MATHEMATICS OF COMPUTATION
Volume 00, Number 0, Pages 000–000
S 0025-5718(XX)0000-0

.

CHEBYSHEV ROOTFINDING VIA COMPUTING EIGENVALUES OF COLLEAGUE MATRICES: WHEN IS IT STABLE?

VANNI NOFERINI AND JAVIER PÉREZ

Abstract. Computing the roots of a scalar polynomial, or the eigenvalues of a matrix polynomial, expressed in the Chebyshev basis $\{T_k(x)\}$ is a fundamental problem that arises in many applications. In this work, we analyze the backward stability of the polynomial rootfinding problem solved with colleague matrices. In other words, given a scalar polynomial $p(x)$ or a matrix polynomial $P(x)$ expressed in the Chebyshev basis, the question is to determine whether the whole set of computed eigenvalues of the colleague matrix, obtained with a backward stable algorithm, like the QR algorithm, are the set of roots of a nearby polynomial or not. In order to do so, we derive a first order backward error analysis of the polynomial rootfinding algorithm using colleague matrices adapting the geometric arguments in [A. Edelman and H. Murakami, *Polynomial roots for companion matrix eigenvalues*, Math. Comp. 210, 763–776, 1995] to the Chebyshev basis. We show that, if the absolute value of the coefficients of $p(x)$ (respectively, the norm of the coefficients of $P(x)$) are bounded by a moderate number, computing the roots of $p(x)$ (respectively, the eigenvalues of $P(x)$) via the eigenvalues of its colleague matrix using a backward stable eigenvalue algorithm is backward stable. This backward error analysis also expands on the very recent work [Y. Nakatsukasa and V. Noferini, *On the stability of computing polynomial roots via confederate linearizations*, To appear in Math. Comp.] that already showed that this algorithm is not backward normwise stable if the coefficients of the polynomial $p(x)$ do not have moderate norms.

1. Introduction

A popular way to compute the roots of a monic polynomial expressed in the monomial basis is via the eigenvalues of its companion matrix. This is, for instance, the way followed by the MATLAB command `roots`, that, after balancing the companion matrix, uses the QR algorithm to get its eigenvalues. The numerical properties of this method for computing roots of polynomials have been extensively studied [8, 9, 15, 25], in particular with respect to conditioning and backward errors. It has been shown that, in practice, if the companion matrix is balanced [21], the rootfinding method using companion matrices is numerically stable, in the sense that the computed roots are the exact roots of a nearby polynomial. However,

Received by the editor December 9, 2015.

2010 *Mathematics Subject Classification.* 65H04, 65H17, 65F15, 65G50.

Key words and phrases. polynomial, roots, Chebyshev basis, matrix polynomial, colleague matrix, backward stability, polynomial eigenvalue problem, Arnold transversality theorem.

The work of Vanni Noferini was supported by European Research Council Advanced Grant MATFUN (267526).

The work of Javier Pérez was supported by Engineering and Physical Sciences Research Council grant EP/I005293.

as it was made famous by Wilkinson [22, 26, 27], polynomial roots that lie on a real interval can be highly sensitive to perturbations in the coefficients when the monomial basis is used. So, even perturbations in the coefficients of order of the machine precision may produce a catastrophically large forward error. In practice, rootfinding on a real interval is a very frequent and important situation, and one way to circumvent this problem is to use, instead, a polynomial basis such that the roots of a polynomial expressed in that basis are better conditioned functions of its coefficients, like the *Chebyshev basis*.

Chebyshev polynomials are a family of polynomials, orthogonal with respect to the weight function $w(x) = (1-x^2)^{-1/2}$ on the interval $[-1,1]$, which may be computed using the following recurrence relation [1, Chapter 22]:

$$\begin{aligned} &T_0(x) = 1, \\ &T_1(x) = x, \quad \text{and}, \\ &T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad \text{for } k \geq 2. \end{aligned} \tag{1.1}$$

Moreover, the Chebyshev polynomials $T_0(x), T_1(x), \ldots, T_n(x)$ form a basis for the vector space of polynomials of degree at most n with real coefficients $\mathbb{R}_n[x]$. Hence, any real polynomial $p(x) \in \mathbb{R}_n[x]$ can be written uniquely as $p(x) = \sum_{k=0}^{n} a_k T_k(x)$.

Chebyshev polynomials are widely used in many areas of numerical analysis, and in particular approximation theory [23]. In fact, a common approach, as done in Chebfun [24], for computing the roots of a nonlinear smooth function $f(x)$ on an interval is to approximate first $f(x)$ by a polynomial $p(x)$ expressed in the Chebyshev basis via Chebyshev interpolation and then compute the roots of $p(x)$ as the eigenvalues of its colleague matrix [11]. Also, computing the eigenvalues of matrix polynomials in the Chebyshev basis is becoming an important problem [10].

In this paper, we are interested in the backward stability of the rootfinding problem (or of the matrix polynomial eigenvalue problem) solved via colleague matrices and a backward stable eigenvalue algorithm. Our work is motivated by [18], which addresses related issues for confederate matrices (the colleague matrix is a particular example of a confederate matrix [4, 17]). Also, similar backward error analysis may be found in [8, 13, 14]. In [8], the authors study the backward stability of rootfinding methods using Fiedler companion matrices of monic polynomials expressed in the monomial basis; in [13], the authors study the backward stability of rootfinding methods using a suitable companion matrix of polynomials expressed in barycentric form; in [14], several bases are analyzed at once, for nonstandard linearizations of larger size with respect to the colleague or the companion.

Given a $p \times p$ *monic matrix polynomial in the Chebyshev basis* of degree n

$$P(x) = I_p T_n(x) + \sum_{k=0}^{n-1} A_k T_k(x), \quad \text{with } A_k \in \mathbb{R}^{p\times p}, \quad \text{for } k = 0, 1, \ldots, n-1, \tag{1.2}$$

where by monic in the Chebyshev basis we mean that the coefficient of $T_n(x)$ is equal to I_p (the $p \times p$ identity matrix), the *polynomial eigenvalue problem* consists of finding the eigenvalues of $P(x)$, that is, finding the roots of the scalar polynomial $\det(P(x))$ (note that the monicity of $P(x)$ implies its regularity, that is, $\det(P(x))$ is not identically zero). For the sake of simplicity of exposition, we focus on polynomials with real coefficients, as they are most common in practice when dealing

with the Chebyshev basis; however, the analysis of this paper can be extended to the complex case.

A common approach to solve the polynomial eigenvalue problem for $P(x)$ is to use the *block colleague matrix*

(1.3)
$$C_T = \frac{1}{2}\begin{bmatrix} -A_{n-1} & -A_{n-2}+I_p & -A_{n-3} & \cdots & -A_2 & -A_1 & -A_0 \\ I_p & 0 & I_p & 0 & \cdots & \cdots & 0 \\ 0 & I_p & 0 & I_p & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & I_p & 0 & I_p \\ 0 & \cdots & \cdots & \cdots & 0 & 2I_p & 0 \end{bmatrix} \in \mathbb{R}^{np\times np},$$

since it is known (see [2]) that the eigenvalues of (1.3) coincide with the eigenvalues of $P(x)$.

The eigenvalues of $P(x)$ may be computed as the eigenvalues of C_T using, for instance, the QR algorithm. The QR algorithm is a backward stable algorithm, this means than the computed eigenvalues are the exact eigenvalues of a matrix $C_T + E$, where E is a (possibly dense) matrix such that

$$\|E\| = O(u)\|C_T\|,$$

for some matrix norm, where u denotes the machine precision. However, the previous equation does not guarantee that the computed eigenvalues are the eigenvalues of a nearby matrix polynomial of $P(x)$ or, in other words, that this polynomial eigensolver is backward stable. In order for the method to be backward stable (in a normwise sense), the computed eigenvalues should be the exact eigenvalues of a polynomial $\widetilde{P}(x) = I_p T_n(x) + \sum_{k=0}^{n-1} \widetilde{A}_k T_k(x)$, such that

$$\frac{\|\widetilde{P} - P\|}{\|P\|} = O(u),$$

for some matrix polynomial norm.

In the scalar polynomial case ($p = 1$), the backward stability of the polynomial rootfinding in degree-graded basis using confederate matrices is studied in [18]. In particular (see [18, Theorem 4.2]), the authors prove that if C_T is the colleague matrix of a polynomial $p(x)$ and $E \in \mathbb{R}^{n\times n}$ is any matrix, then the eigenvalues of $C_T + E$ are the exact roots of a polynomial $\widetilde{p}(x)$ such that

$$\widetilde{p}(x) - p(x) = \sum_{i=0}^{n-1} \delta_i(p, E) T_i(x) + O(\|E\|_2^2), \tag{1.4}$$

where, for $i = 0, 1, \ldots, n-1$, the quantity $\delta_i(p, E)$ is an affine function of the coefficients of $p(x)$, and, separately, of the entries of E.

Equation (1.4) implies that if the roots of $p(x)$ are computed as the eigenvalues of its colleague matrix C_T using a backward stable eigenvalue algorithm, then, the computed roots will be the exact roots of a polynomial $\widetilde{p}(x)$ such that

$$\frac{\|\widetilde{p} - p\|}{\|p\|} = \kappa(n) O(u)\|p\|,$$

for some constant $k(n)$. The previous equation shows, first, that this method is not backward stable if $\|p\| \gg 1$, and, second, that this method is backward stable if the following two conditions are satisfied: (i) the quantity $\kappa(n)$ is a low-degree polynomial in n with moderate coefficients; and, (ii) the norm $\|p\|$ is moderate. As it is observed in [18], writing $\delta_i(p, E) = \sum_{i,j,\ell} \beta_{ij\ell} a_\ell E_{ij}$, since it is not clear what exactly are the constants $\beta_{ij\ell}$ involved, in principle it could happen that $|\beta_{ij\ell}| \gg 1$, implying that $\kappa(n)$ might not be a polynomial in n with moderate coefficients. However, in this work we show that, in fact, $|\beta_{ij\ell}| \leq 4$, and, so,

$$\frac{\|\widetilde{p} - p\|}{\|p\|} = O(u)\|p\|,$$

holds. The previous equation implies that computing the roots of $p(x)$ via the eigenvalues of its colleague matrix using a backward stable eigenvalue algorithm is a backward stable rootfinding algorithm, provided that $\|p\| \lesssim 1$.

Moreover, using some arguments inspired by [3, 9, 15, 16] we will generalized the previous result to the matrix polynomial case, that is, if the eigenvalues of a matrix polynomial $P(x)$ are computed as the eigenvalues of its colleague matrix using a backward stable eigenvalue algorithm, then, we prove that the computed eigenvalues are the exact eigenvalues of a monic matrix polynomial in the Chebyshev basis $\widetilde{P}(x)$ such that

$$\frac{\|\widetilde{P} - P\|}{\|P\|} = O(u)\|P\|.$$

The previous equation implies that this method is backward stable if $\|P\|$ is moderate.

The paper is organized as follows. At the beginning of Section 2 we present Arnold transversality theorem for colleague matrices, which will be the main tool to study the polynomial backward stability of the rootfinding method using colleague matrices. Then, in Section 2.2 we prove Arnold transversality theorem for colleague matrices, and in Section 2.3 we use this theorem to study the backward stability of the rootfinding method using colleague matrices.

Throughout this paper, for a $p \times p$ matrix polynomial $P(x) = \sum_{k=0}^{n} A_k T_k(x)$, non necessarily monic, $\|P\|_F$ is the norm on the vector space of $p \times p$ matrix polynomials of degree less than or equal to n defined as

$$\|P\|_F = \sqrt{\sum_{k=0}^{n} \|A_k\|_F^2}.$$

Notice that, since we are going to deal with monic polynomials in the Chebyshev basis, $A_n = I_p$. Also notice that for a scalar polynomial $p(x) = \sum_{k=0}^{n} a_k T_k(x)$, that is, for $p = 1$, this norm reduces to the usual 2-norm:

$$\|p\|_F = \|p\|_2 = \sqrt{\sum_{k=0}^{n} |a_k|^2}.$$

2. Arnold transversality theorem for colleague matrices and backward error analysis

Arnold transversality theorem will be the main tool in this section to study what kind of polynomial backward stability is provided by matrix backward stability when the roots of scalar polynomials or the eigenvalues of matrix polynomials are computed as the eigenvalues of its colleague matrix with a backward stable eigenvalue algorithm. This theorem was first stated in [3] for companion matrices, and later generalized in [18] to confederate matrices of scalar polynomials.

Following [3, 8, 9, 15, 18], we consider the Euclidian matrix space $\mathbb{R}^{n\times n}$ with the usual Frobenius inner product

$$< A, B >:= \operatorname{tr}(AB^T),$$

where M^T denotes the transpose of $M \in \mathbb{R}^{n\times n}$. In this space, the set of matrices similar to a given matrix $A \in \mathbb{R}^{n\times n}$ is a differentiable manifold in $\mathbb{R}^{n\times n}$. This manifold is called the orbit of A under the action of similarity:

$$\mathcal{O}(A) := \{SAS^{-1} \,:\, S \in \mathbb{R}^{n\times n} \text{ and } \det(S) \neq 0\}.$$

A first-order expansion shows that the tangent space of $\mathcal{O}(A)$ at A is the set

$$T_A\mathcal{O}(A) := \{AX - XA \quad \text{for some } X \in \mathbb{R}^{n\times n}\}.$$

We also consider the vector subspace of *"first block row matrices"*, denoted by $\mathcal{BFR}_{n,p} \subset \mathbb{R}^{np\times np}$, which is defined as those $n \times n$ block matrices $[X_{ij}]$, with $X_{ij} \in \mathbb{R}^{p\times p}$, whose block rows are all zero except (possibly) the first:

$$\mathcal{BFR}_{n,p} := \Big\{ X = \begin{bmatrix} I_P & 0 & \cdots & 0 \end{bmatrix}^T \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \text{ for some } X_1, X_2, \ldots, X_n \in \mathbb{R}^{p\times p} \Big\} \subset \mathbb{R}^{np\times np}.$$

Note that taking $p = 1$ the space $\mathcal{BFR}_{n,p}$ reduces to the vector subspace $\mathcal{FR}_n$ of *"first row matrices"* introduced in [18].

Arnold transversality theorem for a block colleague matrix C_T of a monic matrix polynomial $P(x)$ in the Chebyshev basis states that any matrix $E \in \mathbb{R}^{np\times np}$ may be decomposed as

$$E = F_0 + T,$$

where $F_0 \in \mathcal{BFR}_{n,p}$ is a first block row matrix and $T \in T_{C_T}\mathcal{O}(C_T)$. Notice that taking $p = 1$ in the previous decomposition, this "block" version of Arnold transversality theorem reduces to a special case of [18, Theorem 4.1].

In Section 2.3 we present a proof of Arnold transversality theorem, different to the one in [18], extending (for the important case of the Chebyshev basis) [18, Theorem 4.1] to the more complicated case of matrix polynomials. The new approach allows us to compute explicitly the matrix F_0. Then, using this explicit expression, we study the polynomial backward stability of the rootfinding method using colleague matrices.

2.1. Clenshaw shifts and Clenshaw matrices.

In this section we introduce some matrix polynomials and some matrices, named here as *Clenshaw shifts* and *Clenshaw matrices*, respectively, associated with a monic matrix polynomial in the Chebyshev basis $P(x)$, that will be used through Section 2.3 and will be key in the following developments. Clenshaw shifts are the generalization of the Horner shifts (see [7]) when the polynomial $P(x)$ is expressed in the Chebyshev basis.

Associated with the $p \times p$ monic matrix polynomial in the Chebyshev basis $P(x)$ in (1.2), we define the following $p \times p$ matrix polynomials:

$$\begin{aligned} &H_0(x) = 2I_p, \\ &H_1(x) = 2xH_0(x) + 2A_{n-1}, \\ &H_k(x) = 2xH_{k-1}(x) - H_{k-2}(x) + 2A_{n-k}, \quad \text{for } k = 2, 3, \dots, n-2, \\ &H_{n-1}(x) = xH_{n-2}(x) - H_{n-3}(x)/2 + A_1. \end{aligned} \tag{2.1}$$

We will refer, for $k = 1, 2, \dots, n$, to the matrix polynomial $H_k(x)$ as the degree k *Clenshaw shift* of $P(x)$, since for $p = 1$ they coincide with the well known Clenshaw shifts associated with a scalar polynomial expressed in the Chebyshev basis [6]. Clenshaw shifts are related with the polynomial $P(x)$ through the following equation [6]:

$$2P(x) = 2xH_{n-1}(x) - H_{n-2}(x) + 2A_0. \tag{2.2}$$

In Theorem 2.1, given the Chebyshev polynomial $T_{n-i}(x)$ and the Clenshaw shift $H_{n-k}(x)$, we show how to express $T_{n-i}(x)H_{n-k}(x)$ uniquely as $Q_{ij}(x)+r_{ik}(x)P(x)$, where $Q_{ij}(x)$ is a $p \times p$ matrix polynomial of degree less than or equal to $n-1$ and $r_{ik}(x)$ is a scalar polynomial. The proof of Theorem 2.1 is elementary but rather technical, so we leave it to the appendix. In order to write down a reasonably simple formula for $T_{n-i}(x)H_{n-k}(x)$, we define the following quantities

(2.3)

$$\Gamma_{2k+1} = \Gamma_{2k-1} + 2A_{n-2k-1}, \quad \text{for } k = 1, 2, \dots, \left\lfloor \frac{n}{2} \right\rfloor - 1, \quad \text{with } \Gamma_0 = 2I_p, \quad \text{and}$$

$$\Gamma_{2k} = \Gamma_{2(k-1)} + 2A_{n-2k}, \quad \text{for } k = 1, 2, \dots, \left\lceil \frac{n}{2} \right\rceil - 1, \quad \text{with } \Gamma_1 = 2A_{n-1}.$$

Notice that in Γ_k only appear coefficients of $P(x)$ with indices of the same parity.

Theorem 2.1. *Let $P(x) = I_pT_n(x) + \sum_{k=0}^{n-1} A_kT_k(x)$ be a $p \times p$ monic matrix polynomial in the Chebyshev basis of degree n, let $T_{n-i}(x)$ and $H_{n-k}(x)$ be, respectively, the degree $n-i$ Chebyshev polynomial and the degree $n-k$ Clenshaw shift of $P(x)$, with $i, k \in \{1, 2, \dots, n\}$. Then, there exist a unique $p \times p$ matrix polynomial $Q_{ik}(x)$ of degree less than or equal to $n-1$ and a unique scalar polynomial $r_{ik}(x)$ such that*

$$T_{n-i}(x)H_{n-k}(x) = Q_{ik}(x) + r_{ik}(x)P(x),$$

where,

- *if $i \geq n-k+1$ and $k \geq 2$,*

$$Q_{ik}(x) = \sum_{\ell=0}^{n-k-1} \Gamma_\ell (T_{2n-i-k-\ell}(x) + T_{|k+\ell-i|}(x)) + \Gamma_{n-k}T_{n-i}(x); \tag{2.4}$$

- *if $i = n$ and $k = 1$,*

$$Q_{ik}(x) = \sum_{\ell=0}^{n-2} \Gamma_\ell T_{n-1-\ell}(x) + \frac{\Gamma_{n-1}}{2} T_0(x); \tag{2.5}$$

- *if* $i \le n-k$ *and* $n-1 \ge k \ge 2$,

$$\begin{aligned} Q_{ik}(x) = &\sum_{\ell=0}^{i-2} \Gamma_\ell (T_{i+k-2-\ell}(x) + T_{|k+\ell-i|}(x)) + \Gamma_{i-1} T_{k-1}(x) \\ &- \sum_{\ell=1}^{n-k+1-i} \sum_{r=1}^{k-1+\ell} 2A_{k-1+\ell-r} T_{|n-i+1-\ell-r|}(x); \end{aligned} \tag{2.6}$$

- *if* $i \le n-k$ *and* $k=1$

$$Q_{ik}(x) = \sum_{\ell=0}^{i-2} \Gamma_\ell T_{i-1-\ell}(x) + \frac{\Gamma_{i-1}}{2} T_0(x) - \sum_{\ell=1}^{n-i} \sum_{r=1}^{\ell} A_{\ell-r} T_{|n-i+1-\ell-r|}(x); \tag{2.7}$$

where Γ_ℓ, *for* $\ell = 0, 1, 2, \ldots$, *is defined in* (2.3).

From Theorem 2.1, it is clear that there exists a unique $n \times n$ block matrix $M_k = [(M_k)_{ij}]$, with $(M_k)_{ij} \in \mathbb{R}^{p\times p}$, such that, for $k = 1, 2, \ldots, n$,

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_{n-k}(x) = M_k \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes I_p + \begin{bmatrix} r_{1k}(x) \\ \vdots \\ r_{n-1,k}(x) \\ r_{nk}(x) \end{bmatrix} \otimes P(x), \tag{2.8}$$

where $\otimes$ denotes the Kronecker product, for some scalar polynomials $r_{1k}(x), \ldots, r_{nk}(x)$. We will refer to the matrix M_k in (2.8) as the kth *Clenshaw matrix* of $P(x)$.

By direct multiplication, it may be easily checked that the block colleague matrix C_T satisfies

$$x \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes I_p = C_T \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes I_p + \frac{1}{2} e_1 \otimes P(x). \tag{2.9}$$

Equations (2.8) and (2.9) shows that the Clenshaw matrices and the colleague matrix can be interpreted, respectively, as the multiplication-by-Clenshaw shifts and the multiplication-by-x operators in certain quotient modules (see also [19, Sec. 5]).

Using (2.8) and (2.9), in Proposition 2.2 we show that the Clenshaw matrices $M_1, M_2, \ldots, M_n$ in (2.8) satisfy a simple recurrence relation.

Proposition 2.2. *Let* $P(x) = I_p T_n(x) + \sum_{k=0}^{n-1} A_k T_k(x)$ *be a* $p \times p$ *monic matrix polynomial in the Chebyshev basis of degree* n, *let* C_T *be the block colleague matrix of* $P(x)$, *and let* $M_1, M_2, \ldots, M_n$ *be the Clenshaw matrices in* (2.8). *Then,*

$$\begin{aligned} &M_n = I_n \otimes 2I_p, \\ &M_{n-1} = 2M_n C_T + I_n \otimes 2A_{n-1}, \\ &M_k = 2M_{k+1} C_T - M_{k+2} + I_n \otimes 2A_k, \quad \text{for } k = n-2, \ldots, 3, 2, \quad \text{and} \\ &M_1 = M_2 C_T - M_3/2 + I_n \otimes A_1. \end{aligned} \tag{2.10}$$

Proof. The proof proceeds backwards from $k = n$. First, we prove that the result is true for $k = n$. From (2.1), we have

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_0(x) = \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes 2I_p = \begin{bmatrix} 2I_p & & & \\ & \ddots & & \\ & & 2I_p & \\ & & & 2I_p \end{bmatrix} \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes I_p.$$

Comparing the previous equation with (2.8), we deduce that $M_n = I_n \otimes 2I_p$.

Second, we prove that the result is true for $k = n - 1$. From (2.1), we have

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_1(x) = \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes (2xH_0(x) + 2A_{n-1}) = 2x \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_0(x) + \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes 2A_{n-1}.$$

Using (2.8) with $k = n$, together with (2.9), we get

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_1(x) = (2M_nC_T + I_n \otimes 2A_{n-1}) \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} + \begin{bmatrix} r_{1,n-1}(x) \\ \vdots \\ r_{n-1,n-1}(x) \\ r_{n,n-1}(x) \end{bmatrix} \otimes P(x),$$

for some scalar polynomials $r_{1,n-1}(x), \dots, r_{n,n-1}(x)$. Comparing the previous equation with (2.8), we deduce that $M_{n-1} = 2M_nC_T + I_n \otimes 2A_{n-1}$.

Third, we prove that the result is true for $n - 2 \geq k \geq 2$. From (2.1), we have

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_{n-k}(x) = \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes (2xH_{n-k-1}(x) - H_{n-k-2}(x) + 2A_k)$$

$$= 2x \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_{n-k-1}(x) - \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_{n-k-2}(x) + \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes 2A_k.$$

Using (2.8) with $k+1$ and $k+2$, together with (2.9), we get

$$\begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} \otimes H_{n-k}(x) = (2C_T M_{k+1} - M_{k+2} + I_n \otimes 2A_k) \begin{bmatrix} T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{bmatrix} + \begin{bmatrix} r_{1k}(x) \\ \vdots \\ r_{n-1,k}(x) \\ r_{n,k}(x) \end{bmatrix} \otimes P(x),$$

for some scalar polynomials $r_{1k}(x), \ldots, r_{nk}(x)$. Comparing the previous equation with (2.8), we deduce that $M_k = 2C_T M_{k+1} - M_{k+2} + I_n \otimes 2A_k$.

Finally, the proof of the last case ($k = 1$) is similar to the proof for the previous cases ($n - 2 \geq k \geq 2$), but using $H_{n-1}(x) = xH_{n-2}(x) - H_{n-3}(x)/2 + A_1$, so we omit it. □

Remark 2.3. *Clenshaw matrices are closely related with the so called Leverrier's algorithm for orthogonal polynomial bases* [5]*, which allows the simultaneous determination of the characteristic polynomial of a matrix A and the adjoint matrix of $xI - A$. Indeed, if we consider a scalar polynomial $p(x) = T_n(x) + \sum_{k=0}^{n-1} a_k T_k(x)$, it may be checked that the adjoint of $zI - C_T$ is given by*

$$\text{adj } (xI - C_T) = \frac{1}{2^{n-1}} \sum_{k=0}^{n-1} M_{k+1} T_k(x), \tag{2.11}$$

where $M_1, M_2, \ldots, M_n$ are the Clenshaw matrices of $p(x)$.

The Clenshaw matrices $M_1, M_2, \ldots, M_n$ have a complicated structure. We illustrate this with an example of moderate size. For $n = 6$ and $k = 3$, it is easy to check using (2.10) that the matrix M_k is equal to

$$\begin{bmatrix} 0 & -2A_2 & -2A_3 - 2A_1 & 2I_p - 2A_4 - 2A_2 - 4A_0 & -2A_3 - 4A_1 & -2A_2 - 2A_0 \\ 0 & 0 & 2I_p - 2A_2 & 2A_5 - 2A_3 - 2A_1 & 2I_p - 2A_2 - 4A_0 & -2A_1 \\ 0 & 2I_p & 2A_5 & 2I_p + 2A_4 - 2A_2 & 2A_5 - 2A_1 & 2I_p - 2A_0 \\ 2I_p & 2A_5 & 2I_p + 2A_4 & 2A_5 + 2A_3 & 4I_p + 2A_4 & 2A_5 \\ 0 & 2I_p & 2A_5 & 4I_p + 2A_4 & 4A_5 + 2A_3 & 2A_4 + 2I_p \\ 0 & 0 & 4I_p & 4A_5 & 4I_p + 4A_4 & 2A_5 + 2A_3 \end{bmatrix}.$$

Two observations about the block matrix above are: (i) its first block column is equal to $e_{n-k+1} \otimes 2I_p$, where e_ℓ denotes the ℓth column of the $n \times n$ identity matrix; and, (ii) if we set $A_n := I_p$, each block entry has the form $\sum_{i=0}^{n} c_i A_i$, where $|c_i| \leq 4$. In Theorem 2.4, we show that the two previous observations are true for any n and k. Property (i) will be key to prove Arnold transversality theorem, and property (ii) will be key to study what kind of backward stability of a linearization-based algorithm for the polynomial eigenvalue problem is provided by the backward stability of an eigensolver for the linearized problem.

Theorem 2.4. *Let $P(x) = I_p T_n(x) + \sum_{k=0}^{n-1} A_k T_k(x)$ be a $p \times p$ monic matrix polynomial in the Chebyshev basis of degree n, and let M_k, for $k = 1, 2, \ldots, n$, be the kth Clenshaw matrix in* (2.8)*. Then, the following statements hold:*

(a) *The first block column of M_k is equal to $e_{n-k+1} \otimes 2I_p$, where e_ℓ denotes the ℓth column of the $n \times n$ identity matrix.*
(b) *For $i, j = 1, 2, \dots, n$, the (i,j)th block entry of M_k satisfies $(M_k)_{ij} = \sum_{t=0}^{n} \alpha_{t,ijk} A_t$ with $|\alpha_{t,ijk}| \leq 4$, where we set $A_n := I_p$.*

Proof. From Theorem 2.1 together with (2.8), we have $H_{n-k}(x)T_{n-i}(x) = \sum_{j=1}^{n}(M_k)_{ij}T_{n-j}(x) + r_{ik}(x)P(x)$. Therefore, to prove part (a) it is enough to show that

$$T_{n-i}(x)H_{n-k}(x) = 2I_pT_{n-1}(x) + \cdots + r_{ik}(x)P(x), \tag{2.12}$$

if $i = n - k + 1$, and that

$$T_{n-i}(x)H_{n-k}(x) = (M_k)_{i\nu}T_\nu(x) + \cdots + r_{ik}(x)P(x), \tag{2.13}$$

with $\nu < n-1$, if $i \neq n-k+1$, where the dots correspond to Chebyshev polynomials with lower indices.

First, suppose that $i \geq n - k + 1$. We will prove that $T_{n-i}(x)H_{n-k}(x)$ is of the form (2.12) when $i = n - k + 1$ and it is of the form (2.13) otherwise. We need to distinguish several cases. First, let $k = n$. From (2.4) we get that $T_{n-i}(x)H_0(x) = \Gamma_0 T_{n-i}(x) = 2I_pT_{n-i}(x)$. Since the index $n-i$ is equal to $n-1$ if and only if $i = 1$, the result is true in this case. Then, consider the case $n-1 \geq k \geq 2$. There are three kinds of indices of Chebyshev polynomials in (2.4). The first is $2n-i-k-\ell$, which is equal to $n-1$ if and only if $\ell = 0$ and $i = n-k+1$. This gives a contribution $\Gamma_0 T_{n-1}(x) = 2I_pT_{n-1}(x)$ only when $i = n-k+1$. The second one is $|k+\ell-i|$. Taking into account the possible values that k, ℓ, and i can take in (2.4), it may be easily checked that this index is smaller than or equal to $n-2$. The third index is $n-i$ which necessarily is smaller than or equal to $n-2$, and, hence, the result is true in this case. Finally, consider the case $k = 1$ and $i = n$. There are two kinds on indices of Chebyshev polynomials in (2.5). The first one is $n-1-\ell$, which is equal to $n-1$ if and only if $\ell = 0$. This gives a contribution $\Gamma_0 T_{n-1}(x) = 2I_pT_{n-1}(x)$. The second index is 0, which is smaller than $n-2$. Therefore, the result is also true in this case.

Now suppose that $i \leq n - k$. We will we prove that $T_{n-i}(x)H_{n-k}(x)$ is of the form (2.13). Notice that there are four kinds of indices in (2.6) when $k \geq 2$, namely, $i+k-2-\ell$, $|k+\ell-i|$, $|n-i+1-\ell-r|$ and $k-1$, and three kinds on indices in (2.7) when $k = 1$, namely, $i-1-\ell$, $i-1$ and $|n-i+1-\ell-r|$. Taking into account the possible values that k, ℓ, r, and i can take in (2.13), in both cases ($k \geq 2$ and $k = 2$), it may be checked that these indices do not exceed $n-2$.

Now, we proceed to prove part (b). Again, we need to distinguish several cases. First, suppose that $i \geq n-k+1$ and also assume that $k \geq 2$ (the argument when $k = 1$ is similar and simpler, so we omit it), and consider the three kinds of indices of Chebyshev polynomials that appear in (2.4), namely, $2n-i-k-\ell$, $|k+\ell-i|$, and $n-i$. For $\ell = 0, 1, \dots, n-k$, a careful look at these indices reveals that if $k+\ell-i \geq 0$, then the three of them are different. Therefore, we can write (2.4) as

$$\sum_{\ell=0}^{n-1} B_\ell T_\ell(x) + \sum_{\ell=-(1-n)}^{-1} B_\ell T_{-\ell}(x), \tag{2.14}$$

where B_ℓ is equal to either 0 or Γ_t for some t. It follows that $(M_k)_{ij}$ is equal to either 0, Γ_t for some t, or $\Gamma_{t_1} + \Gamma_{t_2}$ for some t_1, t_2. Finally, recall from (2.3) that

Γ_t is equal to $2I_p + 2A_{n-2} + 2A_{n-4} + \cdots$ if t is even, or to $2A_{n-1} + 2A_{n-3} + \cdots$ if t is odd. Therefore, $(M_k)_{ij} = \sum_{t=0}^{n} \alpha_{t,ijk} A_t$, with $|\alpha_{t,ijk}| \leq 4$.

Then suppose that $i \leq n-k$ and also assume that $k \geq 2$ (again, the argument when $k = 1$ is similar and simpler, so we omit it). First, consider the three kinds of indices of Chebyshev polynomials that appear in the first summand in (2.6), namely, $i + k - 2 - \ell$, $|k + \ell - i|$, $k - 1$. For $\ell = 0, 1, \ldots, i-2$, again, it may be checked that if $k + \ell - i \geq 0$, then these three indices are different. Therefore, the first summand in (2.6) is also of the form (2.14). Finally, consider the index of the Chebyshev polynomials and the index of the coefficients A_i that appear in the second summand in (2.6), namely, $|n - i + 1 - \ell - r|$, and $k + 1 + \ell - r$. If $n - i + 1 - \ell - r \geq 0$, it may be checked that for any two allowed different pairs (ℓ, r) that realize the same value of $n - i + 1 - \ell - r$, then the associate indices $k + 1 + \ell - r$ must be different. Since the same occur when $n - i + 1 - \ell - r < 0$, it follows that (2.6) is of the form

$$\sum_{\ell=0}^{n-2} C_\ell T_\ell(x) + \sum_{\ell=-(2-n)}^{-1} C_\ell T_{-\ell}(x) - 2\sum_{\ell=0}^{n-2} D_\ell T_\ell(x) - 2\sum_{\ell=2-n}^{-1} D_\ell T_{-\ell}(x) + r_{ik}(x)P(x)$$

where C_ℓ is equal to either 0 or Γ_t for some t, and D_ℓ is equal to $\sum_{t=1}^{q_\ell} A_{i_t}$, where $i_{t_1} \neq i_{t_2}$ whenever $t_1 \neq t_2$. Then, it follows that

$$(M_k)_{ij} = \sum_{\ell=0}^{n} \delta_\ell A_\ell - \sum_{\ell=0}^{n} \rho_\ell A_\ell,$$

where δ_ℓ and ρ_ℓ are equal to either 4, or 2 or 0, therefore $(M_k)_{ij} = \sum_{t=0}^{n} \alpha_{t,ijk} A_t$ with $|\alpha_{t,ijk}| \leq 4$. □

If necessary, explicit expressions of the entries of the Clenshaw matrices $M_1, M_2, \ldots, M_n$ may be obtained from Theorem 2.1. However, since Theorem 2.4 is the only information that we will need about them to prove our main results in the following section, we do not pursue that idea.

2.2. **Proof of Arnold transversality theorem for colleague matrices.** In this section we prove Arnold transversality theorem for colleague matrices of monic polynomials in the Chebyshev basis. That is, we show that any matrix $E \in \mathbb{R}^{pn \times pn}$ may be decomposed as

$$E = F_0 + T, \tag{2.15}$$

where $F_0 \in \mathcal{BFR}_{n,p}$ is a first block row matrix and $T \in T_{C_T}\mathcal{O}(C_T)$, constructing the matrix F_0 explicitly.

As in the case of the monomial basis, generically, $\dim\left(T_{C_T}\mathcal{O}(C_T)\right) + \dim\left(\mathcal{BFR}_{n,p}\right) = n^2p^2 - np + np^2 \geq n^2p^2$ (see [9, 16]). In words, the tangent space $T_{C_T}\mathcal{O}(C_T)$ and the vector space of first block row matrices $\mathcal{BFR}_{n,p}$ may have a nontrivial intersection when $p > 1$. For this reason, following [9, 16] we choose a particular subspace of the tangent space that will give a unique decomposition (2.15). This subspace is denoted by $\text{Sub } T_{C_T}\mathcal{O}(C_T)$ and it is given by

$$\text{Sub } T_{C_T}\mathcal{O}(C_T) = \{X \in T_{C_T}\mathcal{O}(C_T) \quad \text{such that} \quad X \text{ has 0 first block column}\}.$$

In order to get the decomposition (2.15) with $T \in \text{Sub } T_{C_T}\mathcal{O}(C_T)$ we will make use of the Clenshaw matrices $M_1, M_2, \ldots, M_n \in \mathbb{R}^{np \times np}$, defined in (2.8), of the matrix polynomial $P(x)$ in (1.2). Though the only information that we need about

Clenshaw matrices are those stated in Theorem 2.4 together with the recurrence relation (2.10).

Following [9], we also define the *block trace* of a $np \times np$ block matrix $Z = [Z_{ij}]$, with $Z_{ij} \in \mathbb{R}^{p\times p}$, as the $p \times p$ matrix

$$\operatorname{tr}_p(Z) := \sum_{i=1}^{n} Z_{ii}.$$

The block trace is used in Theorem 2.5, which provides a characterization of the subspace $\operatorname{Sub} T_{C_T}\mathcal{O}(C_T)$, and is a generalization of [9, Theorem 4.1] when the matrix polynomial $P(x)$ is expressed in the Chebyshev basis.

Theorem 2.5. *For any $Z \in \mathbb{R}^{pn\times pn}$,*

$$\operatorname{tr}_p(M_{k+1}Z) = 0, \quad \text{for } k = 0, 1, \dots, n-1, \tag{2.16}$$

if and only if

$$Z = C_T X - X C_T \quad \text{for some } X \in \mathbb{R}^{np\times np} \text{ with 0 first block column.} \tag{2.17}$$

Moreover, either condition determines the first block row of Z uniquely given the remaining block rows.

Proof. From part (a) in Theorem 2.4, the $(n-k, 1)$ block entry of M_{k+1} is equal to $2I_p$, and the $(i, 1)$ block entry of M_{k+1}, with $i \neq n-k$, is equal to 0. Therefore, $Z_{1,n-k}$, for $k = 0, 1, \dots, n-1$, is uniquely determined from (2.16). Also, if X has 0 first block column, it may be easily checked that the map from X to the last $n-1$ block rows of $C_T X - X C_T$ has a trivial nullspace. Thus, Z is uniquely determined by (2.17).

To finish the proof we need to prove that (2.17) implies (2.16). That is, we need to show that $\operatorname{tr}_p(M_{k+1}(C_T X - X C_T)) = 0$ for any block matrix X with 0 first block column. In order to do this, first, we show that if X has 0 first block column, then $\operatorname{tr}_p(M_{k+1}XC_T) = \operatorname{tr}_p(C_T M_{k+1}X)$. The proof of the previous equation is not completely immediate when $p > 1$ since, in this situation, $\operatorname{tr}_p(AB) = \operatorname{tr}_p(BA)$ does not hold in general. So, consider a block matrix Y that has 0 first block column. Then,

$$\operatorname{tr}_p(C_T Y) = \sum_{i=1}^{p-2} \left(\frac{Y_{i,i+1}}{2} + \frac{Y_{i+2,i+1}}{2} \right) + Y_{p-1,p} = \operatorname{tr}_p(YC_T).$$

Therefore, if X has 0 first block column, then $\operatorname{tr}_p(M_{k+1}XC_T) = \operatorname{tr}_p(C_T M_{k+1}X)$.

Then, we show that $\operatorname{tr}_p(C_T M_{k+1}X) = \operatorname{tr}_p(M_{k+1}C_T X)$. To do this, note that the Clenshaw matrix M_{k+1} is of the form $2^{n-k}C_T^{n-k-1} + \sum_{t=1}^{n-k-1}(I_n \otimes B_k)C_T^{n-k-1-t}$, for some $B_1, B_2, \dots, B_{n-k-1} \in \mathbb{R}^{p\times p}$ (this can be verified by induction using (2.10)). So, we only need to show that $\operatorname{tr}_p(C_T(I_n \otimes B)C_T^j X) = \operatorname{tr}_p((I \otimes B)C_T^j C_T X)$. Indeed, since the matrix $C_T(I_n \otimes B) - (I_n \otimes B)C_T$ is 0 except the first block row, and since $C_T X$ has 0 first block column, it follows that $\operatorname{tr}_p(C_T(I_n \otimes B)C_T^j X - (I \otimes B)C_T^j C_T X) = 0$. Therefore, $\operatorname{tr}_p(C_T M_{k+1}X) = \operatorname{tr}_p(M_{k+1}C_T X)$. Thus, we conclude that $\operatorname{tr}_p(M_{k+1}XC_T) = \operatorname{tr}_p(C_T M_{k+1}X) = \operatorname{tr}_p(M_{k+1}C_T X)$. □

In Theorem 2.6 we present the proof of Arnold transversality theorem for block colleague matrices. Part (a) in Theorem 2.4 will be key here.

Theorem 2.6. *Let $P(x) = I_p T_n(x) + \sum_{k=0}^{n-1} A_k T_k(x)$ be a $p \times p$ monic matrix polynomial in the Chebyshev basis of degree n, and let C_T be its block colleague matrix. Then, any matrix $E \in \mathbb{R}^{np\times np}$ can be expressed as*

$$E = F_0 + T, \tag{2.18}$$

where $F_0 \in \mathcal{BFR}_{n,p}$ is a first block row matrix, and $T \in \operatorname{Sub} T_{C_T}\mathcal{O}(C_T)$. Moreover, if the first block row of F_0 is written as $\begin{bmatrix} F_0^{(n-1)} & \cdots & F_0^{(1)} & F_0^{(0)} \end{bmatrix}$, then

$$F_0^{(k)} = \frac{1}{2}\operatorname{tr}_p(EM_{k+1}), \quad \text{for } k = 0, 1, \ldots, n-1, \tag{2.19}$$

where the matrix M_{k+1} is the $(k+1)$th Clenshaw matrix defined in (2.8).

Proof. Define $F_0^{(k)} = \frac{1}{2}\operatorname{tr}_p(EM_{k+1})$, for $k = 0, 1, \ldots, n-1$, and let $F_0 \in \mathcal{BFR}_{n,p}$ be a first block row matrix such that its first block row is $\begin{bmatrix} F_0^{(n-1)} & \cdots & F_0^{(1)} & F_0^{(0)} \end{bmatrix}$. We may write the matrix $T := E - F_0$. Then, we have to check that $T \in \operatorname{Sub} T_{C_T}\mathcal{O}(C_T)$. From Theorem 2.5, we see that it is sufficient to show that $\operatorname{tr}_p(TM_{k+1}) = 0$, for $k = 0, 1, \ldots, n-1$. Indeed, using part (a) in Theorem 2.4,

$$\begin{aligned}\operatorname{tr}_p(TM_{k+1}) =& \operatorname{tr}_p(EM_{k+1}) - \operatorname{tr}_p(F_0M_{k+1}) = \operatorname{tr}_p(EM_{k+1}) - 2F_0^{(k)} = \\ & \operatorname{tr}_p(EM_{k+1}) - \operatorname{tr}_p(EM_{k+1}) = 0,\end{aligned}$$

for $k = 0, 1, \ldots, n-1$. So, we conclude that $T \in \operatorname{Sub} T_{C_T}\mathcal{O}(C_T)$. □

The norm of the matrix X in $T = C_T X - XC_T$ in (2.18) has the remarkable property that it depends only on the matrix E and not on the coefficients of the matrix polynomial $P(x)$. We prove this fact in the following lemma.

Lemma 2.7. *The matrix X in $T = C_T X - XC_T$ in* (2.18) *can be bounded as $\|X\|_F \leq C\|E\|_F$, for some constant C which does not depend on the coefficients of the matrix polynomial $P(x)$.*

Proof. Recall that the matrix X with 0 first block column is uniquely determined by

$$C_T X - XC_T = \left[\begin{array}{ccc} ? & \cdots & ? \\ \hline E_{21} & \cdots & E_{2n} \\ \vdots & & \vdots \\ E_{n1} & \cdots & E_{nn} \end{array}\right], \tag{2.20}$$

where the "?" blocks are not taken into account. Then, notice that the 0 first block column of X implies that the block entries of the bottom $n-1$ block rows of $C_T X - XC_T$ are just linear combinations of the block entries of X. For example, if $n = 5$, these block rows are

$$\begin{bmatrix} -X_{22} & X_{12}+X_{32}-X_{23} & X_{13}+X_{33}-X_{22}-X_{24} & X_{14}+X_{34}-X_{23}-2X_{25} & X_{15}+X_{35}-X_{24} \\ -X_{32} & X_{22}+X_{42}-X_{33} & X_{23}+X_{43}-X_{32}-X_{34} & X_{24}+X_{44}-X_{33}-2X_{35} & X_{25}+X_{45}-X_{34} \\ -X_{42} & X_{32}+X_{52}-X_{43} & X_{33}+X_{53}-X_{42}-X_{44} & X_{34}+X_{54}-X_{43}-2x_{45} & X_{35}+X_{55}-X_{44} \\ -X_{52} & 2X_{42}-X_{53} & 2X_{43}-X_{52}-X_{54} & 2X_{44}-X_{53}-2X_{55} & 2X_{45}-X_{54} \end{bmatrix}.$$

Thus, (2.20) gives rise to a linear system of equations whose solution does not depend on the coefficients of $P(x)$. This system of equations can be easily solved. For simplicity, we describe the procedure to obtain its solution for $n = 5$: it is immediate to generalize the procedure to any n, and this claim correspond to the fact that the matrix of the coefficients of the linear system is permutation equivalent

to a lower triangular invertible matrix. In this case, the block entries of the matrix X can be obtained in the following order

$$\begin{bmatrix} 0 & (X_{12})_{20} & (X_{13})_{19} & (X_{14})_{18} & (X_{15})_{17} \\ 0 & (X_{22})_{1} & (X_{23})_{16} & (X_{24})_{15} & (X_{25})_{14} \\ 0 & (X_{32})_{2} & (X_{33})_{5} & (X_{34})_{13} & (X_{35})_{12} \\ 0 & (X_{42})_{3} & (X_{43})_{6} & (X_{44})_{8} & (X_{45})_{11} \\ 0 & (X_{52})_{4} & (X_{53})_{7} & (X_{54})_{9} & (X_{55})_{10} \end{bmatrix},$$

where the index outside the parenthesis indicates the order in which each block is obtained while solving the linear system. Each block entry of X is a linear combination of block entries of E, and therefore $\|X\|_F \leq C\|E\|_F$, for some constant C independent of the coefficients of $P(x)$. □

2.3. **Backward error of the Chebyshev rootfinding method using colleague matrices.** An important consequence of the decomposition in Theorem 2.6 and Lemma 2.7, is that if E is a small perturbation of the block colleague matrix C_T, then

$$\begin{aligned} C_T + E =& C_T + F_0 + T = C_T + F_0 + (C_T X - X C_T) = \\ & (I + X)^{-1}(C_T + F_0 + E_1)(I + X), \end{aligned}$$

with $\|E_1\|_F = O(\|E\|_2^2)$, where we have used that T can be written as $C_T X - X C_T$, for some $X \in \mathbb{R}^{np\times np}$ with 0 first block column and $\|X\|_F \leq C\|E\|_F$. Noticing that $C_T + F_0$ is in turn a block colleague matrix of another matrix polynomial, we deduce that a small perturbation of the block colleague matrix of $P(x)$ is similar, to first order in the norm of the perturbation, to a block colleague matrix of a perturbed polynomial $\widetilde{P}(x)$. This observation allows us to formulate the following corollary.

Corollary 2.8. *Let $P(x) = I_p T_n(x) + \sum_{k=0}^{n-1} A_k T_k(x)$ be a $p \times p$ monic matrix polynomial in the Chebyshev basis of degree n, and let C_T be its block colleague matrix. Assume that the eigenvalues of $P(x)$ are computed as the eigenvalues of C_T with a backward stable algorithm, i.e., an algorithm that computes the exact eigenvalues of some matrix $C_T + E$, with $\|E\|_F = O(u)\|C_T\|_F$, where u is the machine precision. Then, to first order in u, the computed roots are the exact roots of a polynomial $\widetilde{P}(x)$ such that*

$$\frac{\|\widetilde{P} - P\|_F}{\|P\|_F} = O(u)\|P\|_F.$$

Proof. If a backward stable eigensolver is given C_T as an input, the computed eigenvalues are the exact eigenvalues of a matrix $C_T + E$, for some E with $\|E\|_F = \epsilon\|C_T\|_F$, where $\epsilon = uh(n)$, for some low degree polynomial h with moderate coefficients. In other words, the computed eigenvalues are the exact roots of the polynomial $\det(xI - C_T - E)$.

Using Theorem 2.6, we can write $E = F_0 + T$, where $T \in \operatorname{Sub} T_{C_T}\mathcal{O}(C_T)$ and F_0 is a first block row matrix with first block row as in (2.19). Therefore, to first order in u, we get

$$C_T + E = C_T + F_0 + C_T X - X C_T = (I + X)^{-1}(C_T + F_0 + O(u^2))(I + X)$$

We now show that we can apply a similarity transformation so that $S(C_T + F_0 + O(u^2))S^{-1}$ is a colleague matrix $C_T + \widehat{F}_0$, with $\|\widehat{F}_0 - F_0\|_F = O(u^2)$. The

construction of this similarity transformation is constructive and algorithmic, along the same lines as the proof of [18, Lemma 3.1]. For simplicity, we describe the procedure to construct it in a moderate case ($n = 5$). The general case can be treated similarly. In this situation, let us write $C_T + F_0 + O(u^2)$ as

$$\frac{1}{2}\begin{bmatrix} -A_4+\widehat{F}_0^{(4)} & I_p - A_3 + \widehat{F}_0^{(3)} & -A_2+\widehat{F}_0^{(2)} & -A_1+\widehat{F}_0^{(1)} & -A_0+\widehat{F}_0^{(0)} \\ \widehat{I_p} & \widehat{0} & \widehat{I_p} & \widehat{0} & \widehat{0} \\ \widehat{0} & \widehat{I_p} & \widehat{0} & \widehat{I_p} & \widehat{0} \\ \widehat{0} & \widehat{0} & \widehat{I_p} & \widehat{0} & \widehat{I_p} \\ \widehat{0} & \widehat{0} & \widehat{0} & 2\widehat{I_p} & \widehat{0} \end{bmatrix},$$

where, following [18], we adopt the following notation. For any matrix A, the matrix $\widehat{A}$ denotes a matrix such that $\|\widehat{A} - A\|_F = O(u^2)$.

The zeros and identity blocks of the matrix above can be recovered via row scaling, and column and row Gaussian operations. The order in which these operations are performed is indicated in the following matrix

$$\frac{1}{2}\begin{bmatrix} -A_4+\widehat{F}_0^{(4)} & I_p - A_3 + \widehat{F}_0^{(3)} & -A_2+\widehat{F}_0^{(2)} & -A_1+\widehat{F}_0^{(1)} & -A_0+\widehat{F}_0^{(0)} \\ I_{p\,(1,\mathrm{rs})} & 0_{\,(11,\mathrm{c})} & I_{p\,(11,\mathrm{c})} & 0_{\,(11,\mathrm{c})} & 0_{\,(11,\mathrm{c})} \\ 0_{\,(2,\mathrm{r})} & I_{p\,(3,\mathrm{rs})} & 0_{\,(10,\mathrm{c})} & I_{p\,(10,\mathrm{c})} & 0_{\,(10,\mathrm{c})} \\ 0_{\,(2,\mathrm{r})} & 0_{\,(4,\mathrm{r})} & I_{p\,(5,\mathrm{rs})} & 0_{\,(9,\mathrm{c})} & I_{p\,(9,\mathrm{c})} \\ 0_{\,(2,\mathrm{r})} & 0_{\,(4,\mathrm{r})} & 0_{\,(6,\mathrm{r})} & 2I_{p\,(7,\mathrm{rs})} & 0_{\,(8,\mathrm{c})} \end{bmatrix},$$

where the first subscript denotes the order in which the $O(u^2)$ perturbations to the zero and identity blocks are annihilate, and the second subscripts denotes whether this is done via a row scaling (rs), or via a row (r) or a column (c) Gaussian operation. Notice that these row and column operations may be obtained, respectively, pre and post multiplying by a matrix of the form $I + S_i$. In order to preserve the eigenvalues, after pre (resp. post) multiplying by $I + S_i$ we need to post (resp. pre) multiply by $(I + S_i)^{-1}$, but notice in addition, that this inverse operation never destroys the already recovered zero and identity blocks.

Finally, writing E as a $np \times np$ block matrix $E = [E_{ij}]$, with $E_{ij} \in \mathbb{R}^{p\times p}$, and noticing that $C_T + \widehat{F}_0$ is the colleague matrix of the matrix polynomial $\widetilde{P}(x) = I_p T_n(x) + \sum_{k=0}^{n-1}(A_k - F_0^{(k)} + O(u^2))T_k(x)$, we have that, to first order in u, the computed eigenvalues are the exact eigenvalues of a matrix polynomial $\widetilde{P}(x) = I_p T_n(x) + \sum_{k=0}^{n-1} \widetilde{A}_k T_k(x)$, with $\|\widetilde{A}_k - A_k\|_F = \|F_0^{(k)}\|_F = \|\mathrm{tr}_p\,(EM_{k+1})\|_F = \|\sum_{i,j=1}^{n} E_{ij}(M_{k+1})_{ji}\|_F$. Therefore, for $k = 0, 1, \ldots, n-1$, we have

$$\begin{aligned}\|\widetilde{A}_k - A_k\|_F \le \sum_{i,j=1}^{n} \|E_{ij}\|_F \|(M_{k+1})_{ji}\|_F &\le \sqrt{\sum_{i,j=1}^{n} \|E_{ij}\|_F^2}\sqrt{\sum_{i,j=1}^{n} \|(M_{k+1})_{ij}\|_F^2} \\ &= \|E\|_F \|M_{k+1}\|_F,\end{aligned}$$

Then, using part (b) of Theorem 2.4, we have

$$\begin{aligned}\|M_{k+1}\|_F &= \sqrt{\sum_{i,j=1}^{n} \|(M_{k+1})_{ij}\|_F^2} = \sqrt{\sum_{i,j=1}^{n} \|\sum_{t=0}^{n} \alpha_{t,ij,k+1} A_t\|_F^2} \\ &\leq \sqrt{\sum_{i,j=1}^{n} \left(\sum_{t=0}^{n} \|\alpha_{t,ij,k+1} A_t\|_F\right)^2} \leq 4\sqrt{\sum_{i,j=1}^{n} \left(\sum_{t=0}^{n} \|A_t\|_F\right)^2} \\ &\leq 4n(n+1)^{1/2}\|P\|_F,\end{aligned}$$

where we have used $\sum_{t=0}^{n} \|A_t\|_F \leq (n+1)^{1/2}\|P\|_F$. Finally, using that $\|E\|_F = \epsilon\|C_T\|_F$, we get that

$$\begin{aligned}\|E\|_F &\leq \frac{\epsilon}{2}\left\|\begin{bmatrix} -A_{n-1} & -A_{n-1} & \cdots & -A_0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\right\|_F + \frac{\epsilon}{2}\left\|\begin{bmatrix} & I_p & & & \\ I_p & & I_p & & \\ & \ddots & & \ddots & \\ & & I_p & 0 & I_p \\ & & & 2I_p & \end{bmatrix}\right\|_F \\ &\leq \frac{\epsilon}{2}\left(\sqrt{\sum_{t=0}^{n-1} \|A_t\|_F^2} + \sqrt{2np}\right) \leq \epsilon\frac{\sqrt{2n}}{2}\left(\sqrt{\sum_{t=0}^{n-1} \|A_t\|_F^2} + \sqrt{p}\right) \leq \epsilon\sqrt{2n}\|P\|_F.\end{aligned}$$

Thus, the computed eigenvalues, to first order in u, are the exact eigenvalues of a monic matrix polynomial in the Chebyshev basis $\widetilde{P}(x)$ such that,

$$\begin{aligned}\|\widetilde{P} - P\|_F =& \sqrt{\sum_{k=0}^{n-1} \|\widetilde{A}_k - A_k\|_F^2} \leq \sum_{k=0}^{n-1} \|\widetilde{A}_k - A_k\|_F \leq \sum_{k=0}^{n-1} \|M_{k+1}\|_F \|E\|_F \\ \leq& 4n^2(n+1)^{1/2}\|P\|_F\|E\|_F \leq \widetilde{\epsilon}\|P\|_F^2,\end{aligned}$$

where $\tilde{\epsilon} = u\widehat{h}(n)$, for some low degree polynomial $\widehat{h}$ with moderate coefficients. □

Remark 2.9. *When the polynomial is scalar ($m = 1$), Corollary 2.8 can be proved without the use of Arnold transversality theorem, using a different argument that is sketched in the following lines. Let us suppose that the roots of a scalar polynomial $p(x) = T_n(x) + \sum_{k=0}^{n} a_k T_k(x)$ are computed as the eigenvalues of its colleague matrix C_T with a backward stable algorithm. Then, the computed roots are the exact eigenvalues of small perturbation of the colleague matrix $C_T + E$, that is, they are the exact roots of $\widetilde{p}(x) = \det(xI - C_T - E)$. In the spirit of* [8]*, combining* (2.11) *with Jacobi's formula for the derivative of a determinant, we get*

$$\begin{aligned}\widetilde{p}(x) =& p(x) - \operatorname{tr}\left(\operatorname{adj}\left(xI - M\right)E\right) + O\left(\|E\|_F^2\right) = \\ & T_n(x) + \sum_{k=0}^{n-1} \left(a_k - \operatorname{tr}(M_{k+1}E)\right) T_k(x) + O\left(\|E\|_F^2\right),\end{aligned}$$

where $M_1, M_2, \ldots, M_n$ are the Clenshaw matrices of $p(x)$. Finally, using the equation above, the norm $\|\widetilde{p} - p\|_2$ may be bounded as we did in the final part of the proof of Corollary 2.8.

In fact, a similar argument may be used when the roots of a nonmonic scalar polynomial $p(x) = \sum_{k=0}^{n} a_k T_k(x)$ are computed as the generalized eigenvalues of its

colleague pencil [18] *using, for example, the QZ algorithm, allowing one to recover the results in* [18, Theorem 3.3]. *Unfortunately, unless the leading coefficient is invertible, this approach does not work in the matrix polynomial case, so we do not pursue these ideas further.*

3. Backward error growth with the norm and the degree of the polynomial

In the previous section we have analyzed the backward stability of polynomial eigenvalue algorithms based on the QR algorithm applied to the colleague matrix (1.3), and we have derived the polynomial backward error upper bound (see Corollary 2.8)

$$\|\widetilde{P} - P\|_F \leq 8\epsilon n^2 (n(n+1))^{1/2} \|P\|_F^2, \tag{3.1}$$

where $\epsilon = O(u)$ is any theoretical bound for the matrix backward error coming from the QR algorithm. In this section, we provide numerical experiments to show whether or not the upper bound (3.1) correctly predict the dependence of the polynomial backward error on the norm and on the degree of the polynomial. For simplicity we focus on scalar polynomials ($m = 1$).

Given a scalar polynomial $p(x) = T_n(x) + \sum_{k=0}^{n-1} a_k T_{k-1}(x)$, to examine the tightness of the bound (3.1) we compute its roots by forming its colleague matrix C_T and computing the eigenvlalues of C_T via the Matlab command `eig(C_T)`. If we denote by $\{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n\}$ the computed eigenvalues of C_T, we then compute the backward error by forming $\widetilde{p}(x) = \prod_{k=1}^{n}(x - \widetilde{x}_k)$ and expanding it in the Chebyshev basis with the help of the Chebfun software [24].

In the first set of numerical experiments we study the dependence of the polynomial backward error $\|\widetilde{p} - p\|_2$ on the norm $\|p\|_2$ (recall that according to Corollary 2.8 this dependence should be quadratic). To this end, we proceed as follows. For each $k = 2, 3, \ldots, 10$, we generate 100 random degree-10 polynomials with 2-norm equal to 10^k. The coefficients of these polynomials are generated via the Matlab commands `p=randn(10)` and `p=10^k*p/norm(p)`. Then, for each polynomial we compute the backward error $\|\widetilde{p} - p\|_2$ when its roots are computed via the eigenvalues of its colleague matrix.

In Figure 1 we plot the maximum backward error obtained for each of the 9 samples of 100 random polynomials against the norm of the polynomials. In addition, we also compare them with the $O(\|p\|_2^2)$ trend predicted by Corollary 2.8. A linear fitting of the data gives, more precisely, a growth as $\|p\|_2^{1.95}$, which is consistent with the theory.

In the second set of numerical experiments we study the dependence of the polynomial backward error $\|\widetilde{p} - p\|_2$ on the degree of $p(x)$, when the norm of the polynomial is fixed to 1. Writing $\epsilon = n^\tau u$, where u is the unit roundoff, notice that (3.1) predicts an upper bound $O(n^{3+\tau})$. To examine the tightness of this bound, for each $n = 10, 12, 14, \ldots, 100$, we generate 100 random degree-n polynomials with 2-norm equal to 1. The coefficients of these polynomials are generated via the Matlab commands `p=randn(n)` and `p=p/norm(p)`. Then, for each polynomial we compute the backward error $\|\widetilde{p} - p\|_2$ when its roots are computed via the eigenvalues of its colleague matrix.

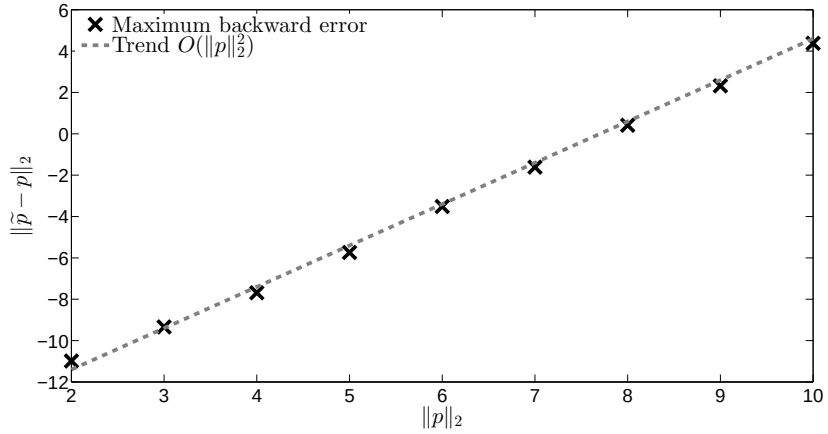


FIGURE 1. Maximum backward errors obtained for each of the 9 samples of 100 random degree-10 polynomials with fixed 2-norm equal to 10^k, when their roots are computed as the eigenvalues of their colleague matrices.

In Figure 2 we plot the maximum backward error obtained for each of the 46 samples of 100 random polynomials against the degree of the polynomials. In addition, we also compute a linear fitting for the logarithms of the maximum backward errors to get the asymptotic dependence with n. As can be seen in Figure 2, these backward errors behave like $n^{1.81}$, which means that our bound (which accounts for the worst case scenario) is overestimating the polynomial backward errors in these cases.

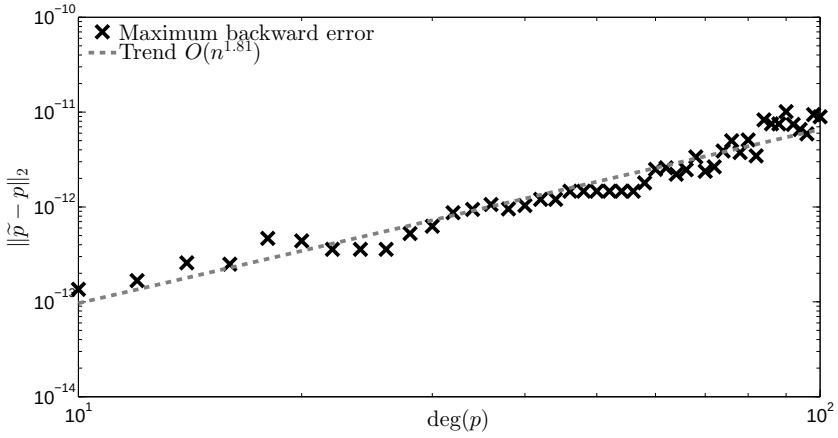


FIGURE 2. Maximum backward errors obtained for each of the 46 samples of 100 random degree-n polynomials with fixed 2-norm equal to 1, when their roots are computed as the eigenvalues of their colleague matrices.

4. Conclusions

In this paper, we have analyzed the backward stability of a Chebyshev-basis polynomial rootfinder (or matrix polynomial eigensolver) based on the solution of the standard eigenvalue problem for the corresponding colleague matrix. More precisely, given a monic scalar polynomial in the Chebyshev basis $p(x)$, we have proved that if the roots of $p(x)$ are computed as the eigenvalues of a colleague matrix using a backward stable eigenvalue algorithm, like the QR algorithm, then the computed roots are the exact roots of a monic polynomial in the Chebyshev basis $\widetilde{p}(x)$ such that

$$\frac{\|\widetilde{p}-p\|_2}{\|p\|_2} = O(u)\|p\|_2,$$

Similarly, if the eigenvalues of a monic matrix polynomial in the Chebyshev basis are computed as the eigenvalues of a block colleague matrix using a backward stable eigenvalue algorithm, then the computed eigenvalues are the exact eigenvalues of a monic matrix polynomial in the Chebyshev basis $\widetilde{P}(x)$ such that

$$\frac{\|\widetilde{P}-P\|_F}{\|P\|_F} = O(u)\|P\|_F,$$

These backward error analysis show that these methods are backward stable when the norms $\|p\|_2$ and $\|P\|_F$ are moderate.

5. Acknowledgements

We would like to thank two anonymous referees for their careful reading of the manuscript that lead to an improved presentation. We are also grateful to Froilán Dopico for pointing out a subtlety in our original proof of Corollary 2.8; following his observation, we have filled this gap.

Appendix A. Proof of Theorem 2.1

In this section we present the proof of Theorem 2.1, that is, given the Clenshaw shift $H_{n-k}(x)$ associated with the matrix polynomial $P(x)$ in (1.2), and the Chebyshev polynomial $T_{n-i}(x)$, we show that

$$T_{n-i}(x)H_{n-k}(x) = Q_{ik}(x) + r_{ik}(x)P(x), \tag{A.1}$$

for some scalar polynomial $r_{ik}(x)$, where $Q_{ik}(x)$ is the matrix polynomial of degree less than or equal to $n-1$ in (2.4)–(2.7). Moreover, we show that the decomposition (A.1) is unique.

Along the proof, quite often products of two of Chebyshev polynomials will occur. For this reason, the following formula [1, Chapter 22] is of fundamental importance here:

$$2T_m(x)T_n(x) = T_{m+n}(x) + T_{|m-n|}(x). \tag{A.2}$$

The first step is to expand the Clenshaw shifts $H_k(x)$, for $k = 0, 1, \dots, n-1$, in the Chebyshev basis. We will prove

$$H_k(x) = \sum_{\ell=0}^{k-1} 2\Gamma_\ell T_{k-\ell}(x) + \Gamma_k T_0(x), \quad \text{for } k = 0, 1, \dots, n-2, \quad \text{and} \tag{A.3}$$

$$H_{n-1}(x) = \sum_{\ell=0}^{n-2} \Gamma_\ell T_{n-1-\ell}(x) + \frac{1}{2}\Gamma_{n-1} T_0(x), \tag{A.4}$$

where Γ_ℓ is defined in (2.3). The proof proceeds by induction on k. From (2.1) we get $H_0(x) = 2I_p = \Gamma_0 T_0(x)$ and $H_1(x) = 4I_p x + 2A_{n-1} = 2\Gamma_0 T_1(x) + \Gamma_1 T_0(x)$, so the result is true for $k = 0$ and $k = 1$. Then, assume that the result is true for $H_0(x), H_1(x), \dots, H_{k-1}(x)$, with $2 \le k \le n-2$. Using the induction hypothesis, together with (2.1), we have

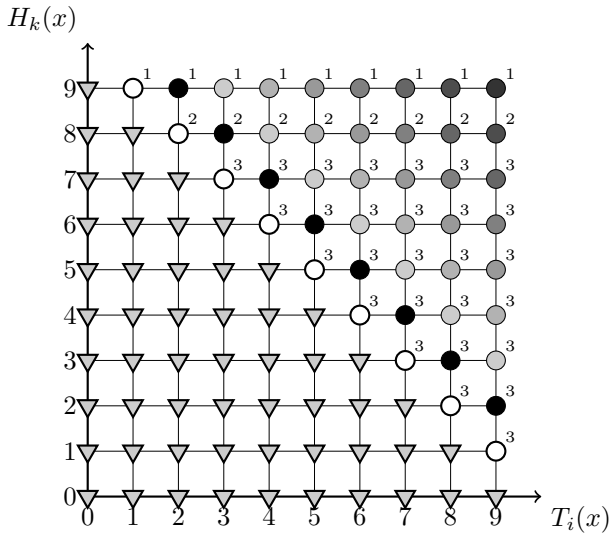
$$\begin{aligned} H_k(x) =& 2xH_{k-1}(x) - H_{k-2}(x) + 2A_{n-k} \\ =& 2x\left(\sum_{\ell=0}^{k-2} 2\Gamma_l T_{k-1-\ell}(x) + \Gamma_{k-1}T_0(x)\right) - \sum_{\ell=0}^{k-3} 2\Gamma_\ell T_{k-2-\ell}(x) - \Gamma_{k-2}T_0(x) + \\ & 2A_{n-k}. \end{aligned}$$

Using $T_0(x) = 1$, $T_1(x) = x$, and (A.2) with $m = 1$ and $n = k$, from the previous equation we get

$$\begin{aligned} H_k(x) =& \sum_{\ell=0}^{k-2} 2\Gamma_\ell \left(T_{k-\ell}(x) + T_{k-2-\ell}(x)\right) + 2\Gamma_{k-1}T_1(x) \\ & - \sum_{\ell=0}^{k-3} 2\Gamma_\ell T_{k-2-\ell}(x) - \Gamma_{k-2}T_0(x) + 2A_{n-k}T_0(x) \\ =& \sum_{\ell=0}^{k-2} 2\Gamma_\ell T_{k-\ell}(x) + 2\Gamma_{k-2}T_0(x) + 2\Gamma_{k-1}T_1(x) - \Gamma_{k-2}T_0(x) + 2A_{n-k}T_0(x) \\ =& \sum_{\ell=0}^{k-1} 2\Gamma_\ell T_{k-\ell}(x) + (\Gamma_{k-2} + 2A_{n-k})T_0(x) = \sum_{\ell=0}^{k-1} 2\Gamma_\ell T_{k-\ell}(x) + \Gamma_k T_0(x), \end{aligned}$$

where in the last equality we have used $\Gamma_{k-2} + 2A_{n-k} = \Gamma_k$. Therefore, the result is also true for $H_k(x)$. Finally, the proof that (A.4) holds is similar to the previous one, but starting with $H_{n-1}(x) = xH_{n-2}(x) - H_{n-3}(x)/2 + A_1$, so we omit the details.

Now we proceed to show that (A.1) holds with $Q_{ik}(x)$ as in (2.4)–(2.7). In order to do that, we will proceed in certain order. To help the reader to follow the steps, we depict all the possible products $T_{n-i}(x)H_{n-k}(x)$ for $n = 10$ in the following 10×10 grid.



The vertices with triangular shape in the previous grid represent the cases in which the degree of $T_{n-i}(x)H_{n-k}(x)$ does not exceed $n-1$, that is, when $i \geq n-k+1$. In this case, the polynomial $Q_{ik}(x)$ coincide with $T_{n-i}(x)H_{n-k}(x)$, so we just need to expand $T_{n-i}(x)H_{n-k}(x)$ in the Chebyshev basis. Indeed, when $i = n$ and $k = 1$, from (A.4), we have

$$T_0(x)H_{n-1}(x) = H_{n-1}(x) = \sum_{\ell=0}^{n-2} \Gamma_\ell T_{n-1-\ell}(x) + \frac{1}{2}\Gamma_{n-1}T_0(x),$$

and when $n-1 \geq i \geq n-k+1$, from (A.2) and (A.3), we have

$$\begin{aligned} T_{n-i}(x)H_{n-k}(x) &= \sum_{\ell=0}^{n-k-1} 2\Gamma_\ell T_{n-i}(x)T_{n-k-\ell}(x) + \Gamma_{n-k}T_{n-i}(x)T_0(x) \\ &= \sum_{\ell=0}^{n-k-1} \Gamma_\ell (T_{2n-i-k-\ell}(x) + T_{|k+\ell-i|}(x)) + \Gamma_{n-k}T_{n-i}(x). \end{aligned}$$

As can be checked, the two previous equations correspond to (2.4) and (2.5), respectively.

Next, we consider the products $T_{n-i}(x)H_{n-k}(x)$ with $i < n-k+1$, represented in the grid by vertices with circular shape. This case is much more involved, since the degree of $T_{n-i}(x)H_{n-k}(x)$ is larger than or equal to n. We will prove that (A.1) holds, with $Q_{ik}(x)$ as in (2.4)–(2.7), each diagonal in the grid at a time (from left to right), showing that each product $T_{n-i}(x)H_{n-k}(x)$ can be computed using, at most, a product represented by a vertex in the same diagonal and two products represented by vertices in the diagonal on its left.

The first step is to consider the products $T_k(x)H_{n-k}(x)$, for $k = 1, 2, \ldots, n-1$, that is, products represented by the diagonal with white circular vertices in the grid. We show that Theorem 2.1 holds for those products from top to bottom. We start with the white circular vertex labeled with 1 in the grid, that is, with the

product $T_1(x)H_{n-1}(x)$. From (2.2) and (A.3), together with $T_1(x) = x$, we have

$$\begin{aligned} T_1(x)H_{n-1}(x) =& xH_{n-1}(x) = \frac{1}{2}H_{n-2}(x) - A_0T_0(x) + \cdots \\ =& \sum_{\ell=0}^{n-3} \Gamma_\ell T_{n-2-\ell}(x) + \frac{1}{2}\Gamma_{n-2}T_0(x) - A_0T_0(x) + \cdots, \end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. As can be easily checked, the previous equation corresponds to (2.7) with $i = n - 1$.

Then, we consider the white circular vertex labeled with 2 in the grid, that is, the product $T_2(x)H_{n-2}(x)$. From (1.1) and (2.1), we have

$$\begin{aligned} H_{n-2}(x)T_2(x) &= H_{n-2}(x)\left(2xT_1(x) - T_0(x)\right) = 2xT_1(x)H_{n-2}(x) - T_0(x)H_{n-2}(x) \\ &= T_1(x)\left(2H_{n-1}(x) + H_{n-3}(x) - 2A_1\right) - T_0(x)H_{n-2}(x) \\ &= 2T_1(x)H_{n-1}(x) + T_1(x)H_{n-3}(x) - T_0(x)H_{n-2}(x) - 2A_1T_1(x). \end{aligned}$$

As can be seen from the previous equation, the product $H_{n-2}(x)T_2(x)$ may be computed from products represented by two triangular vertices: $T_1(x)H_{n-3}(x)$ and $T_0(x)H_{n-2}(x)$, and the product $T_1(x)H_{n-1}(x)$. Then, using (A.3), (A.4), and the result previously obtained for $T_1(x)H_{n-1}(x)$, we get

$$\begin{aligned} T_2(x)H_{n-2}(x) =& \sum_{\ell=0}^{n-4} \Gamma_\ell(T_{n-2-\ell}(x) + T_{|\ell+4-n|}(x)) + \\ & \Gamma_{n-3}T_1(x) - 2A_0T_0(x) - 2A_1T_1(x) + \cdots, \end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. The previous equation corresponds to (2.6) with $i = n - 2$ and $k = 2$.

Finally, we consider the white circular vertices labeled with 3, that is, the products $T_k(x)H_{n-k}(x)$, for $k = 3, 4, \ldots, n$. From (1.1) and (2.1), we have

$$\begin{aligned} T_k(x)H_{n-k}(x) =& (2xT_{k-1}(x) - T_{k-2}(x))H_{n-k}(x) \\ =& 2xT_{k-1}(x)H_{n-k}(x) - T_{k-2}(x)H_{n-k}(x) \\ =& T_{k-1}(x)(H_{n-k+1}(x) + H_{n-k-1}(x) - 2A_{k-1}) - T_{k-2}(x)H_{n-k}(x) \\ =& T_{k-1}(x)H_{n-k+1}(x) + T_{k-1}(x)H_{n-k-1}(x) - T_{k-2}(x)H_{n-k}(x) \\ & - 2A_{k-1}T_{k-1}(x). \end{aligned}$$

As can be seen from the previous equation, $T_k(x)H_{n-k}(x)$ may be computed from $T_{k-1}(x)H_{n-k-1}(x)$ and $T_{k-2}(x)H_{n-k}(x)$, represented in the grid by triangular vertices, and $T_{k-1}(x)H_{n-k+1}(x)$, represented in the grid by the white circular vertex above the white circular vertex corresponding to $T_k(x)H_{n-k}(x)$. Since we have previously seen that Theorem 2.1 holds for $T_1(x)H_{n-1}(x)$ and $T_2(x)H_{n-2}(x)$, and for products represented by triangular vertices, this shows how to prove inductively (from top to bottom) that Theorem 2.1 holds for products represented by white circular vertices labeled with 3. Indeed, assuming that the result holds for

$T_{k-1}(x)H_{n-k+1}(x)$ and using (2.4), we get

$$\begin{aligned} T_k(x)H_{n-k}(x) =& T_{k-1}(x)H_{n-k-1} + \sum_{r=2}^{k}(-2A_{k-r})T_{k-r}(x) - \\ & 2A_{k-1}T_{k-1}(x) + \cdots = \sum_{\ell=0}^{n-k-2} \Gamma_\ell (T_{n-2-\ell}(x) + T_{|\ell+2k-n|}(x)) + \\ & \Gamma_{n-k-1}T_{k-1}(x) + \sum_{r=1}^{k}(-2A_{k-r})T_{k-r}(x) + \cdots , \end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. It is immediate to check that the previous equation corresponds to (2.6) when $i = n - k$.

The second step is to consider the products $T_{k+1}(x)H_{n-k}(x)$, for $k = 2, 3, \ldots, n-2$, that is, the diagonal with black circular vertices in the grid. This step is very similar to the previous one, so we will only sketch the main ideas. We have to distinguish the cases $k = 2$, $k = 3$ and $k > 3$. When $k = 2$, using (1.1), (2.1) and (2.2), it may be proved

$$T_2(x)H_{n-1}(x) = T_1(x)H_{n-2}(x) - T_0(x)H_{n-1}(x) - 2A_0T_1(x) + \cdots ,$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. The previous equation shows that $T_2(x)H_{n-1}(x)$ may be computed from two products represented by triangular vertices in the grid: $T_1(x)H_{n-2}(x)$ and $T_0(x)H_{n-1}(x)$. Since we have seen that Theorem 2.1 holds for products represented by triangular vertices, it may be proved that (2.7) holds for $T_2(x)H_{n-1}(x)$.

Then, from (1.1) and (2.1), it may be proved that, when $k = 2$,

$$T_3(x)H_{n-2}(x) = 2T_2(x)H_{n-1}(x) + T_2(x)H_{n-3}(x) - 2A_1T_2(x),$$

and, when $k > 3$,

$$\begin{aligned} T_{k+1}(x)H_{n-k}(x) =& T_k(x)H_{n-k+1}(x) + T_k(x)H_{n-k-1}(x) - \\ & T_{k-1}(x)H_{n-k}(x) - 2A_{k-1}T_k(x). \end{aligned}$$

These two equations show that $T_{k+1}(x)H_{n-k}(x)$ may be computed from two products represented by triangular vertices, and the product represented by the black circular vertex above the black circular vertex corresponding to $T_{k+1}(x)H_{n-k}(x)$. Assuming that Theorem 2.1 holds for $T_2(x)H_{n-1}(x)$, the previous observation shows how to prove inductively (from top to bottom) that Theorem 2.1 holds for products corresponding to black circular vertices labeled with 2 and 3.

Now, we address the products represented by circular vertices colored with different shades of grey, that is, the products $T_{k+r-1}(x)H_{n-k}(x)$, for $r = 3, 4, \ldots, n-2$ and $k = 1, 2, \ldots, n-1-r$. We will show that Theorem 2.1 holds for products represented by vertices in the same diagonal (same shade of grey) assuming that it holds for products represented by (non-triangular) vertices in the diagonal on its left. Since we have previously proved that Theorem 2.1 holds for products represented by the white and black diagonals, this will imply that Theorem 2.1 holds for all products represented by grey vertices. For each grey diagonal, we have to distinguish the products represented by vertices labeled with 1, 2, and 3.

First, we consider the product $T_r(x)H_{n-1}(x)$, with $r \geq 3$, represented by a grey vertex labeled with 1. From (1.1) and (2.2), we get

$$\begin{aligned} T_r(x)H_{n-1}(x) =& (2xT_{r-1}(x) - T_{r-2}(x))H_{n-1}(x) = 2xT_{r-1}(x)H_{n-1}(x) - \\ & T_{r-2}(x)H_{n-1}(x) \\ =& T_{r-1}(x)H_{n-2}(x) - T_{r-2}(x)H_{n-1}(x) - 2A_0T_{r-1}(x) + \cdots, \end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. The previous equation shows that $T_r(x)H_{n-1}(x)$ may be computed from two products represented by vertices in the diagonal on its left: $T_{r-1}(x)H_{n-2}(x)$ and $T_{r-2}(x)H_{n-1}(x)$. Assuming that (2.6) and (2.7) hold for those products, we have

$$\begin{aligned} T_{r-1}(x)H_{n-2}(x) = & \sum_{\ell=0}^{n-r-1} \Gamma_\ell \left(T_{n-r+1-\ell}(x) + T_{|\ell-n+r+1|}(x)\right) + \Gamma_{n-r}T_1(x) \\ & - \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell+1} 2A_{\ell+1-s}T_{|r-\ell-s|}(x) + \cdots, \end{aligned}$$

and

$$\begin{aligned} T_{r-2}(x)H_{n-1}(x) = & \sum_{\ell=0}^{n-r} \Gamma_\ell T_{n-r+1-\ell}(x) + \frac{1}{2}\Gamma_{n-r+1}T_0(x) - \\ & \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell} A_{\ell-s}T_{|r-\ell-s-1|}(x) + \cdots, \end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. Using

$$\begin{aligned} & \sum_{\ell=0}^{n-r-1} \Gamma_\ell \left(T_{n-r+1-\ell}(x) + T_{|\ell-n+r+1|}(x)\right) + \Gamma_{n-r}T_1(x) - \sum_{\ell=0}^{n-r} \Gamma_\ell T_{n-r+1-\ell}(x) \\ & - \frac{1}{2}\Gamma_{n-r+1}T_0(x) = \sum_{\ell=0}^{n-r-2} \Gamma_\ell T_{n-r-1-\ell}(x) + \frac{1}{2}\Gamma_{n-r-1}T_0(x) - A_{r-1}T_0(x), \end{aligned}$$

where we have used $(\Gamma_{n-r+1} - \Gamma_{n-r-1})/2 = A_{r-1}$, and

$$\begin{aligned} & - \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell+1} 2A_{\ell+1-s}T_{|r-\ell-s|}(x) + \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell} A_{\ell-s}T_{|r-\ell-s-1|}(x) \\ = & - \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell+1} A_{\ell+1-s}T_{|r-\ell-s|}(x) - \sum_{\ell=1}^{r-2} A_\ell T_{|r-\ell-1|}(x) \\ = & - \sum_{\ell=1}^{r-2}\sum_{s=1}^{\ell+1} A_{\ell+1-s}T_{|r-\ell-s|}(x) - \sum_{s=1}^{r} A_{r-s}T_{|s-1|}(x) + A_{r-1}T_0(x) + A_0T_{r-1}(x) \\ = & - \sum_{\ell=0}^{r-1}\sum_{s=1}^{\ell+1} A_{\ell+1-s}T_{|r-\ell-s|}(x) + A_{r-1}T_0(x) + 2A_0T_{r-1}(x) \\ = & - \sum_{\ell=1}^{r}\sum_{s=1}^{\ell} A_{\ell-s}T_{|r+1-\ell-s|}(x) + A_{r-1}T_0(x) + 2A_0T_{r-1}(x), \end{aligned}$$

we get

$$T_{r-1}(x)H_{n-2}(x) = \sum_{\ell=0}^{n-r-2} \Gamma_\ell T_{n-r-1-\ell}(x) + \frac{1}{2}\Gamma_{n-r-1}T_0(x) - \sum_{\ell=1}^{r}\sum_{s=1}^{\ell} 2A_{\ell-s}T_{|r+1-\ell-s|}(x) + \cdots,$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial. As can be checked, the previous equation corresponds to (2.7) with $k = 1$ and $i = n - r$.

The proof that Theorem 2.1 holds for products represented by grey vertices labeled with 2 is very similar to the previous one, so we omit it.

Finally, consider a product $T_{n-i}(x)H_{n-k}(x)$ represented by a grey vertex labeled with 3. From (1.1) and (2.1), we have

$$\begin{aligned} T_{n-i}(x)H_{n-k}(x) =& (2xT_{n-i-1}(x) - T_{n-i-2}(x))H_{n-k}(x) = T_{n-i-1}(x)H_{n-k+1}(x) \\ &+ T_{n-i-1}(x)H_{n-k-1}(x) - T_{n-i-2}(x)H_{n-k}(x) - 2A_{k-1}T_{n-i-1}(x) \end{aligned}$$

The previous equation shows that $T_{n-i}(x)H_{n-k}(x)$ may be computed from two products represented by (non-triangular) vertices in the diagonal on its left: $T_{n-i-1}(x)H_{n-k-1}(x)$ and $T_{n-i-2}(x)H_{n-k}(x)$, and a product represented by a vertex in the same diagonal, above the vertex corresponding to $T_{n-i}(x)H_{n-k}(x)$, that is, the product $T_{n-i-1}(x)H_{n-k+1}(x)$. This observation shows how to prove inductively (from top to bottom) that Theorem 2.1 holds for the grey vertices labeled with 3 in the same diagonal. Assuming that (2.6) holds for $T_{n-i-1}(x)H_{n-k-1}(x)$, $T_{n-i-2}(x)H_{n-k}(x)$ and $T_{n-i-2}(x)H_{n-k}(x)$, and using $\Gamma_{i+1} - \Gamma_{i-1} = 2A_{n-i-1}$,

$$\begin{aligned} &\sum_{\ell=0}^{i-1} \Gamma_\ell (T_{i+k-2-\ell}(x) + T_{|k+\ell-i-2|}(x)) + \Gamma_i T_{k-2}(x) + \\ &\sum_{\ell=0}^{i-1} \Gamma_\ell (T_{i+k-\ell}(x) + T_{|k+\ell-i|}(x)) \\ &+ \Gamma_i T_k(x) - \sum_{\ell=0}^{i} \Gamma_\ell (T_{i+k-\ell}(x) + T_{|k+\ell-i-2|}(x)) - \Gamma_{i+1}T_{k-1}(x) \\ =& \sum_{\ell=0}^{i-2} \Gamma_\ell (T_{i+k-2-\ell}(x) + T_{|k+\ell-i|}(x)) + \Gamma_{i-1}T_{k-1}(x) - 2A_{n-i-1}T_{k-1}, \end{aligned}$$

and

$$\begin{aligned} &- \sum_{\ell=1}^{n-k+1-i}\sum_{r=1}^{k-2+\ell} 2A_{k-2+\ell-r}T_{|n-i-\ell-r|}(x) - \sum_{\ell=1}^{n-k-1-i}\sum_{r=1}^{k+\ell} 2A_{k+\ell-r}T_{|n-i-\ell-r|}(x) \\ &+ \sum_{\ell=1}^{n-k-1-i}\sum_{r=1}^{k-1+\ell} 2A_{k-1+\ell-r}T_{|n-i-1-\ell-r|}(x) \\ =& - \sum_{\ell=1}^{n-k+1-i}\sum_{r=1}^{k-1+\ell} 2A_{k-1+\ell-r}T_{|n-i+1-\ell-r|}(x) + 2A_{k-1}T_{n-i-1}(x) + 2A_{n-i-1}T_{k-1}(x). \end{aligned}$$

we get

$$\begin{aligned}
&T_{n-i}(x)H_{n-k}(x) = T_{n-i-1}(x)H_{n-k+1}(x) + T_{n-i-1}(x)H_{n-k-1}(x) - \\
&T_{n-i-2}(x)H_{n-k}(x) - 2A_{k-1}T_{n-i-1}(x) \\
=&\sum_{\ell=0}^{i-2} \Gamma_\ell (T_{i+k-2-\ell}(x) + T_{|k+\ell-i|}(x)) + \Gamma_{i-1}T_{k-1}(x) - \\
&\sum_{\ell=1}^{n-k+1-i} \sum_{r=1}^{k-1+\ell} 2A_{k-1+\ell-r}T_{|n-i+1-\ell-r|}(x) + \cdots,
\end{aligned}$$

where the dots correspond to something of the form $r(x)P(x)$, with $r(x)$ a scalar polynomial, which shows that (2.6) holds also for $T_{n-i}(x)H_{n-k}(x)$.

The final step of the proof consists in proving the uniqueness of $r_{ik}(x)$ and $Q_{ik}(x)$ in (A.1). For this purpose, assume that there exist two scalar polynomials $r_{ik}(x)$ and $\widetilde{r}_{ik}(x)$, and two matrix polynomials $Q_{ik}(x)$ and $\widetilde{Q}_{ik}(x)$ of degree at most $n-1$ such that $T_{n-i}(x)H_{n-k}(x) = Q_{ik}(x) + r_{ik}(x)P(x) = \widetilde{Q}_{ik}(x) + \widetilde{r}_{ik}(x)P(x)$. Then, $Q_{ik}(x) - \widetilde{Q}_{ik}(x) = (\widetilde{r}_{ik}(x) - r_{ik}(x))P(x)$ is a matrix polynomial of degree at most $n-1$, but, if $r_{ik}(x) \neq \widetilde{r}_{ik}(x)$, the matrix polynomial $(\widetilde{r}_{ik}(x) - r_{ik}(x))P(x)$ has degree larger than or equal to n, hence $r_{ik}(x) = \widetilde{r}_{ik}(x)$ and $Q_{ik}(x) = \widetilde{Q}_{ik}(x)$.

References

1. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Number 55. Courier Dover Publications, 1972.
2. A. Amiraslami, R. M. Corless, and P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*. IMA. J. Numer. Anal. **29** (2009), no. 1, 141–157.
3. V. I. Arnold, *On matrices depending on parameters*, Russian Matehmatical Surveys **26** (1971), no. 2, 29–43.
4. S. Barnett, *Polynomials and Linear Control Systems*. Marcel Dekker Inc., 1983.
5. S. Barnett. *Leverrier's Algorithm for Orthogonal Polynomial Bases*. Linear Algebra Appl. **236** (1996), 245–263, 1996.
6. C. W. Clenshaw, *A note on the summation of Chebyshev series*. Math. Comp. **9** (1955), 118–120.
7. F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations and the recovery of minimal indices*. SIAM J. Matrix Anal. Appl. **31** (2009/2010), no. 4, 2181–2204.
8. F. De Terán, F. M. Dopico, and J. Pérez, *Backward stability of polynomial root-finding using Fiedler companion matrices*. IMA J. Numer. Anal., in press, DOI 10.1093/imanum/dru057.
9. A. Edelman and H. Murakami, *Polynomial roots from companion matrix eigenvalues*. Math. Comp. **64** (1995), 763–776.
10. C. Effenberger and D. Kressner, *Chebyshev interpolation for nonlinear eigenvalue problems*. BIT **52** (2012), 933–951.
11. I. J. Good, *The colleague matrix, a Chebyshev analogue of the companion matrix*. Q. J. Math. **12** (1961), 61–68.
12. P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. Second Ed., Academic Press, San Diego, 1985.
13. P. W. Lawrence and R. M. Corless, *Stability of rootfinding for barycentric Lagrange interpolants*. Numer. Algorithms **65** (2014), 447–464.
14. P. W. Lawrence, M. Van Barel, and P. Van Dooren, *Structured backward error analysis of polynomial eigenvalue problems solved by linearizations*. Preprint, submitted.
15. D. Lemmonier and P. Van Dooren, *Optimal scaling of companion pencils for the QZ-algorithm*. Proceedings SIAM Appl. Lin. Alg. Conference, Paper CP7-4, 2003
16. D. Lemmonier and P. Van Dooren, *Optimal scaling of block companion pencils*. Proceedings of the International Symposium on Mathematical Theory of Networks and Systems, Leuven, Belgium, 2004.

17. J. Maroulas and S. Barnett, *Polynomials with respect to a general basis. I. Theory*. J. Math. Anal. Appl. **72** (1979), no. 1, 177–194.
18. Y. Nakatsukasa and V. Noferini, *On the stability of computing polynomial roots via confederate linearizations*. To appear in Math. Comp.
19. Y. Nakatsukasa, V. Noferini, and A. Townsend, *Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach*. Preprint, submitted.
20. V. Noferini and F. Poloni, *Duality of matrix pencils, Wong chains and linearizations*. Linear Algebra Appl. **471** (2015), 730–767.
21. B. N. Parlett and C. Reinsch, *Balancing a matrix for calculation of eigenvalues and eigenvectors*. Numer. Math. **13** (1963), 293–304.
22. G. Peters and J. H. Wilkinson, *Practical problems arising in the solution of polynomial equations*. J. Inst. Maths. Appl. **8** (1971), 16–35.
23. L. N. Trefethen, *Approximation theory and approximation practice*. SIAM, 2013.
24. L. N. Trefethen et al, Chebfun Version 5. The Chebfun Development Team, 2014. `http://www.maths.ox.ac.uk/chebfun/`.
25. K. -C. Toh and L. N. Trefethen, *Pseudozeros of polynomials and pseudospectra of companion matrices*. Numer. Math. **68** (1994), 403–425.
26. J. H. Wilkinson, *Rounding Errors in Algebraic Processes*. Prentice-Hall, Englewood Cliffs, 1963.
27. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.

School of Mathematics, The University of Manchester, Manchester, England, M13 9PL
E-mail address: `vanni.noferini@maths.manchester.ac.uk`

School of Mathematics, The University of Manchester, Manchester, England, M13 9PL
E-mail address: `javier.perezalvaro@manchester.ac.uk`