SIAM J. Matrix Anal. Appl.
Vol. 38, No. 1, pp. 188–216

A FRAMEWORK FOR STRUCTURED LINEARIZATIONS OF MATRIX POLYNOMIALS IN VARIOUS BASES*

LEONARDO ROBOL†, RAF VANDEBRIL†, AND PAUL VAN DOOREN‡

Abstract. We present a framework for the construction of linearizations for scalar and matrix polynomials based on dual bases which, in the case of orthogonal polynomials, can be described by the associated recurrence relations. The framework provides an extension of the classical linearization theory for polynomials expressed in nonmonomial bases and allows us to represent polynomials expressed in product families, that is, as a linear combination of elements of the form $\phi_i(\lambda)\psi_j(\lambda)$, where $\{\phi_i(\lambda)\}$ and $\{\psi_j(\lambda)\}$ can either be polynomial bases or polynomial families which satisfy some mild assumptions. We show that this general construction can be used for many different purposes. Among them, we show how to linearize sums of polynomials and rational functions expressed in different bases. As an example, this allows us to look for intersections of functions interpolated on different nodes without converting them to the same basis. We then provide some constructions for structured linearizations for $\star$-even and $\star$-palindromic matrix polynomials. The extension of these constructions to $\star$-odd and $\star$-antipalindromic of odd degree is discussed and follows immediately from the previous results.

Key words. matrix polynomials, rational functions, nonmonomial bases, palindromic matrix polynomials, even matrix polynomials, strong linearizations, dual minimal bases

AMS subject classifications. 15A18, 15A22, 65F15, 65H04

DOI. 10.1137/16M106296X

1. Introduction.

In recent years much interest has been devoted to finding linearizations for polynomials and matrix polynomials. The Frobenius linearization, i.e., the classical companion, has been the de-facto standard in polynomial eigenvalue problems and polynomial rootfinding for a long time [21, 23]. Nevertheless, recently much work has been put into developing other families of linearizations. Among these, some linearizations preserve spectral symmetries available in the original problem [25, 29, 31], others linearize matrix polynomials formulated in nonmonomial bases [1, 9], and some variations are based on an idea of Fiedler about decomposing companion matrices into products of simple factors [2, 10, 12, 19].

In this work we take as inspiration the results of Dopico et al. [14] that characterize the structure of some permuted Fiedler linearizations by using dual minimal bases [20]. We extend the results in a way that allows us to deal with many more formulations, and we use it to derive numerous different linearizations. We also use these examples to prove the effectiveness of this result as a tool for constructing structured linearizations (thus preserving spectral symmetries in the spirit of the works cited

*Received by the editors February 23, 2016; accepted for publication (in revised form) by T. Stykel November 9, 2016; published electronically March 16, 2017.

http://www.siam.org/journals/simax/38-1/M106296.html

Funding: The work of the first author was partially supported by INdAM/GNCS Research Project 2016. The work of the first and second authors was supported by the Research Council KU Leuven, projects C14/16/056 (Inverse-free Rational Krylov Methods: Theory and Applications), CREA/13/012 (Can Unconventional Eigenvalue Algorithms Supersede the State-of-the-Art), and PFV/10/002 (Optimization in Engineering, OPTEC). The work of the third author was supported by the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization) and funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office.

†Department of Computer Science, KU Leuven, 3001 Heverlee, Belgium (leonardo.robol@cs.kuleuven.be, raf.vandebril@cs.kuleuven.be).

‡ICTEAM, Universite Catholique de Louvain, Batiment Euler, B-1348 Louvain-la Neuve, Belgium (paul.vandooren@uclouvain.be).

above) and also linearizations for sums of polynomials and rational functions.

In particular, we consider the rootfinding problem for polynomials that are expressed as a linear combination of elements in a so-called product family; the most common case where this can be applied is when considering two different polynomial bases $\{\phi_i\}$ and $\{\psi_j\}$ and representing polynomials as sums of objects of the form $\phi_i(\lambda)\psi_j(\lambda)$. This apparently artificial construction has, however, many interesting applications, such as finding intersections of polynomials and rational functions defined in different bases. This problem arises naturally in computer aided design, where curves are defined locally as polynomials or rational functions using different interpolation bases and their intersections are needed for clipping (see [33] and the references therein). Moreover, functions defined as the sums of polynomials or rational functions in different bases are often found in the analysis of closed loop linear systems. The latter case also involves matrices of rational functions when MIMO systems are considered [26].

In section 2 we give a formal definition of what we call a *product family of polynomials*, denoted by $\phi \otimes \psi$. We define the vector $\pi_{k,\phi}(\lambda)$ to be the one with the elements of the family as entries, and we show that $\pi_{k,\phi\otimes\psi}(\lambda)$ is given by $\pi_{k,\phi}(\lambda) \otimes \pi_{k,\psi}(\lambda)$. We present a theorem that allows us to linearize every polynomial written as a linear combination of elements in a product family, and we also generalize the construction to the product of more than two families in section 2.4. We consider a certain class of *dual polynomial bases* (with the notation of the classical work by Forney [20]) of a polynomial vector $\pi_{k,\phi}(\lambda)$, which we identify with a matrix pencil $L_{k,\phi}(\lambda)$ such that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$, which will be used as a tool to build linearizations. In section 3 we introduce an explicit construction for linearizing polynomial families arising from orthogonal and interpolation bases. We cover the case of every polynomial basis endowed with a recurrence relation, and we provide explicit constructions for the Lagrange, Newton, Hermite, and Bernstein cases. We describe the dual bases for all these cases and, as shown by Theorem 15, they are the only ingredient required to build the linearizations.

The rest of the paper deals with the problem of exploiting this freedom of choice to obtain many interesting results.

In section 4 we show how to linearize the sum of two scalar polynomials or rational functions expressed in different bases, without the need of an explicit basis conversion. This can have important applications in the cases where interpolation polynomials are obtained from experimental data (that cannot be resampled—so there is no choice for the interpolation basis) or in cases where an explicit change of basis is badly conditioned.

Infinite eigenvalues may appear when linearizing the sums of polynomials. We report numerical experiments that show that they do not affect the numerical robustness of the approach in many cases. Moreover, we show that for the rational case, under mild hypotheses, it is possible to construct strong linearizations which do not have spurious infinite eigenvalues.

In section 5 we turn our attention to preserving spectral symmetries and provide explicit constructions for linearizations of $\star$-even/odd and $\star$-palindromic matrix polynomials. We show that a careful choice of the dual bases for use in Theorem 15 yields linearizations with the same spectral symmetries of the original matrix polynomial. Finally, in section 6, we describe a numerical approach to deflate the infinite eigenvalues that are present in some of the constructions, based on the staircase algorithm of Van Dooren [34]. In section 7 we draw some conclusions and propose some possible developments for future research.

2. A general framework to build linearizations.

2.1. Notation. In the following we will often work with the vector space of polynomials of degree at most k on a field $\mathbb{F}$, denoted as $\mathbb{F}_k[\lambda]$. We will denote by $\overline{\mathbb{F}}$ its algebraic closure.

In the study of strong linearizations it is also important to consider the rev operator, which reverses the coefficients of the polynomial when represented in the monomial basis.

Definition 1. *Given a nonzero matrix polynomial $P(\lambda) = \sum_{i=0}^{k} P_i \lambda^i$, we define its* degree *as the largest integer $i \geqslant 0$ such that $P_i \neq 0$, that is, the maximum of all the degrees of the entries of $P(\lambda)$. We denote it by* $\deg P(\lambda)$.

Definition 2. *Given a matrix polynomial $P(\lambda)$, its* reversed polynomial, *denoted by* $\operatorname{rev} P(\lambda)$, *is defined by* $\operatorname{rev} P(\lambda) := x^{\deg P(\lambda)} P(\lambda^{-1})$. *We often refer to* $\operatorname{rev} P(\lambda)$ *as the* reversal *of $P(\lambda)$.*

Intuitively, a linearization for a matrix polynomial $P(\lambda)$ is a matrix pencil $L(\lambda)$ such that $L(\lambda)$ is singular only when $P(\lambda)$ is. However, this is not sufficient in most cases since there is also the need to match eigenvectors and partial multiplicities, so the definition has to be a little more involved. We refer the reader to the work of De Terán, Dopico, and Mackey [13] for a complete overview of the subject.

Definition 3. *An $m \times m$ matrix polynomial $E(\lambda)$ is said to be* unimodular *if it is invertible in the ring of $m \times m$ matrix polynomials or, equivalently, if* $\det E(\lambda)$ *is a nonzero constant of $\mathbb{F}$.*

Definition 4 (extended unimodular equivalence). *Let $P(\lambda)$ and $Q(\lambda)$ be matrix polynomials. We say that they are* extended unimodularly equivalent *if there exist positive integers r, s and two unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ of appropriate dimensions such that*

$$E(\lambda) \begin{bmatrix} I_r & \\ & P(\lambda) \end{bmatrix} F(\lambda) = \begin{bmatrix} I_s & \\ & Q(\lambda) \end{bmatrix}.$$

Definition 5 (linearization). *A matrix pencil $\mathcal{L}(\lambda)$ is a* linearization *for a matrix polynomial $P(\lambda)$ if $P(\lambda)$ is extended unimodularly equivalent to $\mathcal{L}(\lambda)$.*

In order to preserve the complete eigenstructure of a matrix polynomial, it is of interest to also maintain the infinite eigenvalues, which are defined as the zero eigenvalues of the reversed polynomial. To achieve this we have to extend Definition 5.

Definition 6 (spectral equivalence). *Two matrix polynomials $P(\lambda)$ and $Q(\lambda)$ are* spectrally equivalent *if $P(\lambda)$ is extended unimodularly equivalent to $Q(\lambda)$ and* $\operatorname{rev} P(\lambda)$ *is extended unimodularly equivalent to* $\operatorname{rev} Q(\lambda)$.

Definition 7 (strong linearization). *A matrix pencil $\mathcal{L}(\lambda)$ is said to be a* strong linearization *for a matrix polynomial $P(\lambda)$ if it is spectrally equivalent to $\mathcal{L}(\lambda)$.*

2.2. Working with product families of polynomials. The linearizations that we build in this work concern polynomials expressed as linear combinations of elements of a product family. Let us add more details about this concept.

By the term *family of polynomials* (or *polynomial family*) we mean any set of elements in $\mathbb{F}[\lambda]$ indexed on a finite totally ordered set $(I, \leqslant)$. To denote these objects we use the notation $\{\phi_i(\lambda) \mid i \in I\}$ or its more compressed form $\{\phi_i(\lambda)\}$ or even $\{\phi_i\}$ whenever the index set I and the variable λ are clear from the context. Often the set I will be a segment of the natural numbers or a subset of $\mathbb{N}^d$ endowed with the lexicographical order, as in Definition 8.

An important example of such families are the polynomials $\phi_i(\lambda)$ forming a basis for the polynomials of degree up to k. Another extension that deserves our attention is the following.

Definition 8. *Given two families of polynomials $\{\phi_i\}$ for $i = 0, \ldots, \epsilon$ and $\{\psi_j\}$ for $j = 0, \ldots, \eta$, we define the* product family *as the indexed set defined by*

$$\phi \otimes \psi := \{\phi_i(\lambda)\psi_j(\lambda),\ i = 0, \ldots, \epsilon,\ j = 0, \ldots, \eta\},$$

with the lexicographical order (so that $(i, j) \leqslant (i', j')$ if either $i < i'$ or $i = i'$ and $j \leqslant j'$).

We introduce some notation that will make it easier in the following to deal with these product families and their use in linearizations. We use the symbol $\pi_{k,\phi}(\lambda)$ to denote the column vector

$$\pi_{k,\phi}(\lambda) := \begin{bmatrix} \phi_k(\lambda) \\ \vdots \\ \phi_0(\lambda) \end{bmatrix}.$$

We will often identify $\pi_{k,\phi}(\lambda)$ with the family $\{\phi_i \mid i = 0, \ldots, k\}$ since they are just different representations of the same mathematical object.

Notice that Definition 8 is easily extendable to the product of an arbitrary number of families. In this case we always consider the lexicographical order on the new family, which is particularly convenient because then we have

$$\pi_{k,\phi^{(1)}\otimes\cdots\otimes\phi^{(j)}}(\lambda) = \pi_{\epsilon_1,\phi^{(1)}}(\lambda) \otimes \cdots \otimes \pi_{\epsilon_j,\phi^{(j)}}(\lambda).$$

Remark 9. Whenever the family $\{\phi_i\}$ is a basis for the polynomials of degree at most k, every polynomial $p(\lambda) \in \mathbb{F}_k[\lambda]$ can be expressed as

$$p(\lambda) = \sum_{j=0}^{k} a_j \phi_j(\lambda).$$

In particular, the scalar product with $\pi_{k,\phi}(\lambda)$ is a linear isomorphism between $\mathbb{F}^{k+1}$ and the vector space of polynomials of degree at most k. We have

$$\begin{array}{rccc} \Gamma_\phi : & \mathbb{F}^{k+1} & \longrightarrow & \mathbb{F}_k[\lambda], \\ & a & \longmapsto & \Gamma_\phi(a) := a^\top \pi_{k,\phi}(\lambda). \end{array}$$

With the above notation $\Gamma_\phi^{-1}(p(\lambda))$ is the vector of coordinates of $p(\lambda)$ expressed in the basis $\{\phi_i\}$.

We recall the following definitions that can be found in [20].

Definition 10. *A matrix polynomial $G(\lambda) \in \mathbb{F}[\lambda]^{k\times n}$ is a* polynomial basis *if its rows are a basis for a subspace of the vector space of polynomial n-tuples.*

Definition 11 (dual basis). *Two polynomial bases $G(\lambda) \in \mathbb{F}[\lambda]^{k\times n}$ and $H(\lambda) \in \mathbb{F}[\lambda]^{j\times n}$ are* dual *if $G(\lambda)H(\lambda)^\top = 0$ and $j + k = n$.*

We are interested in a particular subclass of dual bases which are relevant for our construction. We will call them *dual linear bases.*

Definition 12 (full row-rank linear dual basis). *We say that a* $k \times (k+1)$ *matrix pencil* $L_{k,\phi}(\lambda)$ *is a* full row-rank linear dual basis *to* $\pi_{k,\phi}(\lambda)$ *(or, analogously, for a polynomial family* $\{\phi_i\}$*) if* $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$*, and* $L_{k,\phi}(\lambda)$ *has full row rank for any* $\lambda \in \overline{\mathbb{F}}$.

Often we just say that $L_{k,\phi}(\lambda)$ is a full row-rank linear dual basis, meaning that it is dual to $\pi_{k,\phi}(\lambda)$. Since the family $\{\phi_i\}$ is reported in the notation that we use for $L_{k,\phi}(\lambda)$, there is no risk of ambiguity. In the context of developing strong linearizations, we also give the following definition (which can again be found in [20]).

Definition 13 (minimal basis). *A polynomial basis* $G(\lambda) \in \mathbb{F}[\lambda]^{k\times n}$ *is said to be* minimal *if the sum of degrees of its rows is minimal among all the possible bases of the vector space that they span.*

We are particularly interested in *dual minimal bases*, that is, bases that are both minimal and dual bases. In [20] it is shown that this is equivalent to asking that $G(\lambda)$ and $H(\lambda)$ be of full row rank for any $\lambda \in \overline{\mathbb{F}}$ and the same holds for the matrices with rows equal to the highest degree coefficient of every row of $G(\lambda)$ and $H(\lambda)$. When the leading coefficients of $G(\lambda)$ and $H(\lambda)$ have only nonzero rows this corresponds to their leading coefficient.

Remark 14. In the rest of the paper we will often consider full row-rank linear dual bases (which will sometimes be minimal) related to polynomial families $\{\phi_i\}$. In order to make the exposition simpler we will call these bases *dual*, without adding the terms *linear* and *full row-rank*. However, it must be noted that these are a very particular kind of dual bases and most of the results could not hold in a more general context.

2.3. Building linearizations using product families. Let $P(\lambda)$ be a polynomial (or a matrix polynomial) expressed as a linear combination of elements of a product family $\phi \otimes \psi$. In this section we provide a way of linearizing it starting from the coefficients of this representation. In order to obtain this construction we rely on the following extension of [14, Theorem 5.2], which covers the case where both $\{\phi_i\}$ and $\{\psi_i\}$ are monomial bases.

Theorem 15. *Let* $L_{\epsilon,\phi}(\lambda) \in \mathbb{C}[\lambda]^{\epsilon\times(\epsilon+1)}$ *and* $L_{\eta,\psi}(\lambda) \in \mathbb{C}[\lambda]^{\eta\times(\eta+1)}$ *be dual linear bases for two polynomial families* $\{\phi_i\}$ *and* $\{\psi_i\}$*. Assume that the elements of each polynomial family have no common divisor, that is, there exists a vector* $w_{k,\star}$ *such that* $\pi_{k,\star}(\lambda)^\top w_{k,\star} = 1$ *for* $(k,\star) \in \{(\epsilon,\phi),(\eta,\psi)\}$*. Then the matrix polynomial*

$$\mathcal{L}(\lambda) := \begin{bmatrix} \lambda M_1 + M_0 & L_{\epsilon,\phi}(\lambda)^\top \otimes I_m \\ L_{\eta,\psi}(\lambda) \otimes I_n & 0_{\eta n\times\epsilon m} \end{bmatrix}, \qquad M_0, M_1 \in \mathbb{C}^{m(\epsilon+1)\times n(\eta+1)}$$

is a linearization for $P(\lambda) = (\pi_{\epsilon,\phi}(\lambda)\otimes I_m)^\top(\lambda M_1 + M_0)(\pi_{\eta,\psi}(\lambda)\otimes I_n)$*, which is an* $m \times n$ *matrix polynomial expressed in the product family* $\phi\otimes\psi$*. Moreover, this linearization is strong*[1] *if the reversals of* $L_{k,\star}(\lambda)$ *have full row rank.*

Proof. We mainly follow the proof given in [14]. Let $(k,\star)$ be either (ϵ,ϕ) or (η,ψ). Then recall that we can find $b_{k,\star}$ such that the matrix polynomial

$$S_{k,\star}(\lambda) := \begin{bmatrix} L_{k,\star}(\lambda) \\ b_{k,\star}^\top \end{bmatrix}$$

[1]Notice that the linearization is guaranteed to be strong for the matrix polynomial formally defined by $(\pi_{\epsilon,\phi}(\lambda)\otimes I_m)^\top(\lambda M_1 + M_0)(\pi_{\eta,\psi}(\lambda)\otimes I_n)$. In particular, this expression might provide a matrix polynomial with leading coefficient zero, but we still need to consider that polynomial and not the one with the leading zero coefficients removed; otherwise the strongness might be lost.

is unimodular [4], and we know that $S_{k,\star}(\lambda)\pi_{k,\star}(\lambda) = \alpha_{k,\star}(\lambda)e_{k+1}$. This can be rewritten as $\pi_{k,\star}(\lambda) = \alpha_{k,\star}(\lambda)S_{k,\star}^{-1}(\lambda)e_{k+1}$. Since the entries of $\pi_{k,\star}(\lambda)$ do not have any common factor, we conclude that $\alpha_{k,\star}(\lambda)$ is a nonzero constant (and so we can drop the dependency on λ). We remark that rescaling the vector $b_{k,\star}$ by a nonzero constant preserves the unimodularity of $S_{k,\star}(\lambda)$ (since it is equivalent to left multiplying by an invertible diagonal matrix). For this reason we can assume that $b_{k,\star}$ is chosen so that $S_{k,\star}(\lambda)\pi_{k,\star}(\lambda) = e_{k+1}$. We define $V_{k,\star}(\lambda) := S_{k,\star}(\lambda)^{-1}$ so that $V_{k,\star}(\lambda)e_{k+1} = \pi_{k,\star}(\lambda)$. With these hypotheses we have that

$$L_{k,\star}(\lambda)V_{k,\star}(\lambda) = \begin{bmatrix} I_k & 0 \end{bmatrix}, \qquad V_{k,\star}^\top(\lambda)L_{k,\star}^\top(\lambda) = \begin{bmatrix} I_k \\ 0 \end{bmatrix}.$$

Now observe that the matrix pencil $\mathcal{L}(\lambda)$ can be transformed by means of a unimodular transformation in the following way:

$$\begin{bmatrix} V_{\epsilon,\phi}^\top(\lambda) \otimes I_m & X(\lambda) \\ 0 & I_{\eta n} \end{bmatrix} \begin{bmatrix} \lambda M_1 + M_0 & L_{\epsilon,\phi}^\top(\lambda) \otimes I_m \\ L_{\eta,\psi}(\lambda) \otimes I_n & 0_{\eta n \times \epsilon m} \end{bmatrix} \begin{bmatrix} V_{\eta,\psi}(\lambda) \otimes I_n & 0 \\ Y(\lambda) & I_{\epsilon m} \end{bmatrix} =: \tilde{P}(\lambda),$$

where $\tilde{P}(\lambda)$ can be chosen as follows:

$$\tilde{P}(\lambda) := \begin{bmatrix} 0 & 0 & I_{\epsilon m} \\ 0 & P(\lambda) & 0 \\ I_{\eta n} & 0 & 0 \end{bmatrix}, \quad P(\lambda) = (\pi_{\epsilon,\phi}(\lambda) \otimes I_m)^\top (\lambda M_1 + M_0)(\pi_{\eta,\psi}(\lambda) \otimes I_n).$$

One can check, by direct substitution, that the following choices for the matrices $X(\lambda)$ and $Y(\lambda)$ provide the above structure:

$$X(\lambda) := -(V_{\epsilon,\phi}^\top(\lambda) \otimes I_m)(\lambda M_1 + M_0)(V_{\eta,\psi}(\lambda) \otimes I_n) \begin{bmatrix} I_{\eta n} \\ 0_{n \times \eta n} \end{bmatrix},$$

$$Y(\lambda) := -\begin{bmatrix} I_{\epsilon m} \\ 0_{m \times \epsilon m} \end{bmatrix}^\top (V_{\epsilon,\phi}(\lambda)^\top \otimes I_m)(\lambda M_1 + M_0)(V_{\eta,\psi}(\lambda) \otimes I_n) \left[(e_{\eta+1}e_{\eta+1}^\top) \otimes I_n\right],$$

which yields the zeros in the block entries $(1,1), (1,2)$, and $(2,1)$. The appearance of $P(\lambda)$ in the block entry $(2,2)$ follows using the relation $V_{k,\star}e_{k+1} = \pi_{k,\star}(\lambda)$.

We now check that the linearization is strong. Similarly to the previous step, we can find a constant vector $u_{k,\star}$ such that

$$\tilde{S}_{k,\star}(\lambda) = \begin{bmatrix} u_{k,\star} \\ \operatorname{rev} L_{k,\star}(\lambda) \end{bmatrix}, \qquad \star \in \{\phi, \psi\},$$

and $\tilde{S}_{k,\star}(\lambda) \operatorname{rev} \pi(\lambda) = \tilde{\alpha}_{k,\star}(\lambda)e_1$. Since the entries of $\pi_{k,\star}(\lambda)$ do not share any common factor, we get that $\tilde{\alpha}_{k,\star}$ is a nonzero constant. As in the previous case, applying a diagonal scaling does not change the unimodularity, so we can assume that $\tilde{\alpha}_{k,\star} = 1$. Define $W_{k,\star}(\lambda) = \tilde{S}_{k,\star}(\lambda)^{-1}$ so that $W_{k,\star}(\lambda)e_1 = \operatorname{rev} \pi_{k,\star}(\lambda)$.

We can perform another unimodular transformation on the reversed polynomial. Let $A(\lambda)$ be defined as follows:

$$\begin{bmatrix} W_{\epsilon,\phi}^\top(\lambda) \otimes I_m & \hat{X}(\lambda) \\ 0 & I_{\eta n} \end{bmatrix} \begin{bmatrix} M_1 + \lambda M_0 & \operatorname{rev} L_{\epsilon,\phi}^\top(\lambda) \otimes I_m \\ \operatorname{rev} L_{\eta,\psi}(\lambda) \otimes I_n & 0_{\eta n \times \epsilon m} \end{bmatrix} \begin{bmatrix} W_{\eta,\psi}(\lambda) \otimes I_n & 0 \\ \hat{Y}(\lambda) & I_{\epsilon m} \end{bmatrix}.$$

Notice that $\operatorname{rev} L_{k,\star}(\lambda)W_{k,\star}(\lambda) = [0 \ I_k]$, so we can write

$$A(\lambda) = \begin{bmatrix} A_{1,1}(\lambda) & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{bmatrix}$$

by appropriately choosing $\hat{X}(\lambda)$ and $\hat{Y}(\lambda)$ as before. In particular we have

$$A_{1,1}(\lambda) = (\operatorname{rev} \pi_{\epsilon,\phi}(\lambda) \otimes I_m)^\top (M_1 + \lambda M_0)(\operatorname{rev} \pi_{\eta,\psi}(\lambda) \otimes I_n) = \operatorname{rev} P(\lambda)$$

if the degree of $P(\lambda)$ is maximum (i.e., if the coefficient that goes in front of the maximum degree term in the previous relation is not zero). ∎

Remark 16. In the following, Theorem 15 will often be applied in the case where $\{\phi_i\}$ and $\{\psi_i\}$ are polynomial bases. It is worth noting that, in this case, the hypothesis on the existence of $w_{k,\star}$ in Theorem 15 is always satisfied, since this is just the vector containing the coefficients of the constant 1 in the prescribed basis.

We emphasize that asking the reversal of $L_{k,\star}(\lambda)$ to have full row rank has a connection with the minimality property. In fact, a sufficient condition for this to hold is that $L_{k,\star}(\lambda)$ is minimal and has all row degrees equal to 1 [20].

Theorem 15, as stated here, holds for $m \times n$ matrix polynomials, but in the following we will mainly deal with $m \times m$ square ones, and we will drop the symbol n for the second dimension.

2.4. An extension to more than two bases. Given the above formulation for a linearization of a polynomial expressed in a product family, it is natural to ask whether the framework can be extended to cover more than two bases, that is, to product families of the form

$$\phi^{(1)} \otimes \cdots \otimes \phi^{(j)} := \{\phi^{(1)}_{i_1} \dots \phi^{(j)}_{i_j} \mid i_s = 0, \dots, k_s,\ s = 1, \dots, j\},$$

where $\{\phi^{(s)}_i \mid i = 0, \dots, k_s\}$ are families of polynomials for $s = 1, \dots, j$.

We show that there is no need to extend Theorem 15, but it is sufficient to construct two appropriate dual bases $L_{\epsilon,\phi}(\lambda)$ and $L_{\eta,\psi}(\lambda)$ to deal with this case. We only need to prove that the hypotheses of Theorem 15 are satisfied.

Definition 17. *Let $L_{\epsilon,\phi}(\lambda) \in \mathbb{C}^{\epsilon \times (\epsilon+1)}[\lambda]$ and $L_{\eta,\psi}(\lambda) \in \mathbb{C}^{\eta \times (\eta+1)}[\lambda]$ be two dual bases for two families $\{\phi_i\}$ and $\{\psi_i\}$. Let w be a constant vector such that $w^\top \pi_{\eta,\psi}(\lambda)$ is a nonzero constant, and let A be an invertible $(\epsilon+1) \times (\epsilon+1)$ matrix. We say that the $k \times (k+1)$ matrix*

$$L_{k,\phi\otimes\psi}(\lambda) = \begin{bmatrix} A \otimes L_{\eta,\psi}(\lambda) \\ L_{\epsilon,\phi}(\lambda) \otimes w^\top \end{bmatrix}, \qquad k := (\epsilon+1)(\eta+1) - 1,$$

is a product dual basis *of $L_{\epsilon,\phi}(\lambda)$ and $L_{\eta,\psi}(\lambda)$. We denote it as $L_{\epsilon,\phi}(\lambda) \underline{\times} L_{\eta,\psi}(\lambda)$.*

The name "dual basis" used in the above definition is justified by Lemma 19, where we show that this matrix pencil is in fact a dual basis for a certain polynomial family.

Notice that, since the product dual basis is not unique, the previous notation actually denotes a family of such matrices, so we should be writing $L_{k,\phi\otimes\psi}(\lambda) \in L_{\epsilon,\phi}(\lambda) \underline{\times} L_{\eta,\psi}(\lambda)$. However, in the following we will often write, by slight abuse of notation, $L_{k,\phi\otimes\psi}(\lambda) = L_{\epsilon,\phi}(\lambda) \underline{\times} L_{\eta,\psi}(\lambda)$.

The above definition can be extended easily to a product of arbitrary families, by means of the following.

Definition 18. *We say that, for any families of polynomials $\{\phi^{(1)}_i\}, \dots, \{\phi^{(j)}_i\}$, the matrix $L_{k,\phi^{(1)}\otimes\cdots\otimes\phi^{(j)}}(\lambda)$ is a* product dual basis *for these families, and we denote it as $L_{\epsilon_1,\phi^{(1)}} \underline{\times} \cdots \underline{\times} L_{\epsilon_j,\phi^{(j)}}(\lambda)$, where*

$$L_{k,\phi^{(1)}\otimes\cdots\otimes\phi^{(j)}}(\lambda) = (L_{\epsilon_1,\phi^{(1)}}(\lambda) \underline{\times} \cdots \underline{\times} L_{\epsilon_{j-1},\phi^{(j-1)}}(\lambda)) \underline{\times} L_{\epsilon_j,\phi^{(j)}}(\lambda).$$

Notice that the above formula provides a recursive manner for computing such product dual bases. In the next lemma we show that they can be used to construct linearizations in the spirit of Theorem 15.

LEMMA 19. *Let $L_{k,\phi\otimes\psi}(\lambda) = L_{\epsilon,\phi}(\lambda)\underline{\times}L_{\eta,\psi}(\lambda)$ be a product dual basis. Then the following hold:*

(i) *If $\pi_{k,\phi\otimes\psi}(\lambda)$ is the vector containing the elements of the product family $\phi\otimes\psi$, then $L_{k,\phi\otimes\psi}(\lambda)\pi_{k,\phi\otimes\psi}(\lambda) = 0$.*

(ii) *$L_{k,\phi\otimes\psi}(\lambda)$ is a rectangular matrix with full row rank for all values of λ.*

In particular, $L_{k,\phi\otimes\psi}(\lambda)$ is a dual basis for the family associated with $\pi_{k,\phi\otimes\psi}(\lambda)$.

Proof. We first check condition (i). Notice that we have $\pi_{k,\phi\otimes\psi}(\lambda) = \pi_{\epsilon,\phi}(\lambda) \otimes \pi_{\eta,\psi}(\lambda)$, according to the ordering specified in Definition 8. For this reason we can write

$$L_{k,\phi\otimes\psi}(\lambda)\pi_{k,\phi\otimes\psi}(\lambda) = \begin{bmatrix} A\pi_{\epsilon,\phi}(\lambda) \otimes L_{\eta,\psi}(\lambda)\pi_{\eta,\psi}(\lambda) \\ L_{\epsilon,\phi}(\lambda)\pi_{\epsilon,\phi}(\lambda) \otimes w^\top \pi_{\eta,\psi}(\lambda) \end{bmatrix} = 0.$$

Regarding the full row rank claim in (ii) we shall check that, for any λ, the only vectors in the right kernel of $L_{k,\phi\otimes\psi}(\lambda)$ are multiples of $\pi_{k,\phi\otimes\psi}(\lambda)$. Let $v(\lambda)$ be such a vector, so that $L_{k,\phi\otimes\psi}(\lambda)v(\lambda) = 0$. We can partition $v(\lambda) = [v_0(\lambda) \ \dots \ v_\epsilon(\lambda)]^\top$ in blocks of size $\eta+1$, according to the block structure of $L_{k,\phi\otimes\psi}(\lambda)$ so that, recalling that A is invertible, we have

$$L_{k,\phi\otimes\psi}(\lambda)v(\lambda) = 0 \iff \begin{cases} L_{\eta,\psi}(\lambda)v_j(\lambda) = 0, \\ (L_{\epsilon,\phi}(\lambda) \otimes w^\top)v(\lambda) = 0, \end{cases} \qquad j = 0, \dots, \epsilon.$$

The first relation tells us that $v_j(\lambda) = \alpha_j(\lambda)\pi_{\eta,\psi}(\lambda)$, since $L_{\eta,\psi}(\lambda)$ has full row rank. If we set $\alpha(\lambda) = [\alpha_0(\lambda) \ \dots \ \alpha_\epsilon(\lambda)]^\top$, we have $v(\lambda) = \alpha(\lambda) \otimes \pi_{\eta,\psi}(\lambda)$, so that the last equation becomes $L_{\epsilon,\phi}(\lambda)\alpha(\lambda) \otimes w^\top \pi_{\eta,\psi}(\lambda) = 0$. Since $w^\top\pi_{\eta,\psi}(\lambda) \neq 0$, the only solution, up to scalar multiples, is given by $\alpha(\lambda) = \pi_{\epsilon,\phi}(\lambda)$. ∎

Remark 20. The proof of Lemma 19 shows that this construction is not the only possible one. As an immediate example, we could have defined $L_{\epsilon,\phi}(\lambda)\underline{\times}L_{\eta,\psi}(\lambda)$ to be the matrix

$$\tilde{L}_{k,\phi\otimes\psi}(\lambda) = \begin{bmatrix} w^\top \otimes L_{\eta,\psi}(\lambda) \\ L_{\epsilon,\phi}(\lambda) \otimes A \end{bmatrix}, \qquad k := (\epsilon+1)(\eta+1) - 1,$$

with the hypotheses of Definition 17, and the proof would have been essentially the same.

Remark 21. Lemma 19 justifies the notation $L_{k,\phi\otimes\psi}(\lambda)$ that we have used until now, since the product dual basis is a dual basis for the product family $\phi\otimes\psi$.

Remark 22. Given the structure of the matrix $L_{\epsilon,\phi}\underline{\times}L_{\eta,\psi}(\lambda)$ that we have defined above, it might be natural to ask whether the more general matrix

$$M(\lambda) = \begin{bmatrix} A \otimes L_{\eta,\psi}(\lambda) \\ L_{\epsilon,\phi}(\lambda) \otimes B \end{bmatrix}, \qquad A \in \mathbb{C}^{k_1\times(\eta+1)}, \ B \in \mathbb{C}^{k_2\times(\epsilon+1)},$$

and such that $k_1\eta + k_2\epsilon = (\eta+1)(\epsilon+1) - 1$, can be a product dual basis when A and B are of full row rank. The answer is negative unless $k_1 = \epsilon+1$ or $k_2 = \eta+1$, and so we are again back in the above two cases, as the next lemma shows.

LEMMA 23. *Let $M(\lambda)$ be a matrix of the form*

$$M(\lambda) = \begin{bmatrix} A \otimes L_{\eta,\psi}(\lambda) \\ L_{\epsilon,\phi}(\lambda) \otimes B \end{bmatrix}, \qquad A \in \mathbb{C}^{k_1 \times (\epsilon+1)},\ B \in \mathbb{C}^{k_2 \times (\eta+1)},$$

with $L_{\epsilon,\phi}(\lambda)$ and $L_{\eta,\psi}(\lambda)$ being dual bases for $\pi_{\epsilon,\phi}(\lambda)$ and $\pi_{\eta,\psi}(\lambda)$, A and B of full row rank, and $k_1\eta + k_2\epsilon = (\epsilon+1)(\eta+1) - 1$. Then, if there exists one λ such that either $A\pi_{\epsilon,\phi}(\lambda) \neq 0$ or $B\pi_{\eta,\psi}(\lambda) \neq 0$, the right kernel of $M(\lambda)$ has dimension at least $1 + (\epsilon + 1 - k_1)(\eta + 1 - k_2)$.

Proof. We start by proving that $k_1 \leqslant \epsilon + 1$ and $k_2 \leqslant \eta + 1$. Assume, by contradiction, that $k_1 > \epsilon + 1$. Then we would have

$$(\eta+1)(\epsilon+1) - 1 = k_1\eta + k_2\epsilon > (\epsilon+1)\eta + k_2\epsilon,$$

which implies $(k_2 - 1)\epsilon < 0$. Since $k_2 \geqslant 1$ and ϵ is positive, this cannot happen and $k_1 \leqslant \epsilon + 1$. The statement for k_2 can be proven in the same way.

Let $S_A = \{v \mid Av = 0\}$ and $S_B = \{w \mid Bw = 0\}$ be the right kernels of A and B which have dimensions $(\epsilon + 1 - k_1)$ and $(\eta + 1 - k_2)$, respectively. We have that the spans of $\pi_{\epsilon,\phi}(\lambda) \otimes \pi_{\eta,\psi}(\lambda)$ and $S_A \otimes S_B$ are included in the kernel of $M(\lambda)$. Since $A\pi_{\epsilon,\phi}(\lambda) \neq 0$ (or, analogously, $B\pi_{\eta,\psi}(\lambda) \neq 0$) for at least one λ, we have that the dimension of the union of these two spaces is at least $1 + (\epsilon + 1 - k_1)(\eta + 1 - k_2)$, which concludes the proof. ∎

Lemma 19 can be generalized to the product of more families of polynomials, yielding the following.

COROLLARY 24. *Let $L_{\epsilon_1,\phi^{(1)}}(\lambda)\underline{\times}\cdots\underline{\times}L_{\epsilon_j,\phi^{(j)}}(\lambda)$ be a product dual basis of j dual bases. Then it has full row rank and the only elements in its right kernel are multiples of $\pi_{k,\phi^{(1)}\otimes\cdots\otimes\phi^{(j)}}(\lambda)$, independently of the construction chosen (either that of Lemma* 19 *or Remark* 20*).*

Proof. Exploit the recursive definition of $L_{\epsilon_1,\phi^{(1)}}(\lambda)\underline{\times}\cdots\underline{\times}L_{\epsilon_j,\phi^{(j)}}(\lambda)$, and apply Lemma 19. ∎

The construction of these product dual bases allows us to formulate the following result, which can be seen as an extension of Theorem 15 that makes it possible to handle more than two bases at once.

THEOREM 25. *Let $\{\phi_i^{(1)}\}, \dots, \{\phi_i^{(j)}\}$ and $\{\psi_i^{(1)}\}, \dots, \{\psi_i^{(l)}\}$ be families of polynomials. Then the matrix pencil*

$$\mathcal{L}(\lambda) = \begin{bmatrix} \lambda M_1 + M_0 & (L_{\epsilon_1,\phi^{(1)}}\underline{\times}\cdots\underline{\times}L_{\epsilon_j,\phi^{(j)}}(\lambda))^\top \\ L_{\eta_1,\psi^{(1)}}\underline{\times}\cdots\underline{\times}L_{\eta_l,\psi^{(l)}}(\lambda) & 0 \end{bmatrix}$$

is a linearization for the polynomial

$$P(\lambda) = (\pi_{\epsilon_1,\phi^{(1)}}(\lambda) \otimes \cdots \otimes \pi_{\epsilon_j,\phi^{(j)}}(\lambda))^\top (\lambda M_1 + M_0)(\pi_{\eta_1,\psi^{(1)}}(\lambda) \otimes \cdots \otimes \pi_{\eta_l,\psi^{(l)}}(\lambda)).$$

Proof. Apply Theorem 15, whose hypotheses are satisfied because of Lemma 19 and Corollary 24. ∎

The above result can be used to linearize a polynomial in the form

$$p(\lambda) = \sum_{i_1,\dots,i_{l+j}} a_{i_1,\dots,i_{l+j}} \phi_{i_1}^{(1)}(\lambda) \dots \phi_{i_j}^{(j)}(\lambda) \psi_{i_{j+1}}^{(1)}(\lambda) \dots \psi_{i_{j+l}}^{(l)}(\lambda), \tag{1}$$

where $a_{i_1,\ldots,i_{l+j}}$ is an $(l+j)$-dimensional tensor with the first l dimensions equal to $\epsilon_1,\ldots,\epsilon_l$ and the remaining ones to $\eta_1,\ldots,\eta_j$. We recall that an (l,j)-flattening of such a tensor is the matrix F obtained by rearranging the elements in an $(\epsilon_1\cdots\epsilon_l)\times(\eta_1\cdots\eta_l)$ matrix maintaining the lexicographical order. This is exactly what the MATLAB function `reshape` does. With this choice one has that, for each choice of vectors $v_1,\ldots,v_l$ and $w_1,\ldots,w_j$ of dimensions $\epsilon_1,\ldots,\epsilon_l$ and $\eta_1,\ldots,\eta_j$, respectively,

$$\sum_{i_1,\ldots,i_{l+j}} a_{i_1,\ldots,i_{l+j}} v_{1,i_1}\ldots v_{l,i_l} w_{1,i_{l+1}}\ldots w_{j,i_{l+j}} = (v_1\otimes\cdots\otimes v_l)F(w_1\otimes\cdots\otimes w_j). \tag{2}$$

Corollary 26. *Let $p(\lambda)$ be a scalar polynomial as in* (1). *Let $M_1\equiv 0$, and let M_0 be the (l,j)-flattening of the tensor $a_{i_1,\ldots,i_{l+j}}$. Then the matrix pencil $\mathcal{L}(\lambda)$ is a linearization for $p(\lambda)$.*

Proof. Applying Theorem 25 and exploiting the relation (2) yields the thesis. ∎

Now we provide an example of the structure that the matrix $L_{\epsilon,\phi}\underline{\times}L_{\eta,\psi}(\lambda)$ can have in a simple case. Let $\{\phi_i\}$ be the Chebyshev basis, and let the family $\{\psi_i\}$ be any degree graded polynomial family. The matrix $L_{\epsilon,\phi}\underline{\times}L_{\eta,\psi}(\lambda)$ can be realized as follows by choosing $A=I$ and $w=e_{\eta+1}$:[2]

$$L_{\epsilon,\phi}\underline{\times}L_{\eta,\psi}(\lambda)=\left[\begin{array}{c|c|c|c|c}
L_{\eta,\psi}(\lambda) & & & & \\
 & L_{\eta,\psi}(\lambda) & & & \\
 & & \ddots & & \\
 & & & L_{\eta,\psi}(\lambda) & \\
 & & & & L_{\eta,\psi}(\lambda) \\ \hline
1\cdot e_{\eta+1}^\top & -2\lambda\cdot e_{\eta+1}^\top & 1\cdot e_{\eta+1}^\top & & \\
 & \ddots & \ddots & \ddots & \\
 & & 1\cdot e_{\eta+1}^\top & -2\lambda\cdot e_{\eta+1}^\top & 1\cdot e_{\eta+1}^\top \\
 & & & 1\cdot e_{\eta+1}^\top & -\lambda\cdot e_{\eta+1}^\top
\end{array}\right].$$

In order to give an example of how these dual bases behave in practice, we consider what happens when taking the product basis of several monomial bases.

The monomial basis, in this setting, is rather special. In fact, the elements of the product family of two monomial bases are of the form $\lambda^i\lambda^j=\lambda^{i+j}$, and so they correspond to elements of a (larger) monomial basis. However, notice that this is not true in general, as, for example, when considering $\phi_i(\lambda)$ belonging to other polynomial bases.

We can exploit this fact by rephrasing any polynomial expressed in the monomial basis as a polynomial in the product family of two monomial bases (like in [14]) or also in the product family of more bases, by using the framework above.

Let $p(\lambda)=\sum_{i=0}^3 p_i\lambda^i$ be a degree 3 polynomial; we consider three different linearizations for it, obtained by rephrasing it in different bases ϕ and ψ in the context of Theorems 15 and 25.

As a first example, choosing $\{\psi_i\}=\{1,\lambda,\lambda^2\}$ and $\{\phi_i\}=\{1\}$ yields the classical Frobenius form:

$$\mathcal{L}(\lambda)=\left[\begin{array}{ccc}
\lambda p_3+p_2 & p_1 & p_0 \\ \hline
1 & -\lambda & \\
 & 1 & -\lambda
\end{array}\right].$$

[2] The definition of $L_{\epsilon,\phi}(\lambda)$ for $\{\phi_i\}$ being the Chebyshev basis will be given in Lemma 27.

We can instead choose $\{\psi_i\} = \{\phi_i\} = \{1, \lambda\}$ and obtain a symmetric linearization (this is only one of the possibilities for distributing the coefficients):

$$\mathcal{L}(\lambda) = \left[\begin{array}{cc|c} \lambda p_3 + p_2 & \frac{1}{2}p_1 & 1 \\ \frac{1}{2}p_1 & p_0 & -\lambda \\ \hline 1 & -\lambda & 0 \end{array}\right].$$

But we can also choose to set $\{\psi_i\} = \{1, \lambda\} \otimes \{1, \lambda\}$ and $\{\phi_i\} = \{1\}$, and we obtain

$$\mathcal{L}(\lambda) = \left[\begin{array}{cccc} \lambda p_3 + p_2 & \frac{1}{2}p_1 & \frac{1}{2}p_1 & p_0 \\ \hline 1 & -\lambda & & \\ & & 1 & -\lambda \\ \hline & 1 & & -\lambda \end{array}\right].$$

One thing can be noticed immediately: we have increased the dimension of the problem. In fact the dual basis to $\{1, \lambda\} \otimes \{1, \lambda\}$ that we have used in the lower part has its dimension increased by 1, since λ is represented two times. This has the consequence that while it has full row rank, its reversal does not, and so the linearization is not strong. In fact, here we have a spurious infinite eigenvalue.

3. Construction of dual bases.

In this section we show how to construct dual bases $L_{k,\phi}(\lambda)$ for many concrete families of polynomials $\{\phi_i\}$. These will be the main ingredient needed in the application of Theorems 15 and 25.

3.1. Handling orthogonal bases.

This section is devoted to studying the different structure of the dual basis $L_{k,\phi}(\lambda)$ when $\{\phi_i\}$ is an orthogonal basis. More precisely, we consider the case where the basis $\{\phi_i\}$ is degree graded and satisfies a three-term recurrence relation of the form

$$\alpha\phi_{j+1}(\lambda) = (\lambda - \beta)\phi_j(\lambda) - \gamma\phi_{j-1}(\lambda), \qquad \alpha \neq 0, \quad j > 0, \tag{3}$$

which includes all the orthogonal polynomials with a constant three-term recurrence (with the possible exception of the first two elements of the basis). Notice, however, that the result can be easily generalized to more general recurrences.

LEMMA 27. *Let $\{\phi_i\}$ be a degree graded basis satisfying the three-term recurrence relation* (3). *Then the matrix pencil $L_{k,\phi}(\lambda)$ of size $k \times (k+1)$ is defined as follows:*

$$L_{k,\phi}(\lambda) := \begin{bmatrix} \alpha & (\beta - \lambda) & \gamma & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha & (\beta - \lambda) & \gamma \\ & & & \phi_0(\lambda) & -\phi_1(\lambda) \end{bmatrix}$$

has full row rank for any $\lambda \in \mathbb{F}$ and is such that

$$L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0, \qquad \text{with } \pi_{k,\phi}(\lambda) := \begin{bmatrix} \phi_k(\lambda) \\ \vdots \\ \phi_0(\lambda) \end{bmatrix}.$$

Moreover, the leading coefficient of $L_{k,\phi}(\lambda)$ has full row rank.

Proof. It is immediate to verify that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$, since each row of $L_{k,\phi}(\lambda)$ but the last one is just the recurrence relation of (3) and the last one yields $\phi_0(\lambda)\phi_1(\lambda) - \phi_1(\lambda)\phi_0(\lambda) = 0$.

We can then check that the matrix has full row rank. Notice that the first k columns of $L_{k,\phi}(\lambda)$ form an upper triangular matrix with determinant $\alpha^{k-1}\phi_0(\lambda)$. The basis is degree graded, so $\phi_0(\lambda)$ is an invertible constant and $L_{k,\phi}(\lambda)$ contains an invertible matrix of order $k \times k$, thereby proving our claim.

It is immediate to verify the last claim, since the leading coefficient of $L_{k,\phi}(\lambda)$ with the first column removed is a diagonal matrix with nonzero elements on the diagonal, and so it is invertible. ∎

We can immediately construct some examples for the application of the lemma. Consider the Chebyshev basis of the first kind $\{T_i(\lambda)\}$, which satisfies a recurrence relation of the form

$$T_{j+1}(\lambda) = 2\lambda T_j(\lambda) - T_{j-1}(\lambda), \qquad T_0(\lambda) := 1, \qquad T_1(\lambda) := \lambda.$$

Then we have that the matrix pencil

$$\mathcal{L}(\lambda) = \begin{bmatrix} \lambda M_1 + M_0 & L_{\epsilon,T}(\lambda)^\top \\ L_{\eta,T}(\lambda) & 0 \end{bmatrix}, \qquad L_{k,T}(\lambda) := \begin{bmatrix} 1 & -2\lambda & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2\lambda & 1 \\ & & & 1 & -\lambda \end{bmatrix}$$

is a linearization for the polynomial $p(\lambda) = \sum_{i=0}^{\epsilon} \sum_{j=0}^{\eta} (\lambda M_1 + M_0)_{i,j} T_i(\lambda) T_j(\lambda)$. As shown in [27], the product $T_i(\lambda)T_j(\lambda)$ can be rephrased in terms of sums of Chebyshev polynomials, and this can be used to build a linearization for polynomials expressed in the Chebyshev basis (without product families involved).

3.2. Handling interpolation bases. The framework covers orthogonal bases, but there are some other interesting cases, for example, interpolation bases such as Lagrange, Newton, and Hermite.

In this section we study their structures. Recall that, by Theorem 15, to construct the dual basis $L_{k,\phi}(\lambda)$ for one of these bases we need to ensure that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$ and that $L_{k,\phi}(\lambda)$ has full row rank. In order to have a strong linearization we also require the reversal of the dual basis to have full row rank.

3.3. The Lagrange basis. Let $\sigma_1^{(1)}, \ldots, \sigma_\epsilon^{(1)}$ and $\sigma_1^{(2)}, \ldots, \sigma_\eta^{(2)}$ be two (not necessarily disjoint) sets of pairwise different nodes in the complex plane. Then we can define the weights and the Lagrange polynomials by

$$t_i^{(s)} := \prod_{j \neq i} (\sigma_i^{(s)} - \sigma_j^{(s)}), \qquad l_i^{(s)}(\lambda) := \frac{1}{t_i^{(s)}} \prod_{j \neq i} (\lambda - \sigma_j^{(s)}), \qquad s \in \{1, 2\}.$$

In the following let $\phi_j(\lambda) = l_j^{(1)}(\lambda)$ and $\psi_j(\lambda) = l_j^{(2)}(\lambda)$, coherently with the notation used before. The linearization for a polynomial expressed in a product family, built according to Theorem 15, has the following structure:

$$\mathcal{L}(\lambda) = \begin{bmatrix} \lambda M_1 + M_0 & L_{\epsilon,\phi}(\lambda)^\top \\ L_{\eta,\psi}(\lambda) & 0 \end{bmatrix},$$

where

$$L_{k,\phi}(\lambda) = \begin{bmatrix} t_1^{(1)}(\lambda - \sigma_1) & -t_2^{(1)}(\lambda - \sigma_2) & & \\ & \ddots & \ddots & \\ & & t_{k-1}^{(1)}(\lambda - \sigma_{k-1}) & -t_k^{(1)}(\lambda - \sigma_k) \end{bmatrix}$$

and $L_{k,\psi}(\lambda)$ can be defined in an analogous way.

LEMMA 28. *The matrix $L_{k,\phi}(\lambda)$ defined above is a dual minimal basis for the Lagrange basis $\{\phi_i\}$ constructed on the nodes $\sigma_1,\dots,\sigma_k$ (that is, it is dual to $\pi_{k,\phi}(\lambda)$).*

Proof. It is easy to verify that $L_{k,\phi}(\lambda) \in \mathbb{C}[\lambda]^{k\times(k+1)}$ and

$$L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0, \qquad \pi_{k,\phi}(\lambda) := \begin{bmatrix} l_k^{(1)}(\lambda) \\ \vdots \\ l_0^{(1)}(\lambda) \end{bmatrix}.$$

We now need to show that the matrix $L_{k,\phi}(\lambda)$ has full row rank for any $\lambda \in \mathbb{F}$. For all values of λ that are not equal to the nodes the first k columns are upper triangular with nonzero elements on the diagonal, and so the hypotheses are satisfied. It remains to deal with the cases where $\lambda = \sigma_i$ for some $i = 1,\dots,k-1$.

We note that in this case one of the columns of the matrix is zero, but removing it yields a square matrix which is block diagonal with only two diagonal blocks. The top-left one is upper triangular and invertible, while the bottom-right one is lower triangular and invertible, since they both have nonzero elements on the diagonal.

Notice that the first k columns of the leading coefficient of $L_{k,\phi}(\lambda)$ are upper triangular with nonzero elements on the diagonal. This implies that the leading coefficient has full row rank, thus proving the minimality of $L_{k,\phi}(\lambda)$. ∎

3.4. Constructing a classical Lagrange linearization. Besides building linearizations for a polynomial expressed in product families of Lagrange bases, the above formulation can be used to linearize a polynomial expressed in a Lagrange basis built on the union of the nodes.

In fact, we observe that if we have two Lagrange polynomials $l_i^{(1)}(\lambda)$ and $l_j^{(2)}(\lambda)$ defined according to the previous notation, then their product is almost a Lagrange polynomial for the union of the nodes. More precisely, assume that we have a set of nodes $\sigma_1,\dots,\sigma_n$, and let $l_i^{(1)}(\lambda)$ and $l_j^{(2)}(\lambda)$ be Lagrange polynomials relative to the nodes $\sigma_1,\dots,\sigma_k$ and $\sigma_{k+1},\dots,\sigma_n$, respectively. Then if $l_i(\lambda)$ are the Lagrange polynomials associated with all the nodes, we have that

$$l_i(\lambda) = \begin{cases} l_i^{(1)}(\lambda)\cdot l_j^{(2)}(\lambda)\cdot \frac{\lambda-\sigma_{j+k}}{\sigma_i-\sigma_{j+k}}\prod_{s\neq j}\frac{\sigma_{j+k}-\sigma_s}{\sigma_i-\sigma_{s+k}}, & i \leqslant k, \\ l_j^{(1)}(\lambda)\cdot l_{i-k_1}^{(2)}(\lambda)\cdot \frac{\lambda-\sigma_j}{\sigma_i-\sigma_j}\prod_{s\neq j}\frac{\sigma_j-\sigma_s}{\sigma_i-\sigma_s}, & i > k. \end{cases}$$

It is worth noting that these formulas become much more straightforward if one considers unscaled Lagrange polynomials by getting rid of the normalization factor, since in that case we obtain

$$l_i(\lambda) = \begin{cases} l_i^{(1)}(\lambda)\cdot l_j^{(2)}(\lambda)\cdot(\lambda-\sigma_{j+k}), & i \leqslant k, \\ l_j^{(1)}(\lambda)\cdot l_{i-k_1}^{(2)}(\lambda)\cdot(\lambda-\sigma_j), & i > k. \end{cases}$$

The part missing from the product of two Lagrange polynomials in order to obtain the one with the union of the nodes is always linear and so can be placed as a coefficient in the top-left matrix pencil $\lambda M_1 + M_0$.

Remark 29. We can choose two equal nodes in $\sigma_1,\dots,\sigma_k$ and $\sigma_{k+1},\dots,\sigma_n$. This allows us to obtain a Lagrange linearization with repeated nodes, which is a special case of Hermite linearization, where it is possible to interpolate a polynomial imposing the value of its first derivative at the nodes. By using the product dual bases it is possible to extend this construction to higher order derivatives. However, such a construction would have redundancy in the polynomial family, thus leading to linearizations which have infinite eigenvalues. In section 3.6 we present a direct construction of the dual basis for the Hermite basis that does not.

3.5. Explicit construction for the Newton basis. Another concrete example is the construction of the Newton basis linearization. We can consider, similarly to the Lagrange case, a set of nodes $\sigma_1, \dots, \sigma_n$ and assume to have two Newton bases, one built using $\sigma_1, \dots, \sigma_k$ and the other built using $\sigma_{k+1}, \dots, \sigma_n$.

To construct the linearization we need to find $L_{k,\phi}(\lambda)$ which satisfies the requirements of Theorem 15. A possible choice is given by the following:

$$L_{k,\phi}(\lambda) := \begin{bmatrix} 1 & \sigma_k - \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \sigma_1 - \lambda \end{bmatrix}, \qquad \pi_{k,\phi}(\lambda) = \begin{bmatrix} \prod_{j=1}^{k}(\lambda - \sigma_j) \\ \vdots \\ \lambda - \sigma_1 \\ 1 \end{bmatrix}.$$

The matrix $L_{k,\phi}(\lambda)$ has the correct dimensions $k \times (k+1)$, has full row rank for any λ, and is such that the product $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$. Moreover, the leading coefficient has full row rank, so we also have the minimality and all the hypotheses of Theorem 15 are satisfied.

3.6. Linearizations in the Hermite basis. Recently a linearization for polynomials expressed in the Hermite basis has been presented by Fassbender, Pérez, and Shayanfar in [18], and previous work on this topic by Lawrence and Corless can be found in [28].

The Hermite basis can be seen as a generalization of the Lagrange basis where not only the values of the functions at the nodes are considered, but also the values of their derivatives.

Assume that we have a set of nodes $\sigma_1, \dots, \sigma_n$, and that we have interpolated a function assigning the derivative up to the s-order, for some $s \geqslant 1$ (the case $s = 1$ gives the Lagrange basis). The order s can also vary depending on the node. We can then consider the basis given by the following vector polynomial:

$$\pi_{k,\phi}(\lambda) = \begin{bmatrix} \frac{\omega(\lambda)}{(\lambda-\sigma_1)^{s_1}} \\ \vdots \\ \frac{\omega(\lambda)}{(\lambda-\sigma_1)} \\ \vdots \\ \frac{\omega(\lambda)}{(\lambda-\sigma_n)^{s_n}} \\ \vdots \\ \frac{\omega(\lambda)}{(\lambda-\sigma_n)} \end{bmatrix}, \qquad \omega(\lambda) := \prod_{j=1}^{n}(\lambda - \sigma_j)^{s_j}, \qquad k = \sum_{j=1}^{n} s_j.$$

A generic polynomial expressed in this basis can be written as $p(\lambda) = p^\top \pi_{k,\phi}(\lambda)$, where p is the column vector with the coefficients in the Hermite basis.

We want to show that it is possible to formulate a linearization for the Hermite basis in our framework. We already have the vector $\pi_{k,\phi}(\lambda)$, so we just need to find a matrix pencil $L_{k,\phi}(\lambda)$ of the correct dimension that has full row rank and such that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$.

LEMMA 30. *The matrix pencil* $L_{k,\phi}(\lambda)$ *defined as*

$$L_{k,\phi}(\lambda) = \begin{bmatrix} J_{\sigma_1}(\lambda) & -(\lambda-\sigma_2)e_{s_1}e_{s_2}^\top & & & \\ & \ddots & \ddots & & \\ & & J_{\sigma_{n-1}}(\lambda) & -(\lambda-\sigma_n)e_{s_{n-1}}e_{s_n}^\top \\ & & & \tilde{J}_{\sigma_n}(\lambda) \end{bmatrix},$$

with

$$J_{\sigma_j}(\lambda) := \begin{bmatrix} \lambda-\sigma_j & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda-\sigma_j \end{bmatrix}, \quad \tilde{J}_{\sigma_j}(\lambda) := \begin{bmatrix} \lambda-\sigma_j & -1 & & \\ & \ddots & \ddots & \\ & & \lambda-\sigma_j & -1 \end{bmatrix},$$

is a dual basis for the Hermite basis $\{\phi_i\}$ *of orders* s_i, $i = 0, \ldots, n$.

Proof. We can check directly that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$, and so it only remains to verify that the row rank is maximum. We notice that for any $\lambda \neq \sigma_j$ the matrix is upper triangular with nonzero elements on the diagonal, and so the condition is obviously satisfied. For $\lambda = \sigma_j$ the diagonal block $J_{\sigma_j}(\lambda)$ is singular. Assume, for simplicity, that $j = 1$, and consider the matrix S obtained by removing the first column of $L_{k,\phi}(\lambda)$. We notice that S has the following structure:

$$S := \begin{bmatrix} -I & & & & \\ 0_{s_1-1}^\top & -(\sigma_1-\sigma_2)e_{s_2}^\top & & & \\ & J_{\sigma_2}(\sigma_1) & -(\sigma_1-\sigma_3)e_{s_2}e_{s_3}^\top & & \\ & & \ddots & \ddots & \\ & & & J_{\sigma_{n-1}}(\sigma_1) & -(\sigma_1-\sigma_n)e_{s_{n-1}}e_{s_n}^\top \\ & & & & \tilde{J}_{\sigma_n}(\sigma_1) \end{bmatrix}.$$

To prove that S is invertible we consider the trailing submatrix $\tilde{S}$ obtained by removing the first block row and column. We can transform $\tilde{S}$ by means of block column operations so that

$$\tilde{S}X = \begin{bmatrix} \tilde{u}^\top & \sigma_1-\sigma_n \\ B(\sigma_1) & -e_{k-\sigma_1-1} \end{bmatrix}, \qquad u(\lambda) = \begin{bmatrix} -(\sigma_1-\sigma_2)e_{s_2} \\ \vdots \\ -(\sigma_1-\sigma_{n-1})e_{s_n} \\ 0_{s_n-1} \end{bmatrix},$$

and $B(\sigma_1)$ is block diagonal with the $J_{\sigma_j}(\sigma_1)$ of size s_j on the block diagonal, except the last one, which is of size $s_n - 1$. Since $u^\top B(\sigma_1)^{-1}e_{k-\sigma_1-1} = 0$, we can write

$$\begin{aligned} \det \tilde{S}X &= \det B(\sigma_1) \cdot \left[(\sigma_1-\sigma_n) + \tilde{u}^\top B(\sigma_1)^{-1}e_{k-\sigma_1-1}\right] \\ &= \prod_{j=1}^{n-1}(\sigma_1-\sigma_j)^{s_j} \cdot (\sigma_1-\sigma_n)^{s_n-1}(\sigma_1-\sigma_n) = \prod_{j=1}^{n}(\sigma_1-\sigma_j)^{s_j} \neq 0. \end{aligned}$$

This proves that $\tilde{S}X$ is invertible, concluding the proof. ∎

The above lemma guarantees the applicability of Theorem 15 for the case of Hermite polynomials (and matrix polynomials), so we have an explicit way of building linearizations in this basis.

3.7. Bernstein basis. A last example that is relevant in the context of computer aided design is the Bernstein basis, which is the building block of Bézier curves [7, 15, 16, 17]. Given an interval $[\alpha, \beta]$, we can define the family of Bernstein polynomials of degree k as follows:

$$\phi_i(\lambda) := \binom{k}{i}(\lambda - \alpha)^i(\beta - \lambda)^{k-i}, \qquad i = 0, \ldots, n.$$

We show that also these polynomials fit in our construction.

LEMMA 31. *The matrix pencil $L_{k,\phi}(\lambda)$ defined as*

$$L_{k,\phi}(\lambda) := \begin{bmatrix} \binom{k}{k-1}(\lambda - \beta) & \binom{k}{k}(\lambda - \alpha) & & \\ & \ddots & \ddots & \\ & & \binom{k}{0}(\lambda - \beta) & \binom{k}{1}(\lambda - \alpha) \end{bmatrix}$$

is a dual minimal basis for the Bernstein polynomials of degree k defined above.

Proof. A direct computation shows that $L_{k,\phi}(\lambda)\pi_{k,\phi}(\lambda) = 0$. Moreover, notice that for any $\lambda \neq \beta$ the first k columns of $L_{k,\phi}(\lambda)$ form a square upper triangular matrix with nonzero diagonal elements, and for any $\lambda \neq \alpha$ the last k columns are an invertible lower triangular matrix. This guarantees that the row rank is maximum for any $\lambda \in \mathbb{F}$. Since the leading coefficient has the first k columns which are upper triangular and invertible, we also have the minimality. ∎

Remark 32. One might be interested, in order to control the size of the coefficients in the interpolation process, to scale the polynomial basis. We notice that this process does not change all the previous results, since it is equivalent to left multiplying $\pi_{k,\phi}(\lambda)$ by an invertible diagonal matrix D. It is then immediate to verify that if $L_{k,\phi}(\lambda)$ is dual to $\pi_{k,\phi}(\lambda)$, then $L_{k,\phi}(\lambda)D^{-1}$ is dual to $D\pi_{k,\phi}(\lambda)$.

4. Linearizing sums of polynomials and rational functions.

4.1. Linearizing the sum and difference of two polynomials. In this section we present another example of linearization which deals with the following problem: assume that we are given two polynomials $p(\lambda)$ and $q(\lambda)$ of which we want to find the intersections, that is, the values of λ such that $q(\lambda) = p(\lambda)$, and assume that $p(\lambda)$ and $q(\lambda)$ are expressed in different bases.

Normally one would solve the problem by considering the polynomial $r(\lambda) = p(\lambda) - q(\lambda)$ and finding its roots, for example, by using a linearization. However, this requires us to perform a change of basis on at least one of the two polynomials, and this operation is possibly ill-conditioned (see [22] for a related analysis).

In the case of interpolation bases, such as Newton or Lagrange, this could also be useful when one wants to find intersections of functions that have been sampled in different data points. In this case it might not even be possible to resample the function (think of measured data). Another application arises from computer aided design, where computing the intersection of polynomials (and rational functions) is an important task [33].

Here we show how to linearize the problem directly.

THEOREM 33. *Let $p(\lambda)$ and $q(\lambda)$ be two polynomials of the form*

$$p(\lambda) := \sum_{j=0}^{\epsilon} p_j \phi_j(\lambda), \qquad q(\lambda) := \sum_{j=0}^{\eta} q_j \psi_j(\lambda),$$

and let $L_{\epsilon,\phi}(\lambda)$ and $L_{\eta,\psi}(\lambda)$ be dual bases for $\{\phi_i\}$ and $\{\psi_i\}$. Let p and q be the vectors containing the coefficients of $p(\lambda)$ and $q(\lambda)$, respectively, so that $p(\lambda) = p^\top \pi_{\epsilon,\phi}(\lambda)$ and $q(\lambda) = q^\top \pi_{\eta,\psi}(\lambda)$. Then the matrix pencil

$$\mathcal{L}(\lambda) := \begin{bmatrix} pw_\psi^\top - w_\phi q^\top & L_{\epsilon,\phi}^\top(\lambda) \\ L_{\eta,\psi}(\lambda) & 0 \end{bmatrix}, \qquad w_\star := \Gamma_\star^{-1}(1), \qquad \star \in \{\phi, \psi\},$$

where $\Gamma_\star^{-1}(1)$ is the vector of the coefficients of the constant 1 in the basis $\star$ (see Remark 9), is a linearization for $r(\lambda) := p(\lambda) - q(\lambda)$.

Proof. $w_\star := \Gamma_\star^{-1}(1)$ means that $w_\star \pi_{k,\star}(\lambda) = 1$ for $\star \in \{\phi, \psi\}$, where k is either ϵ or η depending on the choice of $\star$. By Theorem 15 we know that $\mathcal{L}(\lambda)$ is a linearization for

$$\pi_{\epsilon,\phi}^\top(\lambda)(pw_\psi^\top - w_\phi q^\top)\pi_{\eta,\psi}(\lambda) = p(\lambda) \cdot 1 - 1 \cdot q(\lambda) = r(\lambda).$$

This concludes the proof. ∎

Remark 34. We notice that the linearization above, according to Theorem 15, is a strong linearization for a polynomial of degree $d := \epsilon + \eta$, but $r(\lambda)$ is of degree $\max\{\epsilon, \eta\} \leqslant d$. The reason for this is that this is actually a linearization for a polynomial of *grade* d that could have some leading zero coefficients, thus having degree smaller than d. The grade is defined as the maximum degree of the monomials, while the degree is the maximum of the *nonzero* ones.

The difference between grade and degree causes infinite eigenvalues to appear when we solve the eigenvalue problem obtained through Theorem 33 numerically. However, the finite eigenvalues that we get are still the actual roots of $r(\lambda)$.

The framework of section 2.4 could be used to extend the above result to the sum of an arbitrary number of polynomials (possibly all expressed in different bases). This can be obtained by combining the proof of Theorem 33 with the result of Theorem 25.

We now present some numerical experiments in order to further justify the use of the linearization presented in this section.

Numerical experiment 1. In this example we test the framework on the following example. Let $p_1(\lambda) = \sum_{i=0}^n p_{i,1}\lambda^i$ and $p_2(\lambda) = \sum_{i=0}^n p_{i,2}T_i(\lambda)$ be two polynomials expressed in the monomial and Chebyshev basis of the first kind, respectively. We want to find the roots of their sum $q(\lambda) = p_1(\lambda) + p_2(\lambda)$. The columns of Table 1 represent, in the following order, the result of these different approaches to solve the problem that we tested:

1. Converting $p_2(\lambda)$ to the monomial basis and using the Frobenius linearization to find the roots of the sum (by means of the command `roots` in MATLAB).
2. Converting $p_1(\lambda)$ to the Chebyshev basis and using the colleague linearization [24, 3] to find the roots of the sum of $p_1(\lambda)$ and $p_2(\lambda)$. The colleague pencil's eigenvalues have been approximated using the QZ method in MATLAB.
3. Constructing the linearization of Theorem 33 and computing its eigenvalues with the QZ method (using `eig` in MATLAB).
4. Constructing the linearization of Theorem 33 and deflating the spurious infinite eigenvalues by means of the strategy that will be proposed in section 6.[3]

[3]The approach of section 6 has been moved to the end of this work because it is fairly general and does not add much information about the structure of this linearization. Moreover, it can be applied to other examples that will follow.

Table 1
Numerical errors in the computation of the (finite) roots of the polynomial $p_1(\lambda) + p_2(\lambda)$ of numerical experiment 1.

Degree	Monomial	Chebyshev	Theorem 33	Theorem 33 + deflation
5	7.03e-16	5.19e-16	5.44e-16	8.40e-16
10	1.23e-14	2.07e-15	2.00e-15	2.33e-15
20	1.95e-11	4.48e-15	2.49e-15	4.08e-15
40	1.25e-04	1.15e-14	5.59e-15	6.45e-15
80	1.29e+00	7.62e-09	9.76e-15	1.69e-14
160	4.37e+00	1.05e-01	6.90e-14	3.63e-14
320	9.85e+00	2.97e+00	1.07e-13	7.57e-14
640	1.91e+01	1.52e+01	4.40e-13	1.27e-13

The polynomials have also been, by means of symbolical computations, converted to the monomial basis, and the roots have been computed using `MPSolve` [8] to guarantee 16 accurate digits. These results have been used as a reference to measure the errors, which have been summarized in Table 1 and Figure 1. In all the cases the infinite eigenvalues either have been deflated a priori, or have been detected by the QZ algorithm, and so we could deflate them a posteriori, so the numbers that we report refer to the errors on the finite eigenvalues. In particular, we reported the 2-norm of the vectors containing the absolute errors in the computed roots. The coefficients of the polynomials have been generated by using the `randn` function. Each experiment has been repeated 50 times, and only the average error is reported.

The bad results obtained in the cases where a basis conversion has been performed can be explained by looking at the conditioning of the matrix representing the change of basis between the monomial and the Chebyshev bases.

The conditioning is exponentially growing (see [22] for a related discussion), and as n grows above 50 it cannot be guaranteed to compute even a single correct digit in double precision (see Figure 2, where the exponential growth is clearly visible), and so the results start to deteriorate very quickly.

4.2. Finding intersections of the sum of two rational functions. The results of section 4.1 admit an interesting extension to finding the zeros of a sum of ratios of polynomials. This has the pleasant side effect of mitigating the numerical issues that might be encountered when dealing with a large number of infinite eigenvalues. Let $f(\lambda)$ be a rational function of the form

$$f(\lambda) := \frac{p(\lambda)}{q(\lambda)} + \frac{r(\lambda)}{s(\lambda)},$$

with $p(\lambda)$, $q(\lambda)$, $r(\lambda)$, and $s(\lambda)$ being polynomials, of which we want to find the zeros. We assume, in the following, that the numerators do not share any common factor with the denominators and that the two ratios do not have common poles. With these assumptions, we have that the roots of $f(\lambda)$ are the ones of $f(\lambda)q(\lambda)s(\lambda)$, that is, of the polynomial

$$t(\lambda) := p(\lambda)s(\lambda) + r(\lambda)q(\lambda).$$

In this section we will linearize the polynomial $t(\lambda)$. However, for simplicity we will sometimes inappropriately say that a linearization for $t(\lambda)$ is also a linearization for $f(\lambda)$, since they share the same zeros.

For simplicity we first consider the case in which all the polynomials are given in the monomial basis, and we will handle the case where two different bases are used to define the polynomials $p(\lambda)$, $q(\lambda)$, $r(\lambda)$, and $s(\lambda)$ later.

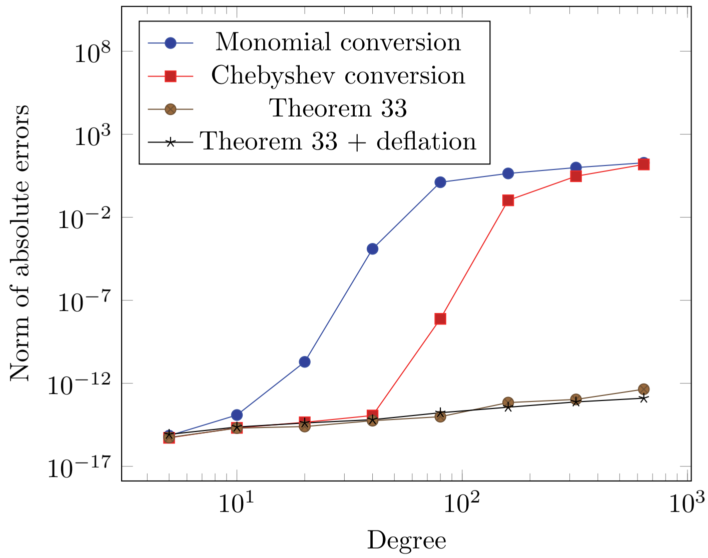

FIG. 1. *Norm of the absolute errors in the computation of the roots of $p_1(\lambda) + p_2(\lambda)$, where $p_1(\lambda)$ is a polynomial expressed in the monomial basis while $p_2(\lambda)$ is one expresesed in the Chebyshev one.*

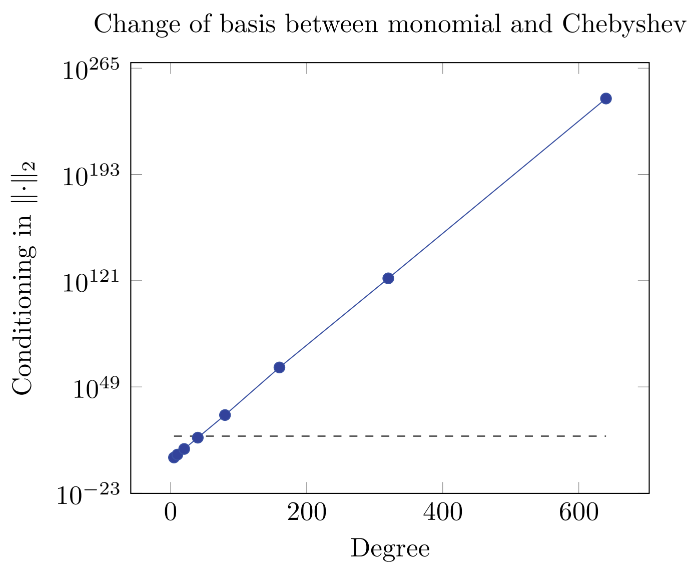

FIG. 2. *Conditioning of the change of basis matrix between the monomial and Chebyshev basis. The dashed line represents the level $\frac{1}{u}$, where u is the machine precision. Beyond that point no correct digits can be guaranteed on the computed coefficients.*

THEOREM 35. *Let $f(\lambda) = \frac{p(\lambda)}{q(\lambda)} + \frac{r(\lambda)}{s(\lambda)}$ be a rational function obtained as the sum of two rational functions expressed in the monomial basis (so that $p(\lambda)$, $q(\lambda)$, $r(\lambda)$, and $s(\lambda)$ are all polynomials). Assume that the numerators and the denominators do not share any common factor, and let $\epsilon := \max\{\deg p(\lambda), \deg q(\lambda)\}$ and $\eta :=$*

$\max\{\deg s(\lambda), \deg r(\lambda)\}$. *Then the matrix pencil*

$$\mathcal{L}(\lambda) = \begin{bmatrix} ps^\top + qr^\top & L_\epsilon^\top(\lambda) \\ L_\eta(\lambda) & 0 \end{bmatrix}$$

is a linearization for $f(\lambda)$*, where* p, q, r*, and* s *are the column vectors containing the coefficients of the polynomials (padded with some leading zeros if the dimensions do not match and ordered according to the basis elements in* $\pi_\epsilon(\lambda)$ *and* $\pi_\eta(\lambda)$*) and* $L_k(\lambda)$ *is the dual basis for the monomial basis of degree* k*.*

Proof. By following the same reasoning of the proof of Theorem 33 we obtain that $\mathcal{L}(\lambda)$ is a linearization for

$$\pi_\epsilon^\top(\lambda)(ps^\top + qr^\top)\pi_\eta(\lambda) = p(\lambda)s(\lambda) + r(\lambda)q(\lambda) = f(\lambda)s(\lambda)q(\lambda),$$

which concludes the proof. ∎

The result can also be extended to the case where different polynomial bases are involved. More precisely, we have the following corollary.

Corollary 36. *Let* $p(\lambda)$*,* $q(\lambda)$*,* $r(\lambda)$*, and* $s(\lambda)$ *be polynomials defined as follows:*

$$p(\lambda) = \sum_{i=0}^{\epsilon} p_i\phi_i(\lambda), \quad q(\lambda) = \sum_{i=0}^{\epsilon} q_i\phi_i(\lambda), \quad r(\lambda) = \sum_{i=0}^{\eta} q_i\psi_i(\lambda), \quad s(\lambda) = \sum_{i=0}^{\eta} s_i\psi_i(\lambda)$$

for some polynomial bases $\{\phi_i\}$ *and* $\{\psi_i\}$*. Let* $\epsilon := \max\{\deg p(\lambda), \deg q(\lambda)\}$ *and* $\eta := \max\{\deg s(\lambda), \deg r(\lambda)\}$ *and* $L_{\epsilon,\phi}(\lambda)$ *and* $L_{\eta,\psi}(\lambda)$ *be dual bases to* $\{\phi_i\}$ *and* $\{\psi_i\}$*, respectively. Assume that* p, q, r*, and* s *are vectors containing the coefficients of the above polynomials in the order coherent with* $\pi_{\epsilon,\phi}(\lambda)$ *and* $\pi_{\eta,\psi}(\lambda)$*. Then the matrix pencil*

$$\mathcal{L}(\lambda) = \begin{bmatrix} ps^\top + qr^\top & L_{\epsilon,\phi}^\top(\lambda) \\ L_{\eta,\psi}(\lambda) & 0 \end{bmatrix}$$

is a linearization for both $f_1(\lambda) = \frac{p(\lambda)}{q(\lambda)} + \frac{r(\lambda)}{s(\lambda)}$ *and* $f_2(\lambda) = \frac{p(\lambda)}{r(\lambda)} + \frac{q(\lambda)}{s(\lambda)}$*, where* p, q, r*, and* s *are the column vectors containing the coefficients of the polynomials (padded with some leading zeros if the dimensions do not match).*

Proof. By following the same proof of Theorem 35 we obtain that $\mathcal{L}(\lambda)$ is a linearization for the polynomial

$$t(\lambda) = \pi_{\epsilon,\phi}^\top(\lambda)(ps^\top + qr^\top)\pi_{\eta,\psi}(\lambda) = p(\lambda)s(\lambda) + q(\lambda)r(\lambda)$$

which has the same roots as the rational functions

$$f_1(\lambda) = \frac{p(\lambda)}{q(\lambda)} + \frac{r(\lambda)}{s(\lambda)}, \qquad f_2(\lambda) = \frac{p(\lambda)}{r(\lambda)} + \frac{q(\lambda)}{s(\lambda)}.$$

This concludes the proof. ∎

Remark 37. The above result shows that we can handle two specific cases. First, the case where each rational function is defined using polynomials in a certain basis, and second, the one where both the denominators and the numerators share a common basis.

An application of the above results is to find the intersection of two rational functions. As in the previous case, according to Theorem 15, $\mathcal{L}(\lambda)$ is a linearization for a polynomial of grade $\epsilon+\eta+1 = \max\{\deg p(\lambda), \deg q(\lambda)\}+\max\{\deg r(\lambda), \deg s(\lambda)\}+1$, while the degree of the polynomial $f(\lambda)s(\lambda)q(\lambda)$ is $\max\{\deg p(\lambda)+\deg s(\lambda), \deg r(\lambda)+\deg q(\lambda)\}$.

Since the first quantity is always larger than the second one, the linearization introduces at least one infinite eigenvalue. However, in many interesting cases, such as when the degree of the numerator and the denominator is the same in each rational function, we only have one spurious infinite eigenvalue, which can be deflated easily.

The result can, however, be improved, and, for these cases, we can build a strong linearization relying on the following.

THEOREM 38. *Let $f(\lambda)$ be a rational function with the same hypotheses and notation of Corollary* 36*, and assume that there exist two $\epsilon \times (\epsilon - 1)$ matrices A and B such that*

$$\pi_{\epsilon,\phi}(\lambda) = (\lambda A + B)\pi_{\epsilon-1,\phi}(\lambda).$$

Then the matrix pencil

$$\mathcal{L}(\lambda) = \begin{bmatrix} (\lambda A + B)^\top ps^\top - (\lambda A + B)^\top qr^\top & L_{\epsilon-1,\phi}^\top(\lambda) \\ L_{\eta,\psi}(\lambda) & 0 \end{bmatrix}$$

is a strong linearization for $f_1(\lambda)$ and $f_2(\lambda)$.

Proof. By again applying Theorem 15 we obtain that $\mathcal{L}(\lambda)$ is a linearization for

$$\begin{aligned} t(\lambda) &= \pi_{\epsilon-1,\phi}^\top(\lambda) \left[(\lambda A + B)^\top ps^\top - (\lambda A + B)^\top qr^\top\right] \pi_{\eta,\psi}(\lambda) \\ &= \pi_{\epsilon,\phi}^\top(\lambda) \left(ps^\top - qr^\top\right) \pi_{\eta,\psi}(\lambda) \\ &= p(\lambda)s(\lambda) + q(\lambda)r(\lambda). \end{aligned}$$

Since $t(\lambda)$ has degree $\epsilon + \eta$, which is the size of $\mathcal{L}(\lambda)$, there are no extra infinite eigenvalues, and so this is a strong linearization. ∎

Remark 39. The hypotheses of Theorem 38 are satisfied in many cases. Some concrete examples are the following:

(i) When $\{\phi_i\}$ is a degree graded basis for $\mathbb{F}_k[\lambda]$, then $\phi_{k+1}(\lambda)$ has degree $k+1$ and we can find a so that $\lambda^k = a^\top \pi_{k,\phi}(\lambda)$. If we choose α to be the leading coefficient of $\phi_{k+1}(\lambda)$, we have

$$\phi_{k+1}(\lambda) - \lambda\alpha a^\top \pi_{k,\phi}(\lambda) = b^\top \pi_{k,\phi}(\lambda)$$

for some $b \in \mathbb{F}^{k+1}$, since the left-hand side has degree k. This implies that

$$\pi_{k+1,\phi}(\lambda) = \left(\lambda \begin{bmatrix} & \alpha a^\top & \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} & b^\top & \\ 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}\right) \pi_{k,\phi}(\lambda).$$

(ii) When $\{\phi_i\}$ is an orthogonal basis, then it is also degree graded, and so the above result applies. In this case, however, it is very easy to get an explicit expression for α, a, and b, since they just contain the coefficients of the recurrence relation that allows one to obtain $\phi_{k+1}(\lambda)$ starting from the previous terms.

(iii) If $\{\phi_i\}$ is the Lagrange basis, we can still find suitable matrices A and B so that the hypotheses are satisfied. Assume that $\pi_{k,\phi}(\lambda)$ is the Lagrange basis on the interpolation nodes $\sigma_1, \dots, \sigma_k$ and that $\pi_{k+1,\phi}(\lambda)$ has the additional node σ_{k+1}. Then we have

$$\pi_{k+1,\phi}(\lambda) = \begin{bmatrix} & \alpha(\lambda-\sigma_k)e_1^\top & \\ \frac{\lambda-\sigma_{k+1}}{\sigma_k-\sigma_{k+1}} & & \\ & \ddots & \\ & & \frac{\lambda-\sigma_{k+1}}{\sigma_1-\sigma_{k+1}} \end{bmatrix} \pi_{k,\phi}(\lambda),$$

where $\alpha = \frac{1}{\sigma_{k+1}-\sigma_k} \prod_{j=1}^{k-1} \frac{\sigma_k-\sigma_j}{\sigma_{k+1}-\sigma_j}$.

Remark 40. Notice that the requirement needs to hold only for one of the two families of polynomials. If the relation holds on $\{\psi_i\}$ instead of $\{\phi_i\}$, the procedure is analogous.

As a concrete example, we report here how the (nonstrong) linearization looks when considering the following rational function:

$$f(\lambda) = \frac{2\lambda^2-1}{\lambda^2+\lambda+3} + \frac{T_1(\lambda)+T_0(\lambda)}{T_1(\lambda)-T_0(\lambda)},$$

where $T_j(\lambda)$ are the Chebyshev polynomials of the first kind and we have chosen $\{\phi_i\} = \{1, \lambda, \lambda^2\}$ and $\{\psi_i\} = \{T_0(\lambda), T_1(\lambda)\}$. We have that p, q, r, and s are given by

$$p = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \quad q = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \quad r = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad s = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We get

$$\mathcal{L}(\lambda) = \begin{bmatrix} ps^\top + qr^\top & L_{\epsilon,\phi}^\top(\lambda) \\ L_{\eta,\psi}(\lambda) & 0 \end{bmatrix} = \left[\begin{array}{cc|cc} 3 & -1 & 1 & 0 \\ 1 & 1 & -\lambda & 1 \\ 2 & 4 & 0 & -\lambda \\ \hline 1 & -\lambda & 0 & 0 \end{array}\right].$$

By Theorem 38 we can also obtain a strong linearization for $f(\lambda)$. In the monomial case the A and B matrices of the hypothesis are given by

$$A = e_1^{(k+1)}(e_1^{(k)})^\top, \qquad B = \begin{bmatrix} 0 & \dots & 0 \\ & I_k & \end{bmatrix},$$

where $e_i^{(k)}$ is the ith column of I_k. A straightforward application of Theorem 38 yields the linearization

$$\mathcal{L}(\lambda) = \left[\begin{array}{cc|c} 3\lambda+1 & 1-\lambda & 1 \\ 2 & 4 & -\lambda \\ \hline 1 & -\lambda & 0 \end{array}\right],$$

which is a strong linearization for the rational function $f(\lambda)$.

In the following we report some numerical experiments that show the effectiveness of the approach.

TABLE 2
Norm of the absolute error on the computed (finite) roots of the rational function $f(\lambda)$.

Degree	Theorem 35	Theorem 38
5	3.29e-16	2.98e-16
10	4.37e-16	4.01e-16
20	5.47e-16	5.07e-16
40	6.87e-16	5.75e-16
80	1.14e-15	7.93e-16
160	1.72e-15	1.40e-15
320	2.57e-15	2.06e-15
640	4.21e-15	3.53e-15

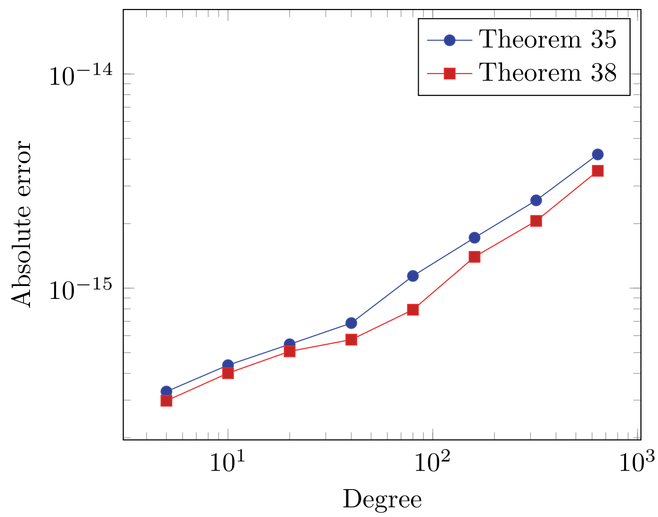

FIG. 3. *Norm of the absolute error on the computed roots of the rational function $f(\lambda)$. Both the strong and the nonstrong versions of the linearization have been tested.*

Numerical experiment 2. Here we test the linearization for the solution of the sums of rational functions. We generate four polynomials $p(\lambda)$, $q(\lambda)$, $r(\lambda)$, and $s(\lambda)$ of the same degree n, and with $p(\lambda), q(\lambda)$ being in the monomial basis and $r(\lambda)$ and $s(\lambda)$ in the Chebyshev one. Their coefficients have been generated using the `randn` function in MATLAB.

We then find the zeros of the rational function

$$f(\lambda) := \frac{p(\lambda)}{q(\lambda)} + \frac{r(\lambda)}{s(\lambda)}$$

by applying Theorems 35 and 38 and using the QZ algorithm on the obtained linearizations. We compare the results with those obtained by symbolically computing the coefficients of the polynomial $t(\lambda) := p(\lambda)s(\lambda) + r(\lambda)q(\lambda)$ and computing its roots with 16 guaranteed digits using `MPSolve` [8]. The experiments have been repeated 50 times, and an average has been taken. The results are reported in Table 2 and Figure 3.

5. Preserving even, odd, and palindromic structures. In this section we deal with the following problem: we consider the case where a matrix polynomial has a $\star$-even, $\star$-odd, or $\star$-palindromic structure. These are often found in applications and are of particular interest, since they induce some symmetries on the spectrum.

For this reason it is important to develop linearizations that enjoy the same structure, so the symmetries in the spectrum will be preserved. Many authors have investigated this problem in recent years, providing different solutions [29, 31, 11, 32]. Linearizations for these structures have been found by exploiting the generality of the $\mathbb{L}_1$ and $\mathbb{L}_2$ spaces of linearizations introduced in [30] or the flexibility offered by Fiedler companion forms. Our approach leads to very similar results (in fact, the linearizations that we build can be recovered following the methods in the previous references), but is instead based on the freedom in choosing the polynomial families $\{\phi_i\}$ and $\{\psi_i\}$.

Here we often use $\star$ in place of the transpose or conjugate transpose operator, since the constructions are valid for both choices. We give the definitions of these structures.

DEFINITION 41. *A matrix polynomial $P(\lambda)$ is* $\star$-even *if $P(\lambda) = P(-\lambda)^\star$. Similarly, we say that $P(\lambda)$ is* $\star$-odd *if $P(\lambda) = -P(-\lambda)^\star$.*

DEFINITION 42. *A matrix polynomial $P(\lambda)$ is said to be* $\star$-palindromic *if $P(\lambda) = \operatorname{rev} P(\lambda)^\star$. Similarly. we say that $P(\lambda)$ is* anti $\star$-palindromic *if $P(\lambda) = -\operatorname{rev} P(\lambda)^\star$.*

Notice that all these relations induce a certain symmetry on the coefficients in the monomial basis. In particular, we have the following, whose proof can be found in [29].

LEMMA 43. *Let $P(\lambda)$ a matrix polynomial. Then the following hold:*

(i) *If the matrix polynomial is $\star$-palindromic or anti $\star$-palindromic, the eigenvalues come in pairs $(\lambda, \frac{1}{\lambda})$ when $\star = \top$ and $(\overline{\lambda}, \frac{1}{\lambda})$ when $\star = H$.*

(ii) *If the matrix polynomial is $\star$-even or $\star$-odd, the eigenvalues come in pairs $(\lambda, -\lambda)$ when $\star = \top$ and $(\lambda, -\overline{\lambda})$ when $\star = H$.*

Moreover, all the paired eigenvalues have the same algebraic and geometric multiplicities.

5.1. Even and odd polynomials. In this section we deal with linearizing even and odd polynomials. In practice we only consider the case of even polynomials since the other is analogous.

THEOREM 44. *Let $P(\lambda) = \sum_{i=0}^{2k+1} P_i \lambda^i$ be a $\star$-even matrix polynomial of grade $2k+1$. Then the even matrix pencil*

$$\mathcal{L}(\lambda) = \left[\begin{array}{cccc|cccc} (-1)^k(\lambda P_{2k+1} + P_{2k}) & & & & I & & & \\ & \ddots & & & -\lambda I & \ddots & & \\ & & \ddots & & & \ddots & I & \\ & & & \lambda P_1 + P_0 & & & -\lambda I & \\ \hline I & \lambda I & & & & & & \\ & \ddots & \ddots & & & & & \\ & & I & \lambda I & & & & \end{array}\right]$$

is a linearization for $P(\lambda)$.

Proof. It is immediate to verify that the pencil is $\star$-even. In order to check that it is a linearization for the correct polynomial we can see that the top-right block and bottom-left block are of the form $L_{k,\phi}(\lambda) \otimes I_m$ and $L_{k,\psi}(\lambda) \otimes I_m$, respectively, with

$L_{k,\star}(\lambda)$ being dual bases for

$$\pi_{k,\phi}(\lambda) = \begin{bmatrix} \lambda^k \\ \vdots \\ \lambda \\ 1 \end{bmatrix}, \qquad \pi_{k,\psi}(\lambda) = \begin{bmatrix} (-1)^k \lambda^k \\ \vdots \\ -\lambda \\ 1 \end{bmatrix}.$$

Applying Theorem 15 guarantees that $\mathcal{L}(\lambda)$ is a linearization for the matrix polynomial

$$(\pi_{k,\phi}(\lambda) \otimes I_m)^\top \operatorname{diag}((-1)^j(\lambda P_{2j+1} + P_{2j}))_{j=0,\ldots,k}(\pi_{k,\psi}(\lambda) \otimes I_m) = P(\lambda),$$

which concludes the proof. ∎

5.2. Palindromic polynomials. A similar procedure can be applied to obtain $\star$-palindromic linearizations for $\star$-palindromic polynomials. However, the construction in this case is slightly more complicated. We first prove the following lemma, which provides linearizations with a $\star$-palindromic structure, and then we show how to choose the top-left block to linearize a concrete $\star$-palindromic matrix polynomial.

THEOREM 45. *Let $\{\phi_i\}$ and $\{\psi_i\}$ be the polynomial bases defined by*

$$\phi_i = \lambda^{k-i}, \qquad \psi_i = \lambda^i, \qquad i = 0, \ldots, k.$$

Then two dual minimal bases $L_{k,\phi}(\lambda)$ and $L_{k,\psi}(\lambda)$ for $\{\phi_i\}$ and $\{\psi_i\}$, respectively, are given by

$$L_{k,\phi}(\lambda) = \begin{bmatrix} 1 & -\lambda & & \\ & \ddots & \ddots & \\ & & 1 & -\lambda \end{bmatrix}, \qquad L_{k,\psi}(\lambda) = \begin{bmatrix} \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \lambda & -1 \end{bmatrix},$$

and the $\star$-palindromic pencil

$$\mathcal{L}(\lambda) = \begin{bmatrix} \lambda M + M^\star & L_{k,\phi}(\lambda)^\star \otimes I_m \\ L_{k,\psi}(\lambda) \otimes I_m & 0 \end{bmatrix}, \quad M = \begin{bmatrix} M_{0,0} & \ldots & M_{0,k} \\ \vdots & & \vdots \\ M_{k,0} & \ldots & M_{k,k} \end{bmatrix},$$

where $M_{ij} \in \mathbb{C}^{m\times m}$, is a linearization for the degree $2k+1$ matrix polynomial defined by

$$P(\lambda) = \sum_{i,j=0}^{k} M_{j,i}^\star \lambda^{k+j-i} + \sum_{i,j=1}^{k} M_{i,j}\lambda^{k+j-i+1}.$$

Proof. It is immediate to verify that the given matrices $L_{k,\phi}(\lambda)$ and $L_{k,\psi}(\lambda)$ are indeed dual minimal bases. By applying Theorem 15 we get $P(\lambda)$ as

$$P(\lambda) = \begin{bmatrix} \lambda^k I_m & \cdots & I_m \end{bmatrix} (\lambda M + M^\star) \begin{bmatrix} I_m \\ \vdots \\ \lambda^k I_m \end{bmatrix}.$$

This concludes the proof. ∎

The result above can be used to construct $\star$-palindromic linearizations for a $\star$-palindromic matrix polynomial $P(\lambda) = \sum_{j=0}^{n} P_j \lambda^j$. We want to describe a procedure to choose the block coefficients of M in Theorem 45 in order to make $\mathcal{L}(\lambda)$ a linearization for $P(\lambda)$.

DEFINITION 46 (block diagonal sum). *Let M be a square matrix of size $mk \times mk$, partitioned in $m \times m$ blocks, denoted by $[M]_{i,j}$. Then we define $X = \mathrm{bds}_m(M, d)$ as the sum of the matrices along the dth block diagonal of the matrix M, that is,*

$$\mathrm{bds}_m(M, d) := \sum_{j-i=d} [M]_{i,j},$$

where we set $[M_{i,j}] = 0$ if i or j is smaller than 1 or if $i \geqslant k$ or $j \geqslant k$. We refer to X as the dth block diagonal sum of M.

Remark 47. The pencil $\mathcal{L}(\lambda)$ defined in Theorem 45 is a linearization for a matrix polynomial $P(\lambda)$ of degree $2k+1$ if and only if the relation

$$P_s = \mathrm{bds}_m(M, s-k-1) + \mathrm{bds}_m(M, k-s)^\star \tag{4}$$

holds for any $s = 0, \dots, 2k+1$, where $P(\lambda) = \sum_{s=0}^{2k+1} P_s \lambda^s$.

Notice that Remark 47 can also be used to build the linearization starting from its coefficients. In fact, the relation (4) for $s \in \{0, 2k+1\}$ simplifies to

$$P_0 = M_{1,k+1}^\star, \qquad P_{2k+1} = M_{1,k+1}.$$

Having determined the term in position $(1, k)$, one can then proceed to fill in the others by imposing the condition of Remark 47.

Here we provide a concrete example of such a construction. However, we stress that is not the only possible choice.

THEOREM 48. *Let $P(\lambda) = \sum_{i=0}^{2k+1} P_i \lambda^i$ be a $\star$-palindromic matrix polynomial with degree $2k+1$. Then the pencil of Theorem 45 with*

$$M = \begin{bmatrix} 0_m & \cdots & 0_m & P_0^\star \\ \vdots & & \vdots & \vdots \\ 0_m & \cdots & 0_m & P_k^\star \end{bmatrix} \tag{5}$$

is a $\star$-palindromic linearization for $P(\lambda)$.

Proof. Notice that in (5), we have that the only nonzero block diagonal elements of M are in the last block column and $M_{i,k+1} = P_{i-1}^\star = P_{2k+2-i}$. We now check that (4) holds for every $s = 0, \dots, 2k+1$. If $0 \leqslant s \leqslant k$, we have $\mathrm{bds}_m(M, s-k-1) = 0$ and $\mathrm{bds}_m(M, k-s)^\star = M_{i,k+1}^\star$, where i is such that $k+1-i = k-s$ (being on the $(k-s)$th block diagonal). This implies that $i = s+1$, and so $M_{i,k+1}^\star = P_s$, as desired. On the other hand, if $k+1 \leqslant s \leqslant 2k+1$, we similarly have $\mathrm{bds}_m(M, k-s) = 0$ and $\mathrm{bds}_m(M, s-k-1) = M_{i,k+1}$ with $k+1-i = s-k-1$ so that $i = 2k+2-s$. This again implies that $M_{i,k+1} = P_{2k+2-i} = P_{2k+2-(2k+2-s)} = P_s$. This concludes the proof. ∎

6. Deflation of infinite eigenvalues. We have observed that in the polynomial sum case of section 4.1 the linearization built according to Theorem 33 is generally not strong and might have many infinite eigenvalues.

In this section we show what the structure of the infinite eigenvalues is and a possible strategy to deflate them based on a simplified approach inspired by [5, 34]. In our case we have the advantage of knowing exactly which eigenvalue we want to deflate, and we can completely characterize the structure of the block in the Kronecker canonical form corresponding to the infinite eigenvalue.

LEMMA 49. *Let $\mathcal{L}(\lambda)$ be the linearization obtained from Theorem* 33 *for the sum of two arbitrary polynomials. Then there exist two unitary bases Q and Z and A_0, A_1, B_0, B_1, with B_1 invertible, such that*

$$Q^H \mathcal{L}(\lambda) Z = \begin{bmatrix} I - \lambda J & \lambda A_1 - A_0 \\ 0 & \lambda B_1 - B_0 \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Proof. Such a decomposition can be obtained by following the deflation procedure for the infinite eigenvalue described in [5] and [34]. We only need to prove that the pencil obtained in the top-left entry of the transformed matrix is exactly $I - \lambda J$. Let A, B be matrices such that $\mathcal{L}(\lambda) = A - \lambda B$. We note that B has nullity equal to 1 in our construction. Recall that the columns of Q and Z are orthogonal bases of the sequence of spaces defined by

$$\mathcal{Z}_i = \begin{cases} \{0\} & \text{if } i = 0, \\ B^{-1} \mathcal{Q}_{i-1} & \text{otherwise}, \end{cases} \qquad \mathcal{Q}_i = A \mathcal{Z}_i,$$

where B^{-1} is the preimage of B. The fact that B has nullity 1 implies that the dimension of $\mathcal{Z}_i$ can increase by at most 1 at each step. This means that there exists a unique diagonal block in the Kronecker canonical form corresponding to the infinite eigenvalue, whose size is exactly equal to its algebraic multiplicity. Since $\lambda B_1 - B_0$ does not have infinite eigenvalues, B_1 is invertible, as requested. ∎

We can use the algorithm described in [5] to compute the matrices Q and Z and then compute the eigenvalues of the pencil $\lambda B_1 - B_0$ instead of $\mathcal{L}(\lambda)$. Experiments using this strategy were reported in section 4.1.

For a more in-depth discussion of the above approach to deflation see the work of Berger and Reis [6], which is based on the analysis originally carried out by Wong [35].

7. Conclusions. We have provided an extension of the main theorem of [14] to construct linearizations. This new result makes it easier to construct many linearizations for, among others, sums of polynomials and rational functions, and allows one to realize structure preserving pencils.

We think that the flexibility offered by the adjustment of the dual basis in the pencil $\mathcal{L}(\lambda)$ allows for even further improvement and for the coverage of more structures. We think that in many cases this construction provides an alternative to other known approaches to find structured linearizations, such as looking in the spaces $\mathbb{L}_1$ and $\mathbb{L}_2$ from [30].

Acknowledgments. We wish to thank Piers Lawrence, who helped us to understand the construction of dual bases for polynomials defined in the Lagrange basis, and Thomas Mach for inspiring discussions. Moreover, we wish to thank the anonymous referees for their comments, which helped to improve the quality and the organization of the paper.

REFERENCES

[1] A. Amiraslani, R. M. Corless, and P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*, IMA J. Numer. Anal., 29 (2009), pp. 141–157, https://doi.org/10.1093/imanum/drm051.

[2] E. N. Antoniou and S. Vologiannidis, *A new family of companion forms of polynomial matrices*, Electron. J. Linear Algebra, 11 (2004), pp. 78–87.

[3] S. Barnett, *A companion matrix analogue for orthogonal polynomials*, Linear Algebra Appl., 12 (1975), pp. 197–202, https://doi.org/10.1016/0024-3795(75)90041-5.

[4] T. Beelen and P. van Dooren, *A pencil approach for embedding a polynomial matrix into a unimodular matrix*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 77–89, https://doi.org/10.1137/0609006.

[5] Th. Beelen and P. van Dooren, *An improved algorithm for the computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl., 105 (1988), pp. 9–65, https://doi.org/10.1016/0024-3795(88)90003-1.

[6] T. Berger and T. Reis, *Controllability of linear differential-algebraic systems—a survey*, in Surveys in Differential-Algebraic Equations I, Springer, Heidelberg, 2013, pp. 1–61, https://doi.org/10.1007/978-3-642-34928-7_1.

[7] D. A. Bini, L. Gemignani, and J. R. Winkler, *Structured matrix methods for CAGD: An application to computing the resultant of polynomials in the Bernstein basis*, Numer. Linear Algebra Appl., 12 (2005), pp. 685–698, https://doi.org/10.1002/nla.444.

[8] D. A. Bini and L. Robol, *Solving secular and polynomial equations: A multiprecision algorithm*, J. Comput. Appl. Math., 272 (2014), pp. 276–292, https://doi.org/10.1016/j.cam.2013.04.037.

[9] R. M. Corless, *On a generalized companion matrix pencil for matrix polynomials expressed in the Lagrange basis*, in Symbolic-Numeric Computation, Trends Math., Birkhäuser, Basel, 2007, pp. 1–15.

[10] 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 (2010), pp. 2181–2204, https://doi.org/10.1137/090772927.

[11] F. De Terán, F. M. Dopico, and D. S. Mackey, *Palindromic companion forms for matrix polynomials of odd degree*, J. Comput. Appl. Math., 236 (2011), pp. 1464–1480, https://doi.org/10.1016/j.cam.2011.09.010.

[12] F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations for rectangular matrix polynomials*, Linear Algebra Appl., 437 (2012), pp. 957–991, https://doi.org/10.1016/j.laa.2012.03.028.

[13] F. De Terán, F. M. Dopico, and D. S. Mackey, *Spectral equivalence of matrix polynomials and the index sum theorem*, Linear Algebra Appl., 459 (2014), pp. 264–333, https://doi.org/10.1016/j.laa.2014.07.007.

[14] F. M. Dopico, P. W. Lawrence, J. A. Pérez, and P. Van Dooren, *Block Kronecker Linearizations of Matrix Polynomials and Their Backward Errors*, MIMS EPrint, 2016.

[15] G. Farin, *A history of curves and surfaces in CAGD*, in Handbook of Computer Aided Geometric Design, North–Holland, Amsterdam, 2002, pp. 1–21.

[16] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide*, Elsevier, Amsterdam, 2014.

[17] R. T. Farouki, T. N. T. Goodman, and T. Sauer, *Construction of orthogonal bases for polynomials in Bernstein form on triangular and simplex domains*, Comput. Aided Geom. Design, 20 (2003), pp. 209–230, https://doi.org/10.1016/S0167-8396(03)00025-6.

[18] H. Fassbender, J. A. Pérez, and N. Shayanfar, *A Sparse Linearization for Hermite Interpolation Matrix Polynomials*, MIMS Eprint, 2015.

[19] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl., 372 (2003), pp. 325–331, https://doi.org/10.1016/S0024-3795(03)00548-2.

[20] G. D. Forney, Jr., *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. Control, 13 (1975), pp. 493–520, https://doi.org/10.1137/0313029.

[21] F. R. Gantmacher, *The Theory of Matrices,* Vol. 1, AMS Chelsea, Providence, RI, 1998, translated from the Russian by K. A. Hirsch, reprint of the 1959 translation.

[22] W. Gautschi, *The condition of Vandermonde-like matrices involving orthogonal polynomials*, Linear Algebra Appl., 52-53 (1983), pp. 293–300, https://doi.org/10.1016/0024-3795(83)80020-2.

[23] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Classics Appl. Math. 58, SIAM, Philadelphia, 2009, https://doi.org/10.1137/1.9780898719024.

[24] I. J. Good, *The colleague matrix, a Chebyshev analogue of the companion matrix*, Quart. J. Math., 12 (1961), pp. 61–68, https://doi.org/10.1093/qmath/12.1.61.

[25] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 143–159, https://doi.org/10.1137/050646202.

[26] T. Kailath, *Linear Systems*, Prentice–Hall Information and System Sciences Series 156, Prentice–Hall, Englewood Cliffs, NJ, 1980.

[27] P. Lawrence and J. Perez, *Constructing Strong Linearizations of Matrix Polynomials Expressed in the Chebyshev Bases*, MIMS EPrint, 2016.

[28] P. W. Lawrence and R. M. Corless, *Numerical stability of barycentric Hermite root-finding*, in Proceedings of the 2011 International Workshop on Symbolic-Numeric Computation, ACM, New York, 2012, pp. 147–148, https://doi.org/10.1145/2331684.2331706.

[29] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Structured polynomial eigenvalue problems: Good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1029–1051, https://doi.org/10.1137/050628362.

[30] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004, https://doi.org/10.1137/050628350.

[31] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Numerical methods for palindromic eigenvalue problems: Computing the anti-triangular Schur form*, Numer. Linear Algebra Appl., 16 (2009), pp. 63–86, https://doi.org/10.1002/nla.612.

[32] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Jordan structures of alternating matrix polynomials*, Linear Algebra Appl., 432 (2010), pp. 867–891, https://doi.org/10.1016/j.laa.2009.10.002.

[33] C. Schulz, *Bézier clipping is quadratically convergent*, Comput. Aided Geom. Design, 26 (2009), pp. 61–74, https://doi.org/10.1016/j.cagd.2007.12.006.

[34] P. Van Dooren, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl., 27 (1979), pp. 103–140, https://doi.org/10.1016/0024-3795(79)90035-1.

[35] K.-T. Wong, *The eigenvalue problem $\lambda Tx + Sx$*, J. Differential Equations, 16 (1974), pp. 270–280, https://doi.org/10.1016/0022-0396(74)90014-X.