The $\mathbb{DL}(P)$ vector space of pencils for singular matrix polynomials

Froilán M. Dopico [*] and Vanni Noferini[†]

December 16, 2022

Abstract

Given a possibly singular matrix polynomial $P(z)$, we study how the eigenvalues, eigenvectors, root polynomials, minimal indices, and minimal bases of the pencils in the vector space $\mathbb{DL}(P)$ introduced in Mackey, Mackey, Mehl, and Mehrmann [SIAM J. Matrix Anal. Appl. 28(4), 971-1004, 2006] are related to those of $P(z)$. If $P(z)$ is regular, it is known that those pencils in $\mathbb{DL}(P)$ satisfying the generic assumptions in the so-called eigenvalue exclusion theorem are strong linearizations for $P(z)$. This property and the block-symmetric structure of the pencils in $\mathbb{DL}(P)$ have made these linearizations among the most influential for the theoretical and numerical treatment of structured regular matrix polynomials. However, it is also known that, if $P(z)$ is singular, then none of the pencils in $\mathbb{DL}(P)$ is a linearization for $P(z)$. In this paper, we prove that despite this fact a generalization of the eigenvalue exclusion theorem holds for any singular matrix polynomial $P(z)$ and that such a generalization allows us to recover all the relevant quantities of $P(z)$ from any pencil in $\mathbb{DL}(P)$ satisfying the eigenvalue exclusion hypothesis. Our proof of this general theorem relies heavily in the representation of the pencils in $\mathbb{DL}(P)$ via Bézoutians by Nakatsukasa, Noferini and Townsend [SIAM J. Matrix Anal. Appl. 38(1), 181-209, 2015].

keyword Bézout matrix, Bézoutian, $\mathbb{DL}(P)$, singular matrix polynomial, minimal basis, minimal indices, root polynomial, eigenvalue exclusion theorem, linearization
2020 MSC: 65F15, 15A18, 15A22, 15A54

1 Introduction

Given an algebraically closed field $\mathbb{F}$ with characteristic[1] 0, this paper deals with a matrix polynomial

$$P(z) = \sum_{i=0}^{k} z^i P_i, \quad P_i \in \mathbb{F}^{m \times n}, \tag{1}$$

[1]The assumption that the base field is closed is for simplicity of exposition, as it ensures for instance that finite eigenvalues all lie in $\mathbb{F}$. The assumption that $\mathrm{char}(\mathbb{F}) = 0$ is only needed in two technical results, Lemma 4.7 and Theorem 4.8, that simplify the proof of our main results. The main results of this paper could also be proved without this assumption, but this would complicate the exposition.

of grade k, since P_k is allowed to be zero, and with one of the most influential families of pencils, i.e., matrix polynomials of grade 1, associated to $P(z)$: the vector space of pencils $\mathbb{DL}(P)$ introduced by Mackey, Mackey, Mehl, and Mehrmann in [29] in the square case $m = n$. The matrix polynomial $P(z)$ is said to be regular if $m = n$ and its determinant is not identically zero. Otherwise, $P(z)$ is said to be singular. Regular and singular matrix polynomials arise in many applications [3, 19, 41, 43] and one of the most reliable methods for computing their eigenvalues, eigenvectors and minimal indices and bases is via *linearization*. See Section 2 for precise definitions of these and other concepts appearing in this introduction.

A linearization of $P(z)$ in (1) is a matrix pencil $L(z)$ with the same finite eigenvalues and associated partial multiplicities as $P(z)$ and also with the same number of right and left minimal indices. The linearization $L(z)$ is strong if, in addition, it has an eigenvalue at infinity whenever $P(z)$ has an eigenvalue at infinity, with the same partial multiplicities. Linearizations and strong linearizations are very important in numerical computations because, when $\mathbb{F} \subseteq \mathbb{C}$, there exist reliable algorithms for approximating eigenvalues and computing minimal indices of matrix pencils, either regular [32] or singular [42]. Applying these algorithms to a strong linearization allows us to get the corresponding magnitudes of $P(z)$. As a consequence many families of strong linearizations of matrix polynomials have been developed and studied in the last two decades. We emphasize in this context the seminal work [29]. See also, for instance, [2, 12, 22, 30, 34] and the references therein among many other references on linearizations.

Among the many existing families of linearizations, the one based on the vector space of pencils $\mathbb{DL}(P)$ introduced in [29] has been particularly relevant and influential for several reasons in the study and numerical treatment of *regular* matrix polynomials $P(z)$. To begin with, for many classes of structured regular matrix polynomials $P(z)$ appearing in applications (symmetric, Hermitian, palindromic, odd, even, ...), the pencils in $\mathbb{DL}(P)$ allow us to construct linearizations with the same structure as $P(z)$ [22, 30], which has important theoretical and numerical consequences. In addition, the linearizations in $\mathbb{DL}(P)$ allow us to recover the left eigenvectors of $P(z)$ from the corresponding ones of the linearization exactly in the same way as the right eigenvectors [22, 30]. Moreover, the numerical properties of the linearizations in $\mathbb{DL}(P)$ were studied in [21, 23], where it was argued that, under some reasonable assumptions, some linearizations in $\mathbb{DL}(P)$ have satisfactory properties with respect to the conditioning of eigenvalues and the backward errors of eigenpairs, though some recent results [5] indicate that such conclusions should be reconsidered. Finally, we mention that the linearizations in $\mathbb{DL}(P)$ have played a fundamental role in the description and parametrization of structure-preserving transformations for matrix polynomials in [18], and that a generalization of the space $\mathbb{DL}(P)$, that allows to go beyond ansatz polynomials of bounded degree by considering ansatz polynomials of any degree and even more generally ansatz functions, was studied in [34].

An important remark is that, given a *regular* matrix polynomial $P(z)$ of grade k, not every pencil in the vector space $\mathbb{DL}(P)$ is a linearization for $P(z)$. In order to explain this fact, let us recall that each nonzero pencil in $\mathbb{DL}(P)$ is uniquely determined by a nonzero vector $\omega = \begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_k \end{bmatrix}^T \in \mathbb{F}^k$, called *ansatz vector* in [29], which gives rise to a scalar polynomial $v(z) = \omega_1 z^{k-1} + \omega_2 z^{k-2} + \cdots + \omega_k$ of grade $k-1$, called *ansatz polynomial* in [34]. It was proved in [29, Theorem 6.7] that, for $P(z)$ regular, the pencil in $\mathbb{DL}(P)$ with ansatz polynomial $v(z)$ is a linearization for $P(z)$ if and only if the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots[2], are disjoint. This neat result was termed in [29] the *eigenvalue exclusion theorem* and allowed to prove in [29, Theorem 6.8] that almost all pencils in $\mathbb{DL}(P)$ are linearizations of $P(z)$ when $P(z)$ is regular. Moreover, it was proved in

---

[2]Recall that the scalar polynomial $v(z) = \omega_1 z^{k-1} + \omega_2 z^{k-2} + \cdots + \omega_k$ of grade $k-1$ has a root at infinity of multiplicity $\ell < k-1$ if $\omega_1 = \cdots = \omega_\ell = 0$ and $\omega_{\ell+1} \neq 0$.

[29, Theorem 4.3] that a pencil in $\mathbb{DL}(P)$ is a linearization for $P(z)$ if and only if is a strong linearization for $P(z)$. On the other hand, neither fact remains true when the matrix polynomial $P(z)$ is singular, since it was proved in [8, Theorem 6.1] that if $P(z)$ is a singular square matrix polynomial, then none of the pencils in $\mathbb{DL}(P)$ is a linearization for $P(z)$.

Theorem 6.1 in [8] is, at a first glance, a very discouraging result, since it seems to indicate that the good properties mentioned above of the pencils in $\mathbb{DL}(P)$ cannot be used at all for singular matrix polynomials. Moreover, it seems to put $\mathbb{DL}(P)$ at disadvantage with respect to other families of linearizations of matrix polynomials, which are linearizations for both regular and singular matrix polynomials [4, 7, 8, 9, 10, 12, 36]. The main purpose of this paper is to prove that the pencils in $\mathbb{DL}(P)$, though they are not linearizations, can still be used for computing the complete eigenstructure of singular matrix polynomials, since they reveal all the structural data of any singular matrix polynomial under exactly the same hypotheses of the eigenvalue exclusion theorem proved in [29] for regular polynomials. Thus, for instance, the pencils in $\mathbb{DL}(P)$ can be used for constructing Hermitian pencils that reveal all the structural data of any, possibly singular, Hermitian matrix polynomial. In this context, the main result of this paper is the following theorem.

Theorem 1.1. *Let $P(z)$ be an $m \times n$, possibly singular, matrix polynomial with coefficients in $\mathbb{F}^{m \times n}$ and of grade $k \geq 2$, right nullity p and left nullity q, and let $v(z)$ be a scalar polynomial of grade $k - 1$ with coefficients in $\mathbb{F}$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots, are disjoint. Denote by $L(z)$ the pencil in $\mathbb{DL}(P)$ with ansatz polynomial $v(z)$. Then:*

1. *The right minimal indices of $P(z)$ are $\gamma_1 \leq \cdots \leq \gamma_p$ if and only if the right minimal indices of $L(z)$ are $\underbrace{0, \ldots, 0}_{p(k-1) \text{ times}}, \gamma_1 \leq \cdots \leq \gamma_p$;*

2. *The left minimal indices of $P(z)$ are $\eta_1 \leq \cdots \leq \eta_q$ if and only if the left minimal indices of $L(z)$ are $\underbrace{0, \ldots, 0}_{q(k-1) \text{ times}}, \eta_1 \leq \cdots \leq \eta_q$;*

3. *$\lambda \in \mathbb{F} \cup \{\infty\}$ is an eigenvalue of $P(z)$ if and only if λ is an eigenvalue of $L(z)$, and its partial multiplicities as an eigenvalue of $P(z)$ and as an eigenvalue of $L(z)$ coincide.*

Note that, if $P(z)$ is regular, then Theorem 1.1 implies the eigenvalue exclusion theorem in [29]: indeed, items 1 and 2 become vacuous as there are no minimal indices (that is, in item 1 $p = 0$ and in item 2 $q = 0$), while item 3 is equivalent to saying that $L(z)$ is a strong linearization of $P(z)$. Hence, we can also consider Theorem 1.1 as a generalization of [29, Theorem 6.7]. Observe that Theorem 1.1 immediately implies that if $P(z)$ is singular, then $L(z)$ is not a linearization for $P(z)$, because the number of left (or right) minimal indices of $P(z)$ is different that the number of left (or right) minimal indices of $L(z)$ [11]. However, the key point is that for practical purposes such a pencil is not worse than a strong linearization: the minimal indices of $L(z)$ are those of $P(z)$ with the addition of a fixed number of extra zero minimal indices. Hence, the additional minimal indices can be trivially identified since the right and left nullities of $P(z)$ are just those of $L(z)$ divided by k. In fact, arguably the recovery of the minimal indices from the pencils in $\mathbb{DL}(P)$ for a singular $P(z)$ is simpler than their recovery from other families of pencils that are strong linearizations, including the classical companion pencils, whose minimal indices are not exactly those of $P(z)$, but shifted sequences of them [4, 8, 9, 12, 36].

We disclose in advance that the proof of Theorem 1.1 requires considerable work and that it relies heavily on the bivariate polynomial approach to $\mathbb{DL}(P)$ pencils introduced in [34], which maps the pencils in $\mathbb{DL}(P)$ into Lerer-Tismenetsky Bézoutian matrices. We emphasize that this

approach also allows us to consider for the first time in the literature $\mathbb{DL}(P)$ pencils for matrix polynomials $P(z)$ that may be rectangular without extra effort.

The results in this paper are connected with those in [14] concerning matrix polynomials, though the approaches in both papers are completely different. Given an $m \times n$ complex matrix polynomial, the authors of [14] consider only one pencil $L(z)$ in $\mathbb{DL}(P)$, the one with ansatz polynomial $v(z) = 1$, i.e., the ansatz polynomial with ∞ as unique root. For this particular pencil $L(z)$, it was proved in [14], without assuming the eigenvalue exclusion condition, that Theorem 1.1 holds with the exception of item 3 for the eigenvalue $\lambda = \infty$, for which it was proved that though it has different partial multiplicities in $L(z)$ and in $P(z)$, the ones in $P(z)$ can be recovered from the ones in $L(z)$. The approach in [14] is *ad hoc* in that sense that it relies on making some matrix manipulations that exploit the very peculiar structure of the pencil in $\mathbb{DL}(P)$ with ansatz polynomial $v(z) = 1$ and on the concept of strongly minimal linearization of a matrix polynomial. The latter is different from the standard definitions of linearization and strong linearization coming from [19]. So, the results in [14] are presented in a form rather different from Theorem 1.1. In this paper, in contrast, we systematically consider all the (uncountably many) pencils in the vector space $\mathbb{DL}(P)$ and we prove that they satisfy Theorem 1.1 working with their representation as Bézoutian matrices via a bivariate polynomial mapping.

It is worth mentioning that, with the exception of the particular pencil considered in [14], this paper presents the first example of a wide family of pencils related to a matrix polynomial $P(z)$ that are not linearizations, according to the standard definition introduced by Gohberg, Lancaster and Rodman [19], but that still allows for the recovery of all the structural data of $P(z)$. Since the recovery of such structural data is what is really wanted in applications, the results we have obtained for the pencils in $\mathbb{DL}(P)$ indicate that the Gohberg, Lancaster and Rodman's concepts of linearization and strong linearization may be unnecessarily rigid.

The paper is organized as follows. Section 2 summarizes the background material that is necessary in the rest of the paper. Some preliminary technical results that are used in the proofs of the main results are presented and proved in Section 3. Items 1 and 2 of Theorem 1.1 are proved in Section 4, while item 3 is proved in Section 5. Section 6 studies how to recover minimal bases, eigenvectors, and root polynomials of a possibly singular matrix polynomial $P(z)$ from those of pencils in the vector space $\mathbb{DL}(P)$. Finally, some conclusions are presented in Section 7.

2 Background material

In this section, we revise the necessary background material. Throughout this paper, we consider an algebraically closed field $\mathbb{F}$ with characteristic 0, the ring of polynomials $\mathbb{F}[z]$ with coefficients in $\mathbb{F}$, and its field of fractions $\mathbb{F}(z)$. The sets of $m \times n$ constant matrices over $\mathbb{F}$, of $m \times n$ matrix polynomials over $\mathbb{F}$, and of $m \times n$ rational matrices over $\mathbb{F}$ are denoted by $\mathbb{F}^{m \times n}$, $\mathbb{F}[z]^{m \times n}$ and $\mathbb{F}(z)^{m \times n}$, respectively.

2.1 Basics on matrix polynomials

The normal rank, or simply the rank, of a matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ is denoted by $\operatorname{rank} P(z)$ and is the size of its largest non-identically-zero minor or, equivalently, its rank over the field $F(z)$. The matrix polynomial $P(z)$ is regular if it is square and $\det P(z)$ is not identically zero. Otherwise, $P(z)$ is said to be singular. An element $\lambda \in \mathbb{F}$ is a finite eigenvalue of $P(z)$ if the rank of the constant matrix $P(\lambda) \in \mathbb{F}^{m \times n}$ is less than the normal rank of $P(z)$, i.e., if $\operatorname{rank} P(\lambda) < \operatorname{rank} P(z)$.

A square matrix polynomial $U(z) \in \mathbb{F}[z]^{n \times n}$ is said to be unimodular if its determinant is a nonzero element of $\mathbb{F}$ or, equivalently, if it is invertible and its inverse is a matrix polynomial.

The Smith canonical form of $P(z)$ with $r = \operatorname{rank} P(z)$ is $S(z) = U(z)P(z)V(z)$, where $U(z), V(z)$ are unimodular matrix polynomials and $S(z) = \operatorname{diag}(d_1(z), d_2(z), \dots, d_r(z), 0, \dots, 0)$ is diagonal and such that each nonzero diagonal entry $d_1(z), d_2(z), \dots, d_r(z)$ divides the next one [17]. The diagonal elements of $S(z)$ are unique up to multiplication by units of $\mathbb{F}[z]$ and the nonzero ones are called invariant factors of $P(z)$. Given $\lambda \in \mathbb{F}$, each invariant factor can be written uniquely as $d_i(\lambda) = (z - \lambda)^{\ell_i} \omega_i(z)$, where $\omega_i(\lambda) \neq 0$ and ℓ_i is a nonnegative integer, for $i = 1, \dots, r$. Thus, $\lambda \in \mathbb{F}$ is an eigenvalue of $P(z)$ if and only if at least one ℓ_i is strictly larger than zero. Those ℓ_i that are positive are called the partial multiplicities of λ as an eigenvalue of $P(z)$ and the corresponding factors $(z - \lambda)^{\ell_i}$ are the elementary divisors of $P(z)$ for the eigenvalue λ.

The partial multiplicities at infinity of $P(z)$ are defined after fixing a grade k for $P(z)$, where k is an integer greater than or equal to the degree of $P(z)$, denoted by $\deg P(z)$ and defined as the largest degree of its entries. Then, the reversal of $P(z)$ with respect to the grade k is defined as $\operatorname{Rev}_k P(z) = z^k P(z^{-1})$. Very often the grade k is chosen to be equal to $\deg P(z)$; if this is the case we will omit the suffix k and simply write $\operatorname{Rev} P(z)$. The partial multiplicities of ∞ in $P(z)$, seen as a grade k polynomial, are then defined to be equal to the partial multiplicities of 0 for $\operatorname{Rev}_k P(z)$. It is clear that different choices of the grade k lead to different partial multiplicities at ∞: more precisely, if $k = \deg P(z) + h$, for some $h > 0$, and $\operatorname{rank} P(z) = r$, then $\operatorname{Rev} P(z)$ has partial multiplicities at 0 equal to $0 < \eta_1 \leq \dots \leq \eta_s$ if and only if $\operatorname{Rev}_k P(z)$ has partial multiplicities at 0 equal to $\underbrace{h = \dots = h}_{r-s \text{ times}} \leq \eta_1 + h \leq \dots \leq \eta_s + h$. Note that here s may be or may be not zero, and accordingly ∞ may be not or may be an eigenvalue of $P(z)$ when $k = \deg P(z)$. However, if $k > \deg P(z)$ and $r > 0$, then ∞ is always an eigenvalue of $P(z)$ with r partial multiplicities.

The right kernel, or right null space, of $P(z) \in \mathbb{F}[z]^{m \times n}$ over the field $\mathbb{F}(z)$ is defined as

$$\ker P(z) = \left\{ x(z) \in \mathbb{F}(z)^{n \times 1} \ : \ P(z)x(z) = 0 \right\}.$$

The left kernel of $P(z)$ over $\mathbb{F}(z)$ is defined as the right kernel of the transpose of $P(z)$, which is denoted as $P(z)^T$. In this paper, we mainly work with right kernels, since the results for left kernels are completely analogous, and for this reason we will for simplicity just write "kernel" to mean "right kernel". It is clear that $\ker P(z)$ has polynomial bases, i.e., bases whose vectors belong to $\mathbb{F}[z]^{n \times 1}$. Among the polynomial bases of $\ker P(z)$, those which have minimal sum of the degrees of their vectors are called (right) minimal bases of $P(z)$. There are infinitely many (right) minimal bases of $P(z)$, but all of them have the same ordered sequence of the degrees of their vectors [16]. Such degrees are called the right minimal indices of $P(z)$. Left minimal bases and indices of $P(z)$ are defined as the right ones of $P(z)^T$.

The index sum theorem [11] states that, for every matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ of grade k and rank r, the sum of all the left and right minimal indices plus the sum of all finite and infinite partial multiplicities is equal to $k\,r$.

2.2 Linearizations of matrix polynomials and $\mathbb{DL}(P)$ pencils

In applications of matrix polynomials, one is usually interested in computing their eigenvalues, together with their associated partial multiplicities, and their minimal indices. Often, this is done by linearizing the matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ one is interested in. Linearizations are pencils that allow to obtain the required information of $P(z)$ by coupling a theoretical analysis that relates this information with its analogue for $L(z)$ together with an algorithm such as, for

instance, those in [32, 42]. More precisely a pencil $L(z) = L_0 + L_1 z$ is a linearization of $P(z)$ if there exist unimodular matrix polynomials $U(z), V(z)$ such that $U(z)L(z)V(z) = I \oplus P(z)$; if, moreover, $\mathrm{Rev}_1 L(z)$ is a linearization for $\mathrm{Rev}_k P(z)$, then $L(z)$ is said to be a strong linearization for $P(z)$ considered as a polynomial of grade k [19]. Strong linearizations preserve all the finite and infinite eigenvalues, together with their partial multiplicities, as well as the dimensions of the left and right kernels. However, they typically do not preserve minimal indices [11], though the minimal indices of the most standard classes of strong linearizations available in the literature allow us to obtain very easily those of the matrix polynomial [8, 12].

In this paper, we will study a family of pencils associated with a matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ that are not linearizations according to the definition above (in particular, they do not preserve the dimensions of the left and right kernels), but that still preserve all the finite and infinite eigenvalues of $P(z)$, together with their partial multiplicities, and moreover allow for the recovery of the minimal indices of $P(z)$ in a way which is (at least) as simple as for the most standard types of linearizations. This family of pencils is described in the rest of this subsection. We emphasize that in this description the considered matrix polynomial is arbitrary, i.e., it may be rectangular, square and singular, or square and regular. This generality is in contrast with previous references [29, 34] that deal with this family but restrict their attention to square matrix polynomials that, moreover, are often also required to be regular. It is easy to check that those properties of the pencils in this family that are needed in this work hold for general matrix polynomials, which motivates us to consider in this work fully general matrix polynomials.

The so-called $\mathbb{DL}(P)$ vector space of pencils associated with a matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ was first defined by Mackey, Mackey, Mehl and Mehrmann in [29]. For a fixed matrix polynomial $P(z)$ of grade k and size $m \times n$, each pencil in $\mathbb{DL}(P)$ has size $km \times kn$ and is uniquely determined by a nonzero ansatz vector ω in $\mathbb{F}^k$. In turn, in this paper we identify nonzero ansatz vectors with nonzero scalar ansatz polynomials[3] of grade $k - 1$ via the bijection

$$\omega \in \mathbb{F}^k \mapsto v(x) = \omega^T \begin{bmatrix} x^{k-1} \\ \vdots \\ x \\ 1 \end{bmatrix} \in \mathbb{F}[x]_{k-1}, \tag{2}$$

where $\mathbb{F}[x]_{k-1}$ denotes the vector space of scalar polynomials in the variable x of grade $k - 1$, or, equivalently, of degree at most $k - 1$. If $L(z)$ is the pencil in $\mathbb{DL}(P)$ associated with the ansatz polynomial $v(z)$, we denote it by $L(z) =: \mathbb{DL}(P, v)$, omitting in the right-hand side of this identity the variable z for brevity. The eigenvalue exclusion theorem [29, Theorem 6.7] states that, when $P(z)$ is a regular matrix polynomial, $\mathbb{DL}(P, v)$ is a strong linearization for $P(z)$ if and only if the set of the roots of $v(z)$ (when seen as a grade $k - 1$ scalar polynomial) and the set of the eigenvalues of $P(z)$ (when seen as a grade k matrix polynomial) are disjoint. For a given regular matrix polynomial $P(z)$, this "disjointness" condition is satisfied for almost all ansatz polynomials $v(z)$, which motivates to label $\mathbb{DL}(P)$ as a vector space of "potential linearizations" for $P(z)$. In contrast, it was proved in [8, Theorem 6.1] that if $P(z)$ is square and singular, then none of the pencils in $\mathbb{DL}(P)$ is a linearization for $P(z)$.

In [34], Nakatsukasa, Noferini and Townsend proposed a convenient interpretation of block matrices as bivariate matrix polynomials. Any $k \times k$ block matrix B of size $km \times kn$ with blocks B_{ij} each of size $m \times n$, $1 \leq i, j \leq k$, is associated with the following $m \times n$ bivariate matrix

---

[3]In [29], what we call the ansatz polynomial was instead called the v-polynomial. However, we prefer to follow instead the nomenclature of [34] because it fits better with the Bézoutian description of $\mathbb{DL}(P, v)$, and the latter is crucial to present relatively simple proofs of our results.

polynomial of grade $k-1$ in both variables

$$F(x,y) = \begin{bmatrix} y^{k-1}I_m & \cdots & yI_m & I_m \end{bmatrix} B \begin{bmatrix} x^{k-1}I_n \\ \vdots \\ xI_n \\ I_n \end{bmatrix} = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} y^i x^j B_{k-i,k-j}, \qquad (3)$$

where I_m and I_n denote the identity matrices of sizes $m \times m$ and $n \times n$, respectively. This map is manifestly a bijection between $m \times n$ bivariate matrix polynomials of grade $k-1$ in both x and y and $km \times kn$ matrices partitioned into $k \times k$ blocks all of size $m \times n$. This framework has the advantage of greatly simplifying many operations that are associated with the pencils in the $\mathbb{DL}(P)$ vector space: see, for example, [34, Tables 1 and 2]. In particular, it was shown in [34] that the pencil $\mathbb{DL}(P,v)$ is mapped by the bijection above to the Lerer-Tismenetsky Bézoutian matrix [1, 25, 26] associated with $P(x)$ and $(x-z)v(x)I$. This statement will be very important in this paper and deserves a detailed explanation in terms of bivariate polynomials and block matrices. Note that given $P(z) \in \mathbb{F}[z]^{m \times n}$ of grade k, the pencil $\mathbb{DL}(P,v) =: L_1 z + L_0 \in \mathbb{F}[z]^{km \times kn}$ can be seen as a $k \times k$ block pencil or, equivalently, as a block matrix whose blocks belong to $\mathbb{F}[z]^{m \times n}$ and have grade 1. Hence, we can view it via the bijection (3) as a trivariate matrix polynomial in x,y,z having grade 1 in z and grade $k-1$ in both x and y. We can express concisely this Lerer-Tismenetsky-Bézout connection as follows: functionally, the bijection (3) maps

$$\mathbb{DL}(P,v) \in \mathbb{F}[z]^{km \times kn} \mapsto \mathcal{B}_{P,v}(x,y,z) := \frac{P(y)(x-z)v(x) - P(x)(y-z)v(y)}{x-y} \in \mathbb{F}[x,y,z]^{m \times n}, \qquad (4)$$

where we observe that $\mathcal{B}_{P,v}(x,y,z) = \mathcal{B}_{P,v}(y,x,z)$. The relation (4) was crucial in [34] to enormously simplify the original proof of the eigenvalue exclusion theorem when $P(z)$ is regular, and to derive some novel results. As said in the introduction (see Theorem 1.1), in this paper, we essentially aim to extend the eigenvalue exclusion theorem to the case where $P(z)$ is singular (square or rectangular). This extension means that if the set of the roots of $v(z)$ (when seen as a grade $k-1$ scalar polynomial) and the set of the eigenvalues of $P(z)$ (when seen as a grade k matrix polynomial) are disjoint, then we will prove that $\mathbb{DL}(P,v)$ has the same eigenvalues and partial multiplicities as $P(z)$, while the minimal indices of $\mathbb{DL}(P,v)$ are precisely those of $P(z)$ together with some extra minimal indices equal to zero. Therefore, it is not surprising that the bijection (4) plays also a central role in this paper. In this context, we emphasize that (4) was derived in [34] without any need to assume that $P(z)$ is regular [34, Remark 3.2], though in many of the results that followed next in [34] it was assumed that $P(z)$ is regular.

Remark 2.1. In principle, the bijection (3) could be defined using *any* polynomial basis for the space of scalar polynomials of degree at most $k-1$, i.e., the restriction to the monomial basis is convenient, but not necessary. This was indeed already observed in [34] and rediscovered in [15]. It leads to an immediate generalization of the space $\mathbb{DL}(P)$. While we restrict to the monomial basis for the sake of a much simpler exposition, we observe that the results of this paper also generalize to $\mathbb{DL}(P)$ in non-monomial bases. This follows easily by observing that (a) a bijection between such a space and the "traditional" $\mathbb{DL}(P)$ in the monomial basis is given by multiplying each pencil in the generalized $\mathbb{DL}(P)$ by $M^T \otimes I_m$ on the left and by $M \otimes I_n$ on the right, where M is the $k \times k$ nonsingular constant change-of-polynomial-basis matrix, and that (b) all the relevant structures, i.e., eigenvalues, partial multiplicities, minimal indices and bases, root polynomials, are well behaved under strict equivalence. This remark could have practical relevance, because $\mathbb{DL}(P)$ in certain non-monomial bases may be useful to exploit the better numerical properties of some basis. One example, albeit limited to the scalar case $m = n = 1$, is

presented in [33] where a method based on the Bézout resultant (i.e., $\mathbb{DL}(P)$) in the Chebyshev basis was used to compute real common zeros of bivariate functions.

2.3 Root polynomials

Maximal sets of root polynomials [13, 35] (see also [19] for the regular case and [38] for an extension to rational functions) are another crucial tool in this paper. They are useful for the analysis of singular matrix polynomials because they carry the relevant information about the partial multiplicities of a matrix polynomial at a given eigenvalue λ. More precisely, suppose that $M(z) \in \mathbb{F}[z]^{n \times p}$ is a right minimal basis for the matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$, which is a short way of stating that the columns of $M(z)$ form a minimal basis of the right kernel of $P(z)$. Then, for any $\lambda \in \mathbb{F}$ we define $\ker_\lambda P(z)$ as the subspace of $\mathbb{F}^n$ spanned by the columns of $M(\lambda)$. A polynomial vector $r(z) \in \mathbb{F}[z]^n$ is said to be a *root polynomial* at λ of order $\ell \geq 1$ for $P(z)$ if

1. $P(z)\, r(z) = (z - \lambda)^\ell w(z)$, for some $w(z) \in \mathbb{F}[z]^m$ such that $w(\lambda) \neq 0$, and

2. $r(\lambda) \notin \ker_\lambda P(z)$.

Suppose that $\{r_i(z)\}_{i=1}^t$ is a set of root polynomials at λ for $P(z)$, with orders $\ell_1 \geq \cdots \geq \ell_t$, and let $R(z) \in \mathbb{F}[z]^{n \times t}$ be the matrix whose ith column is $r_i(z)$. The set $\{r_i(z)\}_{i=1}^t$ is said to be λ-independent if the matrix $\begin{bmatrix} M(\lambda) & R(\lambda) \end{bmatrix}$ has full column rank; it is said to be complete if it is λ-independent and $t = \dim \ker P(\lambda) - \dim \ker_\lambda P(z)$; and it is said to be maximal if it is complete and satisfies the property that, for each $i = 1, \ldots, t$, there is no root polynomial $\hat{r}(z)$ such that $r_1(z), \ldots, r_{i-1}(z), \hat{r}(z)$ are λ-independent and the order of $\hat{r}(z)$ is strictly larger than the order of $r_i(z)$. It was proved in [13, 38] that all the concepts defined above are independent of the particular right minimal basis $M(z)$ of $P(z)$ that is considered and that root polynomials at λ for $P(z)$ exist if and only if λ is a finite eigenvalue of $P(z)$.

All maximal sets of root polynomials of a given matrix polynomial $P(z)$ at a given point λ have the same orders [13, Theorem 4.1-Item 2] and these orders coincide with the partial multiplicities of λ as an eigenvalue of $P(z)$ [13, Theorem 4.2]. Moreover, the following theorem that appeared in [13, Theorem 4.1-Item 3] will be important in this paper.

Theorem 2.2. *Let $\{r_i(z)\}_{i=1}^t$ be a complete set of root polynomials at $\lambda \in \mathbb{F}$ for $P(z) \in \mathbb{F}[z]^{m \times n}$ with orders $\kappa_1 \geq \cdots \geq \kappa_t$. Suppose, moreover, that the partial multiplicities of λ as an eigenvalue of $P(z)$ are $\ell_1 \geq \cdots \geq \ell_t$. Then, $\ell_i \geq \kappa_i$ for all $i = 1, \ldots, t$. Moreover, the following are equivalent:*

1. *$\{r_i(z)\}_{i=1}^t$ is a maximal set of root polynomials at λ for $P(z)$;*

2. *$\ell_i = \kappa_i$ for all $i = 1, \ldots, t$;*

3. *$\sum_{i=1}^t \kappa_i = \sum_{i=1}^t \ell_i$.*

The definitions above of root polynomials and λ-independent, complete and maximal sets of root polynomials correspond to the right versions of these concepts. Left root polynomials and λ-independent, complete and maximal sets of left root polynomials are defined as the right ones for $P(z)^T$. The definition of root polynomials at infinity for $P(z)$ requires certain care, but their properties are analogous to those of root polynomials at finite points. They can be found in [13, Section 6].

3 Preliminary results

This section includes some auxiliary results that are useful to prove the main results in this paper. The lemmata in Subsection 3.1 are essential to study in Section 4 the minimal indices of the pencils in $\mathbb{DL}(P)$. The results in Subsection 3.2 allow us to simplify many of the proofs in the rest of the paper, because they show, via a Möbius transformation, that it is enough to establish them only for matrix polynomials $P(z)$ and ansatz polynomials $v(z)$ in (4) without eigenvalues and roots at infinity, respectively.

3.1 Technical lemmata on certain Vandermonde-structured matrices

In this subsection we provide some results on matrices that have a tensor product Vandermonde structure; these results are crucial for the construction in Section 4 of a minimal basis of $\ker L(z)$, where $L(z) = \mathbb{DL}(P, v)$, for a singular matrix polynomial $P(z)$ and an ansatz polynomial $v(z)$ whose roots are not eigenvalues of $P(z)$.

Let us start with Lemma 3.1, that assumes a tensor product structure with a classical (i.e. not confluent) Vandermonde matrix.

Lemma 3.1. *Let $\ell \leq k$ be two positive integers. For $i = 1, \ldots, \ell$, let $A_i \in \mathbb{F}^{n \times m_i}$ with rank $A_i = m_i$ and let $\alpha_i \in \mathbb{F}$ be such that $\alpha_i \neq \alpha_j$ if $i \neq j$. Let*

$$V(z) = \begin{bmatrix} z^{k-1} \\ \vdots \\ z \\ 1 \end{bmatrix}$$

be the Vandermonde vector of degree $k - 1$ depending on the variable $z \in \mathbb{F}$. Then the matrix $C = \begin{bmatrix} V(\alpha_1) \otimes A_1 & \cdots & V(\alpha_\ell) \otimes A_\ell \end{bmatrix} \in \mathbb{F}^{kn \times m}$ has full column rank $m = \sum_{i=1}^{\ell} m_i$.

Proof. Let $S = \begin{bmatrix} V(\alpha_1) & \cdots & V(\alpha_\ell) \end{bmatrix}$, which can be seen as a $k \times \ell$ submatrix of a $k \times k$ Vandermonde matrix. Then, standard properties of Vandermonde matrices imply that S has full column rank ℓ. Since we can factorize $C = (S \otimes I_n) \operatorname{diag}(A_1, \ldots, A_\ell)$, we have rank $C = \operatorname{rank} \operatorname{diag}(A_1, \ldots, A_\ell) = m$. $\qquad \square$

We have stated Lemma 3.1 mainly for illustrative purposes, as it is simpler to analyze than the general case, but to our goals in Section 4 it is only useful if all the roots of the ansatz polynomial $v(z)$ are simple. This is in fact not typical in practice since the most common choices for $v(z)$ are either 1 or z^{k-1} [14, 22, 29], that have a single root of multiplicity $k - 1$ at ∞ and 0, respectively. However, by looking at the proof, we observe that Lemma 3.1 would still hold if we replaced S by any full column rank matrix, not necessarily Vandermonde. Our next goal is to generalize in Lemma 3.3 the result to a tensor product structure inherited by a *confluent* Vandermonde matrix (see [20, Chapter 22] for information on confluent Vandermonde matrices). Lemma 3.2 is a preliminary step for this goal.

Lemma 3.2. *Let $\ell \leq k$ be two positive integers, let $\alpha \in \mathbb{F}$, let $M(z) \in \mathbb{F}[z]^{n \times p}$ be a matrix polynomial such that $M(\alpha)$ has full column rank, and let $V(z) \in \mathbb{F}[z]^{k \times 1}$ be defined as in Lemma 3.1. Let us denote by $[V(z) \otimes M(z)]'$ and $[V(z) \otimes M(z)]^{(\beta)}$ the derivative and the βth derivative with respect to z, respectively, and by $[\,\cdot\,]|_{z=\alpha}$ evaluation at α. Then the matrix*

$$M_{\alpha,k,\ell,M(z)} = \begin{bmatrix} V(z) \otimes M(z) & [V(z) \otimes M(z)]' & \cdots & [V(z) \otimes M(z)]^{(\ell-1)} \end{bmatrix}\Big|_{z=\alpha} \in \mathbb{F}^{kn \times \ell p}$$

has full column rank.

Proof. The block entry of $M_{\alpha,k,\ell,M(z)}$ (before evaluation at α) in block row $k - \phi$ and block column $\beta + 1$ is

$$[z^\phi M(z)]^{(\beta)} = \sum_{j=0}^{\beta} [z^\phi]^{(j)} \binom{\beta}{j} [M(z)]^{(\beta-j)}.$$

Since the jth derivative of z^ϕ is 0 for $j > \phi$ and in the summation only the first β derivatives of $M(z)$ appear, we can encode these relations by expressing $M_{\alpha,k,\ell,M(z)}$ as the product $M_{\alpha,k,\ell,M(z)} = AB$, where

$$A = \begin{bmatrix} V^{(\ell-1)}(\alpha) & \ldots & V'(\alpha) & V(\alpha) \end{bmatrix} \otimes I_n \in \mathbb{F}^{kn \times \ell n} \tag{5}$$

corresponds to the last columns of a block upper triangular matrix, and

$$B = \begin{bmatrix} 0 & 0 & & \ldots & M(\alpha) \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & M(\alpha) & \ldots & \binom{\ell-1}{2} M^{(\ell-3)}(\alpha) \\ 0 & M(\alpha) & 2M'(\alpha) & \ldots & (\ell-1)M^{(\ell-2)}(\alpha) \\ M(\alpha) & M'(\alpha) & M''(\alpha) & \ldots & M^{(\ell-1)}(\alpha) \end{bmatrix} \in \mathbb{F}^{\ell n \times \ell p}, \tag{6}$$

is block lower antitriangular. Observe that, generally, the block B_{ij} is $\binom{j-1}{\ell-i} M^{(i+j-\ell-1)}(\alpha)$ if $i + j > \ell$ and 0 otherwise. Note that A is a $k \times \ell$ submatrix of a $k \times k$ confluent Vandermonde matrix tensor product the identity, and hence, it has full column rank. As a consequence, $\operatorname{rank} M_{\alpha,k,\ell,M(z)} = \operatorname{rank} B = \ell \operatorname{rank} M(\alpha) = \ell p$. $\qquad\square$

Lemma 3.2 implies the announced generalization of Lemma 3.1.

Lemma 3.3. *Let $\ell_1, \ldots, \ell_s, k$ be positive integers such that $\ell := \sum_{i=1}^{s} \ell_i \leq k$ and let $V(z) \in \mathbb{F}[z]^{k \times 1}$ be defined as in Lemma 3.1. For $i = 1, \ldots, s$, let $\alpha_i \in \mathbb{F}$ be such that $\alpha_i \neq \alpha_j$ if $i \neq j$, let $M_i(z) \in \mathbb{F}[z]^{n \times p_i}$ be a matrix polynomial such that $M_i(\alpha_i)$ has full column rank, and let $M_{\alpha_i,k,\ell_i,M_i(z)} \in \mathbb{F}^{kn \times \ell_i p_i}$ be defined as in Lemma 3.2. Then*

$$C = \begin{bmatrix} M_{\alpha_1,k,\ell_1,M_1(z)} & M_{\alpha_2,k,\ell_2,M_2(z)} & \cdots & M_{\alpha_s,k,\ell_s,M_s(z)} \end{bmatrix} \in \mathbb{F}^{kn \times (\sum_{i=1}^{s} \ell_i p_i)}$$

has full column rank.

Before proving Lemma 3.3, we observe that ℓ_i and $M_i(z)$ are allowed to differ for different values of i and that the matrices $M_i(z)$ may also have different number of columns (but they must have the same number of rows). However, it is a necessary assumption that $M_i(\alpha_i)$ has full column rank for all i. For our scopes in this paper, however, $M_i(z)$ will not depend on the index i (whereas ℓ_i may). Observe also that $\sum_{i=1}^{s} \ell_i p_i \leq (\sum_{i=1}^{s} \ell_i) \max_i p_i \leq kn$. So, the number of columns of C is less than or equal to the number of its rows.

Proof of Lemma 3.3. For each i, let $M_{\alpha_i,k,\ell_i,M_i(z)} = A_i B_i$ be a factorization as in the proof of Lemma 3.2. Then, $C = AB$ with

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_s \end{bmatrix} \in \mathbb{F}^{kn \times \ell n}, \qquad B = B_1 \oplus B_2 \oplus \cdots \oplus B_s \in \mathbb{F}^{\ell n \times (\sum_{i=1}^{s} \ell_i p_i)},$$

where A_i, B_i have the structures of (5) and (6), respectively. Then, A is a $k \times \ell$ submatrix of a $k \times k$, possibly confluent, Vandermonde matrix tensor product the identity. Therefore, it has full colum rank. Hence, $\operatorname{rank} C = \operatorname{rank} B = \sum_{i=1}^{s} \operatorname{rank} B_i = \sum_{i=1}^{s} \ell_i p_i$, proving the statement. $\qquad\square$

3.2 A commuting diagram between $\mathbb{DL}(P)$ pencils and Möbius transformation

Let gcd stand for greatest common divisor. Recall that, given a matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ of degree k and a Möbius function $r(z) = n(z)/d(z)$, where $n(z) = \alpha z + \beta$, $d(z) = \gamma z + \delta$ satisfy $\gcd(n(z), d(z)) = 1$ (or equivalently $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ is invertible), the associated Möbius transformation of grade $g \geq k$ of $P(z)$ is

$$\mathcal{M}_{g,r(z)}(P) = d(z)^g P\left(\frac{n(z)}{d(z)}\right) \in \mathbb{F}[z]^{m \times n},$$

where we omit the dependence on z in the left-hand side for simplicity. These transformations have been studied in papers such as [31, 35]. Crucially, the map $z \mapsto n(z)/d(z)$ is an automorphism of $\mathbb{F} \cup \{\infty\}$ (upon agreeing that, conventionally, for all nonzero $\alpha \in \mathbb{F}$, $\alpha/0 = \infty$; moreover $n(\infty)/d(\infty) = \alpha/\gamma$. In particular, the following results are known (see, for example, [31, Theorems 5.3 and 7.5] and [35, Theorem 4.1]).

Theorem 3.4. *Suppose that $P(z) \in \mathbb{F}[z]^{m \times n}$ has grade g, that $Q(z) = \mathcal{M}_{g,r(z)}(P) \in \mathbb{F}[z]^{m \times n}$ is the Möbius transform of grade g induced by $r(z) = n(z)/d(z)$ of $P(z)$, and assume that $Q(z)$ has also grade g. Let $\lambda, \mu \in \mathbb{F} \cup \{\infty\}$ satisfy $\lambda d(\mu) = n(\mu)$. Then:*

1. *λ is an eigenvalue of $P(z)$ with nonzero partial multiplicities $\delta_1 \leq \cdots \leq \delta_s$ if and only if μ is an eigenvalue of $Q(z)$ with nonzero partial multiplicities $\delta_1 \leq \cdots \leq \delta_s$;*

2. *$m_1(z), \ldots, m_p(z) \in \mathbb{F}[z]^{n \times 1}$ are a right minimal basis of $P(z)$ with right minimal indices $\nu_1 \leq \cdots \leq \nu_p$ if and only if $u_1(z), \ldots, u_p(z) \in \mathbb{F}[z]^{n \times 1}$ are a right minimal basis of $Q(z)$ with right minimal indices $\nu_1 \leq \cdots \leq \nu_p$, where $u_i(z) = \mathcal{M}_{\nu_i, r(z)}(m_i)$ for all $i = 1, \ldots, p$.*

3. *$n_1(z), \ldots, n_q(z) \in \mathbb{F}[z]^{m \times 1}$ are a left minimal basis of $P(z)$ with left minimal indices $\eta_1 \leq \cdots \leq \eta_q$ if and only if $v_1(z), \ldots, v_q(z) \in \mathbb{F}[z]^{m \times 1}$ are a left minimal basis of $Q(z)$ with left minimal indices $\eta_1 \leq \cdots \leq \eta_q$, where $v_i(z) = \mathcal{M}_{\eta_i, r(z)}(n_i)$ for all $i = 1, \ldots, q$.*

With this in mind, we show in Theorem 3.6 how the operations of constructing a $\mathbb{DL}(P)$ pencil of a matrix polynomial and applying Möbius transformations relate. We prove the simple Lemma 3.5 as a preliminary step.

Lemma 3.5. *Let $n(z) = \alpha z + \beta, d(z) = \gamma z + \delta \in \mathbb{F}[z]$ be scalar polynomials of grade 1 such that $\gcd(n(z), d(z)) = 1$ and let $\mathbb{F}[z]_{k-1}$ be the vector space of scalar polynomials of grade $k-1$. Then, $\mathcal{B}_{n,d} = \{n(z)^{k-1}, n(z)^{k-2}d(z), \ldots, n(z)d(z)^{k-2}, d(z)^{k-1}\}$ is a basis of $\mathbb{F}[z]_{k-1}$. Therefore, there exists an invertible change of basis matrix $B \in \mathbb{F}^{k \times k}$ such that*

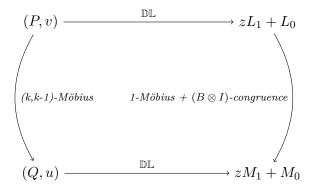$$B \begin{bmatrix} z^{k-1} \\ z^{k-2} \\ \vdots \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} n(z)^{k-1} \\ n(z)^{k-2}d(z) \\ \vdots \\ n(z)d(z)^{k-2} \\ d(z)^{k-1} \end{bmatrix}. \tag{7}$$

Proof. Let $r(z) = n(z)/d(z)$. The statement is a corollary of [31, Theorem 3.18], that says that $p(z) \mapsto \mathcal{M}_{k-1, r(z)}(p)$ is an $\mathbb{F}$-linear automorphism of $\mathbb{F}[z]_{k-1}$. $\square$

Theorem 3.6. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade k and $v(z)$ be an ansatz polynomial of grade $k - 1$. Let $n(z), d(z)$ be coprime linear scalar polynomials and denote by $\mathcal{M}_{g,r(z)}(X)$ the Möbius transform of grade g of the matrix polynomial $X(z)$ induced by $r(z) = n(z)/d(z)$, that is, $\mathcal{M}_{g,r(z)}(X) = d(z)^g X(n(z)/d(z))$. Denote by $L(z) = zL_1 + L_0$ the $\mathbb{DL}(P)$ pencil of $P(z)$ with ansatz polynomial $v(z)$, i.e., $\mathbb{DL}(P,v)$, and by $M(z) = zM_1 + M_0$ the $\mathbb{DL}(Q)$ pencil of $Q(z) := \mathcal{M}_{k,r(z)}(P)$ with ansatz polynomial $u(z) := \mathcal{M}_{k-1,r(z)}(v)$, i.e., $\mathbb{DL}(Q,u)$. Moreover, let B be the $k \times k$ change of basis matrix in Lemma 3.5. Then,*

$$\left(B^T \otimes I_m\right) \mathcal{M}_{1,r(z)}(L) \left(B \otimes I_n\right) = M(z).$$

In other words, the following diagram commutes:

Proof. It is convenient, for the sake of clarity in the proof, to distinguish notationally between the variables associated with the triple (P, v, L) and those associated with (Q, u, M): to this goal, we will put hats on the variables associated with P, v and L. Under the bijection (3) and taking into account (4),

$$L(\hat{z}) \mapsto \frac{P(\hat{y})(\hat{x} - \hat{z})v(\hat{x}) - P(\hat{x})(\hat{y} - \hat{z})v(\hat{y})}{\hat{x} - \hat{y}}.$$

Taking into account (3), (7), and the fact that $\mathcal{M}_{k-1,r(z)}(\sum_{i=0}^{k-1} w_i z^i) = \sum_{i=0}^{k-1} w_i \, n(z)^i \, d(z)^{k-1-i}$, we see that left multiplication of $L(\hat{z})$ by $B^T \otimes I_m$ corresponds functionally to a Möbius transformation of grade $k - 1$ induced by $\hat{y} = r(y)$; similarly, right multiplication of $L(\hat{z})$ by $B \otimes I_n$ corresponds to a Möbius transformation of grade $k - 1$ induced by $\hat{x} = r(x)$; and finally the action of $\mathcal{M}_{1,r(z)}$ implements at the functional level a Möbius transformation of grade 1 induced by $\hat{z} = r(z)$. Hence, using (4), the pencil $(B^T \otimes I)\mathcal{M}_{1,r(z)}(L)(B \otimes I)$ corresponds functionally to

$$d(z)d(x)^{k-1}d(y)^{k-1}\frac{P(r(y))(r(x) - r(z))v(r(x)) - P(r(x))(r(y) - r(z))v(r(y))}{r(x) - r(y)}. \tag{8}$$

On the other hand, once again using (4),

$$M(z) \mapsto \frac{Q(y)(x - z)u(x) - Q(x)(y - z)u(y)}{x - y},$$

or, equivalently after using the definitions of Q and u, $M(z)$ is mapped by (3) to the trivariate matrix polynomial

$$\frac{d(y)^k d(x)^{k-1} P(r(y))(x - z)v(r(x)) - d(x)^k d(y)^{k-1} P(r(x))(y - z)v(r(y))}{x - y}. \tag{9}$$

Bearing in mind that (3) is a bijection, proving the statement is therefore equivalent to showing that (8) and (9) are equal as trivariate polynomial functions in x, y, z. Dividing (8) and (9) by $d(y)^{k-1}d(x)^{k-1}$, the statement is then implied by

$$\frac{P(r(y))(d(z)r(x) - n(z))v(r(x)) - P(r(x))(d(z)r(y) - n(z))v(r(y))}{r(x) - r(y)} =$$
$$= \frac{d(y)P(r(y))(x - z)v(r(x)) - d(x)P(r(x))(y - z)v(r(y))}{x - y}.$$

Breaking each side as the difference of two similar terms, equating each pair of terms, and cancelling common factors, we see that in turn the latter equality is implied by

$$\begin{cases} (x - y)(d(z)r(x) - n(z)) = (r(x) - r(y))d(y)(x - z); \\ (x - y)(d(z)r(y) - n(z)) = (r(x) - r(y))d(x)(y - z). \end{cases} \tag{10}$$

We prove the first functional equation of the system (10), as the second one follows from the first one with the changes $x \mapsto y$ and $y \mapsto x$ and changing the signs of both sides. Multiplying both sides of the first equation in (10) by $d(x)$ we obtain

$$(x - y)(d(z)n(x) - n(z)d(x)) = (n(x)d(y) - n(y)d(x))(x - z),$$

whose correctness can be checked by using $n(s) = \alpha s + \beta, d(s) = \gamma s + \delta$, and then expanding both sides to see that both are equal to $(x - y)(x - z)(\delta\alpha - \beta\gamma)$. $\qquad\square$

A consequence of Theorem 3.6 is that we can assume in the rest of the paper without loss of generality that both the matrix polynomial $P(z)$ and the ansatz polynomial $v(z)$ do not have eigenvalues and roots at infinity, which simplifies considerably the developments. Indeed, if infinity is an eigenvalue of $P(z)$ or a root of $v(z)$, then we can apply a Möbius transformation induced by an appropriate $r(z)$ to both so that neither $Q(z) = \mathcal{M}_{k,r(z)}(P)$ nor $u(z) = \mathcal{M}_{k-1,r(z)}(v)$ has infinite eigenvalues/roots. Note that it is always possible to find such $r(z)$ because the set of eigenvalues/roots of P and v has a finite number of elements. Taking into account Theorem 3.4 and the fact that the strict equivalences $B^T \otimes I_m$ and $B \otimes I_n$ preserve the eigenvalues, the partial multiplicities and the minimal indices, Theorem 3.6 then allows us to obtain statements on the triple $(P, v, L = \mathbb{DL}(P, v))$ from those on the triple $(Q, u, M = \mathbb{DL}(Q, u))$.

4   The minimal indices of $\mathbb{DL}(P, v)$

In this section we will prove items 1 and 2 of Theorem 1.1. We will prove first item 1 and then obtain item 2 as a simple corollary. The proof of item 1 proceeds by building a particular right minimal basis of the pencil $\mathbb{DL}(P, v)$ in (4). Throughout this section, we suppose that the, possibly singular, nonzero matrix polynomial $P(z) \in \mathbb{F}[z]^{m \times n}$ has grade $k \geq 2$, that the scalar ansatz polynomial $v(z)$ has grade $k - 1$, and that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint. We denote $\mathbb{DL}(P, v) =: zL_1 + L_0 =: L(z)$. Moreover, having in mind to exploit Theorem 3.6, we further suppose that $v(z)$ has no infinite roots, which implies that $v(z)$ has degree exactly $k - 1$.

We start by reviewing some basic concepts on minimal bases that are used in this section and that can be found at [16, 28]. Recall that we are working in an algebraically closed field $\mathbb{F}$. For brevity, we say that a polynomial matrix $A(z) \in \mathbb{F}[z]^{n \times p}$, with $n \geq p$, is a minimal basis if its columns form a minimal basis of the rational subspace they span, which is denoted by span $A(z)$. Given a polynomial matrix $P(z) \in \mathbb{F}[z]^{m \times n}$, we say that $A(z)$ is a (right) minimal

basis of $P(z)$ if $A(z)$ is a minimal basis and span $A(z) = \ker P(z)$. Let us write $A(z)$ in terms of its columns as $A(z) = \begin{bmatrix} A_1(z) & \cdots & A_p(z) \end{bmatrix}$. Then, we define the column-wise reversal of $A(z)$ as $(\mathrm{cwRev}\, A)(z) = \begin{bmatrix} \mathrm{Rev}\, A_1(z) & \cdots & \mathrm{Rev}\, A_p(z) \end{bmatrix}$, where we emphasize that each reversal is taken with respect to the degree of the corresponding column, and by convention $\deg 0 = -\infty$. We denote $\{A(z)\} = (\mathrm{cwRev}\, A)(0)$. This constant matrix is termed in [16] the high order coefficient matrix of $A(z)$. The rules in the following lemma follow immediately from the definition of the high order coefficient matrix.

Lemma 4.1. *Let $A(z) \in \mathbb{F}[z]^{n \times p}$, $B(z) \in \mathbb{F}[z]^{n \times t}$ and $C \in \mathbb{F}^{n \times s}$, i.e., C is a constant matrix. Then:*

1. *$\left\{ \begin{bmatrix} A(z) & B(z) \end{bmatrix} \right\} = \begin{bmatrix} \{A(z)\} & \{B(z)\} \end{bmatrix}$;*

2. *$\{C\} = C$;*

3. *for any scalar polynomial $p(z) = p_d z^d + \cdots + p_1 z + p_0$, with $p_d \neq 0$,*

   (a) *$\{p(z)A(z)\} = p_d \{A(z)\}$;*

   (b) *if $\deg p(z) > \deg B(z)$ and no column of $A(z)$ is zero, then $\{p(z)A(z) + B(z)\} = p_d \{A(z)\}$.*

Let us recall a very useful characterization of minimal bases that was first proved in [16].

Lemma 4.2. *Let $A(z) \in \mathbb{F}[z]^{n \times p}$, $n \geq p$. Then $A(z)$ is a minimal basis if and only if*

1. *$\mathrm{rank}\, A(z_0) = p$ for all $z_0 \in \mathbb{F}$ and*

2. *$\mathrm{rank}\{A(z)\} = p$.*

The following simple lemma will be also used in this section.

Lemma 4.3. *Let A, B, C be matrices of adequate sizes with entries in any field. Then $\mathrm{rank} \begin{bmatrix} 0 & C \\ A & B \end{bmatrix} = \mathrm{rank} \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \geq \mathrm{rank}\, A + \mathrm{rank}\, C$.*

In the following proposition, we give a first step towards constructing a right minimal basis of $\mathbb{DL}(P, v)$. Note also that, as a corollary of Proposition 4.4, we obtain an alternative proof to that in [8] for the fact that none of the pencils in $\mathbb{DL}(P)$ is a linearization of $P(z)$ if $P(z)$ is singular, since the dimensions of the right kernels of $L(z)$ and $P(z)$ are different.

Proposition 4.4. *Let $M(z) \in \mathbb{F}[z]^{n \times p}$ be a right minimal basis for $P(z) \in \mathbb{F}[z]^{m \times n}$ and suppose the distinct roots of $v(z)$ are $\mu_1, \ldots, \mu_s \in \mathbb{F}$ with multiplicities $\ell_1, \ldots, \ell_s$, respectively. Let $L(z) = \mathbb{DL}(P, v)$. Using the notation of Lemmata 3.1, 3.2 and 3.3, and taking into account that $\sum_{i=1}^{s} \ell_i = k - 1$, define*

$$C = \begin{bmatrix} M_{\mu_1, k, \ell_1, M(z)} & M_{\mu_2, k, \ell_2, M(z)} & \cdots & M_{\mu_s, k, \ell_s, M(z)} \end{bmatrix} \in \mathbb{F}^{kn \times (k-1)p}, \quad and$$

$$G(z) = \begin{bmatrix} C & D(z) \end{bmatrix} \in \mathbb{F}[z]^{kn \times kp},$$

where $D(z) := V(z) \otimes M(z) \in \mathbb{F}[z]^{kn \times p}$. Then,

1. *C has full column rank and satisfies $L(z)\, C = 0$, that is $\mathrm{span}\, C \subseteq \ker L(z)$. Moreover, $L(z)$ has at least $(k-1)p$ right minimal indices equal to 0, where $p = \dim \ker P(z)$;*

14

2. $G(z)$ has full column rank and satisfies $L(z)\,G(z) = 0$, that is $\operatorname{span} G(z) \subseteq \ker L(z)$.

Proof. Item 1. Taking into account Lemmata 3.2 and 3.3, we consider C partitioned as $C = \begin{bmatrix} C_1 & C_2 & \cdots & C_{k-1} \end{bmatrix}$, where each C_i has a structure of the type $C_i = [V(z) \otimes M(z)]^{(a)}\big|_{z=\mu_j} \in \mathbb{F}^{kn \times p}$ for some root μ_j of $v(z)$ and some a such that $0 \le a \le \ell_j - 1$. Then $L(z)C = \begin{bmatrix} L(z)C_1 & \cdots & L(z)C_{k-1} \end{bmatrix}$. Since, according to (4), $L(z) \mapsto \mathcal{B}_{P,v}(x, y, z)$ via the bivariate matrix polynomial bijection (3), right multiplication of $L(z)$ by the block column C_i corresponds in the functional bivariate formulation to right multiplication of $\mathcal{B}_{P,v}(x, y, z)$ by $M(x)$, followed by taking the a-th partial derivative with respect to x and x-evaluation at $x = \mu_j$ (see [34, Table 1]). Explicitly, this means

$$L(z)C_i \mapsto \frac{\partial^a}{\partial x^a} \frac{P(y)(x - z)v(x) - (y - z)v(y)P(x)}{x - y} M(x)\bigg|_{x=\mu_j}.$$

Since $P(x)M(x) = 0$, we only need to pay attention to the derivative of the first term and, taking into account that μ_j is a root of $v(x)$ of multiplicity $\ell_j > a$ (crucially, the inequality is strict), we then see that $L(z)C_i \mapsto 0$ for all $i = 1, \ldots k - 1$, and hence $L(z)C = 0$ because (3) is a bijection[4].

To prove that C has full column rank, it suffices to use Lemma 4.2-1 on $M(z)$ and then to invoke (a special case of) Lemma 3.3. Finally, since the columns of C are linearly independent constant vectors and belong to $\ker L(z)$, they can be completed to a minimal basis of $\ker L(z)$ via the method described in [17, Ch. XII, §5]. Thus, the $(k-1)p$ columns of the constant matrix C give $(k-1)p$ minimal indices of $L(z)$ equal to zero. Another way to get the same result is through the so-called "Strong Minimality Property of Minimal Indices" proved in [28, Theorem 4.2].

Item 2. Observe that the operation $L(z)D(z)$ in the bivariate matrix polynomial framework amounts to multiply $\mathcal{B}_{P,v}(x, y, z)$ by $M(x)$ on the right followed by x-evaluation at $x = z$. Explicitly, this means

$$L(z)D(z) \mapsto \frac{P(y)(z - z)v(z) - (y - z)v(y)P(z)}{z - y} M(z) = 0,$$

where we have used again that $P(z)M(z) = 0$. Therefore, $L(z)D(z) = 0$, which combined with the result in item 1 implies $L(z)G(z) = 0$. To see that $G(z)$ has full column rank, we apply Lemma 3.3 to $G(\alpha)$, where α is any element of $\mathbb{F}$ such that $\alpha \ne \mu_i$, for $i = 1, \ldots, s$. This proves that the constant matrix $G(\alpha)$ has full column rank, which implies that the matrix polynomial $G(z)$ has full column rank. $\qquad\square$

Proposition 4.4 implies that $\dim \ker L(z) \ge kp$. In fact, we will prove later that $\dim \ker L(z) = kp$. This, and the fact that $G(z)$ is not a minimal basis, motivates us to look for a minimal basis $F(z) \in \mathbb{F}[z]^{kn \times kp}$ such that $\operatorname{span} F(z) = \operatorname{span} G(z)$. Observe that $G(z)$ is not a minimal basis because it does not satisfy the first condition of Lemma 4.2 because $D(\mu_i)$ is equal to the first block column of $M_{\mu_i, k, \ell_i, M(z)}$ for $i = 1, \ldots, s$. The minimal basis $F(z)$ is described in Proposition 4.5 in terms of the Hermite interpolating polynomials (see, for instance, [40, p. 53]) associated with the distinct roots $\mu_1, \ldots, \mu_s$ of $v(z)$ and their multiplicities $\ell_1, \ldots, \ell_s$.

Proposition 4.5. *Let $M(z) \in \mathbb{F}[z]^{n \times p}$ be a right minimal basis for $P(z) \in \mathbb{F}[z]^{m \times n}$ and let $\gamma_1 \le \cdots \le \gamma_p$ be the degrees of the columns of $M(z)$, i.e., $\gamma_1 \le \cdots \le \gamma_p$ are the right minimal*

---

[4]We warn the reader that, as explained in [34, Section 3], the bivariate polynomial bijection (3) can clearly be defined more in general from $k \times h$ block matrices with blocks of size $m \times n$ to $m \times n$ bivariate matrix polynomials of grade $k - 1$ in y and of grade $h - 1$ in x.

indices of $P(z)$. Suppose the distinct roots of $v(z)$ are $\mu_1, \ldots, \mu_s \in \mathbb{F}$ with multiplicities $\ell_1, \ldots, \ell_s$, respectively. Let $L(z) = \mathbb{DL}(P, v)$ and let $C, D(z)$ and $G(z)$ be the matrices defined in Proposition 4.4. Let us partition the matrix $C \in \mathbb{F}^{kn \times (k-1)p}$ as $C = \begin{bmatrix} C_1 & C_2 & \cdots & C_{k-1} \end{bmatrix}$ with $C_i \in \mathbb{F}^{kn \times p}$ for $i = 1, \ldots, k-1$. Let $H_1(z), H_2(z), \ldots, H_{k-1}(z)$ be the scalar Hermite interpolating polynomials associated with $\mu_1, \ldots, \mu_s$ and $\ell_1, \ldots, \ell_s$ ordered as follows: if $C_i = [V(z) \otimes M(z)]^{(a)}\big|_{z=\mu_j} \in \mathbb{F}^{kn \times p}$ for some μ_j and some a such that $0 \le a \le \ell_j - 1$, then

$$H_i^{(b)}(\mu_t) = \begin{cases} 1, & \text{if } t = j \text{ and } b = a, \\ 0, & \text{if } t \ne j \text{ or } b \ne a, \end{cases} \tag{11}$$

where, for each μ_t, the integer b satisfies $0 \le b \le \ell_t - 1$. Define

$$E(z) = D(z) - \sum_{i=1}^{k-1} H_i(z)\, C_i \in \mathbb{F}[z]^{kn \times p} \quad \text{and} \quad F(z) = \begin{bmatrix} C & \frac{1}{v(z)} E(z) \end{bmatrix}.$$

Then,

1. *$F(z)$ has full column rank, satisfies $L(z)\, F(z) = 0$, that is $\operatorname{span} F(z) \subseteq \ker L(z)$, and $\operatorname{span} F(z) = \operatorname{span} G(z)$;*

2. *$F(z) \in \mathbb{F}[z]^{kn \times kp}$, i.e., it is a matrix polynomial, and the degrees of its columns are $\underbrace{0, \ldots, 0}_{(k-1)p}, \gamma_1, \ldots, \gamma_p$; and*

3. *$F(z)$ is a minimal basis.*

Proof. Item 1. Observe that $F(z) = G(z)\, S(z)$, where

$$S(z) = \begin{bmatrix} I_p & & & -H_1(z)I_p \\ & I_p & & -H_2(z)I_p \\ & & \ddots & \vdots \\ & & I_p & -H_{k-1}(z)I_p \\ & & & I_p \end{bmatrix} \begin{bmatrix} I_p & & & \\ & I_p & & \\ & & \ddots & \\ & & & I_p \\ & & & & \frac{1}{v(z)}I_p \end{bmatrix} \tag{12}$$

is an invertible matrix over $\mathbb{F}(z)$. This relation and the properties of $G(z)$ established in item 2 of Proposition 4.4 immediately imply the result.

Item 2. We only need to prove that $\frac{1}{v(z)} E(z)$ is a matrix polynomial and to determine the degrees of its columns. From the definition of $D(z)$, it is obvious that the degrees of its columns are $\gamma_1 + k - 1, \ldots, \gamma_p + k - 1$. On the other hand, the Hermite interpolating polynomials $H_1(z), \ldots, H_{k-1}(z)$ have degrees at most $k - 2$, since each of them is defined through $k - 1$ interpolatory conditions. Thus, the columns of $E(z)$ have degrees $\gamma_1 + k - 1, \ldots, \gamma_p + k - 1$. The definitions of $D(z), H_i(z)$ and C_i imply that

$$E^{(b)}(\mu_t) = D^{(b)}(\mu_t) - \sum_{i=1}^{k-1} H_i^{(b)}(\mu_t)\, C_i = 0, \quad \text{for} \quad t = 1, \ldots s, \ b = 0, \ldots, \ell_t - 1,$$

which in turn implies that $v(z)$ divides each entry of $E(z)$. Therefore, $\frac{1}{v(z)} E(z)$ is a matrix polynomial and the degrees of its columns are $\gamma_1, \ldots, \gamma_p$, since the degree of $v(z)$ is $k - 1$.

Item 3. We will prove that $F(z)$ is a minimal basis by using Lemma 4.2. We start by showing that $F(z)$ satisfies the second condition of Lemma 4.2. For this purpose, denote

$A(z) = \frac{1}{v(z)} E(z)$ and let v_{k-1} be the (nonzero) coefficient of z^{k-1} in $v(z)$. Then, Lemma 4.1-3 implies $\{E(z)\} = \{v(z)\, A(z)\} = v_{k-1}\{A(z)\}$. Using again that the polynomials $H_i(z)$ have degree at most $k-2$, we get that $\{E(z)\} = \{D(z)\} = \{V(z) \otimes M(z)\} = \begin{bmatrix} \{M(z)\} \\ 0 \end{bmatrix}$, due to the structure of $V(z)$ (recall Lemma 3.1). Combining these results and using again Lemma 4.1, we get

$$\{F(z)\} = \left\{ \begin{bmatrix} C & \frac{1}{v(z)} E(z) \end{bmatrix} \right\} = \begin{bmatrix} \{C\} & \frac{1}{v_{k-1}}\{E(z)\} \end{bmatrix} = \begin{bmatrix} \widetilde{C} & \frac{1}{v_{k-1}}\{M(z)\} \\ \widehat{C} & 0 \end{bmatrix}, \tag{13}$$

where $\widetilde{C}$ (resp., $\widehat{C}$) contains the first n rows (resp., the rows from $(n+1)$-th to last) of C. Note that taking into account the definition of C,

$$\widehat{C} = \begin{bmatrix} M_{\mu_1,k-1,\ell_1,M(z)} & M_{\mu_2,k-1,\ell_2,M(z)} & \cdots & M_{\mu_s,k-1,\ell_s,M(z)} \end{bmatrix} \in \mathbb{F}^{(k-1)n \times (k-1)p},$$

which has full column rank by Lemma 3.3, since $\sum_{i=1}^s \ell_i = k-1$. Moreover, $\{M(z)\}$ has full column rank because $M(z)$ is a minimal basis. Hence, (13) implies that $\{F(z)\}$ has full column rank.

For proving that $F(z)$ satisfies the first condition of Lemma 4.2, we distinguish two cases: z_0 is not a root of $v(z)$ and z_0 is a root of $v(z)$. Suppose first that $v(z_0) \neq 0$. Then $G(z_0) = \begin{bmatrix} C & D(z_0) \end{bmatrix}$ has full column rank by Lemma 3.3 (we have already seen this in the proof of item 2 of Proposition 4.4). Moreover, using (12), we get $F(z_0) = G(z_0)\, S(z_0)$, with $S(z_0)$ invertible. This implies that $F(z_0)$ and $G(z_0)$ have the same rank and, hence, that $F(z_0)$ has full column rank.

Next, we assume that $v(z_0) = 0$, i.e., $z_0 = \mu_i$ for some $i = 1, \ldots, s$. Without loss of generality, we can assume $z_0 = \mu_1$, since otherwise we can simply permute the columns of C and place $M_{\mu_i,k,\ell_i,M(z)}$ as the first block column. In this case, we evaluate $E(z)/v(z)$ at μ_1 by applying De L'Hospital rule ℓ_1 times. This yields

$$\left[\frac{E(z)}{v(z)} \right]_{z=\mu_1} = \frac{1}{v^{(\ell_1)}(\mu_1)} \left(D^{(\ell_1)}(\mu_1) - \sum_{i=1}^{k-1} H_i^{(\ell_1)}(\mu_1)\, C_i \right),$$

where note that $v^{(\ell_1)}(\mu_1) \neq 0$. Then,

$$F(\mu_1) = \begin{bmatrix} C & D^{(\ell_1)}(\mu_1) \end{bmatrix} \begin{bmatrix} I_p & & & -H_1^{(\ell_1)}(\mu_1)I_p \\ & I_p & & -H_2^{(\ell_1)}(\mu_1)I_p \\ & & \ddots & \vdots \\ & & I_p & -H_{k-1}^{(\ell_1)}(\mu_1)I_p \\ & & & I_p \end{bmatrix} \begin{bmatrix} I_p & & & \\ & I_p & & \\ & & \ddots & \\ & & & I_p \\ & & & \frac{1}{v^{(\ell_1)}(\mu_1)}I_p \end{bmatrix}.$$

The last two matrices in the equation above are invertible. Thus

$$\operatorname{rank} F(\mu_1) = \operatorname{rank} \begin{bmatrix} C & D^{(\ell_1)}(\mu_1) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} C & [V(z) \otimes M(z)]^{(\ell_1)} \big|_{z=\mu_1} \end{bmatrix}.$$

Note that taking into account the definition of C, we can permute the columns of $\begin{bmatrix} C & D^{(\ell_1)}(\mu_1) \end{bmatrix}$ for moving the block column $D^{(\ell_1)}(\mu_1)$ just after the block $M_{\mu_1,k,\ell_1,M(z)}$ of C to get

$$\begin{bmatrix} M_{\mu_1,k,\ell_1+1,M(z)} & M_{\mu_2,k,\ell_2,M(z)} & \cdots & M_{\mu_s,k,\ell_s,M(z)} \end{bmatrix} \in \mathbb{F}^{kn \times kp}.$$

Since $\ell_1 + 1 + \sum_{i=2}^s \ell_i = k$, this matrix has full column rank by Lemma 3.3, which concludes the proof. □

Observe that if $\dim \ker L(z) = kp$, then Proposition 4.5 implies that $F(z)$ is a right minimal basis for $L(z)$, which proves item 1 in Theorem 1.1 taking into account which are the degrees of the columns of $F(z)$. Note that this proof would work in the case $v(z)$ has no infinite roots , since we are assuming this condition from the beginning of this section. Therefore, our next task is to prove that $\dim \ker L(z) = kp$, which will be a consequence of combining Theorem 4.8 with Proposition 4.5. This will lead to the main result of this section, that is, Theorem 4.9.

Before stating and proving Theorem 4.8, we introduce the auxiliary Definition 4.6 and the technical Lemma 4.7.

Definition 4.6. *Let $w(z) \in \mathbb{F}[z]$ be a scalar polynomial of degree k whose distinct roots are $\mu_0, \mu_1, \ldots, \mu_s \in \mathbb{F}$ with multiplicities $\ell_0, \ell_1, \ldots, \ell_s$. Let $V(z) \in \mathbb{F}[z]^{k \times 1}$ be the Vandermonde vector of degree $k-1$ introduced in Lemma 3.1. Let us define the constant matrices*

$$W_{\mu_i} = \begin{bmatrix} V(\mu_i) & V^{(1)}(\mu_i) & \dfrac{1}{2!}V^{(2)}(\mu_i) & \cdots & \dfrac{1}{(\ell_i - 1)!}V^{(\ell_i - 1)}(\mu_i) \end{bmatrix} \in \mathbb{F}^{k \times \ell_i} \quad \textit{for } i = 0, 1, \ldots, s.$$

The normalized confluent Vandermonde matrix associated with the polynomial $w(z)$ is

$$W = \begin{bmatrix} W_{\mu_0} & W_{\mu_1} & \cdots & W_{\mu_s} \end{bmatrix} \in \mathbb{F}^{k \times k}.$$

The normalized confluent Vandermonde matrix is obviously invertible, since it is just a standard confluent Vandermonde matrix associated with a set of distinct nodes with its columns multiplied by nonzero scalars. In fact, we might use the standard confluent Vandermonde matrix in our developments, but the normalization via the factorials will simplify the algebraic manipulations. In this context, here and below, the jth normalized derivative is defined as

$$\hat{\partial}_{x^j}^j = \frac{1}{j!}\frac{\partial^j}{\partial x^j}.$$

Lemma 4.7. *Let $\mu \in \mathbb{F}$ and let $h_c(x, y)$ be the cth complete homogenous polynomial,*

$$h_c(x, y) = \sum_{h=0}^{c} x^{c-h} y^h.$$

Then, $\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b h_c(x, y)$ evaluated at $x = y = \mu$ only depends on $a + b$, and it is equal to

$$\left[\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b h_c(x, y) \right]_{x=y=\mu} = \begin{cases} 0 & \textit{if } a + b > c; \\ \dbinom{c+1}{a+b+1} \mu^{c-a-b} & \textit{if } a + b \le c. \end{cases}$$

Proof. We may assume that $a + b \le c$, as the statement is otherwise obvious. A direct computation yields

$$\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b h_c(x, y) = \sum_{h=0}^{c} \binom{c-h}{a} \binom{h}{b} x^{c-h-a} y^{h-b},$$

where $\binom{\ell}{k} = 0$ if $\ell < k$, as usual. Thus

$$\left[\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b h_c(x, y) \right]_{x=y=\mu} = \mu^{c-a-b} \sum_{h=0}^{c} \binom{c-h}{a} \binom{h}{b} = \binom{c+1}{a+b+1} \mu^{c-a-b},$$

where the last equality follows from [39, eq. (4)], which is a corollary of the famous Chu–Vandermonde identity. $\qquad\square$

We are now ready to prove a result that will allow us to get a lower bound on rank $L(z)$ or, equivalently, due to the rank-nullity theorem, an upper bound on $\dim \ker L(z)$.

Theorem 4.8. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of degree $k \geq 2$, $v(z)$ be a scalar ansatz polynomial of degree $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Suppose the distinct roots of $v(z)$ are $\mu_1, \ldots, \mu_s$ with multiplicities $\ell_1, \ldots, \ell_s$ and let $\mu_0 \in \mathbb{F}$ be any element such that $\mu_0 \neq \mu_i$ for $i = 1, \ldots, s$. Define the scalar polynomial $w(z) = (z - \mu_0) v(z)$ of degree k and let $W \in \mathbb{F}^{k \times k}$ be the normalized confluent Vandermonde matrix associated with the polynomial $w(z)$, with $\ell_0 = 1$. Then, there exist nonzero constants $c_0, c_1, \ldots, c_s \in \mathbb{F}$ equal to*

$$c_i = \frac{1}{\ell_i!} \, w^{(\ell_i)}(\mu_i),$$

such that

$$(W^T \otimes I_m) L(\mu_0)(W \otimes I_n) = \bigoplus_{i=0}^{s} Q_i, \qquad Q_i = \begin{bmatrix} 0 & 0 & \ldots & 0 & c_i P(\mu_i) \\ 0 & 0 & & c_i P(\mu_i) & \star \\ \vdots & \vdots & \cdot^{\cdot^{\cdot}} & & \\ 0 & c_i P(\mu_i) & \star & \ldots & \star \\ c_i P(\mu_i) & \star & \star & \ldots & \star \end{bmatrix} \in \mathbb{F}^{m\ell_i \times n\ell_i},$$

i.e., each Q_i is block antitriangular with all its antidiagonal blocks equal to $c_i P(\mu_i)$. The symbol $\star$ denotes a block whose precise form is not specified.

Proof. We consider the constant matrix $(W^T \otimes I_m) L(\mu_0)(W \otimes I_n) \in \mathbb{F}^{km \times kn}$ partitioned into $k \times k$ blocks of size $m \times n$. In order to determine these blocks recall that, according to (4), $L(\mu_0) \mapsto \mathcal{B}_{P,v}(x, y, \mu_0)$ via the bijection (3). Moreover, the (α, β) $m \times n$ block of $(W^T \otimes I_m) L(\mu_0)(W \otimes I_n)$ is just the αth block row of $(W^T \otimes I_m)$ (of size $m \times km$) times $L(\mu_0)$ times the βth block column of $(W \otimes I_n)$ (of size $kn \times n$). But taking into account the structure of the columns of W and the equation (3), left multiplication of $L(\mu_0)$ by any block row of $(W^T \otimes I_m)$ corresponds to perform on $\mathcal{B}_{P,v}(x, y, \mu_0)$ a bth normalized derivative with respect to y, $0 \leq b < \ell_j$, followed by evaluation at some $y = \mu_j$. Similarly, right multiplication of $L(\mu_0)$ by any block column of $(W \otimes I_n)$ corresponds to perform on $\mathcal{B}_{P,v}(x, y, \mu_0)$ an ath normalized derivative with respect to x, $0 \leq a < \ell_i$, followed by evaluation at some $x = \mu_i$. Hence, each $m \times n$ block of $(W^T \otimes I_m) L(\mu_0)(W \otimes I_n) \in \mathbb{F}^{km \times kn}$ has the form

$$\left[\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b \frac{P(y)w(x) - w(y)P(x)}{x - y} \right]_{x=\mu_i, y=\mu_j}, \tag{14}$$

for $0 \leq j, i \leq s$, $0 \leq b \leq \ell_j - 1$ and $0 \leq a \leq \ell_i - 1$. Suppose first $i \neq j$ so that $\mu_i \neq \mu_j$. Then, for all fixed $y \neq \mu_i$, the function $(x - y)^{-1} P(y)w(x)$ has a zero in x at μ_i of multiplicity ℓ_i, and hence its ath normalized derivative with respect to x is zero when evaluated at $x = \mu_i$. Analogously, the function $(x - y)^{-1} w(y)P(x)$, for all fixed $x \neq \mu_j$, has a zero in y at μ_j of multiplicity ℓ_j, and therefore its bth normalized derivative with respect to y is zero when evaluated at $y = \mu_j$. We conclude that the corresponding block is 0. Hence

$$(W^T \otimes I_m) L(\mu_0)(W \otimes I_n) = \bigoplus_{i=0}^{s} Q_i, \qquad Q_i = (W_{\mu_i}^T \otimes I_m) L(\mu_0)(W_{\mu_i} \otimes I_n),$$

where we have used the notation in Definition 4.6.

The case $i = j$ corresponds to analyzing the blocks within Q_i with (14) particularized to $x = y = \mu_i$. Each of these blocks corresponds to a pair (b, a) with $0 \leq a, b \leq \ell_i - 1$, where ℓ_i

19

is the multiplicity of μ_i as a root of $w(z)$. Note that we are only interested in the case where $a + b \leq \ell_i - 1$, since the blocks corresponding to $a + b > \ell_i - 1$ are not specified in the statement. Recall that b corresponds to the row-block index and a to the column-block index.

Define

$$\Pi(x, y) = \frac{P(x) - P(y)}{x - y}, \qquad \Omega(x, y) = \frac{w(x) - w(y)}{x - y},$$

so that

$$\frac{P(y)w(x) - w(y)P(x)}{x - y} = P(y)\Omega(x, y) - w(y)\Pi(x, y). \tag{15}$$

Since differentiation is linear, we can consider the two addends in the right-hand side of (15) separately. Note first that

$$\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b [w(y)\Pi(x, y)] = \frac{1}{a!b!} \sum_{h=0}^{b} \binom{b}{h} w^{(h)}(y) \, \partial_{x^a}^a \cdot \partial_{y^{b-h}}^{b-h} \Pi(x, y);$$

since we eventually evaluate at $x = y = \mu_i$, and since $w^{(h)}(\mu_i) = 0$ for all $0 \leq h \leq b < \ell_i$, the contribution of the second addend to any of the $m \times n$ blocks we are investigating is 0. For the first addend in the right-hand side of (15), we observe that

$$\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b [P(y)\Omega(x, y)] = \sum_{h=0}^{b} \frac{1}{(b-h)!} P^{(b-h)}(y) \, \hat{\partial}_{x^a}^a \hat{\partial}_{y^h}^h \Omega(x, y).$$

On the other hand, using the relation $x^{c+1} - y^{c+1} = (x - y)h_c(x, y)$, where $h_c(x, y)$ is the cth complete homegenous polynomial in Lemma 4.7, we see that

$$\Omega(x, y) = \sum_{c=0}^{k-1} w_{c+1} h_c(x, y),$$

where $w(z) = \sum_{t=0}^{k} w_t \, z^t$. Hence, by Lemma 4.7,

$$[\hat{\partial}_{x^a}^a \hat{\partial}_{y^h}^h \Omega(x, y)]_{x=y=\mu_i} = \sum_{c=a+h}^{k-1} \binom{c+1}{a+h+1} w_{c+1} \mu_i^{c-a-h} = \frac{1}{(a+h+1)!} w^{(a+h+1)}(\mu_i).$$

It follows that the (b, a) block of Q_i has the expression

$$\left[\hat{\partial}_{x^a}^a \hat{\partial}_{y^b}^b \frac{P(y)w(x) - w(y)P(x)}{x - y} \right]_{x=y=\mu_i} = \sum_{h=0}^{b} \frac{1}{(b-h)!} P^{(b-h)}(\mu_i) \frac{1}{(a+h+1)!} w^{(a+h+1)}(\mu_i). \tag{16}$$

If $a + b < \ell_i - 1$, then $a + h + 1 \leq a + b + 1 < \ell_i$ and hence $w^{(a+h+1)}(\mu_i) = 0$. This shows that all the blocks above the main antidiagonal are zero. For the blocks on the antidiagonal, the only possibility for the addend in the outer summation to be nonzero is that $h = b$ so that $a + h + 1 = a + b + 1 = \ell_i$. Plugging this information into (16), and taking into account the definition of c_i, we obtain $c_i P(\mu_i)$ as desired. To conclude the proof we need to show that $c_i \neq 0$. This must be true, for if not the multiplicity of μ_i as a root of w would not be ℓ_i. $\qquad \square$

As announced, we are now in the position of stating and proving the main result of this section.

Theorem 4.9. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$ and $v(z)$ be a scalar ansatz polynomial of grade $k - 1$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, and that $v(z)$ has no infinite roots. Let $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$, $p = \dim \ker P(z)$ and $F(z) \in \mathbb{F}[z]^{kn \times kp}$ be the matrix polynomial in Proposition 4.5. Then $F(z)$ is a minimal basis for $\ker L(z)$, i.e., $F(z)$ is a right minimal basis of $L(z)$.*

Proof. In Proposition 4.5, we have already proved that $L(z)F(z) = 0$ (or, equivalently, that $\operatorname{span} F(z) \subseteq \ker L(z)$), that $F(z)$ has full column rank and that $F(z)$ is a minimal basis. Therefore, it only remains to prove that the dimension of $\ker L(z)$ is precisely kp, i.e., that it is equal to the rank of $F(z)$.

From the results above, we obviously have that $\dim \ker L(z) \geq kp = \dim \operatorname{span} F(z)$. We now argue that $\operatorname{rank} L(z) \geq k(n - p)$, which is equivalent to $\dim \ker L(z) \leq kp$ by the rank-nullity theorem, and, so, implies $\dim \ker L(z) = kp$. To this goal, we use Theorem 4.8, choosing an element μ_0 which is not an eigenvalue of $P(z)$, and recall that the roots μ_i, $i = 1, \ldots, s$, of $v(z)$ are not eigenvalues of $P(z)$ by the assumption that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint. Thus, $\operatorname{rank} P(\mu_i) = \operatorname{rank} P(z) = n - p$, for $i = 0, 1, \ldots, s$. Then, by Theorem 4.8, by Lemma 4.3, and by the fact that the matrix W in Theorem 4.8 is nonsingular, we have, using the notation of Theorem 4.8,

$$\operatorname{rank} L(z) \geq \operatorname{rank} L(\mu_0) = \sum_{i=0}^{s} \operatorname{rank} Q_i \geq \sum_{i=0}^{s} \ell_i \operatorname{rank} P(\mu_i) = (n - p) \sum_{i=0}^{s} \ell_i = (n - p)k.$$

$\square$

As a corollary of Theorem 4.9 and the properties of $\mathbb{DL}(P)$ pencils, we prove items 1 and 2 of Theorem 1.1.

Theorem 4.10. (Items 1 and 2 of Theorem 1.1) *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, with $\dim \ker P(z) = p$ and $\dim \ker P(z)^T = q$, and let $v(z) \in \mathbb{F}[z]$ be a scalar polynomial of grade $k - 1$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots, are disjoint. Denote by $L(z) \in \mathbb{F}[z]^{km \times kn}$ the pencil in $\mathbb{DL}(P)$ with ansatz polynomial $v(z)$, i.e., $L(z) = \mathbb{DL}(P, v)$. Then*

1. *The right minimal indices of $P(z)$ are $\gamma_1 \leq \cdots \leq \gamma_p$ if and only if the right minimal indices of $L(z)$ are $\underbrace{0, \ldots, 0}_{p(k-1) \text{ times}}, \gamma_1 \leq \cdots \leq \gamma_p$;*

2. *The left minimal indices of $P(z)$ are $\eta_1 \leq \cdots \leq \eta_q$ if and only if the left minimal indices of $L(z)$ are $\underbrace{0, \ldots, 0}_{q(k-1) \text{ times}}, \eta_1 \leq \cdots \leq \eta_q$.*

Proof. Item 1. Taking into account Theorems 3.4 and 3.6, we can assume without loss of generality that $v(z)$ has no infinite roots, since, otherwise, we can apply a Möbius transformation to $P(z)$ and $v(z)$ for getting a matrix polynomial $Q(z)$ and an ansatz polynomial $u(z)$ without roots at infinity, and work with $Q(z)$ and the pencil $\mathbb{DL}(Q, u)$ without changing the minimal indices of $P(z)$ and of $\mathbb{DL}(P, v)$ (recall the last paragraph in Subsection 3.2). Then, the result follows from Theorem 4.9 and item 2 in Proposition 4.5.

Item 2. The left minimal indices of $P(z)$ and of $L(z)$ are the right ones of $P(z)^T$ and of $L(z)^T$, respectively. With this in mind, the result follows immediately from the result in item 1 and the equality $L(z)^T = (\mathbb{DL}(P, v))^T = \mathbb{DL}(P^T, v)$. When $P(z)$ is square, this is a consequence of the fact that the pencils in $\mathbb{DL}(P)$ are block symmetric [22, 29, 34]. For more general rectangular matrix polynomials, it can be proved as follows. Let B be any $k \times k$ block matrix with blocks

$B_{ij} \in \mathbb{F}^{m \times n}$ and $F(x, y)$ be its associated $m \times n$ bivariate matrix polynomial via the bijection (3). If B^T is also partitioned into $k \times k$ blocks, then its blocks are $(B^T)_{ij} = (B_{ji})^T \in \mathbb{F}^{n \times m}$, which implies that its associated $n \times m$ bivariate matrix polynomial is $F(y, x)^T$. Therefore, (4) implies $(\mathbb{DL}(P, v))^T \mapsto \mathcal{B}_{P,v}(y, x, z)^T = \mathcal{B}_{P,v}(x, y, z)^T$. Moreover, from (4),

$$\mathcal{B}_{P,v}(x, y, z)^T = \frac{P(y)^T (x - z)v(x) - P(x)^T (y - z)v(y)}{x - y} = \mathcal{B}_{P^T,v}(x, y, z).$$

Thus, $(\mathbb{DL}(P, v))^T$ and $\mathbb{DL}(P^T, v)$ are mapped into the same image via the bijection (3) and, therefore, $(\mathbb{DL}(P, v))^T = \mathbb{DL}(P^T, v)$. $\qquad \square$

We conclude this section by using Theorem 4.8 to obtain for each pencil in the $\mathbb{DL}(P)$ vector space another pencil strictly equivalent to it with a remarkable simple block-sparsity pattern. This indicates that Theorem 4.8 is a relevant result that reveals some intrinsic structure of the pencils in $\mathbb{DL}(P)$.

Remark 4.11. As a side product of Theorem 4.8, we obtain a new block-arrowhead pencil which is strictly equivalent to $\mathbb{DL}(P, v)$. Indeed, it is obvious that any pencil $A(z) = A_1 z + A_0$ can be written as $A(z) = A(\mu_0) + (z - \mu_0)A_1$, for any arbitrary choice of μ_0. Assume, for maximal sparsity, that $v(z)$ in Theorem 4.8 has finite distinct roots $\mu_1, \mu_2, \ldots, \mu_{k-1}$ all with multiplicity one and let $\mu_0 \neq \mu_i$ for $i = 1, \ldots, k - 1$. Consider $A(z) = A_1 z + A_0 = (W^T \otimes I_m)L(z)(W \otimes I_n)$, where recall that W is nonsingular. Then, by Theorem 4.8,

$$A(\mu_0) = v(\mu_0)P(\mu_0) \oplus \bigoplus_{i=1}^{k-1} (\mu_i - \mu_0)v'(\mu_i)P(\mu_i)$$

is block diagonal when viewed as a $k \times k$ matrix with blocks of size $m \times n$. Next, we prove that all the blocks of A_1 outside the main block-diagonal, the first block-column and the first block-row are identically zero and, moreover, we find explicit expressions for the possibly non-zero blocks. For this purpose, observe that according to (3) and (4), the (j, i)th block of A_1 is prescribed by the formula

$$\frac{v(y)P(x) - P(y)v(x)}{x - y}$$

followed by evaluation at $x = \mu_i$, $y = \mu_j$, for $0 \leq i, j \leq k - 1$. Suppose first $i = j \neq 0$, then by L'Hospital rule we obtain the block $-v'(\mu_i)P(\mu_i)$. If $0 \neq i \neq j \neq 0$, we obtain the zero block. If $i = j = 0$, we get $v(\mu_0)P'(\mu_0) - P(\mu_0)v'(\mu_0)$. If $i \neq 0 = j$, we obtain

$$\frac{v(\mu_0)P(\mu_i)}{\mu_i - \mu_0},$$

and if $i = 0 \neq j$, we obtain the same expression replacing i by j.

Applying the results above to the case $k = 3$, with $v(z) = z^2 - 1$, thus $\mu_1 = 1$ and $\mu_2 = -1$, and taking $\mu_0 = 0$, we get that the following pencil

$$A(z) = \begin{bmatrix} -P(0) & 0 & 0 \\ 0 & 2P(1) & 0 \\ 0 & 0 & 2P(-1) \end{bmatrix} + z \begin{bmatrix} -P'(0) & -P(1) & P(-1) \\ -P(1) & -2P(1) & 0 \\ P(-1) & 0 & 2P(-1) \end{bmatrix}$$

is strictly equivalent to $\mathbb{DL}(P, v)$ and, therefore, has the same elementary divisors and minimal indices as $\mathbb{DL}(P, v)$. Hence, $A(z)$ is a linearization of $P(z)$ if $P(z)$ is regular and the eigenvalue exclusion theorem holds, or it is a pencil to which the analysis of this paper applies if $P(z)$ is singular and the eigenvalue exclusion theorem holds. The computational cost of constructing this pencil is however nonzero, since it requires to evaluate $P(z)$ and $P'(z)$ at some points.

5 The elementary divisors of $\mathbb{DL}(P, v)$

The goal of this section is to prove item 3 of Theorem 1.1, i.e., to show that $\lambda \in \mathbb{F} \cup \{\infty\}$ is an eigenvalue of $P(z)$ if and only if λ is an eigenvalue of $L(z) = \mathbb{DL}(P, v)$ and that its partial multiplicities as an eigenvalue of $P(z)$ and as an eigenvalue of $L(z)$ are the same, under the assumptions of Theorem 1.1. Thanks to Theorem 3.6, we can assume in the arguments of this section that $P(z)$ has no infinite eigenvalues and that $v(z)$ has no infinite roots. The main result in this section is Theorem 5.3. Its proof requires two preliminary results: Proposition 5.1 about the relationship between the root polynomials of $P(z)$ and $L(z)$ and Lemma 5.2 about the orders of the elements of a λ-independent set of root polynomials of an arbitrary matrix polynomial.

Proposition 5.1. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$ and $v(z)$ be a scalar ansatz polynomial of grade $k-1$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, that $P(z)$ has no infinite eigenvalues and that $v(z)$ has no infinite roots. Let $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$ and $V(z) \in \mathbb{F}[z]^{k \times 1}$ be the Vandermonde vector of degree $k-1$ defined in Lemma 3.1. Then:*

1. *If $r(z) \in \mathbb{F}[z]^{n \times 1}$ is a root polynomial for $P(z)$ of order ℓ at λ, then $\rho(z) = V(z) \otimes r(z) \in \mathbb{F}[z]^{kn \times 1}$ is a root polynomial for $L(z)$ of order ℓ at λ. In particular, if λ is an eigenvalue of $P(z)$, then λ is an eigenvalue of $L(z)$.*

2. *If $\{r_i(z)\}_{i=1}^{t}$ is a maximal set of root polynomials at λ for $P(z)$, then $\{\rho_i(z)\}_{i=1}^{t}$, with $\rho_i(z) = V(z) \otimes r_i(z)$, is a λ-independent set of root polynomials for $L(z)$.*

Proof. Item 1. By assumption, $P(z)r(z) = (z - \lambda)^{\ell}s(z)$ with $s(\lambda) \neq 0$ and, moreover, $r(\lambda) \notin \ker_\lambda P(z) = \text{span } M(\lambda)$, where $M(z) \in \mathbb{F}[z]^{n \times p}$ is a right minimal basis for $P(z)$.

Observe first that (3) and (4) imply

$$L(z)\rho(z) \mapsto \left. \frac{P(y)(x - z)v(x) - P(x)(y - z)v(y)}{x - y} \right|_{x=z} r(z) = v(y)P(z)r(z) = v(y)(z - \lambda)^{\ell}s(z).$$

If $v(y) = \omega^T V(y) = V(y)^T \omega$, with $\omega \in \mathbb{F}^{k \times 1}$, the previous equation implies in turn that

$$L(z)\rho(z) = (z - \lambda)^{\ell} (\omega \otimes s(z)).$$

Since the ansatz polynomial $v(y)$ is not zero, we get that $\omega \neq 0$ and that $\omega \otimes s(\lambda) \neq 0$. Therefore, it only remains to prove that $\rho(\lambda) \notin \ker_\lambda L(z)$. For this purpose, note first that from Theorem 4.9 we get $\ker_\lambda L(z) = \text{span } F(\lambda)$. Moreover, the fact that $r(z)$ is a root polynomial at λ for $P(z)$ implies that λ is a finite eigenvalue of $P(z)$, since root polynomials at λ for $P(z)$ exist if and only if λ is a finite eigenvalue of $P(z)$ [13, Proposition 2.12]. Thus, $v(\lambda) \neq 0$ by assumption and the matrix $S(\lambda)$ in (12) is well-defined and is nonsingular. This implies that $F(\lambda) = G(\lambda) S(\lambda)$, where $G(z) = \begin{bmatrix} C & V(z) \otimes M(z) \end{bmatrix}$ is the matrix polynomial defined in Proposition 4.4. Therefore,

$$\ker_\lambda L(z) = \text{span } F(\lambda) = \text{span } G(\lambda). \tag{17}$$

Next, we proceed by contradiction and assume that $\rho(\lambda) = V(\lambda) \otimes r(\lambda) \in \ker_\lambda L(z) = \text{span } G(\lambda)$. Then there exists a constant vector $d \in \mathbb{F}^{kp \times 1}$ such that

$$\begin{bmatrix} C & V(\lambda) \otimes M(\lambda) \end{bmatrix} d = V(\lambda) \otimes r(\lambda). \tag{18}$$

If we partition d conformably to $\begin{bmatrix} C & V(\lambda) \otimes M(\lambda) \end{bmatrix}$ as $d = \begin{bmatrix} d_C \\ d_M \end{bmatrix}$, then (18) is equivalent to

$$C d_C + V(\lambda) \otimes (M(\lambda) d_M) = V(\lambda) \otimes r(\lambda) \iff \begin{bmatrix} C & V(\lambda) \otimes I_n \end{bmatrix} \begin{bmatrix} d_C \\ M(\lambda) d_M - r(\lambda) \end{bmatrix} = 0.$$

23

The matrix $\begin{bmatrix} C & V(\lambda) \otimes I_n \end{bmatrix}$ has full column rank, as a consequence of the definition of the matrix C in Proposition 4.4, of $\lambda \neq \mu_i$ for $i = 1, \ldots, s$, and of Lemma 3.3 (apply this lemma with $M_{s+1}(z) = I_n$, $\ell_{s+1} = 1$ and $\alpha_{s+1} = \lambda$). Therefore, $d_C = 0$ and $M(\lambda) \, d_M - r(\lambda) = 0$, which implies $r(\lambda) \in \ker_\lambda P(z) = \operatorname{span} M(\lambda)$, contradicting the assumption.

Item 2. By item 1, it is clear that $\rho_i(z)$ are root polynomials at λ for $L(z)$. It remains to check the claim of λ-independence. Let

$$ R(z) = \begin{bmatrix} r_1(z) & \cdots & r_t(z) \end{bmatrix}. $$

Then, taking into account (17) and $G(\lambda) = \begin{bmatrix} C & V(\lambda) \otimes M(\lambda) \end{bmatrix}$, the goal is to prove that the matrix

$$ \begin{bmatrix} C & V(\lambda) \otimes M(\lambda) & V(\lambda) \otimes R(\lambda) \end{bmatrix} = \begin{bmatrix} C & V(\lambda) \otimes \begin{bmatrix} M(\lambda) & R(\lambda) \end{bmatrix} \end{bmatrix} $$

has full column rank. But this follows from Lemma 3.3, with $M_{s+1}(z) = \begin{bmatrix} M(z) & R(z) \end{bmatrix}$, $\ell_{s+1} = 1$ and $\alpha_{s+1} = \lambda$, and the fact that $\begin{bmatrix} M(\lambda) & R(\lambda) \end{bmatrix}$ has full column rank, since $\{r_i(z)\}_{i=1}^t$ is a maximal set of root polynomials at λ for $P(z)$. $\qquad \square$

Lemma 5.2. *Let $\{\ell_i\}_{i=1}^t$ be the partial multiplicities at $\lambda \in \mathbb{F}$ of a matrix polynomial $Q(z) \in \mathbb{F}[z]^{m \times n}$, and let $\{r_i(z)\}_{i=1}^c$ be a λ-independent set of root polynomials at λ for $Q(z)$ of orders $\{\kappa_i\}_{i=1}^c$. Then, $c \leq t$ and if both $\{\ell_i\}_{i=1}^t$ and $\{\kappa_i\}_{i=1}^c$ are listed in non-increasing order, it holds $\kappa_i \leq \ell_i$ for $i = 1, \ldots, c$.*

Proof. The definition and properties of λ-independent and maximal sets of root polynomials at λ for $Q(z)$ imply immediately that $c \leq t$. Let $M(z)$ be a minimal basis for $Q(z)$ and let $v_1, \ldots, v_{t-c}$ be vectors in $\mathbb{F}^n$ that complete the columns of $\begin{bmatrix} M(\lambda) & r_1(\lambda) & \cdots & r_c(\lambda) \end{bmatrix}$ to a basis for $\ker Q(\lambda) \subseteq \mathbb{F}^n$. Then, it is clear that $\{r_i(z)\}_{i=1}^c \cup \{v_i\}_{i=1}^{t-c}$ is a complete set of root polynomials at λ for $Q(z)$, since $Q(\lambda) v_i = 0$ and $v_i \notin \operatorname{span} M(\lambda)$. Denote their orders, listed in a non-increasing manner, by $\{\hat{\ell}_i\}_{i=1}^t$. By Theorem 2.2, we have $\hat{\ell}_i \leq \ell_i$ for $i = 1, \ldots, t$. On the other hand, since the integers $\{\kappa_i\}_{i=1}^c$ are by construction a subsequence of $\{\hat{\ell}_i\}_{i=1}^t$, the inequalities $\kappa_i \leq \hat{\ell}_i \leq \ell_i$ must hold for $i = 1, \ldots, c$. $\qquad \square$

Next, we prove item 3 of Theorem 1.1. The proof is based on Theorem 4.10, the index sum theorem [11, Theorem 6.5], Proposition 5.1 and Lemma 5.2.

Theorem 5.3. (Item 3 of Theorem 1.1) *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$ and let $v(z) \in \mathbb{F}[z]$ be a scalar polynomial of grade $k - 1$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots, are disjoint. Denote by $L(z) \in \mathbb{F}[z]^{km \times kn}$ the pencil in $\mathbb{DL}(P)$ with ansatz polynomial $v(z)$, i.e., $L(z) = \mathbb{DL}(P, v)$. Then, $\lambda \in \mathbb{F} \cup \{\infty\}$ is an eigenvalue of $P(z)$ if and only if λ is an eigenvalue of $L(z)$ and, moreover, the partial multiplicities of λ as an eigenvalue of $P(z)$ and as an eigenvalue of $L(z)$ coincide.*

Proof. Similarly to the proof of Theorem 4.10, taking into account Theorems 3.6 and 3.4, we can assume without loss of generality that $P(z)$ has no infinite eigenvalues and that $v(z)$ has no infinite roots, for if not we can just apply an appropriate Möbius transformation to $P(z)$ and $v(z)$. The proof follows an argument by exhaustion. By Theorem 4.10, and with the notation used there, we obtain that

$$ \operatorname{grade}(L) \operatorname{rank}(L) = kn - kp = k(n - p) = \operatorname{grade}(P) \operatorname{rank}(P), $$

24

since grade$(L) = 1$. Moreover, from Theorem 4.10, we know that the sum of the right and the left minimal indices of $L(z)$ and $P(z)$ are the same. Hence, by the index sum theorem [11, Theorem 6.5],

$$\sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, P) = \sum_{\lambda \in \Lambda(L)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L), \tag{19}$$

where $\Lambda(P)$ denotes the set of distinct eigenvalue of $P(z)$ (and analogously for $\Lambda(L)$), $g(\lambda, P)$ is the geometric multiplicity of λ as an eigenvalue of P (and similarly for $g(\lambda, L)$), and $\ell_i(\lambda, P)$ is the i-th largest partial multiplicity of λ as an eigenvalue of P (and ditto for $\ell_i(\lambda, L)$). Recall that the partial multiplicities have been defined to be positive integers and, so, $g(\lambda, P)$ is equal to the number of partial multiplicities of λ as an eigenvalue of $P(z)$.

Observe that the following results hold:

(a) $\Lambda(P) \subseteq \Lambda(L)$, by item 1 in Proposition 5.1,

(b) $g(\lambda, P) \leq g(\lambda, L)$ for all $\lambda \in \Lambda(P)$, by item 2 in Proposition 5.1 and the properties of maximal sets of root polynomials, and

(c) $\ell_i(\lambda, P) \leq \ell_i(\lambda, L)$ for all $\lambda \in \Lambda(P)$ and for $i = 1, \ldots, g(\lambda, P)$, by Lemma 5.2 and Proposition 5.1.

As a consequence, we have that at each $\lambda \in \Lambda(P)$ it must hold

$$\sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, P) \leq \sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, L) \leq \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L), \tag{20}$$

where the left inequality is strict if and only if $\ell_i(\lambda, P) < \ell_i(\lambda, L)$ for at least one $i = 1, \ldots, g(\lambda, P)$ and the right inequality is strict if and only if $g(\lambda, P) < g(\lambda, L)$.

We prove first that $\Lambda(P) = \Lambda(L)$ by contradiction. Assume that $\Lambda(P) \subset \Lambda(L)$, i.e., the inclusion in (a) above is strict. Then, from (19),

$$\sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, P) = \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L) + \sum_{\lambda \in \Lambda(L) \setminus \Lambda(P)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L)$$

$$> \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L)$$

$$\geq \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, P),$$

where the first strict inequality follows from the strict inclusion $\Lambda(P) \subset \Lambda(L)$ and the last inequality follows from (20). This is a manifest contradiction. Therefore, $\Lambda(P) = \Lambda(L)$.

Secondly, we prove that $g(\lambda, P) = g(\lambda, L)$ for all $\lambda \in \Lambda(P) = \Lambda(L)$. We proceed again by contradiction. If $g(\lambda_0, P) < g(\lambda_0, L)$ for some $\lambda_0 \in \Lambda(P) = \Lambda(L)$, then the second inequality in (20) is strict for λ_0 and $\sum_{i=1}^{g(\lambda_0,P)} \ell_i(\lambda_0, P) < \sum_{i=1}^{g(\lambda_0,L)} \ell_i(\lambda_0, L)$. Combining this with (20), (19), and $\Lambda(P) = \Lambda(L)$, we get

$$\sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L) > \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,P)} \ell_i(\lambda, P) = \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda,L)} \ell_i(\lambda, L),$$

which is not possible. Then, $g(\lambda, P) = g(\lambda, L)$ for all $\lambda \in \Lambda(P) = \Lambda(L)$.

Finally, we prove $\ell_i(\lambda, P) = \ell_i(\lambda, L)$ for all $\lambda \in \Lambda(P) = \Lambda(L)$ and for all $i = 1, \dots, g(\lambda, P) = g(\lambda, L)$. We proceed again by contradiction. Assume that there exist at least one $\lambda_0 \in \Lambda(P) = \Lambda(L)$ and at least one $i = 1, \dots, g(\lambda_0, P) = g(\lambda_0, L)$, such that $\ell_i(\lambda_0, P) < \ell_i(\lambda_0, L)$. Thus, the first inequality in (20) is strict for λ_0, i.e., $\sum_{i=1}^{g(\lambda_0, P)} \ell_i(\lambda_0, P) < \sum_{i=1}^{g(\lambda_0, P)} \ell_i(\lambda_0, L)$. Combining this with (20), (19), and $g(\lambda, P) = g(\lambda, L)$ for all $\lambda \in \Lambda(P) = \Lambda(L)$, we get

$$\sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda, P)} \ell_i(\lambda, L) > \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda, P)} \ell_i(\lambda, P) = \sum_{\lambda \in \Lambda(P)} \sum_{i=1}^{g(\lambda, P)} \ell_i(\lambda, L),$$

which is not possible. Thus, $\ell_i(\lambda, P) = \ell_i(\lambda, L)$ for all $\lambda \in \Lambda(P) = \Lambda(L)$ and for all $i = 1, \dots, g(\lambda, P) = g(\lambda, L)$ and the proof is complete. $\qquad\square$

We finish this section with Theorem 5.4, which is a corollary of Proposition 5.1 and Theorem 5.3.

Theorem 5.4. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$ and $v(z)$ be a scalar ansatz polynomial of grade $k - 1$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, that $P(z)$ has no infinite eigenvalues, and that $v(z)$ has no infinite roots. Let $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$ and $V(z) \in \mathbb{F}[z]^{k \times 1}$ be the Vandermonde vector defined in Lemma 3.1. Let λ be an eigenvalue of $P(z)$ and $\{r_i(z)\}_{i=1}^{t}$ be a maximal set of root polynomials at λ for $P(z)$. Then, $\{\rho_i(z)\}_{i=1}^{t}$, $\rho_i(z) = V(z) \otimes r_i(z)$, is a maximal set of root polynomials at λ for $L(z)$.*

Proof. Proposition 5.1 implies that $\{\rho_i(z)\}_{i=1}^{t}$ is a λ-independent set of root polynomials for $L(z)$ with the same orders as $\{r_i(z)\}_{i=1}^{t}$. From Theorem 5.3 and the properties of maximal sets of root polynomials, we deduce that the orders of $\{\rho_i(z)\}_{i=1}^{t}$ are precisely the partial multiplicites of λ as an eigenvalue of $L(z)$. This implies, in particular, that $t = \dim \ker L(\lambda) - \dim \ker_\lambda L(z)$. Therefore, $\{\rho_i(z)\}_{i=1}^{t}$ is a complete set of root polynomials at λ for $L(z)$. Then, we get that $\{\rho_i(z)\}_{i=1}^{t}$ is maximal from Theorem 2.2. $\qquad\square$

6 Recovery of vectors associated with $P(z)$ from their analogues associated with $\mathbb{DL}(P, v)$

In this section we describe how to recover (right) minimal bases, eigenvectors, and root polynomials of a possibly singular matrix polynomial $P(z)$ from those of $\mathbb{DL}(P, v)$, under the assumptions of Theorem 1.1. We omit the treatment of left minimal bases, eigenvectors, and root polynomials; indeed, it is completely analogous since $\mathbb{DL}(P^T, v) = (\mathbb{DL}(P, v))^T$.

6.1 Recovery of a minimal basis for $P(z)$ from one of $\mathbb{DL}(P, v)$

When studying singular matrix polynomials, often not only the minimal indices are sought, but also a minimal basis. Minimal bases of singular pencils can be computed by postprocessing the output of the staircase algorithm [37, 42]. In this section, we show how a minimal basis for $P(z)$ can be extracted from one of $L(z) = \mathbb{DL}(P, v)$, under the assumption that the sets of roots of $v(z)$ and of eigenvalues of $P(z)$ are disjoint. It is worth reminding that a minimal basis $F(z)$ of $L(z)$ was built in Theorem 4.9 starting from any minimal basis $M(z)$ of $P(z)$, under the additional assumptionthat $v(z)$ has no infinite roots. However, there are two obstacles for recovering $M(z)$ from this construction: first $M(z)$ is really "hidden" inside $F(z)$ and, second and more important, there may be minimal basis of $L(z)$ with a structure different from the one

of $F(z)$. This is easy to see if some of the minimal indices of $P(z)$ are zero. Therefore, we follow in this section a fully different approach.

The vector $\omega \in \mathbb{F}^{k \times 1}$ associated to the scalar polynomial $v(x)$ under the bijection (2) plays a key role in this section. Recall that ω is what was called the *ansatz vector* in [29], and it maps under (2) to what we call, following [34], the *ansatz polynomial* in this paper. We often use in this section that $v(x) = \omega^T V(x)$, where $V(x)$ is the Vandermonde vector of degree $k - 1$ introduced in Lemma 3.1.

The following simple lemma will be often used in this section. It follows directly from the definition of $\mathbb{DL}(P)$ pencils and does not require any assumption on $P(z)$ and on $v(z)$.

Lemma 6.1. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Then*

1. $(V(z)^T \otimes I_m) L(z) = \omega^T \otimes P(z) = P(z) (\omega^T \otimes I_n)$;

2. $L(z) (V(z) \otimes I_n) = \omega \otimes P(z) = (\omega \otimes I_m) P(z)$;

3. $L(z) (V(z) \otimes M(z)) = 0$ *for any minimal basis* $M(z) \in \mathbb{F}[z]^{n \times p}$ *of* $P(z)$;

4. $v(z) S(z) = (\omega^T \otimes I_n)(V(z) \otimes S(z))$ *for any rational matrix (or vector)* $S(z) \in \mathbb{F}(z)^{n \times j}$.

The equalities in items 1, 2, and 3 remain valid if the variable z is replaced by any $\lambda \in \mathbb{F}$, and the one in item 4 remains valid if z is replaced by any λ that is not a pole of the entries of $S(z)$.

Proof. Item 1 follows from (4), which implies that $(V(z)^T \otimes I_m) L(z) \mapsto \mathcal{B}_{P,v}(x, z, z) = v(x) P(z)$ via the bijection (3), while $\omega^T \otimes P(z) \mapsto (\omega^T \otimes P(z))(V(x) \otimes I_n) = v(x) P(z)$. Analogously, item 2 follows from $L(z) (V(z) \otimes I_n) \mapsto \mathcal{B}_{P,v}(z, y, z) = v(y) P(z)$ via the bijection (3), while $\omega \otimes P(z) \mapsto (V(y)^T \otimes I_m)(\omega \otimes P(z)) = v(y) P(z)$. Observe that items 1 and 2 follow also from the fact that the set of $\mathbb{DL}(P)$ pencils is the intersection of the sets of $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ pencils introduced in [29].

Item 3 was already proved in item 2 of Proposition 4.4. It also follows from multiplying on the right the equality in item 2 by $M(z)$ and $P(z) M(z) = 0$. Item 4 is an immediate consequence of the properties of the Kronecker product. $\square$

Lemma 6.1 allows us to prove the next proposition which, again, does not require any assumption on $P(z)$ and on $v(z)$.

Proposition 6.2. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Then, the map*

$$\Omega : \ker L(z) \to \ker P(z), \qquad n(z) \in \mathbb{F}(z)^{kn} \mapsto \Omega(n(z)) = (\omega^T \otimes I_n) n(z) \in \mathbb{F}(z)^n$$

is a surjective vector space homomorphism.

Proof. Lemma 6.1-1 implies that any $n(z) \in \ker L(z)$ satisfies $0 = (V(z)^T \otimes I_m) L(z) n(z) = P(z) ((\omega^T \otimes I_n) n(z))$. So, $(\omega^T \otimes I_n) n(z) \in \ker P(z)$. Moreover, Ω is an $\mathbb{F}(z)$-linear map, and thus a vector space homomorphism from $\ker L(z)$ into $\ker P(z)$. To prove surjectivity, write an arbitrary $m(z) \in \ker P(z)$ as $m(z) = M(z) r(z)$, where $M(z)$ is a minimal basis of $P(z)$ and $r(z)$ is a rational vector. From Lemma 6.1-4, we get $m(z) = v(z) M(z) (r(z)/v(z)) = (\omega^T \otimes I_n)(V(z) \otimes M(z) (r(z)/v(z)))$ and, from Lemma 6.1-3, $L(z) (V(z) \otimes M(z)(r(z)/v(z))) = 0$, and the surjectivity of Ω is proved. $\square$

From Proposition 6.2 and item 1 of Theorem 4.10, we obtain Theorem 6.3, which is the main result of this section.

Theorem 6.3. *(Recovery of minimal bases) Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots, are disjoint. Let $N(z)$ be a minimal basis of $L(z)$ and define the following matrices:*

1. *$\widehat{M}(z) := (\omega^T \otimes I_n)N(z)$;*

2. *Partition, possibly after having permuted its columns, $\widehat{M}(z) = \begin{bmatrix} M_d(z) & M_c \end{bmatrix}$ where no column of $M_d(z)$ is constant in z while M_c is constant;*

3. *$\widetilde{M}(z) := M_d(z)$, if $M_c = 0$, and $\widetilde{M}(z) := \begin{bmatrix} M_d(z) & M_b \end{bmatrix}$, where M_b is any basis of $\mathrm{span}\, M_c$ considered as a subspace over $\mathbb{F}$, otherwise.*

Then, $\widetilde{M}(z)$ is a minimal basis of $P(z)$. Thus, $M_d(z)$ is a minimal basis of $P(z)$ if and only if $M_c = 0$.

Proof. Recall that, according to Theorem 4.10, $p = \dim \ker P(z)$ if and only if $kp = \dim \ker L(z)$ and that the sum of the (right) minimal indices of $P(z)$ is equal to the sum of those of $L(z)$. Proposition 6.2 implies that the polynomial matrix $\widehat{M}(z) := (\omega^T \otimes I_n)N(z) \in \mathbb{F}[z]^{n \times kp}$ spans $\ker P(z)$. Moreover, the degrees of the columns of $\widehat{M}(z)$ are at most equal to the minimal indices of $L(z)$, i.e, to the degrees of the corresponding columns of $N(z)$, which are the minimal indices of $P(z)$ together with $p(k - 1)$ minimal indices equal to zero by Theorem 4.10. Necessarily, $\widehat{M}(z)$ is not a basis of $\ker P(z)$ for dimensional reasons, but since its columns are a spanning set, we can construct a polynomial basis (arranged as a polynomial matrix) $B(z)$ of $\ker P(z)$ by selecting p linearly independent columns of $\widehat{M}(z)$. The fact that any polynomial basis of $\ker P(z)$ must have a sum of the degrees of its vectors greater than or equal to the sum of the minimal indices of $\ker P(z)$ implies that all the positive degrees of the columns of $M_d(z)$ are *equal* to the positive minimal indices of $P(z)$ and that the columns of $M_d(z)$ must be part of the polynomial basis $B(z)$ of $\ker P(z)$, which is in fact minimal due to the value of the sum of the degrees of its vectors. Now, there are two possibilities: (a) $M_d(z)$ has p columns and (b) $M_d(z)$ has less than p columns. In the case (a) $M_d(z)$ is a minimal basis of $\ker P(z)$ for dimensional reasons and each column of M_c is of the form $M_d(z)x(z)$ for some rational vector $x(z)$, which must be zero, because, if not, then [16, Main Theorem, Item 4] would imply that the degree of $M_d(z)x(z)$ is positive. In the case (b) $B(z) = \begin{bmatrix} M_d(z) & S \end{bmatrix}$, with S a subset of columns of M_c, and each column of M_c is of the form $\begin{bmatrix} M_d(z) & S \end{bmatrix} y(z)$ for some rational vector $y(z)$. The entries of $y(z)$ corresponding to $M_d(z)$ must be zero and the ones corresponding to S must be constant, because, if not, then [16, Main Theorem, Item 4] would imply again that the degree of $\begin{bmatrix} M_d(z) & S \end{bmatrix} y(z)$ is positive. Therefore, $\mathrm{span}_{\mathbb{F}} S = \mathrm{span}_{\mathbb{F}} M_c$, which implies that if S is replaced in $B(z)$ by any other basis M_b of $\mathrm{span}_{\mathbb{F}} M_c$, then another minimal basis of $\ker P(z)$ is obtained. $\square$

Theorem 6.3 can be directly translated into a practical algorithm, where the only step that is not immediate is to construct M_b when M_c is not zero. When $\mathbb{F} \subseteq \mathbb{C}$, which is the most important case in practice, M_b can be reliably computed through a singular value decomposition of M_c. For other fields, one can apply Gaussian reduction to M_c to identify its pivot columns, which can be taken as M_b.

Remark 6.4. Note that Proposition 6.2 and Theorem 6.3 are clearly related to the recovery of the left minimal bases of a *singular* matrix polynomial $P(z)$ from those of $\mathbb{L}_1(P)$ linearizations and to the recovery of the right minimal bases of $P(z)$ from those of $\mathbb{L}_2(P)$ linearizations presented in [8, Theorem 5.10], since in both cases such recoveries are obtained just by the product $(\omega^T \otimes I_n)N(z)$, where ω is the ansatz vector defining the linearization and $N(z)$ is any of its corresponding minimal bases. Taking into account that the set of $\mathbb{DL}(P)$ pencils is $\mathbb{L}_1(P) \cap \mathbb{L}_2(P)$, this relation may seem natural, despite the fact that none of the $\mathbb{DL}(P)$ pencils is a linearization of $P(z)$. However, we emphasize that, in contrast, the recovery of the right minimal bases of $P(z)$ from those of $\mathbb{L}_1(P)$ linearizations and the recovery of the left minimal bases of $P(z)$ from those of $\mathbb{L}_2(P)$ linearizations presented in [8, Theorem 5.10] are not useful at all in the context of $\mathbb{DL}(P)$ pencils for singular $P(z)$.

Next, we determine in Theorem 6.5 a basis of the kernel of the map Ω defined in Proposition 6.2. For simplicity, we use in Theorem 6.5 the additional hypotheses that $P(z)$ has no infinite eigenvalues and that $v(z)$ has no infinite roots. These assumptions can be removed via a Möbius transformation, but at the cost of making the statement and the proof of the result more cumbersome.

Theorem 6.5. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$, and Ω be the vector space homomorphism in Proposition 6.2. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, that $P(z)$ has no infinite eigenvalues and that $v(z)$ has no infinite roots. Let $p = \dim \ker P(z)$ and $C \in \mathbb{F}^{kn \times (k-1)p}$ be the matrix defined in Proposition 4.4. Then, $\ker \Omega = \operatorname{span} C$ and the columns of C form a basis of $\ker \Omega$.*

Proof. Theorem 4.10, Proposition 6.2, and the rank-nullity theorem imply that $\dim \ker \Omega = p(k-1)$. Moreover, C has full column rank and $\operatorname{span} C \subseteq \ker L(z)$ by Proposition 4.4. Therefore, it only remains to prove that $(\omega^T \otimes I_n) C = 0$. For that purpose recall that, according to the definition in Lemma 3.2, each submatrix $M_{\mu_i, k, \ell_i, M(z)}$ of C satisfies

$$
\begin{aligned}
(\omega^T \otimes I_n) M_{\mu_i, k, \ell_i, M(z)} &= \left[\omega^T V(z) \otimes M(z) \quad [\omega^T V(z) \otimes M(z)]' \quad \cdots \quad [\omega^T V(z) \otimes M(z)]^{(\ell_i - 1)} \right]\Big|_{z = \mu_i} \\
&= \left[v(z)M(z) \quad [v(z)M(z)]' \quad \cdots \quad [v(z)M(z)]^{(\ell_i - 1)} \right]\Big|_{z = \mu_i} \\
&= 0,
\end{aligned}
$$

because $v^{(a)}(\mu_i) = 0$ for $a = 0, 1, \ldots, \ell_i - 1$. $\qquad\square$

Observe that, according to Theorem 6.5, the minimal indices of the rational subspace $\ker \Omega$ are all equal to zero and that this is the reason why $L(z)$ has $p(k-1)$ minimal indices equal to zero in addition to those of $P(z)$, for any $P(z)$.

We conclude this section with the following result. Again, the assumptions in item 2 of $P(z)$ (resp. $v(z)$) not having infinite eigenvalues (resp. roots) can be removed via a Möbius transformation.

Corollary 6.6. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$, and Ω be the vector space homomorphism in Proposition 6.2. Then,*

1. $\ker P(z) \cong \ker L(z) / \ker \Omega$;

2. *if, moreover, the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, $P(z)$ has no infinite eigenvalues, $v(z)$ has no infinite roots and C is the matrix defined in Proposition 4.4, then $\ker P(z) \cong \ker L(z) / \operatorname{span} C$.*

Proof. Item 1 is an immediate consequence of the first isomorphism theorem [6, Theorem 6.12] and the surjectivity of the vector space homomorphism Ω proved in Proposition 6.2. Item 2 follows immediately from item 1 and Theorem 6.5. □

6.2 Recovery of eigenvectors of $P(z)$ from those of $\mathbb{DL}(P, v)$

Even though it is not common in the literature to discuss eigenvectors of singular matrix polynomials, it was shown in [13, §2.3] and in [38, §3] that these can be defined as nonzero elements of the quotient space $\ker P(\lambda) / \ker_\lambda P(z)$ where $\lambda \in \mathbb{F}$ is an eigenvalue of $P(z)$. In other words, given a vector u such that $P(\lambda)u = 0$, we say that the equivalence class $[u] = \{u + u' : u' \in \ker_\lambda P(z)\}$ is an eigenvector of $P(z)$ associated with λ if $[u] \neq [0]$. This clearly generalizes the notion of an eigenvector of a regular matrix polynomial as a nonzero vector in the null space of the matrix polynomial evaluated at the eigenvalue (indeed, observe that if $P(z)$ is regular then $\ker_\lambda P(z) = \{0\}$). Albeit somewhat abstract, this concept of eigenvectors of singular matrix polynomials has useful applications [24, 27]. In [24, 27], it is also argued that in the (mostly relevant in practice) case $\mathbb{F} \subseteq \mathbb{C}$ then one can make the concept of eigenvectors of a singular matrix polynomial much more concrete. Suppose for example that λ is a simple eigenvalue; then this is done by picking, as a representative of the equivalence class of eigenvectors $[u]$, the unique (up to phase) vector $u_0 \in \mathbb{C}^n$ such that (1) $\|u_0\|_2 = 1$, (2) $u_0 \in (\ker_\lambda P(z))^\perp$ and (3) $[u_0] = [u]$.

In this sense, a recovery of eigenvectors is possible with a similar approach as for minimal bases. This recovery is based on Proposition 6.8 and Corollary 6.9, which are related to Proposition 6.2 and Corollary 6.6.

Remark 6.7. We warn the reader that we are about to commit an abuse of notation, using the same symbol Ω to denote the map defined in Proposition 6.2, the two maps defined in Proposition 6.8 below, and the map defined by (21) below. Indeed, the first is an $\mathbb{F}(z)$-linear map defined on $\mathbb{F}(z)^{kn}$ and restricted on $\ker L(z)$; the second and the third are the restrictions of an $\mathbb{F}$-linear map, defined on $\mathbb{F}^{kn}$, on $\ker_\lambda L(z)$ and $\ker L(\lambda)$, respectively; and the fourth is an $\mathbb{F}$-linear map defined on $\ker L(\lambda) / \ker_\lambda L(z)$ (which formally is a quotient $\mathbb{F}$-vector space of equivalence classes). In practice and when represented in the canonical bases, though, all these maps are realized by left-multiplication times $\omega^T \otimes I_n$; moreover, the context always makes clear which one we are referring to. For these reasons, we opt out of an overwhelmingly baroque notation that uses four different symbols.

Proposition 6.8. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k - 1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Fix any $\lambda \in \mathbb{F}$ such that $v(\lambda) \neq 0$. Then,*

1. *the map*
$$\Omega : \ker_\lambda L(z) \to \ker_\lambda P(z), \qquad u \mapsto (\omega^T \otimes I_n)u$$
is a surjective vector space homomorphism;

2. *the map*
$$\Omega : \ker L(\lambda) \to \ker P(\lambda), \qquad u \mapsto (\omega^T \otimes I_n)u$$
is a surjective vector space homomorphism;

3. *if, moreover, the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint, then the kernel of the map in item 1 is equal to the kernel of the map in item 2, which is denoted by $\ker \Omega$. As a consequence, $\ker_\lambda P(z) \cong \ker_\lambda L(z)/\ker \Omega$ and $\ker P(\lambda) \cong \ker L(\lambda)/\ker \Omega$.*

Proof. Item 1. Let us prove first that Ω indeed maps vectors of $\ker_\lambda L(z)$ to vectors of $\ker_\lambda P(z)$. For this purpose, let $N(z)$ be a minimal basis of $L(z)$ and $M(z)$ be a minimal basis of $P(z)$. Then, Proposition 6.2 implies that $(\omega^T \otimes I_n)N(z) = M(z)R(z)$, where $R(z)$ is a polynomial matrix by [16, Main Theorem, Item 4]. Then, $(\omega^T \otimes I_n)N(\lambda) = M(\lambda)R(\lambda)$. By definition, $u \in \ker_\lambda L(z)$ if $u = N(\lambda)s$ for some constant vector s. Thus, $(\omega^T \otimes I_n)u = (\omega^T \otimes I_n)N(\lambda)s = M(\lambda)(R(\lambda)s) \in \ker_\lambda P(z)$. Since Ω is clearly linear, it only remains to prove its surjectivity. Any $h \in \ker_\lambda P(z)$ can be written as $h = M(\lambda)f$ for some constant vector f. Thus, $h = v(\lambda)M(\lambda)(f/v(\lambda)) = (\omega^T \otimes I_n)(V(\lambda) \otimes M(\lambda))(f/v(\lambda))$, which proves the surjectivity because $(V(\lambda) \otimes M(\lambda))(f/v(\lambda)) \in \ker_\lambda L(z)$. This last step follows from Lemma 6.1-3, which implies $V(z) \otimes M(z) = N(z)T(z)$, for some polynomial matrix $T(z)$. Therefore, $(V(\lambda) \otimes M(\lambda))(f/v(\lambda)) = N(\lambda)(T(\lambda)(f/v(\lambda))) \in \ker_\lambda L(z)$.

Item 2. Lemma 6.1-1 with z replaced by λ immediately implies that Ω maps any vector $u \in \ker L(\lambda)$ into a vector in $\ker P(\lambda)$, because $0 = (V(\lambda)^T \otimes I_m)L(\lambda)u = P(\lambda)(\omega^T \otimes I_n)u$. To prove surjectivity, write any $h \in \ker P(\lambda)$ as $h = v(\lambda)(h/v(\lambda)) = (\omega^T \otimes I_n)(V(\lambda) \otimes I_n)(h/v(\lambda))$. Note that Lemma 6.1-2 with z replaced by λ implies that $L(\lambda)(V(\lambda) \otimes I_n)(h/v(\lambda)) = (\omega \otimes I_m)P(\lambda)(h/v(\lambda)) = 0$, i.e., $(V(\lambda) \otimes I_n)(h/v(\lambda)) \in \ker L(\lambda)$.

Item 3. Theorem 1.1 implies that λ is an eigenvalue of $P(z)$ if and only if λ is an eigenvalue of $L(z)$. If λ is not an eigenvalue, then $\ker_\lambda P(z) = \ker P(\lambda)$ and $\ker_\lambda L(z) = \ker L(\lambda)$ and there is nothing to prove concerning the equality of the kernels. Thus, we assume that λ is an eigenvalue of $P(z)$ and of $L(z)$ and denote the maps in items 1 and 2 by Ω_1 and Ω_2, respectively. Theorem 1.1 implies that $\dim \ker_\lambda P(z) = p$ if and only if $\dim \ker_\lambda L(z) = pk$ and that $\dim \ker P(\lambda) = p+t$ if and only if $\dim \ker L(\lambda) = pk + t$. The rank-nullity theorem and the surjectivity of Ω_1 and Ω_2 imply that $\dim \ker \Omega_1 = \dim \ker \Omega_2 = p(k - 1)$. Moreover, $\ker_\lambda L(z) \subset \ker L(\lambda)$ implies $\ker \Omega_1 \subseteq \ker \Omega_2$ with strict equality by dimensional reasons. The stated isomorphisms follow from the first isomorphism theorem. $\square$

The third isomorphism theorem [6, Theorem 6.18] and item 3 of Proposition 6.8 immediately imply the following corollary.

Corollary 6.9. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z)$ be a scalar ansatz polynomial of grade $k - 1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$ are disjoint and fix any $\lambda \in \mathbb{F}$ such that $v(\lambda) \neq 0$. Then, $\ker P(\lambda)/\ker_\lambda P(z) \cong \ker L(\lambda)/\ker_\lambda L(z)$.*

Our proof of the isomorphism in Corollary 6.9 is abstract in the sense that it reduces to observing that the statement is a special case of deep and basic algebraic results; still, it can be also constructive in the sense of explicitly exhibiting the relevant vector space isomorphism due to items 1 and 2 of Proposition 6.8. Namely, the latter is given by the map

$$\Omega : \ker L(\lambda)/\ker_\lambda L(z) \to \ker P(\lambda)/\ker_\lambda P(z), \qquad [u] \mapsto [(\omega^T \otimes I_n)u]. \tag{21}$$

Finally, observe that, similarly to Theorem 6.5, one can prove that the kernels of the vector space homomorphisms in items 1 and 2 of Proposition 6.8 are both equal to $\operatorname{span} C$.

6.3 Recovery of root polynomials of $P(z)$ from those of $\mathbb{DL}(P, v)$

Finally, in this section we prove that the same approach used in the previous subsections for right minimal bases and right eigenvectors (that we can informally describe as "left multiply

31

times $\omega^T \otimes I_n$") allows us also to recover right root polynomials. For a pencil, maximal sets of root polynomials can be computed starting from the staircase form of the pencil [37]. Therefore, the result in this subsection can be used in practice to compute maximal sets of root polynomials for $P(z)$. For brevity, we focus on root polynomials at finite eigenvalues.

Theorem 6.10. *Let $P(z) \in \mathbb{F}[z]^{m \times n}$ be a matrix polynomial of grade $k \geq 2$, $v(z) = \omega^T V(z)$ be a scalar ansatz polynomial of grade $k-1$, where $V(z) \in \mathbb{F}[z]^{k \times 1}$ is the Vandermonde vector of degree $k-1$, and $L(z) = \mathbb{DL}(P, v) \in \mathbb{F}[z]^{km \times kn}$. Suppose that the sets of eigenvalues of $P(z)$ and of roots of $v(z)$, each including possibly infinite eigenvalues or roots, are disjoint. Let $\{\rho_i(z)\}_{i=1}^t$ be a maximal set of root polynomials at $\lambda \in \mathbb{F}$ for $L(z)$ and define $r_i(z) = (\omega^T \otimes I_n)\rho_i(z)$ for $i = 1, \ldots, t$. Then, $\{r_i(z)\}_{i=1}^t$ is a maximal set of root polynomials at λ for $P(z)$.*

Proof. Let $p = \dim \ker P(z)$. Then, $pk = \dim \ker L(z)$ by Theorem 1.1. Let $N(z) \in \mathbb{F}[\lambda]^{nk \times pk}$ be a minimal basis of $L(z)$ and let $M(z) \in \mathbb{F}[\lambda]^{n \times p}$ be a minimal basis of $P(z)$ obtained by applying Theorem 6.3 to $N(z)$ with M_b a maximal set of linearly independent columns of M_c (in case $M_c \neq 0$). Thus, $M(\lambda)$ is formed by a subset of the columns of $\widehat{M}(z) = (\omega^T \otimes I_n)N(z)$. By definition, the columns of $N(\lambda)$ form a basis of $\ker_\lambda L(z)$ and the columns of

$$N_{ext}(\lambda) = \begin{bmatrix} N(\lambda) & \rho_1(\lambda) & \cdots & \rho_t(\lambda) \end{bmatrix} \in \mathbb{F}^{nk \times (pk+t)}$$

form a basis of $\ker L(\lambda)$. From Proposition 6.8, the columns of $\widehat{M}(\lambda) = (\omega^T \otimes I_n)N(\lambda)$ span $\ker_\lambda P(z)$ and the ones of $(\omega^T \otimes I_n)N_{ext}(\lambda)$ span $\ker P(\lambda)$. But those columns of $\widehat{M}(\lambda)$ corresponding to $M(\lambda)$ also span $\ker_\lambda P(z)$, which implies that the columns of

$$\begin{bmatrix} M(\lambda) & r_1(\lambda) & \cdots & r_t(\lambda) \end{bmatrix} \in \mathbb{F}^{n \times (p+t)}$$

span $\ker P(\lambda)$. Theorem 1.1 guarantees that $\dim \ker P(\lambda) = p + t$. Therefore, the columns of $\begin{bmatrix} M(\lambda) & r_1(\lambda) & \cdots & r_t(\lambda) \end{bmatrix}$ form a basis of $\ker P(\lambda)$, which immediately implies that $r_i(\lambda) \notin \ker_\lambda P(z)$ for $i = 1, \ldots, t$. Therefore, it only remains to prove that $r_1(z), \ldots, r_t(z)$ are root polynomials at λ for $P(z)$ with the same orders as $\rho_1(z), \ldots, \rho_t(z)$, since by Theorem 1.1 these orders are the partial multiplicities of λ as an eigenvalue of $P(z)$ and Theorem 2.2 would imply the maximality of $\{r_i(z)\}_{i=1}^t$. For this purpose, note that from the equations $L(z)\rho_i(z) = (z-\lambda)^{\ell_i} s_i(z)$ with $s_i(\lambda) \neq 0$, corresponding to the fact that $\{\rho_i(z)\}_{i=1}^t$ are root polynomials at λ for $L(z)$ of orders $\{\ell_i\}_{i=1}^t$, and from Lemma 6.1-1, we obtain

$$(z - \lambda)^{\ell_i}(V(z)^T \otimes I_m)s_i(z) = (V(z)^T \otimes I_m)L(z)\rho_i(z) = P(z)(\omega^T \otimes I_n)\rho_i(z)$$

and $(V(\lambda)^T \otimes I_m)s_i(\lambda) \neq 0$. Thus, $r_i(z) = (\omega^T \otimes I_n)\rho_i(z)$ is a root polynomial at λ for $P(z)$ of order ℓ_i, for $i = 1, \ldots, t$. $\square$

Remark 6.11. The recovery procedure for root polynomials described in Theorem 6.10 is, of course, still valid when $P(z)$ is regular. In that case, root polynomials can be seen as generating functions of Jordan chains [13]. Thus, this subsection also describes, as a special case, how to recover Jordan chains for a regular $P(z)$ from those of $\mathbb{DL}(P, v)$ under the eigenvalue exclusion assumption. Note that in the regular case, the pencils in $\mathbb{DL}(P)$ satisfying the eigenvalue exclusion condition are linearizations for $P(z)$, then they are also $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ linearizations [29] and one can also recover maximal sets of root polynomials through a simple block extraction, as it was described in item 1 of [13, Theorem 8.10]. In fact, for $P(z)$ regular, Theorem 6.10 follows from item 2 of [13, Theorem 8.10].

7   Conclusions

We have extended the eigenvalue exclusion theorem for $\mathbb{DL}(P)$ pencils to the case of a singular matrix polynomial $P(z)$. Even if none of the pencils in $\mathbb{DL}(P)$ is a linearization for a singular $P(z)$, we have proved that, under the eigenvalue exclusion theorem assumptions, it is possible from any pencil $\mathbb{DL}(P, v)$ to fully recover all of the relevant magnitudes of $P(z)$, including eigenvalues and their partial multiplicities, (left and right) eigenvectors, (left and right) root polynomials, (left and right) minimal indices, and (left and right) minimal bases. With the exception of the recovery of the minimal indices and bases, this was already known when $P(z)$ is regular, but the arguments for a singular $P(z)$ are much more involved since they cannot be based on the (no longer true) fact that $\mathbb{DL}(P, v)$ is a strong linearization.

Our analysis also raises a question that goes beyond the study of the space of pencils $\mathbb{DL}(P)$. Does one really need strong linearizations to compute the eigenstructure of a matrix polynomial? Or, is any linear pencil that allows us to recover all the relevant quantities possibly as good as a tool? The case study of $\mathbb{DL}(P)$ definitely suggests that strictly restricting to strong linearizations may be an unnecessarily rigid approach. This is also supported by the new concept of "strongly minimal linearizations" introduced in [14].

An open problem in this context is to study what happens when the hypotheses of the generalized eigenvalue exclusion theorem are not satisfied. Based on preliminary analyses and also on the results in [14], we conjecture that a large part of Theorem 1.1 will remain valid, but that those eigenvalues of $P(z)$ that are roots of the ansatz polynomial will require of a different analysis and will satisfy different recovery rules for their partial multiplicities.

References

[1] B. ANDERSON AND E. JURY, *Generalized Bezoutian and Sylvester matrices in multivariable linear control*, IEEE Trans. Autom. Control 21(4), 551–556, 1976.

[2] E.N. ANTONIOU AND S. VOLOGIANNIDIS, *A new family of companion forms of polynomial matrices*, Electron. J. Linear Algebra 11, 78–87, 2004.

[3] T. BETCKE, N.J. HIGHAM, V. MEHRMANN, C. SCHRÖDER AND F. TISSEUR, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software 39(2), 1-28, 2013.

[4] M.I. BUENO, K. CURLETT, AND S. FURTADO, *Structured strong linearizations from Fiedler pencils with repetition I*, Linear Algebra Appl. 460, 51–80, 2014.

[5] M.I. BUENO, J. PÉREZ, AND S. ROGERS, *On why using $\mathbb{DL}(P)$ for the symmetric polynomial eigenvalue problem might need to be reconsidered*, Calcolo 59(4), 1–46, 2022.

[6] S. BURRIS AND H.P. SANKAPPANAVAR, *A Course in Universal Algebra*, Springer, New York, NY, USA, 1981.

[7] B. DAS AND S. BORA, *Vector spaces of generalized linearizations for rectangular matrix polynomials*, Electron. J. Linear Algebra 35, 116–155, 2019.

[8] F. DE TERÁN, F.M. DOPICO, AND D.S. MACKEY, *Linearizations of singular matrix polynomials and the recovery of minimal indices*, Electron. J. Linear Algebra 18, 371–402, 2009.

[9] 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, 2181–2204, 2010.

[10] F. De Terán, F.M. Dopico, and D.S. Mackey, *Fiedler companion linearizations for rectangular matrix polynomials*, Linear Algebra Appl. 437, 957–991, 2012.

[11] F. De Terán, F.M. Dopico, and D.S. Mackey, *Spectral equivalence of matrix polynomials and the index sum theorem*, Linear Algebra Appl. 459, 264–333, 2014.

[12] F.M. Dopico, P.W. Lawrence, J. Pérez, and P. Van Dooren, *Block Kronecker linearizations of matrix polynomials and their backward errors*, Numer. Math. 140, 373–426, 2018.

[13] F.M. Dopico and V. Noferini, *Root polynomials and their role in the theory of matrix polynomials*, Linear Algebra Appl. 584, 37–78, 2020.

[14] F.M. Dopico, M.C. Quintana, and P. Van Dooren, *Strongly minimal self-conjugate linearizations for polynomial and rational matrices*, SIAM J. Matrix Anal. Appl. 43(3), 1354–1381, 2022.

[15] H. Fassbender and P. Saltenberger, *On vector spaces of linearizations for matrix polynomials in orthogonal bases*, Linear Algebra Appl. 525, 59–83, 2017.

[16] G.D. Forney Jr., *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. Control 13, 493–520, 1975.

[17] F.R. Gantmacher, *The Theory of Matrices. Vols. 1, 2*, Chelsea Publishing Co., New York, 1959.

[18] S.D. Garvey, F. Tisseur, and S. Wang, *A parametrization of structure-preserving transformations for matrix polynomials*, Manchester Institute for Mathematical Sciences, MIMS EPrint: 2022.12.

[19] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, PA, USA, 2009, (unabridged republication of book first published by Academic Press in 1982).

[20] N.J. Higham, *Accuracy and Stability of Numerical Algorithms, 2nd Edition*, SIAM, Philadelphia, PA, USA, 2002.

[21] N.J. Higham, R.-C. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl. 29(4), 1218–1241, 2007.

[22] N.J. Higham, D.S. Mackey, N. Mackey, and F. Tisseur, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl. 29(1), 143–159, 2006.

[23] N.J. Higham, D.S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl. 28(4), 1005–1028, 2006.

[24] D. Kressner and I. Šain Giblić, *Singular quadratic eigenvalue problems: Linearization and weak condition numbers*, Preprint, https://arxiv.org/pdf/2204.07424.pdf

[25] L. Lerer and M. Tismenetsky, *The Bezoutian and the eigenvalue-separation problem for matrix polynomials*, Integr. Equ. Oper. Theory 5(1), 386–445, 1982.

[26] L. Lerer and M. Tismenetsky, *Generalized Bezoutian and the inversion problem for block matrices, I. General scheme*, Integr. Equ. Oper. Theory 9(6), 790–819, 1986.

[27] M. LOTZ AND V. NOFERINI, *Wilkinson's bus: Weak condition numbers, with an application to singular polynomial eigenproblems*, Found. Comput. Math. 20(6), 1439–1473, 2020.

[28] D.S. MACKEY, *Minimal indices and minimal bases via filtrations*, Electron. J. Linear Algebra 37, 276–294, 2021.

[29] D.S. MACKEY, N. MACKEY, C. MEHL, AND V. MERHMANN, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl. 28(4), 971–1004, 2006.

[30] 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(4), 1029–1051, 2006.

[31] D.S. MACKEY, N. MACKEY, C. MEHL, AND V. MERHMANN, *Möbius transformations of matrix polynomials*, Linear Algebra Appl. 470, 120–184, 2015.

[32] C. MOLER AND G.W. STEWART, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal. 10(2), 241–256, 1973.

[33] Y. NAKATSUKASA, V. NOFERINI, AND A. TOWNSEND, *Computing the common zeros of two bivariate functions via Bézout resultants*, SIAM Numer. Math. 129(1), 181–209, 2015.

[34] Y. NAKATSUKASA, V. NOFERINI, AND A. TOWNSEND, *Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach*, SIAM J. Matrix Anal. Appl. 38(1), 1–29, 2017.

[35] V. NOFERINI, *The behaviour of the complete eigenstructure of a polynomial matrix under a generic rational transformation*, Electron. J. Linear Algebra 23, 607–624, 2012.

[36] V. NOFERINI AND J. PÉREZ, *Fiedler-comrade and Fiedler–Chebyshev pencils*, SIAM J. Matrix Anal. Appl. 37, 1600–1624, 2016.

[37] V. NOFERINI AND P. VAN DOOREN, *On computing root polynomials and minimal bases of matrix pencils*, Linear Algebra Appl. 658, 86-115, 2023.

[38] V. NOFERINI AND P. VAN DOOREN, *Root vectors of polynomial and rational matrices: theory and computation*, Linear Algebra Appl. 656, 510-540, 2023.

[39] A.D. SOKAL, *A remark on the enumeration of rooted labeled trees*, Discrete Mathematics 343, 111865, 2020.

[40] J. STOER AND R. BULIRSCH, *Introduction to Numerical Analysis*, 2nd Edition, Springer-Verlag, New York, 1993.

[41] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, SIAM Rev. 43(2), 235–286, 2001.

[42] P. VAN DOOREN, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl. 27, 103–141, 1979.

[43] P. VAN DOOREN, *The generalized eigenstructure problem in linear system theory*, IEEE Trans. Aut. Contr. 26(1), 111–129, 1981.