Foundations of Computational Mathematics (2020) 20:423–450
https://doi.org/10.1007/s10208-019-09423-1

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Geometry of Matrix Polynomial Spaces

Andrii Dmytryshyn[1,2] · Stefan Johansson[2] · Bo Kågström[2] · Paul Van Dooren[3]

Received: 27 March 2018 / Revised: 5 February 2019 / Accepted: 29 March 2019 /
Published online: 21 June 2019
© The Author(s) 2019

Abstract
We study how small perturbations of general matrix polynomials may change their elementary divisors and minimal indices by constructing the closure hierarchy (stratification) graphs of matrix polynomials' orbits and bundles. To solve this problem, we construct the stratification graphs for the first companion Fiedler linearization of matrix polynomials. Recall that the first companion Fiedler linearization as well as all the Fiedler linearizations is matrix pencils with particular block structures. Moreover, 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). This geometry coincides with the geometry of the space of matrix polynomials. The novel results are illustrated by examples using the software tool StratiGraph extended with associated new functionality.

Keywords Matrix polynomials · Stratifications · Matrix pencils · Fiedler linearization · Canonical structure information · Orbit · Bundle

Mathematics Subject Classification 15A21 · 15A22 · 65F15 · 47A07

Communicated by Alan Edelman.

Preprint Report UMINF 15.17 (revised), Department of Computing Science, Umeå University.

✉ Andrii Dmytryshyn
andrii.dmytryshyn@oru.se; andrii@cs.umu.se

Stefan Johansson
stefanj@cs.umu.se

Bo Kågström
bokg@cs.umu.se

Paul Van Dooren
paul.vandooren@uclouvain.be

1 School of Science and Technology, Örebro University, 701 82 Örebro, Sweden

2 Department of Computing Science, Umeå University, 901 87 Umeå, Sweden

3 Department of Mathematical Engineering, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium

1 Introduction

For a long-time matrix polynomials

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

have been important objects to investigate. Due to challenging applications [27,28,37, 41,42], matrix polynomials have received much attention in the last decade, resulting in rapid developments of corresponding theories [5–7,19,32,37] and computational techniques [3,27,34,36,39] (see also the recent survey [38]). 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], i.e. matrix polynomials of degree $d = 1$ which are matrix pencils with a particular block structure. However, the canonical structure information is sensitive to perturbations in the coefficient 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 is already known for several matrix problems: matrices under similarity (i.e. Jordan canonical form) [4,21,35,40], matrix pencils (i.e. Kronecker canonical form) [21], skew-symmetric matrix pencils [16], controllability and observability pairs [22], state-space system pencils [15], as well as full (normal)-rank matrix polynomials [32]. Many of these results are already implemented in *StratiGraph* [29,31,33], which is a java-based tool developed to construct and visualize such closure hierarchy graphs. The *Matrix Canonical Structure* (MCS) *Toolbox* for MATLAB [14,29,31] 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 [33].

In this paper, we study how small perturbations of general matrix polynomials, with rectangular matrix coefficients, may change their elementary divisors and minimal indices by constructing the closure hierarchy graphs of the orbits and bundles of matrix polynomial and their Fiedler linearizations. Our new results generalize and extend results from [32], where the study concerned full-rank matrix polynomials. Other recent results that are crucial for this study 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 [6]; the correspondence between perturbations of the linearizations and perturbations of matrix polynomials [32]; as well as the algorithm for the stratification of general matrix pencils [21]. In particular, the results in [6] and [7] allow us to consider polynomials with both left and right minimal indices, in contrast to [32] (recall that full-rank matrix

polynomials may have either left or right minimal indices, not both types); as well as to use any Fiedler linearization in contrast to the fixed choice of either the first or second companion forms (depending on which type of the minimal indices is present).

The rest of the paper is organized as follows. Sections 2–5 present necessary background to matrix polynomials, their linearizations and perturbations, and to matrix pencils. Codimension computation is presented in Sect. 6. Section 7 is devoted to stratifications of Fiedler linearizations of matrix polynomials. Section 7.1 recalls cover relations for complete eigenstructures, a concept frequently used in the results that follow on neighbours in the stratifications. Sections 7.2 and 7.3 provide the results for neighbouring orbits and bundles, respectively. All results are illustrated with examples. Finally, in Sect. 8 stratification results from Sect. 7 are expressed in terms of matrix polynomial invariants. Altogether, we complete the stratification theory for general matrix polynomials and the associated Fiedler linearizations.

All matrices that we consider have complex entries.

2 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 2 (proven in [7]) that explains which canonical structure information a matrix polynomial may have.

Definition 1 Let $P(\lambda)$ and $Q(\lambda)$ be two $m \times n$ matrix polynomials. Then, $P(\lambda)$ and $Q(\lambda)$ are *unimodulary equivalent* if there exist two unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ (i.e. $\det U(\lambda), \det V(\lambda) \in \mathbb{C}\backslash\{0\}$) 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 such transformations is the *Smith form* [24], recalled in the following theorem.

Theorem 1 [24] *Let $P(\lambda)$ be an $m \times n$ matrix polynomial over $\mathbb{C}$. Then, there exists an $r \in \mathbb{N}$, $r \leqslant \min\{m, n\}$ and unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ over $\mathbb{C}$ such that*

$$U(\lambda)P(\lambda)V(\lambda) = \left[\begin{array}{ccc|c} g_1(\lambda) & & 0 & \\ & \ddots & & 0_{r\times(n-r)} \\ 0 & & g_r(\lambda) & \\ \hline & 0_{(m-r)\times r} & & 0_{(m-r)\times(n-r)} \end{array}\right], \tag{2}$$

where $g_j(\lambda)$ is monic for $j = 1, \ldots, r$ and $g_j(\lambda)$ divides $g_{j+1}(\lambda)$ for $j = 1, \ldots, r-1$. Moreover, the canonical form (2) is unique.

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

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

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

where $l_j \geqslant 0,\ \delta_{j1}, \ldots, \delta_{jl_j} > 0$ are integers. If $l_j = 0$, then $g_j(\lambda) = 1$. The numbers $\alpha_1, \ldots, \alpha_{l_j} \in \mathbb{C}$ 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 $\operatorname{rev} P(\lambda) := \lambda^d P(1/\lambda)$. The elementary divisors λ^{γ_k}, $\gamma_k > 0$ for the zero eigenvalue of $\operatorname{rev} P(\lambda)$ are the elementary divisors associated with ∞ of $P(\lambda)$.

Define the *left* and *right null-spaces*, over the field of rational functions $\mathbb{C}(\lambda)$, for an $m \times n$ matrix polynomial $P(\lambda)$ as follows, e.g. see [7]:

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

Every subspace $\mathscr{V}$ of the vector space $\mathbb{C}(\lambda)^n$ has bases consisting entirely of vector polynomials. Recall that a *minimal basis* of $\mathscr{V}$ is a basis of $\mathscr{V}$ consisting of vector polynomials whose sum of degrees is minimal among all bases of $\mathscr{V}$ consisting of vector polynomials. The ordered list of degrees of the vector polynomials in any minimal basis of $\mathscr{V}$ is always the same. These degrees are called the *minimal indices* of $\mathscr{V}$. We use the concepts above in the context of matrix polynomials as follows: let the sets $\{y_1(\lambda)^T, \ldots, y_{m-r}(\lambda)^T\}$ and $\{x_1(\lambda), \ldots, x_{n-r}(\lambda)\}$ be minimal bases of $\mathscr{N}_{\text{left}}(P)$ and $\mathscr{N}_{\text{right}}(P)$, respectively, ordered so that $0 \leqslant \deg(y_1) \leqslant \ldots \leqslant \deg(y_{m-r})$ and $0 \leqslant \deg(x_1) \leqslant \ldots \leqslant \deg(x_{n-r})$. Let $\eta_k = \deg(y_k)$ for $i = 1, \ldots, m - r$ and $\varepsilon_k = \deg(x_k)$ for $i = 1, \ldots, n - r$. Then, the scalars $0 \leqslant \eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_{m-r}$ and $0 \leqslant \varepsilon_1 \leqslant \varepsilon_2 \leqslant \ldots \leqslant \varepsilon_{n-r}$ 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 2 [7] *Let m, n, d, and r, such that $r \leqslant \min\{m, n\}$ be given positive integers. Let $g_1(\lambda), g_2(\lambda), \ldots, g_r(\lambda)$ be r arbitrarily monic polynomials with coefficients in $\mathbb{C}$ and with respective degrees $\delta_1, \delta_2, \ldots, \delta_r$, such that $g_j(\lambda)$ divides $g_{j+1}(\lambda)$ for $j = 1, \ldots, r - 1$. Let $0 \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \ldots \leqslant \gamma_r$, $0 \leqslant \varepsilon_1 \leqslant \varepsilon_2 \leqslant \ldots \leqslant \varepsilon_{n-r}$ and $0 \leqslant \eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_{m-r}$ be given lists of integers. There exists an $m \times n$ matrix polynomial $P(\lambda)$ with rank r, degree d, invariant polynomials $g_1(\lambda), g_2(\lambda), \ldots, g_r(\lambda)$, partial multiplicities at ∞ equal to $\gamma_1, \gamma_2, \ldots, \gamma_r$, and with right and left minimal indices equal to $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-r}$ and $\eta_1, \eta_2, \ldots, \eta_{m-r}$, respectively,* if and only if

$$\sum_{j=1}^{r} \delta_j + \sum_{j=1}^{r} \gamma_j + \sum_{j=1}^{n-r} \varepsilon_j + \sum_{j=1}^{m-r} \eta_j = dr \qquad (\textit{index sum identity}) \tag{3}$$

holds and $\gamma_1 = 0$.

The condition $\gamma_1 = 0$ guarantees that A_d in (1) is a nonzero $m \times n$ matrix.

3 Fiedler Linearizations of Matrix Polynomials

Let us first 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 $\sigma : \{0, 1, \ldots, d-1\} \to \{1, \ldots, d\}$ with inverse σ^{-1}, the Fiedler pencil $\mathscr{F}^{\sigma}_{G(\lambda)}$ of $G(\lambda)$ associated with σ is the $dn \times dn$ matrix pencil

$$\mathscr{F}^{\sigma}_{G(\lambda)} := \lambda M_d - M_{\sigma^{-1}(1)} M_{\sigma^{-1}(2)} \cdots M_{\sigma^{-1}(d)}, \tag{4}$$

where

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

and

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

Note that $\sigma(k)$ describes the position of the factor M_k in the product defining the zero-degree term in (4), i.e. $\sigma(k) = j$ means that M_k is the j^{th} factor in the product. All the non-specified blocks of M_k matrices are conforming size submatrices with zero entries.

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 (i.e. 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 3.

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

$$\mathscr{C}^1_{P(\lambda)} = \lambda \begin{bmatrix} A_d & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} \tag{5}$$

and

$$\mathscr{C}^2_{P(\lambda)} = \lambda \begin{bmatrix} A_d & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix} + \begin{bmatrix} A_{d-1} & -I_m & & 0 \\ A_{d-2} & 0 & \ddots & \\ \vdots & \vdots & \ddots & -I_m \\ A_0 & 0 & \dots & 0 \end{bmatrix} \tag{6}$$

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

Fiedler linearizations preserve finite and infinite elementary divisors but do not, in general, preserve the left and right minimal indices (in some cases, the minimal indices may also be preserved, e.g. for full-rank matrix polynomials [32]). In Theorem 3, proven in [6], we recall the relation between the minimal indices of polynomials and their Fiedler linearizations; see also [5] for the similar results on square matrix polynomials.

We say that a bijection $\sigma : \{0, 1, \dots, d-1\} \to \{1, \dots, d\}$ 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 $\mathrm{i}(\sigma)$ and $\mathrm{c}(\sigma)$ to be the total numbers of inversions and consecutions in σ, respectively. Note that

$$\mathrm{i}(\sigma) + \mathrm{c}(\sigma) = d - 1 \tag{7}$$

for every σ.

Theorem 3 [6] *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of degree $d \geqslant 2$, and let $\mathscr{F}^{\sigma}_{P(\lambda)}$ be its Fiedler linearization. If $0 \leqslant \varepsilon_1 \leqslant \varepsilon_2 \leqslant \ldots \leqslant \varepsilon_s$ and $0 \leqslant \eta_1 \leqslant \eta_2 \leqslant \ldots \leqslant \eta_t$ are the right and left minimal indices of $P(\lambda)$, then*

$$\begin{aligned} &0 \leqslant \varepsilon_1 + \mathrm{i}(\sigma) \leqslant \varepsilon_2 + \mathrm{i}(\sigma) \leqslant \ldots \leqslant \varepsilon_s + \mathrm{i}(\sigma), \quad \textit{and} \\ &0 \leqslant \eta_1 + \mathrm{c}(\sigma) \leqslant \eta_2 + \mathrm{c}(\sigma) \leqslant \ldots \leqslant \eta_t + \mathrm{c}(\sigma) \end{aligned}$$

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

Note also that the Fiedler linearization $\mathscr{F}^{\sigma}_{P(\lambda)}$ has $m\,\mathrm{c}(\sigma) + n\,\mathrm{i}(\sigma) + m$ rows and $m\,\mathrm{c}(\sigma) + n\,\mathrm{i}(\sigma) + n$ columns.

Remark 1 Theorem 3 can straightforwardly be applied to the first and second companion forms. For the first companion form $\mathscr{C}^1_{P(\lambda)}$, we have $\mathrm{i}(\sigma) = d - 1$ and $\mathrm{c}(\sigma) = 0$, and for the second companion form $\mathscr{C}^2_{P(\lambda)}$, we have $\mathrm{i}(\sigma) = 0$ and $\mathrm{c}(\sigma) = d - 1$.

Theorems 2 and 3 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 Perturbations of Matrix Polynomials

Recall that for every matrix $X = [x_{ij}]$, its Frobenius norm is given by $||X|| := ||X||_F = \left(\sum_{i,j} |x_{ij}|^2\right)^{\frac{1}{2}}$. 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 2 Let $P(\lambda)$ and $E(\lambda)$ be two $m \times n$ matrix polynomials, with $\deg P(\lambda) \geq \deg E(\lambda)$. A matrix polynomial $\widetilde{P(\lambda)} := P(\lambda) + E(\lambda)$ is called a *perturbation of an $m \times n$ matrix polynomial* $P(\lambda)$.

Note that, in this paper we are interested in small perturbations, i.e. $||\widetilde{P(\lambda)} - P(\lambda)||$ is small compared to $||P(\lambda)||$ (or equivalently $||E(\lambda)|| << ||P(\lambda)||$). Moreover, we say that there exists an arbitrarily small perturbation $\widetilde{P(\lambda)}$ of $P(\lambda)$ that satisfies a certain property, if for every $\varepsilon > 0$ there exists a perturbation $\widetilde{P(\lambda)}$ such that $||\widetilde{P(\lambda)} - P(\lambda)|| \leqslant \varepsilon$, and $\widetilde{P(\lambda)}$ satisfies the same property.

We remark that Definition 2 is also applicable to matrix pencils and matrices (they are polynomials of degrees one and zero, respectively).

Theorem 4 (proven in [32]) ensures that each perturbation of the linearization of an $m \times n$ matrix polynomial of degree d

$$\begin{aligned}\widetilde{\mathscr{C}^1_{P(\lambda)}} := \lambda &\begin{bmatrix} A_d & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} \\ + \lambda &\begin{bmatrix} E_{11} & E_{12} & E_{13} & \dots & E_{1d} \\ E_{21} & E_{22} & E_{23} & \dots & E_{2d} \\ E_{31} & E_{32} & E_{33} & \dots & E_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{d1} & E_{d2} & E_{d3} & \dots & E_{dd} \end{bmatrix} + \begin{bmatrix} E'_{11} & E'_{12} & E'_{13} & \dots & E'_{1d} \\ E'_{21} & E'_{22} & E'_{23} & \dots & E'_{2d} \\ E'_{31} & E'_{32} & E'_{33} & \dots & E'_{3d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E'_{d1} & E'_{d2} & E'_{d3} & \dots & E'_{dd} \end{bmatrix} \end{aligned} \tag{8}$$

can be smoothly reduced by strict equivalence to the one in which only the blocks $A_i, i = 0, 1, \dots$ are perturbed

$$\mathscr{C}^1_{\widetilde{P(\lambda)}} = \lambda \begin{bmatrix} A_d & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{d-1} & A_{d-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ & & \ddots & \vdots \\ 0 & & -I_n & 0 \end{bmatrix} + \lambda \begin{bmatrix} F_d & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} F_{d-1} & F_{d-2} & \dots & F_0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}. \tag{9}$$

We refer to (8) as a perturbation of the linearization and to (9) as the linearization of a perturbed matrix polynomial. The relation between these two types of perturbations is reflected in the following theorem, which is a slightly adapted formulation of Theorem 2.5 from [10], see also Theorem 5.21 in [19], as well as [32,43].

Theorem 4 *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of degree d, and let $\mathscr{C}^1_{P(\lambda)}$ be its first companion form. If $\widetilde{\mathscr{C}^1_{P(\lambda)}}$ is a perturbation of $\mathscr{C}^1_{P(\lambda)}$ such that*

$$||\widetilde{\mathscr{C}^1_{P(\lambda)}} - \mathscr{C}^1_{P(\lambda)}|| < \frac{\pi}{12\, d^{3/2}},$$

then $\widetilde{\mathscr{C}^1_{P(\lambda)}}$ is strictly equivalent to a pencil $\mathscr{C}^1_{\widetilde{P(\lambda)}}$, i.e. there exist two non-singular matrices X and Y (they are small perturbations of the identity matrices) such that

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

and moreover,

$$||\mathscr{C}^1_{\widetilde{P(\lambda)}} - \mathscr{C}^1_{P(\lambda)}|| \le 4\, d\, (1 + ||P(\lambda)||_F)\, ||\widetilde{\mathscr{C}^1_{P(\lambda)}} - \mathscr{C}^1_{P(\lambda)}||.$$

The following corollary to Theorem 4 shows that the canonical structure information of all pencils that are attainable by perturbations of the form (8) are also attainable by perturbations of the form (9).

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

$$U \cdot \mathscr{C}^1_{\widetilde{P(\lambda)}} \cdot V = \mathscr{C}^1_{Q(\lambda)}, \tag{10}$$

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

$$U' \cdot \widetilde{\mathscr{C}^1_{P(\lambda)}} \cdot V' = \mathscr{C}^1_{Q(\lambda)}. \tag{11}$$

Proof By Theorem 4, we have $X \cdot \widetilde{\mathscr{C}^1_{P(\lambda)}} \cdot Y = \mathscr{C}^1_{\widetilde{P(\lambda)}}$ and substituting $\widetilde{\mathscr{C}^1_{P(\lambda)}}$ in (11) we obtain $U' \cdot X^{-1} \cdot \mathscr{C}^1_{\widetilde{P(\lambda)}} \cdot Y^{-1} \cdot V' = \mathscr{C}^1_{Q(\lambda)}$ which is (10) with $U = U' \cdot X^{-1}$ and $V = Y^{-1} \cdot V'$. The "vice versa" part is obvious. □

An alternative way to derive the results of Corollary 1 is to use the theory of versal deformations [2,12,13] as it was done for state-space system pencils in [15] and skew-symmetric polynomials in [9]. See also Theorem 9, which generalizes the above results for any Fiedler linearization.

5 Matrix Pencils

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

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

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

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

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

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 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:

$$\mathrm{O}^e_{A-\lambda B} = \left\{ Q^{-1}(A - \lambda B)R \ : \ Q \in GL_m(\mathbb{C}),\, R \in GL_n(\mathbb{C}) \right\}. \tag{12}$$

The *dimension* of $\mathrm{O}^e_{A-\lambda B}$ is the dimension of its tangent space

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

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

$$\langle A - \lambda B, C - \lambda D \rangle = \mathrm{trace}(AC^* + BD^*), \tag{13}$$

is called the normal space to this orbit. The dimension of the normal space is the *codimension* of $\mathrm{O}^e_{A-\lambda B}$, denoted $\operatorname{cod} \mathrm{O}^e_{A-\lambda B}$ ($\operatorname{cod} \mathrm{O}^e_{A-\lambda B} = 2mn - \dim \mathrm{O}^e_{A-\lambda B}$). Explicit expressions for the codimensions of strict equivalence orbits are presented in [8].

Theorem 5 [24, 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 $j \geqslant 1$ and $k \geqslant 0$. The j's and k's may be different in each block.

The canonical form defined by the E_j, L_k and L_k^T blocks in Theorem 5 is known as the *Kronecker canonical form* (KCF) of the pencil $A - \lambda B$. 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* $\mathrm{B}^e_{A-\lambda B}$ of a matrix pencil $A - \lambda B$ is a union of orbits $\mathrm{O}^e_{A-\lambda B}$ with the same singular structures and the same regular structures, except that the distinct eigenvalues may be different.

6 Orbits of Linearizations of Matrix Polynomials and Their Codimensions

Let $P(\lambda)$ be an $m \times n$ matrix polynomial of degree d and $C^1_{P(\lambda)}$ be its $(m+n(d-1)) \times nd$ first companion form. The *generalized Sylvester space* at $P(\lambda)$ is defined as (see [32] and references therein)

$$\mathrm{GSYL}^1_{m\times n} = \{\mathscr{C}^1_{P(\lambda)} \ : \ P(\lambda) \text{ are } m \times n \text{ matrix polynomials}\}, \tag{14}$$

where $\mathrm{GSYL}^1_{m\times n}$ is a $(d+1)mn$-dimensional affine subspace in the $(2d^2n^2+2dn(m-n))$-dimensional pencil space; each fixed element in the linearization decreases the degree of freedom by one. If there is no risk of confusion, we write GSYL instead of $\mathrm{GSYL}^1_{m\times n}$. We define the orbit of linearizations of matrix polynomials as

$$\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}} = \left\{ (Q^{-1}\mathscr{C}^1_{P(\lambda)}R) \in \mathrm{GSYL}^1_{m\times n} \ : \ Q \in GL_{m+n(d-1)}(\mathbb{C}), \ R \in GL_{nd}(\mathbb{C}) \right\}. \tag{15}$$

Note that all the elements of $\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}$ have the block structure of $\mathscr{C}^1_{P(\lambda)}$, see (5). By [32, Lemma 9.2], $\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}$ is a manifold in the matrix pencil space.

Codimensions of this manifold are also of our interest, since they define the level of the orbit in the stratification graph: an orbit has only orbits with higher codimensions

in its closure. Recall that $\dim \mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}} := \dim \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}}$ and $\operatorname{cod} \mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}} := \dim \mathrm{N}^e_{\mathscr{C}^1_{P(\lambda)}}$, where N denotes the normal space (see Sect. 5). Define

$$\dim \mathrm{O}_{\mathscr{C}^1_{P(\lambda)}} := \dim(\mathrm{GSYL} \cap \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}}). \tag{16}$$

The following lemma shows that the codimensions of $\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}$ and $\mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}}$ are equal; the latter is computed in [8] (see also [20,25]) and implemented in the MCS Toolbox [31]. We also refer to [32, Section 9] for a slightly different explanation of the analogous results.

Lemma 1 *Let $\mathscr{C}^1_{P(\lambda)}$ be the first companion form for the matrix polynomial $P(\lambda)$, then* $\operatorname{cod} \mathrm{O}_{\mathscr{C}^1_{P(\lambda)}} = \operatorname{cod} \mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}}$.

Proof A general matrix pencil of the same size as $\mathscr{C}^1_{P(\lambda)}$ belongs to the pencil space $\mathscr{P} := \mathbb{C}^{(m+n(d-1))\times nd} \times \mathbb{C}^{(m+n(d-1))\times nd}$. Also, recall that GSYL in (14) is the subspace of all first companion forms of $m \times n$ matrix polynomials. Following the arguments in [32,43], GSYL is an affine subspace in $\mathscr{P}$ that together with the tangent space $\mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}}$ spans the complete $\mathscr{P}$ [32, proof of Lemma 9.2], and since $\mathrm{GSYL} \cap \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}} \neq \emptyset$

$$\dim(\mathscr{P}) = \dim \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}} + \dim \mathrm{GSYL} - \dim(\mathrm{GSYL} \cap \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}}), \tag{17}$$

see also [23, Section 2] for details. Knowing the dimensions of the tangent and the normal spaces and using (16) and (17), we finally get

$$\begin{aligned}
\operatorname{cod} \mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}} &= \dim(\mathscr{P}) - \dim \mathrm{O}^e_{\mathscr{C}^1_{P(\lambda)}} \\
&= \dim \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}} + \dim \mathrm{GSYL} - \dim(\mathrm{GSYL} \cap \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}}) - \dim \mathrm{T}^e_{\mathscr{C}^1_{P(\lambda)}} \\
&= \dim \mathrm{GSYL} - \dim \mathrm{O}_{\mathscr{C}^1_{P(\lambda)}} = \operatorname{cod} \mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}.
\end{aligned}$$

□

We remark that there are other examples where codimension equalities similar to the one in Lemma 1 do hold [22,32] as well as examples where they are not valid [15,17,18].

7 Stratifications of Matrix Polynomial Linearizations

In this section, we start by presenting an algorithm for computing the stratification of the Fiedler linearizations of general $m \times n$ matrix polynomials (1). The algorithm relies on the results presented in Sects. 4–5. Section 7.1 introduces cover relations for complete eigenstructures. Based on these concepts, Sect. 7.2 presents the results for orbit stratifications. Similar results for bundle stratifications are presented in Sect. 7.3.

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 relation, i.e. there is an upward path from a node associated with $\mathscr{F}^{\sigma}_{P(\lambda)}$ to a node associated with $\mathscr{F}^{\sigma}_{Q(\lambda)}$ if and only if $P(\lambda)$ can be transformed by an arbitrarily small perturbation to a matrix polynomial whose canonical structure information coincides 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 6 *Steps 1–3 produce the orbit stratification of the first companion linearizations of $m \times n$ matrix polynomials of degree d.*

Step 1. *Construct the stratification of $(m+n(d-1)) \times nd$ matrix pencil orbits under strict equivalence* [21].

Step 2. *Extract from the stratification obtained in* Step 1 *the orbits* (*nodes*) *that correspond to the first companion linearizations of $m \times n$ matrix polynomials of degree d* (*using Theorems 2 and 3, as well as Remark 1*).

Step 3. *Put an edge between two nodes obtained in* Step 2 *if there is an upward path between these nodes in the graph obtained in* Step 1 *and do not put an edge between these nodes otherwise* (*justified by Theorem 4 and Corollary 1*).

Theorem 7 *The stratification graphs for a matrix polynomial $P(\lambda)$ and any of its Fiedler linearizations $\mathscr{F}^{\sigma}_{P(\lambda)}$ are the same, up to the fact that the nodes in the graph for the Fiedler linearization represent complete eigenstructures with the minimal indices "shifted", see Theorem 3.*

Proof We take the stratification graph for $\mathscr{C}^{1}_{P(\lambda)}$ as a starting point since we know how to construct it using Algorithm 6. Let also $P_1(\lambda)$ and $P_2(\lambda)$ be matrix polynomials belonging to two different orbits in this stratification graph. If there is an arrow from $\mathscr{C}^{1}_{P_1(\lambda)}$ to $\mathscr{C}^{1}_{P_2(\lambda)}$ in the stratification of the first companion forms, then $P_1(\lambda) + E(\lambda)$, for some small perturbation $E(\lambda)$, and $P_2(\lambda)$ have the same canonical structure information. Therefore, there is an arrow from $P_1(\lambda)$ and $P_2(\lambda)$ in the stratification of matrix polynomials. Moreover, for every σ the pencils $\mathscr{F}^{\sigma}_{P_1(\lambda)+E(\lambda)}$ and $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ have the same canonical structure information, and thus, there is an arrow from $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ to $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ in the stratifications of all the Fiedler linearizations of $P_1(\lambda)$ and $P_2(\lambda)$. □

Remark 2 Theorem 7 does not contradict the fact that for a particular matrix polynomial, some linerizations may be better conditioned, more favourable with respect to backward errors, and/or structure preserving, and therefore, the choice of linearization is typically application driven.

7.1 Cover Relation for Complete Eigenstructures

A sequence of integers $\mathscr{N} = (n_1, n_2, n_3, \dots)$ such that $n_1 + n_2 + n_3 + \cdots = n$ and $n_1 \geqslant n_2 \geqslant \dots \geqslant 0$ is called an integer partition of n (for more details and references see [21]). For any $a \in \mathbb{Z}$, we define $\mathscr{N} + a$ as the integer partition

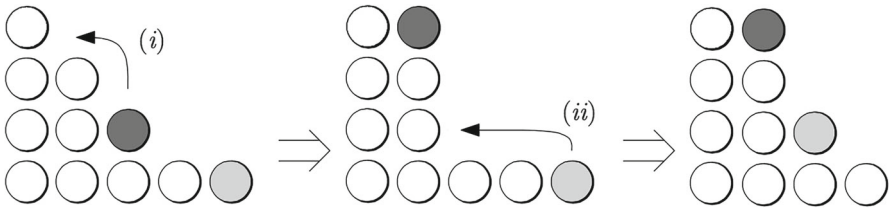

Fig. 1 To the partition (4, 3, 2, 1, 1), on the left, we apply two minimal leftward coin moves: first (i) is a move of a dark grey coin one column leftward, and then, (ii) is a move of a light grey coin one row upward. Note that monotonicity must be preserved. The resulting partition is (4, 4, 2, 1),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

$(n_1+a, n_2+a, n_3+a, \dots)$. The *additive union* of two integer partitions $\mathscr{N}$ and $\mathscr{M}$ is defined as $\mathscr{K} = \mathscr{N} \biguplus \mathscr{M}$ where all the elements from $\mathscr{N}$ and $\mathscr{M}$ are ordered such that $\mathscr{K}$ is monotonically non-increasing (i.e. $\mathscr{K}$ is a multiset sum of $\mathscr{N}$ and $\mathscr{M}$, see, e.g. [26, Chap. 1.2.4], ordered non-increasingly). For example, if $\mathscr{N} = (3, 3, 1)$ and $\mathscr{M} = (7, 3, 2, 2)$, then $\mathscr{K} = \mathscr{N} \biguplus \mathscr{M} = (7, 3, 3, 3, 2, 2, 1)$. We write $\mathscr{N} \succcurlyeq \mathscr{M}$ if and only if $n_1+n_2+\dots+n_i \geqslant m_1+m_2+\dots+m_i$, for $i \geqslant 1$. 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 $\mathscr{R}(W)$, $\mathscr{L}(W)$, and $\{\mathscr{J}_{\mu_i}(W) : j = 1, \dots, q, \mu_i \in \mathbb{C} \cup \{\infty\}\}$, where q is the number of distinct eigenvalues of W (e.g. see [21]). Altogether, these partitions, known as the Weyr characteristics, are constructed as follows:

- For each distinct μ_i, we have $\mathscr{J}_{\mu_i}(W) = (j_1^{\mu_i}, j_2^{\mu_i}, \dots)$, 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).
- $\mathscr{R}(W) = (r_0, r_1, \dots)$, where r_k is the number of L (right singular, see Theorem 5) blocks with the indices ε_i greater than or equal to k (the position numeration starting from 0).
- $\mathscr{L}(W) = (l_0, l_1, \dots)$, where l_k is the number of L^T (left singular, see Theorem 5) blocks with the indices η_i greater than or equal to k (the position numeration starting from 0).

Example 1 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:

$$\mathscr{J}_{\mu_1}(W) = (3, 2, 2), \qquad \mathscr{J}_{\infty}(W) = (2, 2),$$
$$\mathscr{R}(W) = (2, 2, 1, 1, 1), \qquad \mathscr{L}(W) = (1, 1).$$

An integer partition $\mathscr{N} = (n_1, n_2, n_3, \dots)$ 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 $\mathscr{N}$, and keeping $\mathscr{N}$ 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 $\mathscr{N}$, and keeping $\mathscr{N}$ monotonically non-increasing, is called a *minimum leftward coin move*. These two types of coin moves are defined in [21], see also Fig. 1.

By $\overline{\mathscr{X}}$ we denote the *closure* of a set $\mathscr{X}$ in the Euclidean topology. For a matrix polynomial $P(\lambda)$, define $\mathrm{O}_{\mathscr{F}^{\sigma}_{P(\lambda)}}$ to be a set of matrix pencils strictly equivalent to $\mathscr{F}^{\sigma}_{P(\lambda)}$ and with the same block structure as $\mathscr{F}^{\sigma}_{P(\lambda)}$ (this definition is analogous to the definition of $\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}$ for the first companion linearization $\mathscr{C}^1_{P(\lambda)}$). We say that the orbit $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$ *is covered by* $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$ if and only if $\overline{\mathrm{O}}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}} \supset \mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$ and there exists no orbit $\mathrm{O}_{\mathscr{F}^{\sigma}_{Q(\lambda)}}$ such that $\overline{\mathrm{O}}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}} \supset \mathrm{O}_{\mathscr{F}^{\sigma}_{Q(\lambda)}}$ and $\overline{\mathrm{O}}_{\mathscr{F}^{\sigma}_{Q(\lambda)}} \supset \mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$; or equivalently, if and only if there is an edge from $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$ to $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$ in the orbit stratification ($\mathrm{O}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$ is higher up in the graph).

7.2 Neighbouring Orbits in the Stratification

By 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.

In Theorem 8, the rules are formulated for the first companion form $\mathscr{C}^1_{P(\lambda)}$, where $\mathrm{O}_{\mathscr{C}^1_{P(\lambda)}}$ is defined as in (15). Moreover, in Corollary 2 we show that these rules are actually the same for any Fiedler linearization $\mathscr{F}^{\sigma}_{P(\lambda)}$. See also Sect. 8 where the stratification rules for matrix polynomial invariants are presented.

Theorem 8 (Orbit upward rules—matrix polynomial linearizations) *Let $P_1(\lambda)$ and $P_2(\lambda)$ be two $m \times n$ matrix polynomials of degree d with the corresponding Fiedler linearizations $\mathscr{C}^1_{P_1(\lambda)}$ and $\mathscr{C}^1_{P_2(\lambda)}$, respectively.*

The orbit $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$ is covered by $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ if and only if *the canonical structure information of $\mathscr{C}^1_{P_2(\lambda)}$ can be obtained by applying one of the rules below to the structure integer partitions representing the canonical structure information of $\mathscr{C}^1_{P_1(\lambda)}$, (here $\mu_i \in \mathbb{C} \cup \{\infty\}$):*

(a) *Minimum leftward coin move in $\mathscr{R}$ (or $\mathscr{L}$).*[1]
(b) *If $\mathscr{R}$ (or $\mathscr{L}$) is non-empty and the rightmost column in any $\mathscr{J}_{\mu_i}$ is one single coin, move that coin to a new rightmost column of $\mathscr{R}$ (or $\mathscr{L}$).*
(c) *Minimum rightward coin move in any $\mathscr{J}_{\mu_i}$.*
(d) *If both $\mathscr{R}$ and $\mathscr{L}$ are non-empty, Let k denote the total number of coins in the longest (= lowest) rows from both $\mathscr{R}$ and $\mathscr{L}$ 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 $\mathscr{J}_{\mu_i}$. The remaining coins are distributed among new rightmost columns, with one coin per column to any $\mathscr{J}_{\mu_i}$ which may be empty initially (i.e. new partitions for new eigenvalues can be created).*[2,3]

Proof We first show that applying any of the rules (a)–(d) to the structure integer partitions of $\mathscr{C}^1_{P_1(\lambda)}$ for an $m \times n$ matrix polynomial $P_1(\lambda)$ of degree d, there exits

[1] The rule is not allowed to do coin moves that affect r_0 or l_0 (first column in $\mathscr{R}$ or $\mathscr{L}$, respectively).

[2] If $\mu_i = \infty$ for some i, then $j_1^{\mu_i}$ (first column in $\mathscr{J}_{\mu_i}$) has to remain strictly less than the rank of the corresponding matrix polynomial (this restrict the matrix polynomials to those with a nonzero leading coefficient matrix).

[3] Cannot be applied if the total number of nonzero columns of all $\mathscr{J}_{\mu_i}$ is greater than $k-1$.

an $m \times n$ matrix polynomial $P_2(\lambda)$ of degree d such that $\mathscr{C}^1_{P_2(\lambda)}$ has the obtained new partitions. We prove the existence of such polynomial $P_{2(\lambda)}$ by checking that the associated invariants satisfy the index sum identity (3) in Theorem 2. Below, this is shown to hold for each of the rules (a)–(d). Then, we show that if $\mathscr{C}^1_{P_2(\lambda)}$ covers $\mathscr{C}^1_{P_1(\lambda)}$ the partitions of $\mathscr{C}^1_{P_2(\lambda)}$ are obtained from $\mathscr{C}^1_{P_1(\lambda)}$ by one and only one of the rules (a)–(d).

Applying rule (a) either effects the partition $\mathscr{R}$ or $\mathscr{L}$ and does not change the sum of the invariants $\sum \varepsilon_j$ or $\sum \eta_j$, respectively, in (3). Thus, the index sum identity holds for rule (a). Applying rule (b) moves one coin from $\mathscr{J}$ to $\mathscr{R}$ or $\mathscr{L}$, i.e. the rule simultaneously subtracts 1 from either $\sum \delta_j$ or $\sum \gamma_j$ and adds 1 to either $\sum \varepsilon_j$ or $\sum \eta_j$. Thus, the index sum identity holds for rule (b). Proof for rule (c) is analogous to the proof of rule (a). Applying rule (d) removes $\varepsilon+d$ $(= \varepsilon+1+\mathrm{i}(\sigma))$ coins from $\mathscr{R}$ and $\eta+1$ $(= \eta+1+\mathrm{c}(\sigma))$ coins from $\mathscr{L}$ (where $\mathrm{i}(\sigma) = d-1$ and $\mathrm{c}(\sigma) = 0$ are the number of inversions and consecutions, respectively, see Sect. 3 and Remark 1; we also add 1 since the numbering in $\mathscr{R}$ and $\mathscr{L}$ starts from 0). From Theorem 3, this corresponds to the fact that the sum $\sum \varepsilon_j$ in (3) is decreased by ε and $\sum \eta_j$ by η. Furthermore, the rule adds $k-1$ coins to one or several $\mathscr{J}_{\mu_i}$, where now $k = \varepsilon + d + \eta + 1$, which corresponds to that the degrees δ of the new invariant polynomials $g_{r+1}(\lambda)$ in Theorem 2 is $\delta = \varepsilon + d + \eta + 1 - 1 = \varepsilon + \eta + d$, where r is the rank of $P_1(\lambda)$. After applying rule (d) and since the identity (3) holds for $P_1(\lambda)$, the right hand side of the identity (3) loses $\varepsilon + \eta$ but gains $\delta = \varepsilon + \eta + d$; and r increases by 1, and hence, the left hand side changes from rd to $(r+1)d$. Thus, the index sum identity holds for rule (d). Moreover, to ensure that the leading coefficient matrix A_d in (1) is nonzero the condition $j_1^\infty < r$ is added (footnote 2 of rule (d)), where r is the rank of the corresponding matrix polynomial. Summing up, the partitions obtained by applying any of rules (a)–(d) correspond to some $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ that covers $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$.

Now assume that $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ covers $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$ in the stratification of the companion linearizations. By Theorem 4 and Corollary 1 (see also Algorithm 6, Step 3), there is a path from $\mathrm{O}^e_{\mathscr{C}^1_{P_1(\lambda)}}$ to $\mathrm{O}^e_{\mathscr{C}^1_{P_2(\lambda)}}$ in the stratification of equivalence orbits of general matrix pencils of size $(m + n(d-1)) \times nd$. Therefore, the partitions of $\mathscr{C}^1_{P_2(\lambda)}$ are obtained from the partitions of $\mathscr{C}^1_{P_1(\lambda)}$ by a sequence of the rules for general matrix pencils [21,30], which indeed are similar to the rules (a)–(d) (see also Remark 3). If the sequence consists of more than one rule, then we have a contradiction with $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ covering $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$; therefore, $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ must be obtained by *one* of the rules (a)–(d). □

Remark 3 The rules for obtaining the neighbouring orbit above in the stratification graph of a first companion form linearization orbit (and any Fiedler linearization orbit, which is shown in Corollary 2) of a matrix polynomial coincide with the stratification rules for general matrix pencil orbits [30, Table 3(B)] and [21, Theorem 3.2], with the added restriction that the leading coefficient matrix A_d of the matrix polynomial remains nonzero.

Corollary 2 $\mathrm{O}_{\mathscr{F}^\sigma_{P_1(\lambda)}}$ is covered by $\mathrm{O}_{\mathscr{F}^\sigma_{P_2(\lambda)}}$ if and only if *the canonical structure information of* $\mathscr{F}^\sigma_{P_2(\lambda)}$ *can be obtained by applying one of the rules (a)–(d) of Theorem 8*

to the structure integer partitions representing the canonical structure information of $\mathscr{F}^{\sigma}_{P_1(\lambda)}$.

Proof By Theorem 7, there is an arrow from $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$ to $\mathrm{O}_{\mathscr{C}^1_{P_2(\lambda)}}$ if and only if there is an arrow from $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$ to $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$. Now we show that $\mathscr{C}^1_{P_1(\lambda)}$ is obtained from $\mathscr{C}^1_{P_2(\lambda)}$ by rule (x) of Theorem 8 (where $x \in \{a, b, c, d\}$) if and only if $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ is obtained from $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ by applying exactly the same rule (x).

The linearization $\mathscr{C}^1_{P_1(\lambda)}$ is obtained from $\mathscr{C}^1_{P_2(\lambda)}$ by applying rule (a) if and only if the canonical structure information of $\mathscr{C}^1_{P_2(\lambda)}$ and $\mathscr{C}^1_{P_1(\lambda)}$ differs only in two right minimal indices: $\varepsilon_1 + d - 1$ and $\varepsilon_2 + d - 1$ in $\mathscr{C}^1_{P_2(\lambda)}$ versus $\varepsilon_1 + d$ and $\varepsilon_2 + d - 2$ in $\mathscr{C}^1_{P_1(\lambda)}$. Thus, $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ and $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ differ only in two right minimal indices too: $\varepsilon_1 + \mathrm{i}(\sigma)$ and $\varepsilon_2 + \mathrm{i}(\sigma)$ in $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ versus $\varepsilon_1 + \mathrm{i}(\sigma) + 1$ and $\varepsilon_2 + \mathrm{i}(\sigma) - 1$ in $\mathscr{F}^{\sigma}_{P_1(\lambda)}$. The latter is equivalent to the fact that the linearization $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ is obtained from $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ by applying rule (a). The same explanation works for rule (a) applied to the left minimal indices.

Note that all the Fiedler linearizations of the same matrix polynomial (including the first companion form) have the same number of right (left) minimal indices (thus the first column of $\mathscr{R}$ (and $\mathscr{L}$) has the same number of coins for any Fiedler linearization) as well as that the integer partitions for the regular parts are exactly the same for all the Fiedler linearizations. Therefore, we can apply (b) to $\mathscr{C}^1_{P_2(\lambda)}$ if and only if we can apply (b) to $\mathscr{F}^{\sigma}_{P_2(\lambda)}$. Moreover, the change in the complete eigenstructure of $\mathscr{C}^1_{P_2(\lambda)}$ is done by applying rule (b) if and only if the change in the structure of any other Fiedler linearization is done by applying rule (b).

The case of rule (c) follows from the fact that the integer partitions for the regular parts are exactly the same for all the Fiedler linearizations.

Applying rule (d) means that the largest right and left minimal indices of $\mathscr{C}^1_{P_2(\lambda)}$ ($\varepsilon_1 + d - 1$ and η_1) are changed to a regular block of size $\varepsilon_1 + \eta_1 + d$. The corresponding largest indices in a Fiedler linearization $\mathscr{F}^{\sigma}_{P_2(\lambda)}$ are $\varepsilon_1 + 1 + \mathrm{i}(\sigma)$ and $\eta_1 + 1 + \mathrm{c}(\sigma)$. Since $(\varepsilon_1 + 1 + \mathrm{i}(\sigma)) + (\eta_1 + 1 + \mathrm{c}(\sigma)) - 1 = \varepsilon_1 + \eta_1 + (\mathrm{i}(\sigma) + \mathrm{c}(\sigma) + 1) = \varepsilon_1 + \eta_1 + d$, the regular block created by rule (d) is of size $\varepsilon_1 + \eta_1 + d$ in the case of any Fiedler linearization. □

Theorem 8 and Corollary 2 provide the rules to obtain neighbouring pencils in the stratification graphs of $\mathrm{O}_{\mathscr{C}^1_{P_1(\lambda)}}$ and $\mathrm{O}_{\mathscr{F}^{\sigma}_{P(\lambda)}}$, respectively, under block-structure preserving perturbations of these linearizations. The following theorem generalizes Theorem 4 by relating block-structure preserving perturbations and full perturbations of matrix pencils for any Fiedler linearization, see also [19, Theorem 6.23].

Theorem 9 *Let $P(\lambda)$ be an $m \times n$ matrix polynomial. If there exists a matrix pencil R such that $\widetilde{\mathscr{F}^{\sigma}_{P(\lambda)}}$ is strictly equivalent to R, for some arbitrarily small perturbation of $\mathscr{F}^{\sigma}_{P(\lambda)}$, then*

1) *There exists an $m \times n$ matrix polynomial $Q(\lambda)$ such that R is strictly equivalent to $\mathscr{F}^{\sigma}_{Q(\lambda)}$;*
2) *There exists an arbitrarily small perturbation $\widetilde{P(\lambda)}$ of $P(\lambda)$ such that $\mathscr{F}^{\sigma}_{\widetilde{P(\lambda)}}$ is strictly equivalent to $\mathscr{F}^{\sigma}_{Q(\lambda)}$.*

Proof First note that the case when small perturbations do not change the eigenstructure of $\mathscr{F}^{\sigma}_{P(\lambda)}$, i.e. $X \cdot \widetilde{\mathscr{F}^{\sigma}_{P(\lambda)}} \cdot Y = \mathscr{F}^{\sigma}_{P(\lambda)}$, is obvious. For small perturbations that change the complete eigenstructure of $\mathscr{F}^{\sigma}_{P(\lambda)}$, the canonical form of $\widetilde{\mathscr{F}^{\sigma}_{P(\lambda)}}$ is one of the canonical forms in the stratification graph of $(m\,\mathrm{c}(\sigma)+n\,\mathrm{i}(\sigma)+m)\times(m\,\mathrm{c}(\sigma)+n\,\mathrm{i}(\sigma)+n)$ matrix pencils to which there is an upward path from $\mathscr{F}^{\sigma}_{P(\lambda)}$. By [21, Theorem 3.2], the canonical form of $\widetilde{\mathscr{F}^{\sigma}_{P(\lambda)}}$ can be obtained from the canonical form of $\mathscr{F}^{\sigma}_{P(\lambda)}$ by applying a sequence of rules (1)–(4) of [21, Theorem 3.2]. Since rules (1)–(4) of [21, Theorem 3.2] coincide with rules (a)–(d) of Corollary 2 (i.e. they make exactly the same changes in the complete eigenstructure), by Corollary 2 there exists $\mathscr{F}^{\sigma}_{\widetilde{P(\lambda)}}$, such that $\mathscr{F}^{\sigma}_{\widetilde{P(\lambda)}}$ has the same complete eigenstructure as $\widetilde{\mathscr{F}^{\sigma}_{P(\lambda)}}$. □

Remark 4 Theorem 9 justifies that an algorithm similar to Algorithm 6 can be used to construct a stratification of any Fiedler linearization.

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

$$A_3\lambda^3 + A_2\lambda^2 + A_1\lambda + A_0, \qquad A_3 \neq 0. \tag{18}$$

By Theorem 2 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, A_2, A_1, 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 4 and Corollary 1, 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 pencil linearizations are different for different linearizations since Fiedler linearizations preserve elementary divisors but, by Theorem 3, "shift" the minimal indices. In this case, the following shifts are possible: for the first companion form (5), we have $+2$ for the right and no shift for the left minimal indices; for the second companion form (6), we have no shift for the right and $+2$ for the left minimal indices; for the Fiedler linearizations

$$\lambda\begin{bmatrix} A_3 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} A_2 & A_1 & -I \\ -I & 0 & 0 \\ 0 & A_0 & 0 \end{bmatrix} \quad \text{and} \quad \lambda\begin{bmatrix} A_3 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \begin{bmatrix} A_2 & -I & 0 \\ A_1 & 0 & A_0 \\ -I & 0 & 0 \end{bmatrix}, \tag{19}$$

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 Fig. 2 and Theorem 7, otherwise it would mean that different linearizations "behave" generally different under small perturbations, but see also Remark 2.

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)

	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	0	–	–	–	–

	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	–	–	–	–	–	–	–	–	–	–	–	–

Note that δ_j is just the degree of $g_j(\lambda)$ and it gives a few alternatives for the powers δ_{jk} of the elementary divisors. To be exact, the number of these alternatives is the number of ways the integer δ_j can be written as a sum of positive integers, i.e. $\delta_j = \delta_{j1} + \delta_{j2} + \cdots + \delta_{jl_j}$. Thus, some columns in Table 1 correspond to more than one node in the graph in Fig. 2. Since the considered matrix polynomials may have rank at most 2 and $A_3 \neq 0$, by [7, Lemma 2.6] these polynomials may have at most 1 infinite elementary divisor. Therefore, the eigenvalues in the nodes of Fig. 2 which have two Jordan blocks associated with them can not be infinite.

Example 3 Consider rectangular 1×2 matrix polynomials of degree 3. Like in Example 2, we explain how small perturbations of the coefficient matrices of the polynomials may change their canonical structure information. By Theorem 2, 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 that $A_3 \neq 0$. Thus, by [7, Lemma 2.6] we have no infinite elementary divisors in this case.

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 (19) are 4×5. These Fiedler type linearizations "shift" the minimal indices exactly as in Example 2.

The three graphs in Fig. 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 Fig. 3a, L_4 for Fig. 3b, and L_3 for Fig. 3c. Note that each of these graphs is a subgraph of the corresponding general matrix pencil stratification

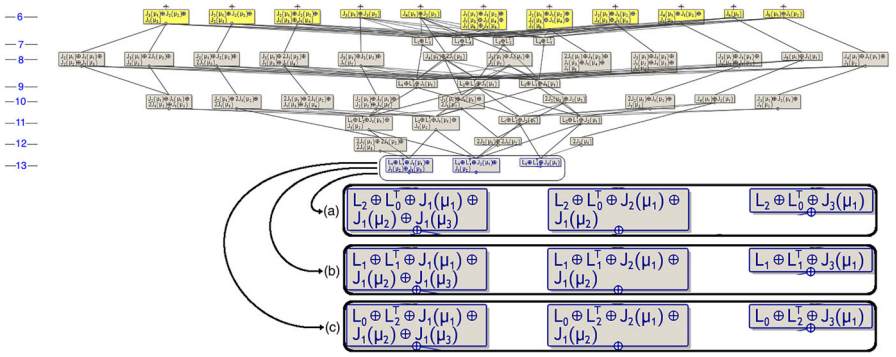

Fig. 2 Orbit stratification of the linearizations of 2×2 matrix polynomials of degree 3 ($A_3 \neq 0$). Only the sizes of the singular canonical blocks depend on the choice of Fiedler linearization, not the numbers of singular blocks, the regular parts, or the closure relations (graph edges). The numbers 6–13, listed on the left, are the codimensions of the orbits in the corresponding level of the graph. The codimensions are computed by Lemma 1. 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 (19), and the second companion form, respectively

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

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

graph; for example, the graph in Fig. 3c is a subgraph of the stratification graph of 3×4 matrix pencils, see Fig. 4.

Note also that the polynomials in this example have full rank. Thus, we can apply the theory from [32] to construct graph (c) in Fig. 3 (but not (a) or (b) since in [32] the choice of the linearization is fixed).

7.3 Neighbouring Bundles in the Stratification

In the orbit stratifications, eigenvalues may appear and disappear but their values cannot change. However, in many applications, see for example [22,32,33], the eigenvalues of the underlying matrices may coalesce or split apart to different eigenvalues, which motivates so-called bundle stratifications. Theories for bundle stratifications are developed along with theories for the orbit stratifications and are known for a number of cases [15,16,20–22,32]. 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, e.g. see [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

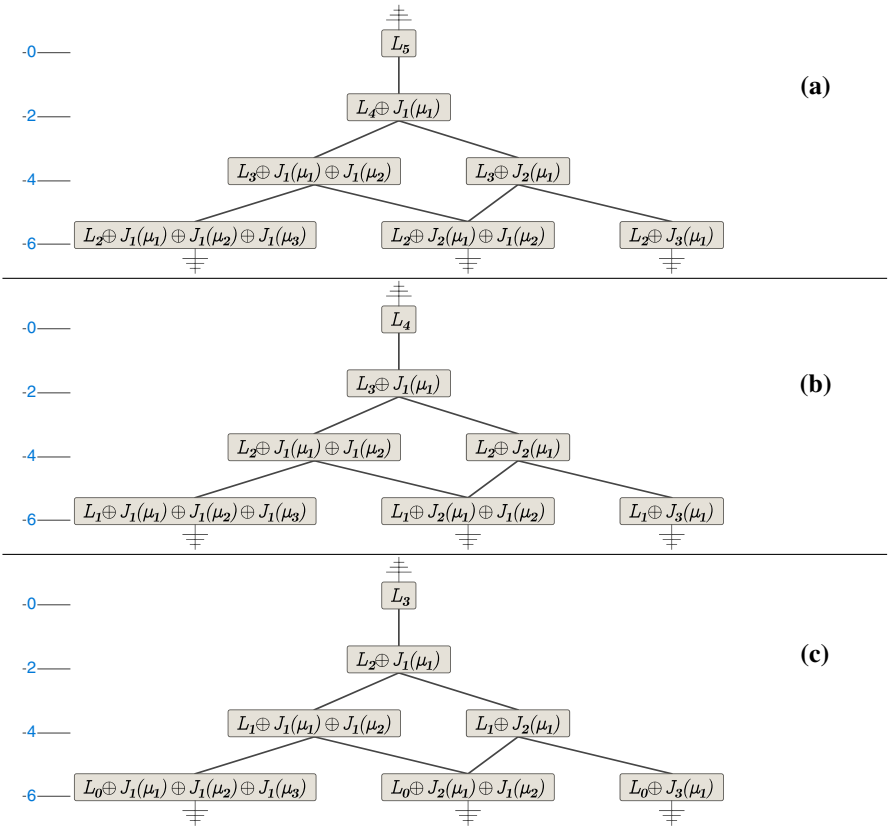

Fig. 3 Orbit stratification of the Fiedler linearizations of 1×2 matrix polynomials of degree 3 ($A_3 \neq 0$). The numbers 0, 2, 4 and 6, listed on the left, are the codimensions of the orbits in the corresponding level of the graph. These codimensions are computed by Lemma 1. Graph **a** is the stratification of the first companion form; the nodes represent 5×6 matrix pencils. Graph **b** is the stratification of the linearizations in (19); the nodes represent 4×5 matrix pencils. Finally, graph **c** is the stratification of the second companion form; the nodes represent 3×4 matrix pencils

in Theorem 8, is coming from our desire to have nonzero 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* $\mathrm{B}_{\mathscr{F}^{\sigma}_{P(\lambda)}}$ *of the matrix polynomial linearization* $\mathscr{F}^{\sigma}_{P(\lambda)}$ to be a union of orbits $\mathrm{O}_{\mathscr{F}^{\sigma}_{P(\lambda)}}$ with the same singular structures and the same regular structures, except that the distinct eigenvalues may be different, see also [32]. Therefore, we have that two Fiedler linearizations $\mathscr{F}^{\sigma}_{P(\lambda)}$ and $\mathscr{F}^{\sigma}_{R(\lambda)}$ 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 bundles of matrix polynomial Fiedler linearizations is analogous to Algorithm 6. 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 $\mathscr{F}^{\sigma}_{P(\lambda)}$ are defined as

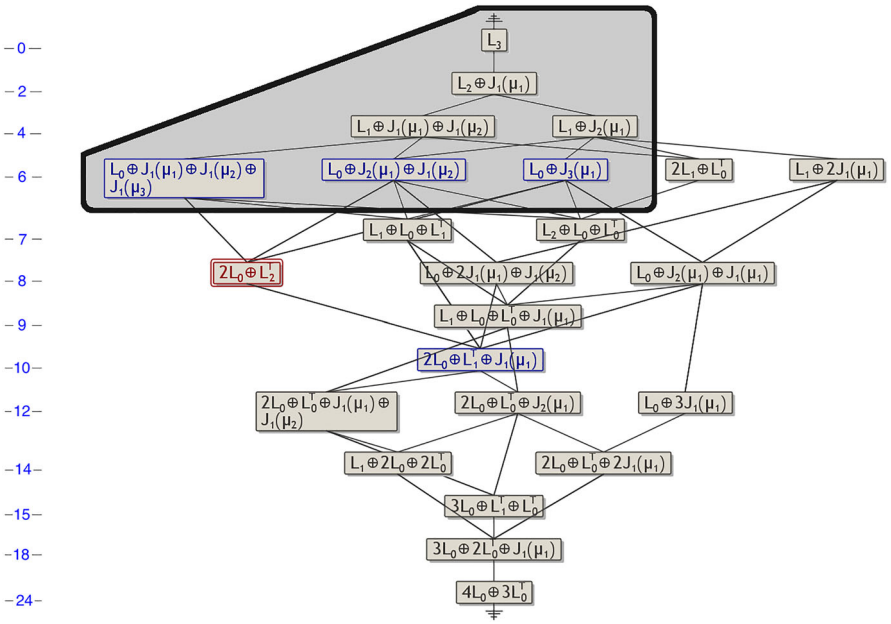

Fig. 4 Orbit stratification for 3×4 matrix pencils. The subgraph in the grey region is exactly the one from Fig. 3c, i.e. it is the stratification of the second companion form of 1×2 matrix polynomials of degree 3 ($A_3 \neq 0$). The numbers 0–24, listed on the left, are the codimensions of the orbits in the corresponding level of the graph. These codimensions are computed by Lemma 1

$$\operatorname{cod} \mathrm{B}_{\mathscr{F}^{\sigma}_{P(\lambda)}} = \operatorname{cod} \mathrm{O}_{\mathscr{F}^{\sigma}_{P(\lambda)}} - \#\left\{\text{distinct eigenvalues of } \mathscr{F}^{\sigma}_{P(\lambda)}\right\}.$$

The definition for the cover relation is analogous to the one for orbits, see Sect. 7.1. The following theorem is the bundle analog of Theorem 8.

Theorem 10 (Bundle upward rules—matrix polynomial linearizations) *Let $P_1(\lambda)$ and $P_2(\lambda)$ be two matrix polynomials with the corresponding Fiedler linearizations $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ and $\mathscr{F}^{\sigma}_{P_2(\lambda)}$, respectively. The bundle $\mathrm{B}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$* is covered by *$\mathrm{B}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$* if and only if *the canonical structure information of $P_2(\lambda)$ can be obtained by applying one of the rules below to the structure integer partitions representing the canonical structure information of $P_1(\lambda)$ (here $\mu_i \in \mathbb{C} \cup \{\infty\}$):*

(a) Same as rule (a) in Theorem 8.
(b) Same as rule (b) in Theorem 8, but only for any $\mathscr{J}_{\mu_i}$ which consists of one single coin.
(c) Same as rule (c) in Theorem 8.
(d) Same as rule (d) in Theorem 8 with the following changes. A new partition $\mathscr{J}_{\mu_i}$ for a new finite eigenvalue may only be created if there does not exist any $\mathscr{J}$ partitions. If so, all coins should be assigned to it and create one row.
(e) For any $\mathscr{J}_{\mu_i}$, split the set of coins into two new non-empty partitions such that their additive union is $\mathscr{J}_{\mu_i}$, i.e. let an eigenvalue separate into two new (different) eigenvalues.

Similarly, to Theorem 8, rules (a)–(e) above in Theorem 10 coincide with the analogous rules for the general matrix pencils presented in Table 3(D) in [30], see also [21, Theorem 3.3]. The proof is essentially the same as the proof of Theorem 8.

Remark 5 Instead of Fiedler linearizations used in this paper, it is also possible to use a broader class of linearizations, namely the block Kronecker linearizations [19]. To do so, we would have to repeat the steps of this paper for the new linearization class, proving all the missing results.

Note also that using any of the Fiedler linearizations, e.g. the first companion form, is enough to describe the changes of the complete eigenstructure of a matrix polynomial under small perturbations, see the Supplementary Materials to this paper for the rules to obtain for a given matrix polynomial, the complete eigenstructures of its neighbouring matrix polynomials, both above and below.

Example 4 In Fig. 5, we stratify the bundles of the Fiedler linerizations (19) of 2×2 matrix polynomials of degree 3. In the graph, each node represents a bundle and each edge a 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 Fig. 2 has eleven most generic orbits (all with codimension 6), marked by yellow colour. In Fig. 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 (with codimension 0).

Example 5 Similarly, to Example 4, we stratify the bundles of the Fiedler linerizations of 1×2 matrix polynomials of degree 3 and present them in Fig. 6. Recall that the orbit stratification graphs are presented in Fig. 3, see Example 3. Notably, for the bundle case there is only one least generic node and one most generic node, the latter corresponds to the same canonical structures for both the orbit and bundle cases.

8 Stratification of Matrix Polynomial Invariants

In this section, we present rules for the orbit and bundle stratifications acting directly on the minimal indices and elementary divisors of the matrix polynomials, see Sect. 2 for the definitions of these invariants. These rules can sometimes be preferable over the rules for the Fiedler linearizations given in Sects. 7.2 and 7.3 since they are independent of any linearization. The rules for orbits are presented in Theorem 11 and the corresponding rules for bundles in Theorem 12. Moreover, these rules also separate the infinite eigenvalues from the finite.

Note that, orbits and bundles of matrix polynomials are defined by analogy with the matrix pencils, i.e. an orbit of a matrix polynomial $P(\lambda)$ is a set of all the matrix polynomials with the same complete eigenstructure as $P(\lambda)$; and a bundle of a matrix polynomial $P(\lambda)$ is a union of orbits of matrix polynomials with the same complete eigenstructure as $P(\lambda)$ but with possibly different values of the eigenvalues. A codimension of the orbit or bundle of a matrix polynomial $P(\lambda)$ is defined to be the

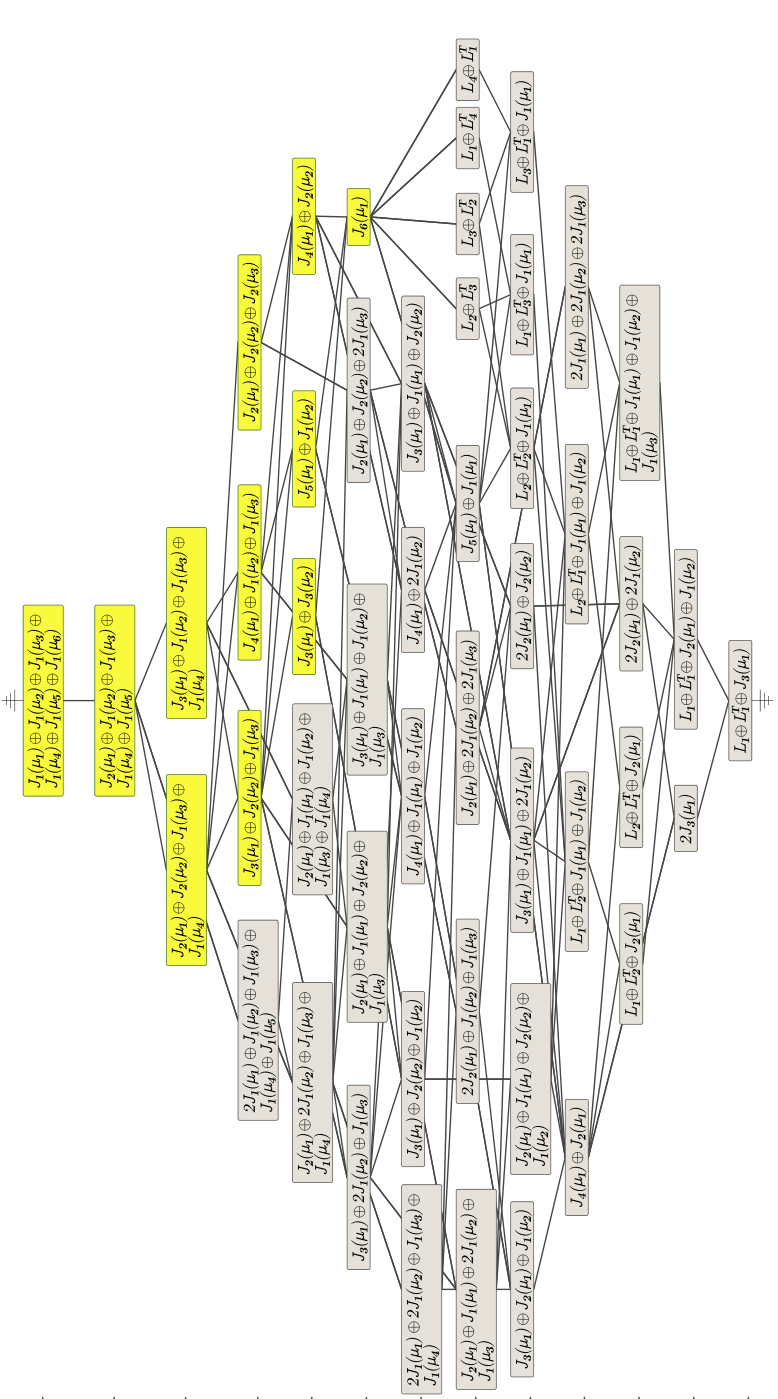

Fig. 5 Bundle stratification of the Fiedler linerizations (19) of 2×2 matrix polynomials of degree 3. The numbers 0–12, listed on the left, are the codimensions of the bundles in the corresponding level of the graph

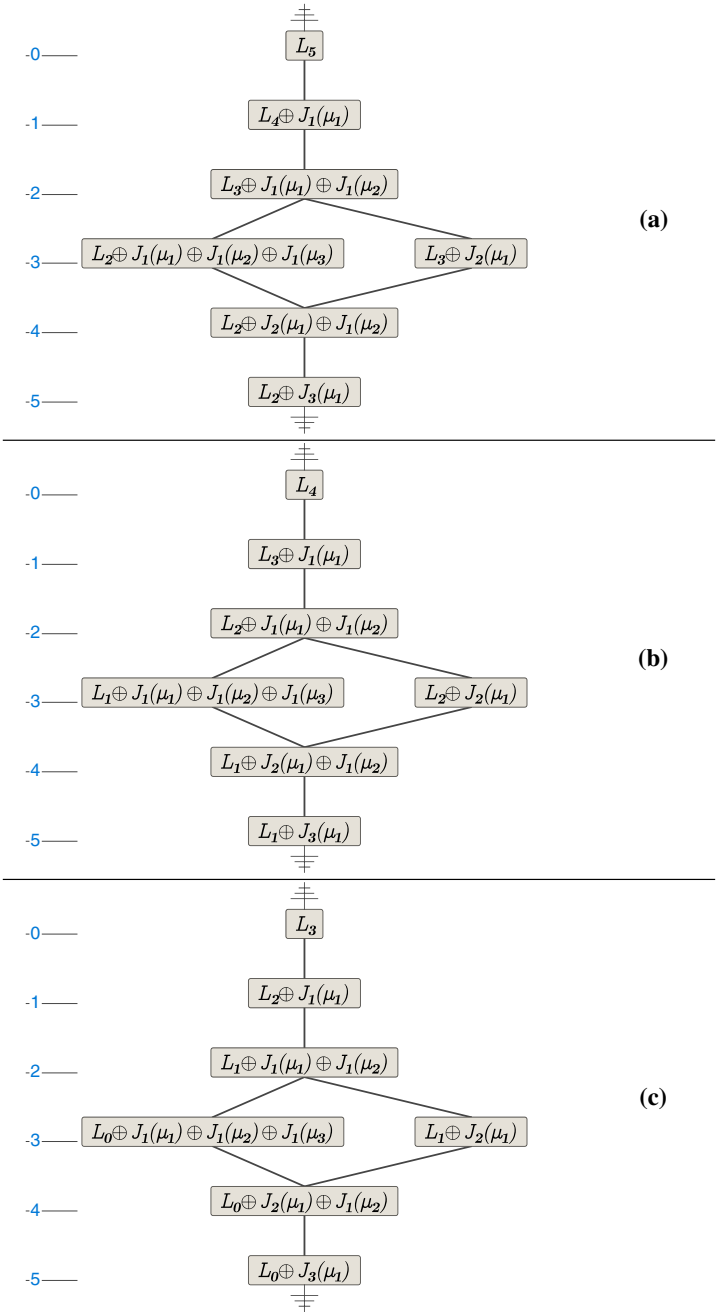

Fig. 6 Bundle stratification of the Fiedler linerizations of 1×2 matrix polynomials of degree 3. The numbers 0–5, listed on the left, are the codimensions of the bundles in the corresponding level of the graph. Similarly, to Figure 3, the graphs **a**, **b**, and **c** are the bundle stratifications of the first companion form (5×6 matrix pencils), linearizations in (19) (4×5 matrix pencils), and second companion form (3×4 matrix pencils), respectively

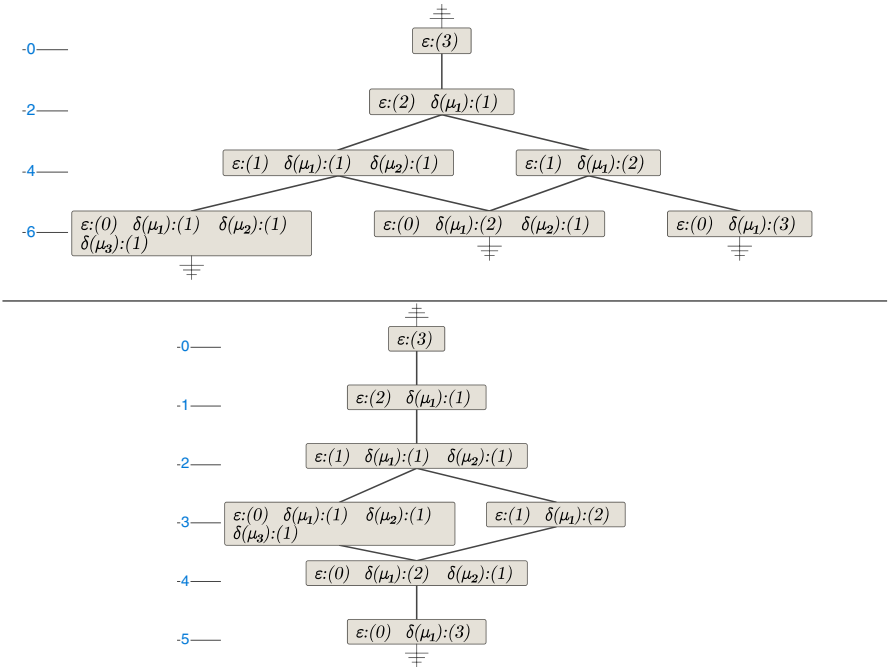

Fig. 7 Orbit (top figure) and bundle (bottom figure) stratifications of 1×2 matrix polynomials of degree 3 ($A_3 \neq 0$). The numbers, listed on the left, are the codimensions of the orbits or bundles in the corresponding level of the graph. The canonical structure information of the matrix polynomials in each node is represented by the set ϵ of right minimal indices and the set $\delta(\mu)$ of the exponents of the elementary divisors for an eigenvalue μ, see Sect. 2

codimension of, respectively, the orbits or bundles of the first companion linearization of the matrix polynomial $\mathrm{O}_{C^1_{P(\lambda)}}$. In Fig. 7, we present the orbit and bundle stratification graphs for 1×2 matrix polynomials of degree 3. Since the stratification now is done on the invariants of matrix polynomials $P(\lambda)$ (not on a linearization), the canonical structure information of the orbits/bundles in the graphs is represented by the set of right ϵ and left η minimal indices and the set $\delta(\mu)$ of exponents of the elementary divisors for an eigenvalue μ. This in contrast to Theorems 8 and 10 where the stratification is done on a Fiedler linearization and the canonical structure information can be represented by Kronecker canonical blocks. Note that the geometry of graphs is the same as the corresponding graphs for the Fiedler linearizations in Figs. 3 and 6.

Theorem 11 (Orbit upward rules—matrix polynomial invariants) *Let $P_1(\lambda)$ and $P_2(\lambda)$ be two matrix polynomials with the corresponding Fiedler linearizations $\mathscr{F}^{\sigma}_{P_1(\lambda)}$ and $\mathscr{F}^{\sigma}_{P_2(\lambda)}$, respectively.*

The orbit $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_1(\lambda)}}$ is covered by $\mathrm{O}_{\mathscr{F}^{\sigma}_{P_2(\lambda)}}$ if and only if *the canonical structure information of $P_2(\lambda)$ can be obtained by applying one of the rules below to the structure integer partitions representing the canonical structure information of $P_1(\lambda)$.*

(a) Same as rule (a) in Theorem 8.
(b) Same as rule (b) in Theorem 8, where $\mu_i = \infty$ or $\mu_i \in \mathbb{C}$.

(c) *Same as rule (c) in Theorem 8, where* $\mu_i = \infty$ *or* $\mu_i \in \mathbb{C}$.
(d) *Same as rule (d) in Theorem 8, but instead distribute* $k+d-2 = (k-1)+(d-1)$ *coins as follows. First add one coin to each nonzero column in* $\mathcal{J}_\infty$ *and then distribute one coin to each nonzero column in all existing* $\mathcal{J}_{\mu_i}$, $\mu_i \in \mathbb{C}$. *The remaining coins are distributed to* $\mathcal{J}_\infty$ *or any* $\mathcal{J}_{\mu_i}$ *which may be empty initially.*[4]

Below follows the stratification rules for bundles. In addition to the differences between the orbit and bundle cases pointed out in Sect. 7.3, the following theorem has the two additional rules (f) and (g) for the specified infinite eigenvalue. The two rules are a direct consequence of that the infinite eigenvalue is treated as a specified eigenvalue.

Theorem 12 (Bundle upward rules—matrix polynomial invariants) *Let* $P_1(\lambda)$ *and* $P_2(\lambda)$ *be two matrix polynomials with the corresponding Fiedler linearizations* $\mathcal{F}^\sigma_{P_1(\lambda)}$ *and* $\mathcal{F}^\sigma_{P_2(\lambda)}$, *respectively.*

The bundle $\mathrm{B}_{\mathcal{F}^\sigma_{P_1(\lambda)}}$ is covered by $\mathrm{B}_{\mathcal{F}^\sigma_{P_2(\lambda)}}$ if and only if *the canonical structure information of* $P_2(\lambda)$ *can be obtained by applying one of the rules below to the structure integer partitions representing the canonical structure information of* $P_1(\lambda)$.

(a) *Same as rule (a) in Theorem 10.*
(b) *Same as rule (b) in Theorem 10, where* $\mu_i = \infty$ *or* $\mu_i \in \mathbb{C}$.
(c) *Same as rule (c) in Theorem 10, where* $\mu_i = \infty$ *or* $\mu_i \in \mathbb{C}$.
(d) *Same as rule (d) in Theorem 10, but instead distribute* $k+d-2 = (k-1)+(d-1)$ *coins as follows. First add one coin to each nonzero column in* $\mathcal{J}_\infty$ *and then distribute one coin to each nonzero column in all existing* $\mathcal{J}_{\mu_i}$, $\mu_i \in \mathbb{C}$. *The remaining coins are distributed to* $\mathcal{J}_\infty$ *(which may be empty initially) or to existing* $\mathcal{J}_{\mu_i}$ *(see footnote 4).*
(e) *Same as rule (e) in Theorem 10, where* $\mu_i \in \mathbb{C}$.
(f) *For* $\mathcal{J}_\infty$, *split the set of coins into one new non-empty partition* $\mathcal{J}_{\mu_i}$ *for a new finite eigenvalue and keep the remaining coins in* $\mathcal{J}_\infty$ *such that* $\mathcal{J}_\infty^{old} = \mathcal{J}_\infty^{new} \biguplus \mathcal{J}_{\mu_i}$.
(g) *If* $\mathcal{J}_\infty$ *consists of one single coin, move that coin to a new* $\mathcal{J}_{\mu_i}$ *for a new finite eigenvalue* μ_i.

Acknowledgements The authors are greatful to Froilán Dopico and Volker Mehrmann for their constructive comments and discussions on an earlier manuscript. The authors also thank the anonymous referees for their helpful suggestions. The work was supported by the Swedish Research Council (VR) under Grant E0485301, and by eSSENCE (essenceofescience.se), a strategic collaborative e-Science programme funded by the Swedish Research Council.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[4] Cannot be applied if the total number of nonzero columns of all $\mathcal{J}_{\mu_i}$ and $\mathcal{J}_\infty$ together is greater than $k+d-2$.

References

1. Antoniou, E., Vologiannidis, S.: A new family of companion forms of polynomial matrices. Electron. J. Linear Algebra **11**(411), 78–87 (2004)
2. Arnold, V.: On matrices depending on parameters. Russian Math. Surveys **26**, 29–43 (1971)
3. Betcke, T., Higham, N., Mehrmann, V., Schröder, C., Tisseur, F.: NLEVP: A Collection of Nonlinear Eigenvalue Problems. ACM Trans. Math. Software **39**(2), 7:1–7:28 (2013). https://doi.org/10.1145/2427023.2427024
4. den Boer, H., Thijsse, P.: Semi-stability of sums of partial multiplicities under additive perturbation. Integral Equations Operator Theory **3**, 23–42 (1980)
5. De Terán, F., Dopico, F., Mackey, D.: Fiedler companion linearizations and the recovery of minimal indices. SIAM J. Matrix Anal. Appl. **31**(4), 2181–2204 (2010)
6. De Terán, F., Dopico, F., Mackey, D.: Fiedler companion linearizations for rectangular matrix polynomials. Linear Algebra Appl. **437**(3), 957–991 (2012)
7. De Terán, F., Dopico, F., Van Dooren, P.: Matrix polynomials with completely prescribed eigenstructure. SIAM J. Matrix Anal. Appl. **36**, 302–328 (2015)
8. Demmel, J., Edelman, A.: The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms. Linear Algebra Appl. **230**, 61–87 (1995)
9. Dmytryshyn, A.: Structure preserving stratification of skew-symmetric matrix polynomials. Linear Algebra Appl. **532**, 266–286 (2017)
10. Dmytryshyn, A., Dopico, F.M.: Generic complete eigenstructures for sets of matrix polynomials with bounded rank and degree. Linear Algebra Appl. **535**, 213–230 (2017). https://doi.org/10.1016/j.laa.2017.09.007.
11. Dmytryshyn, A., Futorny, V., Kågström, B., Klimenko, L., Sergeichuk, V.: 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**, 305–334 (2015). https://doi.org/10.1016/j.laa.2014.11.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0024379514007289
12. Dmytryshyn, A., Futorny, V., Sergeichuk, V.: Miniversal deformations of matrices of bilinear forms. Linear Algebra Appl. **436**, 2670–2700 (2012)
13. Dmytryshyn, A., Futorny, V., Sergeichuk, V.: Miniversal deformations of matrices under *congruence and reducing transformations. Linear Algebra Appl. **446**, 388–420 (2014). https://doi.org/10.1016/j.laa.2014.01.016
14. Dmytryshyn, A., Johansson, S., Kågström, B.: Codimension computations of congruence orbits of matrices, symmetric and skew-symmetric matrix pencils using Matlab. Tech. Rep. UMINF 13.18, Department of Computing Science, Umeå University, Sweden (2013)
15. Dmytryshyn, A., Johansson, S., Kågström, B.: Canonical structure transitions of system pencils. SIAM J. Matrix Anal. Appl. **38**(4), 1249–1267 (2017)
16. Dmytryshyn, A., Kågström, B.: Orbit closure hierarchies of skew-symmetric matrix pencils. SIAM J. Matrix Anal. Appl. **35**(4), 1429–1443 (2014). https://doi.org/10.1137/140956841. URL http://epubs.siam.org/doi/abs/10.1137/140956841?journalCode=sjmael
17. Dmytryshyn, A., Kågström, B., Sergeichuk, V.: Skew-symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations. Linear Algebra Appl. **438**(8), 3375–3396 (2013). https://doi.org/10.1016/j.laa.2012.11.025. URL http://www.sciencedirect.com/science/article/pii/S0024379512008324
18. Dmytryshyn, A., Kågström, B., Sergeichuk, V.: Symmetric matrix pencils: codimension counts and the solution of a pair of matrix equations. Electron. J. Linear Algebra **27**, 1–18 (2014)
19. Dopico, F., Lawrence, P.W., Pérez, J., Van Dooren, P.: Block kronecker linearizations of matrix polynomials and their backward errors. Numer. Math. **140**(2), 373–426 (2018). https://doi.org/10.1007/s00211-018-0969-z.
20. Edelman, A., Elmroth, E., Kågström, B.: A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations. SIAM J. Matrix Anal. Appl. **18**(3), 653–692 (1997)
21. Edelman, A., Elmroth, E., Kågström, B.: 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), 667–669 (1999)
22. Elmroth, E., Johansson, S., Kågström, B.: Stratification of controllability and observability pairs — Theory and use in applications. SIAM J. Matrix Anal. Appl. **31**(2), 203–226 (2009)

23. Gallier, J.: Geometric methods and applications, *Texts in Applied Mathematics*, vol. 38, second edn. Springer, New York (2011). https://doi.org/10.1007/978-1-4419-9961-0. For computer science and engineering
24. Gantmacher, F.: The theory of matrices, Vol. I and II (transl.). Chelsea, New York (1959)
25. García-Planas, M., Sergeichuk, V.: Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils. Linear Algebra Appl. **302–303**, 45–61 (1999)
26. Hein, J.L.: Discrete Mathematics. Jones and Bartlett Publishers (2003)
27. Hilliges, A., Mehl, C., Mehrmann, V.: On the solution of palindromic eigenvalue problems. In: Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS). Jyväskylä, Finland (2004)
28. Ipsen, I.: Accurate eigenvalues for fast trains. SIAM News **37**(9), 1–2 (2004)
29. Johansson, P.: Software Tools for Matrix Canonical Computations and Web-based Software Library Environments. Ph.D. thesis, Department of Computing Science, Umeå University, Sweden (2006)
30. Johansson, S.: Reviewing the closure hierarchy of orbits and bundles of system pencils and their canonical forms. Tech. Rep. UMINF 09.02, Department of Computing Science, Umeå University, Sweden (2009)
31. Johansson, S., Johansson, P.: StratiGraph and MCS Toolbox homepage. Department of Computing Science, Umeå University, Sweden (2019). URL https://www.umu.se/en/stratigraph-mcs
32. Johansson, S., Kågström, B., Van Dooren, P.: Stratification of full rank polynomial matrices. Linear Algebra Appl. **439**, 1062–1090 (2013). https://doi.org/10.1016/j.laa.2012.12.013
33. Kågström, B., Johansson, S., Johansson, P.: StratiGraph Tool: Matrix Stratification in Control Applications. In: L. Biegler, S. Campbell, V. Mehrmann (eds.) Control and Optimization with Differential-Algebraic Constraints, chap. 5. SIAM Publications (2012)
34. Karlsson, L., Tisseur, F.: Algorithms for Hessenberg-Triangular Reduction of Fiedler Linearization of Matrix Polynomials. SIAM Journal on Scientific Computing **37**(3), C384–C414 (2015). https://doi.org/10.1137/140970458.
35. Kraft, H., Procesi, C.: Minimal singularities in gl_n. Invent. Math. **62**(3), 503–515 (1981)
36. Kressner, D., Schröder, C., Watkins, D.: Implicit QR algorithms for palindromic and even eigenvalue problems. Numerical Algorithms **51**(2), 209–238 (2009)
37. Mackey, D., Mackey, N., Mehl, C., Mehrmann, V.: Vector spaces of linearizations for matrix polynomials. SIAM J. Matrix Anal. Appl. **28**(4), 971–1004 (2006)
38. Mackey, D., Mackey, N., Tisseur, F.: Polynomial Eigenvalue Problems: Theory, Computation, and Structure. In: Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory, pp. 319–348. Springer (2015)
39. Mackey, D.S., Mackey, N., Mehl, C., Mehrmann, V.: Numerical methods for palindromic eigenvalue problems: Computing the anti-triangular Schur form. Numerical Linear Algebra with Applications **16**(1), 63–86 (2009)
40. Markus, A., Parilis, E.: The change of the Jordan structure of a matrix under small perturbations. Linear Algebra Appl. **54**, 139–152 (1983). Transl. from Mat. Issled., 54, 98–109 (1980).
41. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. GAMM-Mitteilungen **27**(2), 121–152 (2004)
42. Tisseur, F., Meerbergen, K.: The quadratic eigenvalue problem. SIAM Review **43**(2), 235–286 (2001)
43. Van Dooren, P., Dewilde, P.: The Eigenstructure of a Polynomial Matrix: Computational Aspects. Linear Algebra Appl. **50**, 545–579 (1983)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.