Trimmed linearizations for structured matrix polynomials

Ralph Byers[§†*] Volker Mehrmann[‡*] Hongguo Xu[§¶*]

January 15, 2008

Dedicated to Richard S. Varga on the occasion of his 80th birthday.

Abstract

We discuss the eigenvalue problem for general and structured matrix polynomials which may be singular and may have eigenvalues at infinity. We derive condensed forms that allow (partial) deflation of the infinite eigenvalue and singular structure of the matrix polynomial. The remaining reduced order staircase form leads to new types of linearizations which determine the finite eigenvalues and corresponding eigenvectors. The new linearizations also simplify the construction of structure preserving linearizations.

Keywords matrix polynomial, singular matrix polynomial, Kronecker chain, staircase form, linearization, trimmed linearization, structured trimmed linearization.
AMS subject classification. 65F15, 15A21, 65L80, 65L05, 34A30.

1 Introduction

We study k-th degree matrix polynomials

$$P(\lambda) = \lambda^k A_k + \lambda^{k-1} A_{k-1} + \cdots + \lambda A_1 + A_0, \tag{1.1}$$

with coefficients $A_i \in \mathbb{F}^{m,n}$, where $\mathbb{F}$ is the field of real $\mathbb{R}$ or complex $\mathbb{C}$ numbers.

The main topic of this paper is the reformulation of matrix polynomials as first degree (linear) matrix polynomials of larger dimension.

Definition 1.1 (Linearization) *Let $P(\lambda)$ be an $m \times n$ matrix polynomial of degree k. A pencil $L(\lambda) = \lambda X + Y$ is called a* linearization *of $P(\lambda)$ if there exist unimodular matrix polynomials $E(\lambda), F(\lambda)$ such that*

$$E(\lambda)L(\lambda)F(\lambda) = \left[\begin{array}{c|c} P(\lambda) & 0 \\ \hline 0 & I_s \end{array}\right].$$

[‡]Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG. `mehrmann@math.tu-berlin.de`.

[§]Department of Mathematics, University of Kansas, Lawrence, KS 44045, USA.

[†]`byers@math.ku.edu`. Partially supported the University of Kansas General Research Fund allocation #2301054-003, and by the National Science Foundation grants 0098150, 0112375.

[¶]`xu@math.ku.edu`. Partially supported by the National Science Foundation grant 0314427. *Part of the work was done while this author was visiting TU Berlin whose hospitality is gratefully acknowledged.*

[*]Partially supported by the *Deutsche Forschungsgemeinschaft* through the DFG Research Center MATHEON *Mathematics for key technologies* in Berlin.

(A matrix polynomial $E(\lambda)$ is unimodular if it is square with constant, nonzero determinant independent of λ.)

Note that in contrast to the usual definition of linearization, see e.g. [11, 17], we do not require that the linear pencil $L(\lambda) = \lambda X + Y$, satisfies $X, Y \in \mathbb{F}^{(m+(k-1)n)\times kn}$ or $X, Y \in \mathbb{F}^{(km\times(n+(k-1)m)}$. We allow the dimension to be smaller than this, i.e. we allow $s < (k-1)\min\{m, n\}$, but the usual case is certainly included.

Linearization makes it possible to use mature, well-understood, numerical methods and software developed for linear matrix pencils and the associated differential algebraic equations.

The (first) companion form linearization of the matrix polynomial (1.1) is

$$\lambda \begin{bmatrix} A_k & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_1 & A_0 \\ -I & 0 & \cdots & 0 & 0 \\ 0 & -I & & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & -I & 0 \end{bmatrix}.$$

Companion form linearizations are elegant and successful [11, 16, 17]. However, companion form linearizations may not share the structure of the original matrix polynomial. For example, if the original matrix polynomial is symmetric, skew-symmetric, even or odd, then the companion form linearization is not. Thus, rounding errors in numerical computations on companion form linearizations may destroy vital qualitative aspects of the spectrum like eigenvalue pairing. Companion form linearization may also introduce artificial and unnecessary pathologies. (See Example 1.7 below.) In particular, companion forms are not consistent with the first order formulations for differential-algebraic equations used in multi-body dynamics [7]. (See also [23] for optimal first order formulations in the context of differential-algebraic equations.)

New classes of structure preserving linearizations introduced in [19] and analyzed in [12, 13, 18, 20] hold much promise. Still a different family of linearizations was introduced in [1, 2]. However, in order to use some of these new linearizations, certain eigenvalues must first be deflated from the matrix polynomial in a structure preserving way. Numerically stable procedures for such structured deflation is one of our goals here. To carry out this deflation we will need the following equivalence transformations.

Definition 1.2

(i) Two tuples of matrices $(A_\ell, \dots, A_1, A_0)$ and $(B_\ell, \dots, B_1, B_0)$, $A_i, B_i \in \mathbb{F}^{m,n}$, $i = 0, 1, \dots, \ell$, $0 \le \ell \in \mathbb{N}$, are called strongly equivalent*, denoted by*

$$(A_\ell, \dots, A_1, A_0) \sim (B_\ell, \dots, B_1, B_0),$$

if there exist nonsingular matrices $P \in \mathbb{F}^{m,m}$ and $Q \in \mathbb{F}^{n,n}$ such that

$$B_i = PA_iQ, \quad i = 0, 1, \dots, \ell. \tag{1.2}$$

If both P and Q are unitary (real orthogonal), then the two tuples are called strongly u-equivalent*, denoted by*

$$(A_\ell, \dots, A_1, A_0) \overset{u}{\sim} (B_\ell, \dots, B_1, B_0).$$

ii) Two tuples of matrices $(A_\ell, \ldots, A_1, A_0)$ *and* $(B_\ell, \ldots, B_1, B_0)$, $A_i, B_i \in \mathbb{F}^{n,n}$, $i = 0, 1, \ldots, \ell$, $\ell \in \mathbb{N}_0$, *are called* strongly congruent, *denoted by*

$$(A_\ell, \ldots, A_1, A_0) \stackrel{c}{\sim} (B_\ell, \ldots, B_1, B_0),$$

if there exists a nonsingular matrix $Q \in \mathbb{F}^{n,n}$ *such that*

$$B_i = Q^\star A_i Q, \quad i = 0, 1, \ldots, \ell, \tag{1.3}$$

where $\star$ *is either the transpose or the conjugate transpose depending on the matrix structures of the tuples under consideration.*

If Q *is unitary (real orthogonal), then the two tuples are called* strongly u-congruent, *denoted by*

$$(A_\ell, \ldots, A_1, A_0) \stackrel{uc}{\sim} (B_\ell, \ldots, B_1, B_0).$$

At this writing, a generalization of Kronecker canonical form to matrix polynomials of degree greater than one, i.e. a canonical form under the equivalences of Definition 1.2, is unknown and seems unlikely to exist. However, Jordan and Kronecker chains are partially generalized by the concepts of Jordan triples, see [11].

Another approach is the canonical form for higher order differential-algebraic equations derived in [23] which displays partial information about the Kronecker structure at infinity and the singular structure. This approach, however, uses non-orthogonal (non-unitary) transformations and does not preserve structure, so as a computational method, its numerical stability can not be guaranteed.

In this paper, we present condensed forms under the equivalence transformations in Definition 1.2 that allow the computation of (partial) structural information associated with the eigenvalue infinity and the singular parts of matrix polynomials. If unitary or real orthogonal equivalences are used then such condensed forms are usually called *staircase forms* [6, 25]. Based on these condensed forms we present new first order formulations (linearizations) which we call *trimmed linearizations* although a more apt term might be *trimmed first order formulations*.

We show that these trimmed linearizations properly reflect the structural information about the finite eigenvalues. Hence, on the one hand, trimmed first order formulations generalize the classical concept of linearization, if the matrix polynomial is regular and has no eigenvalue infinity, and, on the other hand, they generalize first order formulations used in constrained multibody dynamics [7] and general higher order differential-algebraic systems [23]. Furthermore, we show that they allow structure preservation under orthogonal/unitary transformations. In all these aspects, our approach differs significantly from the companion form approach.

Let us therefore recall the classical definitions of Jordan/Kronecker chains for matrix polynomials.

Definition 1.3 *A matrix polynomial* $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, *with* $A_0, \ldots, A_k \in \mathbb{F}^{m,n}$, $A_k \neq 0$. *is called* regular *if the coefficients are square matrices and if* $\det P(\lambda)$ *does not vanish identically for all* $\lambda \in \mathbb{C}$, *otherwise it is called* singular.

Definition 1.4 ([17]) *Let* $P(\lambda)$ *be a matrix polynomial as in (1.1).*

A right (left) Jordan chain of length $\ell+1$ *associated with a finite eigenvalue* $\hat{\lambda}$ *of* $P(\lambda)$ *is a sequence of vectors* x_i (y_i), $i = 0, 1, 2, \ldots, \ell$ *with* x_ℓ (y_ℓ) *nonzero and the property that*

$$\begin{aligned} P(\hat{\lambda})x_0 &= 0; \\ P(\hat{\lambda})x_1 + [\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})]x_0 &= 0; \\ &\vdots \\ P(\hat{\lambda})x_\ell + [\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})]x_{\ell-1} + \ldots + [\tfrac{1}{\ell!}\tfrac{d^\ell}{d\lambda^\ell}P(\hat{\lambda})]x_0 &= 0, \end{aligned} \tag{1.4}$$

$$\begin{aligned} y_0^\star P(\hat{\lambda}) &= 0; \\ y_1^\star P(\hat{\lambda}) + y_0^\star[\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})] &= 0; \\ &\vdots \\ y_\ell^\star P(\hat{\lambda}) + y_{\ell-1}^\star[\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})] + \ldots + y_0^\star[\tfrac{1}{\ell!}\tfrac{d^\ell}{d\lambda^\ell}P(\hat{\lambda})] &= 0, \end{aligned} \tag{1.5}$$

respectively.

A right (left) Kronecker chain of length $\ell+1$ *associated with the eigenvalue infinity of* $P(\lambda)$ *is a right (left) Kronecker chain of length* $\ell+1$ *associated with eigenvalue* $\lambda = 0$ *of the* reverse polynomial $\operatorname{rev} P(\lambda) = \sum_{i=0}^{k} \lambda^{k-i} A_i$.

For Kronecker chains associated with the singular parts of the matrix polynomials we extend the classical definition for matrix pencils as in [8, 9, 17].

Definition 1.5 *Let* $P(\lambda)$ *be a matrix polynomial as in (1.1).*

A right singular Kronecker chain of length $\ell+1$ *associated with the right singular part of* $P(\lambda)$ *is defined as the sequence of coefficient vectors* x_i, $i = 0, 1, 2, \ldots, \ell$ *in a nonzero vector polynomial* $x(\lambda) = x_\ell\lambda^\ell + \ldots + x_1\lambda^1 + x_0$ *of minimal degree such that*

$$P(\lambda)x(\lambda) = 0, \tag{1.6}$$

considered as an equation in polynomials in λ.

A left singular Kronecker chain of length $\ell+1$ *associated with the left singular part of* $P(\lambda)$ *is defined analogously as a sequence of coefficient vectors* y_i, $i = 0, 1, 2, \ldots, \ell$ *in a nonzero vector polynomial* $y(\lambda) = y_\ell\lambda^\ell + \ldots + y_1\lambda^1 + y_0$ *of minimal degree such that*

$$y(\lambda)^\star P(\lambda) = 0. \tag{1.7}$$

Here $y(\lambda)^\star = y_\ell^\star\lambda^\ell + \ldots + y_1^\star\lambda^1 + y_0^\star$.

One difficulty with linearizations is that unimodular transformations from the left may alter the lengths of left chains associated with the eigenvalue infinity and the left singular chains, while unimodular transformations from the right may alter the lengths of right chains associated with the eigenvalue infinity and the right singular chains. Accordingly, Definition 1.1 puts different first order formulations in the same class. This observation in the context of infinite eigenvalues led to the definition of strong linearization in [10]. A linear pencil $L(\lambda)$ is a *strong linearization* of a matrix polynomial $P(\lambda)$ if it is a linearization and, at the same time, $\operatorname{rev} L(\lambda)$ is a linearization of $\operatorname{rev} P(\lambda)$. The companion form linearization of a matrix polynomial is a strong linearization. Although strong linearizations avoid some anomalies, we demonstrate in Example 1.8 that a strong linearization of a singular matrix pencil may not preserve the lengths of singular chains. Linearizations like the one in Examples 1.6, 1.7 below correspond to systems of first order differential-algebraic equations that have better computational properties (smaller index) than can be obtained from strong linearizations.

Example 1.6 *Consider the following matrix polynomial which has the structure of a constrained and damped mechanical system [7].*

$$P(\lambda) = \begin{bmatrix} \lambda^2+\lambda+1 & 1 \\ 1 & 0 \end{bmatrix} = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Multiplying $P(\lambda)$ on the left with the unimodular transformation

$$Q(\lambda) = \begin{bmatrix} 0 & 1 \\ 1 & -(\lambda^2+\lambda+1) \end{bmatrix}$$

we obtain the linearization

$$T(\lambda) = Q(\lambda)P(\lambda) = I$$

which has only degree 0. It is not clear whether it is best to treat $T(\lambda)$ as the degree zero polynomial I, as the "degree one" polynomial $\lambda 0 + I$ or as the "degree two" polynomial $\lambda^2 0 + \lambda 0 + I$.

The companion form linearization of $P(\lambda)$ has a chain of length 4 associated with the eigenvalue infinity. Treating $T(\lambda)$ as a degree two polynomial, the companion form linearization is also a linearization of $P(\lambda)$ which has two chains of length 2 associated with infinity. Regarding $T(\lambda) = \lambda 0 + I$ as a degree one matrix pencil, $T(\lambda)$ itself is a linearization of $P(\lambda)$ which has two chains of length 1 associated to infinity.

A even more extreme example is the $k \times k$ identity polynomial which is unimodularly equivalent to upper triangular matrix polynomials of arbitrary degree.

One of the motivations for studying matrix polynomials is the analysis of higher order differential-algebraic equations like those arising in multi-body dynamics. Consider what is done to obtain first order formulations for higher order differential-algebraic equations.

Example 1.7 The Euler-Lagrange equations [7] of a linear constrained and damped mechanical system are given by a differential-algebraic equation of the form

$$\begin{array}{rcl} M\ddot{x} + D\dot{x} + Kx + G^T\mu & = & f(t), \\ Gx & = & 0. \end{array} \tag{1.8}$$

Here M, D, K are mass, damping and stiffness matrices, G describes the constraint, f a forcing function, x is a vector of position variables and μ a Lagrange multiplier.

The associated matrix polynomial is

$$P(\lambda) = \lambda^2 \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} K & G^T \\ G & 0 \end{bmatrix}. \tag{1.9}$$

Under the usual assumptions, i.e. that M is positive definite and that G has full row rank, it can be easily shown that according to Definition 1.4, $P(\lambda)$ has Kronecker chains associated with the eigenvalue infinity of length 4.

The companion linearization is

$$L(\lambda) = \lambda \left[\begin{array}{cc|cc} M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{array}\right] + \left[\begin{array}{cc|cc} D & 0 & K & G^T \\ 0 & 0 & G & 0 \\ \hline -I & 0 & 0 & 0 \\ 0 & -I & 0 & 0 \end{array}\right]. \tag{1.10}$$

It corresponds to extending the two unknowns $[x, \mu]^T$ in (1.8) to four unknowns $[y, \nu, x, \mu]^T$ by introducing new variables $y = \dot{x}$ and $\nu = \dot{\mu}$ (which correspond to $y = \lambda x$ and $\nu = \lambda \mu$ in (1.10)). The derivative of the Lagrange multiplier is intuitively unsatisfying. In contrast, the first order formulation that is used in multibody dynamics introduces only one new variable $y = \dot{x}$ (corresponding to $y = \lambda x$ below in (1.11)) and does not introduce a derivative of the Lagrange multiplier μ. This approach gives the linear matrix pencil

$$\tilde{L}(\lambda) = \lambda \left[\begin{array}{cc|c} M & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{cc|c} D & K & G^T \\ -I & 0 & 0 \\ \hline 0 & G & 0 \end{array}\right], \tag{1.11}$$

which (under the same assumptions) has a Kronecker chain associated with infinity of length 3. Thus, the companion linearization has a longer chain than necessary to obtain the solution of the differential-algebraic equation and this should be avoided, since it is well known that longer chains at infinity create difficulties for numerical solution methods, see e.g. [3, 15, 23].

It has been demonstrated in [23] for general linear high-order differential-algebraic equations that even the formulation used in constrained multibody-dynamics may have unnecessary long chains associated with infinity in the first order formulation. Thus, it would be preferable to have first order formulations where all chains associated with infinity are as short as possible. Finding such linearizations is one of the goals of this paper.

The next example demonstrates that even strong linearizations may not preserve the lengths of singular chains in a singular matrix polynomial.

Example 1.8 For the singular matrix polynomial

$$P(\lambda) = \begin{bmatrix} \lambda^2 + \lambda & 0 \\ 1 & 0 \end{bmatrix},$$

following Definition 1.5, we obtain as right nullspace the vector-polynomial $x(\lambda) = e_2$ which creates a chain of length 1 and from the left $y(\lambda) = \begin{bmatrix} -1 \\ \lambda^2+\lambda \end{bmatrix}$ which gives $y_0 = -e_1,\ y_1 = e_2,\ y_2 = e_2$ and thus the chain has length 3.

Considering the first companion linearization, we get

$$L(\lambda) = \lambda \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] + \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right].$$

The right and left nullspace vector-polynomials are

$$x(\lambda) = \left[\begin{array}{c} 0 \\ \lambda \\ \hline 0 \\ 1 \end{array}\right], \quad y(\lambda) = \left[\begin{array}{c} 1 \\ -\lambda^2 - \lambda \\ \hline \lambda + 1 \\ 0 \end{array}\right]$$

and clearly the right chain does not have the same length as in the original matrix polynomial.

Instead of the companion form we may proceed similarly to the constrained multibody system and introduce only one new variable. This gives the linear pencil

$$\tilde{L}(\lambda) = \lambda \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right] + \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \hline -1 & 0 & 0 \end{array}\right]$$

with right and left nullspace vector-polynomials

$$x(\lambda) = \left[\begin{array}{c} 0 \\ 1 \\ \hline 0 \end{array}\right], \; y(\lambda) = \left[\begin{array}{c} 1 \\ -\lambda^2 - \lambda \\ \hline \lambda + 1 \end{array}\right]$$

and thus both the left and the right chains have the correct length.

Motivated by the above examples, a goal of this paper is to find linearizations that minimize the lengths of chains corresponding to eigenvalue infinity and, in the singular case, minimize the lengths of singular chains. The paper is organized as follows. In section 2 we discuss staircase forms for matrix polynomials and show that some (but maybe not all) of the information associated with the eigenvalue infinity and the singular parts can be obtained from the staircase forms. We use these staircase forms to obtain trimmed linearizations for general matrix polynomials in Section 3 and for structured matrix polynomials in Section 4.

2 Condensed forms for tuples of matrices

In this section we discuss condensed forms for matrix tuples associated with matrix polynomials. As mentioned in the introduction, it is an open problem [11, 24] to find a canonical form for matrix polynomials of degree greater than 1 under strong equivalence.

However, for pairs of matrices, i.e. linear matrix polynomials, although a Kronecker form exists [8, 9] and the information about the invariants can be computed numerically via the so called generalized upper-triangular (GUPTRI) form, see [5, 6, 25], in general one does not need the complete canonical or staircase form to extract the information about the singular blocks and the eigenvalue infinity.

Usually one can reduce the work and the computational difficulties by computing partial canonical or staircase forms, see [3, 15] for such forms in the context of differential-algebraic equations and [4] for a summary of results in case of matrix pairs.

In the context of higher order differential-algebraic equations a partial canonical form has been derived in [23]. But this form uses nonorthogonal transformations and only determines the information about the right eigenvectors and chains associated with the eigenvalue infinity and the singular parts.

In the following we, therefore, derive staircase forms for structured and unstructured matrix tuples under unitary (orthogonal) transformations that display (partial) information about the singular parts and the eigenvalue infinity.

To derive the staircase forms we will need the following Lemma.

Lemma 2.1 *If $N_i \in \mathbb{F}^{m,n}$, $i = 1, \ldots, k$, and $K \in \mathbb{F}^{m,n}$, then the tuple $(N_k, \ldots, N_1, K)$ is strongly u-equivalent to a matrix tuple $(\hat{N}_k, \ldots, \hat{N}_1, \hat{K})$, where all terms $\hat{N}_i$, $i = 1, \ldots, k$, have* *the form*

$$\hat{N}_i = \begin{array}{c} \begin{array}{cccc|c|cccc} q_1 & \cdots & \cdots & q_\tau & t & n_\tau & \cdots & \cdots & n_1 \end{array} \\ \left[\begin{array}{cccc|c|cccc}
N_{11}^{(i)} & \cdots & \cdots & N_{1\tau}^{(i)} & N_{1,\tau+1}^{(i)} & N_{1,\tau+2}^{(i)} & \cdots & N_{1,2\tau}^{(i)} & 0 \\
\vdots & \ddots & & \vdots & \vdots & \vdots & \iddots & \iddots & \\
\vdots & & \ddots & \vdots & \vdots & N_{\tau-1,\tau+2}^{(i)} & \iddots & & \\
N_{\tau 1}^{(i)} & \cdots & \cdots & N_{\tau\tau}^{(i)} & N_{\tau,\tau+1}^{(i)} & 0 & & & \\ \hline
N_{\tau+1,1}^{(i)} & \cdots & \cdots & N_{\tau+1,\tau}^{(i)} & N_{\tau+1,\tau+1}^{(i)} & & & & \\ \hline
N_{\tau+2,1}^{(i)} & \cdots & N_{\tau+2,\tau-1}^{(i)} & 0 & & & & & \\
\vdots & \iddots & \iddots & & & & & & \\
N_{2\tau,1}^{(i)} & \iddots & & & & & & & \\
0 & & & & & & & &
\end{array}\right] \end{array}
\begin{array}{c} m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline p_\tau \\ \vdots \\ p_2 \\ p_1 \end{array}, \qquad (2.1)$$

while the matrix $\hat{K}$ has the form

$$\hat{K} = \begin{array}{c} \begin{array}{cccc|c|cccc} q_1 & \cdots & \cdots & q_\tau & t & n_\tau & \cdots & \cdots & n_1 \end{array} \\ \left[\begin{array}{cccc|c|cccc}
K_{11} & \cdots & \cdots & K_{1\tau} & K_{1,\tau+1} & K_{1,\tau+2} & \cdots & \cdots & K_{1,2\tau+1} \\
\vdots & \ddots & & \vdots & \vdots & \vdots & & \iddots & \\
\vdots & & \ddots & \vdots & \vdots & \vdots & \iddots & & \\
K_{\tau 1} & \cdots & \cdots & K_{\tau\tau} & K_{\tau,\tau+1} & K_{\tau,\tau+2} & & & \\ \hline
K_{\tau+1,1} & \cdots & \cdots & K_{\tau+1,\tau} & K_{\tau+1,\tau+1} & & & & \\ \hline
K_{\tau+2,1} & \cdots & \cdots & K_{\tau+2,\tau} & & & & & \\
\vdots & & \iddots & & & & & & \\
\vdots & \iddots & & & & & & & \\
K_{2\tau+1,1} & & & & & & & &
\end{array}\right] \end{array}
\begin{array}{c} m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline p_\tau \\ \vdots \\ \vdots \\ p_1 \end{array}, \qquad (2.2)$$

where

(i) $p_j \geq q_j$ *and* $n_j \geq m_j$ *for* $j = 1, \ldots, \tau$,

(ii)

$$N_{j,2\tau+1-j}^{(i)} \in \mathbb{F}^{m_j, n_{j+1}}, \qquad 1 \leq j \leq \tau - 1 \quad i = 1, \ldots, k,$$

$$N_{2\tau+1-j,j}^{(i)} \in \mathbb{F}^{p_{j+1}, q_j}, \qquad 1 \leq j \leq \tau - 1 \quad i = 1, \ldots, k,$$

$$K_{j,2\tau+2-j} = \begin{bmatrix} \Sigma_j & 0 \end{bmatrix} \in \mathbb{F}^{m_j, n_j}, \quad \Sigma_j \in \mathbb{F}^{m_j, m_j}, \qquad 1 \leq j \leq \tau,$$

$$K_{2\tau+2-j,j} = \begin{bmatrix} \Gamma_j \\ 0 \end{bmatrix} \in \mathbb{F}^{p_j, q_j}, \quad \Gamma_j \in \mathbb{F}^{q_j, q_j}, \qquad 1 \leq j \leq \tau,$$

Σ_j *and* Γ_j, $j = 1, \ldots, \tau$, *are invertible and can even be chosen diagonal,*

(iii) $N_{\tau+1,\tau+1}^{(i)} = \begin{bmatrix} \tilde{N}_{11}^{(i)} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{s,t}$ *for* $i = 1, \ldots, k$, $K_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix} \in \mathbb{F}^{s,t}$, *and* $\tilde{N}_{11}^{(i)}$, $i = 1, \ldots, k$, *have no nontrivial common left or right nullspace,* $\tilde{K}_{22}$ *is either void (and* $N_{\tau+1,\tau+1}^{(i)} = \tilde{N}_{11}^{(i)}$ *in this case) or is a nonzero scalar.*

Proof. In the following we use unitary (real orthogonal) transformations to compress matrix blocks or determine the left or right nullspaces of matrices. We refrain from depicting these unitary (real orthogonal) transformations and we denote unspecified blocks by N or K.

We first determine the common left and right nullspace of N_i, $i = 1, \ldots, k$, i.e. we transform the tuple to

$$(N_k, \ldots, N_1, K) \sim \left(\begin{bmatrix} N_1^{(k)} & 0 \\ 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} N_1^{(1)} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \right),$$

then by applying the procedure for the case of matrix pairs in Corollary 2.7 of [4] (instead of N alone there, here we are applying the transformation to all N_j's simultaneously) to get

$$\sim \left(\begin{bmatrix} N_{11}^{(k)} & N_{12}^{(k)} & 0 & 0 \\ N_{21}^{(k)} & N_{22}^{(k)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} N_{11}^{(1)} & N_{12}^{(1)} & 0 & 0 \\ N_{21}^{(1)} & N_{22}^{(1)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} & K_{13} & K_{14} \\ K_{21} & K_{22} & K_{23} & 0 \\ K_{31} & K_{32} & \Sigma & 0 \\ K_{41} & 0 & 0 & 0 \end{bmatrix} \right),$$

where $K_{14} = \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} \in \mathbb{F}^{m_1,n_1}$, $K_{41} = \begin{bmatrix} \Gamma_1 \\ 0 \end{bmatrix} \in \mathbb{F}^{p_1,q_1}$, and $\Sigma_1 \in \mathbb{F}^{m_1,m_1}$ and $\Gamma_1 \in \mathbb{F}^{q_1,q_1}$ are invertible and Σ is void or a nonzero scalar. (Hence $n_1 \geq m_1$ and $p_1 \geq q_1$).

We then repeat this process recursively with the middle blocks given by

$$\left(\begin{bmatrix} N_{22}^{(k)} & 0 \\ 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} N_{22}^{(1)} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{22} & K_{23} \\ K_{32} & \Sigma \end{bmatrix} \right)$$

until the first k matrices have no nontrivial common left and right nullspaces. $\square$

In the case that the tuple has extra symmetry structure, we get a structured staircase form. We consider several different structures simultaneously. These are real and complex tuples of matrices with symmetry or skew symmetry under transposition (in the real or complex case) and conjugate transposition (in the complex case). Whenever we consider tuples, we assume that the same operation is used for all coefficients of the tuple. Our transformation matrices are either nonsingular, real orthogonal (in the real case) or unitary (in the complex transposed or conjugate transposed case). We do not consider complex orthogonal transformations.

Corollary 2.2 *If $N_i = \pm N_i^\star \in \mathbb{F}^{n,n}$, $i = 1, \ldots, k$ and $K = \pm K^\star \in \mathbb{F}^{n,n}$, then the tuple $(N_k, \ldots, N_1, K)$ is strongly u-congruent to a matrix tuple $(\hat{N}_k, \ldots, \hat{N}_1, \hat{K})$, where all terms $\hat{N}_i$, $i = 1, \ldots, k$, have the form*

$$\begin{array}{cc}
\begin{array}{cccc|c|ccccc} m_1 & \cdots & \cdots & m_\tau & s & n_\tau & \cdots & \cdots & n_1 & \end{array} & \\
\hat{N}_i = \left[\begin{array}{cccc|c|ccccc}
N^{(i)}_{11} & \cdots & \cdots & N^{(i)}_{1\tau} & N^{(i)}_{1,\tau+1} & N^{(i)}_{1,\tau+2} & \cdots & N^{(i)}_{1,2\tau} & 0 \\
\vdots & \ddots & & \vdots & \vdots & \vdots & \iddots & \iddots & \\
\vdots & & \ddots & \vdots & \vdots & N^{(i)}_{\tau-1,\tau+2} & \iddots & & \\
N^{(i)}_{\tau 1} & \cdots & \cdots & N^{(i)}_{\tau\tau} & N^{(i)}_{\tau,\tau+1} & 0 & & & \\ \hline
N^{(i)}_{\tau+1,1} & \cdots & \cdots & N^{(i)}_{\tau+1,\tau} & N^{(i)}_{\tau+1,\tau+1} & & & & \\ \hline
N^{(i)}_{\tau+2,1} & \cdots & N^{(i)}_{\tau+2,\tau-1} & 0 & & & & & \\
\vdots & \iddots & \iddots & & & & & & \\
N^{(i)}_{2\tau,1} & \iddots & & & & & & & \\
0 & & & & & & & &
\end{array} \right] &
\begin{array}{c} m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline n_\tau \\ \vdots \\ n_2 \\ n_1 \end{array}
\end{array},$$

while the matrix $\hat{K}$ has the form

$$\hat{K} = \begin{array}{c} \begin{array}{cccc|c|cccc} m_1 & \cdots & \cdots & m_\tau & s & n_\tau & \cdots & \cdots & n_1 \end{array} \\ \left[\begin{array}{cccc|c|cccc} K_{11} & \cdots & \cdots & K_{1\tau} & K_{1,\tau+1} & K_{1,\tau+2} & \cdots & \cdots & K_{1,2\tau+1} \\ \vdots & \ddots & & \vdots & \vdots & \vdots & & \iddots & \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \iddots & & \\ K_{\tau 1} & \cdots & \cdots & K_{\tau\tau} & K_{\tau,\tau+1} & K_{\tau,\tau+2} & & & \\ \hline K_{\tau+1,1} & \cdots & \cdots & K_{\tau+1,\tau} & K_{\tau+1,\tau+1} & & & & \\ \hline K_{\tau+2,1} & \cdots & \cdots & K_{\tau+2,\tau} & & & & & \\ \vdots & & \iddots & & & & & & \\ \vdots & \iddots & & & & & & & \\ K_{2\tau+1,1} & & & & & & & & \end{array}\right] \end{array} \begin{array}{c} \\ m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline n_\tau \\ \vdots \\ \vdots \\ n_1 \end{array},$$

where

(i) $n_j \geq m_j$ *for* $j = 1, \ldots, \tau$,

(ii) $N^{(i)}_{j,2\tau+1-j} \in \mathbb{F}^{m_j, n_{j+1}}, \qquad 1 \leq j \leq \tau - 1 \quad i = 1, \ldots, k,$

$$K_{j,2\tau+2-j} = \left[\begin{array}{cc} \Sigma_j & 0 \end{array}\right] \in \mathbb{F}^{m_j,n_j}, \quad \Sigma_j \in \mathbb{F}^{m_j,m_j}, \qquad 1 \leq j \leq \tau,$$

Σ_j, $j = 1, \ldots, \tau$, *are invertible and can even be chosen diagonal, and depending on the symmetry structure we have* $N^{(i)}_{2\tau+1-j,j} = \pm(N^{(i)}_{j,2\tau+1-j})^\star$, $K_{2\tau+2-j,j} = \pm(K_{j,2\tau+2-j})^\star$,

(iii) $N^{(i)}_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{N}^{(i)}_{11} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{s,s}$ *for* $i = 1, \ldots, k$, $K_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix} \in \mathbb{F}^{s,s}$, *and* $\tilde{N}^{(i)}_{11}$, $i = 1, \ldots, k$, *have no nontrivial common left or right nullspace, and depending on the symmetry structure of* K, $\tilde{K}_{22}$ *is either void (and* $N^{(i)}_{\tau+1,\tau+1} = \tilde{N}^{(i)}_{11}$ *in this case), or a nonzero scalar, or Hermitian definite, or* $i\Omega$ *with* Ω *Hermitian definite.*

Furthermore, all coefficients have retained their symmetry structure.

A condensed form for the general case is then as follows.

Theorem 2.3 *If* $A_i \in \mathbb{F}^{m,n}$ *for* $i = 0, \ldots, k$, *then the tuple* $(A_k, \ldots, A_0)$ *is strongly u-equivalent to a matrix tuple* $(\hat{A}_k, \ldots, \hat{A}_0) = (UA_kV, \ldots, UA_0V)$, *where all terms* $\hat{A}_i$, $i = 0, \ldots, k$, *have the form*

$$\begin{array}{c} \begin{array}{ccc|c|ccc} q_1 & \cdots & q_\ell & t & n_\ell & \cdots & n_1 \end{array} \\ \left[\begin{array}{ccc|c|ccc} A & \cdots & A & A & A & \cdots & A^{(i)}_{1,2\ell+1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \\ A & \cdots & A & A & A^{(i)}_{\ell,\ell+2} & & \\ \hline A & \cdots & A & A^{(i)}_{\ell+1,\ell+1} & & & \\ \hline A & \cdots & A^{(i)}_{\ell+2,\ell} & & & & \\ \vdots & \iddots & & & & & \\ A^{(i)}_{2\ell+1,1} & & & & & & \end{array}\right] \end{array} \begin{array}{c} \\ m_1 \\ \vdots \\ m_\ell \\ \hline s \\ \hline p_\ell \\ \vdots \\ p_1 \end{array}, \tag{2.3}$$

each of the blocks $A^{(i)}_{j,2\ell+2-j}$*,* $i = 0,\ldots,k,\ j = 1,\ldots,\ell$ *either has the form* $\begin{bmatrix} \Sigma_j & 0 \end{bmatrix}$ *or* $\begin{bmatrix} 0 & 0 \end{bmatrix}$*, and each of the blocks* $A^{(i)}_{2\ell+2-j,j}$*,* $i = 0,\ldots,k,\ j = 1,\ldots,\ell$ *either has the form* $\left[\begin{smallmatrix} \Gamma_j \\ 0 \end{smallmatrix}\right]$ *or* $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$*. Here* Σ_j *and* Γ_j *again denote a nonsingular (possibly diagonal) matrix of appropriate size. Furthermore, for each* j *only one of the* $A^{(i)}_{j,2\ell+2-j}$ *and one of the* $A^{(i)}_{2\ell+2-j,j}$ *are nonzero.*

All the matrices in the tuple of middle blocks $(A^{(k)}_{\ell+1,\ell+1},\ldots,A^{(0)}_{\ell+1,\ell+1})$ *are* $s \times t$*. These matrices satisfy that*

(i) either no k *matrices from the tuple have a common left and right nullspace,*

(ii) or $A^{(i)}_{\ell+1,\ell+1} = \left[\begin{smallmatrix} \tilde A^{(i)}_{11} & \tilde A^{(i)}_{12} \\ \tilde A^{(i)}_{21} & \tilde A^{(i)}_{22} \end{smallmatrix}\right]$ *for* $i = 0,1,\ldots,k$*, where for* $i_0 \in \{0,1,\ldots,k\}$*,* $\tilde A^{(i_0)}_{22}$ *is a nonzero scalar and* $\tilde A^{(i)}_{12} = 0$*,* $\tilde A^{(i)}_{21} = 0$ *and* $\tilde A^{(i)}_{22} = 0$ *for* $i \neq i_0$*, and no* k *matrices from the tuple including* $A^{(i_0)}_{\ell+1,\ell+1}$ *have a nontrivial common left and right nullspace.*

Proof. The proof follows by the following recursive procedure. First we apply Lemma 2.1 to $N_k = A_k, \ldots N_1 = A_1$ and $K = A_0$ and obtain the u-equivalent tuple of the forms (2.1) and (2.2). Then we continue with the middle block tuple, given by

$$\left(\hat A_k,\ldots,\hat A_1,\hat A_0\right) := \left(N^{(k)}_{\tau+1,\tau+1},\ldots,N^{(1)}_{\tau+1,\tau+1},K_{\tau+1,\tau+1}\right).$$

but we permute the tuple in a cyclic fashion, i.e., we apply Lemma 2.1 to $(N_k,\ldots,N_1) := (\hat A_0,\hat A_k,\ldots \hat A_2)$ and $K = \hat A_1$. We again obtain a middle block that cannot be further reduced which we take as our new tuple $\left(\hat A_k,\ldots,\hat A_1,\hat A_0\right)$. We then proceed again with the cyclically permuted tuple. In each of these steps the middle block gets smaller and we proceed until for the current middle block no cyclic permutation yields a further size reduction in the middle block. Note that in each step the part outside the middle block (the wings) grows by adding structures that have the forms (2.1) and (2.2).

The process stagnates in two cases. The first case is that no k matrices from $\left(\hat A_k,\ldots,\hat A_1,\hat A_0\right)$ have nontrivial common left and right nullspaces. The second case is that the tuple has the block form

$$\left(\hat A_k,\ldots,\hat A_1,\hat A_0\right) =: \left(\begin{bmatrix} \tilde A^{(k)}_{11} & 0 \\ 0 & 0 \end{bmatrix},\ldots,\begin{bmatrix} \tilde A^{(1)}_{11} & 0 \\ 0 & 0 \end{bmatrix},\begin{bmatrix} \tilde A^{(0)}_{11} & \tilde A^{(0)}_{12} \\ \tilde A^{(0)}_{21} & \tilde A^{(0)}_{22} \end{bmatrix}\right),$$

where $\tilde A^{(0)}_{22}$ is a nonzero scalar. Although $\hat A_k,\ldots,\hat A_1$ have a common nullspace, the procedure stops if no k matrices including $\hat A_0$ have a common nullspace. Note that $\hat A_0$ may be in any one of the matrices $A_k,\ldots,A_1,A_0$. □

For the case with symmetry structures we have the following Corollary.

Corollary 2.4 *If* $A_i = \pm A_i^\star \in \mathbb{F}^{n,n}$ *for* $i = 0,\ldots,k$*, then the tuple* $(A_k,\ldots,A_0)$ *is strongly u-congruent to a matrix tuple* $(\hat A_k,\ldots,\hat A_0) = (V^\star A_k V,\ldots,V^\star A_0 V)$*, where all terms* $\hat A_i$*,*

$i = 0, \dots, k$, *have the form*

$$
\begin{array}{c}
\begin{array}{ccc|c|ccc} m_1 & \dots & m_\ell & s & n_\ell & \dots & n_1 \end{array}\\
\left[\begin{array}{ccc|c|ccc}
A & \dots & A & A & A & \dots & A^{(i)}_{1,2\ell+1}\\
\vdots & \ddots & \vdots & \vdots & \vdots & \cdot^{\cdot^{\cdot}} & \\
A & \dots & A & A & A^{(i)}_{\ell,\ell+2} & & \\ \hline
A & \dots & A & A^{(i)}_{\ell+1,\ell+1} & & & \\ \hline
A & \dots & A^{(i)}_{\ell+2,\ell} & & & & \\
\vdots & \cdot^{\cdot^{\cdot}} & & & & & \\
A^{(i)}_{2\ell+1,1} & & & & & &
\end{array}\right]
\end{array}
\begin{array}{c} m_1\\ \vdots \\ m_\ell \\ \hline s \\ \hline n_\ell \\ \vdots \\ n_1\end{array}, \tag{2.4}
$$

and each of the blocks $A^{(i)}_{2\ell+2-j,j}$ $i = 0, \dots, k$, $j = 1, \dots, \ell$ *either has the form* $\begin{bmatrix} \Sigma_j \\ 0 \end{bmatrix}$ *or* $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, *and (depending on the symmetry structure)* $A^{(i)}_{j,2\ell+2-j} = \pm(A^{(i)}_{2\ell+2-j,j})^\star$, $i = 0, \dots, k$, $j = 1, \dots, \ell$. *Here,* Σ_j *again denotes a nonsingular (possibly diagonal) matrix of appropriate size. Furthermore, for each* j *only one of the* $A^{(i)}_{j,2\ell+2-j}$ *is nonzero.*

All the matrices in the tuple of middle blocks $(A^{(k)}_{\ell+1,\ell+1}, \dots, A^{(0)}_{\ell+1,\ell+1})$ *are* $s \times s$. *These matrices satisfy that*

(i) either no k *matrices from the tuple have a common left or right nullspace,*

(ii) or $A^{(i)}_{\ell+1,\ell+1} = \begin{bmatrix} \tilde{A}^{(i)}_{11} & \tilde{A}^{(i)}_{12} \\ \tilde{A}^{(i)}_{21} & \tilde{A}^{(i)}_{22} \end{bmatrix}$ *for* $i = 0, 1, \dots, k$, *where for* $i_0 \in \{0, 1, \dots, k\}$, *(depending on the structure of* A_{i_0}*),* $\tilde{A}^{(i_0)}_{22}$ *is Hermitian definite or* $i\Omega$ *with* Ω *Hermitian definite and* $\tilde{A}^{(i)}_{12} = 0, \tilde{A}^{(i)}_{21} = 0, \tilde{A}^{(i)}_{22} = 0$ *for* $i \neq i_0$, *and no* k *matrices including* $A^{(i_0)}_{\ell+1,\ell+1}$ *from the tuple have a nontrivial common left and right nullspace.*

Proof. The proof is exactly the same as in the general case by applying Corollary 2.2. □

One might hope that it is possible to reduce the middle block tuple further by u-equivalence (u-congruence) even if the reduction procedure stagnates. However, this is not always possible as the following example shows.

Example 2.5 Consider the two 3×4 quadratic matrix polynomials

$$
P(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},
$$

and

$$
Q(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.
$$

For both polynomials, no pair of coefficient matrices has a nontrivial common left and right nullspace.

We reduce $P(\lambda)$ and $Q(\lambda)$ to their Smith forms , see [17],

$$\begin{bmatrix} 1 & -\lambda & \lambda^2 \\ 0 & 1 & -\lambda \\ 0 & 0 & 1 \end{bmatrix} P(\lambda) \begin{bmatrix} 1+\lambda-\lambda^2 & 2-\lambda & 0 & 0 \\ -\lambda^2 & 1-\lambda & 0 & 0 \\ -\lambda+\lambda^3 & \lambda^2-\lambda & 1 & 0 \\ -\lambda+\lambda^3 & -1-\lambda+\lambda^2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \lambda & -\lambda^2 \\ 0 & 1 & -\lambda \\ 0 & 0 & 1 \end{bmatrix} Q(\lambda) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \lambda & \lambda-\lambda^2 & 1 & 0 \\ -1-\lambda & -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \lambda^2(\lambda-1) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

From the Smith forms of $P(\lambda)$ and $\operatorname{rev} P(\lambda)$ (which we do not show) we can determine that the polynomial $P(\lambda)$ has

(a) a right 3×4 singular block with a chain

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix},$$

(b) a simple eigenvalue 0 with a right eigenvector $\begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T$ and a left eigenvector $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$,

(c) a 2×2 Kronecker block associated with the eigenvalue infinity with right (left) chain vectors

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \right).$$

The polynomial $Q(\lambda)$ has

(a) a right 1×2 singular block with a chain

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix},$$

(b) a 2×2 Jordan block associated with the eigenvalue 0 with a right (left) chain

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

(c) a 2×2 Kronecker block associated with the eigenvalue infinity with a right (left) chain

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \qquad \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right),$$

(d) and a simple eigenvalue 1 with a right eigenvector $\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}^T$ and a left eigenvector $\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$.

In both cases, there seems to be no way to use further u-equivalence (u-congruence) transformations to separate the blocks related to the singular part and the eigenvalues 0 and infinity.

If no further reduction is possible by strong equivalence, then as in [23] we may employ unimodular transformations to reduce the tuple further. However, as we have pointed out, unimodular transformations change the length of chains and therefore the structure associated with the singular part and the eigenvalue infinity.

It is an open problem to determine a staircase form under u-congruence for tuples of more than 2 matrices that displays complete information associated with the singular parts and the eigenvalue infinity. There is, however, a particular situation of stagnation in Theorem 2.3 or Corollary 2.4 where the complete information is available. This is the case that (possibly after some further u-equivalence transformations), the tuple of middle blocks in the condensed form (2.3), $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(1)}_{\ell+1,\ell+1}, A^{(0)}_{\ell+1,\ell+1})$, has the form

$$\left(\begin{bmatrix} \Sigma_k & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & 0 & \ldots & 0 \\ \tilde{A}^{(k-1)}_{21} & \Sigma_{k-1} & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, \right. \tag{2.5}$$

$$\left. \begin{bmatrix} \tilde{A}^{(1)}_{11} & \ldots & \tilde{A}^{(1)}_{1,k-1} & \tilde{A}^{(1)}_{1k} & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \tilde{A}^{(1)}_{k-1,1} & \ldots & \tilde{A}^{(1)}_{k-1,k-1} & \tilde{A}^{(1)}_{k-1,k} & 0 \\ \tilde{A}^{(1)}_{k1} & \ldots & \tilde{A}^{(1)}_{k,k-1} & \Sigma_1 & 0 \\ 0 & \ldots & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(0)}_{11} & \ldots & \tilde{A}^{(0)}_{1k} & \tilde{A}^{(0)}_{1,k+1} \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{A}^{(0)}_{k1} & \ldots & \tilde{A}^{(0)}_{kk} & \tilde{A}^{(0)}_{k,k+1} \\ \tilde{A}^{(0)}_{k+1,1} & \ldots & \tilde{A}^{(0)}_{k+1,k} & \Sigma_0 \end{bmatrix} \right),$$

where $\Sigma_0, \Sigma_1, \ldots, \Sigma_k$ are all invertible. It should be noted again that it is not always possible to achieve the form (2.5) as Example 2.5 and the following example show.

Example 2.6 Consider the regular symmetric matrix polynomial

$$P(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix},$$

which has double eigenvalues at $0, \infty$ with (both right and left) Kronecker/Jordan chains

$$x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

associated with infinity and

$$z_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad z_1 = \begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$$

associated with 0. No two coefficients have a common nullspace, and the matrix polynomial is not in the form (2.5). There exist no strong equivalence (congruence) transformations that reduce the matrix polynomial further to get it to the form (2.5). On the other hand, performing unimodular transformations of multiplying from the left and right to $P(\lambda)$ with the matrices

$$\begin{bmatrix} 1 & -\lambda/2 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ -\lambda & 1/2 \end{bmatrix},$$

respectively, yields the Smith form

$$\tilde{P}(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

which is in the form (2.5) (with Σ_1 void) and even symmetric.

In general, the tuple of middle blocks in the condensed form (2.3), $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(1)}_{\ell+1,\ell+1}, A^{(0)}_{\ell+1,\ell+1})$, can only be reduced by strong u-equivalence to the form

$$\left(\begin{bmatrix} \Sigma_k & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & 0 & \ldots & 0 \\ \tilde{A}^{(k-1)}_{21} & \tilde{A}^{(k-1)}_{22} & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots \right. \tag{2.6}$$

$$\left. \begin{bmatrix} \tilde{A}^{(1)}_{11} & \ldots & \tilde{A}^{(1)}_{1,k-1} & \tilde{A}^{(1)}_{1k} & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \tilde{A}^{(1)}_{k-1,1} & \ldots & \tilde{A}^{(1)}_{k-1,k-1} & \tilde{A}^{(1)}_{k-1,k} & 0 \\ \tilde{A}^{(1)}_{k1} & \ldots & \tilde{A}^{(1)}_{k,k-1} & \tilde{A}^{(1)}_{kk} & 0 \\ 0 & \ldots & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(0)}_{11} & \ldots & \tilde{A}^{(0)}_{1k} & \tilde{A}^{(0)}_{1,k+1} \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{A}^{(0)}_{k1} & \ldots & \tilde{A}^{(0)}_{kk} & \tilde{A}^{(0)}_{k,k+1} \\ \tilde{A}^{(0)}_{k+1,1} & \ldots & \tilde{A}^{(0)}_{k+1,k} & \tilde{A}^{(0)}_{k+1,k+1} \end{bmatrix} \right).$$

This form can be obtained by a sequence of unitary equivalence transformations that exploit successively the left and right null spaces of $A^{(k)}_{\ell+1,\ell+1}$, the common left and right null spaces of $A^{(k)}_{\ell+1,\ell+1}, A^{(k-1)}_{\ell+1,\ell+1}$, and eventually $A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(1)}_{\ell+1,\ell+1}$. The matrix Σ_k is still nonsingular, but nothing can be said about other diagonal blocks.

If the original tuple has a symmetry structure as in Corollary 2.4, then the tuple of middle blocks in (2.4) can be transformed via strong u-congruence to a form as (2.6), or possibly even to the form (2.5).

It is again an open problem to characterize when a general tuple of matrices has a condensed form (2.3), where the middle block has the from (2.5). However, most of the matrix polynomials with singular leading term that are encountered in practice are second order. These are either constrained mechanical systems as in Example 1.7 or second order systems arising from optimal control or variational problems such as the palindromic matrix polynomials arising in the vibration analysis of rails [14]. Let us demonstrate this for constrained multibody systems.

Example 2.7 *Consider the matrix polynomial $P(\lambda)$ in (1.9) with M positive definite and G of full row rank. Transforming $P(\lambda)$ to the staircase form (2.3) one obtains a form*

$$\hat{P}(\lambda) = \lambda^2 \begin{bmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{21} & D_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} & G_1^T \\ K_{21} & K_{22} & 0 \\ G_1 & 0 & 0 \end{bmatrix}$$

with M_{22} positive definite and G_1 square nonsingular.

Since M_{22} is invertible, the middle block $\lambda^2 M_{22}+\lambda D_{22}+K_{22}$ is a regular matrix polynomial with only finite eigenvalues in the form (2.5).

The same is true for the palindromic example, see [14].

In order to analyze the tuple in (2.5), we introduce the *strangeness index of a matrix polynomial* analogous to the corresponding concept for a differential-algebraic equation (DAE) with the coefficients A_j, see [23]. Such a DAE has the form

$$A_k x^{(k)} + A_{k-1} x^{(k-1)} + \cdots + A_1 \dot{x} + A_0 x = f(t) \tag{2.7}$$

with some inhomogeneity f. In simple words, the strangeness-index is the highest order of the derivatives of the inhomogeneity f that has to be required so that a continuous solution x exists with the extra property that $x^{(j)}$ is defined in the range space of A_j for $j = 0, \ldots, k$. If a system has strangeness-index 0, then it is called *strangeness-free*. For more details on the strangeness-index see [15]. The strangeness index generalizes the differentiation index [3] to nonsquare and singular systems but the counting is slightly different. Ordinary differential equations as well as purely algebraic equations both are strangeness-free, while ordinary differential equations have differentiation index 0 and algebraic equations have differentiation index 1.

The construction of the staircase forms in this section can in principle be implemented as numerical methods. However, one faces the usual difficulties that already arise in the computation of staircase forms for matrix pencils. First of all it is clear that the methods depend on numerical rank decisions, see [5, 6] for detailed discussion on how to perform these decisions in the context of staircase forms where a sequence of rank decisions (which depend on each other) is needed. For matrix pairs and the analysis of differential-algebraic equations, it has been shown in [21], see also [15], that in case of doubt it is best to assume that the index is higher, i.e. to assume a longer chain. This leads to a kind of regularization procedure. The implementation of numerical methods for the computation of the staircase forms is currently under investigation.

The discussion of this section shows that partial staircase forms under strong u-equivalence (u-congruence) exist, but unfortunately these forms do not always directly display all the structural information about the eigenvalue infinity and the singular part.

3 Polynomial Eigenvalue Problems and Trimmed Linearizations

When solving a polynomial eigenvalue problem $P(\lambda)x = 0$ or $y^\star P(\lambda) = 0$, i.e. if we want to compute eigenvalues, left and right eigenvectors as well as deflating and reducing subspaces associated with the singular parts and the parts associated with the eigenvalue infinity, then we can obtain some of this information directly from the condensed form (2.3). If we partition the matrix polynomial in the form (2.3) as

$$\tilde{P}(\lambda) = \begin{bmatrix} P_{11}(\lambda) & P_{12}(\lambda) & P_{13}(\lambda) \\ P_{21}(\lambda) & P_{22}(\lambda) & 0 \\ P_{31}(\lambda) & 0 & 0 \end{bmatrix}, \tag{3.1}$$

then $P_{31}(\lambda)$ has full column rank and $P_{13}(\lambda)$ has full row rank, when considered as polynomial matrices, i.e. for some value of λ. More specifically, $P_{31}(\lambda)$ and $P_{13}(\lambda)$ are both in a block anti-diagonal form with each anti-diagonal block of $P_{31}(\lambda)$ and $P_{13}(\lambda)$ having form $\lambda^i \begin{bmatrix} \Gamma \\ 0 \end{bmatrix}$ and $\lambda^j \begin{bmatrix} \Sigma & 0 \end{bmatrix}$, respectively, for some integers i and j.

Let $x(\lambda)$ be a polynomial vector such that $P(\lambda)x(\lambda) \equiv 0$. Define $\tilde{x}(\lambda) := Vx(\lambda)$, where V is the transformation matrix from the right, and partition $\tilde{x}(\lambda) = [x_1^T(\lambda), x_2^T(\lambda), x_3^T(\lambda)]^T$ according to the partitioning of (3.1). Then

$$\begin{bmatrix} P_{11}(\lambda) & P_{12}(\lambda) & P_{13}(\lambda) \\ P_{21}(\lambda) & P_{22}(\lambda) & 0 \\ P_{31}(\lambda) & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(\lambda) \\ x_2(\lambda) \\ x_3(\lambda) \end{bmatrix} = 0$$

implies that $x_1(\lambda) \equiv 0$. So, the right singular blocks of the polynomial $P(\lambda)$ are contained in the submatrix polynomial $\begin{bmatrix} P_{12}(\lambda) & P_{13}(\lambda) \\ P_{22}(\lambda) & 0 \end{bmatrix}$. Similarly for $y(\lambda)$ satisfying $y^\star(\lambda)P(\lambda) \equiv 0$, let $\tilde{y}(\lambda) = Uy(\lambda) = [y_1^T(\lambda), y_2^T(\lambda), y_3^T(\lambda)]^T$, where U is the transformation matrix from the left. Then $\tilde{y}^\star(\lambda)\tilde{P}(\lambda) \equiv 0$ implies that $y_1(\lambda) \equiv 0$. So, the left singular blocks of $P(\lambda)$ are contained in $\begin{bmatrix} P_{21}(\lambda) & P_{22}(\lambda) \\ P_{31}(\lambda) & 0 \end{bmatrix}$.

To see where the finite nonzero eigenvalues can be found, suppose that $P(\lambda_0)x = 0$, where x is a nonzero constant vector and λ_0 is a nonzero eigenvalue. Let $\tilde{x} = Vx = [x_1^T, x_2^T, x_3^T]^T$. Then $\tilde{P}(\lambda_0)\tilde{x} = 0$, from which it follows that $P_{31}(\lambda_0)x_1 = 0$. From the block structure of $P_{31}(\lambda)$, it follows that $P_{31}(\lambda_0)$ has full column rank when $\lambda_0 \neq 0$. So, $x_1 = 0$, and hence $\begin{bmatrix} P_{12}(\lambda) & P_{13}(\lambda) \\ P_{22}(\lambda) & 0 \end{bmatrix}$ contains all the eigenvalue information about λ_0. Now let $[y_1^T, y_2^T]^T$ be nonzero and satisfy

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^T \begin{bmatrix} P_{12}(\lambda_0) & P_{13}(\lambda_0) \\ P_{22}(\lambda_0) & 0 \end{bmatrix} = 0.$$

Because $\lambda_0 \neq 0$, $P_{13}(\lambda_0)$ has full row rank. From $y_1^T P_{13}(\lambda_0) = 0$, it follows that $y_1 = 0$. We conclude that all the eigenvalue information associated with finite nonzero eigenvalues is contained in $P_{22}(\lambda)$.

Note that this does not apply to the eigenvalues 0 and infinity. For the eigenvalue 0, $P_{31}(0)$ may not be full column rank unless all the full column rank blocks $\begin{bmatrix} \Gamma_j \\ 0 \end{bmatrix}$ appear in $A^{(0)}_{2\ell+2-j,j}$ $(j = 1, \ldots, \ell)$ in (2.3). This is clearly not always the case.

If the middle block tuple $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(0)}_{\ell+1,\ell+1})$ in the staircase form (2.3) can be reduced further to (2.5), then we can determine the information about the eigenstructure associated with the nonzero finite eigenvalues from this block as follows.

Assume that $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(0)}_{\ell+1,\ell+1})$ is in the form (2.5). Consider the eigenvalue problem $P_{22}(\lambda)\tilde{x} = 0$ with $\tilde{x} = [x_0^T, x_1^T, \ldots, x_k^T]^T$. Then we can turn this into a linear eigenvalue problem by introducing selected new variables (which are different from the usual companion form construction). Let

$$
\begin{aligned}
&z_{0,1} = \lambda x_0, \quad z_{0,2} = \lambda z_{0,1} = \lambda^2 x_0, \ldots, z_{0,k-1} = \lambda z_{0,k-2} = \lambda^{k-1} x_0, \\
&z_{1,1} = \lambda x_1, \quad z_{1,2} = \lambda z_{1,1} = \lambda^2 x_1, \ldots, z_{1,k-2} = \lambda z_{1,k-3} = \lambda^{k-2} x_1, \\
&\vdots \\
&z_{k-2,1} = \lambda x_{k-2}.
\end{aligned}
$$

Define

$$
z = [x_0^T, x_1^T, \ldots, x_k^T, z_{0,1}^T, \ldots, z_{k-2,1}^T, z_{0,2}^T, \ldots, z_{k-3,2}^T, \ldots, z_{0,k-2}^T, z_{1,k-2}^T, z_{0,k-1}^T]^T.
$$

It can be easily verified that z satisfies $L_t(\lambda)z = 0$, with

$$
\begin{aligned}
&L_t(\lambda) = \lambda K_t + N_t = \\
&\lambda \left[\begin{array}{ccccc|cccc|c|cc|c}
\tilde{A}^{(1)}_{11} & \cdots & \tilde{A}^{(1)}_{1,k-1} & \tilde{A}^{(1)}_{1k} & 0 & \tilde{A}^{(2)}_{11} & \cdots & \tilde{A}^{(2)}_{1,k-2} & \tilde{A}^{(2)}_{1,k-1} & & \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & \Sigma_k \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \tilde{A}^{(k-1)}_{21} & \Sigma_{k-1} & 0 \\
\tilde{A}^{(1)}_{k-2,1} & \cdots & \tilde{A}^{(1)}_{k-2,k-1} & \tilde{A}^{(1)}_{k-2,k} & 0 & \tilde{A}^{(2)}_{k-2,1} & \cdots & \tilde{A}^{(2)}_{k-2,k-2} & \tilde{A}^{(2)}_{k-2,k-1} & \cdots & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k-1,1} & \cdots & \tilde{A}^{(1)}_{k-1,k-1} & \tilde{A}^{(1)}_{k-1,k} & 0 & \tilde{A}^{(2)}_{k-1,1} & \cdots & \tilde{A}^{(2)}_{k-1,k-2} & \Sigma_2 & & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k1} & \cdots & \tilde{A}^{(1)}_{k-1,1} & \Sigma_1 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\ \hline
I & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & I & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\ \hline
0 & \cdots & \cdots & 0 & 0 & I & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & I & 0 & & 0 & 0 & 0 \\ \hline
 & & \vdots & & & & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & I & 0 & 0
\end{array}\right] \\
&+ \left[\begin{array}{ccccc|cccc|c|cc|c}
\tilde{A}^{(0)}_{11} & \cdots & \cdots & \tilde{A}^{(0)}_{1k} & \tilde{A}^{(0)}_{1,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \cdots & \vdots & \vdots & \vdots \\
\tilde{A}^{(0)}_{k1} & \cdots & \cdots & \tilde{A}^{(0)}_{kk} & \tilde{A}^{(0)}_{k,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\tilde{A}^{(0)}_{k+1,1} & \cdots & \cdots & \tilde{A}^{(0)}_{k+1,k} & \Sigma_0 & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\ \hline
0 & 0 & \cdots & \cdots & 0 & -I & 0 & \cdots & 0 & & 0 & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & -I & \ddots & \vdots & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 & & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & -I & & 0 & 0 & 0 \\ \hline
 & & \vdots & & & & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & -I & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & -I & 0 \\ \hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 & -I
\end{array}\right].
\end{aligned}
\tag{3.2}
$$

We will now analyze this pencil.

Lemma 3.1 *The pencil $L_t(\lambda) = \lambda K_t + N_t$ in (3.2) is regular and strangeness-free.*

Proof. The off-diagonal blocks in the last row and column of the $(1,1)$ block of N_t can be annihilated by Gaussian elimination with pivot Σ_0. The only blocks that have been changed after the elimination are the remaining blocks (but not Σ_0) in the $(1,1)$ block of N_t. Now, if we delete the last row and column in the first big block of the new $L_t(\lambda)$ (which correspond to the eigenvalue infinity, since Σ_0 is invertible), then we obtain from the positions of Σ_j and

I blocks that the remaining matrix of K_t is nonsingular. This implies, see e.g. [22], that the pencil L_t has no Kronecker blocks associated with the eigenvalue infinity of size bigger than 1, i.e. is strangeness-free. □

Corollary 3.2 *If $P_{22}(\lambda)$ is a matrix polynomial with coefficient matrices in the form (2.5), then the linear pencil $L_t(\lambda)$ in (3.2) is a linearization according to Definition 1.1.*

Proof. Based on the block structure of $L_t(\lambda)$, it is not difficult to annihilate its off-diagonal blocks (subdiagonal blocks first and then the blocks on the first row) by multiplying two unimodular matrix polynomials $E(\lambda)$, $F(\lambda)$ from the left and right, respectively, resulting in

$$E(\lambda)L_t(\lambda)F(\lambda) = \begin{bmatrix} P_{22}(\lambda) & 0 \\ 0 & I \end{bmatrix}.$$

□

Definition 3.3 *For regular strangeness-free matrix polynomials with the coefficient matrices of the form (2.5), the linearization (3.2) is called* trimmed linearization*.*

This terminology is motivated by the fact that in contrast to classical companion like linearizations we have "trimmed" all the chains of the eigenvalue infinity from $\Sigma_1, \ldots, \Sigma_{k-1}$ except for those chains corresponding to Σ_0.

Example 3.4 *Consider the matrix polynomial*

$$P(\lambda) = \lambda^2 \begin{bmatrix} I_{n_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} B_{11} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & I_{n_3} \end{bmatrix}$$

with coefficient matrices already in the form (2.5). The trimmed linearization is

$$L_t(\lambda) = \lambda K_t + N_t = \lambda \left[\begin{array}{ccc|c} B_{11} & 0 & 0 & I_{n_1} \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline I_{n_1} & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|c} C_{11} & C_{12} & 0 & 0 \\ C_{21} & C_{22} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ \hline 0 & 0 & 0 & -I_{n_1} \end{array}\right],$$

which, by interchanging the last two rows and columns, is equivalent to

$$\lambda \left[\begin{array}{ccc|c} B_{11} & 0 & I_{n_1} & 0 \\ 0 & I_{n_2} & 0 & 0 \\ I_{n_1} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|c} C_{11} & C_{12} & 0 & 0 \\ C_{21} & C_{22} & 0 & 0 \\ 0 & 0 & -I_{n_1} & 0 \\ \hline 0 & 0 & 0 & I_{n_3} \end{array}\right].$$

It is easily seen that $\lambda K_t + N_t$ has n_3 chains associated with the eigenvalue infinity of length 1.

In contrast to this, the companion linearization

$$\begin{aligned} \tilde{L}(\lambda) &= \lambda \tilde{K} + \tilde{N} \\ &= \lambda \left[\begin{array}{ccc|ccc} B_{11} & 0 & 0 & I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline I_{n_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{ccc|ccc} C_{11} & C_{12} & 0 & 0 & 0 & 0 \\ C_{21} & C_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -I_{n_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & -I_{n_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & -I_{n_3} \end{array}\right], \end{aligned}$$

is equivalent to

$$\lambda \left[\begin{array}{ccc|c|cc} B_{11} & 0 & I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 & 0 & 0 \\ I_{n_1} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_{n_2} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{n_3} & 0 \end{array}\right] + \left[\begin{array}{ccc|c|cc} C_{11} & C_{12} & 0 & 0 & 0 & 0 \\ C_{21} & C_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & -I_{n_1} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -I_{n_2} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{n_3} & 0 \\ 0 & 0 & 0 & 0 & 0 & -I_{n_3} \end{array}\right].$$

We see that $\lambda\tilde{K} + \tilde{N}$ has n_3 and n_2 chains associated with the eigenvalue infinity of length 2 *and* 1*, respectively.*

If in (3.1) we have $P_{13}(\lambda) \in \mathbb{F}^{r_1,s_1}$ and $r_1 < s_1$, then the matrix polynomial consists of more than a regular strangeness-free part and singular parts and hence there exists further structural invariants associated with the eigenvalue infinity. A similar argument holds for left chains associated with the eigenvalue infinity if $P_{31}(\lambda) \in \mathbb{F}^{r_2,s_2}$ and $r_2 > s_2$. The staircase form, however, does not directly display these further invariants.

In the case of matrix pencils, these invariants, which are the length of the Kronecker chains associated with the eigenvalue infinity and the singular parts, can be read off from the staircase form. If we are interested either only in left or right eigenvectors associated with infinity in $P_{22}(\lambda)$, then the strangeness-free part associated with the eigenvalue infinity can be further reduced by unitary (real orthogonal) transformation from the left (or right). For this consider the matrix polynomial (2.5) in the middle block of (2.3). If we want to split the part associated with the eigenvalue infinity from that associated with the finite eigenvalues with unitary (real orthogonal) strong equivalence transformations, then in general it is not possible to eliminate all blocks above and to the left of the block Σ_0. But with QR factorizations we can eliminate either above or to the left of Σ_0 loosing, however, the structure in the other blocks. This allows deflation of either the left or the right deflating subspace associated with the eigenvalue infinity and reduces the matrix polynomial to one that has only finite eigenvalues. According to [26] it is reasonable to leave this part to the QZ algorithm applied to the trimmed linearization.

It is obvious that the results for the infinite eigenvalue can immediately be transferred to the eigenvalue 0 by considering the reverse polynomial. Thus, we expect that for the eigenvalue 0 the classical companion linearizations may create unnecessary long Jordan chains, and using shifts for that matter also for any other eigenvalue. However, for each finite eigenvalue a different shift and hence also a different trimmed linearization needs to be considered. The procedure to do this is obvious, so we do not present it here.

As mentioned in the last section, the tuple corresponding to $P_{22}(\lambda)$ is not always strongly u-equivalent to (2.5), but always to (2.6). Similarly, using (2.6) one can construct the linear pencil

$$
\tilde{L}_t(\lambda) = \lambda \tilde{K}_t + \tilde{N}_t =
$$

$$
\lambda \left[\begin{array}{ccccc|cccc|c|cc|c}
\tilde{A}^{(1)}_{11} & \cdots & \tilde{A}^{(1)}_{1,k-1} & \tilde{A}^{(1)}_{1k} & 0 & \tilde{A}^{(2)}_{11} & \cdots & \tilde{A}^{(2)}_{1,k-2} & \tilde{A}^{(2)}_{1,k-1} & & \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & \Sigma_k \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \tilde{A}^{(k-1)}_{21} & \tilde{A}^{(k-1)}_{22} & 0 \\
\tilde{A}^{(1)}_{k-2,1} & \cdots & \tilde{A}^{(1)}_{k-2,k-1} & \tilde{A}^{(1)}_{k-2,k} & 0 & \tilde{A}^{(2)}_{k-2,1} & \cdots & \tilde{A}^{(2)}_{k-2,k-2} & \tilde{A}^{(2)}_{k-2,k-1} & \cdots & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k-1,1} & \cdots & \tilde{A}^{(1)}_{k-1,k-1} & \tilde{A}^{(1)}_{k-1,k} & 0 & \tilde{A}^{(2)}_{k-1,1} & \cdots & \tilde{A}^{(2)}_{k-1,k-2} & \tilde{A}^{(2)}_{k-1,k-1} & & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k1} & \cdots & \tilde{A}^{(1)}_{k-1,1} & \tilde{A}^{(1)}_{kk} & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\ \hline
I & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & I & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\ \hline
0 & \cdots & \cdots & 0 & 0 & I & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & I & 0 & & 0 & 0 & 0 \\ \hline
 & & \vdots & & & & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & I & 0 & 0
\end{array}\right]
$$

$$
+ \left[\begin{array}{ccccc|cccc|c|cc|c}
\tilde{A}^{(0)}_{11} & \cdots & \cdots & \tilde{A}^{(0)}_{1k} & \tilde{A}^{(0)}_{1,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \cdots & \vdots & \vdots & \vdots \\
\tilde{A}^{(0)}_{k1} & \cdots & \cdots & \tilde{A}^{(0)}_{kk} & \tilde{A}^{(0)}_{k,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\tilde{A}^{(0)}_{k+1,1} & \cdots & \cdots & \tilde{A}^{(0)}_{k+1,k} & \tilde{A}^{(0)}_{k+1,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\ \hline
0 & 0 & \cdots & \cdots & 0 & -I & 0 & \cdots & 0 & & 0 & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & -I & \ddots & \vdots & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 & & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & -I & & 0 & 0 & 0 \\ \hline
 & & \vdots & & & & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & -I & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & -I & 0 \\ \hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 & -I
\end{array}\right].
$$

The pencil $\tilde{L}_t(\lambda)$ may neither be regular nor strangeness-free. However, one can still show that the linear pencil $\tilde{L}_t(\lambda)$ is a linearization of P_{22} according to Definition 1.1.

We can summarize the procedure that we have described as follows. The staircase form allows partial deflation (and in a special case all) of the singular parts and parts associated with the eigenvalue infinity directly on the matrix polynomial without first performing a linearization. If the resulting middle block is regular and strangeness-free then so is the trimmed linearization.

Since companion linearizations may increase the length of chains associated with the singular part and the eigenvalue infinity, the numerical computation of the corresponding subspaces becomes more ill-conditioned in the classical companion linearization than in the trimmed linearization.

4 Structured linearizations for structured matrix polynomials

If the matrix polynomial under consideration is structured, then we would prefer the staircase form and the trimmed linearization to retain this structure. As we have seen in Corollary 2.4 such a staircase form can be obtained by using strong congruence transformations. So, if the matrix polynomials has all coefficients symmetric (Hermitian), or it has all coefficients skew-symmetric (skew-Hermitian) or if is an even or odd matrix polynomial (which means that the coefficients alternate between symmetric (Hermitian) and skew-symmetric (skew Hermitian)), see [18], then this structure is preserved. Thus, as we have described for a general matrix polynomial, part (or all) of the singular blocks and part (or all) of the chains associated with the eigenvalue infinity can be deflated in a structured way.

In the ideal case the tuple of middle blocks has the form (2.5). If we apply the trimmed linearization to this middle block, however, typically the structure is not preserved. On the other hand some of the structured preserving linearizations derived in [18, 19] cannot be used for (2.5) if it has the eigenvalue infinity, see [18]. We thus have to find structure preserving trimmed linearizations. For this we modify the vector spaces of linearizations that were derived in [19]. These spaces $\mathbb{L}_1(P)$, and $\mathbb{L}_2(P)$ consist of pencils that generalize the classical companion forms and are given by

$$\mathbb{L}_1(P) = \Big\{L(\lambda) = \lambda X + Y : L(\lambda)\cdot(\Lambda \otimes I_n) = v \otimes P(\lambda),\ v \in \mathbb{F}^k\Big\}, \tag{4.1}$$

$$\mathbb{L}_2(P) = \Big\{L(\lambda) = \lambda X + Y : \big(\Lambda^T \otimes I_n\big)\cdot L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{F}^k\Big\} \tag{4.2}$$

where $\Lambda = \begin{bmatrix} \lambda^{k-1} & \lambda^{k-2} & \dots & \lambda & 1 \end{bmatrix}^T$ and $\otimes$ denotes the Kronecker product. The intersection of these spaces is $\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P)$. The vector v in (4.1) is called the *right ansatz vector* of $L(\lambda) \in \mathbb{L}_1(P)$, and the vector w in (4.2) is called the *left ansatz vector* of $L(\lambda) \in \mathbb{L}_2(P)$. For $\mathbb{DL}(P)$ we need that the left and right ansatz vectors are equal, i.e., $v = w$.

It was also shown in [18] how to easily construct structured linear pencils using the *column-shifted sum* and *row-shifted sum*:

Definition 4.1 (Shifted sums) *Let $X = [X_{ij}]$ and $Y = [Y_{ij}]$ be block $k \times k$ matrices in $\mathbb{F}^{kn\times kn}$ with blocks $X_{ij}, Y_{ij} \in \mathbb{F}^{n\times n}$. Then the* column-shifted sum $X \boxplus\!\!\rightarrow Y$, *and* row-shifted sum $X \boxplus\!\!\downarrow Y$ *of X and Y are defined to be*

$$X \boxplus\!\!\rightarrow Y := \begin{bmatrix} X_{11} & \cdots & X_{1k} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ X_{k1} & \cdots & X_{kk} & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y_{11} & \cdots & Y_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & Y_{k1} & \cdots & Y_{kk} \end{bmatrix} \in \mathbb{F}^{kn\times k(n+1)},$$

$$X \boxplus\!\!\downarrow Y := \begin{bmatrix} X_{11} & \cdots & X_{1k} \\ \vdots & \ddots & \vdots \\ X_{k1} & \cdots & X_{kk} \\ 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 \\ Y_{11} & \cdots & Y_{1k} \\ \vdots & \ddots & \vdots \\ Y_{k1} & \cdots & Y_{kk} \end{bmatrix} \in \mathbb{F}^{k(n+1)\times kn}.$$

With $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$, and $L(\lambda) = \lambda X + Y$, it follows that $L(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector v iff $X \boxplus\!\!\rightarrow Y = v \otimes [A_k\ A_{k-1}\ \cdots\ A_0]$ and $L(\lambda) \in \mathbb{L}_2(P)$ with left ansatz vector w iff $X \boxplus\!\!\downarrow Y = w^T \otimes \begin{bmatrix} A_k \\ \vdots \\ A_0 \end{bmatrix}$.

For the described symmetry structures, in [18] then classifications of vectors that lead to structured linear pencils with the same structure, have been derived and it has been shown that structured linear pencils in $\mathbb{DL}(P)$ that are constructed in this way are linearizations iff the polynomial

$$\mathsf{p}(x\,;v) := v_1 x^{k-1} + v_2 x^{k-2} + \cdots + v_{k-1}x + v_k$$

that is constructed from the ansatz vector v has no root that coincides with an eigenvalue of the matrix polynomial. It has also been shown that there exist linear pencils in $\mathbb{DL}(P)$ which are structured linearizations for any of the discussed symmetry structures iff not both 0 and infinity are eigenvalues of $P(\lambda)$.

In the following we will discuss the difficulties that arise if infinity is an eigenvalue of a matrix polynomial

$$P(\lambda) = \lambda^k A_k + \lambda^{k-1} A_{k-1} + \ldots + \lambda A_1 + A_0,$$

with $(A_k, \ldots, A_1, A_0)$ in the form (2.5) and how to obtain structured trimmed linearizations in this case. Analogous constructions can be made for other structured linearizations. An ansatz vector v that leads to an easily constructed pencil in $\mathbb{DL}(P)$ is $v = e_k$, but if $P(\lambda)$ has the eigenvalue infinity then this is not a linearization. In general, if the leading coefficient matrix A_k is singular, then $P(\lambda)$ still has infinity as an eigenvalue, and thus the approach in [18] cannot be used. Let us nevertheless formally construct the resulting structured pencils via the shifted sum approach.

If all the coefficient matrices of $P(\lambda)$ satisfy $A_j = A_j^\star$ or $A_j = -A_j^\star$, $j = 0, 1, \ldots, k$, then the formally constructed linear pencil has the form

$$\lambda X_s + Y_s = \lambda \begin{bmatrix} & & & & A_k \\ & & & A_k & A_{k-1} \\ & & \iddots & \iddots & \vdots \\ & A_k & \iddots & \iddots & A_2 \\ A_k & A_{k-1} & \ldots & A_2 & A_1 \end{bmatrix} + \begin{bmatrix} & & & -A_k & 0 \\ & & \iddots & \vdots & \vdots \\ & \iddots & \iddots & -A_3 & 0 \\ -A_k & \ldots & -A_3 & -A_2 & 0 \\ 0 & \ldots & 0 & 0 & A_0 \end{bmatrix}. \tag{4.3}$$

If $P(\lambda)$ is $\star$-even or $\star$-odd, i.e., $P(\lambda) = P(-\lambda)^\star$ or $P(\lambda) = -P(-\lambda)^\star$, respectively, and $\varPi_k = \mathrm{diag}((-1)^{k-1} I_n, \ldots, (-1)^0 I_n)$ then the formally constructed linear pencil has the form $\lambda X_e + Y_e = \lambda \varPi_k X_s + \varPi_k Y_s$ with X_s, Y_s as in (4.3).

Example 4.2 Consider the ansatz vector $v = e_3$ for $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda C + D$, with one of the properties

i) $A = A^\star$, $B = B^\star$, $C = C^\star$, $D = D^\star$, or $A = -A^\star$, $B = -B^\star$, $C = -C^\star$, $D = -D^\star$,

ii) $A = A^\star$, $B = -B^\star$, $C = C^\star$, $D = -D^\star$, i.e. $P(\lambda) = -P(-\lambda)^\star$ is odd or $A = -A^\star$, $B = B^\star$, $C = -C^\star$, $D = D^\star$, i.e. $P(\lambda) = P(-\lambda)^\star$ is even.

Then the resulting linear pencils in $\mathbb{DL}(P)$ have the structures

$$\text{i)}\ \lambda \begin{bmatrix} 0 & 0 & A \\ 0 & A & B \\ A & B & C \end{bmatrix} + \begin{bmatrix} 0 & -A & 0 \\ -A & -B & 0 \\ 0 & 0 & D \end{bmatrix},$$

$$\text{ii)}\ \lambda \begin{bmatrix} 0 & 0 & A \\ 0 & -A & -B \\ A & B & C \end{bmatrix} + \begin{bmatrix} 0 & -A & 0 \\ A & B & 0 \\ 0 & 0 & D \end{bmatrix},$$

respectively.

If the matrix polynomial has any of the described symmetry structures and is in the form (2.5), then it is obvious that $\lambda X_s + Y_s$ as in (4.3) or in the odd/even case $\lambda X_e + Y_e$ are singular, since in the first $k-1$ block rows both X_s and Y_s have the same zero block rows. Now let $j_1, \ldots, j_k$ be the sizes of the invertible blocks $\varSigma_1, \ldots, \varSigma_k$ in (2.5).

In order to obtain a trimmed linearization, we delete all these zero block rows and columns. Define $\mathcal{I}_p = \begin{bmatrix} I_p \\ 0 \end{bmatrix} \in \mathbb{F}^{n,p}$, and

$$S = \mathrm{diag}(\mathcal{I}_{j_k},\ \mathcal{I}_{j_k+j_{k-1}}, \dots, \mathcal{I}_{j_k+\dots+j_2}, I_n),$$

we obtain the trimmed linear pencils

$$\lambda \hat{X}_s + \hat{Y}_s = \lambda S^T X_s S + S^T Y_s S$$

in the symmetric/skew symmetric/Hermitian/Skew Hermitian case or

$$\lambda \hat{X}_e + \hat{Y}_e = \lambda S^T X_e S + S^T Y_e S$$

in the odd/even case.

Theorem 4.3 *Consider a matrix polynomial (1.1) with the coefficient matrices of the form (2.5). Then the trimmed linear pencils $\lambda \hat{X}_s + \hat{Y}_s$, and $\lambda \hat{X}_e + \hat{Y}_e$, respectively, are linearizations according to Definition 1.1.*

Proof. We only consider the pencil $\lambda \hat{X}_s + \hat{Y}_s$. The results for $\lambda \hat{X}_e + \hat{Y}_e$ can be proved in the same way.

By multiplication with the unimodular matrix

$$U(\lambda) = \left[\begin{array}{c|c|c|c|c|c}
I_{j_k} & 0 & \dots & \dots & \dots & 0 \\ \hline
\begin{matrix} \lambda I_{j_k} \\ 0 \end{matrix} & I_{j_k+j_{k-1}} & 0 & \dots & \dots & 0 \\ \hline
\begin{matrix} \lambda^2 I_{j_k} \\ 0 \end{matrix} & \begin{matrix} \lambda I_{j_k+j_{k-1}} \\ 0 \end{matrix} & I_{j_k+j_{k-1}+j_{k-2}} & \ddots & \ddots & \vdots \\ \hline
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \hline
\vdots & \ddots & \ddots & \begin{matrix} \lambda I_{j_k+\dots+j_3} \\ 0 \end{matrix} & I_{j_k+\dots+j_2} & 0 \\ \hline
\begin{matrix} \lambda^{k-1} I_{j_k} \\ 0 \end{matrix} & \dots & \dots & \begin{matrix} \lambda^2 I_{j_k+\dots+j_3} \\ 0 \end{matrix} & \begin{matrix} \lambda I_{j_k+\dots+j_2} \\ 0 \end{matrix} & I_n
\end{array}\right]$$

we obtain that

$$U(\lambda)(\lambda \hat{X}_s + \hat{Y}_s) = \begin{bmatrix} W_{11} & W_{12}(\lambda) \\ 0 & P(\lambda) \end{bmatrix},$$

where

$$W_{11} = -\hat{S}^T \begin{bmatrix} & & & A_k \\ & & \cdot^{\cdot^{\cdot}} & A_{k-1} \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ A_k & A_{k-1} & \dots & A_2 \end{bmatrix} \hat{S}$$

with $\hat{S} = \mathrm{diag}(\mathcal{I}_{j_k},\ \mathcal{I}_{j_k+j_{k-1}}, \dots, \mathcal{I}_{j_k+\dots+j_2})$, which is constant and invertible. Multiplying with

$$V(\lambda) = \begin{bmatrix} W_{11}^{-1} & -W_{11}^{-1} W_{12}(\lambda) \\ 0 & I_n \end{bmatrix}$$

from the right and performing a block permutation that moves $P(\lambda)$ to the leading diagonal block finishes the proof. $\square$

Example 4.4 Consider $P(\lambda) = \lambda^3 A + \lambda^2 B + \lambda C + D$ with $A = A^\star$, $B = B^\star$, $C = C^\star$, $D = D^\star$, in the form (2.5), i.e.

$$A = \begin{bmatrix} \Sigma_A & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \; B = \begin{bmatrix} B_{11} & B_{12} & & \\ B_{21} & \Sigma_B & & \\ & & 0 & \\ & & & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 \\ C_{21} & C_{22} & C_{23} & 0 \\ C_{31} & C_{32} & \Sigma_C & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; D = \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \\ D_{41} & D_{42} & D_{43} & \Sigma_D \end{bmatrix}.$$

Then the structured trimmed linearization is

$$\begin{aligned} \lambda \hat{X}_s + \hat{Y}_s &= \\ &\lambda \left[\begin{array}{c|cc|cccc} 0 & 0 & 0 & \Sigma_A & 0 & 0 & 0 \\ \hline 0 & \Sigma_A & 0 & B_{11} & B_{12} & 0 & 0 \\ 0 & 0 & 0 & B_{21} & \Sigma_B & 0 & 0 \\ \hline \Sigma_A & B_{11} & B_{12} & C_{11} & C_{12} & C_{13} & 0 \\ 0 & B_{21} & \Sigma_B & C_{21} & C_{22} & C_{23} & 0 \\ 0 & 0 & 0 & C_{31} & C_{32} & \Sigma_C & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \\ &+ \left[\begin{array}{c|cc|cccc} 0 & -\Sigma_A & 0 & 0 & 0 & 0 & 0 \\ \hline -\Sigma_A & -B_{11} & -B_{12} & 0 & 0 & 0 & 0 \\ 0 & -B_{21} & -\Sigma_B & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & D_{11} & D_{12} & D_{13} & D_{14} \\ 0 & 0 & 0 & D_{21} & D_{22} & D_{23} & D_{24} \\ 0 & 0 & 0 & D_{31} & D_{32} & D_{33} & D_{34} \\ 0 & 0 & 0 & D_{41} & D_{42} & D_{43} & \Sigma_D \end{array}\right], \end{aligned}$$

which is obviously strangeness-free. Multiplying with the unimodular matrix

$$U(\lambda) = \left[\begin{array}{c|cc|cccc} I & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \lambda I & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 \\ \hline \lambda^2 I & \lambda I & 0 & I & 0 & 0 & 0 \\ 0 & 0 & \lambda I & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I \end{array}\right]$$

from the left we obtain

$$U(\lambda)(\lambda \hat{X}_s + \hat{Y}_s) = \left[\begin{array}{c|c} W_{11} & W_{12}(\lambda) \\ \hline 0 & P(\lambda) \end{array}\right] = \left[\begin{array}{ccc|c} 0 & -\Sigma_A & 0 & \\ -\Sigma_A & -B_{11} & -B_{12} & W_{12}(\lambda) \\ 0 & -B_{21} & -\Sigma_B & \\ \hline & 0 & & P(\lambda) \end{array}\right].$$

It is easily verified that W_{11} is nonsingular by eliminating the blocks $-B_{11}, -B_{12}, -B_{21}$. So $\lambda \hat{X}_s + \hat{Y}_s$ can be eventually turned into

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I \end{bmatrix}.$$

In this section we have shown how to obtain structured trimmed linearization in the important special case that the middle block has the form (2.5).

For palindromic matrix polynomials, see [18], using Cayley transformation to obtain an even/odd matrix polynomial, applying the described procedure and then inverting the Cayley transformation, a similar procedure can be used to deal with eigenvalues $-1, 1$.

In the general case when the tuple of the coefficient matrices is in the form (2.6), a trimmed linear pencil like $\lambda \hat{X}_s + \hat{Y}_s$ or $\lambda \hat{X}_e + \hat{Y}_e$ can be obtained in the same way. Unfortunately, this is a linearization of $P(\lambda)$ only when $k = 2$, i.e., if $P(\lambda)$ is a quadratic polynomial. In this case,

$$A = \begin{bmatrix} \varSigma_A & & \\ & 0 & \\ & & 0 \end{bmatrix}, \; B = \begin{bmatrix} B_{11} & B_{12} & \\ B_{21} & B_{22} & \\ & & 0 \end{bmatrix}, \; C = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}.$$

For the associated linear pencil $\lambda \hat{X}_s + \hat{Y}_s$, the corresponding matrix W_{11} is $\varSigma_A$ which is nonsingular. If $k = 3$ as in Example 4.4, then

$$W_{11} = \begin{bmatrix} 0 & -\varSigma_A & 0 \\ -\varSigma_A & -B_{11} & -B_{12} \\ 0 & -B_{21} & -B_{22} \end{bmatrix},$$

which is nonsingular only if B_{22} is.

5 Conclusion

We have presented staircase forms for matrix tuples under unitary (real orthogonal) equivalence transformations that display some (but not necessary all) of the structural information associated with the singular parts and the eigenvalue infinity. We have shown how this information may be used to obtain new types of trimmed linearizations that do not create unnecessary long Kronecker chains. We have also shown how these deflations and linearizations can be performed in a structure preserving way. We have mainly dealt with the eigenvalue infinity and the singular part. Using spectral transformations, similar procedures can be derived for any finite eigenvalue, leading to a different staircase like form for each eigenvalue. How to combine staircase forms for several eigenvalues at a time is currently under investigation.

References

[1] E. N. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electr. Journ. Lin. Alg.*, 11:78–87, 2004.

[2] E. N. Antoniou and S. Vologiannidis. Linearizations of polynomial matrices with symmetries and their applications. *Electr. Journ. Lin. Alg.*, 15:107–114, 2006.

[3] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations.* SIAM Publications, Philadelphia, PA, 2nd edition, 1996.

[4] R. Byers, V. Mehrmann, and H. Xu. Staircase forms and trimmed linerizations for structured matrix polynomials. Technical Report 395, DFG Research Center Matheon,

Mathematics for key technologies in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2007. url: http://www.matheon.de/ .

[5] J. W. Demmel and B. Kågström. Stably computing the Kronecker structure and reducing subspaces of singular pencils $A - \lambda B$ for uncertain data. In J. Cullum and R.A. Willoughby, editors, *Large Scale Eigenvalue Problems*, pages 283–323. Elsevier, North-Holland, 1986.

[6] J. W. Demmel and B. Kågström. Computing stable eigendecompositions of matrix pencils. *Linear Algebra Appl.*, 88:139–186, 1987.

[7] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Systems.* B. G. Teubner Stuttgart, 1998.

[8] F. R. Gantmacher. *Theory of Matrices. Vol. 1.* Chelsea, New York, 1959.

[9] F. R. Gantmacher. *Theory of Matrices. Vol. 2.* Chelsea, New York, 1959.

[10] I. Gohberg, M. A. Kaashoek, and P. Lancaster. General theory of regular matrix polynomials and band Toeplitz operators. *Integr. Eq. Operator Theory*, 11:776–882, 1988.

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

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

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

[14] A. Hilliges, C. Mehl, and V. Mehrmann. 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. CD-ROM.

[15] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution.* EMS Publishing House, Zürich, Switzerland, 2006.

[16] P. Lancaster. *Lambda-matrices and Vibrating Systems.* Pergamon Press, Oxford, 1966.

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

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

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

[20] D. Steven Mackey. *Structured Linerizations for matrix polynomials.* PhD thesis, Department of Mathematics, University of Manchester, Manchester, GB, 2006.

[21] R. M. M. Mattheij and P. M. E. J. Wijckmans. Sensitivity of solutions of linear dae to perturbations of the system matrices. *Numer. Algorithms*, 19:159–171, 1998.

[22] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[23] V. Mehrmann and C. Shi. Transformation of high order linear differential-algebraic systems to first order. *Numer. Alg.*, 42:281–307, 2006.

[24] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43:235–286, 2001.

[25] P. Van Dooren. The computation of Kronecker's canonical form of a singular pencil. *Linear Algebra Appl.*, 27:103–121, 1979.

[26] D. S. Watkins. Performance of the QZ algorithm in the presence of infinite eigenvalues. *SIAM J. Matrix Anal. Appl.*, 22:364–375, 2000.