Spectral Equivalence of Matrix Polynomials and the Index Sum Theorem

Fernando De Terán* Froilán M. Dopico† D. Steven Mackey‡

May 16, 2014

Abstract

The concept of linearization is fundamental for theory, applications, and spectral computations related to matrix polynomials. However, recent research on several important classes of structured matrix polynomials arising in applications has revealed that the strategy of using linearizations to develop structure-preserving numerical algorithms that compute the eigenvalues of structured matrix polynomials can be too restrictive, because some structured polynomials do not have any linearization with the same structure. This phenomenon strongly suggests that linearizations should sometimes be replaced by other low degree matrix polynomials in applied numerical computations. Motivated by this fact, we introduce equivalence relations that allow the possibility of matrix polynomials (with coefficients in an arbitrary field) to be equivalent, with the same spectral structure, but have different sizes and degrees. These equivalence relations are directly modeled on the notion of linearization, and consequently inherit the simplicity, applicability, and most relevant properties of linearizations; simultaneously, though, they are much more flexible in the possible degrees of equivalent polynomials. This flexibility allows us to define in a unified way the notions of quadratification and ℓ-ification, to introduce the concept of companion form of arbitrary degree, and to provide concrete and simple examples of these notions that generalize in a natural and smooth way the classical first and second Frobenius companion forms. The properties of ℓ-ifications are studied in depth; in this process a fundamental result on matrix polynomials, the "Index Sum Theorem", is recovered and extended to arbitrary fields. Although this result is known in the systems theory literature for real matrix polynomials, it has remained unnoticed by many researchers. It establishes that the sum of the (finite and infinite) partial multiplicities, together with the (left and right) minimal indices of any matrix polynomial is equal to the rank times the degree of the polynomial. The "Index Sum Theorem" turns out to be a key tool for obtaining a number of significant results: on the possible sizes and degrees of ℓ-ifications and companion forms, on the minimal index preservation properties of companion forms of arbitrary degree, as well as on obstructions to the existence of structured companion forms for structured matrix polynomials of even degree. This paper presents many new results, blended together with results already known in the literature but extended here to the most general setting of matrix polynomials of arbitrary sizes and degrees over arbitrary fields. Therefore we have written the paper in an expository and self-contained style that makes it accessible to a wide variety of readers.

*Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (`fteran@math.uc3m.es`). Supported by Ministerio de Ciencia e Innovación of Spain through grant MTM-2009-09281, and by Ministerio de Economía y Competitividad of Spain through grant MTM2012-32542.

†Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (`dopico@math.uc3m.es`). Supported by Ministerio de Ciencia e Innovación of Spain through grant MTM-2009-09281, and by Ministerio de Economía y Competitividad of Spain through grant MTM2012-32542.

‡Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA (`steve.mackey@wmich.edu`). Supported by National Science Foundation grant DMS-1016224, and by Ministerio de Economía y Competitividad of Spain through grant MTM2012-32542.

Key words. matrix polynomial, matrix pencil, linearization, quadratification, ℓ-ification, singular, regular, unimodular equivalence, spectral equivalence, elementary divisors, partial multiplicity sequence, structural indices, minimal indices, index sum theorem, companion forms, structured matrix polynomials.

AMS subject classification. 65F15, 15A18, 15A21, 15A22, 15A54.

1 Introduction

It is well known that two matrix polynomials of different sizes and degrees can have the same spectral structure, i.e., the same finite and infinite elementary divisors. The classical example of this phenomenon is a polynomial $P(\lambda)$ and any of its Frobenius or Fiedler companion linearizations [2, 10, 19, 20]. Using linearizations has long been the traditional means of solving polynomial eigenvalue problems [20, 58], but recently the conversion of a matrix polynomial into a quadratic polynomial with the same spectral structure (to be called a *quadratification*) has been proposed as a viable alternative method in certain situations [21, 23, 40]. The main motivation for considering quadratifications instead of linearizations stems from the fact that many matrix polynomials arising in applications have particular structures which impose symmetries in the spectrum of the polynomial. These symmetries are essential in applications and should be preserved in numerical computations in finite-precision arithmetic. Therefore the development of structure-preserving algorithms for computing eigenvalues of structured matrix polynomials has become a hot area of research in numerical linear algebra in recent years and, as a consequence, considerable effort has been invested in devising linearizations of matrix polynomials that preserve any structure that the polynomial might possess [2, 4, 11, 22, 37, 42]. However, this research has also shown that in certain important classes of structured matrix polynomials there always exist polynomials $P(\lambda)$ of even degree which do not have any linearization in the same structure class [43, 44, 45]. This fact strongly motivates the consideration of quadratifications and, possibly, other low degree matrix polynomials with the same spectral structure as $P(\lambda)$.

In this context, the main theme of this paper is to further explore the possibilities of matrix polynomials of different sizes and degrees having the same spectral structure in the most general possible setting, that is, matrix polynomials that may be square or rectangular, regular or singular, with coefficients over arbitrary fields. This type of question has been studied before in the literature but, as far as we know, with different goals and motivations than the ones considered in this work. The approach of this paper relies on a simple idea: to generalize as directly as possible the classical definition of linearization [20] from degree one to other degrees ℓ, with the aim of defining low degree polynomials (called ℓ-ifications) that share the simplicity, the applicability for numerical computations, and many other of the relevant properties that linearizations possess. Moreover, we introduce explicit examples of such ℓ-ifications that are easily constructed from the coefficients of any matrix polynomial without any matrix operations. These examples generalize in a natural way the classical first and second Frobenius linearizations [20].

Previous work related to the issues considered in this paper include the study of the theory of isospectral systems, which has attracted much attention recently in connection with the analysis of vibrating systems [18, Chapter 7],[31, 50]. This theory has usually been restricted to regular polynomials, and so is much more particular than the one introduced in this work. Earlier relevant results on matrix polynomials of different sizes and degrees having the same spectral structure have also appeared in the systems and control literature (see [27, 28, 51] and the references therein). However, the main motivation of this work is to investigate when two systems represented by different matrix polynomials share properties that are important in control, e.g., yielding the same transfer function. As a consequence, these works introduce relations between matrix polynomials that are useful and meaningful in systems and control theory, but that apparently have nothing to do with the classic notion of linearization, and hence are not so easy for researchers in numerical and applied linear algebra to use.

Since this paper deals with general matrix polynomials, we need to carefully analyze certain data that are not considered in the study of regular polynomials. Singular matrix polynomials, unlike regular polynomials, possess, in addition to finite and infinite elementary divisors, another important type of spectral-like data, the left and right minimal indices. These play an important role in a number of applications in systems and control theory [14, 26, 52, 53], but they are not preserved in general by linearizations, in contrast to elementary divisors. Therefore it is natural to consider the following questions, which provide some of the main motivation for this paper:

> For two singular matrix polynomials with the same spectral structure, what are the possible relationships between their minimal indices? In particular, is it possible to preserve *both* the spectral structure *and* the minimal indices of a singular polynomial in a polynomial of lower degree, for instance, in a linearization?

Recent work in [9, 10, 12] provides answers to these questions for the special cases of a singular polynomial P and its linearizations in the pencil spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ introduced in [41], as well as for its Frobenius and Fiedler companion linearizations introduced in [2]. We aim to extend this inquiry to much more general situations in this paper; in the process we rescue from the literature a fundamental result on matrix polynomials that we term the *Index Sum Theorem*. This result was presented in [49] for matrix polynomials over the real field, and is proven here over arbitrary fields. Unfortunately, the Index Sum Theorem has not received as much attention in the linear algebra community as it deserves. We will show that it is an important and easy-to-use tool that allows us to address several significant issues in a straightforward manner: preservation of the minimal indices of a polynomial by lower degree polynomials with the same spectral structure, determination of the possible sizes and degrees of ℓ-ifications, and the non-existence of structured linearizations of certain structured matrix polynomials of even degree.

The discussion so far indicates that this work includes many new results motivated by and closely linked to results already known in the literature, but extended here to general matrix polynomials over arbitrary fields. Since we are introducing concepts that we think will be useful for many researchers in matrix polynomials, the paper has been written in an expository style, presenting the results in a self-contained, detailed, and unified way, in order to make the reading of this paper as easy as possible, and to facilitate future references.

We begin in Section 2 by reviewing the preliminary concepts that are needed in this work, then continue in Sections 3 and 4 by introducing and analyzing the basic properties of a new equivalence relation among matrix polynomials that we call *spectral equivalence*. This relation is modeled on, and generalizes, the classical notion of linearization of a matrix polynomial, which we take as the prototype example of two matrix polynomials of different sizes and degrees having very closely related spectral structure. This is in keeping with the underlying point of view of the authors, arising out of the needs, concerns and goals of numerical linear algebra, where the use of linearizations has been a dominant paradigm for dealing with the polynomial eigenproblem. We will see that spectral equivalence is in several senses "intermediate" between the well-known notions of strict equivalence and unimodular equivalence, and also is more flexible to use than either strict or unimodular equivalence. In addition this new concept encompasses in a rigorous way the more recent notion of quadratification, which is currently being actively investigated by several research groups, and allows us to introduce ℓ-ifications, which generalize to arbitrary degree ℓ the concepts of linearization and quadratification. The relationship of this new equivalence relation to some other relations previously introduced by researchers in systems and control theory [27, 28, 51] will also be discussed.

Next, in Section 5 we introduce the definition of companion forms of arbitrary degree ℓ and provide concrete examples of this notion, denoted by $C_1^\ell(\lambda)$ and $C_2^\ell(\lambda)$, that generalize the classical Frobenius companion forms $C_1(\lambda)$ and $C_2(\lambda)$ (of degree one). This is one of the most important results presented in this paper. As a preliminary, we first carefully review the properties of the companion forms $C_1(\lambda)$ and $C_2(\lambda)$ for general (not necessarily square) matrix polynomials $P(\lambda)$ over arbitrary fields, showing that these are always linearizations, and deriving the general relationships between the minimal indices of $P(\lambda)$ and those of $C_1(\lambda)$ and $C_2(\lambda)$. These properties are well known for square polynomials [9, 10, 19, 20]; what is not so well known is the systematic extension to the case of rectangular matrix polynomials. Although these results have appeared previously for real matrix polynomials [47, 49], and can be viewed as special cases of the more general results in [12], we present the proofs here for the sake of completeness, as well as to emphasize their validity for matrix polynomials over arbitrary fields, and, more importantly, because these proofs are the basis for the proofs of the corresponding results for $C_1^\ell(\lambda)$ and $C_2^\ell(\lambda)$.

The results about the Frobenius companion linearizations for general matrix polynomials presented in Section 5 are then used in Section 6 to prove the *Index Sum Theorem for Matrix Polynomials*, a simple but fundamental relation among the elementary divisors, minimal indices, degree, and rank of any matrix polynomial. This result has appeared previously for real matrix polynomials [47, 49], but is not nearly as well known as is warranted by its fundamental nature. We give a proof here that demonstrates its validity for general matrix polynomials over arbitrary fields.

From the Index Sum Theorem follows a basic constraint on how the minimal indices of spectrally equivalent matrix polynomials can ever possibly be related to each other. We conclude in Section 7 by examining the impact of this constraint on a variety of important problems: the determination of the possible sizes and degrees of strong ℓ-ifications and companion forms of arbitrary degrees, the possible preservation of the minimal indices by companion forms of arbitrary degree, and the existence of structured linearizations of structured matrix polynomials.

Finally, we would like to emphasize several important features of this paper. First and foremost is the simplicity of the conceptual framework presented, and the resulting directness of the arguments used to prove basic properties, as compared to previous approaches to the issues considered here. Also noteworthy is the already mentioned generality of the results obtained: valid for all matrix polynomials, square or rectangular, regular or singular, over arbitrary fields, and allowing arbitrary choice of degree.

2 Preliminaries

Let $\mathbb{F}$ be an arbitrary field; then $\mathbb{F}[\lambda]$ denotes the ring of (scalar) polynomials with coefficients from $\mathbb{F}$, and $\mathbb{F}(\lambda)$ the field of rational functions with coefficients from $\mathbb{F}$. A polynomial $p(\lambda) \in \mathbb{F}[\lambda]$ is *identically-zero* if all of its coefficients are zero, and *non-identically-zero* if $p(\lambda)$ has at least one nonzero coefficient. A matrix polynomial

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \quad \text{with} \quad A_i \in \mathbb{F}^{m\times n} \tag{2.1}$$

is said to be *regular* if $m = n$ and $\det P(\lambda)$ is non-identically-zero, equivalently $P(\lambda)$ is regular if it is invertible when viewed as a matrix with entries in the field of rational functions $\mathbb{F}(\lambda)$. Otherwise, $P(\lambda)$ is said to be *singular* (note that this includes rectangular matrix polynomials with $m \neq n$). The *rank* of $P(\lambda)$, denoted $\operatorname{rank} P(\lambda)$, is the size of the largest non-identically-zero minor of $P(\lambda)$, equivalently the rank of $P(\lambda)$ when viewed as a single matrix with entries in the field $\mathbb{F}(\lambda)$. For simplicity, in many cases we drop the dependence of λ when referring to a matrix polynomial.

The matrix polynomial (2.1) is said to have *degree* k if $A_k \neq 0$. Since we wish to allow leading coefficient matrices of P to be zero, as well as allowing P to be singular, some care is needed in order to ensure that the multiplicities and elementary divisors of the eigenvalue at ∞ (see Definition 2.13) are well-defined. Consider, for example, the polynomial $P(\lambda) = A = \lambda 0_n + A$, with nonsingular $A \in \mathbb{F}^{n\times n}$ and where 0_n stands for the zero matrix of size $n \times n$. As we shall see in Example 2.16

and Lemma 2.17, the nature of the eigenvalue at ∞ for P depends on whether $P(\lambda)$ is being viewed as A, or as $\lambda 0_n + A$, or as $\lambda^2 0_n + \lambda 0_n + A$, ... Consequently it must be clarified whether this $P(\lambda)$ is to be viewed as a constant, a pencil, or possibly even as a matrix polynomial of higher "degree". This can be done by specifying the *grade* of $P(\lambda)$, an integer which is at least as large as the degree of P. For example, if $P(\lambda) = A$ is to be viewed as a constant, then we would say that $P(\lambda)$ has grade 0 as well as degree 0. But if $P(\lambda) = \lambda 0_n + A$ is to be viewed as a matrix *pencil*, then we would say that $P(\lambda)$ has grade 1, but still has degree 0. Certainly a polynomial of grade g can also be viewed as a polynomial of any grade higher than g, so the grade must be chosen; the grade of a matrix polynomial $P(\lambda)$ thus constitutes a feature of $P(\lambda)$ in addition to its degree. It does not replace the notion of degree, which retains its usual meaning as the largest m such that the matrix coefficient of λ^m in $P(\lambda)$ is *nonzero*. Throughout this paper, then, a matrix polynomial P must always be accompanied by a *choice* of grade, denoted grade(P). When the grade is not explicitly specified, then it is to be understood that *any* choice of grade$(P) \geq \deg(P)$ will suffice for the truth of the given statement.

Remark 2.1. Considered by Gohberg, Kaashoek, and Lancaster in [19], this notion of the grade of a matrix polynomial was called "extended degree" in [32]. Other recent work [44, 46] has revealed a number of situations where the extra flexibility afforded by the notion of grade leads to a clean and unified theory, whereas restricting to just degree can make results clumsy or very difficult to state. This is especially the case when a set of matrix polynomials rather than just a single individual matrix polynomial is involved. For example, in [46] the properties of Möbius transformations of matrix polynomials are extensively investigated. In this context grade$(P) = k$ indicates that the $m \times n$ matrix polynomial P is to be interpreted as an element of the $\mathbb{F}$-vector space of all matrix polynomials of degree less than or equal to k, equivalently, of all $m \times n$ matrix polynomials of grade k. It is shown in [46] that any Möbius transformation acts in a simple way on this vector space (it is a linear automorphism), even though it changes the degree of many individual matrix polynomials. Other simple properties are shown in [46] to hold when expressed in terms of appropriate grade choices, but fail if restricted to degree. One important use of Möbius transformations is the transferring of results about one class of structured polynomials to analogous results about a different structure class. A primary example of this is the correspondence between T-palindromic polynomials and T-alternating polynomials via a Möbius transformation [42]. It is shown in [46] how the results in [43] and [44] follow from each other in a simple and direct way, but only if everything is expressed in terms of appropriate grade choices. Some other contexts where we can expect grade to be a useful addition to the notion of degree are the study of matrix polynomials expressed in non-standard bases [1] such as a Bernstein or a Lagrange basis, and the perturbation theory of matrix polynomials with small leading term [25].

Example 2.2. The ordered list of coefficients of an $n \times n$ matrix polynomial of degree 1 and grade 3 has the form $(A_3, A_2, A_1, A_0) = (0_n, 0_n, A_1, A_0)$, with $A_1 \neq 0$.

Remark 2.3. Note that throughout the paper we reserve the term *pencil* to refer only to matrix polynomials of grade 1.

2.1 Smith form, partial multiplicity sequences, and elementary divisors

The canonical form of a matrix polynomial $P(\lambda)$ under transformation $E(\lambda)P(\lambda)F(\lambda)$ by unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ is referred to as the *Smith form* of $P(\lambda)$. This form was first developed for integer matrices by H.J.S. Smith [55] in the context of solving linear systems of Diophantine equations [36]. It was then extended by Frobenius in [15] to matrix polynomials; for a more modern treatment see, e.g., [16] or [35]. We use it here as a means to define the finite eigenvalues and associated elementary divisors of a general matrix polynomial $P(\lambda)$. Note that an $m \times m$ polynomial $E(\lambda)$ is said to be *unimodular* if $\det E(\lambda)$ is a *nonzero constant*, equivalently if $E(\lambda)$ has an inverse that is also a matrix polynomial.

Theorem 2.4 (Smith form (Frobenius, 1878)[15]).
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$, and unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ over $\mathbb{F}$ of size $m \times m$ and $n \times n$, respectively, such that

$$E(\lambda)P(\lambda)F(\lambda) = \operatorname{diag}(d_1(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)) =: D(\lambda), \tag{2.2}$$

where $d_i(\lambda) \in \mathbb{F}[\lambda]$, for $i = 1, \ldots, \min\{m, n\}$, $d_1(\lambda), \ldots, d_r(\lambda)$ are monic*, $d_{r+1}(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)$ are identically-zero, and $d_1(\lambda), \ldots, d_r(\lambda)$ form a* divisibility chain*, that is, $d_j(\lambda)$ is a divisor of $d_{j+1}(\lambda)$ for $j = 1, \ldots, r-1$. Moreover, the $m \times n$ diagonal matrix polynomial $D(\lambda)$ is unique, and the number r is equal to the rank of P.*

The nonzero diagonal elements $d_j(\lambda)$, $j = 1, \ldots, r$ in the Smith form $D(\lambda)$ are called the *invariant factors* or *invariant polynomials* of $P(\lambda)$.

Remark 2.5. The uniqueness of $D(\lambda)$ in Theorem 2.4 implies that the Smith form is *insensitive to field extensions*. In other words, suppose $P(\lambda)$ is a matrix polynomial over the field $\mathbb{F}$ and $\mathbb{F} \subseteq \widetilde{\mathbb{F}}$ is any field extension, so that $P(\lambda)$ can also be viewed as a matrix polynomial over the field $\widetilde{\mathbb{F}}$. Then the Smith forms of $P(\lambda)$ over $\mathbb{F}$ and over $\widetilde{\mathbb{F}}$ are identical. An important consequence of this insensitivity to field extensions is that the following notions of the partial multiplicity sequences, eigenvalues, and elementary divisors of $P(\lambda)$ are well-defined.

Definition 2.6 (Partial Multiplicity Sequences).
Let $P(\lambda)$ be an $m \times n$ matrix polynomial of rank r over a field $\mathbb{F}$. For any λ_0 in the algebraic closure $\overline{\mathbb{F}}$, the invariant polynomials $d_i(\lambda)$ of P, for $1 \leq i \leq r$, can each be uniquely factored as

$$d_i(\lambda) = (\lambda - \lambda_0)^{\alpha_i}\, p_i(\lambda) \quad \text{with} \quad \alpha_i \geq 0\,,\ p_i(\lambda_0) \neq 0\,. \tag{2.3}$$

The sequence of exponents $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ for any given $\lambda_0 \in \overline{\mathbb{F}}$ satisfies the condition $0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_r$ by the divisibility chain property of the Smith form, and is called the *partial multiplicity sequence of P at $\lambda_0 \in \overline{\mathbb{F}}$*.

Remark 2.7. It is common practice to refer only to *nonzero* exponents α_i occurring in (2.3) as *partial multiplicities*. However, note that in Definition 2.6 we allow any, even all, of the α_i's in a partial multiplicity sequence $\big(\alpha_1(\lambda_0), \alpha_2(\lambda_0), \ldots, \alpha_r(\lambda_0)\big)$ to be zero. Indeed, having $\alpha_i(\lambda_0) = 0$ for all $i = 1, \ldots, r$ occurs for all but a finite number of $\lambda_0 \in \overline{\mathbb{F}}$. These exceptional λ_0 with at least one nonzero exponent $\alpha_i(\lambda_0)$ are of course just the eigenvalues of $P(\lambda)$. Note that the numbers $\big(\alpha_1(\lambda_0), \alpha_2(\lambda_0), \ldots, \alpha_r(\lambda_0)\big)$ have also been called (see [47, 58]) the *structural indices of P at λ_0*.

Definition 2.8 (Eigenvalues and Elementary Divisors).
A scalar $\lambda_0 \in \overline{\mathbb{F}}$ is a (finite) *eigenvalue* of a matrix polynomial P whenever its partial multiplicity sequence $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ is not the zero sequence. The *elementary divisors* for an eigenvalue λ_0 of P are the collection of factors $(\lambda - \lambda_0)^{\alpha_i}$ with $\alpha_i \neq 0$, including repetitions. The *algebraic multiplicity* of an eigenvalue λ_0 is the sum $\alpha_1 + \alpha_2 + \cdots + \alpha_r$ of the terms in its partial multiplicity sequence, while the *geometric multiplicity* is the number of nonzero terms in this sequence.

It is worth noting that defining the eigenvalues of a matrix polynomial via the Smith form subsumes the more restrictive notion of eigenvalues as the roots of $\det P(\lambda)$, which is completely inadequate for singular matrix polynomials.

Remark 2.9. Observe that the Smith form of an $m \times n$ polynomial $P(\lambda)$ is completely (and uniquely) determined by three features of $P(\lambda)$: its size, its rank, and the elementary divisors of its finite eigenvalues.

Another quantity associated with a matrix polynomial $P(\lambda)$ that plays an important role later in this paper can also be conveniently defined from the Smith form of P.

Definition 2.10. Let $P(\lambda)$ be an $m \times n$ matrix polynomial over a field $\mathbb{F}$ with rank r and invariant polynomials $d_1(\lambda), d_2(\lambda), \ldots, d_r(\lambda)$ as in Theorem 2.4. Then $\delta_{\text{fin}}(P)$ is the sum of the degrees of all the invariant polynomials, i.e.,

$$\delta_{\text{fin}}(P) \;:=\; \sum_{i=1}^{r} \deg\bigl[d_i(\lambda)\bigr]\,. \tag{2.4}$$

Remark 2.11. Since the Smith form is insensitive to field extensions, then clearly so is the quantity $\delta_{\text{fin}}(P)$. Note also that there are a number of other equivalent ways to define $\delta_{\text{fin}}(P)$:

$$\begin{aligned} \delta_{\text{fin}}(P) \;&=\; \text{sum of the degrees of all (finite) elementary divisors of } P \\ &=\; \text{sum of the algebraic multiplicities of all (finite) eigenvalues of } P \\ &=\; \text{sum of all the (finite) structural indices of } P\,. \end{aligned}$$

Observe that these alternative descriptions of $\delta_{\text{fin}}(P)$ all require passing to the algebraic closure $\overline{\mathbb{F}}$, whereas Definition 2.10 describes $\delta_{\text{fin}}(P)$ more intrinsically, within the original field $\mathbb{F}$ itself.

Matrix polynomials may also have infinite eigenvalues, with a corresponding notion of elementary divisors at ∞. In order to define the elementary divisors at ∞ we need one more preliminary concept, that of the reversal of a matrix polynomial.

Definition 2.12 (j-reversal)**.**
Let $P(\lambda)$ be a nonzero matrix polynomial of degree $d \geq 0$. For $j \geq d$, the *j-reversal* of P is the matrix polynomial $\operatorname{rev}_j P$ given by

$$(\operatorname{rev}_j P)(\lambda) \;:=\; \lambda^j P(1/\lambda). \tag{2.5}$$

In the special case when $j = d$, the j-reversal of P is called the *reversal* of P and is sometimes denoted by just $\operatorname{rev} P$.

Definition 2.13 (Elementary divisors at ∞ / Structural indices at ∞)**.**
Let $P(\lambda)$ be a nonzero matrix polynomial of grade g and rank r. We say that $\lambda_0 = \infty$ is an eigenvalue of P if and only if 0 is an eigenvalue of $\operatorname{rev}_g P$, and the partial multiplicity sequence of P at $\lambda_0 = \infty$ is defined to be the same as that of the eigenvalue 0 for $\operatorname{rev}_g P$. If this partial multiplicity sequence is $(\alpha_1, \alpha_2, \ldots, \alpha_r)$, then for each $\alpha_i \neq 0$ we say there is an elementary divisor[1] of degree α_i for the eigenvalue $\lambda_0 = \infty$ of P. The numbers in the partial multiplicity sequence $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ are also known [47, 58] as the *structural indices of P at ∞*.

Remark 2.14. Let $P(\lambda) = \sum_{i=0}^{g} \lambda^i A_i$ be a matrix polynomial of grade g and rank r. Then P has an eigenvalue at ∞ if and only if the rank of the leading coefficient matrix A_g is strictly less than r. For a regular polynomial P this just means that A_g is singular. Observe that if $g > \deg P$, then $A_g = 0$ and P necessarily has r elementary divisors at ∞.

To the eigenvalue at ∞ we associate an important quantity, analogous to the number $\delta_{\text{fin}}(P)$ defined for the finite eigenvalues of $P(\lambda)$.

Definition 2.15. Suppose $P(\lambda)$ is a nonzero matrix polynomial of grade g, with partial multiplicity sequence $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ at ∞. Then

$$\delta_\infty(P) \;:=\; \alpha_1 + \alpha_2 + \cdots + \alpha_r \tag{2.6}$$

is the algebraic multiplicity of the eigenvalue at ∞, equivalently the sum of the degrees of the elementary divisors at ∞, or the sum of the structural indices of P at ∞.

[1] Such an elementary divisor is conventionally denoted by μ^{α_i}, although we will not use this notation anywhere in this paper.

Example 2.16. Let $A \in \mathbb{F}^{n\times n}$ be invertible. Then the constant polynomial $P(\lambda) = A$ of grade 0 has *no* eigenvalues at all — there are no finite eigenvalues because $P(\lambda)$ has the trivial Smith form I_n, and since $(\mathrm{rev}_0 P)(\lambda) = A$ also has Smith form I_n, P does not have ∞ as an eigenvalue either. On the other hand, when viewed as a matrix polynomial $P(\lambda) = \lambda 0_n + A$ of grade 1, we have $(\mathrm{rev}_1 P)(\lambda) = \lambda A + 0_n$ with Smith form λI_n, so $\mathrm{rev}_1 P$ has the eigenvalue 0 with partial multiplicity sequence $(1, 1, \dots, 1)$ and algebraic multiplicity n. Thus when viewed as a pencil, P still has no finite eigenvalues, but it does have the eigenvalue ∞ with $\delta_\infty(P) = n$. In general, if $P(\lambda) = A = \lambda^g 0_n + \cdots + \lambda 0_n + A$ is regarded as a polynomial of grade $g \ge 1$, then P has the eigenvalue ∞ with partial multiplicity sequence $(g, g, \dots, g)$, and hence $\delta_\infty(P) = gn$.

Example 2.16 illustrates how the infinite elementary divisors of P depend in an essential way on the grade chosen for P. More generally, the following result describes the effect of the choice of grade on the elementary divisor structure at ∞. The proof is straightforward and so is omitted.

Lemma 2.17. *Suppose $P(\lambda)$ is a matrix polynomial with rank r, with* $\operatorname{grade} P = \deg P = d$, *and with partial multiplicity sequence $(\alpha_1, \alpha_2, \dots, \alpha_r)$ at $\lambda_0 = \infty$, and hence $\delta_\infty(P) = \alpha_1 + \alpha_2 + \cdots + \alpha_r$. Then $P(\lambda)$ regarded as a matrix polynomial with grade $g > d$ has partial multiplicity sequence*

$$\big(\,\alpha_1 + (g-d)\,,\ \alpha_2 + (g-d)\,,\ \dots\,,\ \alpha_r + (g-d)\,\big)$$

at $\lambda_0 = \infty$, and hence $\delta_\infty(P) = (\alpha_1 + \alpha_2 + \cdots + \alpha_r) + r(g-d)$.

Definition 2.18 (Spectral Structure of a Matrix Polynomial)**.**
The collection of all the eigenvalues of a matrix polynomial $P(\lambda)$, both finite and infinite, is the *spectrum* of P. The collection of all the elementary divisors of P, both finite and infinite, including repetitions, constitutes the *spectral structure* of P. Equivalently, the spectral structure of P may be viewed as comprised of the spectrum of P together with all of its structural indices. The partial multiplicity sequences (structural indices) of the finite and infinite eigenvalues of P are also referred to, separately, as the *finite* and *infinite Jordan structures* of P.

2.2 Minimal Indices of Singular Polynomials.

From now on, we will denote by $\mathbb{F}(\lambda)^{m\times n}$ the vector space of $m \times n$ matrices with entries from the field of rational functions over $\mathbb{F}$. An $m \times n$ singular matrix polynomial $P(\lambda)$ has nontrivial *right* (*column*) and/or *left* (*row*) *null vectors*, that is, nonzero vectors $x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1}$ and/or $y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m}$ such that $P(\lambda)x(\lambda) \equiv 0$ and $y(\lambda)^T P(\lambda) \equiv 0$, where $y(\lambda)^T$ denotes the transpose of $y(\lambda)$. Equivalently, $P(\lambda)$ is singular when at least one of the subspaces

$$\begin{aligned} \mathcal{N}_r(P) &:= \left\{ x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1} \,:\, P(\lambda)x(\lambda) \equiv 0 \right\}, \\ \mathcal{N}_\ell(P) &:= \left\{ y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m} \,:\, y(\lambda)^T P(\lambda) \equiv 0^T \right\} \end{aligned}$$

is nontrivial. We will refer to these subspaces, respectively, as the *right* and *left nullspaces* of $P(\lambda)$.

A *vector polynomial* is a vector whose entries are polynomials in the variable λ. For any subspace of $\mathbb{F}(\lambda)^n$, it is always possible to find a basis consisting entirely of vector polynomials; simply take an arbitrary basis and multiply each vector by the denominators of its entries. The *degree* of a vector polynomial is the largest degree of its components, and the *order* of a polynomial basis is defined as the sum of the degrees of its vectors [14, p. 494]. Then the following definition picks out a natural class of polynomial bases.

Definition 2.19. (Minimal basis [14]).
Let $\mathcal{V}$ be a subspace of $\mathbb{F}(\lambda)^n$. A *minimal basis* of $\mathcal{V}$ is any polynomial basis of $\mathcal{V}$ with least order among all polynomial bases of $\mathcal{V}$.

It can be shown [14] that for any given subspace $\mathcal{V} \subseteq \mathbb{F}(\lambda)^n$, the ordered list of degrees of the vector polynomials in any minimal basis of $\mathcal{V}$ is always the same. These degrees are then called the *minimal indices* of $\mathcal{V}$. Specializing $\mathcal{V}$ to be the left and right nullspaces of a singular matrix polynomial gives the following definition, where $\deg\big(v(\lambda)\big)$ denotes the degree of the vector polynomial $v(\lambda)$.

Definition 2.20. (Minimal indices)
Let $P(\lambda)$ be a singular matrix polynomial over a field $\mathbb{F}$, and let the sets $\big\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\big\}$ and $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ be minimal bases of, respectively, the left and right nullspaces of $P(\lambda)$, ordered so that $0 \le \deg(y_1) \le \deg(y_2) \le \cdots \le \deg(y_q)$ and $0 \le \deg(x_1) \le \deg(x_2) \le \cdots \le \deg(x_p)$. Let $\eta_i = \deg(y_i)$ for $i = 1, \ldots, q$ and $\varepsilon_j = \deg(x_j)$ for $j = 1, \ldots, p$. Then the scalars $\eta_1 \le \eta_2 \le \cdots \le \eta_q$ and $\varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_p$ are, respectively, the *left* and *right minimal indices of* $P(\lambda)$.

The definitions given above are due to Forney [14], but there are several other ways to define minimal indices. The classical approach is due to Kronecker, and can be found in [16]; another approach uses the notion of a filtration of a vector space [39]. Arguments that these three approaches all produce the same values for minimal indices can be found in [7] and [39].

The sum of all the minimal indices of a given $P(\lambda)$ will be an important quantity for us later, so we introduce some convenient notation for it here.

Definition 2.21. Let $P(\lambda)$ be a singular matrix polynomial with minimal indices $\eta_1 \le \cdots \le \eta_q$ and $\varepsilon_1 \le \cdots \le \varepsilon_p$ as in Definition 2.20. Then

$$\mu(P) \;:=\; \sum_{i=1}^{q} \eta_i \;+\; \sum_{j=1}^{p} \varepsilon_j \tag{2.7}$$

denotes the sum of all the minimal indices of P. If P is regular, then we define $\mu(P) := 0$.

Note that $\mu(P) = 0$ may also occur with a singular matrix polynomial P in Definition 2.21. Observe also that $\mu(P)$ may be viewed as the sum of the *order* of a left minimal basis together with the *order* of a right minimal basis for P.

Definition 2.22 (Singular Structure of a Matrix Polynomial)**.**
The collection of all the minimal indices of a matrix polynomial $P(\lambda)$, both left and right, including repetitions, constitutes the *singular structure* of P. If P is regular, then the singular structure is empty.

2.3 Kronecker Canonical Form

For matrix pencils over algebraically closed fields, the canonical form under strict equivalence (see Definition 3.1) is the *Kronecker Canonical Form*, or KCF for short. This form is particularly valuable because it explicitly displays both the spectral structure and the singular structure of the pencil. We briefly recall this classical result for the convenience of the reader.

Theorem 2.23 (Kronecker Canonical Form [29])**.**
Any matrix pencil over an algebraically closed field $\mathbb{F}$ is strictly equivalent to a direct sum of blocks of the following types:

(i) $(\lambda - \lambda_0)I_k + N_k\,,$ *where* $N_k := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$ *is the standard $k \times k$ nilpotent matrix,*

(ii) $I_m + \lambda N_m\,,$

(iii) $S_d(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}_{d\times(d+1)}$ *or* $S_d^T(\lambda)$ *with* $d \geq 1$,

(iv) $0_{p\times q}$, *where* $p, q \geq 0$.

Note that each type (i) block records a finite elementary divisor of degree k, while each block of type (ii) records an infinite elementary divisor of degree m. By contrast, each type (iii) block $S_d(\lambda)$ (respectively, $S_d^T(\lambda)$) records a right (respectively, left) minimal index with value $d \geq 1$, while a type (iv) block records any zero minimal indices. For further details see [16, Ch. XII, Section 4].

Since this paper is developing theorems that hold for matrix polynomials over an *arbitrary field*, we will never be able to invoke the full strength of Theorem 2.23. Instead (in Section 6) we use only a weaker version of the KCF, stated in Lemma 6.2, that is valid for matrix pencils over arbitrary fields. Nevertheless, it will certainly be helpful for the reader to keep the KCF in mind as part of the background context for all the developments to follow.

For the sake of simplicity, all matrix polynomials in the rest of the paper will be assumed to have entries in an arbitrary field $\mathbb{F}$, unless specifically stated otherwise.

3 Equivalence Relations on Matrix Polynomials

3.1 Classical Equivalence Relations

Let us begin by recalling the two classical equivalence relations on matrix polynomials; these relations preserve all or part of the spectral structure and singular structure of matrix polynomials.

Definition 3.1 (Classical Equivalence Relations on Matrix Polynomials)**.**
Two $m \times n$ matrix polynomials P and Q over a fixed but arbitrary field $\mathbb{F}$ are said to be

(a) *unimodularly equivalent*, denoted $P \sim Q$, if there exist unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ over $\mathbb{F}$ such that $E(\lambda)P(\lambda)F(\lambda) = Q(\lambda)$,

(b) *strictly equivalent*, denoted $P \cong Q$, if there exist invertible (constant) matrices E and F over $\mathbb{F}$ such that $E \cdot P(\lambda) \cdot F = Q(\lambda)$.

Note that both of the equivalence relations in Definition 3.1 apply only to matrix polynomials P and Q of the *same size*, although unimodular equivalence does at least allow P and Q to have different degrees.

It is worthwhile comparing these equivalence relations in terms of the invariants associated with them. Since two matrix polynomials are unimodularly equivalent if and only if they have the same Smith form, we see from Remark 2.9 that the complete set of invariants for matrix polynomials with respect to unimodular equivalence consists of just the size, the rank, and the finite elementary divisors. By contrast, strict equivalence preserves the size, rank, and degree of matrix polynomials, all finite and infinite elementary divisors, as well as all left and right minimal indices. In other words, strict equivalence preserves both the spectral structure and singular structure of matrix polynomials.

As described in Section 2.3, the canonical form for matrix pencils (over an algebraically closed field) under strict equivalence is the Kronecker Canonical Form. However, there is no known canonical form for higher degree (or higher grade) matrix polynomials under strict equivalence. It is rather tempting to think that two matrix polynomials will be strictly equivalent if and only if they have the same size, rank, degree, and the same spectral structure and singular structure, since the KCF shows this to be the case for matrix pencils. Unfortunately this is not true. Consider, for instance, the quadratic matrix polynomials

$$P(\lambda) = \begin{bmatrix} \lambda^2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad Q(\lambda) = \begin{bmatrix} \lambda^2 & \lambda \\ 0 & 1 \end{bmatrix}.$$

Both $P(\lambda)$ and $Q(\lambda)$ have the same size, rank, and degree, the same finite and infinite elementary divisors, and the same minimal indices (none, because they are regular). But they are not strictly equivalent, since $Q(\lambda)$ has a nonzero first degree term and $P(\lambda)$ does not.

Unimodular equivalence does not in general preserve either the infinite elementary divisors or the (left or right) minimal indices. Indeed, it is shown in [33] that for regular pencils $L(\lambda)$ with singular leading term, unimodular equivalence leaves the quantity $\delta_\infty(L)$ unchanged, but can otherwise *arbitrarily alter* the number and degrees of the infinite elementary divisors.

3.2 New Equivalence Relations

The following two new equivalence relations allow the possibility of matrix polynomials of different sizes *and* different degrees being equivalent, motivated by (and modeled on) the well-known notions of linearization and strong linearization [19, 20].

Definition 3.2 (Two New Equivalence Relations)**.**
Suppose P and Q are two matrix polynomials, not necessarily of the same size or degree, with $g = \operatorname{grade} P$ and $h = \operatorname{grade} Q$.

(a) P and Q are said to be *extended unimodularly equivalent*, denoted $P \smile Q$, if for some $r, s \geq 0$ we have $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$.

(b) P and Q are said to be *spectrally equivalent*, denoted $P \asymp Q$, if $P \smile Q$ *and* $\operatorname{rev}_g P \smile \operatorname{rev}_h Q$.

Typically when using Definition 3.2 in practice we have either $r = 0$ or $s = 0$; however, we *allow* them both to be nonzero. Indeed, we will see later in Corollary 4.3 that $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$ holds for some nonzero r and s if and only if $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$ also holds with at least one of r or s being zero. Thus there is no loss or gain in generality in allowing both r and s to be nonzero, only a gain in convenience and flexibility.

Why have the relations in Definition 3.2 been given these particular names? The name for the relation $P \smile Q$ is completely natural, since it is just unimodular equivalence enhanced by the possibility of extending the size of a matrix polynomial by adjoining an identity. Justification for the name of the relation $P \asymp Q$ is provided by the results of Theorem 4.1, which characterize the relations $P \smile Q$ and $P \asymp Q$ in terms of data from the spectral and singular structures of P and Q.

The following definition establishes terminology for several important special cases of Definition 3.2, and in particular makes a connection back to the prototype examples (linearization and strong linearization) on which Definition 3.2 is based.

Definition 3.3 (Linearization, Quadratification, and ℓ-ification)**.**
Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade g.

(a) A matrix pencil $L(\lambda)$ is said to be a *linearization* of $P(\lambda)$ if $L(\lambda) \smile P(\lambda)$. A linearization is said to be strong if, in addition, $\operatorname{rev}_1 L(\lambda) \smile \operatorname{rev}_g P(\lambda)$. Equivalently, a pencil $L(\lambda)$ is a *strong linearization* for $P(\lambda)$ if

$$L(\lambda) \asymp P(\lambda)\,.$$

(b) A quadratic matrix polynomial $Q(\lambda)$, that is, a polynomial with $\operatorname{grade} Q = 2$, is said to be a *quadratification* of $P(\lambda)$ if $Q(\lambda) \smile P(\lambda)$. A quadratification is said to be strong if, in addition, $\operatorname{rev}_2 Q(\lambda) \smile \operatorname{rev}_g P(\lambda)$. Equivalently, a quadratic $Q(\lambda)$ is a *strong quadratification* for $P(\lambda)$ if

$$Q(\lambda) \asymp P(\lambda)\,.$$

(c) More generally, a matrix polynomial $R(\lambda)$ of grade ℓ is said to be an *ℓ-ification* of $P(\lambda)$ if $R(\lambda) \smile P(\lambda)$, and a *strong ℓ-ification* of $P(\lambda)$ if

$$R(\lambda) \asymp P(\lambda)\,.$$

The classical notions of linearization and strong linearization [19, 20, 33] are certainly included in Definition 3.3(a) as special cases. For a matrix polynomial $P(\lambda)$ of degree k and size $n \times n$, recall that the classical definition says that a linearization for $P(\lambda)$ is a pencil $L(\lambda)$ with the *specific fixed size $kn \times kn$* such that

$$L(\lambda) \sim \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_s \end{bmatrix} \tag{3.1}$$

with $s = (k-1)n$; $L(\lambda)$ is said to be a (classical) strong linearization if, in addition, we have $\mathrm{rev}_1 L(\lambda) \sim \mathrm{diag}[\,\mathrm{rev}_k P(\lambda),\, I_s\,]$, again with $s = (k-1)n$. Non-standard sizes for linearizations (i.e., using other values for s in (3.1)) have been considered in [6] and [8]. However, in [8] it was shown that if P is regular, then any strong linearization for P in the extended sense of (3.1) allowing non-classical values for s, can in fact *only* have the classical size $kn \times kn$. Thus we see that the classical definition is particularly well-suited for the regular case. For singular polynomials, though, allowing other values for s in (3.1) leads to many other viable sizes for linearizations. In [8], the smallest possible $s = s_{\min}$ for a given singular P is determined, and it is shown that every $s \geq s_{\min}$ will support a linearization for P.

However, we want to stress that the new Definition 3.3(a) extends not only the classical definition, but also the definition using eqn. (3.1) with non-standard values for s. No particular size for linearizations, quadratifications, or ℓ-ifications is specified in Definition 3.3; *any size that works* is allowed. Indeed, a linearization in the sense of Definition 3.3 may now even have a *smaller* size than $P(\lambda)$, as illustrated by the following example.

Example 3.4. Consider the regular quadratic matrix polynomial $P(\lambda) = \left[\begin{smallmatrix} \lambda & \lambda^2 \\ 0 & 1 \end{smallmatrix}\right]$ and the regular pencil $L(\lambda) = [\,\lambda\,]$. Then it is easy to see that

$$P(\lambda) \sim \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} L(\lambda) & 0 \\ 0 & 1 \end{bmatrix},$$

so $L(\lambda) \smile P(\lambda)$, and hence L is a linearization of P in the sense of Definition 3.3(a). Clearly, though, L is *not* a linearization of P in the extended classical sense of (3.1) with non-standard values for s, since any relation of the form (3.1) with $s \geq 0$ is impossible. Note that any pencil of the form $\widetilde{L}(\lambda) = \mathrm{diag}\big[L(\lambda), I_r\big]$ is also easily seen to be a linearization of P in the sense of Definition 3.3(a), so this regular quadratic P has a linearization of *every* possible size. On the other hand, we can also readily see that L is *not* a strong linearization of P, even in the sense of Definition 3.3. Since $\mathrm{rev}_1 L(\lambda) = [\,1\,]$ and $\mathrm{rev}_2 P(\lambda) = \left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda^2 \end{smallmatrix}\right]$, it is clear from taking determinants that no relation of the form $\mathrm{diag}[\,\mathrm{rev}_1 L, I_r\,] \sim \mathrm{diag}[\,\mathrm{rev}_2 P, I_s\,]$ could ever hold, so $\mathrm{rev}_1 L(\lambda) \not\smile \mathrm{rev}_2 P(\lambda)$, and hence $L(\lambda) \not\asymp P(\lambda)$.

We will show later in Theorem 4.11 that any *strong* linearization (in the sense of Definition 3.3) of a regular P can still *only* have the classical size, despite the freedom in size possible for ("weak") linearizations afforded by Definition 3.3, as illustrated by this example.

Remark 3.5. Observe that in Example 3.4 the finite elementary divisors of L (or the finite elementary divisors of any of the $\widetilde{L}$ pencils defined there) are exactly the same as those of P, as we would surely want to be the case for anything worthy of the name "linearization". This is not a coincidence, as we will see in Theorem 4.1, where we characterize the relations $P \smile Q$ and $P \asymp Q$ in terms of the spectral and singular structures of P and Q.

In Section 5 we will consider the Frobenius companion forms, which are the classical examples of (strong) linearizations. Other families of (strong) linearizations have been introduced in [2, 41, 59] for regular polynomials, and later extended in [3, 9, 10] to square singular polynomials and in [12] to rectangular polynomials. Still more families of strong linearizations are constructed in [4, 11], with the extra property of preserving T-palindromic structure. So in addition to being a well-known concept, linearizations are well-represented in the literature by a wide variety of concrete examples. This is not the case for quadratifications or ℓ-ifications, but this will be remedied in Section 5.2, where four families of simple examples illustrating these notions are presented and analyzed in Theorems 5.7 – 5.11. In particular, it will follow as an immediate consequence of Theorem 5.7 (or Theorem 5.8) that every matrix polynomial of even grade has a strong quadratification.

We close this section by stating two basic properties of the two new relations defined in Definition 3.2. The proofs of both follow in a completely straightforward manner from the definitions, and so are omitted.

Lemma 3.6 (Basic properties of $P \smile Q$ and $P \asymp Q$).
For matrix polynomials P and Q:

(a) *The relations $P \smile Q$ and $P \asymp Q$ are both* equivalence *relations.*

(b) *Let P be $m \times n$ and Q be $p \times q$, with $P \smile Q$. Then $p - m = q - n$, i.e., the row size difference between P and Q is the same as the column size difference.*

3.3 Some Other Equivalence Relations

There is a significant body of work in the systems and control theory literature investigating various equivalence relations on matrix polynomials of different sizes and degrees. The main goal of these relations is to guarantee that systems represented by equivalent matrix polynomials share properties that are important in control; e.g., systems represented by equivalent polynomials should have the same matrix transfer function [51, Theorem 4]. As a consequence of guaranteeing such properties, these relations also ensure that equivalent matrix polynomials have the same finite and/or infinite elementary divisors. However, the fact that these equivalence relations are aiming to capture "equivalence of systems" leads to them being expressed in a way that is difficult for researchers in numerical linear algebra to use; they tend not to be amenable for numerical computations, and do not resemble the familiar notion of linearization, which is fundamental in numerical and applied problems in matrix polynomials. We quote here from the classic paper [51], summarizing some of the difficulties with these equivalence relations:

> "The real disadvantage appears to be that extended strict system equivalence cannot be described completely in terms of elementary row and column operations."

This is in contrast with our approach in Definition 3.2 that is explicitly based on unimodular transformations, which are nothing but a sequence of elementary row and column operations on matrix polynomials [16]. This feature of extended unimodular equivalence and spectral equivalence will be used extensively in the proofs concerning the concrete ℓ-ifications presented in Section 5.

Much of the research on equivalence relations of matrix polynomials in systems and control theory stems from the work of Pugh and Shelton [51], who generalized, clarified, and summarized previous research on this topic. Among the relations introduced after [51] are strong equivalence, $\{s_0\}$-equivalence, factor equivalence, complete equivalence, full equivalence, fundamental equivalence, and divisor equivalence, to name a few; an overview of these equivalence relations and their inter-relationships can be found in [27, 28]. Since many of these relations are variations or enhanced versions of the equivalence relation introduced by Pugh and Shelton in [51], we focus first on that relation, and then on the one introduced later by Karampetakis and Vologiannidis in [27]. This second relation takes information about the infinite eigenvalue into account, something that is missing in the relation of Pugh and Shelton. Note that this is analogous to the situation with the concepts of extended unimodular equivalence and spectral equivalence. Before introducing these two relations in Definition 3.8, we need first to recall in Definition 3.7 some well-known notions from the literature (see, for instance, [54, Ch. 2, Sec. 6]). For the sake of brevity, the notation $A(\lambda)_{m\times n}$ is sometimes used to indicate that $A(\lambda)$ has size $m \times n$.

Definition 3.7 (Coprime matrix polynomials)**.**
Suppose $A(\lambda)_{m\times n}$ and $B(\lambda)_{m\times p}$ are matrix polynomials over a field $\mathbb{F}$, with the *same number of rows*. An $m \times m$ matrix polynomial $L(\lambda)$ over $\mathbb{F}$ is said to be a *left common divisor* of $A(\lambda)$ and $B(\lambda)$ if $A(\lambda) = L(\lambda)C(\lambda)$ and $B(\lambda) = L(\lambda)D(\lambda)$ for some matrix polynomials $C(\lambda)$ and $D(\lambda)$ over $\mathbb{F}$. Then $A(\lambda)$ and $B(\lambda)$ are said to be *left coprime* (or relatively left prime) if every left common divisor of $A(\lambda)$ and $B(\lambda)$ is unimodular.

(Analogous definitions for *right common divisor* and *right coprime* can be given for pairs of matrix polynomials with the *same number of columns*.)

Definition 3.8. Suppose $P(\lambda)_{m\times n}$ and $Q(\lambda)_{p\times q}$ are matrix polynomials such that $p - m = q - n$. Then P and Q are

(a) *extended unimodularly equivalent in the sense of Pugh and Shelton* ("PS-eue" for short) if there exist matrix polynomials $M(\lambda)_{p\times m}$ and $N(\lambda)_{q\times n}$ such that

$$M(\lambda)P(\lambda) \;=\; Q(\lambda)N(\lambda)\,,$$

where M, Q are left coprime, and P, N are right coprime.

(b) *strongly equivalent in the sense of Karampetakis and Vologiannidis* ("KV-se" for short) if the following two conditions hold:

(b1) P and Q are PS-eue, and

(b2) there are rational matrices $\widetilde{M}(\lambda)_{p\times m}, \widetilde{N}(\lambda)_{q\times n}$ with no poles at $\lambda = 0$, such that

$$\widetilde{M}(\lambda) \cdot \operatorname{rev} P(\lambda) \;=\; \operatorname{rev} Q(\lambda) \cdot \widetilde{N}(\lambda)\,,$$

where

$$\begin{bmatrix} \widetilde{M}(\lambda) & \operatorname{rev} Q(\lambda) \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} \operatorname{rev} P(\lambda) \\ -\widetilde{N}(\lambda) \end{bmatrix}$$

each have full rank at $\lambda = 0$.

For the definition of the *poles* of a rational matrix we refer the reader to [54, Ch. 3, Sec. 4]. The reader should bear in mind that matrix polynomials do not have any finite poles at all. Here we have stated the definition of PS-eue exactly as it appears in the original paper [51]. However, in [27] this notion is introduced in a slightly different (but equivalent) form. From the point of view of system theory it is important to note that the transformation induced by PS-eue on the class of polynomial system matrices is identical to the transformation of *extended strict system equivalence* [51, p. 666].

Despite the very different appearance of the pair of relations PS-eue and KV-se as compared to extended unimodular equivalence ($\smile$) and spectral equivalence ($\asymp$), it turns out, rather surprisingly, that they respectively partition the set of matrix polynomials into exactly the same equivalence classes; this will be shown in Section 4.2. Thus from a rigid set-theoretic point of view they are identical as pairs of equivalence relations. However, there are a number of features that distinguish these two pairs of relations from numerical, practical, and theoretical points of view, making $\smile$ and $\asymp$ much more convenient for applied and numerical linear algebra, while PS-eue and KV-se are more convenient for systems theory. Let us briefly mention some of these contrasting features. In the first place, to prove that $\smile$ and $\asymp$ are actually equivalence relations is completely straightforward from their definitions, but this is not the case for PS-eue and KV-se; indeed, to prove that PS-eue is an equivalence relation requires considerable effort [51]. A second distinguishing feature concerns the more "implicit" nature of the definitions of PS-eue and KV-se as compared to $\smile$ and $\asymp$. In particular, given a matrix polynomial $P(\lambda)$, it is very easy to explicitly construct many other polynomials $Q(\lambda)$ such that $P(\lambda) \smile Q(\lambda)$: begin by appending *any* I_r to P, then do *any* unimodular transformations (i.e., any sequence of elementary row and column operations) to $\operatorname{diag}(P, I_r)$ to obtain such a Q. By contrast, for PS-eue it is in general not so clear how to systematically generate even a single other polynomial Q that is PS-eue to a given P, since the allowed transformations to "convert" P into Q are not free, but are jointly constrained by both P and Q themselves. On the other hand, the definitions of extended unimodular and spectral equivalence do not allow us to check directly if two systems represented by equivalent matrix polynomials have, for instance, the same transfer function [51, Theorem 4].

4 Comparative Analysis of Equivalence Relations

4.1 Spectral Characterization of $P \smile Q$ and $P \asymp Q$

In this section we characterize the relations $P \smile Q$ and $P \asymp Q$ in terms of data from the spectral and singular structures of P and Q. This result extends the characterization of (classical) linearizations and strong linearizations of square matrix polynomials given in [9, Lemma 2.3] to the more general setting of the relations $P \smile Q$ and $P \asymp Q$. Note that a somewhat different criterion (based on ideas related to the "local Smith form") for a matrix pencil to be a (classical) linearization or strong linearization of a given regular matrix polynomial was provided in [32].

Theorem 4.1 (Characterization of $P \smile Q$ and $P \asymp Q$)**.**
Consider matrix polynomials $P(\lambda)_{m\times n}$ with grade g and $Q(\lambda)_{p\times q}$ with grade h, and the following three conditions on P and Q:

(a) $\dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(Q)$ *and* $\dim \mathcal{N}_\ell(P) = \dim \mathcal{N}_\ell(Q)$
(*i.e., P and Q have the same* number *of right* (*respectively, left*) *minimal indices*).

(b) *P and Q have exactly the same* finite *elementary divisors.*

(c) *P and Q have exactly the same* infinite *elementary divisors.*

Then:

(1) *$P \smile Q$ if and only if conditions* (a) *and* (b) *hold.*

(2) *$P \asymp Q$ if and only if conditions* (a), (b), *and* (c) *hold.*

Proof. (1): ($\Rightarrow$) Assuming that $P \smile Q$, then from Definition 3.2(a) we have

$$\dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(\operatorname{diag}(P, I_r)) = \dim \mathcal{N}_r(\operatorname{diag}(Q, I_s)) = \dim \mathcal{N}_r(Q)\,,$$

and similarly for the left nullspaces, so that condition (a) holds. That P and Q have the same finite elementary divisors follows from the equality of the Smith forms of $\operatorname{diag}(P, I_r)$ and $\operatorname{diag}(Q, I_s)$, so condition (b) holds.
(1): ($\Leftarrow$) We first show that condition (a) implies that $p - m = q - n$, thus guaranteeing the existence of an $r \geq 0$ and an $s \geq 0$ such that $m + r = p + s$ and $n + r = q + s$, i.e., such that $\operatorname{diag}(P, I_r)$ and $\operatorname{diag}(Q, I_s)$ have the same size. Observe that the equalities $\dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(Q)$ and $\dim \mathcal{N}_\ell(P) = \dim \mathcal{N}_\ell(Q)$ respectively imply

$$n - \operatorname{rank} P \;=\; q - \operatorname{rank} Q \qquad \text{and} \qquad m - \operatorname{rank} P \;=\; p - \operatorname{rank} Q\,,$$

and these in turn imply

$$p - m \;=\; \operatorname{rank} Q - \operatorname{rank} P \;=\; q - n\,.$$

Thus any $r, s \geq 0$ satisfying $r - s = p - m = q - n$ will make $\operatorname{diag}(P, I_r)$ and $\operatorname{diag}(Q, I_s)$ have the same size. With any such choice of r and s, condition (a) now further implies that $\operatorname{rank}\operatorname{diag}(P, I_r) = \operatorname{rank}\operatorname{diag}(Q, I_s)$. Together with condition (b), we see that $\operatorname{diag}(P, I_r)$ and $\operatorname{diag}(Q, I_s)$ have the same size, the same rank, and the same finite elementary divisors, and hence the same Smith form (see Remark 2.9). Consequently $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$, and hence $P \smile Q$.

(2): ($\Rightarrow$) If $P \asymp Q$, then a fortiori $P \smile Q$, so conditions (a) and (b) follow from part (1). Condition (c) also follows by applying part (1), but this time to $\operatorname{rev}_g P \smile \operatorname{rev}_h Q$.

(2): ($\Leftarrow$) Conversely, conditions (a) and (b) immediately imply that $P \smile Q$ by part (1). All that remains is to see why $\operatorname{rev}_g P \smile \operatorname{rev}_h Q$, and then we will have $P \asymp Q$. Again this can be done by using the converse direction of part (1), applied to the polynomials $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$. First, since taking the reversal of a matrix polynomial is a particular case of a Möbius transformation, note that conditions (b) and (c) for P and Q together with results in [46, 57] imply that $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$ have the same finite elementary divisors, so condition (b) holds for $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$.

To see that condition (a) holds for $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$, first observe that a polynomial and its reversal always have the same rank. This follows directly from the definition of reversal, which implies that a minor of $\operatorname{rev}_g P$ is identically zero if and only if the corresponding minor of P is identically zero. Consequently we have

$$\dim \mathcal{N}_r(\operatorname{rev}_g P) \;=\; n - \operatorname{rank}(\operatorname{rev}_g P) \;=\; n - \operatorname{rank}(P) = \dim \mathcal{N}_r(P)\,,$$

and similarly $\dim \mathcal{N}_\ell(\operatorname{rev}_g P) = \dim \mathcal{N}_\ell(P)$, for any matrix polynomial P and any choice of grade P. From condition (a) for P and Q we can now conclude that

$$\dim \mathcal{N}_r(\operatorname{rev}_g P) = \dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(Q) = \dim \mathcal{N}_r(\operatorname{rev}_h Q)\,,$$

and similarly $\dim \mathcal{N}_\ell(\operatorname{rev}_g P) = \dim \mathcal{N}_\ell(\operatorname{rev}_h Q)$, so condition (a) holds for $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$. Since both (a) and (b) hold for $\operatorname{rev}_g P$ and $\operatorname{rev}_h Q$, we have $\operatorname{rev}_g P \smile \operatorname{rev}_h Q$, and hence $P \asymp Q$. $\square$

Remark 4.2. Note that condition (a) of Theorem 4.1 is *not* equivalent to saying that $\operatorname{rank}(P) = \operatorname{rank}(Q)$, since P and Q are not assumed to have the same size.

Observe that part (2) of Theorem 4.1 may be concisely summarized:

Spectral equivalence preserves spectral structure, plus a little bit of singular structure.

This observation provides justification for the term "spectral equivalence" as an appropriate name for the relation $P \asymp Q$. Note that if P and Q are regular, then condition (a) in Theorem 4.1 is trivially satisfied, since all four nullspaces have dimension zero. The term *isospectral* is in common use [17, 31, 34, 50] for *regular* matrix polynomials P and Q (perhaps of different size and degree) that satisfy conditions (b) and (c); see also [18] for a somewhat weaker notion of isospectral system. Thus Theorem 4.1 shows that spectral equivalence generalizes this notion of isospectrality to all matrix polynomials.

The following result shows that allowing both r and s in Definition 3.2 to be nonzero is not really essential, and could have been left out. However, the flexibility of allowing both to be nonzero has been retained in the definition for the sake of convenience.

Corollary 4.3.

(a) *$P \smile Q$ if and only if* $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$ *can be achieved with either $r = 0$ or $s = 0$.*

(b) $P \asymp Q$ *if and only if* $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$ *can be achieved with either* $r = 0$ *or* $s = 0$, *and* $\operatorname{diag}(\operatorname{rev}_g P, I_t) \sim \operatorname{diag}(\operatorname{rev}_h Q, I_u)$ *can be achieved with either* $t = 0$ *or* $u = 0$.

Proof. It suffices just to modify the proof of part (1):($\Leftarrow$) in Theorem 4.1 to use the unique nonnegative integers $r, s \geq 0$ that solve the Diophantine equation $r - s = p - m = q - n$, where at least one of r or s is zero. Then the rest of the proof goes through as before to show that $\operatorname{diag}(P, I_r) \sim \operatorname{diag}(Q, I_s)$ with this particular choice of r and s. $\square$

Remark 4.4 (Möbius Transformations Preserve Spectral Equivalence)**.**
The classical notion of Möbius transformation can be extended to define a transformation of matrix polynomials: for any nonsingular matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ and matrix polynomial $P(\lambda)$ of grade k, the Möbius transform of P with respect to A is the grade k matrix polynomial

$$[\mathbf{M}_A(P)]\,(\mu) := (c\mu + d)^k P\left(\frac{a\mu + b}{c\mu + d}\right).$$

The basic properties of these transformations are explored in [46, 48, 57]; examples include Cayley transformations and the reversal operation rev_k described in Definition 2.12.

Based on the characterization in Theorem 4.1, it is shown in [46] that spectral equivalence is *preserved* by *any* Möbius transformation; that is, if $P \asymp Q$ then $\mathbf{M}_A(P) \asymp \mathbf{M}_A(Q)$. By contrast, extended unimodular equivalence is not always preserved; indeed, if $P \smile Q$ but $P \not\asymp Q$, then for "almost all" Möbius transformations $\mathbf{M}_A$ we will have $\mathbf{M}_A(P) \not\smile \mathbf{M}_A(Q)$. As a consequence it follows that Möbius transformations preserve all *strong* linearizations and quadratifications (in the sense of Definition 3.3), but usually do not preserve "weak" ones.

4.2 Comparison with Other Equivalence Relations

In Section 3.3 we recalled two equivalence relations on matrix polynomials coming from the systems and control literature, PS-eue and KV-se. In addition, their origins, motivations, and definitions were compared with those of the two relations introduced in this work. Several ways in which these two pairs of equivalence relations are very different were highlighted in that earlier section, but we also noted there the surprising fact that they partition the set of matrix polynomials into the same equivalence classes. To prove this fact is the purpose of this section. More precisely, it is proved that PS-eue defines the same equivalence classes as extended unimodular equivalence, while KV-se defines the same equivalence classes as spectral equivalence. Since our proofs rely on nontrivial results presented in [51] and [27], we are forced in this section to restrict attention to matrix polynomials with grade equal to degree.

Theorem 4.5. *Let* $P(\lambda)_{m\times n}$ *and* $Q(\lambda)_{p\times q}$ *be two matrix polynomials over an arbitrary field. Then* $P(\lambda) \smile Q(\lambda)$ *if and only if* $P(\lambda)$ *and* $Q(\lambda)$ *are PS-eue.*

Proof. Let us first assume that P and Q are PS-eue and assume, without loss of generality, that $s := m - p = n - q \geq 0$. Then Theorem 2 in [51] implies that the Smith form of P is equal to the Smith form of $\operatorname{diag}(Q, I_s)$. Therefore $P \sim \operatorname{diag}(Q, I_s)$, and hence $P \smile Q$. Conversely, if $P \smile Q$, then we may assume without loss of generality that $P \sim \operatorname{diag}(Q, I_s)$, by Corollary 4.3. As a consequence, the Smith form of P is $\operatorname{diag}(I_s, D(Q))$, where $D(Q)$ is the Smith form of Q. Now, Theorem 3 in [51] implies that P and Q are PS-eue. $\square$

Theorem 4.6. *Let* $P(\lambda)$ *and* $Q(\lambda)$ *be two matrix polynomials over an arbitrary field* $\mathbb{F}$*, with* $\operatorname{grade} P = \operatorname{degree} P$ *and* $\operatorname{grade} Q = \operatorname{degree} Q$*. Then* $P(\lambda) \asymp Q(\lambda)$ *if and only if* $P(\lambda)$ *and* $Q(\lambda)$ *are KV-se.*

Proof. Assume first that $P(\lambda) \asymp Q(\lambda)$. Then, by definition, $P(\lambda) \smile Q(\lambda)$ and $\operatorname{rev} P(\lambda) \smile \operatorname{rev} Q(\lambda)$. Theorem 4.5 applied to $P(\lambda)$ and $Q(\lambda)$ implies that $P(\lambda)$ and $Q(\lambda)$ are PS-eue, which is precisely condition (b1) in Definition 3.8(b). In addition, Theorem 4.5 applied to $\operatorname{rev} P(\lambda)$ and $\operatorname{rev} Q(\lambda)$ implies that $\operatorname{rev} P(\lambda)$ and $\operatorname{rev} Q(\lambda)$ are PS-eue, i.e., that there exist two matrix polynomials $\widetilde{M}(\lambda)$ and $\widetilde{N}(\lambda)$ such that

$$\widetilde{M}(\lambda) \cdot \operatorname{rev} P(\lambda) \;=\; \operatorname{rev} Q(\lambda) \cdot \widetilde{N}(\lambda)\,, \tag{4.1}$$

where $\widetilde{M}$, $\operatorname{rev} Q$ are left coprime, and $\operatorname{rev} P$, $\widetilde{N}$ are right coprime. These two "coprime-ness" conditions are equivalent [54, Ch. 2, Sect. 6, p. 70-71] to requiring that

$$\begin{bmatrix} \widetilde{M}(\lambda) & \operatorname{rev} Q(\lambda) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \operatorname{rev} P(\lambda) \\ -\widetilde{N}(\lambda) \end{bmatrix} \quad \text{each have full rank for all } \lambda \in \overline{\mathbb{F}}. \tag{4.2}$$

Note that (4.1) and (4.2) imply condition (b2) in Definition 3.8(b), since the *polynomial* matrices $\widetilde{M}$ and $\widetilde{N}$ do not have any finite poles at all. Therefore, we have proved that $P(\lambda)$ and $Q(\lambda)$ are KV-se.

Next assume that $P(\lambda)$ and $Q(\lambda)$ are KV-se. According to condition (b1) in Definition 3.8(b), we have that P and Q are PS-eue, and so Theorem 4.5 implies in turn that $P(\lambda) \smile Q(\lambda)$. Part (1) in Theorem 4.1 then guarantees that $P(\lambda)$ and $Q(\lambda)$ have the same numbers of left and right minimal indices, and exactly the same finite elementary divisors. In addition, $P(\lambda)$ and $Q(\lambda)$ have the same infinite elementary divisors, as stated immediately after the proof of Theorem 2 in [27]. Finally, part (2) of Theorem 4.1 allows us to conclude that $P(\lambda) \asymp Q(\lambda)$. □

4.3 Infinite Jordan Structure and Singular Structure

This section explores the effect of three equivalence relations — unimodular, extended unimodular, and spectral equivalence — on the infinite Jordan structure and singular structure of matrix polynomials. We especially emphasize the case of matrix pencils, both to make the presentation more concrete, as well as to focus on the properties of linearizations and strong linearizations defined via extended unimodular and spectral equivalence.

Let us first consider unimodular equivalence. As mentioned earlier, it was shown in [33] that in a linearization of a *regular* matrix polynomial, the infinite elementary divisors can be arbitrarily altered by unimodular equivalence, subject only to the condition that δ_∞ (i.e., the algebraic multiplicity of the infinite eigenvalue) is preserved.

For *singular* matrix polynomials, the minimal indices may also be changed by unimodular equivalence. Even worse, the Jordan structure at infinity can get *mixed together* with the singular structure by unimodular transformations. This phenomenon is starkly illustrated by the following result, which shows that for matrix pencils this mixing together of singular structure with infinite Jordan structure can occur in an essentially arbitrary way. Note that the argument given here makes use of a result (Lemma 6.3) from later in this paper; however, since this result is proved in Section 6 in a manner that is completely independent of Theorem 4.7, there is no logical circularity.

Theorem 4.7. *Consider two $m \times n$ matrix pencils $L_1(\lambda)$ and $L_2(\lambda)$ over a field $\mathbb{F}$, and the four properties:*

(a) *$L_1(\lambda)$ and $L_2(\lambda)$ have exactly the same* finite *elementary divisors,*

(b) $\operatorname{rank}(L_1) \;=\; \operatorname{rank}(L_2)$,

(c) *$L_1(\lambda)$ and $L_2(\lambda)$ have the same* numbers *of left and right minimal indices,*

(d) $\delta_\infty(L_1) + \mu(L_1) \;=\; \delta_\infty(L_2) + \mu(L_2)$.

Then the following five statements are equivalent:

(1) $L_1(\lambda) \sim L_2(\lambda)$, *i.e.,* $L_1(\lambda)$ *and* $L_2(\lambda)$ *are unimodularly equivalent.*

(2) $L_1(\lambda)$ *and* $L_2(\lambda)$ *have the same Smith form.*

(3) $L_1(\lambda)$ *and* $L_2(\lambda)$ *satisfy properties* (a) *and* (b).

(4) $L_1(\lambda)$ *and* $L_2(\lambda)$ *satisfy properties* (a) *and* (c).

(5) $L_1(\lambda)$ *and* $L_2(\lambda)$ *satisfy properties* (a) *and* (d).

Proof. The equivalence of (1) and (2) is well known, and follows from the uniqueness of the Smith form. The equivalence of (2) and (3) follows from the observation in Remark 2.9, that a Smith form is uniquely determined by its size, rank, and finite elementary divisors. That (3) and (4) are equivalent is a consequence of the characterizations

$$\begin{aligned} \#\text{ of left minimal indices of } L_i(\lambda) &= \dim \mathcal{N}_\ell(L_i) = m - \operatorname{rank}(L_i)\,, \\ \#\text{ of right minimal indices of } L_i(\lambda) &= \dim \mathcal{N}_r(L_i) = n - \operatorname{rank}(L_i)\,, \end{aligned}$$

for $i = 1, 2$, that follow directly from Definition 2.20 and the rank/nullity theorem. Finally, the equivalence of (3) and (5) follows from the relationship

$$\operatorname{rank}(L) = \delta_{\text{fin}}(L) + \delta_\infty(L) + \mu(L)\,,$$

proved in Lemma 6.3 for pencils L of any size, over an arbitrary field. □

Remark 4.8. Theorem 4.7 remains valid if $L_1(\lambda)$ and $L_2(\lambda)$ are any two $m \times n$ matrix polynomials of the *same grade*. The proof of this more general result is the same except that Theorem 6.5 has to be used in the last two lines in place of Lemma 6.3. We have emphasized the particular case of pencils, i.e., grade equal to 1, since it is the only one needed in this paper.

Remark 4.9. For matrix pencils $L(\lambda)$ over an algebraically closed field, the issues addressed in Theorem 4.7 may be viewed in a very concrete way using the Kronecker canonical form. It is convenient to partition the KCF $K(\lambda)$ of $L(\lambda)$ as $K(\lambda) = \operatorname{diag}\big[\, F(\lambda), \Omega(\lambda), S(\lambda) \,\big]$, where $F(\lambda)$ contains the blocks corresponding to the finite elementary divisors of $L(\lambda)$, $\Omega(\lambda)$ contains the blocks corresponding to the infinite elementary divisors of $L(\lambda)$, and $S(\lambda)$ contains the singular blocks corresponding to the minimal indices of $L(\lambda)$. Observe that $\Omega(\lambda)$ is itself unimodular, hence has a unimodular inverse, so we immediately see that $\Omega(\lambda) \sim I_{\delta_\infty(L)}$. Each right singular block in $S(\lambda)$ corresponding to a right minimal index d is of the form

$$S_d(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \end{bmatrix}$$

where S_d has size $d \times (d+1)$, and is easily seen by elementary column operations to be unimodularly equivalent to the $d \times (d+1)$ matrix $\big[\, 0 \mid I_d \,\big]$. Similarly each left singular block in $S(\lambda)$ corresponding to a left minimal index ℓ can be concretely shown to be unimodularly equivalent to the $(\ell + 1) \times \ell$ matrix $\begin{bmatrix} 0 \\ I_\ell \end{bmatrix}$. By some additional row and column permutations we then see that $S(\lambda)$ is unimodularly equivalent to the direct sum of an identity matrix and a zero matrix. More careful counting shows that each singular block (left or right) with minimal index e contributes e ones to this identity, and exactly one zero row (for a left minimal index) or one zero column (for a right minimal index) to the zero matrix. Altogether, then, we see that $S(\lambda) \sim \operatorname{diag}\big[\, I_{\mu(L)}, 0_{q\times p} \,\big]$, where q and p are respectively the number of left and right minimal indices of L. Hence

$$L(\lambda) \cong K(\lambda) \sim \operatorname{diag}\big[\, F(\lambda), I_{\delta_\infty + \mu}, 0_{q\times p} \,\big]\,,$$

from which we can recover the results of Theorem 4.7, at least for matrix pencils over an algebraically closed field, and we can see explicitly how unimodular equivalence in pencils only preserves, for the infinite spectral structure and the singular structure, the joint magnitude $\delta_\infty + \mu$ together with the numbers of left and right minimal indices.

The key part of Theorem 4.7 for our current discussion is the equivalence of statements (1) and (5). This equivalence shows that in addition to the finite elementary divisors, the only independent structural feature of matrix pencils preserved by unimodular equivalence is the *sum* $\delta_\infty + \mu$. Thus unimodular equivalence can change the individual values of δ_∞ and μ, in effect creating or destroying Jordan structure at ∞ in exchange for a compensating alteration of the order of left and/or right minimal bases. In the extreme case, unimodular equivalence can wipe out *all* of the infinite Jordan structure (making $\delta_\infty = 0$) and convert it all into singular structure, or conversely can zero out all the minimal indices and create elementary divisors at ∞ that weren't previously there. (Note, however, that the *numbers* of left and right minimal indices cannot be changed by unimodular equivalence, even though their *values* can all be decreased to zero.) This can be seen very concretely for matrix pencils over any algebraically closed field, based on the discussion in Remark 4.9, and can be viewed as an extension of the discussion for regular polynomials in [33] to the case of singular polynomials.

Just as for unimodular equivalence, the extended notion $P \smile Q$ of unimodular equivalence is too weak to preserve all the spectral structure of a matrix polynomial, and can also mix infinite Jordan structure together with singular structure. The rest of this section explores some of the effects that the relations $P \smile Q$ and $P \asymp Q$ can have on infinite Jordan structure and singular structure, in the specific context of investigating the properties of the new notions of linearization and strong linearization. In particular we address the following questions:

- For any given matrix polynomial P, what are the possible *sizes* of linearizations and strong linearizations for P in the sense of Definition 3.3?
- What are the possible *combinations* of singular structure and infinite Jordan structure that can appear in these linearizations and strong linearizations?

These are natural extensions of the main issues considered in [8] and [33]. Indeed, the results presented here can be viewed as extending the discussion of those papers in several senses, not only from regular to singular polynomials, and from matrix polynomials over $\mathbb{C}$ to matrix polynomials over an arbitrary field $\mathbb{F}$, but also from classical (strong) linearizations to the more general formulation of Definition 3.3.

Theorem 4.10 (Size range of linearizations).
Suppose $P(\lambda)$ is an $m \times n$ matrix polynomial over a field $\mathbb{F}$ with $\operatorname{rank} P = r$. Let $q := m - r$ and $p := n - r$ be the number of left and right minimal indices for P, respectively. Then:

(a) *There is an $s_1 \times s_2$ linearization for P in the sense of Definition 3.3, i.e., an $s_1 \times s_2$ matrix pencil $L(\lambda)$ over $\mathbb{F}$ such that $L(\lambda) \smile P(\lambda)$, if and only if*

$$s_1 \geq \delta_{\text{fin}}(P) + q\,, \quad s_2 \geq \delta_{\text{fin}}(P) + p\,, \quad \textit{and} \quad s_1 - s_2 = q - p = m - n\,. \tag{4.3}$$

In particular, the minimum-size linearization for P has $s_1 = \delta_{\text{fin}}(P) + q$ and $s_2 = \delta_{\text{fin}}(P) + p$.

(b) *For* any *choice of q left minimal indices $0 \leq \eta_1 \leq \cdots \leq \eta_q$ and p right minimal indices $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$, and* any *finite (possibly empty) list of partial multiplicities $0 < t_1 \leq \cdots \leq t_\ell$ for the eigenvalue at ∞, there is a linearization $\widetilde{L}(\lambda)$ for $P(\lambda)$ having the specified singular structure and infinite Jordan structure. The size of this $\widetilde{L}(\lambda)$ is $s_1 \times s_2$, where*

$$\begin{aligned} s_1 &= \delta_{\text{fin}}(P) + q + \mu + \omega\,, \\ s_2 &= \delta_{\text{fin}}(P) + p + \mu + \omega\,. \end{aligned} \tag{4.4}$$

Here $\mu := \sum_i \eta_i + \sum_j \varepsilon_j \geq 0$ is the sum of all the specified minimal indices, and $\omega := \sum_k t_k \geq 0$ is the sum of the specified partial multiplicities at ∞.

Proof. (a) ($\Rightarrow$): If $L \smile P$, then from Lemma 3.6(b) we have $s_1 - m = s_2 - n$, which is the third condition in (4.3). Now viewing L as a pencil over the algebraic closure $\overline{\mathbb{F}}$, we may transform L into Kronecker canonical form K by strict equivalence. Since Smith forms are invariant under field extension (see Remark 2.5), we know that L, and hence also K, has the same finite elementary divisors as P. Thus there must be blocks in K corresponding to these finite elementary divisors, occupying exactly $\delta_{\text{fin}}(P)$ rows and $\delta_{\text{fin}}(P)$ columns. But L, and hence also K, has the same number of left and right minimal indices as P, so there must be at least q further rows and p further columns in K. Thus K, and hence also L, must have at least $\delta_{\text{fin}}(P) + q$ rows and $\delta_{\text{fin}}(P) + p$ columns, and so all of (4.3) must hold.

(a) ($\Leftarrow$): To see that there exists a linearization of P for each size allowed by (4.3), first recall that for any monic degree k scalar polynomial $p(\lambda) = \lambda^k + a_{k-1}\lambda^{k-1} + \cdots + a_0$ with coefficients in $\mathbb{F}$, the associated Frobenius companion pencil is defined to be the $k \times k$ pencil

$$\mathcal{C}_p(\lambda) := \lambda I_k + \begin{bmatrix} a_{k-1} & a_{k-2} & \cdots & a_0 \\ -1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -1 & 0 \end{bmatrix} \in \mathbb{F}[\lambda]^{k\times k}.$$

It is well known [20] that $\mathcal{C}_p(\lambda)$ is unimodularly equivalent to $\operatorname{diag}[\, p(\lambda), I_{k-1}\,]$; explicit transformations to achieve this can be obtained as special cases of the arguments in the proof of Theorem 5.3.

Now if $d_j(\lambda), \ldots, d_r(\lambda)$ are the *nontrivial* invariant polynomials of $P(\lambda)$ (i.e., those with positive degree), then the direct sum of companion blocks

$$F(\lambda) := \mathcal{C}_{d_j}(\lambda) \oplus \cdots \oplus \mathcal{C}_{d_r}(\lambda) \tag{4.5}$$

defines a matrix pencil over $\mathbb{F}$ with exactly the same finite elementary divisors as $P(\lambda)$. Note that $\deg(d_j) + \cdots + \deg(d_r) = \delta_{\text{fin}}(P)$, so $F(\lambda)$ has size $\delta_{\text{fin}}(P) \times \delta_{\text{fin}}(P)$. The direct sum

$$L(\lambda) := F(\lambda) \oplus 0_{q\times p} \oplus I_\alpha$$

then gives, for any[2] $\alpha \in \mathbb{N}$, a pencil over $\mathbb{F}$ with exactly q left minimal indices (all zero) and p right minimal indices (again all zero), and exactly the same finite elementary divisors as $P(\lambda)$. Hence $L(\lambda) \smile P(\lambda)$ by Theorem 4.1, thus providing a linearization of size $s_1 \times s_2$ for every pair of integers (s_1, s_2) satisfying the conditions (4.3).

(b) To construct a linearization $\widetilde{L}(\lambda)$ for P that has the specified singular structure and infinite Jordan structure, consider the pencil

$$S(\lambda) := S_{\varepsilon_1}(\lambda) \oplus \cdots \oplus S_{\varepsilon_p}(\lambda) \oplus S^T_{\eta_1}(\lambda) \oplus \cdots \oplus S^T_{\eta_q}(\lambda)\,, \tag{4.6}$$

where

$$S_d(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \end{bmatrix} \tag{4.7}$$

denotes a *right singular block* of size $d \times (d+1)$ [16, Ch. XII], and the pencil

$$\Omega(\lambda) := I_\omega + \lambda N\,, \tag{4.8}$$

[2] Here we include 0 as an element of $\mathbb{N}$, with I_0 denoting the empty matrix.

where $N = J_{t_1}(0) \oplus \cdots \oplus J_{t_\ell}(0)$ is a direct sum of Jordan nilpotent blocks. Then

$$\widetilde{L}(\lambda) \;:=\; F(\lambda) \oplus S(\lambda) \oplus \Omega(\lambda) \tag{4.9}$$

is a pencil over $\mathbb{F}$ with the specified singular structure and infinite Jordan structure, and is a linearization for $P(\lambda)$ by Theorem 4.1, since by construction $\widetilde{L}(\lambda)$ has the same finite elementary divisors as P, and the same number of left and right minimal indices as P. The size of $S(\lambda)$ is $(q+\mu) \times (p+\mu)$, and the size of $\Omega(\lambda)$ is $\omega \times \omega$, so the total size of $\widetilde{L}(\lambda)$ is as described in (4.4). $\square$

It is worth noting that for any *fixed* size $s_1 \times s_2$ consistent with (4.3), all the linearizations of this size that are constructed in the proof of Theorem 4.10 have the same value for $\delta_\infty + \mu$, in accordance with Theorem 4.7. Also note that the linearizations constructed for Theorem 4.10, with all their various sizes, provide a large supply of examples concretely illustrating how matrix polynomials of many different sizes can be extended unimodular equivalent to each other.

Finally let us consider spectral equivalence. Theorem 4.1 tells us that this relation preserves spectral structure, but for singular structure the only invariants are the *number* of left and the *number* of right minimal indices. This means that given a singular matrix polynomial $P(\lambda)$ with p right and q left minimal indices, we might expect there to be a $Q(\lambda)$ spectrally equivalent to $P(\lambda)$, but having p right and q left minimal indices with arbitrarily different values from those of $P(\lambda)$. The following result, which is the analog for spectral equivalence of Theorem 4.10, shows that such a polynomial does indeed always exist, and that it can be chosen to be a matrix pencil. Theorem 4.11 achieves this by characterizing the possible sizes and singular structures that can occur in any strong linearization for a matrix polynomial P.

Theorem 4.11 (Size range of strong linearizations)**.**
Suppose $P(\lambda)$ is an $m \times n$ matrix polynomial over a field $\mathbb{F}$ with $\operatorname{rank} P = r$. Let $q := m - r$ and $p := n - r$ be the number of left and right minimal indices for P, respectively. If $P(\lambda)$ is singular, i.e., if at least one of q or p is nonzero, then:

(a) *There is an $s_1 \times s_2$ strong linearization for P in the sense of Definition 3.3, i.e., an $s_1 \times s_2$ matrix pencil $L(\lambda)$ over $\mathbb{F}$ such that $L(\lambda) \asymp P(\lambda)$, if and only if*

$$s_1 \geq \delta_{\text{fin}}(P) + \delta_\infty(P) + q\,, \quad s_2 \geq \delta_{\text{fin}}(P) + \delta_\infty(P) + p\,, \quad \text{and} \quad s_1 - s_2 = q - p = m - n\,. \tag{4.10}$$

In particular, the minimum-size strong linearization for P has $s_1 = \delta_{\text{fin}}(P) + \delta_\infty(P) + q$ and $s_2 = \delta_{\text{fin}}(P) + \delta_\infty(P) + p$.

(b) *For any choice of q left minimal indices $0 \leq \eta_1 \leq \cdots \leq \eta_q$ and p right minimal indices $0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_p$, there is a strong linearization $\widetilde{L}(\lambda)$ for $P(\lambda)$ having the specified singular structure. The size of this $\widetilde{L}(\lambda)$ is $s_1 \times s_2$, where*

$$\begin{aligned} s_1 &= \delta_{\text{fin}}(P) + \delta_\infty(P) + q + \mu\,, \\ s_2 &= \delta_{\text{fin}}(P) + \delta_\infty(P) + p + \mu\,. \end{aligned} \tag{4.11}$$

Here $\mu := \sum_i \eta_i + \sum_j \varepsilon_j \geq 0$ is the sum of all the specified minimal indices.

On the other hand, if $P(\lambda)$ is regular (so $q = p = \mu = 0$), then there is an $s_1 \times s_2$ strong linearization for P in the sense of Definition 3.3 if and only if $s_1 = s_2 = \delta_{\text{fin}}(P) + \delta_\infty(P)$.

Proof. The proof of this theorem has considerable overlap with the proof of Theorem 4.10, so we only outline the main points. The proof of (a) ($\Rightarrow$) follows the same line as in Theorem 4.10, but with $L \asymp P$ as assumption in place of $L \smile P$; the additional feature is that infinite elementary divisors are now also preserved, so that there must be blocks in the Kronecker canonical form of L

occupying $\delta_\infty(P)$ rows and columns. This implies the lower bounds for s_1 and s_2 in (4.10). The reverse implication (a) ($\Leftarrow$) follows from part (b), since the specified value for μ in (4.11) may be any number in $\mathbb{N}$; thus we turn immediately to the argument for part (b).

The construction of the desired strong linearization $\widetilde{L}(\lambda)$ for $P(\lambda)$ with specified singular structure is essentially the same as given in the proof of Theorem 4.10(b). Take $\widetilde{L}(\lambda) := F(\lambda) \oplus S(\lambda) \oplus \Omega(\lambda)$ as in (4.9), with $F(\lambda)$, $S(\lambda)$ and $\Omega(\lambda)$ as in (4.5), (4.6), and (4.8), respectively. The only difference in this context is that the (possibly empty) list of partial multiplicities $0 < t_1 \leq \cdots \leq t_\ell$ for the eigenvalue at ∞ is not freely chosen as in Theorem 4.10, but rather is fixed to be the same as the partial multiplicity sequence for the infinite eigenvalue of P. Then $\widetilde{L}$ has size as in (4.11), and by Theorem 4.1 we have $\widetilde{L} \asymp P$, which completes the proof of part (b) of the singular case.

Finally, if P is regular and L is any pencil such that $L \asymp P$, then by Theorem 4.1 we know that L is regular with $\delta_{\text{fin}}(L) = \delta_{\text{fin}}(P)$ and $\delta_\infty(L) = \delta_\infty(P)$. Now use the same argument as for part (a) ($\Rightarrow$) of Theorem 4.10. Viewing L as a pencil over the algebraic closure $\overline{\mathbb{F}}$, the Kronecker form of L (and hence also L itself) must have size $s_1 \times s_2$, where $s_1 = s_2 = \delta_{\text{fin}}(L) + \delta_\infty(L) = \delta_{\text{fin}}(P) + \delta_\infty(P)$. Thus strong linearizations of *regular* matrix polynomials have a uniquely determined size. $\square$

Remark 4.12. Alternative definitions for linearizations and strong linearizations of *regular* polynomials have been given in [32], and then used extensively in [1]. For example, a strong linearization for a regular $n \times n$ P of grade k is defined in [32] to be a $kn \times kn$ pencil with exactly the same finite and infinite elementary divisors as P. Note that this definition is completely consistent with the definitions and results in this paper based on spectral equivalence, at least in the regular case, in particular with Theorem 4.11 on the size of strong linearizations, and with Theorem 4.1 on the characterization of spectral equivalence in terms of elementary divisors. To see this consistency, note that Lemma 6.1 proves that for regular polynomials P of size $n \times n$ and grade k the equality $\delta_{\text{fin}}(P) + \delta_\infty(P) = kn$ holds, and so by Theorem 4.11 the unique possible size $s \times s$ of any strong linearization of P in the sense of Definition 3.3 must be with $s = kn$.

However, the definition given in [32] is not appropriate for singular polynomials, as illustrated by the following example. Consider the 2×2 singular quadratic polynomial $P(\lambda) = \operatorname{diag}[\,\lambda^2, 0\,]$, and the 4×4 pencils $L_1(\lambda) = \operatorname{diag}[\,J_2(\lambda), 0_2\,]$ and $L_2(\lambda) = \operatorname{diag}[\,J_2(\lambda), R_2(\lambda)\,]$, where $J_2(\lambda) = \left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right]$, $R_2(\lambda) = \left[\begin{smallmatrix} 1 & \lambda \\ 0 & 0 \end{smallmatrix}\right]$, and 0_2 is the 2×2 zero matrix. Then $P(\lambda)$, $L_1(\lambda)$, and $L_2(\lambda)$ all have the same finite and infinite elementary divisors (just λ^2), and both L_1 and L_2 have the classical size ($kn = 4$), but only L_2 is a strong linearization for P, either in the sense of Definition 3.3(a) or in the classical sense defined in the paragraph containing equation (3.1). To see that L_1 is not even a linearization for P, despite having the same finite and infinite elementary divisors as P, observe that $\operatorname{diag}[\,P, I_2\,]$ has rank 3 while L_1 is only rank 2, making it impossible for L_1 and $\operatorname{diag}[\,P, I_2\,]$ to be unimodularly equivalent.

At this point, it is worth summarizing the main conclusions of the developments so far in the paper:

- Strict equivalence is *too rigid*. It preserves all spectral and singular structure, but does not allow any change of size or degree. Indeed $P \cong Q$ implies that $\deg P = \deg Q$ (consider the effect of strict equivalence on the leading coefficient).

- Extended unimodular equivalence $P \smile Q$ is *too loose*, since it can destroy almost all information about the singular structure and the Jordan structure at infinity, by altering and mixing these structures together. It does, however, at least allow for a change of degree and/or a change of size (as in a linearization).

- Spectral equivalence is in between strict and extended unimodular equivalence, allowing change of both size and degree, while preserving all Jordan structure, both finite and infinite.

We have seen from Theorem 4.1 that the only matrix polynomials that can possibly share the same spectral and singular structures are ones that at least are *spectrally equivalent*. Therefore for the rest of the paper we restrict attention to spectrally equivalent polynomials.

Of course the nicest possible scenario for a linearization, or more generally for an ℓ-ification of a matrix polynomial P, would be to preserve both the spectral structure and the singular structure of P, i.e., to preserve *all* finite and infinite elementary divisors, as well as *all* left and right minimal indices. For *regular* polynomials P, with no minimal indices at all, this can certainly always be done with any strong linearization, e.g., the first or second Frobenius companion form, or any of the Fiedler pencils [2, 10]. But for *singular* P, Theorem 4.11 shows that this is now problematic, since even a strong linearization can have a completely different singular structure from that of P. This puts a premium, then, on identifying those situations where *recovery* of minimal indices of P from those of a linearization or an ℓ-ification of P is simple and a priori predictable. This in turn highlights the importance of results like those in [9, 10, 12], where minimal index recovery has been shown to be extremely simple for important classes of strong linearizations, and like those for the rectangular Frobenius-like companion forms of arbitrary grade introduced in Section 5 of this paper, where minimal index recovery is again extremely simple.

Despite the freedom expressed in Theorem 4.11 on the individual values of minimal indices under spectral equivalence, the results of that theorem also make it reasonable to suspect that there may be some general relationship constraining the sizes, degrees and minimal index sums of spectrally equivalent polynomials. To investigate the existence of such a relationship is one of the main goals of the following sections. For this purpose, we begin in the next section by establishing some basic facts about the classical Frobenius companion linearizations that are well known for square polynomials, but not so well known for rectangular polynomials. These facts provide some key tools for all that follows.

Our development of the properties of the classical Frobenius companion forms also provides a pattern for both designing and analyzing the first-ever class of strong ℓ-ifications, whose structure and properties parallel those of the classical Frobenius linearizations. In addition, this class of strong ℓ-ifications supplies us with many new examples illustrating the phenomenon of spectrally equivalent polynomials having minimal indices that are related by means of a simple uniform shift.

5 Companion Forms for Matrix Polynomials

In this section several concrete examples of strong ℓ-ifications are presented, and their basic properties are developed. First we consider $\ell = 1$, the most important case both from a historical and a practical point of view; in particular we thoroughly explore the properties of the best-known strong linearizations, the classical *first and second Frobenius companion forms* [20]. We include results and proofs in a unified framework about the Frobenius companion forms that are not as widely known as they deserve, although all of them have been presented in various ways in previous references. New strong ℓ-ifications for grades $\ell > 1$ are then introduced and analyzed. We will see that the structure and properties of these new ℓ-ifications smoothly generalize those of the classical Frobenius companion forms. A key remark is in order here: although it is natural to expect to be able to convert a matrix polynomial of *any* grade k into one of *any* other grade ℓ via a strong ℓ-ification, the reader is forewarned that such strong ℓ-ifications do not exist in general. Later results in Section 7 show that there are fundamental restrictions on which (k, ℓ)-pairs can support a strong ℓ-ification. Of course, such restrictions are satisfied by the new strong ℓ-ifications introduced in this section.

The most important linearizations from either a theoretical or computational point of view, such as the Frobenius companion forms, have a number of desirable properties that motivate us to introduce the following definition.

Definition 5.1 (Companion form of grade ℓ).
Consider the space $\mathcal{P}(g,\, m\times n,\, \mathbb{F})$ of all matrix polynomials of fixed grade g and fixed size $m\times n$, over an arbitrary field $\mathbb{F}$. Then a *companion form of grade* ℓ for matrix polynomials $P(\lambda)$ in $\mathcal{P}(g,\, m\times n,\, \mathbb{F})$ is a *uniform template* for building a grade ℓ matrix polynomial $\mathcal{C}_P(\lambda)$ in the space $\mathcal{P}(\ell,\, p\times q,\, \mathbb{F})$ from the entries in the coefficient matrices of P, in such a way that $\mathcal{C}_P(\lambda)$ is a *strong* ℓ*-ification for* P for *every* $P\in\mathcal{P}(g, m\times n, \mathbb{F})$, no matter whether P is regular or singular. The construction of the coefficient matrices of $\mathcal{C}_P(\lambda)$ from the coefficient matrices of P should involve *no matrix operations* other than scalar multiplication.

To make the notion of uniform template more precise, view each entry of each coefficient matrix of $P(\lambda)=\sum_{i=0}^{g}\lambda^i A_i$ as an independent variable x_j, so that altogether we have $(g+1)mn$ variables $x_1, x_2, \ldots, x_{(g+1)mn}$; thus, we identify $\mathcal{P}(g,\, m\times n,\, \mathbb{F})$ with $\mathbb{F}^{(g+1)mn}$. Then a uniform template for building $\mathcal{C}_P(\lambda)$ is simply a *function*

$$\begin{array}{rcl} \mathcal{P}(g, m\times n, \mathbb{F}) & \longrightarrow & \mathcal{P}(\ell,\, p\times q,\, \mathbb{F}) \\ (x_1, x_2, \ldots, x_{(g+1)mn}) & \longmapsto & \mathcal{C}_P(\lambda)=\sum_{i=0}^{\ell}\lambda^i X_i\,, \end{array} \tag{5.1}$$

where each entry of each coefficient matrix X_i is a scalar-valued function of the variables x_j with $1\le j\le (g+1)mn$ that is one of the following two types: either a constant $\alpha\in\mathbb{F}$, or a constant multiple of *just one* of the variables, i.e., a function βx_j for some $\beta\in\mathbb{F}$ and $1\le j\le (g+1)mn$.

The most common way to build a companion form of grade ℓ is with a uniform template for $\mathcal{C}_P(\lambda)=\sum_{i=0}^{\ell}\lambda^i X_i$ that consists simply of block-partitioning each coefficient X_i for $i=0,\ldots,\ell$, such that each nonzero block of X_i is either $\pm I_r$ for some $r>0$, or $\pm A_i$ for $i=0,1,\ldots,g$. This kind of uniform template has the significant advantage of being simultaneously applicable to matrix polynomials over *any field at all.* Although Definition 5.1 allows the possibility of companion forms that are specific to some particular field $\mathbb{F}$ or some special class of fields, note that all examples in the literature that are known to these authors are of the block-partitioning type described here, and thus are insensitive to the underlying field. In particular, all of the examples of companion forms of grade ℓ discussed in this paper, beginning with the Frobenius companion forms in Section 5.1, are of this type, as well as all Fiedler pencils [10, 12].

Remark 5.2. Companion forms of grade ℓ may sometimes have other valuable properties in addition to those specified in Definition 5.1, such as:

(a) The eigenvectors, minimal indices, and/or minimal bases of P can be easily recovered from those of $\mathcal{C}_P$.

(b) "Structure is preserved", in the sense that for some given class $\mathcal{S}$ of structured matrix polynomials, $\mathcal{C}_P\in\mathcal{S}$ whenever $P\in\mathcal{S}$. Such a template is a *structured companion form of grade* ℓ for the class $\mathcal{S}$.

Since the companion forms that appear most often in theory and applications are the companion forms of grade 1, we reserve the simple name *"companion form"* to mean exactly "companion form of grade 1". Any other grade different from $\ell=1$ will be stated explicitly.

Companion forms of grade 2, providing uniform templates for strong quadratifications, are also of interest in applications and can be easily constructed; in Section 5.2 we will show how to do this for matrix polynomials of even grade over arbitrary fields. Other ways of converting even grade matrix polynomials over $\mathbb{C}$ into quadratic polynomials that are equivalent in a certain sense, and have the additional property of preserving palindromic structure, have recently been presented in [23]. However, it is important to emphasize that the notion of equivalence used in [23] is much weaker than spectral equivalence, since in general it does not preserve the partial multiplicities of eigenvalues; thus it is not clear whether these examples are quadratifications in the sense of Definition 3.3 or not. In addition, some of the constructions in [23] require some matrix multiplications, i.e., are not operation free, and hence are not companion forms of grade 2 in the sense of Definition 5.1.

Concrete examples of some new companion forms of grade $\ell \geq 2$ will be constructed in Section 5.2. See also [2, 4, 10, 11, 12, 38, 40, 43, 44, 45, 59] for more on these issues, as well as for many concrete examples of structured companion forms (of grade 1) for matrix polynomial classes $\mathcal{P}(g, n \times n)$ with arbitrary odd grade g.

5.1 Frobenius Companion Forms for Rectangular Polynomials

The best known and most commonly used companion form is the *first Frobenius companion form*, defined as follows. Let $P(\lambda) = \lambda^k A_k + \lambda^{k-1} A_{k-1} + \cdots + \lambda A_1 + A_0$ denote a general $m \times n$ matrix polynomial of grade k over and arbitrary field $\mathbb{F}$. Define

$$X_1 = \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix}_{s_1 \times t_1} \quad \text{and} \quad Y_1 = \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix}_{s_1 \times t_1}, \tag{5.2}$$

where $s_1 = m + (k-1)n$ and $t_1 = kn$. Then the first Frobenius companion form of $P(\lambda)$ is the $s_1 \times t_1$ pencil $C_1(\lambda) := \lambda X_1 + Y_1$.

For regular polynomials $P(\lambda)$, it is very well known [19, 20] that $C_1(\lambda) \asymp P(\lambda)$, i.e., that $C_1(\lambda)$ is a *strong linearization* for $P(\lambda)$. Indeed $C_1(\lambda)$ is even a companion form in the sense of Definition 5.1. Not nearly as well known is the fact that $C_1(\lambda) \asymp P(\lambda)$ is still true for singular matrix polynomials, including the rectangular ($m \neq n$) case, and that the relations between the minimal indices and bases of $P(\lambda)$ and $C_1(\lambda)$ are extremely simple. All these properties are stated and proved in Theorem 5.3. This result is a special case of much more general results presented recently in [12], valid for arbitrary Fiedler pencils, which are a wide class of pencils that include the Frobenius companion forms. We present the proof here both for the sake of clarity, since the proof for the Frobenius companion forms is much simpler than the proof for general Fiedler pencils, and also because it is very similar to the arguments presented in Section 5.2 for the new strong ℓ-ifications. The interested reader may find in Remark 5.5 a detailed account of references in the literature containing results closely related to those in Theorem 5.3.

Theorem 5.3. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with grade $k \geq 2$, over an arbitrary field $\mathbb{F}$, and let $C_1(\lambda)$ be its first Frobenius companion form. Then:*

(a) *$C_1(\lambda) \asymp P(\lambda)$, i.e., $C_1(\lambda)$ is a strong linearization for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is any right minimal basis of $C_1(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $C_1(\lambda)$, and let $x_j(\lambda)$ be the* kth $(n \times 1)$ *block of $z_j(\lambda)$ for $j = 1, \ldots, p$. Then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose $\{w_1(\lambda)^T, \ldots, w_q(\lambda)^T\}$ is any left minimal basis of $C_1(\lambda)$, with vectors partitioned into blocks conformable to the block rows of $C_1(\lambda)$, and let $y_j(\lambda)^T$ be the* first $(1 \times m)$ *block of $w_j(\lambda)^T$ for $j = 1, \ldots, q$. Then $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.*

(c) (c1) *If $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then*

$$\varepsilon_1 + k - 1 \leq \varepsilon_2 + k - 1 \leq \cdots \leq \varepsilon_p + k - 1$$

are the right minimal indices of $C_1(\lambda)$.

(c2) *If* $0 \leq \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$ *are the left minimal indices of* $P(\lambda)$*, then*

$$\eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$$

are also the left minimal indices of $C_1(\lambda)$*.*

Proof. Let $P_i(\lambda) = \lambda^i A_k + \lambda^{i-1} A_{k-1} + \cdots + \lambda A_{k-i+1} + A_{k-i}$ denote the ith *Horner shift* of $P(\lambda)$, for $i = 0, 1, \ldots, k$. Note that

$$P_0(\lambda) = A_k\,, \quad P_k(\lambda) = P(\lambda)\,, \quad \text{and} \quad P_{i+1}(\lambda) - \lambda P_i(\lambda) = A_{k-(i+1)} \quad \text{for} \quad i = 0, \ldots, k-1\,. \tag{5.3}$$

Set

$$S(\lambda) = \begin{bmatrix} I_m & P_1(\lambda) & P_2(\lambda) & \cdots & P_{k-1}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix}$$

and

$$R(\lambda) = \begin{bmatrix} \lambda^{k-1} I_n & -I_n & -\lambda I_n & \cdots & -\lambda^{k-2} I_n \\ \lambda^{k-2} I_n & 0 & -I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\lambda I_n \\ \lambda I_n & 0 & \cdots & 0 & -I_n \\ I_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

It is easy to see that both $R(\lambda)$ and $S(\lambda)$ are unimodular matrices. Then straightforward computations using the relations in (5.3) show that

$$S(\lambda)\, C_1(\lambda) \;=\; \begin{bmatrix} 0 & 0 & \cdots & 0 & P(\lambda) \\ -I_n & \lambda I_n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -I_n & \lambda I_n & 0 \\ 0 & \cdots & 0 & -I_n & \lambda I_n \end{bmatrix},$$

and then

$$S(\lambda)\, C_1(\lambda)\, R(\lambda) \;=\; \begin{bmatrix} P(\lambda) & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix}, \tag{5.4}$$

so that $C_1(\lambda) \smile P(\lambda)$. Now set

$$\widetilde{S}(\lambda) = \begin{bmatrix} I_m & -\widetilde{P}_{k-2}(\lambda) & -\widetilde{P}_{k-3}(\lambda) & \cdots & -\widetilde{P}_0(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix},$$

where $\widetilde{P}_i(\lambda) := \lambda A_i + \lambda^2 A_{i-1} + \cdots + \lambda^{i+1} A_0$ for $i = 0, \ldots, k-2$, and

$$\widetilde{R}(\lambda) = \begin{bmatrix} I_n & 0 & \cdots & 0 \\ \lambda I_n & I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \lambda^{k-1} I_n & \cdots & \lambda I_n & I_n \end{bmatrix};$$

note that $\widetilde{R}(\lambda)$ is block-Toeplitz. Here it is immediate that both $\widetilde{R}(\lambda)$ and $\widetilde{S}(\lambda)$ are unimodular. Now observe that the polynomials $\widetilde{P}_i(\lambda)$ satisfy the relations

$$\begin{aligned} \widetilde{P}_0(\lambda) \;=\; \lambda A_0\,, \quad \operatorname{rev}_k P(\lambda) \;&=\; \lambda \widetilde{P}_{k-2}(\lambda) \,+\, \lambda A_{k-1} \,+\, A_k\,, \\ \text{and} \qquad \widetilde{P}_{i+1}(\lambda) - \lambda \widetilde{P}_i(\lambda) \;&=\; \lambda A_{i+1} \quad \text{for } \; i = 0, \ldots, k-3\,. \end{aligned} \tag{5.5}$$

Then using (5.5) it is straightforward to show that

$$\widetilde{S}(\lambda)\,(\operatorname{rev}_1 C_1)(\lambda) \;=\; \begin{bmatrix} \operatorname{rev}_k P(\lambda) & 0 & \cdots & 0 & 0 \\ -\lambda I_n & I_n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\lambda I_n & I_n & 0 \\ 0 & \cdots & 0 & -\lambda I_n & I_n \end{bmatrix},$$

and then that

$$\widetilde{S}(\lambda)\,(\operatorname{rev}_1 C_1)(\lambda)\,\widetilde{R}(\lambda) \;=\; \begin{bmatrix} \operatorname{rev}_k P(\lambda) & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix}.$$

So $\operatorname{rev}_1 C_1 \smile \operatorname{rev}_k P$, which completes the proof of (a).

Let us now prove (b) and (c), starting with the right minimal indices and bases. First observe that the structure of the bottom $k-1$ block rows of $C_1(\lambda)$ implies that any $z(\lambda) \in \mathcal{N}_r(C_1)$ must be of the form

$$z(\lambda) = \begin{bmatrix} \lambda^{k-1} x(\lambda)^T & \cdots & \lambda x(\lambda)^T & x(\lambda)^T \end{bmatrix}^T \tag{5.6}$$

for some $x(\lambda) \in \mathbb{F}(\lambda)^n$. Then from the first block row of $C_1(\lambda)$ we see that $z(\lambda) \in \mathcal{N}_r(C_1)$ if and only if $x(\lambda) \in \mathcal{N}_r(P)$. As a consequence of the structure of $z(\lambda)$ in (5.6), it is clear that $z(\lambda)$ is a vector *polynomial* if and only if $x(\lambda)$ is a vector polynomial. The structure of $z(\lambda)$ also implies that a list of vectors $z_1(\lambda), \ldots, z_j(\lambda) \in \mathcal{N}_r(C_1)$ is linearly independent if and only if the corresponding list $x_1(\lambda), \ldots, x_j(\lambda) \in \mathcal{N}_r(P)$ is linearly independent. Thus the correspondence $z(\lambda) \leftrightarrow x(\lambda)$ from (5.6) induces a one-to-one correspondence between vector polynomial bases of $\mathcal{N}_r(C_1)$ and vector polynomial bases of $\mathcal{N}_r(P)$. Next observe that for corresponding nonzero vector polynomials $z(\lambda)$ and $x(\lambda)$ we have $\deg z(\lambda) = \deg x(\lambda) + (k-1)$, so that the order of every corresponding pair of vector polynomial bases of $\mathcal{N}_r(C_1)$ and $\mathcal{N}_r(P)$ differs by exactly $p(k-1)$, where $p = \dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(C_1)$. Thus we conclude that $z_1(\lambda), \ldots, z_p(\lambda) \in \mathcal{N}_r(C_1)$ is a right *minimal* basis for $C_1(\lambda)$ if and only if $x_1(\lambda), \ldots, x_p(\lambda) \in \mathcal{N}_r(P)$ is a right *minimal* basis for P, which completes the proof of (b1) and (c1).

All that remains is to prove the corresponding results (b2) and (c2) for the left minimal bases and indices. The structure of vectors in the left nullspace $\mathcal{N}_\ell(C_1)$ is somewhat more complicated, but can be inferred from (5.4), rewritten in the form

$$C_1(\lambda) \;=\; S^{-1}(\lambda)\,\operatorname{diag}[\,P(\lambda), I_n, \ldots, I_n\,]\,R^{-1}(\lambda)\,. \tag{5.7}$$

Note that $S(\lambda)$ and $R(\lambda)$ are unimodular, so $S^{-1}(\lambda)$ and $R^{-1}(\lambda)$ are also unimodular polynomials. Letting $T(\lambda) := S^{-1}(\lambda)\,\operatorname{diag}[\,P(\lambda), I_n, \ldots, I_n\,]$, then (5.7) implies that $\mathcal{N}_\ell(C_1) = \mathcal{N}_\ell(T)$. But

$$T(\lambda) \;=\; \begin{bmatrix} P(\lambda) & -P_1(\lambda) & -P_2(\lambda) & \cdots & -P_{k-1}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix},$$

so we see that $w(\lambda)^T \in \mathcal{N}_\ell(C_1)$ if and only if $w(\lambda)^T$ is of the form

$$w(\lambda)^T \;=\; \left[\; y(\lambda)^T \quad y(\lambda)^T P_1(\lambda) \quad \cdots \quad y(\lambda)^T P_{k-1}(\lambda) \;\right] \tag{5.8}$$

for some $y(\lambda)^T \in \mathcal{N}_\ell(P) \subseteq \mathbb{F}(\lambda)^{1\times m}$. (Note that $y(\lambda)^T P_i(\lambda) \in \mathbb{F}(\lambda)^{1\times n}$ for $i = 1, \ldots, k-1$.) The same kind of argument as used previously for right minimal bases now shows that the correspondence $w(\lambda)^T \leftrightarrow y(\lambda)^T$ from (5.8) induces a one-to-one correspondence between vector polynomial bases of $\mathcal{N}_\ell(C_1)$ and vector polynomial bases of $\mathcal{N}_\ell(P)$.

To complete the argument for (b2) and (c2), we need to establish a connection between the degrees of corresponding left null vector polynomials $w(\lambda)^T$ and $y(\lambda)^T$. At first sight it appears that $\deg w(\lambda)^T$ could be larger than $\deg y(\lambda)^T$ by as much as $k-1$, due to the presence of the block $y(\lambda)^T P_{k-1}(\lambda)$ in $w(\lambda)^T$. However, we will show that in fact

$$\deg w(\lambda)^T \;=\; \deg y(\lambda)^T \tag{5.9}$$

holds for all vector polynomials $w(\lambda)^T \in \mathcal{N}_\ell(C_1)$. Once this is established, then it follows that every corresponding pair of vector polynomial bases of $\mathcal{N}_\ell(C_1)$ and $\mathcal{N}_\ell(P)$ have the *same order*, and so every *minimal* basis of one nullspace induces a minimal basis of the other nullspace via the correspondence $w(\lambda)^T \leftrightarrow y(\lambda)^T$ from (5.8), thus completing the proof of (b2) and (c2).

Consider the Horner shift $P_i(\lambda)$ for any $1 \le i \le k-1$. Then clearly we have

$$\lambda^{k-i} P_i(\lambda) \;+\; \widehat{P}_{k-i-1}(\lambda) \;=\; P(\lambda)\,, \tag{5.10}$$

where $\widehat{P}_{k-i-1}(\lambda) \;=\; \lambda^{k-i-1}A_{k-i-1} + \cdots + \lambda A_1 + A_0$ is the degree $k-i-1$ truncation of $P(\lambda)$. Now suppose $y(\lambda)^T$ is any vector polynomial in $\mathcal{N}_\ell(P)$ such that $y(\lambda)^T P_i(\lambda) \neq 0$. Then from (5.10) we see that

$$\lambda^{k-i} y(\lambda)^T P_i(\lambda) \;=\; -y(\lambda)^T \widehat{P}_{k-i-1}(\lambda)\,,$$

and taking degree of both sides, we have

$$(k-i) + \deg\big(y(\lambda)^T P_i(\lambda)\big) \;=\; \deg\big(y(\lambda)^T \widehat{P}_{k-i-1}(\lambda)\big) \;\le\; (k-i-1) + \deg y(\lambda)^T\,,$$

from which it follows that $\deg\big(y(\lambda)^T P_i(\lambda)\big) < \deg y(\lambda)^T$. Thus in (5.8) with any vector polynomial $y(\lambda)^T \in \mathcal{N}_\ell(P)$, we have either $y(\lambda)^T P_i(\lambda) = 0$ or $\deg\big(y(\lambda)^T P_i(\lambda)\big) < \deg y(\lambda)^T$ for $1 \le i \le k-1$, so that (5.9) is established, and the proof is complete. □

Similar results also hold for the *second Frobenius companion form* $C_2(\lambda) := \lambda X_2 + Y_2$, where

$$X_2 = \begin{bmatrix} A_k & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix}_{s_2\times t_2} \qquad \text{and} \qquad Y_2 = \begin{bmatrix} A_{k-1} & -I_m & & 0 \\ A_{k-2} & 0 & \ddots & \\ \vdots & \vdots & \ddots & -I_m \\ A_0 & 0 & \cdots & 0 \end{bmatrix}_{s_2\times t_2} \tag{5.11}$$

with $s_2 = km$ and $t_2 = n + (k-1)m$. Specifically we have

Theorem 5.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m\times n$ matrix polynomial with grade $k \ge 2$, over an arbitrary field $\mathbb{F}$, and let $C_2(\lambda)$ be its second Frobenius companion form. Then:*

(a) *$C_2(\lambda) \asymp P(\lambda)$, i.e., $C_2(\lambda)$ is a strong linearization for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is any right minimal basis of $C_2(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $C_2(\lambda)$, and let $x_j(\lambda)$ be the* first *$(n\times 1)$ block of $z_j(\lambda)$ for $j = 1, \ldots, p$. Then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose* $\{w_1(\lambda)^T, \ldots, w_q(\lambda)^T\}$ *is any left minimal basis of* $C_2(\lambda)$*, with vectors partitioned into blocks conformable to the block rows of* $C_2(\lambda)$*, and let* $y_j(\lambda)^T$ *be the* kth $(1 \times m)$ *block of* $w_j(\lambda)^T$ *for* $j = 1, \ldots, q$*. Then* $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ *is a left minimal basis of* $P(\lambda)$*.*

(c) (c1) *If* $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$ *are the right minimal indices of* $P(\lambda)$*, then*

$$\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$$

are also the right minimal indices of $C_2(\lambda)$*.*

(c2) *If* $0 \leq \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$ *are the left minimal indices of* $P(\lambda)$*, then*

$$\eta_1 + k - 1 \leq \eta_2 + k - 1 \leq \cdots \leq \eta_q + k - 1$$

are the left minimal indices of $C_2(\lambda)$*.*

Proof. The proof is similar to that for Theorem 5.3, and so will be omitted. □

The minimal index results of Theorems 5.3 and 5.4 will be very important for us in Section 6; a convenient summary of these results can be expressed in terms of "index shifts":

- For the first companion form C_1, the left minimal indices of P and C_1 are *identical*, while the right minimal indices of C_1 are *uniformly shifted* from those of P by $k-1$, where $k = \operatorname{grade} P$.
- For the second companion form C_2, now it is the right minimal indices of P and C_2 that are *identical*, while the left minimal indices of C_2 are the ones that are *uniformly shifted* from those of P by $k-1$, where $k = \operatorname{grade} P$.

Remark 5.5. The fact that $C_1(\lambda)$ and $C_2(\lambda)$ are strong linearizations in the rectangular case is already known to many researchers. It is implicit for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, for example, in [6, 49, 58]. We have generalized it here for matrix polynomials over arbitrary fields. Similar arguments to the ones used in the proof of Theorem 5.3(a) have recently appeared in [24]. The minimal index and minimal basis results contained in Theorems 5.3 and 5.4 are proved in [49] for matrix polynomials over the real field and with grade chosen to be equal to the degree. We have shown that these results are valid over arbitrary fields and with arbitrary choice of grade. It was proved in [9] that exactly the same minimal index recovery rules that hold for $C_1(\lambda)$ and $C_2(\lambda)$ also hold for the linearizations in the vector spaces introduced in [41], valid for square matrix polynomials over arbitrary fields; simple recovery rules for minimal bases were also presented in [9]. The results in Theorems 5.3 and 5.4 have been extended even further to the rectangular versions of the Fiedler companion pencils in [12].

Remark 5.6. Note that $C_1(\lambda)$ and $C_2(\lambda)$ provide very concrete examples (in addition to the more artificially constructed examples of Theorem 4.11) showing that *pencils* of *different* sizes may be spectrally equivalent to the *same* polynomial, and hence also to each other. These examples show once again that strong linearizations of rectangular matrix polynomials do not have a single fixed size, in contrast to the situation for regular polynomials. The rectangular versions of the Fiedler pencils [12] provide further concrete examples of strong linearizations (even companion forms for a rectangular $P(\lambda)$) with an even wider variety of sizes.

5.2 Companion Forms of grade $\ell \geq 1$

In order to define the new strong ℓ-ifications, we need some preliminary definitions. Given our usual $m \times n$ matrix polynomial $P(\lambda) = \lambda^k A_k + \lambda^{k-1} A_{k-1} + \cdots + \lambda A_1 + A_0$ of grade k over an arbitrary field $\mathbb{F}$, the strong ℓ-ifications we build require $1 \leq \ell < k$ to be a *divisor of* k, that is, $k = \ell s$ for some positive integer s. At first glance this constraint may seem unnatural, but later results presented in Theorem 7.5 and Remark 7.7 show that, in fact, it cannot be avoided when constructing companion forms of grade ℓ possessing the additional property of being valid for matrix polynomials of arbitrary sizes. This size-invariance property holds, for instance, for the first and second Frobenius companion forms and, more generally, for all Fiedler pencils [12]. Based on the coefficients of $P(\lambda)$ and the factorization $k = \ell s$, the following matrix polynomials of grade ℓ are well defined:

$$B_1(\lambda) := \lambda^\ell A_\ell + \lambda^{\ell-1} A_{\ell-1} + \cdots + \lambda A_1 + A_0\,, \tag{5.12}$$

$$B_j(\lambda) := \lambda^\ell A_{\ell j} + \lambda^{\ell-1} A_{\ell j-1} + \cdots + \lambda A_{\ell(j-1)+1}\,, \qquad \text{for } j = 2, \ldots, s\,. \tag{5.13}$$

Observe that each of $B_2(\lambda)$ through $B_s(\lambda)$ has exactly ℓ terms, and *no degree zero term*, while $B_1(\lambda)$ has $\ell + 1$ terms, including the degree zero term A_0. The key property of the polynomials $B_1, B_2, \ldots, B_s$ is that they satisfy the equality

$$P(\lambda) = \lambda^{\ell(s-1)}\, B_s(\lambda) + \lambda^{\ell(s-2)}\, B_{s-1}(\lambda) + \cdots + \lambda^\ell\, B_2(\lambda) + B_1(\lambda)\,. \tag{5.14}$$

By using the polynomials $B_1, B_2, \ldots, B_s$ as blocks, we define the following two matrix polynomials of grade ℓ:

$$C_1^\ell(\lambda) := \begin{bmatrix} B_s(\lambda) & B_{s-1}(\lambda) & B_{s-2}(\lambda) & \cdots & B_1(\lambda) \\ -I_n & \lambda^\ell I_n & 0 & \cdots & 0 \\ & -I_n & \lambda^\ell I_n & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & -I_n & \lambda^\ell I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+(s-1)n)\times sn} \tag{5.15}$$

and

$$C_2^\ell(\lambda) := \begin{bmatrix} B_s(\lambda) & -I_m & & & \\ B_{s-1}(\lambda) & \lambda^\ell I_m & -I_m & & \\ B_{s-2}(\lambda) & 0 & \lambda^\ell I_m & \ddots & \\ \vdots & \vdots & \ddots & \ddots & -I_m \\ B_1(\lambda) & 0 & \cdots & 0 & \lambda^\ell I_m \end{bmatrix} \in \mathbb{F}[\lambda]^{sm\times(n+(s-1)m)}\,. \tag{5.16}$$

Theorems 5.7 and 5.8 will prove, respectively, that $C_1^\ell(\lambda)$ and $C_2^\ell(\lambda)$ are strong ℓ-ifications for every matrix polynomial P of grade $k = \ell s$. In addition, the particular block structures of $C_1^\ell(\lambda)$ and $C_2^\ell(\lambda)$ imply that they are indeed companion forms of grade ℓ in the sense of Definition 5.1. It is worth observing how similar the structures of $C_1^\ell(\lambda)$ and $C_1(\lambda)$ (resp., of $C_2^\ell(\lambda)$ and $C_2(\lambda)$) are. These similarities make it reasonable to refer to $C_1^\ell(\lambda)$ and $C_2^\ell(\lambda)$ as *Frobenius-like companion forms of grade* ℓ. Furthermore, the similarities between $C_1^\ell(\lambda)$ and $C_1(\lambda)$ (resp., between $C_2^\ell(\lambda)$ and $C_2(\lambda)$) extend also to the proofs of Theorems 5.7 and 5.8.

Theorem 5.7. *Let $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ be any $m \times n$ matrix polynomial with grade $k \geq 2$, over an arbitrary field $\mathbb{F}$. Suppose $1 \leq \ell < k$ is any divisor of k with $k = \ell s$, and let $C_1^\ell(\lambda)$ be the matrix polynomial of grade ℓ defined in* (5.15)*. Then:*

(a) *$C_1^\ell(\lambda) \asymp P(\lambda)$, i.e., $C_1^\ell(\lambda)$ is a strong ℓ-ification for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is any right minimal basis of $C_1^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $C_1^\ell(\lambda)$, and let $x_j(\lambda)$ be the* sth $(n \times 1)$ *block of $z_j(\lambda)$ for $j = 1, \ldots, p$. Then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose $\{w_1(\lambda)^T, \ldots, w_q(\lambda)^T\}$ is any left minimal basis of $C_1^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block rows of $C_1^\ell(\lambda)$, and let $y_j(\lambda)^T$ be the* first $(1 \times m)$ *block of $w_j(\lambda)^T$ for $j = 1, \ldots, q$. Then $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.*

(c) (c1) *If* $0 \le \varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_p$ *are the right minimal indices of* $P(\lambda)$*, then*

$$\varepsilon_1 + k - \ell \;\le\; \varepsilon_2 + k - \ell \;\le\; \cdots \;\le\; \varepsilon_p + k - \ell$$

are the right minimal indices of $C_1^{\ell}(\lambda)$*.*

(c2) *If* $0 \le \eta_1 \le \eta_2 \le \cdots \le \eta_q$ *are the left minimal indices of* $P(\lambda)$*, then*

$$\eta_1 \le \eta_2 \le \cdots \le \eta_q$$

are also the left minimal indices of $C_1^{\ell}(\lambda)$*.*

Proof. The proof proceeds in parallel with that of Theorem 5.3, so many of the details are omitted. We start by defining the polynomials

$$Q_0(\lambda) := B_s(\lambda) \quad \text{and} \quad Q_j(\lambda) := \lambda^{\ell} Q_{j-1}(\lambda) + B_{s-j}(\lambda), \quad \text{for } j = 1, \ldots, s-1\,, \tag{5.17}$$

which will play roles analogous to those of the Horner shifts in the proof of Theorem 5.3. Note that (5.14) can be simply restated as

$$Q_{s-1}(\lambda) \;=\; P(\lambda)\,. \tag{5.18}$$

Consider the unimodular matrices

$$S_{\ell}(\lambda) = \begin{bmatrix} I_m & Q_0(\lambda) & Q_1(\lambda) & \cdots & Q_{s-2}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+(s-1)n)\times(m+(s-1)n)}$$

and

$$R_{\ell}(\lambda) = \begin{bmatrix} \lambda^{(s-1)\ell} I_n & -I_n & -\lambda^{\ell} I_n & \cdots & -\lambda^{(s-2)\ell} I_n \\ \lambda^{(s-2)\ell} I_n & 0 & -I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\lambda^{\ell} I_n \\ \lambda^{\ell} I_n & 0 & \cdots & 0 & -I_n \\ I_n & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{F}[\lambda]^{\,sn\times sn}\,.$$

A straightforward computation using (5.18) shows that

$$S_{\ell}(\lambda)\, C_1^{\ell}(\lambda)\, R_{\ell}(\lambda) = \operatorname{diag}[\,P(\lambda), I_n, \ldots, I_n\,]\,, \tag{5.19}$$

so $C_1^{\ell}(\lambda) \smile P(\lambda)$.

Next, we prove that $\operatorname{rev}_{\ell} C_1^{\ell} \smile \operatorname{rev}_k P$. To this end, first note that

$$(\operatorname{rev}_{\ell} C_1^{\ell})(\lambda) \;=\; \begin{bmatrix} \operatorname{rev}_{\ell} B_s & \operatorname{rev}_{\ell} B_{s-1} & \operatorname{rev}_{\ell} B_{s-2} & \cdots & \operatorname{rev}_{\ell} B_1 \\ -\lambda^{\ell} I_n & I_n & 0 & \cdots & 0 \\ & -\lambda^{\ell} I_n & I_n & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & & & -\lambda^{\ell} I_n & I_n \end{bmatrix},$$

and define

$$\begin{aligned} \widetilde{Q}_0(\lambda) &:= (\operatorname{rev}_{\ell} B_1)(\lambda) \quad \text{and} \\ \widetilde{Q}_j(\lambda) &:= \lambda^{\ell} \widetilde{Q}_{j-1}(\lambda) + (\operatorname{rev}_{\ell} B_{j+1})(\lambda)\,, \quad \text{for } j = 1, \ldots, s-1\,, \end{aligned} \tag{5.20}$$

so that

$$\widetilde{Q}_{s-1}(\lambda) \;=\; (\operatorname{rev}_k P)(\lambda)\,. \tag{5.21}$$

Now consider the unimodular matrices

$$\widetilde{S}_\ell(\lambda) = \begin{bmatrix} I_m & -\widetilde{Q}_{s-2}(\lambda) & -\widetilde{Q}_{s-3}(\lambda) & \cdots & -\widetilde{Q}_0(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+(s-1)n)\times(m+(s-1)n)}$$

and

$$\widetilde{R}_\ell(\lambda) = \begin{bmatrix} I_n & 0 & 0 & \cdots & 0 \\ \lambda^\ell I_n & I_n & \ddots & \ddots & \vdots \\ \lambda^{2\ell} I_n & \lambda^\ell I_n & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \lambda^{(s-1)\ell} I_n & \cdots & \lambda^{2\ell} I_n & \lambda^\ell I_n & I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{sn\times sn};$$

note that $\widetilde{R}_\ell(\lambda)$ is block-Toeplitz. Again, a straightforward computation using (5.20) and (5.21) shows that

$$\widetilde{S}_\ell(\lambda)\,(\mathrm{rev}_\ell C_1^\ell)(\lambda)\,\widetilde{R}_\ell(\lambda) \;=\; \mathrm{diag}[\,(\mathrm{rev}_k P)(\lambda), I_n, \ldots, I_n\,].$$

So $\mathrm{rev}_\ell C_1^\ell \smile \mathrm{rev}_k P$, which completes the proof of (a).

Let us now sketch the proofs of (b1) and (c1), which are very similar to the proofs of (b1) and (c1) of Theorem 5.3. The structure of $C_1^\ell(\lambda)$ implies that $z(\lambda) \in \mathcal{N}_r(C_1^\ell)$ if and only if

$$z(\lambda) = \begin{bmatrix} \lambda^{(s-1)\ell}\, x(\lambda)^T & \cdots & \lambda^{2\ell}\, x(\lambda)^T & \lambda^\ell\, x(\lambda)^T & x(\lambda)^T \end{bmatrix}^T \quad \text{for some } x(\lambda) \in \mathcal{N}_r(P). \qquad (5.22)$$

It is clear from (5.22) that $z(\lambda)$ is a vector *polynomial* if and only if $x(\lambda)$ is a vector polynomial, and also that a list of vectors $z_1(\lambda), \ldots, z_j(\lambda) \in \mathcal{N}_r(C_1^\ell)$ is linearly independent if and only if the corresponding list $x_1(\lambda), \ldots, x_j(\lambda) \in \mathcal{N}_r(P)$ is linearly independent. Thus the correspondence $z(\lambda) \leftrightarrow x(\lambda)$ from (5.22) induces a one-to-one correspondence between vector polynomial bases of $\mathcal{N}_r(C_1^\ell)$ and vector polynomial bases of $\mathcal{N}_r(P)$. Next observe that for corresponding nonzero vector polynomials $z(\lambda)$ and $x(\lambda)$ we have $\deg z(\lambda) = \deg x(\lambda) + (s-1)\ell = \deg x(\lambda) + (k-\ell)$. The rest of the argument is identical to the one used in the proof of (b1) and (c1) of Theorem 5.3; simply replace $(k-1)$ by $(k-\ell)$.

It only remains to prove (b2) and (c2). Again the proof is similar to the one of (b2) and (c2) of Theorem 5.3. From (5.19) the left nullspace satisfies $\mathcal{N}_\ell(C_1^\ell) = \mathcal{N}_\ell(T_\ell)$, where $T_\ell(\lambda) := S_\ell^{-1}(\lambda)\ \mathrm{diag}[\,P(\lambda), I_n, \ldots, I_n\,]$. Note that

$$T_\ell(\lambda) \;=\; \begin{bmatrix} P(\lambda) & -Q_0(\lambda) & -Q_1(\lambda) & \cdots & -Q_{s-2}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix}.$$

So $w(\lambda)^T \in \mathcal{N}_\ell(C_1^\ell)$ if and only if $w(\lambda)^T$ is of the form

$$w(\lambda)^T \;=\; \begin{bmatrix} y(\lambda)^T & y(\lambda)^T Q_0(\lambda) & \cdots & y(\lambda)^T Q_{s-2}(\lambda) \end{bmatrix} \qquad (5.23)$$

for some $y(\lambda)^T \in \mathcal{N}_\ell(P) \subseteq \mathbb{F}(\lambda)^{1\times m}$. The same kind of argument used in the proof of Theorem 5.3 now shows that the correspondence $w(\lambda)^T \leftrightarrow y(\lambda)^T$ from (5.23) induces a one-to-one correspondence between vector polynomial bases of $\mathcal{N}_\ell(C_1^\ell)$ and vector polynomial bases of $\mathcal{N}_\ell(P)$. To complete the argument for (b2) and (c2), we need to establish that $\deg w(\lambda)^T \;=\; \deg y(\lambda)^T$ holds for all vector polynomials $w(\lambda)^T \in \mathcal{N}_\ell(C_1^\ell)$. Once this is established, the same argument as in the proof of Theorem 5.3 allows us to conclude that (b2) and (c2) hold. From (5.14) and (5.17), we obtain the following analogs of (5.10):

$$\lambda^{\ell(s-(j+1))} Q_j(\lambda) \,+\, \widehat{Q}_j(\lambda) \;=\; P(\lambda) \quad \text{for each } \; j = 0, 1, \ldots, s-2\,, \tag{5.24}$$

where $\widehat{Q}_j(\lambda) := \lambda^{\ell(s-(j+2))} B_{s-(j+1)}(\lambda) \,+\, \cdots + \lambda^{\ell} B_2(\lambda) + B_1(\lambda)$ is the grade $\big(\ell(s-(j+2)) + \ell\big)$ truncation of $P(\lambda)$. So if $y(\lambda)^T \in \mathcal{N}_\ell(P)$ and $y(\lambda)^T\, Q_j(\lambda) \neq 0$, then

$$\lambda^{\ell(s-(j+1))} y(\lambda)^T\; Q_j(\lambda) \;=\; -y(\lambda)^T\, \widehat{Q}_j(\lambda)\,.$$

Since the grade of $B_{s-(j+1)}(\lambda)$ is ℓ, then taking degrees of both sides gives

$$\ell(s-(j+1)) + \deg\big(y(\lambda)^T Q_j(\lambda)\big) \;=\; \deg\big(y(\lambda)^T\, \widehat{Q}_j(\lambda)\big) \;\leq\; \deg y(\lambda)^T + \ell(s-(j+2)) + \ell\,,$$

from which it follows that $\deg\big(y(\lambda)^T Q_j(\lambda)\big) \leq \deg y(\lambda)^T$ for $j = 0, 1, \ldots, s-2$. Thus in (5.23) with any vector polynomial $y(\lambda)^T \in \mathcal{N}_\ell(P)$, we have either $\deg\big(y(\lambda)^T Q_j(\lambda)\big) \leq \deg y(\lambda)^T$ or $y(\lambda)^T Q_j(\lambda) = 0$ for $0 \leq j \leq s-2$, so $\deg w(\lambda)^T \;=\; \deg y(\lambda)^T$, and the proof is complete. □

Theorem 5.8 is the counterpart of Theorem 5.7 for $C_2^\ell(\lambda)$. The proof is very similar to that of Theorem 5.7, and so is omitted.

Theorem 5.8. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with grade $k \geq 2$, over an arbitrary field $\mathbb{F}$. Suppose $1 \leq \ell < k$ is any divisor of k with $k = \ell s$, and let $C_2^\ell(\lambda)$ be the matrix polynomial of grade ℓ defined in* (5.16)*. Then:*

(a) *$C_2^\ell(\lambda) \asymp P(\lambda)$, i.e., $C_2^\ell(\lambda)$ is a strong ℓ-ification for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is any right minimal basis of $C_2^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $C_2^\ell(\lambda)$, and let $x_j(\lambda)$ be the* first *$(n \times 1)$ block of $z_j(\lambda)$ for $j = 1, \ldots, p$. Then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose $\{w_1(\lambda)^T, \ldots, w_q(\lambda)^T\}$ is any left minimal basis of $C_2^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block rows of $C_2^\ell(\lambda)$, and let $y_j(\lambda)^T$ be the* sth *$(1 \times m)$ block of $w_j(\lambda)^T$ for $j = 1, \ldots, q$. Then $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.*

(c) (c1) *If $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then*

$$\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$$

are also the right minimal indices of $C_2^\ell(\lambda)$.

(c2) *If $0 \leq \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$ are the left minimal indices of $P(\lambda)$, then*

$$\eta_1 + k - \ell \;\leq\; \eta_2 + k - \ell \;\leq\; \cdots \;\leq\; \eta_q + k - \ell$$

are the left minimal indices of $C_2^\ell(\lambda)$.

After linearizations, quadratifications are probably the most important type of ℓ-ification from the practical point of view. As an immediate consequence of Theorem 5.7 (or of Theorem 5.8) with $\ell = 2$ and any even $k \geq 2$, we have the following existence result.

Corollary 5.9. *Every matrix polynomial of even grade has a strong quadratification.*

Furthermore, these theorems provide two explicit examples of companion forms of grade 2 for matrix polynomials of any even grade k.

For the construction of the particular companion forms C_1^ℓ and C_2^ℓ of grade ℓ, it is essential that ℓ be a divisor of the grade k of the polynomial P being ℓ-ified; thus for a given P, the ℓ in C_1^ℓ and C_2^ℓ cannot be arbitrarily specified. We close this section by showing a simple way to construct ℓ-ifications for any $m \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ of grade k, where the grade ℓ with $1 \le \ell < k$ *can* be arbitrarily chosen. However, we want to point out that these ℓ-ifications are not companion forms of grade ℓ, since they are not strong ℓ-ifications in general. Nonetheless, they provide uniform templates for building ℓ-ifications for any $P(\lambda)$. In fact, we will see in Theorem 7.5, as a consequence of the "Index Sum Theorem", that for matrix polynomials P with grade k, there can be certain values $\ell < k$ such that P has no strong ℓ-ification at all.

The ℓ-ifications we consider are again inspired by the classical Frobenius companion forms, and are defined as follows:

$$W_1^\ell(\lambda) := \begin{bmatrix} P_\ell(\lambda) & A_{k-\ell-1} & \cdots & A_1 & A_0 \\ -I_n & \lambda I_n & \cdots & 0 & 0 \\ & \ddots & \ddots & \vdots & \vdots \\ & & -I_n & \lambda I_n & 0 \\ 0 & & & -I_n & \lambda I_n \end{bmatrix} \in \mathbb{F}[\lambda]^{(m+n(k-\ell))\times n(k-\ell+1)} \tag{5.25}$$

and

$$W_2^\ell(\lambda) := \begin{bmatrix} P_\ell(\lambda) & -I_m & & & 0 \\ A_{k-\ell-1} & \lambda I_m & -I_m & & \\ \vdots & & \ddots & \ddots & \\ A_1 & 0 & \dots & \lambda I_m & -I_m \\ A_0 & 0 & \dots & 0 & \lambda I_m \end{bmatrix} \in \mathbb{F}[\lambda]^{(\,m(k-\ell+1)\,)\times(\,n+m(k-\ell)\,)}, \tag{5.26}$$

where $P_\ell(\lambda) = \lambda^\ell A_k + \lambda^{\ell-1} A_{k-1} + \cdots + A_{k-\ell}$ is the ℓth Horner shift of $P(\lambda)$. Theorems 5.10 and 5.11 prove that $W_1^\ell(\lambda)$ and $W_2^\ell(\lambda)$ are indeed ℓ-ifications for $P(\lambda)$, and show the relations between the minimal indices and bases of $P(\lambda)$ and those of $W_1^\ell(\lambda)$ and $W_2^\ell(\lambda)$. We only sketch the proof of Theorem 5.10 very briefly, since it is similar to that of Theorem 5.7. The proof of Theorem 5.11 is omitted.

Theorem 5.10. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with grade $k \ge 2$, over an arbitrary field $\mathbb{F}$. Suppose ℓ is any integer such that $1 \le \ell < k$, and let $W_1^\ell(\lambda)$ be the matrix polynomial of grade ℓ defined in* (5.25). *Then:*

(a) *$W_1^\ell(\lambda) \smile P(\lambda)$, i.e., $W_1^\ell(\lambda)$ is an ℓ-ification for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \dots, z_p(\lambda)\}$ is any right minimal basis of $W_1^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $W_1^\ell(\lambda)$, and let $x_j(\lambda)$ be the $(k-\ell+1)$th $(n \times 1)$ block of $z_j(\lambda)$ for $j = 1, \dots, p$. Then $\{x_1(\lambda), \dots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose $\{w_1(\lambda)^T, \dots, w_q(\lambda)^T\}$ is any left minimal basis of $W_1^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block rows of $W_1^\ell(\lambda)$, and let $y_j(\lambda)^T$ be the* first *$(1 \times m)$ block of $w_j(\lambda)^T$ for $j = 1, \dots, q$. Then $\{y_1(\lambda)^T, \dots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.*

(c) (c1) *If $0 \le \varepsilon_1 \le \varepsilon_2 \le \cdots \le \varepsilon_p$ are the right minimal indices of $P(\lambda)$, then*

$$\varepsilon_1 + k - \ell \le \varepsilon_2 + k - \ell \le \cdots \le \varepsilon_p + k - \ell$$

are the right minimal indices of $W_1^\ell(\lambda)$.

(c2) *If $0 \le \eta_1 \le \eta_2 \le \cdots \le \eta_q$ are the left minimal indices of $P(\lambda)$, then*

$$\eta_1 \le \eta_2 \le \cdots \le \eta_q$$

are also the left minimal indices of $W_1^\ell(\lambda)$.

Proof. Set

$$U_\ell(\lambda) = \begin{bmatrix} I_m & P_\ell(\lambda) & P_{\ell+1}(\lambda) & \cdots & P_{k-1}(\lambda) \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n & 0 \\ 0 & \cdots & \cdots & 0 & I_n \end{bmatrix}$$

and

$$V_\ell(\lambda) = \begin{bmatrix} \lambda^{k-\ell} I_n & -I_n & -\lambda I_n & \cdots & -\lambda^{k-\ell-1} I_n \\ \lambda^{k-\ell-1} I_n & 0 & -I_n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\lambda I_n \\ \lambda I_n & 0 & \cdots & 0 & -I_n \\ I_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Both $U_\ell(\lambda)$ and $V_\ell(\lambda)$ are unimodular matrices, and direct computations analogous to those used in the proof of Theorem 5.3(a) give

$$U_\ell(\lambda)\, W_1^\ell(\lambda)\, V_\ell(\lambda) \;=\; \mathrm{diag}[\, P(\lambda), I_n, \ldots, I_n\,],$$

so W_1^ℓ is an ℓ-ification of $P(\lambda)$.

The proofs of (b1) and (c1) follow from the structure of $W_1^\ell(\lambda)$, which implies that $z(\lambda) \in \mathcal{N}_r(W_1^\ell)$ if and only if

$$z(\lambda) = \begin{bmatrix} \lambda^{(k-\ell)}\, x(\lambda)^T & \cdots & \lambda^2\, x(\lambda)^T & \lambda\, x(\lambda)^T & x(\lambda)^T \end{bmatrix}^T \quad \text{for some } x(\lambda) \in \mathcal{N}_r(P). \tag{5.27}$$

From (5.27), we obtain (b1) and (c1) via the same argument used in the proof of (b1) and (c1) of Theorem 5.7.

The proofs of (b2) and (c2) follow from the fact that $\mathcal{N}_\ell(W_1^\ell) = \mathcal{N}_\ell(T_\ell)$, where $T_\ell(\lambda) := U_\ell^{-1}(\lambda)\,\mathrm{diag}[\,P(\lambda), I_n, \ldots, I_n\,]$, which allows us to prove that $w(\lambda)^T \in \mathcal{N}_\ell(W_1^\ell)$ if and only if $w(\lambda)^T$ is of the form

$$w(\lambda)^T \;=\; \begin{bmatrix} y(\lambda)^T & y(\lambda)^T P_\ell(\lambda) & y(\lambda)^T P_{\ell+1}(\lambda) & \cdots & y(\lambda)^T P_{k-1}(\lambda) \end{bmatrix} \tag{5.28}$$

for some $y(\lambda)^T \in \mathcal{N}_\ell(P) \subseteq \mathbb{F}(\lambda)^{1\times m}$. From here, the argument is completely analogous to the proof of (b2) and (c2) in Theorem 5.3. ☐

Theorem 5.11. *Let $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ be an $m \times n$ matrix polynomial with grade $k \ge 2$, over an arbitrary field $\mathbb{F}$. Suppose ℓ is any integer such that $1 \le \ell < k$, and let $W_2^\ell(\lambda)$ be the matrix polynomial of grade ℓ defined in* (5.26)*. Then:*

(a) *$W_2^\ell(\lambda) \smile P(\lambda)$, i.e., $W_2^\ell(\lambda)$ is an ℓ-ification for $P(\lambda)$.*

(b) (b1) *Suppose $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is any right minimal basis of $W_2^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block columns of $W_2^\ell(\lambda)$, and let $x_j(\lambda)$ be the* first *$(n \times 1)$ block of $z_j(\lambda)$ for $j = 1, \ldots, p$. Then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

(b2) *Suppose $\{w_1(\lambda)^T, \ldots, w_q(\lambda)^T\}$ is any left minimal basis of $W_2^\ell(\lambda)$, with vectors partitioned into blocks conformable to the block rows of $W_2^\ell(\lambda)$, and let $y_j(\lambda)^T$ be the $(k-\ell+1)$th $(1 \times m)$ block of $w_j(\lambda)^T$ for $j = 1, \ldots, q$. Then $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.*

(c) (c1) *If* $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$ *are the right minimal indices of* $P(\lambda)$*, then*

$$\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_p$$

are also the right minimal indices of $W_2^\ell(\lambda)$*.*

(c2) *If* $0 \leq \eta_1 \leq \eta_2 \leq \cdots \leq \eta_q$ *are the left minimal indices of* $P(\lambda)$*, then*

$$\eta_1 + k - \ell \leq \eta_2 + k - \ell \leq \cdots \leq \eta_q + k - \ell$$

are the left minimal indices of $W_2^\ell(\lambda)$*.*

Despite the fact that W_1^ℓ and W_2^ℓ are just ℓ-ifications and not strong ℓ-ifications, note that as a consequence of Theorem 4.1 they still may be useful for studying polynomials $P(\lambda)$ when the presence or behavior of any eigenvalues at ∞ is not of any interest.

Remark 5.12 (Recovery of eigenvectors from ℓ-ifications)**.**
In this section, we have established how to obtain minimal bases of P from minimal bases of the ℓ-ifications C_1^ℓ, C_2^ℓ, W_1^ℓ, and W_2^ℓ. This is achieved almost without effort simply by extracting the appropriate block (sometimes the first, sometimes the last) of each minimal basis vector of the ℓ-ification. Regular polynomials P of size $n \times n$ do not have minimal bases, so the vectors of interest in this important case are the left and right eigenvectors corresponding to the eigenvalues of P, more generally bases for the nullspaces $\mathcal{N}_\ell(P(\lambda_0)) \subseteq \overline{\mathbb{F}}^{1\times n}$ and $\mathcal{N}_r(P(\lambda_0)) \subseteq \overline{\mathbb{F}}^{n\times 1}$ for each eigenvalue $\lambda_0 \in \overline{\mathbb{F}}$ of $P(\lambda)$. Thus in the regular case we would like to know how to recover bases of $\mathcal{N}_\ell(P(\lambda_0))$ from bases of $\mathcal{N}_\ell(C_1^\ell(\lambda_0))$, $\mathcal{N}_\ell(C_2^\ell(\lambda_0))$, $\mathcal{N}_\ell(W_1^\ell(\lambda_0))$, or $\mathcal{N}_\ell(W_2^\ell(\lambda_0))$ [respectively, bases of $\mathcal{N}_r(P(\lambda_0))$ from bases of $\mathcal{N}_r(C_1^\ell(\lambda_0))$, $\mathcal{N}_r(C_2^\ell(\lambda_0))$, $\mathcal{N}_r(W_1^\ell(\lambda_0))$, or $\mathcal{N}_r(W_2^\ell(\lambda_0))$]. It turns out that for each finite eigenvalue λ_0, this can be achieved by extracting exactly the same blocks (either first or last) as in the corresponding recoveries of minimal bases described in Theorems 5.7, 5.8, 5.10, and 5.11. The proofs of these facts have the same flavor as the arguments presented in this section, and so are omitted for the sake of brevity.

For $\lambda_0 = \infty$, the eigenvector recovery problem makes sense only for C_1^ℓ and C_2^ℓ, since W_1^ℓ and W_2^ℓ are not strong linearizations. In this case, the nullspaces of interest are those of the matrix coefficient of λ^ℓ in C_1^ℓ and C_2^ℓ, and those of the matrix coefficient of λ^k in $P(\lambda)$. A simple argument then shows that recovery of eigenvectors for P can always be achieved by extracting the *first* block of each basis vector for the left and right eigenspaces at ∞ of C_1^ℓ and C_2^ℓ.

6 The Index Sum Theorem for Matrix Polynomials

Using the minimal index shift results for the Frobenius companion forms from Section 5.1, we can establish a simple but fundamental relationship between the structural indices (i.e., the elementary divisor degrees) and minimal indices of a general matrix polynomial over an arbitrary field. This relationship is already known for matrix polynomials over the real field [49], although it has remained unnoticed by many researchers in the linear algebra community. We first consider this relationship for two special types of matrix polynomials (regular polynomials and general pencils) in Lemmas 6.1 and 6.3, then use these special cases to derive the result for general matrix polynomials.

Recall some notation introduced earlier in Definitions 2.10, 2.15, and 2.21:

$$\begin{aligned} \delta_{\text{fin}}(P) &:= \text{ sum of all finite structural indices of } P(\lambda)\,, \\ \delta_{\infty}(P) &:= \text{ sum of all infinite structural indices of } P(\lambda)\,, \\ \mu(P) &:= \text{ sum of all minimal indices of } P(\lambda)\,. \end{aligned}$$

The following result appeared in [8] for regular pencils over $\mathbb{C}$, with essentially the same proof as given here. We include that proof not only for the convenience of the reader, but also to highlight the validity of the argument for regular polynomials over an arbitrary field, with an arbitrary choice of grade.

Lemma 6.1 (Index sum lemma for regular polynomials).
Suppose $P(\lambda)$ is a regular *$n \times n$ matrix polynomial of grade k, over an arbitrary field. Then*

$$\delta_{\text{fin}}(P) \;+\; \delta_{\infty}(P) \;=\; kn\,. \tag{6.1}$$

Proof. From the Smith form of $P(\lambda)$ it is easy to see that $\delta_{\text{fin}}(P)$ is equal to $d := \deg\big(\det P(\lambda)\big)$, and consequently that $d \leq kn$. On the other hand, from the definition of the elementary divisors at ∞ and the fact that $\operatorname{rev}_k P(\lambda)$ is also regular, we see that $\delta_{\infty}(P)$ is the same as the multiplicity of λ as a factor of $\det\big(\operatorname{rev}_k P(\lambda)\big)$, which we now proceed to compute. Letting

$$\det P(\lambda) \;=\; a_d\lambda^d + a_{d-1}\lambda^{d-1} + \cdots + a_1\lambda + a_0$$

with $a_d \neq 0$, we find that

$$\begin{aligned}
\det\big(\operatorname{rev}_k P(\lambda)\big) = \det\left[\lambda^k P\Big(\frac{1}{\lambda}\Big)\right] \;&=\; \lambda^{kn}\det P\Big(\frac{1}{\lambda}\Big)\\
&=\; \lambda^{kn}\big(a_d\lambda^{-d} + a_{d-1}\lambda^{-(d-1)} + \cdots + a_1\lambda^{-1} + a_0\big)\\
&=\; \lambda^{kn-d}(a_d + a_{d-1}\lambda + \cdots + a_1\lambda^{d-1} + a_0\lambda^d)\,.
\end{aligned}$$

Since $a_d \neq 0$, we have $\delta_{\infty}(P) = kn - d$, and thus $\delta_{\text{fin}}(P) + \delta_{\infty}(P) \;=\; d + (kn - d) \;=\; kn$, as desired. □

As described in Section 2.3, the Kronecker canonical form is the canonical form under strict equivalence for matrix pencils over an *algebraically closed field*. However, even for matrix pencils over arbitrary fields the following "partial" Kronecker form can always be attained.

Lemma 6.2 (Partial Kronecker form).
Suppose $L(\lambda)$ is an $m \times n$ pencil over an arbitrary field $\mathbb{F}$. Then L is strictly equivalent over $\mathbb{F}$ to a pencil of the form

$$\begin{bmatrix} R(\lambda) & 0 \\ 0 & S(\lambda) \end{bmatrix}, \tag{6.2}$$

where $R(\lambda)$ is a regular *$r \times r$ pencil, and $S(\lambda)$ is a* completely singular *$(m-r) \times (n-r)$ pencil in Kronecker canonical form. In other words, $S(\lambda)$ has* no *spectrum, finite or infinite, and is the direct sum of blocks of the form $0_{k\times\ell}$, $S_d(\lambda)$, and $S_d^T(\lambda)$, where $S_d(\lambda)$ is a canonical $d \times (d+1)$ right singular block as in* (4.7).

Proof. The proof given by Gantmacher in [16, Ch. XII, Sect. 4] for the Kronecker canonical form first shows that any pencil can be reduced by strict equivalence to the form (6.2), and then as a final step reduces the "regular part" $R(\lambda)$ to Weierstrass form. This initial reduction to (6.2) is valid over an arbitrary field; only the further reduction of $R(\lambda)$ to Weierstrass form requires that the underlying field be algebraically closed. □

Lemma 6.3 (Index sum lemma for general matrix pencils).
For any matrix pencil $L(\lambda)$, square or rectangular, over an arbitrary field,

$$\delta_{\text{fin}}(L) \;+\; \delta_{\infty}(L) \;+\; \mu(L) \;=\; \operatorname{rank}(L)\,. \tag{6.3}$$

Proof. Since strict equivalence preserves all finite and infinite elementary divisors, as well as the rank and all left and right minimal indices, then by Lemma 6.2 we may assume without loss of generality that $L(\lambda)$ is in the partial Kronecker form (6.2), with *regular part* $R(\lambda)$ and *singular part* $S(\lambda)$. Let $\mu := \mu(L) = \mu(S)$. Then a straightforward examination of the blocks in $S(\lambda)$ shows that $\operatorname{rank} S(\lambda) = \mu$. In $R(\lambda)$ we see from Lemma 6.1 that $\delta_{\text{fin}}(R) \;+\; \delta_{\infty}(R) \;=\; \operatorname{rank}(R)$. Thus we have

$$\operatorname{rank}(L) \;=\; \operatorname{rank}(R) + \operatorname{rank}(S) \;=\; \delta_{\text{fin}}(R) + \delta_{\infty}(R) + \mu(S) \;=\; \delta_{\text{fin}}(L) + \delta_{\infty}(L) + \mu(L)\,,$$

and the proof is complete. □

Remark 6.4. An alternative proof of Lemma 6.3 can be fashioned which makes use of the *full* KCF for pencils over algebraically closed fields, in place of Lemmas 6.1 and 6.2. We describe it here in outline only.

This argument is based on the invariance of *both* the elementary divisors and the minimal indices of a matrix polynomial under field extension. That is, if P is a matrix polynomial over an arbitrary field $\mathbb{F}$, and $\widetilde{\mathbb{F}} \supseteq \mathbb{F}$ is any extension field, then the elementary divisors and the minimal indices of P are the same regardless of whether P is viewed as a polynomial over $\mathbb{F}$ or as a polynomial over $\widetilde{\mathbb{F}}$. The invariance of elementary divisors follows from the uniqueness of the Smith form, and was discussed previously in Remarks 2.5 and 2.11. The invariance of minimal indices under field extension, although a very basic and natural-sounding property, does not seem to have appeared before in the literature. A very recent proof of this property can be found, though, in [39].

Given these two invariance properties, a proof of Lemma 6.3 may then proceed by taking the given matrix pencil $L(\lambda)$ over an arbitrary field $\mathbb{F}$, viewing it as a pencil over the algebraic closure $\overline{\mathbb{F}}$, and then reducing it by strict equivalence to Kronecker canonical form. Since this process does not change either the structural or minimal indices of the pencil, we may without loss of generality assume that L is in Kronecker canonical form to begin with. Then a straightforward counting up of the ranks of all the blocks in the KCF gives the result (6.3).

With Lemma 6.3 in hand, we are now in a position to prove one of the featured results of this paper, the *Index Sum Theorem for Matrix Polynomials*, describing a fundamental relationship between the rank, the grade, and the total size of the spectral and singular structures of an arbitrary matrix polynomial.

Theorem 6.5 (Index Sum Theorem for Matrix Polynomials)**.**
Suppose $P(\lambda)$ is an arbitrary $m \times n$ matrix polynomial over an arbitrary field. Then

$$\delta_{\text{fin}}(P) \,+\, \delta_{\infty}(P) \,+\, \mu(P) \;=\; \text{grade}(P) \cdot \text{rank}(P). \tag{6.4}$$

Proof. The result follows by counting up the rank of the first Frobenius companion form $C_1(\lambda)$ of $P(\lambda)$ in two different ways. The first way uses Lemma 6.3, the index shift result of Theorem 5.3(c), and the fact that C_1 and P have exactly the same finite and infinite elementary divisors, since C_1 is a strong linearization for P. More precisely, letting $k = \text{grade}\, P$ we have

$$\text{rank}(C_1) \;=\; \delta_{\text{fin}}(C_1) + \delta_{\infty}(C_1) + \mu(C_1) \;=\; \delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) + p(k-1)\,,$$

where p is the number of right minimal indices of $P(\lambda)$, equivalently $p = \dim \mathcal{N}_r(P)$.

The second way of counting up $\text{rank}(C_1)$ uses the definition of C_1 being a linearization for P more directly, i.e., that C_1 is unimodularly equivalent to $\text{diag}\big[\,P, I_{n(k-1)}\,\big]$ as in (5.4), so that

$$\text{rank}(C_1) \;=\; \text{rank}(P) + n(k-1) \;=\; r + n(k-1)\,,$$

where $r = \text{rank}\, P$. Equating these two counts yields

$$\delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) + p(k-1) \;=\; r + n(k-1)\,,$$

so that

$$\delta_{\text{fin}}(P) + \delta_{\infty}(P) + \mu(P) \;=\; r + (n-p)(k-1) \;=\; r + r(k-1) \;=\; kr\,,$$

as desired. □

Remark 6.6. It is important to emphasize that (6.4) continues to hold for any given matrix polynomial P, irrespective of our choice of grade for P. Observe that although $\delta_{\text{fin}}(P)$, $\mu(P)$, and $\text{rank}\, P$ are *independent* of the choice of grade, $\delta_{\infty}(P)$ is *not*. Indeed it can be shown that changing $\text{grade}\, P$ will alter the value of $\delta_{\infty}(P)$ by exactly the right amount to compensate for the change in the right-hand side of (6.4). See Lemma 2.17 for further details about this phenomenon.

Remark 6.7. The use of $C_1(\lambda)$ is not essential to prove Theorem 6.5. An alternative proof can instead use the second Frobenius companion form $C_2(\lambda)$ of $P(\lambda)$, and its index shift properties as described in Theorem 5.4(c), or any other companion form whose index shift properties are known. In particular, any Fiedler pencil [12] can be used in the proof.

Despite the fundamental nature of the Index Sum Theorem, it is not nearly as well known in the linear algebra community as it deserves. We have shown here that it holds for any matrix polynomial over an arbitrary field, but it is already known in the control and systems theory literature [27, 28, 49] for matrix polynomials over $\mathbb{R}$. To the best of our knowledge, the first appearance of this result (for real matrix polynomials) is in the conference proceedings [49], and then soon after in the paper [47]. In the recent paper [24, Th. 5.2], several special cases of the Index Sum Theorem have been independently proved (and used), in particular for matrix polynomials of either full row rank or full column rank.

7 Properties of Companion Forms and Structured Linearizations

The Index Sum Theorem imposes a mutual constraint on four of the basic properties of any matrix polynomial — its spectral structure, singular structure, grade and rank. A fifth basic property, the *size* of a matrix polynomial, is also mutually constrained by the other four properties. In particular, the rank/nullity theorem implies that the size is determined by the rank and the number of left and right minimal indices. This final section focuses on these five properties for *pairs* (P, Q) of *spectrally equivalent* matrix polynomials, and investigates the possible relationships between these properties for P and those for Q. Note that many of the results in this section are rather easy consequences of results from *other* parts of the paper; for ease of exposition they will simply be labelled as "Corollary".

As a first example, let us consider the size, rank, and grade of spectrally equivalent matrix polynomials. In the singular case, there is no a priori relationship between these three quantities, as can be seen by the polynomials

$$P(\lambda) = \begin{bmatrix} 1 & \lambda \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad Q(\lambda) = \begin{bmatrix} 1 & \lambda^k \\ 0 & 0 \end{bmatrix},$$

with $\operatorname{grade} P = 1$ and $\operatorname{grade} Q = k \geq 2$. Then using either Theorem 4.1 or Definition 3.2, it is not hard to see that P and Q are spectrally equivalent with the same size and the same rank, but with arbitrarily different grades. By contrast, the regular case provides enough extra rigidity to give us the following relationship.

Corollary 7.1. *Let $P(\lambda)$ and $Q(\lambda)$ be spectrally equivalent matrix polynomials over an arbitrary field, with P regular. Then* $\operatorname{grade} P = \operatorname{grade} Q$ *if and only if P and Q have the same size.*

Proof. Note that since $P \asymp Q$ and P is regular, Q must also be regular by Theorem 4.1. Hence the ranks of P and Q coincide with their respective sizes, $\delta_{\text{fin}}(P) = \delta_{\text{fin}}(Q)$, $\delta_\infty(P) = \delta_\infty(Q)$ and $\mu(P) = \mu(Q) = 0$. The result now follows immediately from the Index Sum Theorem. $\square$

Next let us consider the minimal indices of spectrally equivalent matrix polynomials. In order to more conveniently describe the relationship between these quantities, we introduce the following notion.

Definition 7.2. (Total index shift).
Let P and Q be two spectrally equivalent matrix polynomials over an arbitrary field. Then the *total index shift* from P to Q is the difference

$$\mathcal{S}(P, Q) := \mu(Q) - \mu(P).$$

Notice that the total index shift is not symmetric; in particular, $\mathcal{S}(Q, P) = -\mathcal{S}(P, Q)$.

Example 7.3. In [12], it was shown that every $m \times n$ polynomial P is spectrally equivalent to each of its associated Fiedler pencils F_σ, and the relationships between the minimal indices of P and those of any of the associated F_σ were determined. In particular it was shown that if P has grade k with q left and p right minimal indices, then for each F_σ there is an integer $\mathfrak{c}$ with $0 \leq \mathfrak{c} \leq k-1$ such that the q left minimal indices of F_σ are each exactly $\mathfrak{c}$ larger than those of P, while the p right minimal indices of F_σ are each exactly $(k-1-\mathfrak{c})$ larger than those of P. Thus we can easily compute the total index shift:

$$\mathcal{S}(P, F_\sigma) \;=\; q\mathfrak{c} + p\,(k-1-\mathfrak{c}) = (q-p)\,\mathfrak{c} + p\,(k-1)\,. \tag{7.1}$$

For square polynomials $p = q$, so (7.1) simplifies to

$$\mathcal{S}(P, F_\sigma) \;=\; p(k-1)\,,$$

which is independent of the Fiedler pencil being considered, and shows that $\mathcal{S}(P, F_\sigma) > 0$ for any square singular P with grade $k > 1$; thus no Fiedler pencil can ever preserve all the minimal indices of such a P. In Corollary 7.12, we will see that these properties are not special to the Fiedler companion forms of square polynomials, but hold more generally.

$\mathcal{S}(P, F_\sigma) > 0$ also holds for any *rectangular* P with size $m \times n$, grade $k > 1$, and $\operatorname{rank} P < \min\{m, n\}$, i.e., *for any P without full rank.* However, the proof requires a bit more work. Note first that [12, Corollary 4.6] implies

$$\begin{aligned} \operatorname{rank} F_\sigma &= \operatorname{rank} P + m\mathfrak{c} + n(k-1-\mathfrak{c}) && (7.2)\\ &\geq \operatorname{rank} P + \min\{m, n\}(k-1) \\ &> k \operatorname{rank} P. && (7.3) \end{aligned}$$

Next, from (7.1) and (7.2) we get

$$\begin{aligned} \mathcal{S}(P, F_\sigma) &= (m-n)\,\mathfrak{c} + \,(n - \operatorname{rank} P)\,(k-1) \\ &= \operatorname{rank} P + m\,\mathfrak{c} + n\,(k-1-\mathfrak{c}) - k \operatorname{rank} P \\ &= \operatorname{rank} F_\sigma - k \operatorname{rank} P, \end{aligned}$$

and so, $\mathcal{S}(P, F_\sigma) > 0$ by (7.3). Observe that the equality $\mathcal{S}(P, F_\sigma) = \operatorname{rank} F_\sigma - k \operatorname{rank} P$ that holds for Fiedler pencils is just a particular case of the general result presented in Corollary 7.4.

The previous example shows that the total index shift $\mathcal{S}(P, Q)$ may (at least sometimes) be simply related to the basic properties of P and Q. However, using the Index Sum Theorem it is easy to see that the total index shift can in general be expressed in terms of the ranks and grades of P and Q, as in the following result.

Corollary 7.4. *Let P and Q be spectrally equivalent matrix polynomials of grades k and ℓ, respectively, over an arbitrary field. Then the total index shift from P to Q is*

$$\mathcal{S}(P, Q) \;=\; \ell \operatorname{rank} Q - k \operatorname{rank} P\,. \tag{7.4}$$

Proof. From the Index Sum Theorem we know that

$$\begin{aligned} \mu(Q) &= \ell \operatorname{rank} Q - \delta_{\text{fin}}(Q) - \delta_\infty(Q) \\ \text{and} \qquad \mu(P) &= k \operatorname{rank} P - \delta_{\text{fin}}(P) - \delta_\infty(P)\,. \end{aligned}$$

But $P \asymp Q$, so $\delta_{\text{fin}}(P) = \delta_{\text{fin}}(Q)$ and $\delta_\infty(P) = \delta_\infty(Q)$, and (7.4) now follows immediately. $\square$

Corollary 7.4 provides a tool with which we can address the following natural question related to the results in Section 5.2:

Given a matrix polynomial P of grade k,
for what values $\ell < k$ do there exist *strong* ℓ-ifications of P?

In other words, for which values $\ell < k$ is there some Q with lower grade ℓ such that $P \asymp Q$? Theorem 7.5 answers this question for regular polynomials and makes clear that strong ℓ-ifications exist only for particular values of ℓ.

Theorem 7.5. *Let $k > 0$ be an integer, let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of grade k over an arbitrary field $\mathbb{F}$, and let $1 \leq \ell \leq k$.*

(a) *If Q is a strong ℓ-ification for P, then ℓ is a divisor of kn, and Q has size $s \times s$ with $s = kn/\ell$.*

(b) *Suppose, in addition, that $\mathbb{F}$ is algebraically closed. Then there exists a strong ℓ-ification for P if and only if ℓ is a divisor of kn.*

Proof. (a): Assume that Q is a strong ℓ-ification of size $s \times s$ for P. Since P is regular and $Q \asymp P$, Q must also be regular by Theorem 4.1. Hence the ranks of both P and Q coincide with their respective sizes. In addition, $\mathcal{S}(P, Q) = 0$ since both P and Q are regular. Therefore, Corollary 7.4 implies $\ell\, s = k\, n$, which means that ℓ is a divisor of kn and $s = kn/\ell$.

(b): The ($\Rightarrow$) direction is part of (a), so only ($\Leftarrow$) remains to be shown. We begin by recalling some recent results from [46], on which the proof of ($\Leftarrow$) relies. Consider the following scenario: suppose a finite list $\mathcal{L}$ of elementary divisors is given (including any desired repetitions, as well as any elementary divisors at ∞), where the underlying field $\mathbb{F}$ is algebraically closed. Let $T := \delta_{\text{fin}}(\mathcal{L}) + \delta_\infty(\mathcal{L})$ be the sum of the degrees of all the elementary divisors in the given list $\mathcal{L}$, and let M denote the maximum geometric multiplicity of any eigenvalue appearing in $\mathcal{L}$ (equivalently, M is the largest number of elementary divisors in $\mathcal{L}$ associated with any particular eigenvalue). Then building on ideas in Section 1.3 of [20], it is shown in [46] that if $T = kn$ and $M \leq n$, then there exists an $n \times n$ regular matrix polynomial $R(\lambda)$ of grade k whose elementary divisors are exactly the same as those in $\mathcal{L}$. Furthermore, $R(\lambda)$ can always be taken to be upper triangular. (Note that results closely related to the ones described here from [46] can also be found in [56].)

Now suppose that $\mathcal{L}$ is the elementary divisor list of the given $n \times n$ regular matrix polynomial P of grade k, so that $T = kn$ (by Lemma 6.1) and $M \leq n$ for the list $\mathcal{L}$. Further suppose that $1 \leq \ell \leq k$ and ℓ divides kn, so that $kn = \ell s$. Then clearly $s \geq n$, so the list $\mathcal{L}$ also satisfies $T = \ell s$ and $M \leq n \leq s$. Hence by the result from [46] described above there exists a regular $s \times s$ polynomial $Q(\lambda)$ of grade ℓ with exactly the same elementary divisors as $\mathcal{L}$. Thus $Q \asymp P$ by Theorem 4.1, providing a strong ℓ-ification for P, and the proof is complete. $\square$

Remark 7.6. It is natural to conjecture that Theorem 7.5(b) remains true for regular matrix polynomials over *arbitrary* fields. After this paper was submitted, this conjecture has been proved for arbitrary *infinite* fields in [13, Section 4.2]. To establish whether the result is true or not for finite fields remains an open problem.

Remark 7.7. Under the assumption that ℓ is a divisor of k, i.e., $\ell\,|\,k$, we proved in Theorems 5.7 and 5.8 that C_1^ℓ and C_2^ℓ are strong ℓ-ifications for $P(\lambda)$. Clearly $\ell\,|\,k$ implies $\ell\,|\,kn$, but not conversely. So the assumption in Theorems 5.7 and 5.8 is more restrictive than the necessary and sufficient condition established in Theorem 7.5. This is because C_1^ℓ and C_2^ℓ have an additional property apart from just being strong ℓ-ifications; they are companion forms of grade ℓ for matrix polynomials of *any size*, even rectangular. If this additional property were required of the strong ℓ-ifications in Theorem 7.5, then we would need "$\ell\,|\,kn$ for *all* n", or equivalently $\ell\,|\,k$.

The following result is a corollary of Theorem 7.5. It delineates a large class of polynomials of grade larger than 2 for which there are no strong quadratifications.

Corollary 7.8. *Let n and k be odd positive integers, and let $P(\lambda)$ be any regular matrix polynomial with size $n \times n$ and grade k, over an arbitrary field. Then P has no strong quadratification.*

Proof. Suppose Q was any strong quadratification of P. Then 2 would be a divisor of kn, by Theorem 7.5(a). But this is impossible since kn is an odd integer. $\square$

As a second illustration of how Theorem 7.5 puts constraints on the existence of strong ℓ-ifications, let us reconsider the ℓ-ifications W_1^ℓ and W_2^ℓ introduced in (5.25) and (5.26). Recall that if P is $n \times n$ with grade k, then W_1^ℓ and W_2^ℓ are both $s \times s$ with $s = (k - \ell + 1)n$. Then the following result shows, in particular, that except for $\ell = 1$, *neither* of these two ℓ-ifications is ever a strong ℓ-ification, for any regular polynomial P.

Corollary 7.9. *Let $k > 0$ be an integer, $P(\lambda)$ be an $n \times n$ regular matrix polynomial of grade k over an arbitrary field, and $1 < \ell < k$. Then there are no strong ℓ-ifications of P of size $s \times s$ with $s = (k - \ell + 1)n$.*

Proof. Suppose that there did exist such an $s \times s$ strong ℓ-ification for P; call it Q. Then Theorem 7.5(a) implies that s must also be equal to kn/ℓ. Therefore $(k-\ell+1)n = kn/\ell$, or equivalently $0 = \ell(k - \ell + 1)n - kn$, which implies $(\ell - 1)(k - \ell) = 0$, i.e., $\ell = 1$ or $\ell = k$. But this contradicts the hypothesis $1 < \ell < k$, hence no such Q can exist. $\square$

The Index Sum Theorem and Corollary 7.4 can be used to obtain even further insight into the properties of strong ℓ-ifications. In the remainder of Section 7 we see how these results constrain the properties of any possible companion form of grade ℓ, as well as reveal subtle obstructions to the existence of strong linearizations that preserve structure.

7.1 Properties of Companion Forms of grade ℓ

A companion form of grade ℓ as defined in Definition 5.1 is the most useful type of ℓ-ification, since it provides a uniform template for constructing a strong ℓ-ification of *every* matrix polynomial (regular and singular) in a given size/grade class. Companion forms of grade $\ell = 1$, termed simply "companion forms", correspond to the best possible linearizations and are by far the most important type of ℓ-ifications. The results included in this section are valid for companion forms of grades ℓ that may be different from 1, but they are particularly simple and relevant when $\ell = 1$.

We have seen in Theorem 4.11 that strong linearizations of any given singular polynomial can have many different sizes and the same is expected to be true for strong ℓ-ifications with $\ell \geq 2$, although an explicit result in that sense is not yet available in the literature. But a *companion form of grade ℓ* for square matrix polynomials can only have one size, as we see in the next result.

Corollary 7.10 (Size of square companion forms of grade ℓ)**.**
Consider the class $\mathcal{P}(k, n \times n, \mathbb{F})$ of all matrix polynomials of grade k and size $n \times n$, over an arbitrary field $\mathbb{F}$, and suppose $P(\lambda) \in \mathcal{P}(k, n \times n, \mathbb{F})$. Then:

(a) *$P(\lambda)$ is regular if and only if $\delta_{\text{fin}}(P) + \delta_\infty(P) = kn$.*

(b) *If $P(\lambda)$ is regular, and $R(\lambda)$ is a strong ℓ-ification for $P(\lambda)$ in the sense of Definition 3.3 with $\ell \leq k$, then ℓ must divide kn and $R(\lambda)$ must have size $(kn/\ell) \times (kn/\ell)$.*

(c) *Any companion form of grade $\ell \leq k$ (in the sense of Definition 5.1) for the class $\mathcal{P}(k, n \times n, \mathbb{F})$ must have size $(kn/\ell) \times (kn/\ell)$.*

Proof. (a) If P is regular, then $\delta_{\text{fin}}(P) + \delta_\infty(P) = kn$ by Lemma 6.1. In the other direction, the Index Sum Theorem (6.4) together with $\delta_{\text{fin}}(P) + \delta_\infty(P) = kn$ implies that $kn + \mu(P) = kr$, or

$\mu(P) = k(r-n)$, where $r = \operatorname{rank} P$. But $\mu(P) \geq 0$, so $r \geq n$. However, $r \leq n$ by the definition of rank. Hence $r = n$, i.e., P is regular.

(b) This is just Theorem 7.5(a).

(c) This follows immediately from (b), since any companion form of grade ℓ must in particular be a strong ℓ-ification for any regular polynomial in $\mathcal{P}(k,\, n \times n,\, \mathbb{F})$. ☐

Observe that when the results in Corollary 7.10 are specialized to the case $\ell = 1$, the "classical size" $kn \times kn$ of companion forms is recovered. In addition, Corollary 7.10 establishes that the "standard" (not yet classical) size of companions forms of grade ℓ is $(kn/\ell) \times (kn/\ell)$, since that is the unique possible size for strong ℓ-ifications of regular matrix polynomials.

We are now in a position to resolve one of the fundamental questions posed at the beginning of this paper — can *both* the spectral structure *and* the minimal indices of a singular matrix polynomial be preserved *a priori* in a polynomial of lower degree, for instance, in a linearization? Here *a priori* means "without knowing the minimal indices of the singular polynomial", which is a natural requirement given that the typical goal is to use the polynomial of lower degree to compute the spectral structure and the minimal indices of the original singular polynomial[3]. Such a complete preservation of spectral and singular structure turns out to be impossible, at least for the most useful type of ℓ-ifications, that is, for companion forms of grade ℓ. This is proved in Corollary 7.12. The following two results are also relevant to understanding various important subclasses of square polynomials, such as those with Hermitian or palindromic structure, and the restrictions on the existence of structured linearizations and structured companion forms stemming from minimal index considerations, issues to be addressed in the next section.

Corollary 7.11 (Total index shift of standard-sized strong ℓ-ifications).
Consider the class $\mathcal{P}(k,\, n \times n,\, \mathbb{F})$ of square matrix polynomials of fixed grade k and fixed size $n \times n$, over an arbitrary field $\mathbb{F}$. Let $1 \leq \ell \leq kn$ be a divisor of kn, and let $P(\lambda)$ be any polynomial in $\mathcal{P}(k,\, n \times n,\, \mathbb{F})$ such that

$$q := \dim \mathcal{N}_\ell(P) = \#\,(\textit{left minimal indices of } P) = \dim \mathcal{N}_r(P) = \#\,(\textit{right minimal indices of } P)\,.$$

If $R(\lambda)$ is any $(kn/\ell) \times (kn/\ell)$ matrix polynomial of grade ℓ such that $R \asymp P$, then the total index shift is

$$\mathcal{S}(P,R) := \mu(R) - \mu(P) \;=\; (k-\ell)q\,.$$

Proof. Since $R \asymp P$, we have by Theorem 4.1 that R also has q left minimal indices. Therefore the rank/nullity theorem implies that $\operatorname{rank} P = n - q$ and $\operatorname{rank} R = (kn/\ell) - q$. Then from Corollary 7.4,

$$\mathcal{S}(P,R) \;=\; \ell \cdot \operatorname{rank} R - k \cdot \operatorname{rank} P \;=\; \ell\,(\,(kn/\ell) - q) - k(n-q) \;=\; (k-\ell)q\,,$$

and the proof is complete. ☐

Note that we immediately recover the index shift result for (square) Fiedler pencils described in Example 7.3 as a special case of Corollary 7.11. Also observe that in Corollary 7.11, the total index shift $\mathcal{S}(P,R)$ is zero if and only if either $q = 0$ or $k = \ell$, i.e., either for regular polynomials ($q = 0$) or for strong ℓ-ifications of the same grade as the polynomial ($k = \ell$). Finally, it is worth noting that the total index shift can be *negative*, but only if $\ell > k$, i.e., only if the ℓ-ification has grade higher than that of the given polynomial P.

[3]We recall that Theorem 4.11(b) and its proof show, in particular, how to construct a strong linearization for any singular polynomial P having the same minimal indices as P. However, note that this construction requires complete prior knowledge of the spectral structure and the minimal indices of P, and therefore is not useful in practice.

Corollary 7.12 (Non-preservation of minimal indices by companion forms of grade ℓ).
Consider the class $\mathcal{P}(k,\, m\times n,\, \mathbb{F})$ of all matrix polynomials of fixed grade k and fixed size $m\times n$, over an arbitrary field $\mathbb{F}$. Let ℓ be an integer such that $1\le \ell < k$, and let $\mathcal{C}_P$ be any companion form of grade ℓ for the class $\mathcal{P}(k,\, m\times n,\, \mathbb{F})$.

(a) *Suppose $m=n$. Then for* every *singular $P\in\mathcal{P}(k,\, n\times n,\, \mathbb{F})$, the set of minimal indices of P is different from the set of minimal indices of $\mathcal{C}_P$.*

(b) *Suppose $m\neq n$, and $\mathcal{C}_P$ has size $s_1\times s_2$. Let $\widetilde{r} := \frac{\ell(s_1-m)}{k-\ell} = \frac{\ell(s_2-n)}{k-\ell}$. Then for every $P\in\mathcal{P}(k,\, m\times n,\, \mathbb{F})$ with $\operatorname{rank} P\neq \widetilde{r}$, the set of minimal indices of P is different from the set of minimal indices of $\mathcal{C}_P$.*

Proof. (a) By Corollary 7.10(c) the size of $\mathcal{C}_P$ must be $(kn/\ell)\times(kn/\ell)$, so by Corollary 7.11 we have $\mathcal{S}(P,\mathcal{C}_P) = (k-\ell)q$ for every polynomial P in $\mathcal{P}(k,\, n\times n,\, \mathbb{F})$ with q left minimal indices. Then for any singular $P\in\mathcal{P}(k,\, n\times n,\, \mathbb{F})$, we have $q>0$ and $\ell<k$, so the total index shift in passing from P to $\mathcal{C}_P$ is strictly positive.

(b) Suppose that P and $\mathcal{C}_P$ have exactly the same minimal indices. Then certainly $\mathcal{S}(P,\mathcal{C}_P)=0$, so that by Corollary 7.4 we would have

$$\ell \operatorname{rank}\mathcal{C}_P \;=\; k\operatorname{rank} P\,. \tag{7.5}$$

Then from the rank/nullity theorem and Theorem 4.1(a) combined with (7.5) we see that

$$\begin{aligned}\ell\, s_2 \;=\; \ell\,(\operatorname{rank}\mathcal{C}_P + \dim\mathcal{N}_r(\mathcal{C}_P)\,) \;&=\; k\operatorname{rank} P + \ell\,\dim\mathcal{N}_r(P)\\ &=\; k\operatorname{rank} P + \ell\,(\,n-\operatorname{rank} P\,) \;=\; (k-\ell)\,\operatorname{rank} P + \ell\, n\,.\end{aligned}$$

Solving for $\operatorname{rank} P$ now shows that it is only possible for P and $\mathcal{C}_P$ to have the same minimal indices if $\operatorname{rank} P = \widetilde{r} = \frac{\ell(s_2-n)}{k-\ell}$. Note that the equality of the two expressions given for $\widetilde{r}$ follows from Lemma 3.6(b). ☐

Remark 7.13. Note that part (b) allows the possibility that $\mathcal{C}_P$ might preserve all minimal indices for *some* polynomials $P\in\mathcal{P}(k,\, m\times n,\, \mathbb{F})$; however, this is only possible for P with the *exceptional rank* $\widetilde{r}$. Surprisingly, though, it turns out that for some important companion forms of grade ℓ the preservation of all minimal indices does in fact occur for *almost all* $P\in\mathcal{P}(k,\, m\times n,\, \mathbb{F})$. For example, if $m<n$, then most $P\in\mathcal{P}(k,\, m\times n,\, \mathbb{F})$ have $\operatorname{rank} P = m$, and thus do not have any left minimal indices at all. For these full rank polynomials both the second Frobenius companion form C_2 and the strong ℓ-ification C_2^ℓ defined in (5.16) preserve all minimal indices as a consequence of Theorems 5.4 and 5.8. Note that for both C_2 and C_2^ℓ the exceptional rank is indeed $\widetilde{r}=m$. Similar results also hold for C_1 and C_1^ℓ when $m>n$.

Remark 7.14. Corollary 7.12(b) allows us to prove very easily that a very important class of companion forms never preserve all minimal indices of a rectangular polynomial. These companion forms are those Fiedler pencils that are different from the first and second Frobenius companion forms, termed in this remark “non-Frobenius” Fiedler pencils for short. Corollaries 5.4 and 5.7 in [12] provide simple rules to recover the minimal indices of any rectangular matrix polynomial P from those of any of its associated Fiedler pencils. These rules imply immediately that “non-Frobenius” Fiedler pencils never preserve all minimal indices of P. Unfortunately, the proofs of Corollaries 5.4 and 5.7 in [12] are very long and difficult. However, simply by looking at the *sizes* of “non-Frobenius” Fiedler pencils [12, Def. 3.8], it is straightforward to show that $\widetilde{r} > \min\{m,n\} \geq \operatorname{rank} P$ always holds, and so Corollary 7.12(b) implies directly that “non-Frobenius” Fiedler pencils never preserve all minimal indices of P.

7.2 Structured Companion Forms of Structured Polynomials

The notion of a *structured companion form of grade* ℓ was introduced in Remark 5.2(b). In this final section we investigate a phenomenon closely linked to this concept, indeed a phenomenon providing one of the main motivations for studying strong ℓ-ifications and companion forms of grade $\ell > 1$ in the first place — the *non-existence* of structured companion forms (of grade $\ell = 1$) for several important classes of structured matrix polynomials. It should be stressed that this issue has been studied previously in [43, 44], but from a perspective completely different from the one considered here. More precisely, this section investigates the restrictions on the existence of structured companion forms arising from the special properties of the *minimal indices* of structured matrix polynomials; by contrast the analysis in [43, 44] is based instead on the special properties of *elementary divisors* induced by matrix polynomial structure. We will see that the new "minimal indices approach" recovers all known non-existence-of-structured-companion-form results derived from the "elementary divisors approach", but also establishes several additional such results for structure classes not amenable to arguments based on elementary divisor properties. In particular, we prove for the first time that *structured companion forms cannot exist for Hermitian or symmetric matrix polynomials of even grade*, as a consequence of minimal index considerations. For these two classes of matrix polynomials this cannot be deduced from any elementary divisor consideration.

Many matrix polynomials arising in applications have some kind of extra structure, a fact that was noticed in very early work on matrix polynomials [30]. As a consequence, structured matrix polynomials have received and continue to receive a lot of attention in the literature; see [22, 37, 42] for many further references on this subject. One of the main avenues in modern research on structured matrix polynomials is the development of *structured numerical methods* for computing the eigenvalues of such polynomials, i.e., methods which preserve the symmetries imposed on the spectrum by the various structures arising in practice. One very natural approach to developing such a method begins by devising *structured companion forms*, since for unstructured matrix polynomials the use of strong linearizations, in particular the Frobenius companion forms, combined with well-established eigenvalue algorithms for pencils has been, and still is, the preferred numerical approach to computing eigenvalues. This natural approach soon meets important obstacles, though, because examples of structured matrix polynomials not having any structured linearization in the same class can be easily constructed [42]. This has motivated further fundamental research, both on the spectral structure [43, 44, 45] and the singular structure [9] of structured matrix polynomials. The results in this section can be viewed as new contributions in this area that follow from the Index Sum Theorem.

The structures we consider are introduced in Definition 7.15, all of them for square matrix polynomials only. For the sake of conciseness, we use the symbol $\star$ as an abbreviation to denote either the transpose T or the conjugate transpose $*$ when $\mathbb{F} = \mathbb{C}$, but when working over arbitrary fields $\mathbb{F}$ to denote just the transpose T.

Definition 7.15. For a square matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ of grade k over the field $\mathbb{F}$, define the associated polynomial $P^\star(\lambda) := \sum_{i=0}^{k} \lambda^i A_i^\star$. Then $P(\lambda)$ is said to be

(a) $\star$-*symmetric* if $P^\star(\lambda) = P(\lambda)$. (Note: for $\mathbb{F} = \mathbb{C}$ and $\star = *$, "$\star$-symmetric" = "Hermitian")

(b) $\star$-*alternating* if $P^\star(\lambda) = \pm P(-\lambda)$.

(c) $\star$-*palindromic* if $P^\star(\lambda) = \pm \mathrm{rev}_k\, P(\lambda)$.

(d) $\star$-*skew-symmetric* if $P^\star(\lambda) = -P(\lambda)$, and all diagonal entries of $P(\lambda)$ are zero when $\star = T$.

Remark 7.16. Note that the second condition in Definition 7.15(d) is redundant for matrix polynomials over many fields, but is included to make the notion of skew-symmetry behave seamlessly for matrix polynomials over *all* fields. Further discussion of this point can be found in [45]. The plus and minus signs in parts (b) and (c) in Definition 7.15 correspond to slightly different structures; in fact, the structured polynomials corresponding to the minus signs are often known in the literature as anti-$\star$-*alternating* and anti-$\star$-*palindromic* polynomials [42], respectively. However, in the interest of conciseness we use a single name in each case to include both the plus and minus signs.

Let us summarize the results already available in the literature on the existence of structured companion forms for the classes of structured matrix polynomials introduced in Definition 7.15.

- For any *odd grade* k, structured companion forms exist for each of these classes.

This has been proved by explicit construction; especially simple examples can be found in [43, pp. 884–887], [44, pp. 81–84], and [45, pp. 4646–4647]. All of these simple examples are based on a pencil previously introduced in [2] only for regular polynomials. An even wider variety of structured companion forms for structured polynomials of odd grade can be found in [4, 11].

For even grade structured polynomials, the situation is more complicated and indeed a bit surprising. In [43] and [44] the following has been shown:

- For any *even grade* $k \geq 2$, structured companion forms *do not exist* for any of the $\star$-alternating or $\star$-palindromic structure classes.

The proofs of these facts follow from the existence of mismatches between the allowed multiplicities of elementary divisors associated with certain critical eigenvalues for even grade $\star$-alternating and $\star$-palindromic polynomials, versus what is possible for pencils of the same structure type. These mismatches make it possible to construct explicit examples of $\star$-alternating and $\star$-palindromic polynomials of even grade which do not have *any* structured strong linearization; the non-existence of a structure-preserving uniform template then follows. For brevity, we will refer to these mismatches as *elementary divisor obstructions* to the existence of structured companion forms. It is important to remark that it is possible to build up structured strong linearizations for almost all (regular) $\star$-alternating and $\star$-palindromic polynomials [42] of even grade, and that these linearizations are useful in practice, but the reliability of their use is always threatened by the potential existence of a nearby structured polynomial which does not have any structured strong linearization.

For $\star$-symmetric and T-skew-symmetric matrix polynomials of *even grade* the situation is different, since for these structure classes there are no elementary divisor obstructions [45]. Nonetheless, as a consequence of the special rank properties of $n \times n$ T-skew-symmetric polynomials, it has been proved in [45, Example 6.15] that structured companion forms for T-skew-symmetric polynomials with *odd size* n and even grade *do not exist.* However, the existence of structured companion forms for *even grade* $\star$-symmetric, $*$-skew-symmetric, and *even size* T-skew-symmetric matrix polynomials remained as open problems before this work.

In this section we introduce a completely new approach to the existence question for structured companion forms, an approach based on certain *minimal index obstructions* that are consequences of the Index Sum Theorem together with the special properties of the structured matrix polynomials in Definition 7.15. Using this approach we are then able to give new proofs for all the previously known non-existence results, as well as solving all the open problems posed in the previous paragraph except one: on the possible existence of structured companion forms for T-skew-symmetric matrix polynomials with even size and even grade. Although we do not settle this problem, it is shown in Theorem 7.21 that this issue cannot be resolved by the construction of a structured matrix polynomial which does not have any structured strong linearization.

We now introduce some convenient terminology to describe a key property that is encountered in many classes of structured singular polynomials.

Definition 7.17 (Left and right minimal indices coincide)**.**
Suppose $P(\lambda)$ is an $n \times n$ singular matrix polynomial over an arbitrary field, with left and right minimal indices $\eta_1 \leq \cdots \leq \eta_p$ and $\varepsilon_1 \leq \cdots \leq \varepsilon_p$, respectively. Then we say that *the left and right minimal indices of $P(\lambda)$ coincide* if $\eta_i = \varepsilon_i$ for $i = 1, \ldots, p$.

Theorem 7.18. *Let $P(\lambda)$ be a structured singular matrix polynomial from any of the structure classes introduced in* Definition 7.15*. Then the left and right minimal indices of $P(\lambda)$ coincide.*

Proof. The proofs for all the structures described in Definition 7.15(a),(b), and (c) can be found in [9, Thms. 3.4–3.6]. For $\star$-skew-symmetric polynomials, first observe that if P is $\star$-skew-symmetric, then $x(\lambda) \in \mathcal{N}_r(P)$ if and only if $x^\star(\lambda) \in \mathcal{N}_\ell(P)$, so the left and right minimal bases of P coincide (up to a conjugation if $\mathbb{F} = \mathbb{C}$ and $\star = *$). Thus the left and right minimal indices also coincide. $\square$

Corollary 7.19 delineates a subtle constraint on the minimal indices of any even grade singular structured matrix polynomial that has a structured strong linearization of "classical size". Recall that this is the size established in Corollary 7.10 (with $\ell = 1$) to be the unique possible size of any companion form.

Corollary 7.19. *Let $P(\lambda)$ be any $n \times n$ matrix polynomial over an arbitrary field with* even *grade k, whose left and right minimal indices coincide. Suppose there exists some $kn \times kn$ strong linearization $L(\lambda)$ of $P(\lambda)$ such that the left and right minimal indices of $L(\lambda)$ also coincide. Then $P(\lambda)$ must have an* even *number of left minimal indices and an* even *number of right minimal indices.*

Proof. Let q be the number of left minimal indices of P; since left and right minimal indices coincide this is also the number of right minimal indices of P. Then from Corollary 7.11 with $\ell = 1$ we have $\mathcal{S}(P, L) = (k - 1)q$. On the other hand, observe that coincidence of left and right minimal indices means that both $\mu(P)$ and $\mu(L)$ are even, so the total index shift $\mathcal{S}(P, L) := \mu(L) - \mu(P) = (k-1)q$ must also be even. Then k being even implies that q must be even. $\square$

For our purposes, a more cogent way to view the result of Corollary 7.19 is the following: if an even grade structured polynomial P in any of the classes described in Definition 7.15 has an *odd* number of left minimal indices and an *odd* number of right minimal indices, then P *cannot* have any structured strong linearization of classical size. Therefore, Corollary 7.19 constitutes a *minimal index obstruction* to the existence of structured strong linearizations of classical size for many even grade structured matrix polynomials, and hence also an obstruction to the existence of structured companion forms. This complements the *elementary divisor obstructions* discovered in [43, 44], also for *even grade* structured polynomials. However, observe that the minimal index obstruction is present for some classes of structured polynomials for which there are no elementary divisor obstructions. Both types of obstruction establish a remarkable *even/odd grade dichotomy* in the behavior of certain classes of structured polynomials, in particular for the existence of structured companion forms.

With the help of Corollary 7.19, we now prove the most important results in this section, Theorems 7.20 and 7.21. Observe that the class of T-skew-symmetric polynomials is considered separately in Theorem 7.21, since for that structure class the minimal index obstruction of Corollary 7.19 is only relevant for polynomials of odd size.

Theorem 7.20 (Non-existence of Structured Companion Forms for Even Grades)**.**
Let $\mathcal{S}$ denote any of the following structure classes of $n \times n$ matrix polynomials described in Definition 7.15*: $\star$-symmetric, $\star$-alternating, or $\star$-palindromic over an arbitrary field, $*$-skew-symmetric over the complex field. Then for* any even grade k*, there is* no structured companion form *for the class $\mathcal{S} \subset \mathcal{P}(k,\, n \times n,\, \mathbb{F})$.*

Proof. Assume first that $n > 1$, and observe that there exist *regular* matrix polynomials of size $(n - 1) \times (n - 1)$ and grade k in each structure class $\mathcal{S}$. Indeed, these regular polynomials may be easily constructed to be diagonal or antidiagonal. Now let $Q(\lambda) \in \mathcal{S}$ be any one of these regular $(n-1) \times (n-1)$ structured polynomials, and define the $n \times n$ polynomial $P(\lambda) = \mathrm{diag}\big[\, Q(\lambda), 0_{1\times 1} \,\big]$. Observe that $P(\lambda)$ is also in $\mathcal{S}$, and has *exactly one* left minimal index and *exactly one* right minimal index (both equal to zero). Then by Corollary 7.19, $P(\lambda)$ does not have any strong linearization of size $kn \times kn$ with the same structure as $P(\lambda)$. Thus there does not exist any structured companion form for even grade matrix polynomials in any of the structure classes $\mathcal{S}$, since any companion form must have size $kn \times kn$ by Corollary 7.10(c).

In the case $n = 1$, we can use the zero scalar polynomial for $P(\lambda)$, which again has *only one* left minimal index and *only one* right minimal index, and is in each of the structure classes $\mathcal{S}$. Then the same argument as before proves the result. $\square$

Our final result is Theorem 7.21, which considers the existence of structured strong linearizations of classical size for *even grade* T-skew-symmetric matrix polynomials. This result deserves a few preliminary comments. First, it establishes the first-ever example of an even/odd *size* (rather than even/odd grade) dichotomy for the existence of structured strong linearizations. Second, although Theorem 7.21(a) certainly proves that there are no structured companion forms for the class of even grade/odd size T-skew-symmetric matrix polynomials, it actually shows much more, since it proves that no polynomial at all in this class has a structured strong linearization of the classical size. This is in stark contrast with the strategies used in [43, 44] and Theorem 7.20, which prove the non-existence of structured companion forms by showing just that there exist *some* (in fact, just "a few") matrix polynomials in the corresponding structure classes that do not have any structured strong linearization. Consequently, there is a strong need to develop structured companion forms of grade $\ell > 1$ for even grade/odd size T-skew-symmetric polynomials. Finally, note that Theorem 7.21(b) implicitly poses the open problem of the existence (and construction) of structured companion forms (of grade 1) for even grade/even size T-skew-symmetric polynomials, since the structured strong linearizations that we construct in the proof are certainly not companion forms[4].

Theorem 7.21 (Structured strong linearizations for T-skew-symmetric polynomials)**.**
Let $P(\lambda)$ be any T-skew-symmetric matrix polynomial of even grade k and size $n \times n$ over an arbitrary field.

(a) *If n is odd, then $P(\lambda)$* never *has any T-skew-symmetric strong linearization of size $kn \times kn$.*

(b) *If n is even, then $P(\lambda)$* always *has a T-skew-symmetric strong linearization of size $kn \times kn$.*

Proof. (a) Any T-skew-symmetric matrix polynomial P has even rank [45, Theorem 4.2], so the number of left minimal indices of P is the odd number $q = n - \operatorname{rank} P$. Therefore Corollary 7.19 implies that P does not have any T-skew-symmetric strong linearization of size $kn \times kn$.

(b) The argument given here follows along the lines of the proof of [45, Theorem 6.16], which considers only regular polynomials. According to [45, Theorem 4.2], the Smith form of any T-skew-symmetric P has the form

$$P(\lambda) \sim \operatorname{diag}\big[\, d_1(\lambda), d_1(\lambda), d_2(\lambda), d_2(\lambda), \ldots, d_s(\lambda), d_s(\lambda), \underbrace{0, \ldots, 0}_{n-2s} \big]\,. \tag{7.6}$$

Observe that the number of left (or right) minimal indices of P is $q = n - 2s$, and so is even. Let the left minimal indices of P be

$$\eta_1 \le \cdots \le \eta_q \,, \tag{7.7}$$

which coincide with the right minimal indices of P by Theorem 7.18. Another property of any T-skew-symmetric P [45, Theorem 6.4] is that if ∞ is an eigenvalue of P, then the partial multiplicity sequence of P at ∞ is of the form

$$0 = \beta_1 = \beta_1 = \cdots = \beta_{j-1} = \beta_{j-1} < \beta_j = \beta_j \le \cdots \le \beta_s = \beta_s \,,$$

[4] It is worth noting that some additional structured strong linearizations of size $kn \times kn$ for even grade-(k)/even size-(n) T-skew-symmetric matrix polynomials $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ have been introduced in the recent paper [5]. These linearizations are constructed directly from the coefficients of the polynomial without involving any matrix operations, and so are interesting from an applied point of view. However, they are strong linearizations only when either A_k or A_0 is nonsingular, so they are not structured companion forms in the sense of Definition 5.1 and Remark 5.2(b).

with the same s as in (7.6). To complete the argument, we use the techniques presented in the proof of Theorem 4.10. If $d_1(\lambda) = d_2(\lambda) = \cdots = d_{i-1}(\lambda) = 1$ in (7.6), then define the pencil

$$F(\lambda) \;:=\; \mathcal{C}_{d_i}(\lambda) \oplus \mathcal{C}_{d_{i+1}}(\lambda) \oplus \cdots \oplus \mathcal{C}_{d_s}(\lambda), \tag{7.8}$$

analogous to the construction in (4.5). Let $\omega = \beta_j + \beta_{j+1} + \cdots + \beta_s$, and define also the pencil

$$\Omega(\lambda) \;:=\; I_\omega + \lambda\,(\,J_{\beta_j}(0) \oplus J_{\beta_{j+1}}(0) \oplus \cdots \oplus J_{\beta_s}(0)\,)\,, \tag{7.9}$$

as done in (4.8). Note that $\Omega(\lambda)$ is defined as the empty matrix if ∞ is not an eigenvalue of P. Now recalling that q is even, use the minimal indices from (7.7) to define

$$\eta_i' := \begin{cases} \eta_i, & \text{for } i = 1, \ldots, q/2\,, \\ \eta_i + (k-1), & \text{for } i = q/2+1, \ldots, q\,, \end{cases} \tag{7.10}$$

and use the singular block $S_d(\lambda)$ introduced in (4.7) to define the pencil

$$S(\lambda) := S_{\eta_1'}(\lambda) \oplus \cdots \oplus S_{\eta_q'}(\lambda)\,. \tag{7.11}$$

Note that $S(\lambda)$ is taken to be the empty matrix if $q = 0$, i.e., if P is regular. Finally, define the skew-symmetric pencil

$$L(\lambda) := \begin{bmatrix} 0 & F(\lambda) \\ -F^T(\lambda) & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & \Omega(\lambda) \\ -\Omega^T(\lambda) & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & S(\lambda) \\ -S^T(\lambda) & 0 \end{bmatrix}, \tag{7.12}$$

and observe that $L(\lambda)$ has the same number of left and same number of right minimal indices as $P(\lambda)$, exactly the same finite elementary divisors as $P(\lambda)$, and exactly the same infinite elementary divisors as $P(\lambda)$. Therefore $L(\lambda)$ is a skew-symmetric strong linearization of $P(\lambda)$, by Theorem 4.1. Finally, the size of $L(\lambda)$ is $t \times t$, where $t = \delta_{\text{fin}}(P) + \delta_\infty(P) + \mu(P) + q\,(k-1) + q$. Using the Index Sum Theorem we see that this simplifies to $t = k \;\text{rank}\, P + q\,k = kn$, where the last equality follows from the rank/nullity theorem. □

8 Conclusions

This paper has developed a framework that generalizes the fundamental notion of linearization of matrix polynomials in a direct and simple way from degree one to other low degrees. We expect that the theory presented here will be useful for many researchers in matrix polynomials, providing a consistent and rigorous way to deal with linearizations, quadratifications, and the more general concept of ℓ-ification of arbitrary degree ℓ, as well as their "strong" counterparts, within a unified pattern. The introduced framework is based on two new equivalence relations among matrix polynomials, *extended unimodular equivalence* and *spectral equivalence*. Many properties of these two relations have been established, including their characterization via spectral data. Concrete and simple examples of ℓ-ifications that generalize the classical Frobenius companion forms in a natural way have been introduced, and we have shown how the minimal indices of the original polynomial can be easily recovered from any of these Frobenius-like ℓ-ifications. The study of the properties of ℓ-ifications and companion forms of arbitrary degree has led us to the *Index Sum Theorem*, a fundamental result already known in the literature, but which has been extended here to general matrix polynomials over arbitrary fields. The Index Sum Theorem has provided the key tool enabling us to determine the possible sizes and degrees of strong ℓ-ifications and companion forms, to prove that no companion form can preserve all the minimal indices of every polynomial, and to establish the non-existence of structured companion forms for several classes of structured matrix polynomials of even degree.

Acknowledgments. We wish to thank S. Vologiannidis for pointing out references [27] and [28], and V. Perović for useful discussions on several aspects of this paper. We also thank two anonymous referees for many useful suggestions and constructive comments that have helped us to considerably improve this manuscript.

References

[1] A. Amiraslani, R. M. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29:141–157, 2009.

[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] M. I. Bueno, F. De Terán, and F. M. Dopico. Recovery of eigenvectors and minimal bases of matrix polynomials from generalized Fiedler linearizations. *SIAM J. Matrix Anal. Appl.*, 32:463–483, 2011.

[4] M. I. Bueno and S. Furtado. Palindromic linearizations of a matrix polynomial of odd degree obtained from Fiedler pencils with repetition. *Electron. J. Linear Algebra*, 23:562–577, 2012.

[5] M. I. Bueno and S. Furtado. Structured linearizations from Fiedler pencils with repetition II. *Submitted*. Available at
`http://www.math.ucsb.edu/~mbueno/PUBLICATIONS.html`

[6] R. Byers, V. Mehrmann, and H. Xu. Trimmed linearizations for structured matrix polynomials. *Linear Algebra Appl.*, 429:2373–2400, 2008.

[7] F. De Terán and F. M. Dopico. Low rank perturbation of Kronecker structures without full rank. *SIAM J. Matrix Anal. Appl.*, 29:496–529, 2007.

[8] F. De Terán and F. M. Dopico. Sharp lower bounds for the dimension of linearizations of matrix polynomials. *Electron. J. Linear Algebra*, 17:518–531, 2008.

[9] 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.

[10] F. De Terán, F. M. Dopico, and D. S. Mackey. Fiedler companion linearizations and the recovery of minimal indices. *SIAM J. Matrix Anal. Appl.*, 31:2181–2204, 2010.

[11] F. De Terán, F. M. Dopico, and D. S. Mackey. Palindromic companion forms for matrix polynomials of odd degree. *J. Comp. Appl. Math.*, 236:1464–1480, 2011.

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

[13] F. De Terán, F. M. Dopico, and P. Van Dooren. Matrix polynomials with completely prescribed eigenstructure. *Submitted*. Available at
`http://gauss.uc3m.es/web/personal_web/fdopico/index_sp.html`

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

[15] G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten. *J. Reine Angew. Math. (Crelle)*, 86:146–208, 1878.

[16] F. R. Gantmacher. *Theory of Matrices*. Chelsea, New York, 1959.

[17] S. D. Garvey, P. Lancaster, A. A. Popov, U. Prells, and I. Zaballa. Filters connecting isospectral quadratic systems. *Linear Algebra Appl.*, 438:1497–1516, 2013.

[18] G. M. L. Gladwell. *Inverse problems in vibration*. Kluwer Academic Publishers, Dordrecht, 2nd edition, 2004.

[19] I. Gohberg, M. A. Kaashoek, and P. Lancaster. General theory of regular matrix polynomials and band Toeplitz operators. *Integral Equations Operator Theory*, 11(6):776–882, 1988.

[20] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[21] L. Grammont, N. J. Higham, and F. Tisseur. A framework for analyzing nonlinear eigenproblems and parametrized linear systems. *Linear Algebra Appl.*, 435:623–640, 2011.

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

[23] T.-M. Huang, W.-W. Lin, and W.-S. Su. Palindromic quadratization and structure-preserving algorithm for palindromic matrix polynomials of even degree. *Numer. Math.*, 118:713–735, 2011.

[24] S. Johansson, B. Kågström, and P. Van Dooren. Stratification of full rank polynomial matrices. *Linear Algebra Appl.*, 439(4):1062–1090, 2013.

[25] G. F. Jónsson and S. Vavasis. Solving polynomials with small leading coefficients. *SIAM J. Matrix Anal. Appl.*, 26:400–414, 2005.

[26] T. Kailath. Linear Systems. Prentice-Hall, Englewood Cliffs, NJ, 1980.

[27] N. P. Karampetakis and S. Vologiannidis. Infinite elementary divisor structure-preserving transformations for polynomial matrices. *International J. Appl. Math. Comput. Sci.*, 13:493–503, 2003.

[28] N. P. Karampetakis, S. Vologiannidis, and A. I. G. Vardulakis. A new notion of equivalence for discrete time AR representations. *International J. Control*, 77:584–597, 2004.

[29] L. Kronecker. Über die congruenten Transformationen der bilinearen Formen, 1874. In *Collected Works I*, pages 423–483. Chelsea, New York, 1968.

[30] P. Lancaster. *Lambda-matrices and vibrating systems*. Pergamon Press, Oxford, 1966.

[31] P. Lancaster. Isospectral vibrating systems. Part 1. The spectral method. *Linear Algebra Appl.*, 409:51–69, 2005.

[32] P. Lancaster. Linearizations of regular matrix polynomials. *Electron. J. Linear Algebra*, 17:21–27, 2008.

[33] P. Lancaster and P. Psarrakos. A note on weak and strong linearizations of regular matrix polynomials. MIMS EPrint 2006.72, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2006.

[34] P. Lancaster and U. Prells. Isospectral families of high-order systems. *ZAMM J. Appl. Math. Mech.*, 87(3):219–234, 2007.

[35] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, 2nd edition, 1985.

[36] F. Lazebnik. On systems of linear diophantine equations. *Math. Mag.*, 69(4):261–266, 1996.

[37] D. S. Mackey. Structured Linearizations for Matrix Polynomials. PhD Thesis, The University of Manchester, Manchester, UK, 2006. Available as MIMS EPrint 2006.68, Manchester Institute for Mathematical Sciences.

[38] D. S. Mackey. The continuing influence of Fiedler's work on companion matrices. *Linear Algebra Appl.*, 439(4):810–817, 2013.

[39] D. S. Mackey. Minimal indices and minimal bases via filtrations. Submitted for publication. Also available as MIMS EPrint 2012.82, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012.

[40] D. S. Mackey. A quasi-canonical form for quadratic matrix polynomials. In preparation.

[41] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28:971–1004, 2006.

[42] 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: 1029–1051, 2006.

[43] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Jordan structures of alternating matrix polynomials. *Linear Algebra Appl.*, 432(4):867–891, 2010.

[44] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Smith forms of palindromic matrix polynomials. *Electron. J. Linear Algebra*, 22:53–91, 2011.

[45] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Skew-symmetric matrix polynomials and their Smith forms. *Linear Algebra Appl.*, 438(12):4625–4653, 2013.

[46] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Möbius transformations of matrix polynomials. To appear in *Linear Algebra Appl.* Also available as MIMS EPrint 2014.2, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2014.

[47] W. H. L. Neven and C. Praagman. Column reduction of polynomial matrices. *Linear Algebra Appl.*, 188/189:569–589, 1993.

[48] V. Noferini. The behavior of the complete eigenstructure of a polynomial matrix under a generic rational transformation. *Electron. J. Linear Algebra*, 23:607–624, 2012.

[49] C. Praagman. Invariants of polynomial matrices. *Proceedings of the First European Control Conference, Grenoble 1991.* (I. Landau, Ed.) INRIA, 1274–1277, 1991.

[50] U. Prells and S. D. Garvey. On the class of strictly isospectral systems. *Mechanical Systems and Signal Processing*, 23:2000–2007, 2009.

[51] A. C. Pugh and A. K. Shelton. On a new definition of strict system equivalence. *International J. Control*, 27:657–672, 1978.

[52] H. H. Rosenbrock. Minimal indices in dynamical systems. *Electronics Letters*, 5(25): 658–659, 1969.

[53] H. H. Rosenbrock. Further properties of minimal indices. *Electronics Letters*, 6(14): 450, 1970.

[54] H. H. Rosenbrock. *State-Space and Multivariable Theory.* John Wiley, New York, 1970.

[55] H. J. S. Smith. On systems of linear indeterminate equations and congruences. *Phil. Trans. Roy. Soc. London*, 151:293–326, 1861.

[56] L. Taslaman, F. Tisseur, and I. Zaballa. Triangularizing matrix polynomials. *Linear Algebra Appl.*, 439(7):1679–1699, 2013.

[57] F. Tisseur and I. Zaballa. Finite and infinite elementary divisors of matrix polynomials: A global approach. Available as MIMS EPrint 2012.78, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012.

[58] P. Van Dooren and P. Dewilde. The eigenstructure of an arbitrary polynomial matrix: computational aspects. *Linear Algebra Appl*, 50:545–579, 1983.

[59] S. Vologiannidis and E. N. Antoniou. A permuted factors approach for the linearization of polynomial matrices. *Math. Control Signals Syst.*, 22:317–342, 2011.