MANCHESTER 1824

The University of Manchester

Triangularizing matrix polynomials

Taslaman, Leo and Tisseur, Francoise and Zaballa, Ion

2012

MIMS EPrint: **2012.77**

ISSN 1749-9097

Triangularizing matrix polynomials ☆

Leo Taslaman[a,1], Françoise Tisseur [a,1,*], Ion Zaballa[b,2]

[a] *School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK*
[b] *Departamento de Matemática Aplicada y EIO, Euskal Herriko Univertsitatea (UPV/EHU), Apdo. Correos 644, 48080 Bilbao, Spain*

Abstract

For an algebraically closed field $\mathbb{F}$, we show that any matrix polynomial $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$, $n \le m$, can be reduced to triangular form, preserving the degree and the finite and infinite elementary divisors. We also characterize the real matrix polynomials that are triangularizable over the real numbers and show that those that are not triangularizable are quasi-triangularizable with diagonal blocks of sizes 1×1 and 2×2. The proofs we present solve the structured inverse problem of building up triangular matrix polynomials starting from lists of elementary divisors.

Keywords: matrix polynomial, triangular, quasi-triangular, equivalence, elementary divisors, Smith form, majorization, Segre characteristic, Weyr characteristic, polynomial eigenvalue problem, inverse polynomial eigenvalue problem
2000 MSC: 15A18, 15A21, 15A54, 65F15

---

1. Introduction

Let $\mathbb{F}[\lambda]$ be the ring of polynomials in one variable with coefficients in a field $\mathbb{F}$. We are concerned with the problem of reducing over $\mathbb{F}[\lambda]$ (when possible) a matrix polynomial $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j \in \mathbb{F}[\lambda]^{n\times m}$ with $A_\ell \neq 0$, to triangular or trapezoidal form preserving its degree ℓ and eigenstructure. If such a reduction exists then $P(\lambda)$ is said to be triangularizable. By the eigenstructure of a matrix polynomial it is meant the eigenvalues and their partial multiplicities or, equivalently, the elementary divisors (or invariant factors) of the matrix polynomial, including those at infinity.

We show that when $\mathbb{F}$ is algebraically closed, any $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ with $n \le m$ is triangularizable, thereby extending an earlier but little-known result by Gohberg, Lancaster and Rodman for square monic matrix polynomials with complex coefficient matrices [3,

---

☆Version of November 2, 2012.
*Corresponding author.
Email addresses: `Leo.Taslaman@ma.man.ac.uk` (Leo Taslaman), `ftisseur@ma.man.ac.uk` (Françoise Tisseur), `ion.zaballa@ehu.es` (Ion Zaballa)
[1] These authors were supported by Engineering and Physical Sciences Research Council grant EP/I005293.
[2] This author was supported by the Dirección General de Investigación, Proyecto de Investigación MTM2010-19356-C02-01 and Gobierno Vasco, GIC10/IT-361-10 and UPV/EHU UFI11/52.

proof of Thm. 1.7]. Over the real numbers, however, not all matrix polynomials are triangularizable. We characterize those that are triangularizable and show that the real matrix polynomials that are not triangularizable over $\mathbb{R}[\lambda]$ are quasi-triangularizable with diagonal blocks of size 2×2 and 1×1. Our results hold for $n \times m$ real matrix polynomials with $n \leq m$ and extend in a non trivial way some recent results by Tisseur and Zaballa for square regular (i.e., $\det P(\lambda) \not\equiv 0$) quadratic matrix polynomials [10]. Our proofs concerning the reduction to triangular and quasi-triangular forms are constructive provided that the elementary divisors (finite and at infinity) of the original matrix $P(\lambda)$ are available. Since this is the only information that are used, we are solving an inverse problem: given a list of scalar polynomials, determine under what conditions they can be the elementary divisors (finite and at infinity) of a real or complex triangular matrix polynomial of fixed degree and, in that case, design a constructive procedure to obtain it.

The paper is organized as follows. In Section 2 we review basic materials used in the proofs of our main results. These include the Smith form, Möbius transforms, majorization of sequences of real numbers and conjugate sequences. Section 3 is concerned with the triangularization of matrix polynomials over algebraically closed fields and Section 4 treats the case of real matrix polynomials. In Section 5 the inverse polynomial problems solved in the previous sections are identified and the case of self-adjoint matrix polynomials is considered.

2. Background material

A matrix polynomial $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ is an $n \times m$ matrix with entries in $\mathbb{F}[\lambda]$, or equivalently, a polynomial with coefficient matrices,

$$P(\lambda) = \lambda^\ell A_\ell + \cdots + \lambda A_1 + A_0, \quad A_j \in \mathbb{F}^{n\times m}. \tag{1}$$

We assume that $A_\ell \neq 0$ and so the number ℓ is the *degree* of $P(\lambda)$.

A matrix polynomial $P(\lambda)$ is *regular* if it is square and its determinant is not identically zero and *singular* otherwise. The *rank* (also called *normal rank*) of a matrix polynomial $P(\lambda)$ is the order of the largest nonzero minor of $P(\lambda)$. We recall that a minor of order k of $P(\lambda)$ is the determinant of a $k \times k$ submatrix of $P(\lambda)$.

2.1. Smith form, partial multiplicity sequence and strong equivalence

A matrix polynomial is called *unimodular* if it is square and its determinant is a nonzero constant polynomial. Any $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ can be transformed to the canonical form

$$D(\lambda) = \left[\begin{array}{ccc|c} d_1 & & & \\ & \ddots & & 0_{r,m-r} \\ & & d_r & \\ \hline & 0_{n-r,r} & & 0_{n-r,m-r} \end{array}\right] = U(\lambda)P(\lambda)V(\lambda), \tag{2}$$

where $U(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$, $V(\lambda) \in \mathbb{F}[\lambda]^{m\times m}$ are unimodular, $r = \operatorname{rank} P(\lambda)$ and $d_1|\cdots|d_r$ are monic polynomials. Here, "$|$" stands for divisibility, thus $d_j|d_{j+1}$ means that d_j is a divisor of d_{j+1}. The diagonal matrix $D(\lambda)$ in (2) is called the *Smith form* of $P(\lambda)$ and it is unique. The nonzero scalar polynomials d_j on the diagonal of $D(\lambda)$ are called the *invariant factors* of $P(\lambda)$. They can be decomposed into irreducible factors over $\mathbb{F}[\lambda]$ as follows [1, Chap. VI, §3]:

$$\begin{array}{rcl} d_1 & = & \phi_1^{m_{11}} \dots \phi_s^{m_{1s}}, \\ d_2 & = & \phi_1^{m_{21}} \dots \phi_s^{m_{2s}}, \\ \vdots & & \vdots \\ d_r & = & \phi_1^{m_{r1}} \dots \phi_s^{m_{rs}}, \end{array} \tag{3}$$

where the ϕ_i are distinct monic polynomials irreducible over $\mathbb{F}[\lambda]$, and

$$0 \le m_{1j} \le m_{2j} \le \cdots \le m_{rj}, \quad j = 1{:}\, s, \tag{4}$$

are nonnegative integers. The factors $\phi_j^{m_{ij}}$ with $m_{ij} > 0$ are the *finite elementary divisors* of $P(\lambda)$ with *partial multiplicity* m_{ij}. Notice that when $\mathbb{F} = \mathbb{C}$, $\phi_j = \lambda - \lambda_j$ is linear and when $\mathbb{F} = \mathbb{R}$, ϕ_j is either linear or quadratic.

We denote by $\overline{\mathbb{F}}$ the algebraic closure of $\mathbb{F}$. A finite eigenvalue of a $P(\lambda)$ with rank r is any scalar $\lambda_0 \in \overline{\mathbb{F}}$ such that $\operatorname{rank} P(\lambda_0) < r$, or equivalently, a root of some elementary divisors ϕ_j in (3). The sequence m_{ij} $(i = 1{:}\, r)$ in (3)–(4) is called the *Segre characteristic* of the eigenvalue λ_0 as a root of ϕ_j. The *geometric multiplicity* of λ_0 is the number of nonzero m_{ij} and the *algebraic multiplicity* of λ_0 is $\sum_{i=1}^{r} m_{ij}$.

The *elementary divisors at infinity* of the matrix polynomial $P(\lambda)$ in (1) are defined as the elementary divisors of $\mathrm{rev}P(\lambda)$ at 0, where

$$\mathrm{rev}P(\lambda) = \lambda^{\ell} P\left(\lambda^{-1}\right) = A_0 \lambda^{\ell} + A_1 \lambda^{\ell-1} + \cdots + A_{\ell}$$

is the reversal of $P(\lambda)$. For a regular $P(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ of degree ℓ, the Smith form of $P(\lambda)$ provides the algebraic multiplicity of the eigenvalues at infinity via the degree deficiency in $\det P(\lambda)$, that is, $\ell n - \sum_{j=1}^{r} \deg\, d_j$ but it does not display their partial multiplicities. For singular polynomials, the Smith form does not detect the presence of elementary divisors at infinity but if $\operatorname{rank} P(\lambda) = r > \operatorname{rank} \mathrm{rev}P(0)$ then $P(\lambda)$ has elementary divisors at infinity.

For any $\lambda_0 \in \overline{\mathbb{F}}$, the invariant factors d_i of $P(\lambda)$ can be uniquely factored as

$$d_i = (\lambda - \lambda_0)^{a_i} p_i, \quad a_i \ge 0, \quad p_i(\lambda_0) \ne 0.$$

The sequence of exponents $a_1, a_2, \dots, a_r$ with $0 \le a_1 \le \cdots \le a_r$ is called the *partial multiplicity sequence of* P *at* λ_0 and is denoted by

$$\mathcal{J}(P, \lambda_0) = (a_1, a_2, \dots, a_r).$$

This sequence is usually all zeros unless λ_0 is an eigenvalue of $P(\lambda)$. If $\phi_i(\lambda_0) = 0$ then $\mathcal{J}(P, \lambda_0) = (m_{i1}, \dots, m_{ir})$ is the Segre characteristic corresponding to λ_0. The partial multiplicity sequence for $\lambda_0 = \infty$ is defined to be

$$\mathcal{J}(P, \infty) = \mathcal{J}(\mathrm{rev}(P), 0).$$

A matrix polynomial $P_1(\lambda) \in \mathbb{F}[\lambda]^{n \times m}$ is *equivalent* to $P_2(\lambda) \in \mathbb{F}[\lambda]^{n \times m}$ if it satisfies any one (and hence all) of the following equivalent properties.

(i) There exist unimodular matrices $U(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ and $V(\lambda) \in \mathbb{F}[\lambda]^{m \times m}$ such that $P_2(\lambda) = U(\lambda) P_1(\lambda) V(\lambda)$.

(ii) $P_1(\lambda)$ and $P_2(\lambda)$ have the same invariant factors, or equivalently, the same finite elementary divisors .

(iii) $\mathcal{J}(P_1, \lambda_0) = \mathcal{J}(P_2, \lambda_0)$ for any $\lambda_0 \in \overline{\mathbb{F}}$.

Moreover, $P_1(\lambda)$ is said to be *strongly equivalent* to $P_2(\lambda)$ if it is equivalent to $P_2(\lambda)$ and $\mathcal{J}(P_1, \infty) = \mathcal{J}(P_2, \infty)$. That (i) is equivalent to (ii) is shown in [1, p. 141] or [3, Thm. S1.11], whereas (ii) equivalent to (iii) follows from the definition of $\mathcal{J}(P, \lambda_0)$.

Unimodular transformations have the property that they preserve the partial multiplicities of the finite eigenvalues but not necessarily those of infinite eigenvalues as illustrated in the following example.

Example 2.1 The regular matrix polynomial $\mathrm{diag}(1, 1, \lambda)$ has one finite elementary divisor at zero and two linear elementary divisors at infinity. Multiplying $P(\lambda)$ on the left by the unimodular matrix $I_3 + \lambda e_1 e_2^T$, where e_i denotes the ith column of the identity matrix, yields

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

which is easily seen to have one finite elementary divisor at zero and one quadratic elementary divisor at infinity with partial multiplicity two. Hence $P(\lambda)$ and $\widetilde{P}(\lambda)$ are equivalent but not strongly equivalent. □

2.2. Möbius transformations

To any nonsingular matrix $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathbb{F}^{2\times 2}$ is associated a *Möbius function* $m_A : \mathbb{F} \cup \{\infty\} \to \mathbb{F} \cup \{\infty\}$ of the form

$$m_A(z) = \frac{az + b}{cz + d}, \qquad ad - bc \neq 0,$$

where

$$m_A(\infty) = \begin{cases} a/c & \text{if } c \neq 0, \\ \infty & \text{if } c = 0, \end{cases} \qquad m_A(-d/c) = \infty \ \text{ if } \ c \neq 0.$$

Let $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j \in \mathbb{F}[\lambda]^{n\times m}$ with $A_\ell \neq 0$, and let $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathbb{F}^{2\times 2}$ be nonsingular. Then the *Möbius transform of* $P(\lambda)$ *with respect to* A is the $n \times m$ matrix polynomial $\mathcal{M}_A(P)$ defined by

$$\mathcal{M}_A(P) = (c\lambda + d)^\ell P\big(m_A(\lambda)\big) = \sum_{j=0}^{\ell} A_j (a\lambda + b)^j (c\lambda + d)^{\ell - j}. \tag{5}$$

Note that the coefficient matrices of $\mathcal{M}_A(P)$ are simply linear combinations of those of $P(\lambda)$, so any zero structure—and in particular the triangular structure—is preserved.

Remark 2.2 For a given matrix polynomial $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ we will use the following technique to prove that it is strongly equivalent to a triangular or quasi-triangular matrix polynomial of the same degree:

(i) If $P(\lambda)$ has elementary divisors at infinity, we apply a Möbius transform to $P(\lambda)$ with Möbius function m_A such that $\mathcal{M}_A(P)$ only has finite elementary divisors and $(\mathcal{M}_{A^{-1}} \circ \mathcal{M}_A)(P) = P$ up to a product by a nonzero scalar. If $P(\lambda)$ has no eigenvalues at infinity then we take $A = I_2$ so that $\mathcal{M}_A(P) = P$.

(ii) We then show that $\mathcal{M}_A(P)$ is equivalent to a triangular or quasi-triangular matrix polynomial, $T(\lambda)$ say.

Notice that $\mathcal{M}_{A^{-1}}(T)$ is triangular or quasi-triangular according as $T(\lambda)$ is triangular or quasi-triangular, respectively. We claim that, provided that $\mathbb{F}$ is an infinite field, a Möbius function always exists that satisfies the two conditions of item (i) and $\mathcal{M}_{A^{-1}}(T)$ is strongly equivalent to $P(\lambda)$. In fact, let $a, c \in \mathbb{F}$, $c \neq 0$, such that a/c is not an eigenvalue of $P(\lambda)$ and take $b, d \in \mathbb{F}$ such that $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is nonsingular. Write $P(\lambda)$ as

$$P(\lambda) = \widetilde{A}_\ell (c\lambda - a)^\ell + \widetilde{A}_{\ell-1}(c\lambda - a)^{\ell-1} + \cdots + \widetilde{A}_1(c\lambda - a) + \widetilde{A}_0.$$

Given that a/c is not an eigenvalue of $P(\lambda)$ it follows that $\operatorname{rank} \widetilde{A}_0 = \operatorname{rank} P(a/c) = \operatorname{rank} P(\lambda)$. Also,

$$\mathcal{M}_A(P) = \widetilde{A}_0(c\lambda + d)^\ell + (bc - ad)\widetilde{A}_1(c\lambda + d)^{\ell-1} + \cdots + (bc - ad)^\ell \widetilde{A}_\ell$$

so that the leading coefficient of $\mathcal{M}_A(P)$ is $c^\ell \widetilde{A}_0$. Hence $\operatorname{rank} \operatorname{rev}\big(\mathcal{M}_A(P)\big)(0) = \operatorname{rank} \widetilde{A}_0 = \operatorname{rank} P(\lambda)$ and $\mathcal{M}_A(P)$ has no eigenvalues at infinity. Now, $m_{A^{-1}}(z) = (-dz+b)/(cz-a)$ and

$$\mathcal{M}_{A^{-1}}\big(\mathcal{M}_A(P)\big) = (c\lambda - a)^\ell \big(\mathcal{M}_A(P)\big)(m_{A^{-1}}(\lambda)) = (bc - ad)^\ell P(\lambda).$$

This proves our claim about the existence of a Möbius function that satisfies the two conditions of item (i) for each matrix polynomial $P(\lambda)$. The second part of the claim (that $P(\lambda)$ and $\mathcal{M}_{A^{-1}}(T)$ are strongly equivalent) is an immediate consequence of the following result which, in turn, is a particular case of [5, Thm. 5.3].

Theorem 2.3 *Let $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ and let $A \in \mathbb{F}^{2\times 2}$ be nonsingular such that $c \neq 0$ and $P(a/c) \neq 0$. Then, for any $\lambda_0 \in \mathbb{F} \cup \{\infty\}$,*

$$\mathcal{J}\big(\mathcal{M}_A(P), m_A^{-1}(\lambda_0)\big) = \mathcal{J}(P, \lambda_0).$$

This ends Remark 2.2. □

2.3. Majorization and conjugate sequence

When dealing with real matrix polynomials in Section 4 we will need the notion of *majorization* of sequences of real numbers.

Recall that if $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$ are two sequences of real numbers and $a_{[1]} \geq \cdots \geq a_{[m]}$ and $b_{[1]} \geq \cdots \geq b_{[m]}$ are the given sequences arranged in nonincreasing order then b is majorized by a (or a majorizes b) and we write $b \prec a$ if

$$\sum_{i=1}^{j} b_{[i]} \leq \sum_{i=1}^{j} a_{[i]}, \quad j = 1{:}\, m,$$

with equality for $j = m$.

We will need the following properties of majorization.

Lemma 2.4 (i) *If* $(b_1,\ldots,b_m)\prec(a_1,\ldots,a_m)$ *and* $(d_1,\ldots,d_m)\prec(c_1,\ldots,c_m)$ *then* $(b_1+d_1,\ldots,b_m+d_m)\prec(a_1+c_1,\ldots,a_m+c_m)$ *provided that* $a_1\geq\cdots\geq a_m$ *and* $c_1\geq\cdots\geq c_m$.

(ii) *Let* $(a_1,\ldots,a_m)$ *be a sequence of nonnegative integers and let* $|a|=mq+r$ *with* $0\leq r<m$ *be the Euclidean division of* $|a|=a_1+\cdots+a_m$ *by* m. *Then*

$$(\underbrace{q+1,\ldots,q+1}_{r\ times},q,\ldots,q)\prec(a_1,\ldots,a_m).$$

Proof. Item (i) is an immediate consequence of [9, Ch. 6, Prop. A.1.b] and item (ii) is trivial. □

Given a sequence of nonnegative integers $(a_1,\ldots,a_m)$, its *conjugate sequence* is $(a_1^*,\ldots,a_p^*)$, where

$$a_i^*=\#\{j\geq 1: a_j\geq i\}$$

and # stands for "cardinality of". We denote by $(a_1,\ldots,a_m)^*$ the conjugate sequence of $(a_1,\ldots,a_m)$. Notice that although the sequence $(a_1,\ldots,a_m)$ may be in no specific order, its conjugate sequence $(a_1,\ldots,a_m)^*$ always satisfies $a_1^*\geq a_2^*\geq\cdots$. Let $(m_{1j},\ldots,m_{nj})$ be the Segre characteristic of an eigenvalue λ_j of $P(\lambda)\in\mathbb{F}[\lambda]^{n\times m}$. Then the conjugate sequence $(m_{1j},\ldots,m_{nj})^*$ is called the *Weyr characteristic* corresponding to the eigenvalue λ_j.

If $(a_1,\ldots,a_m)$ and $(b_1,\ldots,b_m)$ are two sequences of real numbers we will denote by $(a_1,\ldots,a_m)\cup(b_1,\ldots,b_m)$ the sequence obtained by putting together the elements of both sequences, that is,

$$(a_1,\ldots,a_m)\cup(b_1,\ldots,b_m)=(a_1,\ldots,a_m,b_1,\ldots,b_m).$$

A proof of the following lemma can be found, for example, in [4, p. 5-6].

Lemma 2.5 *Let* $(a_1,\ldots,a_m)$ *and* $(b_1,\ldots,b_m)$ *be sequences of nonnegative integers. Then*

$$(a_1,\ldots,a_m)\prec(b_1,\ldots,b_m)\quad\Longleftrightarrow\quad(b_1,\ldots,b_m)^*\prec(a_1,\ldots,a_m)^*$$

and if $a_1\geq\cdots\geq a_m$ *and* $b_1\geq\cdots\geq b_m$ *then*

$$(a_1+b_1,\ldots,a_m+b_m)^*=(a_1,\ldots,a_m)^*\cup(b_1,\ldots,b_m)^*.\tag{6}$$

Let $a=(a_1,\ldots,a_m)$ and $b=(b_1,\ldots,b_m)$ be sequences of nonnegative integers. Sequence a is said to be obtained from sequence b by an elementary transformation if the sequences are identical except for indices j and k ($j<k$) where $a_j=b_j+1$ and $a_k=b_k-1$. The following result will be used in the proof of Lemma 4.6. An explicit proof can be found in [6] (see also [4, p. 9] or [9, Sec. 2B]).

Lemma 2.6 *Let* $a=(a_1,\ldots,a_m)$ *and* $b=(b_1,\ldots,b_m)$ *be sequences of nonnegative integers arranged in nonincreasing order* ($a_1\geq a_2\geq\cdots$, $b_1\geq b_2\geq\cdots$) *such that* $a\neq b$. *Then* $b\prec a$ *if and only if* b *can be obtained from* a *by a finite number of elementary transformations.*

3. Triangularization of matrix polynomials over algebraically closed fields

To make the text more readable, we say that a matrix is triangular if all entries (i, j) such that $i > j$ are zero; this includes trapezoidal matrices. We say that $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ is *triangularizable* over $\mathbb{F}[\lambda]$ if it is strongly equivalent to a triangular matrix polynomial $T(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ of the same degree.

We start with a deflation procedure which will be used to construct upper triangular matrix polynomials with diagonal entries of a specified degree.

Lemma 3.1 *Let $d_1|\cdots|d_n$ be monic polynomials with coefficients in $\mathbb{F}$ and define $\ell_j :=$ $\deg d_j$. Assume that for a given positive integer q and a pair of indices (i, j) such that $\ell_i \le q < \ell_j$, there is a polynomial s with $\deg s < \ell_j$ such that $d_{k-1}|sd_i|d_k$ for some index $k \le j$. Then $D(\lambda) = \operatorname{diag}(d_1, \ldots, d_n)$ is equivalent to $\widetilde{D}(\lambda) + d_i e_{k-1} e_j^T$, where e_i denotes the ith column of the $n \times n$ identity matrix and*

$$\widetilde{D}(\lambda) = \operatorname{diag}(\underbrace{d_1, \ldots, d_{i-1}, d_{i+1}, \ldots d_{k-1}}_{k-2 \text{ terms}}, sd_i, \underbrace{d_k, \ldots, d_{j-1}}_{j-k \text{ terms}}, -d_j/s, \underbrace{d_{j+1}, \ldots, d_n}_{n-j \text{ terms}}).$$

Proof. We obtain $\widetilde{D}(\lambda) + d_i e_{k-1} e_j^T$ by performing the following elementary transformations on $D(\lambda)$:

(i) add to column j column i multiplied by s,
(ii) add to row j row i multiplied by $-d_j/(sd_i)$,
(iii) permute columns i and j,
(iv) successively interchange rows t and $t + 1$ for $t = i{:}\,k - 1$, so that rows i, $i + 1$, ..., $k - 2$, $k - 1$ of the new matrix are rows $i + 1$, $i + 2$, ..., $k - 1$ and i, respectively, of the former one.
(v) permute columns i to $k - 1$ in the same way as the rows in (iv). □

A version of the next result appears for $\mathbb{F} = \mathbb{C}$ in the proof of a theorem by Gohberg, Lancaster and Rodman on the inverse problem for linearizations [3, proof of Thm. 1.7].

Lemma 3.2 *Let $d_1|\cdots|d_n$ be monic polynomials with coefficients in an algebraically closed field $\mathbb{F}$. There exists a monic triangular matrix polynomial $P(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ of degree ℓ and with $d_1, \ldots, d_n$ as invariant factors if and only if $\sum_{j=1}^{n} \deg d_j = \ell n$.*

Proof. The "only if" part is trivial.

Suppose that there are monic polynomials $d_1|\cdots|d_n$ such that $\sum_{j=1}^{n} \ell_j = \ell n$, where $\ell_j = \deg d_j$ and let $D(\lambda) = \operatorname{diag}(d_1, \ldots, d_n)$.

If $\ell_1 = \ell$ then $\ell_i = \ell$ for $i = 2{:}\,n$. Hence $D(\lambda)$ is a monic triangular (in fact diagonal) matrix polynomial of degree ℓ and the construction is done.

If $\ell_1 < \ell$ then $\ell_n > \ell$ and so $\ell_1 < \ell_1 + \ell_n - \ell < \ell_n$, then there is a monic polynomial s of degree $\ell_n - \ell$ such that $d_{k-1}|sd_1|d_k$ for some index k, $1 < k \le n$. By Lemma 3.1,

$D(\lambda)$ is equivalent to

$$T_1(\lambda) = \begin{bmatrix} d_2 & & & & & & & \\ & \ddots & & & & & & \\ & & d_{k-1} & & & & & \\ & & & d_1 s & \cdots & & \cdots & d_1 \\ & & & & d_k & & & \vdots \\ & & & & & \ddots & & \\ & & & & & & d_{n-1} & \vdots \\ & & & & & & & -d_n/s \end{bmatrix} = \begin{bmatrix} D_1(\lambda) & t_1 \\ 0 & -d_n/s \end{bmatrix},$$

where $D_1(\lambda) = \operatorname{diag}(d_2, \ldots, d_{k-1}, d_1 s, \ldots, d_{n-1}) = \operatorname{diag}(d_1^{(1)}, \ldots, d_{n-1}^{(1)})$ is in Smith form. If $\deg D_1 > \ell$ then we look for a new index k and a monic polynomial s_1 of degree $\deg d_{n-1}^{(1)} - \ell$ such that $d_{k-1}^{(1)} | s_1 d_1^{(1)} | d_k^{(1)}$. Apply the elementary transformations of Lemma 3.1 to the whole matrix $T_1(\lambda)$ so that the obtained matrix has the degree ℓ polynomial $-d_{n-1}^{(1)}/s_1$ as the $(n-1)$st diagonal entry. Notice that these elementary transformations can modify the off-diagonal elements of the nth column of $T_1(\lambda)$ but the degree ℓ polynomial $-d_n/s$ in position (n, n) remains unchanged. We repeat this deflation process until all diagonal entries are of degree ℓ. The resulting matrix polynomial $T(\lambda)$ is upper triangular but not necessarily monic or of degree ℓ. We can, however, use the diagonal entries to eliminate off-diagonal terms of degree larger than $\ell - 1$ in the following way: if $\deg t_{ij} > \ell - 1$ for $i < j$ then $t_{ij} = t_{ii} b_{ij} + c_{ij}$ for some b_{ij} and c_{ij} such that $\deg c_{ij} < \deg t_{jj} = \ell$ or $c_{ij} = 0$. Then adding to the ith column of $T(\lambda)$ the jth column multiplied by $-b_{ij}$ reduces the (i, j) entry to zero or to some polynomial with degree strictly less than ℓ. Note that to reduce the degree of all the off-diagonal entries of $T(\lambda)$ we must work bottom up. We then end up with a monic upper triangular matrix polynomial of degree ℓ. □

We notice that for a given list of invariant factors $d_1 | \cdots | d_n$ such that $\sum_{j=1}^{n} \deg d_j = \ell n$, there may be more than one monic triangular matrix polynomial of degree ℓ having $d_1, \ldots, d_n$ as invariant factors as illustrated by the following example.

Example 3.3 Let $\mathbb{F} = \mathbb{C}$ and $d_1(\lambda) = 1$, $d_2(\lambda) = (\lambda^2+1)\lambda^2$ and $d_3(\lambda) = (\lambda^2+1)\lambda^2(\lambda-1)$ be given. Note that $d_1 | d_2 | d_3$ and $\sum_{j=1}^{3} \deg d_j = 9$ so that by Lemma 3.2 there is a 3×3 monic triangular matrix polynomial $T(\lambda)$ of degree 3 with d_1, d_2, d_3 as invariant factors. To construct $T(\lambda)$ we first look for a degree $\deg d_3 - \ell = 2$ polynomial s and an index k such that $d_{k-1} | s d_1 | d_k$. We have to take $k = 2$ and for s we can choose either $s = \lambda^2$ or $s = \lambda^2 + 1$. Both choices yield a different upper triangular cubic matrix polynomial.

- If $s = \lambda^2$ then by Lemma 3.1 $D(\lambda) = \operatorname{diag}(d_1, d_2, d_3)$ is equivalent to $T_1(\lambda) = \operatorname{diag}\big(\lambda^2, (\lambda^2+1)\lambda^2, -(\lambda^2+1)(\lambda-1)\big) + e_2 e_3^T$. Then we look for a degree 1 polynomial s_1 such that $\lambda^2 | s_1 \lambda^2 | \lambda^2(\lambda^2 + 1)$. We can take either $s_1 = \lambda - i$ or $s_1 = \lambda + i$. The latter choice yields

$$T_{(a)}(\lambda) = \begin{bmatrix} \lambda^2(\lambda+i) & \lambda^2 & 1 \\ & -\lambda^2(\lambda - i) & -(\lambda - i) \\ & & -(\lambda^2+1)(\lambda-1) \end{bmatrix}.$$

- The choice $s = \lambda^2 + 1$ in (b) leads to

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

Note that $T_{(b)}(\lambda)$ is real. $\square$

We are now ready to state the main result of this section, which generalizes [5, Thm. 9.3].

Theorem 3.4 *For an algebraically closed field $\mathbb{F}$, any $P(\lambda) \in \mathbb{F}[\lambda]^{n\times m}$ with $n \le m$ is triangularizable.*

Proof. Assume $P(\lambda)$ has degree ℓ and rank r. Since $\mathbb{F}$ is algebraically closed, it is infinite and by Remark 2.2 there is a Möbius function m_A induced by a nonsingular matrix $A \in \mathbb{F}^{2\times 2}$ such that if $\mathcal{M}_A(P)$ is triangularizable then $P(\lambda)$ is strongly equivalent to a degree ℓ triangular matrix. We will show that $\mathcal{M}_A(P)$ is triangularizable. By [5, Thm. 6.6], $\operatorname{rank} \mathcal{M}_A(P) = \operatorname{rank} P = r$. Let

$$D(\lambda) = \operatorname{diag}(d_1, \ldots, d_r, 0, \ldots, 0) \in \mathbb{F}[\lambda]^{n\times m}$$

be the Smith form of $\mathcal{M}_A(P)$. Because all minors of $\mathcal{M}_A(P)$ of order r are of degree at most $r\ell$, and because the greatest common divisor of all such minors is invariant under unimodular transformations ([1, p. 140],[3, Thm. S1.2]), we know that $\sum_{j=1}^{r} \deg d_j \le r\ell$. We consider two cases.

Case 1. $\sum_{j=1}^{r} \deg d_j = r\ell$. By Lemma 3.2, the regular part $\operatorname{diag}(d_1, \ldots, d_r) \in \mathbb{F}[\lambda]^{r\times r}$ of $D(\lambda)$ is equivalent to an $r \times r$ upper triangular matrix polynomial of degree ℓ. Hence $\mathcal{M}_A(P)$ is triangularizable.

Case 2. $\sum_{j=1}^{r} \deg d_j < r\ell$. If $r = m$, $\mathcal{M}_A(P)$ is square and regular, that is, $\mathcal{M}_A(P)$ has $m\ell$ eigenvalues, a contradiction. Thus $r < m$. Starting with $\widetilde{T}_0(\lambda) = \operatorname{diag}(d_1, \ldots, d_r) \in \mathbb{F}[\lambda]^{r\times r}$, we follow the construction in Lemma 3.2 until we reach a step, say $r - k$, such that the matrix polynomial has the form

$$\widetilde{T}_{r-k}(\lambda) = \begin{bmatrix} \widetilde{d}_1 & & & & * & \cdots & * \\ & \widetilde{d}_2 & & & \vdots & \ddots & \vdots \\ & & \ddots & & \vdots & \ddots & \vdots \\ & & & \widetilde{d}_k & * & \cdots & * \\ & & & & * & \cdots & * \\ & & & & & \ddots & \vdots \\ & & & & & & * \end{bmatrix}, \qquad (7)$$

where $\deg \widetilde{d}_j < \ell$ for $j = 1\colon k$ and the $*$'s on the diagonal denote polynomials of degree ℓ. Now, as in the proof of Lemma 3.2, suppose we have applied an appropriate sequence of elementary transformations to reduce the degree of the off-diagonal entries of $\widetilde{T}_{r-k}(\lambda)$

to polynomials of degree strictly less than ℓ. Then $\mathcal{M}_A(P)$ is equivalent to the upper triangular matrix polynomial of degree ℓ,

$$T_{r-k}(\lambda) = \begin{bmatrix} \widetilde{T}_{r-k}(\lambda) & 0_{r,m-r} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{bmatrix}.$$

Note that $\widetilde{T}_{r-k}(\lambda)$ has a singular leading coefficient implying that $\operatorname{rank}(\operatorname{rev} T_{r-k}(0)) < r$. This means that $T_{r-k}(\lambda)$ has elementary divisors at infinity, and it is hence not strongly equivalent to $\mathcal{M}_A(P)$. We now show how to remove the elementary divisors at infinity while maintaining the upper triangular form. Note that since $r < m$, the last column of $T_{r-k}(\lambda)$ is a zero column. Thus permuting the columns according to $(1, 2, \ldots, n)$ (cycle notation), preserves the triangular structure. Define g_i through $\deg(\lambda^{g_i}\widetilde{d}_i) = \ell$. Using a sequence of k right elementary operations we obtain the equivalent matrix polynomial

$$T(\lambda) = \begin{bmatrix} \lambda^{g_1}\widetilde{d}_1 & \widetilde{d}_1 & & & & & * & \cdots & * & & & \\ & \lambda^{g_2}\widetilde{d}_2 & \widetilde{d}_2 & & & & \vdots & \ddots & \vdots & & & \\ & & \ddots & & & & \vdots & \ddots & \vdots & & & \\ & & & \lambda^{g_{k-1}}\widetilde{d}_{k-1} & \widetilde{d}_{k-1} & & * & \cdots & * & & & \\ & & & & \lambda^{g_k}\widetilde{d}_k & \widetilde{d}_k & * & \cdots & * & & & \\ & & & & & & * & \cdots & * & & & \\ & & & & & & & \ddots & \vdots & & & \\ & & & & & & & & * & & & \\ & & & & & & & & & 0 & & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & 0 \end{bmatrix}, \quad (8)$$

which is still upper triangular and of degree ℓ. It now remains to show that $\operatorname{rev} T(\lambda) = \lambda^{\ell} T(\frac{1}{\lambda})$ has no elementary divisor at zero. For this we write $\widetilde{d}_i$ in factorized form

$$\widetilde{d}_i = \begin{cases} \prod_{j=1}^{\ell-g_i}(\lambda - \lambda_{ij}) & \text{if } \ell > g_i, \\ 1 & \text{otherwise} \end{cases}$$

and let

$$c_i = \begin{cases} \prod_{j=1}^{\ell-g_i}(1 - \lambda\lambda_{ij}) & \text{if } \ell > g_i, \\ 1 & \text{otherwise.} \end{cases}$$

Then

$$
\text{rev } T(\lambda) = \begin{bmatrix}
c_1 & \lambda^{g_1} c_1 & & & & & & * & \cdots & * & & & \\
& c_2 & \lambda^{g_2} c_2 & & & & & \vdots & \ddots & \vdots & & & \\
& & & \ddots & & & & \vdots & \ddots & \vdots & & & \\
& & & c_{k-1} & \lambda^{g_{k-1}} c_{k-1} & & & * & \cdots & * & & & \\
& & & & c_k & \lambda^{g_k} c_k & & * & \cdots & * & & & \\
& & & & & & & * & \cdots & * & & & \\
& & & & & & & & \ddots & \vdots & & & \\
& & & & & & & & & * & & & \\
& & & & & & & & & & 0 & & \\
& & & & & & & & & & & \ddots & \\
& & & & & & & & & & & & 0
\end{bmatrix}.
$$

By construction, the polynomials represented as $*$'s on the diagonal of rev $T(\lambda)$ do not annihilate when evaluated at zero, and similarly, $c_i(0) \neq 0$. Therefore rank $\big((\text{rev } T)(0)\big) = r$ and so rev $T(\lambda)$ has no elementary divisors at zero. Hence, the upper triangular matrix polynomial $T(\lambda)$ in (8) is strongly equivalent to $\mathcal{M}_A(P)$, that is, $\mathcal{M}_A(P)$ is triangularizable. □

Remark 3.5 If $n > m$, we cannot always triangularize; see Example 3.7. The construction fails when we can no longer guarantee that $r < m$, implying that we cannot permute the nonzero part of the matrix one step to the right. However, using similar arguments, we can in this case ensure that $r < n$. By permuting the nonzero part of the matrix one step down instead of one step to the right, we can still build a matrix polynomial with the correct elementary divisors; the matrix will have Hessenberg structure (all entries (i, j) are zero for $i + 1 > j$). □

Let us illustrate Theorem 3.4 with the following example taken from [11, Ex. 1].

Example 3.6 The quadratic matrix polynomial

$$
Q(\lambda) = \begin{bmatrix} \lambda^2 + \lambda & 4\lambda^2 + 3\lambda & 2\lambda^2 \\ \lambda & 4\lambda - 1 & 2\lambda - 2 \\ \lambda^2 - \lambda & 4\lambda^2 - \lambda & 2\lambda^2 - 2\lambda \end{bmatrix}
$$

has Smith form

$$
D(\lambda) = \begin{bmatrix} 1 & -1 & -1 \\ -\lambda & 1+\lambda & \lambda \\ 0 & -\lambda & 1 \end{bmatrix} Q(\lambda) \begin{bmatrix} 1 & -3 & 6 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},
$$

and since $\det(Q(\lambda)) = \det(D(\lambda)) \equiv 0$, $Q(\lambda)$ is singular. This matrix polynomial has only one finite elementary divisor, $\lambda - 1$. Note that rank(rev $Q(0)) = 1 < 2 = \operatorname{rank} Q(\lambda)$ so $Q(\lambda)$ has elementary divisors at infinity. Now the Smith form of rev $Q(\lambda)$, given by

$$
\widetilde{D}(\lambda) = \operatorname{diag}(1, \lambda^2(\lambda - 1), 0),
$$

reveals an elementary divisors at infinity for $Q(\lambda)$ with partial multiplicity 2.

As 0 is not eigenvalue of $Q(\lambda)$, bearing in mind Remark 2.2, we can take $a = 0$ and $c = 1$. In fact, if $A = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ then $\mathcal{M}_A(Q) = \text{rev } Q(\lambda)$ has no elementary divisors at infinity and we can follow the proof of Theorem 3.4 with this matrix. We start the triangularization process with the submatrix $\text{diag}(1, \lambda^2(\lambda-1))$. Lemma 3.1 with $(i,j) = (1,2)$, $k = 2$ and $s(\lambda) = \lambda - 1$ yields $\widetilde{T}(\lambda) = \text{diag}(\lambda - 1, -\lambda^2) + e_1 e_2^T$. Hence, $\widetilde{D}(\lambda)$ is equivalent to

$$T_1(\lambda) = \begin{bmatrix} \lambda - 1 & 1 & 0 \\ 0 & -\lambda^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrix polynomial $T_1(\lambda)$, which is quadratic and upper triangular, has a singular leading coefficient indicating that $T_1(\lambda)$ and $\mathcal{M}_A(Q)$ are not strongly equivalent. Its elementary divisors at infinity can be removed as described in the proof of Theorem 3.4. First we permute the columns according to (1,2,3), to obtain:

$$T_2(\lambda) = \begin{bmatrix} 0 & \lambda - 1 & 1 \\ 0 & 0 & -\lambda^2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Second, multiply the second column by λ and add it to the first one. This yields:

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

which is strongly equivalent to $\mathcal{M}_A(Q)$. Finally,

$$\mathcal{M}_{A^{-1}}(T) = \text{rev } T(\lambda) = \begin{bmatrix} \lambda + 1 & -\lambda^2 + \lambda & \lambda^2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

is quadratic, triangular and strongly equivalent to $Q(\lambda)$. □

The next example shows that we cannot generalize Theorem 3.4 to include the case $n > m$.

Example 3.7 The quadratic $Q(\lambda) = \left[\begin{smallmatrix} \lambda \\ \lambda^2 \end{smallmatrix}\right]$ has for Smith form $D(\lambda) = \left[\begin{smallmatrix} \lambda \\ 0 \end{smallmatrix}\right]$. Since $\text{rank}((\text{rev } Q)(0)) = 1 = \text{rank}\, Q(\lambda)$, there is no elementary divisor at infinity and $Q(\lambda)$ has only one finite elementary divisor λ. Now the triangular matrix polynomial $\left[\begin{smallmatrix} q(\lambda) \\ 0 \end{smallmatrix}\right]$ has for elementary divisor λ if and only if $q(\lambda) = \lambda$ but then $\text{rev } T(\lambda) = \left[\begin{smallmatrix} \lambda \\ 0 \end{smallmatrix}\right]$, which shows that $T(\lambda)$ has an elementary divisor at infinity. Hence $Q(\lambda)$ is not triangularizable. □

4. Real matrix polynomials

We now concentrate on the non-algebraically closed field $\mathbb{R}$. Although some real matrix polynomials are triangularizable over $\mathbb{R}[\lambda]$ (see for instance Examples 2.1 and 3.3) it is shown in [10] that not all quadratic real matrix polynomials are triangularizable over $\mathbb{R}[\lambda]$.

We say that $P(\lambda) \in \mathbb{R}[\lambda]^{n\times m}$ is *quasi-triangularizable* over $\mathbb{R}[\lambda]$ if it is strongly equivalent to a quasi-triangular matrix polynomial $T(\lambda) \in \mathbb{R}[\lambda]^{n\times m}$ of the same degree and with 1×1 and 2×2 diagonal blocks. We begin by showing that any $P(\lambda) \in \mathbb{R}[\lambda]^{n\times m}$ with $n \le m$ is quasi-triangularizable. Section 4.2 is devoted to characterize those real matrix polynomials that are triangularizable over $\mathbb{R}[\lambda]$.

4.1. Quasi-triangularizable real matrix polynomials

We start with an analog of Lemma 3.2 for real matrix polynomials.

Theorem 4.1 *Let $d_1|\cdots|d_n$ be monic polynomials with coefficients in $\mathbb{R}$. There exists a quasi-triangular matrix polynomial $T(\lambda) \in \mathbb{R}[\lambda]^{n\times n}$ of degree ℓ with nonsingular leading matrix coefficient and with $d_1,\dots, d_n$ as invariant factors if and only if $\sum_{j=1}^{n} \deg d_j = \ell n$.*

Proof. We only prove the"if" part as the "only if" part is trivial. Suppose that there are monic polynomials $d_1|\cdots|d_n$ such that $\sum_{j=1}^{n} \ell_j = \ell n$, where $\ell_j = \deg d_j$ and $\ell_1 < \ell$. Let $T_n(\lambda) = \operatorname{diag}(d_1,\dots,d_n)$. We start by constructing an upper triangular matrix polynomial equivalent to $T_n(\lambda)$ whose diagonal entries have degree either ℓ or $\ell+1$ or $\ell-1$.

Assume that there is a pair of indices (i,j) such that $\ell_i < \ell < \ell_j$ for which there is a real polynomial s of degree $\ell_j - \ell$ such that $d_{k-1}|sd_i|d_k$ for some index $k \le j$. It is important for us to remark that the existence of such a real polynomial s is equivalent to the existence of a real polynomial r of degree $\ell_j-\ell-\deg\,(d_{k-1}/d_i)$ such that $r|(d_k/d_{k-1})$. Then by Lemma 3.1, $T_n(\lambda)$ is equivalent to a matrix polynomial of the form

$$\begin{bmatrix} \ddots & & & & & & \\ & sd_i & & d_i & & & \\ & & \ddots & & & & \\ & & & -d_j/s & & & \\ & & & & d_{j+1} & & \\ & & & & & \ddots & \\ & & & & & & d_n \end{bmatrix}.$$

By permuting rows and columns we can move the degree ℓ polynomial $-d_j/s$ to the lower right corner while keeping the upper triangular form and the $(n-1)\times(n-1)$ leading submatrix in Smith form. Specifically,

$$\begin{bmatrix} \ddots & & & & & & \\ & sd_i & & & & & d_i \\ & & \ddots & & & & \\ & & & d_{j+1} & & & \\ & & & & \ddots & & \\ & & & & & d_n & \\ & & & & & & -d_j/s \end{bmatrix}.$$

We say that polynomial $-d_j/s$ has been *deflated* to the (n,n) position. We repeat this deflation procedure for all possible pairs of indices (i,j) satisfying the above conditions.

Also, by means of appropriate permutations of rows and columns which do not introduce nonzero entries in the lower triangular part of the matrix, we can move (deflate) all the diagonal entries of degree ℓ down to the lower right part of the matrix. We end up with a matrix polynomial of the following form

$$T_p(\lambda) = \begin{bmatrix} c_1 & & & & * & \cdots & * \\ & c_2 & & & \vdots & \ddots & \vdots \\ & & \ddots & & \vdots & \ddots & \vdots \\ & & & c_p & * & \cdots & * \\ & & & & * & \cdots & * \\ & & & & & \ddots & \vdots \\ & & & & & & * \end{bmatrix},$$

where the c_i's are polynomials such that $c_1|c_2|\cdots|c_p$ (i.e., the $p \times p$ leading principal submatrix of $T_p(\lambda)$ is in Smith form), $\sum_{i=1}^{p} \deg\, c_i = p\ell$, and the $*$'s on the diagonal denote polynomials of degree ℓ. We redefine ℓ_i to be the degree of c_i. Note that if $\ell_1 = \ell$ then $T_p(\lambda)$ has all its diagonal entries of degree ℓ.

Suppose that $\ell_1 < \ell$, which implies that $p \geq 2$. We show that if we cannot deflate a degree ℓ polynomial, then we can consecutively deflate two polynomials of degree $\ell+1$ and $\ell-1$, respectively. If $p = 2$ and there is no real polynomial s of degree $\ell_2 - \ell$ such that $c_1|sc_1|c_2$ then there is no real polynomial r of degree $\ell_2 - \ell$ such that $r|(c_2/c_1)$. This implies that c_2/c_1 has no linear factor and $\ell_2 - \ell$ is odd. Thus there is a degree $\ell_2 - (\ell+1)$ polynomial s_1 such that $c_1|s_1c_1|c_2$. Then using the procedure described in Lemma 3.1 with $(i,j) = (1,2)$, $k = 2$ and s_1, we deflate a degree $\ell+1$ polynomial in position $(2,2)$ leaving s_1c_1 of degree $\ell-1$ in position $(1,1)$.

We now assume that $p > 2$ and that for any pair of indices (i,j) with $\ell_i < \ell < \ell_j$ we cannot find a real polynomial s of degree $\ell_j - \ell$ such that $c_{k-1}|sc_i|c_k$ for any index k, $i < k \leq j$. Then there is no real polynomial r of degree $\ell_j - \ell - \deg(c_{k-1}/c_i)$ such that $r|(c_k/c_{k-1})$. It follows then that c_k/c_{k-1} contains no linear factors and $\ell_j - \ell$ $(= \deg\, s)$ and $\ell_{k-1} - \ell_i$ $(= \deg(c_{k-1}/c_i))$ have different parity. We consider three cases.

Case 1. $\ell_2 < \ell < \ell_{p-1}$. Then $\ell_1 < \ell < \ell_p$ and there is a degree $\ell_p - (\ell+1)$ polynomial s_1 such that for some index $k \leq p$, $c_{k-1}|s_1c_1|c_k$. We use the procedure described in Lemma 3.1 with $(i,j) = (1,p)$, s_1 and the index k to deflate the degree $\ell+1$ polynomial $-c_p/s_1$ to position (p,p). This produces a matrix $T_{p-1}(\lambda)$, whose $(p-1)\times(p-1)$ leading principal submatrix is a Smith form still having c_2 and c_{p-1} as diagonal elements. We then repeat the argument using c_2 and c_{p-1} and a polynomial s_2 of degree $\ell_{p-1} - (\ell-1)$ such that $c_{k-1}|s_2c_2|c_k$ for some index $k \leq p-1$ to deflate the degree $\ell-1$ polynomial $-c_{p-1}/s_2$ to position $(p-1,p-1)$.

Case 2. $\ell_2 > \ell$. As explained above, there are no linear factors in c_2/c_1, so ℓ_1 and ℓ_2 have the same parity. Further, $\ell_3 - \ell$ is odd and c_3/c_2 contains no linear factors (otherwise there would be a real polynomial r of degree $\ell_3 - \ell - \deg(c_3/c_1)$ such that $r|(c_3/c_2)$ and so $c_2|sc_1|c_3$ for some polynomial of degree $\ell_3 - \ell$; a contradiction). Hence ℓ_1, ℓ_2 and ℓ_3 have the same parity. Using c_1, c_2 and a polynomial s_1 of even degree $\ell_2 - \ell - 1$ such that $c_1|s_1c_1|c_2$, we apply the procedure describe in Lemma 3.1 to deflate the degree $\ell+1$ polynomial $-c_2/s_1$ to the (p,p) diagonal entry. This produces a triangular matrix whose $(p-1)\times(p-1)$ leading principal submatrix is $\mathrm{diag}(s_1c_1, c_3, \ldots, c_p)$. Note that $\deg(c_3/(s_1c_1))$ is even and $\ell_3 - \ell$ is odd. Note also that $\ell_2 > \ell$ implies that $2\ell \ge \ell_1 + \ell_2$ and so $\ell_3 - \deg(s_1c_1) = \ell_3 - \ell_2 + \ell + 1 - \ell_1 \ge \ell_3 - \ell + 1$. Hence, we can always find a polynomial s_2 of degree $\ell_3 - \ell + 1$ such that $s_1c_1|s_2s_1c_1|c_3$ and deflate the degree $\ell - 1$ polynomial $-c_3/s_2$ to position $(p-1, p-1)$.

Case 3. $\ell_{p-1} < \ell$. The condition $\sum_{i=1}^{p} \deg\, c_i = p\ell$ implies $\ell_p - \ell > 0$. Furthermore, $\ell_p - \ell$ is odd and $\ell_p - \ell_{p-1}$ is even because otherwise there would be a real polynomial s of degree $\ell_p - \ell$ such that $s|(c_p/c_{p-1})$ and so $c_{p-1}|sc_{p-1}|c_p$. This would imply that the degree ℓ polynomial $-c_p/s$ could be deflated by using Lemma 3.1. Now we use c_{p-1}, c_p and a polynomial s_1 of even degree $\ell_p - \ell - 1$ satisfying $c_{p-1}|s_1c_{p-1}|c_p$ to deflate the degree $\ell + 1$ polynomial $-c_p/s_1$ to position (p, p). We are left with $\mathrm{diag}(c_1, \ldots, c_{p-2}, s_1c_{p-1})$. Notice that $\ell_p + \ell_{p-1} \ge 2\ell$. If $\deg(s_1c_{p-1}) = \ell - 1$ (i.e., $\ell_p + \ell_{p-1} = 2\ell$) we have already deflated two polynomials of degrees $\ell + 1$ and $\ell - 1$ to positions (p, p) and $(p-1, p-1)$. Otherwise, we look for a real polynomial s_2 of degree $\deg(s_1c_{p-1}) - \ell + 1 = \ell_{p-1} + \ell_p - 2\ell$ such that $c_{p-2}|s_2c_{p-2}|s_1c_{p-1}$. Note that $\deg\, s_2$ is even so we can always construct it. Using the procedure described in Lemma 3.1 with $(i, j) = (p-2, p-1)$, $k = p - 1$ and s_2, we deflate the degree $\ell - 1$ polynomial $-s_1c_{p-1}/s_2$ to the $(p-1, p-1)$ position.

We repeat these processes until all diagonal entries of the matrix polynomials are of degree either ℓ or $\ell + 1$ or $\ell - 1$. It now remains to transform the resulting upper triangular matrix polynomial to quasi-triangular form with entries of degree at most ℓ. We assume that the diagonal entries have been scaled to become monic. Now suppose that all the entries below row i are of degree ℓ or less. If the ith diagonal entry is of degree ℓ then we use the procedure described at the end of Lemma 3.2 to reduce the entries in row i except the (i, i) entry to polynomials of degree strictly less than ℓ. If the ith diagonal entry is of degree $\ell + 1$ then the $(i-1)$th diagonal entry is of degree $\ell - 1$. We use the procedure described at the end of Lemma 3.2 to reduce the entries in rows i and $i - 1$ except those on the diagonal to polynomials of degree at most ℓ for row i and polynomials of degree at most $\ell - 2$ for row $i - 1$. Hence rows $i - 1$ and i look like

$$\begin{bmatrix} 0 & \cdots & 0 & \tilde{d}_{i-1} & \diamondsuit & \diamondsuit & \cdots & \diamondsuit \\ 0 & \cdots & 0 & 0 & \tilde{d}_i & \times & \cdots & \times \end{bmatrix}, \qquad \begin{array}{ll} \deg \tilde{d}_{i-1} = \ell - 1, & \deg \diamondsuit \le \ell - 2, \\ \deg \tilde{d}_i = \ell + 1, & \deg \times \le \ell. \end{array}$$

Next, we add λ times row $i - 1$ to row i, and then $-\lambda$ times column $i - 1$ to column i leading to

$$\begin{bmatrix} 0 & \cdots & 0 & \tilde{d}_{i-1} & * & * & \cdots & * \\ 0 & \cdots & 0 & \lambda\tilde{d}_{i-1} & e_i & * & \cdots & * \end{bmatrix},$$

where $\deg\, e_i \le \ell$ and no entry hiding behind the asterisks is of degree larger than ℓ. Moving upwards through the matrix in this way we end up with an $n \times n$ real upper quasi-triangular matrix polynomial $T(\lambda) = \sum_{j=0}^{\ell} \lambda^j T_j$ of degree ℓ with $d_1| \ldots |d_n$ as invariant factors. Since $\sum_{j=1}^{n} \deg\, d_j = \ell n$, $T(\lambda)$ has ℓn finite eigenvalues, implying that leading matrix coefficient T_ℓ is nonsingular. □

Example 4.2 Let $D(\lambda) = \mathrm{diag}\big(1, (\lambda^2+1)^2, (\lambda^2+1)^2, (\lambda^2+1)^2\big) = \mathrm{diag}(d_1, d_2, d_3, d_4)$ be the Smith form of a 4×4 cubic polynomial $P(\lambda)$. We follow the proof of Lemma 4.1 to construct a quasi-triangular polynomial of degree $\ell = 3$ with Smith form $D(\lambda)$. Notice that $\ell_1 < \ell < \ell_2 = \ell_3 = \ell_4$, where $\ell_i = \deg d_i$ and that there is no real polynomial s of degree $\ell_2 - \ell = 1$ such that $1|s|d_2$. This corresponds to *Case 2*. Following the instructions yields $s_1 = 1$, so the first part of *Case 2* is simply a permutation of d_2 to the lower right corner. Because $d_2 = d_3 = d_4$, this does not modify $D(\lambda)$ but to follow *Case 2* in detail, we now consider the matrix $\mathrm{diag}(d_1, d_3, d_4, d_2)$. Next, we look for a degree $\ell_3 - \ell + 1 = 2$ real polynomial s_2 such that $1|s_2|d_3$. We have to take $s_2 = \lambda^2 + 1$. Then by Lemma 3.1 $D(\lambda)$ is equivalent to $T_1(\lambda) = \mathrm{diag}(\lambda^2 + 1, (\lambda^2 + 1)^2, -(\lambda^2 + 1), (\lambda^2 + 1)^2) + e_1 e_3^T$. It remains to apply the last step of the proof of Theorem 4.1 to block triangularize the polynomial. This leads to

$$T(\lambda) = \begin{bmatrix} \lambda^2+1 & -\lambda(\lambda^2+1) & 1 & -\lambda \\ \lambda(\lambda^2+1) & \lambda^2+1 & \lambda & -\lambda^2 \\ & & \lambda^2+1 & -\lambda(\lambda^2+1) \\ & & \lambda(\lambda^2+1) & \lambda^2+1 \end{bmatrix}. \qquad \square$$

We can now state the analog of Theorem 3.4 for real polynomials.

Theorem 4.3 *Any $P(\lambda) \in \mathbb{R}[\lambda]^{n \times m}$ with $n \leq m$ is quasi-triangularizable.*

Proof. The proof is along the same line as that presented for Theorem 3.4. We only sketch it and point out the differences.

We apply a Möbius transform $\mathcal{M}_A$ to $P(\lambda)$ induced by a real 2×2 nonsingular matrix A such that $\mathcal{M}_A(P)$ has no elementary divisors at infinity. We compute the Smith form $D(\lambda)$ of $\mathcal{M}_A(P)$, and let $\mathrm{diag}(d_1, \ldots, d_r)$ denote the regular part of $D(\lambda)$, where $r = \mathrm{rank}(P)$. Starting with $\mathrm{diag}(d_1, \ldots, d_r)$, we follow the triangularization procedure in the proof of Theorem 4.1 with two small modification if $\sum_{j=1}^{r} \deg d_j < \ell r$:

(i) We stop the induction procedure when the remaining (non-deflated) diagonal elements are of degrees strictly less than ℓ.

(ii) If the induction procedure reach "case 3", item (i) assures that $\ell_p > \ell$. We might, however, have $\ell_p + \ell_{p-1} < 2\ell - 1$ ($\ell_p + \ell_{p-1}$ is even so $\ell_p + \ell_{p-1} = 2\ell - 1$ is not possible), in which case we deflate a polynomial of degree $\ell - 1$ to position (p, p). The remaining diagonal elements are of degrees strictly less than ℓ so we stop the induction.

Now, all diagonal elements of degree $\ell + 1$ are preceded by a diagonal element of degree $\ell - 1$. Hence, we can perform the block-triangularization as in Theorem 4.1. Finally, we remove unwanted elementary divisors at infinity using the procedure described in Theorem 3.4. $\square$

Remark 4.4 In the singular case $n \leq m$, $r < m$, the procedure for removing elementary divisors at infinity moves the nonzero quasi-triangular part of the matrix polynomial one column to the right. This means that the resulting matrix polynomial is in fact triangular! $\square$

4.2. Triangularizable real matrix polynomials

For $n \leq m$, we give now a characterization of all $P(\lambda) \in \mathbb{R}[\lambda]^{n\times m}$ that are triangularizable over the real numbers. By Remark 4.4 all singular real $n \times m$ matrix polynomials are triangularizable, so we consider only the regular $n \times n$ case. The conditions for triangularizability that will be derived depend only on the partial multiplicities of the real eigenvalues and the matrix size. These quantities, as well as the triangular structure, are preserved under real invertible Möbius transformations. We can always find a real Möbius transformation $\mathcal{M}_A$ such that $\mathcal{M}_A(P)$ has no elementary divisors at infinity, that is, the partial multiplicities at infinity of $P(\lambda)$ get represented as the partial multiplicities of some real eigenvalue of $\mathcal{M}_A(P)$. This allow us to focus on real matrix polynomials with nonsingular leading coefficients.

The main results in this subsection are based on the following theorem.

Theorem 4.5 *[8] Let $d_1|\cdots|d_n$ and $t_1, \dots, t_n$ be monic polynomials in $\mathbb{F}[\lambda]$. Then there exists a triangular matrix polynomial in $\mathbb{F}[\lambda]^{n\times n}$ with diagonal entries $t_1, \dots, t_n$ and $d_1, \dots, d_n$ as invariant factors if and only if*

$$\prod_{j=1}^{k} d_j \Big| \gcd\Big\{\prod_{j=1}^{k} t_{i_j} : i_1 < \cdots < i_k\Big\}, \quad k = 1{:}\,n, \quad \textit{and} \quad \prod_{j=1}^{n} d_j = \prod_{j=1}^{n} t_j. \tag{9}$$

We remark that the proof of Theorem 4.5 is constructive. The next result will be used in the proof of Theorem 4.8. Its proof is inspired by that of [12, Cor. 4.3]. Recall the notion of majorization (Section 2.3).

Lemma 4.6 *Let $p_1, \dots, p_s \in \mathbb{F}[\lambda]$ be irreducible and of degree g. For $i = 1{:}\,n$, let $d_i = p_1^{a_{i1}} p_2^{a_{i2}} \dots p_s^{a_{is}}$, $a_i = a_{i1} + \cdots + a_{is}$, and b_i be a nonnegative integer. If*

$$(b_1, \dots, b_n) \prec (a_1, \dots, a_n), \tag{10}$$

then there exist n polynomials $t_i \in \mathbb{F}[\lambda]$ with $\deg t_i = gb_i$ such that

$$\gcd\Big\{\prod_{j=1}^{k} d_{i_j} : i_1 < \cdots < i_k\Big\} \Big| \gcd\Big\{\prod_{j=1}^{k} t_{i_j} : i_1 < \cdots < i_k\Big\}, \quad k = 1{:}\,n, \tag{11}$$

and

$$\prod_{j=1}^{n} d_j = \prod_{j=1}^{n} t_j. \tag{12}$$

Proof. By Lemma 2.6 condition (10) implies that if $(a_{[1]}, \dots, a_{[n]})$ is an arrangement of $(a_1, \dots, a_n)$ in nonincreasing order $(a_{[1]} \geq \cdots \geq a_{[n]})$ and the same for $(b_{[1]}, \dots, b_{[n]})$ then $(b_{[1]}, \dots, b_{[n]})$ can be obtained from $(a_{[1]}, \dots, a_{[n]})$ by a finite number of elementary transformations. That is to say, there is a nonnegative integer r and sequences $(c_1^{(i)}, \dots, c_n^{(i)})$, $i = 1{:}\,r$, such that

- $(c_1^{(1)}, \dots, c_n^{(1)}) = (a_{[1]}, \dots, a_{[n]})$ and $(c_1^{(r)}, \dots, c_n^{(r)}) = (b_{[1]}, \dots, b_{[n]})$,
- $c_1^{(i)} \geq \cdots \geq c_n^{(i)}$, $i = 1{:}\,r$, and

- for $i = 2{:}\,r$, $(c_1^{(i)}, \ldots, c_n^{(i)})$ is obtained from $(c_1^{(i-1)}, \ldots, c_n^{(i-1)})$ by an elementary transformation.

It is enough to prove that if $(b_{[1]}, \ldots, b_{[n]})$ is obtained from $(a_{[1]}, \ldots, a_{[n]})$ by an elementary transformation then there exist polynomials $t_1, \ldots, t_n \in \mathbb{F}[\lambda]$, with $\deg t_i = gb_{[i]}$, satisfying conditions (11)–(12). For in that case, bearing in mind that $(c_1^{(2)}, \ldots, c_n^{(2)})$ can be obtained from $(a_{[1]}, \ldots, a_{[n]})$ by an elementary transformation, there are polynomials $t_1^{(2)}, \ldots, t_n^{(2)}$ with $\deg t_j^{(2)} = gc_j^{(2)}$ such that (11)–(12) hold true with t_{i_j} replaced by $t_{i_j}^{(2)}$. Similarly, $(c_1^{(3)}, \ldots, c_n^{(3)})$ is obtained from $(c_1^{(2)}, \ldots, c_n^{(2)})$ by an elementary transformation, so there are polynomials $t_1^{(3)}, \ldots, t_n^{(3)}$ with $\deg t_j^{(3)} = gc_j^{(3)}$ such that (11)–(12) hold true with d_{i_j} and t_{i_j} replaced by $t_{i_j}^{(2)}$ and $t_{i_j}^{(3)}$, respectively. By taking $r-1$ steps in this way we obtain polynomials $t_1 = t_1^{(r)}, \ldots, t_n = t_n^{(r)}$, of desired degrees, $\deg t_i = gb_{[i]}$, and conditions (11) and (12) are satisfied.

Assume that $(b_{[1]}, \ldots, b_{[n]})$ is obtained from $(a_{[1]}, \ldots, a_{[n]})$ by an elementary transformation. This means that there are indices $[j] < [m]$ such that

$$b_{[j]} = a_{[j]} - 1, \quad b_{[m]} = a_{[m]} + 1, \quad b_{[i]} = a_{[i]} + 1, \quad i \neq j, m. \tag{13}$$

As $b_{[j]} \geq b_{[m]}$, we have $a_{[j]} \geq a_{[m]} + 2$ and so there is an index u, $1 \leq u \leq s$ such that $a_{[j]u} \geq a_{[m]u} + 1$ (otherwise $a_{[m]u} \geq a_{[j]u}$ for all u would imply $a_{[m]} \geq a_{[j]}$). Define

$$t_{[m]} := d_{[m]} p_u, \quad t_{[j]} := \frac{d_{[j]}}{p_u}, \quad t_i := d_i, \quad i \neq [j], [m].$$

We claim that with the polynomials t_i defined in this way, conditions (11) and (12) are fulfilled. It is plain that the latter holds true. On the other hand,

$$\gcd\Big\{ \prod_{j=1}^{k} d_{i_j} : i_1 < \cdots < i_k \Big\} = \prod_{v=1}^{s} p_v^{\min\{a_{i_1 v} + \cdots + a_{i_k v} : i_1 < \cdots < i_k\}}$$

and

$$\gcd\Big\{ \prod_{j=1}^{k} t_{i_j} : i_1 < \cdots < i_k \Big\} = \prod_{v=1}^{s} p_v^{\min\{b_{i_1 v} + \cdots + b_{i_k v} : i_1 < \cdots < i_k\}},$$

where $b_{iv} = a_{iv}$, except for $v = u$ and $i = [j], [m]$ for which $b_{[j]u} = a_{[j]u]} - 1$ and $b_{[m]u} = a_{[m]u} + 1$. Thus, in proving condition (11) we can assume without lost of generality that $d_i = p_u^{a_i}$ and $t_i = p_u^{b_i}$ for $i = 1{:}\,n$. We have

$$\gcd\Big\{ \prod_{j=1}^{k} d_{i_j} : i_1 < \cdots < i_k \Big\} = p_u^{a_{[n]} + \cdots + a_{[n-k+1]}},$$

$$\gcd\Big\{ \prod_{j=1}^{k} t_{i_j} : i_1 < \cdots < i_k \Big\} = p_u^{b_{[n]} + \cdots + b_{[n-k+1]}},$$

and it is an immediate consequence of (13) that

$$p_u^{a_{[n]} + \cdots + a_{[n-k+1]}} \Big| p_u^{b_{[n]} + \cdots + b_{[n-k+1]}}, \quad k = 1{:}\,n. \qquad \square$$

If $d_1, \ldots, d_n$ in Lemma 4.6 are ordered by divisibility as in Theorem 4.5 and

$$t_i = p_1^{b_{i1}} p_2^{b_{i2}} \cdots p_s^{b_{is}}, \quad i = 1:n,$$

then condition (9) is equivalent to

$$\sum_{i=1}^{k} a_{ij} \leq \min\Big\{\sum_{r=1}^{k} b_{i_r j} : i_1 < \cdots < i_k\Big\}, \quad k = 1\colon n \quad \text{and} \quad j = 1\colon s$$

with equality for $k = n$. Hence, using the notion of majorization condition (9) is equivalent to $(b_{1j}, \ldots, b_{nj}) \prec (a_{1j}, \ldots, a_{nj})$ for $j = 1\colon s$. Now, condition (10) follows at once from Lemma 2.4 (i). We have shown the following corollary.

Corollary 4.7 *Under the same conditions as in Lemma* 4.6 *assume that* $d_1|\cdots|d_n$*. Then condition* (9) *is equivalent to the majorization condition* (10)*.*

Now suppose that $P(\lambda) \in \mathbb{R}[\lambda]^{n\times n}$ of degree ℓ and with nonsingular leading coefficient has for invariant factors $d_1|\cdots|d_n$, each with a factorization

$$d_i = \prod_{j=1}^{s} (\lambda^2 + a_j\lambda + b_j)^{q_{ij}} \prod_{j=1}^{t} (\lambda - \lambda_j)^{m_{ij}}, \tag{14}$$

with

$$0 \leq q_{1j} \leq q_{2j} \leq \cdots \leq q_{nj}, \quad j = 1\colon s,$$

and

$$0 \leq m_{1j} \leq m_{2j} \leq \cdots \leq m_{nj}, \quad j = 1\colon t,$$

and $\lambda^2 + a_j\lambda + b_j$ are irreducible polynomials in $\mathbb{R}[\lambda]$. Exponents equal to zero may have been included for notational convenience.

Suppose $P(\lambda)$ is equivalent to a real $n \times n$ triangular matrix polynomial $T(\lambda)$ of degree ℓ. We have that $\prod_{i=1}^{n} d_i = \prod_{i=1}^{n} t_i$, where t_i is the ith diagonal entry of $T(\lambda)$ and has the form

$$t_i = \prod_{j=1}^{s} (\lambda^2 + a_j\lambda + b_j)^{g_{ij}} \prod_{j=1}^{t} (\lambda - \lambda_j)^{h_{ij}}. \tag{15}$$

We cannot assume, however, any specific order for the exponents g_{ij} and h_{ij}.

Let us define

$$q_i := \sum_{j=1}^{s} q_{ij}, \quad m_i := \sum_{j=1}^{t} m_{ij},$$

and notice that $q_n \geq \cdots \geq q_1$ and $m_n \geq \cdots \geq m_1$.

In what follows we will adopt the following notation: if

$$p = \prod_{j=1}^{u} (\lambda^2 + a_j\lambda + b_j)^{c_j} \prod_{j=1}^{v} (\lambda - \lambda_j)^{d_j}$$

is the prime factorization of $p \in \mathbb{R}[\lambda]$ then

$$p_c = \prod_{j=1}^{u} (\lambda^2 + a_j\lambda + b_j)^{c_j}, \quad p_r = \prod_{j=1}^{v} (\lambda - \lambda_j)^{d_j}.$$

We are now ready to state and prove the main results of this section.

Theorem 4.8 *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n\times n}$ with nonsingular leading coefficient have invariant factors $d_1|\cdots|d_n$ with prime factorizations* (14). *Let $g_1,\dots,g_n$ and $h_1,\dots,h_n$ be nonnegative integers. Then $P(\lambda)$ is equivalent to a real triangular matrix polynomial $T(\lambda)$ with diagonal elements $t_i = t_{ic}t_{ir}$ such that* $\deg t_{ic} = 2g_i$ *and* $\deg t_{ir} = h_i$*, if and only if*

$$(g_1,\dots,g_n) \prec (q_n,\dots,q_1), \qquad (h_1,\dots,h_n) \prec (m_n,\dots,m_1). \tag{16}$$

Proof. ($\Rightarrow$) Assume that $T(\lambda)$ is a triangular matrix equivalent to $P(\lambda)$ and with diagonal elements $t_i = t_{ic}t_{ir}$ such that $\deg t_{ic} = 2g_i$ and $\deg t_{ir} = h_i$. Assume also that the prime factorization of t_i is given by (15). Then $g_i = \sum_{j=1}^{s} g_{ij}$, $h_i = \sum_{j=1}^{t} h_{ij}$ and (16) follow from Theorem 4.5 and Corollary 4.7.

($\Leftarrow$) Write $d_i = d_{ic}d_{ir}$, and notice that $\deg d_{ic} = q_i$, $\deg d_{ir} = m_i$. By Lemma 4.6 and $(h_1,\dots,h_n) \prec (m_n,\dots,m_1)$, there exist polynomials $t_{1r},\dots, t_{nr}$ such that $\deg t_{ir} = h_i$,

$$\prod_{j=1}^{k} d_{jr} \Big| \gcd\Big\{\prod_{j=1}^{k} t_{i_j r} : i_1 < \cdots < i_k\Big\}, \quad k = 1\colon n, \quad \text{and} \quad \prod_{j=1}^{n} d_{jr} = \prod_{j=1}^{n} t_{jr}.$$

Similarly, by $(g_1,\dots,g_n) \prec (q_n,\dots,q_1)$, there exist polynomials $t_{1c},\dots, t_{nc}$ such that $\deg t_{ci} = 2g_i$,

$$\prod_{j=1}^{k} d_{jc} \Big| \gcd\Big\{\prod_{j=1}^{k} t_{i_j c} : i_1 < \cdots < i_k\Big\}, \quad k = 1\colon n, \quad \text{and} \quad \prod_{j=1}^{n} d_{jc} = \prod_{j=1}^{n} t_{jc}.$$

Defining $t_i := t_{ic}t_{ir}$, we have that $\deg t_i = 2g_i + h_i$ and condition (9) is satisfied. The result now follows from Theorem 4.5. □

The next theorem is a consequence of Theorem 4.8.

Theorem 4.9 *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n\times n}$ of degree ℓ and with nonsingular leading coefficient have invariant factors $d_1|\cdots|d_n$ with prime factorizations* (14). *Let $n_c = qn + s$ be the Euclidean division of $n_c = q_1+\cdots+q_n$ by n. Then $P(\lambda)$ is equivalent to a real triangular matrix with diagonal elements of degree ℓ if and only if*

$$(\underbrace{\ell - 2q,\dots,\ell - 2q}_{n-s \text{ times}},\underbrace{\ell - 2q - 2,\dots,\ell - 2q - 2}_{s \text{ times}}) \prec (m_1,\dots,m_n) \tag{17}$$

and all integers in the left-hand side of the majorization inequality (17) *are nonnegative.*

Proof. Notice that since $\deg\big(\det P(\lambda)\big) = n\ell$ we have

$$0 \le n\ell - 2n_c = n\ell - 2qn - 2s = n(\ell - 2q) - 2s \quad \Rightarrow \quad \ell - 2q \ge 0. \tag{18}$$

Hence the sum of the elements in the left hand side of (17) is always nonnegative. It may happen however that $\ell - 2q - 2 < 0$. This is the case, for example, if $\ell = n = 3$ and the elementary divisors of $P(\lambda)$ are $(\lambda - 1)$ and $(\lambda^2 + 1)^4$.

($\Rightarrow$) Assume now that $P(\lambda)$ is equivalent to a real triangular matrix $T(\lambda)$ with diagonal elements t_j of degree ℓ and prime factorization (15). Bearing in mind that $n_c = nq+s$, if $s > 0$ then at least one element in the diagonal of $T(\lambda)$ has to be of degree $2q+2$ or bigger. Hence $\ell \geq 2q+2$ when $s > 0$ and if $P(\lambda)$ is equivalent to a real triangular matrix all integers in left hand side of (17) are nonnegative. Also, $2g_i + h_i = \ell$, where $g_i = \sum_{j=1}^{s} g_{ij}$ and $h_i = \sum_{j=1}^{t} h_{ij}$. By Lemma 2.4 (ii) it follows from $n_c = qn + s$, $0 \leq s < n$, that

$$(\underbrace{q+1,\ldots,q+1}_{s \text{ times}},\underbrace{q,\ldots,q}_{n-s \text{ times}}) \prec (g_1,\ldots,g_n).$$

Arrange $g_1,\ldots,g_n$ in nonincreasing order $g_{i_1} \geq \cdots \geq g_{i_n}$. Then

$$\begin{aligned} kq &\geq g_{i_n} + \cdots + g_{i_{n-k+1}}, & k &= 1\colon n-s, \\ (n-s)q + k(q+1) &\geq g_{i_n} + \cdots + g_{i_{s-k+1}}, & k &= 1\colon s, \end{aligned}$$

and $g_1 + \cdots + g_n = nq + s$. Since $h_{i_j} = \ell - 2g_{i_j}$ we have

$$\begin{aligned} h_{i_n} + \cdots + h_{i_{n-k+1}} &\geq k(\ell - 2q), & k &= 1\colon n-s, \\ h_{i_n} + \cdots + h_{i_{s-k+1}} &\geq (n-s)(\ell - 2q) + k(\ell - 2q - 2), & k &= 1\colon s, \end{aligned}$$

and $h_1 + \cdots + h_n = n\ell - 2nq - 2s$. This implies

$$(\underbrace{\ell-2q,\ldots,\ell-2q}_{n-s \text{ times}},\underbrace{\ell-2q-2,\ldots,\ell-2q-2}_{s \text{ times}}) \prec (h_1,\ldots,h_n)$$

and (17) follows from the second majorization in (16).

Conversely, assume that (17) holds with all integers in left-hand side of (17) nonnegative. Define

$$\begin{aligned} h_k &= \ell - 2q, & g_k &= q, & k &= 1\colon n-s, \\ h_k &= \ell - 2q - 2, & g_k &= q+1, & k &= n-s+1\colon n. \end{aligned}$$

Then $h_i \geq 0$ and $h_i + 2gi = \ell$ for $i = 1\colon n$. By (17) $(h_1,\ldots,h_n) \prec (m_n,\ldots,m_1)$. In addition, as $q_1 + \cdots + q_n = n_c = qn + s$, $0 \leq s < n$, by Lemma 2.4 (ii) we also have $(g_1,\ldots,g_n) \prec (q_n,\ldots,q_1)$. Hence, by Theorem 4.8 there is a real triangular matrix $T(\lambda)$ equivalent to $P(\lambda)$ with polynomials of degree ℓ on the diagonal. $\square$

It follows from Theorem 4.9 that when $\ell = 1$, q and s must be zero for $P(\lambda)$ to be triangularizable because $\ell \geq 2q$ and $\ell \geq 2q+2$ if $s > 0$. Thus, we get the well-known result that a real pencil is triangularizable over $\mathbb{R}[\lambda]$ if and only if it has only real eigenvalues.

In Example 4.2, $\ell = 3$, $n = 4$, $n_c = 6$, $q = 1$, $s = 2$ and $\ell < 2q+2$. Hence, there is no real triangular cubic matrix polynomial with Smith form $D(\lambda) = \operatorname{diag}(1, (\lambda^2+1)^2, (\lambda^2+1)^2, (\lambda^2+1)^2)$.

It has been seen in [10, Th. 3.6] that a necessary and sufficient condition for a real regular quadratic matrix polynomial to be triangularizable over $\mathbb{R}[\lambda]$ is $n \geq n_c + p$ where

p is the geometric multiplicity of the real eigenvalue with largest geometric multiplicity. That this condition is equivalent to (17) for quadratic matrix polynomials is better visualized if the Weyr characteristic rather than the Segre characteristic is used (see Section 2.3 for the definition of Weyr characteristic). The following result is a rephrasing of Theorem 4.9 in terms of the union of the elements in the Weyr characteristic of $P(\lambda)$ associated with the real eigenvalues.

Theorem 4.10 *Let $P(\lambda) \in \mathbb{R}[\lambda]^{n\times n}$ be of degree ℓ and with nonsingular leading coefficient. Let n_c be the sum of the exponents of the elementary divisors of $P(\lambda)$ associated to irreducible polynomials of degree 2. Let $n_c = qn + s$ be the Euclidean division of n_c by n and let $(p_1, p_2, \ldots)$, with $p_1 \geq p_2 \geq \ldots$, be the union of the elements in the Weyr characteristic of $P(\lambda)$ corresponding to the real eigenvalues, where $p_1 = 0$ if $P(\lambda)$ has no real eigenvalue and $p_k = 0$ for $k \leq 0$. Then $P(\lambda)$ is triangularizable if and only if either $s = 0$, or $s > 0$ and*

$$n(\ell - q - 1) - n_c \geq p_1 + \cdots + p_{\ell-2q-1}. \tag{19}$$

Proof. By Theorem 4.9, $P(\lambda)$ is triangularizable over $\mathbb{R}[\lambda]$ if and only if (17) holds with all integers in its left hand side nonnegative. Now, by Lemma 2.5, (17) is equivalent to

$$(m_n, \ldots, m_1)^* \prec (\underbrace{\ell - 2q, \ldots, \ell - 2q}_{n-s \text{ times}}, \underbrace{\ell - 2q - 2, \ldots, \ell - 2q - 2}_{s \text{ times}})^*.$$

Let us compute these conjugate sequences:

$$(\underbrace{\ell - 2q, \ldots, \ell - 2q}_{n-s \text{ times}}, \underbrace{\ell - 2q - 2, \ldots, \ell - 2q - 2}_{s \text{ times}})^* = \begin{cases} (\underbrace{n, \ldots, n}_{\ell-2q-2 \text{ times}}, n-s, n-s) & \text{if } s > 0, \\ (\underbrace{n, \ldots, n}_{\ell-2q \text{ times}}) & \text{if } s = 0. \end{cases}$$

Recall that $m_i = \sum_{j=1}^{t} m_{ij}$ and $m_{nj} \geq \ldots \geq m_{1j} \geq 0$ for $j = 1{:}t$. Hence by (6),

$$(m_n, \ldots, m_1)^* = (m_{n1}, \ldots, m_{11})^* \cup \cdots \cup (m_{nt}, \ldots, m_{1t})^*.$$

Now, $(m_{nj}, \ldots, m_{1j})^*$ is the Weyr characteristic corresponding to the real eigenvalue λ_j, so $(m_n, \ldots, m_1)^* = (p_1, p_2, \ldots)$. Notice that p_1 is the geometric multiplicity of the real eigenvalue of $P(\lambda)$ with largest geometric multiplicity. Hence (17) is equivalent to

$$(p_1, p_2, \ldots) \prec (\underbrace{n, \ldots, n}_{\ell-2q-2 \text{ times}}, n-s, n-s) \tag{20}$$

if $s > 0$ and to

$$(p_1, p_2, \ldots) \prec (\underbrace{n, \ldots, n}_{\ell-2q \text{ times}}) \tag{21}$$

if $s = 0$. Since the geometric multiplicity of any real eigenvalue of $P(\lambda)$ is at most n, it follows that $p_i \leq n$ and condition (21) is always satisfied. In addition,

$$\sum_{i\geq 1} p_i = n\ell - 2n_c = n\ell - 2(qn + s) = n(\ell - 2q) - 2s = n(\ell - 2q - 2) + 2(n - s).$$

Since $p_i \leq n$, the only nontrivial inequality that must be satisfied for condition (20) to be fulfilled is

$$n(\ell - 2q - 2) + n - s \geq p_1 + \cdots + p_{\ell-2q-1},$$

which is equivalent to (19). Agreeing that $p_k = 0$ for $k \leq 0$, (19) implies that $\ell \geq 2q + 2$. For if $\ell - 2q - 1 \leq 0$ then we get from (19) that $n(\ell - 2q - 2) + n - s \geq 0$ and so $\ell - 2q - 1 \geq (s/n) > 0$, a contradiction. □

To conclude, if n_c is a multiple of n then $P(\lambda)$ is always triangularizable. The reverse is not always true, as the case of a pencil with nonreal complex conjugate eigenvalues shows. In this case, $\ell = 1$ and $q = 0$. If $n_c > 0$ then $n(\ell - q - 1) - n_c < 0 = p_0$.

For $\ell = 2$ we have $2n \geq 2n_c$ and so $q \leq 1$ with $q = 1$ if and only if $n_c = n$. Thus $P(\lambda)$ is triangularizable if and only if $n_c = 0$, $n_c = n$ (i.e., $q = 1$, $s = 0$ and $\ell = 2q$) or $q = 0$, $s > 0$ and $n - n_c = (n(\ell - q - 1) - n_c \geq p_1$. Taking into account that $p_1 = 0$ means that $P(\lambda)$ has no real eigenvalues (i.e., $n_c = n$), we conclude that $P(\lambda)$ is triangularizable if and only if $n - n_c \geq p_1$. This is Theorem 3.6 in [10].

For $\ell = 3$, we must have $q \leq 1$ since $\ell \geq 2q$ (see (18)). Now $P(\lambda)$ is not triangularizable over $\mathbb{R}[\lambda]$ if and only if

$$n(2 - q) - n_c < p_1 + \cdots + p_{2-2q}, \quad q = 0, 1,$$

or equivalently, if and only if $q = 1$ and $n < n_c$ or $q = 0$ and $2n - n_c < p_1 + p_2$. An example where $q = 1$ and $n < n_c$ is $n = \ell = 3$ and the elementary divisors are $(\lambda - 1)$, $(\lambda^2 + 1)$ and $(\lambda^2 + 1)^3$. Example 4.2 also corresponds to that case. An example where $q = 0$ and $2n - n_c < p_1 + p_2$ is $n = \ell = 3$ and the elementary divisors are λ, λ^2, λ^2, $(\lambda^2 + 1)$ and $(\lambda^2 + 1)$. Finally, in Example 3.3, $n_c = 2 < n = 3$, $q = 0$ and $p_1 = 2$, $p_2 = 2$. Hence, $2n - n_c = p_1 + p_2$, which confirms that the matrix polynomial in Example 3.3 is triangularizable.

5. Inverse problems

The main motivation of this paper is the characterization of the real and complex matrices that can be reduced to triangular or trapezoidal form preserving the degree and the finite and infinite elementary divisors. However, as a by-product, we are solving a structured inverse polynomial eigenvalue problem. Recall that problems concerning the construction of matrix polynomials having certain eigenvalues or elementary divisors, are called inverse polynomial eigenvalue problems. In [3, Thm. 1.7] a monic inverse polynomial eigenvalue problem is solved over $\mathbb{C}$ (in fact over any algebraically closed field). Since monic matrix polynomials have no elementary divisors at infinity, the Smith form contains all the information about elementary divisors. It is shown in the above reference that in order to build such an $n \times n$ matrix polynomial of degree ℓ, the only constraints on the list of its elementary divisors are

(i) the geometric multiplicities are bounded by n (because any regular $n \times n$ matrix polynomial has n invariant factors), and

(ii) the sum of the partial multiplicities of all the eigenvalues is $n\ell$.

This is generalized to matrices with nonsingular leading coefficients over arbitrary fields in [7, Thm. 5.2]. From Theorem 3.4 and Remark 3.5 it follows that we can realize a list of finite and infinite elementary divisors by an $n \times m$ matrix polynomial of degree ℓ over an algebraically closed field if and only if condition (i) above and

(iii) the sum of all partial multiplicities, including those at infinity, is at most $\ell \min(m, n)$,

are satisfied, thereby extending the result in [3, Thm. 1.7] and [7, Thm. 5.2]. Furthermore, from Theorem 4.3, we get the solution to the corresponding inverse problem over $\mathbb{R}[\lambda]$. As one could expect, the only additional constraint is that nonreal elementary divisors must come in complex conjugate pairs.

Constraints on the structure of matrix polynomials often impose constraints on the elementary divisors. We have described these constraints in the case of real triangular matrix polynomials and have shown that there are no constraints for complex ones other that (i) and (iii).

Recall that a matrix polynomial is called *self-adjoint* or *Hermitian* if all the coefficient matrices are Hermitian. If the leading coefficient is nonsingular it is well-known that all nonreal elementary divisors comes in complex conjugate pairs [2, Lem. 1.2]. Given a regular self-adjoint matrix polynomial $P(\lambda)$, we can always find a real Möbius transformation m_A such that $\mathcal{M}_A(P)$ is self-adjoint and has nonsingular leading coefficient. Hence, it follows from Theorem 2.3 that also the nonreal elementary divisors of $P(\lambda)$ must come in complex conjugate pairs. This constraint on the list of elementary divisors is exactly the same constraint as in the inverse polynomial eigenvalue problem over $\mathbb{R}[\lambda]$. We have proved the following result.

Theorem 5.1 *Any regular self-adjoint matrix polynomial is strongly equivalent to a real matrix polynomial.*

We conjecture that the theorem is true in the other direction too.

Conjecture 5.2 *Any regular real matrix polynomial is strongly equivalent to a regular self-adjoint matrix polynomial, and vice versa.*

6. Concluding remarks

In the quasi-triangularization process described in Theorem 4.1 we are faced with some choices. If a real matrix polynomial is triangularizable, the construction in Theorem 4.1 does not necessarily produce a triangular matrix polynomial. It is not known if the freedom can be exploited to obtain a triangular form. Tisseur and Zaballa [10, Thm. 4.3] have shown that for regular real quadratic matrix polynomials $Q(\lambda)$ that are not triangularizable, the minimum number of 2×2 diagonal blocks is $\max\{0, p_1 + n_c - n\}$, where n_c is the sum of the exponents of the elementary divisors of $P(\lambda)$ associated to irreducible polynomials of degree 2 and p_1 is the geometric multiplicity of the real eigenvalue of $P(\lambda)$ with largest geometric multiplicity. Extending their results to arbitrary degree matrix polynomials is left as an open problem.

References

[1] F. R. Gantmacher. *The Theory of Matrices*, volume one. Chelsea, New York, 1959. ISBN 0-8284-0131-4. x+374 pp.

[2] I. Gohberg, P. Lancaster, and L. Rodman. Spectral analysis of selfadjoint matrix polynomials. *Ann. of Math. (2)*, 112(1):33–71, 1980.

[3] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009. ISBN 0-898716-81-8. xxiv+409 pp. Unabridged republication of book first published by Academic Press in 1982.

[4] I. Macdonald. *Symmetric Functions and Hall Polynomials.* Oxford University Press, New York, second edition. 488 pp.

[5] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Möbius transformations of matrix polynomials. Technical report, 2012. In preparation.

[6] A. S. Markus and E. E. Parilis. The change of the Jordan structure of a matrix under small perturbations. *Linear Algebra Appl.*, 54:139–152, 1983.

[7] E. Marques de Sá. Imbedding conditions for λ-matrices. *Linear Algebra Appl.*, 24: 33–50, 1979.

[8] E. Marques de Sá. On the diagonal of integer matrices. *Czechoslovak Mathematics*, 5(30):207–212, 1980.

[9] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and Its Applications.* pub-Springer, New York, second edition, 2011. ISBN 978-0-387-40087-7.

[10] F. Tisseur and I. Zaballa. Triangularizing quadratic matrix polynomials. MIMS EPrint 2012.30, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2012.

[11] P. M. Van Dooren and P. Dewilde. The eigenstructure of an arbitrary polynomial matrix: Computational aspects. *Linear Algebra Appl.*, 50:545–579, 1983.

[12] I. Zaballa. Controllability and Hermite indices of matrix pairs. *Internat. J. Control*, 68(1):61–68, 1997.