigenvalue of $\mathcal{A}$. We have:*

$$\frac{1}{\sqrt{d}} \leqslant \frac{\kappa(\mathcal{C}_\mathcal{B}, \mathfrak{m})}{\kappa(\mathcal{B}, \mathfrak{m})} \leqslant C_d n \mathcal{T}(\mathfrak{m})^d \leqslant \sqrt{d} C_d n \mathcal{T}(\mathfrak{m})^d \frac{\kappa(\mathcal{C}_\mathcal{A}, \mathfrak{l})}{\kappa(\mathcal{A}, \mathfrak{l})} \ .$$

with C_d as in Theorem 4.31.

Proof. From Lemma 4.24, we have $b := \max(\|\mathcal{B}_0\|, \dots, \|\mathcal{B}_d\|) \leqslant n$, and, for all $(i,j) \in G$, there exists $k \in \{0, \dots, d\}$ such that $|(\mathcal{B}_k)_{ij}| = 1$. Since G contains at least one arc (it contains a permutation), we get that $b \geqslant \|\mathcal{B}_k\| \geqslant 1$. These bounds on b imply that $\mathcal{T}(b) \leqslant n$, and applying the first assertion of Corollary 4.33 to $\mathcal{B}$ and $\mathfrak{m}$, we get the two first inequalities of the present corollary. Then applying the lower bound of Corollary 4.33 to $\mathcal{A}$ and $\mathfrak{l}$, we deduce the last inequality of the present corollary. □

The above bound can be improved by adding a multiplicative factor to the diagonal scaling as follows. Note that this factor should be crucial when n is large.

Corollary 4.35. *Let us use the notations of Lemma 4.24, and let $\mathcal{B}' = (\max(\|\mathcal{B}_k\|, k = 0, \dots, d))^{-1}\mathcal{B}$. Denote by $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_{\mathcal{B}'}$ the companion linearizations of $\mathcal{A}$ and $\mathcal{B}'$, respectively. Let $\mathfrak{m}$ be a simple eigenvalue of $\mathcal{B}'$ and $\mathfrak{l} = e^\gamma \mathfrak{m}$ be the corresponding eigenvalue of $\mathcal{A}$. We have:*

$$\frac{1}{\sqrt{d}} \leqslant \frac{\kappa(\mathcal{C}_{\mathcal{B}'}, \mathfrak{m})}{\kappa(\mathcal{B}', \mathfrak{m})} \leqslant C_d \mathcal{T}(\mathfrak{m})^d \leqslant \sqrt{d} C_d \mathcal{T}(\mathfrak{m})^d \frac{\kappa(\mathcal{C}_\mathcal{A}, \mathfrak{l})}{\kappa(\mathcal{A}, \mathfrak{l})} \ .$$

with C_d as in Theorem 4.31. □

Remark 4.8. The inequalities in Corollaries 4.34 and 4.35 show that if e^γ is a good approximation of $|\mathfrak{l}|$, so that $|\mathfrak{m}|$ is close to 1, then $\kappa(\mathcal{C}_\mathcal{B}, \mathfrak{m})$ is in the same order as $\kappa(\mathcal{B}, \mathfrak{m})$ and the ratio between the eigenvalue condition number of the linearized matrix polynomial and the eigenvalue condition number of the matrix polynomial cannot increase too much after diagonal scaling. Combined with the result of Corollary 4.27 and the comments which follow it and which remain true for the scaling of Corollary 4.35, we obtain under the same conditions described there, that the condition number of $\mathfrak{m}$ with respect to the linearization $\mathcal{C}_\mathcal{B}$ of $\mathcal{B}$ is improved. Combined with the property that QZ is backward stable, this implies that the eigenvalue computation is improved by the diagonal scaling.

The following result shows that the condition that e^γ is a good approximation of $|\mathfrak{l}|$, or that $|\mathfrak{m}|$ is close to 1, is in general necessary to obtain that $\kappa(\mathcal{C}_\mathcal{B}, \mathfrak{m})$ is in the same order as $\kappa(\mathcal{B}, \mathfrak{m})$.

Theorem 4.36. *Let $\mathcal{A}$ be a matrix polynomial, and let $\mathcal{C}_\mathcal{A}$ be the companion linearization of (4.25). Let $a = \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_d\|)$. Then, for any simple nonzero eigenvalue $\mathfrak{l}$ of $\mathcal{A}$, we have:*

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \mathfrak{l})}{\kappa(\mathcal{A}, \mathfrak{l})} \geqslant \frac{1}{a} \max\left(\frac{1}{|\mathfrak{l}|\sqrt{d} + \frac{\|\mathcal{A}_0\|}{a}}, \frac{1}{\frac{\sqrt{d}}{|\mathfrak{l}|} + \frac{\|\mathcal{A}_d\|}{a}}, \frac{1}{\sqrt{d}} \right) \ .$$

Moreover, if $a > 1$ *and* $a = \max(\|\mathcal{A}_1\|, \cdots, \|\mathcal{A}_d\|)$*, then*

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \max\left(\frac{1}{\frac{\sqrt{d}}{|\ell|} + \frac{\|\mathcal{A}_d\|}{a}}, \frac{1}{\sqrt{d}} \right) \ .$$

Proof. From the arguments of the proof of Theorem 4.31, we have the lower bound (4.31). which implies in particular:

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \lambda \frac{1 + |\ell|}{\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k} \ , \tag{4.33}$$

with λ as in Proposition 4.30. Now

$$\sum_{k=0}^{d-1} \|\mathcal{A}_k\| \ |\ell|^k \leqslant \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_{d-1}\|)\lambda\sqrt{d} \leqslant a\lambda\sqrt{d} \ ,$$

and

$$\|\mathcal{A}_d\| \ |\ell|^d \leqslant \|\mathcal{A}_d\| \ |\ell|\lambda \ .$$

With (4.33), this implies:

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \frac{1 + |\ell|}{a\sqrt{d} + \|\mathcal{A}_d\| \ |\ell|} \geqslant \frac{1}{a} \ \frac{1}{\frac{\sqrt{d}}{|\ell|} + \frac{\|\mathcal{A}_d\|}{a}} \ .$$

Similarly,

$$\sum_{k=1}^{d} \|\mathcal{A}_k\| \ |\ell|^k \leqslant \max(\|\mathcal{A}_1\|, \cdots, \|\mathcal{A}_d\|) \ |\ell|\lambda\sqrt{d} \leqslant a|\ell|\lambda\sqrt{d} \ ,$$

hence, using $1 \leqslant \lambda$ and (4.33), this implies:

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \frac{1 + |\ell|}{a|\ell|\sqrt{d} + \|\mathcal{A}_0\|} \geqslant \frac{1}{a} \frac{1}{|\ell|\sqrt{d} + \frac{\|\mathcal{A}_0\|}{a}} \ .$$

Gathering the above inequalities, we obtain the first lower bound of the theorem. If now $a > 1$ and $a = \max(\|\mathcal{A}_1\|, \cdots, \|\mathcal{A}_d\|)$, then using (4.31), we can replace (4.33) by

$$\frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \geqslant \lambda \frac{1 + a|\ell|}{\sum_{k=0}^{d} \|\mathcal{A}_k\| \ |\ell|^k} \ ,$$

which with the above arguments gives the second lower bound of the theorem. □

Corollary 4.37. *Let* $\mathcal{A}$ *be a matrix polynomial, and* $\mathcal{C}_\mathcal{A}$ *be the companion linearization of* (4.25)*. Assume that* $a = \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_d\|) \leqslant 1$*, and let* $b = \min(\|\mathcal{A}_0\|, \|\mathcal{A}_d\|)$*. If* ℓ *is a nonzero simple eigenvalue such that either* $|\ell| \leqslant b/a$ *and* $b = \|\mathcal{A}_0\|$*, or* $1/|\ell| \leqslant b/a$ *and* $b = \|\mathcal{A}_d\|$*, then*

$$\frac{1}{(\sqrt{d}+1)b} \leqslant \frac{\kappa(\mathcal{C}_\mathcal{A}, \ell)}{\kappa(\mathcal{A}, \ell)} \leqslant \frac{C_d}{b} \ .$$

Proof. The lower bound follows from Theorem 4.36, and the upper bound follows from Corollary 4.32. □

Remark 4.9. The previous result shows that if b is much smaller than 1, it may happen that the eigenvalue condition number of the linearized polynomial is much larger than the one of the matrix polynomial. This holds in particular in the following situation. Let $\mathcal{A}$ be a matrix polynomial such that $a = \max(\|\mathcal{A}_0\|, \cdots, \|\mathcal{A}_d\|) \leqslant 1$ and such that $\mathbb{1}$ is a root of the tropical polynomial $p := \oplus_{i=0}^{d} p_k \mathsf{Y}^k \in \mathbb{R}_{\max}[\mathsf{Y}]$ with coefficients $p_k = \log \|\mathcal{A}_k\|$, $k = 0, \ldots, d$ (this holds with $a = 1$, if we already applied the tropical scaling of [GS09]). Assume that $b = \min(\|\mathcal{A}_0\|, \|\mathcal{A}_d\|) < a$, which holds if and only if $\mathbb{1}$ is not the only root of p, that is if its multiplicity is $< d$. Then $-\log(b/a)$ is a lower bound of the distance between the root $\mathbb{1}$ and the minimal and maximal roots of p. It is also an upper bound of the separation between the root $\mathbb{1}$ and the roots of p which are closest to $\mathbb{1}$. So b is much smaller than 1 if either a is much smaller than 1, or $\mathbb{1}$ is well separated from the other roots of p.

Assume now that $b = \|\mathcal{A}_0\|$ and $\mathcal{A}_0$ is invertible. Since $b < a$, we get that $a = \max(\|\mathcal{A}_1\|, \cdots, \|\mathcal{A}_d\|)$. Consider in that case the companion linearization $\mathcal{C}'$ of $\mathcal{A}_0^{-1}\mathcal{A}$ as in (4.25). We have that $\mathcal{C}' = I + \mathsf{Y}\mathcal{C}'_1$, and so ℓ is an eigenvalue of $\mathcal{A}$ if and only if $-1/\ell$ is an eigenvalue of the matrix $\mathcal{C}'_1$. Moreover, using the formula of $\mathcal{C}'_1$, and $a = \max(\|\mathcal{A}_1\|, \cdots, \|\mathcal{A}_d\|)$, we get that

$$\|\mathcal{C}'_1\| \geqslant \max(\|\mathcal{A}_0^{-1}\mathcal{A}_k\|, k = 1, \ldots, d) \geqslant \max(\frac{\|\mathcal{A}_k\|}{\|\mathcal{A}_0\|}, k = 1, \ldots, d) = \frac{a}{b} \ .$$

Hence, if $\|\mathcal{C}'_1\|$ is equal to the spectral radius of $\mathcal{C}'_1$, then there exists an eigenvalue ℓ of $\mathcal{A}$ such that $1/|\ell| = \|\mathcal{C}'_1\| \geqslant a/b$, or equivalently $|\ell| \leqslant b/a$, and if ℓ is a simple eigenvalue of $\mathcal{A}$ (or equivalently $-1/\ell$ is a simple eigenvalue of $\mathcal{C}'_1$), the assumption of Corollary 4.37 is satisfied. Moreover, a similar situation holds in the case where $b = \|\mathcal{A}_d\|$, $\mathcal{A}_d$ is invertible, and the matrix $\mathcal{C}'_1$ obtained as above with $\mathcal{A}_k$ replaced by $\mathcal{A}_{d-k}$, is such that $\|\mathcal{C}'_1\|$ is equal to the spectral radius of $\mathcal{C}'_1$.

4.7 Numerical examples

Given an asymptotic matrix polynomial $\mathcal{A}_\epsilon$ as in (4.1), we shall compute its eigenvalues for several values of $\epsilon \to 0$ and study the eigenvalue condition numbers and backward errors of their approximations.

We will use three different scalings on the matrix polynomial before computing the eigenvalues: first no scaling at all, then an "eigenvalue scaling" and finally the diagonal scaling described in Theorem 4.13, that we shall also call Hungarian scaling. The eigenvalue scaling goes as follows. As in Section 4.3, we consider the tropical matrix polynomial A in (4.3) with coefficients as in (4.2). For each tropical eigenvalue γ of A with multiplicity m_γ, we expect m_γ eigenvalues of order $\epsilon^{-\gamma}$ (see Theorem 4.1 or 4.2); we then build the matrix polynomial $\beta\tilde{\mathcal{A}}_\epsilon(\mathsf{Y}) = \mathcal{A}_\epsilon(\epsilon^{-\gamma}\mathsf{Y})$, where β is a normalization parameter. Each eigenvalue ℓ_ϵ of order $\epsilon^{-\gamma}$ of $\mathcal{A}_\epsilon$ corresponds to an eigenvalue $\ell_\epsilon\epsilon^\gamma$ of order 1 of $\tilde{\mathcal{A}}$. So, for each fixed $\epsilon > 0$, we take the m_γ eigenvalues of $\tilde{\mathcal{A}}$ the modulus of which are the closest to 1 (for the Thompson distance) and multiply them back by $\epsilon^{-\gamma}$ to recover the corresponding original eigenvalues of $\mathcal{A}_\epsilon$. Once we have done this for every γ, we expect to recover all the $n \cdot d$ eigenvalues.

The diagonal scaling follows the same idea, except this time instead of $\tilde{\mathcal{A}}_\epsilon$ we make use of the Hungarian variables and build the matrix polynomial $\mathcal{B}_\epsilon$ described in Theorem 4.13.

After scaling the matrix polynomial, we apply a customized version of MATLAB's `polyeig` function (which applies the companion linearization of Section 4.6 followed by QZ algorithm)

to compute its eigenvalues. The three methods should theoretically give the same results. We will test them against our best approximation of the true eigenvalues: the zeros of $\det \mathcal{A}_\epsilon(\lambda)$ as computed with the help of MATLAB's symbolic toolbox MuPAD.

Example 4.10. We choose $n = 5$, $d = 4$, and for 50 logarithmically equally spaced values of ϵ between 10^{-1} and 10^{-16} we generate the matrix polynomial

$$\mathcal{A}_\epsilon = \mathcal{A}_{\epsilon,0} + \mathsf{Y}\mathcal{A}_{\epsilon,1} + \cdots + \mathsf{Y}^d\mathcal{A}_{\epsilon,d}, \qquad \text{with } (\mathcal{A}_{\epsilon,k})_{ij} = a_k(J + \mathcal{E}_{\epsilon,k})_{ij}\epsilon^{-(A_k)_{ij}}. \tag{4.34}$$

The entries of the matrices A_k are either random small-denominator rationals between -2 and 2, or $-\infty$. All tropical eigenvalues of A are finite. The entries of the matrices a_k are complex, and both their real and imaginary parts are chosen from a Gaussian distribution $\mathcal{N}(0, 1)$. J is the matrix of all ones, and the entries of the matrices $\mathcal{E}_{\epsilon,k}$ are perturbations of 0 sampled independently from a normal distribution with mean 0 and standard deviation $-1/\log(\epsilon)$ (so that $\lim_{\epsilon\to 0} \mathcal{E}_{\epsilon,k} = 0$). The choice of a slowly convergent standard deviation was made to keep the perturbation noticeable. We assume that Condition (4.6) is also satisfied.

In Figures 4.1 to 4.3, we plot the modulus of all the eigenvalues obtained for the three types of scalings and compare them to the true ones obtained by symbolic computation. For each value of ϵ we expect $n \cdot d = 20$ distinct eigenvalues; nonetheless, each vertical slice in the following pictures will appear to account for less than that. This is actually a consequence of Theorem 4.2: eigenvalues corresponding to the same tropical eigenvalue are clustered in magnitude, and thus we are not able to tell them apart at the scale of the picture.

We only show one random instance of the problem; the features it exhibits are consistent throughout all the simulations we have run.

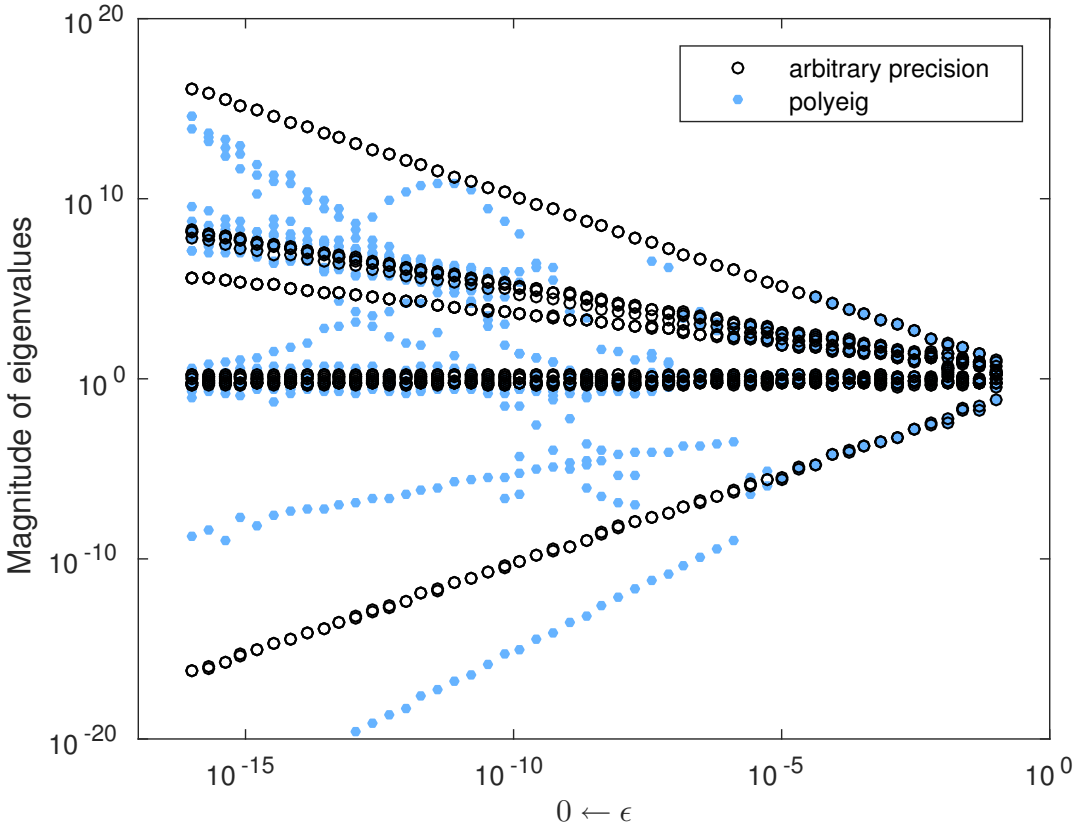


Figure 4.1: Comparison between the moduli of the true eigenvalues (computed with arbitrary precision as roots of $\det \mathcal{A}_\epsilon$, represented by hollow circles) and the moduli of the eigenvalues computed with `polyeig` (solid dots) of the matrix polynomial (4.34). Each vertical slice represents one particular value of ϵ. Clusters of eigenvalues with the exponential behavior predicted by Theorem 4.2 are noticeable when computing with arbitrary precision. It is clear from this picture that when ϵ is small, so that the entries of $\mathcal{A}$ vary a lot in magnitude, polyeig is not dependable.

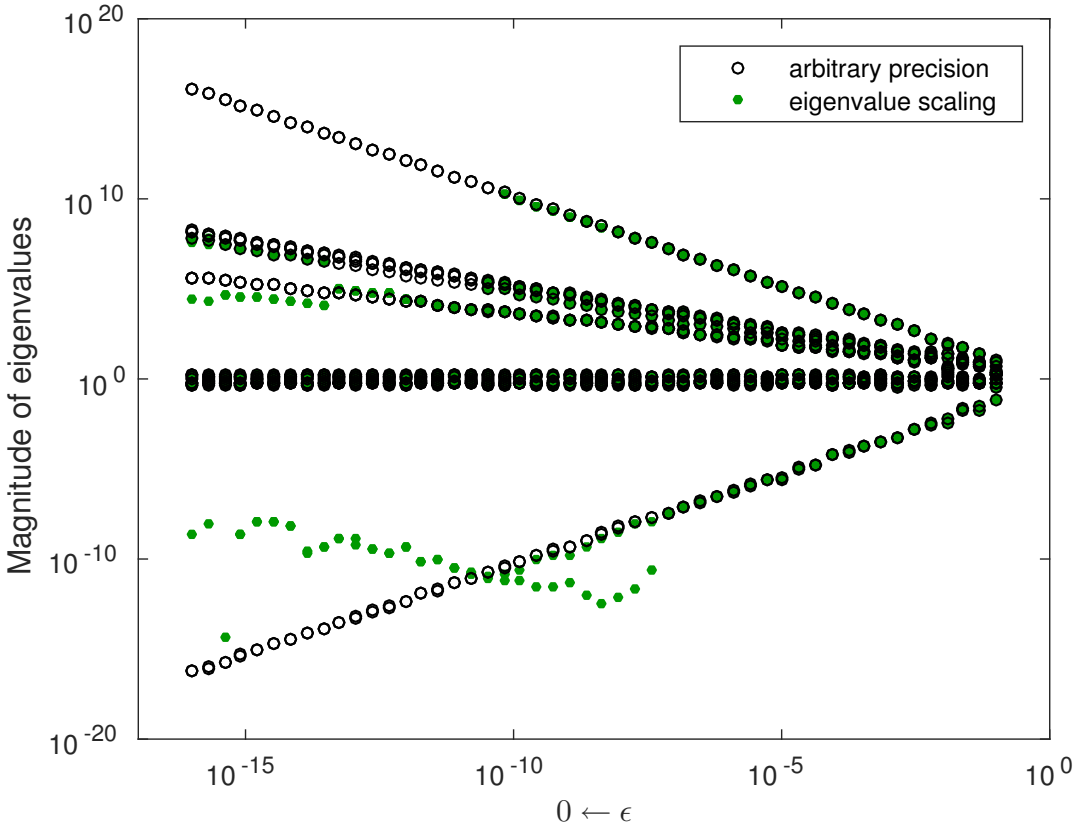


Figure 4.2: Comparison between the moduli of the true eigenvalues of (4.34) and the moduli of the eigenvalues obtained with `polyeig` after performing an "eigenvalue scaling". The resulting figure is less chaotic than the previous one, but still far from satisfying, as approaching zero some spurious eigenvalues appear, while some other ones are computed as infinite by `polyeig` (and thus they are not plotted).

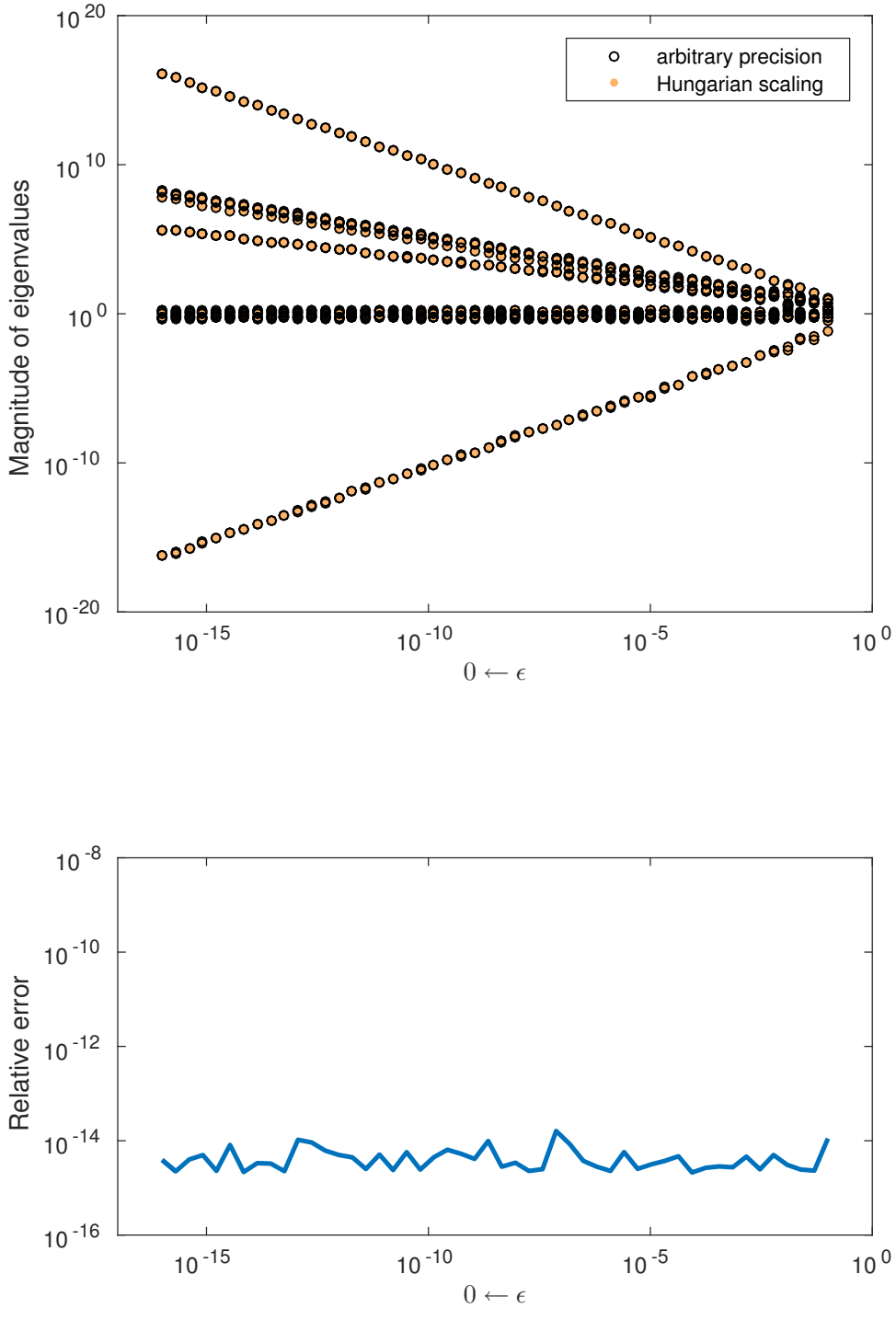


Figure 4.3: Comparison between the moduli of the true eigenvalues of (4.34) and the moduli of the eigenvalues obtained with `polyeig` after performing a diagonal scaling. Here the correspondence between true and computed eigenvalues is much better. The second graph shows, for each value of ϵ, the largest relative error among the nd computed eigenvalues.

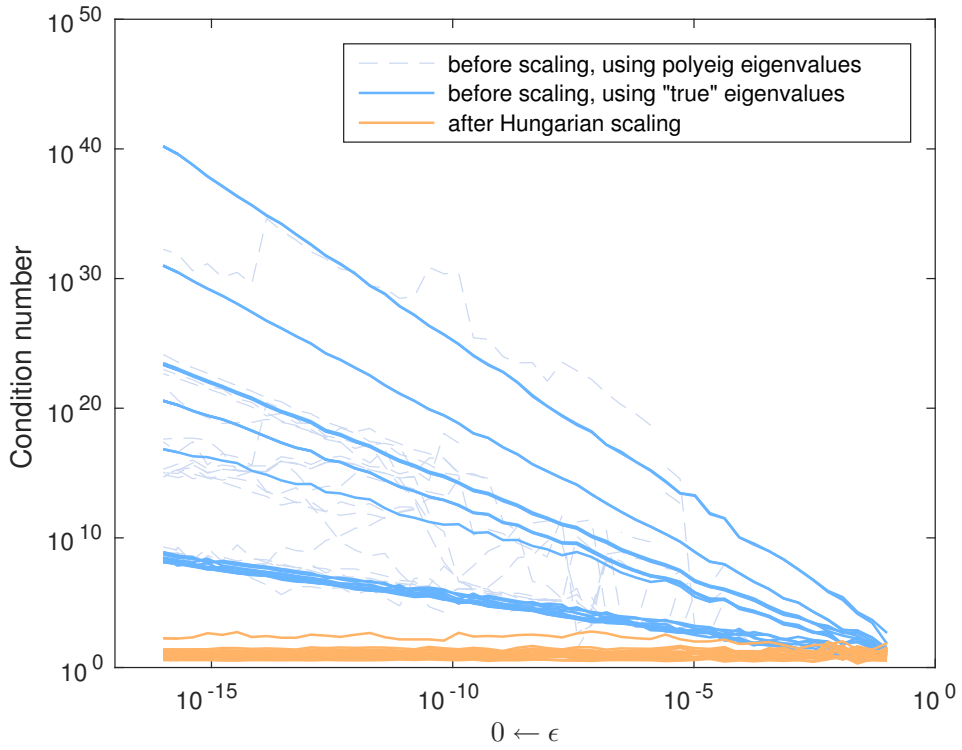


Figure 4.4: Condition number of the eigenvalues before and after diagonal scaling. The condition numbers are relative to the eigenvalues of the polynomial in its original form (4.1) (not the linearized form), and they are computed using formula (4.11). Note that since the formula depends on eigenvalues and eigenvectors, applying it with the low-accuracy eigenpairs obtained from polyeig before scaling can yield poor approximations of the true condition numbers (see the dashed lines in the picture). Plugging in the arbitrary precision values gives more consistent results (solid blue lines).

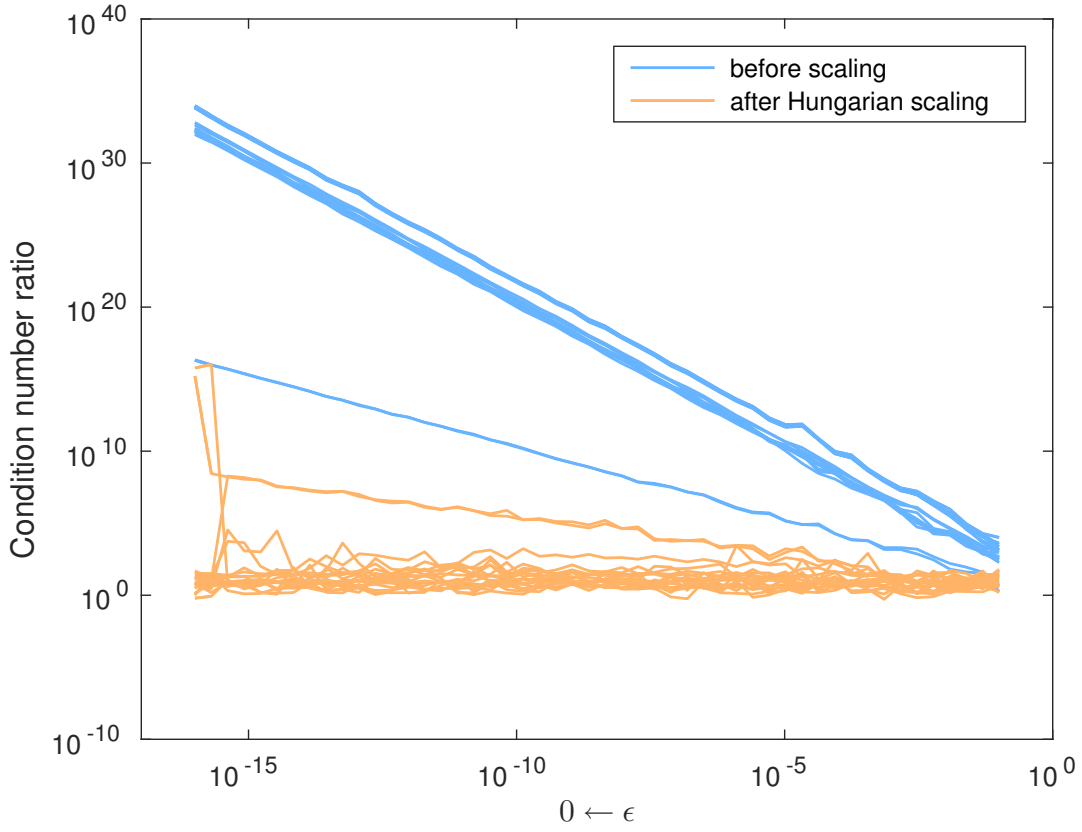


Figure 4.5: Ratios between the condition numbers of the eigenvalues of the linearization and those of the corresponding eigenvalues of the original matrix polynomial.

Bibliography

[ABG04] M. Akian, R. Bapat, and S. Gaubert. Perturbation of eigenvalues of matrix pencils and optimal assignment problem. *C. R. Acad. Sci. Paris, Série I*, 339:103–108, 2004.

[ABG15] M. Akian, R. Bapat, and S. Gaubert. Non-archimedean valuations of eigenvalues of matrix polynomials, 2015. Preprint.

[AGH+15] M. Akian, S. Gaubert, J. Hook, A. Marchesini, and F. Tisseur. Tropical diagonal scaling of matrix polynomials. Technical report, 2015. In preparation.

[AGM14] M. Akian, S. Gaubert, and A. Marchesini. Tropical bounds for eigenvalues of matrices. *Linear Algebra Appl.*, 446:281–303, 2014. See also arXiv:1309.7319.

[AGS13] M. Akian, S. Gaubert, and M. Sharify. Log-majorization of the moduli of the eigenvalues of a matrix polynomial by tropical roots, 2013. See also arXiv:1304.2967.

[BB03] R. E. Burkard and P. Butkovic. Finding all essential terms of a characteristic maxpolynomial. *Discrete Applied Mathematics*, 130(3):367–380, 2003.

[BCOQ92] F. Baccelli, G. Cohen, G.J. Olsder, and J.P. Quadrat. *Synchronization and Linearity — an Algebra for Discrete Event Systems*. Wiley, 1992.

[Bet08] T. Betcke. Optimal scaling of generalized and polynomial eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 30(4):1320–1338, 2008.

[BL07] P. Butkovič and S. Lewis. On the job rotation problem. *Discrete Optim.*, 4(2):163–174, 2007.

[BN13] Dario A. Bini and Vanni Noferini. Solving polynomial eigenvalue problems by means of the ehrlich–aberth method. *Linear Algebra and its Applications*, 439(4):1130 – 1149, 2013. 17th Conference of the International Linear Algebra Society, Braunschweig, Germany, August 2011.

[BNS13] D.A. Bini, V. Noferini, and M. Sharify. Locating the Eigenvalues of Matrix Polynomials. *SIAM J. Matrix Anal. Appl.*, 34(4):1708–1727, 2013.

[BR91] R. R. Brualdi and H. J. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, 1991.

[BR97] R.B. Bapat and T.E.S. Raghavan. *Nonnegative Matrices and Application*. Cambridge University Press, 1997.

[But10] Peter Butkovič. *Max-linear systems : theory and algorithms*. Springer monographs in mathematics. Springer, London, Dordrecht, Heidelberg, 2010.

[CG79] R. A. Cuninghame-Green. *Minimax algebra*. Lecture notes in economics and mathematical systems. Springer-Verlag, 1979.

[CG83] R. A. Cuninghame-Green. The characteristic maxpolynomial of a matrix. *Journal of Mathematical Analysis and Applications*, 95(1):110 – 116, 1983.

[CG94] R.A. Cuninghame-Green. Minimax algebra and applications. volume 90 of *Advances in Imaging and Electron Physics*, pages 1 – 121. Elsevier, 1994.

[CGM80] R. A. Cuninghame-Green and P. F. J. Meijer. An algebra for piecewise-linear minimax problems. *Discrete Appl. Math.*, 2(4):267–294, 1980.

[DTDM14] Fernando De Terán, Froilán M Dopico, and D Steven Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. *Linear Algebra and its Applications*, 459:264–333, 2014.

[EJD88] L. Elsner, Charles R. Johnson, and J.A. Dias da Silva. The perron root of a weighted geometric mean of nonneagative matrices. *Linear and Multilinear Algebra*, 24(1):1–13, 1988.

[EKL06] M. Einsiedler, M. Kapranov, and D. Lind. Non-Archimedean amoebas and tropical varieties. *J. Reine Angew. Math.*, 601:139–157, 2006.

[FLVD04] H.-Y. Fan, W-W. Lin, and P. Van Dooren. Normwise scaling of second order polynomial matrices. *SIAM J. Matrix Anal. Appl.*, 26(1):252–256, 2004.

[Fri86] S Friedland. Limit eigenvalues of nonnegative matrices. *Linear Algebra and Its Applications*, 74:173–178, 1986.

[GHT11] Laurence Grammont, Nicholas J. Higham, and Françoise Tisseur. A framework for analyzing nonlinear eigenproblems and parametrized linear systems. *Linear Algebra and its Applications*, 435(3):623 – 640, 2011. Special Issue: Dedication to Pete Stewart on the occasion of his 70th birthday.

[GK10] E. Gassner and B. Klinz. A fast parametric assignment algorithm with applications in max-algebra. *Networks*, 55(2):61–77, 2010.

[GKZ94] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Birkhäuser, 1994.

[GM84] M. Gondran and M. Minoux. Linear algebra in dioids: a survey of recent results. In *Algebraic and combinatorial methods in operations research*, volume 95 of *North-Holland Math. Stud.*, pages 147–163. North-Holland, Amsterdam, 1984.

[Gra72] R. L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Inf. Proc. Lett.*, 1(4):132–133, 1972.

[GS09] S. Gaubert and M. Sharify. Tropical scaling of polynomial matrices. In Rafael Bru and Sergio Romero-Vivó, editors, *Proceedings of the third Multidisciplinary International Symposium on Positive Systems: Theory and Applications (POSTA 09)*, volume 389 of *LNCIS*, pages 291–303, Valencia, Spain, 2009. Springer. arXiv:0905.0121.

[GT88] H.N. Gabow and R.E. Tarjan. Almost-optimum speed-ups of algorithms for bipartite matching and related problems. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 514–527, New York, 1988. The Association for Computing Machinery.

[Hal98] M. Hall. *Combinatorial Theory*. Wiley Interscience Series in Discrete Mathematics. Wiley, 1998.

[HJ90] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[HLT07] Nicholas J. Higham, Ren-Cang Li, and Françoise Tisseur. Backward error of polynomial eigenproblems solved by linearization. *SIAM J. Matrix Anal. Appl.*, 29(4):1218–1241, 2007.

[HMT06] Nicholas J. Higham, D. Steven Mackey, and Françoise Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.

[HMT13] S. Hammarling, C. J. Munro, and F. Tisseur. An algorithm for the complete solution of quadratic eigenvalue problems. *ACM Trans. Math. Software*, 39(3):Art. 18, 19, 2013.

[Hoo13] J. Hook, 2013. Privately circulated notes.

[HOW06] Bernd Heidergott, Geert Jan Olsder, and Jacob W. van der Woude. *Max Plus at work : modeling and analysis of synchronized systems : a course on Max-Plus algebra and its applications*. Princeton series in applied mathematics. Princeton University Press, Princeton (N.J.), 2006.

[IMS07] I. Itenberg, G. Mikhalkin, and E. Shustin. *Tropical algebraic geometry*. Oberwolfach seminars. Birkhäuser, 2007.

[Kin61] J. F. C. Kingman. A convexity property of positive matrices. *The Quarterly Journal of Mathematics*, 12(1):283–284, 1961.

[LD06] Damien Lemonnier and Paul Van Dooren. Balancing regular matrix pencils. *SIAM Journal on Matrix Analysis and Applications*, 28(1):253–263, 2006.

[Mik04] G. Mikhalkin. Amoebas of algebraic varieties and tropical geometry. In *Different faces of geometry*, volume 3 of *Int. Math. Ser. (N. Y.)*, pages 257–300. Kluwer/Plenum, New York, 2004.

[MM92] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities*. Number vol. 14 in Phoenix Edition Series. Dover Publ., 1992.

[MMMM06] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.

[Mou97] O. Moussa. An inequality about the largest roots of a polynomial. *Applicable Algebra in Engineering, Communication and Computing*, 8:497–504, 1997.

[MS73] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973.

[NST14] V. Noferini, M. Sharify, and F. Tisseur. Tropical roots as approximations to eigenvalues of matrix polynomials. http://eprints.ma.man.ac.uk/2126/01/covered/MIMS_ep2014_16.pdf, to appear in SIAM J. Matrix Anal. Appl., 2014.

[Osb60] E. E. Osborne. On pre-conditioning of matrices. *J. ACM*, 7(4):338–345, October 1960.

[Ost40] A. Ostrowski. Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent. *Acta Math.*, 72:99–155, 1940.

[Pin98] Jean-Éric Pin. Tropical semirings. In Jeremy Gunawardena, editor, *Idempotency*, pages 50–69. Cambridge University Press, 1998.

[PR69] B.N. Parlett and C. Reinsch. Balancing a matrix for calculation of eigenvalues and eigenvectors. *Numerische Mathematik*, 13(4):293–304, 1969.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton mathematical series. Princeton University Press, 1970.

[Sch03] Alexander Schrijver. *Combinatorial Optimization*, volume A. Springer, 2003.

[Sim78] Imre Simon. Limited subsets of a free monoid. In *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*, pages 143–150, Washington, DC, USA, 1978. IEEE Computer Society.

[SS04] D. Speyer and B. Sturmfels. The tropical Grassmannian. *Adv. Geom.*, 4(3):389–411, 2004.

[ST14] M. Sharify and F. Tisseur. Effect of tropical scaling on linearizations of matrix polynomials: backward error and conditioning. t-mims-eprint, inst-MIMS, inst-MIMS:adr, 2014. In preparation.

[Tis00] F. Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309(1-3):339–361, 2000.

[vdDS98] L. van den Dries and P. Speissegger. The real field with convergent generalized power series. *Trans. Amer. Math. Soc.*, 350(11):4377–4421, 1998.

[Vir01] O. Viro. Dequantization of real algebraic geometry on logarithmic paper. In *European Congress of Mathematics, Vol. I (Barcelona, 2000)*, volume 201 of *Progr. Math.*, pages 135–146. Birkhäuser, Basel, 2001.

[War81] Robert C. Ward. Balancing the generalized eigenvalue problem. *SIAM Journal on Scientific and Statistical Computing*, 2(2):141–152, 1981.