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

On why using $\mathbb{DL}(P)$ for the symmetric polynomial eigenvalue problem might need to be reconsidered.

M. I. Bueno · J. Pérez

Received: date / Accepted: date

Abstract In the literature it is common to use the first and last pencils $D_1(\lambda, P)$ and $D_k(\lambda, P)$ in the "standard basis" for the vector space $\mathbb{DL}(P)$ of block-symmetric pencils to solve the symmetric/Hermitian polynomial eigenvalue problem $P(\lambda)x = 0$. When the polynomial $P(\lambda)$ has odd degree, it was proven in recent years that the use of an alternative linearization $\mathcal{T}_P$ is more convenient because it has better numerical properties and its use is more universal since $\mathcal{T}_P$ is a strong linearization of any matrix polynomial $P(\lambda)$, while $D_1(\lambda; P)$ and $D_k(\lambda; P)$ are not. However, $\mathcal{T}_P$ is not defined for even degree matrix polynomials. In this paper we consider the case when $P(\lambda)$ has even degree. It is believed that the backward errors of eigenpairs computed with the use of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ are "small" based on the computed theoretical bounds for the backward errors of eigenpairs of $P(\lambda)$ computed from eigenpairs of these linearizations. We show that this is not the case, even when the polynomial $P(\lambda)$ is well-scaled because of the ill-conditioning of the eigenvectors of $D_1(\lambda; P)$ and $D_k(\lambda; P)$. We introduce two block-symmetric linearizations for even degree matrix polynomials that overcome this problem and become an appropriate alternative to the traditional use of $D_1(\lambda; P)$ and $D_k(\lambda; P)$.

Keywords symmetric/Hermitian matrix polynomial · eigenvalue · backward errors · eigenvector · polynomial eigenvalue problem

Mathematics Subject Classification (2000) 65F05 · 65F15 · 65F35 · 65H04 · 65H17 · 15A12 · 15A21

The research of M. I. Bueno was also partially supported by "Ministerio de Economía, Industria y Competitividad (MINECO)" of Spain and "Fondo Europeo de Desarrollo Regional (FEDER)" of EU through grant MTM2015-65798-P and by the "Proyecto financiado por la Agencia Estatal de Investigación de España" (PID2019-106362GB-I00 / AEI / 10.13039/501100011033)." .

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.

J. Pérez
Department of Mathematical Sciences, University of Montana, Missoula, MT 59812, USA E-mail: javier.perez-alvaro@mso.umt.edu.

1 Introduction

A square *matrix polynomial* takes the form

$$P(\lambda) = \lambda^k A_k + \cdots + \lambda A_1 + A_0, \qquad A_0, \ldots, A_k \in \mathbb{F}^{n \times n}, \tag{1.1}$$

where $\mathbb{F}$ denotes any field. In this paper, we consider the field of real or complex numbers. We say that $P(\lambda)$ has degree k if $A_k \neq 0$ and we say that $P(\lambda)$ has grade k, otherwise. In this work, we are interested in symmetric and Hermitian matrix polynomials. We say that $P(\lambda)$ is *symmetric* if $A_i^T = A_i$, for $i = 0, 1, \ldots, k$, and we say that $P(\lambda)$ is *Hermitian* if $\mathbb{F} = \mathbb{C}$ and $A_i^* = A_i$, for $i = 0, 1, \ldots, k$, where $(\cdot)^*$ denotes the complex conjugate transpose operation.

Throughout this paper, we assume that the matrix polynomial $P(\lambda)$ in (1.1) is regular, this is, the scalar polynomial $\det P(\lambda)$ is not the zero polynomial. We also assume $A_k \neq 0$ and $A_0 \neq 0$ in order to avoid some trivialities. The *polynomial eigenvalue problem (PEP)* associated with a regular matrix polynomial $P(\lambda)$ consists in finding scalars λ_0 for which the equations

$$P(\lambda_0)x = 0 \quad \text{and} \quad y^* P(\lambda_0) = 0 \tag{1.2}$$

have nontrivial solutions $x, y \in \mathbb{F}^n$. The scalar λ_0 is called an *eigenvalue* of $P(\lambda)$, and the vectors x and y are associated *right and left eigenvectors*. The set of all eigenvalues of the matrix polynomial $P(\lambda)$ is called the *spectrum* of $P(\lambda)$. The eigenvalue/eigenvector pair (λ_0, x) (resp. (y, λ_0)) is called a right (resp. left) *eigenpair* of $P(\lambda)$. When the matrix polynomial $P(\lambda)$ is symmetric (resp. Hermitian), we refer to (1.2) as the *symmetric (resp. Hermitian) polynomial eigenvalue problem*. When $P(\lambda)$ is symmetric or Hermitian, the sets of left and right eigenvectors coincide.

Structured PEPs, that is, PEP in which the matrix coefficients of the matrix polynomial present some type of structure, arise from many applications. For instance, symmetric and Hermitian PEPs arise in the classical problem of vibration analysis [9, 15, 23]. When solving numerically a structured PEP it is well-recognized the importance of using structure preserving eigenvalue algorithms [13]. For example, symmetric or Hermitian matrix polynomials have a spectrum that is symmetric with respect to the real axis. In a finite precision environment, an algorithm that ignores the structure of the polynomial may lose this symmetry [16]. For this reason, one of the most common approaches for numerically solving structured PEPs is to use structure-preserving linearizations (see Section 2.1 for the definition of linearization). This process replaces the original structured PEP with a generalized eigenvalue problem with the same structure. Standard methods for structured generalized eigenvalue problems can then be applied; see, e.g., [11] and the references therein.

The landmark paper [16] introduced a family of candidate linearizations for matrix polynomials as in (1.1), the so-called $\mathbb{DL}(P)$ vector space. It was proved in [16] that almost all matrix pencils in $\mathbb{DL}(P)$ are linearizations of the matrix polynomial $P(\lambda)$, and that $\mathbb{DL}(P)$ is a rich source of structure-preserving linearizations for structured matrix polynomials. Moreover, among all the linearizations in $\mathbb{DL}(P)$, the pencils $D_1(\lambda; P)$ and $D_k(\lambda; P)$ (see (4.1) and (4.2)) were identified in [**?**,10] as those with almost optimal numerical properties (in terms of eigenvalue conditioning and backward errors). These optimality results have led several authors to propose the use of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ (or small variations of $D_1(\lambda; P)$ and $D_k(\lambda; P)$) in the task of solving numerically structured PEPs from applications. These structure-preserving linearizations have been used, for example, to solve palindromic and even PEPs [14], Hamiltonian (alternating) PEPs [18], to solve complex-symmetric PEPs [8], to solve symmetric or Hermitian rational eigenvalue problems [24], as the starting point to build trimmed linearizations for structured matrix polynomials [6], to develop a backward stable algorithm for symmetric or Hermitian quadratic eigenvalue problems [25], to estimate the distance to uncontrollability of higher order dynamical systems [20], to compute the H_∞ norm [3], and to solve nonlinear eigenvalue problems by using the infinity Lanczos method [19], to name some recent works.

Although the numerical properties of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ are good enough for certain applications, one of the key findings of this work is the extreme sensitivity of the eigenvectors of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ to small perturbations. Hence, the computation of accurate eigenvalues and eigenvectors of structured matrix polynomials requires to find structure-preserving linearizations with better numerical properties. Steps in this direction can be found in [4], where the authors compare the numerical properties of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ with the block-tridiagonal linearization introduced in [2], in the case when the matrix polynomial has *odd degree*. Their analysis reveals that the block-symmetric linearization from [2] has much better numerical properties than the linearizations in $\mathbb{DL}(P)$. In this work, we address the case when the matrix polynomial has *even degree*. This case is different from the odd degree case because there are symmetric (resp. Hermitian) matrix polynomials of even degree that do not have symmetric (resp. Hermitian) linearizations while they always exist for odd degree polynomials. To guarantee the existence of structure-preserving linearizations for even degree matrix polynomials, one has to impose some conditions on the matrix polynomial coefficients. For example, symmetric and Hermitian matrix polynomials with nonsingular leading and/or trailing matrix coefficients always present structure-preserving linearizations. These conditions make the numerical analysis more challenging.

In this paper we analyze different strategies for solving PEPs associated with even-degree structured matrix polynomials and propose the combined use of two linearizations $\mathcal{H}_P^{A_k}$ and $\mathcal{G}_P^{A_k}$ introduced in (5.6) and (5.9) (using $S = A_k$) as an alternative to the use of the linearizations $D_1(\lambda; P)$ and $D_k(\lambda; P)$ because it is numerically superior and avoids the problem with the sensitivity of the eigenvectors.

The structure of the paper is as follows: In Section 2 we introduce the mathematical background necessary for the rest of the paper. In Section 3 we recall the definition of (normwise) eigenvalue condition number and backward error of an eigenpair of a matrix polynomial as well as convenient formulas to compute these quantities. In Section 4, we recall the definition and properties of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ and provide theoretical and numerical evidence of the sensitivity of the eigenvectors of these pencils to small perturbations in the coefficients of the polynomial $P(\lambda)$. In Section 5, we show how to use the pencil $\mathcal{T}_P(\lambda)$ to construct a family of pencils $\mathcal{H}_P^S$ (resp. $\mathcal{G}_P^S$) that are strong linearizations of an even degree matrix polynomial $P(\lambda)$ as in (1.1) with nonsingular A_k (resp. A_0). In Sections 6 and 7, we provide a numerical analysis of the eigenvalue condition number and backward errors of the pencils in the families $\mathcal{H}_P^S$ and $\mathcal{G}_P^S$, and show that optimal behavior is attained when $S = A_k$. Finally, in Section 8, we present the proofs of the main results in Sections 6 and 7.

2 Definitions and technical results

We review in this section the notions of linearization and strong linearization of a matrix polynomial. For a more detailed introduction on these concepts, we refer the reader to the classical book [9] and to the more recent reference [7].

Additionally, we present some technical results that will be used in the proofs of the main theorems of this manuscript.

2.1 Linearizations of matrix polynomials

A matrix polynomial $U(\lambda)$ is said to be *unimodular* if $\det U(\lambda)$ is a nonzero constant (i.e., independent of λ). A grade-1 matrix polynomial $L(\lambda) = \lambda B + A$ is called a *matrix pencil*, or *pencil* for short. A matrix pencil $L(\lambda) = \lambda B + A$ is called a *linearization* of a matrix polynomial $P(\lambda)$ if there exist unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ such that

$$L(\lambda) = U(\lambda) \begin{bmatrix} I_s & 0 \\ 0 & P(\lambda) \end{bmatrix} V(\lambda),$$

for some s, where I_s denotes de $s \times s$ identity matrix. Linearizations preserve the finite eigenvalues of the polynomial $P(\lambda)$ and their multiplicities.

Given a matrix polynomial $P(\lambda)$ as in (1.1), its *reversal matrix polynomial* is defined by

$$\mathrm{rev}P(\lambda) = \lambda^k P(\lambda^{-1}) = \lambda^k A_0 + \cdots + \lambda A_{k-1} + A_k.$$

We say that $P(\lambda)$ has an eigenvalue at infinity if 0 is an eigenvalue of $\mathrm{rev}P(\lambda)$. A linearization $L(\lambda)$ of $P(\lambda)$ is said to be *strong* if $\mathrm{rev}(L)$ is a linearization of $\mathrm{rev}(P)$. Strong linearizations preserve both the finite and infinite eigenvalues of $P(\lambda)$ and their multiplicities.

2.2 Some auxiliary results

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

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

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

Lemma 2.1 *Let m be a positive integer and let a be a positive real number. Then,*

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

Next we provide an upper and lower bound on the norm of a block-matrix in terms of the norms of its blocks.

Proposition 2.1 [10, Lemma 3.5] *For any complex $\ell \times m$ block-matrix $B = (B_{ij})$ we have*

$$\max_{i,j} \|B_{ij}\|_2 \leq \|B\|_2 \leq \sqrt{\ell m} \max_{i,j} \|B_{ij}\|_2. \tag{2.1}$$

Some of our main results require the systematic use of the Horner shifts of a matrix polynomial $P(\lambda)$.

Definition 2.1 (Horner shifts) Given a matrix polynomial $P(\lambda)$ of degree k as in (1.1), the ith Horner shift of $P(\lambda)$, for $i = 0 : k$, is given by

$$P_i(\lambda) := \lambda^i A_k + \lambda^{i-1} A_{k-1} + \cdots + \lambda A_{k-i+1} + A_{k-i}. \tag{2.2}$$

Notice that $P_0(\lambda) = A_k$ and $P_k(\lambda) = P(\lambda)$. Moreover, Horner shifts satisfy the recurrence relation

$$P_{i+1}(\lambda) - A_{k-i-1} = \lambda P_i(\lambda), \quad \text{for} \quad i = 0 : k-1. \tag{2.3}$$

We also denote

$$P^i(\lambda) := \lambda^i A_i + \cdots + \lambda A_1 + A_0, \quad \text{for} \quad i = 0 : k. \tag{2.4}$$

Notice that $P^0(\lambda) = A_0$ and $P^k(\lambda) = P(\lambda)$. Furthermore, the two families of polynomials (2.2) and (2.4) are related as follows

$$P(\lambda) = \lambda^{k-i} P_i(\lambda) + P^{k-i-1}(\lambda), \quad i = 0 : k-1. \tag{2.5}$$

Lemma 2.2 provides another relation between the two families of Horner shifts.

Lemma 2.2 [4, 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* (2.2) *and* (2.4)*, respectively. Let λ_0 be a nonzero and 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 $i = 0 : k-1$,*

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

The proof of Lemma 2.3 can be easily verified.

Lemma 2.3 *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* (2.2) *and* (2.4)*, respectively. Then, for any $n \times n$ matrix M and for $i = 0 : k$, we have*

$$\begin{aligned}
&\|MP_i(\lambda_0)\|_2 \leq \max_{j=0:k}\{\|MA_j\|_2\} \sum_{j=0}^{i} |\lambda_0|^j, \\
&\|MP^i(\lambda_0)\|_2 \leq \max_{j=0:k}\{\|MA_j\|_2\} \sum_{j=0}^{i} |\lambda_0|^j, \quad \textit{and} \\
&\|P_i(\lambda_0)\|_2 \geq \max_{j=0:i}\{|\lambda_0|^j \|A_j\|_2\}.
\end{aligned}$$

3 Eigenvalue condition numbers and backward errors of approximate eigenpairs

In this section, we review the notions of relative eigenvalue condition number and backward error of approximate eigenpairs of a matrix polynomial, and state some of their basic properties.

Definition 3.1 (Eigenvalue condition number) [22] Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1). If λ_0 is a *simple, finite, nonzero* eigenvalue of $P(\lambda)$ with corresponding right eigenvector x, then the *relative condition number of* λ_0 is defined by

$$\kappa_{\mathrm{r}}(\lambda_0; P) := \lim_{\epsilon\to 0} \sup \left\{ \frac{|\Delta\lambda_0|}{\epsilon|\lambda_0|} \,:\, \left(P(\lambda_0+\Delta\lambda_0) + \Delta P(\lambda_0+\Delta\lambda_0)\right)(x+\Delta x) = 0, \right.$$
$$\left. \text{with } \|\Delta A_i\|_2 \le \epsilon\, \omega_i, \text{ for } i = 0:k \right\},$$

where ω_i are some previously selected nonnegative weights.

Definition 3.2 (Backward error of an approximate eigenpair) [22] Let $P(\lambda)$ be a regular matrix polynomial of degree k as in (1.1). For a given approximate right eigenpair $(\widetilde{\lambda}_0, \widetilde{x})$ of $P(\lambda)$, the *backward error of* $(\widetilde{\lambda}_0, \widetilde{x})$ is

$$\eta(\widetilde{\lambda}_0, \widetilde{x}; P) := \min \left\{ \epsilon \,:\, (P(\widetilde{\lambda}_0) + \Delta P(\widetilde{\lambda}_0))\widetilde{x} = 0, \text{ with } \|\Delta A_i\|_2 \le \epsilon\, \omega_i, \right.$$
$$\left. \text{for } i = 0:k, \right\},$$

where ω_i are some previously selected nonnegative weights.

Explicit formulas for the condition number $\kappa_{\mathrm{r}}(\lambda_0; P)$ and the backward error $\eta(\widetilde{\lambda}_0, \widetilde{x}; P)$ were obtained in [22].

Theorem 3.1 [22, Theorem 5] *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1). *If λ_0 is a simple, finite, nonzero eigenvalue of $P(\lambda)$ with corresponding right and left eigenvectors x and y, then*

$$\kappa_{\mathrm{r}}(\lambda_0; P) = \frac{\left(\sum_{i=0}^k |\lambda_0|^i \omega_i\right) \|y\|_2 \|x\|_2}{|\lambda_0| \cdot |y^* P'(\lambda_0) x|}, \tag{3.1}$$

where $P'(\lambda)$ denotes the derivative of $P(\lambda)$ with respect to λ.

Theorem 3.2 [22, Theorem 1] *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1). *For a given approximate right eigenpair $(\widetilde{\lambda}_0, \widetilde{x})$ of $P(\lambda)$, the backward error of $(\widetilde{\lambda}_0, \widetilde{x})$ is given by*

$$\eta(\widetilde{\lambda}_0, \widetilde{x}; P) = \frac{\|P(\widetilde{\lambda}_0)\widetilde{x}\|_2}{\left(\sum_{i=0}^k |\widetilde{\lambda}_0|^i \omega_i\right) \|\widetilde{x}\|_2}. \tag{3.2}$$

The following two lemmas will be useful in later sections. Before stating them, we recall that if λ_0 is a simple, finite, nonzero eigenvalue of a matrix polynomial $P(\lambda)$ with associated right eigenvector x, then λ_0^{-1} is a simple eigenvalue of $\mathrm{rev}P(\lambda)$ with associated right eigenvector x.

The immediate proofs of Lemmas 3.1 and 3.2 are omitted.

Lemma 3.1 *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1)*. Let λ_0 be a simple, finite, nonzero eigenvalue of $P(\lambda)$. Then, $\kappa_{\rm r}(\lambda_0; P) = \kappa_{\rm r}(\lambda_0^{-1}; {\rm rev}P)$, when the weights used for* rev$P(\lambda)$ *are equal to those used for $P(\lambda)$ but in reversed order.*

Lemma 3.2 *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1)*. Let $(\widetilde{\lambda}_0, \widetilde{x})$ be an approximate right eigenpair of $P(\lambda)$. If we consider $(\widetilde{\lambda}_0^{-1}, \widetilde{x})$ as an approximate eigenpair of* rev$P(\lambda)$*, then we have $\eta(\widetilde{\lambda}_0, \widetilde{x}; P) = \eta(\widetilde{\lambda}_0^{-1}, \widetilde{x}; {\rm rev}P)$, when the weights used for* rev$P(\lambda)$ *are those used for $P(\lambda)$ but in reversed order.*

The nonnegative weights ω_i in the definitions of $\kappa_r(\lambda_0; P)$ and $\eta(\lambda_0, x; P)$ allow flexibility in how perturbations of $P(\lambda)$ are measured. Typically, one is interested in either coefficient-wise or norm-wise perturbations of $P(\lambda)$. Norm-wise perturbations are obtained by choosing

$$\omega_i := \max_{i=0:k}\{\|A_i\|_2\} \quad \text{for } i = 0 : k.$$

Coefficient-wise perturbations are obtained by choosing

$$\omega_i := \|A_i\|_2 \quad \text{for } i = 0 : k.$$

In this work, we study both norm-wise and coefficient-wise perturbations. When norm-wise perturbations are considered, we write

$$\kappa_{\rm ra}(\lambda_0; P) := \frac{\max_{i=0:k}\{\|A_i\|_2\}\left(\sum_{i=0}^k |\lambda_0|^i\right)\|x\|_2\|y\|_2}{|\lambda_0|\,|y^* P'(\lambda_0)x|}, \quad \text{and}$$
$$\eta_{\rm ra}(\widetilde{\lambda}_0, \widetilde{x}; P) := \frac{\|P(\widetilde{\lambda}_0)\widetilde{x}\|_2}{\max_{i=0:k}\{\|A_i\|_2\}\left(\sum_{i=0}^k |\widetilde{\lambda}_0|^i\right)\|\widetilde{x}\|_2},$$

and refer to $\kappa_{\rm ra}(\lambda_0; P)$ and $\eta_{\rm ra}(\widetilde{\lambda}_0, \widetilde{x}; P)$, respectively, as the *relative-absolute eigenvalue condition number and backward error.* When coefficient-wise perturbations are considered, we write

$$\kappa_{\rm rr}(\lambda_0; P) := \frac{\left(\sum_{i=0}^k |\lambda_0|^i \|A_i\|_2\right)\|x\|_2\|y\|_2}{|\lambda_0|\,|y^* P'(\lambda_0)x|}, \quad \text{and}$$
$$\eta_{\rm rr}(\widetilde{\lambda}_0, \widetilde{x}; P) := \frac{\|P(\widetilde{\lambda}_0)\widetilde{x}\|_2}{\left(\sum_{i=0}^k |\widetilde{\lambda}_0|^i \|A_i\|_2\right)\|\widetilde{x}\|_2},$$

and refer to $\kappa_{\rm rr}(\lambda_0; P)$ and $\eta_{\rm rr}(\widetilde{\lambda}_0, \widetilde{x}; P)$, respectively, as the *relative-relative eigenvalue condition number and backward error.*

Remark 3.1 When the matrix polynomial $P(\lambda)$ is symmetric (resp. Hermitian), it is natural to consider symmetric (resp. Hermitian) perturbations in the definition of condition numbers and backward errors. This leads to the notions of structured condition numbers and structured backward errors. However, as it has been shown in [1], the structured and unstructured condition numbers and backward errors are nearly the same. This is why we only focus on the unstructured ones.

3.1 Sensitivity of the eigenvectors of a matrix pencil

In the next section we will explore the sensitivity of the eigenvectors of the block-symmetric linearizations $D_1(\lambda; P)$ and $D_k(\lambda; P)$ to small perturbations of their matrix coefficients. Theorem 3.3 in this section will be used to provide some intuition behind the fact that the eigenvectors of these two linearizations can be very ill-conditioned even when the corresponding eigenvalue is well-conditioned.

We first introduce an auxiliary lemma that generalizes a well-known result for eigenvectors of matrices.

Lemma 3.3 *Let $L(\lambda) = \lambda B - A$ be a regular matrix pencil. Let λ_1 and λ_2 be two distinct finite eigenvalues of $L(\lambda)$ and let z_1 and w_2 be a right and a left eigenvector of $L(\lambda)$ associated with λ_1 and λ_2, respectively. Then, $w_2^* B z_1 = 0$.*

Proof By definition of right and left eigenvector, we have

$$\lambda_1 B z_1 = A z_1 \quad \text{and} \quad w_2^* \lambda_2 B = w_2^* A.$$

Multiplying the first equality by w_2^* on the left, multiplying the second equality by z_1 on the right and subtracting both expressions, we get $(\lambda_1 - \lambda_2) w_2^* B z_1 = 0$. Since $\lambda_1 \neq \lambda_2$, the result follows.

Theorem 3.3 *Let $L(\lambda) = \lambda B - A$ and $L(\lambda) + \Delta L(\lambda) = \lambda(B + \Delta B) - (A + \Delta A)$ be two $m \times m$ regular matrix pencils, where $\|\Delta B\| \leq \epsilon \|B\|_2$ and $\|\Delta A\| \leq \epsilon \|A\|_2$ for some $\epsilon > 0$ so that L and ΔL have the same number of eigenvalues. Assume that all the eigenvalues of $L(\lambda)$ are simple and finite. Let $\lambda_1, \ldots, \lambda_m$ denote the eigenvalues of $L(\lambda)$, and, for $i = 1 : m$, let z_i be a right eigenvector associated with the eigenvalue λ_i. Let $\lambda_1 + \Delta\lambda_1, \ldots, \lambda_m + \Delta\lambda_m$ denote the eigenvalues of $L(\lambda) + \Delta L(\lambda)$. If $\widetilde{z}_i = z_i + \Delta z_i$ denotes a right eigenvector of $L(\lambda) + \Delta L(\lambda)$ associated with $\lambda_i + \Delta\lambda_i$, then, to first order in ϵ, we have*

$$\operatorname{dist}(\widetilde{z}_i, \operatorname{spam}\{z_i\}) \leq \left[\epsilon \sum_{\ell \neq i} \frac{|\lambda_\ell|}{1 + |\lambda_\ell|} \frac{1 + |\lambda_i|}{|\lambda_i - \lambda_\ell|} \kappa_{\mathrm{ra}}(\lambda_\ell; L) \right] \|z_i\|_2, \tag{3.3}$$

where dist *denotes the Euclidean distance, and $\kappa_{\mathrm{ra}}(\lambda_\ell; L)$ denotes the relative-absolute eigenvalue condition number of λ_ℓ.*

Proof Since the vectors $z_1, \ldots, z_m$ form a basis for $\mathbb{F}^m$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, we have $\widetilde{z}_i = z_i + \sum_{\ell=1}^m c_\ell z_\ell$, for some constants c_ℓ. Then, notice that $\operatorname{dist}(\widetilde{z}_i, \operatorname{spam}\{z_i\}) \leq \|\widetilde{z}_i - v\|_2$, for any vector $v \in \operatorname{spam}\{z_i\}$. Hence, taking $v = z_i + c_i z_i$, we get

$$\operatorname{dist}(\widetilde{z}_i, \operatorname{spam}\{z_i\}) \leq \| \sum_{\ell \neq i} c_\ell z_\ell \|_2 \leq \sum_{\ell \neq i} |c_\ell| \, \|z_\ell\|_2. \tag{3.4}$$

To finish the proof, we need to bound the scalars $|c_\ell|$. By Lemma 3.3, denoting by w_ℓ a left eigenvector of $L(\lambda)$ associated with λ_ℓ,

$$w_\ell^* B z_i = 0 \quad \text{for } \ell \neq i.$$

The B-orthogonality of left and right eigenvectors implies that the scalars c_ℓ are given by

$$c_\ell = \frac{w_\ell^* B \Delta z_i}{w_\ell^* B z_\ell} \quad \text{for } \ell = 1 : m,$$

where $w_\ell^* B z_\ell \neq 0$ because the eigenvalues of $L(\lambda)$ are simple.

Expanding to first order in ϵ the equality

$$(\lambda_i + \Delta\lambda_i)(B + \Delta B)(z_i + \Delta z_i) = (A + \Delta A)(z_i + \Delta z_i),$$

we find

$$\Delta\lambda_i B z_i + \lambda_i \Delta B z_i + \lambda_i B \Delta z_i = A \Delta z_i + \Delta A z_i. \tag{3.5}$$

Multiplying (3.5) on the left by w_ℓ^*, with $\ell \neq i$, and taking into account that $w_\ell^* A = \lambda_\ell w_\ell^* B$ and $w_\ell^* B z_i = 0$, yields

$$\lambda_i w_\ell^* \Delta B z_i + \lambda_i w_\ell^* B \Delta z_i = \lambda_\ell w_\ell^* B \Delta z_i + w_\ell^* \Delta A z_i.$$

Hence,

$$w_\ell^* B \Delta z_i = \frac{\lambda_i w_\ell^* \Delta B z_i - w_\ell^* \Delta A z_i}{\lambda_\ell - \lambda_i},$$

and, so,

$$\frac{w_\ell^* B \Delta z_i}{w_\ell^* B z_\ell} = \frac{1}{\lambda_\ell - \lambda_i} \frac{\lambda_i w_\ell^* \Delta B z_i - w_\ell^* \Delta A z_i}{w_\ell^* B z_\ell}. \tag{3.6}$$

Plugging (3.6) into (3.4), taking norms, using the triangle inequality, and using $\|\Delta B\| \leq \epsilon \|B\|_2$ and $\|\Delta A\| \leq \epsilon \|A\|_2$, we get

$$\sum_{\ell \neq i} |c_\ell| \, \|z_\ell\|_2 \leq \epsilon \sum_{\ell \neq i} \frac{1}{|\lambda_\ell - \lambda_i|} \frac{\|w_\ell\|_2 \|z_\ell\|_2 \|z_i\| (|\lambda_i| \, \|B\|_2 + \|A\|_2)}{|w_\ell^* B z_\ell|}.$$

The result now readily follows from the formula for the relative-absolute condition number $\kappa_{\mathrm{ra}}(\lambda_\ell; L)$ taking into account that

$$\frac{(|\lambda_i| \, \|B\|_2 + \|A\|_2)}{\max\{\|A\|_2, \|B\|_2\}(1 + |\lambda_\ell)} \leq \frac{1 + |\lambda_i|}{1 + |\lambda_\ell|}.$$

Remark 3.2 We note that Theorem 3.3 implies that the relative error

$$\frac{\mathrm{dist}\,(\widetilde{z}_i, \mathrm{spam}\{z_i\})}{\|z_i\|_2}$$

in the eigenvector z_i associated with the eigenvalue λ_i can potentially be large when λ_i is close to be a multiple eigenvalue or if any of the eigenvalues other than λ_i is ill-conditioned. It is well-known that the eigenvalues of $D_1(\lambda; P)$ (resp. $D_k(\lambda; P)$) with small modulus (resp. large modulus) tend to be very ill-conditioned which can potentially be a reason why, as we will show numerically in the next section, the eigenvectors of $D_1(\lambda; P)$ (resp. $D_k(\lambda; P)$) associated with eigenvalues of large modulus (resp. small modulus) can be very ill-conditioned.

We would like to mention that there have been other attempts in the literature to study the sensitivity of the eigenvectors of a matrix polynomial to small changes in its matrix coefficients. See for example [21].

4 Using the linearizations $D_1(\lambda; P)$ and $D_k(\lambda; P)$.

In this section we debunk the common belief that the pencils $D_1(\lambda; P)$ and $D_k(\lambda; P)$ in the vector space $\mathbb{DL}(P)$ [16] given by

$$D_1(\lambda; P) := \lambda \begin{bmatrix} A_k & 0 & \cdots & \cdots & 0 \\ 0 & -A_{k-2} & -A_{k-3} & \cdots & -A_0 \\ \vdots & -A_{k-3} & \cdots & -A_0 & 0 \\ \vdots & \vdots & \iddots & \iddots & \vdots \\ 0 & -A_0 & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_1 & A_0 \\ A_{k-2} & \cdots & A_1 & A_0 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ A_1 & A_0 & 0 & \cdots & 0 \\ A_0 & 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{4.1}$$

and

$$D_k(\lambda; P) := \lambda \begin{bmatrix} 0 & \cdots & \cdots & 0 & A_k \\ 0 & \cdots & 0 & A_k & A_{k-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 0 & A_k & A_{k-1} & \cdots & A_2 \\ A_k & A_{k-1} & \cdots & A_2 & A_1 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & -A_k & 0 \\ \vdots & \iddots & \iddots & \vdots & \vdots \\ 0 & -A_k & \cdots & -A_3 & \vdots \\ -A_k & \cdots & -A_3 & -A_2 & 0 \\ 0 & \cdots & \cdots & 0 & A_0 \end{bmatrix} \tag{4.2}$$

are "good" linearizations of a symmetric (resp. Hermitian) matrix polynomial $P(\lambda)$ as in (1.1). This belief is based on the following two results for the relative-absolute conditioning of eigenvalues and backward errors of approximate eigenpairs, as well as on analogous results for the relative-relative case that can be obtained just multiplying the bounds in Theorems 4.1 and 4.2 by the constant

$$\rho := \frac{\max_{i=0:k}\{\|A_i\|_2\}}{\min\{\|A_0\|_2, \|A_k\|_2\}}. \tag{4.3}$$

Theorem 4.1 (Conditioning of $D_1(\lambda; P)$ and $D_k(\lambda; P)$) [4, Theorem 6.1] *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1). *Assume 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 $\ell = k$. Then,*

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

Remark 4.1 As we mentioned in Remark 3.2, the eigenvalues of $D_1(\lambda; P)$ (resp. $D_k(\lambda; P)$) with small modulus (resp. large modulus) tend to be very ill-conditioned. The previous theorem shows that if $|\lambda_0| < 1$, then

$$\kappa_{\mathrm{ra}}(\lambda_0; D_1) \ge |\lambda_0|^{1-k} \kappa_{\mathrm{ra}}(\lambda_0; P)$$

and if $|\lambda_0| > 1$, then

$$\kappa_{\mathrm{ra}}(\lambda_0; D_k) \ge |\lambda_0|^{k-1} \kappa_{\mathrm{ra}}(\lambda_0; P).$$

Thus, even if the condition number of λ_0 as an eigenvalue of P is relatively small, the condition number of λ_0 as an eigenvalue of $D_1(\lambda; P)$ (resp. $D_k(\lambda; P)$) can grow significantly.

Theorem 4.2 (Backward errors of $D_1(\lambda;P)$ and $D_k(\lambda;P)$) [4, Theorem 6.2] *Let $P(\lambda)$ be a regular matrix polynomial of degree k as in* (1.1)*. Let $\ell \in \{1,k\}$ and suppose that A_0 is nonsingular if $\ell=1$, and A_k is nonsingular if $\ell=k$. Let $(\widetilde{\lambda}_0,\widetilde{z})$ be an approximate right eigenpair of $D_\ell(\lambda,P)$, with $\widetilde{\lambda}_0$ nonzero and finite, and let $\widetilde{z}_\ell := (e_\ell^T \otimes I_n)\widetilde{z}$. If $(\widetilde{\lambda}_0,\widetilde{z}_\ell)$ is considered an approximate right eigenpair for $P(\lambda)$, then*

$$\frac{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z}_\ell;P)}{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z};D_\ell)} \le k^{3/2}\frac{\|\widetilde{z}\|_2}{\|\widetilde{z}_\ell\|_2}.$$

Remark 4.2 It is well-known that any (right) eigenvector of $D_1(\lambda;P)$ or $D_k(\lambda;P)$ associated with λ_0 is of the form

$$z = \begin{bmatrix}\lambda_0^{k-1} & \cdots & \lambda_0 & 1\end{bmatrix}^T x, \tag{4.4}$$

for some (right) eigenvector x of $P(\lambda)$ associated with λ_0. This implies that for exact z and z_ℓ, we get

$$\frac{\|z\|_2}{\|z_\ell\|_2} \le \begin{cases} \sqrt{k}\max\{1,|\lambda_0|^{1-k}\} & \text{if } \ell=1, \text{ and} \\ \sqrt{k}\max\{1,|\lambda_0|^{k-1}\} & \text{if } \ell=k. \end{cases} \tag{4.5}$$

Assuming that (4.5) holds for the computed eigenpairs in Theorem 4.2, we get the following upper bounds

$$\frac{\eta_{\rm ra}(\lambda_0,z_\ell;P)}{\eta_{\rm ra}(\lambda_0,z;D_\ell)} \le \begin{cases} k^{5/2}\max\{1,|\lambda_0|^{k-1}\} & \text{if } \ell=k, \text{ and} \\ k^{5/2}\max\{1,|\lambda_0|^{1-k}\} & \text{if } \ell=1. \end{cases}$$

which are in accordance with the conditioning results in Theorem 4.1.

Based on the ideas discussed in Remark 4.2, the following strategy for computing eigenpairs of a matrix polynomial $P(\lambda)$ with small backward errors (at least in the relative-absolute sense, or in the relative-relative sense when the polynomial is well-scaled, i.e., $\rho \approx 1$) has been proposed.

1 Apply a backward stable eigenvalue algorithm, like the QZ algorithm, to the linearizations $D_1(\lambda;P)$ and $D_k(\lambda;P)$.
2 For the computed eigenvalues with modulus less than or equal to one, recover the eigenvectors of $P(\lambda)$ from the kth block z_k of the corresponding eigenvectors of $D_k(\lambda;P)$.
3 For the computed eigenvalues with modulus greater than one, recover the eigenvectors of $P(\lambda)$ from the first block z_1 of the corresponding eigenvectors of $D_1(\lambda;P)$.

Next, we argue that this strategy does not always guarantee small backward errors due to the extreme sensitivity of the eigenvectors of $D_1(\lambda;P)$ and $D_k(\lambda;P)$ to small perturbations of the coefficients of these pencils. Our explanation focuses on $D_k(\lambda;P)$ (since similar comments can be made for $D_1(\lambda;P)$). We will also illustrate these facts with numerical experiments.

Let us assume that a polynomial eigenvalue problem associated with a symmetric/Hermitian matrix polynomial $P(\lambda)$ is solved by using the linearization $D_k(\lambda;P)=\lambda B - A$. Assume $P(\lambda)$ has been scaled so that $\max_{i=0:k}\{\|A_i\|_2\}=1$. Theorem 4.2 and Remark 4.2 suggest that, if $|\lambda_0|\le 1$, one should be able to compute an approximate eigenpair $(\widetilde{\lambda}_0,\widetilde{z}_k)$ of $P(\lambda)$ from a computed eigenpair $(\widetilde{\lambda}_0,\widetilde{z})$

of $D_k(\lambda;P)$ with a small backward error $\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z}_k;P)$. However, in the numerical examples that we show next, we will see that this is not necessarily true. This does not imply that there is something wrong with the results in Theorem 4.2 and Remark 4.2. The problem is that, in floating point arithmetic, we cannot assume that the ratio $\|\widetilde{z}\|_2/\|\widetilde{z}_k\|_2$ is bounded by a moderate constant, the reason being the potentially large sensitivity of the eigenvectors of $D_k(\lambda;P)$ to small perturbations in the coefficients of the linearization. We give an intuitive explanation for this sensitivity to perturbations as follows. Let z and z_k denote, respectively, the exact eigenvector of $D_k(\lambda_0;P)$ associated with the eigenvalue λ_0 and its kth block. Let $\widetilde{z}$ denote the computed eigenvector of $D_k(\lambda;P)$ associated with the computed eigenvalue $\widetilde{\lambda}_0$. Then, there exists a positive constant α such that

$$\frac{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z}_k;P)}{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z};D_k)} \le k^{3/2}\frac{\|\widetilde{z}\|_2}{\|\widetilde{z}_k\|_2} \le k^{3/2}\alpha\frac{\|z\|_2}{\|\widetilde{z}_k\|2} \le \alpha k^2\frac{\|z_k\|_2}{\|\widetilde{z}_k\|_2}, \tag{4.6}$$

where the last inequality follows from (4.5).

As the numerical experiments will show, the ratio $\mu := \frac{\|z_k\|_2}{\|\widetilde{z}_k\|_2}$ is very large for some eigenvectors and, surprisingly, it is a very accurate predictor of $\frac{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z}_k;P)}{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z};D_\ell)}$ when $|\lambda_0| \le 1$. We must point out that both z_k and $\tilde{z}_k$ in our experiments are the eigenvectors computed by Matlab. The exact eigenvector was computed transforming the constructed matrix polynomial to a symbolic object. Moreover, we have observed that, in the cases when the ratio μ is very large, $\widetilde{z}_k$ is very close to 0. This implies that the small backward errors introduced by the QZ algorithm may destroy the exact structure (4.4) of the eigenvectors of $D_k(\lambda;P)$. Conclusively, in floating point arithmetic we cannot assume computed eigenvectors of the form (4.4) and, thus, we cannot assume that $\frac{\|\widetilde{z}\|_2}{\|\widetilde{z}_k\|_2}$ is small.

Next we present two numerical examples illustrating that the strategy of solving a PEP with the combined use of $D_1(\lambda;P)$ and $D_k(\lambda;P)$ is potentially unstable. In particular, we show that using $D_k(\lambda;P)$ for computing the eigenvalues with modulus less than 1 can increase the backward error of a computed eigenpair up to the point in which most of the accuracy is lost.

In the first numerical experiment, we consider a random matrix polynomial of degree 4 and size $n = 20$. The matrix polynomial is constructed in MATLAB as follows:

$$\begin{aligned}
A_0 &= \texttt{1e2} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_1 &= \texttt{1e1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_2 &= \texttt{1e2} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_3 &= \texttt{1e7} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_4 &= \texttt{1e1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});
\end{aligned} \tag{4.7}$$

and then, we computed $A_i' := A_i + A_i^T$ so that the matrix polynomial is symmetric. This matrix polynomial has 60 out of its 80 eigenvalues with modulus between 1 and 10^{-2} while the rest of the eigenvalues have modulus larger than 10^2. Moreover, the eigenvalues with modulus larger than one have condition number larger than 10^{23} (recall Remark 3.2).

In Figure 4.1, we plot the ratio of backward errors $\frac{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z}_k;P)}{\eta_{\rm ra}(\widetilde{\lambda}_0,\widetilde{z};D_k)}$ for all the eigenvalues λ_0 of $P(\lambda)$ ordered in increasing order of modulus. This graph is denoted

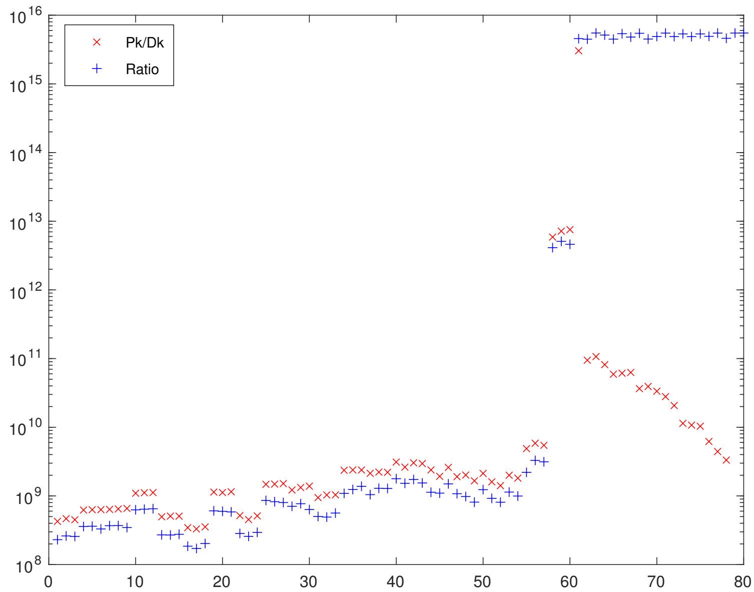

Fig. 4.1 Relative-absolute ratio of backward errors using $D_k(\lambda; P)$ and bound when P is not well scaled

by "Pk/Dk" in the legend of the figure. We also plot the ratio $\frac{\|z_k\|_2}{\|\widetilde{z}_k\|_2}$, denoted by "Ratio" in the legend. We observe that the exact ratio of backward errors for the eigenvalues of modulus less than 1 range between 10^7 and 10^{13}. Moreover, we observe that the function"Ratio" fully predicts the values of these ratios. This indicates that, in the computation of the right eigenvectors of $P(\lambda)$ associated with the "small" eigenvalues, the norm of the last block of the exact eigenvector is very sensitive to changes in the coefficients of $P(\lambda)$ and therefore, using $D_k(\lambda; P)$ to compute these eigenvalues is not a good strategy.

The problems in the backward errors observed in this numerical experiment could be attributed to the fact that the polynomial $P(\lambda)$ is not well scaled. In our second numerical experiment, we show that this problem can be observed also in the case in which $P(\lambda)$ is well scaled although in this case fewer eigenvalues have large ratio of backward errors. In this example, we consider again a random matrix polynomial of degree 4 and size $n = 20$. The matrix polynomial is constructed in MATLAB as follows:

$$\begin{aligned}
A_0 &= (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_1 &= (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_2 &= (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_3 &= (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_4 &= (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});
\end{aligned} \tag{4.8}$$

and then, we computed $A_i' := A_i + A_i^T$ so that the matrix polynomial is symmetric. Moreover, we changed the singular values of A_0' and A_k' so that these two matrix coefficients keep their norm but so that the matrix polynomial has 6 eigenvalues with modulus between 10^{-7} and 10^{-5}. The first 46 eigenvalues have modulus less than or equal to 1 and all the eigenvalues have modulus less than 10. In this case, six of the eigenvalues with modulus larger than 1 have condition number larger than 10^{21}.

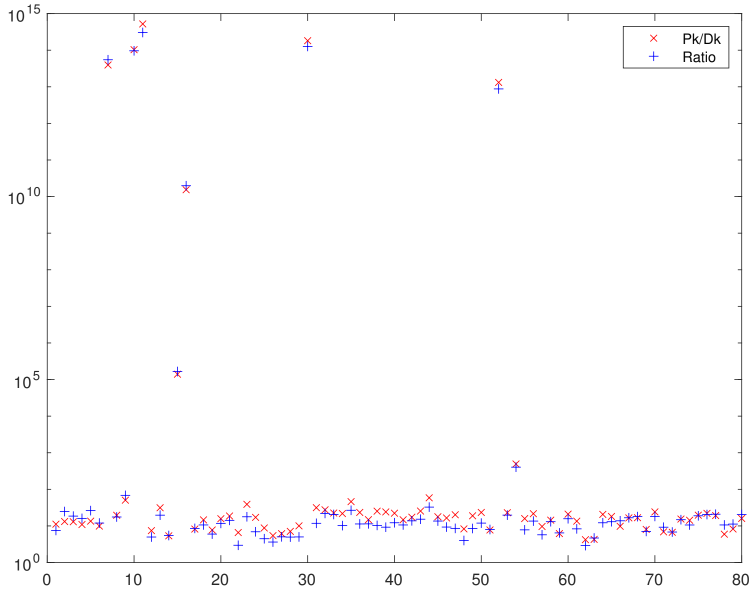

Fig. 4.2 Relative-absolute ratio of backward errors using $D_k(\lambda; P)$ and bound when P is well scaled.

In Figure 4.2, we plot the functions "Pk/Dk", and "Ratio" as we did in the first numerical experiment. We observe that, for some of the eigenvalues with modulus less than one, the ratio of backward errors is of order 10^{15} and that the behavior of the ratio of backward errors can also be fully predicted by the value of the ratio $\frac{\|z_k\|_2}{\|\tilde{z}_k\|_2}$, as happened in the first experiment. We must point out that this behavior is not unique to the two numerical experiments presented here but that it was observed in a multitude of different numerical experiments.

In conclusion, we cannot guarantee that the eigenpairs associated with eigenvalues of small modulus of a matrix polynomial can be computed accurately from $D_k(\lambda; P)$, specially when $P(\lambda)$ is not well scaled. Similar conclusions can be obtained for eigenvalues of large modulus when the linearization $D_1(\lambda; P)$ is used.

We also want to point out that, when $D_k(\lambda; P)$ (resp. $D_1(\lambda; P)$) does not compute eigenpairs associated with small (resp. large) modulus eigenvalues accurately, $D_1(\lambda; P)$ (resp. $D_k(\lambda; P)$), in general, does not either, as we show next. In Figures 4.3 and 4.4 we present two examples in which the ratio of backward errors is plotted when $D_1(\lambda; P)$ is used as a linearization of a matrix polynomial $P(\lambda)$ (blue graph) and when $D_k(\lambda; P)$ is used as a linearization of $P(\lambda)$ (red graph). In both cases, there are eigenvalues that are not accurately computed by neither $D_1(\lambda; P)$ nor by $D_k(\lambda; P)$. In Figure 4.3, the eigenvalues of small modulus are not accurately computed while in Figure 4.4, the eigenvalues of large modulus are not accurately computed.

For Figure 4.3, we constructed a matrix polynomial using the same strategy as in the first experiment but using the coefficients:

$$\begin{aligned}
A_0 &= \texttt{1e1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_1 &= \texttt{1e2} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_2 &= \texttt{1e} - \texttt{1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_3 &= \texttt{1e8} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_4 &= \texttt{1e1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});
\end{aligned} \tag{4.9}$$

For Figure 4.4, we used the coefficients

$$\begin{aligned}
A_0 &= \texttt{1e1} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_1 &= \texttt{1e11} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_2 &= \texttt{1e18} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_3 &= \texttt{1e10} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});\\
A_4 &= \texttt{1e12} * (\texttt{randn(n)} + \texttt{sqrt}(-1) * \texttt{randn(n)});
\end{aligned} \tag{4.10}$$

We have not observed such pathological behavior from the alternative linearizations that we propose in this work. As an illustration, we show in Figures 4.5 and 4.6 the relative-absolute backward error ratios for the linearizations $D_k(\lambda; P)$, $D_1(\lambda; P)$ and the linearizations that we denote for now as DH and DG but we formally introduce in (5.7) and (5.9), with $S = A_k$. The two experiments are the same as those presented in Figures 4.3 and 4.4 but adding now the ratios for DH and DG. Note that the combined used of the linearizations DH and DG allow to compute all eigenpairs accurately.

5 Using $\mathcal{T}_P(\lambda)$ for even-degree matrix polynomials.

A well-known block-symmetric strong linearization for odd degree matrix polynomials $P(\lambda)$ as in (1.1) is the pencil

$$\mathcal{T}_P^k(\lambda) := \begin{bmatrix}
\lambda A_k + A_{k-1} & -I_n & & & & & \\
-I_n & 0 & \lambda I_n & & & & \\
& \lambda I_n & \lambda A_{k-2} + A_{k-3} & -I_n & & & \\
& & -I_n & 0 & & & \\
& & & & \ddots & & \\
& & & & & 0 & \lambda I_n \\
& & & & & \lambda I_n & \lambda A_1 + A_0
\end{bmatrix}, \tag{5.1}$$

introduced in [2]. The missing blocks in this matrix and in any other matrices in the sequel, as usual, represent zero blocks. The pencil $\mathcal{T}_P^k(\lambda)$ was proven to

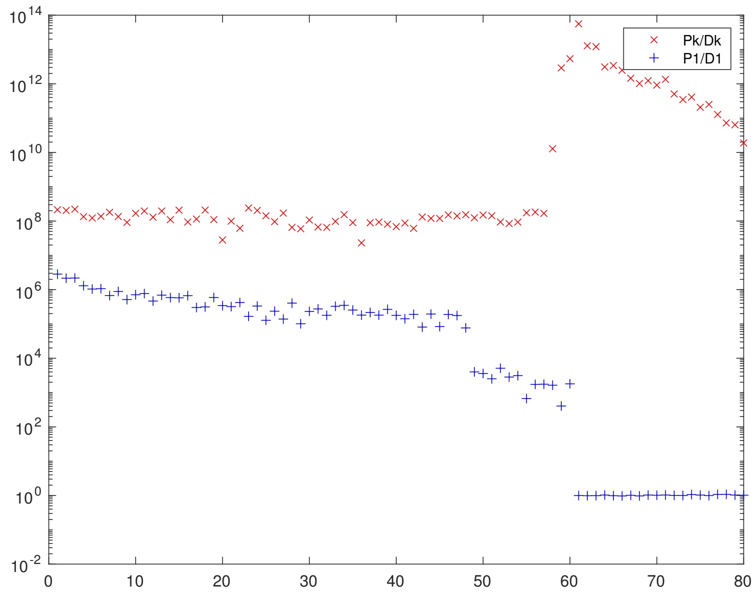

Fig. 4.3 Relative-absolute ratio of backward errors using D_k and D_1, and small eigenvalues cannot be accurately computed.

enjoy excellent numerical properties in terms of conditioning of eigenvalues and backward errors in [4]. Our goal in this paper is to find structured linearizations of even-degree matrix polynomials and, unfortunately, this pencil cannot be used as a linearization of such matrix polynomials since its structure requires odd degree.

One possible strategy to construct a (symmetric or Hermitian) strong linearization of an even-degree (symmetric or Hermitian) matrix polynomial and, at the same time, try to take advantage of the good numerical properties of $\mathcal{T}_P^k$ is to transform our matrix polynomial of even degree k into an odd grade matrix polynomial by adding the term $0 \cdot \lambda^{k+1}$, that is, to consider the matrix polynomial

$$\widetilde{P}(\lambda) = 0 \cdot \lambda^{k+1} + \lambda^k A_k + \cdots + \lambda A_1 + A_0.$$

By applying the linearization (5.1) to $\widetilde{P}(\lambda)$, we obtain the pencil

$$\mathcal{H}_P(\lambda) := \mathcal{T}_P^{k+1}(\lambda) = \begin{bmatrix} A_k & -I_n & & & & & \\ -I_n & 0 & \lambda I_n & & & & \\ & \lambda I_n & \lambda A_{k-1} + A_{k-2} & -I_n & & & \\ & & -I_n & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \lambda I_n \\ & & & & & \lambda I_n & \lambda A_1 + A_0 \end{bmatrix}, \quad (5.2)$$

which is a strong linearization of the matrix polynomial $P(\lambda)$ when seen as a polynomial of grade $k+1$. We must observe though that the linearization (5.2)

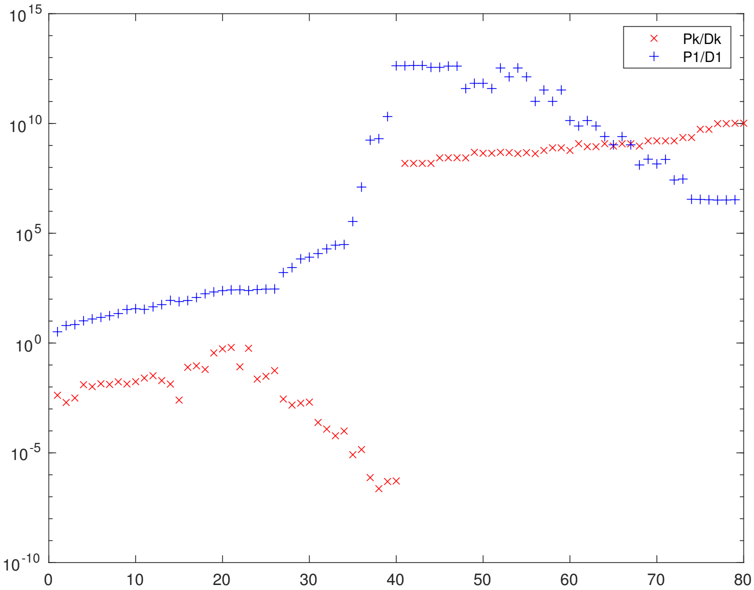

Fig. 4.4 Relative-absolute ratio of backward errors using D_k and D_1, and large eigenvalues cannot be accurately computed.

has n eigenvalues at infinity that were not present in the original polynomial eigenvalue problem. Therefore, before we try to compute the eigenvalues of $P(\lambda)$ from $\mathcal{T}_P^{k+1}(\lambda)$, it is necessary to deflate the n extra eigenvalues at infinitiy. In the following section, we show how the deflation can be done.

5.1 Deflating the spurious eigenvalues of $\mathcal{H}_P(\lambda)$

Next we show how to deflate the n spurious eigenvalues at infinity of $\mathcal{H}_P(\lambda)$, assuming that A_k is nonsingular and symmetric/Hermitian. In Section 5.2 we present an alternative to $\mathcal{H}_P(\lambda)$ when A_k is singular but A_0 is not.

In order to deflate the n spurious eigenvalues of $\mathcal{H}_P(\lambda)$ while preserving the symmetric or Hermitian structure, we need to find a nonsingular matrix U such that

$$U^*\mathcal{H}_P(\lambda)U = \begin{bmatrix} \mathcal{H}_1(\lambda) & 0 \\ 0 & \mathcal{H}_2(\lambda) \end{bmatrix},$$

where $\mathcal{H}_1(\lambda)$ is a pencil whose eigenvalues are exactly the n extra eigenvalues at infinity, (recall that, for any matrix A, A^* denotes the conjugate transpose of A). Note that, since $\mathcal{H}_P(\lambda)$ and $U^*\mathcal{H}_P(\lambda)U$ are strictly equivalent, both matrices have the same eigenvalues. Thus, $\mathcal{H}_2(\lambda)$ is a pencil with the same eigenvalues as $P(\lambda)$. Moreover, since $\mathcal{H}_P(\lambda)$ and $U^*\mathcal{H}_P(\lambda)U$ are congruent, one of these pencils is symmetric (resp. Hermitian) if and only if the other is.

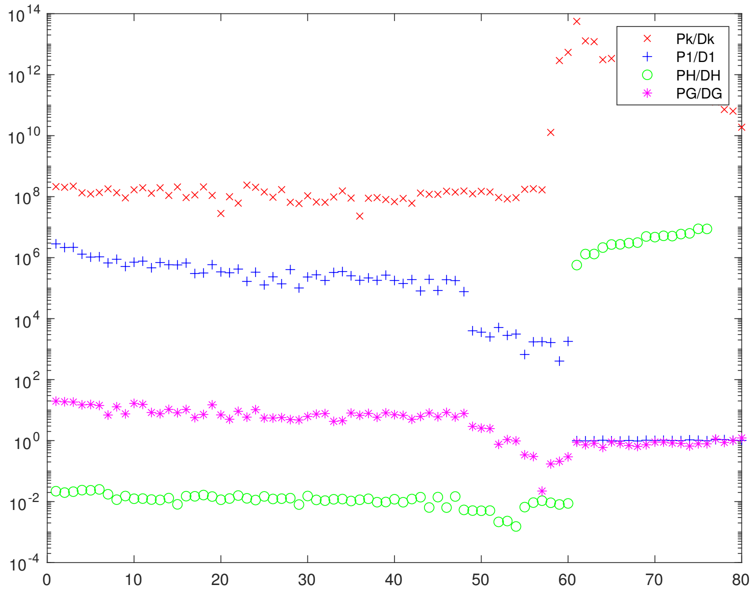

Fig. 4.5 Relative-absolute ratio of backward errors using D_k, D_1, DH and DG.

As we will show, in order to construct the nonsingular matrix U, we only need to find a matrix whose columns form a basis for the nullspace of

$$M = \begin{bmatrix} A_k & -I_n \end{bmatrix}.$$

Notice that $\dim(\mathrm{null}(M)) = n$ because M has full row rank. Obvious choices for matrices whose columns span the nullspace of M are

$$\begin{bmatrix} I_n \\ A_k \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} A_k^{-1} \\ I_n \end{bmatrix}. \tag{5.3}$$

Another alternative for constructing a basis for the nullspace of M is via a rank revealing factorization of M (via the QR factorization with column pivoting or the SVD, for example).

Let $V = \left[\begin{smallmatrix} T \\ S \end{smallmatrix}\right]$ be any $2n \times n$ full-column-rank matrix such that $MV = 0$, or equivalently, $A_k T - S = 0$. Since A_k is nonsingular, the product by A_k preserves the linearly independency of the columns of T, that is, the set of indices corresponding to the linearly independent columns of T is the same as that of S. Since V has full rank, it follows that the matrix S is nonsingular. Hence, the following pencil is strictly equivalent and congruent to $\mathcal{H}_P(\lambda)$, and therefore, has the same eigenvalues as $\mathcal{H}_P(\lambda)$:

$$\left[\begin{array}{cc|c} I_n & 0 & \\ T^* & S^* & \\ \hline & & I_{(k-1)n} \end{array}\right] \mathcal{H}_P(\lambda) \left[\begin{array}{cc|c} I_n & T & \\ 0 & S & \\ \hline & & I_{(k-1)n} \end{array}\right]. \tag{5.4}$$

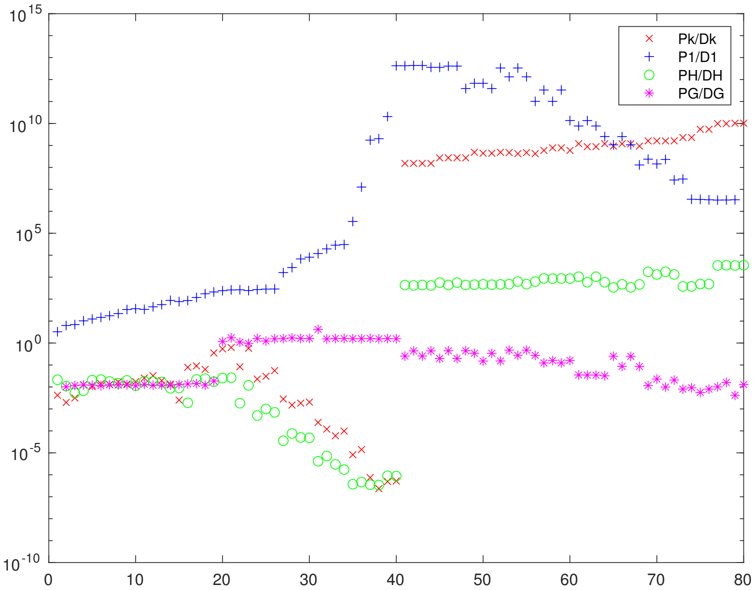

Fig. 4.6 Relative-absolute ratio of backward errors using D_k, D_1, DH and DG.

Moreover, since $A_kT - S = 0$ and $T^*A_k - S^* = 0$, the pencil in (5.4) can be expressed as

$$\left[\begin{array}{c|cccccc} A_k & & & & & & \\ \hline & -S^*A_k^{-1}S & \lambda S^* & & & & \\ & \lambda S & \lambda A_{k-1} + A_{k-2} & -I_n & & & \\ & & -I_n & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \lambda I_n \\ & & & & & \lambda I_n & \lambda A_1 + A_0 \end{array}\right]. \tag{5.5}$$

We note that the pencil $A_k = 0 \cdot \lambda + A_k$ has exactly n eigenvalues at infinity since $\mathrm{rev}_1(A_k) = A_k\lambda + 0$ has n zero eigenvalues. Conclusively, the deflation of the spurious eigenvalues at infinity produces the pencil

$$\mathcal{H}_P^S(\lambda) := \begin{bmatrix} -S^*A_k^{-1}S & \lambda S^* & & & & \\ \lambda S & \lambda A_{k-1} + A_{k-2} & -I_n & & & \\ & -I_n & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \lambda I_n \\ & & & & \lambda I_n & \lambda A_1 + A_0 \end{bmatrix}, \tag{5.6}$$

for any nonsingular matrix S.

Remark 5.1 When any of the matrices in (5.3) are employed in the deflation procedure (i.e. when we choose $T = I_n$ and $S = A_k$, or when we choose $T = A_k^{-1}$ and $S = I_n$), the corresponding pencil $\mathcal{H}_P^S$ has already appeared in the literature. More precisely, for $V = \left[\begin{smallmatrix} I_n \\ A_k \end{smallmatrix}\right]$, we get

$$\mathcal{H}_P^{A_k} = \begin{bmatrix} -A_k & \lambda A_k & & & & \\ \lambda A_k & \lambda A_{k-1} + A_{k-2} & -I_n & & & \\ & -I_n & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \lambda I_n \\ & & & & \lambda I_n & \lambda A_1 + A_0 \end{bmatrix}, \tag{5.7}$$

which is a permuted version of the extended block Kronecker pencil $\mathcal{E}_2^P(\lambda)$ in [5, Section 4.4]. For $V = \left[\begin{smallmatrix} A_k^{-1} \\ I_n \end{smallmatrix}\right]$, we get

$$\mathcal{H}_P^{I_n}(\lambda) = \begin{bmatrix} -A_k^{-1} & \lambda I_n & & & & \\ \lambda I_n & \lambda A_{k-1} + A_{k-2} & -I_n & & & \\ & -I_n & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \lambda I_n \\ & & & & \lambda I_n & \lambda A_1 + A_0 \end{bmatrix},$$

which was originally introduced in [2].

Remark 5.2 When V is chosen so that its columns are an orthonormal basis for the nullspace of $\begin{bmatrix} A_k & -I_n \end{bmatrix}$, we refer to the resulting matrix S as S_{MX}. The reason for this is that our structured deflation procedure coincides with the structured deflation procedure proposed by Mehrmann and Xu [17] when their method is applied to the pencil $\mathcal{H}_P(\lambda)$.

In the next result we show that the pencil $\mathcal{H}_P^S(\lambda)$ is a strong linearization of even degree matrix polynomials $P(\lambda)$ even if A_k is not symmetric/Hermitian.

Theorem 5.1 *Let $P(\lambda)$ be an even-degree regular matrix polynomial as in* (1.1) *with nonsingular A_k and let S be a nonsingular matrix. Then, the pencil $\mathcal{H}_P^S(\lambda)$ as in* (5.6) *is a strong linearization of $P(\lambda)$.*

Proof First, we note that the pencil $\mathcal{H}_P^{A_k}(\lambda)$ is permutationally equivalent to the pencil $\mathcal{E}_2^P(\lambda)$ defined in [5, Section 4.4]. More precisely, there exists a block permutation matrix

$$\Pi_1 := \Pi_{(1,2,\frac{k}{2}+2,3,\frac{k}{2}+3,\ldots,\frac{k}{2},k,\frac{k}{2}+1)}$$

such that $\mathcal{H}_P^{A_k}(\lambda) = \Pi_1 \mathcal{E}_2^P(\lambda) \Pi_1^{\mathcal{B}}$. Since the pencil $\mathcal{E}_2^P(\lambda)$ is a strong linearization of $P(\lambda)$ if A_k is nonsingular (see [5, Theorem 4.15]), we deduce that $\mathcal{H}_P^{A_k}(\lambda)$ is a strong linearization of $P(\lambda)$ as well. Second, observe that

$$\mathcal{H}_P^S(\lambda) = \begin{bmatrix} S^* A_k^{-1} & 0 \\ 0 & I_{n(k-1)} \end{bmatrix} \mathcal{H}_P^{A_k}(\lambda) \begin{bmatrix} A_k^{-1} S & 0 \\ 0 & I_{n(k-1)} \end{bmatrix}.$$

Since both A_k and S are nonsingular, this is an equivalence transformation. Thus, the pencil $\mathcal{H}_P^S(\lambda)$ is a strong linearization of $P(\lambda)$. □

Theorem 5.2 establishes two right-sided factorizations of the linearization $\mathcal{H}_P^S(\lambda)$. These factorizations will be key for studying the numerical properties (conditioning and backward errors) of this pencil.

Theorem 5.2 *Let $P(\lambda)$ be an even degree matrix polynomial as in* (1.1)*, let S be an $n \times n$ nonsingular matrix, let $\mathcal{H}_P^S(\lambda)$ be as in* (5.6)*, and let $P_i(\lambda)$ and $P^i(\lambda)$, $i = 0 : k$, be the matrix polynomials defined in* (2.2) *and* (2.4)*. Define the $kn \times n$ matrix polynomials*

$$\Delta_1(\lambda) := \begin{bmatrix} \lambda^{\frac{k}{2}} S^{-1} A_k \\ \lambda^{\frac{k-2}{2}} I_n \\ \lambda^{\frac{k-2}{2}} P_2(\lambda) \\ \vdots \\ \lambda^2 P_{k-4}(\lambda) \\ \lambda I_n \\ \lambda P_{k-2}(\lambda) \\ I_n \end{bmatrix} \quad \text{and} \quad \Delta_2(\lambda) := \begin{bmatrix} \lambda^{k-1} S^{-1} A_k \\ \lambda^{k-2} I_n \\ -P^{k-3}(\lambda) \\ \lambda^{k-3} I_n \\ -\lambda P^{k-5}(\lambda) \\ \vdots \\ \lambda^{\frac{k-4}{2}} P^1(\lambda) \\ \lambda^{\frac{k-2}{2}} I_n \end{bmatrix}. \tag{5.8}$$

Then, the following right-sided factorizations hold

$$\mathcal{H}_P^S(\lambda)\Delta_1(\lambda) = e_k \otimes P(\lambda) \quad \text{and} \quad \mathcal{H}_P^S(\lambda)\Delta_2(\lambda) = e_2 \otimes P(\lambda),$$

where e_i denotes the ith column of the $k \times k$ identity matrix.

Proof For simplicity, we omit the dependence on λ in the Horner polynomials $P_i(\lambda)$ and $P^i(\lambda)$. Let $\mathcal{H}_P^S(\lambda) =: \lambda\mathcal{H}_1 - \mathcal{H}_0$. A direct computation shows that

$$\mathcal{H}_1\Delta_1(\lambda) = \begin{bmatrix} \lambda^{\frac{k-2}{2}} S^* \\ \lambda^{\frac{k}{2}} A_k + \lambda^{\frac{k-2}{2}} A_{k-1} \\ \lambda^{\frac{k-4}{2}} I_n \\ \lambda^{\frac{k-2}{2}} P_2 + \lambda^{\frac{k-4}{2}} A_{k-3} \\ \vdots \\ I_n \\ \lambda P_{k-2} + A_1 \end{bmatrix} \quad \text{and} \quad \mathcal{H}_0\Delta_1(\lambda) = \begin{bmatrix} -\lambda^{\frac{k}{2}} S^* \\ \lambda^{\frac{k-2}{2}} A_{k-2} - \lambda^{\frac{k-2}{2}} P_2 \\ -\lambda^{\frac{k-2}{2}} I_n \\ \lambda^{\frac{k-4}{2}} A_{k-4} - \lambda^{\frac{k-4}{2}} P_4 \\ \vdots \\ -\lambda I_n \\ A_0 \end{bmatrix}.$$

It is clear that the first claim follows for the block entries of $\mathcal{H}_P^S(\lambda)\Delta_1(\lambda)$ in odd positions. In order to prove that the claim also follows for the block entries in even positions, we notice that, for $i = 0, 2, \ldots, k-2$,

$$\begin{aligned} &\lambda[\lambda^{\frac{k-i}{2}} P_i + \lambda^{\frac{k-i-2}{2}} A_{k-i-1}] + \lambda^{\frac{k-i-2}{2}} A_{k-i-2} - \lambda^{\frac{k-i-2}{2}} P_{i+2}] = \\ &\lambda^{\frac{k-i}{2}} [\lambda P_i + A_{k-i-1}] + \lambda^{\frac{k-i-2}{2}} [A_{k-i-2} - P_{i+2}] = \\ &\lambda^{\frac{k-i}{2}} P_{i+1} + \lambda^{\frac{k-i-2}{2}} [A_{k-i-2} - P_{i+2}] = \\ &\lambda^{\frac{k-i-2}{2}} [\lambda P_{i+1} + A_{k-i-2} - P_{i+2}] = 0, \end{aligned}$$

where the second and fourth equalities follow from $\lambda P_i + A_{k-i-1} = P_{i+1}$, $i = 0 : k-1$. Recall that $A_k = P_0$. Moreover, for the kth block-entry of $\mathcal{H}_P^S\Delta_1$ we have

$$\lambda(\lambda P_{k-2} + A_1) + A_0 = \lambda P_{k-1} + A_0 = P_k = P(\lambda).$$

which proves the first claim. The second claims can be proven similarly.

Theorem 5.3 provides explicit formulas for the eigenvectors of the pencil $\mathcal{H}_P^S(\lambda)$ in terms of the eigenvectors of the matrix polynomial $P(\lambda)$. Its proof is similar to the proof of [4, Theorem 4.1], so we omit it.

Theorem 5.3 *Let $P(\lambda)$ be a regular matrix polynomial of even degree k as in* (1.1) *whose leading coefficient A_k is nonsingular, and let S be a nonsingular $n \times n$ matrix. Let λ_0 be a finite eigenvalue of $P(\lambda)$. Then, v is a right eigenvector of $\mathcal{H}_P^S(\lambda)$ with eigenvalue λ_0 if and only if $v = \Delta_1(\lambda_0)x$, for some right eigenvector x of $P(\lambda)$ with eigenvalue λ_0.*

5.2 The case when A_k is singular but A_0 is not

In Section 5.1, we assumed in all our discussions that the leading coefficient A_k of $P(\lambda)$ was nonsingular. In this section we consider the case in which A_0 is nonsingular. The case when both A_k and A_0 are singular is an open question.

As an alternative to the linearization $\mathcal{H}_P^S(\lambda)$ in (5.6), we can consider the pencil $\mathcal{G}_P^S(\lambda) := \operatorname{rev} \mathcal{H}_{\operatorname{rev} P}^S(\lambda)$, which takes the form

$$\mathcal{G}_P^S(\lambda) := \begin{bmatrix} -\lambda S^* A_0^{-1} S & S^* & & & & \\ S & \lambda A_2 + A_1 & -\lambda I_n & & & \\ & -\lambda I_n & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & I_n \\ & & & & I_n & \lambda A_k + A_{k-1} \end{bmatrix}. \tag{5.9}$$

Theorem 5.4 *Let $P(\lambda)$ be an even-degree matrix polynomial as in* (1.1) *with nonsingular matrix coefficient A_0, and let S be a nonsingular matrix. Then, the pencil $\mathcal{G}_P^S(\lambda)$ as in* (5.9) *is a strong linearization of $P(\lambda)$.*

Proof Noticing that the pencil $\mathcal{G}_P^S(\lambda)$ when $S = A_0$ is permutationally equivalent to the pencil $\mathcal{E}_1^P(\lambda)$ defined in [5, Section 4.3], the proof is identical to that of Theorem 5.1. □

The following lemma is easy to prove. Note that the claim follows from the definition of $\mathcal{G}_P^S(\lambda)$ and the definition of reversal of a matrix polynomial.

Lemma 5.1 *Let $P(\lambda)$ be an even degree regular matrix polynomial as in* (1.1) *with nonsingular A_0. Let S be a nonsingular matrix. If λ_0 is a nonzero eigenvalue of $P(\lambda)$, then the vectors z and w are, respectively, right and left eigenvectors of $\mathcal{G}_P^S(\lambda)$ associated with λ_0 if and only if z and w are, respectively, right and left eigenvectors of $\mathcal{H}_{\operatorname{rev} P}^S(\lambda)$ associated with $\frac{1}{\lambda_0}$.*

6 Eigenvalue condition numbers ratio bounds

In this section, we compare the eigenvalue condition numbers of a matrix polynomial $P(\lambda)$ and its linearization $\mathcal{H}_P^S(\lambda)$ for different nonsingular matrices S. The comparison is done by providing upper and lower bounds on the ratios of the two condition numbers. In all our results, we assume that the leading coefficient A_k of

$P(\lambda)$ is nonsingular as this condition guarantees that $\mathcal{H}_P^S(\lambda)$ is a strong linearization of $P(\lambda)$. We also assume that $P(\lambda)$ is symmetric/Hermitian, although many of our results don't require this assumption, for simplicity.

Theorem 6.1 will allow us to address the case when A_k is singular but A_0 is nonsingular, by translating all the results obtained for $\mathcal{H}_P^S(\lambda)$ to $\mathcal{G}_P^S(\lambda)$ just by replacing $P(\lambda)$ by $\mathrm{rev}P(\lambda)$ and λ_0 by $\frac{1}{\lambda_0}$.

Theorem 6.1 [4, Lemmas 2.1 and 2.2] *Let $P(\lambda)$ be an even degree regular matrix polynomial as in* (1.1) *and let λ_0 be a finite, nonzero, and simple eigenvalue of $P(\lambda)$. Assume that A_0 is nonsingular. Then,*

$$\kappa_{\mathrm{ra}}(\lambda_0;\mathcal{G}_P^S)=\kappa_{\mathrm{ra}}\left(\frac{1}{\lambda_0};\mathcal{H}_{\mathrm{rev}P}^S\right) \quad \text{and} \quad \kappa_{\mathrm{rr}}(\lambda_0;\mathcal{G}_P^S)=\kappa_{\mathrm{rr}}\left(\frac{1}{\lambda_0};\mathcal{H}_{\mathrm{rev}P}^S\right).$$

Moreover, if $(\widetilde{z},\widetilde{\lambda}_0)$ is an approximate right eigenpair of $\mathcal{G}_P^S(\lambda)$, then $(\widetilde{z},\frac{1}{\widetilde{\lambda}_0})$ is an approximate eigenpair of $\mathcal{H}_{\mathrm{rev}P}^S(\lambda)$ and

$$\eta_{\mathrm{ra}}(\widetilde{z},\widetilde{\lambda}_0;\mathcal{G}_P^S)=\eta_{\mathrm{ra}}(\widetilde{z},\frac{1}{\widetilde{\lambda}_0};\mathcal{H}_{\mathrm{rev}P}^S) \quad \text{and} \quad \eta_{\mathrm{rr}}(\widetilde{z},\widetilde{\lambda}_0;\mathcal{G}_P^S)=\eta_{\mathrm{rr}}(\widetilde{z},\frac{1}{\widetilde{\lambda}_0};\mathcal{H}_{\mathrm{rev}P}^S).$$

In what follows we will use the following notation

$$\zeta := \max\{1,\|S\|_2,\|S^*A_k^{-1}S\|_2\}, \tag{6.1}$$
$$\mu_a := \zeta\ \max\{1,\|S^{-1}A_k\|_2^2\}, \tag{6.2}$$
$$\mu_b := \zeta\ \min\{1,\|A_k^{-1}S\|_2^{-2}\}, \quad \text{and} \tag{6.3}$$
$$\mu_c := \zeta\ \max\{1,\max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\}. \tag{6.4}$$

Next we include the main result of this section. Its proof will be presented in Section 8 since it is very involved.

Theorem 6.2 (Relative-absolute conditioning bounds) *Let $P(\lambda)$ be a regular $n\times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and $\max_{i=0:k}\{\|A_i\|_2\}=1$. Assume that λ_0 is a simple, finite, and nonzero eigenvalue of $P(\lambda)$. Let S be an $n\times n$ nonsingular matrix and let $\mathcal{H}_P^S(\lambda)$ be as in* (5.6).

(i) If $|\lambda_0|\le 1$, then

$$\max\left\{\mu_b,\frac{\zeta}{2}\right\}\le\frac{\kappa_{\mathrm{ra}}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0;P)}\le k^3\mu_a;$$

Moreover, if $|\lambda_0|$ is close to 0, then

$$\max\left\{\mu_b,\frac{\zeta}{2}\right\}\le\frac{\kappa_{\mathrm{ra}}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0;P)}\lessapprox k\mu_a.$$

(ii) If $|\lambda_0|>1$, then

$$\max\left\{\left(1+\frac{|\lambda_0|}{k+1}\right)\mu_b,\frac{\zeta}{2|\lambda_0|}\right\}\le\frac{\kappa_{\mathrm{ra}}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0;P)}\le 2\min\left\{k^3|\lambda_0|\mu_a,\frac{k+k^3}{|\lambda_0|}\mu_c\right\};$$

where the constants ζ, μ_a, μ_b and μ_c have been defined in (6.1)–(6.4).

We note that, for $\mathcal{H}_P^S(\lambda)$ to be a "good" linearization of $P(\lambda)$ in terms of conditioning, we would like the upper bounds on the ratios of condition numbers provided in Theorem 6.2 to be "small". This will happen if μ_a and μ_c are "small". Notice that these constants depend on our selection of the matrix S. Next we consider the particular cases $S = A_k$, $S = I_n$ and $S = S_{MX}$, where S_{MX} is the matrix from the Mehrmann-Xu deflation process discussed in Remark 5.2. As in Theorem 6.2, the factor k^3 in the upper bounds for the ratios of condition numbers can be replaced by k when $|\lambda_0|$ is close to zero.

Theorem 6.3 *Let $P(\lambda)$ be a regular $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and* $\max_{i=0:k}\{\|A_i\|_2\} = 1$. *Assume that λ_0 is a simple, finite, and nonzero eigenvalue of $P(\lambda)$.*

(i) If $S = A_k$, then

$$\left.\begin{array}{ll} 1 & \text{if } |\lambda_0| \leq 1 \\ 1 + \frac{|\lambda_0|}{k+1} & \text{if } |\lambda_0| > 1 \end{array}\right\} \leq \frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} \leq \begin{cases} k^3 & \text{if } |\lambda_0| \leq 1 \\ 2k^3|\lambda_0| & \text{if } |\lambda_0| > 1. \end{cases}$$

(ii) If $S = I_n$, then

$$\left.\begin{array}{ll} \frac{\max\{1, \|A_k^{-1}\|_2\}}{2} & \text{if } |\lambda_0| \leq 1 \\ \frac{\max\{1, \|A_k^{-1}\|_2\}}{2|\lambda_0|} & \text{if } |\lambda_0| > 1 \end{array}\right\} \leq \frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} \leq \begin{cases} k^3 \max\{1, \|A_k^{-1}\|_2\} & \text{if } |\lambda_0| \leq 1 \\ 4k^3 \frac{\max\{1, \|A_k^{-1}\|_2\}}{|\lambda_0|} & \text{if } |\lambda_0| > 1. \end{cases}$$

(iii) If $S = S_{MX}$, then

$$\left.\begin{array}{ll} 1 & \text{if } |\lambda_0| \leq 1 \\ 1 + \frac{|\lambda_0|}{k+1} & \text{if } |\lambda_0| > 1 \end{array}\right\} \leq \frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} \leq \begin{cases} 2k^3 & \text{if } |\lambda_0| \leq 1 \\ 4k^3|\lambda_0| & \text{if } |\lambda_0| > 1. \end{cases}$$

Proof Observe that $\mu_a = \mu_b = 1$ when $S = A_k$ since $\|A_k\|_2 \leq 1$, and $\zeta = \mu_a = \mu_c = \max\{1, \|A_k^{-1}\|_2\}$ when $S = I_n$. Then, when $S = A_k$ or $S = I_n$, the lower and upper bounds follow immediately from Theorem 6.2.

Next, we obtain the bounds when $S = S_{MX}$. Recall that this matrix is obtained from an orthonormal basis for the nullspace of $M = \begin{bmatrix} A_k & -I_n \end{bmatrix}$. Let $V = \begin{bmatrix} T \\ S_{MX} \end{bmatrix}$ be one such basis. From $MV = 0$, we obtain $A_k T = S_{MX}$. Since A_k is nonsingular, we have $A_k^{-1} S_{MX} = T$. Hence,

$$\mu_b = \max\{1, \|S_{MX}\|_2, \|S_{MX}^* T\|_2\} \min\{1, \|T\|_2^{-2}\}.$$

Since V has orthonormal columns, we have $\|T\|_2 \leq 1$, $\|S_{MX}\|_2 \leq 1$, and $\|S_{MX}^* T\|_2 \leq 1$. This readily implies $\mu_b = 1$. Then, observe that

$$W = \begin{bmatrix} I_n \\ A_k \end{bmatrix} (I_n + A_k^2)^{-1/2}, \tag{6.5}$$

where $(I_n + A_k^2)^{1/2}$ denotes the unique positive definite square root of $I_n + A_k^2$, is another orthonormal basis for the nullspace of $M = \begin{bmatrix} A_k & -I_n \end{bmatrix}$. Thus, $V = WU$, for some $n \times n$ unitary matrix U. Hence, $T = (I_n + A_k^2)^{-1/2} U$ and, so, $T^{-1} = U^*(I_n + A_k^2)^{1/2}$. Finally, notice

$$\|T^{-1}\|_2 = \|(I_n + A_k^2)\|_2^{1/2} \leq \sqrt{2} \max\{1, \|A_k\|_2\} = \sqrt{2},$$

which implies

$$\mu_a = \max\{1, \|S_{MX}\|_2, \|S_{MX}^* T\|_2\} \max\{1, \|T^{-1}\|_2^2\} \leq 2.$$

Conclusively, if $S = S_{MX}$, then $\mu_a \leq 2$ and $\mu_b = 1$, and, thus, the bounds readily follow from Theorem 6.2. □

Remark 6.1 From the previous theorem, we conclude that, from the relative-absolute condition number point of view, $\mathcal{H}_P^{A_k}$ and $\mathcal{H}_P^{S_{MX}}$ are comparable and have an optimal behavior for matrix polynomials $P(\lambda)$ with "small" degree and for eigenvalues λ_0 with "small" modulus.

The optimality in this context means that the sensitivity of λ_0 as an eigenvalue of P is approximately the same as the sensitivity of λ_0 as an eigenvalue of $\mathcal{H}_P^S$.

Note that the lower bounds for these two linearizations show that if $|\lambda_0| \gg 1$, then neither of the two linearizations will be a good choice.

If A_k is a matrix whose absolute condition number $\|A_k^{-1}\|_2$ is "small", then $\mathcal{H}_P^{I_n}$ has optimal condition number regardless of the modulus of λ_0 for moderate k. Nonetheless, every eigenvalue of $P(\lambda)$ satisfies

$$|\lambda_0| \leq 1 + \|A_k^{-1}\|_2 \sum_{i=0}^{k-1} \|A_i\|_2 \leq 1 + k\|A_k^{-1}\|_2,$$

see [12, Lemma 2.2]. Hence, if $\|A_k^{-1}\|_2$ is moderate, then $P(\lambda)$ does not have eigenvalues with large modulus and, so, $\mathcal{H}_P^{A_k}$ and $\mathcal{H}_P^{S_{MX}}$ also have optimal condition numbers for all eigenvalues of $P(\lambda)$.

Now, using Theorem 6.1, we can also conclude that, when A_0 is nonsingular, $\mathcal{G}_P^{A_k}$ and $\mathcal{G}_P^{S_{MX}}$ are comparable and have an optimal behavior for matrix polynomials $P(\lambda)$ with "small" degree and for eigenvalues with "large" modulus. Thus, if $P(\lambda)$ is a matrix polynomial with A_k and A_0 nonsingular, in order to compute all the eigenvalues accurately, the use of two linearizations ($\mathcal{H}_P^{A_k}$ and $\mathcal{G}_P^{A_k}$, for example) would be necessary. This strategy is similar to the one used in the literature with the linearizations $D_1(\lambda; P)$ and $D_k(\lambda; P)$ given in (4.1) and (4.2), respectively. We note that the linearizations DH and DG used in the numerical experiments in Section 4 are precisely the linearizations $\mathcal{H}_P^{A_k}$ and $\mathcal{G}_P^{A_k}$, respectively, discussed here.

Remark 6.2 So far we have shown that the combined used of $\mathcal{H}_P^{A_k}$ and $\mathcal{G}_P^{A_k}$ ensures optimal eigenvalue conditioning for eigenvalues of any modulus. But the same holds for $D_k(\lambda; P)$ and $D_1(\lambda; P)$. So, what is the advantage of using these two linearizations compared to $D_k(\lambda; P)$ and $D_1(\lambda; P)$? In Remark 3.2 we argued that one of the possible reasons why the eigenvectors of $D_k(\lambda; P)$ and $D_1(\lambda; P)$ are so sensitive to changes in the coefficients of these two pencils is the fact that both linearizations tend to have very ill-conditioned eigenvalues. In Remark 4.1 we showed that this is due to the fact that the condition number of the eigenvalues λ_0 with large (resp. small) modulus of $D_k(\lambda; P)$ (resp. $D_1(\lambda; P)$) is bounded below by the product of the corresponding condition number when λ_0 is seen as an eigenvalue of $P(\lambda)$ and $|\lambda_0|^{k-1}$ (resp. $|\lambda_0|^{1-k}$). Theorems 6.3 and 6.1 show, however, that the condition number of the eigenvalues λ_0 with large (resp. small) modulus of $\mathcal{H}_P^{A_k}$ (resp. $\mathcal{G}_P^{A_k}$) is bounded above by a multiple of the product of the corresponding condition number when λ_0 is seen as an eigenvalue of P and $|\lambda_0|$

(resp. $|\lambda_0|^{-1}$). Thus, if the eigenvalue is well-conditioned in $P(\lambda)$, its condition number in the linearization is not much worse as long as $|\lambda_0|$ is moderate. This might be the reason for the good behavior of the backward error ratios when $\mathcal{H}_P^{A_k}$ and $\mathcal{G}_P^{A_k}$ were used in the numerical experiments showed in Section 4.

The next theorem provides bounds for the relative-relative condition numbers ratio. Its proof will also be presented in Section 8. We note that, when finding the bounds presented in this theorem, our main goal was to obtain bounds as sharp as possible. For less tight but easier to interpret bounds, see Remark 6.3.

Theorem 6.4 (Relative-relative conditioning bounds) *Let $P(\lambda)$ be a regular $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and $\max_{i=0:k}\{\|A_i\|_2\} = 1$. Assume that λ_0 is a simple, finite, and nonzero eigenvalue of $P(\lambda)$. Let S be an $n \times n$ nonsingular matrix and let $\mathcal{H}_P^S(\lambda)$ be as in* (5.2).

(i) If $|\lambda_0| \le 1$, then

$$\frac{\max\{1, \|S^* A_k^{-1} S\|_2\}}{(k+1)\max\limits_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} \le \frac{\kappa_{\mathrm{rr}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{rr}}(\lambda_0; P)} \le \frac{2k^3 \mu_a}{\max\limits_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}};$$

(ii) If $|\lambda_0| > 1$, then

$$\frac{|\lambda_0|^k \max\{1, \|S\|_2\} \max\left\{|\lambda_0| \, \|A_k^{-1} S\|_2^{-2}, \frac{1}{|\lambda_0|}\right\}}{(k+1)\max\limits_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} \le$$

$$\frac{\kappa_{\mathrm{rr}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{rr}}(\lambda_0; P)} \le \frac{2|\lambda_0|^k \min\left\{k^3 |\lambda_0| \mu_a, \frac{(k+k^3)\mu_c}{|\lambda_0|}\right\}}{\max\limits_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}},$$

where the constants ζ, μ_a, μ_b and μ_c are as in (6.1)–(6.4).

As with Theorem 6.2, the upper bounds presented in the previous theorem depend on μ_a and μ_c. However, in this case, the bounds also depend on the norm of each monomial of the polynomial $P(\lambda)$. Theorem 6.5 interprets these bounds in the cases when $S = A_k$, $S = I_n$, and $S = S_{MX}$.

Theorem 6.5 *Let $P(\lambda)$ be a regular $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and $\max_{i=0:k}\{\|A_i\|_2\} = 1$. Assume that λ_0 is a simple, finite, and nonzero eigenvalue of $P(\lambda)$.*

(i) If $S = A_k$, then

$$\left.\begin{array}{ll} \frac{1}{(k+1)\max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} & \text{if } |\lambda_0| \le 1 \\ \frac{|\lambda_0|^{k+1}}{(k+1)\max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} & \text{if } |\lambda_0| > 1 \end{array}\right\} \le$$

$$\frac{\kappa_{\mathrm{rr}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{rr}}(\lambda_0; P)} \le \begin{cases} \frac{2k^3}{\max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} & \text{if } |\lambda_0| \le 1 \\ \frac{2k^3 |\lambda_0|^{k+1}}{\max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}} & \text{if } |\lambda_0| > 1. \end{cases}$$

(ii) *If* $S = I_n$, *then*

$$\left.\begin{array}{ll}\frac{\max\{1,\|A_k^{-1}\|_2\}}{(k+1)\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|\le 1\\ \frac{|\lambda_0|^{k-1}}{(k+1)\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|>1\end{array}\right\}\le$$

$$\frac{\kappa_{\rm rr}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\rm rr}(\lambda_0;P)}\le\left\{\begin{array}{ll}\frac{2k^3\max\{1,\|A_k^{-1}\|_2\}}{\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|\le 1\\ \frac{2k^3|\lambda_0|^{k-1}\max\{1,\|A_k^{-1}\|_2\}}{\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|>1.\end{array}\right.$$

(iii) *If* $S = S_{MX}$, *then*

$$\left.\begin{array}{ll}\frac{1}{(k+1)\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|\le 1\\ \frac{|\lambda_0|^{k+1}}{(k+1)\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|>1\end{array}\right\}\le$$

$$\frac{\kappa_{\rm rr}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\rm rr}(\lambda_0;P)}\le\left\{\begin{array}{ll}\frac{4k^3}{\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|\le 1\\ \frac{4k^3|\lambda_0|^{k+1}}{\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}} & \textit{if } |\lambda_0|>1.\end{array}\right.$$

Proof Recall from the proof of Theorem 6.3 that $\mu_a = \mu_b = 1$ when $S = A_k$, $\mu_a = \mu_c = \max\{1, \|A_k^{-1}\|_2\}$ when $S = I_n$, and $\mu_a \le 2$, $\|A_k^{-1}S\|_2 \le 1$ and $\|S^*A_k^{-1}S\|_2 \le 1$ when $S = S_{MX}$. All the bounds, then, readily follow from Theorem 6.4.

Remark 6.3 In order to give an easy interpretation of the upper bounds obtained in Theorem 6.5, we use the following fact

$$\max\{\|A_0\|_2, |\lambda_0|^k\|A_k\|_2\}\le\max_{i=0:k}\{|\lambda_0|^i\|A_i\|_2\}\le\left\{\begin{array}{ll}1 & \text{if } |\lambda_0|\le 1\\ |\lambda_0|^k & \text{if } |\lambda_0|>1.\end{array}\right.$$

Then, from Theorem 6.5, we get the following simpler bounds for the relative-relative condition numbers ratio.
If $S = A_k$, then

$$\left.\begin{array}{ll}\frac{1}{(k+1)} & \text{if } |\lambda_0|\le 1\\ \frac{|\lambda_0|}{(k+1)} & \text{if } |\lambda_0|>1\end{array}\right\}\le\frac{\kappa_{\rm rr}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\rm rr}(\lambda_0;P)}\le\left\{\begin{array}{ll}\frac{2k^3}{\|A_0\|_2} & \text{if } |\lambda_0|\le 1\\ \frac{2k^3|\lambda_0|}{\|A_k\|_2} & \text{if } |\lambda_0|>1.\end{array}\right.$$

If $S = I_n$, then

$$\left.\begin{array}{ll}\frac{\max\{1,\|A_k^{-1}\|_2\}}{(k+1)} & \text{if } |\lambda_0|\le 1\\ \frac{1}{(k+1)|\lambda_0|} & \text{if } |\lambda_0|>1\end{array}\right\}\le\frac{\kappa_{\rm rr}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\rm rr}(\lambda_0;P)}\le\left\{\begin{array}{ll}\frac{2k^3\max\{1,\|A_k^{-1}\|_2\}}{\|A_0\|_2} & \text{if } |\lambda_0|\le 1\\ \frac{2k^3\max\{1,\|A_k^{-1}\|_2\}}{|\lambda_0|\|A_k\|_2} & \text{if } |\lambda_0|>1.\end{array}\right.$$

If $S = S_{MX}$, then

$$\left.\begin{array}{ll}\frac{1}{(k+1)} & \text{if } |\lambda_0|\le 1\\ \frac{|\lambda_0|}{(k+1)} & \text{if } |\lambda_0|>1\end{array}\right\}\le\frac{\kappa_{\rm rr}(\lambda_0;\mathcal{H}_P^S)}{\kappa_{\rm rr}(\lambda_0;P)}\le\left\{\begin{array}{ll}\frac{4k^3}{\|A_0\|_2} & \text{if } |\lambda_0|\le 1\\ \frac{4k^3|\lambda_0|}{\|A_k\|_2} & \text{if } |\lambda_0|>1.\end{array}\right.$$

These bounds are less tight than those in Theorem 6.5 but easier to interpret. From these bounds we conclude that $\mathcal{H}_P^{A_k}$ and $\mathcal{H}_P^{S_{MX}}$ have also a comparable behavior in terms of relative-relative conditioning. The behavior is optimal if $|\lambda_0| \le 1$, k is moderate and $\|A_0\|_2 \approx 1$ (recall that we have scaled $P(\lambda)$ so that $\max_{i=0:k}\{\|A_i\|_2\} = 1$); or if $|\lambda_0| > 1$ is moderate, k is moderate, and $\|A_k\|_2 \approx 1$. The lower bounds for these two linearizations show that if $|\lambda_0| \gg 1$, then neither of the two linearizations will be a good choice. In this case, $\mathcal{H}_P^{I_n}$ could potentially be a good choice if both $\|A_k\|_2$ and $\|A_k^{-1}\|_2$ have approximately the same norm. But in this case, as we argued in Remark 6.1, $P(\lambda)$ does not have eigenvalues with large modulus and, thus, $\mathcal{H}_P^{A_k}$, $\mathcal{H}_P^{S_{MX}}$ and $\mathcal{H}_P^{I_n}$ are all optimally conditioned.

A comment regarding $\mathcal{G}_P^S$ similar to that in Remark 6.1 is appropriate here as well.

7 Backward errors ratio bounds

In this section, we compare the backward errors of approximate eigenpairs of a matrix polynomial $P(\lambda)$ and its linearization $\mathcal{H}_P^S(\lambda)$ for different nonsingular matrices S. The comparison is done by providing upper on the ratio of the two backward errors. In all our results, we assume that the leading coefficient A_k of $P(\lambda)$ is nonsingular as this condition guarantees that $\mathcal{H}_P^S(\lambda)$ is a strong linearization of $P(\lambda)$.

Theorem 6.1 allows us to address the case when A_k is singular but A_0 is nonsingular, by translating all the results obtained for $\mathcal{H}_P^S(\lambda)$ to $\mathcal{G}_P^S(\lambda)$ just by replacing $P(\lambda)$ by $\text{rev}P(\lambda)$ and λ_0 by $\frac{1}{\lambda_0}$.

The proof of Theorem 7.1 is omitted because it is very involved but similar to the proof of Theorems 6.2 and 6.4. It is also similar to the proof of Theorem 5.2 in [4]. The block-vector $\Delta_2(\lambda)$ defined in Theorem 5.2 would be necessary in this case.

Theorem 7.1 (Backward error bounds) *Let $P(\lambda)$ be a regular $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and $\max_{i=0:k}\{\|A_i\|_2\} = 1$. Let S be an $n \times n$ nonsingular matrix and let $\mathcal{H}_P^S(\lambda)$ be as in* (5.6). *Let $(\widetilde{z}, \widetilde{\lambda}_0)$ be an approximate right eigenpair of $\mathcal{H}_P^S(\lambda)$, and define the vector*

$$\widetilde{x} := \begin{cases} (e_k \otimes I_n)\widetilde{z} & \text{if } |\widetilde{\lambda}_0| \leq 1, \quad \text{and} \\ (e_2 \otimes I_n)\widetilde{z} & \text{if } |\widetilde{\lambda}_0| > 1. \end{cases} \tag{7.1}$$

Then,

$$\frac{\eta_{\mathrm{ra}}(\widetilde{x}, \widetilde{\lambda}_0; P)}{\eta_{\mathrm{ra}}(\widetilde{z}, \widetilde{\lambda}_0; \mathcal{H}_P^S)} \leq 4k^{3/2}\, \zeta\, \max\{1, \|S^{-1}A_k\|_2\} \frac{\|\widetilde{z}\|_2}{\|\widetilde{x}\|_2},$$

and

$$\frac{\eta_{\mathrm{rr}}(\widetilde{x}, \widetilde{\lambda}_0; P)}{\eta_{\mathrm{rr}}(\widetilde{z}, \widetilde{\lambda}_0; \mathcal{H}_P^S)} \leq 4k^{3/2}\, \zeta\, \max\{1, \|S^{-1}A_k\|_2\} \frac{\max\{1, |\widetilde{\lambda}_0|^k\}}{\max_{i=0:k}\{|\widetilde{\lambda}_0|^i \|A_i\|_2\}} \frac{\|\widetilde{z}\|_2}{\|\widetilde{x}\|_2},$$

where ζ is as in (6.1).

The following result follows from Theorem 7.1 and the fact that $\|S_{MX}\|_2$, $\|S_{MX}^* A_k^{-1} S_{MX}\|_2 \leq 1$ and $\|S_{MX}^{-1} A_k\|_2 = \sqrt{2}$ as shown in the proof of Theorem 6.3.

Corollary 7.1 *Let $P(\lambda)$ be a regular $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with nonsingular A_k and* $\max_{i=0:k}\{\|A_i\|_2\} = 1$.

Let $(\widetilde{z}, \widetilde{\lambda}_0)$ *be an approximate right eigenpair of* $\mathcal{H}_P^S(\lambda)$*, and let* $\widetilde{x}$ *be as in* (7.1)*. Then*

$$\frac{\eta_{\mathrm{ra}}(\widetilde{x}, \widetilde{\lambda}_0; P)}{\eta_{\mathrm{ra}}(\widetilde{z}, \widetilde{\lambda}_0; \mathcal{H}_P^S)} \leq 4k^{3/2} \frac{\|\widetilde{z}\|_2}{\|\widetilde{x}\|_2} \times \begin{cases} 2^{1/2} & \text{if } S = S_{MX}, \\ 1 & \text{if } S = A_k, \\ \max\{1, \|A_k^{-1}\|_2\} & \text{if } S = I_n. \end{cases}$$

and

$$\frac{\eta_{\mathrm{rr}}(\widetilde{x}, \widetilde{\lambda}_0; P)}{\eta_{\mathrm{rr}}(\widetilde{z}, \widetilde{\lambda}_0; \mathcal{H}_P^S)} \leq 4k^{3/2} \frac{\max\{1, |\widetilde{\lambda}_0|^k\}}{\max_{i=0:k}\{|\widetilde{\lambda}_0|^i \|A_i\|_2\}} \frac{\|\widetilde{z}\|_2}{\|\widetilde{x}\|_2} \times \begin{cases} 2^{1/2} & \text{if } S = S_{MX}, \\ 1 & \text{if } S = A_k, \\ \max\{1, \|A_k^{-1}\|_2\} & \text{if } S = I_n. \end{cases}$$

Remark 7.1 From the previous theorem, we conclude that, from the relative-absolute backward error point of view, $\mathcal{H}_P^{A_k}$ and $\mathcal{H}_P^{S_{MX}}$ are comparable and have an optimal behavior for matrix polynomials $P(\lambda)$ with "small" degree and for eigenvalues λ_0 with "small" modulus, as happened with the eigenvalue condition number.

The optimality in this context means that the backward error of approximate eigenpairs $(\widetilde{\lambda}_0, \widetilde{x})$ of P is not much worse than the backward error of approximate eigenpairs $(\widetilde{\lambda}_0, \widetilde{z})$ of $\mathcal{H}_P^S$ when $\widetilde{x}$ is recovered from $\widetilde{z}$ as explained in Corollary 7.1.

Moreover, if λ_0 (with $|\lambda_0| \leq 1$) is an exact eigenvalue of $\mathcal{H}_P^S(\lambda)$ with corresponding right eigenvector z, then according to Theorem 5.3, $z = \Delta_1(\lambda_0)x$ for some eigenvector x of $P(\lambda)$. Because of the structure of $\Delta_1(\lambda)$, we have that $x = (e_k \otimes I_n)z$. This implies, as we will show in (8.16), that

$$\frac{\|z\|_2}{\|x\|_2} = \frac{\|\Delta_1(\lambda_0)x\|_2}{\|x\|_2} \leq \left(\frac{k^3}{2}\right)^{1/2} \max\{1, \|S^{-1}A_k\|_2\}.$$

Thus, for $S \in \{A_k, I_n, S_{MX}\}$ and $|\lambda_0| \leq 1$, we have

$$\frac{\|z\|_2}{\|x\|_2} \leq k^{3/2}.$$

So, if the computed eigenvector $\widetilde{z}$ has the same structure as the exact eigenvector z, we know that the upper bound for $\frac{\eta_{\mathrm{ra}}(\widetilde{x}, \widetilde{\lambda}_0; P)}{\eta_{\mathrm{ra}}(\widetilde{z}, \widetilde{\lambda}_0; \mathcal{H}_P^S)}$ is moderate for moderate values of k, for eigenvalues $|\widetilde{\lambda}_0| \leq 1$ and for $S \in \{A_k, S_{MX}\}$. Although we cannot guarantee that this is the case, in all our numerical experiments this seems to be the case, in stark contrast with what happened with $D_1(\lambda; P)$ and $D_k(\lambda; P)$. Recall our comments in Remark 6.2 for a possible explanation.

As with the eigenvalue condition number, in order to guarantee small backward errors for $|\widetilde{\lambda}_0| > 1$, it is necessary to assume that A_0 is also nonsingular and use the linearization $\mathcal{G}_P^{A_k}$ or $\mathcal{G}_P^{S_{MX}}$.

8 Proof of the eigenvalue condition bounds

The next lemma is the key result that leads to the proofs of Theorems 6.2 and 6.4.

Lemma 8.1 *Let $P(\lambda)$ be a regular $n\times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1). *Assume that λ_0 is a simple, finite, and nonzero eigenvalue of $P(\lambda)$ with corresponding right eigenvector x. Let S be an $n \times n$ nonsingular matrix and let $\mathcal{H}_P^S(\lambda) =: \lambda\mathcal{H}_1 - \mathcal{H}_0$. Then,*

$$\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S) = \frac{(|\lambda_0|+1)\max\{\|\mathcal{H}_1\|_2, \|\mathcal{H}_0\|_2\}\|\Delta_1(\lambda_0)x\|_2^2}{|\lambda_0|\cdot|x^*P'(\lambda_0)x|} \quad \textit{and}$$
$$\kappa_{\mathrm{rr}}(\lambda_0; \mathcal{H}_P^S) = \frac{(|\lambda_0|\|\mathcal{H}_1\|_2 + \|\mathcal{H}_0\|_2)\|\Delta_1(\lambda_0)x\|_2^2}{|\lambda_0|\cdot|x^*P'(\lambda_0)x|},$$

where $\Delta_1(\lambda)$ is as in (5.8).

Proof By Theorem 5.3, the vector $\Delta_1(\lambda_0)x$ is a right eigenvector of $\mathcal{H}_P^S(\lambda)$ with eigenvalue λ_0. Since $P(\lambda)$ is symmetric, so is $\mathcal{H}_P^S(\lambda)$. Hence, $\Delta_1(\lambda_0)x$ is also a left eigenvector of $\mathcal{H}_P^S(\lambda)$ with eigenvalue λ_0. By Theorem 5.2, we have the following right-sided factorization

$$\mathcal{H}_P^S(\lambda)\Delta_1(\lambda) = e_k \otimes P(\lambda).$$

Differentiating this expression with respect to λ, we get

$$\mathcal{H}_P^S(\lambda)'\Delta_1(\lambda) + \mathcal{H}_P^S(\lambda)\Delta_1'(\lambda) = e_k \otimes P'(\lambda).$$

Now, we evaluate this expression at λ_0 and multiply it by x on the right and by $(\Delta_1(\lambda_0)x)^*$ on the left. We get

$$(\Delta_1(\lambda_0)x)^*\mathcal{H}_P^S(\lambda_0)'\Delta_1(\lambda_0)x = (\Delta_1(\lambda_0)x)^*(e_k \otimes P'(\lambda_0))x = x^*P'(\lambda_0)x,$$

and the results readily follow from the eigenvalue condition number formulas in Theorem 3.1.

Next we bound the norm of the matrix coefficients of the linearization $\mathcal{H}_P^S(\lambda)$ in terms of the norms of the matrix coefficients of the matrix polynomial $P(\lambda)$ and the marix S.

Lemma 8.2 *Let $P(\lambda)$ be an $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1) *with* $\max\{\|A_i\|_2\} = 1$, *let S be an $n \times n$ nonsingular matrix and let $\mathcal{H}_P^S(\lambda) =: \lambda\mathcal{H}_1 - \mathcal{H}_0$. Then,*

$$\|\mathcal{H}_1\|_2 \le 2\max\{1, \|S\|_2\},$$
$$\|\mathcal{H}_0\|_2 \le 2\max\{1, \|S^*A_k^{-1}S\|_2\}.$$

Proof When $k = 2$, we have

$$\mathcal{H}_P^S(\lambda) = \begin{bmatrix} -S^*A_2^{-1}S & \lambda S^* \\ \lambda S & \lambda A_1 + A_0 \end{bmatrix} = \lambda \begin{bmatrix} 0 & S^* \\ S & A_1 \end{bmatrix} - \begin{bmatrix} S^*A_2^{-1}S & 0 \\ 0 & -A_0 \end{bmatrix},$$

and the result thus follows from Proposition 2.1.

Next, assume $k \geq 4$. Let $z = \begin{bmatrix} z_1^T & \cdots & z_k^T \end{bmatrix}^T$ be a nonzero vector partitioned into k blocks of size $n \times 1$. Then, defining $z_0 := 0$, we have

$$\|\mathcal{H}_1 z\|_2^2 = \|S^* z_2\|_2^2 + \|Sz_1 + A_{k-1}z_2\|_2^2 + \sum_{i=2}^{\frac{k}{2}} \|z_{2i}\|_2^2 + \sum_{i=1}^{\frac{k-2}{2}} \|z_{2i+1} + A_{k-2i-1}z_{2i+2}\|_2^2.$$

Using the triangle inequality, we get

$$\|\mathcal{H}_1 z\|_2^2 \leq \|S\|_2^2\|z_2\|_2^2 + (\|S\|_2\|z_1\|_2 + \|A_{k-1}\|_2\|z_2\|_2)^2 + \sum_{i=2}^{\frac{k}{2}} \|z_{2i}\|_2^2 + \sum_{i=1}^{\frac{k-2}{2}} (\|z_{2i+1}\|_2 + \|A_{k-2i-1}\|_2\|z_{2i+2}\|_2)^2.$$

Finally, some simple inequalities and manipulations yield

$$\begin{aligned}
\|\mathcal{H}_1 z\|_2^2 &\leq \max\{1, \|S\|_2^2\} \Bigg[\sum_{i=1}^{\frac{k}{2}} \|z_{2i}\|_2^2 + \\
&\qquad \sum_{i=0}^{\frac{k-2}{2}} \|z_{2i+1}\|_2^2 + \sum_{i=0}^{\frac{k-2}{2}} \|z_{2i+2}\|_2^2 + 2\sum_{i=0}^{\frac{k-2}{2}} \|z_{2i+1}\|_2\|z_{2i+2}\|_2\Bigg] \\
&\leq \max\{1, \|S\|_2^2\} \left[2\sum_{i=1}^{k} \|z_i\|_2^2 + 2\sum_{i=0}^{\frac{k-2}{2}} \max\{\|z_{2i+1}\|_2^2, \|z_{2i+2}\|_2^2\}\right] \\
&\leq 4\max\{1, \|S\|_2^2\}\|z\|_2^2,
\end{aligned}$$

which implies the result for $\mathcal{H}_1$. The result for $\mathcal{H}_0$ can be obtained similarly.

We now need to prove some technical lemmas.

Lemma 8.3 *Let $P(\lambda)$ be an $n \times n$ symmetric/Hermitian matrix polynomial of even degree k as in* (1.1)*, let S be an $n \times n$ nonsingular matrix, and let $\Delta_1(\lambda)$be as in* (5.8)*. Define the following three functions*

$$d_1(\lambda) = \sum_{r=0}^{\frac{k}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-2}{2}} \left[|\lambda|^{2r}(k - 2r + 1) \sum_{j=0}^{k-2r} |\lambda|^{2j}\right], \tag{8.1}$$

$$d_2(\lambda) = \sum_{r=0}^{\frac{k}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k-2}{2}}(2r) \sum_{i=-r}^{r-1} |\lambda|^{2i}, \qquad \textit{and} \tag{8.2}$$

$$d_3(\lambda) = \sum_{r=0}^{\frac{k-2}{2}} |\lambda|^{2r} + \sum_{r=1}^{\frac{k}{2}}(2r) \sum_{i=-r}^{r-1} |\lambda|^{2i}. \tag{8.3}$$

Then, the following inequality holds

$$\|\Delta_1(\lambda)\|_2 \leq \sqrt{d_1(\lambda)}\max\{1, \|S^{-1}A_k\|_2, \max_{i=0:k} \|A_i\|_2\}, \tag{8.4}$$

for any $\lambda \in \mathbb{C}$. Moreover, if λ_0 is a finite eigenvalue of $P(\lambda)$ with corresponding right eigenvector x, then the following inequalities hold

$$\frac{\|\Delta_1(\lambda_0)x\|_2}{\|x\|_2} \leq \min\{\sqrt{d_1(\lambda_0)}, \sqrt{d_2(\lambda_0)}\} \max\{1, \|S^{-1}A_k\|_2, \max_{i=0:k} \|A_i\|_2\}, \tag{8.5}$$

$$\frac{\|\Delta_1(\lambda_0)x\|_2}{\|x\|_2} \leq \sqrt{d_3(\lambda_0)} \max\{1, \max_{i=0:k}\{\|A_i\|_2\}, \max_{i=0:k}\{\|S^{-1}A_i\|_2\}\}. \tag{8.6}$$

Proof Let x be an arbitrary nonzero vector conformable with $\Delta_1(\lambda)$ for multiplication. From (5.8), together with the first inequality in Lemma 2.3, we get

$$\begin{aligned}
\|\Delta_1(\lambda)x\|_2^2 &= |\lambda|^k \|S^{-1}A_k x\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}} |\lambda|^{2r} \|x\|_2^2 + \sum_{r=1}^{\frac{k-2}{2}} |\lambda|^{2r} \|P_{k-2r}(\lambda)x\|_2^2 \\
&\leq |\lambda|^k \|S^{-1}A_k\|_2^2 \|x\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}} |\lambda|^{2r} \|x\|_2^2 + \\
&\qquad \sum_{r=1}^{\frac{k-2}{2}} |\lambda|^{2r} \max_{i=0:k}\{\|A_i\|_2^2\} \left(\sum_{j=0}^{k-2r} |\lambda|^j\right)^2 \|x\|_2^2 \\
&\leq \max\{1, \|S^{-1}A_k\|_2^2, \max_{i=0:k}\{\|A_i\|_2^2\}\} \left[|\lambda|^k + \sum_{r=0}^{\frac{k-2}{2}} |\lambda|^{2r} + \right. \\
&\qquad \left. \sum_{r=1}^{\frac{k-2}{2}} |\lambda|^{2r} \left(\sum_{j=0}^{k-2r} |\lambda|^j\right)^2 \right] \|x\|_2^2.
\end{aligned} \tag{8.7}$$

Using Lemma 2.1, we obtain

$$\begin{aligned}
\|\Delta_1(\lambda)x\|_2^2 \leq \max\{1, \|S^{-1}A_k\|_2^2, \max_{i=0:k}\{\|A_i\|_2^2\}\} &\left[\sum_{r=0}^{\frac{k}{2}} |\lambda|^{2r} + \right. \\
&\left. \sum_{r=1}^{\frac{k-2}{2}} |\lambda|^{2r} (k-2r+1) \sum_{j=0}^{k-2r} |\lambda|^{2j} \right] \|x\|_2^2.
\end{aligned}$$

Thus, we have found the upper bound

$$\frac{\|\Delta_1(\lambda)x\|_2}{\|x\|_2} \leq \sqrt{d_1(\lambda)} \max\{1, \|S^{-1}A_k\|_2, \max_{i=0:k}\{\|A_i\|_2\}\},$$

which does not depend on x. Since $\|\Delta_1(\lambda)\|_2 = \max_{x \neq 0} \frac{\|\Delta_1(\lambda)x\|_2}{\|x\|_2}$, this is also an upper bound for $\|\Delta_1(\lambda)\|_2$, which establishes the inequality (8.4).

Now let us consider an eigenvalue λ_0 of $P(\lambda)$ with corresponding right eigenvector x. The computations above give

$$\frac{\|\Delta_1(\lambda_0)x\|_2}{\|x\|_2} \leq \sqrt{d_1(\lambda_0)} \max\{1, \|S^{-1}A_k\|_2^2, \max_{i=0:k}\{\|A_i\|_2^2\}\}.$$

Furthermore, by (8.7) and Lemma 2.2, we also have

$$\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} \leq |\lambda_0|^k\|S^{-1}A_k\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-2}{2}}|\lambda_0|^{-2r}\|P^{2r-1}(\lambda_0)\|_2^2. \quad (8.8)$$

From Lemma 2.1 and the second inequality in Lemma 2.3, we thus obtain

$$\begin{aligned}\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} &\leq |\lambda_0|^k\|S^{-1}A_k\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r} + \\ &\qquad \max_{i=0:k}\{\|A_i\|_2^2\}\sum_{r=1}^{\frac{k-2}{2}}|\lambda_0|^{-2r}\left(\sum_{i=0}^{2r-1}|\lambda_0|^i\right)^2 \\ &\leq |\lambda_0|^k\|S^{-1}A_k\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r} + \max_{i=0:k}\{\|A_i\|_2^2\}\sum_{r=1}^{\frac{k-2}{2}}2r\sum_{i=-r}^{r-1}|\lambda_0|^{2i} \\ &\leq d_2(\lambda_0)\max\{1, \|S^{-1}A_k\|_2^2, \max_{i=0:k}\{\|A_i\|_2^2\}\}.\end{aligned}$$

This establishes inequality (8.5).

We now prove inequality (8.6). Recall $A_k = P_0(\lambda)$. Hence, from (8.7) and Lemma 2.2, we get

$$\|\Delta_1(\lambda_0)x\|_2^2 \leq |\lambda_0|^{-k}\|S^{-1}P^{k-1}(\lambda_0)x\|_2^2 + \sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r}\|x\|_2^2 + \sum_{r=1}^{\frac{k-2}{2}}|\lambda_0|^{-2r}\|P^{2r-1}(\lambda_0)x\|_2^2.$$

Using the second inequality in Lemma 2.3, we get

$$\begin{aligned}\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} \leq &\max\{1, \max_{i=0:k}\{\|A_i\|_2^2\}, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\}\left[|\lambda_0|^{-k}\left(\sum_{j=0}^{k-1}|\lambda_0|^j\right)^2\right. \\ &\left.+\sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r} + \sum_{r=1}^{\frac{k-2}{2}}|\lambda_0|^{-2r}\left(\sum_{i=0}^{2r-1}|\lambda_0|^i\right)^2\right].\end{aligned}$$

Using Lemma 2.1, we finally obtain

$$\begin{aligned}\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} \leq &\max\{1, \max_{i=0:k}\{\|A_i\|_2^2\}, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\}\left[\sum_{r=0}^{\frac{k-2}{2}}|\lambda_0|^{2r} + \right. \\ &\left.\sum_{r=1}^{\frac{k}{2}}2r\sum_{i=-r}^{r-1}|\lambda_0|^{2i}\right] \\ \leq &d_3(\lambda)\max\{1, \max_{i=0:k}\{\|A_i\|_2^2\}, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\},\end{aligned}$$

which is the desired result. □

Lemma 8.4 *Let $\lambda_0 \in \mathbb{C}$ be nonzero and let $k \geq 2$ be a positive even integer. Let $d_1(\lambda)$, $d_2(\lambda)$ and $d_3(\lambda)$ be the functions in* (8.1), (8.2) *and* (8.3).

(a) *If* $|\lambda_0| \leq 1$, *then*

$$d_1(\lambda_0) \leq \frac{k+2}{2} + \frac{k(k^2-1)}{6}|\lambda_0|^2. \tag{8.9}$$

(b) *If* $|\lambda_0| > 1$, *then*

$$d_2(\lambda_0) \leq \left(\frac{k}{2}+1\right)|\lambda_0|^k + \frac{k(k-1)(k-2)}{6}|\lambda_0|^{k-4}, \quad \textit{and} \tag{8.10}$$

$$d_3(\lambda_0) \leq \frac{k(k+3)^2}{6}|\lambda_0|^{k-2}. \tag{8.11}$$

Proof We first prove inequality (8.9). So, assume $|\lambda_0| \leq 1$. Using $|\lambda_0|^i \leq |\lambda_0|^j$ when $i \geq j$, we get from (8.1)

$$\begin{aligned} d_1(\lambda_0) \leq & \frac{k}{2} + 1 + |\lambda_0|^2 \sum_{r=1}^{\frac{k-2}{2}} (k-2r+1)^2 \\ = & \frac{k+2}{2} + |\lambda_0|^2 \frac{k(k^2-1)}{6} \end{aligned}$$

where the equality follows from $1^2 + 3^2 + \cdots + (2n-1)^2 = \frac{n(2n+1)(2n-1)}{3}$, which implies (8.9).

Next, assume $|\lambda_0| \geq 1$. Let us prove (8.10) and(8.11). Using $|\lambda_0|^i \leq |\lambda_0|^j$ when $i \leq j$, we get from (8.2)

$$\begin{aligned} d_2(\lambda_0) &\leq \left(\frac{k}{2}+1\right)|\lambda_0|^k + |\lambda_0|^{k-4} \sum_{r=1}^{\frac{k-2}{2}} (2r)^2 \\ &= \left(\frac{k}{2}+1\right)|\lambda_0|^k + \frac{k(k-1)(k-2)}{6}|\lambda_0|^{k-4}. \end{aligned}$$

where the equality follows from

$$2^2 + \cdots + (2n)^2 = \frac{2n(n+1)(2n+1)}{3}. \tag{8.12}$$

Analogously, from (8.3), we obtain

$$\begin{aligned} d_3(\lambda_0) &\leq \frac{k}{2}|\lambda_0|^{k-2} + |\lambda_0|^{k-2} \sum_{r=1}^{\frac{k}{2}} (2r)^2 \\ &= \frac{k}{2}|\lambda_0|^{k-2} + \frac{k(k+1)(k+2)}{6}|\lambda_0|^{k-2}, \end{aligned}$$

where the equality follows also from (8.12), which are the desired results. □

We are finally in a position to prove Theorems 6.2 and 6.4.

Proof (of Theorem 6.2) By Lemma 8.1 and the definition of $\kappa_{\mathrm{ra}}(\lambda_0, P)$, we have

$$\frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} = \max\{\|\mathcal{H}_0\|_2, \|\mathcal{H}_1\|_2\} \frac{|\lambda_0|+1}{\sum_{i=0}^k |\lambda_0|^i} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2}, \tag{8.13}$$

where we have used $\max_{i=0:k}\{\|A_i\|_2\} = 1$.

We start by proving the relative-absolute upper bounds. Notice that

$$\frac{|\lambda_0|+1}{\sum_{i=0}^{k}|\lambda_0|^i} \leq \begin{cases} 1, & \text{if } |\lambda_0| \leq 1, \quad \text{and} \\ \frac{2}{|\lambda_0|^{k-1}}, & \text{if } |\lambda_0| > 1. \end{cases} \tag{8.14}$$

Moreover, by Lemma 8.2, we have

$$\max\{\|\mathcal{H}_0\|_2, \|\mathcal{H}_1\|_2\} \leq 2\zeta, \tag{8.15}$$

where ζ has been defined in (6.1). Hence, to finish the proof, we need to bound the square of the ratio $\|\Delta_1(\lambda_0)x\|_2/\|x\|_2$.

If $|\lambda_0| \leq 1$, by inequalities (8.5) and (8.9), we have

$$\begin{aligned} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} &\leq d_1(\lambda_0)\max\{1, \|S^{-1}A_k\|_2^2\} \\ &\leq \left(\frac{k+2}{2} + \frac{k(k^2-1)}{6}|\lambda_0|^2\right)\max\{1, \|S^{-1}A_k\|_2^2\} \\ &\leq \frac{k^3}{2}\max\{1, \|S^{-1}A_k\|_2^2\}. \end{aligned} \tag{8.16}$$

We notice that, if $|\lambda_0|$ is close to 0, then the previous upper bound is close to

$$\frac{k+2}{2}\max\{1, \|S^{-1}A_k\|_2^2\}.$$

If $|\lambda_0| > 1$, then, by the inequalities (8.5) and (8.10), we get

$$\begin{aligned} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} &\leq d_2(\lambda_0)\max\{1, \|S^{-1}A_k\|_2^2\} \\ &\leq \left(\left(\frac{k}{2}+1\right)|\lambda_0|^k + \frac{k(k-1)(k-2)}{6}|\lambda_0|^{k-4}\right)\max\{1, \|S^{-1}A_k\|_2^2\} \\ &\leq \frac{|\lambda_0|^k}{2}\left(k+2+\frac{k(k-1)(k-2)}{3}\right)\max\{1, \|S^{-1}A_k\|_2^2\} \\ &\leq |\lambda_0|^k\frac{k^3}{3}\max\{1, \|S^{-1}A_k\|_2^2\}, \end{aligned} \tag{8.17}$$

and, by the inequalities (8.6) and (8.11),

$$\begin{aligned} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} \leq & d_3(\lambda_0)\max\{1, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\} \\ \leq & \frac{k+k^3}{2}|\lambda_0|^{k-2}\max\{1, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\}. \end{aligned} \tag{8.18}$$

Hence, if $|\lambda_0| > 1$,

$$\begin{aligned} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} &\leq |\lambda_0|^k \times \\ &\min\left\{\frac{k^3}{3}\max\{1, \|S^{-1}A_k\|_2^2\}, \frac{k+k^3}{2|\lambda_0|^2}\max\{1, \max_{i=0:k}\{\|S^{-1}A_i\|_2^2\}\}\right\}. \end{aligned} \tag{8.19}$$

The desired upper bounds are obtained by combining the inequalities (8.14), (8.15), (8.16) and (8.19) with (8.13). □

Now, we prove the relative-absolute lower bounds. First, a direct application of the lower bound in Proposition 2.1 to $\mathcal{H}_0$ and $\mathcal{H}_1$ yields

$$\max\{\|\mathcal{H}_0\|_2, \|\mathcal{H}_1\|_2\} \geq \zeta. \tag{8.20}$$

Second, from (5.8), we get

$$\begin{aligned}\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} &\geq \sum_{i=0}^{\frac{k-2}{2}} |\lambda_0|^{2i} + |\lambda_0|^k \frac{\|S^{-1}A_k x\|_2^2}{\|x\|_2^2} \geq \sum_{i=0}^{\frac{k-2}{2}} |\lambda_0|^{2i} + |\lambda_0|^k \|A_k^{-1}S\|_2^{-2} \geq \\ &\geq \sum_{i=0}^{\frac{k}{2}} |\lambda_0|^{2i} \min\{1, \|A_k^{-1}S\|_2^{-2}\},\end{aligned} \tag{8.21}$$

where we have used that $\|x\|_2 \leq \|A^{-1}\|_2 \|Ax\|_2$ for any vector x and invertible matrix A. Hence, combining (8.20) and (8.21) with (8.13) yields

$$\begin{aligned}\frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} \geq &\mu_b \frac{(1+|\lambda_0|)\sum_{i=0}^{\frac{k}{2}} |\lambda_0|^{2i}}{\sum_{i=0}^{k} |\lambda_0|^i} = \mu_b \frac{\sum_{i=0}^{k+1} |\lambda_0|^i}{\sum_{i=0}^{k} |\lambda_0|^i} \geq \\ &\begin{cases} \mu_b & \text{if } |\lambda_0| \leq 1, \text{and} \\ \mu_b \left(1 + \frac{|\lambda_0|}{k+1}\right) & \text{if } |\lambda_0| > 1, \end{cases}\end{aligned}$$

where μ_b has been defined in (6.3). Furthermore, by (5.8), we also have

$$\frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2} \geq \sum_{i=0}^{\frac{k-2}{2}} |\lambda_0|^{2i}. \tag{8.22}$$

Thus, from (8.13), (8.20) and (8.22), we obtain

$$\begin{aligned}\frac{\kappa_{\mathrm{ra}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{ra}}(\lambda_0; P)} \geq &\zeta \frac{(|\lambda_0|+1)\sum_{i=0}^{\frac{k-2}{2}} |\lambda_0|^{2i}}{\sum_{i=0}^{k} |\lambda_0|^i} \geq \zeta \frac{\sum_{i=0}^{k-1} |\lambda_0|^i}{\sum_{i=0}^{k} |\lambda_0|^i} \geq \\ &\begin{cases} \dfrac{\zeta}{2} & \text{if } |\lambda_0| \leq 1, \text{ and} \\ \dfrac{\zeta}{2|\lambda_0|} & \text{if } |\lambda_0| > 1, \end{cases}\end{aligned}$$

and the desired lower bounds have been established. □

Proof (of Theorem 6.4) By Lemma 8.1 and the definition of $\kappa_{\mathrm{rr}}(\lambda_0, P)$, we have

$$\frac{\kappa_{\mathrm{rr}}(\lambda_0; \mathcal{H}_P^S)}{\kappa_{\mathrm{rr}}(\lambda_0; P)} = \frac{(|\lambda_0|\|\mathcal{H}_1\|_2 + \|\mathcal{H}_0\|_2)}{\sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2} \frac{\|\Delta_1(\lambda_0)x\|_2^2}{\|x\|_2^2}. \tag{8.23}$$

Notice also

$$\max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\} \leq \sum_{i=0}^{k} |\lambda_0|^i \|A_i\|_2 \leq (k+1) \max_{i=0:k}\{|\lambda_0|^i \|A_i\|_2\}. \tag{8.24}$$

Furthermore, from Lemma 8.2, we readily obtain

$$|\lambda_0|\,\|\mathcal{H}_1\|_2 + \|\mathcal{H}_0\|_2 \leq \begin{cases} 4\zeta & \text{if } |\lambda_0| \leq 1, \quad \text{and} \\ 4|\lambda_0|\zeta & \text{if } |\lambda_0| > 1, \end{cases} \tag{8.25}$$

where ζ has been defined in (6.1).

When $|\lambda_0| \leq 1$, the desired upper bound follows by combining (8.16), (8.24), and (8.25) with (8.23). When $|\lambda_0| > 1$, the desired upper bound follows by combining (8.19), (8.24), and (8.25) with (8.23).

Now, we prove the lower bounds. First, from Proposition 2.1, we get

$$|\lambda_0|\,\|\mathcal{H}_1\|_2 + \|\mathcal{H}_0\|_2 \geq \begin{cases} \max\{1, \|S^* A_k^{-1} S\|_2\} & \text{if } |\lambda_0| \leq 1, \quad \text{and} \\ |\lambda_0| \max\{1, \|S\|_2\} & \text{if } |\lambda_0| > 1. \end{cases} \tag{8.26}$$

Then, notice

$$\frac{\|\Delta_1(\lambda_0)x\|_2}{\|x\|_2^2} \geq \begin{cases} 1 & \text{if } |\lambda_0| \leq 1, \quad \text{and} \\ \max\{|\lambda_0|^k \|A_k^{-1} S\|_2^{-2}, |\lambda_0|^{k-2}\} & \text{if } |\lambda_0| > 1. \end{cases} \tag{8.27}$$

which readily follows from (5.8). The lower bounds are obtained by combining (8.24), (8.26) and (8.27) with (8.23).

9 Conclusions

In this paper, we propose a new strategy to solve the even degree symmetric/Hermitian polynomial eigenvalue problem. We have shown evidence that, the traditional approach of using the linearizations $D_1(\lambda; P)$ and $D_k(\lambda; P)$ is, in many occasions, risky due to the fact that the eigenvectors of these two linearizations are too sensitive to small perturbations in their matrix coefficients. This sensitivity leads to large backward errors for the computed eigenpairs. We propose instead the use of the linearizations $\mathcal{H}_P^{A_k}(\lambda)$ and $\mathcal{G}_P^{A_k}(\lambda)$ introduced in (5.7) and (5.9) (when S is replaced by A_k). We have proven that the condition numbers of the eigenvalues with small (resp. large) modulus of $D_k(\lambda; P)$ (resp. $D_1(\lambda; P)$) and $\mathcal{H}_P^{A_k}(\lambda)$ (resp. $\mathcal{G}_P^{A_k}(\lambda)$) are comparable. But we have also shown that the condition number of the eigenvalues with large (resp. small) modulus of $D_k(\lambda; P)$ (resp. $D_1(\lambda; P)$) is significantly worse than that of the eigenvalues of $\mathcal{H}_P^{A_k}(\lambda)$ (resp. $\mathcal{G}_P^{A_k}(\lambda)$), specially for moderate to large values of the degree k of $P(\lambda)$. In future work we intend to determine if the sensitivity of the eigenvectors of $D_1(\lambda; P)$ and $D_k(\lambda; P)$ truly depends on the existence of ill-conditioned eigenvalues or if it depends on any other factors. We would also like to determine how this sensitivity changes the structure of the computed eigenvectors, in particular, the structure of the blocks from which the eigenvectors of the polynomial $P(\lambda)$ are recovered, and how this change affects the backward errors of the computed eigenpairs.

A Structure preserving deflation

In this section we consider the even-degree matrix polynomial (1.1) as an odd-grade matrix polynomial by adding an extra zero matrix coefficient, that is,

$$Q(\lambda) := \lambda^{k-1} 0_n + P(\lambda). \tag{A.1}$$

We observe that the pencil $\mathcal{T}_Q^{k+1}(\lambda)$ (see (5.2)) is a "weak" linearization for $Q(\lambda)$, i.e., it is not a strong linearization, since $\mathcal{T}_Q^{k+1}(\lambda)$ has n extra spurious eigenvalues at infinity. Nonetheless, the Kronecker structure of these eigenvalues at infinity is very simple, as we show in the next lemma.

Lemma A.1 *Let $P(\lambda)$ be an $n \times n$ even-degree matrix polynomial as in* (1.1)*, and let Q be as in* (A.1)*. Then, the spectrum of $\mathcal{T}_Q^{k+1}(\lambda)$ consists of the spectrum of $P(\lambda)$ together with n eigenvalues at infinity of index one, i.e., with Kronecker blocks of size 1.*

Proof Notice $\mathrm{rev}_{k+1}Q(\lambda) = \lambda \mathrm{rev}_k P(\lambda)$. Hence,

$$\det\left(\mathrm{rev}_{k+1}Q(\lambda)\right) = \lambda^n \det\left(\mathrm{rev}_k P(\lambda)\right),$$

and, so, all the extra n eigenvalues at infinity of $Q(\lambda)$ have algebraic and geometric multiplicity equal to one.

Lemma A.1 allows us to apply to $\mathcal{T}_Q^{k+1}(\lambda)$ the structure preserving deflation developed by Mehrmann and Xu [17], provided that A_k is nonsingular. Hence, we can deflate the n spurious eigenvalues at infinity of $\mathcal{T}_Q^{k+1}(\lambda)$ preserving the symmetric structure of the pencil. Surprisingly, the result of applying the deflation procedure to $\mathcal{T}_Q^{k+1}(\lambda)$ is essentially a pencil of the form (5.6). The overall goal of this section is to prove this fact.

Let $\lambda\mathcal{T}_1 - \mathcal{T}_0 := \mathcal{T}_Q^{k+1}(\lambda)$. The first step of the deflation consists in finding a unitary matrix U such that

$$U^*\mathcal{T}_1 = \begin{bmatrix} N \\ 0 \end{bmatrix}$$

where N is of full row rank. Notice that $\mathcal{T}_1$ has n zero rows (its first n rows), so the unitary matrix can be chosen as the permutation matrix.

$$U^* = \begin{bmatrix} U_1^* \\ \hline U_2^* \end{bmatrix} := \begin{bmatrix} 0 & I_{kn} \\ \hline I_n & 0 \end{bmatrix}.$$

With this choice for U, the resulting N is of full row rank since it contains a $kn \times kn$ nonsingular matrix.

The second step of the deflation procedure consists in finding a unitary matrix V such that

$$U_2^*\mathcal{T}_0 V = \begin{bmatrix} A_k & -I_n & 0 & \cdots & 0 \end{bmatrix} V = \begin{bmatrix} 0_{kn} & M \end{bmatrix},$$

where M is nonsingular. We can find such unitary matrix V by using a rank revealing factorization (via a QR decomposition with partial pivoting or the SVD decomposition, for example). Let

$$\begin{bmatrix} A_k & -I_n \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \begin{bmatrix} 0_n & M \end{bmatrix},$$

which implies, in particular, $A_k V_{11} = V_{21}$. Then, set

$$V = \left[\begin{array}{c|cc} 0 & V_{11} & V_{12} \\ & V_{21} & V_{22} \\ \hline I_{(k-1)n} & \multicolumn{2}{c}{0} \end{array}\right].$$

Using MATLAB notation for submatrices, the deflated pencil is the $kn \times kn$ pencil

$$V(:, 1:kn)^* \mathcal{T}_Q^{k+1}(\lambda) V(:, 1:kn),$$

which is permutationally equivalent to

$$\left[\begin{array}{cc|c} V_{11}^* & V_{21}^* & 0 \\ \hline \multicolumn{2}{c|}{0} & I_{(k-1)n} \end{array}\right] \mathcal{T}_Q^{k+1}(\lambda) \left[\begin{array}{c|c} V_{11} & 0 \\ V_{21} & \\ \hline 0 & I_{(k-1)n} \end{array}\right],$$

which equals

$$\begin{bmatrix} -V_{21}^*V_{11} & \lambda V_{21}^* & & & & & \\ \lambda V_{21} & \lambda A_{k-1}+A_{k-2} & -I_n & & & & \\ & -I_n & 0 & \lambda I_n & & & \\ & & \lambda I_n & \lambda A_{k-3}+A_{k-4} & & & \\ & & & & \ddots & & \\ & & & & & & \lambda I_n \\ & & & & & \lambda I_n & \lambda A_1 + A_0 \end{bmatrix} = \mathcal{H}_P^{V_{21}}(\lambda),$$

as we wanted to show, where we have used $A_k V_{11} = V_{21}$.

References

1. B. Adhikari, R. Alam, D. Kressner. Structured eigenvalue condition numbers and linearizations for matrix polynomials. Linear Algebra Appl., 435(9), pp. 2193–2221, 2011
2. E. N. Antoniou, and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electron. J. Linear Algebra*, 11, pp. 78–87, 2004.
3. M. N. Belur, C. Praagman. An efficient algorithm for computing the H_∞ norm. *IEEE Transactions on Automatic Control*, 56(7), pp. 1656–1660, 2011.
4. M. I. Bueno, F. M. Dopico, S. Furtado, and L. Medina. A block-symmetric linearization of odd-degree matrix polynomials with optimal eigenvalue condition number and backward error, Calcolo, pp. 55–32, 2018.
5. M. I. Bueno, M. Marting, J. Pérez, A. Song, and I. Viviano. Explicit Block-Structures for Block-Symmetric Fiedler-Like Pencils. *Electronic Journal of Linear Algebra*, 34, pp. 472–499, 2018.
6. R. Byers, V. Mehrmann, H. Xu. Trimmed linearizations for structured matrix polynomials. *Linear Algebra Appl.*, 429, pp. 2373–2400, 2008.
7. F. De Terán, F. M. Dopico, and D. S. Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. *Linear Algebra Appl.*, 459, pp. 264–333, 2014.
8. N. A. Dumont. On the solution of generalized non-linear complex-symmetric eigenvalue problems. *Int. J. Numer. Meth. Engng*, 71, pp. 1534–1568, 2007.
9. I. Gohberg, P. Lancaster, L. Rodman. *Matrix Polynomials.* Academic Press, New York-London, 1982.
10. N. J. Higham, R. -C. Li and F. Tisseur. Backward error of polynomial eigenproblems solved by linearizations. *SIAM J. Matrix Anal. Appl.*, 29, pp. 143–159, 2006.
11. N. J. Higham, D. S. Mackey, N. Mackey, F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1), pp. 143–159, 2006.
12. N. J. Higham, and F. Tisseur. Bounds for eigenvalues of matrix polynomials. *Linear Algebra Appl.*, 258, pp. 5–22, 2003.
13. A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*, Jyväskylä, Finland, 2004. CD-ROM.
14. D. Kressner, C. Schröder, D. S. Watkins. Implicit QR algorithm for palindromic and even eigenvalue problems. *Numer. Algor.*, 51, pp. 209–238, 2009.
15. P. Lancaster. *Lambda-Matrices and Vibrating Systems.* Pergamon, Oxford, 1966.
16. 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, pp. 867–891, 2006.
17. V. Mehrmann, and H. Hongguo Xu. Structure preserving deflation of infinite eigevanlues in structured pencils. *ETNA*, 44, pp. 1–24, 2015.
18. V. Mehrmann, D. S. Watkins. Polynomial eigenvalue problems with Hamiltonian structure. *ETNA*, 13, pp. 106–118, 2002.
19. G. Mele. The infinite Lanczos method for symmetric nonlinear eigenvalue problems. arXiv:1812.07557, 2019.
20. E. Mengi. On the estimation of the distance to uncontrollability for higher order systems. *SIAM J. Matrix Anal. Appl.*, 30(1), pp. 154–172, 2008.
21. Y. Nakatsukasa, F. Tisseur Eigenvector error bound and perturbation for polynomial and rational eigenvalue problems Technical report.

22. F. Tisseur. Backward error and condition of polynomial eigenvalue problems. Linear Algebra Appl., 309(1–3), pp. 339–361, 2000.
23. F. Tisseur, K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43, pp. 235–286, 2001.
24. Y. Su, Z. Bai. Solving Rational Eigenvalue Problems via Linearization. *SIAM J. Matrix Anal. Appl.*, 32(1), pp. 201–216, 2011.
25. L. Zeng, Y. Su. A backward stable algorithm for quadratic eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 35(2), pp. 499–516, 2014.