

**MATHEMATICAL ENGINEERING
TECHNICAL REPORTS**

**Eigenvector Error Bound and Perturbation for
Polynomial and Rational Eigenvalue Problems**

Yuji NAKATSUKASA and Françoise TISSEUR

METR 2016-04

April 2016

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: <http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html>

The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may not be reposted without the explicit permission of the copyright holder.

EIGENVECTOR ERROR BOUND AND PERTURBATION FOR POLYNOMIAL AND RATIONAL EIGENVALUE PROBLEMS

YUJI NAKATSUKASA* AND FRANÇOISE TISSEUR†

Abstract. A computed approximate eigenpair $(\widehat{\lambda}_*, \widehat{x})$ of a polynomial eigenvalue problem $P(\lambda)x = 0$ is usually considered acceptable if the normalized residual $P(\widehat{\lambda}_*)\widehat{x}$ is small, as it indicates the computed solution has a small backward error. The error in the approximate eigenvalue $\widehat{\lambda}_*$ can be bounded for example as the error of $\widehat{\lambda}_*$ as an eigenvalue of a linearization of P . However, no previous result exists that rigorously bound the error in the computed eigenvector \widehat{x} . We derive a posteriori upper bounds for the angle between \widehat{x} and an exact eigenvector x of $P(\lambda)$ by showing that the desired angle $\angle(\widehat{x}, x)$ is bounded by $\angle(\widehat{y}, y)$, where y is an exact eigenvectors of a linearization and \widehat{y} is its approximation, constructed using \widehat{x} and $\widehat{\lambda}_*$. The result extends to other nonlinear eigenvalue problems such as rational matrix eigenproblems, and the bounds can be obtained not only from linearizations but also from other formulations such as quadratifications and ℓ -ifications. This work shows that just like for linear eigenvalue problems, the eigenvector error is proportional to the residual and inversely proportional to the separation between the eigenvalues of $P(\lambda)$. One implication of our result is that an eigenvector can be computed accurately even when it corresponds to several distinct eigenvalues.

Key words. polynomial eigenvalue problem, eigenvector, error bounds, perturbation, nonlinear eigenvalue problem, rational eigenvalue problem, linearization, quadratification, ℓ -ification

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

1. Introduction. Polynomial eigenvalue problems $P(\lambda)x = 0$ where

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

arise in many areas of scientific computing. A pair (λ_*, x) with $x \neq 0$ is called an eigenpair if $P(\lambda_*)x = 0$. A number of numerical methods for solving such problems are known, such as forming a companion linearization and invoking standard algorithms for linear eigenproblems [10, 20, 31], Jacobi-Davidson type methods [25], and Krylov subspace methods [2, 22].

We are concerned with quantifying the errors in a computed eigenpair $(\widehat{\lambda}_*, \widehat{x})$ that approximates an exact eigenpair (λ_*, x) , such that the residual $r := P(\widehat{\lambda}_*)\widehat{x}$ is small but nonzero. In this paper we focus on the forward error of the eigenvector measured by the angle $\angle(x, \widehat{x})$ rather than the backward errors as discussed in [30]. By angle between two nonzero vectors x and \widehat{x} , we mean

$$\angle(x, \widehat{x}) = \cos^{-1} \frac{|\widehat{x}^* x|}{\|x\|_2 \|\widehat{x}\|_2}$$

For eigenvalues, bounding the error in $\widehat{\lambda}_*$ can be done by regarding $\widehat{\lambda}_*$ as an approximate eigenvalue of a linearization and then invoke well-established eigenvalue

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan (nakatsukasa@mist.i.u-tokyo.ac.jp). Supported by JSPS grant No. 26870149, and EPSRC grant EP/I005293/1.

† School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (ftisseur@ma.man.ac.uk). Supported by Engineering and Physical Sciences Research Council grant EP/I005293 by a Royal Society-Wolfson Research Merit Award.

perturbation for linear problems [27]. For eigenvectors, however, such techniques cannot be used directly because the eigenvector y for a linearization is not equal to the eigenvector x of $P(\lambda)$.

The main result of this paper is that the error $\angle(\hat{x}, x)$ of an approximate eigenvector \hat{x} can nonetheless be bounded rigorously and accurately by the error $\angle(\hat{y}, y)$ for an appropriately constructed approximate eigenvector \hat{y} of a linearization (and more general reductions of $P(\lambda)$, as we explain in Section 2). Bounding $\angle(\hat{y}, y)$ can then be done by applying existing perturbation theory for linear eigenproblems.

Previous work exist on the sensitivity and condition number of eigenvectors of matrix polynomials [4, 7]. However, such studies have dealt only with first-order perturbation analysis, leading to the condition numbers of the eigenvector. The resulting first-order bounds are not applicable to give rigorous error bounds for \hat{x} in a practical setting in which the residual r is small but not infinitesimally so. The bounds we derive are simple involving only the residual and a quantity that measures the separation between eigenvalues, often denoted “sep”. These are precisely the quantities used to bound eigenvector errors for linear eigenproblems.

There are infinitely many linearizations of a matrix polynomial $P(\lambda)$, and many types that can be constructed explicitly from the coefficients of $P(\lambda)$ have been proposed. For the standard companion linearization, and more generally for the \mathbb{L}_1 linearizations introduced in [20], the corresponding eigenvector of $L(\lambda)$ such that $L(\lambda)y = 0$ has the Vandermonde structure

$$y = \begin{bmatrix} \lambda^{k-1}x \\ \lambda^{k-2}x \\ \vdots \\ \lambda x \\ x \end{bmatrix} = \begin{bmatrix} \lambda^{k-1} \\ \lambda^{k-2} \\ \vdots \\ \lambda \\ 1 \end{bmatrix} \otimes x, \quad (1.2)$$

where \otimes denotes the Kronecker product [11, § 1.3.6], [14, App. B].

A simple but key result of this paper is that for an approximate (computed) eigenpair $(\hat{\lambda}_*, \hat{x})$ of $P(\lambda)$, the accuracy of \hat{x} , measured by the angle $\angle(x, \hat{x})$ between x and an exact eigenvector x , is at least as good as that of \hat{y} as an approximation to an exact eigenvector y of a linearization $L(\lambda)$, that is,

$$\angle(x, \hat{x}) \leq \angle(y, \hat{y}). \quad (1.3)$$

Here \hat{y} is defined using the computed pair $(\hat{\lambda}_*, \hat{x})$ so as to inherit the Vandermonde form that y has in (1.2):

$$\hat{y} = [\hat{\lambda}_*^{k-1}, \hat{\lambda}_*^{k-2}, \dots, \hat{\lambda}_*, 1]^T \otimes \hat{x}. \quad (1.4)$$

The inequality (1.3) shows that an error bound of a computed eigenvector \hat{x} of $P(\lambda)$ can be obtained by quantifying the error of \hat{y} as an eigenvector of the linearization $L(\lambda)$, and bounding $\angle(y, \hat{y})$ can be done using classical results on eigenvector perturbation theory for linear eigenvalue problems, such as those in the classical book [27]. Note that \hat{y} is not equal to the computed eigenvector of $L(\lambda)$, although \hat{x} is usually obtained by extracting a part of it: it is this construction of \hat{y} forcing it to be in Vandermonde form that leads to the crucial inequality (1.3) to connect the eigenvector angles with respect to $P(\lambda)$ and $L(\lambda)$.

Our results hold not only for \mathbb{L}_1 -type linearizations, but in a more general setting. For example, sometimes $P(\lambda)$ is expressed in a polynomial basis $\{\phi_i(z)\}_{i=0}^k$ that is not necessarily the standard monomials. Such matrix polynomials have been the subject of several recent studies, see for example [1, 8, 21]. In this case the exact and approximate eigenvectors for the companion-like¹ (sometimes called comrade) linearization are in the forms

$$y = [\phi_{k-1}(\lambda), \phi_{k-2}(\lambda), \dots, \phi_0(\lambda)]^T \otimes x, \quad \hat{y} = [\phi_{k-1}(\hat{\lambda}_*), \phi_{k-2}(\hat{\lambda}_*), \dots, \phi_0(\hat{\lambda}_*)]^T \otimes \hat{x}. \quad (1.5)$$

The inequality (1.3) still holds with such x, \hat{x}, y, \hat{y} .

More generally, our results are applicable not only for linearizations, but for any reduction of $P(\lambda)$ into another eigenproblem $L(\lambda)y = 0$ as long as the eigenvector y can be expressed in the Kronecker form

$$y = [\phi_{k-1}(\lambda), \phi_i(\lambda), \dots, \phi_0(\lambda)]^T \otimes x = \Lambda(\lambda) \otimes x, \quad (1.6)$$

and $\Lambda(\lambda) = [\phi_{k-1}(\lambda), \phi_{k-2}(\lambda), \dots, \phi_0(\lambda)]^T$ is a vector whose elements $\phi_i(\lambda)$ are functions of λ . The vectors in (1.2), (1.4) and (1.5) are special cases of (1.6). Note that $\phi_i(\lambda)$ are not necessarily polynomials, and the only assumption is that the vector $\Lambda(\lambda)$ is nonzero $\|\Lambda(\lambda)\| \neq 0$ for every λ . For an approximate eigenpair $(\hat{\lambda}_*, \hat{x})$ of $P(\lambda)$ we always construct the approximate eigenpair $(\hat{\lambda}_*, \hat{y})$ for a reduction $L(\lambda)$ by forming the Vandermonde-like vector

$$\hat{y} = \Lambda(\hat{\lambda}_*) \otimes \hat{x}. \quad (1.7)$$

Examples other than linearizations of matrix polynomials for which our results hold include companion-like linearizations for rational matrix eigenvalue problems and ℓ -ifications of matrix polynomials introduced in [6], as we discuss in Section 2.

For the sake of generality we present the results mainly in the general form (1.6), and it often helps to think of the special but important case where $L(\lambda)$ is the standard companion linearization of a matrix polynomial $P(\lambda)$. Generally $\phi_i(\lambda)$ is determined from the structure of the eigenvectors of the reduction one chooses, and is allowed to be different from the basis in which $P(\lambda)$ is represented.

In Section 2 we review some reductions $L(\lambda)$ of $P(\lambda)$ for which our results are applicable. Section 3 derives a key lemma connecting $\angle(\hat{x}, x)$ and $\angle(\hat{y}, y)$, and in Section 4 we derive eigenvector error bounds for matrix polynomials. We discuss eigenvector perturbation bounds in Section 5, and treat eigenspace corresponding to a multiple eigenvalue in Section 6. We summarize the results and discuss in Section 7, and present numerical experiments in Section 8.

Notation. $P(\lambda)$ represents a matrix whose entries are nonlinear functions of λ (such as a matrix polynomial), and $(\hat{\lambda}_*, \hat{x})$ is an approximate eigenpair such that the residual $P(\hat{\lambda}_*)\hat{x}$ is small but nonzero. $L(\lambda)$ is a reduced problem of $P(\lambda)$ (such as linearization) whose eigenvectors y satisfy (1.6) for some vector function $\Lambda(\hat{\lambda}_*)$. The approximate eigenvector \hat{y} for $L(\lambda)$ is always constructed from $\hat{x}, \hat{\lambda}_*$ as in (1.7). We assume throughout that $P(\lambda)$ is regular, that is, $\det P(\lambda)$ is not identically zero. We scale the eigenvector x and its computed approximation \hat{x} to have unit norm $\|x\|_2 = \|\hat{x}\|_2 = 1$.

¹These linearizations can actually be regarded as an \mathbb{L}_1 -type linearization in a nonstandard polynomial basis $\phi_i(\lambda)$ [24].

2. Reductions $L(\lambda)$ of $P(\lambda)$ and their eigenvector structures. For most linearizations $L(\lambda)$ used in practice and some other reductions for solving nonlinear eigenproblems $P(\lambda)$, the eigenvector y of the reduced problem $L(\lambda)$ is related to the eigenvector x of $P(\lambda)$ in a highly structured manner, written as $y = \Lambda \otimes x$ as in (1.5). Each $\phi_i(\lambda)$ is a function of λ , usually but not necessarily a polynomial. Our results hold for any such linearizations, and below we list important examples.

The most well known linearization of a matrix polynomial $P(\lambda)$ as in (1.1) is the companion linearization for $P(\lambda)$ expressed in the monomial basis $P(\lambda) = \sum_{i=0}^k A_i \lambda^i$, which yields the linear eigenvalue problem

$$L(\lambda) = C_0 + \lambda C_1 = \begin{bmatrix} A_{k-1} & A_{k-2} & \dots & A_0 \\ -I & & & \\ & \ddots & & \\ & & -I & \end{bmatrix} + \lambda \begin{bmatrix} A_k & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix}. \quad (2.1)$$

For each eigenpair (λ, x) such that $P(\lambda)x = 0$, we have $L(\lambda)y = 0$ where the eigenvector y is in Vandermonde form $y = [\lambda^{k-1}, \lambda^{k-2}, \dots, \lambda, 1]^T \otimes x$, which clearly has the form (1.6) with $\phi_i(\lambda) = \lambda^i$.

In some cases $P(\lambda)$ is expressed as $P(\lambda) = \sum_{i=0}^k \varphi_i(\lambda) A_i$, where $\varphi_i(\lambda)$ is a polynomial basis that is not necessarily the standard monomials, for example orthogonal polynomials such as Chebyshev [8] or Legendre [1]. Analogues of companion linearization in the Chebyshev basis is known as the colleague matrix [33, Ch. 18], and as comrade matrices in other orthogonal bases [3]. In such cases the eigenvector y of the comrade matrices can be written as (1.5) with $\phi_i(\lambda) = \varphi_i(\lambda)$, which is in Vandermonde form with respect to the basis $\{\varphi_i(x)\}_{i=0}^k$.

A recent work [34] introduces linearization for matrix polynomials obtained by Lagrange or Hermite interpolation. The linearization there also has the Vandermonde eigenvector structure (1.6), with $\phi_i(\lambda)$ being Lagrange polynomials.

Fiedler linearizations [9] can also be seen as an \mathbb{L}_1 pencil for monic matrix polynomials, where the basis is not the standard monomials. Extensions of the $\mathbb{L}_1, \mathbb{DL}$ linearization spaces to nonstandard bases is developed in [32]. The right eigenvectors y of such linearizations are again in the form (1.6).

The remainder of this section discusses examples of reductions $L(\lambda)$ that are not standard linearizations of matrix polynomials but still satisfy the Vandermonde eigenvector structure (1.6).

2.1. Rational eigenvalue problem. Suppose $R(\lambda)$ is a rational matrix function, expressed in partial fraction form as

$$R(\lambda) = P(\lambda) + \sum_{i=1}^{\ell} \left(\sum_{j=1}^{m_i} \frac{R_{ij}}{(\lambda - p_i)^j} \right), \quad (2.2)$$

where $P(x) = \sum_{i=0}^d A_i \lambda^i \in \mathbb{C}[\lambda]^{n \times n}$ is a matrix polynomial, $R_{ij} \in \mathbb{C}^{n \times n}$, p_i are known scalars and m_i are integers. The goal of a rational eigenvalue problem is to find λ and $x \neq 0$ such that $R(\lambda)x = 0$. A naive approach is to form the polynomial $r(x) \prod_i (x - p_i)^{m_i}$ and find its roots, say by the companion linearization. However, multiplying $\prod_i (x - p_i)^{m_i}$ can change the conditioning of the solutions, resulting in

unnecessary numerical error. A more efficient linearization is introduced in [28] under the assumption that the rational terms are of low-rank, but that linearization does not have eigenvectors in the form (1.6), see Section 2.3 below.

An alternative linearization with the eigenvector property (1.6) can be formed directly from the coefficients A_i and R_{ij} [23]. These can be regarded as companion-like linearizations [19], in the sense that they can be formed by appropriately arranging the coefficient matrices and fixed constants 1s and p_i s.

For simplicity we assume monicity, i.e., the leading coefficient A_d is I . Otherwise we can left-multiply $R(\lambda)$ by A_d^{-1} ; we need to work with a generalized eigenproblem when A_d is singular. The companion-like matrix C is constructed as follows:

$$C = \begin{bmatrix} -A_{d-1} & -A_{d-2} & \dots & -A_0 & -R_{11} & \dots & -R_{1m_1} & -R_{21} & \dots & -R_{\ell m_\ell} \\ I & & & & & & & & & \\ & I & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & I & p_1 I & & & & & \\ & & & & \ddots & \ddots & & & & \\ & & & & & I & p_1 I & & & \\ & & & I & & & p_2 I & & & \\ & & & & & & & I & \ddots & \\ & & & I & & & & & & p_\ell I \end{bmatrix}. \quad (2.3)$$

One can verify that if $R(\lambda)x = 0$ then $Cy = \lambda y$, with

$$y = \left[\lambda^{d-1}x, \lambda^{d-2}x, \dots, \lambda x, x, \frac{x}{\lambda-p_1}, \dots, \frac{x}{(\lambda-p_1)^{m_1}}, \frac{x}{\lambda-p_2}, \dots, \frac{x}{(\lambda-p_\ell)^{m_\ell}} \right]^T.$$

This is still in the form (1.6) in which $\phi_i(\lambda)$ are rational functions

$$[\phi_{k-1}(\lambda), \dots, \phi_0(\lambda)] = \left[\lambda^{d-1}, \lambda^{d-2}, \dots, \lambda, 1, \frac{1}{\lambda-p_1}, \dots, \frac{1}{(\lambda-p_1)^{m_1}}, \frac{1}{\lambda-p_2}, \dots, \frac{1}{(\lambda-p_\ell)^{m_\ell}} \right],$$

where $k = d + \sum_{i=1}^{\ell} m_i$.

2.2. Quadratication and ℓ -ifications. Linearizations are not the only method to reduce a polynomial eigenproblem to a lower degree eigenproblem. For example, an $n \times n$ polynomial eigenproblem $P(\lambda)$ of degree $k\ell$ can be reduced to a $nk \times nk$ polynomial eigenproblem of degree ℓ , called ℓ -ifications [6]. The eigenvectors are in the form (1.6) in which $\Lambda(\lambda)$ is a $k \times 1$ vector polynomial. We illustrate this with an example of a quartic polynomial eigenvalue problem

$$P(\lambda) = A\lambda^4 + B\lambda^3 + C\lambda^2 + D\lambda + E.$$

We can construct a quadratication $Q(\lambda)$ of size $2n \times 2n$ defined by

$$Q(\lambda) = \begin{bmatrix} A & C \\ 0 & I \end{bmatrix} \lambda^2 + \begin{bmatrix} B & D \\ 0 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0 & E \\ -I & 0 \end{bmatrix}. \quad (2.4)$$

We can verify that for every eigenpair (λ, x) of $P(\lambda)$ such that $P(\lambda)x = 0$ we have

$$Q(\lambda)y = 0, \quad y = \begin{bmatrix} \lambda^2 \\ 1 \end{bmatrix} \otimes x.$$

Observe that y is in the form (1.6) with $\Lambda = \begin{bmatrix} \lambda^2 \\ 1 \end{bmatrix}$.

2.3. Inapplicable reductions. It is worth noting that there are reductions for which our results cannot be used to give bounds for $\angle(x, \hat{x})$.

One example is the quadratification of an even-degree matrix polynomial given in [12, Example 2.6], [17], which reduces a polynomial eigenvalue problem to a quadratic eigenvalue problem in a form different from (2.4). The resulting quadratic eigenproblem has eigenvectors of the form $y = [(\lambda^2 I + A_0)x]^\top$ where A_0 is the coefficient matrix for the constant term. Note that y does not have the form $\Lambda \otimes x$, so our results are inapplicable unless A_0 is a multiple of I .

Another inapplicable example is the linearization for rational eigenproblems introduced in [28], in which the eigenvector structure is $y = \begin{bmatrix} \Lambda(\lambda) \otimes x \\ C(\lambda)x \end{bmatrix}$ for some matrix function $C(\lambda)$, which is not in the form (1.6) unless $C(\lambda)$ is a multiple of I .

3. Bounding $\angle(\hat{x}, x)$ via $\angle(\hat{y}, y)$ of a reduction. Since a linearization $L(\lambda) = \lambda M + N$ is a linear eigenproblem there are several known ways to bound $\angle(\hat{y}, y)$, where \hat{y} is an approximate eigenvector of $L(\lambda)$. A natural question is, therefore, how is $\angle(\hat{x}, x)$, the angle between the approximate and exact eigenvectors of $P(\lambda)$, related to $\angle(\hat{y}, y)$, that of a linearization $\angle(\hat{x}, x)$? This section addresses this question.

We start with a simple but important result that relates $\angle(\hat{x}, x)$ with $\angle(\hat{y}, y)$ where y, \hat{y} are as defined in (1.5) for any x and λ .

LEMMA 3.1. *Let $\phi_i(z)$, $i = 0, \dots, k-1$ be functions of z such that for any z , at least one $\phi_i(x)$ is nonzero. Then for any vectors x, \hat{x} and scalars $\lambda, \hat{\lambda}_*$, defining $y = \Lambda(\lambda) \otimes x$ and $\hat{y} = \Lambda(\hat{\lambda}_*) \otimes \hat{x}$ as in (1.6) and (1.7), we have*

$$\angle(\hat{x}, x) \leq \angle(\hat{y}, y). \quad (3.1)$$

Proof. We prove that $\cos \angle(\hat{y}, y) \leq \cos \angle(\hat{x}, x)$. Recalling $y = \Lambda(\lambda) \otimes x = \begin{bmatrix} \phi_{k-1}(\lambda)x \\ \vdots \\ \phi_0(\lambda)x \end{bmatrix}$

and $\hat{y} = \Lambda(\hat{\lambda}_*) \otimes \hat{x} = \begin{bmatrix} \phi_{k-1}(\hat{\lambda}_*)\hat{x} \\ \vdots \\ \phi_0(\hat{\lambda}_*)\hat{x} \end{bmatrix}$, we have

$$\begin{aligned} \cos \angle(\hat{y}, y) &= \frac{|\hat{y}^* y|}{\|\hat{y}\|_2 \|y\|_2} = \frac{|\Lambda(\hat{\lambda}_*)^* \Lambda(\lambda) \|\hat{x}^* x|}{\|\hat{\Lambda}(\lambda)\|_2 \|\hat{x}\|_2 \|\Lambda(\lambda)\|_2 \|x\|_2} = \frac{|\Lambda(\hat{\lambda}_*)^* \Lambda(\lambda)|}{\|\Lambda(\hat{\lambda}_*)\|_2 \|\Lambda(\lambda)\|_2} \cos \angle(\hat{x}, x) \\ &= \cos \angle(\Lambda(\lambda), \Lambda(\hat{\lambda}_*)) \cos \angle(\hat{x}, x) \leq \cos \angle(\hat{x}, x). \end{aligned} \quad (3.2)$$

□

We immediately have the following theorem.

THEOREM 3.2. *For a nonlinear matrix function $P(\lambda)$, let $(\hat{\lambda}_*, \hat{x})$ be an approximate eigenpair with nonzero residual $\|P(\hat{\lambda}_*)\hat{x}\|$. Let $L(\lambda)$ be another eigenproblem such that for each eigenpair (λ, x) of $P(\lambda)$ we have $L(\lambda)y = 0$, where y is in the form $y = \Lambda(\lambda) \otimes x$ for some vector function $\Lambda(\lambda)$, as in (1.5). Then*

$$\angle(\hat{x}, x) \leq \angle(\hat{y}, y).$$

Theorem 3.2 shows the accuracy of the approximate eigenvector $\angle(\hat{x}, x)$ of $P(\lambda)$ is always at least as good as $\angle(\hat{y}, y)$ of a reduction $L(\lambda)$. The theorem is stated in

general form, and it is helpful to think of the important case where $P(\lambda)$ is a matrix polynomial and $L(\lambda)$ is a linearization such as the companion form.

In practice $\widehat{\lambda}_*$ is an approximation to λ so we usually have $\widehat{\lambda}_* \simeq \lambda$, which means typically $\cos \angle(\widehat{\Lambda}, \Lambda) \simeq 1$ and the last inequality in (3.2) becomes nearly an equality, and hence so is (3.1). It follows that using Theorem 3.2 to obtain a bound for $\angle(\widehat{x}, x)$ is a reliable approach that gives sharp bounds. The same holds for other reductions as described in Section 2.2.

Lemma 3.1 and Theorem 3.2 hold for any reduction mentioned in Section 2. In what follows we mainly treat linearizations for its practical importance and because known results for bounding $\angle(y, \widehat{y})$ are available only in that context.

3.1. Eigenvector corresponding to eigenvalue at infinity. In the above discussion we treated $\lambda_*, \widehat{\lambda}_*$ as finite scalars, but the result readily extends to the case where $\lambda_* = \infty$. Eigenvalues at ∞ and their partial multiplicities for a matrix polynomial $P(\lambda)$ of degree k are defined as those of the eigenvalues of the reversal $\lambda^k P(1/\lambda)$. Similarly, for rational eigenvalue problems (2.2) we can define them as the eigenvalues at infinity of a polynomialized function obtained by multiplying out the denominators. In this case y is still in the form (1.5), but by normalizing such that $\|y\| = \|\Lambda(\infty)\| = 1$, all the elements of $\Lambda(\infty)$ becomes zero except the leading terms in the limit $\lambda \rightarrow \infty$. For example, in the companion linearization in the monomial basis (1.4) we have $\Lambda(\infty) = [1, 0, \dots, 0]^T$. No other complication arises for eigenvalues at infinity for the statement of Theorem 3.2.

4. Eigenvector error bound for the linearization. The results in the previous section indicate that we can obtain a bound for $\angle(\widehat{x}, x)$ in terms of $P(\lambda)$ if we have an upper bound for $\angle(\widehat{y}, y)$ in terms of the linearization $L(\lambda)$. Here we discuss deriving upper bounds for $\angle(\widehat{y}, y)$ via existing perturbation theory for eigenvectors of linear eigenvalue problems.

Since the bounds for $\angle(\widehat{y}, y)$ are generally expressed roughly in the form “(residual $\|L(\widehat{\lambda}_*)\widehat{y}\|$)/(eigenvalue separation)” as we make precise below, we first quantify the residual in terms of the linearization $\|L(\widehat{\lambda}_*)\widehat{y}\|$ with respect to the residual for the original $\|P(\widehat{\lambda}_*)\widehat{x}\|$.

4.1. Relation between residuals $\|P(\widehat{\lambda}_*)\widehat{x}\|$ and $\|L(\widehat{\lambda}_*)\widehat{y}\|$. For virtually all the linearizations (and reductions) that we have discussed, the norm of the residuals $\|P(\widehat{\lambda}_*)\widehat{x}\|$ and $\|L(\widehat{\lambda}_*)\widehat{y}\|$ have a direct connection with many choices of reductions $L(\lambda)$. Since the particular connection depends on the particular choice of the reduction $L(\lambda)$, we specify them for each example we gave in Section 2.

For many linearizations, such as the companion linearization, colleague linearization, the linearization (2.3) for rational functions, we have $L(\widehat{\lambda}_*)\widehat{y} = [(P(\widehat{\lambda}_*)\widehat{x})^T, 0, 0, \dots, 0]^T$, therefore the simplest connection between $\|P(\widehat{\lambda}_*)\widehat{x}\|$ and $\|L(\widehat{\lambda}_*)\widehat{y}\|$ holds:

$$\|L(\widehat{\lambda}_*)\widehat{y}\| = \|P(\widehat{\lambda}_*)\widehat{x}\|. \quad (4.1)$$

For the quadratification (2.4), writing $L(\lambda) = Q(\lambda)$ it can be verified that $L(\widehat{\lambda}_*)\widehat{y} = [P(\widehat{\lambda}_*)\widehat{x}]$, hence $\|L(\widehat{\lambda}_*)\widehat{y}\| = \|P(\widehat{\lambda}_*)\widehat{x}\|$. Similarly, with the linearization for the rational eigenproblem (2.3) we have $L(\widehat{\lambda}_*)\widehat{y} = [0, 0, \dots, 0]^T \otimes P(\widehat{\lambda}_*)\widehat{x}$, so again $\|L(\widehat{\lambda}_*)\widehat{y}\| = \|P(\widehat{\lambda}_*)\widehat{x}\|$.

For \mathbb{L}_1 linearizations with ansatz vector v we have [20, 32]

$$L(\lambda)(\Lambda(\lambda) \otimes x) = v \otimes P(\lambda)x$$

for any λ and x . Since $\hat{y} = \Lambda(\lambda) \otimes \hat{x}$, it immediately follows that

$$\|L(\hat{\lambda}_*)\hat{y}\| = \|v\| \|P(\hat{\lambda}_*)\hat{x}\|. \quad (4.2)$$

In summary, for each $L(\lambda)$ the residual $\|L(\hat{\lambda}_*)\hat{y}\|$ is proportional to $\|P(\hat{\lambda}_*)\hat{x}\|$ and can be computed without forming $L(\hat{\lambda}_*)\hat{y}$.

4.1.1. Bounding $\angle(\hat{y}, y)$. We now discuss how to obtain a bound for $\angle(\hat{y}, y)$. The derivation is a straightforward application of the existing theory, but we include it for completeness and to shed light on when an eigenvector of $P(\lambda)$ is difficult to compute accurately.

THEOREM 4.1. *Let $P(\lambda)$ be a matrix polynomial and $(\hat{\lambda}_*, \hat{x})$ its approximate eigenpair with residual $P(\hat{\lambda}_*)\hat{x} \neq 0$. Suppose $L(\lambda) = \lambda M + N$ is a linearization of $P(\lambda)$ with eigenvector structure (1.6), and $L(\lambda)$ has a generalized Schur form [11, Ch. 7.7.2]*

$$Q^*MZ = \begin{bmatrix} \alpha & v_1^* \\ 0 & M_1 \end{bmatrix}, \quad Q^*NZ = \begin{bmatrix} \beta & v_2^* \\ 0 & N_1 \end{bmatrix}, \quad \lambda_* = -\frac{\beta}{\alpha}, \quad (4.3)$$

in which M_1, N_1 are both upper-triangular and $P(\lambda_*)x = 0$ with $x \neq 0$. Denote $Q = [q_1 \ Q_2]$, $Z = [y \ Z_2]$ where $y = \Lambda(\lambda_*) \otimes x$, $\hat{y} = \Lambda(\hat{\lambda}_*) \otimes \hat{x}$ as in (1.5). Then we have

$$\sin \angle(\hat{x}, x) \leq \sin \angle(\hat{y}, y) \leq \frac{\|L(\hat{\lambda}_*)\hat{y}\|_2}{\text{sep}(\hat{\lambda}_*, (N_1, -M_1))\|\hat{y}\|_2}, \quad (4.4)$$

where $\text{sep}(\hat{\lambda}_*, (N_1, -M_1)) = \|(\hat{\lambda}_*M_1 + N_1)^{-1}\|_2 = \sigma_{\min}((\hat{\lambda}_*M_1 + N_1)^{-1})$.

Proof. Following [16, proof of Lem. 2.2] we have

$$\begin{aligned} Q_2^*L(\hat{\lambda}_*)\hat{y} &= Q_2^*(\hat{\lambda}_*M + N)\hat{y} \\ &= Q_2^*Q \left(\hat{\lambda}_* \begin{bmatrix} \alpha & v_1^* \\ 0 & M_1 \end{bmatrix} + \begin{bmatrix} \beta & v_2^* \\ 0 & N_1 \end{bmatrix} \right) Z^*\hat{y} \\ &= \begin{bmatrix} 0 & I_{kn-1} \end{bmatrix} \left(\hat{\lambda}_* \begin{bmatrix} \alpha & v_1^* \\ 0 & M_1 \end{bmatrix} + \begin{bmatrix} \beta & v_2^* \\ 0 & N_1 \end{bmatrix} \right) \begin{bmatrix} y^* \\ Z_2^* \end{bmatrix} \hat{y} \\ &= (\hat{\lambda}_*M_1 + N_1)Z_2^*\hat{y}. \end{aligned}$$

Hence, using $\|Q_2^*L(\hat{\lambda}_*)\hat{y}\|_2 \leq \|L(\hat{\lambda}_*)\hat{y}\|_2$ we obtain

$$\|Z_2^*\hat{y}\|_2 \leq \|(\hat{\lambda}_*M_1 + N_1)^{-1}\|_2 \|L(\hat{\lambda}_*)\hat{y}\|_2 \leq \|(\hat{\lambda}_*M_1 + N_1)^{-1}\|_2 \|L(\hat{\lambda}_*)\hat{y}\|_2.$$

Since $\|Z_2^*\hat{y}\|_2 / \|\hat{y}\|_2 = \sin \angle(y, \hat{y})$, we conclude that

$$\sin \angle(y, \hat{y}) \leq \|(\hat{\lambda}_*M_1 + N_1)^{-1}\|_2 \frac{\|L(\hat{\lambda}_*)\hat{y}\|_2}{\|\hat{y}\|_2} = \frac{\|L(\hat{\lambda}_*)\hat{y}\|_2}{\text{sep}(\hat{\lambda}_*, (N_1, -M_1))\|\hat{y}\|_2}, \quad (4.5)$$

completing the proof. \square

We note that $\text{sep}(\widehat{\lambda}_*, (N_1, -M_1)) = \|(\widehat{\lambda}_* M_1 + N_1)^{-1}\|_2^{-1}$ is a quantity that roughly measures the separation between the eigenvalue $\widehat{\lambda}_*$ and those of the pencil $\lambda M_1 + N_1$ (the eigenvalues of $L(\lambda)$ and $P(\lambda)$ other than $\lambda_* = -\beta/\alpha$), as commonly used in the literature for eigenvector perturbation theory, see for example [26, 27]. In the special case where $\lambda M + N$ is a standard symmetric eigenproblem $\lambda I + N$ with $N = N^T$, sep reduces to the difference between $\widehat{\lambda}_*$ and $\lambda I + N_1$, known as the gap.

Two remarks regarding Theorem 4.1 are in order.

- The theorem indicates that an approximate eigenvector \widehat{x} is accurate if two conditions are satisfied:
 1. The residual $\|P(\widehat{\lambda}_*)\widehat{x}\|_2$ is small, and hence so is $\|L(\widehat{\lambda}_*)\widehat{y}\|_2$.
 2. $\widehat{\lambda}_*$ is well separated from the rest of the eigenvalues of $P(\lambda)$.

Note that these are precisely the conditions required for an approximate eigenvector to be accurate in a standard linear eigenproblem, see e.g. [27, Ch. 4].

- The theorem holds regardless of which generalized Schur form is chosen, and indeed the top-left corner can be any eigenvalue of $P(\lambda)$, not necessarily the $\lambda_* = -\alpha/\beta$ closest to $\widehat{\lambda}_*$. Of course, useful information is obtained only when the eigenvalue $-\alpha/\beta$ is the eigenvalue close to $\widehat{\lambda}_*$, and if one starts with a “wrong” generalized Schur form then (4.5) still gives a bound on $\sin \angle(y, \widehat{y})$, but then the denominator $\text{sep}(\widehat{\lambda}_*, (N_1, -M_1))$ in (4.4) is small, so the bound is likely to be larger than 1, a useless bound.

For ease of reference, we state the implication of Theorem 4.1 when applied to companion and \mathbb{L}_1 linearizations of a matrix polynomial $P(\lambda)$, for which the bounds can be stated purely in terms of $P(\lambda)$.

COROLLARY 4.2. *Under the assumptions in Theorem 4.1,*

1. *When $L(\lambda) = \lambda M + N$ is the companion linearization,*

$$\sin \angle(\widehat{x}, x) \leq \frac{\|P(\widehat{\lambda}_*)\widehat{x}\|_2}{\|\Lambda(\widehat{\lambda}_*)\|_2 \text{sep}(\widehat{\lambda}_*, (N_1, -M_1))}. \quad (4.6)$$

2. *When $L(\lambda) = \lambda M + N$ is a \mathbb{L}_1 linearizations with ansatz vector v ,*

$$\sin \angle(\widehat{x}, x) \leq \frac{\|v\|_2 \|P(\widehat{\lambda}_*)\widehat{x}\|_2}{\|\Lambda(\widehat{\lambda}_*)\|_2 \text{sep}(\widehat{\lambda}_*, (N_1, -M_1))}. \quad (4.7)$$

In both (4.6) and (4.7), $\Lambda(\widehat{\lambda}_*) = \Lambda(\lambda) = [\widehat{\lambda}_*^{k-1}, \widehat{\lambda}_*^{k-2}, \dots, \widehat{\lambda}_*, 1]^T$ as in (1.6) with monomials, hence $\|\Lambda(\widehat{\lambda}_*)\|_2 = \sqrt{1 + \widehat{\lambda}_*^2 + \dots + \widehat{\lambda}_*^{2(k-1)}}$.

Note that (4.7) appears to be proportional to $\|v\|_2$, and in particular taking $v \leftarrow cv$ for a scalar $c > 0$ appears to give a different bound. The explanation is that $\text{sep}(\widehat{\lambda}_*, (N_1, -M_1))$ also gets multiplied by c .

5. Eigenvector perturbation bounds. We now turn our attention to a different but closely related problem: suppose (λ_*, x) is an exact eigenpair of $P(\lambda)$. Can we bound the angular perturbation $\angle(x, \widehat{x})$ where \widehat{x} is an exact eigenvector of a perturbed matrix polynomial $\widetilde{P}(\lambda) = P(\lambda) + E(\lambda)$ of the same degree as $P(\lambda)$, where $E(\lambda)$ is small?

We note that in the limit $E \rightarrow 0$ the problem reduces to conditioning analysis of the eigenvector x , which has been studied previously for matrix polynomials [4, 7].

Here the focus is to obtain rigorous bounds on $\angle(x, \tilde{x})$ when $E(\lambda)$ is small but not infinitesimally so.

The basic idea is to follow the argument of Stewart [26, § 4.2.3], which is to view (λ_*, x) as an approximate eigenpair of $\tilde{P}(\lambda)$ with corresponding residual $\tilde{P}(\lambda_*)x$, and invoke the residual error bounds in the previous sections to give bounds on $\angle(x, \tilde{x})$. Note that the residuals are related by

$$\tilde{P}(\lambda_*)x = (P(\lambda_*) + E(\lambda_*))x = E(\lambda_*)x,$$

Hence

$$\|\tilde{P}(\lambda_*)x\|_2 = \|E(\lambda_*)x\|_2 \leq \|E(\lambda_*)\|_2,$$

where we suppose x is normalized so that $\|x\|_2 = 1$. Therefore it follows that we can directly invoke the previous results, such as Theorem 4.1, with residual $\|E(\lambda_*)\|_2$. Recall that the residuals $\|L(\hat{\lambda}_*)\hat{y}\|$ and $\|v\|P(\hat{\lambda}_*)\hat{x}\|$ are usually related as in (4.1) and (4.2). Below we assume $L(\lambda)$ and $\tilde{L}(\lambda)$ are linearizations, obtained via the same linearizations process (e.g., companion or \mathbb{DL} with the same ansatz vector) applied to $P(\lambda)$ and $\tilde{P}(\lambda)$ respectively. We summarize these and derive two bounds for the perturbation in x in the next theorem.

THEOREM 5.1. *Let $P(\lambda), E(\lambda)$ be matrix polynomials of the same degrees and let (λ, x) be an eigenpair such that $P(\lambda)x = 0$. Let $L(\lambda) = \lambda M + N$ and $\tilde{L}(\lambda) = \lambda \tilde{M} + \tilde{N} = \lambda(M + \Delta M) + (N + \Delta N)$ be linearizations of $P(\lambda)$ and $\tilde{P}(\lambda) = P(\lambda) + E(\lambda)$ with the same Vandermonde eigenvector structures $y = \Lambda(\lambda) \otimes x$ and $\tilde{y} = \Lambda(\tilde{\lambda}) \otimes \tilde{x}$ as in (1.5), and a generalized Schur form of $L(\lambda)$ with the eigenvalue λ_* located at the top is*

$$QMZ = \begin{bmatrix} \alpha & v_1^* \\ 0 & M_1 \end{bmatrix}, \quad QNZ = \begin{bmatrix} \beta & v_2^* \\ 0 & N_1 \end{bmatrix}, \quad \lambda_* = -\beta/\alpha. \quad (5.1)$$

Similarly, let a generalized Schur form of $\lambda \tilde{M} + \tilde{N}$ be

$$\tilde{Q}\tilde{M}\tilde{Z} = \begin{bmatrix} \tilde{\alpha} & \tilde{v}_1^* \\ 0 & \tilde{M}_1 \end{bmatrix}, \quad \tilde{Q}\tilde{N}\tilde{Z} = \begin{bmatrix} \tilde{\beta} & \tilde{v}_2^* \\ 0 & \tilde{N}_1 \end{bmatrix}, \quad \tilde{\lambda} = -\tilde{\beta}/\tilde{\alpha}. \quad (5.2)$$

Then the eigenpair $(\tilde{\lambda}, \tilde{x})$ of $\tilde{P}(\lambda)$ satisfies

$$\angle(\tilde{x}, x) \leq \angle(\tilde{y}, y), \quad (5.3)$$

and $\sin \angle(\tilde{x}, x)$ is bounded by

$$\sin \angle(x, \tilde{x}) \leq \frac{\|(\lambda_* \tilde{M} + \tilde{N})y\|_2}{\text{sep}(\lambda_*, (\tilde{N}_1, -\tilde{M}_1))\|y\|_2}, \quad (5.4)$$

and

$$\sin \angle(x, \tilde{x}) \leq \frac{\|(\lambda_* \Delta M + \Delta N)\|_2 + \|(\lambda_* - \tilde{\lambda})\tilde{M}\|_2}{\text{sep}(\lambda_*, (N_1, -M_1))}. \quad (5.5)$$

Proof. The inequality (5.3) is a direct consequence of Lemma 3.1, applied with $\hat{\lambda}_* \leftarrow \tilde{\lambda}, \hat{x} \leftarrow \tilde{x}$.

To obtain (5.4) and (5.5) we shall derive a bound for $\angle(\tilde{y}, y)$, which by (5.3) is also a bound for $\angle(\tilde{x}, x)$. For (5.4), by the same argument as Theorem 4.1 with $P \leftarrow \tilde{P}$,

$$\begin{aligned}\tilde{Q}_2^*(\lambda_*\tilde{M} + \tilde{N})y &= \tilde{Q}_2^*\tilde{Q}\left(\lambda_*\begin{bmatrix}\tilde{\alpha} & \tilde{v}_1^* \\ 0 & \tilde{M}_1\end{bmatrix} + \begin{bmatrix}\tilde{\beta} & \tilde{v}_2^* \\ 0 & \tilde{N}_1\end{bmatrix}\right)\tilde{Z}^*y \\ &= [0 \quad I_{kn-1}]\left(\lambda_*\begin{bmatrix}\tilde{\alpha} & \tilde{v}_1^* \\ 0 & \tilde{M}_1\end{bmatrix} + \begin{bmatrix}\tilde{\beta} & \tilde{v}_2^* \\ 0 & \tilde{N}_1\end{bmatrix}\right)\begin{bmatrix}\tilde{y}^* \\ \tilde{Z}_2^*\end{bmatrix}y \\ &= (\lambda_*\tilde{M}_1 + \tilde{N}_1)\tilde{Z}_2^*y.\end{aligned}$$

Since $\sin \angle(y, \tilde{y}) = \|\tilde{Z}_2^*y\|_2/\|y\|_2$, we conclude that

$$\begin{aligned}\sin \angle(y, \tilde{y}) &\leq \|(\lambda_*\tilde{M}_1 + \tilde{N}_1)^{-1}\|_2\|(\lambda_*\tilde{M} + \tilde{N})y\|_2/\|y\|_2 \\ &= \frac{\|(\lambda_*\tilde{M} + \tilde{N})y\|_2}{\text{sep}(\lambda_*, (\tilde{N}_1, -\tilde{M}_1))\|y\|_2},\end{aligned}$$

which is (5.4).

To obtain (5.5), we have

$$\begin{aligned}Q_2^*(\lambda_*\tilde{M} + \tilde{N})\tilde{y} &= Q_2^*(\lambda_*M + N + (\lambda_*\Delta M + \Delta N))\tilde{y} \\ &= Q_2^*Q\left(\lambda_*\begin{bmatrix}\alpha & v_1^* \\ 0 & M_1\end{bmatrix} + \begin{bmatrix}\beta & v_2^* \\ 0 & N_1\end{bmatrix}\right)Z^*\tilde{y} + Q_2^*(\lambda_*\Delta M + \Delta N)\tilde{y} \\ &= [0 \quad I_{kn-1}]\left(\lambda_*\begin{bmatrix}\alpha & v_1^* \\ 0 & M_1\end{bmatrix} + \begin{bmatrix}\beta & v_2^* \\ 0 & N_1\end{bmatrix}\right)\begin{bmatrix}y^* \\ Z_2^*\end{bmatrix}\tilde{y} + Q_2^*(\lambda_*\Delta M + \Delta N)\tilde{y} \\ &= (\lambda_*M_1 + N_1)Z_2^*\tilde{y} + Q_2^*(\lambda_*\Delta M + \Delta N)\tilde{y}.\end{aligned}$$

Hence

$$\|Z_2^*\tilde{y}\|_2 \leq \|(\lambda_*M_1 + N_1)^{-1}\|_2(\|(\lambda_*\Delta M + \Delta N)\tilde{y}\|_2 + \|(\lambda_*\tilde{M} + \tilde{N})\tilde{y}\|_2).$$

Since $\sin \angle(y, \tilde{y}) = \|Z_2^*\tilde{y}\|_2/\|\tilde{y}\|_2$, we conclude that

$$\begin{aligned}\sin \angle(y, \tilde{y}) &\leq \frac{\|(\lambda_*M_1 + N_1)^{-1}\|_2(\|(\lambda_*\Delta M + \Delta N)\tilde{y}\|_2 + \|(\lambda_*\tilde{M} + \tilde{N})\tilde{y}\|_2)}{\|\tilde{y}\|_2} \\ &= \frac{\|(\lambda_*\Delta M + \Delta N)\tilde{y}\|_2 + \|(\lambda_*\tilde{M} + \tilde{N})\tilde{y}\|_2}{\text{sep}(\lambda_*, (N_1, -M_1))\|\tilde{y}\|_2}.\end{aligned}$$

Now since $(\tilde{\lambda}\tilde{M} + \tilde{N})\tilde{y} = 0$, we have $\|(\lambda_*\tilde{M} + \tilde{N})\tilde{y}\|_2 = \|(\lambda_* - \tilde{\lambda})\tilde{M}\tilde{y}\|_2$. Hence

$$\begin{aligned}\sin \angle(y, \tilde{y}) &\leq \frac{\|(\lambda_*\Delta M + \Delta N)\tilde{y}\|_2 + \|(\lambda_* - \tilde{\lambda})\tilde{M}\tilde{y}\|_2}{\text{sep}(\lambda_*, (N_1, -M_1))\|\tilde{y}\|_2} \\ &\leq \frac{\|(\lambda_*\Delta M + \Delta N)\|_2 + \|(\lambda_* - \tilde{\lambda})\tilde{M}\|_2}{\text{sep}(\lambda_*, (N_1, -M_1))},\end{aligned}$$

as required. \square

Between the two bounds (5.4) and (5.5), (5.4) looks simpler, but (5.4) involves $\text{sep}(\lambda_*, (\widetilde{M}_1, -\widetilde{N}_1))$, the separation of λ_* from the spectrum of $\widetilde{M}_1 + \widetilde{N}_1$, which is a quantity in terms of $\widetilde{P}(\lambda)$. The separation with respect to $\widetilde{P}(\lambda)$ is nontrivial to bound (see [26, eq. (4.2.29)], [27, Thm. VI.2.13] for related discussions). By contrast, the sep in (5.5) is defined for $\lambda M_1 + N_1$ of the original $P(\lambda)$, and although the bound involves the perturbation in the eigenvalue $\lambda_* - \widetilde{\lambda}$, this can be bounded or estimated using existing results on linear eigenvalue perturbation.

Sometimes one might want to obtain perturbation bounds without specifying E other than a rough estimate such as the norm $\|E\|$. Such bounds are treated in the next subsection.

We note that the bounds (5.5) and (5.4) are not necessarily attainable: in particular, the bounds depend on the choice of the linearization used to obtain bounds on $\angle(x, \widetilde{x})$. One approach would be to try many linearizations and take the smallest bound.

For simplicity we have focused on the case where $P(\lambda)$ and $P(\lambda) + E(\lambda)$ are matrix polynomials, but it is straightforward to extend the argument to rational eigenproblems as in Section 2.1, as long as the original and perturbed problems are of the same forms so that their linearization are of the same size.

5.1. Error bounds without \widetilde{y} or sep in $\widetilde{P}(\lambda)$. As indicated by (5.3), any bound for $\angle(y, \widetilde{y})$ of the linear matrix pencils $L(\lambda)$ and $L(\widetilde{\lambda})$ is a bound for the desired quantity $\angle(x, \widetilde{x})$. A number of results are available for bounding the eigenvector perturbation of linear eigenvalue problems in the literature, and below we treat the case where $L(\lambda)$ is a monic linearization $L(\lambda) = \lambda I + N$ (this restriction is due solely to apparent lack of results in the literature for bounding the angle of the eigenvector perturbation for generalized eigenproblems that are straightforward to apply in our context; for example, the bounds in [27] examine the perturbation in all eigenvectors, instead of a specified one x). The result applies for example when a companion linearization is used for a monic matrix polynomial.

THEOREM 5.2. *Let $P(\lambda), E(\lambda)$ be matrix polynomials of the same degrees and let (λ, x) be an eigenpair such that $P(\lambda)x = 0$. Let $L(\lambda) = \lambda I + N$ and $\widetilde{L}(\lambda) = \lambda I + \widetilde{N}$ be linearizations of $P(\lambda)$ and $\widetilde{P}(\lambda) = P(\lambda) + E(\lambda)$ respectively, both with Vandermonde eigenvector structure (1.5). Let y, z be the right and left eigenvectors of N corresponding to $-\lambda_*$ such that $Ny = -\lambda_*y$, $z^T N = -\lambda_*z^T$, and let N have spectral decomposition $N = -\lambda_*yz^* - Y_2N_2Z_2^*$ such that*

$$\begin{bmatrix} z^* \\ Y_2^* \end{bmatrix} N \begin{bmatrix} y \\ Y_2 \end{bmatrix} = - \begin{bmatrix} \lambda_* & \\ & N_2 \end{bmatrix}, \quad \begin{bmatrix} z^* \\ Y_2^* \end{bmatrix} \begin{bmatrix} y \\ Y_2 \end{bmatrix} \begin{bmatrix} z^* \\ Y_2^* \end{bmatrix} = I.$$

Define $\theta = \arccos \frac{y^T z}{\|y\| \|z\|}$. Then there exists an exact eigenpair $(\widetilde{\lambda}, \widetilde{x})$ of $\widetilde{P}(\lambda)$ such that

$$\tan \angle(x, \widetilde{x}) \leq \tan \angle(y, \widetilde{y}) \leq \frac{\|N - \widetilde{N}\|_2}{\text{sep}(\lambda_*, (-N_2, I)) - 2(\tan \theta + \sec \theta) \|N - \widetilde{N}\|}, \quad (5.6)$$

where $\text{sep}(\lambda_*, (-N_2, I)) = \|(\lambda_* I + N_2)^{-1}\|_2^{-1}$.

Proof. The first inequality is (5.3), and the second is a direct consequence of [26, Cor. 2.14] applied to the matrix N and its perturbation \widetilde{N} . \square

Note that the bound (5.6) is defined purely in terms of the original $P(\lambda)$ and the norm of the perturbation, unlike those in Theorem 5.1.

6. Angles between subspaces. When λ_* is a multiple eigenvalue of $L(\lambda)$ (and hence of $P(\lambda)$) or λ_* belongs to a cluster of eigenvalues, the corresponding eigenvector y of $L(\lambda)$ is known [26, § 4.2] to be sensitive to perturbation in the coefficient matrices, and hence so is x , the corresponding eigenvector of $P(\lambda)$. This generally makes it difficult or impossible to numerically compute x or y accurately. This difficulty is present even in the simplest case of standard symmetric eigenvalue problems, and manifests itself as a small denominator sep in (5.4). A well known remedy to such difficulty is to instead compute a subspace corresponding to the cluster. Hence it is of interest to compare two subspaces spanned by the columns of $n \times k$ matrices X and \hat{X} , where X contains a set of exact eigenvectors and \hat{X} contains approximate eigenvectors of $P(\lambda)$. As in the previous sections, we wish to relate $\angle(\hat{X}, X)$ to $\angle(\hat{Y}, Y)$, the canonical angles between approximate and exact eigenspaces of a linearization $\lambda M + N$.

Nonetheless, when a target eigenpair is exactly multiple a direct analogue of Lemma 3.1 holds:

LEMMA 6.1. *Let the columns of $X \in \mathbb{R}^{n \times \ell}$ form an orthonormal subspace corresponding to a multiple eigenvalue λ_* . Then for an approximate set of eigenpairs $(\hat{\lambda}_*, \hat{X})$ where $\hat{X} \in \mathbb{R}^{n \times \ell}$ is orthonormal, taking $\Lambda = \lambda_* I_\ell$ and $\hat{\Lambda} = \hat{\lambda}_* I_\ell$ we have*

$$\angle_i(\hat{X}, X) \leq \angle_i(\hat{Y}, Y), \quad i = 1, \dots, \ell. \quad (6.1)$$

where

$$Y = \begin{bmatrix} X\phi_{k-1}(\Lambda) \\ \vdots \\ X\phi_1(\Lambda) \\ X\phi_0(\Lambda) \end{bmatrix}, \quad \hat{Y} = \begin{bmatrix} \hat{X}\phi_{k-1}(\hat{\Lambda}) \\ \vdots \\ \hat{X}\phi_1(\hat{\Lambda}) \\ \hat{X}\phi_0(\hat{\Lambda}) \end{bmatrix} \quad (6.2)$$

Proof. Recall that the angles between two subspaces X_1, X_2 are defined by the singular values of $\text{orth}(X_1)^T \text{orth}(X_2)$, where orth denotes the orthonormal column space (e.g., the Q factor in the QR factorization).

By the assumptions, we have

$$\text{orth}(Y) = \frac{1}{\sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda)^2}} Y, \quad \text{orth}(\hat{Y}) = \frac{1}{\sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\hat{\lambda}_*)^2}} \hat{Y}.$$

Therefore we have

$$\begin{aligned} \text{orth}(Y)^T \text{orth}(\hat{Y}) &= \frac{Y^T \hat{Y}}{\sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda_*)^2} \sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\hat{\lambda}_*)^2}} \\ &= \frac{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda_*) \phi_{k-1}(\hat{\lambda}_*)}{\sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda_*)^2} \sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\hat{\lambda}_*)^2}} X^T \hat{X}. \end{aligned}$$

Now by Cauchy-Schwarz we have

$$\frac{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda_*) \phi_{k-1}(\hat{\lambda}_*)}{\sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\lambda_*)^2} \sqrt{\sum_{i=0}^{k-1} \phi_{k-1}(\hat{\lambda}_*)^2}} \leq 1,$$

so it follows that $\sigma_i(\text{orth}(Y)^T \text{orth}(\widehat{Y})) \leq \sigma_i(X^T \widehat{X})$ for all i . This implies $\angle_i(\widehat{Y}, Y) \geq \angle_i(\widehat{X}, X)$, as required. \square

THEOREM 6.2. *Let $P(\lambda) = \sum_{i=0}^k \phi_i(\lambda)A_i \in \mathbb{C}[\lambda]^{n \times n}$ and $\lambda M + N$ be its linearization with Vandermonde eigenvector structure (1.6). Let the columns of $\widehat{X} \in \mathbb{C}^{n \times \ell}$ be approximations to the eigenvectors corresponding to an eigenpair (λ, X) of multiplicity ℓ , with approximate eigenvalue $\widehat{\lambda}_*$. Define Y, \widehat{Y} as in Lemma 6.1, so that $(\widehat{\Lambda}, \widehat{Y})$ and (Λ, Y) are approximate and exact sets of ℓ eigenpairs of $L(\lambda)$. Let a generalized Schur form [11, § 7.7.2] of $L(\lambda) = \lambda M + N$ be*

$$Q^*MZ = \begin{bmatrix} M_0 & V_1^* \\ 0 & M_1 \end{bmatrix}, \quad Q^*NZ = \begin{bmatrix} N_0 & V_2^* \\ 0 & N_1 \end{bmatrix}, \quad (6.3)$$

in which the eigenvalues of $\lambda M_0 + N_0$ are all λ_* , and $M_1, N_1 \in \mathbb{C}^{(nk-\ell) \times (nk-\ell)}$. Denote $Q = [Q_1 \ Q_2]$, $Z = [Y \ Y_2]$ where $Y = \Lambda(\lambda I_\ell) \otimes X$ as in (6.2). Then we have

$$\sin \angle_i(\widehat{X}, X) \leq \sin \angle_i(\widehat{Y}, Y) \leq \frac{\|L(\widehat{\lambda}_*)\widehat{Y}\|_2}{\text{sep}(\widehat{\lambda}_*, (N_1, -M_1))}, \quad (6.4)$$

where $\text{sep}(\widehat{\lambda}_*, (N_1, -M_1)) = \|(\widehat{\lambda}_* M_1 + N_1)^{-1}\|_2 = \sigma_{\min}((\widehat{\lambda}_* M_1 + N_1)^{-1})$.

Proof. The proof is much the same as in Theorem 4.1. Using the block residual $L(\widehat{\lambda}_*)\widehat{Y}$ we have

$$\begin{aligned} Q_2^*L(\widehat{\lambda}_*)\widehat{Y} &= Q_2^*(\widehat{\lambda}_* M + N)\widehat{Y} \\ &= Q_2^*Q \left(\widehat{\lambda}_* \begin{bmatrix} M_0 & V_1^* \\ 0 & M_1 \end{bmatrix} + \begin{bmatrix} N_0 & V_2^* \\ 0 & N_1 \end{bmatrix} \right) Z^* \widehat{Y} \\ &= \begin{bmatrix} 0 & I_{kn-\ell} \end{bmatrix} \left(\widehat{\lambda}_* \begin{bmatrix} M_0 & V_1^* \\ 0 & M_1 \end{bmatrix} + \begin{bmatrix} N_0 & V_2^* \\ 0 & N_1 \end{bmatrix} \right) \begin{bmatrix} y^* \\ Z_2^* \end{bmatrix} \widehat{y} \\ &= (\widehat{\lambda}_* M_1 + N_1) Z_2^* \widehat{Y}. \end{aligned}$$

Hence

$$\|Z_2^* \widehat{Y}\|_2 \leq \|(\widehat{\lambda}_* M_1 + N_1)^{-1}\|_2 \|Q_2^*L(\widehat{\lambda}_*)\widehat{Y}\|_2 \leq \|(\widehat{\lambda}_* M_1 + N_1)^{-1}\|_2 \|L(\widehat{\lambda}_*)\widehat{Y}\|_2.$$

Since $\sigma_i(Z_2^* \widehat{Y}) = \sin \angle_i(Y, \widehat{Y})$ for $i = 1, \dots, \ell$, we conclude that

$$\sin \angle_i(Y, \widehat{Y}) \leq \|(\widehat{\lambda}_* M_1 + N_1)^{-1}\|_2 \|L(\widehat{\lambda}_*)\widehat{Y}\|_2 = \frac{\|L(\widehat{\lambda}_*)\widehat{Y}\|_2}{\text{sep}(\widehat{\lambda}_*, (N_1, -M_1))}. \quad (6.5)$$

\square

It is worth noting that the results of this subsection are not as general as one might like. Most importantly, they assume that λ is an exactly multiple (nondefective) eigenvalue, excluding the case where λ is simple but belongs to a cluster. Unfortunately, we are currently unable to derive bounds (e.g., an analogue of Lemma 3.1) that allow for such situations.

Another restriction is that the results here assume that the computed approximate eigenvalue $\widehat{\lambda}_*$ is also multiple, although this is a somewhat minor assumption as we can perturb the approximants $(\widehat{\lambda}_i, \widehat{x}_i)$ so that $\widehat{\lambda}_i = \widehat{\lambda}_*$ for $\widehat{\lambda}_i$ corresponding to a cluster.

7. Discussion. Several remarks are in order regarding the eigenvector error and perturbation bounds we obtained.

- It is worth recalling that matrix polynomials $P(\lambda)$ can have distinct eigenvalues λ_* and μ belonging to a single eigenvector x , that is, $P(\lambda_*)x = P(\mu)x$. One might imagine that this may cause complications in the behavior of λ_* , μ and x , in particular the accuracy of their approximation. However, our analysis above makes no special reference to μ when assessing the accuracy of (λ_*, x) . The (perhaps surprising) implication is that as far as the eigenpair (λ_*, x) is concerned, the presence of an eigenvalue μ sharing the same eigenvector, or even the sensitivity of the pair (μ, x) , plays no role. It may well be that (λ_*, x) is insensitive to perturbation and so can be computed accurately, while μ belongs to a cluster of eigenvalues and so (μ, x) is highly sensitive to perturbations.

On the other hand, when (μ, y) and (λ_*, x) are eigenpairs with $\mu \approx \lambda_*$ and $\angle(x, y) \simeq 0$, that is, if both the eigenvalues and eigenvector are close, the corresponding eigenvectors in terms of the linearization are also nearly parallel, so computing them numerically necessarily becomes a challenge.

- Kressner and Betcke [5, 18] deal with invariant pairs instead of a subspace spanned by eigenvectors for nonlinear eigenvalue problems (including matrix polynomials). Szyld and Xue [29] develop first-order perturbation of invariant pairs. One advantage of invariant pairs is that they are distinct for eigenpairs (λ_1, x) , (λ_2, x) that share an eigenvector. Here we took the more classical approach of subspaces spanned by eigenvectors, given that two distinct eigenvalues having the same eigenvector does not seem to cause any difficulty as we just discussed.
- In practice a common way of obtaining \hat{x} is to solve a linearized eigenproblem and take a $n \times 1$ part of the $nk \times 1$ computed eigenvector \hat{y} of the linearization $L(\lambda)$. Note that the above result does not mean the accuracy of the computed eigenvector \hat{x} of $P(\lambda)$ is at least as good as that of \check{y} of $L(\lambda)$ (note that \check{y} is different from both \hat{y} and \tilde{y}), because \check{y} is generally not in the form of \hat{y} as in (1.4). In particular, the residual $\|L(\hat{\lambda}_*)\check{y}\|_2$ of $(\hat{\lambda}_*, \check{y})$ as an approximate eigenpair of the linearization can be smaller than $\|L(\hat{\lambda}_*)\hat{x}\|_2$.
- From a practical viewpoint, the bounds derived here depend on the choice of linearization $L(\lambda)$. It would be of interest to examine finding an optimal choice of linearization in terms of minimizing the bounds, e.g. in Theorem 4.1. Generally, given a polynomial eigenproblem, there is a plethora of linearizations available today, and the choice of linearization depends on many aspects, including efficiency and eigenvalue conditioning [15].

8. Numerical experiments. We present numerical experiments to illustrate our results. All experiments were performed in MATLAB version R2013a on a desktop machine with an Intel Xeon 3.20GHz Processor with four cores, and 16GB RAM, using IEEE double precision arithmetic.

8.1. $\angle(x, \hat{x})$ and $\angle(y, \hat{y})$. To illustrate our error bounds, we generate 10 sets of 10×10 quadratic matrix polynomials $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$, and obtained the nk approximate eigenpairs $(\hat{\lambda}_*, \hat{x})$ computed by MATLAB's `polyeig`, and `quadeig` developed in [13]. We then form $\hat{y} = \begin{bmatrix} \hat{\lambda}_* \hat{x} \\ \hat{x} \end{bmatrix}$ as in (1.4).

We then obtained the “exact” eigenpairs (λ_*, x) by computing the eigenvalue decomposition of the companion form C of $P(\lambda)$ in variable precision arithmetic using MATLAB’s `vpa` command. From the eigenpairs (λ_*, y) of C we obtain x as the larger part (in norm) of the two $n \times 1$ blocks of y ; this selection is made to minimize the effect of numerical errors.

Randomly generating the coefficient matrices A_i gives computed results that have accuracy in the order of machine precision, so in order to examine the results also with difficult cases we also formed a set of 10×10 quadratic matrix polynomials that have near-multiple eigenvalues with clusters of seven eigenvalues, and also containing $\lambda_* = \infty$.

Figure 8.1 shows the scatterplot of $\angle(x, \hat{x})$ and $\angle(y, \hat{y})$, along with the boundary of $\angle(x, \hat{x}) \leq \angle(y, \hat{y})$ in Lemma 3.1. The left plot is for randomly generated A_i , and the right shows the results with difficult matrices. Observe that scatterplot lies just below the dashed line, illustrating that $\angle(x, \hat{x}) \leq \angle(y, \hat{y})$, and that this bound is often a remarkably sharp bound. This indicates that assessing the error in \hat{x} by that of \hat{y} is reliable and does not result in severe overestimation. We observed the same behavior for matrix polynomials of higher degree $k \geq 3$.

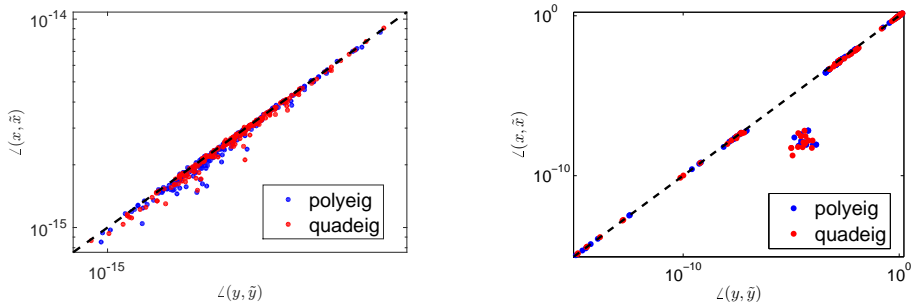


FIG. 8.1. $\angle(x, \hat{x})$ vs. $\angle(y, \hat{y})$ log-log plot for 10 sets of quadratic eigenvalue problems, random-easy case (left) and difficult case with multiple and infinite eigenvalues (right). The dashed black line shows $\angle(x, \hat{x}) = \angle(y, \hat{y})$.

Shared eigenvalue or eigenvector. In Section 7 we noted that the possible presence of a (nearly) shared eigenvector should not impair the accuracy of the computed eigenpair $(\hat{\lambda}_*, \hat{x})$, while a shared eigenvalue does indicate an ill-conditioned eigenvector. Here we verify this observation by constructing quadratic eigenvalue problems $P(\lambda)$ with prescribed eigenvalues and eigenvectors as follows.

We first form a diagonal matrix $D = \text{diag}(d_1, \dots, d_{nk})$, whose diagonal elements will be the eigenvalues of $P(\lambda)$. Then generate $X \in \mathbb{C}^{n \times nk}$, whose columns (which are clearly not linearly independent) form the eigenvectors. We then form

$$C = \begin{bmatrix} XD^{k-1} \\ \vdots \\ XD \\ X \end{bmatrix} D \begin{bmatrix} XD^{k-1} \\ \vdots \\ XD \\ X \end{bmatrix}^{-1}. \quad (8.1)$$

It can be easily verified that C must be in the companion form $C = \begin{bmatrix} -A_{k-1} & -A_{k-2} & \dots & -A_0 \\ I & & & \\ & \ddots & & \\ & & I & \end{bmatrix}$,

from which the coefficient matrices of the monic matrix polynomial $P(\lambda) = I\lambda^k + A_{k-1}\lambda^{k-1} + \dots + A_0$ can be obtained directly. $P(\lambda)$ has prescribed eigenvalues d_i and eigenvectors $X = [x_1, \dots, x_{kn}]$.

With such process for degree $k = 2$, we perform two sets of experiments:

- Shared eigenvector. We set $d_i = i$ for $i = 1, \dots, nk$ and generate X randomly except for the last column of X , which is equal to its first column. The resulting $P(\lambda)$ has distinct eigenvalues, with d_1 and d_n sharing an eigenvector $X(:, 1)$ (using MATLAB notation).
- Shared eigenvalue. We set X randomly and set $d_i = i$ for $i = 1, \dots, nk - 1$ and $d_{nk} = d_1$. The resulting $P(\lambda)$ has one non-defective multiple eigenvalue of multiplicity two.

For each set of experiments Figure 8.2 shows the scatterplots of $\angle(x, \hat{x})$ and $\angle(y, \hat{y})$. Observe that in the left plot (shared eigenvector) all the eigenvectors are computed

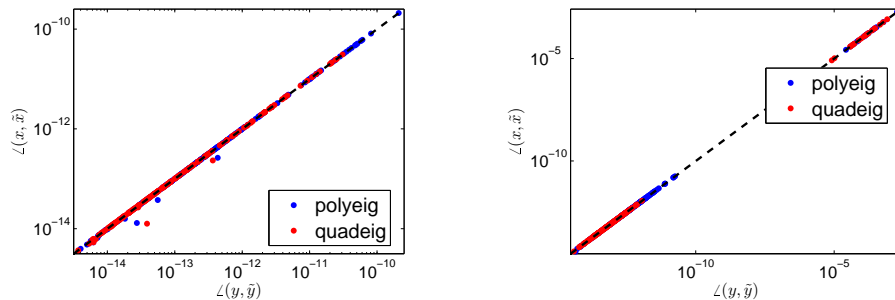


FIG. 8.2. $\angle(x, \hat{x})$ vs. $\angle(y, \hat{y})$ plots in log-log scale for 10 sets of quadratic eigenvalue problems with a shared eigenvector (left) or eigenvalue (right). The dashed black line shows $\angle(x, \hat{x}) = \angle(y, \hat{y})$.

with smaller than 10^{-10} accuracy. This is the same order of error that we get without the constraint $X(:, 1) = X(:, nk)$, verifying that a shared eigenvector does not cause numerical issues. By contrast, the right plot shows that in the presence of shared eigenvalues, some eigenvectors are not computed accurately. We confirmed that all the inaccurate eigenpairs with $\angle(x, \hat{x}) > 10^{-10}$ correspond to $(\hat{\lambda}_*, \hat{x})$ with $\hat{\lambda}_* \approx d_1 = d_n$, verifying that eigenvectors are ill-conditioned if the eigenvalue is (nearly) multiple, which is true also in linear eigenvalue problems. In any case, the bound $\angle(x, \hat{x}) \leq \angle(y, \hat{y})$ by Lemma 3.1 is always satisfied.

On a sidenote, we observe that `quadeig` tends to give more accurate results than `polyeig` (since the blue dots tend to lie above the red), reflecting the claim in [13].

8.2. Error bounds. We now examine the error bounds obtained in Theorem 4.1 and Corollary 4.2. We generate quadratic matrix polynomials as in the above example using (8.1).

To obtain the bounds for $\angle(x, \hat{x})$ we compute the Schur form of the companion matrix C , then compute the bound (4.6) for $(\hat{\lambda}_*, \hat{x})$ approximating the $2n$ eigenpairs.

Figure 8.3 shows the results for 10 sets of quadratic eigenvalue problems, again with shared eigenvector (left) and eigenvalue (right). We confirm that (4.6) is always an upper bound for $\angle(x, \hat{x})$, although they are not as tight as the bound $\angle(x, \hat{x}) \leq \angle(y, \hat{y})$ that we saw above.

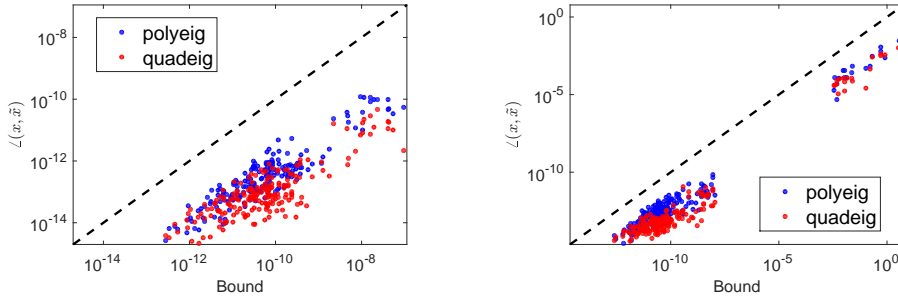


FIG. 8.3. Error bound (4.6) vs. $\angle(x, \hat{x})$ in log-log scale, quadratic $P(\lambda)$ with a shared eigenvector (left) and shared eigenvalue (right). The dashed black line shows $\angle(x, \hat{x}) = \text{Bound}$.

8.3. Perturbation bounds. To illustrate the perturbation bounds in Section 5 we test with a randomly generated quadratic matrix polynomial $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ and its perturbed variant $\hat{P}(\lambda) = \sum_{i=0}^k \lambda^i \hat{A}_i$ with $k = 2$. We then normalized each coefficient matrix A_i so that $\|A_i\|_2 = 1$ and defined $\hat{A}_i = A_i + E_i$ where E_i are randomly generated matrices. We then took x, \hat{x} to be the eigenvector of $P(\lambda), \hat{P}(\lambda)$ respectively, corresponding to the eigenvalue of largest absolute value, which were computed by `quadeig`. The vector \hat{x} is therefore an eigenvector of a perturbed matrix polynomial. We then obtain the corresponding y, \hat{y} by forming a companion linearization $C - \lambda I$ of $\hat{P}(\lambda)$, form the Schur form $C = QTQ^*$ and compute the bound (5.4) in Theorem 5.1, which in this case reduces to $\frac{\|P(\hat{\lambda}_*)\hat{x}\|_2}{\sqrt{1 + \hat{\lambda}_*^2 \|(\tilde{N}_1 - \hat{\lambda}_* I)^{-1}\|_2}}$, where N_1 is obtained from T by removing the first column and row.

We tested with eight such examples, in which in the j th example the perturbation norm $\|E_i\|_2$ was set to $\|E_i\|_2 = 10^{-j-2}$ for $j = 1, \dots, 8$. Figure 8.4 shows the log-log scatterplots of the perturbation $\angle(x, \hat{x})$ and its bound (5.4). There are eight clusters of points, and the j th cluster from the right corresponds to the j th test example. Again, we verify that the perturbation bounds correctly (and quite sharply) bound the actual perturbation in the eigenvector.

REFERENCES

- [1] A. Amiraslani, R. M. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29(1):141–157, 2009.
- [2] Z. Bai and Y. Su. SOAR: A second-order Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 26(3):640–659, 2005.
- [3] S. Barnett. *Polynomials and Linear Control Systems*. Marcel Dekker Inc., 1983.
- [4] M. Berhanu. *The Polynomial Eigenvalue Problem*. PhD thesis, University of Manchester, 2005.
- [5] T. Betcke and D. Kressner. Perturbation, extraction and refinement of invariant pairs for matrix polynomials. *Linear Algebra Appl.*, 435(3):514–536, Aug. 2011.
- [6] F. De Terán, F. M. Dopico, and D. S. Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. *Linear Algebra Appl.*, 459:264–333, 2014.
- [7] J. P. Dedieu and F. Tisseur. Perturbation theory for homogeneous polynomial eigenvalue problems. *Linear Algebra Appl.*, 358(1):71–94, 2003.

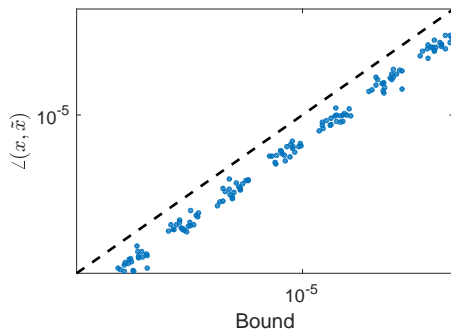


FIG. 8.4. The bound $\frac{\|P(\hat{\lambda}_*)\hat{x}\|_2}{\sqrt{1+\hat{\lambda}_*^2\|(\tilde{N}_1-\hat{\lambda}_*I)^{-1}\|_2}}$ by Theorem 5.1 vs. $\angle(x, \hat{x})$.

- [8] C. Effenberger and D. Kressner. Chebyshev interpolation for nonlinear eigenvalue problems. *BIT*, 52:933–951, 2012.
- [9] M. Fiedler. A note on companion matrices. *Linear Algebra Appl.*, 372:325–332, 2003.
- [10] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. SIAM, Philadelphia, USA, 2009. Unabridged republication of book first published by Academic Press in 1982.
- [11] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 4th edition, 2012.
- [12] L. Grammont, N. J. Higham, and F. Tisseur. A framework for analyzing nonlinear eigenproblems and parametrized linear systems. *Linear Algebra Appl.*, 435(3):623–640, 2011.
- [13] S. Hammarling, C. J. Munro, and F. Tisseur. An algorithm for the complete solution of quadratic eigenvalue problems. *ACM Trans. Math. Soft.*, 39(3):18, 2013.
- [14] N. J. Higham. *Functions of Matrices: Theory and Computation*. SIAM, Philadelphia, PA, USA, 2008.
- [15] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. 28(4):1005–1028, 2006.
- [16] T.-M. Huang, Z. Jia, and W.-W. Lin. On the Convergence of q-Ritz Pairs and Refined q-Ritz Vectors for Quadratic Eigenvalue Problems. *ArXiv e-prints*, Sept. 2011.
- [17] T.-M. Huang, W.-W. Lin, and W.-S. Su. Palindromic quadratization and structure-preserving algorithm for palindromic matrix polynomials of even degree. *Numer. Math.*, 118(4):713–735, 2011.
- [18] D. Kressner. A block Newton method for nonlinear eigenvalue problems. *Numer. Math.*, 114(2):355–372, 2009.
- [19] D. S. Mackey. The continuing influence of Fiedler’s work on companion matrices. *Linear Algebra Appl.*, 439(4):810–817, 2012.
- [20] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28:971–1004, 2005.
- [21] D. S. Mackey, N. Mackey, and F. Tisseur. Polynomial eigenvalue problems: Theory, computation, and structure. In P. Benner, M. Bollhöfer, D. Kressner, C. Mehl, and T. Stykel, editors, *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 319–348. Cham Heidelberg New York Dordrecht London, 2015.
- [22] K. Meerbergen. The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 30(4):1463–1482, 2008.
- [23] Y. Nakatsukasa, V. Nofefini, and A. Townsend. Linearizations for computing roots of rational functions. in preparation.
- [24] Y. Nakatsukasa, V. Nofefini, and A. Townsend. Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach. MIMS EPrint 2012.118, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012.
- [25] G. L. G. Sleijpen, A. G. L. Booten, D. R. Fokkema, and H. A. Van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36(3):595–633, 1996.
- [26] G. W. Stewart. *Matrix Algorithms Volume II: Eigensystems*. SIAM, Philadelphia, 2001.
- [27] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory (Computer Science and Scientific Computing)*. Academic Press, 1990.

- [28] Y. Su and Z. Bai. Solving rational eigenvalue problems via linearization. *SIAM J. Matrix Anal. Appl.*, 32(1):201–216, 2011.
- [29] D. B. Szyld and F. Xue. Several properties of invariant pairs of nonlinear algebraic eigenvalue problems. *IMA J. Numer. Anal.*, 2013.
- [30] F. Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309(1):339–361, 2000.
- [31] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43:235–286, 2001.
- [32] A. Townsend, V. Noferini, and Y. Nakatsukasa. Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach. *The Mathematical Institute, University of Oxford, Eprints Archive 1638*, 2013.
- [33] L. N. Trefethen. *Approximation Theory and Approximation Practice*. SIAM, Philadelphia, 2013.
- [34] R. Van Beeumen, W. Michiels, and K. Meerbergen. Linearization of Lagrange and Hermite interpolating matrix polynomials. *IMA J. Numer. Anal.*, 35(2):909–930, 2015.