



Geometry of spaces for matrix polynomial Fiedler linearizations

by

Andrii Dmytryshyn, Stefan Johansson,
Bo Kågström, and Paul Van Dooren

UMINF 15.17

UMEÅ UNIVERSITY
DEPARTMENT OF COMPUTING SCIENCE
SE-901 87 UMEÅ
SWEDEN

Geometry of spaces for matrix polynomial Fiedler linearizations*

Andrii Dmytryshyn[†] Stefan Johansson[†]
Bo Kågström[†] Paul Van Dooren[‡]

Abstract

We study how small perturbations of matrix polynomials may change their elementary divisors and minimal indices by constructing the closure hierarchy graphs (stratifications) of orbits and bundles of matrix polynomial Fiedler linearizations. We show that the stratification graphs do not depend on the choice of Fiedler linearization which means that all the spaces of the matrix polynomial Fiedler linearizations have the same geometry (topology). The results are illustrated by examples using the software tool StratiGraph.

1 Introduction

For a long time matrix polynomials

$$P(\lambda) = \lambda^d A_d + \cdots + \lambda A_1 + A_0, \quad A_i \in \mathbb{C}^{m \times n}, i = 0, \dots, d, \text{ and } A_d \neq 0, \quad (1)$$

have been important objects to investigate. Due to challenging applications [23, 24, 31, 34, 36], matrix polynomials have received much attention in the last decade, resulting in rapid developments of corresponding theories [5, 6, 7, 27, 31] and computational techniques [3, 23, 28, 29, 32] (see also the

*Preprint Report UMINF 15.17, Department of Computing Science, Umeå University

[†]Department of Computing Science and HPC2N, Umeå University, SE-901 87 Umeå, Sweden. E-mails: andrii@cs.umu.se, stefanj@cs.umu.se, bokg@cs.umu.se

[‡]Department of Mathematical Engineering, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. E-mail: paul.vandooren@uclouvain.be

recent survey [33]). In a number of cases, the canonical structure information, i.e. elementary divisors and minimal indices of the matrix polynomials are the actual objects of interest. This information is usually computed via linearizations [3], in particular, Fiedler linearizations [1]. However, the canonical structure information is sensitive to perturbations in the coefficients matrices of the polynomial. How small perturbations may change the canonical structure information can be studied through constructing the orbit and bundle closure hierarchy (or stratification) graphs. Each node of such a graph represents a set of matrix polynomials with a certain canonical structure information, and there is an edge from one node to another if we can perturb any matrix polynomial associated with the first node such that its canonical structure information becomes equal to one of the matrix polynomials associated with the second node. The theory to compute and construct the stratification graphs are already known for several matrix problems: matrices under similarity (i.e., Jordan canonical form) [18], matrix pencils (i.e., Kronecker canonical form) [18], skew-symmetric matrix pencils [14], controllability and observability pairs [19], state-space system pencils [13], as well as full (normal) rank matrix polynomials [27]. Many of these results are already implemented in the *StratiGraph* software [26, 30, 35], which is a java-based tool developed to construct and visualize such closure hierarchy graphs. The *Matrix Canonical Structure (MCS) Toolbox* for Matlab [12, 26, 35] was also developed for simplifying the work with the matrices in canonical forms and connecting Matlab with StratiGraph. For more details on each of these cases we recommend to check the corresponding papers and their references; some control applications are discussed in [30].

In this paper, we study how small perturbations of (rectangular) matrix polynomials may change their elementary divisors and minimal indices by constructing the closure hierarchy graphs of the orbits and bundles of matrix polynomial Fiedler linearizations. Our results use and generalize the results of [27] where the same problem is solved for full-rank matrix polynomials. Other recent results that are crucial for the paper include necessary and sufficient conditions for a matrix polynomial with certain degree and canonical structure information to exist [7]; the strong linearization templates and how the minimal indices of such linearizations are related to the minimal indices of the polynomials [5]; the correspondence between perturbations of the linearizations and perturbations of matrix polynomials [27]; as well as the algorithm for the stratification of general matrix pencils [18]. In particular, the results in [5] and [7] allow us to consider polynomials with both left and

right minimal indices, in contrast to [27] (recall that full-rank polynomials may have either left or right minimal indices, not both types); as well as to use any Fiedler linearizations in contrast to the fixed choice of either the first or second companion forms (depending on which type of the minimal indices is present).

All matrices that we consider have complex entries.

2 Matrix pencils

We start by recalling the Kronecker canonical form of general matrix pencils $A - \lambda B$ (a matrix polynomial of degree one) under strict equivalence.

For each $k = 1, 2, \dots$, define the $k \times k$ matrices

$$J_k(\mu) := \begin{bmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{bmatrix}, \quad I_k := \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

where $\mu \in \mathbb{C}$, and for each $k = 0, 1, \dots$, define the $k \times (k + 1)$ matrices

$$F_k := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \end{bmatrix}, \quad G_k := \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}.$$

All non-specified entries of $J_k(\mu)$, I_k , F_k , and G_k are zeros.

An $m \times n$ matrix pencil $A - \lambda B$ is called *strictly equivalent* to $C - \lambda D$ if and only if there are non-singular matrices Q and R such that $Q^{-1}AR = C$ and $Q^{-1}BR = D$. The set of matrix pencils strictly equivalent to $A - \lambda B$ forms a manifold in the complex $2mn$ dimensional space. This manifold is the *orbit* of $A - \lambda B$ under the action of the group $GL_m(\mathbb{C}) \times GL_n(\mathbb{C})$ on the space of all matrix pencils by strict equivalence:

$$O_{A-\lambda B}^e = \{Q^{-1}(A - \lambda B)R : Q \in GL_m(\mathbb{C}), R \in GL_n(\mathbb{C})\}. \quad (2)$$

The *dimension* of $O_{A-\lambda B}^e$ is the dimension of its tangent space

$$T_{A-\lambda B}^e := \{(XA - AY) - \lambda(XB - BY) : X \in \mathbb{C}^{m \times m}, Y \in \mathbb{C}^{n \times n}\}$$

at the point $A - \lambda B$, $\dim T_{A-\lambda B}^e$. The orthogonal complement to $T_{A-\lambda B}^e$, with respect to the Frobenius inner product

$$\langle A - \lambda B, C - \lambda D \rangle = \text{trace}(AC^* + BD^*), \quad (3)$$

is called the normal space to this orbit. The dimension of the normal space is the *codimension* of $O_{A-\lambda B}^e$, denoted $\text{cod } O_{A-\lambda B}^e$, and is equal to $2mn$ minus the dimension of $O_{A-\lambda B}^e$. Explicit expressions for the codimensions of strict equivalence orbits are presented in [4].

Theorem 1. [20, Sect. XII, 4] *Each $m \times n$ matrix pencil $A - \lambda B$ is strictly equivalent to a direct sum, uniquely determined up to permutation of summands, of pencils of the form*

$$E_j(\mu) := J_j(\mu) - \lambda I_j, \text{ in which } \mu \in \mathbb{C}, \quad E_j(\infty) := I_j - \lambda J_j(0), \\ L_k := F_k - \lambda G_k, \quad \text{and} \quad L_k^T := F_k^T - \lambda G_k^T,$$

where $j \geq 1$ and $k \geq 0$.

The canonical form in Theorem 1 is known as the *Kronecker canonical form* (KCF). The blocks $E_j(\mu)$ (with up to $\min\{m, n\}$ different eigenvalues μ_i) and $E_j(\infty)$ correspond to the finite and infinite eigenvalues, respectively, and altogether form the regular part of $A - \lambda B$. The blocks L_k and L_k^T correspond to the right (column) and left (row) minimal indices, respectively, and form the singular part of the matrix pencil.

A *bundle* $B_{A-\lambda B}^e$ of a general matrix pencil $A - \lambda B$ is a union of orbits $O_{A-\lambda B}^e$ with the same singular structures and the same regular structures, except that the distinct eigenvalues may be different.

Computing the Kronecker canonical form is an ill-posed problem, i.e., small perturbations in the matrix entries may lead to completely different KCFs [17, 18]. This problem can be investigated by constructing a closure hierarchy (stratification) graph for orbits or bundles of matrix pencils [18], see, for example, the graph in Figure 4.

3 Matrix polynomials with prescribed invariants

In this section, we consider matrix polynomials (1) and recall the definitions of the canonical structure information for matrix polynomials, i.e., the elementary divisors and minimal indices, and state Theorem 4 (proven in [7]) that explains which canonical structure information a matrix polynomial may have.

Definition 2. Let $P(\lambda)$ and $Q(\lambda)$ be two $m \times n$ matrix polynomials. Then $P(\lambda)$ and $Q(\lambda)$ are *unimodular equivalent* if there exist two unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ (i.e., $\det U(\lambda), \det V(\lambda) \in \mathbb{C} \setminus \{0\}$) such that

$$U(\lambda)P(\lambda)V(\lambda) = Q(\lambda).$$

The transformation $P(\lambda) \mapsto U(\lambda)P(\lambda)V(\lambda)$ is called a unimodular equivalence transformation and the canonical form with respect to this transformation is the *Smith form* [20], recalled in the following theorem.

Theorem 3. [20] *Let $P(\lambda)$ be an $m \times n$ matrix polynomial over \mathbb{C} . Then there exists $r \in \mathbb{N}$, $r \leq \min\{m, n\}$ and unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ over \mathbb{C} such that*

$$U(\lambda)P(\lambda)V(\lambda) = \left[\begin{array}{ccc|c} g_1(\lambda) & & 0 & 0_{r \times (n-r)} \\ & \ddots & & \\ 0 & & g_r(\lambda) & \\ \hline & 0_{(m-r) \times r} & & 0_{(m-r) \times (n-r)} \end{array} \right], \quad (4)$$

where $g_j(\lambda)$ is monic for $j = 1, \dots, r$ and $g_j(\lambda)$ divides $g_{j+1}(\lambda)$ for $j = 1, \dots, r-1$. Moreover, the canonical form (4) is unique.

The integer r is the (normal) *rank* of the matrix polynomial $P(\lambda)$ and $P(\lambda)$ is called *full rank* if $r = \min\{m, n\}$.

Every $g_j(\lambda)$ is called an *invariant polynomial* of $P(\lambda)$, and can be uniquely factored as

$$g_j(\lambda) = (\lambda - \alpha_1)^{\delta_{j1}} \cdot (\lambda - \alpha_2)^{\delta_{j2}} \cdot \dots \cdot (\lambda - \alpha_{l_j})^{\delta_{jl_j}},$$

where $l_j \geq 0$, $\delta_{j1}, \dots, \delta_{jl_j} > 0$ are integers. If $l_j = 0$ then $g_j(\lambda) = 1$. The numbers $\alpha_1, \dots, \alpha_{l_j} \in \mathbb{C}$ are finite eigenvalues (zeros) of $P(\lambda)$. The *elementary divisors* of $P(\lambda)$ associated with the finite eigenvalue α_k is the collection of factors $(\lambda - \alpha_k)^{\delta_{jk}}$, including repetitions.

We say that $\lambda = \infty$ is an eigenvalue of the matrix polynomial $P(\lambda)$ if zero is an eigenvalue of $\text{rev } P(\lambda) := \lambda^d P(1/\lambda)$. The elementary divisors λ^{γ_k} , $\gamma_k > 0$, for the zero eigenvalue of $\text{rev } P(\lambda)$ are the elementary divisors associated with ∞ of $P(\lambda)$.

Define the left and right null-spaces, over the field $\mathbb{C}(\lambda)$, for an $m \times n$ matrix polynomial $P(\lambda)$ as follows:

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

Every subspace \mathcal{V} of the vector space $\mathbb{C}(\lambda)^n$ has bases consisting entirely of vector polynomials. Recall that, a minimal basis of \mathcal{V} is a basis of \mathcal{V} consisting of vector polynomials whose sum of degrees is minimal among all bases of \mathcal{V} consisting of vector polynomials. The ordered list of degrees of the vector polynomials in any minimal basis of \mathcal{V} is always the same. These degrees are called the minimal indices of \mathcal{V} . More formally, let the sets $\{y_1(\lambda)^T, \dots, y_{m-r}(\lambda)^T\}$ and $\{x_1(\lambda), \dots, x_{n-r}(\lambda)\}$ be minimal bases of $\mathcal{N}_{\text{left}}(P)$ and $\mathcal{N}_{\text{right}}(P)$, respectively, ordered so that $0 \leq \deg(y_1) \leq \dots \leq \deg(y_{m-r})$ and $0 \leq \deg(x_1) \leq \dots \leq \deg(x_{n-r})$. Let $\eta_k = \deg(y_k)$ for $i = 1, \dots, m-r$ and $\varepsilon_k = \deg(x_k)$ for $i = 1, \dots, n-r$. Then the scalars $0 \leq \eta_1 \leq \eta_2 \leq \dots \leq \eta_{m-r}$ and $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_{n-r}$ are, respectively, the *left* and *right minimal indices* of $P(\lambda)$.

To understand which combinations of the elementary divisors and minimal indices a matrix polynomial of certain degree may have, we use the following theorem.

Theorem 4. [7] *Let m, n, d , and r , such that $r \leq \min\{m, n\}$ be given positive integers. Let $g_1(\lambda), g_2(\lambda), \dots, g_r(\lambda)$ be r arbitrarily monic polynomials with coefficients in \mathbb{C} and with respective degrees $\delta_1, \delta_2, \dots, \delta_r$, such that $g_j(\lambda)$ divides $g_{j+1}(\lambda)$ for $j = 1, \dots, r-1$. Let $0 \leq \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_r$, $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_{n-r}$, and $0 \leq \eta_1 \leq \eta_2 \leq \dots \leq \eta_{m-r}$ be given lists of integers. There exists an $m \times n$ matrix polynomial $P(\lambda)$ with rank r , degree d , invariant polynomials $g_1(\lambda), g_2(\lambda), \dots, g_r(\lambda)$, partial multiplicities at ∞ equal to $\gamma_1, \gamma_2, \dots, \gamma_r$, and with right and left minimal indices equal to $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{n-r}$ and $\eta_1, \eta_2, \dots, \eta_{m-r}$, respectively, if and only if*

$$\sum_{j=1}^r \delta_j + \sum_{j=1}^r \gamma_j + \sum_{j=1}^{n-r} \varepsilon_j + \sum_{j=1}^{m-r} \eta_j = dr \quad (\text{index sum identity}) \quad (5)$$

holds and $\gamma_1 = 0$.

The condition $\gamma_1 = 0$ guarantees that $A_d \neq 0$ in (1).

4 Fiedler linearizations of matrix polynomials

Let us define *Fiedler linearizations* [1], with all the details, for the square matrix polynomials ($m = n$). Let $G(\lambda) = \sum_{k=0}^d \lambda^k A_k$ be an $n \times n$ matrix

polynomial. Given any bijection $\sigma : \{0, 1, \dots, d-1\} \rightarrow \{1, \dots, d\}$ with inverse σ^{-1} , the Fiedler pencil $\mathcal{F}_{G(\lambda)}^\sigma$ of $G(\lambda)$ associated with σ is the $dn \times dn$ matrix pencil

$$\mathcal{F}_{G(\lambda)}^\sigma := \lambda M_d - M_{\sigma^{-1}(1)} M_{\sigma^{-1}(2)} \dots M_{\sigma^{-1}(d)}, \quad (6)$$

where

$$M_d := \begin{bmatrix} A_d & & \\ & I_{(d-1)n} & \\ & & \ddots \\ & & & I_n \end{bmatrix}, \quad M_0 := \begin{bmatrix} I_{(d-1)n} & & \\ & & \\ & & \\ & & & -A_0 \end{bmatrix},$$

and

$$M_k := \begin{bmatrix} I_{(d-k-1)n} & & & & \\ & -A_k & I_n & & \\ & & I_n & 0 & \\ & & & & \\ & & & & I_{(k-1)n} \end{bmatrix}, \quad k = 1, \dots, d-1.$$

Note that $\sigma(k)$ describes the position of the factor M_k in the product defining the zero-degree term in (6), i.e. $\sigma(k) = j$ means that M_k is the j^{th} factor in the product.

By using bijections σ we can construct Fiedler linearizations via a ‘‘multiplication free’’ algorithm (i.e., by avoiding multiplying the matrices M_k) [6]. The advantage of such an algorithm is that it can be adapted to rectangular matrix polynomials. Note that the ‘‘shapes’’ of the linearizations (\equiv positions of the coefficient-matrices in the linearization pencils) for the rectangular matrix polynomials are the same as for the square matrix polynomials [6]. Moreover, different linearizations of rectangular matrix polynomials have different sizes, see Example 17.

Probably, the most known Fiedler linearizations are the first and second companion forms. For an $m \times n$ matrix polynomial $P(\lambda)$ of degree d they can be expressed as the matrix pencils

$$\mathcal{C}_{P(\lambda)}^1 = \lambda \begin{bmatrix} A_d & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} \quad (7)$$

and

$$\mathcal{C}_{P(\lambda)}^2 = \lambda \begin{bmatrix} A_d & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix} + \begin{bmatrix} A_{d-1} & -I_m & & 0 \\ A_{d-2} & 0 & \ddots & \\ \vdots & \vdots & \ddots & -I_m \\ A_0 & 0 & \dots & 0 \end{bmatrix} \quad (8)$$

of the sizes $(m + n(d-1)) \times nd$ and $md \times (n + m(d-1))$, respectively.

Fiedler linearizations preserve finite and infinite elementary divisors but do not usually preserve the left and right minimal indices (in some cases the minimal indices may also be preserved, e.g., for full rank polynomials [27]). In Theorem 5, proven in [6], we recall the relation between the minimal indices of polynomials and their Fiedler linearizations; see also [5] for the same results on square matrix polynomials.

We say that a bijection $\sigma : \{0, 1, \dots, d-1\} \rightarrow \{1, \dots, d\}$ has a consecution at k if $\sigma(k) < \sigma(k+1)$, and that σ has an inversion at k if $\sigma(k) > \sigma(k+1)$, where $k = 0, \dots, d-2$. Define $i(\sigma)$ and $c(\sigma)$ to be the total numbers of inversions and consecutions in σ , respectively. Note that

$$i(\sigma) + c(\sigma) = d - 1 \tag{9}$$

for every σ .

Theorem 5. [6] *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of degree $d \geq 2$, and let $\mathcal{F}_{P(\lambda)}^\sigma$ be its Fiedler linearization. If $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_s$ and $0 \leq \eta_1 \leq \eta_2 \leq \dots \leq \eta_t$ are the right and left minimal indices of $P(\lambda)$ then*

$$0 \leq \varepsilon_1 + i(\sigma) \leq \varepsilon_2 + i(\sigma) \leq \dots \leq \varepsilon_s + i(\sigma)$$

and

$$0 \leq \eta_1 + c(\sigma) \leq \eta_2 + c(\sigma) \leq \dots \leq \eta_t + c(\sigma),$$

are the right and left minimal indices of $\mathcal{F}_{P(\lambda)}^\sigma$.

Remark 6. Theorem 5 can be used for the first and second companion forms by putting the corresponding values of $i(\sigma)$ and $c(\sigma)$. For the first companion form $\mathcal{C}_{P(\lambda)}^1$, we have $i(\sigma) = d - 1$ and $c(\sigma) = 0$, and for the second companion form $\mathcal{C}_{P(\lambda)}^2$, we have $i(\sigma) = 0$ and $c(\sigma) = d - 1$.

Theorems 4 and 5 allow us to describe all the possible combinations of elementary divisors and minimal indices that the Fiedler linearizations of matrix polynomials of certain degree may have. In other words, we can identify those orbits of general matrix pencils which contain pencils that are the linearizations of some $m \times n$ matrix polynomials of certain degree.

4.1 Orbits of linearizations of matrix polynomials and their codimensions

The definitions and results in this section will be stated for the first companion form $\mathcal{C}_{P(\lambda)}^1$ but are valid for all Fiedler linearizations.

Define the *generalized Sylvester space* at $P(\lambda)$ as follows (see [27] and references therein)

$$\text{GSYL}(\mathcal{C}_{P(\lambda)}^1) = \{\mathcal{C}_{P(\lambda)}^1 : P(\lambda) \text{ are } m \times n \text{ matrix polynomials}\}. \quad (10)$$

If there is no risk of confusion we will write GSYL instead of $\text{GSYL}(\mathcal{C}_{P(\lambda)}^1)$. Now we define the orbit of linearizations of matrix polynomials

$$\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1} = \{(Q^{-1}\mathcal{C}_{P(\lambda)}^1 R) \in \text{GSYL}(\mathcal{C}_{P(\lambda)}^1) : Q \in GL_m(\mathbb{C}), R \in GL_n(\mathbb{C})\}. \quad (11)$$

Note that every element in $\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}$ is a linearization of $P(\lambda)$ in contrast with $\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e$ which also contains matrix pencils that are not linearizations of $P(\lambda)$ (or any other polynomial). By [27, Lemma 9.2], $\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}$ is a manifold in the matrix pencil space. Codimensions of this manifold are also of our interest, since they provide a coarse stratification: An orbit has only orbits with lower codimensions in its closure. Recall that $\dim \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e := \dim \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e$ and $\text{cod } \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e := \dim \mathbb{N}_{\mathcal{C}_{P(\lambda)}^1}^e$, where \mathbb{N} denotes the normal space (see Section 2). Define $\dim \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1} := \dim(\text{GSYL} \cap \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e)$. The following lemma shows that the codimensions of $\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}$ and $\mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e$ are the same; the latter is computed in [4] (see also [17, 21]) and implemented in the MCS Toolbox [35]. We also refer to [27] for a slightly different explanation of the analogous results.

Lemma 7. *Let $\mathcal{C}_{P(\lambda)}^1$ be the first companion form for the matrix polynomial $P(\lambda)$ then $\text{cod } \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1} = \text{cod } \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e$.*

Proof. Note that $\mathbb{C}^{(m+n(d-1)) \times nd} \times \mathbb{C}^{(m+n(d-1)) \times nd}$ is the least affine space containing $\mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e$ and GSYL, and since $\mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e \cap \text{GSYL} \neq \emptyset$ we have

$$\begin{aligned} \dim(\mathbb{C}^{(m+n(d-1)) \times nd} \times \mathbb{C}^{(m+n(d-1)) \times nd}) \\ = \dim \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e + \dim \text{GSYL} - \dim(\text{GSYL} \cap \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e), \end{aligned}$$

see [22, Section 2] for more details. Therefore

$$\begin{aligned} \text{cod } \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e &= \dim(\mathbb{C}^{(m+n(d-1)) \times nd} \times \mathbb{C}^{(m+n(d-1)) \times nd}) - \dim \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}^e \\ &= \dim \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e + \dim \text{GSYL} - \dim(\text{GSYL} \cap \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e) - \dim \mathbb{T}_{\mathcal{C}_{P(\lambda)}^1}^e \\ &= \dim \text{GSYL} - \dim \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1} = \text{cod } \mathcal{O}_{\mathcal{C}_{P(\lambda)}^1}. \end{aligned}$$

□

We remark that there are other examples where codimension equalities similar to the one in Lemma 7 do hold [19, 27] as well as examples where they are not valid [13, 15, 16].

5 Perturbations of matrix polynomials

Recall that for every matrix $X = [x_{ij}]$ its Frobenius norm is given by $\|X\| := \|X\|_F = (\sum_{i,j} x_{ij}^2)^{\frac{1}{2}}$. Define a norm of a matrix polynomial $P(\lambda) = \sum_{k=0}^d \lambda^k A_k$ as follows

$$\|P(\lambda)\| := \left(\sum_{k=0}^d \|A_k\|_F^2 \right)^{\frac{1}{2}}.$$

Definition 8. Let $P(\lambda)$ and $E(\lambda)$ be two $m \times n$ matrix polynomials, with $\deg P(\lambda) \geq \deg E(\lambda)$, and $\|E(\lambda)\|$ is arbitrarily small (in particular, $\|E(\lambda)\| \ll \|P(\lambda)\|$). A matrix polynomial $\widetilde{P(\lambda)} := P(\lambda) + E(\lambda)$ is a *perturbation of an $m \times n$ matrix polynomial $P(\lambda)$* .

We remark that Definition 8 is also applicable to matrix pencils and matrices (they are polynomials of degrees one and zero, respectively).

As in Section 4.1, the results are stated for $\mathcal{C}_{P(\lambda)}^1$ and the analogous results are valid for all Fiedler linearizations.

Theorem 9 (proven in [27]) ensures that each perturbation of the linearization of an $m \times n$ matrix polynomial of degree d

$$\begin{aligned} \widetilde{\mathcal{C}_{P(\lambda)}^1} := & \lambda \begin{bmatrix} A_d & & & & \\ & I_n & & & \\ & & \ddots & & \\ & & & I_n & \\ & & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} \\ & + \lambda \begin{bmatrix} E_{11} & E_{12} & E_{13} & \dots & E_{1d} \\ E_{21} & E_{22} & E_{23} & \dots & E_{2d} \\ E_{31} & E_{32} & E_{33} & \dots & E_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{d1} & E_{d2} & E_{d3} & \dots & E_{dd} \end{bmatrix} + \begin{bmatrix} E'_{11} & E'_{12} & E'_{13} & \dots & E'_{1d} \\ E'_{21} & E'_{22} & E'_{23} & \dots & E'_{2d} \\ E'_{31} & E'_{32} & E'_{33} & \dots & E'_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E'_{d1} & E'_{d2} & E'_{d3} & \dots & E'_{dd} \end{bmatrix} \quad (12) \end{aligned}$$

can be smoothly reduced by strict equivalence to the one in which only the

blocks $A_i, i = 0, 1, \dots$ are perturbed

$$\begin{aligned} \mathcal{C}_{\widetilde{P(\lambda)}}^1 &= \lambda \begin{bmatrix} A_d & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} \\ &+ \lambda \begin{bmatrix} F_d & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} F_{d-1} & F_{d-2} & \dots & F_0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}. \end{aligned} \quad (13)$$

We refer to (12) as a perturbation of the linearization and to (13) as the linearization of a perturbed matrix polynomial.

Theorem 9. [27] *Let $P(\lambda)$ be an $m \times n$ matrix polynomial, $\|P(\lambda)\| \gg \varepsilon$, and $\mathcal{C}_{P(\lambda)}^1$ be its first companion form. For every small $\varepsilon > 0$ such that $\|\widetilde{\mathcal{C}_{P(\lambda)}^1} - \mathcal{C}_{P(\lambda)}^1\| < \varepsilon$ and $\|\mathcal{C}_{\widetilde{P(\lambda)}}^1 - \mathcal{C}_{P(\lambda)}^1\| < \varepsilon$ there exist two nonsingular matrices X and Y (they are small perturbations of the identity matrices) such that*

$$X \cdot \widetilde{\mathcal{C}_{P(\lambda)}^1} \cdot Y = \mathcal{C}_{\widetilde{P(\lambda)}}^1.$$

The following corollary to Theorem 9 shows that all pencils that are attainable by perturbations of the form (12), are also attainable by perturbations of the form (13).

Corollary 10. Let $P(\lambda)$ and $Q(\lambda)$ be two $m \times n$ matrix polynomials, and $\mathcal{C}_{P(\lambda)}^1$ and $\mathcal{C}_{Q(\lambda)}^1$ be their first companion linearizations. There exists an arbitrarily small perturbation of $P(\lambda)$, denoted $\widetilde{P(\lambda)}$, and non-singular matrices U, V , such that

$$U \cdot \mathcal{C}_{\widetilde{P(\lambda)}}^1 \cdot V = \mathcal{C}_{Q(\lambda)}^1, \quad (14)$$

if and only if there exist an arbitrarily small perturbation of the linearization of the matrix polynomial $P(\lambda)$, $\widetilde{\mathcal{C}_{P(\lambda)}^1}$, and non-singular matrices U', V' , such that

$$U' \cdot \widetilde{\mathcal{C}_{P(\lambda)}^1} \cdot V' = \mathcal{C}_{Q(\lambda)}^1. \quad (15)$$

Proof. By Theorem 9 we have $X \cdot \widetilde{\mathcal{C}_{P(\lambda)}^1} \cdot Y = \mathcal{C}_{\widetilde{P(\lambda)}}^1$ and substituting $\widetilde{\mathcal{C}_{P(\lambda)}^1}$ in (15) we obtain $U' \cdot X^{-1} \cdot \mathcal{C}_{\widetilde{P(\lambda)}}^1 \cdot Y^{-1} \cdot V' = \mathcal{C}_{Q(\lambda)}^1$ which is (14) with $U = U' \cdot X^{-1}$ and $V = Y^{-1} \cdot V'$. The ‘‘vice versa’’ part is obvious. \square

Note that it is also possible to prove Theorem 9 using the theory of versal deformations [2, 9, 10] as it was done for state-space system pencils [13] or skew-symmetric polynomials in [8].

6 Orbit stratifications of the matrix polynomial linearizations

In this section, we present an algorithm for the stratification of the Fiedler linearizations of $m \times n$ matrix polynomials. The algorithm relies on the results presented in Sections 2–5.

Stratifications or closure hierarchy graphs for orbits of the matrix polynomial linearizations are defined as follows: Each node (vertex) of the graph represents the orbit of a matrix polynomial linearization and each edge represents a cover/closure relation, i.e., there is an upward path from a node associated with $\mathcal{F}_{P(\lambda)}^\sigma$ to a node associated with $\mathcal{F}_{Q(\lambda)}^\sigma$ if and only if $P(\lambda)$ can be transformed by an arbitrarily small perturbation to a matrix polynomial whose canonical structure information coincide with the one for $Q(\lambda)$.

The closure hierarchy graph obtained by the following algorithm is the orbit stratification of the first companion form of $m \times n$ matrix polynomials of degree d .

Algorithm 11. *Steps 1–3 produce the orbit stratification of the first companion linearizations of $m \times n$ matrix polynomials.*

Step 1. *Construct the stratification of $(m+n(d-1)) \times nd$ matrix pencil orbits under strict equivalence [18].*

Step 2. *Extract from the stratification obtained at Step 1 the nodes that correspond to the first companion linearizations of $m \times n$ matrix polynomials (using Theorems 4 and 5, as well as Remark 6).*

Step 3. *Put an edge between two nodes obtained at Step 2 if there is an upward path between these nodes in the graph obtained at Step 1 and do not put an edge otherwise (justified by Theorem 9 and Corollary 10).*

Analogous algorithms are valid for all Fiedler linearizations.

Theorem 12. *The stratification graphs for all the Fiedler linearizations $\mathcal{F}_{P(\lambda)}^\sigma$ of a polynomial $P(\lambda)$ are the same.*

Proof. Assume that there is an arrow from $\mathcal{C}_{P(\lambda)}^1$ to $\mathcal{C}_{Q(\lambda)}^1$ in the stratification of the first companion forms then $P(\lambda) + E(\lambda)$ and $Q(\lambda)$ have the same canonical structure information. Therefore for every σ the pencils $\mathcal{F}_{P(\lambda)+E(\lambda)}^\sigma$ and $\mathcal{F}_{Q(\lambda)}^\sigma$ have the same canonical structure information and thus there is an arrow from $\mathcal{F}_{P(\lambda)}^\sigma$ to $\mathcal{F}_{Q(\lambda)}^\sigma$ in the stratifications of all the Fiedler linearizations of $P(\lambda)$ and $Q(\lambda)$. \square

Remark 13. Note that Theorem 12 does not contradict the fact that for a particular matrix polynomial some linerizations may be better conditioned and/or structure preserving and therefore the choice of linearization is typically application driven.

6.1 Neighbouring orbits in the stratification

A sequence of integers $\mathcal{N} = (n_1, n_2, n_3, \dots)$ such that $n_1 + n_2 + n_3 + \dots = n$ and $n_1 \geq n_2 \geq \dots \geq 0$ is called an integer partition of n (for more details and references see [18]). For any $a \in \mathbb{Z}$ we define $\mathcal{N} + a$ as the integer partition $(n_1 + a, n_2 + a, n_3 + a, \dots)$. We write $\mathcal{N} \succcurlyeq \mathcal{M}$ if and only if $n_1 + n_2 + \dots + n_i \geq m_1 + m_2 + \dots + m_i$, for $i \geq 1$. The set of all integer partitions forms a poset (even a lattice) with respect to the order “ \succcurlyeq ”.

With every matrix pencil $W \equiv A - \lambda B$ (with eigenvalues $\mu_i \in \mathbb{C} \cup \infty$) we associate the set of integer partitions $\mathcal{R}(W), \mathcal{L}(W)$, and $\{\mathcal{J}_{\mu_i}(W) : j = 1, \dots, q, \mu_i \in \mathbb{C} \cup \infty\}$, where q is the number of distinct eigenvalues of W (e.g., see [18]). Altogether these partitions, known as the Weyr characteristics, are constructed as follows:

- For each distinct μ_i we have $\mathcal{J}_{\mu_i}(W) = (j_1^{\mu_i}, j_2^{\mu_i}, \dots)$, where $j_k^{\mu_i}$ is the number of Jordan blocks of size δ_{ij} greater than or equal to k (the position numeration starting from 1).
- $\mathcal{R}(W) = (r_0, r_1, \dots)$, where r_k is the number of L (right singular) blocks with the indices ε_i greater than or equal to k (the position numeration starting from 0).
- $\mathcal{L}(W) = (l_0, l_1, \dots)$, where l_k is the number of L^T (left singular) blocks with the indices η_i greater than or equal to k (the position numeration starting from 0).

Example 14. Let $W = 2E_3(\mu_1) \oplus E_1(\mu_1) \oplus 2E_2(\infty) \oplus L_4 \oplus L_1 \oplus L_1^T$ be an 18×19 matrix pencil in KCF. The associated partitions are:

$$\begin{aligned} \mathcal{J}_{\mu_1}(W) &= (3, 2, 2), & \mathcal{J}_{\infty}(W) &= (2, 2), \\ \mathcal{R}(W) &= (2, 2, 1, 1, 1), & \mathcal{L}(W) &= (1, 1). \end{aligned}$$

An integer partition $\mathcal{N} = (n_1, n_2, n_3, \dots)$ can also be represented by n piles of coins, where the first pile has n_1 coins, the second n_2 coins and so on. Moving one coin one column rightwards or one row downwards in the integer partition \mathcal{N} , and keep \mathcal{N} monotonically non-increasing, is called a *minimum rightward coin move*. Similarly, moving one coin one column leftwards or one row upwards in the integer partition \mathcal{N} , and keep \mathcal{N} monotonically non-increasing, is called a *minimum leftward coin move*. These two types of coin moves are defined in [18], see also Figure 1.

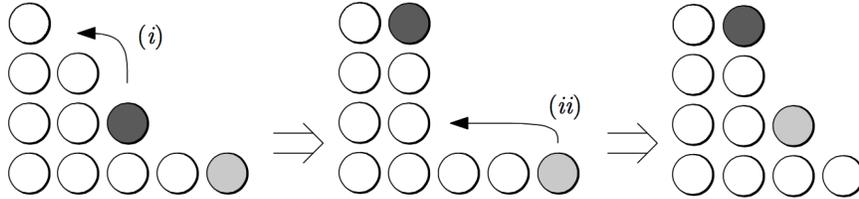


Figure 1: To the partition $(4, 3, 2, 1, 1)$, on the left, we apply two minimal leftward coin moves: first (i) is a move of a dark-grey coin one column leftward and then (ii) is a move of a light-grey coin one row upward. Note that monotonicity must be preserved. The resulting partition is $(4, 4, 2, 1)$, on the right.

By $\overline{\mathcal{X}}$ we denote the *closure* of a set \mathcal{X} in the Euclidean topology. We say that the orbit $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ is covered by $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ if and only if $\overline{O_{\mathcal{F}_{P_2(\lambda)}^\sigma}} \supset O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ and there exists no orbit $O_{\mathcal{F}_{Q(\lambda)}^\sigma}$ such that $\overline{O_{\mathcal{F}_{P_2(\lambda)}^\sigma}} \supset O_{\mathcal{F}_{Q(\lambda)}^\sigma}$ and $\overline{O_{\mathcal{F}_{Q(\lambda)}^\sigma}} \supset O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$; or equivalently, if and only if there is an edge from $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ to $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ in the orbit stratification ($O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ is higher up in the graph).

Representing the canonical structure information as integer partitions we can express the cover relations between two orbits by utilizing minimal coin moves and combinatorial rules on these integer partitions. The main idea of Theorem 15 is, starting from the corresponding sets of rules for general matrix pencils, to construct the rules that preserve the linearization structure, i.e., if the rules are applied to the linearization of an $m \times n$ matrix polynomial of

degree d then the resulting pencil is also the same linearization of another $m \times n$ matrix polynomial of degree d .

Theorem 15. $\mathcal{O}_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ is covered by $\mathcal{O}_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ if and only if $P_2(\lambda)$ can be obtained by applying one of the rules (a)–(d) to the structure integer partitions of $P_1(\lambda)$, (here $\mu_i \in \mathbb{C} \cup \infty$):

- (a) Minimum leftward coin move in \mathcal{R} (or \mathcal{L}).
- (b) If \mathcal{R} (or \mathcal{L}) is non-empty and the rightmost column in \mathcal{J}_{μ_i} is one single coin, move that coin to a new rightmost column of \mathcal{R} (or \mathcal{L}).
- (c) Minimum rightward coin move in any \mathcal{J}_{μ_i} .
- (d) If both \mathcal{R} and \mathcal{L} are non-empty: Let k denote the total number of coins in all of the longest (= lowest) rows from both \mathcal{R} and \mathcal{L} together. Remove these k coins, subtract one coin from the set, and distribute $k - 1$ coins as follows. First distribute one coin to each nonzero column in all existing \mathcal{J}_{μ_i} . The remaining coins are distributed among new rightmost columns, with one coin per column to any \mathcal{J}_{μ_i} which may be empty initially (i.e., new partitions for new eigenvalues can be created).

If $\mu_i = \infty$ for some i then $j_1^{\mu_i}$ has to remain strictly less than the rank of the corresponding polynomial (this restriction is due to consideration of the polynomials with the non-zero leading coefficients). Rules (a)–(b) are not allowed to do coin moves that affect r_0 or l_0 (first column in \mathcal{R} or \mathcal{L} , respectively). Rule (d) cannot be applied if the total number of nonzero columns of \mathcal{J}_{μ_i} is greater than $k - 1$.

Proof. First note that rules (a)–(d) coincide with the analogous rules for general matrix pencils (see Table 3(B) in [25], and [18, Theorem 3.2]). Therefore, applying any of the rules (a)–(d) to the partitions of $\mathcal{F}_{P_1(\lambda)}^\sigma$ for a matrix polynomial $P_1(\lambda)$, we get the partitions of the closest orbit in the general matrix pencil hierarchy. We need to show that there is a matrix polynomial $P_2(\lambda)$ such that $\mathcal{F}_{P_2(\lambda)}^\sigma$ has the obtained partitions. The canonical structure information of $P_2(\lambda)$ must then satisfy (5) in Theorem 4. It is obvious for rules (a)–(c) since they result in simultaneously adding 1 to some invariant and subtracting 1 from another invariant. Applying rule (d) to the integer partitions of $\mathcal{F}_{P_1(\lambda)}^\sigma$, we remove $\varepsilon + 1 + i(\sigma)$ coins from \mathcal{R} and $\eta + 1 + c(\sigma)$ coins from \mathcal{L} (hence $i(\sigma)$ and $c(\sigma)$ are the “linearization shifts”, see Theorem 5, and we add 1 since the numbering starts from 0). Thus, from rule (d)

and (9), the new degree δ of the corresponding invariant polynomial is equal to $\varepsilon + 1 + i(\sigma) + \eta + 1 + c(\sigma) - 1 = \varepsilon + \eta + d$. For $P_1(\lambda)$ the equality (5) holds and after applying rule (d): the right hand side of the equality (5) loses $\varepsilon + \eta$ but gains $\delta = \varepsilon + \eta + d$; r increases by 1; and the left hand side changes from rd to $(r+1)d$. Thus (5) holds for the canonical structure information associated with the obtained partitions too and there is a polynomial that has this canonical structure information by Theorem 4. Summing up, the partitions obtained by applying any of rules (a)–(d) correspond to some $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ that covers $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$.

Now assume that $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ covers $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ in the stratification of the linearizations. By Corollary 10 there is a path from $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}^e$ to $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}^e$ in the stratification of general matrix pencils. Therefore the partitions of $\mathcal{F}_{P_2(\lambda)}^\sigma$ are obtained from the partitions of $\mathcal{F}_{P_1(\lambda)}^\sigma$ by a sequence of rules (a)–(d) (recall that they coincide with the rules for the general matrix pencils). If the sequence has more than one rule then we have a contradiction with $O_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ covering $O_{\mathcal{F}_{P_1(\lambda)}^\sigma}$. □

Example 16. Consider a 2×2 matrix polynomial of degree 3, i.e.,

$$A_3\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0, \quad A_3 \neq 0. \quad (16)$$

By Theorem 4 such a matrix polynomial has the canonical structure information $\delta_1, \delta_2, \gamma_1, \gamma_2, \varepsilon_1$, and η_1 presented in one of the columns of Table 1 ($\delta_1, \delta_2, \gamma_1$ and γ_2 form the regular part; ε_1 and η_1 form the singular part). We now explain how small perturbations of the coefficient matrices, A_3, \dots, A_0 , of the polynomial may change this canonical structure information. For example, if a polynomial has the canonical structure information $\delta_1 = 1, \gamma_1 = 0, \varepsilon_1 = 0$, and $\eta_1 = 2$ (column 7 of Table 1) and if we perturb this polynomial its canonical structure information may change to $\delta_1 = 0, \gamma_1 = 0, \varepsilon_1 = 0$, and $\eta_1 = 3$ (column 4 of Table 1).

By Theorem 9 and Corollary 10, perturbations of Fiedler linearization pencils correspond to perturbations in the matrix coefficients of the underlying matrix polynomials. Thus we can investigate changes of the canonical structure information of the corresponding matrix pencil linearizations. Notably, the sets of the corresponding matrix pencils are different for different linearizations since Fiedler linearizations preserve elementary divisors but

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
δ_1	0	0	0	0	1	1	1	2	2	3	0	1	2	3
δ_2	-	-	-	-	-	-	-	-	-	-	6	5	4	3
γ_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
γ_2	-	-	-	-	-	-	-	-	-	-	0	0	0	0
ε_1	3	2	1	0	2	1	0	1	0	0	-	-	-	-
η_1	0	1	2	3	0	1	2	0	1	-	-	-	-	-
	15	16	17	18	19	20	21	22	23	24	25	26		
δ_1	0	1	2	0	1	2	0	1	0	1	0	0		
δ_2	5	4	3	4	3	2	3	2	2	1	1	0		
γ_1	0	0	0	0	0	0	0	0	0	0	0	0		
γ_2	1	1	1	2	2	2	3	3	4	4	5	6		
ε_1	-	-	-	-	-	-	-	-	-	-	-	-		
η_1	-	-	-	-	-	-	-	-	-	-	-	-		

Table 1: There exists a 2×2 matrix polynomial of degree 3 ($A_3 \neq 0$) with the canonical structure information $\delta_1, \delta_2, \gamma_1, \gamma_2, \varepsilon_1$, and η_1 if and only if $\delta_1, \delta_2, \gamma_1, \gamma_2, \varepsilon_1$, and η_1 are those in one of the columns of this table. Columns 1–10 correspond to singular polynomials and columns 11–26 to regular polynomials. (The table is split into two parts just to fit on the page).

“shift” the minimal indices, see Theorem 5. In this case, the following shifts are possible: for the first companion form (7), we have +2 for the right and no shift for the left minimal indices; for the second companion form (8), we have no shift for the right and +2 for the left minimal indices; for the linearizations

$$\lambda \begin{bmatrix} A_3 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} A_2 & A_1 & -I \\ -I & 0 & 0 \\ 0 & A_0 & 0 \end{bmatrix} \quad \text{and} \quad \lambda \begin{bmatrix} A_3 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} A_2 & -I & 0 \\ A_1 & 0 & A_0 \\ -I & 0 & 0 \end{bmatrix}, \quad (17)$$

with 1 inversion and 1 consecution, we have +1 for the right and +1 for the left minimal indices. We obtain the same stratification graph for all the linearizations, see Figure 2 and Theorem 12, otherwise it would mean that different linearizations “behave” generally different under small perturbations, but see also Remark 13.

Note that δ_j is just the degree of $g_j(\lambda)$ and it gives a few possibilities for the powers δ_{jk} of the elementary divisors. To be exact, the number of such possibilities is the number of ways the integer δ_j can be written as a sum of positive integers, i.e., $\delta_j = \delta_{j1} + \delta_{j2} + \dots + \delta_{jl_j}$. Thus some columns in Table 1 correspond to more than one node in the graph in Figure 2. Since the considered matrix polynomials may have rank at most 2 and $A_3 \neq 0$, by

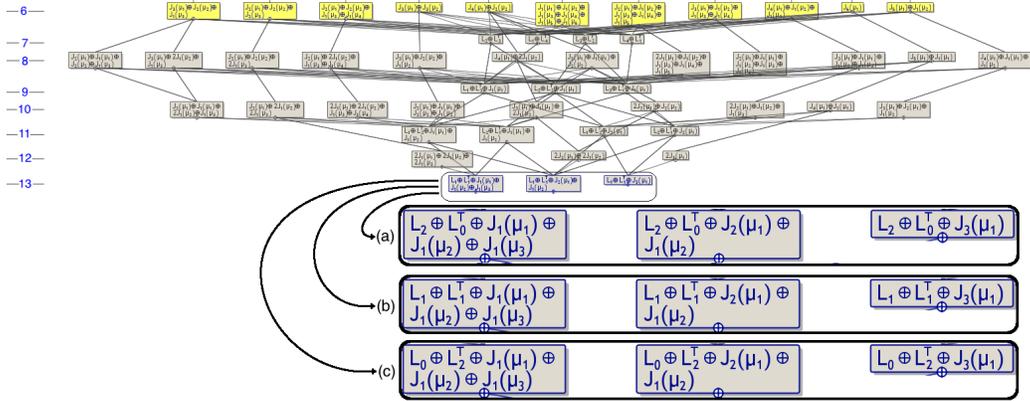


Figure 2: Orbit stratification of the linearizations of 2×2 matrix polynomials of degree 3 ($A_3 \neq 0$). Only the sizes of the singular canonical blocks depend on the choice of Fiedler linearization, not the numbers of singular blocks, the regular parts, or the closure relations (graph edges). In (a), (b), and (c) we show the three most degenerate structures (the bottom nodes of the graphs) for the first companion form, the linearizations (17), and the second companion form, respectively.

[7, Lemma 2.6] these polynomials may have at most 1 infinite elementary divisor. Therefore the eigenvalues in the nodes of Figure 2 which have two Jordan blocks associated with them can not be infinite.

Example 17. Consider rectangular 1×2 matrix polynomials of degree 3. Like in Example 16, we explain how small perturbations of the coefficient matrices of the polynomials may change their canonical structure information. By Theorem 4 such a polynomial has the canonical structure information δ_1, γ_1 , and ε_1 , presented in one of the four columns of Table 2. Note that the ranks of these polynomials are 1 and $A_3 \neq 0$. Thus by [7, Lemma 2.6] we have no infinite elementary divisors in this case.

	1	2	3	4
δ_1	0	1	2	3
γ_1	0	0	0	0
ε_1	3	2	1	0

Table 2: There exists a 1×2 matrix polynomial of degree 3 ($A_3 \neq 0$) with the canonical structure information δ_1, γ_1 , and ε_1 , if and only if δ_1, γ_1 , and ε_1 take the values in one of the columns of this table.

Since the polynomials are rectangular the Fiedler linearizations are of different sizes: the first companion form is 5×6 , the second companion form is 3×4 , and both linearizations in (17) are 4×5 . These Fiedler type linearizations “shift” the minimal indices exactly as in Example 16. The three graphs

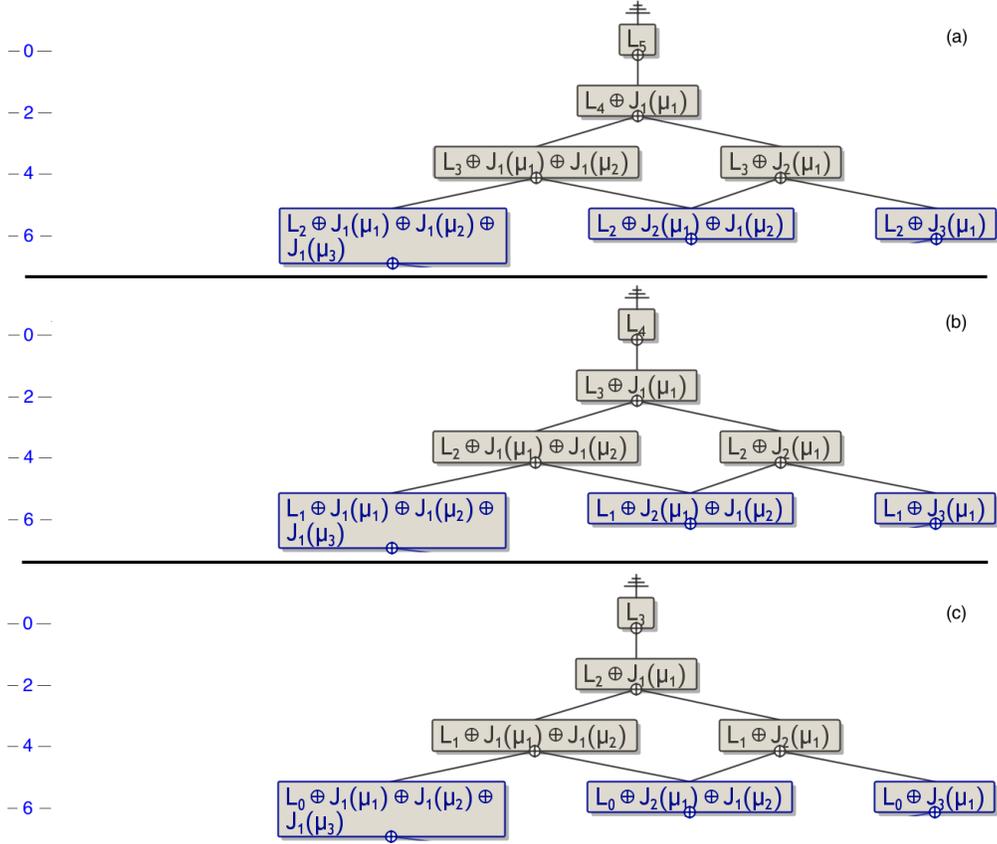


Figure 3: Orbit stratification of the Fiedler linearizations of 1×2 matrix polynomials of degree 3 ($A_3 \neq 0$). Graph (a) is the stratification of the first companion form, its nodes represent 5×6 matrix pencils. Graph (b) is the stratification of the linearizations in (17), its nodes represent 4×5 matrix pencils. Finally, graph (c) is the stratification of the second companion form, its nodes represent 3×4 matrix pencils.

in Figure 3 have the same set of edges that connect nodes corresponding to matrix pencil orbits with the same regular structures ($J_k(\mu)$ blocks) but that differ in the sizes of the singular structure (L_k blocks). For example,

the most generic nodes are L_5 for Figure 3(a), L_4 for Figure 3(b), and L_3 for Figure 3(c). Note that each of these graphs is a subgraph of the corresponding general matrix pencil stratification graph, for example, the graph in Figure 3(c) is a subgraph of the stratification graph of 3×4 matrix pencils, see Figure 4.

Note also that the polynomials in this example have full ranks. Thus we can apply the theory from [27] to construct graph (c) in Figure 3 (but not (a) or (b) since in [27] the choice of the linearization is fixed).

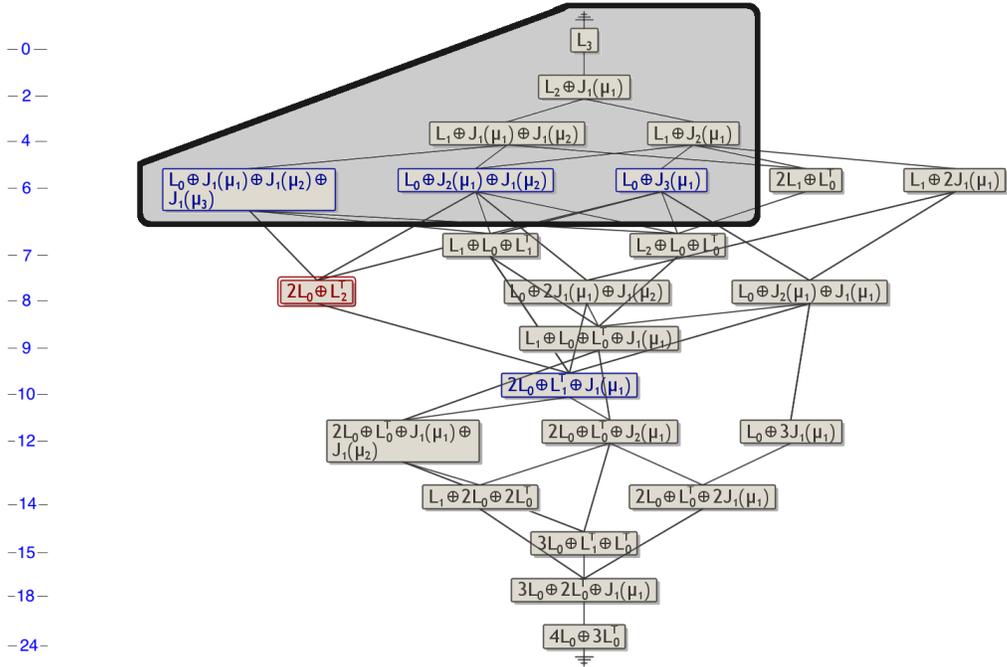


Figure 4: Orbit stratification for 3×4 matrix pencils. The subgraph in the grey region is exactly the one from Figure 3(c), i.e., it is the stratification of the second companion form of 1×2 matrix polynomials of degree 3 ($A_3 \neq 0$).

7 Bundle stratifications of the matrix polynomial linearizations

In the orbit stratifications, the eigenvalues may appear and disappear but their values cannot change. However in many applications, see for example

[19, 27, 30], the eigenvalues of the underlying matrices may coalesce or split apart to different eigenvalues, which motivates so called bundle stratifications. The theories for bundle stratifications are developed along with the theories for the orbit stratifications and are known for a number of cases [13, 14, 17, 18, 19, 27]. Similarly, we consider stratifications of the bundles of matrix polynomial Fiedler linearizations. Defining a bundle may be a problem by itself, in particular, for the cases where the behaviour of an eigenvalue depends on its value, see e.g. [11, Section 6]. Nevertheless, in our case of the matrix polynomial Fiedler linearizations all the eigenvalues have the same behaviour and the restriction on the number of Jordan blocks associated with the infinite eigenvalue, for example in Theorem 15, are coming from our desire to have non-zero leading coefficient matrices of the polynomials but not from the geometrical properties.

Following the definition of bundles for general matrix pencils, we define a bundle $B_{\mathcal{F}_{P(\lambda)}^\sigma}$ of the matrix polynomial linearization $\mathcal{F}_{P(\lambda)}^\sigma$ to be a union of orbits $O_{\mathcal{F}_{P(\lambda)}^\sigma}$ with the same singular structures and the same regular structures, except that the distinct eigenvalues may be different, see also [27]. Therefore we have that two Fiedler linearizations $\mathcal{F}_{P(\lambda)}^\sigma$ and $\mathcal{F}_{R(\lambda)}^\sigma$ are in the same bundle if and only if they are in the same bundle as general matrix pencils. This ensures that the stratification algorithm for the bundles of the matrix polynomial Fiedler linearizations is analogous to Algorithm 11. So we extract the bundles that correspond to the linearizations from the stratification of the general matrix pencil bundles and put an edge between two of them if there is a path between them in the stratification graph for the general matrix pencils. In addition, the codimensions of the bundles of $\mathcal{F}_{P(\lambda)}^\sigma$ are defined as

$$\text{cod } B_{\mathcal{F}_{P(\lambda)}^\sigma} = \text{cod } O_{\mathcal{F}_{P(\lambda)}^\sigma} - \# \left\{ \text{distinct eigenvalues of } \mathcal{F}_{P(\lambda)}^\sigma \right\}.$$

The definition for the cover relation is analogous to the one for orbits, see Section 6.1. The following theorem is the bundle analog of Theorem 15.

Theorem 18. $B_{\mathcal{F}_{P_1(\lambda)}^\sigma}$ is covered by $B_{\mathcal{F}_{P_2(\lambda)}^\sigma}$ if and only if $P_2(\lambda)$ can be obtained by applying one of the rules (a)–(e) to the structure integer partitions of $P_1(\lambda)$, (here $\mu_i \in \mathbb{C} \cup \infty$):

- (a) Minimum leftward coin move in \mathcal{R} (or \mathcal{L}).
- (b) If \mathcal{R} (or \mathcal{L}) is non-empty and \mathcal{J}_{μ_i} consist of one single coin, move that coin to a new rightmost column of \mathcal{R} (or \mathcal{L}).

- (c) *Minimum rightward coin move in any \mathcal{J}_{μ_i} .*
- (d) *If both \mathcal{R} and \mathcal{L} are non-empty: Let k denote the total number of coins in all of the longest (= lowest) rows from both \mathcal{R} and \mathcal{L} together. Remove these k coins, subtract one coin from the set, and distribute $k - 1$ coins as follows. First distribute one coin to each nonzero column in all existing \mathcal{J}_{μ_i} . The remaining coins are distributed among new rightmost columns, with one coin per column to any \mathcal{J}_{μ_i} which may be empty initially. New partitions for new finite eigenvalues may only be created if there exist no \mathcal{J}_{μ_i} . If a new set is created, all coins should be assigned to it and create one row.*
- (e) *Split \mathcal{J}_{μ_i} into two new partitions corresponding to two different eigenvalues.*

If $\mu_i = \infty$ for some i then $j_1^{\mu_i}$ has to remain strictly less than the rank of the corresponding polynomial (this restriction is since we consider the polynomials with the non-zero leading coefficients). Rules (a)–(b) are not allowed to do coin moves that affect r_0 or l_0 (first column in \mathcal{R} and \mathcal{L} , respectively). Rule (d) cannot be applied if the total number of nonzero columns of \mathcal{J}_{μ_i} is greater than $k - 1$.

Proof. Similarly to Theorem 15, rules (a)–(e) presented here coincide with the analogous rules for the general matrix pencils presented in Table 3(D) in [25], see also [18, Theorem 3.3]. The proof is essentially the same as the proof of Theorem 15. \square

Example 19. In Figure 5, we stratify the bundles of the Fiedler linearizations (17) of 2×2 matrix polynomials of degree 3. In the graph, each node represents a bundle and each edge a closure/cover relation. An arbitrarily small perturbation of coefficient matrices of matrix polynomials, in any bundle, may change the canonical structure to any more generic node that we have an upward path to.

We recall that the orbit stratification of the polynomials presented in Figure 2 has eleven most generic orbits, marked by yellow colour. In Figure 5 these eleven orbits are marked by yellow colour again but since eigenvalues are allowed to split apart in the bundle case only one of them is the most generic.

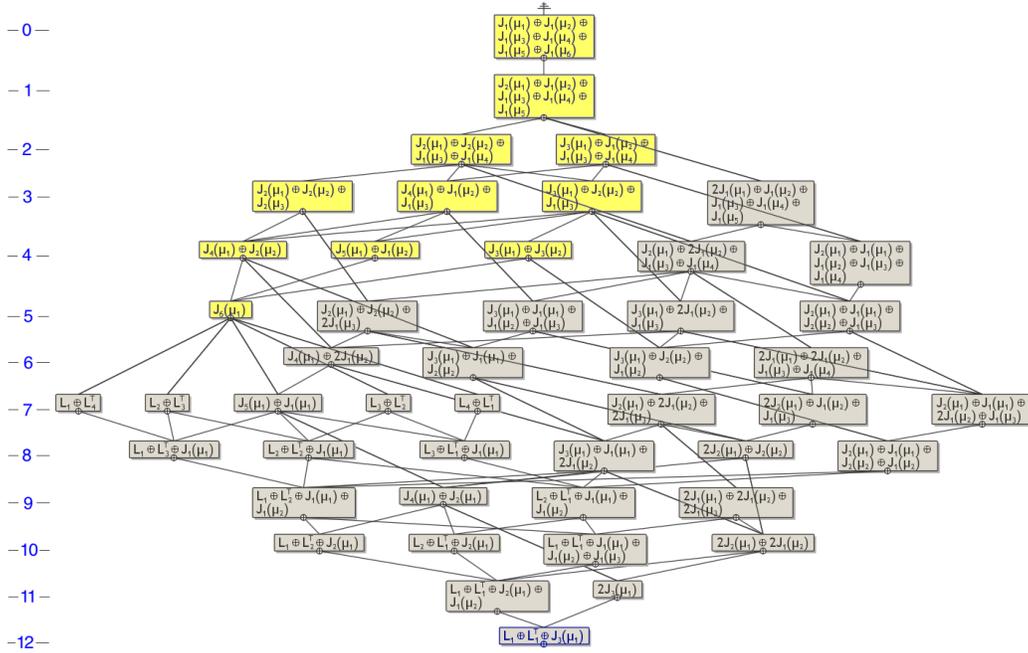


Figure 5: Bundle stratification of the Fiedler linearizations (17) of 2×2 matrix polynomials of degree 3.

Example 20. Similarly to Example 19, we stratify the bundles of the Fiedler linearizations of 1×2 matrix polynomials of degree 3 and present them in Figure 6. Recall that the orbit stratification graphs are presented in Figure 3, see Example 17. Notably, for the bundle case there is only one least generic node and one most generic node, the latter correspond to the same canonical structures for both the orbit and bundle cases.

Acknowledgements

The work was supported by the Swedish Research Council (VR) under grant E0485301, and by eSSANCE (essenceofscience.se), a strategic collaborative e-Science programme funded by the Swedish Research Council.

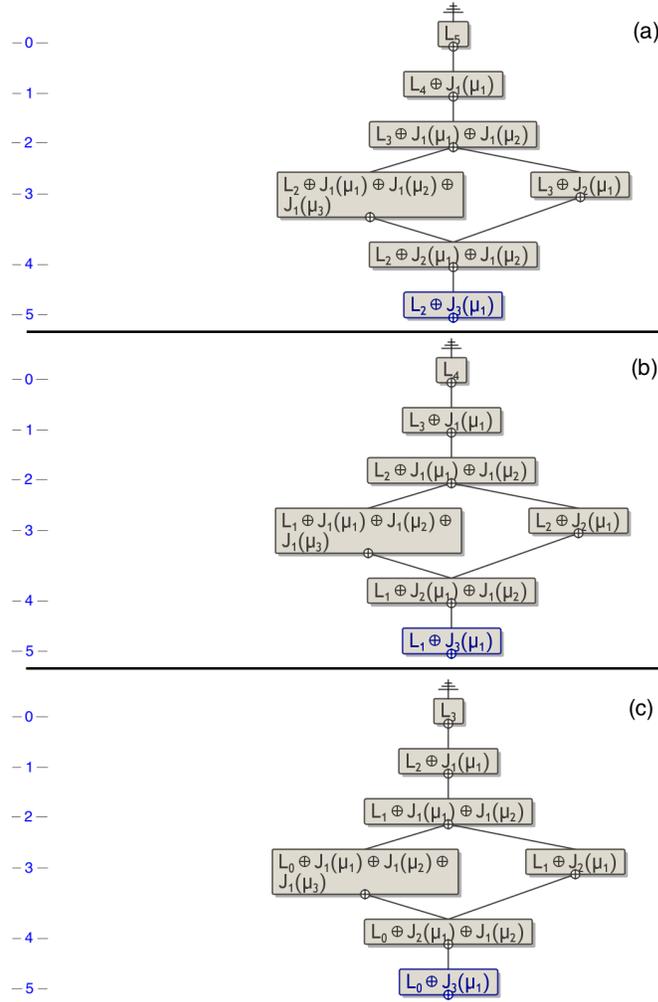


Figure 6: Bundle stratification of the Fiedler linearizations of 1×2 matrix polynomials of degree 3. Like in Figure 3, the graphs (a), (b), and (c) are the bundle stratifications of the first companion form (5×6 matrix pencils), linearizations in (17) (4×5 matrix pencils), and second companion form (3×4 matrix pencils), respectively.

References

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

- [2] V.I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys, 26 (2) (1971) 29–43.
- [3] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software, 39(2) (2013) 1–28.
- [4] J. Demmel and A. Edelman, *The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms*, Linear Algebra Appl., 230 (1995) 61–87.
- [5] 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(4) (2010) 2181–2204.
- [6] F. De Terán, F.M. Dopico, and D.S. Mackey, *Fiedler companion linearizations for rectangular matrix polynomials*, Linear Algebra Appl., 437(3) (2012) 957–991.
- [7] F. De Terán, F.M. Dopico, and P. Van Dooren, *Matrix polynomials with completely prescribed eigenstructure*, SIAM J. Matrix Anal. Appl., 36(1) (2015) 302–328.
- [8] A. Dmytryshyn, *Structure preserving stratification of skew-symmetric matrix polynomials*, Report UMINF 15.16, Dept. of Computing Science, Umeå University, Sweden, 2015.
- [9] A.R. Dmytryshyn, V. Futorny, and V.V. Sergeichuk, *Miniversal deformations of matrices of bilinear forms*, Linear Algebra Appl., 436(7) (2012) 2670–2700.
- [10] A. Dmytryshyn, V. Futorny, and V.V. Sergeichuk, *Miniversal deformations of matrices under \ast -congruence and reducing transformations*, Linear Algebra Appl., 446 (2014) 388–420.
- [11] A. Dmytryshyn, V. Futorny, B. Kågström, L. Klimenko, and V.V. Sergeichuk, *Change of the congruence canonical form of 2-by-2 and 3-by-3 matrices under perturbations and bundles of matrices under congruence*, Linear Algebra Appl., 469 (2015) 305–334.
- [12] A. Dmytryshyn, S. Johansson, and B. Kågström, *Codimension computations of congruence orbits of matrices, skew-symmetric and symmetric matrix pencils using Matlab*, Report UMINF 13.18, Dept. of Computing Science, Umeå University, Sweden, 2013.
- [13] A. Dmytryshyn, S. Johansson, and B. Kågström, *Canonical structure transitions of system pencils*, Report UMINF 15.15, Dept. of Computing Science, Umeå University, Sweden, 2015.
- [14] A. Dmytryshyn and B. Kågström, *Orbit closure hierarchies of skew-symmetric matrix pencils*, SIAM J. Matrix Anal. Appl., 35(4) (2014) 1429–1443.

- [15] A. Dmytryshyn, B. Kågström, and V.V. Sergeichuk, *Skew-symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations*, Linear Algebra Appl., 438 (2013) 3375–3396.
- [16] A. Dmytryshyn, B. Kågström, and V.V. Sergeichuk, *Symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations*, Electron. J. Linear Algebra, 27 (2014) 1–18.
- [17] A. Edelman, E. Elmroth, and B. Kågström, *A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations*, SIAM J. Matrix Anal. Appl., 18(3) (1997) 653–692.
- [18] 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(3) (1999) 667–699.
- [19] E. Elmroth, S. Johansson, and B. Kågström, *Stratification of controllability and observability pairs – theory and use in applications*, SIAM J. Matrix Anal. Appl., 31(2) (2009) 203–226.
- [20] F.R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.
- [21] M.I. Garcia-Planas and V.V. Sergeichuk, *Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils*, Linear Algebra Appl., 302–303 (1999) 45–61.
- [22] J. Gallier, *Geometric Methods and Applications*, Springer, New York, 2011.
- [23] A. Hilliges, C. Mehl, and V. Mehrmann, *On the solution of palindromic eigenvalue problems*, Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS), Jyväskylä, Finland, 2004.
- [24] I.C.F. Ipsen, *Accurate eigenvalues for fast trains*, SIAM News, 37(9) (2004) 1–2.
- [25] S. Johansson, *Reviewing the closure hierarchy of orbits and bundles of system pencils and their canonical forms*, Report UMINF 09.02, Umeå University, Department of Computing Science, 2009.
- [26] P. Johansson, *Software Tools for Matrix Canonical Computations and Web-based Software Library Environments*, PhD thesis, Department of Computing Science, Umeå University, Sweden, 2006.
- [27] S. Johansson, B. Kågström, and P. Van Dooren, *Stratification of full rank polynomial matrices*, Linear Algebra Appl., 439 (2013) 1062–1090.
- [28] L. Karlsson and F. Tisseur, *Algorithms for Hessenberg-triangular reduction of Fiedler linearization of matrix polynomials*, SIAM J. Sci. Comput., 37(3) (2015) C384–C414.

- [29] D. Kressner, C. Schröder, and D.S. Watkins, *Implicit QR algorithms for palindromic and even eigenvalue problems*, Numer. Algorithms, 51(2) 2009 209–238.
- [30] B. Kågström, S. Johansson, and P. Johansson, *StratiGraph Tool: Matrix Stratification in Control Applications*. In L. Biegler, S.L. Campbell, and V. Mehrmann, editors, Control and Optimization with Differential-Algebraic Constraints, chapter 5. SIAM Publications, 2012.
- [31] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Structured polynomial eigenvalue problems: Good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28(4) (2006) 1029–1051.
- [32] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form*, Numer. Linear Algebra Appl., 16 (2009) 63–86.
- [33] D.S. Mackey, N. Mackey, and F. Tisseur, *Polynomial Eigenvalue Problems: Theory, Computation, and Structure*, in P. Benner et. al., Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory, chapter 12, Springer, 2015.
- [34] V. Mehrmann and H. Voss, *Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27(2) (2004) 121–152.
- [35] StratiGraph and MCS Toolbox Homepage, Department of Computing Science, Umeå University, Sweden, 2015. Available from: <http://www.cs.umu.se/english/research/groups/matrix-computations/stratigraph>.
- [36] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43(2) (2001) 235–286.