

CONDITIONING AND BACKWARD ERROR OF BLOCK-SYMMETRIC BLOCK-TRIDIAGONAL LINEARIZATIONS OF MATRIX POLYNOMIALS

M. I. BUENO*, F. M. DOPICO †, S. FURTADO ‡, AND L. MEDINA §

Abstract. For each square matrix polynomial $P(\lambda)$ of odd degree, a block-symmetric block-tridiagonal pencil $\mathcal{T}_P(\lambda)$, in the family of generalized Fiedler pencils, was introduced by Antoniou and Vologiannidis in 2004, and a variation $\mathcal{R}_P(\lambda)$ of this pencil was introduced by Mackey et al. in 2010. These two pencils have several appealing properties, namely they are always strong linearizations of $P(\lambda)$, they are easy to construct from the coefficients of $P(\lambda)$, the eigenvectors of $P(\lambda)$ can be recovered easily from those of $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$, the two pencils are symmetric (resp. Hermitian) when $P(\lambda)$ is, and they preserve the (classical) sign characteristic of $P(\lambda)$ when $P(\lambda)$ is Hermitian. In this paper we study the numerical behavior of $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$. We compare the conditioning of a finite, nonzero, simple eigenvalue δ of $P(\lambda)$, when considered an eigenvalue of $P(\lambda)$ and an eigenvalue of $\mathcal{T}_P(\lambda)$. We also compare the backward error of an approximate eigenpair (z, δ) of $\mathcal{T}_P(\lambda)$ with the backward error of an approximate eigenpair (x, δ) of $P(\lambda)$, where x was recovered from z in an appropriate way. We show that analogous results are obtained for $\mathcal{R}_P(\lambda)$. When the matrix coefficients of $P(\lambda)$ have similar norms and $P(\lambda)$ is scaled so that the largest norm of the matrix coefficients of $P(\lambda)$ is one, we conclude that $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ have good numerical properties in terms of eigenvalue conditioning and backward error. Moreover, we compare the numerical behavior of $\mathcal{T}_P(\lambda)$ (and $\mathcal{R}_P(\lambda)$) with that of other well-studied linearizations in the literature, namely, the first companion linearization $C_1(\lambda)$, and the first and last linearizations in the standard basis of the vector space $\mathbb{DL}(P)$, introduced by Mackey et al. (2006), and conclude that $\mathcal{T}_P(\lambda)$ performs better than these linearizations when $P(\lambda)$ has odd degree and has been scaled. The theoretical results obtained in the paper are illustrated by numerical experiments, which confirm the advantage of using $\mathcal{T}_P(\lambda)$ (and $\mathcal{R}_P(\lambda)$) over the other linearizations considered in the paper, in particular in symmetric and Hermitian problems.

Key words. Backward error of an approximate eigenpair; block-symmetric generalized Fiedler pencil; conditioning of an eigenvalue; eigenvalue; eigenvector; first companion linearization; standard basis of $\mathbb{DL}(P)$; strong linearization.

AMS subject classifications. 65F15, 65F35, 15A18, 15A22.

1. Introduction. This paper focus on the study of numerical properties of some specific linearizations of a matrix polynomial. We start by introducing some well-known concepts that will be fundamental in the understanding of the problem that we address.

Let $P(\lambda)$ be an $n \times n$ matrix polynomial of grade k with complex coefficients, i.e.,

$$P(\lambda) = \sum_{i=0}^k A_i \lambda^i, \quad A_i \in \mathbb{C}^{n \times n}. \quad (1.1)$$

If $A_k \neq 0$, we say that $P(\lambda)$ has degree k . We say that $P(\lambda)$ is a *regular matrix polynomial* if $\det(P(\lambda))$ is not identically zero. Otherwise, $P(\lambda)$ is said to be *singular*. In this paper we will focus on regular matrix polynomials.

A matrix pencil $L(\lambda) = \lambda L_1 - L_0$, with $L_1, L_0 \in \mathbb{C}^{kn \times kn}$, is a *linearization* of $P(\lambda)$ of grade k (see [13, 14]) if there exist two unimodular matrix polynomials (i.e. matrix polynomials with constant nonzero determinant), $U(\lambda)$ and $V(\lambda)$, such that

$$U(\lambda)L(\lambda)V(\lambda) = \begin{bmatrix} I^{(k-1)n} & 0 \\ 0 & P(\lambda) \end{bmatrix}.$$

*Department of Mathematics and College of Creative Studies, University of California, Santa Barbara, CA 93106, USA (mbueno@math.ucsb.edu). The research of M. I. Bueno was partially supported by NSF grant DMS-1358884 and partially supported by “Ministerio de Economía, Industria y Competitividad of Spain” and “Fondo Europeo de Desarrollo Regional (FEDER) of EU” through grants MTM-2015-68805-REDT and MTM-2015-65798-P (MINECO/FEDER, UE).

†Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es). The research of F. M. Dopico was partially supported by “Ministerio de Economía, Industria y Competitividad of Spain” and “Fondo Europeo de Desarrollo Regional (FEDER) of EU” through grants MTM-2015-68805-REDT and MTM-2015-65798-P (MINECO/FEDER, UE).

‡Centro de Análise Funcional, Estruturas Lineares e Aplicações da Universidade de Lisboa and Faculdade de Economia do Porto, Rua Dr. Roberto Frias 4200-464 Porto, Portugal (sbf@fep.up.pt). The research of S. Furtado was partially supported by project UID/MAT/04721/2013.

§Boston University. One Silber Way. Boston, MA 02215, USA (medina1@bu.edu). The research of L. Medina was partially supported by NSF grant DMS-1358884.

Here and hereafter I_m denotes the $m \times m$ identity matrix. Clearly, a linearization of a regular matrix polynomial is also regular.

We denote by $revP(\lambda)$ the matrix polynomial obtained from $P(\lambda)$ by reversing the order of the matrix coefficients, that is, if $P(\lambda)$ is as in (1.1), then

$$revP(\lambda) := \lambda^k P\left(\frac{1}{\lambda}\right) = \sum_{i=0}^k A_{k-i} \lambda^i.$$

The matrix polynomial $revP(\lambda)$ is known as the *reversal* of $P(\lambda)$. When $L(\lambda)$ is a linearization of $P(\lambda)$ and, in addition, $revL(\lambda)$ is a linearization of $revP(\lambda)$, we say that $L(\lambda)$ is a *strong linearization* of $P(\lambda)$.

Suppose that $P(\lambda)$ is regular. If x and y are nonzero vectors such that $P(\delta)x = 0$ and $y^*P(\delta) = 0$ for some complex number δ , then x and y are said to be, respectively, a *right* and a *left eigenvector* of $P(\lambda)$ associated with the *eigenvalue* δ . The polynomial eigenvalue problem consists in finding the eigenvalues and (right and left) eigenvectors of a given matrix polynomial $P(\lambda)$ and appears in many applications [5, 18, 23, 20, 21, 25, 29], which has motivated an intense research activity on its numerical solution in the last fifteen years [15, 16, 28, 30]. The standard way to solve this problem numerically is to construct a linearization $L(\lambda)$ of $P(\lambda)$ and solve the corresponding generalized eigenproblem $L(\lambda)z = 0$ using well-known algorithms, like the QZ algorithm for moderate size problems [26], or a projection method for large sparse problems [4]. It is clear from the definition of linearization that $P(\lambda)$ and $L(\lambda)$ share the finite elementary divisors and, thus, the finite eigenvalues. If $L(\lambda)$ is a strong linearization of $P(\lambda)$, $P(\lambda)$ and $L(\lambda)$ also share the infinite elementary divisors (that is, $revP(\lambda)$ and $revL(\lambda)$ share the elementary divisors associated with the eigenvalue 0). However, (strong) linearizations of $P(\lambda)$ do not share the eigenvectors of $P(\lambda)$. Thus, for a linearization of $P(\lambda)$ to be useful in solving the polynomial eigenvalue problem for $P(\lambda)$, among other aspects, it is important that the eigenvectors of $P(\lambda)$ can be recovered from the eigenvectors of $L(\lambda)$ in an easy way. Additionally, if $P(\lambda)$ is a structured matrix polynomial (symmetric, Hermitian, palindromic, etc.), it is important that the linearization $L(\lambda)$ shares the same structure so that it preserves numerically, i.e., in the presence of rounding errors, the spectral properties imposed by this structure. In this paper, we will be particularly interested in Hermitian matrix polynomials, which arise very often in applications to model systems. A set of signs, called the sign characteristic, can be associated with the eigenvalues of such type of matrix polynomials. This set of signs is crucial for determining the behavior of the systems described by them, namely, by helping to understand the difference in behavior of the eigenvalues of their eigenvalues under structured and unstructured perturbations. Thus, when solving a polynomial eigenvalue problem using a linearization, it is important to know how the sign characteristic of the linearization and of the matrix polynomial are related [1, 7, 9, 24]. In particular, it may be convenient to use Hermitian linearizations of a Hermitian matrix polynomial that preserve the sign characteristic of $P(\lambda)$.

When calculating the eigenvalues and eigenvectors of a linearization of a matrix polynomial $P(\lambda)$, rounding errors occur. Thus, it is important to know how these errors affect an approximate eigenpair of $P(\lambda)$ obtained from an approximate eigenpair of the linearization. In this context, it is desirable that the condition number of an eigenvalue of a linearization $L(\lambda)$ of $P(\lambda)$ is similar to the condition number of the same eigenvalue of $P(\lambda)$ and, since the algorithms to solve the generalized eigenvalue problem produce small backward errors on the linearization, that the backward error of an eigenpair of $P(\lambda)$ is close to that of $L(\lambda)$. The goal of this paper is to study for the first time the numerical properties of the block-tridiagonal block-symmetric linearizations in the family of generalized Fiedler pencils (GFP) [3] in terms of conditioning of eigenvalues and backward errors of approximate eigenpairs, and to compare their behavior with the one of other well-known linearizations in the literature. We will consider the classical relative normwise condition number of a simple, finite, nonzero eigenvalue of $P(\lambda)$ and the classical normwise backward error of an approximate eigenpair of $P(\lambda)$ (see Section 2 and [28]; see also Section 6). We note that it is not our objective here to provide new linearizations nor to study known linearizations from an algebraic point of view.

In [15, 16], the relative condition number of eigenvalues and the backward error of approximate eigenpairs of the companion linearizations $C_1(\lambda)$ and $C_2(\lambda)$ of a regular matrix polynomial $P(\lambda)$, and of

the linearizations of $P(\lambda)$ in the vector space $\mathbb{DL}(P)$, were studied (for a description of $C_1(\lambda)$ and $C_2(\lambda)$, see [13], and for a description of the pencils in $\mathbb{DL}(P)$, see [17, 20]).

One of the conclusions in [16] is that, given a simple, finite, nonzero eigenvalue δ of $P(\lambda)$ as in (1.1), if A_0 is nonsingular and $|\delta| \geq 1$, then $D_1(\lambda, P)$, the first pencil in the standard basis of $\mathbb{DL}(P)$, has a condition number close to optimal among the linearizations of $P(\lambda)$ in $\mathbb{DL}(P)$, while if A_k is nonsingular and $|\delta| \leq 1$, $D_k(\lambda, P)$, the last pencil in the standard basis of $\mathbb{DL}(P)$, has the same property (assuming that the matrix coefficients of $P(\lambda)$ have similar norms). Moreover, each of these condition numbers is close to that of $P(\lambda)$ for the same eigenvalue. Recall that $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is a linearization of a regular $P(\lambda)$ if and only if A_0 (resp. A_k) is nonsingular ([8, Theorem 5.5] and [20, Theorem 6.7]). Similarly, in [15], it was shown that the linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ have optimal properties with respect to backward errors when $|\delta| \geq 1$ and $|\delta| \leq 1$, respectively, assuming that the matrix coefficients of $P(\lambda)$ have similar norms. We observe that, when $P(\lambda)$ is Hermitian, $D_1(\lambda, P)$ and $D_k(\lambda, P)$ are also Hermitian and have attractive properties regarding the sign characteristic. In fact, in [9] we showed that, if $P(\lambda)$ as in (1.1) is Hermitian and A_k is nonsingular, then $D_k(\lambda, P)$ is a strong linearization of $P(\lambda)$ that preserves the sign characteristic of $P(\lambda)$ associated with the real eigenvalues. If k is odd, then $D_1(\lambda, P)$ preserves the sign characteristic of $P(\lambda)$ as well.

Regarding the Frobenius companion forms $C_1(\lambda)$ and $C_2(\lambda)$, it is shown in [16] that their eigenvalues are potentially more ill conditioned than the same eigenvalues of $P(\lambda)$, though if the spectral norms of the matrix coefficients of $P(\lambda)$ are approximately 1, the eigenvalues of $P(\lambda)$ and of the companion linearizations have similar conditions numbers. An analogous behavior holds with respect to backward errors of approximate eigenpairs of the companion linearizations, when using an appropriate algorithm to recover the eigenpairs of $P(\lambda)$ from those of the linearizations (see [15]).

In this paper, we focus on two block-symmetric block-tridiagonal pencils associated with a matrix polynomial $P(\lambda)$ of odd degree, which we denote by $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$. These pencils were introduced in [3, 22]. When $P(\lambda)$ (regular or singular) has odd degree, these pencils are strong linearizations of $P(\lambda)$ (with no nonsingularity restrictions on the coefficients of $P(\lambda)$). The pencil $\mathcal{T}_P(\lambda)$ is in the family GFP while $\mathcal{R}_P(\lambda)$ is obtained from $\mathcal{T}_P(\lambda)$ by performing some operations on the block-rows and block-columns, namely permutations and multiplications by -1 . These pencils have many relevant properties that may make them very useful in practice. They are companion forms, that is, for $P(\lambda)$ as in (1.1), they are always strong linearizations and their matrix coefficients are block matrices whose blocks are of the form 0 , $\pm I_n$ or $\pm A_i$. Since they are block-symmetric pencils, they are symmetric (Hermitian) when $P(\lambda)$ is. The eigenvectors of $P(\lambda)$ can be easily recovered from the eigenvectors of these two linearizations (see Section 4). Moreover, they preserve the sign characteristic associated with the real eigenvalues of $P(\lambda)$, when $P(\lambda)$ is Hermitian with nonsingular leading coefficient [9, Theorem 5.3].

We will show that, in terms of conditioning and backward errors, $\mathcal{T}_p(\lambda)$ and $\mathcal{R}_p(\lambda)$ also have attractive properties when $P(\lambda)$ has odd degree. More precisely, for a simple, finite, nonzero eigenvalue δ of a regular matrix polynomial $P(\lambda)$ of odd degree k , we study the condition number of δ as an eigenvalue of $\mathcal{T}_P(\lambda)$ and of $\mathcal{R}_P(\lambda)$. We show that, when it is possible to scale $P(\lambda)$ so that all its matrix coefficients have norm approximately equal to 1, the condition number of δ as an eigenvalue of $\mathcal{T}_P(\lambda)$ and of $\mathcal{R}_P(\lambda)$ is comparable to the condition number of δ as an eigenvalue of $P(\lambda)$ regardless of the modulus of δ , which is in stark contrast with the behavior of the linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ described above. Similarly, we study the relationship between the backward error of an approximate eigenpair of $\mathcal{T}_P(\lambda)$ and of $\mathcal{R}_P(\lambda)$ with the backward error of an approximate eigenpair of $P(\lambda)$ associated with the same eigenvalue and show that both backward errors are comparable when the approximate eigenvector in the eigenpair of $P(\lambda)$ is recovered from the approximate eigenvector of the linearization in a convenient way. Due to all their attractive properties, the pencils $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ can be considered one of the most useful linearizations of $P(\lambda)$ known in the literature, when $P(\lambda)$ has odd degree, and, even more, when $P(\lambda)$ is symmetric or Hermitian.

In fact, the pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ have certain disadvantages when used to compute eigenvalues and eigenvectors of a symmetric (Hermitian) regular matrix polynomial $P(\lambda)$. First of all, when A_0 (resp. A_k) is close to be singular, it is unlikely that $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) will exhibit a good numerical behavior since $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is not a linearization of $P(\lambda)$ when A_0 (resp. A_k)

is singular. Additionally, the optimality of their condition number and backward error depends on the modulus of the eigenvalue that needs to be computed which forces the use of the two linearizations when the polynomial has eigenvalues with modulus less than 1 and eigenvalues with modulus larger than 1. These problems can be solved if, instead of using $D_1(\lambda, P)$ and/or $D_k(\lambda, P)$ as linearizations of $P(\lambda)$, we use $C_1(\lambda)$, which is a strong linearization of $P(\lambda)$ (regular and singular). However, the main problem with using $C_1(\lambda)$ is that it is not symmetric (Hermitian) when $P(\lambda)$ is. In contrast with these linearizations, when $P(\lambda)$ has odd degree, the pencil $\mathcal{T}_P(\lambda)$ is symmetric (Hermitian) when $P(\lambda)$ is and it is always a strong linearization of $P(\lambda)$ (regular or singular), i.e. without requiring any conditions on the matrix coefficients of the polynomial. Moreover, as will be shown, it has a numerical behavior similar to $C_1(\lambda)$.

This paper is organized as follows. In Section 2 we introduce the concept of condition number $\kappa_P(\delta)$ of a simple, finite, nonzero eigenvalue δ of a regular matrix polynomial $P(\lambda)$ and present the explicit formula for it. We also recall the definition of backward error $\eta_P(x, \delta)$ of an eigenpair (x, δ) of a matrix polynomial $P(\lambda)$ and provide the explicit formula for it given in [28]. In Section 3 we introduce some auxiliary concepts and results that will be helpful in proving the main results of the paper. In Section 4 we present a block-symmetric block-tridiagonal GFP associated with a matrix polynomial $P(\lambda)$ of odd degree, denoted by $\mathcal{T}_P(\lambda)$, and a block-symmetric block-tridiagonal pencil strictly equivalent to $\mathcal{T}_P(\lambda)$, denoted $\mathcal{R}_P(\lambda)$. In Section 4 we also explain how to construct the eigenvectors of these linearizations of $P(\lambda)$ from eigenvectors of $P(\lambda)$ associated with the same eigenvalues and viceversa. Section 5 is the main section of the paper. We study the conditioning of eigenvalues and backward error of approximate eigenpairs of the linearizations $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$. More precisely, we provide bounds on the ratio $\frac{\kappa_L(\delta)}{\kappa_P(\delta)}$, where δ is a simple, nonzero, finite eigenvalue of $P(\lambda)$, and $L(\lambda)$ denotes any of the linearizations $\mathcal{T}_P(\lambda)$ or $\mathcal{R}_P(\lambda)$. Upper bounds on the ratio $\frac{\eta_P(x, \delta)}{\eta_L(z, \delta)}$ are also presented, where (z, δ) is an approximate eigenpair of $L(\lambda)$ and (x, δ) is an approximate eigenpair of $P(\lambda)$, with x recovered from z in a convenient way to minimize the upper bounds. In Section 6 we recall the definition of the pencils $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$ and improve some results already known in the literature regarding the comparison of the conditioning of eigenvalues and backward error of eigenpairs of these pencils with those of $P(\lambda)$, when they are linearizations of $P(\lambda)$. Finally, in Section 7, we present some numerical experiments that illustrate the theoretical results presented in previous sections and compare the performance of the different linearizations considered in the paper, emphasizing the attractive behavior of $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$.

2. Condition number and backward error of matrix polynomials. In this section we recall the concepts of relative condition number of an eigenvalue and of backward error of an approximate eigenpair of a regular matrix polynomial of degree k as in (1.1).

We note that, in this paper, we only consider simple eigenvalues since these are essentially the only ones appearing in numerical practice. The reason is that they are by far the most common and in the rare occasions where exact multiple eigenvalues are present, they become almost always simple clustered eigenvalues by the effect of rounding errors.

Given a complex vector x , we denote by $\|x\|_2$ the Euclidean norm of x . For $A \in \mathbb{C}^{n \times n}$, we denote by $\|A\|_2$ the spectral norm of A , that is, the matrix norm of A induced by the Euclidean norm.

Let δ be a simple, finite, nonzero eigenvalue of a regular matrix polynomial $P(\lambda)$ of degree k as in (1.1), and let x be a right eigenvector of $P(\lambda)$ associated with δ . A (relative) normwise condition number $\kappa_P(\delta)$ of δ can be defined by

$$\kappa_P(\delta) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\delta|}{\epsilon|\delta|} : [P(\delta + \Delta\delta) + \Delta P(\delta + \Delta\delta)](x + \Delta x) = 0, \right. \\ \left. \|\Delta A_i\|_2 \leq \epsilon \omega_i, i = 0, \dots, k \right\},$$

where $\Delta P(\lambda) = \sum_{i=0}^k \lambda^i \Delta A_i$ and ω_i , $i = 0, \dots, k$, are nonnegative weights that allow flexibility in how the perturbations are measured. This condition number is an immediate generalization of the well-known Wilkinson condition number for the standard eigenvalue problem and measures the relative change in an eigenvalue.

The next theorem gives an explicit formula for this condition number. For a matrix polynomial $P(\lambda)$, we denote by $P'(\lambda)$ the first derivative of $P(\lambda)$ with respect to λ .

THEOREM 2.1. [28, Theorem 5] *Let $P(\lambda)$ be a regular matrix polynomial of degree k . Let δ be a simple, finite, nonzero eigenvalue of $P(\lambda)$, and let x and y be a right and a left eigenvector of $P(\lambda)$ associated with δ . The normwise condition number $\kappa_P(\delta)$ is given by*

$$\kappa_P(\delta) = \frac{(\sum_{i=0}^k |\delta|^i \omega_i) \|y\|_2 \|x\|_2}{|\delta| \|y^* P'(\delta) x|}. \quad (2.1)$$

Note that, since δ is nonzero and simple, the denominator of the expression for $\kappa_P(\delta)$ given in (2.1) is nonzero [2, Theorem 3.2].

In calculating $\kappa_P(\delta)$, we will use the weights $\omega_i = \|A_i\|_2$. We will call these weights *the natural weights* for $P(\lambda)$. In particular, for a pencil $\lambda L_1 - L_0$, the natural weights are $\omega_1 = \|L_1\|_2$ and $\omega_0 = \|L_0\|_2$.

Note that, if $\delta \neq 0$ is an eigenvalue of $P(\lambda)$, then the (left and right) eigenvectors of $P(\lambda)$ and $\text{rev}P(\lambda) = \lambda^k P(\frac{1}{\lambda})$ associated with δ and $\frac{1}{\delta}$, respectively, coincide. Thus, when considering the natural weights for $P(\lambda)$ and $\text{rev}P(\lambda)$, we obtain the following result, which is a simple consequence of Theorem 2.1.

Note that, if $P(\lambda)$ as in (1.1) has degree k and $A_0 \neq 0$, then $\text{rev}P(\lambda)$ has degree k as well.

LEMMA 2.2. *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let δ be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Then*

$$\kappa_P(\delta) = \kappa_{\text{rev}P}\left(\frac{1}{\delta}\right).$$

In order to find a strong linearization $L(\lambda)$ of $P(\lambda)$ such that δ has a comparable condition number when considered as an eigenvalue of $P(\lambda)$ and of $L(\lambda)$, it may be convenient to scale $P(\lambda)$. Let us consider the eigenvalue parameter scaling given by $\tilde{P}(\mu) := \beta P(\gamma\mu)$, where β, γ are nonzero complex scaling parameters. We have

$$\tilde{P}(\mu) = \mu^k \tilde{A}_k + \cdots + \mu \tilde{A}_1 + \tilde{A}_0, \quad (2.2)$$

where $\tilde{A}_i = \beta \gamma^i A_i$, $i = 1, \dots, k$. It is clear that δ is an eigenvalue of $P(\lambda)$ if and only if δ/γ is an eigenvalue of $\tilde{P}(\mu)$. Moreover, $P(\lambda)$ and $\tilde{P}(\mu)$ have the same eigenvectors for the eigenvalues δ and δ/γ , respectively. Thus, by considering the natural weights for $P(\lambda)$ and $\tilde{P}(\mu)$, it follows that $\kappa_P(\delta)$ is invariant under a scaling of the type described above, that is, $\kappa_P(\delta) = \kappa_{\tilde{P}}(\delta/\gamma)$. We will show that this type of scaling can be used on $P(\lambda)$ to improve the bounds of the ratio of the condition numbers of an eigenvalue of $P(\lambda)$ and of a linearization. More explicitly, when the norms of the matrix coefficients of $P(\lambda)$ do not vary too much, we will use the scaling given by $\gamma = 1$ and $\beta^{-1} = \max_{i=0, \dots, k} \{\|A_i\|_2\}$ to improve such bounds. Otherwise, we will consider scalings with $\gamma \neq 1$.

A disadvantage of the condition number $\kappa_P(\delta)$ defined above is that it is not valid for zero or infinite eigenvalues. In [11, Theorem 4.2], Dedieu and Tisseur defined a relative eigenvalue condition number available for all eigenvalues, including 0 and infinity. In this definition, the matrix polynomial $P(\lambda)$ is rewritten in homogeneous form, that is,

$$P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} A_i,$$

and an eigenvalue δ of $P(\lambda)$ is identified with any pair $(\alpha, \beta) \neq (0, 0)$ for which $\delta = \alpha/\beta$. We observe that, contrarily to what happens with the condition number in (2.1), which is invariant under some scalings of the matrix polynomial (a procedure that we apply to improve the condition number of linearizations), the condition number of Dedieu and Tisseur is scale-dependent, that is, it can be changed by a scaling of the original matrix polynomial [30]. Since we anticipate that the ratio of the Dedieu and Tisseur condition

number of a zero or infinite eigenvalue, when considered as an eigenvalue of a matrix polynomial $P(\lambda)$ and of each of the linearizations considered in this paper, may not be good without a previous scaling of $P(\lambda)$, we do not consider such eigenvalues in our study and focus on the condition number in (2.1). In [30] the homogeneous formulation was avoided for similar reasons.

The normwise backward error of an approximate (right) eigenpair (x, δ) of $P(\lambda)$, where δ is finite, is defined by

$$\eta_P(x, \delta) = \min\{\epsilon : (P(\delta) + \Delta P(\delta))x = 0, \quad \|\Delta A_i\|_2 \leq \epsilon \|A_i\|_2, \quad i = 0, \dots, k\},$$

where $\Delta P(\lambda) = \sum_{i=0}^k \lambda^i \Delta A_i$. Similarly, for an approximate left eigenpair (y^*, δ) , we have

$$\eta_P(y^*, \delta) := \min\{\epsilon : y^*(P(\delta) + \Delta P(\delta)) = 0, \quad \|\Delta A_i\|_2 \leq \epsilon \|A_i\|_2, \quad i = 0, \dots, k\}.$$

The following result provides explicit formulas for $\eta_P(x, \delta)$ and $\eta_P(y^*, \delta)$.

THEOREM 2.3. [28, Theorem 1] *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1). For a given approximate right eigenpair (x, δ) of $P(\lambda)$, where $x \in \mathbb{C}^{n \times 1}$ and $\delta \in \mathbb{C}$, the normwise backward error $\eta_P(x, \delta)$ is given by*

$$\eta_P(x, \delta) = \frac{\|P(\delta)x\|_2}{\left(\sum_{i=0}^k |\delta|^i \|A_i\|_2\right) \|x\|_2}.$$

For an approximate left eigenpair (y^*, δ) , where $y \in \mathbb{C}^{n \times 1}$ and $\delta \in \mathbb{C}$, we have

$$\eta_P(y^*, \delta) = \frac{\|y^* P(\delta)\|_2}{\left(\sum_{i=0}^k |\delta|^i \|A_i\|_2\right) \|y\|_2}.$$

The result we present next is a consequence of Theorem 2.3 and can be easily checked.

LEMMA 2.4. *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. For a given approximate right (resp. left) eigenpair (x, δ) (resp. (y^*, δ)) of $P(\lambda)$, where $x, y \in \mathbb{C}^{n \times 1}$ and $0 \neq \delta \in \mathbb{C}$, we have that $(x, \frac{1}{\delta})$ (resp. $(y^*, \frac{1}{\delta})$) is an approximate right (resp. left) eigenpair of $\text{rev}P(\lambda)$ and*

$$\eta_P(x, \delta) = \eta_{\text{rev}P}\left(x, \frac{1}{\delta}\right) \quad \text{and} \quad \eta_P(y^*, \delta) = \eta_{\text{rev}P}\left(y^*, \frac{1}{\delta}\right).$$

A scaling on the eigenvalue parameter can also be used to improve the normwise backward error of an approximate eigenpair of a linearization of a matrix polynomial. It is easy to show that $\eta_{\tilde{P}}(x, \mu) = \eta_P(x, \gamma\mu)$, where $\tilde{P}(\mu)$ is as in (2.2), while the corresponding backward errors of the (same) linearizations of $P(\lambda)$ and $\tilde{P}(\mu)$ can be quite different. We will use this fact in our numerical experiments to show that a scaling of the polynomial can be applied to decrease the backward error of the block-symmetric linearizations that we are studying in this paper.

REMARK 2.5. *To compare the conditioning of eigenvalues and backward error of approximate eigenpairs of a matrix polynomial $P(\lambda)$ with those of the linearizations considered in this paper, it will be convenient to assume that $P(\lambda)$ as in (1.1) has a nonzero constant term A_0 . However, if $A_0 = 0$, we have that $P(\lambda) = \lambda^s P_1(\lambda)$ for some s , where the constant term of $P_1(\lambda)$ is nonzero, and we may consider $P_1(\lambda)$ instead of $P(\lambda)$ to compute the nonzero eigenvalues of $P(\lambda)$. Note that $P(\lambda)$ and $P_1(\lambda)$ have the same nonzero finite eigenvalues.*

The sign characteristic of a Hermitian matrix polynomial can be obtained from the sign characteristic of one of its Hermitian linearizations, as long as we know how they are related. In [7, 9], the authors considered the classical definition of sign characteristic of a Hermitian matrix polynomial $P(\lambda)$ with nonsingular leading coefficient, presented in [13, 14], and studied the (classical) sign characteristic of its linearizations in the vector space $\mathbb{DL}(P)$ and in the family of block-symmetric generalized Fiedler pencils with repetition (introduced in [8]), as well as of the linearizations $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$, to be introduced in

Section 4 (see also the Introduction). Besides these papers and [1], where some partial results on the sign characteristic of linearizations in $\mathbb{DL}(P)$ are given, the authors do not know of any other papers in the literature studying the sign characteristic of specific linearizations of matrix polynomials. (We observe that a generalization of the concept of sign characteristic to general matrix polynomials was published recently [24] but we are not aware of any results regarding the sign characteristic of Hermitian linearizations, using this more general definition). Since a scaling of $P(\lambda)$ may be used to improve conditioning or backward errors of its linearizations, we note in the following theorem that the procedure of scaling a Hermitian $P(\lambda)$ with nonsingular leading coefficient preserves the sign characteristic of $P(\lambda)$, when the scaling parameters β and γ are positive real numbers. Based on this result, the (classical) sign characteristic of the original matrix polynomial can be obtained from the sign characteristic of a linearization $L(\mu)$ of the scaled matrix polynomial $\tilde{P}(\mu)$, as long as we know how the sign characteristic of $\tilde{P}(\mu)$ and $L(\mu)$ are related. The proof of the next theorem appears in Appendix A.

THEOREM 2.6. *Let $P(\lambda)$ be a Hermitian matrix polynomial of degree k as in (1.1), with A_k nonsingular, and let $\tilde{P}(\mu) := \beta P(\gamma\mu)$, where β and γ are positive real numbers. Then, $P(\lambda)$ and $\tilde{P}(\mu)$ have the same (classical) sign characteristic.*

3. Some definitions and auxiliary results. We introduce some concepts and technical results that will be used in the proofs of our main results.

If a and b are two positive integers such that $a \leq b$, we define

$$a : b := a, a + 1, \dots, b.$$

The following result is an immediate consequence of the Cauchy-Schwarz inequality when the standard inner product is considered in \mathbb{C}^n .

LEMMA 3.1. *Let m be a positive integer and let a be a positive real number. Then,*

$$\left(\sum_{j=0}^m a^j \right)^2 \leq (m+1) \sum_{j=0}^m a^{2j}.$$

The following property is well known (see Lemma 3.5 in [15] for a proof of the second inequality).

PROPOSITION 3.2. *For any complex $l \times m$ block-matrix $B = (B_{ij})$ we have*

$$\max_{i,j} \|B_{ij}\|_2 \leq \|B\|_2 \leq \sqrt{lm} \max_{i,j} \|B_{ij}\|_2. \quad (3.1)$$

Given a matrix polynomial $P(\lambda)$ of degree k as in (1.1), the i th Horner shift of $P(\lambda)$, $i = 0 : k$, is given by

$$P_i(\lambda) := \lambda^i A_k + \lambda^{i-1} A_{k-1} + \dots + \lambda A_{k-i+1} + A_{k-i}. \quad (3.2)$$

Notice that $P_0(\lambda) = A_k$, $P_k(\lambda) = P(\lambda)$, and

$$P_{i+1}(\lambda) - A_{k-i-1} = \lambda P_i(\lambda), \quad i = 0 : k-1. \quad (3.3)$$

When convenient, we write P_i to denote $P_i(\lambda)$. We also denote

$$P^i(\lambda) := \lambda^i A_i + \dots + \lambda A_1 + A_0, \quad i = 0 : k. \quad (3.4)$$

Notice that $P^0(\lambda) = A_0$ and $P^k(\lambda) = P(\lambda)$.

LEMMA 3.3. *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1). Let $P_i(\lambda)$ and $P^i(\lambda)$, $i = 0 : k$, be the matrix polynomials defined in (3.2) and (3.4). Let δ be a nonzero, finite eigenvalue of $P(\lambda)$, and let x and y be, respectively, a right and left eigenvector of $P(\lambda)$ associated with δ . Then,*

$$P_i(\delta)x = -\delta^{i-k} P^{k-i-1}(\delta)x, \quad \text{and} \quad y^* P_i(\delta) = -\delta^{i-k} y^* P^{k-i-1}(\delta), \quad i = 0 : k-1.$$

Proof. Note that, for $i = 0 : k - 1$, we have $P(\delta) = \delta^{k-i}P_i(\delta) + P^{k-i-1}(\delta)$. Thus, the result follows taking into account that δ is nonzero, $P(\delta)x = 0$, and $y^*P(\delta) = 0$, since x and y are, respectively, a right and a left eigenvector of $P(\lambda)$ associated with δ . \square

The next lemma can be easily verified.

LEMMA 3.4. *Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1), let $\delta \in \mathbb{C}$, and let $P_i(\lambda)$ and $P^i(\lambda)$, $i = 0 : k$, be the matrix polynomials defined in (3.2) and (3.4). Then,*

$$\|P_i(\delta)\|_2, \|P^i(\delta)\|_2 \leq \max_{j=0:k} \{\|A_j\|_2\} \sum_{j=0}^i |\delta|^j, \quad i = 0 : k.$$

We close this section with two combinatorial lemmas that will be used later in the proofs of our main results.

LEMMA 3.5. *Let $P(\lambda)$ be an $n \times n$ matrix polynomial of odd degree k as in (1.1), let $P_i(\lambda)$, $i = 0 : k$, be the Horner shifts defined in (3.2), and let*

$$\Delta^{\mathcal{B}}(\lambda) := [\lambda^{\frac{k-1}{2}} I_n, \lambda^{\frac{k-1}{2}} P_1, \lambda^{\frac{k-3}{2}} I_n, \lambda^{\frac{k-3}{2}} P_3, \dots, \lambda I_n, \lambda P_{k-2}, I_n], \quad (3.5)$$

where $\Delta^{\mathcal{B}}(\lambda)$ denotes the block-transpose of $\Delta(\lambda)$ when viewed as a $k \times 1$ block-matrix whose blocks are $n \times n$. Let $\delta \in \mathbb{C}$. Then,

$$\|\Delta(\delta)\|_2 \leq \sqrt{d_1(\delta)} \max_{i=0:k} \{1, \|A_i\|_2\}, \quad (3.6)$$

where

$$d_1(\delta) = \sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\delta|^{2s} \right). \quad (3.7)$$

Proof. Let $\delta \in \mathbb{C}$ and $0 \neq x \in \mathbb{C}^n$. Taking into account the definition of $\Delta(\lambda)$, we get

$$\|\Delta(\delta)x\|_2^2 = \sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r} \|x\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\delta|^{2r} \|P_{k-2r}(\delta)x\|_2^2 \quad (3.8)$$

$$\leq \left[\sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} |\delta|^{2r} \|P_{k-2r}(\delta)\|_2^2 \right] \|x\|_2^2. \quad (3.9)$$

By Lemmas 3.4 and 3.1, we obtain

$$\begin{aligned} \left(\frac{\|\Delta(\delta)x\|_2}{\|x\|_2} \right)^2 &\leq \sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} |\delta|^{2r} \left(\max_{i=0:k} \{\|A_i\|_2\} \sum_{s=0}^{k-2r} |\delta|^s \right)^2 \\ &\leq \left[\sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\delta|^{2s} \right) \right] \max_{i=0:k} \{1, \|A_i\|_2\}^2 \\ &= d_1(\delta) \max_{i=0:k} \{\|A_i\|_2, 1\}^2. \end{aligned}$$

Since $\|\Delta(\delta)\|_2 = \sup_{x \neq 0} \frac{\|\Delta(\delta)x\|_2}{\|x\|_2}$, (3.6) follows. \square

LEMMA 3.6. *Let $\delta \in \mathbb{C}$ be nonzero and let $k \geq 3$ be a positive odd integer. Let $d_1(\lambda)$ be as in (3.7). If $|\delta| \leq 1$, then*

$$d_1(\delta) \leq \frac{k+1}{2} + \frac{(k-1)^3}{2} |\delta|^2. \quad (3.10)$$

Proof. Assume that $|\delta| < 1$. Then,

$$\begin{aligned} \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\delta|^{2s} \right) &\leq \frac{(k-1)^2}{2} \sum_{s=1}^{k-1} |\delta|^{2s} \\ &= \frac{(k-1)^2}{2} |\delta|^2 \sum_{s=1}^{k-1} |\delta|^{2(s-1)} \leq \frac{(k-1)^3}{2} |\delta|^2. \end{aligned}$$

Thus, from (3.7), (3.10) follows. The case $|\delta| = 1$ follows by using a continuity argument. \square

4. Block-symmetric block-tridiagonal generalized Fiedler pencils. In this section we present the block-symmetric linearizations of a matrix polynomial $P(\lambda)$ that are the focus of this paper. Here and in the next section we only consider matrix polynomials $P(\lambda)$ of odd degree k and we assume that $k > 2$.

The family of generalized Fiedler pencils (GFP) associated with a matrix polynomial $P(\lambda)$ was introduced in [3, 10]. The following block-symmetric block-tridiagonal GFP associated with an $n \times n$ $P(\lambda)$ of odd degree k was presented in [3, Theorem 3.1]:

$$\mathcal{T}_P(\lambda) = \begin{bmatrix} \lambda A_k + A_{k-1} & -I_n & & & & & & 0 \\ & -I_n & 0 & \lambda I_n & & & & \\ & & \lambda I_n & \lambda A_{k-2} + A_{k-3} & -I_n & & & \\ & & & -I_n & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & \lambda A_3 + A_2 & -I_n \\ & & & & & & -I_n & 0 & \lambda I_n \\ & & & & & & & \lambda I_n & \lambda A_1 + A_0 \\ 0 & & & & & & & & & 0 \end{bmatrix}. \quad (4.1)$$

Also, the following block-symmetric block-tridiagonal pencil associated with an $n \times n$ $P(\lambda)$ of odd degree k was introduced in [22, formula (5.1)] :

$$\mathcal{R}_P(\lambda) := S \mathcal{T}_P(\lambda) R S \quad (4.2)$$

where

$$R := \begin{bmatrix} & & I_n \\ & \ddots & \\ I_n & & \end{bmatrix} \in \mathbb{C}^{nk \times nk} \quad (4.3)$$

and S is the $k \times k$ block-diagonal matrix whose (i, i) th block-entry is given by

$$S(i, i) = \begin{cases} -I_n, & \text{if } i \equiv 0, 1 \pmod{4}, \\ I_n, & \text{otherwise.} \end{cases} \quad (4.4)$$

It is well-known [3, 22] that, for any matrix polynomial $P(\lambda)$ of odd degree k , the pencils $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ are strong linearizations of $P(\lambda)$. As mentioned in the introduction, these linearizations have several attractive properties. In particular, it is easy to recover an eigenvector of $P(\lambda)$ associated with an eigenvalue δ from an eigenvector of the linearizations associated with the same eigenvalue, as we show next. We start with a technical lemma that will be useful for this purpose.

Here and in the next sections, we denote by e_i the i th column of the identity matrix of appropriate size for the context.

LEMMA 4.1. *Let $P(\lambda)$ be a matrix polynomial of odd degree k as in (1.1) and let $\mathcal{T}_P(\lambda)$ be as in (4.1). Then,*

$$\mathcal{T}_P(\lambda) \Delta(\lambda) = e_k \otimes P(\lambda), \quad \text{and} \quad \Delta^{\mathcal{B}}(\lambda) \mathcal{T}_P(\lambda) = e_k^T \otimes P(\lambda), \quad (4.5)$$

where $\Delta^{\mathcal{B}}(\lambda)$ is as in (3.5).

Proof. Let $\mathcal{T}_P(\lambda) := \lambda\mathcal{T}_1 - \mathcal{T}_0$. Using (3.3), a direct computation shows that

$$\mathcal{T}_1\Delta(\lambda) = \begin{bmatrix} \lambda^{\frac{k-1}{2}}P_0 \\ \lambda^{\frac{k-1}{2}-1}I_n \\ \lambda^{\frac{k-1}{2}-1}P_2 \\ \lambda^{\frac{k-1}{2}-2}I_n \\ \vdots \\ \lambda P_{k-3} \\ I_n \\ P_{k-1} \end{bmatrix} \quad \text{and} \quad \mathcal{T}_0\Delta(\lambda) = \begin{bmatrix} \lambda^{\frac{k+1}{2}}P_0 \\ \lambda^{\frac{k-1}{2}}I_n \\ \lambda^{\frac{k-1}{2}}P_2 \\ \lambda^{\frac{k-1}{2}-1}I_n \\ \vdots \\ \lambda^2P_{k-3} \\ \lambda I_n \\ -A_0 \end{bmatrix},$$

where P_i , $i = 0 : k$, are the Horner shifts defined in (3.2).

Taking into account (3.3) again and the fact that $P_k(\lambda) = P(\lambda)$, the first claim in (4.5) follows. The second claim in (4.5) follows easily from the first claim by noting that $(\mathcal{T}_P(\lambda)\Delta(\lambda))^{\mathcal{B}} = \Delta^{\mathcal{B}}(\lambda)\mathcal{T}_P(\lambda)$, as the i th block-row of $\mathcal{T}_P(\lambda)$, with i even, just contains blocks of the form 0 , I_n , and λI_n , and this type of blocks commute with P_j . \square

Note that the equations in (4.5) are a particular case of equation (2.11) in [15] (where the homogeneous approach is considered).

The following theorem follows from Lemma 4.1 using arguments similar to those in the proof of Theorem 3.8 in [20].

THEOREM 4.2. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1). Assume that δ is a finite eigenvalue of $P(\lambda)$. Let $\Delta(\lambda)$ be as in (3.5). A vector z is a right (resp. left) eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ if and only if $z = \Delta(\delta)x$ (resp. $z = \overline{\Delta(\delta)}y$), for some right (resp. left) eigenvector x (resp. y) of $P(\lambda)$ associated with δ .*

Proof. Taking into account Lemma 4.1, we have

$$\mathcal{T}_P(\lambda)\Delta(\lambda)x = (e_k \otimes P(\lambda))x = e_k \otimes P(\lambda)x.$$

Clearly, if $\{x_1, \dots, x_m\}$ is a basis of the eigenspace associated with the eigenvalue δ of $P(\lambda)$, then $\{\Delta(\delta)x_1, \dots, \Delta(\delta)x_k\}$ is linearly independent and, since δ has the same geometric multiplicity as an eigenvalue of $P(\lambda)$ and as eigenvalue of $\mathcal{T}_P(\lambda)$, because $\mathcal{T}_P(\lambda)$ is a linearization of $P(\lambda)$, then $\{\Delta(\delta)x_1, \dots, \Delta(\delta)x_k\}$ is a basis of the eigenspace associated with the eigenvalue δ of $\mathcal{T}_P(\lambda)$. Thus, a vector z is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ if and only if it is a linear combination of $\Delta(\delta)x_1, \dots, \Delta(\delta)x_k$, that is, it is of the form $\Delta(\delta)x$ for some eigenvector x of $P(\lambda)$ associated with δ .

The proof for left eigenvectors is similar. \square

Let R be as in (4.3), S be as in (4.4), and $\Delta(\lambda)$ be as in (3.5). From (4.2) and Lemma 4.1, we obtain

$$\mathcal{R}_P(\lambda)(SR\Delta(\lambda)) = (SRe_k) \otimes P(\lambda). \quad (4.6)$$

Note that $R^2 = I$ and $S^2 = I$. Based on this fact, we give next an expression for the right and left eigenvectors of $\mathcal{R}_P(\lambda)$.

THEOREM 4.3. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1). Assume that δ is a finite eigenvalue of $P(\lambda)$. Let $\Delta(\lambda)$ be as in (3.5). A vector z is a right (resp. left) eigenvector of $\mathcal{R}_P(\lambda)$ associated with δ if and only if $z = SR\Delta(\delta)x$ (resp. $z = SR\overline{\Delta(\delta)}y$), for some right (resp. left) eigenvector x (resp. y) of $P(\lambda)$ associated with δ .*

We close this section with two results that will help us studying the numerical performance of the linearizations $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ in the next section. The next lemma, which can be easily verified, allows us to focus on eigenvalues of modulus not greater than 1.

LEMMA 4.4. *Let $P(\lambda)$ be a matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$. Then,*

$$\mathcal{T}_P(\lambda) = \lambda D R T_{revP} \begin{pmatrix} 1 \\ \lambda \end{pmatrix} R D,$$

where R is as in (4.3) and

$$D = \text{diag}(I_n, -I_n, I_n, -I_n, \dots, I_n). \quad (4.7)$$

Moreover, $z \in \mathbb{C}^{nk}$ is a right (resp. left) eigenvector of $\mathcal{T}_P(\lambda)$ associated with a nonzero eigenvalue δ if and only if RDz is a right (resp. left) eigenvector of $\mathcal{T}_{revP}(\lambda)$ associated with the eigenvalue $\frac{1}{\delta}$.

The following result gives an upper bound on the spectral norm of the matrix coefficients of $\mathcal{T}_P(\lambda)$, improving the one obtained by using Proposition 3.2.

PROPOSITION 4.5. *Let $P(\lambda) = \sum_{i=0}^k A_i \lambda^i$ be an $n \times n$ matrix polynomial of odd degree k and $\mathcal{T}_P(\lambda) := \lambda \mathcal{T}_1 - \mathcal{T}_0$ be as in (4.1). Then,*

$$\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2 \leq 2 \max_{i=0:k} \{1, \|A_i\|_2\}.$$

Proof. Let $z = [z_1, \dots, z_k]^B$ be a nonzero $k \times 1$ block-vector partitioned into $n \times 1$ blocks. Then, defining $z_0 := 0$, we have

$$\begin{aligned} \|\mathcal{T}_1 z\|_2^2 &= \sum_{i=1}^{\frac{k-1}{2}} \|z_{2i+1}\|_2^2 + \sum_{i=0}^{\frac{k-1}{2}} \|z_{2i} + A_{k-2i} z_{2i+1}\|_2^2 \\ &\leq \sum_{i=1}^{\frac{k-1}{2}} \|z_{2i+1}\|_2^2 + \sum_{i=0}^{\frac{k-1}{2}} (\|z_{2i}\|_2^2 + \|A_{k-2i} z_{2i+1}\|_2^2 + 2\|z_{2i}\|_2 \|A_{k-2i} z_{2i+1}\|_2) \\ &\leq \max_{i=0:k} \{1, \|A_i\|_2^2\} \left(\sum_{i=1}^k \|z_i\|_2^2 + \sum_{i=0}^{\frac{k-1}{2}} (\|z_{2i+1}\|_2^2 + 2 \max\{\|z_{2i}\|_2^2, \|z_{2i+1}\|_2^2\}) \right) \\ &\leq \max_{i=0:k} \{1, \|A_i\|_2^2\} (2\|z\|_2^2 + 2 \sum_{i=0}^{\frac{k-1}{2}} \max\{\|z_{2i}\|_2^2, \|z_{2i+1}\|_2^2\}) \leq 4 \max_{i=0:k} \{1, \|A_i\|_2^2\} \|z\|_2^2. \end{aligned}$$

The proof for \mathcal{T}_0 is analogous. \square

5. Conditioning and backward error of $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$. Let $P(\lambda)$ be a matrix polynomial of odd degree k with $A_0 \neq 0$ and let $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ be the linearizations of $P(\lambda)$ defined in (4.1) and (4.2), respectively. In this section we present the main results in this paper, Theorems 5.1 and 5.3, concerned with the comparison of the conditioning of eigenvalues and the backward error of approximate eigenpairs of $\mathcal{T}_P(\lambda)$ with, respectively, the conditioning of eigenvalues and the backward error of approximate eigenpairs of $P(\lambda)$. As will follow from Remarks 5.2 and 5.4, Theorems 5.1 and 5.3 apply to $\mathcal{R}_P(\lambda)$ as well.

We next state Theorems 5.1 and 5.3, which will be proven in Sections 5.1 and 5.2, respectively. We use the notation for $P(\lambda)$ in (1.1) and we denote

$$\rho_1 := \frac{\max_{i=0:k} \{1, \|A_i\|_2^3\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}, \quad \rho_2 := \frac{\min\{\max\{1, \|A_k\|_2\}, \max\{1, \|A_0\|_2\}\}}{\max_{i=0:k} \{\|A_i\|_2\}} \quad (5.1)$$

and

$$\rho' := \frac{\max_{i=0:k} \{1, \|A_i\|_2^2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}. \quad (5.2)$$

THEOREM 5.1. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$. Assume that δ is a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let ρ_1 and ρ_2 be as in (5.1). Then,*

$$\rho_2 \leq \frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)} \leq 2k^3 \rho_1. \quad (5.3)$$

Moreover, if $|\delta| \neq 1$ and $\min\{|\delta|, \frac{1}{|\delta|}\} \leq \frac{1}{k-1}$, then

$$\frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)} \leq 4(k+1)\rho_1.$$

REMARK 5.2. Taking into account the definition of $\mathcal{R}_P(\lambda)$ and Theorem 4.3, it follows easily from Theorem 2.1 applied to $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ that $\kappa_{\mathcal{T}_P}(\delta) = \kappa_{\mathcal{R}_P}(\delta)$. Recall that the spectral norm is unitarily invariant. Thus, Theorem 5.3 holds for $\mathcal{R}_P(\lambda)$ as well.

From Theorem 5.1 we can conclude that, if the norms of the matrix coefficients of $P(\lambda)$ have similar magnitudes and $\max_{i=0:k}\{\|A_i\|_2\} = 1$ (which can be obtained by scaling $P(\lambda)$ as explained in Section 2), the condition number of any simple, nonzero, finite eigenvalue of $P(\lambda)$ is close to the condition number of the same eigenvalue of $\mathcal{T}_P(\lambda)$ with no restriction on the modulus of δ , in contrast with what happens to $D_1(\lambda, P)$ and $D_k(\lambda, P)$.

Since, for not extremely large values of nk , the algorithm QZ, combined with adequate methods for computing eigenvectors [26], is used to compute the eigenvalues and eigenvectors of a linearization of $P(\lambda)$ and this algorithm produces small backward errors of order unit-roundoff, if we prove that $\eta_P(x, \delta)$ is not much larger than $\eta_{\mathcal{T}_P}(z, \delta)$ (where (z, δ) denotes an approximate eigenpair of \mathcal{T}_P and (x, δ) denotes an approximate eigenpair of P obtained from (z, δ) in an appropriate way), we ensure small backward errors for the approximate eigenpairs of $P(\lambda)$ as well. This motivates the next theorem, in which an upper bound for the ratio $\frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)}$ is established.

THEOREM 5.3. Let $P(\lambda)$ be a matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$. Let (z, δ) be an approximate right eigenpair of $\mathcal{T}_P(\lambda)$, $x = (e_1^T \otimes I_n)z$ if $|\delta| > 1$ and $x = (e_k^T \otimes I_n)z$ if $|\delta| \leq 1$. Then, (x, δ) is an approximate right eigenpair of $P(\lambda)$ and

$$\frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)} \leq 4k^{\frac{3}{2}} \frac{\|z\|_2}{\|x\|_2} \rho'. \quad (5.4)$$

Moreover, if $|\delta| \neq 1$ and $\min\{|\delta|, \frac{1}{|\delta|}\} \leq \frac{1}{k-1}$, then

$$\frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)} \leq 4\sqrt{k+1} \frac{\|z\|_2}{\|x\|_2} \rho'. \quad (5.5)$$

REMARK 5.4. Theorem 5.3 also holds for $\mathcal{R}_P(\lambda)$ taking into consideration the following observations. Let (z, δ) be an approximate eigenpair of $\mathcal{R}_P(\lambda)$. It is easy to see that

$$\eta_{\mathcal{R}_P}(z, \delta) = \eta_{\mathcal{T}_P}(RSz, \delta).$$

Since, for $t \in \{1, k\}$, $(e_t^T \otimes I_n)z = \pm(e_{k-t+1}^T \otimes I_n)RSz := \pm x_t$, we have

$$\frac{\eta_P((e_t^T \otimes I_n)z, \delta)}{\eta_{\mathcal{R}_P}(z, \delta)} = \frac{\eta_P(x_t, \delta)}{\eta_{\mathcal{T}_P}(RSz, \delta)}.$$

Thus, it is clear that, if proven for $\mathcal{T}_P(\lambda)$, Theorem 5.3 also holds for $\mathcal{R}_P(\lambda)$.

In Theorem 5.3 we considered right eigenpairs of $P(\lambda)$ and $\mathcal{T}_P(\lambda)$. The case of left eigenpairs can be easily reduced to the previous case. Indeed, (y^*, δ) is a left eigenpair of $P(\lambda)$ if and only if $(y, \bar{\delta})$ is a right eigenpair of $P^*(\lambda)$. Similarly, (z^*, δ) is a left eigenpair of $\mathcal{T}_P(\lambda)$ if and only if $(z, \bar{\delta})$ is a right eigenpair of $\mathcal{T}_{P^*}(\lambda) = \mathcal{T}_{P^*}(\lambda)$. Thus, it can be easily seen that

$$\eta_P(y^*, \delta) = \eta_{P^*}(y, \bar{\delta}) \quad \text{and} \quad \eta_{\mathcal{T}_P}(z^*, \delta) = \eta_{\mathcal{T}_{P^*}}(z, \bar{\delta}),$$

and the result for approximate left eigenpairs follows from Theorem 5.3 applied to $P^*(\lambda)$ and $\mathcal{T}_{P^*}(\lambda)$.

Note that Theorems 5.1 and 5.3 hold with no nonsingularity restrictions on the matrix coefficients of $P(\lambda)$, in contrast with the analogous results for the block-symmetric linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ (see [15, 16] and Section 6).

REMARK 5.5. From Theorem 4.2, x' is an eigenvector of $P(\lambda)$ associated with the eigenvalue δ if and only if $z = \Delta(\delta)x'$ is an eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ . As in the proof of Lemma 3.5, for $z = \Delta(\delta)x'$ (which implies $x' = (e_k^T \otimes I_n)z$), we obtain

$$\frac{\|z\|_2^2}{\|x'\|_2^2} \leq d_1(\delta) \max_{i=0:k} \{\|A_i\|_2, 1\}^2,$$

where $d_1(\delta)$ is as in (3.7). If $|\delta| \leq 1$, from Lemma 3.6, we have

$$\frac{\|z\|_2^2}{\|x'\|_2^2} \leq \left(\frac{k+1}{2} + \frac{(k-1)^3}{2} \right) \max_{i=0:k} \{1, \|A_i\|_2\}^2,$$

implying

$$\frac{\|z\|_2}{\|x'\|_2} \leq \frac{1}{2} k^{\frac{3}{2}} \max_{i=0:k} \{1, \|A_i\|_2\}.$$

If $|\delta| > 1$, taking into account Lemma 4.4, the eigenvectors of $\mathcal{T}_{revP}(\lambda)$ associated with the eigenvalue $\frac{1}{\delta}$ are of the form RDz . From Theorem 4.2, $(e_1^T \otimes I_n)RDz$ is an eigenvector of $revP(\lambda)$ associated with the eigenvalue $\frac{1}{\delta}$ and, from the calculations above, we also have

$$\frac{\|z\|_2}{\|(e_1^T \otimes I_n)z\|_2} = \frac{\|RDz\|_2}{\|(e_k^T \otimes I_n)RDz\|_2} \leq \frac{1}{2} k^{\frac{3}{2}} \max_{i=0:k} \{1, \|A_i\|_2\}.$$

Thus, if $P(\lambda)$ is scaled so that $\max_{i=0:k} \{\|A_i\|_2\} = 1$, we obtain

$$\frac{\|z\|_2}{\|x\|_2} \lesssim \frac{1}{2} k^{\frac{3}{2}},$$

where $x = (e_k^T \otimes I_n)z$ if $|\delta| \leq 1$ and $x = (e_1^T \otimes I_n)z$ if $|\delta| > 1$. If we assume that an approximate eigenvector z has a block structure similar to that of an eigenvector of $\mathcal{T}_P(\lambda)$, then we can expect that the bound above for the quotient $\frac{\|z\|_2}{\|x\|_2}$ appearing in Theorem 5.3 still holds and it is close to 1 for moderate k .

Note that, when the matrix coefficients of $P(\lambda)$ have similar norms and $P(\lambda)$ is scaled so that $\max_{i=0:k} \{\|A_i\|_2\} = 1$, then $\rho' \approx 1$ and the bound of the quotient of backward errors in Theorem 5.3 is expected to depend only on k .

In the rest of this section we use the notation introduced in Section 4. In particular, $\mathcal{T}_P(\lambda) := \lambda\mathcal{T}_1 - \mathcal{T}_0$ denotes the block symmetric pencil defined in (4.1).

5.1. Proof of Theorem 5.1. We start with a lemma in which we give an explicit expression for the condition number of a simple, finite, nonzero eigenvalue δ of the linearization $\mathcal{T}_P(\lambda)$ of a matrix polynomial $P(\lambda)$ of odd degree.

LEMMA 5.6. Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1). Assume that δ is a simple, finite, nonzero eigenvalue of $P(\lambda)$ with left and right eigenvectors x_1 and x_2 , respectively. Let $\mathcal{T}_P(\lambda) := \lambda\mathcal{T}_1 - \mathcal{T}_0$. Then,

$$\kappa_{\mathcal{T}_P}(\delta) = \frac{(|\delta|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)\|\Delta(\delta)x_1\|_2\|\Delta(\delta)x_2\|_2}{|\delta|\|x_1^*P'(\lambda)x_2|}, \quad (5.6)$$

where $\Delta(\lambda)$ is defined as in (3.5).

Proof. Differentiating the first equality in (4.5), we get

$$\mathcal{T}'(\lambda)\Delta(\lambda) + \mathcal{T}(\lambda)\Delta'(\lambda) = e_k \otimes P'(\lambda). \quad (5.7)$$

By Theorem 4.2, the vector $z_1 = \overline{\Delta}(\delta)x_1$ is a left eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ and $z_2 = \Delta(\delta)x_2$ is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ . Evaluating the expression (5.7) at δ , premultiplying by z_1^* , and postmultiplying by x_2 , we get

$$z_1^*\mathcal{T}'(\delta)(\Delta(\delta)x_2) = z_1^*(e_k \otimes P'(\delta)x_2),$$

or, equivalently,

$$z_1^* \mathcal{T}'(\delta) z_2 = x_1^* \Delta^T(\delta) (e_k \otimes P'(\delta) x_2) = x_1^* P'(\delta) x_2.$$

Now (5.6) follows from Theorem 2.1. \square

The following lemma will allow us to only consider eigenvalues δ such that $|\delta| \leq 1$ when proving Theorem 5.1.

LEMMA 5.7. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$, and let δ be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Then,*

$$\kappa_{\mathcal{T}_P}(\delta) = \kappa_{\mathcal{T}_{revP}}\left(\frac{1}{\delta}\right).$$

Proof. Let $\mathcal{T}_P(\lambda) := \lambda \mathcal{T}_1 - \mathcal{T}_0$. Then, from Lemma 4.4,

$$\mathcal{T}_{revP}(\lambda) = \lambda R D \mathcal{T}_P\left(\frac{1}{\lambda}\right) D R = R D (\mathcal{T}_1 - \lambda \mathcal{T}_0) D R =: \lambda \tilde{\mathcal{T}}_1 - \tilde{\mathcal{T}}_0. \quad (5.8)$$

Moreover, if y and x are, respectively, a left and a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with δ , then $\tilde{y} = R D y$ and $\tilde{x} = R D x$ are, respectively, a left and a right eigenvector of $\mathcal{T}_{revP}(\lambda)$ associated with $1/\delta$. Then, taking into account Theorem 2.1 and the fact that the Frobenius norm and the spectral norm are unitarily invariant, we have

$$\begin{aligned} \kappa_{\mathcal{T}_{revP}}\left(\frac{1}{\delta}\right) &= \frac{(\frac{1}{|\delta|} \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2) \|\tilde{y}\|_2 \|\tilde{x}\|_2}{|\frac{1}{\delta}| |\tilde{y}^* \tilde{\mathcal{T}}_1 \tilde{x}|} \\ &= \frac{(\|\tilde{\mathcal{T}}_1\|_2 + |\delta| \|\tilde{\mathcal{T}}_0\|_2) \|\tilde{y}\|_2 \|\tilde{x}\|_2}{|\tilde{y}^* \tilde{\mathcal{T}}_1 \tilde{x}|} = \frac{(\|\mathcal{T}_0\|_2 + |\delta| \|\mathcal{T}_1\|_2) \|y\|_2 \|x\|_2}{|y^* \mathcal{T}_0 x|} \\ &= \frac{(\|\mathcal{T}_0\|_2 + |\delta| \|\mathcal{T}_1\|_2) \|y\|_2 \|x\|_2}{|\delta| |y^* \mathcal{T}_1 x|} = \kappa_{\mathcal{T}_P}(\delta), \end{aligned}$$

where the last equality follows from the fact that $(\delta \mathcal{T}_1 - \mathcal{T}_0)x = 0$. \square

Proof of Theorem 5.1. Let x_1 and x_2 be, respectively, a left and a right eigenvector of $P(\lambda)$ associated with δ .

We compare $\kappa_{\mathcal{T}_P}(\delta)$ with $\kappa_P(\delta)$. Taking into account Lemmas 2.2 and 5.7, we assume $|\delta| \leq 1$, as otherwise we replace $P(\lambda)$ by $revP(\lambda)$ and δ by $\frac{1}{\delta}$. Note that both ρ_1 and ρ_2 in (5.1) take the same value when considering $P(\lambda)$ and $revP(\lambda)$. From Theorem 2.1 and Lemma 5.6, we have

$$\frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)} = \frac{(|\delta| \|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2) \|\Delta(\delta) x_1\|_2 \|\Delta(\delta) x_2\|_2}{(\sum_{i=0}^k |\delta|^i \|A_i\|_2) \|x_1\|_2 \|x_2\|_2}. \quad (5.9)$$

By Propositions 3.2 and 4.5, and taking into account that each of the matrix coefficients of $\mathcal{T}_P(\lambda)$ contains an identity block, we get

$$\begin{aligned} (|\delta| + 1) \min\{\max\{1, \|A_k\|_2\}, \max\{1, \|A_0\|_2\}\} &\leq \\ &|\delta| \|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2 \leq 2(|\delta| + 1) \max_{i=0:k} \{1, \|A_i\|_2\}. \end{aligned} \quad (5.10)$$

We also have

$$\sum_{i=0}^k |\delta|^i \|A_i\|_2 \geq \|A_k\|_2 |\delta|^k + \|A_0\|_2 \geq \min\{\|A_k\|_2, \|A_0\|_2\} (|\delta|^k + 1). \quad (5.11)$$

Thus, from (5.9) and the inequalities above, we get

$$\rho_2 \frac{|\delta| + 1}{\sum_{i=0}^k |\delta|^i} Q_T \leq \frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)} \leq 2 \frac{\max_{i=0:k} \{1, \|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{|\delta| + 1}{|\delta|^k + 1} Q_T, \quad (5.12)$$

where

$$Q_T = \frac{\|\Delta(\delta)x_1\|_2 \|\Delta(\delta)x_2\|_2}{\|x_1\|_2 \|x_2\|_2} \leq \|\Delta(\delta)\|_2^2 \leq d_1(\delta) \max_{i=0:k} \{1, \|A_i\|_2^2\}, \quad (5.13)$$

with $d_1(\delta)$ as in (3.7). Note that the last inequality follows from Lemma 3.5.

We now focus on the upper bound for $\frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)}$. Taking into account Lemma 3.6, and since $|\delta| \leq 1$, we obtain, for $k > 2$,

$$\frac{|\delta| + 1}{|\delta|^{k+1}} d_1(\delta) \leq 2d_1(\delta) \leq k^3.$$

A better bound for $\frac{|\delta|+1}{|\delta|^{k+1}} d_1(\delta)$ can be obtained when $|\delta| \ll 1$. More precisely, taking into account Lemma 3.6, if

$$(k-1)^3 |\delta|^2 \leq k+1,$$

then, $d_1(\delta) \leq k+1$, implying that

$$\frac{|\delta| + 1}{|\delta|^{k+1}} d_1(\delta) \leq 2(k+1).$$

Taking into account (5.12), (5.13) and the two previous upper bounds for $\frac{|\delta|+1}{|\delta|^{k+1}} d_1(\delta)$, the upper bound part of the theorem follows.

Next we show the lower bound for $\frac{\kappa_{\mathcal{T}_P}(\delta)}{\kappa_P(\delta)}$. From (3.8) with x replaced by x_1 , we have

$$\|\Delta(\delta)x_1\|_2^2 \geq \|x_1\|_2^2 \sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r}.$$

An analogous inequality holds for x_2 . Thus,

$$\frac{|\delta| + 1}{\sum_{i=0}^k |\delta|^i} Q_T \geq \frac{(|\delta| + 1) \sum_{r=0}^{\frac{k-1}{2}} |\delta|^{2r}}{\sum_{i=0}^k |\delta|^i} = 1,$$

and the result follows from (5.12).

5.2. Proof of Theorem 5.3. The following lemma will allow us to only consider eigenvalues δ such that $|\delta| \leq 1$ when proving Theorem 5.3.

LEMMA 5.8. *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$, and let (z, δ) be an approximate right eigenpair of $\mathcal{T}_P(\lambda)$, with $\delta \neq 0$. Then, $(RDz, \frac{1}{\delta})$ is an approximate right eigenpair of $\mathcal{T}_{revP}(\lambda)$ and*

$$\eta_{\mathcal{T}_P}(z, \delta) = \eta_{\mathcal{T}_{revP}}\left(RDz, \frac{1}{\delta}\right),$$

where R and D are as in (4.3) and (4.7), respectively.

Proof. Let $\mathcal{T}_P(\lambda) := \lambda \mathcal{T}_1 - \mathcal{T}_0$ and let $\mathcal{T}_{revP}(\lambda) := \lambda \tilde{\mathcal{T}}_1 - \tilde{\mathcal{T}}_0$. Note that (5.8) holds. Then, from Theorem 2.3 and Lemma 4.4,

$$\begin{aligned} \eta_{\mathcal{T}_{revP}}\left(RDz, \frac{1}{\delta}\right) &= \frac{\|\mathcal{T}_{revP}(\frac{1}{\delta})RDz\|_2}{\left(\frac{1}{|\delta|} \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2\right) \|RDz\|_2} = \frac{\|\frac{1}{\delta} \mathcal{T}_P(\delta)z\|_2}{\left(\frac{1}{|\delta|} \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2\right) \|z\|_2} \\ &= \frac{\|\mathcal{T}_P(\delta)z\|_2}{\left(\|\tilde{\mathcal{T}}_0\|_2 + |\delta| \|\tilde{\mathcal{T}}_1\|_2\right) \|z\|_2} = \eta_{\mathcal{T}_P}(z, \delta). \end{aligned}$$

□

Proof of Theorem 5.3. Let (z, δ) be an approximate right eigenpair of $\mathcal{T}_P(\lambda) := \lambda\mathcal{T}_1 - \mathcal{T}_0$ and assume that $|\delta| \leq 1$. Let $x := (e_k^T \otimes I_n)z$. Note that (x, δ) can be seen as an approximate right eigenpair of $P(\lambda)$. First we show the upper bound in (5.4). We have

$$P(\delta)x = P(\delta)(e_k^T \otimes I_n)z = (e_k^T \otimes P(\delta))z = \Delta^{\mathcal{B}}(\delta)\mathcal{T}_P(\delta)z,$$

where the last equality follows from Lemma 4.1. Thus,

$$\begin{aligned} \frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)} &= \frac{\|P(\delta)x\|_2}{(\sum_{i=0}^k |\delta|^i \|A_i\|_2) \|x\|_2} \cdot \frac{(|\delta| \|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2) \|z\|_2}{\|\mathcal{T}_P(\delta)z\|_2} \\ &\leq \frac{\|\Delta^{\mathcal{B}}(\delta)\|_2 \|\mathcal{T}_P(\delta)z\|_2}{(\sum_{i=0}^k |\delta|^i \|A_i\|_2) \|x\|_2} \cdot \frac{(|\delta| \|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2) \|z\|_2}{\|\mathcal{T}_P(\delta)z\|_2} \\ &= \frac{\|\Delta(\delta)\|_2 (|\delta| \|\mathcal{T}_1\|_2 + \|\mathcal{T}_2\|_2)}{\sum_{i=0}^k |\delta|^i \|A_i\|_2} \cdot \frac{\|z\|_2}{\|x\|_2} \\ &\leq 2 \frac{\max_{i=0:k} \{1, \|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{|\delta| + 1}{|\delta|^k + 1} \frac{\|z\|_2}{\|x\|_2} \|\Delta(\delta)\|_2, \end{aligned} \quad (5.14)$$

where the last inequality follows from (5.11) and the second inequality in (5.10).

By Lemma 3.5, we have

$$\|\Delta(\delta)\|_2 \leq \sqrt{d_1(\delta)} \max_{i=0:k} \{1, \|A_i\|_2\},$$

where $d_1(\lambda)$ is as in (3.7). Taking into account Lemma 3.6, since $|\delta| \leq 1$, we have, for $k > 2$,

$$\frac{|\delta| + 1}{|\delta|^k + 1} \sqrt{d_1(\delta)} \leq \frac{|\delta| + 1}{|\delta|^k + 1} \sqrt{\frac{k+1}{2} + \frac{(k-1)^3}{2} |\delta|^2} \leq 2k^{3/2}.$$

Thus, the upper bound in (5.4) follows by combining the two previous bounds with (5.14).

If $|\delta| \ll 1$, a better bound can be obtained. More precisely, if

$$(k-1)^3 |\delta|^2 \leq k+1,$$

then

$$\frac{|\delta| + 1}{|\delta|^k + 1} \sqrt{d_1(\delta)} \leq 2\sqrt{k+1},$$

implying the upper bound for $\frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)}$ in (5.5).

Now suppose that $|\delta| > 1$. From Lemmas 2.4 and 5.8, we have

$$\frac{\eta_P((e_1^T \otimes I_n)z, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)} = \frac{\eta_{revP}((e_k^T \otimes I_n)RDz, \frac{1}{\delta})}{\eta_{\mathcal{T}_{revP}}(RDz, \frac{1}{\delta})}.$$

Taking into account Lemma 4.4 and since $|\frac{1}{\delta}| < 1$, by the part of the theorem already proved, we have that (5.4) and (5.5) hold with $\frac{\eta_P(x, \delta)}{\eta_{\mathcal{T}_P}(z, \delta)}$ replaced by

$$\frac{\eta_{revP}((e_k^T \otimes I_n)RDz, \frac{1}{\delta})}{\eta_{\mathcal{T}_{revP}}(RDz, \frac{1}{\delta})}.$$

Note that

$$\frac{\|z\|_2}{\|(e_1^T \otimes I_n)z\|_2} = \frac{\|RDz\|_2}{\|(e_k^T \otimes I_n)RDz\|_2}.$$

6. Conditioning and backward error of $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$. In this section we present results analogous to Theorems 5.1 and 5.3 for the linearizations $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$. We note that these results were previously obtained in [15] and [16]. We include them for completeness with the goal of comparing the conditioning and backward error of $\mathcal{T}_P(\lambda)$ and those of the linearizations $D_1(\lambda, P)$, $D_k(\lambda, P)$, and $C_1(\lambda)$. With respect to the conditioning, we give some improvements in the bounds, which allow us to obtain a more accurate comparison of the different linearizations.

We start by recalling the definition of the pencils that we are considering in this section. The reader can find more details in [17, 20].

Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1) and assume that $k \geq 2$. We have

$$D_1(\lambda, P) := \lambda \left[\begin{array}{c|cccc} A_k & & & & \\ \hline & -A_{k-2} & -A_{k-3} & \cdots & -A_0 \\ & -A_{k-3} & -A_{k-4} & \cdots & 0 \\ & \vdots & \ddots & & \vdots \\ & -A_0 & 0 & \cdots & 0 \end{array} \right] - \left[\begin{array}{ccccc} -A_{k-1} & -A_{k-2} & \cdots & -A_1 & -A_0 \\ -A_{k-2} & -A_{k-3} & \cdots & -A_0 & 0 \\ \vdots & \ddots & & & \vdots \\ -A_0 & 0 & \cdots & \cdots & 0 \end{array} \right],$$

$$D_k(\lambda, P) := \lambda \left[\begin{array}{cccc} 0 & \cdots & 0 & A_k \\ 0 & \cdots & A_k & A_{k-1} \\ \vdots & \ddots & \vdots & \vdots \\ A_k & \cdots & A_2 & A_1 \end{array} \right] - \left[\begin{array}{cccc|c} 0 & \cdots & 0 & A_k & \\ 0 & \cdots & A_k & A_{k-1} & \\ \vdots & \ddots & \vdots & \vdots & \\ \hline A_k & \cdots & A_3 & A_2 & \\ \hline & & & & -A_0 \end{array} \right],$$

$$C_1(\lambda) := \lambda \left[\begin{array}{cccc} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{array} \right] - \left[\begin{array}{cccc} -A_{k-1} & -A_{k-2} & \cdots & -A_0 \\ I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I_n & 0 \end{array} \right].$$

We emphasize that $C_1(\lambda)$ is the very well-known first Frobenius companion form, which is fundamental in the theory and in the numerical computations of matrix polynomials [14]. The block-symmetric pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ have been thoroughly studied recently in [7, 15, 16, 17, 20, 21], although they were introduced as early as in [19]. Recall that $D_1(\lambda, P)$, $D_k(\lambda, P) \in \mathbb{DL}(P)$, where $\mathbb{DL}(P)$ denotes the vector space of pencils defined in [20].

When comparing the conditioning of eigenvalues and the backward error of approximate eigenpairs of a regular $P(\lambda)$ with those of any of its linearizations $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$, the next two lemmas will be useful since they provide a way to recover an eigenvector of a matrix polynomial $P(\lambda)$ from an eigenvector of $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$. We will use the following notation:

$$\Lambda(\lambda) = [\lambda^{k-1} \ \cdots \ \lambda \ 1]^T. \quad (6.1)$$

LEMMA 6.1. [20, Theorem 3.8] *Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of degree k and δ be a finite eigenvalue of $P(\lambda)$. Let $L(\lambda) \in \mathbb{DL}(P)$ be a linearization of $P(\lambda)$. A vector z is a right eigenvector of $L(\lambda)$ associated with δ if and only if $z = \Lambda(\delta) \otimes x$ for some right eigenvector x of $P(\lambda)$ associated with δ . Similarly, a vector ω is a left eigenvector of $L(\lambda)$ associated with δ if and only if $\omega = \Lambda(\delta) \otimes y$ for some left eigenvector y of $P(\lambda)$ associated with δ .*

LEMMA 6.2. [16, Section 1 and Lemma 7.2] *Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of degree k and δ be a finite eigenvalue of $P(\lambda)$. A vector z is a right eigenvector of $C_1(\lambda)$ associated with δ if and only if $z = \Lambda(\delta) \otimes x$ for some right eigenvector x of $P(\lambda)$ associated with δ .*

A vector ω is a left eigenvector of $C_1(\lambda)$ associated with δ if and only if $\omega^ = y^* [I_n, P_1(\delta), \dots, P_{k-2}(\delta), P_{k-1}(\delta)]$, for some left eigenvector y of $P(\lambda)$ associated with δ , where P_i is as in (3.2). Thus, if ω is a left eigenvector of $C_1(\lambda)$ with eigenvalue δ , then $y = (e_1^T \otimes I_n) \omega$ is a left eigenvector of $P(\lambda)$ with eigenvalue δ . Moreover, any left eigenvector y of $P(\lambda)$ with eigenvalue δ can be recovered from some left eigenvector ω of $C_1(\lambda)$ by taking $y = (e_1^T \otimes I_n) \omega$.*

6.1. Conditioning and backward error of $D_1(\lambda, P)$ and $D_k(\lambda, P)$. Next we recall some well-known results on the conditioning of eigenvalues and backward error of approximate eigenpairs of the linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ of a regular matrix polynomial $P(\lambda)$ introduced previously in [15, 16]. Moreover, we sharpen one of the results in [16]. Our goal is to compare the results in Section 5 for $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ with those for $D_1(\lambda, P)$ and $D_k(\lambda, P)$.

Recall that $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is a linearization of a regular matrix polynomial $P(\lambda)$ as in (1.1) if and only if A_0 (resp. A_k) is nonsingular.

In [16], it was shown that, for a simple, finite, nonzero eigenvalue δ of a regular $P(\lambda)$ as in (1.1) with $A_0 \neq 0$, the condition number of δ as an eigenvalue of $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$), when A_0 (resp. A_k) is nonsingular, is close to optimal among the linearizations of $P(\lambda)$ in $\mathbb{DL}(P)$, when $|\delta| \geq 1$ (resp. $|\delta| \leq 1$), provided that

$$\rho := \frac{\max_{i=0:k} \{\|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}. \quad (6.2)$$

is of order 1, which implies, in particular, that all matrix coefficients of $P(\lambda)$ must have similar norm. Notice that $\rho \geq 1$.

The combination of Theorems 4.4 and 4.5 in [16] provides a lower and an upper bound for $\frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)}$, when either $t = 1$ and $|\delta| \geq 1$, or $t = k$ and $|\delta| \leq 1$, namely,

$$\left(\frac{2\sqrt{k}}{k+1}\right) \frac{1}{\rho} \leq \frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} \leq \sqrt{k} \rho^2.$$

Using the same techniques as those applied to compute the bounds in Theorem 5.1, next we deduce sharper bounds for the quotient $\frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)}$ than those provided in [16]. Based on the results obtained, we can provide a fair comparison of the linearizations $\mathcal{T}_P(\lambda)$, $D_1(\lambda, P)$ and $D_k(\lambda, P)$ with respect to conditioning and explain the numerical experiments in Section 7 appropriately.

THEOREM 6.3. *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Assume that δ is a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $t \in \{1, k\}$ and suppose that A_0 is nonsingular if $t = 1$, and A_k is nonsingular if $t = k$. Let ρ be as in (6.2). Then, if either $t = 1$ and $|\delta| \geq 1$, or $t = k$ and $|\delta| \leq 1$, we have*

$$\frac{1}{\rho} \leq \frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} \leq k^2 \rho.$$

Proof. Let x and y be a right and a left eigenvector of $P(\lambda)$ associated with δ , respectively. Let $\Lambda(\lambda)$ be as in (6.1). Let $t \in \{1, k\}$, and define $D_t(\lambda, P) := L_1^t \lambda - L_0^t$. Taking into account Theorem 3.2 in [16], and using the natural weights for $\kappa_{D_t}(\delta)$, we have

$$\frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} = \frac{(|\delta| \|L_1^t\|_2 + \|L_0^t\|_2) \|\Lambda(\delta)\|_2^2}{|\delta|^{k-t} \sum_{i=0}^k |\delta|^i \|A_i\|_2}. \quad (6.3)$$

By Proposition 3.2 and taking into account that all block-entries of L_1^t and L_0^t are either 0 or matrix coefficients of $P(\lambda)$, we obtain

$$|\delta| \|L_1^t\|_2 + \|L_0^t\|_2 \leq (|\delta| + 1) k \max_{i=0:k} \{\|A_i\|_2\}. \quad (6.4)$$

Thus, from (5.11) and (6.3), we obtain

$$\frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} \leq \frac{k \max_{i=0:k} \{\|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{(|\delta| + 1) \sum_{i=0}^{k-1} |\delta|^{2i}}{(|\delta|^k + 1) |\delta|^{k-t}} = k \rho \frac{\sum_{i=0}^{2k-1} |\delta|^i}{(|\delta|^k + 1) |\delta|^{k-t}}. \quad (6.5)$$

Clearly, if $|\delta| = 1$, the upper bound in the statement follows. Now suppose that $t = 1$ and $|\delta| > 1$, or $t = k$ and $|\delta| < 1$. Since $|\delta|^{2k} - 1 = (|\delta| - 1) \sum_{i=0}^{2k-1} |\delta|^i$, from (6.5), we have

$$\begin{aligned} \frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} &\leq k\rho \frac{|\delta|^{2k} - 1}{(|\delta| - 1)(|\delta|^k + 1)|\delta|^{k-t}} = k\rho \frac{|\delta|^k - 1}{(|\delta| - 1)|\delta|^{k-t}} \\ &\leq k\rho \frac{|\delta|^{k-1} + \dots + 1}{|\delta|^{k-t}} \leq k^2\rho \frac{\max\{1, |\delta|^{k-1}\}}{|\delta|^{k-t}} = k^2\rho. \end{aligned}$$

Now we show the lower bound in the statement. From Proposition 3.2 and taking into account the block-structure of $D_t(\lambda, P)$, we obtain

$$(|\delta| + 1) \min\{\|A_k\|_2, \|A_0\|_2\} \leq |\delta| \|L_1^t\|_2 + \|L_0^t\|_2.$$

Thus, from (6.3),

$$\begin{aligned} \frac{\kappa_{D_t}(\delta)}{\kappa_P(\delta)} &\geq \frac{(|\delta| + 1) \min\{\|A_k\|_2, \|A_0\|_2\} \sum_{i=0}^{k-1} |\delta|^{2i}}{\max_{i=0:k} \{\|A_i\|_2\} |\delta|^{k-t} \sum_{i=0}^k |\delta|^i} \\ &= \frac{1}{\rho} \frac{\sum_{i=0}^{2k-1} |\delta|^i}{|\delta|^{k-t} \sum_{i=0}^k |\delta|^i} \geq \frac{1}{\rho}. \end{aligned}$$

□

Notice that the previous theorem implies that, if the norms of the matrix coefficients of $P(\lambda)$ have similar magnitudes, that is, $\rho \approx 1$, and k is moderate, then the conditioning of the eigenvalues of $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is close to that of the corresponding eigenvalues of $P(\lambda)$ when $|\delta| \geq 1$ and A_0 is nonsingular (resp. when $|\delta| \leq 1$ and A_k is nonsingular).

REMARK 6.4. *Based on Theorems 5.1 and 6.3, we compare the conditioning of the nonzero finite simple eigenvalues of the linearizations $\mathcal{T}_P(\lambda)$, $D_1(\lambda, P)$ and $D_k(\lambda, P)$. We note first that the bounds on the quotient of condition numbers associated with $\mathcal{T}_P(\lambda)$ are valid for all eigenvalues, regardless of their modulus, while the bound for $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is only valid in a certain range of eigenvalues. This forces the use of both linearizations, $D_1(\lambda, P)$ and $D_k(\lambda, P)$, when the matrix polynomial has both eigenvalues with modulus larger than 1 and eigenvalues with modulus less than 1. We note also that, when $P(\lambda)$ is scaled by dividing all matrix coefficients by $\max_{i=0:k} \{\|A_i\|_2\}$, the parameters ρ and ρ_1 that appear in the bounds of the quotients of condition numbers are equal. We must also point out that, although the general upper bound for these quotients has degree 2 on k for $D_1(\lambda, P)$ and $D_k(\lambda, P)$, and degree 3 for $\mathcal{T}_P(\lambda)$, when the modulus of the eigenvalue is not close to 1, the bound for $\mathcal{T}_P(\lambda)$ has degree 1 on k . Finally, observe that $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is a linearization of $P(\lambda)$ if and only if A_0 (resp. A_k) is nonsingular, contrarily to $\mathcal{T}_P(\lambda)$, which is always a linearization of $P(\lambda)$. Thus, it is clear that the use of $\mathcal{T}_P(\lambda)$ presents clear advantages over the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ for conditioning purposes.*

We now recall a result that compares the backward error of an approximate eigenpair of the linearization $D_t(\lambda, P)$, $t \in \{1, k\}$, with the backward error of a certain approximate eigenpair of $P(\lambda)$.

THEOREM 6.5. [15, Corollary 3.11] *Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let $t \in \{1, k\}$ and suppose that A_0 is nonsingular if $t = 1$, and A_k is nonsingular if $t = k$. Let ρ be as in (6.2). Let (z, δ) be an approximate right eigenpair of $D_t(\lambda, P)$, with δ nonzero and finite. Then, for $z_t = (e_t^T \otimes I_n)z$, we have that (z_t, δ) is an approximate right eigenpair for $P(\lambda)$ and*

$$\frac{\eta_P(z_t, \delta)}{\eta_{D_t}(z, \delta)} \leq k^{3/2} \frac{\|z\|_2}{\|z_t\|_2} \rho. \quad (6.6)$$

An analogue result holds for left eigenpairs (w^, δ) of $D_t(\lambda, P)$ by simply replacing z by w and z_t by $w_t := (e_t^T \otimes I_n)w$.*

Taking into account the form of the right eigenvectors of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ (see Lemma 6.1), and assuming that the approximate eigenvector z has a similar block-structure, it is expected that, if

$t = 1$ and $|\delta| \geq 1$, or if $t = k$ and $|\delta| \leq 1$, the approximate eigenvector z_t for $P(\lambda)$, recovered from the approximate eigenvector z of $D_t(\lambda, P)$ as in Theorem 6.5, makes the quotient $\frac{\|z\|_2}{\|z_t\|_2}$ in (6.6) close to 1.

REMARK 6.6. *Based on Theorems 5.3 and 6.5, we compare the backward errors of $\mathcal{T}_P(\lambda)$, $D_1(\lambda, P)$ and $D_k(\lambda, P)$. First note that, although Theorem 6.5 is valid for any value of δ , an argument similar to the one in Remark 5.5 shows that, for $D_1(\lambda, P)$, the quotient $\frac{\|z\|_2}{\|z_1\|_2}$ in (6.6) is expected to be close to one only if $|\delta| \geq 1$, while for $D_k(\lambda, P)$, the quotient $\frac{\|z\|_2}{\|z_k\|_2}$ is expected to be close to 1 only if $|\delta| \leq 1$. In contrast, according to Remark 5.5, the quotient $\frac{\|z\|_2}{\|x\|_2}$ appearing in Theorem 5.3 is expected to be close to one for any value of δ . In addition, observe that, when $P(\lambda)$ is scaled by dividing all the matrix coefficients by $\max_{i=0:k}\{\|A_i\|_2\}$, the parameters ρ and ρ' that appear in the bounds in Theorem 5.3 and 6.5, respectively, have the same value, which is approximately one for polynomials whose coefficients have similar norms. Finally, observe that the bounds in Theorem 5.3 and 6.5 have the same dependence on k , that is, $k^{3/2}$, in general, and that the bound corresponding to $\mathcal{T}_p(\lambda)$ is improved if $|\delta|$ is not close to 1, depending on $k^{1/2}$ in this case. Therefore, once $P(\lambda)$ is divided by $\max_{i=0:k}\{\|A_i\|_2\}$, the use of $\mathcal{T}_p(\lambda)$ presents clear advantages with respect to $D_1(\lambda, P)$ and $D_k(\lambda, P)$ in terms of backward errors, since by using $\mathcal{T}_p(\lambda)$, for all approximate eigenvalues, we will get similar backward errors as using $D_1(\lambda, P)$ for computing the eigenvalues with $|\delta| \geq 1$ and $D_k(\lambda, P)$, for computing the eigenvalues with $|\delta| \leq 1$.*

6.2. Conditioning and backward error of $C_1(\lambda)$. Next we focus on the first Frobenius companion linearization $C_1(\lambda)$ of $P(\lambda)$. Note that, since $C_2(\lambda) = [C_1(\lambda)]^B$, where $C_1(\lambda)^B$ denotes the block-transpose of $C_1(\lambda)$ and $C_2(\lambda)$ denotes the second Frobenius companion form of $P(\lambda)$ [14], any result that we produce for $C_1(\lambda)$ has an immediate counterpart for $C_2(\lambda)$ (see [16, Lemma 7.1]).

We start by comparing the conditioning of the eigenvalues of $C_1(\lambda)$ with the conditioning of the corresponding eigenvalues of $P(\lambda)$. As far as we know, an explicit result, valid for any k , is not given in the literature, though the quadratic case was studied in [16].

THEOREM 6.7. *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let δ be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $C_1(\lambda)$ be the first Frobenius companion linearization of $P(\lambda)$. Let*

$$\rho' := \frac{\max_{i=0:k}\{1, \|A_i\|_2^2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \quad \text{and} \quad \nu := \frac{\min\{\max\{1, \|A_k\|_2\}, \max_{i=0:k-1}\{1, \|A_i\|_2\}\}}{\max_{i=0:k}\{\|A_i\|_2\}}. \quad (6.7)$$

Then,

$$\frac{\nu}{k+1} \leq \frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} \leq 2\sqrt{2}k^3\rho'.$$

Moreover, if $|\delta| \geq \sqrt{(k-1)^3}$ or $|\delta| \leq \frac{1}{2}$, then

$$\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} \leq \frac{4}{3}k(1+k)\rho'.$$

Proof. Let $C_1(\lambda) := \lambda X_1 + Y_1$. By Lemma 6.2, if y is a left-eigenvector of $P(\lambda)$ associated with δ , then

$$w = \begin{bmatrix} I_n \\ P_1^*(\delta) \\ \vdots \\ P_{k-1}^*(\delta) \end{bmatrix} y,$$

is a left eigenvector of $C_1(\lambda)$ associated with δ . Taking into account [16, Theorem 7.3], we obtain

$$\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} = \frac{\|w\|_2 (\|\delta\|X_1\|_2 + \|Y_1\|_2)\|\Lambda(\delta)\|_2}{\|y\|_2 \sum_{i=0}^k |\delta|^i \|A_i\|_2},$$

where $\Lambda(\lambda)$ is as in (6.1). By [16, Theorem 7.4], we have

$$\|X_1\|_2 = \max\{1, \|A_k\|_2\} \quad (6.8)$$

and

$$\max_{i=0:k-1} \{1, \|A_i\|_2\} \leq \|Y_1\|_2 \leq k \max_{i=0:k-1} \{1, \|A_i\|_2\}. \quad (6.9)$$

Thus, using (5.11), we get

$$\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} \leq \frac{\|w\|_2}{\|y\|_2} \frac{(\|\delta\| + k) \max_{i=0:k} \{1, \|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{\sqrt{\sum_{i=0}^{k-1} |\delta|^{2i}}}{|\delta|^k + 1}. \quad (6.10)$$

Assume that $|\delta| \leq 1$. Then,

$$\begin{aligned} \|w\|_2^2 &= \|y\|_2^2 + \sum_{i=1}^{k-1} \|P_i^*(\delta)y\|_2^2 \leq \left(1 + \sum_{i=1}^{k-1} \|P_i^*(\delta)\|_2^2\right) \|y\|_2^2 \\ &\leq \max_{i=0:k} \{1, \|A_i\|_2^2\} \|y\|_2^2 \left(1 + \sum_{i=1}^{k-1} \left(\sum_{j=0}^i |\delta|^j\right)^2\right) \\ &\leq \max_{i=0:k} \{1, \|A_i\|_2^2\} \|y\|_2^2 \left(1 + (k-1)k \sum_{j=0}^{k-1} |\delta|^{2j}\right), \end{aligned} \quad (6.11)$$

where the second and third inequalities follow taking into account Lemma 3.4 and Lemma 3.1, respectively. Thus,

$$\|w\|_2^2 \leq \max_{i=0:k} \{1, \|A_i\|_2^2\} \|y\|_2^2 [1 + (k-1)k^2] \leq k^3 \max_{i=0:k} \{1, \|A_i\|_2^2\} \|y\|_2^2.$$

From (6.10), we obtain $\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} \leq k^2(1+k)\rho' \leq 2\sqrt{2}k^3\rho'$, when $|\delta| \leq 1$.

A better bound for $\kappa_{C_1}(\delta)/\kappa_P(\delta)$ can be found when $|\delta| \ll 1$. Assume that $|\delta| \leq \frac{1}{2}$. Taking into account (6.10) and the upper bound for $\|w\|_2$ in (6.11), we have

$$\begin{aligned} \frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} &\leq (1+k)\rho' \frac{\sqrt{(1+k(k-1) \sum_{j=0}^{k-1} |\delta|^{2j}) \sum_{i=0}^{k-1} |\delta|^{2i}}}{|\delta|^k + 1} \\ &\leq (1+k)k\rho' \frac{\sqrt{\sum_{j=0}^{k-1} |\delta|^{2j} \sum_{i=0}^{k-1} |\delta|^{2i}}}{|\delta|^k + 1} \\ &\leq (1+k)k\rho' \frac{\sum_{j=0}^{k-1} |\delta|^{2j}}{|\delta|^k + 1} = (1+k)k\rho' \frac{(1-|\delta|^{2k})}{(1-|\delta|^2)(1+|\delta|^k)} \\ &\leq (1+k)k\rho' \frac{1}{(1-|\delta|^2)} \leq \frac{4}{3}(1+k)k\rho', \end{aligned}$$

where the second inequality follows because

$$1 + k(k-1) \sum_{j=0}^{k-1} |\delta|^{2j} \leq \sum_{j=0}^{k-1} |\delta|^{2j} + (k^2 - k) \sum_{j=0}^{k-1} |\delta|^{2j} \leq k^2 \sum_{j=0}^{k-1} |\delta|^{2j},$$

and the last inequality follows from the assumption $|\delta| \leq \frac{1}{2}$.

If $|\delta| > 1$, taking into account Lemma 3.1, Lemma 3.3, and (3.4), we have

$$\begin{aligned}
\|w\|_2^2 &= \|y\|_2^2 + \sum_{i=1}^{k-1} \|y^* P_i(\delta)\|_2^2 = \|y\|_2^2 + \sum_{i=1}^{k-1} \|\delta^{i-k} y^* P^{k-i-1}(\delta)\|_2^2 \\
&\leq \|y\|_2^2 + \sum_{i=1}^{k-1} |\delta|^{2(i-k)} \|P^{k-i-1}(\delta)\|_2^2 \|y\|_2^2 \\
&\leq \|y\|_2^2 \max_{i=0:k} \{1, \|A_i\|_2^2\} \left[1 + \sum_{i=1}^{k-1} |\delta|^{2(i-k)} \left(\sum_{j=0}^{k-i-1} |\delta|^j \right)^2 \right] \\
&\leq \|y\|_2^2 \max_{i=0:k} \{1, \|A_i\|_2^2\} \left[1 + \sum_{i=1}^{k-1} (k-i) \sum_{j=0}^{k-i-1} |\delta|^{2i-2k+2j} \right] \\
&\leq \|y\|_2^2 \max_{i=0:k} \{1, \|A_i\|_2^2\} \left[1 + (k-1)^2 \sum_{j=0}^{k-2} |\delta|^{2j+2-2k} \right] \\
&\leq \|y\|_2^2 \max_{i=0:k} \{1, \|A_i\|_2^2\} [1 + (k-1)^3 |\delta|^{-2}]. \tag{6.12}
\end{aligned}$$

Since $|\delta| + k \leq (|\delta| + 1)k$, taking into account (6.10) and the upper bound for $\|w\|_2^2/\|y\|_2^2$ in (6.12), we have

$$\begin{aligned}
\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} &\leq \frac{\|w\|_2 \max_{i=0:k} \{1, \|A_i\|_2\}}{\|y\|_2 \min\{\|A_k\|_2, \|A_0\|_2\}} \frac{(1 + |\delta|)k\sqrt{k}|\delta|^{k-1}}{1 + |\delta|^k} \\
&= \rho' k \sqrt{k} \sqrt{1 + (k-1)^3 |\delta|^{-2}} \frac{(1 + |\delta|)}{|\delta|} \\
&\leq \begin{cases} 2\sqrt{2}k^{3/2}\rho' \leq \frac{4}{3}k(k+1)\rho', & \text{if } |\delta| \geq \sqrt{(k-1)^3}, \\ 2\sqrt{2}k^3\rho', & \text{otherwise.} \end{cases}
\end{aligned}$$

Now we find a lower bound for $\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)}$. Notice that $\|w\|_2 \geq \|y\|_2$. Thus, from (6.8) and (6.9), we have

$$\begin{aligned}
\frac{\kappa_{C_1}(\delta)}{\kappa_P(\delta)} &\geq \frac{(|\delta|\|X\|_2 + \|Y\|_2)\|\Lambda(\delta)\|_2}{\sum_{i=0}^k |\delta|^i \|A_i\|_2} \\
&\geq \frac{\min\{\max\{1, \|A_k\|_2\}, \max_{i=0:k-1}\{1, \|A_i\|_2\}\}}{\max_{i=0:k} \{ \|A_i\|_2 \}} \frac{(|\delta| + 1) \sqrt{\sum_{i=0}^{k-1} |\delta|^{2i}}}{\sum_{i=0}^k |\delta|^i} \geq \frac{\nu}{k+1}
\end{aligned}$$

where the last inequality holds since

$$\frac{(|\delta| + 1) \sqrt{\sum_{i=0}^{k-1} |\delta|^{2i}}}{\sum_{i=0}^k |\delta|^i} \geq \frac{1}{k+1}.$$

In fact, to see this, note that, if $|\delta| \geq 1$,

$$(|\delta| + 1) \sqrt{\sum_{i=0}^{k-1} |\delta|^{2i}} \geq |\delta|^k \quad \text{and} \quad \sum_{i=0}^k |\delta|^i \leq (k+1)|\delta|^k,$$

and, if $|\delta| \leq 1$,

$$(|\delta| + 1) \sqrt{\sum_{i=0}^{k-1} |\delta|^{2i}} \geq 1 \quad \text{and} \quad \sum_{i=0}^k |\delta|^i \leq k+1.$$

□

We now recall a result obtained in [15] regarding the comparison of the backward errors of approximate eigenpairs of $C_1(\lambda)$ and $P(\lambda)$.

THEOREM 6.8. [15, Theorems 3.6 and 3.8] *Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let (z, δ) be an approximate right eigenpair of $C_1(\lambda)$. Then, for $z_t = (e_t^T \otimes I_n)z$, $t = 1 : k$, we have that (z_t, δ) is an approximate right eigenpair of $P(\lambda)$ and*

$$\frac{\eta_P(z_t, \delta)}{\eta_{C_1}(z, \delta)} \leq k^{5/2} \frac{\|z\|_2}{\|z_t\|_2} \rho', \quad (6.13)$$

where ρ' is as in (6.7). Let (w^*, δ) be an approximate left eigenpair of $C_1(\lambda)$. Then, for $w_1 = (e_1^T \otimes I_n)w$, we have that (w_1^*, δ) is an approximate left eigenpair of $P(\lambda)$ and

$$\frac{\eta_P(w_1^*, \delta)}{\eta_{C_1}(w^*, \delta)} \leq k^{3/2} \frac{\|w\|_2}{\|w_1\|_2} \tau,$$

where $\tau = \frac{\max_{i=0:k} \{1, \|A_i\|_2\}}{\min\{\|A_0\|_2, \|A_k\|_2\}}$.

REMARK 6.9. *Taking into account the form of the right eigenvectors of $C_1(\lambda)$ (see Lemma 6.2), and assuming that the approximate eigenvector z has a similar block structure, it is expected that the quotient $\frac{\|z\|_2}{\|z_t\|_2}$ in (6.13) is close to 1 if $t = 1$ and $|\delta| \geq 1$, or if $t = k$ and $|\delta| \leq 1$. Thus, depending on the modulus of the approximate eigenvalue δ , it is convenient to recover the approximate eigenvector x of $P(\lambda)$ from the eigenvector z of $C_1(\lambda)$ as follows: we take $x = z_1$ if $|\delta| \geq 1$ and $x = z_k$ if $|\delta| \leq 1$. Moreover, $\rho' \approx 1$ if all matrix coefficients have norms close to 1. In practice, we can scale $P(\lambda)$ by dividing each of its matrix coefficients by $\max_{i=0:k} \{\|A_i\|_2\}$. The scaled polynomial $\tilde{P}(\lambda)$ has the same eigenvalues and eigenvectors as $P(\lambda)$ and, additionally, the conditioning of the eigenvalues and the backward error of the approximate eigenpairs is not changed by the scaling. Thus, we can work with the pencil $C_1(\lambda)$ associated with $\tilde{P}(\lambda)$. In this case, for ρ' to be close to 1, it is enough that all matrix coefficients of $\tilde{P}(\lambda)$ have similar norms.*

REMARK 6.10. *Based on Theorems 5.1, 5.3, 6.7 and 6.8, we conclude that $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$ have similar behavior in terms of conditioning and backward errors. Note that the bounds on the quotients of condition numbers and backward errors are comparable when $P(\lambda)$ is scaled by dividing its matrix coefficients by $\max_{i=0:k} \{\|A_i\|_2\}$. Moreover, both pencils are always linearizations of $P(\lambda)$ (regular and singular). However, when $P(\lambda)$ is symmetric (Hermitian), $\mathcal{T}_P(\lambda)$ has an advantage over $C_1(\lambda)$ since, in this case, $\mathcal{T}_P(\lambda)$ is also symmetric (Hermitian) while $C_1(\lambda)$ is not and, for numerical reasons, it is more convenient that the linearizations preserve the structure of the original matrix polynomial in order to preserve any symmetries in the spectrum.*

7. Numerical experiments. In this section, we run some numerical experiments to illustrate the theoretical results presented in previous sections concerning the conditioning of eigenvalues and backward errors of approximate eigenpairs of the linearizations $\mathcal{T}_P(\lambda)$, $D_1(\lambda, P)$, $D_k(\lambda, P)$ and $C_1(\lambda)$ of a matrix polynomial $P(\lambda)$ and to compare the behavior of the different linearizations analyzed. In some examples, $P(\lambda)$ is a random matrix polynomial while in others, $P(\lambda)$ is one of the matrix polynomials connected with some applications discussed in [5].

The experiments were run on MATLAB-R2016a, for which the unit roundoff is 2^{-53} . When calculating the condition numbers, we computed the eigenvalues, eigenvectors, and the condition numbers themselves using variable precision arithmetic with 40 digits of precision. When computing the backward error, we considered only right eigenpairs. The function `polyeig` is used to compute approximate eigenvalues and eigenvectors of the linearization. If z denotes a computed eigenvector for the linearization associated with a computed eigenvalue δ , an eigenvector for $P(\lambda)$ was recovered from z as described in Remark 6.9 and Theorems 5.3 and 6.5.

In order to improve the ratio of the condition numbers and backward errors of the linearizations $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$ of the matrix polynomial $P(\lambda)$, a scaling of the polynomial $P(\lambda)$ of the type discussed in Section 2 is applied in some cases (recall that the condition number and the backward error of $P(\lambda)$ are invariant under such scalings). More specifically, in our experiments with random polynomials we

scale $P(\lambda)$ by dividing each matrix coefficient by $\max_{i=0:k}\{\|A_i\|_2\}$, a scaling that leaves the parameter λ invariant. As pointed out in Remarks 6.4, 6.6, and 6.10, this scaling ensures that the parameters ρ , ρ' , ρ_1 and τ appearing in the bounds of the quotients of condition numbers and backward errors are of similar magnitudes and that the quotient of the norms of approximate eigenvectors in the upper bounds of the ratios of backward errors are close to 1. We note that this type of scaling does not produce any improvements in the ratios associated with $D_1(\lambda, P)$ and $D_k(\lambda, P)$. The numerical results that we present next show both cases, when $P(\lambda)$ is scaled and when $P(\lambda)$ is not. In our experiments with the applied problems, we also apply some eigenvalue scalings, as we explain later.

To start with, we examine a 20×20 random matrix polynomial $P(\lambda)$ of degree 3. The polynomial $P(\lambda)$ was generated by producing random matrix coefficients with entries between $[-50, 50]$. Note that the norms of the matrix coefficients of $P(\lambda)$ are similar as can be seen in Table 1. Also, in this table the smallest and the largest modulus of the eigenvalues are displayed.

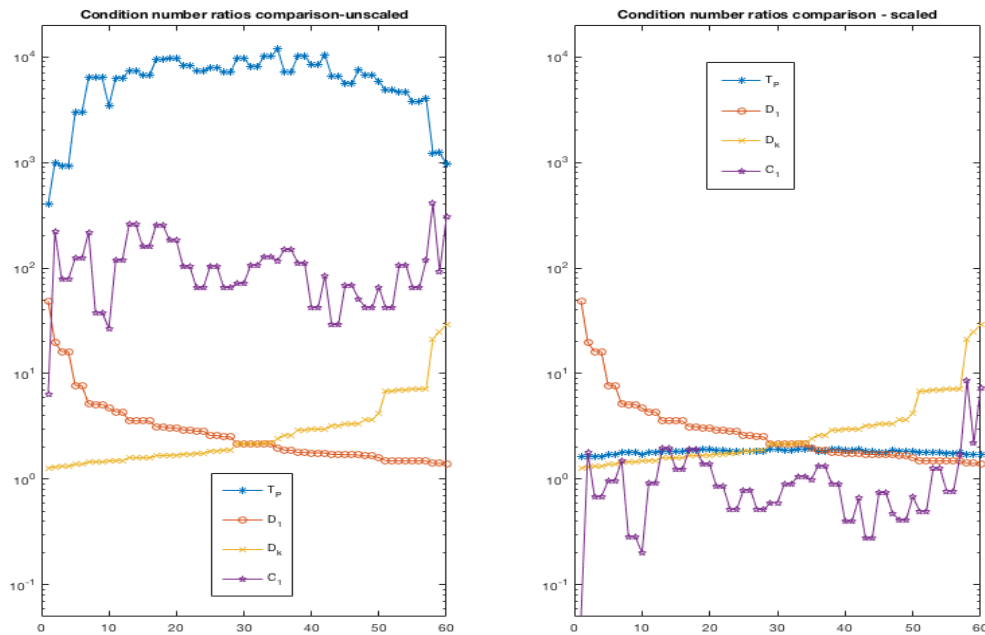


FIG. 1. Condition number ratio for unscaled (left) and scaled (right) random 20×20 matrix polynomial $P(\lambda)$ of degree 3.

In Figures 1 and 2, the x -axis has the indices $1, 2, \dots, 60$ to represent the non-zero, simple finite eigenvalues of $P(\lambda)$, which are sorted in increasing order by modulus (i.e. 1 represents the smallest eigenvalue in this order while 60 represents the eigenvalue with largest modulus). On the y -axis we give the ratio $\frac{\kappa_L(\delta)}{\kappa_P(\delta)}$, where $L(\lambda)$ denotes any of the linearizations, $D_1(\lambda, P)$, $D_k(\lambda, P)$, $C_1(\lambda)$ or $\mathcal{T}_P(\lambda)$.

In Figure 1, we present the results for conditioning while in Figure 2 we present the results for backward errors. The conditioning results associated with $\mathcal{T}_P(\lambda)$ in this example are not ideal when $P(\lambda)$ is not scaled (figure on the left), with ratios mostly close to 10^4 compared to ratios close to 1 corresponding to the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$. Note that we can guess what eigenvalues have modulus less than 1 or larger than 1 by inspecting where the graphs for $D_1(\lambda, P)$ and $D_k(\lambda, P)$ intersect. A behavior similar to that of $\mathcal{T}_P(\lambda)$ can be observed for $C_1(\lambda)$, which is natural since, in contrast with $D_1(\lambda, P)$ and $D_k(\lambda, P)$, $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$ have both identity blocks that will lead to undesirable behaviors either when $\max_{i=0:k}\{\|A_i\|_2\} \ll 1$ or when $\max_{i=0:k}\{\|A_i\|_2\} \gg 1$. In this example, since $\max_{i=0:k}\{\|A_i\|_2\} \approx 250$, it can be expected that the behavior of $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$ is penalized by a factor of $\approx 10^4$ with respect to

that of $D_1(\lambda, P)$ and $D_k(\lambda, P)$. We emphasize that these naturally expected behaviors are fully supported by the theoretical results in Section 5 and 6.

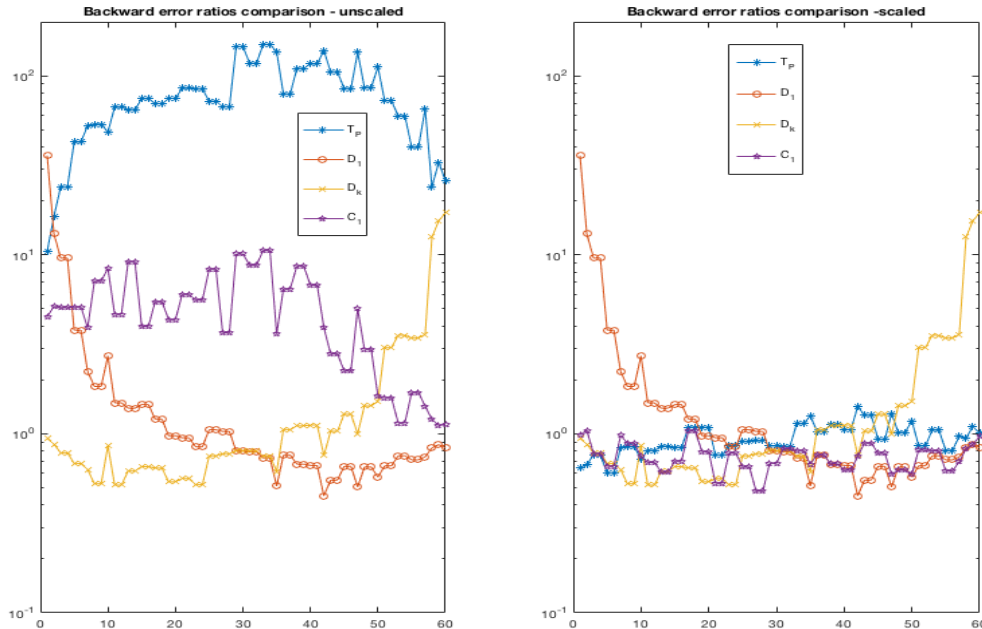


FIG. 2. Backward error ratio for unscaled (left) and scaled (right) random 20×20 matrix polynomial $P(\lambda)$ of degree 3.

In the figure on the right in Figure 1, we repeat the experiment with $P(\lambda)$ scaled by dividing all its matrix coefficients by $\max_{i=0:k} \{\|A_i\|_2\}$. Note that now $\mathcal{T}_P(\lambda)$ yields remarkably better results, comparable to those of the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ and much better than any of them considered individually. It is worth noting that while the scaling improved the behavior of $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$, it had no effect on the condition number ratios for $D_1(\lambda, P)$ and $D_k(\lambda, P)$, as the theory predicts and it is naturally expected as discussed in the previous paragraph. In Table 1, the value of the maximum ratio of condition numbers for each of the four linearizations has been highlighted in boldface.

In Figure 2 we present the values of the ratio $\frac{\eta_P(x, \delta)}{\eta_L(z, \delta)}$ of backward errors associated with $P(\lambda)$ and the linearizations that we are considering, before scaling (on the left) and after scaling (on the right). These figures show the improvement in the backward error ratio after scaling the same matrix polynomial $P(\lambda)$ again by dividing all its matrix coefficients by $\max_{i=0:k} \{\|A_i\|_2\}$. Once more, we notice that the backward error ratios for $\mathcal{T}_P(\lambda)$ and $C_1(\lambda)$ are worse than those of the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ before scaling, but become much closer to 1 after scaling $P(\lambda)$. In Table 1, the value of the maximum ratio of backward errors for each of the four linearizations has been highlighted in boldface. Notice the good behavior of $\mathcal{T}_P(\lambda)$ compared to the other linearizations. We can state, informally, that after scaling $\mathcal{T}_P(\lambda)$ is the winner among the block symmetric linearizations.

All data pertaining to Figures 1-2 is summarized in Table 1. In the table we also add the data for a 2×2 random matrix polynomial of degree 3. For this example we do not display the corresponding graphs, which are similar to the ones previously shown. Note that the results do not seem to be influenced by the size of the matrix polynomial.

One might expect that, for large k , $\mathcal{T}_P(\lambda)$ has a poor behavior in terms of conditioning and backward error, especially for eigenvalues of modulus close to 1, since the upper bounds for the ratio of the condition numbers and the ratios of backward errors increase with k and, in fact, the bound for the ratio of condition

TABLE 1
Experiment results for random polynomials: $k = 3, n = 2$ and $k = 3, n = 20$.

Problem:	k=3 n=2		k=3 n=20	
	Unscaled	Scaled	Unscaled	Scaled
$ \delta_{min} $	0.12	0.12	0.178	0.178
$ \delta_{max} $	1.36	1.36	4.31	4.31
$\ A_3\ _2$	41.7	0.75	228.07	0.94
$\ A_2\ _2$	24.1	0.43	243.23	1
$\ A_1\ _2$	45.2	0.81	235.7	0.97
$\ A_0\ _2$	55.5	1	242.8	0.99
ρ	1.3	1.3	1.066	1.066
ρ_1	4117.02	1.3	63093.5	1.066
ρ'	74.1	1.3	259.4	1.066
D_1, D_k upper bound for cond	12	12	9.6	9.6
D_1 upper bound for back	483.5	483.5	176.6	176.6
D_k upper bound for back	17.35	17.35	105.99	105.99
\mathcal{T}_P upper bound for cond.	222319.3	72	3407051.27	57.59
\mathcal{T}_P upper bound for back.	55764.9	41.8	815295.5	33.84
C_1 upper bound for cond.	5658.99	101.86	19809.6	81.44
C_1 upper bound for back.	1990.5	35.8	6845.6	28.1
$\min\{\kappa_{D_1}/\kappa_P\}$	2.1	2.1	1.4	1.4
$\max\{\kappa_{D_1}/\kappa_P\}$	91.5	91.5	48.8	48.8
$\min\{\kappa_{D_k}/\kappa_P\}$	1.04	1.04	1.28	1.28
$\max\{\kappa_{D_k}/\kappa_P\}$	3.4	3.4	28.96	28.96
$\min\{\kappa_{\mathcal{T}_P}/\kappa_P\}$	4.86	1.3	400.18	1.63
$\max\{\kappa_{\mathcal{T}_P}/\kappa_P\}$	1439.6	2.65	11812.15	1.97
$\min\{\kappa_{C_1}/\kappa_P\}$	17.16	0.63	6.45	0.05
$\max\{\kappa_{C_1}/\kappa_P\}$	78.5	22.38	409.8	8.69
$\min\{\eta_P/\eta_{D_1}\}$	0.27	0.27	0.45	0.45
$\max\{\eta_P/\eta_{D_1}\}$	74.12	74.12	35.8	35.8
$\min\{\eta_P/\eta_{D_k}\}$	0.43	0.43	0.52	0.52
$\max\{\eta_P/\eta_{D_k}\}$	0.99	0.99	17.14	17.14
$\min\{\eta_P/\eta_{\mathcal{T}_P}\}$	1.15	0.64	10.46	0.6
$\max\{\eta_P/\eta_{\mathcal{T}_P}\}$	42.5	1.74	149.7	1.41
$\min\{\eta_P/\eta_{C_1}\}$	2.5	0.41	1.11	0.48
$\max\{\eta_P/\eta_{C_1}\}$	5.08	1.22	10.6	1.05

numbers increases faster than the one for $D_t(\lambda)$, $t = 0, k$. However, optimal behavior is still observed for random matrix polynomials of large odd degree k whose matrix coefficients have similar norms. Figure 3 shows the results for a 3×3 random matrix polynomial of degree $k = 21$ after scaling the matrix coefficients of the matrix polynomial by dividing by $\max_{i=0:k}\{\|A_i\|_2\}$.

So far results for randomly generated matrix polynomials have been presented. Next we show the results of several experiments on matrix polynomials related with some applied problems given in [5]. We investigate the backward error ratios of two specific problems: 1) ‘‘Relative Pose 5pt Problem’’ (five point relative pose problem in computer vision), which generates a 10×10 matrix polynomial of degree 3; 2) ‘‘Plasma-drift Problem’’ (modeling of drift instabilities in the plasma edge inside a Tokamak reactor), which generates a 128×128 matrix polynomial of degree 3. The results for condition number ratios are

similar to the backward error results and, therefore, are omitted for brevity.

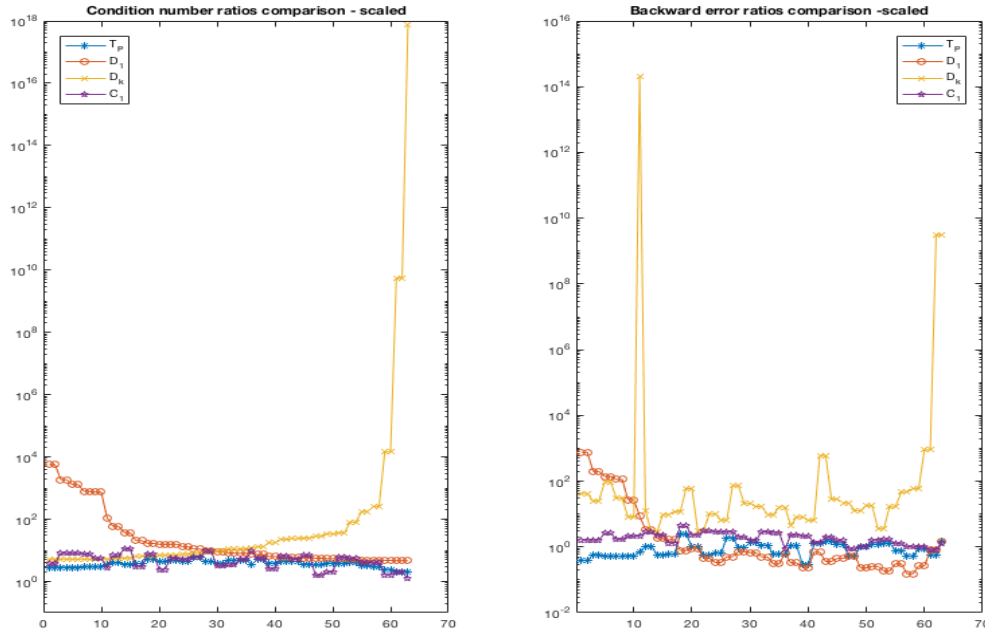


FIG. 3. Condition number ratio (left) and backward error ratio (right) for scaled random 3×3 $P(\lambda)$ of degree 21.

For the Relative Pose 5pt problem, since the polynomial has a singular leading coefficient, we only consider the linearizations $D_1(\lambda, P)$, $\mathcal{T}_P(\lambda)$, and $C_1(\lambda)$, since $D_k(\lambda, P)$ is not a linearization of the polynomial in this case. We note that $P(\lambda)$ only has 10 finite eigenvalues. Moreover, the norms of the matrix coefficients of $P(\lambda)$ are similar. In this case the quotients of backward errors obtained associated with the three linearizations are close to 1 or less than 1, even when the polynomial is not scaled. Figure 4 shows the results after scaling the polynomial by dividing its matrix coefficients by $\max_{i=0:k}\{\|A_i\|_2\}$. In Table 2 we show data relative to this problem before and after scaling.

The plasma-drift problem is more complex. As Table 2 shows, the norms of the matrix coefficients of the matrix polynomial are not similar. Figure 5 shows the ratios of backward errors corresponding to the unscaled polynomial (left figure) and the polynomial scaled by dividing its matrix coefficients by $\max_{i=0:k}\{\|A_i\|_2\}$ (right figure). This scaling is clearly not enough for $\mathcal{T}_P(\lambda)$ to outperform the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ although the maximal backward error ratio for $\mathcal{T}_P(\lambda)$ is 35.28, which is very satisfactory. Thus, we consider two tropical scalings of the eigenvalue [6, 12, 27]. More explicitly, we consider the scaled polynomials $\beta P(\gamma\mu)$, where, for the tropical scaling 1, we use the parameters $\beta \approx 0.0081$ and $\gamma \approx 0.1$; and for the tropical scaling 2, we use the parameters $\beta \approx 0.0001$ and $\gamma \approx 9.8541$ (see Sections 2.1 and 2.2 in [27] on how to calculate these parameters). The obtained results are presented in Figure 6. We note that, in this case, using the tropical scaling 1, $\mathcal{T}_P(\lambda)$ outperforms the combined use of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ at one third of the eigenvalues (the ones with smallest modulus) but for large eigenvalues the behavior of $\mathcal{T}_P(\lambda)$ is not satisfactory and $D_1(\lambda, P)$ is clearly the winner; using the tropical scaling 2, $\mathcal{T}_P(\lambda)$ produces ratios close to 1 for the rest of the eigenvalues.

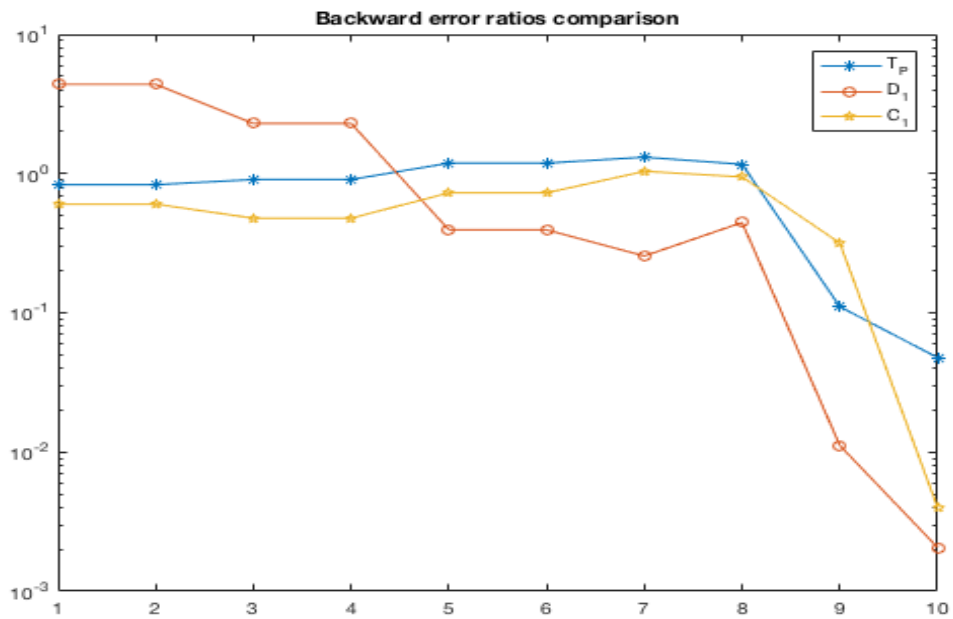


FIG. 4. Backward error ratios after scaling $P(\lambda)$ in Relative Pose 5pt Problem.

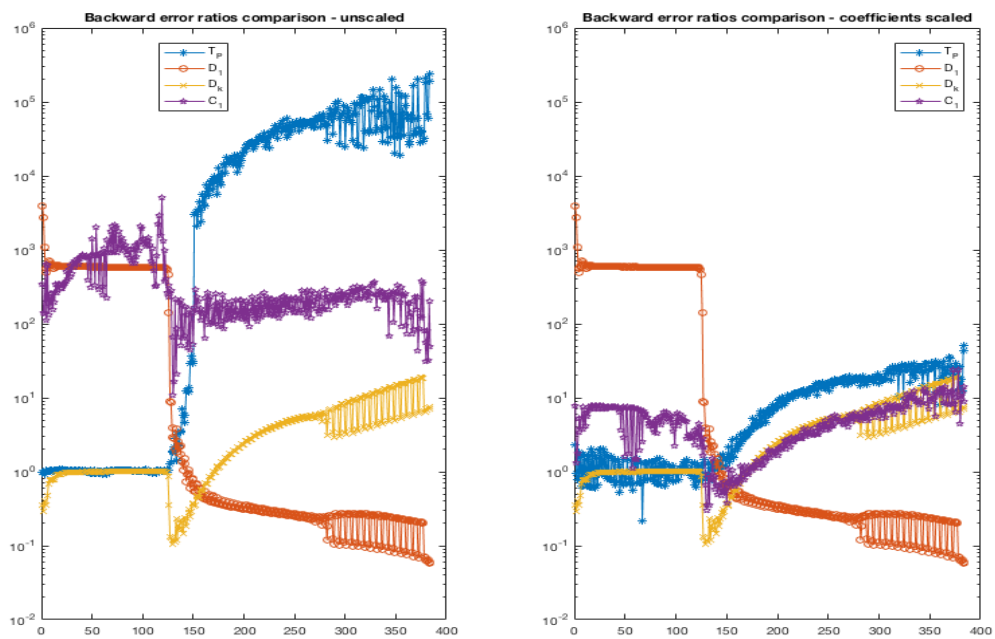


FIG. 5. Backward error ratios before scaling (left) and scaling only coefficients (right) in Plasma Drift Problem.

TABLE 2
 Backward error results for two problems in NLEVP: A collection of nonlinear eigenvalue problems.

Problem:	Relative Pose 5pt		Plasma-drift			
	Unscaled	Scaled	Unscaled	Scaled	Tropical 1	Tropical 2
$ \delta_{min} $	0.573	0.573	0.028	0.028	0.28	0.0028
$ \delta_{max} $	29.71	29.71	11.745	11.745	117.45	1.1919
$\ A_3\ _2$	0.708	0.336	12.698	0.0103	0.0001	1
$\ A_2\ _2$	1.29	0.614	5.304	0.0043	0.0004	0.042
$\ A_1\ _2$	1.88	0.891	1233.03	1.0	1	1
$\ A_0\ _2$	2.11	1	123.23	0.0999	1	0.01
$\min\{\eta_P/\eta_{D_1}\}$	0.002	0.002	0.0577	0.0577	0.61	0.011
$\max\{\eta_P/\eta_{D_1}\}$	4.372	4.372	3980.4	3980.4	23.52	406578.4
$\min\{\eta_P/\eta_{D_k}\}$	N/A	N/A	0.104	0.104	1.027	0.018
$\max\{\eta_P/\eta_{D_k}\}$	N/A	N/A	18.749	18.749	13278.8	41.74
$\min\{\eta_P/\eta_{T_P}\}$	0.033	0.072	0.857	0.359	0.34	0.66
$\max\{\eta_P/\eta_{T_P}\}$	1.18	1.45	190738.36	35.28	2764.58	30.59
$\min\{\eta_P/\eta_{C_1}\}$	0.112	0.044	13.79	0.341	0.28	0.2
$\max\{\eta_P/\eta_{C_1}\}$	1.19	1.28	1883.16	14.24	2682.44	29.7

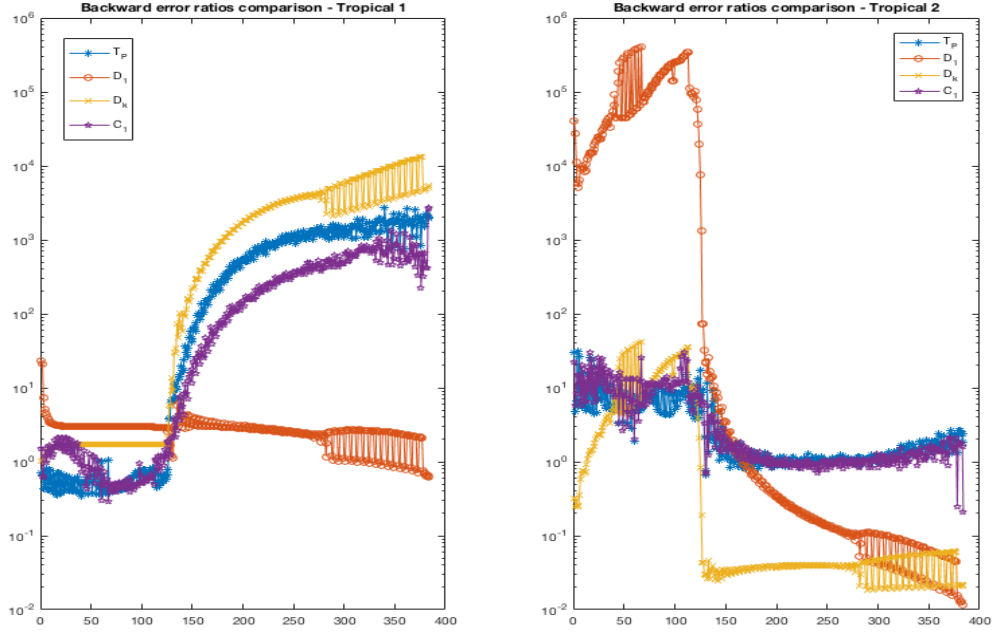


FIG. 6. Backward error ratios after Tropical scaling 1 (left) and after Tropical scaling 2 (right) in Plasma Drift Problem.

As a conclusion, in all the numerical experiments that we have run, we were able to find appropriate scalings of the matrix polynomial $P(\lambda)$ that allowed the use of $\mathcal{T}_P(\lambda)$ to produce small ratios of condition numbers for all nonzero finite simple eigenvalues of $P(\lambda)$, and small ratios of backward errors for approximate eigenpairs corresponding to all such eigenvalues. When the norms of the matrix coefficients of the matrix polynomial were similar, just scaling the matrix coefficients so that the largest one had norm 1 was enough for the numerical behavior of $\mathcal{T}_P(\lambda)$ to be comparable with the combined used of $D_1(\lambda, P)$

and $D_k(\lambda, P)$, i.e., of the other two block symmetric linearizations considered in this work. When the norms of the matrix coefficients of $P(\lambda)$ were not similar, an eigenvalue tropical scaling was necessary.

8. Conclusions. In this paper, we have studied for the first time the eigenvalue conditioning and backward errors of the linearization $\mathcal{T}_P(\lambda)$ (and $\mathcal{R}_P(\lambda)$) of an odd degree regular matrix polynomial $P(\lambda)$ and compared them with those of $D_1(\lambda, P)$, $D_k(\lambda, P)$, and $C_1(\lambda)$. We compared the condition number of a finite, nonzero, simple eigenvalue δ of $P(\lambda)$, when considered an eigenvalue of $P(\lambda)$ and an eigenvalue of each of the linearizations mentioned above. We have also compared the backward error of an approximate eigenpair (z, δ) of each of the previous linearizations with the backward error of an approximate eigenpair (x, δ) of $P(\lambda)$, where x has been recovered from z in a convenient way. As follows from the discussion below, when $P(\lambda)$ is symmetric or Hermitian and has odd degree, among the four linearizations studied, the linearization $\mathcal{T}_P(\lambda)$ (and $\mathcal{R}_P(\lambda)$) seems to be the most convenient to use in practice.

For every regular matrix polynomial $P(\lambda)$ as in (1.1), the pencil $C_1(\lambda)$ is a strong linearization of $P(\lambda)$. According to [15, 16] and our results in Section 6, this linearization exhibits good numerical behavior, both in terms of conditioning and backward errors, when $P(\lambda)$ has matrix coefficients with similar norm, specially after the matrix coefficients of $P(\lambda)$ are scaled by dividing each of them by $\max_{i=0:k} \{\|A_i\|_2\}$. The main disadvantage of this linearization is that it is not symmetric (Hermitian) when $P(\lambda)$ is, and preserving the structure of $P(\lambda)$ is convenient in order to preserve symmetries in the spectrum in the presence of rounding errors (see, for instance, Section 3 in [9]).

The pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ are both symmetric (Hermitian) when $P(\lambda)$ is. However, $D_1(\lambda, P)$ and $D_k(\lambda, P)$ are not always a linearization of a regular matrix polynomial $P(\lambda)$. It is well-known that $D_1(\lambda, P)$ is a linearization of $P(\lambda)$ if and only if A_0 is nonsingular and $D_k(\lambda, P)$ is a linearization of $P(\lambda)$ if and only if A_k is nonsingular. According to [16] and our results in Section 6, when $D_1(\lambda, P)$ and $D_k(\lambda, P)$ are linearizations of $P(\lambda)$, they exhibit good numerical behavior in terms of condition number but only in a certain range of eigenvalues, and as long as the matrix coefficients of $P(\lambda)$ have similar norms. More precisely, the eigenvalue δ has condition number ratio close to 1, when considered an eigenvalue of $D_1(\lambda, P)$, if $|\delta| \geq 1$, and when considered an eigenvalue of $D_k(\lambda, P)$, if $|\delta| \leq 1$. Similar conclusions can be drawn regarding the backward error, as shown in [15].

For every matrix polynomial $P(\lambda)$ of odd degree, the pencils $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ are strong linearizations of $P(\lambda)$. Moreover, they are symmetric (Hermitian) when $P(\lambda)$ is. These pencils exhibit good numerical behavior in terms of conditioning and backward error, for eigenvalues of any modulus, as long as the norms of the matrix coefficients are similar. Thus, we can conclude that, for regular symmetric or Hermitian matrix polynomials of odd degree, $\mathcal{T}_P(\lambda)$ and $\mathcal{R}_P(\lambda)$ are the ideal choice among the linearizations that we have considered in this paper. It is an open question though if we can find linearizations for matrix polynomials of even degree with properties similar to those of $\mathcal{T}_P(\lambda)$ (and $\mathcal{R}_P(\lambda)$).

9. Appendix A. Next we give the proof of Theorem 2.6.

Given a matrix polynomial $P(\lambda)$ as in (1.1), let C_P and B_P be the matrices defined in [9, Section 2]. The sign characteristic of a matrix polynomial is the sign characteristic of the pair (C_P, B_P) or any other pair unitarily similar to it (See Definition 2.3 in [9]). We recall that unitarily similar pairs have the same sign characteristic.

Let us denote $\Gamma_l := \text{diag}(1, \gamma, \gamma^2, \dots, \gamma^{l-1})$. It is easy to show that

$$C_{\tilde{P}} = \frac{1}{\gamma} \Gamma_k^{-1} C_P \Gamma_k, \quad B_{\tilde{P}} = \gamma \beta (\Gamma_k B_P \Gamma_k),$$

Denote by J_P and $J_{\tilde{P}}$ the Jordan forms of C_P and $C_{\tilde{P}}$, respectively, and let S_P and $S_{\tilde{P}}$ be nonsingular matrices such that

$$J_P = S_P^{-1} C_P S_P, \quad J_{\tilde{P}} = S_{\tilde{P}}^{-1} C_{\tilde{P}} S_{\tilde{P}}.$$

Let $N_P = S_P^* B_P S_P$ and $N_{\tilde{P}} = S_{\tilde{P}}^* B_{\tilde{P}} S_{\tilde{P}}$. Note that (J_P, N_P) is unitarily similar to (C_P, B_P) , and $(J_{\tilde{P}}, N_{\tilde{P}})$ is unitarily similar to $(C_{\tilde{P}}, B_{\tilde{P}})$. This implies that the sign characteristic of P and \tilde{P} is given

by the sign characteristic of (J_P, N_P) and $(J_{\bar{P}}, N_{\bar{P}})$, respectively. We have

$$J_{\bar{P}} = \frac{1}{\gamma}(S_{\bar{P}}^{-1}\Gamma_k^{-1})C_P(\Gamma_k S_{\bar{P}}) = \frac{1}{\gamma}(S_{\bar{P}}^{-1}\Gamma_k^{-1}S_P)J_P(S_P^{-1}\Gamma_k S_{\bar{P}}) = (S_P^{-1}\Gamma_k S_{\bar{P}})^{-1} \left(\frac{1}{\gamma}J_P \right) (S_{\bar{P}}^{-1}\Gamma_k S_{\bar{P}}).$$

Also,

$$N_{\bar{P}} = \gamma\beta(S_{\bar{P}}^*\Gamma_k)B_P(\Gamma_k S_{\bar{P}}) = \gamma\beta(S_{\bar{P}}^*\Gamma_k S_P^*)N_P(S_P^{-1}\Gamma_k S_{\bar{P}}) = (S_P^{-1}\Gamma_k S_{\bar{P}})^*(\gamma\beta N_P)(S_{\bar{P}}^{-1}\Gamma_k S_{\bar{P}}).$$

Therefore, $(J_{\bar{P}}, N_{\bar{P}})$ is unitarily similar to $(\frac{1}{\gamma}J_P, \gamma\beta N_P)$. Thus, in order to prove our claim, it is enough to see that $(\frac{1}{\gamma}J_P, \gamma\beta N_P)$ and (J_P, N_P) have the same sign characteristic.

Let T be such that $T^{-1}J_P T = J$ and $T^*N_P T = P_{\varepsilon, J}$, where J and $P_{\varepsilon, J}$ are as in Theorem 2.2 in [9]. Using the notation for α , β , and l_i , $i = 1, \dots, \beta$, in that theorem, let

$$L = \bigoplus_{i=1}^{\alpha} \sqrt{(\beta\gamma^{l_i})^{-1}}\Gamma_{l_i} \oplus \bigoplus_{i=\alpha+1}^{\beta} \left[\sqrt{(\beta\gamma^{l_i})^{-1}} \left(\Gamma_{\frac{l_i}{2}} \oplus \Gamma_{\frac{l_i}{2}} \right) \right].$$

Then,

$$(TL)^{-1} \left(\frac{1}{\gamma}J_P \right) TL = L^{-1} \left(\frac{1}{\gamma}J \right) L = J',$$

where J' has the same structure as J , with each eigenvalue in J multiplied by $\frac{1}{\gamma}$. Also,

$$(TL)^*(\gamma\beta N_P)TL = \gamma\beta L^*P_{\varepsilon, J}L = P_{\varepsilon, J}.$$

Thus, $(\frac{1}{\gamma}J_P, \gamma\beta N_P)$ and (J_P, N_P) have the same sign characteristic associated with the corresponding eigenvalues.

REFERENCES

- [1] M. AL-AMMARI AND F. TISSEUR, *Hermitian matrix polynomials with real eigenvalues of definite type. Part I: classification*, Linear Algebra Appl., 436 (2012), 3954-3973.
- [2] A. L. ANDREW, K. - W. ERIC CHU, AND P. LANCASTER, *Derivatives of eigenvalues and eigenvectors of matrix functions*, SIAM J. Matrix Anal. Appl. 14(4) (1993), 903-926.
- [3] E. N. ANTONIOU AND S. VOLOGIANNIDIS, *A new family of companion forms of polynomial matrices*, Electron. J. Linear Algebra, 11 (2004), 78-87.
- [4] Z. BAI, J. W. DEMMEL, J. J. DONGARRA, A. RUHE, AND H. A. VAN DER VORST, EDs., *Templates for the solution of algebraic eigenvalue problems: A practical guide*, Software Environ. Tools 11, SIAM, Philadelphia, 2000.
- [5] T. BETCKE, N.J. HIGHAM, V. MEHRMANN, C. SCHRÖDER, AND F. TISSEUR, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software 39(2), (2013), Art 7, 28 pp.
- [6] D. A. BINI, V. NOFERINI, AND M. SHARIFY, *Locating the eigenvalues of matrix polynomials*, SIAM J. Matrix Anal. Appl., 34(4) (2013), 1708-1727.
- [7] M. I. BUENO, J. BREEN, S. FORD, AND S. FURTADO, *On the sign characteristic of Hermitian linearizations in $\mathbb{DL}(P)$* , Linear Algebra Appl, 519 (2017), 73-101.
- [8] M.I. BUENO, F.M. DOPICO, S. FURTADO, AND M. RYCHNOVSKY, *Large vector spaces of block-symmetric strong linearizations*. Linear Algebra Appl., 477 (2015), 165-210.
- [9] M.I. BUENO, F.M. DOPICO, AND S. FURTADO, *Linearizations of Hermitian matrix polynomials preserving the sign characteristic*, SIAM J. Matrix Anal. Appl., 38 (2017), 249-272.
- [10] M.I. BUENO, F. DE TERÁN, AND F. M. DOPICO, *Recovery of eigenvectors and minimal bases of matrix polynomials from generalized Fiedler linearizations*, SIAM J. Matrix Anal. Appl., 32 (2011), 463-483.
- [11] J. P. DEDIEU AND F. TISSEUR, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), 71-94.
- [12] S. GAUBERT AND M. SHARIFY, *Tropical scaling of polynomial matrices*, in Positive systems, vol. 389 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 2009, 291-303.
- [13] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Indefinite Linear Algebra and Applications*, Springer Verlag, Basel, Switzerland, 2005.
- [14] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, SIAM, Philadelphia, 2009.

- [15] N.J. HIGHAM, R-C. LI, AND F. TISSEUR, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), 1218–1241.
- [16] N. J. HIGHAM, D. S. MACKEY, AND F. TISSEUR, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), 1005–1028.
- [17] N. J. HIGHAM, D. S. MACKEY, N. MACKEY, AND F. TISSEUR, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 29 (2006), 143–159.
- [18] T. KAILATH, *Linear Systems*, Prentice Hall, Inc., Englewood Cliffs, N. J., 1980.
- [19] P. LANCASTER, *Symmetric transformations of the companion matrix*, NABLA: Bulletin of the Malayan Math. Soc., 8 (1961), 146–148.
- [20] D. S. MACKEY, N. MACKEY, C. MEHL, AND V. MEHRMANN, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), 867–891.
- [21] D. S. MACKEY, N. MACKEY, C. MEHL, AND V. MEHRMANN, *Structured polynomial eigenvalue problems: good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28 (2006), 1029–1051.
- [22] D. S. MACKEY, N. MACKEY, C. MEHL, AND V. MEHRMANN, *Jordan structures of alternating matrix polynomials*, Linear Algebra Appl., 432 (2010), 971–1004.
- [23] D. S. MACKEY, *The continuing influence of Fiedler’s work on companion matrices*, Linear Algebra Appl. 439(4) (2013), 810–817.
- [24] V. MEHRMANN, V. NOFERINI, F. TISSEUR, AND H. XU, *On the sign characteristics of Hermitian polynomials*, Linear Algebra Appl., 511 (2016), 328–364.
- [25] , V. MEHRMANN AND H. VOSS, *Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27 (2004) 121–152.
- [26] C. B. MOLER AND G. W. STEWART, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), 241–256.
- [27] V. NOFERINI, M. SHARIFY, AND F. TISSEUR, *Tropical roots as approximations to eigenvalues of matrix polynomials*, SIAM J. Matrix Anal. Appl. 36 (2015), 138–157.
- [28] F. TISSEUR, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), 339–361.
- [29] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, SIAM Review, 43 (2001), 235–286.
- [30] L. ZENG AND Y. SU, *A backward stable algorithm for quadratic eigenvalue problems*, SIAM J. Matrix. Anal. Appl., 35 (2014), 499–516.