Fiedler companion linearizations for rectangular matrix polynomials[☆]

Fernando De Terán[a,*], Froilán M. Dopico[b], D. Steven Mackey[c]

[a] *Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain.*
[b] *Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain.*
[c] *Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA*

Abstract

The development of new classes of linearizations of square matrix polynomials that generalize the classical first and second Frobenius companion forms has attracted much attention in the last decade. Research in this area has two main goals: finding linearizations that retain whatever structure the original polynomial might possess, and improving properties that are essential for accurate numerical computation, such as eigenvalue condition numbers and backward errors. However, all recent progress on linearizations has been restricted to square matrix polynomials. Since rectangular polynomials arise in many applications, it is natural to investigate if the new classes of linearizations can be extended to rectangular polynomials. In this paper, the family of Fiedler linearizations is extended from square to rectangular matrix polynomials, and it is shown that minimal indices and bases of polynomials can be recovered from those of any linearization in this class via the same simple procedures developed previously for square polynomials. Fiedler linearizations are one of the most important classes of linearizations introduced in recent years, but their generalization to rectangular polynomials is nontrivial, and requires a completely different approach to the one used in the square case. To the best of our knowledge, this is the first class of new linearizations that has been generalized to rectangular polynomials.

Keywords: matrix polynomials, Fiedler pencils, linearizations, minimal indices, minimal bases
2010 MSC: 65F15, 15A18, 15A22, 47H60

1. Introduction

We consider in this paper $m \times n$ matrix polynomials with degree $k \geq 2$ of the form

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \,, \qquad A_0, \ldots, A_k \in \mathbb{F}^{m\times n}, \quad A_k \neq 0 \,, \tag{1}$$

where $\mathbb{F}$ is an arbitrary field and λ is a scalar variable in $\mathbb{F}$. Our main focus is on rectangular matrix polynomials, i.e., with $m \neq n$, although new results for square polynomials are also presented. A matrix polynomial $P(\lambda)$ is said to be *singular* either if it is rectangular, or it is square and $\det P(\lambda)$ is identically zero, i.e., if all the coefficients of $\det P(\lambda)$ are zero; otherwise $P(\lambda)$ is *regular*.

Matrix polynomials arise in many applications like systems of differential-algebraic equations, vibration analysis of structural systems, acoustics, fluid-structure interaction problems, computer graphics, signal processing, control theory, and linear system theory [4, 18, 24, 25, 29, 30, 31, 32]. Rectangular matrix polynomials appear mainly in control theory and linear system theory. The magnitudes that are

[☆]This work was partially supported by the Ministerio de Ciencia e Innovación of Spain through grant MTM-2009-09281 (Fernando De Terán and Froilán M. Dopico), by the Comunidad Autónoma de Madrid and Universidad Carlos III de Madrid through grant CCG10-UC3M/ESP-5019 (Fernando De Terán), and by National Science Foundation grants DMS-0713799 and DMS-1016224 (D. S. Mackey).

[*]Corresponding author
Email addresses: fteran@math.uc3m.es (Fernando De Terán), dopico@math.uc3m.es (Froilán M. Dopico), steve.mackey@wmich.edu (D. Steven Mackey)

usually relevant in the applications of regular matrix polynomials are their finite and infinite eigenvalues and the corresponding eigenvectors [18], while in applications of singular polynomials their minimal indices and bases also play an important role [15, 24].

A standard way of dealing, both theoretically and numerically, with a matrix polynomial $P(\lambda)$ is to convert it into an equivalent matrix pencil. This process is known as linearization [18], and is explained in Section 2. The classical approach uses the *first and second Frobenius companion forms* (4) and (5) as linearizations. However, these companion forms usually do not share any algebraic structure that $P(\lambda)$ might have, and their use in numerical computations, via well-established algorithms for pencils [3, 7, 8, 19, 33], may destroy important qualitative features of the eigenvalues/eigenvectors and minimal indices/bases as a consequence of rounding errors. In addition, the condition numbers of the eigenvalues in the Frobenius companion linearizations may be much larger than in $P(\lambda)$, and small eigenvalue backward errors in the linearization do not guarantee small backward errors in the polynomial [21, 22].

These difficulties have motivated intense activity in the last decade towards the development of new classes of linearizations. At first, only linearizations for regular matrix polynomials were considered [1, 2, 23, 27, 28], while more recently square singular polynomials have also received attention [10, 11, 12, 34]. However, all this recent progress on linearizations has been restricted to *square* matrix polynomials. The main goal of this paper is to extend one of the most relevant new classes of linearizations from square to rectangular matrix polynomials. This is the family of Fiedler pencils, which was originally introduced by Fiedler for scalar polynomials in [14], generalized to regular matrix polynomials over $\mathbb{C}$ in [2], and then extended and further analyzed in [11] for both regular and singular square matrix polynomials over an arbitrary field $\mathbb{F}$.

Fiedler pencils of square matrix polynomials $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ enjoy a number of important properties that make them attractive candidates for generalization to rectangular polynomials. They are strong linearizations for any square polynomial, regular or singular, over an arbitrary field, and the coefficients of these pencils are simply constructed as block partitioned matrices whose blocks are either 0, $\pm I$, or $\pm A_i$, $i = 0, 1, \ldots, k$ [11]. This means that they are all companion forms in the sense of [12, Definition 1.1]. Fiedler pencils allow us to very easily recover not only the eigenvalues, but also the eigenvectors, minimal indices, and minimal bases of $P(\lambda)$ from the corresponding magnitudes of the pencil [11]. These pencils can also be generalized to preserve structures of polynomials that are important in applications, like symmetry and palindromicity [2, 12, 34]. No other class of linearizations introduced in recent years simultaneously satisfy all these properties. In fact, for other important classes of new linearizations [27], it is very easy to find pencils that cannot be extended to rectangular matrix polynomials as a consequence of obvious size constraints.

We remark that the extension of Fiedler pencils from square to rectangular matrix polynomials is not direct, since the original definition cannot be applied to rectangular polynomials. This issue is discussed in Section 3.2. Consequently we follow an approach completely different than the one considered in [2, 11, 14] for square polynomials. This approach is based on the construction presented in `Algorithm 2`, which provides the foundation for the main Definition 3.8. With this definition in hand, and after considerable technical effort, we prove in Theorem 4.5 that Fiedler pencils of rectangular matrix polynomials are always strong linearizations over arbitrary fields, again using new techniques. Finally, simple recovery procedures for minimal indices and bases are presented in Corollaries 5.4 and 5.7. These recovery rules are essentially the same as the ones derived for square polynomials in [11]. Although the new proofs and definitions may seem complicated, we emphasize that the key idea is very simple: we perform the same operations that we would do in the square case, but proving that the rectangular matrices that appear are always conformable for multiplication. This requires a substantial amount of care. Another essential difference between Fiedler pencils for rectangular and square polynomials is that when a polynomial $P(\lambda)$ is rectangular, there are always associated Fiedler pencils of several different sizes. Indeed the two Frobenius companion forms are always the Fiedler pencils with largest and smallest sizes, while the other Fiedler pencils have intermediate sizes. This always makes one of the two Frobenius companion forms a privileged choice to use when working with rectangular matrix polynomials, although the low band-width structure of some other Fiedler pencils might make them preferable in certain situations.

The paper is organized as follows. In Section 2 we introduce the basic definitions and notation used throughout the paper. In Section 3 we recall first the notion of Fiedler pencils for square matrix polynomials, then present an algorithm to construct these pencils in a manner that readily generalizes to rectangular matrix polynomials. It is by means of this algorithm that we are then able to extend the notion of Fiedler pencils to rectangular polynomials. In the last part of Section 3, we establish the relationship between the reversal of a polynomial and the reversal of any of its Fiedler pencils (Theorem 3.14). This relationship is needed to prove that Fiedler pencils of rectangular polynomials are always strong linearizations in Section 4. Section 5 establishes very simple formulae for the recovery of the minimal indices and bases of a matrix polynomial from the minimal indices and bases of any of its Fiedler pencils. Finally, Section 6 gives some conclusions and describes possible future work motivated by the results in this paper.

2. Basic notation and definitions

We present in this section some basic concepts related to rectangular matrix polynomials. The reader can find more information in [10, Section 2] and [11, Section 2], where these concepts are presented in greater detail for square polynomials. In the rest of the paper we adopt the following notation: 0_d and I_d are used to denote the $d \times d$ zero and identity matrices, respectively. If there is no risk of confusion, then the sizes are not indicated and we simply write 0 or I. Two $m \times n$ matrix polynomials $P(\lambda)$ and $Q(\lambda)$ are *strictly equivalent* if there exist two constant nonsingular matrices E and F such that $P(\lambda) = E\,Q(\lambda)\,F$. We emphasize that any equation in this paper involving expressions in λ is to be understood as a formal algebraic identity, and not just as an equality of functions on the field $\mathbb{F}$. For finite fields $\mathbb{F}$ this distinction is important, and we will always intend the stronger meaning of a formal algebraic identity.

Let $\mathbb{F}(\lambda)$ denote the field of rational functions with coefficients in $\mathbb{F}$, so that $\mathbb{F}(\lambda)^{n\times 1}$ is the vector space of column n-tuples with entries in $\mathbb{F}(\lambda)$. The *normal rank* of a matrix polynomial $P(\lambda)$, denoted $\operatorname{nrank} P(\lambda)$, is the rank of $P(\lambda)$ considered as a matrix with entries in $\mathbb{F}(\lambda)$, or equivalently, the size of the largest non-identically zero minor of $P(\lambda)$ [16]. A *finite eigenvalue* of $P(\lambda)$ is an element $\lambda_0 \in \mathbb{F}$ such that

$$\operatorname{rank} P(\lambda_0) < \operatorname{nrank} P(\lambda)\,.$$

We say that $P(\lambda)$ with degree k has an *infinite eigenvalue* if the *reversal polynomial*

$$\operatorname{rev} P(\lambda) := \lambda^k P(1/\lambda) = \sum_{i=0}^{k} \lambda^i A_{k-i} \tag{2}$$

has zero as an eigenvalue.

An $m \times n$ singular matrix polynomial $P(\lambda)$ may have *right* (*column*) and/or *left* (*row*) *null vectors*, that is, vectors $x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1}$ and $y(\lambda)^T \in \mathbb{F}(\lambda)^{1\times m}$ such that $P(\lambda)x(\lambda) \equiv 0$ and $y(\lambda)^T P(\lambda) \equiv 0$, respectively, where $y(\lambda)^T$ denotes the transpose of $y(\lambda)$. This leads to the following definition.

Definition 2.1. *The* right *and* left nullspaces *of the $m \times n$ matrix polynomial $P(\lambda)$, denoted by $\mathcal{N}_r(P)$ and $\mathcal{N}_\ell(P)$, respectively, are the following subspaces:*

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

Note that the identities $\operatorname{nrank} P(\lambda) = n - \dim \mathcal{N}_r(P) = m - \dim \mathcal{N}_\ell(P)$ hold.

It is well known that the *elementary divisors* of $P(\lambda)$ corresponding to its finite eigenvalues, as well as the dimensions of $\mathcal{N}_r(P)$ and $\mathcal{N}_\ell(P)$, are invariant under *unimodular equivalence* [16], i.e., under pre- and post-multiplication of $P(\lambda)$ by *unimodular matrices* (square matrix polynomials with nonzero constant determinant). The elementary divisors of $P(\lambda)$ corresponding to the infinite eigenvalue are defined as the elementary divisors corresponding to the zero eigenvalue of the reversal polynomial [20, Definition 1], and may be altered by unimodular equivalence [26].

Next we define linearizations and strong linearizations of matrix polynomials.

Definition 2.2. *A matrix pencil $L(\lambda) = \lambda X + Y$ is a* linearization *of an $m \times n$ matrix polynomial $P(\lambda)$, if for some $s \geq 0$ there exist unimodular matrices $U(\lambda)$ and $V(\lambda)$ such that*

$$U(\lambda)L(\lambda)V(\lambda) = \left[\begin{array}{c|c} I_s & 0 \\ \hline 0 & P(\lambda) \end{array}\right], \tag{3}$$

i.e., if $L(\lambda)$ *is* unimodularly equivalent *to* $\operatorname{diag}\big[I_s, P(\lambda)\big]$. *A linearization* $L(\lambda)$ *is called a* strong linearization *if* $\operatorname{rev} L(\lambda)$ *is also a linearization of* $\operatorname{rev} P(\lambda)$.

The definition of linearization was introduced in [18], while the notion of strong linearization was introduced in [17] and later named in [26]. In [17, 18, 26] only regular (square) matrix polynomials were considered. These definitions were extended to any matrix polynomial in [9], that is, including rectangular and square (regular or singular) polynomials. The original definition in [18, p. 12] for $n \times n$ regular polynomials considers linearizations with sizes $(n+s)\times(n+s)$ and $s \geq 0$ arbitrary. However, for $n \times n$ matrix polynomials with degree k, the definition of linearization presented in most references fixes the size of the linearizations to be $nk \times nk$, which corresponds to $s = (k-1)n$ in Definition 2.2. Perhaps the reason for this commonly encountered size restriction lies in the fact that all linearizations of a matrix polynomial with nonsingular leading coefficient have sizes at least $nk \times nk$ and that, moreover, all strong linearizations of regular matrix polynomials have size exactly $nk \times nk$ [9]. However, if $P(\lambda)$ is an $n \times n$ singular polynomial with degree k, then there are strong linearizations with size strictly less than $nk \times nk$ [9] that have interest in applications [6]. For these and other reasons, the size of the matrix pencil $L(\lambda)$ in Definition 2.2 is not fixed. In fact, when $P(\lambda)$ is *rectangular* there always exist strong linearizations for $P(\lambda)$ with *different sizes*. This is illustrated by the two most common linearizations used in practice, i.e., the *first* and *second Frobenius companion forms*, which for the $n \times n$ polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ are

$$C_1(\lambda) := \lambda \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{bmatrix} \tag{4}$$

and

$$C_2(\lambda) := \lambda \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{k-1} & -I_n & \cdots & 0 \\ A_{k-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \cdots & 0 \end{bmatrix}, \tag{5}$$

and both have size $nk \times nk$. However, if $P(\lambda)$ is rectangular with size $m \times n$, then the identity matrices in $C_2(\lambda)$ must now have size $m \times m$. Thus $C_1(\lambda)$ has size $\big(m + (k-1)n\big) \times kn$, while $C_2(\lambda)$ has size $km \times \big((k-1)m + n\big)$. Clearly these sizes are different when $m \neq n$.

It is well known that strong linearizations are relevant in the study of both regular and singular square matrix polynomials, because they are the only matrix pencils preserving the dimension of the left and right nullspaces as well as the finite and infinite elementary divisors of $P(\lambda)$ [10, Lemma 2.3]. Since the arguments used to prove this fact do not depend on $P(\lambda)$ being square or rectangular (see the proof of Lemma 2.3 in [10]), the same result holds in the rectangular case. Thus for rectangular matrix polynomials we have the following analogue of Lemma 2.3 in [10].

Lemma 2.3. *Let* $P(\lambda)$ *be an* $m \times n$ *matrix polynomial and let* $L(\lambda)$ *be an* $(m+s)\times(n+s)$ *matrix pencil for some* $s \geq 0$, *and consider the following conditions on* $L(\lambda)$ *and* $P(\lambda)$:

(a) $\dim \mathcal{N}_r(L) = \dim \mathcal{N}_r(P)$,

(b) $L(\lambda)$ *and* $P(\lambda)$ *have exactly the same finite elementary divisors,*

(c) $L(\lambda)$ *and* $P(\lambda)$ *have exactly the same infinite elementary divisors.*

Then $L(\lambda)$ *is*

- *a linearization of* $P(\lambda)$ *if and only if conditions* (a) *and* (b) *hold,*
- *a strong linearization of* $P(\lambda)$ *if and only if conditions* (a), (b) *and* (c) *hold.*

Note that condition (a) in Lemma 2.3 is equivalent to $\dim \mathcal{N}_\ell(L) = \dim \mathcal{N}_\ell(P)$.

A *vector polynomial* is a vector whose entries are polynomials in the variable λ. For any subspace of $\mathbb{F}(\lambda)^{n\times 1}$, it is always possible to find a basis consisting entirely of vector polynomials. The *degree* of a vector polynomial is the greatest degree of its components, and the *order* of a polynomial basis is defined as the sum of the degrees of its vectors [15, p. 494]. Then the following definition makes sense.

Definition 2.4. [15] *Let $\mathcal{V}$ be a subspace of $\mathbb{F}(\lambda)^{n\times 1}$. 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 [15] that for any given subspace $\mathcal{V}$ of $\mathbb{F}(\lambda)^{n\times 1}$, 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 $\mathcal{V}$. Given a matrix polynomial $P(\lambda)$, the minimal indices and bases of the subspace $\mathcal{N}_r(P)$ are called the right minimal indices and bases of $P(\lambda)$, while the minimal indices and bases of $\mathcal{N}_\ell(P)$ are called the left minimal indices and bases of $P(\lambda)$. These magnitudes have important applications in Linear System Theory [24].

The left (right) minimal indices of a matrix pencil can be read off from the sizes of the left (right) singular blocks of the Kronecker canonical form of the pencil [16, Chap. XII]. Consequently, the minimal indices of a pencil can be stably computed via the GUPTRI form [7, 8, 13, 33]. Therefore it is natural to look for relationships between the minimal indices of a singular matrix polynomial $P(\lambda)$ and the minimal indices of a given linearization of $P(\lambda)$, since this would lead to a numerical method for computing the minimal indices of $P(\lambda)$. In the case of square singular matrix polynomials, such relationships were found in [10] for the pencils introduced in [27], in [11] for Fiedler pencils, and in [5] for generalized Fiedler pencils. In the case of Fiedler pencils of rectangular polynomials, we will develop analogous relationships in Section 5.

3. Fiedler pencils: definition and structural properties

In this section we first recall the notion of *Fiedler pencils* for square matrix polynomials, introduced in [2] and named later in [11], and in Section 3.1 we present `Algorithm 1` to construct these pencils. In Section 3.2 we extend the notion of Fiedler pencils to rectangular $m \times n$ matrix polynomials by means of `Algorithm 2`, which is a slight modification of `Algorithm 1`. This motivates the main definition in this paper, Definition 3.8, which includes the one for the square case by just considering $n = m$. Also in Section 3.2 we present some structural properties of Fiedler pencils that will be used later. Finally in Section 3.3 we show the connection between the reversal of a Fiedler pencil and the reversal of its associated polynomial.

To introduce the Fiedler pencils of an $n \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, we need the following block-partitioned matrices:

$$M_k := \begin{bmatrix} A_k & \\ & I_{(k-1)n} \end{bmatrix}, \quad M_0 := \begin{bmatrix} I_{(k-1)n} & \\ & -A_0 \end{bmatrix}, \tag{6}$$

and

$$M_i := \begin{bmatrix} I_{(k-i-1)n} & & & \\ & -A_i & I_n & \\ & I_n & 0 & \\ & & & I_{(i-1)n} \end{bmatrix}, \quad i = 1, \ldots, k-1\,. \tag{7}$$

Notice that

$$M_i M_j = M_j M_i \quad \text{for } \ |i - j| \neq 1. \tag{8}$$

Now we introduce Fiedler pencils in the same way as in [11].

Definition 3.1 (Fiedler Pencils for square matrix polynomials)**.** *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$, and let M_i, $i = 0, 1, \ldots, k$, be the matrices defined in* (6) *and* (7)*. Given any bijection $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$, the* Fiedler pencil of $P(\lambda)$ associated with σ *is the $nk \times nk$ matrix pencil*

$$F_\sigma(\lambda) := \lambda M_k - M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(k)}\,. \tag{9}$$

Note that $\sigma(i)$ describes the position of the factor M_i in the product $M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(k)}$ defining the zero-degree term in (9)*: i.e., $\sigma(i) = j$ means that M_i is the jth factor in the product. For brevity, we denote this product by*

$$M_\sigma := M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(k)}\,, \tag{10}$$

so that $F_\sigma(\lambda) := \lambda M_k - M_\sigma$.

As in [11], sometimes we will write the bijection σ using the array notation $\sigma = (\sigma(0), \sigma(1), \ldots, \sigma(k-1))$. Unless otherwise stated, the matrices M_i, $i = 0, \ldots, k$, M_σ, and the Fiedler pencil $F_\sigma(\lambda)$ refer to the matrix polynomial $P(\lambda)$ in (1). When necessary, we will explicitly indicate the dependence on a certain polynomial $Q(\lambda)$ by writing $M_i(Q)$, $M_\sigma(Q)$ and $F_\sigma(Q)$. In this situation, the dependence on λ is dropped in the Fiedler pencil $F_\sigma(Q)$ for simplicity. Since matrix polynomials will always be denoted by capital letters, there is no risk of confusion between $F_\sigma(\lambda)$ and $F_\sigma(Q)$.

The set of Fiedler pencils includes the first and second companion forms [18, 11]. More precisely, the first companion form corresponds to the bijection $\sigma_1 = (k, k-1, \ldots, 2, 1)$ and the second to the bijection $\sigma_2 = (1, 2, \ldots, k-1, k)$. Other significant Fiedler pencils are the pentadiagonal Fiedler pencils that are described in detail in [11, Example 3.2].

It is shown in [11] that the relative positions of the matrices M_i and M_{i+1}, for $i = 0, 1, \ldots, k-2$, in the product M_σ determine most of the relevant properties of the Fiedler pencil $F_\sigma(\lambda)$. This motivates Definition 3.2, that was introduced in [11, Definition 3.3].

Definition 3.2. *Let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection.*

(a) *For $i = 0, \ldots, k-2$, we say that σ has a* consecution *at i if $\sigma(i) < \sigma(i+1)$, and that σ has an* inversion *at i if $\sigma(i) > \sigma(i+1)$.*

(b) *Denote by $\mathfrak{c}(\sigma)$ the total number of consecutions in σ, and by $\mathfrak{i}(\sigma)$ the total number of inversions in σ.*

(c) *For $i \leq j$, we denote by $\mathfrak{c}(\sigma(i:j))$ the total number of consecutions that σ has at $i, i+1, \ldots, j$, and by $\mathfrak{i}(\sigma(i:j))$ the total number of inversions that σ has at $i, i+1, \ldots, j$. Observe that $\mathfrak{c}(\sigma) = \mathfrak{c}(\sigma(0:k-2))$ and $\mathfrak{i}(\sigma) = \mathfrak{i}(\sigma(0:k-2))$.*

(d) *The consecution-inversion structure sequence of σ, denoted by* CISS(σ)*, is the tuple $(c_1, i_1, c_2, i_2, \ldots, c_\ell, i_\ell)$, where σ has c_1 consecutive consecutions at $0, 1, \ldots, c_1 - 1$; i_1 consecutive inversions at $c_1, c_1+1, \ldots, c_1+i_1-1$ and so on, up to i_ℓ inversions at $k-1-i_\ell, \ldots, k-2$.*

We want to point out that, though the notions introduced in Definition 3.2 depend only on the bijection σ and not on the Fiedler pencil $F_\sigma(\lambda)$, they are closely related to the definition of $F_\sigma(\lambda)$, as is shown in the following remark.

Remark 3.3. The following simple observations on Definition 3.2 will be used freely.

1. σ has a consecution at i if and only if M_i is to the *left* of M_{i+1} in M_σ, while σ has an inversion at i if and only if M_i is to the right of M_{i+1} in M_σ.
2. Either c_1 or i_ℓ in CISS(σ) may be zero (in the first case σ has an inversion at 0, in the second it has a consecution at $k-2$), but $i_1, c_2, i_2, \ldots, i_{\ell-1}, c_\ell$ are all strictly positive. These conditions uniquely determine CISS(σ) and, in particular, the parameter ℓ.
3. $\mathfrak{c}(\sigma) = \sum_{j=1}^{\ell} c_j$, $\mathfrak{i}(\sigma) = \sum_{j=1}^{\ell} i_j$, and $\mathfrak{c}(\sigma) + \mathfrak{i}(\sigma) = k-1$.

The reader may find in [11, Example 3.5] explicit examples of CISS(σ) for some relevant Fiedler pencils.

3.1. A multiplication free algorithm to construct Fiedler pencils of square matrix polynomials

We focus only on how to construct the zero-degree term M_σ in the Fiedler pencil (9), since the first-degree term is already known. The obvious option is to directly perform the multiplication of all factors, but this is not convenient if the degree is large.[1] Theorem 3.4 shows how to construct Fiedler pencils without performing any arithmetic operation. *Throughout this paper, we will use MATLAB notation for submatrices on block indices*; that is, if A is a matrix partitioned into blocks, then $A(i:j,:)$ indicates the submatrix of A consisting of *block rows* i through j and $A(:,k:l)$ indicates the submatrix of A consisting of *block columns* k through l.

[1]Polynomials with large degrees may appear, for instance, in the computation of the roots of scalar polynomials as the eigenvalues of a Fiedler pencil [14].

Theorem 3.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $n \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection, and let M_σ be the zero-degree term of the Fiedler pencil of $P(\lambda)$ associated with σ. Consider the matrices $W_0, W_1, \ldots, W_{k-2}$ constructed by* `Algorithm 1` *below, partitioned, respectively, into $2 \times 2, 3 \times 3, \ldots, k \times k$ blocks of size $n \times n$. Then* `Algorithm 1` *constructs M_σ, more precisely, $M_\sigma = W_{k-2}$.*

`Algorithm 1.` *Given $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ with size $n \times n$ and a bijection σ, the following algorithm constructs M_σ.*

if σ has a consecution at 0 then

$$W_0 = \begin{bmatrix} -A_1 & I_n \\ -A_0 & 0 \end{bmatrix}$$

else

$$W_0 = \begin{bmatrix} -A_1 & -A_0 \\ I_n & 0 \end{bmatrix}$$

endif

for $i = 1 : k-2$

if σ has a consecution at i then

$$W_i = \begin{bmatrix} -A_{i+1} & I_n & 0 \\ W_{i-1}(:,1) & 0 & W_{i-1}(:,2:i+1) \end{bmatrix}$$

else

$$W_i = \begin{bmatrix} -A_{i+1} & W_{i-1}(1,:) \\ I_n & 0 \\ 0 & W_{i-1}(2:i+1,:) \end{bmatrix}$$

endif

endfor

$M_\sigma = W_{k-2}$

Proof. The proof proceeds by induction on the degree k. The result is obvious for $k = 2$, because in this case there are only two possible options for M_σ, namely, $M_\sigma = M_0 M_1$ if σ has a consecution at 0 or $M_\sigma = M_1 M_0$ if σ has an inversion at 0. A direct computation shows that

$$M_0 M_1 = \begin{bmatrix} -A_1 & I_n \\ -A_0 & 0 \end{bmatrix} \quad \text{and} \quad M_1 M_0 = \begin{bmatrix} -A_1 & -A_0 \\ I_n & 0 \end{bmatrix} \quad \text{for } k = 2, \tag{11}$$

and the result follows.

Assume now that the result is valid for matrix polynomials of degree $k - 1 \geq 2$, and let us prove it for the polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ and the bijection $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$. Note first that the matrices $M_i(P)$ defined in (6) and (7) for $P(\lambda)$ satisfy

$$M_i(P) = \mathrm{diag}(I_n, M_i(Q)), \quad \text{for } i = 0, \ldots, k-2, \tag{12}$$

where $M_i(Q)$ are the $n(k-1) \times n(k-1)$ matrices corresponding to the polynomial $Q(\lambda) = \sum_{i=0}^{k-1} \lambda^i A_i$. We need to distinguish two cases in the proof.

Case 1. If σ has a consecution at $k-2$, then the commutativity relations (8) of the M_i matrices allow us to write

$$M_\sigma(P) = M_{i_0}(P) \cdots M_{i_{k-2}}(P) M_{k-1}(P),$$

where $(i_0, i_1, \ldots, i_{k-2})$ is a permutation of $(0, 1, \ldots, k-2)$. By using (12), we can write

$$M_\sigma(P) = \mathrm{diag}(I_n, M_{\widetilde{\sigma}}(Q))\, M_{k-1}(P), \tag{13}$$

where $\widetilde{\sigma} : \{0, 1, \ldots, k-2\} \to \{1, \ldots, k-1\}$ is a bijection such that, for $i = 0, \ldots, k-3$, $\widetilde{\sigma}$ has a consecution (resp., inversion) at i if and only if σ has a consecution (resp., inversion) at i. Therefore, by the induction hypothesis, $M_{\widetilde{\sigma}}(Q) = W_{k-3}$. Finally, we perform the simple block product in (13) as follows

$$\begin{aligned} M_\sigma(P) &= \begin{bmatrix} I_n & 0_n & 0 \\ 0 & W_{k-3}(:,1) & W_{k-3}(:,2:k-1) \end{bmatrix} \begin{bmatrix} -A_{k-1} & I_n & \\ I_n & 0_n & \\ & & I_{(k-2)n} \end{bmatrix} \\ &= \begin{bmatrix} -A_{k-1} & I_n & 0 \\ W_{k-3}(:,1) & 0 & W_{k-3}(:,2:k-1) \end{bmatrix}, \end{aligned}$$

which is precisely the matrix W_{k-2} constructed by `Algorithm 1` when σ has a consecution at $k-2$.

Case 2. If σ has an inversion at $k-2$ the proof is similar, but with $M_{k-1}(P)$ placed on the left, i.e.,

$$M_\sigma(P) = M_{k-1}(P) M_{i_0}(P) \cdots M_{i_{k-2}}(P) = M_{k-1}(P)\,\mathrm{diag}(I_n, M_{\widetilde{\sigma}}(Q)).$$

□

3.2. Fiedler pencils of rectangular matrix polynomials

The extension of equation (9) to a rectangular $m \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ presents difficulties, because it is not clear how to define the sizes of all the identity blocks on the main block diagonal of the factors M_i. A tentative approach is simply to try to choose the sizes of the diagonal identities in both (6) and (7) so that all the factors in (10) are conformal for multiplication (notice that the non-diagonal identity blocks in the central 2×2 block submatrix of (7) are determined by the size of $A_i \in \mathbb{F}^{m \times n}$). This can be done, but it is not immediate and is cumbersome, because the presence of the block $-A_i$ in the matrix M_i imposes restrictions on the sizes of the diagonal identity blocks of the factors to both the left and the right of M_i in the product defining M_σ. To proceed in this way requires a very careful determination of the sizes of the M_i matrices, as well as the sizes of all the identity blocks in each M_i. Furthermore, these sizes all depend on the position of the M_i factor in the product defining M_σ. In other words, the M_i factors themselves are dependent on the choice of bijection σ. These issues are better explained with an example.

Example 3.5. *With $A_i \in \mathbb{F}^{m \times n}$, let $P(\lambda) = A_0 + \lambda A_1 + \lambda^2 A_2 + \lambda^3 A_3$ be a matrix polynomial of degree 3, and let $\sigma_1 = (1, 3, 2)$ and $\sigma_2 = (2, 3, 1)$ be bijections from $\{0, 1, 2\}$ to $\{1, 2, 3\}$. Let us see how to give a meaning to the symbolic expressions*

$$F_{\sigma_1}(\lambda) = \lambda M_3 - M_0 M_2 M_1 \quad \text{and} \quad F_{\sigma_2}(\lambda) = \lambda M_3 - M_2' M_0' M_1'$$

by properly defining the factors in the Fiedler pencils for $P(\lambda)$ associated with the bijections σ_1 and σ_2. When $P(\lambda)$ is square ($n = m$), the commutativity relations (8) immediately imply that $F_{\sigma_1}(\lambda) = F_{\sigma_2}(\lambda)$. However, if $m \neq n$, then the factors in the zero degree term of $F_{\sigma_1}(\lambda)$ will be conformal for multiplication if and only if they are

$$M_0 = \begin{bmatrix} I_m & 0 & 0 \\ 0 & I_n & 0 \\ 0 & 0 & -A_0 \end{bmatrix}, \quad M_1 = \begin{bmatrix} I_n & 0 & 0 \\ 0 & -A_1 & I_m \\ 0 & I_n & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} -A_2 & I_m & 0 \\ I_n & 0 & 0 \\ 0 & 0 & I_n \end{bmatrix},$$

while the factors in the zero degree term of $F_{\sigma_2}(\lambda)$ are conformal for multiplication if and only if they are

$$M_0' = \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_m & 0 \\ 0 & 0 & -A_0 \end{bmatrix}, \quad M_1' = \begin{bmatrix} I_n & 0 & 0 \\ 0 & -A_1 & I_m \\ 0 & I_n & 0 \end{bmatrix}, \quad M_2' = \begin{bmatrix} -A_2 & I_m & 0 \\ I_n & 0 & 0 \\ 0 & 0 & I_m \end{bmatrix}.$$

Note that the size of M_2 is different than the size of M_2'. However, the reader is invited to check that $F_{\sigma_1}(\lambda) = F_{\sigma_2}(\lambda)$ still holds. This example shows that defining Fiedler pencils for rectangular polynomials in a similar way as in the square case would force the sizes of the M_i matrices to depend on the specific choice of bijection σ. It is easy to devise examples of rectangular matrix polynomials of degree higher than 3 and bijections σ_1 and σ_2 where the sizes differ for more than one factor M_i.

A first option to extend Fiedler pencils from square to rectangular polynomials that is not affected by the difficulties illustrated in Example 3.5 would be the following. In the square case, use the commutativity relations (8) to order the factors M_i in M_σ (10) in a certain canonical order that is exactly the same for all Fiedler pencils with the same CISS(σ). (Note that two Fiedler pencils with the same CISS(σ) are in fact the same pencil, by Theorem 3.4). One possible order may be found in [2, eqn. (2.9)]. Then use this order and force the conformability of all M_i factors for multiplication in this canonical order, by properly choosing the sizes of their identity blocks, to extend the Fiedler pencil to rectangular matrix polynomials. Again, this can be done, but it requires one to prove, for each different CISS(σ), that the sizes of the M_i factors can always be properly chosen, and to determine these sizes. This is not obvious, and is certainly tedious. In addition, the reader may easily check that the sizes of the M_i factors may be different for different CISS(σ), so that this unpleasant issue still remains.

Another option for extending Fiedler pencils from square to rectangular polynomials, bypassing all the difficulties mentioned above, is simply to avoid the use of the factors M_i in the rectangular case altogether. To this end, we might start by symbolically performing the product defining M_σ in (10) in the square case, in order to obtain an explicit expression for the block-entries of M_σ in terms of the coefficients A_i of the polynomial $P(\lambda)$. This can be done by using CISS(σ), although it is rather complicated and requires a cumbersome notation. Once this explicit expression is known, we would then replace the square $n \times n$ blocks A_i, $i = 0, 1, \ldots, k-1$, by rectangular $m \times n$ blocks A_i, and check that the sizes of the block rows and block columns fit together properly with an appropriate assignment of either a size $n \times n$ or $m \times m$ to every identity block that appears in M_σ. Unfortunately this again requires a tedious proof.

Therefore, we will follow a simpler approach based on adapting `Algorithm 1` to rectangular matrix polynomials. This approach is developed in Theorem 3.6 and Definition 3.8, and is, in fact, equivalent to the process described above of obtaining an explicit expression of the block-entries of M_σ in terms of the coefficients A_i. Note that in `Algorithm 2` *we again use MATLAB notation for submatrices on block indices.*

Theorem 3.6. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection, and consider the following algorithm:*

> `Algorithm 2.` *Given $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ with size $m \times n$ and a bijection σ, the following algorithm constructs a sequence of matrices $\{W_0, W_1, \ldots, W_{k-2}\}$, where each matrix W_i for $i = 1, 2, \ldots, k-2$ is partitioned into blocks in such a way that the blocks of W_{i-1} are blocks of W_i.*
>
> if σ has a consecution at 0 then
>
> $$W_0 = \begin{bmatrix} -A_1 & I_m \\ -A_0 & 0 \end{bmatrix}$$
>
> else
>
> $$W_0 = \begin{bmatrix} -A_1 & -A_0 \\ I_n & 0 \end{bmatrix}$$
>
> endif
>
> for $i = 1 : k-2$
>
>   if σ has a consecution at i then
>
> $$W_i = \begin{bmatrix} -A_{i+1} & I_m & 0 \\ W_{i-1}(:,1) & 0 & W_{i-1}(:, 2:i+1) \end{bmatrix}$$
>
>   else
>
> $$W_i = \begin{bmatrix} -A_{i+1} & W_{i-1}(1,:) \\ I_n & 0 \\ 0 & W_{i-1}(2:i+1,:) \end{bmatrix}$$
>
>   endif
>
> endfor

Then the matrices $W_0, W_1, \ldots, W_{k-2}$ are partitioned in $2 \times 2, 3 \times 3, \ldots, k \times k$ blocks, respectively, and satisfy the following properties:

(a) *The size of W_i is*

$$\big(m + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big) \times \big(n + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big).$$

(b) *The first diagonal block of W_i is $-A_{i+1}$, and so has size $m \times n$. The rest of the diagonal blocks of W_i are square zero matrices, more precisely*

$$W_i(i+2-j, i+2-j) = \begin{cases} 0_m & \text{if } \sigma \text{ has a consecution at } j \\ 0_n & \text{if } \sigma \text{ has an inversion at } j \end{cases}, \quad \text{for } j = 0, 1, \ldots, i.$$

Proof. The proof is elementary. We simply sketch the main points. First, notice that the matrix W_0 is well-defined in `Algorithm 2`. Therefore W_1 is also well-defined, either when σ has a consecution at 1 or an inversion at 1, because in both cases $W_1(1,1) = -A_2$, $W_0(:,1)$ has n columns, and $W_0(1,:)$ has m rows. The same argument can be applied inductively to show that $W_2, \ldots, W_{k-2}$ are also well-defined. The fact that W_i is partitioned into $(i+2)\times(i+2)$ blocks is true by definition for W_0, and for the rest of the matrices in the sequence it follows from the fact that one block row and one block column are added in each step of the "for" loop of `Algorithm 2`. Part (a) is again true for W_0, and for obtaining the result for the rest of the matrices in the sequence note that: (1) if σ has a consecution at i, then W_i has m rows and m columns more than W_{i-1}; (2) if σ has an inversion at i, then W_i has n rows and n columns more than W_{i-1}. Finally, let us prove part (b). The result is true for W_0. For the rest of the matrices in the sequence, assume that it is true for W_{i-1} and let us prove it for W_i. Note that by construction $W_i(1,1) = -A_{i+1}$ and

$$W_i(2,2) = \begin{cases} 0_m & \text{if } \sigma \text{ has a consecution at } i \\ 0_n & \text{if } \sigma \text{ has an inversion at } i \end{cases},$$

which is part (b) for $j = i$. Observe also that

$$W_i(3:i+2, 3:i+2) = W_{i-1}(2:i+1, 2:i+1),$$

which implies $W_i(i+2-j, i+2-j) = W_{i-1}((i-1)+2-j, (i-1)+2-j)$ for $j = 0, 1, \ldots, i-1$. This proves the result since we are assuming that the result is true for W_{i-1}. ☐

Remark 3.7. In part (b) of Theorem 3.6 we assume, as in the rest of the paper, that the block indices of W_i run from 1 to $i+2$. Thus the diagonal blocks of W_i are $W_i(1,1), \ldots, W_i(i+2, i+2)$. If we let the block indices of W_i run from $k-i-1$ to k, the result in part (b) is expressed as

$$W_i(k-j, k-j) = \begin{cases} 0_m & \text{if } \sigma \text{ has a consecution at } j \\ 0_n & \text{if } \sigma \text{ has an inversion at } j \end{cases}, \quad \text{for } j = 0, 1, \ldots, i,$$

which shows that the sizes of these blocks only depend on j and not on i, as long as $0 \le j \le i$.

Observe that `Algorithm 2` differs from `Algorithm 1` only in the sizes of the identity blocks added at each step of the construction, that are chosen to fit the size $m \times n$ of the coefficients of the polynomial $P(\lambda)$. This fact and Theorem 3.4 motivate Definition 3.8, which is the main definition in this paper.

Definition 3.8 (Fiedler Pencils for rectangular matrix polynomials)**.** *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m\times n$ matrix polynomial with degree $k \ge 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection, and denote by M_σ the last matrix of the sequence constructed by* `Algorithm 2`*, that is,*

$$M_\sigma := W_{k-2}\,.$$

The Fiedler pencil of $P(\lambda)$ associated with σ is the $\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big)$ matrix pencil

$$F_\sigma(\lambda) := \lambda \begin{bmatrix} A_k & \\ & I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)} \end{bmatrix} - M_\sigma\,. \tag{14}$$

Remark 3.9. Some remarks on Definition 3.8 may be useful for the reader.

1. The leading coefficient $\left[\begin{smallmatrix} A_k & \\ & I \end{smallmatrix}\right]$ of the Fiedler pencil $F_\sigma(\lambda)$ introduced in Definition 3.8 has the same structure as the matrix M_k in (6), but the size of the block diagonal identity is different when $m \ne n$.
2. If $m \ne n$, then there are Fiedler pencils associated with $P(\lambda)$ with several different sizes, because the sum $\mathfrak{c}(\sigma) + \mathfrak{i}(\sigma) = k-1$ is fixed for all σ, and so different pairs $(\mathfrak{c}(\sigma), \mathfrak{i}(\sigma))$ produce different sizes of $F_\sigma(\lambda)$. For instance, if $m > n$, then the Fiedler pencil with smallest size corresponds to $\mathfrak{c}(\sigma) = 0$, i.e., to the first companion form, and the one with largest size corresponds to $\mathfrak{i}(\sigma) = 0$, i.e., to the second companion form [11]. If $n > m$, then the opposite situation holds.

3. In Theorem 3.6 and Definition 3.8 we use a bijection σ only for the purpose of keeping a parallelism with the standard definition of Fiedler pencils for square polynomials. However, a bijection is not really needed, since we do not actually use the factors M_i anywhere in our definition. Observe that `Algorithm 2` only needs a sequence of decisions that we have identified with σ having a consecution or inversion.
4. A comparison between `Algorithms 1` and `2` shows that, for the same bijection σ, Fiedler pencils of square and rectangular matrix polynomials look symbolically the same, except for the sizes of the identity blocks. Thus a fundamental consequence of Theorems 3.4 and 3.6 is that in every Fiedler pencil of a square $n \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, the identity blocks are positioned in such a way that if the polynomial $P(\lambda)$ becomes rectangular with size $m \times n$, then these identity blocks may always be consistently transformed into I_m or I_n matrices so as to produce a Fiedler pencil for the $m \times n$ polynomial $P(\lambda)$. Let us examine a specific example to see how this works. Consider $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ with degree $k = 6$ and size $n \times n$, and the bijection $\tau = (1, 2, 5, 3, 6, 4)$. In this case
$$M_\tau = M_0 M_1 M_3 M_5 M_2 M_4 = \begin{bmatrix} -A_5 & -A_4 & I_n & 0 & 0 & 0 \\ I_n & 0 & 0 & 0 & 0 & 0 \\ 0 & -A_3 & 0 & -A_2 & I_n & 0 \\ 0 & I_n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -A_1 & 0 & I_n \\ 0 & 0 & 0 & -A_0 & 0 & 0 \end{bmatrix}, \tag{15}$$
which can be constructed by direct multiplication of the factors M_i or via `Algorithm 1`, since τ has consecutions at $0, 1, 3$ and inversions at $2, 4$. If the size of $P(\lambda)$ becomes $m \times n$, then `Algorithm 2` produces
$$M_\tau = \begin{bmatrix} -A_5 & -A_4 & I_m & 0 & 0 & 0 \\ I_n & 0 & 0 & 0 & 0 & 0 \\ 0 & -A_3 & 0 & -A_2 & I_m & 0 \\ 0 & I_n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -A_1 & 0 & I_m \\ 0 & 0 & 0 & -A_0 & 0 & 0 \end{bmatrix}, \tag{16}$$
which is nothing other than (15), but with the sizes of three of the identity blocks modified in order to be compatible with the size $m \times n$ of the coefficients $-A_i$.

Theorem 3.10 is a direct consequence of Theorem 3.6, and establishes that the zero-degree term M_σ of any Fiedler pencil of $P(\lambda)$ has as non-zero blocks exactly one copy of each of $-A_0, -A_1, \dots, -A_{k-1}$, as well as $(k-1)$ identities of size $n \times n$ or $m \times m$. This property is very well known in the case of the first and second companion forms, that are particular cases of Fiedler pencils. Theorem 3.10 also includes additional information on the structure of M_σ that will be used later.

Theorem 3.10. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$ and let $\sigma : \{0, 1, \dots, k-1\} \to \{1, \dots, k\}$ be a bijection. Suppose that $F_\sigma(\lambda) = \lambda \left[\begin{smallmatrix} A_k & \\ & I \end{smallmatrix}\right] - M_\sigma$ is the Fiedler pencil of $P(\lambda)$ associated with σ, and consider M_σ partitioned into $k \times k$ blocks according to* `Algorithm 2`. *Then:*

(a) *M_σ has k blocks equal to $-A_0, -A_1, \dots, -A_{k-1}$, with exactly one copy of each.*

(b) *M_σ has $k-1$ identity blocks: $\mathfrak{c}(\sigma)$ blocks equal to I_m, and $\mathfrak{i}(\sigma)$ blocks equal to I_n.*

(c) *The rest of the blocks of M_σ are equal to 0 matrices of size $n \times n$, $m \times m$, $n \times m$, or $m \times n$.*

(d) *The $k-1$ identity blocks in part (b) satisfy the following:*
   1. *None of them is on the main block diagonal of M_σ.*
   2. *Two of these blocks are never in the same block row (or in the same block column) of M_σ.*
   3. *If an identity block is in the (i, j) block-entry of M_σ, then one and only one of the following two properties holds: (a) the rest of the blocks in the ith block row of M_σ are equal to 0 and at least one of the matrices $-A_0, -A_1, \dots, -A_{k-1}$ is in the jth block column of M_σ; (b) the rest of the blocks in the jth block column of M_σ are equal to 0 and at least one of the matrices $-A_0, -A_1, \dots, -A_{k-1}$ is in the ith block row M_σ.*

4. *If* $\{i_1, i_2, \ldots, i_t\}$ *(resp.,* $\{j_{t+1}, j_{t+2}, \ldots, j_{k-1}\}$*) are the block indices of the block rows (resp., block columns) of* M_σ *containing one identity block and having the remaining blocks equal to zero, then the (unordered) set* $\{i_1, i_2, \ldots, i_t, j_{t+1}, j_{t+2}, \ldots, j_{k-1}\}$ *is equal to* $\{2, 3, \ldots, k\}$.

Proof. Parts (a), (b), and (c) are obvious from `Algorithm 2`. Part (d)-1 follows from Theorem 3.6(b). The proofs of parts (d)-2, (d)-3, and (d)-4 proceed by induction on the matrices $W_0, \ldots, W_{k-2} (= M_\sigma)$ constructed by `Algorithm 2`. A direct inspection shows that parts (d)-2, (d)-3, and (d)-4 hold for W_0 with $k = 2$. Let us assume that they hold for W_{k-3} with $k-1$ instead of k. Next partition W_{k-3} as follows:

$$W_{k-3} = \begin{bmatrix} -A_{k-2} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}.$$

Then `Algorithm 2` gives for $W_{k-2} = M_\sigma$ either

$$W_{k-2} = \begin{bmatrix} -A_{k-1} & I_m & 0 \\ -A_{k-2} & 0 & Z_{12} \\ Z_{21} & 0 & Z_{22} \end{bmatrix} \quad \text{or} \quad W_{k-2} = \begin{bmatrix} -A_{k-1} & -A_{k-2} & Z_{12} \\ I_n & 0 & 0 \\ 0 & Z_{21} & Z_{22} \end{bmatrix}, \tag{17}$$

and observe that the main block diagonal of Z_{22} is on the main block diagonal of W_{k-2}. The structure of M_σ in (17) and the fact that W_{k-3} satisfies parts (d)-2 and (d)-3 make evident that M_σ also satisfies parts (d)-2 and (d)-3. The block indices of the identity blocks of W_{k-3} in part (d)-4 are $\{2, 3, \ldots, k-1\}$, and observe that the induction hypothesis implies that if an identity block is a block-entry of Z_{12} (resp., Z_{21}) then the corresponding block column (resp., block row) in Z_{22} is zero. This fact and the structure of M_σ in (17) imply that the block indices in part (d)-(4) of the identity blocks of W_{k-3} as block entries of M_σ are $\{3, 4, \ldots, k\}$. Finally, note that the identity block that is added to construct W_{k-2} from W_{k-3} always has index 2 in the set of indices in part (d)-4. □

3.3. The reversal of a Fiedler pencil

The main result in this section is Theorem 3.14, which establishes that for a rectangular matrix polynomial $P(\lambda)$, the reversal of any of its Fiedler pencils is *strictly equivalent* to a Fiedler pencil of $\operatorname{rev} P(\lambda)$. We think that this result is interesting in its own right; in this paper, though, its main role will be to help prove in Section 4 that every Fiedler pencil of a rectangular matrix polynomial $P(\lambda)$ is a strong linearization of $P(\lambda)$. The proof of Theorem 3.14 is long and can be skipped on a first reading. The proof is based on the technical Lemmas 3.11, 3.12, and 3.13 that are presented next.

Lemma 3.11. *Let* $P(\lambda)$ *be an* $m \times n$ *matrix polynomial with degree* $k \geq 2$ *and let* $F_\sigma(P)$ *be the Fiedler pencil of* $P(\lambda)$ *associated with the bijection* σ*. Then the Fiedler pencil* $F_\sigma(-P)$ *of* $-P(\lambda)$ *is strictly equivalent to* $F_\sigma(P)$.

Proof. Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$. Throughout this proof, we view $F_\sigma(P) = \lambda \operatorname{diag}(A_k, I) - M_\sigma(P)$ and $F_\sigma(-P) = \lambda \operatorname{diag}(-A_k, I) - M_\sigma(-P)$ as $k \times k$ block-matrices, with the sizes of the blocks determined by the way `Algorithm 2` constructs $M_\sigma(P)$ and $M_\sigma(-P)$. In particular, we consider the block I in $\operatorname{diag}(A_k, I)$ and $\operatorname{diag}(-A_k, I)$ as $I = \operatorname{diag}(I_{r_2}, I_{r_3}, \ldots, I_{r_k})$, where $r_i = m$ or n by Theorem 3.6(b). Note, in the first place, that $F_\sigma(-P)$ is strictly equivalent to $-F_\sigma(-P)$. On the other hand, according to `Algorithm 2` and Theorem 3.10, the only difference between the pencils $-F_\sigma(-P) = \lambda \operatorname{diag}(A_k, -I) - (-M_\sigma(-P))$ and $F_\sigma(P) = \lambda \operatorname{diag}(A_k, I) - M_\sigma(P)$ are the signs of the $k-1$ identity blocks of $M_\sigma(P)$ and the signs of the $k-1$ diagonal identity blocks of $\operatorname{diag}(A_k, I)$. Let $\{i_1, i_2, \ldots, i_t\}$ and $\{j_{t+1}, j_{t+2}, \ldots, j_{k-1}\}$ be the indices defined in Theorem 3.10(d-4), and define now the matrices

$$U := \operatorname{diag}(I_m, \eta_2 I_{r_2}, \eta_3 I_{r_3}, \ldots, \eta_k I_{r_k}), \quad \text{where } \eta_i = \begin{cases} -1 & \text{if } \eta_i \in \{i_1, i_2, \ldots, i_t\} \\ 1 & \text{otherwise} \end{cases},$$

and

$$V := \operatorname{diag}(I_n, \alpha_2 I_{r_2}, \alpha_3 I_{r_3}, \ldots, \alpha_k I_{r_k}), \quad \text{where } \alpha_i = \begin{cases} -1 & \text{if } \alpha_i \in \{j_{t+1}, j_{t+2}, \ldots, j_{k-1}\} \\ 1 & \text{otherwise} \end{cases}.$$

According to Theorem 3.10(d-4) and the previous discussion

$$UF_\sigma(P)V = \lambda\, U \operatorname{diag}(A_k, I)V - UM_\sigma(P)V = \lambda \operatorname{diag}(A_k, -I) - (-M_\sigma(-P)) = -F_\sigma(-P),$$

which concludes the proof. □

Fiedler pencils for $\operatorname{rev} P(\lambda)$ can be easily constructed by applying `Algorithm 2` to the reversal polynomial. Lemma 3.12 shows us another way to construct Fiedler pencils for $\operatorname{rev} P(\lambda)$ that is useful in proving Theorem 3.14. According to Definition 3.8, we only need to pay attention in Lemma 3.12 to the construction of the zero-degree term of the pencil. In addition, for technical reasons that will be clear later, we construct pencils for the polynomial $-\operatorname{rev} P(\lambda)$.

Lemma 3.12 (Construction of Fiedler pencils for $-\operatorname{rev} P(\lambda)$ via block reverse identities)**.** *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection, and consider the following algorithm:*

> `Algorithm 3.` *Given $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ with size $m \times n$ and a bijection σ, the following algorithm constructs a sequence of matrices $\{Y_0, Y_1, \ldots, Y_{k-2}\}$, where each matrix Y_i for $i = 1, 2, \ldots, k-2$ is partitioned into blocks in such a way that the blocks of Y_{i-1} are blocks of Y_i.*
>
> if σ has a consecution at 0 then
> $\quad Y_0 = \begin{bmatrix} 0 & A_k \\ I_m & A_{k-1} \end{bmatrix}$
> else
> $\quad Y_0 = \begin{bmatrix} 0 & I_n \\ A_k & A_{k-1} \end{bmatrix}$
> endif
> for $i = 1 : k-2$
> $\quad$ if σ has a consecution at i then
> $\quad\quad Y_i = \begin{bmatrix} Y_{i-1}(:, 1:i) & 0 & Y_{i-1}(:, i+1) \\ 0 & I_m & A_{k-i-1} \end{bmatrix}$
> $\quad$ else
> $\quad\quad Y_i = \begin{bmatrix} Y_{i-1}(1:i, :) & 0 \\ 0 & I_n \\ Y_{i-1}(i+1, :) & A_{k-i-1} \end{bmatrix}$
> $\quad$ endif
> endfor

Then the matrices $Y_0, Y_1, \ldots, Y_{k-2}$ are partitioned in $2\times 2, 3\times 3, \ldots, k \times k$ blocks, respectively, and satisfy the following properties:

(a) *The size of Y_i is*

$$\big(m + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big) \times \big(n + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big).$$

(b) *The last diagonal block of Y_i is A_{k-i-1}, and so has size $m \times n$. The rest of the diagonal blocks of Y_i are square zero matrices, more precisely*

$$Y_i(j,j) = \begin{cases} 0_m & \text{if } \sigma \text{ has a consecution at } j-1 \\ 0_n & \text{if } \sigma \text{ has an inversion at } j-1 \end{cases}, \quad \text{for } j = 1, 2, \ldots, i+1.$$

(c) *Let $d_j \times d_j$ be the size of $Y_i(j,j)$ for $j = 1, 2, \ldots, i+1$, and define the $(i+2) \times (i+2)$ block reverse identities*

$$R_\ell^{(i)} := \begin{bmatrix} & & & I_m \\ & & I_{d_{i+1}} & \\ & \iddots & & \\ I_{d_1} & & & \end{bmatrix} \quad \text{and} \quad R_r^{(i)} := \begin{bmatrix} & & & I_{d_1} \\ & & \iddots & \\ & I_{d_{i+1}} & & \\ I_n & & & \end{bmatrix}.$$

Then we have

$$R_\ell^{(i)}\, Y_i\, R_r^{(i)} = W_i(-\operatorname{rev} P), \quad \textit{for } i = 0, 1, \ldots, k-2, \tag{18}$$

where $W_i(-\operatorname{rev} P)$ are the matrices constructed by `Algorithm 2` *for the polynomial $-\operatorname{rev} P(\lambda)$ and the bijection σ. In particular, according to Definition* 3.8,

$$R_\ell^{(k-2)}\, Y_{k-2}\, R_r^{(k-2)} = M_\sigma(-\operatorname{rev} P).$$

Proof. The proof of the lemma up through part (b) is analogous to the inductive proof of Theorem 3.6, and so is omitted. We only indicate that part (b) for the block $Y_i(i+1, i+1)$ is a direct consequence of the way `Algorithm 3` constructs Y_i, while the expressions for the remaining $Y_i(j, j)$ blocks follow from $Y_i(1:i, 1:i) = Y_{i-1}(1:i, 1:i)$ via induction. It is important to note that the size of $Y_i(j, j)$ only depends on j and not on i, whenever $1 \le j \le i+1$.

Before proving part (c), it is convenient to pay close attention to the structure of the matrices $R_\ell^{(i)}$ and $R_r^{(i)}$. First, note that the upper-right (resp., lower-left) block of $R_\ell^{(i)}$ (resp., $R_r^{(i)}$) is special because it is always equal to I_m (resp., I_n), independently of the consecutions/inversions that σ may have. The reason for the presence of these special blocks is to make the product in (18) conformable, since the last diagonal block of Y_i has size $m \times n$. This motivates the definition of two matrices, $\widehat{R}_\ell^{(i)}$ and $\widehat{R}_r^{(i)}$, obtained from $R_\ell^{(i)}$ and $R_r^{(i)}$ by removing these special blocks and the corresponding rows/columns, that is,

$$R_\ell^{(i)} =: \begin{bmatrix} & I_m \\ \widehat{R}_\ell^{(i)} & \end{bmatrix} \quad \text{and} \quad R_r^{(i)} =: \begin{bmatrix} & \widehat{R}_r^{(i)} \\ I_n & \end{bmatrix}. \tag{19}$$

Observe that the matrices $\widehat{R}_\ell^{(i)}$ and $\widehat{R}_r^{(i)}$ enjoy the following embedding properties,

$$\widehat{R}_\ell^{(i)} = \begin{bmatrix} & I_{d_{i+1}} \\ \widehat{R}_\ell^{(i-1)} & \end{bmatrix} \quad \text{and} \quad \widehat{R}_r^{(i)} = \begin{bmatrix} & \widehat{R}_r^{(i-1)} \\ I_{d_{i+1}} & \end{bmatrix}, \tag{20}$$

that do not hold for the un-hatted matrices $R_\ell^{(i)}$ and $R_r^{(i)}$.

We are now in a position to prove (18) by induction on i. The definitions of $R_\ell^{(i)}$ and $R_r^{(i)}$ guarantee that the three factors in the left-hand side of (18) are conformal for multiplication. The initial step $i = 0$ is proved directly, because for $i = 0$ we have:

- If σ has a consecution at 0, then

$$R_\ell^{(0)} Y_0 R_r^{(0)} = \begin{bmatrix} & I_m \\ I_m & \end{bmatrix} \begin{bmatrix} 0 & A_k \\ I_m & A_{k-1} \end{bmatrix} \begin{bmatrix} & I_m \\ I_n & \end{bmatrix} = \begin{bmatrix} A_{k-1} & I_m \\ A_k & 0 \end{bmatrix} = W_0(-\operatorname{rev} P).$$

- If σ has an inversion at 0, then

$$R_\ell^{(0)} Y_0 R_r^{(0)} = \begin{bmatrix} & I_m \\ I_n & \end{bmatrix} \begin{bmatrix} 0 & I_n \\ A_k & A_{k-1} \end{bmatrix} \begin{bmatrix} & I_n \\ I_n & \end{bmatrix} = \begin{bmatrix} A_{k-1} & A_k \\ I_n & 0 \end{bmatrix} = W_0(-\operatorname{rev} P).$$

Assume now that (18) is true for some $i-1$, such that $0 \le (i-1) \le k-3$, and we will prove it for i. We need to distinguish two cases according to whether σ has a consecution or an inversion at i.
Case 1: σ has a consecution at i. In this case $d_{i+1} = m$. Then (19) and (20) imply

$$\begin{aligned} R_\ell^{(i)} Y_i R_r^{(i)} &= \begin{bmatrix} & I_m \\ \widehat{R}_\ell^{(i)} & \end{bmatrix} \begin{bmatrix} Y_{i-1}(:, 1:i) & 0 & Y_{i-1}(:, i+1) \\ 0 & I_m & A_{k-i-1} \end{bmatrix} \begin{bmatrix} & & \widehat{R}_r^{(i-1)} \\ & I_m & \\ I_n & & \end{bmatrix} \\ &= \begin{bmatrix} A_{k-i-1} & I_m & 0 \\ \widehat{R}_\ell^{(i)} Y_{i-1}(:, i+1) & 0 & \widehat{R}_\ell^{(i)} Y_{i-1}(:, 1:i) \widehat{R}_r^{(i-1)} \end{bmatrix}. \end{aligned} \tag{21}$$

Observe that $d_{i+1} = m$, together with (19) and (20), imply that $R_\ell^{(i-1)} = \widehat{R}_\ell^{(i)}$. Now we use the induction assumption, that is, that (18) is true for $(i-1)$.

$$W_{i-1}(-\operatorname{rev} P) = R_\ell^{(i-1)} Y_{i-1} R_r^{(i-1)} = \widehat{R}_\ell^{(i)} \begin{bmatrix} Y_{i-1}(:,1:i) & Y_{i-1}(:,i+1) \end{bmatrix} \begin{bmatrix} & \widehat{R}_r^{(i-1)} \\ I_n & \end{bmatrix}$$
$$= \begin{bmatrix} \widehat{R}_\ell^{(i)} Y_{i-1}(:,i+1) & \widehat{R}_\ell^{(i)} Y_{i-1}(:,1:i) \widehat{R}_r^{(i-1)} \end{bmatrix}. \quad (22)$$

We substitute equation (22) in (21) to get

$$R_\ell^{(i)} Y_i R_r^{(i)} = \begin{bmatrix} A_{k-i-1} & I_m & 0 \\ [W_{i-1}(-\operatorname{rev} P)](:,1) & 0 & [W_{i-1}(-\operatorname{rev} P)](:,2:i+1) \end{bmatrix} = W_i(-\operatorname{rev} P),$$

where the last step follows from applying `Algorithm 2` to $-\operatorname{rev} P(\lambda)$ and σ. This concludes the proof of *Case 1*.

Case 2: σ has an inversion at i. In this case $d_{i+1} = n$. Then (19) and (20) imply

$$R_\ell^{(i)} Y_i R_r^{(i)} = \begin{bmatrix} & & I_m \\ & I_n & \\ \widehat{R}_\ell^{(i-1)} & & \end{bmatrix} \begin{bmatrix} Y_{i-1}(1:i,:) & 0 \\ 0 & I_n \\ Y_{i-1}(i+1,:) & A_{k-i-1} \end{bmatrix} \begin{bmatrix} & \widehat{R}_r^{(i)} \\ I_n & \end{bmatrix}$$
$$= \begin{bmatrix} A_{k-i-1} & Y_{i-1}(i+1,:)\widehat{R}_r^{(i)} \\ I_n & 0 \\ 0 & \widehat{R}_\ell^{(i-1)} Y_{i-1}(1:i,:) \widehat{R}_r^{(i)} \end{bmatrix}. \quad (23)$$

Observe that $d_{i+1} = n$, together with (19) and (20), imply that $R_r^{(i-1)} = \widehat{R}_r^{(i)}$. Now we use the induction assumption.

$$W_{i-1}(-\operatorname{rev} P) = R_\ell^{(i-1)} Y_{i-1} R_r^{(i-1)} = \begin{bmatrix} & I_m \\ \widehat{R}_\ell^{(i-1)} & \end{bmatrix} \begin{bmatrix} Y_{i-1}(1:i,:) \\ Y_{i-1}(i+1,:) \end{bmatrix} \widehat{R}_r^{(i)}$$
$$= \begin{bmatrix} Y_{i-1}(i+1,:)\widehat{R}_r^{(i)} \\ \widehat{R}_\ell^{(i-1)} Y_{i-1}(1:i,:) \widehat{R}_r^{(i)} \end{bmatrix}. \quad (24)$$

We substitute equation (24) in (23) to get

$$R_\ell^{(i)} Y_i R_r^{(i)} = \begin{bmatrix} A_{k-i-1} & [W_{i-1}(-\operatorname{rev} P)](1,:) \\ I_n & 0 \\ 0 & [W_{i-1}(-\operatorname{rev} P)](2:i+1,:) \end{bmatrix} = W_i(-\operatorname{rev} P).$$

This concludes the proof of *Case 2*. □

Lemma 3.13 shows the result of certain matrix multiplications that are used in the proof of Theorem 3.14 to perform strict equivalences on the reversals of Fiedler pencils when the degree k of the polynomial satisfies $k \geq 3$.

Lemma 3.13. *Let $\sigma, \tau : \{0, 1, \ldots, k-1\} \to \{1, \ldots, k\}$ be two bijections such that σ has a consecution at $i-1$ if and only if τ has a consecution at $k-i-1$ for $i = 1, \ldots, k-1$. Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 3$, let $\{W_i\}_{i=0}^{k-2}$ be the sequence of block partitioned matrices constructed by* `Algorithm 2` *for $P(\lambda)$ and σ, and let $\{Y_i\}_{i=0}^{k-2}$ be the sequence of block partitioned matrices constructed by* `Algorithm 3` *for $P(\lambda)$ and τ. Also define $W_{-1} := -A_0$ and $Y_{-1} := A_k$. Let us define two sequences, $\{\widetilde{I}_i\}_{i=0}^{k-1}$ and $\{\widehat{I}_i\}_{i=0}^{k-1}$, of partitioned matrices as follows: $\widetilde{I}_0$ and $\widehat{I}_0$ are 0×0 empty matrices, and*

$$\widetilde{I}_i := \begin{bmatrix} I_{s_1} & & & \\ & I_{s_2} & & \\ & & \ddots & \\ & & & I_{s_i} \end{bmatrix} \quad \text{and} \quad \widehat{I}_i := \begin{bmatrix} I_{t_{k-i+1}} & & & \\ & I_{t_{k-i+2}} & & \\ & & \ddots & \\ & & & I_{t_k} \end{bmatrix}, \quad \text{for } i = 1, \ldots, k-1,$$

where $\{s_j\}_{j=1}^{i}$ *are the sizes of the square diagonal blocks* $\{Y_{i-1}(j,j)\}_{j=1}^{i}$*, and* $\{t_j\}_{j=k-i+1}^{k}$ *are the sizes of the square diagonal blocks* $\{W_{i-1}(j,j)\}_{j=2}^{i+1}$*. Then the following statements hold.*

(a) *For each* $i = 0, 1, \ldots, k-1$*, the matrices*

$$\widetilde{W}_{i-1} := \begin{bmatrix} \widetilde{I}_{k-i-1} & \\ & W_{i-1} \end{bmatrix} \quad \text{and} \quad \widetilde{Y}_{k-i-2} := \begin{bmatrix} Y_{k-i-2} & \\ & \widehat{I}_i \end{bmatrix}$$

are partitioned into $k \times k$ *blocks and the size of the block* $\widetilde{W}_{i-1}(p,q)$ *is equal to the size of the block* $\widetilde{Y}_{k-i-2}(p,q)$ *for all* $1 \le p, q \le k$*. In addition,* $\widetilde{W}_{i-1}$ *and* $\widetilde{Y}_{k-i-2}$ *both have size*

$$\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big),$$

that is, the same size as the Fiedler pencil of $P(\lambda)$ *associated with* σ*.*

(b) *Define a sequence of matrices* $\{S_i\}_{i=1}^{k-1}$ *as follows:*

$$S_1 := \begin{bmatrix} \widetilde{I}_{k-2} & & \\ & 0 & I_n \\ & I_m & A_1 \end{bmatrix}, \quad S_i := \begin{bmatrix} \widetilde{I}_{k-i-1} & & & & & \\ & 0 & I_n & & & \\ & I_m & A_i & & & \\ & & & I_{t_{k-i+2}} & & \\ & & & & \ddots & \\ & & & & & I_{t_k} \end{bmatrix}, \ i = 2, \ldots, k-1.$$

Then for each $i = 1, \ldots, k-1$*, the following statements hold:*

(b1) *If* σ *has a consecution at* $i-1$*, then* S_i *has size* $\big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big)$*, and*

$$\widetilde{W}_{i-1} S_i = \widetilde{W}_{i-2} \quad \text{and} \quad \widetilde{Y}_{k-i-2} S_i = \widetilde{Y}_{k-i-1}\,.$$

(b2) *If* σ *has an inversion at* $i-1$*, then* S_i *has size* $\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big)$*, and*

$$S_i \widetilde{W}_{i-1} = \widetilde{W}_{i-2} \quad \text{and} \quad S_i \widetilde{Y}_{k-i-2} = \widetilde{Y}_{k-i-1}\,.$$

Proof. Part (a). $\widetilde{I}_{k-i-1}$ has $(k-i-1) \times (k-i-1)$ blocks, by definition, and by Theorem 3.6, W_{i-1} has $(i+1) \times (i+1)$ blocks. So $\widetilde{W}_{i-1}$ has $k \times k$ blocks. Analogously, by Lemma 3.12, Y_{k-i-2} has $(k-i) \times (k-i)$ blocks and, by definition, $\widehat{I}_i$ has $i \times i$ blocks. So $\widetilde{Y}_{k-i-2}$ has $k \times k$ blocks.

Next we prove that the sizes of the blocks of $\widetilde{W}_{i-1}$ are equal to the sizes of the corresponding blocks of $\widetilde{Y}_{k-i-2}$. Assume first that $1 \le i \le k-2$, and recall that $W_{i-1}(1,1) = -A_i$ and $Y_{k-i-2}(k-i, k-i) = A_{i+1}$. Then

$$\widetilde{W}_{i-1} = \begin{bmatrix} \widetilde{I}_{k-i-1} & & \\ & -A_i & * \\ & * & W_{i-1}(2:i+1, 2:i+1) \end{bmatrix}, \tag{25}$$

$$\widetilde{Y}_{k-i-2} = \begin{bmatrix} Y_{k-i-2}(1:k-i-1, 1:k-i-1) & * & \\ * & A_{i+1} & \\ & & \widehat{I}_i \end{bmatrix}. \tag{26}$$

By definition, $\widetilde{I}_{k-i-1}$ is partitioned into blocks exactly as $Y_{k-i-2}(1:k-i-1, 1:k-i-1)$, and $\widehat{I}_i$ is partitioned exactly as $W_{i-1}(2:i+1, 2:i+1)$. Therefore $\widetilde{W}_{i-1}$ and $\widetilde{Y}_{k-i-2}$ have corresponding blocks with equal sizes. For $i = 0$ we have $\widetilde{W}_{-1} = \operatorname{diag}(\widetilde{I}_{k-1}, -A_0)$, $\widetilde{Y}_{k-2} = Y_{k-2}$, and the definition of $\widetilde{I}_{k-1}$ together with Lemma 3.12(b) guarantee that the sizes of corresponding blocks are equal. For $i = k-1$ we have $\widetilde{W}_{k-2} = W_{k-2}$, $\widetilde{Y}_{-1} = \operatorname{diag}(A_k, \widehat{I}_{k-1})$, and the definition of $\widehat{I}_{k-1}$ together with Theorem 3.6(b) imply the result.

We consider now the total size of the matrices $\widetilde{W}_{i-1}$ and $\widetilde{Y}_{k-i-2}$. Note first that $\mathfrak{c}(\sigma) = \mathfrak{c}(\tau)$ and $\mathfrak{i}(\sigma) = \mathfrak{i}(\tau)$. For $i = 0$, we get from the previous discussion that $\widetilde{W}_{-1}$ and $\widetilde{Y}_{k-2} = Y_{k-2}$ both have the size of Y_{k-2}, that is $\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big)$ according to Lemma 3.12(a). For $i = k-1$, we get from the previous discussion that $\widetilde{W}_{k-2} = W_{k-2}$ and $\widetilde{Y}_{-1}$ both have the size of W_{k-2}, that is $\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times \big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big)$ according to Theorem 3.6(a). For $1 \le i \le k-2$, we get again from the previous discussion that $\widetilde{W}_{i-1}$ and $\widetilde{Y}_{k-i-2}$ both have the same size. This size is the sum of the sizes of the three diagonal blocks in (25) (or (26)), which according to Theorem 3.6 and Lemma 3.12 is $(m+r) \times (n+r)$ with

$$r = m\big[\,\mathfrak{c}\big(\tau(0:k-i-2)\big) + \mathfrak{c}\big(\sigma(0:i-1)\big)\,\big] + n\big[\,\mathfrak{i}\big(\tau(0:k-i-2)\big) + \mathfrak{i}\big(\sigma(0:i-1)\big)\,\big].$$

From the definition of τ, we see that r is equal to

$$\begin{aligned} r &= m\big[\,\mathfrak{c}\big(\sigma(i:k-2)\big) + \mathfrak{c}\big(\sigma(0:i-1)\big)\,\big] + n\big[\,\mathfrak{i}\big(\sigma(i:k-2)\big) + \mathfrak{i}\big(\sigma(0:i-1)\big)\,\big] \\ &= m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma). \end{aligned}$$

This concludes the proof of Part (a).
Part (b). For brevity, we prove only (b1). The proof of (b2) is similar and is omitted. Let us establish the size of S_i, which is clearly a square matrix for each i. So we only pay attention to the number of rows. Consider first the number of rows of S_1. Note that if σ has a consecution at 0, then

$$\widetilde{W}_0 = \begin{bmatrix} \widetilde{I}_{k-2} & \\ & W_0 \end{bmatrix} = \begin{bmatrix} \widetilde{I}_{k-2} & & \\ & -A_1 & I_m \\ & -A_0 & 0 \end{bmatrix}.$$

This makes evident that the number of columns of $\widetilde{W}_0$ is equal to the number of rows of S_1, and therefore this number is $(\,n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma))$ by part (a). Note that we have also shown that the partitions of $\widetilde{W}_0$ and S_1 are conformal for the product $\widetilde{W}_0 S_1$; by part (a), the same happens for the product $\widetilde{Y}_{k-3}S_1$. Consider next the number of columns of S_i, for $i = 2, \ldots, k-1$. Note that if σ has a consecution at $i-1$, then

$$\widetilde{W}_{i-1} = \begin{bmatrix} \widetilde{I}_{k-i-1} & \\ & W_{i-1} \end{bmatrix} = \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & -A_i & I_m & 0 \\ & W_{i-2}(:,1) & 0 & W_{i-2}(:,2:i) \end{bmatrix}.$$

By definition, $t_{k-i+2}, \ldots, t_k$ are the number of columns of the block columns of $W_{i-1}(:,3:i+1)$, which have the same number of columns as the block columns of $W_{i-2}(:,2:i)$. Therefore the number of columns of $\widetilde{W}_{i-1}$ is equal to the number of rows of S_i, and this number is $(\,n{+}m\,\mathfrak{c}(\sigma){+}n\,\mathfrak{i}(\sigma))$ by part (a). Observe that we have also proved that the partitions of $\widetilde{W}_{i-1}$ and S_i are conformal for the product $\widetilde{W}_{i-1}S_i$. This implies, by part (a), that the partitions of $\widetilde{Y}_{k-i-2}$ and S_i are conformal for the product $\widetilde{Y}_{k-i-2}S_i$.

To prove that $\widetilde{W}_{i-1}S_i = \widetilde{W}_{i-2}$ for $i = 1, \ldots, k-1$, we first need to deal separately with the case $i = 1$, because $\widetilde{W}_{-1}$ has a structure different from $\widetilde{W}_i$ for $i > -1$. A direct block multiplication shows that

$$\widetilde{W}_0 S_1 = \begin{bmatrix} \widetilde{I}_{k-2} & & \\ & -A_1 & I_m \\ & -A_0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{I}_{k-2} & & \\ & 0 & I_n \\ & I_m & A_1 \end{bmatrix} = \begin{bmatrix} \widetilde{I}_{k-2} & & \\ & I_m & 0 \\ & 0 & -A_0 \end{bmatrix} = \widetilde{W}_{-1},$$

where we have used that $\widetilde{I}_{k-1} = \operatorname{diag}(\widetilde{I}_{k-2}, I_m)$. This latter fact holds because, according to Lemma 3.12(b), the sizes of the blocks $\{Y_{k-3}(j,j)\}_{j=1}^{k-2}$ are equal to the sizes of the blocks $\{Y_{k-2}(j,j)\}_{j=1}^{k-2}$, and $Y_{k-2}(k-1,k-1) = 0_m$ because σ has a consecution at 0, that is, τ has a consecution at $k-2$. Let us now consider $i = 2, \ldots, k-1$. Then

$$\begin{aligned} \widetilde{W}_{i-1}S_i &= \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & -A_i & I_m & 0 \\ & W_{i-2}(:,1) & 0 & W_{i-2}(:,2:i) \end{bmatrix} \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & 0 & I_n & \\ & I_m & A_i & \\ & & & I_{t_{k-i+2}+\cdots+t_k} \end{bmatrix} \\ &= \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & I_m & 0 & 0 \\ & 0 & W_{i-2}(:,1) & W_{i-2}(:,2:i) \end{bmatrix} = \begin{bmatrix} \widetilde{I}_{k-i} & \\ & W_{i-2} \end{bmatrix} = \widetilde{W}_{i-2}, \end{aligned}$$

where we have used that $\widetilde{I}_{k-i} = \operatorname{diag}(\widetilde{I}_{k-i-1}, I_m)$. This holds because, according to Lemma 3.12(b), the sizes of the blocks $\{Y_{k-i-2}(j,j)\}_{j=1}^{k-i-1}$ are equal to the sizes of the blocks $\{Y_{k-i-1}(j,j)\}_{j=1}^{k-i-1}$, and $Y_{k-i-1}(k-i,k-i) = 0_m$ because σ has a consecution at $i-1$, that is, τ has a consecution at $k-i-1$.

Next we proceed to show that $\widetilde{Y}_{k-i-2}S_i = \widetilde{Y}_{k-i-1}$. Here we need to deal separately at the end with the case $i = k-1$, because $\widetilde{Y}_{-1}$ has a structure different from the remaining $\widetilde{Y}_i$. We consider first $i = 1, \ldots, k-2$. Since σ has a consecution at $i-1$, we have that $W_{i-1}(2,2) = 0_m$ by Theorem 3.6(b), and the first block of $\widehat{I}_i$ is $I_{t_{k-i+1}} = I_m$. In addition, note that $\widehat{I}_i = \operatorname{diag}(I_m, \widehat{I}_{i-1})$ because the sizes of the blocks $\{W_{i-1}(j,j)\}_{j=3}^{i+1}$ are equal to the sizes of the blocks $\{W_{i-2}(j,j)\}_{j=2}^{i}$ by Theorem 3.6(b) (recall also Remark 3.7). Therefore

$$\begin{aligned}
\widetilde{Y}_{k-i-2}S_i &= \begin{bmatrix} Y_{k-i-2} & & \\ & I_m & \\ & & I_{t_{k-i+2}+\cdots+t_k} \end{bmatrix} \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & 0 & I_n & \\ & I_m & A_i & \\ & & & I_{t_{k-i+2}+\cdots+t_k} \end{bmatrix} \\
&= \begin{bmatrix} Y_{k-i-2}(:,1:k-i-1) & Y_{k-i-2}(:,k-i) & & \\ & & I_m & \\ & & & \widehat{I}_{i-1} \end{bmatrix} \begin{bmatrix} \widetilde{I}_{k-i-1} & & & \\ & 0 & I_n & \\ & I_m & A_i & \\ & & & \widehat{I}_{i-1} \end{bmatrix} \\
&= \begin{bmatrix} Y_{k-i-2}(:,1:k-i-1) & & Y_{k-i-2}(:,k-i) & \\ & I_m & A_i & \\ & & & \widehat{I}_{i-1} \end{bmatrix} = \begin{bmatrix} Y_{k-i-1} & \\ & \widehat{I}_{i-1} \end{bmatrix} = \widetilde{Y}_{k-i-1},
\end{aligned}$$

where we have used `Algorithm 3`, taking into account that τ has a consecution at $k-i-1$. We finally cover the case $i = k-1$. Since σ has a consecution at $k-2$, an argument similar to the one above shows that $\widehat{I}_{k-1} = \operatorname{diag}(I_m, \widehat{I}_{k-2})$. Therefore,

$$\begin{aligned}
\widetilde{Y}_{-1}S_{k-1} &= \begin{bmatrix} A_k & \\ & \widehat{I}_{k-1} \end{bmatrix} \begin{bmatrix} 0 & I_n & \\ I_m & A_{k-1} & \\ & & I_{t_3+\cdots+t_k} \end{bmatrix} = \begin{bmatrix} A_k & & \\ & I_m & \\ & & \widehat{I}_{k-2} \end{bmatrix} \begin{bmatrix} 0 & I_n & \\ I_m & A_{k-1} & \\ & & \widehat{I}_{k-2} \end{bmatrix} \\
&= \begin{bmatrix} 0 & A_k & \\ I_m & A_{k-1} & \\ & & \widehat{I}_{k-2} \end{bmatrix} = \begin{bmatrix} Y_0 & \\ & \widehat{I}_{k-2} \end{bmatrix} = \widetilde{Y}_0,
\end{aligned}$$

where we have used `Algorithm 3`, taking into account that τ has a consecution at 0. This concludes the proof of (b1). $\square$

Now we are in a position to prove the main result of this section.

Theorem 3.14. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, and let $F_\sigma(P)$ be the Fiedler pencil of $P(\lambda)$ associated with a bijection σ. Then $\operatorname{rev} F_\sigma(P)$ is strictly equivalent to a Fiedler pencil of $\operatorname{rev} P(\lambda)$. More precisely, $\operatorname{rev} F_\sigma(P)$ is strictly equivalent to $F_\tau(\operatorname{rev} P)$, where $\tau : \{0,1,\ldots,k-1\} \to \{1,\ldots,k\}$ is any bijection such that τ has a consecution (resp., inversion) at $k-i-1$ if σ has a consecution (resp., inversion) at $i-1$, for $i = 1,\ldots,k-1$.*

Proof. Let $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$. If the degree is $k = 2$, then `Algorithm 2` shows that there are only two different Fiedler pencils. These are the two companion forms

$$C_1(P) = \lambda \begin{bmatrix} A_2 & 0 \\ 0 & I_n \end{bmatrix} + \begin{bmatrix} A_1 & A_0 \\ -I_n & 0 \end{bmatrix} \quad \text{and} \quad C_2(P) = \lambda \begin{bmatrix} A_2 & 0 \\ 0 & I_m \end{bmatrix} + \begin{bmatrix} A_1 & -I_m \\ A_0 & 0 \end{bmatrix}.$$

For $k = 2$, direct matrix multiplications show that

$$\begin{bmatrix} I_m & A_1 \\ 0 & -I_n \end{bmatrix} (\operatorname{rev} C_1(P)) \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} = C_1(\operatorname{rev} P) \quad \text{and}$$
$$\begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix} (\operatorname{rev} C_2(P)) \begin{bmatrix} I_n & 0 \\ A_1 & -I_m \end{bmatrix} = C_2(\operatorname{rev} P),$$

which proves the result because the matrices multiplying $\operatorname{rev} C_1(P)$ and $\operatorname{rev} C_2(P)$ are always nonsingular. Observe that for $k = 2$ the bijections σ and τ are identical.

For $k \geq 3$, the proof relies on Lemma 3.13, so for the rest of the proof we use exactly the same definitions and notation as in Lemma 3.13. Note that $\widetilde{W}_{k-2} = W_{k-2} = M_\sigma$, where $-M_\sigma$ is the zero degree term of $F_\sigma(P)$ according to Definition 3.8. Then

$$F_\sigma(P) = \lambda \begin{bmatrix} A_k & \\ & I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)} \end{bmatrix} - M_\sigma = \lambda \widetilde{Y}_{-1} - \widetilde{W}_{k-2} \quad \text{and} \quad \operatorname{rev} F_\sigma(P) = \widetilde{Y}_{-1} - \lambda \widetilde{W}_{k-2}.$$

Next we use the *nonsingular* matrices $S_{k-1}, S_{k-2}, \ldots, S_1$ introduced in Lemma 3.13(b), and multiply $\operatorname{rev} F_\sigma(P)$ first by S_{k-1}, second by S_{k-2}, and so on until we multiply by S_1. The multiplications are performed on the right or on the left according to the consecutions or inversions of σ as indicated in Lemma 3.13(b1)-(b2). So we obtain

$$\operatorname{rev} F_\sigma(P) = (\widetilde{Y}_{-1} - \lambda\widetilde{W}_{k-2}) \sim_s (\widetilde{Y}_0 - \lambda\widetilde{W}_{k-3}) \sim_s (\widetilde{Y}_1 - \lambda\widetilde{W}_{k-4}) \sim_s \cdots \sim_s (\widetilde{Y}_{k-2} - \lambda\widetilde{W}_{-1}),$$

where the symbol $\sim_s$ denotes that we are performing strict equivalences, because the matrices S_i are always nonsingular. From Lemma 3.13 we see that $\widetilde{Y}_{k-2} = Y_{k-2}$ and $\widetilde{W}_{-1} = \operatorname{diag}(\widetilde{I}_{k-1}, -A_0)$. Therefore

$$\operatorname{rev} F_\sigma(P) \sim_s \big(Y_{k-2} - \lambda \operatorname{diag}(\widetilde{I}_{k-1}, -A_0) \big).$$

We now apply Lemma 3.12(c) to get

$$\begin{aligned} \operatorname{rev} F_\sigma(P) &\sim_s \big(R_\ell^{(k-2)} Y_{k-2} R_r^{(k-2)} - \lambda R_\ell^{(k-2)} \operatorname{diag}(\widetilde{I}_{k-1}, -A_0) R_r^{(k-2)} \big) \\ &= M_\tau(-\operatorname{rev} P) - \lambda \operatorname{diag}(-A_0, \widetilde{I}_{k-1}) = -F_\tau(-\operatorname{rev} P). \end{aligned}$$

Finally, by Lemma 3.11, $-F_\tau(-\operatorname{rev} P)$ is strictly equivalent to $-F_\tau(\operatorname{rev} P)$, which in turn is strictly equivalent to $F_\tau(\operatorname{rev} P)$. Hence we conclude that $\operatorname{rev} F_\sigma(P)$ is strictly equivalent to $F_\tau(\operatorname{rev} P)$. □

4. Fiedler pencils of rectangular matrix polynomials are strong linearizations

We will prove in this section that all Fiedler pencils $F_\sigma(\lambda)$ of a rectangular matrix polynomial $P(\lambda)$ are strong linearizations for $P(\lambda)$. This is proved in Theorem 4.5, which generalizes [11, Thm 4.6] in a nontrivial way. The approach we follow is constructive, in the sense that we will show how to construct appropriate unimodular matrices $U(\lambda)$ and $V(\lambda)$ satisfying (3) for every $F_\sigma(\lambda)$. The construction of these matrices is accomplished via the construction of sequences of block partitioned matrix polynomials in `Algorithms 4` and `5`, which follow the spirit of Definition 3.8 of Fiedler pencils for rectangular polynomials. The properties of the unimodular transformations generated by these sequences are then further studied in Lemma 4.4. In this section, we make systematic use of the *Horner shifts* introduced in Definition 4.1.

Definition 4.1. *Let $P(\lambda) = A_0 + \lambda A_1 + \cdots + \lambda^k A_k$ be an $m \times n$ matrix polynomial of degree k. For $d = 0, \ldots, k$, the* degree d Horner shift of $P(\lambda)$ *is the matrix polynomial $P_d(\lambda) := A_{k-d} + \lambda A_{k-d+1} + \cdots + \lambda^d A_k$.*

Observe that the Horner shifts of $P(\lambda)$ satisfy the following recurrence relation

$$P_0(\lambda) = A_k, \;\; P_{d+1}(\lambda) = \lambda P_d(\lambda) + A_{k-d-1}, \text{ for } 0 \leq d \leq k-1, \;\text{ and }\; P_k(\lambda) = P(\lambda).$$

Lemma 4.2. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, 2, \ldots, k\}$ be a bijection, and consider the following algorithms:*

> `Algorithm 4.` *Given $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ with size $m \times n$ and a bijection σ, the following algorithm constructs a sequence of matrix polynomials $\{N_0, N_1, \ldots, N_{k-2}\}$, where each matrix N_i for $i = 1, 2, \ldots, k-2$ is partitioned into blocks in such a way that the blocks of N_{i-1} are blocks of N_i. Note that P_d denotes the degree d Horner shift of $P(\lambda)$, and that the dependence on λ is dropped for simplicity, both in P_d and in $\{N_i\}_{i=0}^{k-2}$.*

if σ has a consecution at 0 then

$$N_0 = \begin{bmatrix} I_m & 0 \\ \lambda I_m & I_m \end{bmatrix}$$

else

$$N_0 = \begin{bmatrix} 0 & -I_n \\ I_m & P_{k-1} \end{bmatrix}$$

endif

for $i = 1 : k-2$

if σ has a consecution at i then

$$N_i = \begin{bmatrix} I_m & 0 \\ \lambda N_{i-1}(:,1) & N_{i-1} \end{bmatrix}$$

else

$$N_i = \begin{bmatrix} 0 & -I_n & 0 \\ N_{i-1}(:,1) & N_{i-1}(:,1)P_{k-i-1} & N_{i-1}(:,2:i+1) \end{bmatrix}$$

endif

endfor

`Algorithm 5.` *Given* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ *with size* $m \times n$ *and a bijection* σ*, the following algorithm constructs a sequence of matrix polynomials* $\{H_0, H_1, \ldots, H_{k-2}\}$*, where each matrix* H_i *for* $i = 1, 2, \ldots, k-2$ *is partitioned into blocks in such a way that the blocks of* H_{i-1} *are blocks of* H_i*. Note that* P_d *denotes the degree* d *Horner shift of* $P(\lambda)$*, and that the dependence on* λ *is dropped for simplicity, both in* P_d *and in* $\{H_i\}_{i=0}^{k-2}$*.*

if σ has a consecution at 0 then

$$H_0 = \begin{bmatrix} 0 & I_n \\ -I_m & P_{k-1} \end{bmatrix}$$

else

$$H_0 = \begin{bmatrix} I_n & \lambda I_n \\ 0 & I_n \end{bmatrix}$$

endif

for $i = 1 : k-2$

if σ has a consecution at i then

$$H_i = \begin{bmatrix} 0 & H_{i-1}(1,:) \\ -I_m & P_{k-i-1} H_{i-1}(1,:) \\ 0 & H_{i-1}(2:i+1,:) \end{bmatrix}$$

else

$$H_i = \begin{bmatrix} I_n & \lambda H_{i-1}(1,:) \\ 0 & H_{i-1} \end{bmatrix}$$

endif

endfor

Then the matrix polynomials $N_0, N_1, \ldots, N_{k-2}$ *constructed by* `Algorithm 4` *are partitioned into* $2\times 2, 3\times 3, \ldots, k\times k$ *blocks, respectively, and the matrix polynomials* $H_0, H_1, \ldots, H_{k-2}$ *constructed by* `Algorithm 5` *are partitioned into* $2\times 2, 3\times 3, \ldots, k\times k$ *blocks, respectively. Moreover, if* $\{W_i\}_{i=0}^{k-2}$ *is the sequence of block partitioned matrices constructed by* `Algorithm 2` *for* $P(\lambda)$ *and* σ*, then the matrix polynomials* $\{N_i\}_{i=0}^{k-2}$ *and* $\{H_i\}_{i=0}^{k-2}$ *satisfy the following properties:*

(a) *For* $0 \le i \le k-2$ *and* $1 \le j \le i+2$*, the number of columns of* $N_i(:,j)$ *is equal to the number of rows of* $W_i(j,:)$*. This means that the matrix product* $N_i(:,j)\, W_i(j,:)$ *is well-defined.*

(b) *For* $0 \le i \le k-2$ *and* $1 \le j \le i+2$*, the number of columns of* $W_i(:,j)$ *is equal to the number of rows of* $H_i(j,:)$*. This means that the matrix product* $W_i(:,j)\, H_i(j,:)$ *is well-defined.*

(c) *The size of* N_i *is* $\big(m + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big) \times \big(m + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big)$*.*

(d) *The size of* H_i *is* $\big(n + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big) \times \big(n + m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))\big)$*.*

(e) *The matrix polynomials* N_i *and* H_i *are unimodular. In fact,* $\det(N_i) = \pm 1$ *and* $\det(H_i) = \pm 1$*.*

(f) *$N_i(i+2,:)$ has m rows and $H_i(:,i+2)$ has n columns; that is, the last block row of N_i has m rows and the last block column of H_i has n columns.*

Proof. The proof is trivial by induction. We indicate only the main points. Starting with N_0 and H_0, it is obvious to see by induction that $N_i(:,1)$ has m columns and $H_i(1,:)$ has n rows for $i = 0, 1, \ldots, k-2$. Therefore the sequences $\{N_i\}_{i=0}^{k-2}$ and $\{H_i\}_{i=0}^{k-2}$ are well-defined. Since, for $i \geq 1$, N_i and H_i are obtained from N_{i-1} and H_{i-1}, respectively, by adding one block row and one block column, then we see that N_i and H_i are partitioned into $(i+2) \times (i+2)$ blocks.

To prove part (a), note that the result is true for N_0 and W_0, and make the inductive assumption that the result holds for N_{i-1} and W_{i-1} with $(i-1) \geq 0$. Then the constructions of N_i in `Algorithm 4` and W_i in `Algorithm 2` make evident that the result is true for N_i and W_i. Part (b) follows from a similar inductive argument.

To prove parts (c) and (d), we note first that all matrices in the sequences $\{N_i\}_{i=0}^{k-2}$ and $\{H_i\}_{i=0}^{k-2}$ are square; by definition this is true for N_0 and W_0, and for $i \geq 1$, N_i is obtained from N_{i-1} by adding the same number of rows as columns, and H_i is also obtained from H_{i-1} by adding the same number of rows as columns. Then (c) follows from (a), and (d) from (b), by using Theorem 3.6(a). Finally parts (e) and (f) follow again by induction. □

Since the matrices N_{k-2} and H_{k-2} in Lemma 4.2 play a key role in the rest of the paper, we give them each a special name and notation.

Definition 4.3. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, 1, \ldots, k-1\} \to \{1, 2, \ldots, k\}$ be a bijection, and let N_{k-2} and H_{k-2} be, respectively, the last matrices of the sequences constructed by* `Algorithms 4` *and* `5` *for $P(\lambda)$ and σ. Then*

- *$U_\sigma(\lambda) := N_{k-2}$ is the* left unimodular equivalence matrix *associated with $P(\lambda)$ and σ,*
- *$V_\sigma(\lambda) := H_{k-2}$ is the* right unimodular equivalence matrix *associated with $P(\lambda)$ and σ.*

Lemma 4.4 further studies the unimodular transformations generated by the sequences $\{N_i\}_{i=0}^{k-2}$ and $\{H_i\}_{i=0}^{k-2}$.

Lemma 4.4. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, \ldots, k-1\} \to \{1, \ldots, k\}$ be a bijection, and let $\{W_i\}_{i=0}^{k-2}, \{N_i\}_{i=0}^{k-2}, \{H_i\}_{i=0}^{k-2}$ be the sequences of block partitioned matrices constructed, respectively, by* `Algorithms 2, 4` *and* `5` *for $P(\lambda)$ and σ. Also consider the numbers $\alpha_i = m\,\mathfrak{c}(\sigma(0:i)) + n\,\mathfrak{i}(\sigma(0:i))$ and $\beta_i = m\,\mathfrak{c}(\sigma(i)) + n\,\mathfrak{i}(\sigma(i))$; note that $\beta_i = m$ if σ has a consecution at i, and $\beta_i = n$ if σ has an inversion at i, for $i = 0, 1, \ldots, k-2$. Then the following two identities hold.*

(a) *For $1 \leq i \leq k-2$,*

$$N_i \left(\begin{bmatrix} \lambda P_{k-i-2} & \\ & \lambda I_{\alpha_i} \end{bmatrix} - W_i \right) H_i = \begin{bmatrix} I_{\beta_i} & \\ & N_{i-1} \left(\begin{bmatrix} \lambda P_{k-i-1} & \\ & \lambda I_{\alpha_{i-1}} \end{bmatrix} - W_{i-1} \right) H_{i-1} \end{bmatrix},$$

(b) *and, for $i = 0$,*

$$N_0 \left(\begin{bmatrix} \lambda P_{k-2} & \\ & \lambda I_{\alpha_0} \end{bmatrix} - W_0 \right) H_0 = \begin{bmatrix} I_{\beta_0} & \\ & P \end{bmatrix},$$

where P_d is the degree d Horner shift of P. Here the dependences on λ have been dropped for simplicity.

Proof. Observe first that, for $0 \leq i \leq k-2$, parts (a) and (b) of Lemma 4.2 guarantee that the products $N_i W_i H_i$ are well-defined and that the block partitions of N_i, W_i, and H_i are conformal for this matrix product. Moreover, from Theorem 3.6, the block $W_i(1,1)$ always has size $m \times n$, and so the size of $W_i(2:i{+}2, 2:i{+}2)$ is $\alpha_i \times \alpha_i$. Therefore $\operatorname{diag}(\lambda P_{k-i-2}, \lambda I_{\alpha_i})$ has the same size as W_i, and can be partitioned into blocks exactly in the same way as W_i, where recall that the diagonal blocks $W_i(2,2), \ldots, W_i(i+2, i+2)$ are square. As a consequence, the products $N_i \operatorname{diag}(\lambda P_{k-i-2}, \lambda I_{\alpha_i})\, H_i$ are also well-defined.

The rest of the proof consists in carefully performing block matrix products. We start with the proof of part (a), which must be split into two cases.

(a1) Part (a), case 1: *σ has a consecution at i.* In this case we have

$$N_i = \begin{bmatrix} I_m & 0 \\ \lambda N_{i-1}(:,1) & N_{i-1} \end{bmatrix}, \quad W_i = \begin{bmatrix} -A_{i+1} & I_m & 0 \\ W_{i-1}(:,1) & 0 & W_{i-1}(:,2:i+1) \end{bmatrix}, \quad H_i = \begin{bmatrix} 0 & H_{i-1}(1,:) \\ -I_m & P_{k-i-1}H_{i-1}(1,:) \\ 0 & H_{i-1}(2:i+1,:) \end{bmatrix}.$$

This implies

$$N_i \begin{bmatrix} \lambda P_{k-i-2} & \\ & \lambda I_{\alpha_i} \end{bmatrix} H_i = \begin{bmatrix} 0 & \lambda P_{k-i-2}H_{i-1}(1,:) \\ -\lambda N_{i-1}(:,1) & (S_i)_{22} \end{bmatrix}, \tag{27}$$

where

$$(S_i)_{22} = N_{i-1}(:,1)(\lambda^2 P_{k-i-2} + \lambda P_{k-i-1})H_{i-1}(1,:) + \lambda N_{i-1}(:,2:i+1)H_{i-1}(2:i+1,:),$$

and also

$$N_i W_i H_i = \begin{bmatrix} -I_m & (-A_{i+1} + P_{k-i-1})H_{i-1}(1,:) \\ -\lambda N_{i-1}(:,1) & (T_i)_{22} \end{bmatrix}, \tag{28}$$

where

$$\begin{aligned}(T_i)_{22} &= N_{i-1}(:,1)(-\lambda A_{i+1} + \lambda P_{k-i-1})H_{i-1}(1,:) \\ &\quad + N_{i-1}(W_{i-1}(:,1)H_{i-1}(1,:) + W_{i-1}(:,2:i+1)H_{i-1}(2:i+1,:)).\end{aligned}$$

Now we use (27), (28), and $-A_{i+1} + P_{k-i-1} = \lambda P_{k-i-2}$ to get

$$N_i \left(\begin{bmatrix} \lambda P_{k-i-2} & \\ & \lambda I_{\alpha_i} \end{bmatrix} - W_i \right) H_i = \begin{bmatrix} I_m & 0 \\ 0 & (Z_i)_{22} \end{bmatrix},$$

where

$$\begin{aligned}(Z_i)_{22} &= \lambda N_{i-1}(:,1)P_{k-i-1}H_{i-1}(1,:) + \lambda N_{i-1}(:,2:i+1)H_{i-1}(2:i+1,:) \\ &\quad - N_{i-1}(W_{i-1}(:,1)H_{i-1}(1,:) + W_{i-1}(:,2:i+1)H_{i-1}(2:i+1,:)) \\ &= N_{i-1} \begin{bmatrix} \lambda P_{k-i-1} & \\ & \lambda I_{\alpha_{i-1}} \end{bmatrix} H_{i-1} - N_{i-1}W_{i-1}H_{i-1} \\ &= N_{i-1} \left(\begin{bmatrix} \lambda P_{k-i-1} & \\ & \lambda I_{\alpha_{i-1}} \end{bmatrix} - W_{i-1} \right) H_{i-1}.\end{aligned}$$

(a2) Part (a), case 2: *σ has an inversion at i.* In this case

$$N_i \begin{bmatrix} \lambda P_{k-i-2} & \\ & \lambda I_{\alpha_i} \end{bmatrix} H_i = \begin{bmatrix} 0 & -\lambda H_{i-1}(1,:) \\ \lambda N_{i-1}(:,1)P_{k-i-2} & (S_i)_{22} \end{bmatrix}, \tag{29}$$

where $(S_i)_{22}$ is the same as in (a1), and

$$N_i W_i H_i = \begin{bmatrix} -I_n & -\lambda H_{i-1}(1,:) \\ \lambda N_{i-1}(:,1)P_{k-i-2} & (\widetilde{T}_i)_{22} \end{bmatrix}, \tag{30}$$

where now

$$\begin{aligned}(\widetilde{T}_i)_{22} &= N_{i-1}(:,1)(-\lambda A_{i+1} + \lambda P_{k-i-1})H_{i-1}(1,:) \\ &\quad + (N_{i-1}(:,1)W_{i-1}(1,:) + N_{i-1}(:,2:i+1)W_{i-1}(2:i+1,:))H_{i-1}.\end{aligned}$$

Subtracting (30) from (29) and reasoning as in (a1), we again obtain the desired identity.

The proof of part (b) is again a direct block matrix product and is omitted. We only remark that one has to again consider two separate cases: σ has a consecution at 0, and σ has an inversion at 0. Also one must use that $\lambda P_{k-2} = P_{k-1} - A_1$ and $P(\lambda) = \lambda^2 P_{k-2} + \lambda A_1 + A_0$. □

Now we are ready to prove the main result of this section.

Theorem 4.5. *Let $P(\lambda)$ be an $m \times n$ matrix polynomial with degree larger than or equal to 2. Then any Fiedler companion pencil $F_\sigma(\lambda)$ of $P(\lambda)$ is a strong linearization for $P(\lambda)$.*

Proof. We denote as usual $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, and adopt the notation used in Lemma 4.4. Moreover, recall from Definition 4.3 that $U_\sigma(\lambda) = N_{k-2}$ and $V_\sigma(\lambda) = H_{k-2}$, from Definition 3.8 that $M_\sigma = W_{k-2}$, that $P_0 = A_k$, and that $\alpha_{k-2} = m\mathfrak{c}(\sigma) + n\mathfrak{i}(\sigma)$. Then part (a) in Lemma 4.4 for $i = k-2$ implies

$$\begin{aligned} U_\sigma(\lambda)F_\sigma(\lambda)V_\sigma(\lambda) &= N_{k-2}\left(\lambda\begin{bmatrix} P_0 & \\ & I_{\alpha_{k-2}}\end{bmatrix} - W_{k-2}\right)H_{k-2} \\ &= \begin{bmatrix} I_{\beta_{k-2}} & \\ & N_{k-3}\left(\begin{bmatrix} \lambda P_1 & \\ & \lambda I_{\alpha_{k-3}}\end{bmatrix} - W_{k-3}\right)H_{k-3}\end{bmatrix}. \end{aligned}$$

Now apply part (a) in Lemma 4.4 for $i = k-3$ to the lower right block in the previous equation to get

$$U_\sigma(\lambda)F_\sigma(\lambda)V_\sigma(\lambda) = \begin{bmatrix} I_{\beta_{k-2}} & & \\ & I_{\beta_{k-3}} & \\ & & N_{k-4}\left(\begin{bmatrix} \lambda P_2 & \\ & \lambda I_{\alpha_{k-4}}\end{bmatrix} - W_{k-4}\right)H_{k-4}\end{bmatrix}.$$

Next, we apply part (a) in Lemma 4.4 consecutively for $i = k-4, k-5, \dots, 1$ until we get

$$U_\sigma(\lambda)F_\sigma(\lambda)V_\sigma(\lambda) = \begin{bmatrix} I_{\beta_{k-2}+\beta_{k-3}+\dots+\beta_1} & \\ & N_0\left(\begin{bmatrix} \lambda P_{k-2} & \\ & \lambda I_{\alpha_0}\end{bmatrix} - W_0\right)H_0\end{bmatrix}.$$

Finally apply part (b) in Lemma 4.4, and use $\beta_{k-2} + \dots + \beta_1 + \beta_0 = \alpha_{k-2}$ to obtain

$$U_\sigma(\lambda)F_\sigma(\lambda)V_\sigma(\lambda) = \begin{bmatrix} I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)} & \\ & P(\lambda)\end{bmatrix}, \tag{31}$$

which proves that $F_\sigma(\lambda)$ is a linearization of $P(\lambda)$, since $U_\sigma(\lambda)$ and $V_\sigma(\lambda)$ are unimodular.

To show that $F_\sigma(\lambda)$ is a strong linearization of $P(\lambda)$, we invoke Theorem 3.14, which states that $\operatorname{rev} F_\sigma(P)$ is strictly equivalent to $F_\tau(\operatorname{rev} P)$, where τ is a bijection (defined in the statement of Theorem 3.14) with the same total number of consecutions and the same total number of inversions as σ. We can now apply (31) to $F_\tau(\operatorname{rev} P)$ and $\operatorname{rev} P$ to see that $F_\tau(\operatorname{rev} P)$ is unimodularly equivalent to $\operatorname{diag}(I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)}, \operatorname{rev} P)$, and hence that $\operatorname{rev} F_\sigma(P)$ is unimodularly equivalent to $\operatorname{diag}(I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)}, \operatorname{rev} P)$. This proves that $F_\sigma(\lambda)$ is indeed a strong linearization of $P(\lambda)$. □

From the proof of Theorem 4.5, we get Corollary 4.6, which will be fundamental in Section 5.

Corollary 4.6. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $\sigma : \{0, \dots, k-1\} \to \{1, \dots, k\}$ be a bijection, and let $U_\sigma(\lambda)$ and $V_\sigma(\lambda)$ be, respectively, the left and right unimodular equivalence matrices associated with $P(\lambda)$ and σ that were introduced in Definition 4.3. Then*

$$U_\sigma(\lambda)F_\sigma(\lambda)V_\sigma(\lambda) = \begin{bmatrix} I_{m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma)} & \\ & P(\lambda)\end{bmatrix}.$$

5. Recovery of minimal indices and bases of rectangular polynomials from Fiedler pencils

In this section we show how to recover in a very simple way the minimal indices and bases of a rectangular matrix polynomial from those of any of its Fiedler pencils. The results and most of the proofs in this section are very similar to the ones presented for *singular square* matrix polynomials in Section 5 of [11]. Therefore, in order to avoid some repetitions, we omit all proofs that the reader can deduce easily from [11, Section 5], and pay close attention only to those arguments where the fact that the polynomial is rectangular makes a significant difference. Simultaneously, in order to keep the reading of the paper self-contained, we present careful statements of the main results and give exact pointers to the results in [11] where the proofs can be found. In this sense, this section has a different character than Sections 3 and 4, where most proofs have been presented in detail since the approaches followed in Sections 3 and 4 are new and very different from those in [11].

The main recovery results in this section are Corollaries 5.4 and 5.7, which are consequences of Theorems 5.3 and 5.6, respectively. These results extend to rectangular matrix polynomials what was previously proved in [11] *only for square singular polynomials*, specifically in Corollaries 5.8 and 5.11 and Theorems 5.7 and 5.9 in [11].

Corollaries 5.4 and 5.7 and Theorems 5.3 and 5.6 in this paper rely on a series of previous results. The first one is Lemma 5.1, which is an extension to rectangular matrix polynomials of [11, Lemma 5.1]. The proof is an obvious modification of the one given in [11, Lemma 5.1], and so is omitted.

Lemma 5.1. *Let the pencil $L(\lambda)$ be a linearization of an $m \times n$ matrix polynomial $P(\lambda)$ of degree $k \geq 2$, and let $U(\lambda)$ and $V(\lambda)$ be two unimodular matrix polynomials such that*

$$U(\lambda)L(\lambda)V(\lambda) = \begin{bmatrix} I_s & 0 \\ 0 & P(\lambda) \end{bmatrix}.$$

Let $U^L = U^L(\lambda)$ and $V^R = V^R(\lambda)$ be, respectively, the matrices comprising the last m rows of $U(\lambda)$ and the last n columns of $V(\lambda)$. Then

(a) *the linear map*

$$\begin{array}{rccc} \mathcal{L}: & \mathcal{N}_\ell(P) & \longrightarrow & \mathcal{N}_\ell(L) \\ & w^T(\lambda) & \longmapsto & w^T(\lambda) \cdot U^L \end{array}$$

is an isomorphism of $\mathbb{F}(\lambda)$-vector spaces;

(b) *the linear map*

$$\begin{array}{rccc} \mathcal{R}: & \mathcal{N}_r(P) & \longrightarrow & \mathcal{N}_r(L) \\ & v(\lambda) & \longmapsto & V^R \cdot v(\lambda) \end{array}$$

is an isomorphism of $\mathbb{F}(\lambda)$-vector spaces.

Clearly an immediate consequence of Lemma 5.1 is that the rational bases of $\mathcal{N}_r(P)$ and $\mathcal{N}_r(L)$ are in one-to-one correspondence via the map $\mathcal{R}$. But for an arbitrary linearization $L(\lambda)$, the map $\mathcal{R}$ does not necessarily establish a bijection between vector polynomial bases of $\mathcal{N}_r(P)$ and $\mathcal{N}_r(L)$, let alone between minimal bases of $\mathcal{N}_r(P)$ and $\mathcal{N}_r(L)$. A key point in our development is to show that for each Fiedler pencil $F_\sigma(\lambda)$ of an $m \times n$ matrix polynomial $P(\lambda)$, if $V_\sigma(\lambda)$ is the right unimodular equivalence matrix appearing in Corollary 4.6, then the map $\mathcal{R}_\sigma$ associated with $V_\sigma(\lambda)$ actually provides a one-to-one correspondence between the right minimal bases of $P(\lambda)$ and those of $F_\sigma(\lambda)$. This correspondence is so simple that it allows us to very easily obtain the right minimal bases of $P(\lambda)$ from the right minimal bases of $F_\sigma(\lambda)$, and hence also the right minimal indices of $P(\lambda)$ from those of $F_\sigma(\lambda)$. Analogous results hold for left minimal indices and bases.

Lemma 5.1 and Corollary 4.6 indicate that we need to determine the last block column of the matrix $V_\sigma(\lambda) = H_{k-2}$ introduced in Definition 4.3, since this last block column has precisely n columns according to Lemma 4.2(f). For this purpose, we need to define as in [11] some additional magnitudes and matrices, which are based on the consecution-inversion structure sequence of σ introduced in Definition 3.2, i.e., $\mathrm{CISS}(\sigma) = (c_1, i_1, \ldots, c_\ell, i_\ell)$. So we define

$$s_0 := 0\,, \quad s_j := \sum_{p=1}^{j} (c_p + i_p) \quad \text{for } j = 1, \ldots, \ell, \tag{32}$$

$$m_0 := 0\,, \quad m_j := \sum_{p=1}^{j} i_p \quad \text{for } j = 1, \ldots, \ell. \tag{33}$$

Note that $s_\ell = k-1$ and $m_\ell = \mathfrak{i}(\sigma)$. We also need to define some matrices, denoted $\Lambda_{\sigma,j}(P)$ for $j = 1, \ldots, \ell$ and $\widehat{\Lambda}_{\sigma,j}(P)$ for $j = 1, \ldots, \ell - 1$, associated with the $m \times n$ matrix polynomial $P(\lambda)$ and the bijection σ.

These matrices are defined in terms of the Horner shifts of $P(\lambda)$ as follows:

$$\Lambda_{\sigma,j}(P) := \begin{bmatrix} \lambda^{i_j} I_n \\ \vdots \\ \lambda I_n \\ I_n \\ P_{k-s_{j-1}-c_j} \\ \vdots \\ P_{k-s_{j-1}-2} \\ P_{k-s_{j-1}-1} \end{bmatrix} \quad \text{and} \quad \widehat{\Lambda}_{\sigma,j}(P) := \begin{bmatrix} \lambda^{i_j-1} I_n \\ \vdots \\ \lambda I_n \\ I_n \\ P_{k-s_{j-1}-c_j} \\ \vdots \\ P_{k-s_{j-1}-2} \\ P_{k-s_{j-1}-1} \end{bmatrix} \quad \text{if } c_1 \geq 1, \tag{34}$$

but if $c_1 = 0$, then $\Lambda_{\sigma,1}(P) := [\lambda^{i_1} I_n, \ldots, \lambda I_n, I_n]^T$, $\widehat{\Lambda}_{\sigma,1}(P) := [\lambda^{i_1-1} I_n, \ldots, \lambda I_n, I_n]^T$, with $\Lambda_{\sigma,j}(P)$, $\widehat{\Lambda}_{\sigma,j}(P)$ as in (34) for $j > 1$. Here for simplicity we omit λ from the Horner shifts $P_d(\lambda)$. With all these definitions we are now able to state and prove Lemma 5.2, which describes explicitly the last block-column of $V_\sigma(\lambda)$. Note that Lemma 5.2 generalizes [11, Lemma 5.3] to rectangular polynomials. However, the proof is completely different than that given in [11, Lemma 5.3].

Lemma 5.2. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $F_\sigma(\lambda)$ be the Fiedler pencil of $P(\lambda)$ associated with the bijection σ, and let $V_\sigma(\lambda)$ be the right unimodular equivalence matrix associated with $P(\lambda)$ and σ introduced in Definition* 4.3*. Consider $V_\sigma(\lambda)$ partitioned into $k \times k$ blocks according to* `Algorithm 5`*. Then the last block-column $V^R(\lambda)$ of $V_\sigma(\lambda)$, i.e., the last n columns of $V_\sigma(\lambda)$, is*

$$\Lambda_\sigma^R(P) := \begin{bmatrix} \lambda^{m_{\ell-1}} \Lambda_{\sigma,\ell}(P) \\ \lambda^{m_{\ell-2}} \widehat{\Lambda}_{\sigma,\ell-1}(P) \\ \vdots \\ \lambda^{m_1} \widehat{\Lambda}_{\sigma,2}(P) \\ \widehat{\Lambda}_{\sigma,1}(P) \end{bmatrix} \quad \text{if } \ell > 1, \tag{35}$$

and $V^R(\lambda) = \Lambda_{\sigma,1}(P) =: \Lambda_\sigma^R(P)$ if $\ell = 1$.

Proof. The last block-column of $V_\sigma(\lambda) = H_{k-2}$ can be determined by using `Algorithm 5`, and just looking at the last block-column at each step of the algorithm. Set $\text{CISS}(\sigma) = (c_1, i_1, \ldots, c_\ell, i_\ell)$. Assume first that $c_1 > 0$, which means that σ has consecutions at $0, 1, \ldots, c_1 - 1$. Then the last block-column of the matrix H_{c_1-1} constructed after steps $0, 1, \ldots, c_1 - 1$ of `Algorithm 5` is of the form

$$H_{c_1-1}(:, c_1+1) = \begin{bmatrix} I_n \\ P_{k-c_1} \\ \vdots \\ P_{k-1} \end{bmatrix}.$$

Next, σ has inversions at $c_1, c_1+1, \ldots, c_1+i_1-1$, so then the last block-column of $H_{c_1+i_1-1}$ is

$$H_{c_1+i_1-1}(:, c_1+i_1+1) = \begin{bmatrix} \lambda^{i_1} I_n \\ \vdots \\ \lambda I_n \\ I_n \\ P_{k-c_1} \\ \vdots \\ P_{k-1} \end{bmatrix} = \Lambda_{\sigma,1}(P).$$

The reader may check that if $c_1 = 0$, then $H_{i_1-1}(:, i_1+1) = [\lambda^{i_1} I_n, \ldots, \lambda I_n, I_n]^T = \Lambda_{\sigma,1}(P)$. The proof is complete here if $\ell = 1$, because in this case $c_1 + i_1 = k - 1$ and $H_{c_1+i_1-1}(:, c_1+i_1+1) = H_{k-2}(:, k)$ is the last block column of $V_\sigma(\lambda)$.

If $\ell > 1$, then the next c_2 consecutions of σ at $c_1 + i_1, c_i + i_1 + 1, \ldots, c_1 + i_1 + c_2 - 1$ give, according to `Algorithm 5`,

$$H_{c_1+i_1+c_2-1}(:, c_1 + i_1 + c_2 + 1) = \begin{bmatrix} \lambda^{i_1} I_n \\ \lambda^{i_1} P_{k-s_1-c_2} \\ \vdots \\ \lambda^{i_1} P_{k-s_1-1} \\ \lambda^{i_1-1} I_n \\ \vdots \\ \lambda I_n \\ I_n \\ P_{k-c_1} \\ \vdots \\ P_{k-1} \end{bmatrix} = \begin{bmatrix} \lambda^{i_1} I_n \\ \lambda^{i_1} P_{k-s_1-c_2} \\ \vdots \\ \lambda^{i_1} P_{k-s_1-1} \\ \widehat{\Lambda}_{\sigma,1}(P) \end{bmatrix}.$$

Notice that this produces the "truncated" block matrix $\widehat{\Lambda}_{\sigma,1}(P)$ at the bottom of the last block-column of $H_{c_1+i_1+c_2-1}$. The next i_2 inversions of σ produce

$$H_{s_2-1}(:, s_2 + 1) = \begin{bmatrix} \lambda^{m_1} \Lambda_{\sigma,2}(P) \\ \widehat{\Lambda}_{\sigma,1}(P) \end{bmatrix}.$$

The rest of the proof follows by induction, with arguments similar to the ones given above. Assume that for j such that $2 \le j < \ell$ we have

$$H_{s_j-1}(:, s_j + 1) = \begin{bmatrix} \lambda^{m_{j-1}} \Lambda_{\sigma,j}(P) \\ \lambda^{m_{j-2}} \widehat{\Lambda}_{\sigma,j-1}(P) \\ \vdots \\ \lambda^{m_1} \widehat{\Lambda}_{\sigma,2}(P) \\ \widehat{\Lambda}_{\sigma,1}(P) \end{bmatrix},$$

and prove via `Algorithm 5` that the corresponding result holds for $j + 1$. To finish the proof, note that $H_{s_\ell-1}(:, s_\ell + 1) = H_{k-2}(:, k)$ is precisely the last block column of $V_\sigma(\lambda)$. □

A fundamental fact in Lemma 5.2 is that $\Lambda_\sigma^R(P)$ always has *exactly one block equal to* I_n *at block index* $k - c_1$. This is the block that allows us to easily recover the minimal bases of $P(\lambda)$ from those of $F_\sigma(\lambda)$. To this purpose we first state Theorem 5.3, whose proof is omitted since it is essentially the same as the one of [11, Theorem 5.7].

Theorem 5.3. *Let* $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ *be an* $m \times n$ *matrix polynomial with degree* $k \ge 2$*, let* $F_\sigma(\lambda)$ *be the Fiedler pencil of* $P(\lambda)$ *associated with a bijection* σ*, let* $\mathfrak{i}(\sigma)$ *be the total number of inversions of* σ *and* $\mathfrak{c}(\sigma)$ *the total number of consecutions of* σ*, and let* $\Lambda_\sigma^R(P)$ *be the* $\big(n + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times n$ *matrix defined in* (35)*. Then the linear map*

$$\begin{array}{cccc} \mathcal{R}_\sigma : & \mathcal{N}_r(P) & \longrightarrow & \mathcal{N}_r(F_\sigma) \\ & v & \longmapsto & \Lambda_\sigma^R(P)\, v \end{array}$$

is an isomorphism of $\mathbb{F}(\lambda)$*-vector spaces with uniform degree-shift* $\mathfrak{i}(\sigma)$ *on the vector polynomials in* $\mathcal{N}_r(P)$*. More precisely,* $\mathcal{R}_\sigma$ *induces a bijection between the subsets of vector polynomials in* $\mathcal{N}_r(P)$ *and* $\mathcal{N}_r(F_\sigma)$*, with the property that*

$$\deg \mathcal{R}_\sigma(v) = \mathfrak{i}(\sigma) + \deg v$$

for every nonzero vector polynomial v*. Furthermore, for any nonzero vector polynomial* v*,* $\deg \mathcal{R}_\sigma(v)$ *is attained only in the topmost* $n \times 1$ *block of* $\mathcal{R}_\sigma(v)$*.*

An immediate consequence of Theorem 5.3 is Corollary 5.4, that establishes a very simple relationship between the right minimal indices and bases of $P(\lambda)$ and $F_\sigma(\lambda)$. The proof of Corollary 5.4 is also omitted since it is almost the same as the one of [11, Corollary 5.8].

Corollary 5.4 (Recovery of right minimal indices and bases). *Let $P(\lambda)$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, and let $F_\sigma(\lambda)$ be the Fiedler pencil of $P(\lambda)$ associated with a bijection σ having* $\text{CISS}(\sigma) = (c_1, i_1, \ldots, c_\ell, i_\ell)$ *and total number of consecutions and inversions $\mathfrak{c}(\sigma)$ and $\mathfrak{i}(\sigma)$, respectively. Suppose that each vector $z(\lambda) \in \mathcal{N}_r(F_\sigma) \subset \mathbb{F}(\lambda)^{(n+m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma))\times 1}$ is partitioned into $k \times 1$ blocks which are conformal for multiplication with the partition of $F_\sigma(\lambda)$ given by* `Algorithm 2`*.*

(a) *If $z(\lambda) \in \mathcal{N}_r(F_\sigma)$, and $x(\lambda) \in \mathbb{F}(\lambda)^{n\times 1}$ is the $(k - c_1)$*th *block of $z(\lambda)$, then $x(\lambda) \in \mathcal{N}_r(P)$.*

(b) *If $\{z_1(\lambda), \ldots, z_p(\lambda)\}$ is a right minimal basis of $F_\sigma(\lambda)$, and $x_j(\lambda)$ is the $(k - c_1)$*th *block of $z_j(\lambda)$ for each $j = 1, \ldots, p$, then $\{x_1(\lambda), \ldots, x_p(\lambda)\}$ is a right minimal basis of $P(\lambda)$.*

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

$$\varepsilon_1 + \mathfrak{i}(\sigma) \leq \varepsilon_2 + \mathfrak{i}(\sigma) \leq \cdots \leq \varepsilon_p + \mathfrak{i}(\sigma)\,,$$

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

Note that these results hold for the first companion form of $P(\lambda)$ using $(c_1, i_1) = (0, k-1)$ and $\mathfrak{i}(\sigma) = k-1$, and for the second companion form using $(c_1, i_1) = (k-1, 0)$ and $\mathfrak{i}(\sigma) = 0$.

For the recovery of left minimal indices and bases, it is possible to take a similar approach to the one we have used for right minimal indices and bases; that is, based on the results of Lemma 5.1 and Corollary 4.6, focus on the last m rows of $U_\sigma(\lambda)$, and determine them via `Algorithm 4`. However, we follow here a different strategy, based on the fact that the left minimal indices and bases of a matrix polynomial (and in particular, of a matrix pencil) coincide with the right minimal indices and bases of its transpose, since $y(\lambda)^T \in \mathcal{N}_\ell(P)$ if and only if $y(\lambda) \in \mathcal{N}_r(P^T)$. Then we relate the right minimal indices and bases of $P(\lambda)^T$ with the ones of $F_\sigma(\lambda)^T$, using the fact that $F_\sigma(\lambda)^T$ is a Fiedler pencil for $P(\lambda)^T$, as the following lemma shows.

Lemma 5.5. *Let $P(\lambda) = \sum_{i=0}^k \lambda^i A_i$ be an $m \times n$ matrix polynomial of degree $k \geq 2$, and $\sigma : \{0, \ldots, k-1\} \to \{1, \ldots, k\}$ a bijection. Define the* reversal bijection *of σ as follows: $\operatorname{rev}\sigma(i) := k + 1 - \sigma(i)$ for $i = 0, 1, \ldots, k-1$. Then*

$$[F_\sigma(P)]^T = F_{\operatorname{rev}\sigma}(P^T)\,.$$

Proof. This lemma can be easily proved by induction using `Algorithm 2`. Let $\{W_i\}_{i=0}^{k-2}$ be the sequence of matrices constructed by `Algorithm 2` for $P(\lambda)$ and σ, and let $\{W_i'\}_{i=0}^{k-2}$ be the sequence of matrices constructed by `Algorithm 2` for $P(\lambda)^T = \sum_{i=0}^k \lambda^i A_i^T$ and $\operatorname{rev}\sigma$. Note that $\operatorname{rev}\sigma$ has a consecution (resp., inversion) at i if and only if σ has an inversion (resp., consecution) at i. First notice that either

$$W_0^T = \begin{bmatrix} -A_1^T & -A_0^T \\ I_m & 0 \end{bmatrix} \quad \text{or} \quad W_0^T = \begin{bmatrix} -A_1^T & I_n \\ -A_0^T & 0 \end{bmatrix},$$

depending on whether σ has a consecution or an inversion at 0. Observe that, in both cases, we get $W_0^T = W_0'$. Now we proceed by induction: assume $W_{i-1}^T = W_{i-1}'$ for some $0 \leq (i-1) < k-2$, and prove that $W_i^T = W_i'$. For this purpose, use `Algorithm 2` to see that

$$W_i^T = \begin{bmatrix} -A_{i+1}^T & W_{i-1}(:,1)^T \\ I_m & 0 \\ 0 & W_{i-1}(:,2:i+1)^T \end{bmatrix} \quad \text{or} \quad W_i^T = \begin{bmatrix} -A_{i+1}^T & I_n & 0 \\ W_{i-1}(1,:)^T & 0 & W_{i-1}(2:i+1,:)^T \end{bmatrix},$$

depending on whether σ has a consecution or an inversion at i. Using the induction hypothesis, this can be seen to be precisely the same as the matrix W_i'. The conclusion of Lemma 5.5 is now just a restatement of $W_{k-2}^T = W_{k-2}'$. □

Lemma 5.5 allows us to prove Theorem 5.6 in essentially the same way as Theorem 5.9 in [11] was proved. Therefore we omit the proof of Theorem 5.6, although we remark that we cannot use here the block-transpose operation $(\cdot)^{\mathcal{B}}$, see [11, Definition 3.6], because the blocks of $\Lambda_{\operatorname{rev}\sigma}^R(P^T)$ do not all have the same sizes when $P(\lambda)$ is rectangular. This motivates a minor modification[2] in the statement of Theorem 5.6 as compared to the statement of [11, Theorem 5.9]. Note also that $\mathfrak{i}(\operatorname{rev}\sigma) = \mathfrak{c}(\sigma)$ and $\mathfrak{c}(\operatorname{rev}\sigma) = \mathfrak{i}(\sigma)$.

Theorem 5.6. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, let $F_\sigma(\lambda)$ be the Fiedler pencil of $P(\lambda)$ associated with bijection σ, let $\mathfrak{c}(\sigma)$ be the total number of consecutions of σ and $\mathfrak{i}(\sigma)$ the total number of inversions of σ, and let $\Lambda^R_{\operatorname{rev}\sigma}(P^T)$ be, for the $n \times m$ polynomial $P(\lambda)^T$ and the reversal bijection $\operatorname{rev}\sigma$, the $\big(m + m\,\mathfrak{c}(\sigma) + n\,\mathfrak{i}(\sigma)\big) \times m$ matrix defined in Lemma 5.2 . Then the linear map*

$$\begin{array}{rccc} \mathcal{L}_\sigma : & \mathcal{N}_\ell(P) & \longrightarrow & \mathcal{N}_\ell(F_\sigma) \\ & u^T & \longmapsto & u^T \Lambda^L_\sigma(P)\,, \end{array}$$

where $\Lambda^L_\sigma(P) := \left[\Lambda^R_{\operatorname{rev}\sigma}(P^T)\right]^T$, is an isomorphism of $\mathbb{F}(\lambda)$-vector spaces with uniform degree-shift $\mathfrak{c}(\sigma)$ on the vector polynomials in $\mathcal{N}_\ell(P)$. More precisely, $\mathcal{L}_\sigma$ induces a bijection between the subsets of vector polynomials in $\mathcal{N}_\ell(P)$ and $\mathcal{N}_\ell(F_\sigma)$, with the property that

$$\deg \mathcal{L}_\sigma(u^T) = \mathfrak{c}(\sigma) + \deg(u^T) \tag{36}$$

for every nonzero vector polynomial u^T. Furthermore, for any nonzero vector polynomial u^T, $\deg \mathcal{L}_\sigma(u^T)$ is attained only in the leftmost $1 \times m$ block of $\mathcal{L}_\sigma(u^T)$.

An immediate consequence of Theorem 5.6 is Corollary 5.7, which establishes a very simple relationship between the left minimal indices and bases of $P(\lambda)$ and $F_\sigma(\lambda)$. The easy proof is also omitted. We only indicate that the fact "$\operatorname{rev}\sigma$ has a consecution (resp., inversion) at i if and only if σ has an inversion (resp., consecution) at i" implies that $\Lambda^R_{\operatorname{rev}\sigma}(P^T)$ has exactly one block equal to I_m at block index k if $c_1 > 0$, and at block index $k - i_1$ if $c_1 = 0$.

Corollary 5.7 (Recovery of left minimal indices and bases). *Let $P(\lambda)$ be an $m \times n$ matrix polynomial with degree $k \geq 2$, and let $F_\sigma(\lambda)$ be the Fiedler pencil of $P(\lambda)$ associated with a bijection σ having* $\operatorname{CISS}(\sigma) = (c_1, i_1, \ldots, c_\ell, i_\ell)$ *and total number of consecutions and inversions $\mathfrak{c}(\sigma)$ and $\mathfrak{i}(\sigma)$, respectively. Suppose that each vector $z(\lambda)^T \in \mathcal{N}_\ell(F_\sigma) \subset \mathbb{F}(\lambda)^{1\times(m+m\mathfrak{c}(\sigma)+n\mathfrak{i}(\sigma))}$ is partitioned into $1 \times k$ blocks which are conformal for multiplication with the partition of $F_\sigma(\lambda)$ given by* `Algorithm 2`*.*

(a) *If $z(\lambda)^T \in \mathcal{N}_\ell(F_\sigma)$, and*

$$y(\lambda)^T \text{ is the } \left\{ \begin{array}{cl} k\text{th block of } z(\lambda)^T & \text{if } c_1 > 0, \\ (k-i_1)\text{th block of } z(\lambda)^T & \text{if } c_1 = 0, \end{array} \right.$$

then $y(\lambda)^T \in \mathcal{N}_\ell(P)$.

(b) *If $\{z_1(\lambda)^T, \ldots, z_q(\lambda)^T\}$ is a left minimal basis of $F_\sigma(\lambda)$, and*

$$y_j(\lambda)^T \text{ is the } \left\{ \begin{array}{cl} k\text{th block of } z_j(\lambda)^T & \text{if } c_1 > 0, \\ (k-i_1)\text{th block of } z_j(\lambda)^T & \text{if } c_1 = 0, \end{array} \right.$$

for $j = 1, \ldots, q$, then $\{y_1(\lambda)^T, \ldots, y_q(\lambda)^T\}$ is a left minimal basis of $P(\lambda)$.

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

$$\eta_1 + \mathfrak{c}(\sigma) \leq \eta_2 + \mathfrak{c}(\sigma) \leq \cdots \leq \eta_q + \mathfrak{c}(\sigma)\,,$$

are the left minimal indices of $F_\sigma(\lambda)$.

Note that these results hold for the first companion form of $P(\lambda)$ using $(c_1, i_1) = (0, k-1)$ and $\mathfrak{c}(\sigma) = 0$, and for the second companion form using $(c_1, i_1) = (k-1, 0)$ and $\mathfrak{c}(\sigma) = k-1$.

---

[2]In [11, Theorem 5.9] the matrix $\left[\Lambda^R_{\operatorname{rev}\sigma}(P)\right]^{\mathcal{B}}$ was used, while in Theorem 5.6 we use $\left[\Lambda^R_{\operatorname{rev}\sigma}(P^T)\right]^T$. Note that both expressions coincide for square matrix polynomials, but that $\left[\Lambda^R_{\operatorname{rev}\sigma}(P)\right]^{\mathcal{B}}$ is not defined for rectangular polynomials.

Next we include an example that illustrates the results presented in this section. This example extends to rectangular matrix polynomials what appears in [11, Example 5.12] only for square singular polynomials, which allows the reader to appreciate the strong similarities and the really minor differences between square and rectangular polynomials.

Example 5.8. *Let us consider an* $m \times n$ *matrix polynomial* $P(\lambda) = \sum_{i=0}^{6} \lambda^i A_i$ *with degree* 6*, and the Fiedler pencil* $F_\tau(\lambda)$ *of* $P(\lambda)$ *associated with the bijection* $\tau = (1, 2, 5, 3, 6, 4)$*. Recall that the zero degree term* M_τ *of this pencil was considered in* (16)*, and so*

$$F_\tau(\lambda) = \lambda \operatorname{diag}(A_6, I_n, I_m, I_n, I_m, I_m) - M_\tau.$$

Observe that $\mathrm{CISS}(\tau) = (2, 1, 1, 1)$*. So for* τ*, the parameters in* (32) *and* (33) *are* $\ell = 2$*,* $s_{\ell-1} = s_1 = 3$*, and* $m_{\ell-1} = m_1 = 1$*. In addition,* $\operatorname{rev}\tau = (6, 5, 2, 4, 1, 3)$*, hence* $\mathrm{CISS}(\operatorname{rev}\tau) = (0, 2, 1, 1, 1, 0)$*; also for* $\operatorname{rev}\tau$ *we have* $\ell = 3$*,* $s_1 = 2$*,* $s_{\ell-1} = s_2 = 4$*, and* $m_1 = 2$*,* $m_{\ell-1} = m_2 = 3$*. Therefore*

$$\begin{aligned} \Lambda_\tau^L(P) &= \left[\Lambda_{\operatorname{rev}\tau}^R(P^T)\right]^T = \left[\ \lambda^3 I_m \ \ \lambda^3 P_1(\lambda) \ \ \lambda^2 I_m \ \ \lambda^2 P_3(\lambda) \ \ \lambda I_m \quad I_m\ \right] \\ \textit{and} \quad \Lambda_\tau^R(P) &= \left[\ \lambda^2 I_n \quad \lambda I_n \quad \lambda P_2(\lambda)^T \quad I_n \quad P_4(\lambda)^T \ \ P_5(\lambda)^T\ \right]^T. \end{aligned}$$

The relationships between the minimal indices and bases of $F_\tau(\lambda)$ *and those of* $P(\lambda)$ *may now be summarized as follows:*

- *Right minimal indices of* $F_\tau(\lambda)$ *are shifted from those of* $P(\lambda)$ *by* $\mathfrak{i}(\tau) = 2$*.*
- *Left minimal indices of* $F_\tau(\lambda)$ *are shifted from those of* $P(\lambda)$ *by* $\mathfrak{c}(\tau) = 3$*.*
- *A right minimal basis of* $P(\lambda)$ *is recovered from the* $4\text{th} = (k - c_1)\text{th}$ *blocks* (*of size* $n \times 1$) *of any right minimal basis of* $F_\tau(\lambda)$*.*
- *A left minimal basis of* $P(\lambda)$ *is recovered from the* $6\text{th} = k\text{th}$ *blocks* (*of size* $1 \times m$) *of any left minimal basis of* $F_\tau(\lambda)$*.* $\square$

6. Conclusions and future work

In the last decade several new classes of linearizations for *square* matrix polynomials have been introduced by various authors [1, 2, 11, 12, 23, 27, 28, 34]. Among them, the class of Fiedler companion linearizations, which includes the classical first and second Frobenius companion forms, is a privileged class as a consequence of possessing the many valuable properties described in the Introduction. In this paper, we have extended Fiedler linearizations from square to *rectangular* matrix polynomials. To achieve this we have followed a completely different approach than the one followed in [2, 11] for regular and singular square polynomials, which cannot be easily generalized to the rectangular case. This new approach is based on a constructive definition via `Algorithm 2`, and has allowed us to prove that Fiedler pencils of rectangular matrix polynomials satisfy the same properties as Fiedler pencils of square matrix polynomials. More precisely, we have proved that every Fiedler pencil of a given rectangular polynomial $P(\lambda)$ is always a strong linearization for $P(\lambda)$, and that Fiedler pencils of rectangular matrix polynomials allow us to recover minimal indices and bases of matrix polynomials with essentially the same extremely simple rules as for Fiedler pencils of square polynomials. As far as we know, the class of Fiedler linearizations is the first of the new classes of linearizations introduced in the last decade that has been extended from square to rectangular polynomials. The most natural open problem in this context is to try to extend other classes of linearizations from square to rectangular matrix polynomials, e.g., the classes related to Fiedler pencils considered in [2, 5, 12, 34], or the vector spaces of linearizations introduced in [27]. Investigating the possibility of such extensions will be the subject of future work.

References

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

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

[3] T. G. Beelen and G. W. Veltkamp, Numerical computation of a coprime factorization of a transfer function matrix, Systems Control Lett. 9 (1987) 281–288.

[4] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, NLEVP: A collection of nonlinear eigenvalue problems, MIMS-EPrints 2010.98 and 2010.99, Manchester Institute for Mathematical Sciences, Manchester, UK, 2010.

[5] 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 (2011) 463–483.

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

[7] J. W. Demmel and B. Kågström, The generalized Schur decomposition of an arbitrary pencil $A-\lambda B$: Robust software with error bounds and applications. Part I: Theory and Algorithms, ACM Trans. Math. Software 19 (1993) 160–174.

[8] J. W. Demmel and B. Kågström, The generalized Schur decomposition of an arbitrary pencil $A-\lambda B$: Robust software with error bounds and applications. Part II: Software and Applications, ACM Trans. Math. Software 19 (1993) 175–201.

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

[10] 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 (2009) 371–402.

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

[12] F. De Terán, F. M. Dopico, and D. S. Mackey, Palindromic companion forms for matrix polynomials of odd degree, J. Comput. Appl. Math. 236 (2011) 1464–1480.

[13] A. Edelman, E. Elmroth, and B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part II. A stratification-enhanced staircase algorithm, SIAM J. Matrix Anal. Appl. 20 (1999) 667–699.

[14] M. Fiedler, A note on companion matrices, Linear Algebra Appl. 372 (2003) 325–331.

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

[16] F. R. Gantmacher, The Theory of Matrices, AMS Chelsea, Providence, RI, 1998.

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

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

[19] G. H. Golub and C. F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.

[20] G. E. Hayton, A. C. Pugh, and P. Fretwell, Infinite elementary divisors of a matrix polynomial and implications, Int. J. Control 47 (1988) 53–64.

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

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

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

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

[25] P. Lancaster, Lambda-matrices and vibrating systems, Pergamon Press, Oxford, 1966 (reprinted by Dover in 2002).

[26] 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, Manchester, UK, 2006.

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

[28] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, Structured polynomial eigenvalue problems: Good vibrations from good linearizations, SIAM J. Matrix Anal. Appl. 28 (2006) 1029–1051.

[29] D. Manocha and J. Demmel, Algorithms for intersecting parametric and algebraic curves I: simple intersections, ACM Trans. on Graphics 13 (1994) 73–100.

[30] D. Manocha and J. Demmel, Algorithms for intersecting parametric and algebraic curves II: mutiple intersections, Graphical Models and Image Processing 57 (1995) 81–100.

[31] L. Olson and T. Vandini, Eigenproblems from finite element analysis of fluid-structure interaction problems, Computers & Structures 33 (1989) 679–687.

[32] F. Tisseur and K. Meerbergen, The quadratic eigenvalue problem, SIAM Rev. 43 (2001) 235–286.

[33] P. Van Dooren, The computation of Kronecker's canonical form of a singular pencil, Linear Algebra Appl. 27 (1979) 103–140.

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