Calcolo manuscript No.
(will be inserted by the editor)

A block-symmetric linearization of odd degree matrix polynomials with optimal eigenvalue condition number and backward error

M. I Bueno · F. M. Dopico ·
S. Furtado · L. Medina

Received: date / Accepted: date

Abstract The standard way of solving numerically a polynomial eigenvalue problem (PEP) is to use a linearization and solve the corresponding generalized eigenvalue problem (GEP). In addition, if the PEP possesses one of the structures arising very often in applications, then the use of a linearization that preserves such structure combined with a structured algorithm for the GEP presents considerable numerical advantages. Block-symmetric linearizations have proven to be very useful for constructing structured linearizations of structured matrix polynomials. In this scenario, we analyze the eigenvalue condition numbers and backward errors of approximated eigenpairs of a block symmetric linearization that was introduced by Fiedler in 2003 for scalar polynomials and generalized to matrix polynomials by Antoniou and Vologiannidis in 2004. This analysis reveals that such linearization has much better numerical properties than any other block-symmetric linearization analyzed so far in the literature, including those in the well known vector

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 grant MTM-2015-65798-P (MINECO/FEDER, UE). 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). The research of S. Furtado was partially supported by project UID/MAT/04721/2013. The research of L. Medina was partially supported by NSF grant DMS-1358884.

M. I. Bueno
Department of Mathematics and College of Creative Studies, University of California Santa Barbara, Santa Barbara, CA, USA. E-mail: mbueno@ucsb.edu.

F. M. Dopico
Department of Mathematics, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28911 Leganés, Spain. E-mail: dopico@math.uc3m.es.

S. Furtado
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. E-mail: sbf@fep.up.pt.

L. Medina
Boston University, One Silber Way, Boston, MA 02215, USA. E-mail: medinal@bu.edu.

space $\mathbb{DL}(P)$ of block-symmetric linearizations. The main drawback of the analyzed linearization is that it can be constructed only for matrix polynomials of odd degree, but we believe that it will be possible to extend its use to even degree polynomials via some strategies in the near future.

Keywords backward error of an approximate eigenpair · block-symmetric linearization · conditioning of an eigenvalue · eigenvalue · eigenvector · linearization · matrix polynomial · strong linearization.

Mathematics Subject Classification (2000) 65F15 · 65F35.

1 Introduction

This paper deals with the polynomial eigenvalue problem (PEP), that is, with the problem of computing scalars $\lambda_0 \in \mathbb{C}$ and nonzero vectors $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ such that $P(\lambda_0)x = 0$ and $y^* P(\lambda_0) = 0$, where

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

is a matrix polynomial. We assume throughout the paper that $P(\lambda)$ is regular (that is, $\det(P(\lambda)) \not\equiv 0$), $A_k \neq 0$ (thus $P(\lambda)$ has degree k), and, in addition, that $A_0 \neq 0$ in order to avoid trivialities. The vectors x and y are said to be right and left eigenvectors of $P(\lambda)$ associated with the eigenvalue λ_0 of $P(\lambda)$. A matrix polynomial as in (1.1) may have also an eigenvalue at infinity, which, by definition, happens when zero is an eigenvalue of the reversal polynomial $\text{rev}P(\lambda) := \lambda^k P(1/\lambda)$.

PEPs arise in many applications, either directly or as approximations of other nonlinear eigenvalue problems [4,20,21,32,36,41]. However, the numerical algorithms currently available for the solution of PEPs are not completely satisfactory, despite of recent impressive advances in this area, both for small to medium size dense PEPs as well as for large-scale sparse PEPs [6,12,22,27,30,42,43,45]. The key drawback of current algorithms is that they do not guarantee *a priori* that the computed eigenvalues and eigenvectors are the exact eigenvalues and eigenvectors of a nearby matrix polynomial $P(\lambda) + \Delta P(\lambda) = \sum_{i=0}^{k} (A_i + \Delta A_i)\lambda^i$ that satisfies $\|\Delta A_i\|_2 / \|A_i\|_2 = O(\mathbf{u})$ for $i = 0, 1, \ldots, k$, where $\|\cdot\|_2$ denotes the spectral matrix norm and $\mathbf{u}$ is the unit roundoff of the computer or some tolerance fixed by the user. In contrast, current algorithms only guarantee that $\|\Delta A_i\|_2 / \max_i\{\|A_i\|_2\} = O(\mathbf{u})$, for $i = 0, 1, \ldots, k$, holds, which is not satisfactory for PEPs having matrix coefficients with very different norms. A remarkable exception is the algorithm for small to medium size *quadratic* PEPs in [45], which is heavily influenced by [22], and which guarantees $\|\Delta A_i\|_2 / \|A_i\|_2 = O(\mathbf{u})$ for $i = 0, 1, 2$ under some nontrivial assumptions on the computed eigenvectors and *at the cost of solving the problem twice in two different ways.* As a consequence of this discussion, the PEP is nowadays a very active area of research.

The standard way of solving the PEP is to construct a linear matrix polynomial, or matrix pencil, $L(\lambda)$ with the same eigenvalues (including multiplicities) as $P(\lambda)$, and solve the generalized eigenproblem for $L(\lambda)$, which is done with the QZ algorithm for moderate size problems [37] or with Krylov projection methods that take advantage of the particular structure of the considered $L(\lambda)$ for large-scale sparse problems [12,30,43]. That $L(\lambda)$ has the same finite eigenvalues and multiplicities as $P(\lambda)$ is guaranteed [19,20] if there exist two unimodular matrix polynomials (i.e., matrix polynomials with constant nonzero determinant), $U(\lambda)$ and $V(\lambda)$, such that $U(\lambda)L(\lambda)V(\lambda) = \mathrm{diag}(I_{(k-1)n}, P(\lambda))$, where I_m denotes the $m \times m$ identity matrix. Such $L(\lambda)$ is called a *linearization* of $P(\lambda)$. If, in addition, $\mathrm{rev}L(\lambda)$ is a linearization of $\mathrm{rev}P(\lambda)$, it is said that $L(\lambda)$ is a *strong linearization* of $P(\lambda)$ and, in this case, $L(\lambda)$ has the same finite and infinite eigenvalues and multiplicities as $P(\lambda)$ [13].

The importance of linearizations in solving PEPs numerically, the drawbacks of the numerical algorithms for PEPs, and the convenience of preserving in the linearizations the structures that the PEPs arising in applications often possess, have motivated in the last years an intense research on linearizations of matrix polynomials that can be constructed very easily, that allow an easy recovery of the eigenvectors of the polynomial from those of the linearization, and that preserve interesting structures. Among the references that triggered such recent research, we highlight [3,18,25,31,32]. For the sake of brevity, we omit a detailed list of the many references published recently on linearizations of matrix polynomials and, instead, we invite the reader to check the references included in [10,14,43]. Unfortunately, this explosion of new classes of linearizations has not been followed by the corresponding analyses of their numerical properties, i.e., by the study of the errors they produce when they are used for solving numerically a PEP, and the number of papers analyzing this question is still low [1,44,14,15,23,24,39,45].

Currently, there exist three complementary ways of measuring how "good" a linearization is for solving PEPs in terms of the errors produced by its use. The first one compares the condition numbers of individual (simple) eigenvalues in the linearization and in the original polynomial [1,24,39,45]. For this purpose (as a consequence of the discussion above) two types of condition numbers should be considered according whether the magnitude of the perturbations is measured in an absolute sense, i.e., as $\|\Delta A_i\|_2 / \max_i\{\|A_i\|_2\}$ for $i = 0, 1, \ldots, k$, or in a relative sense, i.e., as $\|\Delta A_i\|_2 / \|A_i\|_2$ for $i = 0, 1, \ldots, k$. The second way compares the residual backward errors of approximated individual right (x, λ_0), or left (y^*, λ_0), eigenpairs in the polynomial and in the linearization, where again the perturbations can be measured in the absolute and relative senses mentioned above [23,39, 45]. Finally, one can consider global backward error analyses [44,14,15,29], which try to prove that the whole set of computed eigenvalues is the whole set of exact eigenvalues of a matrix polynomial very close to the original one, assuming that this happens for the linearization. The global analyses have the advantage that they do not involve eigenvectors or any assumption on them, and they are valid *a priori* for all the eigenvalues, but have the drawback that, so far, they guarantee small absolute backward errors but do not guarantee small relative backward errors. Moreover, the bounds that are obtained in this way are less precise than those obtained with the other approaches.

When the preservation of structures is important for computational efficiency and/or for preserving the finite arithmetic symmetries in the spectrum, it is well-known [25,32] that block symmetric linearizations of matrix polynomials are fundamental, since they preserve directly the symmetric and Hermitian structures and can be modified to cope with other structures such as alternating, palindromic, or skew-symmetric [32].

In this scenario, this paper studies for the first time condition numbers of individual eigenvalues and residual backward errors of individual approximated eigenpairs of a block symmetric strong linearization, that we denote by $\mathcal{T}_P(\lambda)$, and prove that, under the standard mild condition $\max_i\{\|A_i\|_2\} \approx 1$ [44] (that can always be obtained by dividing $P(\lambda)$ by a number), $\mathcal{T}_P(\lambda)$ has in these respects much better properties than any other block symmetric linearization analyzed so far in the literature. The linearization $\mathcal{T}_P(\lambda)$ was originally introduced by Fiedler for scalar monic polynomials [18, p. 330], extended to regular matrix polynomials in [3], and slightly modified for dealing with other structured matrix polynomials apart from Hermitian ones in [33–35]. Its excellent properties with respect to global backward error analyses follow from the general results in [14,15]. Unfortunately, the linearization $\mathcal{T}_P(\lambda)$ is only valid for matrix polynomials with odd degree, and, although such polynomials appear in practice [4], this clearly limits the use of $\mathcal{T}_P(\lambda)$. A strategy to overcome this limitation, that has been already used in [29] for some non-block symmetric linearizations, is to introduce a zero leading coefficient $A_{k+1} = 0$ in $P(\lambda)$, construct $\mathcal{T}_P(\lambda)$ including such zero coefficient, and then to deflate exactly the extra eigenvalues introduced at infinity. We plan to investigate in the future the application of this strategy to $\mathcal{T}_P(\lambda)$.

The linearization $\mathcal{T}_P(\lambda)$ and its properties are revised in Section 4 and its eigenvalue condition numbers and backward errors of approximated eigenpairs are studied in Section 5, which includes the main results of this paper in Theorems 5.1 and 5.2. We emphasize that these results prove that, when the perturbations are measured in the absolute sense (as explained above) *the linearization $\mathcal{T}_P(\lambda)$ is optimal*, since its condition numbers and backward errors are "always equal up to a moderate constant depending on k" to those of the original matrix polynomial (under some assumptions in the case of backward errors). When the perturbations are measured in a relative sense, such equality holds, up to a moderate constant, provided that $\min\{\|A_k\|_2, \|A_0\|_2\} \approx \max_i\{\|A_i\|_2\} \approx 1$. The properties of $\mathcal{T}_P(\lambda)$ described in this paragraph are the same as those of any of the Frobenius companion forms [20], which are the standard linearizations used for solving numerically PEPs, but which do not preserve any of the structures of the PEPs arising in applications. The analysis of the eigenvalue condition numbers and backward errors of approximated eigenpairs of the Frobenius companion forms for perturbations measured in a relative sense can be found in [23,24]. We revise this analysis, improve some bounds, and complete it for perturbations measured in an absolute sense in Section 6.2.

We emphasize that $\mathcal{T}_P(\lambda)$ has much better properties concerning eigenvalue condition numbers and backward errors of approximated eigenpairs than the famous block symmetric linearizations in the space $\mathbb{DL}(P)$ introduced in [25,31], and whose condition numbers and backward errors for perturbations measured in a relative sense have been carefully analyzed in [23,24]. *These analyses have revealed a very inconvenient feature of the linearizations in $\mathbb{DL}(P)$ since, if one sticks to linearizations in $\mathbb{DL}(P)$, it is needed, in general, to solve the PEP twice with two linearizations in order to get reliable solutions with the same quality as those obtained from solving the PEP once with $\mathcal{T}_P(\lambda)$ or with the Frobenius companion forms.* On the other hand, as far as we know, the linearizations in $\mathbb{DL}(P)$ are the only block symmetric linearizations whose condition numbers and backward errors have been analyzed so far, which, among other reasons, have motivated considerable activity on these linearizations (see [2, 9, 16, 17, 38] and the references therein).

In Section 6.1, we revise the analysis presented in [23, 24] of the linearizations in $\mathbb{DL}(P)$, improve some bounds, and complete the analysis considering perturbations measured in an absolute sense. The reader can find the main results on eigenvalue condition numbers and backward errors of approximated eigenpairs for the linearizations in $\mathbb{DL}(P)$ in Theorems 6.1 and 6.2. These theorems confirm the conclusions established in [23, 24]: given a simple, finite, nonzero eigenvalue λ_0 of $P(\lambda)$ as in (1.1), if A_0 is nonsingular and $|\lambda_0| \geq 1$, then the first pencil in the standard basis of $\mathbb{DL}(P)$, denoted in this paper as $D_1(\lambda, P)$, has a condition number and backward error close to optimal among the linearizations of $P(\lambda)$ in $\mathbb{DL}(P)$, while if A_k is nonsingular and $|\lambda_0| \leq 1$, then the last pencil in the standard basis of $\mathbb{DL}(P)$, denoted in this paper as $D_k(\lambda, P)$, has the same property. Moreover, when $|\lambda_0| \gg 1$, $D_k(\lambda, P)$ behaves much worse than $P(\lambda)$ with respect to eigenvalue condition numbers, and the approximated eigenpairs recovered from $D_k(\lambda, P)$ lead to large backward errors in $P(\lambda)$. The same happens when $|\lambda_0| \ll 1$ and $D_1(\lambda, P)$ is used. Thus, if a regular matrix polynomial has eigenvalues with modulus less than and larger than 1, the use of these two pencils is mandatory when calculating eigenvalues and eigenvectors.

In addition to the sections and results described above, this paper contains the following material. Section 2 revises the definitions of eigenvalue condition numbers and backward errors of approximated eigenpairs of matrix polynomials. In Section 3 we introduce some technical results that are needed in proving the main results of the paper. Section 7 includes numerical tests that confirm the theoretical results presented in previous sections and illustrate the advantages of $\mathcal{T}_P(\lambda)$ with respect to the linearizations in $\mathbb{DL}(P)$. Finally some conclusions and lines of future research are discussed in Section 8.

Note that although the main applications of $\mathcal{T}_P(\lambda)$, and of some modified versions of this pencil [33–35], are for solving structured PEPs (symmetric, Hermitian, alternating, palindromic,..), we do not incorporate structure in the analysis. There are three reasons for that: for such structures very often the structured and unstructured condition numbers and backward errors are very similar [1, 7], to consider structures would make the analysis more complicated, and, finally, the study of all these structures would make the analysis very long.

2 Eigenvalue 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 also note that, in this paper, for the sake of brevity, we do not consider the homogeneous formulation of the polynomial eigenvalue problem and the corresponding condition number. Additionally, we only consider simple eigenvalues since these are essentially the only ones appearing in numerical practice.

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 λ_0 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 λ_0. A (relative) normwise condition number $\kappa_r(\lambda_0, P)$ of λ_0 can be defined by

$$\kappa_r(\lambda_0, P) := \lim_{\epsilon \to 0} \sup \left\{ \frac{|\Delta\lambda_0|}{\epsilon|\lambda_0|} : [P(\lambda_0 + \Delta\lambda_0) + \Delta P\,(\lambda_0 + \Delta\lambda_0)](x + \Delta x) = 0, \right.$$
$$\left. \|\Delta A_i\|_2 \leq \epsilon\, \omega_i,\ 0 \leq i \leq k \right\},$$

where $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ and ω_i, $0 \leq i \leq k$, are nonnegative weights that allow flexibility in how the perturbations are measured [24]. 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 with respect to perturbations in the matrix polynomial. We will measure these perturbations in two ways: 1) $\omega_i = \|A_i\|_2$ (relative perturbations); 2) $\omega_i = \max\limits_{0 \leq i \leq k} \{\|A_i\|_2\}$ (absolute perturbations).

Theorem 2.1 gives an explicit formula for $\kappa_r(\lambda_0, P)$. For a matrix polynomial $P(\lambda)$, we denote by $P'(\lambda)$ the first derivative of $P(\lambda)$ with respect to λ.

Theorem 2.1 *[39, Theorem 5] Let $P(\lambda)$ be a regular matrix polynomial of degree k. Let λ_0 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 λ_0. Then*

$$\kappa_r(\lambda_0, P) = \frac{(\sum_{i=0}^{k} |\lambda_0|^i \omega_i)\|y\|_2\|x\|_2}{|\lambda_0||y^* P'(\lambda_0)x|}. \tag{2.1}$$

Attending to the choice of weights ω_i mentioned above, we consider two versions of the normwise relative condition number:

– Relative-absolute condition number: $\omega_i = \max\limits_{0 \leq j \leq k} \{\|A_j\|_2\}$, $i = 0, \ldots, k$,

$$\kappa_{ra}(\lambda_0, P) = \frac{\max\limits_{0 \leq i \leq k} \{\|A_i\|_2\}(\sum_{i=0}^{k} |\lambda_0|^i)\|x\|_2\|y\|_2}{|\lambda_0||y^* P'(\lambda_0)x|}. \tag{2.2}$$

– Relative-relative condition number: $\omega_i = \|A_i\|_2$, $i = 0, \ldots, k$,

$$\kappa_{rr}(\lambda_0, P) = \frac{(\sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2)\|x\|_2\|y\|_2}{|\lambda_0||y^* P'(\lambda_0)x|}. \tag{2.3}$$

Remark 2.1 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. In fact, 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.

Observe that, if $\lambda_0 \neq 0$ is an eigenvalue of $P(\lambda)$, then the (left and right) eigenvectors of $P(\lambda)$ and $\mathrm{rev}P(\lambda) = \lambda^k P\left(\frac{1}{\lambda}\right)$ associated with λ_0 and $\frac{1}{\lambda_0}$, respectively, coincide. Thus, when using the two types of weights considered above for $P(\lambda)$ and $\mathrm{rev}P(\lambda)$, we obtain the following result, which is a simple consequence of (2.2) and (2.3). We should point out that, if $P(\lambda)$ as in (1.1) has degree k and $A_0 \neq 0$, then $\mathrm{rev}P(\lambda)$ has degree k as well.

Lemma 2.1 *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let λ_0 be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Then*

$$\kappa_{ra}(\lambda_0, P) = \kappa_{ra}\left(\frac{1}{\lambda_0}, \mathrm{rev}P\right) \quad \text{and} \quad \kappa_{rr}(\lambda_0, P) = \kappa_{rr}\left(\frac{1}{\lambda_0}, \mathrm{rev}P\right).$$

In some occasions, we will need to scale the matrix polynomial $P(\lambda)$ by dividing each of its matrix coefficients by $\max_{0\le i\le k}\{\|A_i\|_2\}$ in order to improve the bounds of the ratio of the condition numbers of an eigenvalue of $P(\lambda)$ and of a linearization. Notice that $\kappa_{ra}(\lambda_0, P)$ and $\kappa_{rr}(\lambda_0, P)$ are invariant under a scaling of this type.

The (relative) normwise backward error of an approximate (right) eigenpair (x, λ_0) of $P(\lambda)$, where λ_0 is finite, is defined by

$$\eta_r(x, \lambda_0, P) := \min\{\epsilon : (P(\lambda_0) + \Delta P(\lambda_0))x = 0, \quad \|\Delta A_i\|_2 \le \epsilon\, \omega_i,\ 0 \le i \le k\},$$

where $\Delta P(\lambda) = \sum_{i=0}^{k} \lambda^i \Delta A_i$ and ω_i, $0 \le i \le k$, are arbitrary nonnegative numbers that represent tolerances against which the perturbations of the matrix coefficients of $P(\lambda)$ will be measured [23].

Similarly, for an approximate left eigenpair (y^*, λ_0), we have

$$\eta_r(y^*, \lambda_0, P) := \min\{\epsilon : y^*(P(\lambda_0) + \Delta P(\lambda_0)) = 0, \quad \|\Delta A_i\|_2 \le \epsilon\, \omega_i,\ 0 \le i \le k\}.$$

The following result provides explicit formulas for the normwise backward error of approximate eigenpairs.

Theorem 2.2 *[39, Theorem 1] Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1). For a given approximate right eigenpair (x, λ_0) of $P(\lambda)$, where $x \in \mathbb{C}^n$ and $\lambda_0 \in \mathbb{C}$, the normwise backward error $\eta_r(x, \lambda_0, P)$ is given by*

$$\eta_r(x, \lambda_0, P) = \frac{\|P(\lambda_0)x\|_2}{(\sum_{i=0}^{k} |\lambda_0|^i \omega_i)\|x\|_2}.$$

For an approximate left eigenpair (y^*, λ_0), where $y \in \mathbb{C}^n$ and $\lambda_0 \in \mathbb{C}$, we have

$$\eta_r(y^*, \lambda_0, P) = \frac{\|y^* P(\lambda_0)\|_2}{(\sum_{i=0}^{k} |\lambda_0|^i \omega_i)\|y\|_2}.$$

As with the condition number, we will consider two cases, depending of the selection of the weights ω_i. We only display the expressions for right eigenpairs. The expressions for left eigenpairs can be obtained similarly.

– Relative-absolute backward error: $\omega_i = \max\limits_{0\le j\le k}\{\|A_j\|_2\}$, $i = 0, \ldots, k$,

$$\eta_{ra}(x, \lambda_0, P) = \frac{\|P(\lambda_0)x\|_2}{\max\limits_{i=0:k}\{\|A_i\|_2\}(\sum_{i=0}^{k} |\lambda_0|^i)\|x\|_2}.$$

– Relative-relative backward error: $\omega_i = \|A_i\|_2$, $i = 0, \ldots, k$,

$$\eta_{rr}(x, \lambda_0, P) = \frac{\|P(\lambda_0)x\|_2}{(\sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2)\|x\|_2}. \tag{2.4}$$

The next result follows from Theorem 2.2 and can be easily checked. A similar result can be obtained for left eigenvectors.

Lemma 2.2 *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 eigenpair (x, λ_0) of $P(\lambda)$, where $x \in \mathbb{C}^n$ and $0 \neq \lambda_0 \in \mathbb{C}$, we have that $(x, \frac{1}{\lambda_0})$ is an approximate right eigenpair of* $\mathrm{rev}P(\lambda)$ *and*

$$\eta_{ra}(x, \lambda_0, P) = \eta_{ra}\left(x, \frac{1}{\lambda_0}, \mathrm{rev}P\right) \quad \text{and} \quad \eta_{rr}(x, \lambda_0, P) = \eta_{rr}\left(x, \frac{1}{\lambda_0}, \mathrm{rev}P\right).$$

The scaling of the polynomial $P(\lambda)$ by dividing by $\max_{0 \leq i \leq k}\{\|A_i\|_2\}$ 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 $\eta_{ra}(x, \lambda_0, \tilde{P}) = \eta_{ra}(x, \lambda_0, P)$ and $\eta_{rr}(x, \lambda_0, \tilde{P}) = \eta_{rr}(x, \lambda_0, P)$, where $\tilde{P}(\mu)$ is the scaled polynomial, 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 linearization $\mathcal{T}_P(\lambda)$ that we are studying.

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

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

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 [23] for a proof of the second inequality).

Proposition 3.1 *For any complex $\ell \times m$ block-matrix $B = (B_{ij})$ we have*

$$\max_{i,j} \|B_{ij}\|_2 \leq \|B\|_2 \leq \sqrt{\ell m}\ \max_{i,j} \|B_{ij}\|_2. \tag{3.1}$$

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

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

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

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

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

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

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

Lemma 3.2 *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 λ_0 be a nonzero, finite eigenvalue of $P(\lambda)$, and let x and y be, respectively, a right and a left eigenvector of $P(\lambda)$ associated with λ_0. Then, for $i = 0 : k-1$,*

$$P_i(\lambda_0)x = -\lambda_0^{i-k} P^{k-i-1}(\lambda_0)x \quad \textit{and} \quad y^* P_i(\lambda_0) = -\lambda_0^{i-k} y^* P^{k-i-1}(\lambda_0).$$

Proof Note that, for $i = 0 : k-1$, we have $P(\lambda_0) = \lambda_0^{k-i} P_i(\lambda_0) + P^{k-i-1}(\lambda_0)$. Thus, the result follows taking into account that λ_0 is nonzero, $P(\lambda_0)x = 0$, and $y^* P(\lambda_0) = 0$. □

The next lemma can be easily verified.

Lemma 3.3 *Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1), let $\lambda_0 \in \mathbb{C}$, and let $P_i(\lambda)$ and $P^i(\lambda)$, $i = 0 : k$, be the matrix polynomials defined in (3.2) and (3.4). Then,*

$$\|P_i(\lambda_0)\|_2, \|P^i(\lambda_0)\|_2 \le \max_{j=0:k}\{\|A_j\|_2\} \sum_{j=0}^{i} |\lambda_0|^j, \quad i = 0 : k.$$

We close this section with two combinatorial lemmas that will be used later in the proofs of our main results. In the statement of the first lemma we use the following concept: Let $H = (H_{ij})$ be a $k \times k$ block-matrix with $H_{ij} \in \mathbb{C}^{n \times n}$. The block-transpose of H is the matrix $H^{\mathcal{B}} := (H_{ji})$, i.e., the block-transpose of H is the block-matrix whose block-entry in position (i,j) is H_{ji}.

In the following lemma and throughout the paper we will use the following notation:

$$\max_{i=0:k}\{1, \|A_i\|_2\} := \max\{1, \max_{0 \le i \le k}\{\|A_i\|_2\}\}.$$

Lemma 3.4 *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*

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

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 $\lambda_0 \in \mathbb{C}$. Then,

$$\|\Delta(\lambda_0)\|_2, \|\Delta^{\mathcal{B}}(\lambda_0)\|_2 \le \sqrt{d_1(\lambda_0)} \max_{i=0:k}\{1, \|A_i\|_2\}, \tag{3.6}$$

where

$$d_1(\lambda_0) = \sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\lambda_0|^{2s} \right). \tag{3.7}$$

Proof Let $\lambda_0 \in \mathbb{C}$ and $0 \neq x \in \mathbb{C}^n$. Taking into account the definition of $\Delta(\lambda)$, we get

$$\|\Delta(\lambda_0)x\|_2^2 = \sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r} \|x\|_2^2 + \sum_{r=1}^{\frac{k-1}{2}} |\lambda_0|^{2r} \|P_{k-2r}(\lambda_0)x\|_2^2 \tag{3.8}$$

$$\leq \left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} |\lambda_0|^{2r} \|P_{k-2r}(\lambda_0)\|_2^2 \right] \|x\|_2^2. \tag{3.9}$$

By Lemmas 3.3 and 3.1, we obtain

$$\begin{aligned} \left(\frac{\|\Delta(\lambda_0)x\|_2}{\|x\|_2} \right)^2 &\leq \sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} |\lambda_0|^{2r} \left(\max_{i=0:k} \{\|A_i\|_2\} \sum_{s=0}^{k-2r} |\lambda_0|^s \right)^2 \\ &\leq \left[\sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\lambda_0|^{2s} \right) \right] \max_{i=0:k} \{1, \|A_i\|_2\}^2 \\ &= d_1(\lambda_0) \max_{i=0:k} \{\|A_i\|_2, 1\}^2. \end{aligned}$$

Since $\|\Delta(\lambda_0)\|_2 = \sup_{x \neq 0} \frac{\|\Delta(\lambda_0)x\|_2}{\|x\|_2}$, the result for $\Delta(\lambda_0)$ in (3.6) follows. In order to show that the result is also true for $\Delta^{\mathcal{B}}(\lambda_0)$, notice that

$$\|\Delta^{\mathcal{B}}(\lambda_0)\|_2 = \|\Delta^{\mathcal{B}}(\lambda_0)^*\|_2$$

and $\Delta^{\mathcal{B}}(\lambda_0)^*$ is the block-vector $\Delta(\overline{\lambda_0})$ associated with $P(\lambda)^* = \sum_{i=0}^k \lambda^i A_i^*$. Thus,

$$\|\Delta^{\mathcal{B}}(\lambda_0)\|_2 \leq \sqrt{d_1(\overline{\lambda_0})} \max_{i=0:k} \{1, \|A_i\|_2\}.$$

Since $d_1(\overline{\lambda_0}) = d_1(\lambda_0)$, the result follows. □

Lemma 3.5 *Let $\lambda_0 \in \mathbb{C}$ be nonzero and let $k \geq 3$ be a positive odd integer. Let $d_1(\lambda)$ be as in (3.7). If $|\lambda_0| \leq 1$, then*

$$d_1(\lambda_0) \leq \frac{k+1}{2} + \frac{(k-1)^3}{2} |\lambda_0|^2. \tag{3.10}$$

Proof Assume that $|\lambda_0| \leq 1$. Then,

$$\begin{aligned} \sum_{r=1}^{\frac{k-1}{2}} \left((k-2r+1) \sum_{s=r}^{k-r} |\lambda_0|^{2s} \right) &\leq \frac{(k-1)^2}{2} \sum_{s=1}^{k-1} |\lambda_0|^{2s} \\ &= \frac{(k-1)^2}{2} |\lambda_0|^2 \sum_{s=1}^{k-1} |\lambda_0|^{2(s-1)} \leq \frac{(k-1)^3}{2} |\lambda_0|^2. \end{aligned}$$

Thus, from (3.7), (3.10) follows. □

4 The block-symmetric linearization $\mathcal{T}_P(\lambda)$.

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 $k \geq 3$.

The family of generalized Fiedler pencils (GFP) associated with a matrix polynomial $P(\lambda)$ was introduced in [3,11]. 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]:

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

Notice that, after permuting some block-rows and block-columns, the pencil $\mathcal{T}_P(\lambda)$ becomes the block Kronecker pencil $\mathcal{O}_1^P(\lambda)$ in [8] given by

$$\mathcal{O}_1^P(\lambda) := \left[\begin{array}{c|c} M(\lambda; P) & K_s(\lambda)^T \\ \hline K_s(\lambda) & 0 \end{array}\right],$$

where

$$M(\lambda; P) = \begin{bmatrix} \lambda A_k + A_{k-1} & & \\ & \ddots & \\ & & \lambda A_1 + A_0 \end{bmatrix}, \quad K_s(\lambda) = \begin{bmatrix} -I_n & \lambda I_n & & & \\ & -I_n & \lambda I_n & & \\ & & \ddots & \ddots & \\ & & & -I_n & \lambda I_n \end{bmatrix}.$$

Also, the following block-symmetric block-tridiagonal pencil associated with an $n \times n$ $P(\lambda)$ of odd degree k was introduced in [33, formula (5.1)] :

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

where

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

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

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

Thus, $\mathcal{R}_P(\lambda)$ is just $\mathcal{T}_P(\lambda)$ after reordering some block-rows and block-columns and adding some minus signs.

Because of the simple relationships between $\mathcal{T}_P(\lambda)$, $\mathcal{R}_P(\lambda)$ and $\mathcal{O}_1^P(\lambda)$, it can be proven that these three pencils have exactly the same properties with respect to conditioning of eigenvalues and backward error of approximate eigenpairs. Thus, in the rest of the paper, we focus only on $\mathcal{T}_P(\lambda)$.

It is known [11] that, for any matrix polynomial $P(\lambda)$ of odd degree k, the pencil $\mathcal{T}_P(\lambda)$ is a strong linearization of $P(\lambda)$. As mentioned in the introduction, this linearization has several attractive properties. In particular, it is easy to recover an eigenvector of $P(\lambda)$ associated with an eigenvalue λ_0 from an eigenvector of $\mathcal{T}_P(\lambda)$ 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 ith 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,*

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

where $\Delta^{\mathcal{B}}(\lambda)$, as in (3.5), denotes the block-transpose of $\Delta(\lambda)$.

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

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

where $P_i(\lambda)$, $i = 0 : k$, 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 ith 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. □

Note that the equations in (4.5) are a particular case of equation (2.11) in [23] (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 [31].

Theorem 4.1 *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1). Assume that λ_0 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 λ_0 if and only if $z = \Delta(\lambda_0)x$ (resp. $z = (\Delta^{\mathcal{B}}(\lambda_0))^* y$), for some right (resp. left) eigenvector x (resp. y) of $P(\lambda)$ associated with λ_0.*

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 $\{x_1, \ldots, x_m\}$ is a basis of the eigenspace associated with the eigenvalue λ_0 of $P(\lambda)$, then the vectors $\Delta(\lambda_0)x_1, \ldots, \Delta(\lambda_0)x_m$ are linearly independent and, since λ_0 has the same geometric multiplicity as an eigenvalue of $P(\lambda)$

and as an eigenvalue of $\mathcal{T}_P(\lambda)$, because $\mathcal{T}_P(\lambda)$ is a linearization of $P(\lambda)$, then $\{\Delta(\lambda_0)x_1, \ldots, \Delta(\lambda_0)x_m\}$ is a basis of the eigenspace associated with the eigenvalue λ_0 of $\mathcal{T}_P(\lambda)$. Thus, a vector z is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0 if and only if it is a linear combination of $\Delta(\lambda_0)x_1, \ldots, \Delta(\lambda_0)x_m$, that is, it is of the form $\Delta(\lambda_0)x$ for some eigenvector x of $P(\lambda)$ associated with λ_0.

The proof for the left eigenvectors is similar. □

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

Lemma 4.2 *Let $P(\lambda)$ be a matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$. Then,*

$$\mathcal{T}_P(\lambda) = \lambda D R \mathcal{T}_{\mathrm{rev}P}\left(\frac{1}{\lambda}\right) RD,$$

where R is as in (4.3) and

$$D = \mathrm{diag}(I_n, -I_n, I_n, -I_n, \ldots, I_n). \tag{4.6}$$

Moreover, $z \in \mathbb{C}^{nk}$ is a right (resp. left) eigenvector of $\mathcal{T}_P(\lambda)$ associated with a finite, nonzero eigenvalue λ_0 if and only if RDz is a right (resp. left) eigenvector of $\mathcal{T}_{\mathrm{rev}P}(\lambda)$ associated with the eigenvalue $\frac{1}{\lambda_0}$.

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 3.1.

Proposition 4.1 *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 $z = [z_1, \ldots, z_k]^{\mathcal{B}}$ be a nonzero $k \times 1$ block-vector partitioned into $n \times 1$ blocks. Then, defining $z_0 := 0$, we have

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

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

5 Conditioning and backward error of eigenvalues of $\mathcal{T}_P(\lambda)$.

Let $P(\lambda)$ be a matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$ and let $\mathcal{T}_P(\lambda)$ be the linearization of $P(\lambda)$ defined in (4.1). In this section we present the main results in this paper, Theorems 5.1 and 5.2, 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 $P(\lambda)$.

In the statements of our main results for $\mathcal{T}_P(\lambda)$, we assume that the polynomial $P(\lambda)$ has been scaled by dividing all the matrix coefficients of $P(\lambda)$ by $\max_{i=0:k}\{\|A_i\|_2\}$. To scale a polynomial in this way is a standard practice when using linearizations whose matrix coefficients, seen as block-matrices, have nonzero blocks equal to the matrix coefficients of $P(\lambda)$ and blocks equal to $\pm I_n$ to avoid the unbalance in terms of norms of the $\pm I_n$ blocks and the blocks equal to matrix coefficients of $P(\lambda)$ [44]. Linearizations of this type include the famous first Frobenius companion form and also the $\mathcal{T}_P(\lambda)$ linearization considered in this work. From the point of view of computational cost, it would be more convenient to divide by $\max_{i=0:k}\{\|A_i\|_F\}$, which leads to the same bounds up to some moderate dimensional constants.

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

$$\rho_1 := \frac{1}{\min\{\|A_k\|_2, \|A_0\|_2\}}. \tag{5.1}$$

Theorem 5.1 *Let* $P(\lambda)$ *be a regular matrix polynomial of odd degree* $k \geq 3$ *as in (1.1) with* $A_0 \neq 0$ *and* $\max_{i=0:k}\{\|A_i\|_2\} = 1$. *Assume that* λ_0 *is a simple, finite, nonzero eigenvalue of* $P(\lambda)$. *Let* ρ_1 *be as in (5.1). Then,*

$$1 \leq \frac{\kappa_{ra}(\lambda_0, \mathcal{T}_P)}{\kappa_{ra}(\lambda_0, P)} \leq (k-1)^3, \quad 1 \leq \frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)} \leq 2k^3\rho_1. \tag{5.2}$$

Moreover, if $\min\{|\lambda_0|, \frac{1}{|\lambda_0|}\} \leq \frac{1}{k-1}$, *then*

$$\frac{\kappa_{ra}(\lambda_0, \mathcal{T}_P)}{\kappa_{ra}(\lambda_0, P)} \leq 2k, \quad \frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)} \leq 4k\rho_1.$$

From Theorem 5.1 we conclude that, up to a dimensional constant, $\kappa_{ra}(\lambda_0, \mathcal{T}_P) = \kappa_{ra}(\lambda_0, P)$, which implies that $\mathcal{T}_P(\lambda)$ presents optimal behavior in terms of eigenvalue conditioning in the relative-absolute case since the sensitivity of eigenvalues is the same in $P(\lambda)$ and in $\mathcal{T}_P(\lambda)$. We note that 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* λ_0, in contrast with what happens to the first and last pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ in the standard basis of $\mathbb{DL}(P)$, which only present good condition number ratios when $|\lambda_0| \geq 1$ and $|\lambda_0| \leq 1$, respectively. In the relative-relative case, if the norms of the matrix coefficients of $P(\lambda)$ have similar magnitudes, all the comments above for the relative-absolute case apply as well.

Remark 5.1 As we will see from the proof of Theorem 5.1, if we do not assume that $P(\lambda)$ is scaled so that $\max_{i=0:k}\{\|A_i\|_2\} = 1$, we have

$$1 \leq \frac{\kappa_{ra}(\lambda_0, \mathcal{T}_P)}{\kappa_{ra}(\lambda_0, P)} \leq (k-1)^3 \mu_1', \quad \rho_2 \leq \frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)} \leq 2k^3 \mu_1, \tag{5.3}$$

with

$$\mu_1' := \frac{\max\{1, \max_{i=0:k}\{\|A_i\|_2^3\}\}}{\max_{i=0:k}\{\|A_i\|_2\}}, \quad \mu_1 := \frac{\max\{1, \max_{i=0:k}\{\|A_i\|_2^3\}\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}, \tag{5.4}$$

$$\rho_2 := \frac{\min\{\max\{1, \|A_k\|_2\}, \max\{1, \|A_0\|_2\}\}}{\max_{i=0:k}\{\|A_i\|_2\}}. \tag{5.5}$$

Next we consider the behavior of $\mathcal{T}_P(\lambda)$ in terms of backward errors of approximate right eigenpairs. Since, for not extremely large values of nk, the algorithm QZ, combined with adequate methods for computing eigenvectors [37], 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 $\eta_{ra}(x, \lambda_0, P)$ (resp. $\eta_{rr}(x, \lambda_0, P)$) is not much larger than $\eta_{ra}(z, \lambda_0, \mathcal{T}_P)$ (resp. $\eta_{rr}(z, \lambda_0, \mathcal{T}_P)$) (where (z, λ_0) denotes an approximate right eigenpair of $\mathcal{T}_P(\lambda)$ and (x, λ_0) denotes an approximate right eigenpair of $P(\lambda)$ obtained from (z, λ_0) in an appropriate way), we ensure small backward errors for the approximate right eigenpairs of $P(\lambda)$ as well. Note also that, for eigenvalues and eigenvectors computed with some structured algorithms, $\eta_{ra}(z, \lambda_0, \mathcal{T}_P)$ (resp. $\eta_{rr}(z, \lambda_0, \mathcal{T}_P)$) is also of order unit-roundoff most of the times, although there is not a formal proof of this fact. This happens, for instance, if $P(\lambda)$ is real symmetric and a combination of the algorithms in [5, 40] is used for computing the eigenvalues/eigenvectors of the real symmetric pencil $\mathcal{T}_P(\lambda)$. In these situations if we prove that $\eta_r(x, \lambda_0, P)/\eta_r(z, \lambda_0, \mathcal{T}_P)$ is moderate we ensure small backward errors for the approximate eigenpairs of $P(\lambda)$ as well. This discussion motivates the next theorem, in which an upper bound for the ratios $\frac{\eta_{ra}(x, \lambda_0, P)}{\eta_{ra}(z, \lambda_0, \mathcal{T}_P)}$ and $\frac{\eta_{rr}(x, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, \mathcal{T}_P)}$ is established. Note that Theorem 5.2 only deals with right eigenpairs. The result for left eigenpairs is similar and is omitted for brevity.

Theorem 5.2 *Let $P(\lambda)$ be a matrix polynomial of odd degree $k \geq 3$ as in (1.1) with $A_0 \neq 0$ and $\max_{i=0:k}\{\|A_i\|_2\} = 1$. Let (z, λ_0) be an approximate right eigenpair of $\mathcal{T}_P(\lambda)$, $x = (e_1^T \otimes I_n)z$ if $|\lambda_0| > 1$ and $x = (e_k^T \otimes I_n)z$ if $|\lambda_0| \leq 1$. Then, (x, λ_0) is an approximate right eigenpair of $P(\lambda)$ and*

$$\frac{\eta_{ra}(x, \lambda_0, P)}{\eta_{ra}(z, \lambda_0, \mathcal{T}_P)} \leq \sqrt{2(k-1)^3}\frac{\|z\|_2}{\|x\|_2}, \quad \frac{\eta_{rr}(x, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, \mathcal{T}_P)} \leq 2\sqrt{2}k^{\frac{3}{2}}\frac{\|z\|_2}{\|x\|_2}\rho_1. \tag{5.6}$$

Moreover, if $\min\{|\lambda_0|, \frac{1}{|\lambda_0|}\} \leq \frac{1}{k-1}$*, then*

$$\frac{\eta_{ra}(x, \lambda_0, P)}{\eta_{ra}(z, \lambda_0, \mathcal{T}_P)} \leq 2\sqrt{k}\frac{\|z\|_2}{\|x\|_2}, \quad \frac{\eta_{rr}(x, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, \mathcal{T}_P)} \leq 4\sqrt{k}\frac{\|z\|_2}{\|x\|_2}\rho_1. \tag{5.7}$$

Remark 5.2 In this remark, we investigate the quotient $\frac{\|z\|_2}{\|x\|_2}$ appearing in Theorem 5.2 *assuming that z and x are exact eigenvectors*. The goal is to provide some guidance on the expected values of the bounds obtained in Theorem 5.2, although this guidance needs to be taken with extreme caution since the approximate eigenvectors may be very different than the exact eigenvectors, especially in ill-conditioned problems.

From Theorem 4.1, x' is a right eigenvector of $P(\lambda)$ associated with the eigenvalue λ_0 if and only if $z = \Delta(\lambda_0)x'$ is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0. From Lemma 3.4, for $z = \Delta(\lambda_0)x'$ (which implies $x' = (e_k^T \otimes I_n)z$), we obtain

$$\frac{\|z\|_2^2}{\|x'\|_2^2} \leq d_1(\lambda_0) \max_{i=0:k}\{\|A_i\|_2, 1\}^2,$$

where $d_1(\lambda_0)$ is as in (3.7). If $|\lambda_0| \leq 1$, from Lemma 3.5, 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

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

If $|\lambda_0| > 1$, taking into account Lemma 4.2, the right eigenvectors of $\mathcal{T}_{\mathrm{rev}P}(\lambda)$ associated with the eigenvalue $\frac{1}{\lambda_0}$ are of the form RDz, where z is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0. From Theorem 4.1, $(e_k^T \otimes I_n)RDz$ is an eigenvector of $\mathrm{rev}P(\lambda)$ associated with the eigenvalue $\frac{1}{\lambda_0}$, that is, an eigenvector of $P(\lambda)$ associated with λ_0. Thus, from the calculations above, we also have

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

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

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

where $x = (e_k^T \otimes I_n)z$ if $|\lambda_0| \leq 1$ and $x = (e_1^T \otimes I_n)z$ if $|\lambda_0| > 1$. If we assume that an approximate right eigenvector z has a block structure similar to that of an exact right eigenvector of $\mathcal{T}_p(\lambda)$, then we can expect that the bound given above for the quotient $\frac{\|z\|_2}{\|x\|_2}$ appearing in Theorem 5.2 still holds and it is close to 1 for moderate k.

Note that, when the matrix coefficients of $P(\lambda)$ have similar norms, then $\rho_1 \approx 1$ and the bound on the quotient of relative-relative backward errors in Theorem 5.2 is expected to depend only on k. This happens for the quotient of relative-absolute backward errors for any $P(\lambda)$ with $\max\limits_{i=0:k}\{\|A_i\|_2\} = 1$.

From Theorem 5.2 and Remark 5.2 we conclude that, up to a constant depending on the degree of the matrix polynomial, we can expect $\eta_{ra}(z, \lambda_0, \mathcal{T}_P) \approx \eta_{ra}(x, \lambda_0, P)$, when x is recovered from z as explained in Theorem 5.2, which implies that $\mathcal{T}_P(\lambda)$ presents optimal behavior in terms of backward error in the relative-absolute case, that is, the same behavior as the polynomial itself. In the relative-relative case, if the norms of the matrix coefficients of $P(\lambda)$ have similar magnitudes, assuming that $\max\limits_{i=0:k}\{\|A_i\|_2\} = 1$, the backward error of an approximate eigenpair of $P(\lambda)$ is close to the backward error of the corresponding eigenpair of $\mathcal{T}_P(\lambda)$.

Remark 5.3 As we will see from the proof of Theorem 5.2, if we do not assume that $P(\lambda)$ is scaled so that $\max\limits_{i=0:k}\{\|A_i\|_2\} = 1$, we have

$$\frac{\eta_{ra}(x,\lambda_0,P)}{\eta_{ra}(z,\lambda_0,\mathcal{T}_P)} \leq \sqrt{2(k-1)^3}\nu_1' \frac{\|z\|_2}{\|x\|_2}, \quad \frac{\eta_{rr}(x,\lambda_0,P)}{\eta_{rr}(z,\lambda_0,\mathcal{T}_P)} \leq 2\sqrt{2}k^{\frac{3}{2}}\nu_1 \frac{\|z\|_2}{\|x\|_2},$$

where

$$\nu_1' := \frac{\max\{1, \max_{i=0:k}\{\|A_i\|_2^2\}\}}{\max\limits_{i=0:k}\{\|A_i\|_2\}}, \quad \nu_1 := \frac{\max\{1, \max_{i=0:k}\{\|A_i\|_2^2\}\}}{\min\{\|A_k\|_2, \|A_0\|_2\}}.$$

Note that Theorems 5.1 and 5.2 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 [23,24] and Section 6).

In the rest of this section we use the notation introduced in Sections 3 and 4.

5.1 Proof of Theorem 5.1

We start with a lemma in which we give explicit expressions for the two considered condition numbers (relative-absolute and relative-relative) of a simple, finite, nonzero eigenvalue λ_0 of the linearization $\mathcal{T}_P(\lambda)$ of a matrix polynomial $P(\lambda)$ of odd degree.

Lemma 5.1 *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1). Assume that λ_0 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,*

$$\kappa_{ra}(\lambda_0, \mathcal{T}_P) = \frac{\max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\}(|\lambda_0|+1)\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0)x_2\|_2}{|\lambda_0||x_1^* P'(\lambda_0)x_2|}, \quad (5.8)$$

$$\kappa_{rr}(\lambda_0, \mathcal{T}_P) = \frac{(|\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0)x_2\|_2}{|\lambda_0||x_1^* P'(\lambda_0)x_2|}, \quad (5.9)$$

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

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

$$\mathcal{T}_P'(\lambda)\Delta(\lambda) + \mathcal{T}_P(\lambda)\Delta'(\lambda) = e_k \otimes P'(\lambda). \quad (5.10)$$

By Theorem 4.1, the vector $z_1 = \Delta^{\mathcal{B}}(\lambda_0)^* x_1$ is a left eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0 and $z_2 = \Delta(\lambda_0)x_2$ is a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0. Evaluating the expression (5.10) at λ_0, premultiplying by z_1^*, and postmultiplying by x_2, we get

$$z_1^* \mathcal{T}_P'(\lambda_0)(\Delta(\lambda_0)x_2) = z_1^*(e_k \otimes P'(\lambda_0)x_2),$$

or, equivalently,

$$z_1^* \mathcal{T}_P'(\lambda_0)z_2 = x_1^* \Delta^{\mathcal{B}}(\lambda_0)(e_k \otimes P'(\lambda_0)x_2) = x_1^* P'(\lambda_0)x_2.$$

Now (5.8) and (5.9) follow from (2.2) and (2.3). $\square$

The following lemma will allow us to only consider eigenvalues λ_0 such that $|\lambda_0| \leq 1$ when proving Theorem 5.1.

Lemma 5.2 *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$, and let λ_0 be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Then,*

$$\kappa_{ra}(\lambda_0, \mathcal{T}_P) = \kappa_{ra}\left(\frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right), \quad \kappa_{rr}(\lambda_0, \mathcal{T}_P) = \kappa_{rr}\left(\frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right).$$

Proof Let $\mathcal{T}_P(\lambda) =: \lambda\mathcal{T}_1 - \mathcal{T}_0$. We prove the second claim. The first one can be proven similarly. From Lemma 4.2,

$$\mathcal{T}_{\mathrm{rev}P}(\lambda) = \lambda RD\mathcal{T}_P\left(\frac{1}{\lambda}\right)DR = RD(\mathcal{T}_1 - \lambda\mathcal{T}_0)DR =: \lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{T}}_0. \tag{5.11}$$

Moreover, if y and x are, respectively, a left and a right eigenvector of $\mathcal{T}_P(\lambda)$ associated with λ_0, then $\tilde{y} = RDy$ and $\tilde{x} = RDx$ are, respectively, a left and a right eigenvector of $\mathcal{T}_{\mathrm{rev}P}(\lambda)$ associated with $1/\lambda_0$. Then, taking into account (2.3) and the fact that the spectral norm is unitarily invariant, we have

$$\begin{aligned}
\kappa_{rr}\left(\frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right) &= \frac{\left(\left|\frac{1}{\lambda_0}\right| \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2\right)\|\tilde{y}\|_2\|\tilde{x}\|_2}{\left|\frac{1}{\lambda_0}\right| |\tilde{y}^*\tilde{\mathcal{T}}_1\tilde{x}|} \\
&= \frac{(\|\tilde{\mathcal{T}}_1\|_2 + |\lambda_0|\|\tilde{\mathcal{T}}_0\|_2)\|\tilde{y}\|_2\|\tilde{x}\|_2}{|\tilde{y}^*\tilde{\mathcal{T}}_1\tilde{x}|} = \frac{(\|\mathcal{T}_0\|_2 + |\lambda_0|\|\mathcal{T}_1\|_2)\|y\|_2\|x\|_2}{|y^*\mathcal{T}_0 x|} \\
&= \frac{(\|\mathcal{T}_0\|_2 + |\lambda_0|\|\mathcal{T}_1\|_2)\|y\|_2\|x\|_2}{|\lambda_0||y^*\mathcal{T}_1 x|} = \kappa_{rr}(\lambda_0, \mathcal{T}_P),
\end{aligned}$$

where the equality before the last equality follows from the fact that $(\lambda_0\mathcal{T}_1 - \mathcal{T}_0)x = 0$. □

Proof of Theorem 5.1. Let x_1 and x_2 be, respectively, a left and a right eigenvector of $P(\lambda)$ associated with λ_0. First we compare $\kappa_{ra}(\lambda_0, \mathcal{T}_P)$ with $\kappa_{ra}(\lambda_0, P)$. Taking into account Lemmas 2.1 and 5.2, we assume $|\lambda_0| \leq 1$, as otherwise we replace $P(\lambda)$ by $\mathrm{rev}P(\lambda)$ and λ_0 by $\frac{1}{\lambda_0}$. From (2.2) and Lemma 5.1, we have

$$\frac{\kappa_{ra}(\lambda_0, \mathcal{T}_P)}{\kappa_{ra}(\lambda_0, P)} = \frac{\max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\}(|\lambda_0| + 1)\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0)x_2\|_2}{\max\limits_{i=0:k}\{\|A_i\|_2\}(\sum_{i=0}^k |\lambda_0|^i)\|x_1\|_2\|x_2\|_2}. \tag{5.12}$$

By Proposition 4.1, we get

$$\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2 \leq 2\max_{i=0:k}\{1, \|A_i\|_2\}. \tag{5.13}$$

Moreover, by Proposition 3.1,

$$\max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\} \geq \max_{i=0:k}\{\|A_i\|_2\}.$$

Thus, from (5.12) and the inequalities above, we get

$$\frac{|\lambda_0| + 1}{\sum_{i=0}^k |\lambda_0|^i} Q_T \leq \frac{\kappa_{ra}(\lambda_0, \mathcal{T}_P)}{\kappa_{ra}(\lambda_0, P)} \leq \frac{2\max\limits_{i=0:k}\{1, \|A_i\|_2\}}{\max_{i=0:k}\{\|A_i\|_2\}} \frac{|\lambda_0| + 1}{\sum_{i=0}^k |\lambda_0|^i} Q_T, \tag{5.14}$$

where

$$Q_T := \frac{\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0) x_2\|_2}{\|x_1\|_2 \|x_2\|_2} \leq \|\Delta(\lambda_0)\|_2 \|\Delta^{\mathcal{B}}(\lambda_0)\|_2 \leq d_1(\lambda_0) \max_{i=0:k}\{1, \|A_i\|_2^2\}, \tag{5.15}$$

with $d_1(\lambda_0)$ as in (3.7). Note that the last inequality follows from Lemma 3.4.

We continue by computing the upper bound for $\frac{\kappa_{ra}(\lambda_0,\mathcal{T}_P)}{\kappa_{ra}(\lambda_0,P)}$. Taking into account Lemma 3.5, and since $|\lambda_0| \leq 1$, we obtain, for $k \geq 3$,

$$\begin{aligned}
\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} d_1(\lambda_0) &\leq \frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} \left(\frac{k+1}{2} + \frac{(k-1)^3}{2} |\lambda_0|^2 \right) \\
&\leq \frac{(k-1)^3}{2} \frac{(|\lambda_0|+1)(1+|\lambda_0|^2)}{\sum_{i=0}^k |\lambda_0|^i} \\
&= \frac{(k-1)^3}{2} \frac{\sum_{i=0}^3 |\lambda_0|^i}{\sum_{i=0}^k |\lambda_0|^i} \leq \frac{(k-1)^3}{2}.
\end{aligned}$$

A better bound for $\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} d_1(\lambda_0)$ can be obtained when $|\lambda_0| \leq \frac{1}{k-1}$. Taking into account Lemma 3.5, we have $d_1(\lambda_0) \leq k$, implying that

$$\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} d_1(\lambda_0) \leq k.$$

Taking into account (5.14), (5.15), the previous upper bounds for $\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} d_1(\lambda_0)$ and the fact that we are assuming that $\max_{i=0:k}\{\|A_i\|_2\} = 1$, the upper bound part of the theorem follows for the ratio of relative-absolute condition numbers.

Now we compare $\kappa_{rr}(\lambda_0, \mathcal{T}_P)$ with $\kappa_{rr}(\lambda_0, P)$. Again we can assume $|\lambda_0| \leq 1$. From 2.3 and Lemma 5.1, we have

$$\frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)} = \frac{(|\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0) x_2\|_2}{(\sum_{i=0}^k |\lambda_0|^i \|A_i\|_2)\|x_1\|_2\|x_2\|_2}. \tag{5.16}$$

By Propositions 3.1 and 4.1, and taking into account that each of the matrix coefficients of $\mathcal{T}_P(\lambda)$ contains an identity block, we get

$$\begin{aligned}
&(|\lambda_0|+1)\min\{\max\{1, \|A_k\|_2\}, \max\{1, \|A_0\|_2\}\} \leq \\
&\qquad |\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2 \leq 2(|\lambda_0|+1) \max_{i=0:k}\{1, \|A_i\|_2\}.
\end{aligned} \tag{5.17}$$

We also have

$$\sum_{i=0}^k |\lambda_0|^i \|A_i\|_2 \geq \|A_k\|_2 |\lambda_0|^k + \|A_0\|_2 \geq \min\{\|A_k\|_2, \|A_0\|_2\}(|\lambda_0|^k + 1). \tag{5.18}$$

Thus, from (5.16) and the inequalities above, we get

$$\rho_2 \frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} Q_T \leq \frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)} \leq 2 \frac{\max_{i=0:k}\{1, \|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{|\lambda_0|+1}{|\lambda_0|^k+1} Q_T, \tag{5.19}$$

where ρ_2 is as in (5.5) and Q_T is as in (5.15).

We now focus on the upper bound for $\frac{\kappa_{rr}(\lambda_0,\mathcal{T}_P)}{\kappa_{rr}(\lambda_0,P)}$. Taking into account Lemma 3.5, and since $|\lambda_0| \leq 1$, we obtain, for $k > 2$,

$$\frac{|\lambda_0|+1}{|\lambda_0|^k+1} d_1(\lambda_0) \leq 2d_1(\lambda_0) \leq k^3.$$

A better bound for $\frac{|\lambda_0|+1}{|\lambda_0|^k+1} d_1(\lambda_0)$ can be obtained when $|\lambda_0| \leq \frac{1}{k-1}$. Taking into account Lemma 3.5, $d_1(\lambda_0) \leq k$, implying that

$$\frac{|\lambda_0|+1}{|\lambda_0|^k+1} d_1(\lambda_0) \leq 2k.$$

Taking into account (5.19), (5.15), the previous upper bounds for $\frac{|\lambda_0|+1}{|\lambda_0|^k+1} d_1(\lambda_0)$ and the fact that $\max_{i=0:k}\{1, \|A_i\|_2\} = 1$, the upper bound part of the theorem follows for the ratio of relative-relative condition numbers.

Next we show the lower bounds for both ratios $\kappa_{ra}(\lambda_0,\mathcal{T}_P)/\kappa_{ra}(\lambda_0,P)$ and $\kappa_{rr}(\lambda_0,\mathcal{T}_P)/\kappa_{rr}(\lambda_0,P)$. From the definition of $\Delta(\lambda)$ in (3.5) we get

$$\|\Delta^{\mathcal{B}}(\lambda_0)^* x_1\|_2 \|\Delta(\lambda_0)x_2\|_2 = \|x_1^* \Delta^{\mathcal{B}}(\lambda_0)\|_2 \|\Delta(\lambda_0)x_2\|_2 \geq \|x_1\|_2 \|x_2\|_2 \sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r}.$$

Thus,

$$\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} Q_T \geq \frac{(|\lambda_0|+1)\sum_{r=0}^{\frac{k-1}{2}} |\lambda_0|^{2r}}{\sum_{i=0}^k |\lambda_0|^i} = 1,$$

and the result follows from (5.14) and (5.19). □

5.2 Proof of Theorem 5.2

The following lemma will allow us to only consider eigenvalues λ_0 such that $|\lambda_0| \leq 1$ when proving Theorem 5.2.

Lemma 5.3 *Let $P(\lambda)$ be a regular matrix polynomial of odd degree k as in (1.1) with $A_0 \neq 0$, and let (z, λ_0) be an approximate right eigenpair of $\mathcal{T}_P(\lambda)$, with $\lambda_0 \neq 0$. Then, $(RDz, \frac{1}{\lambda_0})$ is an approximate right eigenpair of $\mathcal{T}_{\mathrm{rev}P}(\lambda)$ and*

$$\eta_{ra}(z,\lambda_0,\mathcal{T}_P) = \eta_{ra}\left(RDz, \frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right), \quad \eta_{rr}(z,\lambda_0,\mathcal{T}_P) = \eta_{rr}\left(RDz, \frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right),$$

where R and D are as in (4.3) and (4.6), respectively.

Proof Let $\mathcal{T}_P(\lambda) =: \lambda\mathcal{T}_1 - \mathcal{T}_0$ and let $\mathcal{T}_{\mathrm{rev}P}(\lambda) := \lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{T}}_0$. Note that (5.11) holds. Then, from (2.4) and Lemma 4.2,

$$\begin{aligned}\eta_{rr}\left(RDz, \frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P}\right) &= \frac{\|\mathcal{T}_{\mathrm{rev}P}(\frac{1}{\lambda_0})RDz\|_2}{(\left|\frac{1}{\lambda_0}\right| \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2)\|RDz\|_2} = \frac{\|\frac{1}{\lambda_0}\mathcal{T}_P(\lambda_0)z\|_2}{(\left|\frac{1}{\lambda_0}\right| \|\tilde{\mathcal{T}}_1\|_2 + \|\tilde{\mathcal{T}}_0\|_2)\|z\|_2} \\ &= \frac{\|\mathcal{T}_P(\lambda_0)z\|_2}{(\|\mathcal{T}_0\|_2 + |\lambda_0|\|\mathcal{T}_1\|_2)\|z\|_2} = \eta_{rr}(z,\lambda_0,\mathcal{T}_P).\end{aligned}$$

The result for the relative-absolute case can be proven similarly. □

Proof of Theorem 5.2. Let (z, λ_0) be an approximate right eigenpair of $\mathcal{T}_P(\lambda) := \lambda\mathcal{T}_1 - \mathcal{T}_0$ and assume that $|\lambda_0| \leq 1$. Let $x := (e_k^T \otimes I_n)z$. Note that (x, λ_0) can be seen as an approximate right eigenpair of $P(\lambda)$. First we show the upper bounds in (5.6). We have

$$P(\lambda_0)x = P(\lambda_0)(e_k^T \otimes I_n)z = (e_k^T \otimes P(\lambda_0))z = \Delta^{\mathcal{B}}(\lambda_0)\mathcal{T}_P(\lambda_0)z,$$

where the last equality follows from Lemma 4.1. Thus,

$$\begin{aligned}
\frac{\eta_{ra}(x,\lambda_0,P)}{\eta_{ra}(z,\lambda_0,\mathcal{T}_P)} &= \frac{\|P(\lambda_0)x\|_2}{\max\limits_{i=0:k}\{\|A_i\|_2\}(\sum_{i=0}^k |\lambda_0|^i)\|x\|_2} \cdot \frac{\max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\}(1+|\lambda_0|)\|z\|_2}{\|\mathcal{T}_P(\lambda_0)z\|_2} \\
&\leq \frac{\|\Delta^{\mathcal{B}}(\lambda_0)\|_2\|\mathcal{T}_P(\lambda_0)z\|_2}{\max\limits_{i=0:k}\{\|A_i\|_2\}(\sum_{i=0}^k |\lambda_0|^i)\|x\|_2} \cdot \frac{\max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\}(1+|\lambda_0|)\|z\|_2}{\|\mathcal{T}_P(\lambda_0)z\|_2} \\
&= \frac{\|\Delta^{\mathcal{B}}(\lambda_0)\|_2 \max\{\|\mathcal{T}_1\|_2, \|\mathcal{T}_0\|_2\}(1+|\lambda_0|)}{\max\limits_{i=0:k}\{\|A_i\|_2\}\sum_{i=0}^k |\lambda_0|^i} \cdot \frac{\|z\|_2}{\|x\|_2} \\
&\leq 2\frac{\max\limits_{i=0:k}\{1, \|A_i\|_2\}}{\max\limits_{i=0:k}\{\|A_i\|_2\}} \frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} \frac{\|z\|_2}{\|x\|_2}\|\Delta^{\mathcal{B}}(\lambda_0)\|_2,
\end{aligned} \tag{5.20}$$

where the last inequality follows from Proposition 4.1. By Lemma 3.4, we have

$$\|\Delta^{\mathcal{B}}(\lambda_0)\|_2 \leq \sqrt{d_1(\lambda_0)} \max_{i=0:k}\{1, \|A_i\|_2\},$$

where $d_1(\lambda)$ is as in (3.7). Taking into account Lemma 3.5, since $|\lambda_0| \leq 1$, we have, for $k \geq 3$,

$$\begin{aligned}
\frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i}\sqrt{d_1(\lambda_0)} &\leq \frac{\sqrt{(|\lambda_0|+1)^2\left(\frac{k+1}{2} + \frac{(k-1)^3}{2}|\lambda_0|^2\right)}}{\sum_{i=0}^k |\lambda_0|^i} \\
&\leq \sqrt{\frac{(k-1)^3}{2}} \frac{\sqrt{(|\lambda_0|+1)^2(1+|\lambda_0|^2)}}{\sum_{i=0}^k |\lambda_0|^i} \\
&\leq \sqrt{\frac{(k-1)^3}{2}} \frac{\sum_{i=0}^2 |\lambda_0|^i}{\sum_{i=0}^k |\lambda_0|^i} \leq \sqrt{\frac{(k-1)^3}{2}}.
\end{aligned}$$

Thus, the first upper bound in (5.6) follows by combining the two previous bounds with (5.20).

Similarly,

$$\begin{aligned}
\frac{\eta_{rr}(x,\lambda_0,P)}{\eta_{rr}(z,\lambda_0,\mathcal{T}_P)} &= \frac{\|P(\lambda_0)x\|_2}{(\sum_{i=0}^k |\lambda_0|^i\|A_i\|_2)\|x\|_2} \cdot \frac{(|\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)\|z\|_2}{\|\mathcal{T}_P(\lambda_0)z\|_2} \\
&\leq \frac{\|\Delta^{\mathcal{B}}(\lambda_0)\|_2\|\mathcal{T}_P(\lambda_0)z\|_2}{(\sum_{i=0}^k |\lambda_0|^i\|A_i\|_2)\|x\|_2} \cdot \frac{(|\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)\|z\|_2}{\|\mathcal{T}_P(\lambda_0)z\|_2} \\
&= \frac{\|\Delta^{\mathcal{B}}(\lambda_0)\|_2(|\lambda_0|\|\mathcal{T}_1\|_2 + \|\mathcal{T}_0\|_2)}{\sum_{i=0}^k |\lambda_0|^i\|A_i\|_2} \cdot \frac{\|z\|_2}{\|x\|_2} \\
&\leq 2\frac{\max_{i=0:k}\{1, \|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{|\lambda_0|+1}{|\lambda_0|^k+1} \frac{\|z\|_2}{\|x\|_2}\|\Delta^{\mathcal{B}}(\lambda_0)\|_2,
\end{aligned} \tag{5.21}$$

where the last inequality follows from (5.18) and the second inequality in (5.17). Moreover, since $|\lambda_0| \leq 1$, we have, for $k \geq 3$, using again Lemma 3.5,

$$\frac{|\lambda_0|+1}{|\lambda_0|^k+1}\sqrt{d_1(\lambda_0)} \leq \frac{|\lambda_0|+1}{|\lambda_0|^k+1}\sqrt{\frac{k+1}{2}+\frac{(k-1)^3}{2}|\lambda_0|^2} \leq \sqrt{2}k^{3/2}.$$

Thus, the second upper bound in (5.6) follows by combining the previous bound with (5.21) and the bound for $\|\Delta^{\mathcal{B}}(\lambda_0)\|_2$ in Lemma 3.4.

If $|\lambda_0| \leq \frac{1}{k-1}$, a better bound can be obtained. In this case,

$$\frac{|\lambda_0|+1}{\sum_{i=0}^{k}|\lambda_0|^i}\sqrt{d_1(\lambda_0)} \leq \sqrt{k}, \quad \frac{|\lambda_0|+1}{|\lambda_0|^k+1}\sqrt{d_1(\lambda_0)} \leq 2\sqrt{k},$$

implying the upper bounds for the ratios of backward errors in (5.7).

Now suppose that $|\lambda_0| > 1$. From Lemmas 2.2 and 5.3, we have

$$\frac{\eta_{rr}((e_1^T \otimes I_n)z, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, \mathcal{T}_P)} = \frac{\eta_{rr}((e_k^T \otimes I_n)RDz, \frac{1}{\lambda_0}, \mathrm{rev}P)}{\eta_{rr}(RDz, \frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P})}.$$

Taking into account Lemma 4.2 and since $|\frac{1}{\lambda_0}| < 1$, by the part of Theorem 5.2 already proven, we have that

$$\begin{aligned}\frac{\eta_{rr}((e_k^T \otimes I_n)RDz, \frac{1}{\lambda_0}, \mathrm{rev}P)}{\eta_{rr}(RDz, \frac{1}{\lambda_0}, \mathcal{T}_{\mathrm{rev}P})} &= 2\sqrt{2}k^{3/2}\frac{\|RDz\|_2}{\|(e_k^T \otimes I_n)RDz\|_2}\rho_1 \\ &= 2\sqrt{2}k^{3/2}\frac{\|z\|_2}{\|(e_1^T \otimes I_n)z\|_2}\rho_1.\end{aligned}$$

A similar observation applies when the relative-absolute backward error is considered.

6 Conditioning and backward error of eigenvalues 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 5.2 for the linearizations $D_1(\lambda, P)$, $D_k(\lambda, P)$ (which are the first and the last pencils in the standard basis of the vector space of block-symmetric pencils $\mathbb{DL}(P)$ introduced in [31]), and for $C_1(\lambda)$ (called the first Frobenius companion form, and which is the linearization used by default by MATLAB when solving a polynomial eigenvalue problem). We introduce formally these pencils below. We note that some of the claims in the results presented in this section were previously obtained in [23] and [24]. We include them for completeness with the goal of comparing the conditioning and backward error of the eigenvalues of $\mathcal{T}_P(\lambda)$ and those of the linearizations $D_1(\lambda, P)$, $D_k(\lambda, P)$, and $C_1(\lambda)$, and of discussing the advantages of $\mathcal{T}_P(\lambda)$ over these linearizations. With respect to the conditioning, we give some improvements on the bounds proven in [24], 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 [25,31].

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

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

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

$$C_1(\lambda) := \lambda \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} - \begin{bmatrix} -A_{k-1} & -A_{k-2} & \cdots & -A_0 \\ I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I_n & 0 \end{bmatrix}.$$

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 [20]. The block-symmetric pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ have been thoroughly studied recently in [9, 23–25, 31, 32], although they were introduced as early as in [28].

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:

$$\Lambda(\lambda) = \begin{bmatrix} \lambda^{k-1} & \cdots \lambda & 1 \end{bmatrix}^T. \tag{6.1}$$

Lemma 6.1 *[31, Theorems 3.8 and 3.14] Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of degree k and λ_0 be a finite eigenvalue of $P(\lambda)$. Let $L(\lambda) = D_i(\lambda, P)$, with $i \in \{1, k\}$, be a linearization of $P(\lambda)$. A vector z is a right eigenvector of $L(\lambda)$ associated with λ_0 if and only if $z = \Lambda(\lambda_0) \otimes x$ for some right eigenvector x of $P(\lambda)$ associated with λ_0. Similarly, a vector ω is a left eigenvector of $L(\lambda)$ associated with λ_0 if and only if $\omega = \overline{\Lambda(\lambda_0)} \otimes y$ for some left eigenvector y of $P(\lambda)$ associated with λ_0.*

Lemma 6.2 *[24, Section 1 and Lemma 7.2] Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of degree k and λ_0 be a finite eigenvalue of $P(\lambda)$. A vector z is a right eigenvector of $C_1(\lambda)$ associated with λ_0 if and only if $z = \Lambda(\lambda_0) \otimes x$ for some right eigenvector x of $P(\lambda)$ associated with λ_0.*

A vector ω is a left eigenvector of $C_1(\lambda)$ associated with λ_0 if and only if $\omega^* = y^*[I_n, P_1(\lambda_0), \ldots, P_{k-2}(\lambda_0), P_{k-1}(\lambda_0)]$, for some left eigenvector y of $P(\lambda)$ associated with λ_0, where $P_i(\lambda)$ is as in (3.2). Thus, if ω is a left eigenvector of $C_1(\lambda)$ with eigenvalue λ_0, then $y = (e_1^T \otimes I_n)\omega$ is a left eigenvector of $P(\lambda)$ with eigenvalue λ_0. Moreover, any left eigenvector y of $P(\lambda)$ with eigenvalue λ_0 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 eigenvalues 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 [23,24]. Moreover, we sharpen one of the results in [24].

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

In [24], it was shown that, for a simple, finite, nonzero eigenvalue λ_0 of a regular $P(\lambda)$ as in (1.1) with $A_0 \neq 0$, the relative-relative condition number of λ_0 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 $|\lambda_0| \geq 1$ (resp. $|\lambda_0| \leq 1$), provided that

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

is of order 1, which happens, in particular, when all matrix coefficients of $P(\lambda)$ have similar norms. In this case, these optimal condition numbers are close to the relative-relative condition number of the polynomial itself.

The combination of Theorems 4.4 and 4.5 in [24] provides a lower and an upper bound for $\frac{\kappa_{rr}(\lambda_0, D_t)}{\kappa_{rr}(\lambda_0, P)}$, when either $t = 1$ and $|\lambda_0| \geq 1$, or $t = k$ and $|\lambda_0| \leq 1$, namely,

$$\left(\frac{2\sqrt{k}}{k+1}\right)\frac{1}{\rho} \leq \frac{\kappa_{rr}(\lambda_0, D_t)}{\kappa_{rr}(\lambda_0, P)} \leq \sqrt{k^7}\rho^2.$$

Using the same techniques as those applied to compute the bounds in Theorem 5.1, next we deduce sharper bounds for the quotient $\frac{\kappa_{rr}(\lambda_0, D_t)}{\kappa_{rr}(\lambda_0, P)}$ than those provided in [24] and compute lower and upper bounds for $\frac{\kappa_{ra}(\lambda_0, D_t)}{\kappa_{ra}(\lambda_0, P)}$. 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 6.1 *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Assume that λ_0 is a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $\ell \in \{1, k\}$ and suppose that A_0 is nonsingular if $\ell = 1$, and A_k is nonsingular if $\ell = k$. Let ρ be as in (6.2). Then,*

$$\left.\begin{array}{l} \max\{1, |\lambda_0|^{k-1}\}, \text{ if } \ell = k \\ \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}, \text{ if } \ell = 1 \end{array}\right\} \leq \frac{\kappa_{ra}(\lambda_0, D_\ell)}{\kappa_{ra}(\lambda_0, P)} \leq \left\{\begin{array}{l} k^2 \max\{1, |\lambda_0|^{k-1}\}, \text{ if } \ell = k, \\ k^2 \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}, \text{ if } \ell = 1. \end{array}\right.$$

$$\left.\begin{array}{l} \max\{1, |\lambda_0|^{k-1}\}\frac{1}{\rho}, \text{ if } \ell = k, \\ \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}\frac{1}{\rho}, \text{ if } \ell = 1. \end{array}\right\} \leq \frac{\kappa_{rr}(\lambda_0, D_\ell)}{\kappa_{rr}(\lambda_0, P)} \leq \left\{\begin{array}{l} k^2\rho \max\{1, |\lambda_0|^{k-1}\}, \text{ if } \ell = k, \\ k^2\rho \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}, \text{ if } \ell = 1. \end{array}\right.$$

Proof We only prove the bounds for the quotient of relative-relative condition numbers. The relative-absolute case can be proven similarly. Let x and y be a right and a left eigenvector of $P(\lambda)$ associated with λ_0, respectively. Let $\Lambda(\lambda)$ be as in (6.1). Let $\ell \in \{1, k\}$, and define $D_\ell(\lambda, P) := L_1^\ell \lambda - L_0^\ell$. Taking into account Theorem 3.2 in [24], we have

$$\frac{\kappa_{rr}(\lambda_0, D_\ell)}{\kappa_{rr}(\lambda_0, P)} = \frac{(|\lambda_0| \|L_1^\ell\|_2 + \|L_0^\ell\|_2)\|\Lambda(\lambda_0)\|_2^2}{|\lambda_0|^{k-\ell} \sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2}. \tag{6.3}$$

By Proposition 3.1 and taking into account that all block-entries of L_1^ℓ and L_0^ℓ are either 0 or matrix coefficients of $P(\lambda)$, we obtain

$$|\lambda_0| \|L_1^\ell\|_2 + \|L_0^\ell\|_2 \le (|\lambda_0| + 1)k \max_{i=0:k}\{\|A_i\|_2\}. \tag{6.4}$$

Thus, from (5.18) and (6.3), we obtain

$$\frac{\kappa_{rr}(\lambda_0, D_\ell)}{\kappa_{rr}(\lambda_0, P)} \le \frac{k \max_{i=0:k}\{\|A_i\|_2\}}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{(|\lambda_0| + 1)\sum_{i=0}^{k-1} |\lambda_0|^{2i}}{(|\lambda_0|^k + 1)|\lambda_0|^{k-\ell}} = k\rho \frac{\sum_{i=0}^{2k-1} |\lambda_0|^i}{(|\lambda_0|^k + 1)|\lambda_0|^{k-\ell}}. \tag{6.5}$$

Clearly, if $|\lambda_0| = 1$, the upper bound in the statement follows. Now suppose that $|\lambda_0| \ne 1$. We have

$$\begin{aligned}
\frac{\kappa_{rr}(\lambda_0, D_\ell)}{\kappa_{rr}(\lambda_0, P)} &\le k\rho \frac{|\lambda_0|^{2k} - 1}{(|\lambda_0| - 1)(|\lambda_0|^k + 1)|\lambda_0|^{k-\ell}} = k\rho \frac{|\lambda_0|^k - 1}{(|\lambda_0| - 1)|\lambda_0|^{k-\ell}} \\
&\le k\rho \frac{|\lambda_0|^{k-1} + \cdots + 1}{|\lambda_0|^{k-\ell}} \le k^2 \rho \frac{\max\{1, |\lambda_0|^{k-1}\}}{|\lambda_0|^{k-\ell}} \\
&= \begin{cases} k^2 \rho \max\{1, |\lambda_0|^{k-1}\}, & \text{if } \ell = k, \\ k^2 \rho \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}, & \text{if } \ell = 1. \end{cases}
\end{aligned}$$

Thus, the result for the upper bound in the relative-relative case follows.

Now we show the lower bound in the statement. From Proposition 3.1 and taking into account the block-structure of $D_\ell(\lambda, P)$, we obtain

$$(|\lambda_0| + 1)\min\{\|A_k\|_2, \|A_0\|_2\} \le |\lambda_0| \|L_1^\ell\|_2 + \|L_0^\ell\|_2.$$

Thus, from (6.3),

$$\begin{aligned}
\frac{\kappa_{rr}(\lambda_0, D_\ell)}{\kappa_{rr}(\lambda_0, P)} &\ge \frac{(|\lambda_0| + 1)\min\{\|A_k\|_2, \|A_0\|_2\}\sum_{i=0}^{k-1} |\lambda_0|^{2i}}{\max_{i=0:k}\{\|A_i\|_2\}|\lambda_0|^{k-\ell}\sum_{i=0}^{k} |\lambda_0|^i} \\
&= \frac{1}{\rho} \frac{\sum_{i=0}^{2k-1} |\lambda_0|^i}{|\lambda_0|^{k-\ell}\sum_{i=0}^{k} |\lambda_0|^i} \ge \frac{1}{\rho} \begin{cases} \max\{1, |\lambda_0|^{k-1}\}, & \text{if } \ell = k, \\ \max\{1, \frac{1}{|\lambda_0|^{k-1}}\}, & \text{if } \ell = 1. \end{cases} \quad \square
\end{aligned}$$

We notice that the previous theorem implies that, if k is moderate, $|\lambda_0| \ge 1$, and A_0 is nonsingular (resp. $|\lambda_0| \le 1$ and A_k is nonsingular), then $\kappa_{ra}(\lambda_0, D_1) = \kappa_{ra}(\lambda_0, P)$ (resp. $\kappa_{ra}(\lambda_0, D_k) = \kappa_{ra}(\lambda_0, P)$), up to some dimensional constant. In the relative-relative case, the same conclusion is obtained if, in addition, $\rho \approx 1$. However, if $|\lambda_0| < 1$ (resp. $|\lambda_0| > 1$), the ratio of condition numbers when λ_0 is considered an eigenvalue of $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) grows as $|\lambda_0|$ decreases (resp. increases) in both the relative-absolute and the relative-relative case, even if $\rho \approx 1$. Thus, $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) cannot be used to compute the eigenvalues with small (resp. large) modulus.

Remark 6.1 Based on Theorems 5.1 and 6.1, 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, while we can use $\mathcal{T}_P(\lambda)$ to compute any eigenvalues of $P(\lambda)$, regardless of their modulus, we are forced to use 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. This is the case since $\kappa_{ra}(\lambda_0, \mathcal{T}_P) = \kappa_{ra}(\lambda_0, P)$, up to a moderate constant (which is often very pessimistic) for every λ_0, while $\kappa_{ra}(\lambda_0, D_1)$ (resp. $\kappa_{ra}(\lambda_0, D_k)$) can be much larger than $\kappa_{ra}(\lambda_0, P)$ if $|\lambda_0| < 1$ (resp. $|\lambda_0| > 1$). A similar observation applies in the relative-relative case if, in addition, we assume $\rho \approx 1$ (notice that $\rho = \rho_1$ when $\max_{i=0:k}\{\|A_i\|_2\} = 1$). We also observe that $D_1(\lambda, P)$ (resp. $D_k(\lambda, P)$) is a strong 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 strong 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 $D_\ell(\lambda, P)$, $\ell \in \{1, k\}$, with the backward error of a certain approximate eigenpair of $P(\lambda)$. The result for the relative-relative case was obtained in [23, Corollary 3.11]. A few extra computations show the relative-absolute case.

Theorem 6.2 *Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$. Let $\ell \in \{1, k\}$ and suppose that A_0 is nonsingular if $\ell = 1$, and A_k is nonsingular if $\ell = k$. Let ρ be as in (6.2). Let (z, λ_0) be an approximate right eigenpair of $D_\ell(\lambda, P)$, with λ_0 nonzero and finite. Then, for $z_\ell = (e_\ell^T \otimes I_n)z$, we have that (z_ℓ, λ_0) is an approximate right eigenpair for $P(\lambda)$ and*

$$\frac{\eta_{ra}(z_\ell, \lambda_0, P)}{\eta_{ra}(z, \lambda_0, D_\ell)} \leq k^{3/2}\frac{\|z\|_2}{\|z_\ell\|_2}, \quad \frac{\eta_{rr}(z_\ell, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, D_\ell)} \leq k^{3/2}\frac{\|z\|_2}{\|z_\ell\|_2}\rho. \tag{6.6}$$

An analogous result holds for left eigenpairs (w^*, λ_0) of $D_\ell(\lambda, P)$ by simply replacing z by w and z_ℓ by $w_\ell := (e_\ell^T \otimes I_n)w$.

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 $\ell = 1$ and $|\lambda_0| \geq 1$, or if $\ell = k$ and $|\lambda_0| \leq 1$, the approximate eigenvector z_ℓ for $P(\lambda)$, recovered from the approximate eigenvector z of $D_\ell(\lambda, P)$ as in Theorem 6.2, makes the quotient $\frac{\|z\|_2}{\|z_\ell\|_2}$ in (6.6) not larger than $\sqrt{k}$.

Remark 6.2 Based on Theorems 5.2 and 6.2, we compare the backward errors of $\mathcal{T}_P(\lambda)$, $D_1(\lambda, P)$ and $D_k(\lambda, P)$. First note that, although Theorem 6.2 is valid for any value of λ_0, an argument similar to the one in Remark 5.2 shows that, for $D_1(\lambda, P)$, the quotient $\frac{\|z\|_2}{\|z_1\|_2}$ in (6.6) is expected to be not larger than $\sqrt{k}$ only if $|\lambda_0| \geq 1$, while for $D_k(\lambda, P)$, the quotient $\frac{\|z\|_2}{\|z_k\|_2}$ is expected to be not larger than $\sqrt{k}$ only if $|\lambda_0| \leq 1$. In contrast, according to Remark 5.2, the quotient $\frac{\|z\|_2}{\|x\|_2}$ appearing in Theorem 5.2 is expected to be bounded by $\frac{1}{\sqrt{2}}k^{3/2}$ for any value of λ_0. This implies that in the relative-absolute case, $\mathcal{T}_P(\lambda)$ presents optimal behavior in terms of backward error in contrast with $D_1(\lambda, P)$ and $D_k(\lambda, P)$. 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_1 and ρ that appear in the bounds of the quotients of relative-relative backward errors in Theorems 5.2 and 6.2, respectively, have the same value, which is approximately one for polynomials whose coefficients have similar norms. Finally, observe that the bounds in Theorem 5.2 and 6.2 have the same dependence on k, that is, $k^{3/2}$. 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 $|\lambda_0| \geq 1$ and $D_k(\lambda, P)$, for computing the eigenvalues with $|\lambda_0| \leq 1$.

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 $C_2(\lambda) = [C_1(\lambda)]^{\mathcal{B}}$, where $C_1(\lambda)^{\mathcal{B}}$ denotes the block-transpose of $C_1(\lambda)$ and $C_2(\lambda)$ denotes the second Frobenius companion form of $P(\lambda)$ [20], any result that we produce for $C_1(\lambda)$ has an immediate counterpart for $C_2(\lambda)$ (see [24, 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 [24] in the relative-relative case.

Theorem 6.3 *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$ and $\max_{i=0:k}\{\|A_i\|_2\} = 1$. Let λ_0 be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Let $C_1(\lambda)$ be the first Frobenius companion linearization of $P(\lambda)$. Let ρ_1 be as in (5.1).*

Then,

$$\frac{1}{\sqrt{k+1}} \leq \frac{\kappa_{ra}(\lambda_0, C_1)}{\kappa_{ra}(\lambda_0, P)} \leq 2k^3, \quad \frac{1}{\sqrt{k+1}} \leq \frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} \leq 2k^3 \rho_1.$$

Proof We prove the result for the relative-relative case. The relative-absolute case can be proven similarly. Let $C_1(\lambda) := \lambda X_1 + Y_1$. By Lemma 6.2, if y is a left-eigenvector of $P(\lambda)$ associated with λ_0, then

$$w = \begin{bmatrix} I_n \\ (P_1(\lambda_0))^* \\ \vdots \\ (P_{k-1}(\lambda_0))^* \end{bmatrix} y,$$

is a left eigenvector of $C_1(\lambda)$ associated with λ_0. Taking into account [24, Theorem 7.3], we obtain

$$\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} = \frac{\|w\|_2}{\|y\|_2} \frac{(|\lambda_0| \|X_1\|_2 + \|Y_1\|_2) \|\Lambda(\lambda_0)\|_2}{\sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2},$$

where $\Lambda(\lambda)$ is as in (6.1). By [24, Lemma 7.4], we have

$$\|X_1\|_2 = \max\{1, \|A_k\|_2\} = 1 \tag{6.7}$$

and

$$1 = \max_{i=0:k-1}\{1, \|A_i\|_2\} \leq \|Y_1\|_2 \leq k \max_{i=0:k-1}\{1, \|A_i\|_2\} = k. \tag{6.8}$$

Thus, using (5.18), we get

$$\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} \leq \frac{\|w\|_2}{\|y\|_2} \frac{|\lambda_0| + k}{\min\{\|A_k\|_2, \|A_0\|_2\}} \frac{\sqrt{\sum_{i=0}^{k-1} |\lambda_0|^{2i}}}{|\lambda_0|^k + 1}. \tag{6.9}$$

Assume that $|\lambda_0| \leq 1$. Then,

$$\begin{aligned}
\|w\|_2^2 = \|y\|_2^2 + \sum_{i=1}^{k-1} \|(P_i(\lambda_0))^* y\|_2^2 &\leq \left(1 + \sum_{i=1}^{k-1} \|(P_i(\lambda_0))^*\|_2^2\right) \|y\|_2^2 \\
&\leq \max_{i=0:k}\{1, \|A_i\|_2^2\}\|y\|_2^2 \left(1 + \sum_{i=1}^{k-1} \left(\sum_{j=0}^{i} |\lambda_0|^j\right)^2\right) \\
&\leq \max_{i=0:k}\{1, \|A_i\|_2^2\}\|y\|_2^2 \left(1 + (k-1)k \sum_{j=0}^{k-1} |\lambda_0|^{2j}\right),
\end{aligned} \tag{6.10}$$

where the second and third inequalities follow taking into account Lemma 3.3 and Lemma 3.1, respectively. Thus,

$$\|w\|_2^2 \leq \max_{i=0:k}\{1, \|A_i\|_2^2\}\|y\|_2^2 \left[1 + (k-1)k^2\right] \leq k^3 \|y\|_2^2. \tag{6.11}$$

From (6.9), we obtain $\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} \leq 2k^3 \rho_1$, when $|\lambda_0| \leq 1$.

If $|\lambda_0| > 1$, taking into account Lemmas 3.1, 3.2 and 3.3, we have

$$\begin{aligned}
\|w\|_2^2 = \|y\|_2^2 + \sum_{i=1}^{k-1} \|y^* P_i(\lambda_0)\|_2^2 &= \|y\|_2^2 + \sum_{i=1}^{k-1} \|\lambda_0^{i-k} y^* P^{k-i-1}(\lambda_0)\|_2^2 \\
&\leq \|y\|_2^2 + \sum_{i=1}^{k-1} |\lambda_0|^{2(i-k)} \|P^{k-i-1}(\lambda_0)\|_2^2 \|y\|_2^2 \\
&\leq \|y\|_2^2 \max_{i=0:k}\{1, \|A_i\|_2^2\} \left[1 + \sum_{i=1}^{k-1} |\lambda_0|^{2(i-k)} \left(\sum_{j=0}^{k-i-1} |\lambda_0|^j\right)^2\right] \\
&\leq \|y\|_2^2 \left[1 + \sum_{i=1}^{k-1} (k-i) \sum_{j=0}^{k-i-1} |\lambda_0|^{2i-2k+2j}\right] \\
&\leq \|y\|_2^2 \left[1 + (k-1)^2 \sum_{j=0}^{k-2} |\lambda_0|^{2j+2-2k}\right] \\
&\leq \|y\|_2^2 \left[1 + (k-1)^3 |\lambda_0|^{-2}\right].
\end{aligned} \tag{6.12}$$

Since $|\lambda_0|+k \leq (|\lambda_0|+1)k$, taking into account (6.9) and the upper bound for $\|w\|_2^2/\|y\|_2^2$ in (6.12), we have

$$\begin{aligned}\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} &\leq \frac{\|w\|_2}{\|y\|_2}\frac{1}{\min\{\|A_k\|_2, \|A_0\|_2\}}\frac{(1+|\lambda_0|)k\sqrt{k}|\lambda_0|^{k-1}}{1+|\lambda_0|^k} \\ &= \rho_1 k\sqrt{k}\sqrt{1+(k-1)^3|\lambda_0|^{-2}}\frac{(1+|\lambda_0|)}{|\lambda_0|} \leq 2k^3\rho_1.\end{aligned}$$

Now we find a lower bound for $\frac{\kappa_{rr}(\lambda_0,C_1)}{\kappa_{rr}(\lambda_0,P)}$. Notice that $\|w\|_2 \geq \|y\|_2$. Thus, from (6.7) and (6.8), and since $\max_{i=0:k}\{\|A_i\|_2\} = 1$, we have

$$\begin{aligned}&\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, P)} \geq \frac{(|\lambda_0|\|X_1\|_2 + \|Y_1\|_2)\|A(\lambda_0)\|_2}{\sum_{i=0}^{k}|\lambda_0|^i\|A_i\|_2} \geq \frac{(|\lambda_0|+1)\sqrt{\sum_{i=0}^{k-1}|\lambda_0|^{2i}}}{\sum_{i=0}^{k}|\lambda_0|^i} \\ &= \frac{\sqrt{(|\lambda_0|+1)^2\sum_{i=0}^{k-1}|\lambda_0|^{2i}}}{\sum_{i=0}^{k}|\lambda_0|^i} \geq \frac{\sqrt{(|\lambda_0|^2+1)\sum_{i=0}^{k-1}|\lambda_0|^{2i}}}{\sum_{i=0}^{k}|\lambda_0|^i} \\ &\geq \frac{\sqrt{\sum_{i=0}^{k}|\lambda_0|^{2i}}}{\sum_{i=0}^{k}|\lambda_0|^i} = \sqrt{\frac{\sum_{i=0}^{k}|\lambda_0|^{2i}}{(\sum_{i=0}^{k}|\lambda_0|^i)^2}} \geq \frac{1}{\sqrt{k+1}}.\end{aligned}$$

Notice that the last inequality holds by Lemma 3.1. □

We now recall a result obtained in [23] regarding the comparison of the relative-relative backward errors of approximate eigenpairs of $C_1(\lambda)$ and $P(\lambda)$. We also include the relative-absolute result, which can be obtained similarly.

Theorem 6.4 *[23, Theorems 3.6 and 3.8] Let $P(\lambda)$ be a matrix polynomial of degree k as in (1.1) with $A_0 \neq 0$ and $\max\limits_{i=0:k}\{\|A_i\|_2\} = 1$. Let (z, λ_0) be an approximate right eigenpair of $C_1(\lambda)$. Then, for $z_\ell = (e_\ell^T \otimes I_n)z$, $\ell = 1 : k$, we have that (z_ℓ, λ_0) is an approximate right eigenpair of $P(\lambda)$ and*

$$\frac{\eta_{ra}(z_\ell, \lambda_0, P)}{\eta_{ra}(z, \lambda_0, C_1)} \leq k^{5/2}\frac{\|z\|_2}{\|z_\ell\|_2}, \quad \frac{\eta_{rr}(z_\ell, \lambda_0, P)}{\eta_{rr}(z, \lambda_0, C_1)} \leq k^{5/2}\frac{\|z\|_2}{\|z_\ell\|_2}\rho_1, \tag{6.13}$$

where ρ_1 is as in (5.1). Let (w^*, λ_0) 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^*, λ_0) is an approximate left eigenpair of $P(\lambda)$ and

$$\frac{\eta_{ra}(w_1^*, \lambda_0, P)}{\eta_{ra}(w^*, \lambda_0, C_1)} \leq k^{3/2}\frac{\|w\|_2}{\|w_1\|_2}, \quad \frac{\eta_{rr}(w_1^*, \lambda_0, P)}{\eta_{rr}(w^*, \lambda_0, C_1)} \leq k^{3/2}\frac{\|w\|_2}{\|w_1\|_2}\rho_1. \tag{6.14}$$

Remark 6.3 Taking into account the form of the right eigenvectors of $C_1(\lambda)$ (see Lemma 6.2), and assuming that the approximate right eigenvector z has a similar block structure as the exact ones, it is expected that the quotient $\frac{\|z\|_2}{\|z_\ell\|_2}$ in (6.13) is not larger than $\sqrt{k}$ if $\ell = 1$ and $|\lambda_0| \geq 1$, or if $\ell = k$ and $|\lambda_0| \leq 1$. Thus, depending on the modulus of the approximate eigenvalue λ_0, 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 $|\lambda_0| \geq 1$ and $x = z_k$ if $|\lambda_0| \leq 1$. Regarding the left eigenvectors, assuming that the approximate left eigenvector ω has a block structure similar to the exact one, from (6.11) and (6.12) it is expected that the quotient $\frac{\|\omega\|_2}{\|\omega_1\|_2}$ in (6.14) is not larger than $k^{3/2}$. Finally, we note that, in the bounds for the quotients of relative-relative backward errors, for ρ_1 to be close to 1, it is enough that all matrix coefficients of $P(\lambda)$ have similar norms.

Remark 6.4 Based on Theorems 5.1, 5.2, 6.3 and 6.4, we conclude that $\mathcal{T}_p(\lambda)$ and $C_1(\lambda)$ have similar behavior in terms of conditioning and backward errors. 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 and save operations.

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. The goal is to show that indeed in practice $\mathcal{T}_P(\lambda)$ is a block symmetric linearization with much better numerical properties than the block symmetric linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ and with properties similar to those of the standard unstructured first Frobenius companion form $C_1(\lambda)$. In some examples, $P(\lambda)$ is a random matrix polynomial, sometimes prepared to illustrate situations in which the advantage of $\mathcal{T}_P(\lambda)$ over the other linearizations are really striking, while in another, $P(\lambda)$ is one of the matrix polynomials connected with an application discussed in [4]. But in all the examples that we have run, the numerical experiments show that the linearization $\mathcal{T}_P(\lambda)$ outperforms $D_1(\lambda, P)$ and $D_k(\lambda, P)$ numerically.

The experiments were run on MATLAB-R2016a, for which the unit roundoff is 2^{-53}. When calculating the condition numbers in Experiments 1 and 2, we computed the eigenvalues, eigenvectors, and the condition numbers themselves using variable precision arithmetic with 40 decimal digits of precision. This was not possible in Experiments 3 and 4 since either the matrix polynomial is very large (degree 3 and size 128×128) or we run a large number of experiments (we used 50 and 100 matrix polynomials, respectively). The standard double precision of MATLAB was used in this case. When computing the backward errors, we considered only right eigenpairs and the computations were done in the double precision floating point arithmetic of MATLAB in all the experiments. The function eig is used to compute approximate eigenvalues and eigenvectors of the linearizations. If z denotes a computed eigenvector of a linearization, associated with a computed eigenvalue λ_0, an eigenvector for $P(\lambda)$ was recovered from z as described in Remark 6.3 and Theorems 5.2 and 6.2, for $C_1(\lambda)$, $\mathcal{T}_P(\lambda)$, and $D_1(\lambda, P)$ and $D_k(\lambda, P)$, respectively.

According to the statements of the main results in Section 5 and 6, and for the reasons explained there, in all the numerical tests, we scale the considered matrix polynomial $P(\lambda)$ by dividing each of its coefficients by $\max_{i=0:k}\{\|A_i\|_2\}$. Recall

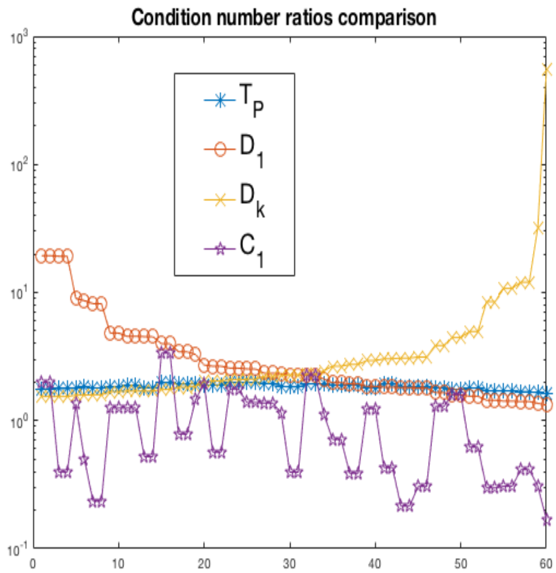

Fig. 7.1: Condition numbers ratio for random 20×20 matrix polynomial $P(\lambda)$ of degree 3.

that, with this scaling, the ratios $\kappa_{ra}(\lambda_0, \mathcal{T}_P)/\kappa_{ra}(\lambda_0, P)$ and $\eta_{ra}(x, \lambda_0, P)/\eta_{ra}(z, \lambda_0, \mathcal{T}_P)$ have been proven to be a moderate number of order 1 (for moderate k) and so $\mathcal{T}_P(\lambda)$ is optimal in the relative-absolute sense. Therefore, our experiments focus only on comparing the relative-relative condition numbers and backward errors. We note that the scaling mentioned above does not affect the conditioning and the backward errors of $D_1(\lambda, P)$ and $D_k(\lambda, P)$.

Experiment 1. 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 the MATLAB function `randn`. The matrix coefficients of $P(\lambda)$ had similar spectral norms. More precisely, $\|A_3\|_2 = 1$, $\|A_2\|_2 = 0.94$, $\|A_1\|_2 = 0.93$ and $\|A_0\|_2 = 0.96$. The smallest modulus of the eigenvalues of $P(\lambda)$ was 0.22 and the largest was 3.25, i.e., all the moduli of the eigenvalues of $P(\lambda)$ were close to 1. The main results for Experiment 1 are shown in Figures 7.1 and 7.2, where the x-axis has the indices $1, 2, ..., 60$ 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). The y-axis in the graph in Figure 7.1 corresponds to the ratio of relative-relative condition numbers $\frac{\kappa_{rr}(\lambda_0, L)}{\kappa_{rr}(\lambda_0, P)}$, where L is any of the linearizations, $D_1(\lambda, P)$, $D_k(\lambda, P)$, $C_1(\lambda)$ or $\mathcal{T}_P(\lambda)$. The y-axis in the graph in Figure 7.2 corresponds to the ratio of relative-relative backward errors $\eta_{rr}(x, \lambda_0, P)/\eta_{rr}(z, \lambda_0, L)$.

Figures 7.1 and 7.2 show that $\mathcal{T}_P(\lambda)$ has, in terms of conditioning and backward errors, better behavior than each of the block symmetric linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$ and similar behavior to the combined used of $D_1(\lambda, P)$ and $D_k(\lambda, P)$, when $D_1(\lambda, P)$ is used to compute the eigenvalues with moduli larger than 1 and

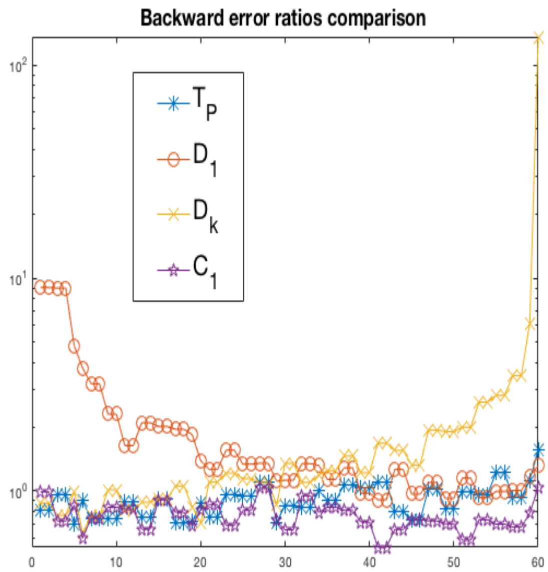

Fig. 7.2: Backward errors ratio for random 20×20 matrix polynomial $P(\lambda)$ of degree 3.

$D_k(\lambda, P)$ is used to compute the eigenvalues with moduli smaller than 1. 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. In addition, the behavior of $\mathcal{T}_P(\lambda)$ is similar to the behavior of the unstructured Frobenius companion form $C_1(\lambda)$.

We emphasize that we have repeated this random experiment for several sizes and odd degrees (up to 21) and we have always obtained similar results. Moreover, since the matrix coefficients of $P(\lambda)$ have very similar norms, the results in these examples remain essentially unchanged for the relative-absolute ratios $\frac{\kappa_{ra}(\lambda_0,L)}{\kappa_{ra}(\lambda_0,P)}$ and $\eta_{ra}(x, \lambda_0, P)/\eta_{ra}(z, \lambda_0, L)$, illustrating the optimality of $\mathcal{T}_P(\lambda)$, which behaves essentially as the polynomial $P(\lambda)$, i.e., producing ratios essentially equal to 1. Finally, we remark that the improvements of $\mathcal{T}_P(\lambda)$ with respect to $D_1(\lambda, P)$ and $D_k(\lambda, P)$ shown in Figures 7.1 and 7.2 are very moderate, as a consequence of the fact that all the eigenvalues of $P(\lambda)$ have moduli close to 1 (recall Theorem 6.1 in this respect). This led us to devise the following experiment.

Experiment 2. In this experiment, we examine another random matrix polynomial of size 20×20 and degree 3. The coefficients of $P(\lambda)$ were again generated with the command `randn` of MATLAB, with the exception of A_0 and A_3 that were both constructed in such a way that each of them has six very small singular values equal to $10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}$. The matrix coefficients of $P(\lambda)$ had similar norms. After scaling $P(\lambda)$ by dividing all matrix coefficients by $\max\{\|A_i\|_2, i = 0 : 3\}$, the scaled matrix coefficients $\tilde{A}_i$ of $P(\lambda)$ satisfy $\min_{i=0:3}\{\|\tilde{A}_i\|_2\} = 0.43$ and $\max_{i=0:3}\{\|\tilde{A}_i\|_2\} = 1$. The smallest modulus of the eigenvalues of $P(\lambda)$ was $3.6 \cdot 10^{-8}$ and the largest $6.9 \cdot 10^7$. Thus, in this case,

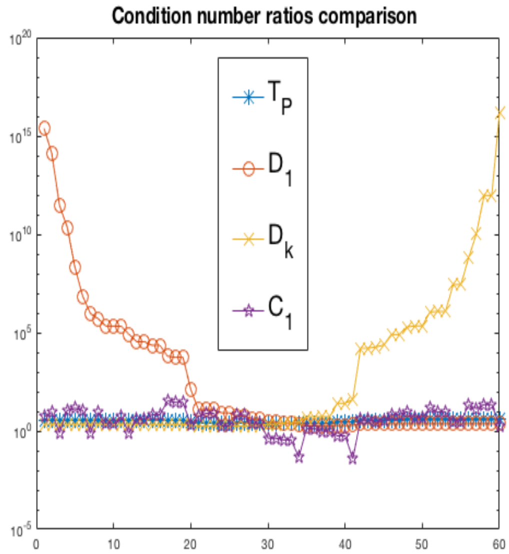

Fig. 7.3: Condition numbers ratio for random 20×20 matrix polynomial $P(\lambda)$ of degree 3 with very small and very large eigenvalues.

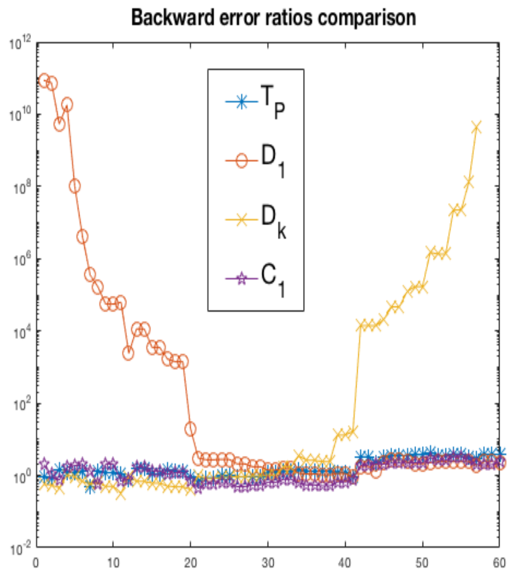

Fig. 7.4: Backward errors ratio for random 20×20 matrix polynomial $P(\lambda)$ of degree 3 with very small and very large eigenvalues.

there are eigenvalues with moduli very different from 1. The results are shown in Figures 7.3 and 7.4 and similar comments to those for Figures 7.1 and 7.2 hold with respect to the comparison of the behavior of the different linearizations, although, in this case, $\mathcal{T}_P(\lambda)$ outperforms the individual behavior of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ by several orders of magnitude in the smallest and the largest eigenvalues, respectively.

We would like to point out that, in the graph for the backward errors, the ratio corresponding to the largest eigenvalue of $D_k(\lambda, P)$ does not appear. This phenomenon, that we have observed in numerous experiments, is related with the fact that the leading matrix coefficient of $D_k(\lambda, P)$ is very ill-conditioned and its condition number is much larger than that of the leading matrix coefficients of the other linearizations and $P(\lambda)$. When computing this eigenvalue of $D_k(\lambda, P)$, MATLAB outputs an infinite eigenvalue. We would like to highlight that this behavior only affects $D_k(\lambda, P)$ and that this problem does not occur with the ratio of condition numbers since these have been computed with extended accuracy.

Experiment 3. The first part of this numerical experiment comes from an applied problem discussed in [4]: the "plasma-drift problem" (modeling of drift instabilities in the plasma edge inside a Tokamak reactor). This problem corresponds to a 128×128 matrix polynomial $P(\lambda)$ of degree 3 that we have scaled in such a way that $\max_{i=0:k}\{\|A_i\|_2\} = 1$. In contrast with previous experiments, in this example the matrix coefficients of $P(\lambda)$ have norms of different magnitudes. More specifically, $\|A_3\|_2 = 0.0103$, $\|A_2\|_2 = 0.0043$, $\|A_1\|_2 = 1$ and $\|A_0\|_2 = 0.0999$. The smallest modulus of the eigenvalues of $P(\lambda)$ is 0.028 and the largest is 11.745. The results for the ratios of relative-relative condition numbers and backward errors are shown in Figures 7.5 and 7.6. We observe again that $\mathcal{T}_P(\lambda)$ has a better behavior than each of the block symmetric linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$, in this case by several orders of magnitude, a similar behavior to the combined used of $D_1(\lambda, P)$ and $D_k(\lambda, P)$ (never differing by more than a factor 10), and also similar to the behavior of the first Frobenius form $C_1(\lambda)$. In this test, the ratios for $\mathcal{T}_P(\lambda)$ and the other linearizations increase with the moduli of the eigenvalues, which did not happen in Experiments 1 and 2. This is related to the differences in norms of the matrix coefficients of $P(\lambda)$ and the fact that we are considering relative-relative condition numbers and backward errors. For their relative-absolute counterparts, the ratios for $\mathcal{T}_P(\lambda)$ are always essentially equal to 1, i.e., optimal, as we proved in Theorems 5.1 and 5.2.

Although in the plasma-drift problem, where the matrix coefficients of the matrix polynomial vary widely in norm, $C_1(\lambda)$ and $\mathcal{T}_P(\lambda)$ behave very similarly, this is not always the case. To show this, we constructed 50 random matrix polynomials of degree 3 and size 40 and multiplied the matrix coefficients of each polynomial by constants so that the norm of the leading coefficient was approximately 10^6 times the norm of the matrix coefficient of the term of degree 0. For each polynomial, we computed the maximum of the ratios $\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}$ and the maximum of the ratios $\frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, C_1)}$, among all the eigenvalues of the matrix polynomial. Then, we constructed two vectors, u and v, whose ith entry is, respectively, the maximum ratio $\frac{\kappa_{rr}(\lambda_0, C_1)}{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}$ and $\frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, C_1)}$ associated with the ith matrix polynomial. Finally, we computed the maximum, mode, average, and median of the entries of those two vectors and got the following results: $\max(u) = 1.486 \cdot 10^4$, $\text{mode}(u) = 12.52$, $\text{average}(u) = 722.84$, $\text{median}(u) = 113.78$, $\max(v) = 3.185 \cdot 10^3$, $\text{mode}(v) = 5.19$, average$(v) = 109.18$, and median$(v) = 20.46$. Thus, we conclude that, if the matrix coefficients of a matrix polynomial vary widely in norm, then $\kappa_{rr}(\lambda_0, C_1)$ can be much larger than $\kappa_{rr}(\lambda_0, \mathcal{T}_P)$ for some eigenvalues and polynomials, but also $\kappa_{rr}(\lambda_0, \mathcal{T}_P)$ can be much larger than $\kappa_{rr}(\lambda_0, C_1)$ for other eigenvalues and polynomials. We do not have yet any explanation for these rather different behaviors.

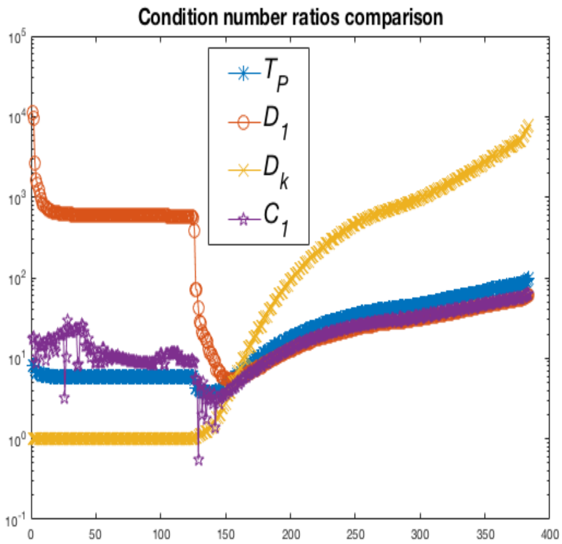

Fig. 7.5: Condition number ratios for Plasma-drift problem.

Experiment 4. Theorem 5.1 provides upper bounds on the ratio of condition numbers associated with $\mathcal{T}_P(\lambda)$ that depend on k^3, unless the modulus of the eigenvalue is smaller than $\frac{1}{k-1}$ or is larger than $k-1$. If this bound was sharp, for k moderate to large, say 10^2, the bound would be very large (a multiple of 10^6 in this example). In this experiment we study if the ways those bounds depend on k are pessimistic or if it can be expected that those bounds are attained for eigenvalues with modulus close to 1. In the experiment, we first generated 100 random matrix polynomials of degree 101 and size 2 and, for each of the polynomials, we computed the maximum ratio $\frac{\kappa_{rr}(\lambda_0, \mathcal{T}_P)}{\kappa_{rr}(\lambda_0, P)}$ among the eigenvalues λ_0 of the polynomial. Then, we constructed a vector u whose ith entry is that maximum ratio for the ith random matrix polynomial and computed the maximum, mode, average and median of the entries of u. We got that the maximum entry in u was $2.012 \cdot 10^3$, the mode was 22.108, the average was 100.13 and the median, 29.2811. Additionally, we split the range of values of u into 11 intervals of equal length 200 and considered the frequencies of the entries of u in those intervals. We got the vector of frequencies $N = [92, 1, 4, 0, 1, 0, 0, 0, 0, 1, 1]$. We conclude that, in general, the bounds in Theorem 5.1 are pessimistic and bounds depending on $k^{3/2}$ more than on k^3 can be expected.

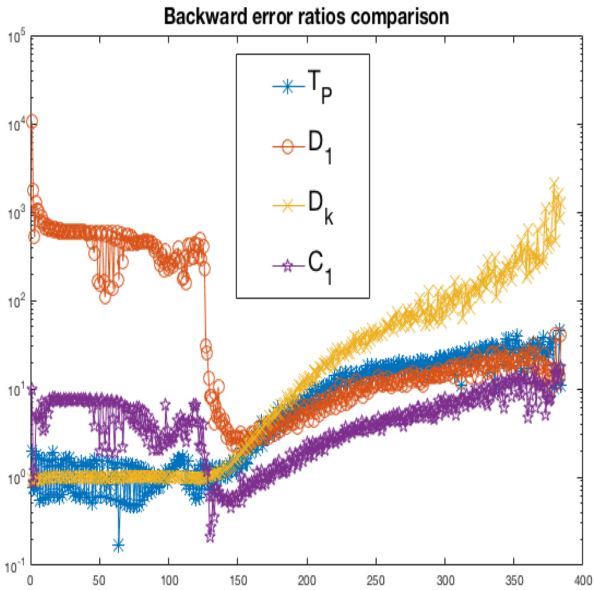

Fig. 7.6: Backward error ratios for Plasma-drift problem.

As a conclusion, except for a few matrix polynomials of degree 101 and for a very few matrix polynomials whose coefficients vary widely in norm, in all the numerical experiments that we have run, after scaling the matrix polynomial $P(\lambda)$, the linerization $\mathcal{T}_P(\lambda)$ produced 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. Moreover, the numerical behavior of $\mathcal{T}_P(\lambda)$ was comparable with the combined used of the block symmetric linearizations $D_1(\lambda, P)$ and $D_k(\lambda, P)$. Also, $\mathcal{T}_P(\lambda)$ presented a similar numerical behavior to that of the unstructured linearization $C_1(\lambda)$.

8 Conclusions

In this paper, we have studied for the first time eigenvalue condition numbers and backward errors of approximated eigenpairs of the block symmetric linearization $\mathcal{T}_P(\lambda)$ introduced by Fiedler in [18] for scalar polynomials with odd degree and extended by Antoniou and Vologiannidis to regular matrix polynomials with odd degree in [3]. The theoretical analysis and the numerical tests in this paper show that $\mathcal{T}_P(\lambda)$ has much better properties with respect to condition numbers and backward errors than any other block symmetric linearization analyzed so far in the literature, and it has similar properties to the *unstructured* Frobenius companion forms, which are the linearizations used more often for solving polynomial eigenvalue problems. Moreover, we have seen that, when the perturbations are measured in an absolute sense, the condition numbers and backward errors of $\mathcal{T}_P(\lambda)$ are always essentially equal to those of the polynomial $P(\lambda)$ and, so, $\mathcal{T}_P(\lambda)$ can be considered optimal in this respect. Future work in this area includes the development of strategies for extending the excellent properties of $\mathcal{T}_P(\lambda)$ to matrix polynomials with even degree and to design and analyze other block symmetric linearizations of even degree polynomials with better numerical properties than those of the linearizations analyzed so far in the literature, in particular, better than those of the linearizations in $\mathbb{DL}(P)$ [25,31]. We are currently working on these ideas. Another line of potential future research is to incorporate different structures, as symmetric, Hermitian, palindromic, alternating, etc, in the analysis performed in this paper.

Acknowledgements. We would like to thank the two anonymous referees for reading our manuscript so thoroughly and providing such constructive feedback. They asked very interesting questions that helped us improve our paper significantly.

References

1. B. Adhikari, R. Alam, and D. Kressner, *Structured eigenvalue condition numbers and linearizations for matrix polynomials*, Linear Algebra Appl., 435 (2011), 2193–2221.
2. 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.
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. 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.
5. D. A. Bini, L. Gemignani, and F. Tisseur, *The Ehrlich-Aberth method for the nonsymmetric tridiagonal eigenvalue problem*, SIAM J. Matrix Anal. Appl., 27 (2005), 153–175.
6. D. A. Bini and V. Noferini, *Solving polynomial eigenvalue problems by means of the Ehrlich–Aberth method*, Linear Algebra Appl., 439 (2013), 1130–1149.
7. S. Bora, *Structured eigenvalue condition number and backward error of a class of polynomial eigenvalue problems*, SIAM J. Matrix Anal. Appl., 31(3) (2009), 900–917.
8. M. I. Bueno, M. Martin, J. Pérez, A. Song, and I. Viviano, *Explicit block-structures for block-symmetric Fiedler-like pencils.* To appear in ELA.
9. 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.
10. 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.
11. 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.
12. C. Campos and J. Román, *Parallel Krylov solvers for the polynomial eigenvalue problem in SLEPc*, SIAM J. Sci. Comput., 38 (2016), S385–S411.
13. F. De Terán, F. M. Dopico, and D. S. Mackey, *Spectral equivalence of matrix polynomials and the Index Sum Theorem*, Linear Algebra Appl., 459 (2014), 264–333.
14. F. M. Dopico, P. Lawrence, J. Pérez, and P. Van Dooren, *Block Kronecker linearizations of matrix polynomials and their backward errors.* To appear in Numerische Mathematik. Available as arXiv:1707.04843v1.

15. F. M. Dopico, J. Pérez, and P. Van Dooren, *Structured backward error analysis of linearized structured polynomial eigenvalue problems.* To appear in Mathematics of Computation. Available as arXiv:1612.07011v1.
16. H. Fassbender and P. Saltenberger, *On vector spaces of linearizations for matrix polynomials in orthogonal bases*, Linear Algebra Appl., 525 (2017), 59–83.
17. H. Fassbender and P. Saltenberger, *Block Kronecker ansatz spaces for matrix polynomials*, Linear Algebra Appl. (2017), `http://dx.doi.org/10.1016/j.laa.2017.03.019`
18. M. Fiedler, *A note on companion matrices*, Linear Algebra Appl., 372 (2003), 325–331.
19. I. Gohberg, P. Lancaster, and L. Rodman, *Indefinite Linear Algebra and Applications*, Springer Verlag, Basel, Switzerland, 2005.
20. I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, 2009.
21. S. Güttel and F. Tisseur, *The nonlinear eigenvalue problem*, Acta Numerica, (2017), 1–94.
22. S. Hammarling, C. J. Munro, and F. Tisseur, *An algorithm for the complete solution of quadratic eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), Art. 18, 19.
23. 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.
24. 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.
25. 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.
26. T. Kailath, *Linear Systems*, Prentice Hall, Inc., Englewood Cliffs, N. J., 1980.
27. D. Kressner and J. Roman, *Memory-efficient Arnoldi algorithms for linearizations of matrix polynomials in Chebyshev basis*, Numer. Linear Algebra Appl., 21 (2014), 569–588.
28. P. Lancaster, *Symmetric transformations of the companion matrix*, NABLA: Bulletin of the Malayan Math. Soc., 8 (1961), 146–148.
29. P. Lawrence, M. Van Barel, and P. Van Dooren, *Backward error analysis of polynomial eigenvalue problems solved by linearization*, SIAM J. Matrix Anal. Appl., 37 (2016), 123–144.
30. D. Lu, Y. Su, and Z. Bai, *Stability analysis of the two-level orthogonal Arnoldi procedure*, SIAM J. Matrix Anal. Appl., 37 (2016), 192–214.
31. 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.
32. 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.
33. D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Jordan structures of alternating matrix polynomials*, Linear Algebra Appl., 432 (2010), 971–1004.
34. D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Smith forms of palindromic matrix polynomials*, Electron. J. Linear Algebra, 22 (2011), 53–91.
35. D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Skew-symmetric matrix polynomials and their Smith forms*, Linear Algebra Appl., 438 (2013), 4625–4653.
36. V. Mehrmann and H. Voss, *Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods*, GAMM Mitt. Ges. Anqew. Math. Mech., 27 (2004) 121–152.
37. C. B. Moler and G. W. Stewart, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), 241–256.
38. Y. Nakatsukasa, V. Noferini, and A. Townsend, *Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach*, SIAM J. Matrix Anal. Appl., 38 (2017), 1–29.
39. F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), 339–361.
40. F. Tisseur, *Tridiagonal-diagonal reduction of symmetric indefinite pairs*, SIAM J. Matrix Anal. Appl., 26 (2004), 215–232.
41. F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Review, 43 (2001), 235–286.
42. M. Van Barel and F. Tisseur, *Polynomial eigenvalue solver based on tropically scaled Lagrange linearization*, Linear Algebra Appl., (2017), `http://dx.doi.org/10.1016/j.laa.2017.04.025`
43. R. Van Beeumen, K. Meerbergen, and W. Michiels, *Compact rational Krylov methods for nonlinear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 36 (2015), 820–838.

44. P. Van Dooren and P. Dewilde, *The eigenstructure of an arbitrary polynomial matrix: computational aspects*, Linear Algebra Appl., 50 (1983), 545–579.
45. L. Zeng and Y. Su, *A backward stable algorithm for quadratic eigenvalue problems*, SIAM J. Matrix. Anal. Appl., 35 (2014), 499–516.