Staircase forms and trimmed linearizations for structured matrix polynomials

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

May 25, 2007

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 staircase condensed forms that allow 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 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 construction of staircase forms and the reformulation of matrix polynomials as first degree (linear) matrix polynomials of larger dimension which expose spectral information including eigenvalues, eigenvectors, Jordan-Kronecker chains, and singular chains (if any).

[‡]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.

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*

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

(A matrix pencil $E(\lambda)$ 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. [12, 19], 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. Also, linear pencils have a Kronecker canonical form that displays properties like index at infinity, Jordan structure and singular structure–information that identifies the behavior and pathologies of associated linear, constant coefficient differential algebraic equations. (Even for linear, constant coefficient differential algebraic equations there is more to the story than the Kronecker canonical form. See for example [3, 17].)

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 [12, 18, 19]. 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 implicitly differentiate Lagrange multipliers or otherwise introduce artificial and unnecessary pathologies. (See Example 1.6 below.) In particular, companion forms are not consistent with the first order formulations for differential-algebraic equations used in multi-body dynamics [8]. (See also [25] for optimal first order formulations in the context of differential-algebraic equations.)

New classes of structure preserving linearizations introduced in [22] and analyzed in [14, 15, 23, 21] hold much promise. Still a different family of linearizations was introduced in [1, 2]. However, in order to use some of these new linearization, 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.

At this writing, a generalization of Kronecker canonical form to matrix polynomials of degree greater than one is unknown and seems unlikely to exist. However, Jordan and Kronecker chains are partially generalized by the concepts of Jordan triples, see [12]. In the context of index reduction methods for differential-algebraic equations, a recent canonical form derived in [25] displays the 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 staircase forms under orthogonal transformations that allow computation of the structural information associated with the eigenvalue infinity and the singular parts of matrix polynomials. These results generalize concepts studied in the context of differential-algebraic equations in [25]. 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 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 correspond to first order formulations in constrained multibody dynamics [8] and general higher order differential-algebraic systems [25]. 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 1.2 *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i$ be a matrix polynomial of degree k, where $A_0, \ldots, A_k \in \mathbb{F}^{m,n}$, $A_k \neq 0$.*

i) The matrix polynomial $P(\lambda)$ 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.

- *Define the* adjoint $P^\star(\lambda)$ *and the* reversal $\operatorname{rev} P(\lambda)$ *of $P(\lambda)$, respectively, by*

$$P^\star(\lambda) := \sum_{i=0}^{k} \lambda^i A_i^\star \quad \text{and} \quad \operatorname{rev} P(\lambda) := \lambda^k P(1/\lambda) = \sum_{i=0}^{k} \lambda^{k-i} A_i. \quad (1.2)$$

Here and throughout this paper * stands for either ^{T}, the transpose when $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$, or *, the complex conjugate transpose when $\mathbb{F} = \mathbb{C}$.

Definition 1.3 ([19]) *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*

$$\begin{aligned} P(\hat{\lambda})x_0 &= 0; \\ P(\hat{\lambda})x_1 + [\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})]x_0 &= 0; \\ &\vdots \\ P(\hat{\lambda})x_\ell + [\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})]x_{\ell-1} + \ldots + [\tfrac{1}{\ell!}\tfrac{d^\ell}{d\lambda^\ell}P(\hat{\lambda})]x_0 &= 0, \end{aligned} \tag{1.3}$$

$$\begin{aligned} y_0^\star P(\hat{\lambda}) &= 0; \\ y_1^\star P(\hat{\lambda}) + y_0[\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})] &= 0; \\ &\vdots \\ y_\ell^\star P(\hat{\lambda}) + y_{\ell-1}^\star[\tfrac{1}{1!}\tfrac{d}{d\lambda}P(\hat{\lambda})] + \ldots + y_0^\star[\tfrac{1}{\ell!}\tfrac{d^\ell}{d\lambda^\ell}P(\hat{\lambda})] &= 0, \end{aligned} \tag{1.4}$$

respectively.

A right Kronecker chain of length $\ell+1$ *associated with the eigenvalue infinity of* $P(\lambda)$ *is a right Kronecker chain of length* $\ell+1$ *associated with eigenvalue* $\lambda = 0$ *of* $\operatorname{rev} P(\lambda)$. *A left* Kronecker chain of length $\ell+1$ *associated with the eigenvalue infinity of* $P(\lambda)$ *is a left Kronecker chain of length* $\ell+1$ *associated with eigenvalue* $\lambda = 0$ *of* $\operatorname{rev} P(\lambda)$.

For Kronecker chains associated with the singular parts of the matrix polynomials we extend the classical definition for matrix pencils as in [9, 10, 19].

Definition 1.4 *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*

$$P(\lambda)x(\lambda) = 0, \tag{1.5}$$

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*

$$y(\lambda)^\star P(\lambda) = 0. \tag{1.6}$$

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 [11]. 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, rev $L(\lambda)$ is a linearization of 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 1.7 that a strong linearization of a singular matrix pencil may not preserve the lengths of singular chains. Linearizations like the one in Example 1.6 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 1.5 *Consider the following matrix polynomial which has the structure of a constrained and damped mechanical system [8].*

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

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 an index 4 infinite eigenvalue. Treating $T(\lambda)$ as a degree two polynomial, the companion form linearization is also a linearization of $P(\lambda)$ which has two index 2 eigenvalues at 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 index 0 infinite eigenvalues.

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 1.6 The Euler-Lagrange equations [8] of a linear constrained and damped mechanical system are given by a differential-algebraic equation of the form

$$\begin{aligned} M\ddot{x} + D\dot{x} + Kx + G^T\mu &= f(t), \\ Gx &= 0. \end{aligned} \tag{1.7}$$

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

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

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 1.3, $P(\lambda)$ has Kronecker chains associated with the eigenvalue infinity of length 4.

The companion linearization is

$$L(\lambda) = \lambda \left[\begin{array}{cc|cc} M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{array}\right] + \left[\begin{array}{cc|cc} D & 0 & K & G^T \\ 0 & 0 & G & 0 \\ \hline -I & 0 & 0 & 0 \\ 0 & -I & 0 & 0 \end{array}\right]. \tag{1.8}$$

It corresponds to extending the two unknowns $[x, \mu]^T$ in (1.7) 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 (1.8)). 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 (1.9)) and does not introduce a derivative of the Lagrange multiplier μ. This approach gives the linear matrix pencil

$$\tilde{L}(\lambda) = \lambda \left[\begin{array}{cc|c} M & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{cc|c} D & K & G^T \\ -I & 0 & 0 \\ \hline 0 & G & 0 \end{array}\right], \tag{1.9}$$

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, 17, 25].

It has been demonstrated in [25] 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 1.7 For the singular matrix polynomial

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

following Definition 1.4, 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 recall some results on (structured) staircase forms for matrix pairs under orthogonal/unitary transformations. In section 3 we then extend these results to matrix polynomials and show that some (but not all) of the information associated with the eigenvalue infinity and the singular parts can be obtained from the staircase forms. In section 4 we use these staircase forms to obtain new (structured) trimmed linearizations.

2 Staircase forms for matrix pairs

In this section we discuss different types of equivalence relations and recall and extend some staircase forms for structured pairs of matrices. 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, then 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.

Definition 2.1 *We denote the set of nonnegative integers by* $\mathbb{N}_0$.

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

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

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

$$B_i = PA_iQ, \quad i = 0, 1, \ldots, \ell. \tag{2.1}$$

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

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

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*

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

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

$$B_i = Q^\star A_i Q, \quad i = 0, 1, \ldots, \ell. \tag{2.2}$$

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

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

At this writing, canonical forms under strong equivalence are only known for the case $\ell < 2$, by the Kronecker canonical form, [9, 10]. Condensed forms which present partial information about the invariants associated with the eigenvalue infinity and the right singular chains are studied in [25]. We are particularly interested in such forms under strong u-equivalence transformations.

Lemma 2.2 *([4]) Let* $A, B \in \mathbb{F}^{m,n}$. *The matrix pair* (A, B) *is strongly u-equivalent to a matrix pair of the form*

$$\left(\begin{bmatrix} A_{11} & A_{12} & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} & B_{13} & \Sigma_{14} & 0 \\ B_{21} & B_{22} & B_{23} & 0 & 0 \\ B_{31} & B_{32} & \Sigma_{33} & 0 & 0 \\ \Sigma_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right), \tag{2.3}$$

where the matrices

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \Sigma_{14}, \quad \Sigma_{41}, \quad \Sigma_{33}$$

are square nonsingular. (The matrices Σ_{ij} *can be even chosen to be diagonal.) Any of the block rows or block columns may be void.*

This condensed form displays some but not all of the singular and eigenvalue infinity structure of the matrix pair. In order to derive the complete set of invariants, the compression process has to be continued and leads to the so called generalized upper-triangular (GUPTRI) form or staircase form, see [6, 7, 29].

If the matrices A, B have some symmetry structure, then this structure is preserved under appropriate u-congruence transformations.

Corollary 2.3 *If $A, B \in \mathbb{F}^{n,n}$ satisfy $A = \pm A^\star$, $B = \pm B^\star$, where $\star$ is either T or $*$, then the matrix pair (A, B) is strongly u-congruent to a matrix pair $(Q^\star AQ, Q^\star BQ)$ of the form (2.3), where Q is unitary, and the matrices*

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \pm \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^\star, \quad \Sigma_{14} = \pm \Sigma_{41}^\star, \quad \Sigma_{33} = \pm \Sigma_{33}^\star$$

are square nonsingular. Furthermore, the symmetry structure of the original pair is preserved. When both A and B are real, then the transformation matrix Q can be chosen real, and (2.3) is in real condensed form as well.

Proof. We present a constructive proof for the case with $A = A^T, B = B^T \in \mathbb{F}^{n,n}$. The other cases can be proved in a similar way.

Step 1. Use a unitary (or real orthogonal) matrix Q_1 to compress A to

$$Q_1^T A Q_1 = \begin{bmatrix} \hat{A} & 0 \\ 0 & 0 \end{bmatrix},$$

where $\hat{A}$ is nonsingular (i.e. compute the Takagi decomposition, see e.g. [16] when A is complex, or the symmetric Schur form when A is real). Apply a congruence transformation to B with the same Q_1 to get

$$Q_1^T B Q_1 := \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ \hat{B}_{12}^T & \hat{B}_{22} \end{bmatrix},$$

partitioned conformable with $Q_1^T A Q_1$.

Step 2. In the same way as for A, transform $\hat{B}_{22}$ with a unitary (or real orthogonal) matrix $\hat{Q}_2$ to

$$\hat{Q}_2^T \hat{B}_{22} \hat{Q}_2 = \begin{bmatrix} \Sigma_{33} & 0 \\ 0 & 0 \end{bmatrix},$$

where Σ_{33} is nonsingular. Let $Q_2 = Q_1 \mathrm{diag}(I, \hat{Q}_2)$. So,

$$(Q_2^T A Q_2, Q_2^T B Q_2) = \left(\left[\begin{array}{c|cc} \hat{A} & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right], \left[\begin{array}{c|cc} \hat{B}_{11} & \tilde{B}_{12} & \tilde{B}_{13} \\ \hline \tilde{B}_{12}^T & \Sigma_{33} & 0 \\ \tilde{B}_{13}^T & 0 & 0 \end{array}\right] \right).$$

Step 3. Determine unitary (or real orthogonal) matrices $\hat{Q}_3, \hat{Q}_4$ such that

$$\hat{Q}_3^T \tilde{B}_{13} \hat{Q}_4 = \begin{bmatrix} \Sigma_{14} & 0 \\ 0 & 0 \end{bmatrix},$$

where Σ_{14} is nonsingular, (e.g. using the singular value decomposition, see e.g. [13]). Let $Q = Q_2 \mathrm{diag}(\hat{Q}_3, I, \hat{Q}_4)$. Then the matrix pair $(Q^T A Q, Q^T B Q)$ has the form of (2.3). Obviously, the symmetry structure is preserved under these congruence transformations. □

It is possible to compress the middle blocks

$$\begin{bmatrix} A_{22} & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} B_{22} & B_{23} \\ B_{32} & \Sigma_{33} \end{bmatrix}$$

in the pair (2.3) even further.

Lemma 2.4 *The matrix pair*

$$(N, K) = \left(\begin{bmatrix} N_{11} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \right)$$

with $K_{22} \in \mathbb{F}^{r,r}$ invertible and $N_{11} \in \mathbb{F}^{n-r,n-r}$, is u-equivalent to

$$\left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} & \tilde{K}_{14} \\ \tilde{K}_{21} & \tilde{K}_{22} & \tilde{K}_{23} & 0 \\ \tilde{K}_{31} & \tilde{K}_{32} & \tilde{K}_{33} & 0 \\ \tilde{K}_{41} & 0 & 0 & 0 \end{bmatrix} \right),$$

where $\tilde{K}_{14}$ and $\tilde{K}_{41}$ have the same size and are both invertible, $\tilde{N}_{22} \in \mathbb{F}^{n-r,n-r}$, and $\tilde{K}_{33}$ is a nonzero scalar if r is odd or is void when r is even.

Moreover, rank N_{11} = rank $\tilde{N}_{22}$*, and*

$$\begin{array}{rcl} (\tilde{N}_{22}, \tilde{K}_{22} - \tilde{K}_{23}\tilde{K}_{33}^{-1}\tilde{K}_{32}) & \sim & (N_{11}, K_{11} - K_{12}K_{22}^{-1}K_{21}) \text{ if } r \text{ is odd }, \\ (\tilde{N}_{22}, \tilde{K}_{22}) & \sim & (N_{11}, K_{11} - K_{12}K_{22}^{-1}K_{21}) \text{ if } r \text{ is even.} \end{array} \tag{2.4}$$

Proof. Since K_{22} is nonsingular, there exist unitary (real orthogonal) matrices $\hat{U}_1, \hat{V}_1$ such that

$$\hat{U}_1 K_{22} \hat{V}_1 = \begin{bmatrix} \hat{K}_{22} & \Sigma_{12} & 0 \\ \Sigma_{21} & 0 & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}, \tag{2.5}$$

where Σ_{33} is a nonzero scalar when r is odd, or it is void when r is even, and Σ_{12} and Σ_{21} are both $s \times s$ ($s = \lfloor r/2 \rfloor$) and invertible. This form can be achieved e.g. by using the singular value decomposition followed by appropriate permutations, and then we even have $\hat{K}_{22} = 0$. Let $U_1 = \mathrm{diag}(I_{n-r}, \hat{U}_1)$ and $V_1 = \mathrm{diag}(I_{n-r}, \hat{V}_1)$. We obtain

$$(U_1 N V_1, U_1 K V_1) := \left(\begin{bmatrix} N_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & \hat{K}_{12} & \hat{K}_{13} & \hat{K}_{14} \\ \hat{K}_{21} & \hat{K}_{22} & \Sigma_{12} & 0 \\ \hat{K}_{31} & \Sigma_{21} & 0 & 0 \\ \hat{K}_{41} & 0 & 0 & \Sigma_{33} \end{bmatrix} \right).$$

Let $\hat{U}_2$ and $\hat{V}_2$ be unitary (real orthogonal) such that

$$\begin{bmatrix} \hat{K}_{13} \\ \Sigma_{12} \end{bmatrix} = \hat{U}_2^* \begin{bmatrix} \tilde{K}_{14} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \hat{K}_{31} & \Sigma_{21} \end{bmatrix} = \begin{bmatrix} \tilde{K}_{41} & 0 \end{bmatrix} \hat{V}_2^*,$$

where $\tilde{K}_{14}$ and $\tilde{K}_{41}$ are both $s \times s$. Note that $\hat{K}_{13}$ is $(n-r) \times s$ and $\hat{K}_{31}$ is $s \times (n-r)$. Partition

$$\hat{U}_2 := \begin{matrix} \\ s \\ n-r \end{matrix} \begin{matrix} \begin{matrix} n-r & s \end{matrix} \\ \begin{bmatrix} \hat{U}_{11} & \hat{U}_{12} \\ \hat{U}_{21} & \hat{U}_{22} \end{bmatrix} \end{matrix}, \quad \hat{V}_2 := \begin{matrix} \\ n-r \\ s \end{matrix} \begin{matrix} \begin{matrix} s & n-r \end{matrix} \\ \begin{bmatrix} \hat{V}_{11} & \hat{V}_{12} \\ \hat{V}_{21} & \hat{V}_{22} \end{bmatrix} \end{matrix}.$$

Since Σ_{12} and Σ_{21} are invertible, it follows from $\Sigma_{12} = \hat{U}_{12}^* \tilde{K}_{14}$ and $\Sigma_{21} = \tilde{K}_{41} \hat{V}_{21}^*$ that the matrices $\hat{U}_{12}$, $\hat{V}_{21}$ and $\tilde{K}_{14}$, $\tilde{K}_{41}$ are all invertible. Since $\hat{U}_2, \hat{V}_2$ are unitary (real orthogonal) it follows from their CS-decomposition, see e.g. [13], that $\hat{U}_{21}$ and $\hat{V}_{12}$ are also nonsingular. Define $U_2 = U_1 \mathrm{diag}(\hat{U}_2, I)$ and $V_2 = V_1 \mathrm{diag}(\hat{V}_2, I)$. Then

$$(U_2 N V_2, U_2 K V_2) := \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{14} & \tilde{K}_{13} \\ \tilde{K}_{21} & \tilde{K}_{22} & 0 & \tilde{K}_{23} \\ \tilde{K}_{41} & 0 & 0 & 0 \\ \tilde{K}_{31} & \tilde{K}_{32} & 0 & \Sigma_{33} \end{bmatrix} \right). \tag{2.6}$$

Let P be a permutation that interchanges the last two columns of $U_2 K V_2$ when it is post-multiplied, and let $U = PU_2$, $V = V_2 P$. Then U, V are unitary (real orthogonal), and

$$(UNV, UKV) = \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} & \tilde{K}_{14} \\ \tilde{K}_{21} & \tilde{K}_{22} & \tilde{K}_{23} & 0 \\ \tilde{K}_{31} & \tilde{K}_{32} & \tilde{K}_{33} & 0 \\ \tilde{K}_{41} & 0 & 0 & 0 \end{bmatrix} \right)$$

with $\tilde{K}_{33} = \Sigma_{33}$.

We now prove (2.4) for the case that $r = 2s + 1$ is odd. The even case follows analogously. With the factorization (2.5),

$$\begin{aligned} K_{11} - K_{12} K_{22}^{-1} K_{21} &= K_{11} - (K_{12} \hat{V}_1)(\hat{U}_1 K_{22} \hat{V}_1)^{-1} (\hat{U}_1 K_{21}) \\ &= K_{11} - \begin{bmatrix} \hat{K}_{12} & \hat{K}_{13} & \hat{K}_{14} \end{bmatrix} \begin{bmatrix} \hat{K}_{22} & \Sigma_{12} & 0 \\ \Sigma_{21} & 0 & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{K}_{21} \\ \hat{K}_{31} \\ \hat{K}_{41} \end{bmatrix} \\ &= K_{11} - \hat{K}_{13} \Sigma_{12}^{-1} \hat{K}_{21} - \hat{K}_{12} \Sigma_{21}^{-1} \hat{K}_{31} + \hat{K}_{13} \Sigma_{12}^{-1} \hat{K}_{22} \Sigma_{21}^{-1} \hat{K}_{31} - \hat{K}_{14} \Sigma_{33}^{-1} \hat{K}_{41} \\ &= \begin{bmatrix} I & -\hat{K}_{13} \Sigma_{12}^{-1} \end{bmatrix} \begin{bmatrix} K_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} I \\ -\Sigma_{21}^{-1} \hat{K}_{31} \end{bmatrix} - \hat{K}_{14} \Sigma_{33}^{-1} \hat{K}_{41}. \end{aligned}$$

From (2.6), we obtain that

$$\begin{aligned} &\hat{U}_{21} \hat{K}_{13} + \hat{U}_{22} \Sigma_{12} = 0, \\ &\hat{K}_{31} \hat{V}_{12} + \Sigma_{21} \hat{V}_{22} = 0, \\ &\tilde{K}_{32} = K_{41} \hat{V}_{12}, \\ &\tilde{K}_{23} = \hat{U}_{21} K_{14}, \\ &\tilde{K}_{22} = \begin{bmatrix} \hat{U}_{21} & \hat{U}_{22} \end{bmatrix} \begin{bmatrix} K_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} \hat{V}_{12} \\ \hat{V}_{22} \end{bmatrix}, \\ &\tilde{N}_{22} = \hat{U}_{21} N_{11} \hat{V}_{12}. \end{aligned} \tag{2.7}$$

Since $\hat{U}_{21}$ and $\hat{V}_{12}$ are invertible, from the first two equations of (2.7), we have that

$$-\hat{K}_{13}\Sigma_{12}^{-1} = \hat{U}_{21}^{-1}\hat{U}_{22}, \qquad -\Sigma_{21}^{-1}\hat{K}_{31} = \hat{V}_{22}\hat{V}_{12}^{-1}.$$

Since $\Sigma_{33} = \tilde{K}_{33}$, we can combine this with the other equations of (2.7) and obtain

$$\begin{aligned}
&\hat{U}_{21}(K_{11} - K_{12}K_{22}^{-1}K_{21})\hat{V}_{12} \\
&= \hat{U}_{21}\left(\begin{bmatrix} I & -\hat{K}_{13}\Sigma_{12}^{-1} \end{bmatrix} \begin{bmatrix} K_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} I \\ -\Sigma_{21}^{-1}\hat{K}_{31} \end{bmatrix} - \hat{K}_{14}\Sigma_{33}^{-1}\hat{K}_{41} \right)\hat{V}_{12} \\
&= \hat{U}_{21}\left(\begin{bmatrix} I & \hat{U}_{21}^{-1}\hat{U}_{22} \end{bmatrix} \begin{bmatrix} K_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} I \\ \hat{V}_{22}\hat{V}_{12}^{-1} \end{bmatrix} - \hat{K}_{14}\Sigma_{33}^{-1}\hat{K}_{41} \right)\hat{V}_{12} \\
&= \begin{bmatrix} \hat{U}_{21} & \hat{U}_{22} \end{bmatrix} \begin{bmatrix} K_{11} & \hat{K}_{12} \\ \hat{K}_{21} & \hat{K}_{22} \end{bmatrix} \begin{bmatrix} \hat{V}_{12} \\ \hat{V}_{22} \end{bmatrix} - \tilde{K}_{23}\Sigma_{33}^{-1}\tilde{K}_{32} \\
&= \tilde{K}_{22} - \tilde{K}_{23}\tilde{K}_{33}^{-1}\tilde{K}_{32}.
\end{aligned}$$

Since $\tilde{N}_{22} = \hat{U}_{21}N_{11}\hat{V}_{12}$, we have that (2.4) holds with invertible matrices $\hat{U}_{21}$ and $\hat{V}_{12}$. Finally, that $\mathrm{rank}\tilde{N}_{22} = \mathrm{rank}N_{11}$ follows from $\tilde{N}_{22} = \hat{U}_{21}N_{11}\hat{V}_{12}$. □

If the matrices K, N in (2.4) have symmetries, then the symmetry structures can be preserved.

Lemma 2.5 *Consider a pair of matrices*

$$(N, K) = \left(\begin{bmatrix} N_{11} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \right)$$

with $K_{22} \in \mathbb{F}^{r,r}$ invertible, and $N_{11} \in \mathbb{F}^{n-r,n-r}$.

(a) If $K = -K^T, N = \pm N^T \in \mathbb{F}^{n,n}$, then there exists a unitary matrix Q such that

$$(Q^T N Q, Q^T K Q) = \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} \\ -\tilde{K}_{12}^T & \tilde{K}_{22} & 0 \\ -\tilde{K}_{13}^T & 0 & 0 \end{bmatrix}, \right)$$

where $\tilde{K}_{13}$ is invertible and $\tilde{N}_{22} \in \mathbb{F}^{n-r,n-r}$.

Moreover, $\mathrm{rank}N_{11} = \mathrm{rank}\tilde{N}_{22}$, *and*

$$(\tilde{N}_{22}, \tilde{K}_{22}) \overset{c}{\sim} (N_{11}, K_{11} + K_{12}K_{22}^{-1}K_{12}^T).$$

(b) If $K = K^T, N = \pm N^T \in \mathbb{F}^{n,n}$, then there exist a unitary matrix Q such that

$$(Q^T N Q, Q^T K Q) = \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} & \tilde{K}_{14} \\ \tilde{K}_{12}^T & \tilde{K}_{22} & \tilde{K}_{23} & 0 \\ \tilde{K}_{13}^T & \tilde{K}_{23}^T & \tilde{K}_{33} & 0 \\ \tilde{K}_{14}^T & 0 & 0 & 0 \end{bmatrix} \right),$$

where $\tilde{K}_{14}$ *is invertible,* $\tilde{N}_{22} \in \mathbb{C}^{n-r,n-r}$*, and* $\tilde{K}_{33}$ *is a nonzero scalar if* r *is odd or is void if* r *is even.*

Moreover, $\mathrm{rank} N_{11} = \mathrm{rank} \tilde{N}_{22}$ *and*

$$(\tilde{N}_{22}, \tilde{K}_{22} - \tilde{K}_{23}\tilde{K}_{33}^{-1}\tilde{K}_{23}^T) \overset{c}{\sim} (N_{11}, K_{11} - K_{12}K_{22}^{-1}K_{12}^T),$$

where the pair on the left is just $(\tilde{N}_{22}, \tilde{K}_{22})$ *if* r *is even.*

In general Q *is complex unitary no matter whether* $\mathbb{F}$ *is* $\mathbb{C}$ *or* $\mathbb{R}$*.*

(c) If $K = K^*, N = \pm N^* \in \mathbb{F}^{n,n}$ *and* $K_{22} \overset{c}{\sim} \mathrm{diag}(I_\omega, -I_\mu)$*, then there exist a unitary matrix* Q *such that*

$$(Q^*NQ, Q^*KQ) = \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} & \tilde{K}_{14} \\ \tilde{K}_{12}^* & \tilde{K}_{22} & \tilde{K}_{23} & 0 \\ \tilde{K}_{13}^* & \tilde{K}_{23}^* & \tilde{K}_{33} & 0 \\ \tilde{K}_{14}^* & 0 & 0 & 0 \end{bmatrix} \right),$$

where $\tilde{K}_{14}$ *is invertible,* $\tilde{N}_{22} \in \mathbb{F}^{n-r,n-r}$*, and* $\tilde{K}_{33} = \tilde{K}_{33}^*$ *is definite and of size* $|\omega - \mu| \times |\omega - \mu|$*.*

Moreover, $\mathrm{rank} N_{11} = \mathrm{rank} \tilde{N}_{22}$*, and*

$$(\tilde{N}_{22}, \tilde{K}_{22} - \tilde{K}_{23}\tilde{K}_{33}^{-1}\tilde{K}_{23}^*) \overset{c}{\sim} (N_{11}, K_{11} - K_{12}K_{22}^{-1}K_{12}^*).$$

(d) If $K = -K^*, N = \pm N^* \in \mathbb{F}^{n,n}$ *and* $K_{22} \overset{c}{\sim} \mathrm{diag}(iI_\omega, -iI_\mu)$*, then there exist a unitary matrix* Q *such that*

$$(Q^*NQ, Q^*KQ) = \left(\begin{bmatrix} \tilde{N}_{11} & \tilde{N}_{12} & 0 & 0 \\ \tilde{N}_{21} & \tilde{N}_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} & \tilde{K}_{13} & \tilde{K}_{14} \\ -\tilde{K}_{12}^* & \tilde{K}_{22} & \tilde{K}_{23} & 0 \\ -\tilde{K}_{13}^* & -\tilde{K}_{23}^* & \tilde{K}_{33} & 0 \\ -\tilde{K}_{14}^* & 0 & 0 & 0 \end{bmatrix} \right),$$

where $\tilde{K}_{14}$ *is invertible,* $\tilde{N}_{22} \in \mathbb{F}^{n-r,n-r}$*, and* $\tilde{K}_{33} = i\Omega$ *with* Ω *being definite of size* $|\omega - \mu| \times |\omega - \mu|$*.*

Moreover, $\mathrm{rank} N_{11} = \mathrm{rank} \tilde{N}_{22}$*, and*

$$(\tilde{N}_{22}, \tilde{K}_{22} + \tilde{K}_{23}\tilde{K}_{33}^{-1}\tilde{K}_{23}^*) \overset{c}{\sim} (N_{11}, K_{11} + K_{12}K_{22}^{-1}K_{12}^*).$$

In all cases the symmetry structure of the original pair is preserved. When both N, K *are real, in all cases but (b),* Q *can be chosen to be real orthogonal, and the staircase forms are real as well.*

Proof. The proof is essentially the same as that of Lemma 2.4. Note that unitary (or real orthogonal) matrices U, V in Lemma 2.4 can be chosen to be the same (which is Q here) due to the symmetry. The only part that still needs a proof for all cases is the factorization form for K_{22} as in (2.5). We will prove this case by case.

(a) If $K = -K^T$ then $K_{22} = -K_{22}^T$. In this case r is even, since $\det K_{22} = 0$ if r is odd.

If K is real skew-symmetric, then K_{22} is also real skew-symmetric and thus there exits an orthogonal matrix $\hat{Q}_1$ such that K_{22} has the form

$$\hat{Q}_1^T K_{22} \hat{Q}_1 = \begin{bmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{bmatrix},$$

where Λ is real diagonal and nonsingular.

If $K = -K^T$ is complex, then K_{22} is also skew-symmetric and complex. In this case one may use U_1, a product of Householder reflections to reduce K_{22} to $U_1^T K_{22} U_1$ which is tridiagonal. One then applies the permutation $P_1 = [e_1, e_3, \ldots, e_{r-1}, e_2, e_4, \ldots, e_r]$ such that

$$(U_1 P_1)^T K_{22} U_1 P_1 = \begin{bmatrix} 0 & \Sigma \\ -\Sigma^T & 0 \end{bmatrix},$$

where Σ is upper bidiagonal. One may reduce Σ further to a diagonal matrix by using the singular value decomposition.

(b) If K is complex symmetric, then K_{22} has the Takagi form

$$U_1^T K_{22} U_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r),$$

where U_1 is unitary and $\sigma_1, \ldots, \sigma_r$ are positive scalars. If $r = 2p$ is even, then for $D = \mathrm{diag}(I_p, iI_p)$ one has

$$(U_1 D)^T K_{22} (U_1 D) = \mathrm{diag}(\sigma_1, \ldots, \sigma_p, -\sigma_{p+1}, \ldots, -\sigma_{2p}).$$

For each $j = 1, \ldots, p$ we pair σ_j with $-\sigma_{p+j}$ and consider the diagonal matrix $\begin{bmatrix} \sigma_j & 0 \\ 0 & -\sigma_{p+j} \end{bmatrix}$. Considering the orthogonal matrix $G_j = \begin{bmatrix} c_j & s_j \\ -s_j & c_j \end{bmatrix}$ with $c_j = \sqrt{\sigma_j/(\sigma_j + \sigma_{p+j})}$, $s_j = \sqrt{\sigma_{p+j}/(\sigma_j + \sigma_{p+j})}$, we obtain

$$G_j^T \begin{bmatrix} \sigma_j & 0 \\ 0 & -\sigma_{p+j} \end{bmatrix} G_j = \begin{bmatrix} \sigma_j - \sigma_{p+j} & \sqrt{\sigma_j \sigma_{p+j}} \\ \sqrt{\sigma_j \sigma_{p+j}} & 0 \end{bmatrix}.$$

Applying these transformations, we can determine a real orthogonal matrix U_2 such that, with the unitary matrix $Q_1 := U_1 D U_2$, we have

$$Q_1^T K_{22} Q_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & 0 \end{bmatrix},$$

where $\Sigma_{12} = \mathrm{diag}(\sqrt{\sigma_1 \sigma_{p+1}}, \ldots, \sqrt{\sigma_p \sigma_{2p}})$ is nonsingular and $\Sigma_{11} = \mathrm{diag}(\sigma_1 - \sigma_{p+1}, \ldots, \sigma_p - \sigma_{2p})$.

If r is odd, then one of the diagonal elements, say σ_r, will remain unpaired. However, with the same procedure as in the even case we still can get a form as (2.5).

If K is real symmetric, we begin with the Schur form of K_{22} and then continue as before.

(c) By the assumptions, there exists a unitary matrix U_1 such that

$$U_1^* K_{22} U_1 = \mathrm{diag}(\Sigma_1, -\Sigma_2),$$

where $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_\omega)$, $\Sigma_2 = \mathrm{diag}(\sigma_{\omega+1}, \ldots, \sigma_{\omega+\mu})$, both with positive diagonal elements. Suppose that $\mu \geq \omega$, otherwise we may switch Σ_1 and $-\Sigma_2$ by a block permutation. We may pair $\sigma_j, -\sigma_{\omega+j}$ for $j = 1, \ldots, \omega$ and apply the same transformations used in (b) to the matrix $\left[\begin{smallmatrix} \sigma_j & 0 \\ 0 & -\sigma_{\omega+j} \end{smallmatrix}\right]$. In this way we obtain a unitary matrix Q_1 such that

$$Q_1^* K_{22} Q_1 = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{12}^* & 0 & 0 \\ 0 & 0 & \Sigma_{33} \end{bmatrix},$$

which is similar to (2.5) but $\Sigma_{33} = -\mathrm{diag}(\sigma_{2\omega+1}, \ldots, \sigma_{\omega+\mu})$, which is $|\mu - \omega| \times |\mu - \omega|$.

(d) In this case a factorization like (2.5) can be obtained by considering iK_{22} and applying the proof of part in (c). The other statements are obvious. □

Remark 2.6 It should be noted that in case (b) of Lemma 2.5 we use unitary transformations even in the case that K, N are real, since otherwise we cannot perform the transformation with the diagonal matrix D and we will get a bigger size block K_{33} as that in case (c).

Using Lemmas 2.4 and 2.5 we obtain the following corollaries.

Corollary 2.7 *If $A, B \in \mathbb{F}^{m,n}$, then the matrix pair (A, B) is strongly u-equivalent to a matrix pair of the form*

$$\left(\begin{bmatrix} A_{11} & A_{12} & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} B_{11} & B_{12} & B_{13} & \Sigma_{14} & 0 \\ B_{21} & B_{22} & B_{23} & 0 & 0 \\ B_{31} & B_{32} & \Sigma_{33} & 0 & 0 \\ \Sigma_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right),$$

where the matrix A_{22} is square, Σ_{14} and Σ_{41} are square and invertible, Σ_{33} is either void or a nonzero scalar, $\left[\begin{smallmatrix} A_{21} & A_{22} \end{smallmatrix}\right]$ has full row rank and $\left[\begin{smallmatrix} A_{12} \\ A_{22} \end{smallmatrix}\right]$ has full column rank.

Proof. We first apply Lemma 2.2 to (A, B) to obtain the strongly u-equivalent pair

$$\left(\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} & 0 & 0 & 0 \\ \hat{A}_{21} & \hat{A}_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} & \hat{B}_{13} & \hat{\Sigma}_{14} & 0 \\ \hat{B}_{21} & \hat{B}_{22} & \hat{B}_{23} & 0 & 0 \\ \hat{B}_{31} & \hat{B}_{32} & \hat{\Sigma}_{33} & 0 & 0 \\ \hat{\Sigma}_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right).$$

We then apply Lemma 2.4 to

$$(N, K) := \left(\begin{bmatrix} \hat{A}_{22} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \hat{B}_{22} & \hat{B}_{23} \\ \hat{B}_{32} & \hat{\Sigma}_{33} \end{bmatrix} \right),$$

and obtain that the original pair is strongly u-equivalent to

$$\left(\left[\begin{array}{ccc|c|cc|c} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} & 0 & 0 & 0 & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} & \tilde{A}_{23} & 0 & 0 & 0 & 0 \\ \hline \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_{33} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc|c|cc|c} \tilde{B}_{11} & \tilde{B}_{12} & \tilde{B}_{13} & \tilde{B}_{14} & \tilde{B}_{15} & \hat{\Sigma}_{14} & 0 \\ \tilde{B}_{21} & \tilde{B}_{22} & \tilde{B}_{23} & \tilde{B}_{24} & \tilde{\Sigma}_{25} & 0 & 0 \\ \hline \tilde{B}_{31} & \tilde{B}_{32} & \tilde{B}_{33} & \tilde{B}_{34} & 0 & 0 & 0 \\ \hline \tilde{B}_{41} & \tilde{B}_{42} & \tilde{B}_{43} & \tilde{\Sigma}_{44} & 0 & 0 & 0 \\ \hline \tilde{B}_{51} & \tilde{\Sigma}_{52} & 0 & 0 & 0 & 0 & 0 \\ \hat{\Sigma}_{41} & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]\right).$$

With some re-partitioning we then obtain the desired condensed form.

Because $\left[\begin{array}{cc} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{array}\right]$ is invertible, and

$$\begin{aligned} \left[\begin{array}{cc} A_{21} & A_{22} \end{array}\right] &:= \left[\begin{array}{ccc} \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_{33} \end{array}\right] = \hat{U}_{21} \left[\begin{array}{ccc} \hat{A}_{21} & \hat{A}_{22} & 0 \end{array}\right] \operatorname{diag}(I, \hat{V}_2), \\ \left[\begin{array}{c} A_{12} \\ A_{22} \end{array}\right] &= \left[\begin{array}{c} \tilde{A}_{13} \\ \tilde{A}_{23} \\ \tilde{A}_{33} \end{array}\right] = \operatorname{diag}(I, \hat{U}_2) \left[\begin{array}{c} \hat{A}_{12} \\ \hat{A}_{22} \\ 0 \end{array}\right] \hat{V}_{12}, \end{aligned}$$

where $\hat{U}_{21}$ and $\hat{V}_{12}$ are invertible as in the proof of Lemma 2.4 along with unitary matrices $\hat{U}_2, \hat{V}_2$, it follows that $\left[\begin{array}{cc} A_{21} & A_{22} \end{array}\right]$ has full row rank and that $\left[\begin{smallmatrix} A_{12} \\ A_{22} \end{smallmatrix}\right]$ has full column rank. □

We also have a corresponding result for structured pairs.

Corollary 2.8 *If $A, B \in \mathbb{F}^{n,n}$ satisfy $A = \pm A^\star$ and $B = \pm B^\star$, then the matrix pair (A, B) is strongly u-congruent to a matrix pair $(Q^\star AQ, Q^\star BQ)$ of the form*

$$\left(\left[\begin{array}{ccccc} A_{11} & A_{12} & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccccc} B_{11} & B_{12} & B_{13} & \Sigma_{14} & 0 \\ B_{21} & B_{22} & B_{23} & 0 & 0 \\ B_{31} & B_{32} & \Sigma_{33} & 0 & 0 \\ \Sigma_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]\right),$$

where Q is unitary, A_{22} is square, Σ_{14}, Σ_{41} are square and invertible of the same size, and Σ_{33} is

(a) void if $B = -B^T$ and $A = \pm A^T$,

(b) a nonzero scalar or void if $B = B^T$ and $A = \pm A^T$,

(c) symmetric (Hermitian) definite or void if $B = B^*$ and $A = \pm A^*$,

(d) void or $i\Omega$, where Ω is Hermitian and definite if $B = -B^*$ and $A = \pm A^*$.

Furthermore, the symmetry structure of the original pair is preserved, and $\left[\begin{array}{cc} A_{21} & A_{22} \end{array}\right]$ has full row rank. If A, B are real, then the condensed form can be chosen to be real in all cases but (b). In case (b), in general, Q can only be complex unitary.

Proof. The condensed forms follow from Corollary 2.3 and Lemma 2.5. □

For pairs of the discussed symmetry structures the canonical forms are known, [26, 27]. The staircase form for matrix pairs where both matrices are symmetric (Hermitian) has recently been discussed in [20] and for the case that one of the matrices is symmetric (Hermitian) and the other skew-symmetric (skew-Hermitian) in [5]. We present here all these results in a combined way.

Theorem 2.9 *For a matrix pair (N, K) with $N, K \in \mathbb{F}^{n,n}$, $N = \pm N^\star$, and $K = \pm K^\star$, there exists a unitary (or real orthogonal) matrix $U \in \mathbb{F}^{n,n}$, such that*

$$U^\star N U = \left[\begin{array}{cccc|c|ccccc|c}
N_{11} & \cdots & \cdots & N_{1,m} & N_{1,m+1} & N_{1,m+2} & \cdots & N_{1,2m} & 0 & n_1 \\
\vdots & \ddots & & \vdots & \vdots & \vdots & \iddots & \iddots & & \vdots \\
\vdots & & \ddots & \vdots & \vdots & N_{m-1,m+2} & \iddots & & & \vdots \\
N_{m,1} & \cdots & \cdots & N_{m,m} & N_{m,m+1} & 0 & & & & n_m \\
\hline
N_{m+1,1} & \cdots & \cdots & N_{m+1,m} & N_{m+1,m+1} & & & & & l \\
\hline
N_{m+2,1} & \cdots & N_{m+2,m-1} & 0 & & & & & & q_m \\
\vdots & \iddots & \iddots & & & & & & & \vdots \\
N_{2m,1} & \iddots & & & & & & & & q_2 \\
0 & & & & & & & & & q_1
\end{array}\right]$$

$$U^\star K U = \left[\begin{array}{cccc|c|ccccc|c}
K_{11} & \cdots & \cdots & K_{1,m} & K_{1,m+1} & K_{1,m+2} & \cdots & \cdots & K_{1,2m+1} & n_1 \\
\vdots & \ddots & & \vdots & \vdots & \vdots & & \iddots & & \vdots \\
\vdots & & \ddots & \vdots & \vdots & \vdots & \iddots & & & \vdots \\
K_{m,1} & \cdots & \cdots & K_{m,m} & K_{m,m+1} & K_{m,m+2} & & & & n_m \\
\hline
K_{m+1,1} & \cdots & \cdots & K_{m+1,m} & K_{m+1,m+1} & & & & & l \\
\hline
K_{m+2,1} & \cdots & \cdots & K_{m+2,m} & & & & & & q_m \\
\vdots & & \iddots & & & & & & & \vdots \\
\vdots & \iddots & & & & & & & & \vdots \\
K_{2m+1,1} & & & & & & & & & q_1
\end{array}\right], \tag{2.8}$$

where $q_j \geq n_j$ for $j = 1, 2, \ldots, m$,

$$\begin{aligned}
N_{j,2m+1-j} &\in \mathbb{F}^{n_j,q_{j+1}}, \qquad 1 \leq j \leq m-1, \\
K_{j,2m+2-j} &= \begin{bmatrix} \Sigma_j & 0 \end{bmatrix} \in \mathbb{F}^{n_j,q_j}, \quad \Sigma_j \in \mathbb{F}^{n_j,n_j}, \qquad 1 \leq j \leq m,
\end{aligned}$$

and the blocks Σ_j, $j = 1, \ldots, m$ are nonsingular.

The pair $(N_{m+1,m+1}, K_{m+1,m+1})$ has the following block forms depending on the structure of K.

(a) If $K = -K^T, N = \pm N^T \in \mathbb{F}^{n,n}$, then $N_{m+1,m+1}, K_{m+1,m+1} \in \mathbb{F}^{l,l}$ and $N_{m+1,m+1}$ is invertible.

(b) If $K = K^T, N = \pm N^T \in \mathbb{F}^{n,n}$, then

$$(N_{m+1,m+1}, K_{m+1,m+1}) = \left(\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} \right),$$

Δ_{11} is invertible and Φ_{22} is either void or a nonzero scalar.

(c) If $K = K^*, N = \pm N^* \in \mathbb{F}^{n,n}$, *then*

$$(N_{m+1,m+1}, K_{m+1,m+1}) = \left(\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\star & \Phi_{22} \end{bmatrix} \right),$$

where Δ_{11} *is invertible and* Φ_{22} *is either void or Hermitian (or real symmetric) definite.*

(d) If $K = -K^*, N = \pm N^* \in \mathbb{F}^{n,n}$, *then*

$$(N_{m+1,m+1}, K_{m+1,m+1}) = \left(\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ -\Phi_{12}^* & i\Phi_{22} \end{bmatrix} \right),$$

where Δ_{11} *is invertible and* Φ_{22} *is either void or Hermitian definite.*

In all cases the pencil $\lambda N_{m+1,m+1} + K_{m+1,m+1}$ *is regular and all chains associated with the eigenvalue infinity for this pencil have length at most* 1.

The form (2.8) can be chosen to be real when N, K *are real in all cases but (b).*

Proof. We apply Corollary 2.8 to (N, K), so that it is strongly u-congruent to

$$\left(\left[\begin{array}{c|cc|cc} N_{11} & N_{12} & 0 & 0 & 0 \\ \hline N_{21} & N_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{c|ccc|cc} K_{11} & K_{12} & K_{13} & \Sigma_{14} & 0 \\ \hline K_{21} & K_{22} & K_{23} & 0 & 0 \\ K_{31} & K_{32} & \Sigma_{33} & 0 & 0 \\ \hline \Sigma_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \right).$$

We then reduce the middle block pair $\left(\begin{bmatrix} N_{22} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{22} & K_{23} \\ K_{32} & \Sigma_{33} \end{bmatrix} \right)$ successively by using Corollary 2.8, until the block in N_{22} position becomes invertible. □

It has been shown in [5] how to obtain the full information about the lengths of chains associated with the eigenvalues infinity and the singular chains of $\lambda N + K$ from a similar staircase form in the case of even pencils, where $N = -N^T$ and $K = K^T$. This information can be obtained in a similar way for the other structured pencils with the only difference in the block structure of $(N_{m+1,m+1}, K_{m+1,m+1})$. Also, if one still needs the structure information for the eigenvalue zero of $\lambda N + K$, one may continue the reduction on the middle block pair but in reversed order $(K_{m+1,m+1}, N_{m+1,m+1})$.

In this section we have studied (structured) staircase forms for matrix pairs. In the next section we partially extend these results to matrix tuples.

3 Condensed forms for tuples of matrices

In this section we discuss staircase form for matrix tuples associated with matrix polynomials. As mentioned in the introduction, it is an open problem [12, 28] to find a canonical form for quadratic matrix polynomials under strong equivalence. However, for pairs of matrices, i.e. linear matrix polynomials, in general we do not need the complete canonical form to extract the information about the singular blocks and the eigenvalue infinity. In the context of

differential-algebraic equations, such a form has been derived in [25]. But this form uses nonorthogonal transformations and only determines the information about the right eigenvectors and chains. Here we derive staircase forms under unitary (orthogonal) transformations that display the information about the singular parts and the eigenvalue infinity.

The first step is given by the following Lemma.

Lemma 3.1 *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*

$$\hat{N}_i = \left[\begin{array}{c|c|c} \begin{matrix} N_{11}^{(i)} & \cdots & \cdots & N_{1\tau}^{(i)} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ N_{\tau 1}^{(i)} & \cdots & \cdots & N_{\tau\tau}^{(i)} \end{matrix} & \begin{matrix} N_{1,\tau+1}^{(i)} \\ \vdots \\ \vdots \\ N_{\tau,\tau+1}^{(i)} \end{matrix} & \begin{matrix} N_{1,\tau+2}^{(i)} & \cdots & N_{1,2\tau}^{(i)} & 0 \\ \vdots & \iddots & \iddots & \\ N_{\tau-1,\tau+2}^{(i)} & \iddots & & \\ 0 & & & \end{matrix} \\ \hline \begin{matrix} N_{\tau+1,1}^{(i)} & \cdots & \cdots & N_{\tau+1,\tau}^{(i)} \end{matrix} & \begin{matrix} N_{\tau+1,\tau+1}^{(i)} \end{matrix} & \\ \hline \begin{matrix} N_{\tau+2,1}^{(i)} & \cdots & N_{\tau+2,\tau-1}^{(i)} & 0 \\ \vdots & \iddots & \iddots & \\ N_{2\tau,1}^{(i)} & \iddots & & \\ 0 & & & \end{matrix} & & \end{array}\right] \begin{matrix} m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline p_\tau \\ \vdots \\ p_2 \\ p_1 \end{matrix}, \tag{3.1}$$

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

$$\hat{K} = \left[\begin{array}{c|c|c} \begin{matrix} K_{11} & \cdots & \cdots & K_{1\tau} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ K_{\tau 1} & \cdots & \cdots & K_{\tau\tau} \end{matrix} & \begin{matrix} K_{1,\tau+1} \\ \vdots \\ \vdots \\ K_{\tau,\tau+1} \end{matrix} & \begin{matrix} K_{1,\tau+2} & \cdots & \cdots & K_{1,2\tau+1} \\ \vdots & & \iddots & \\ \vdots & \iddots & & \\ K_{\tau,\tau+2} & & & \end{matrix} \\ \hline \begin{matrix} K_{\tau+1,1} & \cdots & \cdots & K_{\tau+1,\tau} \end{matrix} & \begin{matrix} K_{\tau+1,\tau+1} \end{matrix} & \\ \hline \begin{matrix} K_{\tau+2,1} & \cdots & \cdots & K_{\tau+2,\tau} \\ \vdots & & \iddots & \\ \vdots & \iddots & & \\ K_{2\tau+1,1} & & & \end{matrix} & & \end{array}\right] \begin{matrix} m_1 \\ \vdots \\ \vdots \\ m_\tau \\ \hline s \\ \hline p_\tau \\ \vdots \\ \vdots \\ p_1 \end{matrix}, \tag{3.2}$$

where

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

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

$$N^{(i)}_{2\tau+1-j,j} \in \mathbb{F}^{p_{j+1},q_j}, \qquad 1 \le j \le \tau-1 \quad i=1,\dots,k,$$

(ii)

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

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

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

(iii) $N^{(i)}_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{N}^{(i)}_{11} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{s,t}$ *for* $i=1,\dots,k$, $K_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix} \in \mathbb{F}^{s,t}$, *and* $\tilde{N}^{(i)}_{11}$, $i=1,\dots,k$, *have no nontrivial common left or right nullspace,* $\tilde{K}_{22}$ *is either void (and* $N^{(i)}_{\tau+1,\tau+1} = \tilde{N}^{(i)}_{11}$ *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, $i=1,\dots,k$, i.e. we transform the tuple to

$$(N_k,\dots,N_1,K) \sim \left(\begin{bmatrix} N^{(k)}_1 & 0 \\ 0 & 0 \end{bmatrix}, \dots, \begin{bmatrix} N^{(1)}_1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \right),$$

by applying the same procedure for the pair case as in Corollary 2.7 (instead of N alone there, here we are applying the transformation to all N_j's simultaneously) to get

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

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 $n_1 \ge m_1$ and $p_1 \ge q_1$).

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

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

until the first k matrices have no nontrivial common left and right nullspaces. □

In the case that the tuple has extra symmetry structure, we get a structured staircase form.

Corollary 3.2 *If $N_i = \pm N_i^\star \in \mathbb{F}^{n,n}$, $i = 1,\dots,k$ and $K = \pm K^\star \in \mathbb{F}^{n,n}$, then the tuple $(N_k,\dots,N_1,K)$ is strongly u-congruent to a matrix tuple $(\hat{N}_k,\dots,\hat{N}_1,\hat{K})$, where all terms $\hat{N}_i$, $i = 1,\dots,k$, have the form*

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

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

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

where

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

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

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

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

(iii) $N_{\tau+1,\tau+1}^{(i)} = \begin{bmatrix} \tilde{N}_{11}^{(i)} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{s,s}$ *for* $i = 1,\dots,k$, $K_{\tau+1,\tau+1} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix} \in \mathbb{F}^{s,s}$, *and* $\tilde{N}_{11}^{(i)}$, $i = 1,\dots,k$, *have no nontrivial common left or right nullspace, and depending on the symmetry structure of K, $\tilde{K}_{22}$ is either void (and* $N_{\tau+1,\tau+1}^{(i)} = \tilde{N}_{11}^{(i)}$ *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 3.3 *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*

$$\left[\begin{array}{ccc|c|ccc} q_1 & \ldots & q_\ell & t & n_\ell & \ldots & n_1 \\ \hline A & \ldots & A & A & A & \ldots & A^{(i)}_{1,2\ell+1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \\ A & \ldots & A & A & A^{(i)}_{\ell,\ell+2} & & \\ \hline A & \ldots & A & A^{(i)}_{\ell+1,\ell+1} & & & \\ \hline A & \ldots & A^{(i)}_{\ell+2,\ell} & & & & \\ \vdots & \iddots & & & & & \\ A^{(i)}_{2\ell+1,1} & & & & & & \end{array}\right] \begin{array}{c} m_1 \\ \vdots \\ m_\ell \\ \hline s \\ \hline p_\ell \\ \vdots \\ p_1 \end{array}, \tag{3.3}$$

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 $\begin{bmatrix} \Gamma_j \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Here Σ_j and Γ_j again denote a non-singular (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 $A^{(i)}_{\ell+1,\ell+1} = \begin{bmatrix} \tilde{A}^{(i)}_{11} & \tilde{A}^{(i)}_{12} \\ \tilde{A}^{(i)}_{21} & \tilde{A}^{(i)}_{22} \end{bmatrix}$ 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 3.1 to $N_k = A_k, \ldots N_1 = A_1$ and $K = A_0$ and obtain the u-equivalent tuple of the forms (3.1) and (3.2). 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 3.1 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 (3.1) and (3.2).

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

$$\left(\hat{A}_k, \ldots, \hat{A}_1, \hat{A}_0\right) =: \left(\begin{bmatrix} \tilde{A}_{11}^{(k)} & 0 \\ 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} \tilde{A}_{11}^{(1)} & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}_{11}^{(0)} & \tilde{A}_{12}^{(0)} \\ \tilde{A}_{21}^{(0)} & \tilde{A}_{22}^{(0)} \end{bmatrix}\right),$$

where $\tilde{A}_{22}^{(0)}$ 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 3.4 *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 $\hat{A}_i$, $i = 0, \ldots, k$, have the form*

$$\begin{array}{c|ccc|c|ccc|c}
 & m_1 & \ldots & m_\ell & s & n_\ell & \ldots & n_1 & \\
\hline
 & A & \ldots & A & A & A & \ldots & A_{1,2\ell+1}^{(i)} & m_1 \\
 & \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & & \vdots \\
 & A & \ldots & A & A & A_{\ell,\ell+2}^{(i)} & & & m_\ell \\
\hline
 & A & \ldots & A & A_{\ell+1,\ell+1}^{(i)} & & & & s \\
\hline
 & A & \ldots & A_{\ell+2,\ell}^{(i)} & & & & & n_\ell \\
 & \vdots & \iddots & & & & & & \vdots \\
 & A_{2\ell+1,1}^{(i)} & & & & & & & n_1
\end{array} , \tag{3.4}$$

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

All the matrices in the tuple of middle blocks $(A_{\ell+1,\ell+1}^{(k)}, \ldots, A_{\ell+1,\ell+1}^{(0)})$ 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 $A_{\ell+1,\ell+1}^{(i)} = \begin{bmatrix} \tilde{A}_{11}^{(i)} & \tilde{A}_{12}^{(i)} \\ \tilde{A}_{21}^{(i)} & \tilde{A}_{22}^{(i)} \end{bmatrix}$ for $i = 0, 1, \ldots, k$, where for $i_0 \in \{0, 1, \ldots, k\}$, (depending on the structure of A_{i_0}), $\tilde{A}_{22}^{(i_0)}$ is Hermitian definite or $i\Omega$ with Ω Hermitian definite and $\tilde{A}_{12}^{(i)} = 0, \tilde{A}_{21}^{(i)} = 0, \tilde{A}_{22}^{(i)} = 0$ for $i \neq i_0$, and no k matrices including $A_{\ell+1,\ell+1}^{(i_0)}$ 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 3.2. □

Example 3.5 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 [19],

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

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

From the Smith forms of $P(\lambda)$ and $\operatorname{rev} P(\lambda)$ (which we didn't 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

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

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

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

(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 in [25] unimodular transformations are employed, that allow a combination of the different coefficients 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 3.3 or Corollary 3.4 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

(3.3), $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(1)}_{\ell+1,\ell+1}, A^{(0)}_{\ell+1,\ell+1})$, has the form

$$\left(\begin{bmatrix} \Sigma_k & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & 0 & \ldots & 0 \\ \tilde{A}^{(k-1)}_{21} & \Sigma_{k-1} & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, \right. \tag{3.5}$$

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

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 (3.5) as Example 3.5 and the following example show.

Example 3.6 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

$$x_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

associated with infinity and

$$z_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad z_2 = \begin{bmatrix} 0 \\ -1/2 \end{bmatrix}$$

associated with 0. No two coefficients have a common nullspace, and the matrix polynomial is not in the form (3.5). There exist no strong equivalence (congruence) transformations that reduce the matrix polynomial further to get it to the form (3.5). 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 (3.5) (with Σ_1 void) and even symmetric.

In general, the tuple of middle blocks in the condensed form (3.3), $(A^{(k)}_{\ell+1,\ell+1}, \ldots, A^{(1)}_{\ell+1,\ell+1}, A^{(0)}_{\ell+1,\ell+1})$, can only be reduced by strong u-equivalence to the form

$$\left(\begin{bmatrix} \Sigma_k & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix}, \begin{bmatrix} \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & 0 & \ldots & 0 \\ \tilde{A}^{(k-1)}_{21} & \tilde{A}^{(k-1)}_{22} & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots \right. \tag{3.6}$$

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

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 3.4, then the tuple of middle blocks in (3.4) can be transformed via strong u-congruence to a form as (3.6), or possibly even to the form (3.5). In order to analyze the tuple in (3.5), 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 [25]. Such a DAE has the form

$$A_k x^{(k)} + A_{k-1} x^{(k-1)} + \cdots + A_1 \dot{x} + A_0 x = f(t) \tag{3.7}$$

with some imhomogeneity 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 [17]. 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 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.

4 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 (3.3). If we partition the matrix polynomial in the form (3.3) as

$$\tilde{P}(\lambda) = \begin{bmatrix} P_{11}(\lambda) & P_{12}(\lambda) & P_{13}(\lambda) \\ P_{21}(\lambda) & P_{22}(\lambda) & 0 \\ P_{31}(\lambda) & 0 & 0 \end{bmatrix}, \tag{4.1}$$

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 (4.1). 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 $\begin{bmatrix} I_j \\ 0 \end{bmatrix}$ appear in $A^{(0)}_{2\ell+2-j,j}$ $(j = 1, \ldots, \ell)$ in (3.3). 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 (3.3) can be reduced further to (3.5), 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 (3.5). 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

$$
L_t(\lambda) = \lambda K_t + N_t =
$$

$$
\lambda \left[\begin{array}{cccccc|cccc|c|ccc}
\tilde{A}^{(1)}_{11} & \cdots & \tilde{A}^{(1)}_{1,k-1} & \tilde{A}^{(1)}_{1k} & 0 & \tilde{A}^{(2)}_{11} & \cdots & \tilde{A}^{(2)}_{1,k-2} & \tilde{A}^{(2)}_{1,k-1} & & \tilde{A}^{(k-1)}_{11} & \tilde{A}^{(k-1)}_{12} & \Sigma_k \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & & & \\
 & & & & & & & & & & \tilde{A}^{(k-1)}_{21} & \Sigma_{k-1} & 0 \\
\tilde{A}^{(1)}_{k-2,1} & \cdots & \tilde{A}^{(1)}_{k-2,k-1} & \tilde{A}^{(1)}_{k-2,k} & 0 & \tilde{A}^{(2)}_{k-2,1} & \cdots & \tilde{A}^{(2)}_{k-2,k-2} & \tilde{A}^{(2)}_{k-2,k-1} & \cdots & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k-1,1} & \cdots & \tilde{A}^{(1)}_{k-1,k-1} & \tilde{A}^{(1)}_{k-1,k} & 0 & \tilde{A}^{(2)}_{k-1,1} & \cdots & \tilde{A}^{(2)}_{k-1,k-2} & \Sigma_2 & & 0 & 0 & 0 \\
\tilde{A}^{(1)}_{k1} & \cdots & \tilde{A}^{(1)}_{k-1,1} & \Sigma_1 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\hline
I & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & I & 0 & 0 & 0 & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\hline
0 & \cdots & \cdots & 0 & 0 & I & \cdots & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & I & 0 & & 0 & 0 & 0 \\
\hline
 & & \vdots & & & & & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\hline
0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & I & 0 & 0
\end{array}\right]
$$

$$
+ \left[\begin{array}{ccccc|ccccc|c|ccc}
\tilde{A}^{(0)}_{11} & \cdots & \cdots & \tilde{A}^{(0)}_{1k} & \tilde{A}^{(0)}_{1,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \cdots & \vdots & \vdots & \vdots \\
\tilde{A}^{(0)}_{k1} & \cdots & \cdots & \tilde{A}^{(0)}_{kk} & \tilde{A}^{(0)}_{k,k+1} & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\tilde{A}^{(0)}_{k+1,1} & \cdots & \cdots & \tilde{A}^{(0)}_{k+1,k} & \Sigma_0 & 0 & \cdots & \cdots & 0 & & 0 & 0 & 0 \\
\hline
0 & 0 & \cdots & \cdots & 0 & -I & 0 & \cdots & 0 & & 0 & 0 & 0 \\
 & & & & & & \ddots & & \vdots & & & & \\
0 & 0 & \cdots & \cdots & 0 & 0 & -I & \ddots & & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 & & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & -I & & 0 & 0 & 0 \\
\hline
 & & \vdots & & & & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\
\hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & -I & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & -I & 0 \\
\hline
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 & -I
\end{array}\right]. \tag{4.2}
$$

We will now analyze this pencil.

Lemma 4.1 *The pencil $L_t(\lambda) = \lambda K_t + N_t$ in (4.2) 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. [24], 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 4.2 *If $P_{22}(\lambda)$ is a matrix polynomial with coefficient matrices in the form (3.5), then the linear pencil $L_t(\lambda)$ in (4.2) 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 4.3 *For regular strangeness-free matrix polynomials with the coefficient matrices of the form (3.5), the linearization (4.2) 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 4.4 *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 (3.5). 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.*

Remark 4.5 The extra zero blocks in the coefficient matrices of $P(\lambda)$ in Example 4.4 do not make $P(\lambda)$ anything special. In fact, the tuple (3.5) is strongly equivalent to the tuple

$$\begin{aligned} &\left(\left[\begin{array}{ccccc} \Sigma_k & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{array}\right], \left[\begin{array}{ccccc} \hat{A}_{11}^{(k-1)} & 0 & 0 & \dots & 0 \\ 0 & \Sigma_{k-1} & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{array}\right], \dots \right. \qquad (4.3) \\ &\left. \left[\begin{array}{ccccc} \hat{A}_{11}^{(1)} & \dots & \hat{A}_{1,k-1}^{(1)} & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \hat{A}_{k-1,1}^{(1)} & \dots & \hat{A}_{k-1,k-1}^{(1)} & 0 & 0 \\ 0 & \dots & 0 & \Sigma_1 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccccc} \hat{A}_{11}^{(0)} & \hat{A}_{12}^{(0)} & \dots & \hat{A}_{1k}^{(0)} & 0 \\ \hat{A}_{21}^{(0)} & \hat{A}_{22}^{(0)} & \dots & \hat{A}_{2k}^{(0)} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{A}_{k1}^{(0)} & \hat{A}_{k2}^{(0)} & \dots & \hat{A}_{kk}^{(0)} & 0 \\ 0 & 0 & \dots & 0 & \Sigma_0 \end{array}\right] \right). \end{aligned}$$

This form can be obtained by performing block Gaussian elimination using the blocks Σ_j in (3.5) to eliminate above and to the left. We may carry out

the elimination starting with Σ_0, i.e. the last diagonal block of the coefficient $A^{(0)}_{\ell+1,\ell+1}$, and then proceed inductively.

Based on (4.3), we can even construct a trimmed linearization without the eigenvalue infinity (as partially described in the proof of Lemma 4.1). However, since the last step of removing all the blocks associated with chains associated with infinity of length 1 cannot be performed by a strong u-equivalence, the computation of such a trimmed linearization may be numerically unstable.

If in (4.1) 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 (3.5) in the middle block of (3.3). We have seen that with the help of block Gaussian elimination we can split the part associated with the eigenvalue infinity form the finite eigenvalues. If we want to do this 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 [30] 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 associated with the eigenvalue 0, 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 (3.5), but always to (3.6). Similarly, using (3.6)

one can construct the linear pencil

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

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

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

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 of the singular parts and parts those 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.

5 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 3.4 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 [21], 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 longer than 1 associated with the eigenvalue infinity can be deflated in a structured way.

In the ideal case the tuple of middle blocks has the form (3.5). 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 [22, 21] cannot be used for (3.5) if it has the eigenvalue infinity, see [21]. We thus have to find structure preserving trimmed linearizations. For this we modify the vector spaces of linearizations that were derived in [22]. These spaces $\mathbb{L}_1(P)$, and $\mathbb{L}_2(P)$ consist of pencils that generalize the classical companion forms and are given by

$$\mathbb{L}_1(P) = \Big\{ L(\lambda) = \lambda X + Y : L(\lambda) \cdot (\Lambda \otimes I_n) = v \otimes P(\lambda),\, v \in \mathbb{F}^k \Big\}, \quad (5.1)$$

$$\mathbb{L}_2(P) = \Big\{ L(\lambda) = \lambda X + Y : \big(\Lambda^T \otimes I_n\big) \cdot L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{F}^k \Big\} \quad (5.2)$$

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 (5.1) is called the *right ansatz vector* of $L(\lambda) \in \mathbb{L}_1(P)$, and the vector w in (5.2) 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 [21] how to easily construct structured linear pencils using the *column-shifted sum* and *row-shifted sum*:

Definition 5.1 (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 [21], 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

$$\mathsf{p}(x\,;v) := v_1x^{k-1} + v_2x^{k-2} + \cdots + v_{k-1}x + v_k.$$

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 (3.5) 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 [21] 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

$$\lambda X_s + Y_s = \lambda \begin{bmatrix} & & & & A_k \\ & & & A_k & A_{k-1} \\ & & \iddots & \iddots & \vdots \\ & A_k & \iddots & \iddots & A_2 \\ A_k & A_{k-1} & \ldots & A_2 & A_1 \end{bmatrix} + \begin{bmatrix} & & & -A_k & 0 \\ & & \iddots & \vdots & \vdots \\ & \iddots & \iddots & -A_3 & 0 \\ -A_k & \ldots & -A_3 & -A_2 & 0 \\ 0 & \ldots & 0 & 0 & A_0 \end{bmatrix}. \tag{5.3}$$

If $P(\lambda)$ is $\star$-even or $\star$-odd, i.e., $P(\lambda) = P(-\lambda)^\star$ or $P(\lambda) = -P(-\lambda)^\star$, respectively, and $\mathit{\Pi}_k = \operatorname{diag}((-1)^{k-1}I_n, \ldots, (-1)^0 I_n)$ then the formally constructed linear pencil has the form $\lambda X_e + Y_e = \lambda \mathit{\Pi}_k X_s + \mathit{\Pi}_k Y_s$ with X_s, Y_s as in (5.3).

Example 5.2 Consider the ansatz vector $v = e_2$ for $P(\lambda) = \lambda^2 A + \lambda B + C$, with one of the properties

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

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

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

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

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

respectively.

Example 5.3 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

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

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 (3.5), then it is obvious that $\lambda X_s + Y_s$ as in (5.3) 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 $\Sigma_1, \ldots, \Sigma_k$ in (3.5).

In order to obtain a trimmed linearization, we delete all these zero block rows and columns. Setting $\mathcal{I}_{j_l} = \begin{bmatrix} I_{j_l} & 0 \end{bmatrix} \in \mathbb{F}^{j_l, n}$,

$$S = \operatorname{diag}(\mathcal{I}_{j_1},\ \mathcal{I}_{j_1+j_2}, \ldots, \mathcal{I}_{j_1+\ldots+j_{k-2}}, I_n),$$

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 5.4 *Consider a matrix polynomial (1.1) with the coefficient matrices of the form (3.5). 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

$$U(\lambda) = \left[\begin{array}{c|c|c|c|c|c} I_{j_1} & 0 & \dots & \dots & \dots & 0 \\ \hline \begin{matrix}\lambda I_{j_1} \\ 0\end{matrix} & I_{j_1+j_2} & 0 & \dots & \dots & 0 \\ \hline \begin{matrix}\lambda^2 I_{j_1} \\ 0\end{matrix} & \begin{matrix}\lambda I_{j_1+j_2} \\ 0\end{matrix} & I_{j_1+\dots+j_3} & \ddots & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \hline \vdots & \ddots & \ddots & \begin{matrix}\lambda I_{j_1+\dots+j_{k-3}} \\ 0\end{matrix} & I_{j_1+\dots+j_{k-2}} & 0 \\ \hline \begin{matrix}\lambda^{k-1} I_{j_1} \\ 0\end{matrix} & \dots & \dots & \begin{matrix}\lambda^2 I_{j_1+\dots+j_{k-3}} \\ 0\end{matrix} & \begin{matrix}\lambda I_{j_1+\dots+j_{k-2}} \\ 0\end{matrix} & I_n \end{array}\right]$$

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

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

with $\hat{S} = \operatorname{diag}(\mathcal{I}_{j_1},\ \mathcal{I}_{j_1+j_2}, \dots, \mathcal{I}_{j_1+\dots+j_{k-2}})$, 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. □

Example 5.5 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 (3.5), 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

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

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

$$U(\lambda)(\lambda\hat{X}_s+\hat{Y}_s) = \left[\begin{array}{c|c} W_{11} & W_{12}(\lambda) \\ \hline 0 & P(\lambda) \end{array}\right] = \left[\begin{array}{ccc|c} 0 & -\Sigma_A & 0 & \\ -\Sigma_A & -B_{11} & -B_{12} & W_{12}(\lambda) \\ 0 & -B_{21} & -\Sigma_B & \\ \hline \multicolumn{3}{c|}{0} & P(\lambda) \end{array}\right].$$

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

$$\left[\begin{array}{cc} P(\lambda) & 0 \\ 0 & I \end{array}\right].$$

In this section we have shown how to obtain structured trimmed linearization in the important special case that the middle block has the form (3.5).

For palindromic matrix polynomials, see [21], 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 (3.6), 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,

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

For the associated linear pencil $\lambda\hat{X}_s + \hat{Y}_s$, the corresponding matrix W_{11} is Σ_A which is nonsingular. If $k = 3$ as in Example 5.5, then

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

which is nonsingular only if B_{22} is.

6 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] A. Bunse-Gerstner, V. Mehrmann, and N. K. Nichols. Regularization of descriptor systems by derivative and proportional state feedback. *SIAM J. Matrix Anal. Appl.*, 13:46–67, 1992.

[5] R. Byers, V. Mehrmann, and H. Xu. A structured staircase algorithm for skew-symmetric/symmetric pencils. *Electr. Trans. Num. Anal.*, 26:1–33, 2007.

[6] 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.

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

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

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

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

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

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

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, 3rd edition, 1996.

[14] 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.

[15] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.

[16] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

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

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

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

[20] C. Lucas. *Algorithms for Cholesky and QR Factorizations, and the Semidefinite Generalized Eigenvalue Problem.* PhD thesis, University of Manchester, Manchester, England, 2004.

[21] 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.

[22] 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.

[23] D. Steven Mackey. *Structured Linearizations for Matrix Polynomials*. PhD thesis, Department of Mathematics, University of Manchester, Manchester, GB, 2006.

[24] 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.

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

[26] R. C. Thompson. The characteristic polynomial of a principal submatrix of a Hermitian pencil. *Linear Algebra Appl.*, 14:135–177, 1976.

[27] R. C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.

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

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

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