Shifted Normal Forms of Polynomial Matrices

Bernhard Beckermann
Laboratoire d'Analyse Numérique et d'Optimisation
UFR IEEA M3, USTL Flandres Artois
F–59655 Villeneuve d'Ascq CEDEX, France
bbecker@ano.univ-lille1.fr

George Labahn
Department of Computer Science
University of Waterloo, Ontario, Canada
glabahn@daisy.uwaterloo.ca

Gilles Villard
LMC-IMAG, BP 53
F-38041 Grenoble cedex 9
Gilles.Villard@imag.fr

Abstract

In this paper we study the problem of transforming, via invertible column operations, a matrix polynomial into a variety of *shifted* forms. Examples of forms covered in our framework include a column reduced form, a triangular form, a Hermite normal form or a Popov normal form along with their shifted counterparts.

By obtaining degree bounds for unimodular multipliers of shifted Popov forms we are able to embed the problem of computing a normal form into one of determining a shifted form of a minimal polynomial basis for an associated matrix polynomial. Shifted minimal polynomial bases can be computed via sigma bases [2, 3] and in Popov form via Mahler systems [4]. The latter method gives a fraction-free algorithm for computing matrix normal forms. **Key words:** Popov Form, Hermite Normal Form

1 Introduction

Matrix polynomial arithmetic is fundamental to many applications in science and engineering. It is encountered in linear systems theory [12], determining minimal partial realizations of matrix sequences [20] and solving linear diophantine equations [14, 16]. Not surprisingly, the arithmetic of matrix polynomials has substantial differences to that found with scalar polynomials. Such fundamental operations as determining a degree, a leading coefficient, and normal and canonical forms [10] have numerous variations in the matrix case. For example, the degree of a scalar polynomial has equivalents such as row degree, column degree, degree, degree of the determinant, MacMillan Degree, and others, each with their own usefulness. Leading coefficients of a matrix polynomial can mean leading column or leading row matrices and are not necessarily invertible, a problem when attempting division-like computations.

In order to use many analogous concepts from scalar polynomial arithmetic one often transforms a given matrix polynomial into an equivalent matrix polynomial having better properties for fundamental operations. Formally, two matrix polynomials $\mathbf{A}(z)$ and $\mathbf{B}(z)$ are column equivalent if there exists a unimodular polynomial matrix $\mathbf{U}(z)$ such that $\mathbf{A}(z) \cdot \mathbf{U}(z) = \mathbf{B}(z)$. The matrix $\mathbf{U}(z)$ corresponds to a sequence of elementary column operations. For a given $\mathbf{A}(z)$ one has a number of equivalent forms that are useful for applications. These include column reduced forms, where the leading column coefficient matrix is nonsingular (a property useful for division algorithms), triangular forms (useful for solving systems of linear equations) and Hermite normal forms or Popov normal forms (useful for determining when two matrix polynomials are not equivalent).

In this paper we study the problem of computing normal forms for full column rank matrix polynomials. In particular we are interested in computing Popov and *shifted* Popov normal forms for such matrices. Roughly speaking, a Popov normal form [17, 22] is a form having a "good" structure for leading coefficient matrices on both row and column sides. It has the important property that it always reduces column degrees of the input matrix. This is different from other forms such as the classical Hermite normal form which is an upper triangular matrix with additional degree constraints with respect to diagonals but which typically has degrees increased during the reduction to normal form.

The notion of a shifted form is basically one of altering the degree structures of the rows of a matrix and then computing forms of the resulting matrix. It is a simple process but very powerful. For example, the classical Hermite normal form can be obtained by determining a shifted Popov normal form for a shift determined from the degree structure of the input matrix polynomial (cf. Example 2.5). Shifted Popov Forms for square nonsingular matrix polynomials were first introduced in [4] as a convenient normal form for describing the properties of Mahler systems. Mahler systems were used as a basic building block for recursively computing solutions to module bases for matrix rational approximation and matrix rational interpolation problems (see also [1] in the case of matrix Padé systems). The (vector) shift was useful in this application for keeping track of a path of computation that allowed one to avoid singular situations in order to recursively compute along a path of closest nor-

mal points to a singularity. A (scalar) shift has also been used by Beelen et al [7] in determining a reduced column form of a full column rank matrix polynomial by computing a minimal polynomial basis of a shifted stacked rectangular matrix. The use of a shift ensured that the nonsingular leading coefficient matrix be isolated in specific (in this case last) rows.

By obtaining degree bounds for unimodular multipliers of shifted Popov forms, we are able to embed the problem of computing a normal form into one of determining a shifted form of a minimal polynomial basis for an associated matrix polynomial kernel. Shifted minimal polynomial bases can be computed via sigma bases [2, 3] and in Popov form via Mahler systems [4]. The last named algorithm has an important property: if the entries in the original matrix are polynomials having coefficients from an integral domain (for example a matrix with entries from $\mathbf{Z}[z]$ or $\mathbb{Q}[a_1, \ldots, a_k][z]$) then it computes a minimal polynomial bases using only fraction-free arithmetic.

The results in this paper are part of a larger research program: the efficient computation of matrix normal forms for arbitrary matrix polynomials. In particular we are interested in efficient fraction-free computation of such normal forms for nonsingular, singular and rectangular matrices. Our results give a first step in this direction. Additional results are available in the manuscript [6].

The remainder of the paper is organized as follows. Section 2 gives the basic definitions of shifted reduced and normal forms in the case of nonsingular square matrices while the next section looks at the equivalent problem for rectangular matrix polynomials of full column rank. Section 4 gives degree bounds on the (unique) unimodular multiplier. Section 5 shows how to embed the problem of computing a shifted Popov Normal Form and the associated unimodular multiplier into one of computing a minimal polynomial basis in normal form. This allows shifted normal forms to be computed using the algorithm of [4]. The last section includes a conclusion along with a discussion of future research directions.

2 Shifted Popov Forms of Nonsingular Matrices

In this section we give some of the basic definitions and properties required for the remainder of the paper for the case where the input matrix polynomial is square and nonsingular. Various shifted forms are introduced along with the concept of shifted column reduction. We remark that for any multi-index $\vec{a}$ (i.e. a vector of integers) we denote by $|\vec{a}|$ the sum of its components, $\max \vec{a}$ its maximum component and $\mathrm{perm}(\vec{a})$ a permutation of the components. In addition the multi-index $\vec{e}$ denotes the vector $(1, \ldots, 1)$.

Definition 2.1 ($\vec{a}$-Shifted forms, square matrices) *An $m \times m$ matrix polynomial $\mathbf{T}(z) \in \mathbb{Q}^{m \times m}[z]$ is $\vec{a}$–column reduced with $\vec{a}$-column degree $\vec{\alpha}$ – if there exists a multi-index $\vec{\alpha}$ such that*

$$z^{-\vec{a}} \cdot \mathbf{T}(z) \cdot z^{\vec{a}-\vec{\alpha}} = \mathbf{T}' + \mathcal{O}(z^{-1})_{z \to \infty}, \quad \text{with } \mathbf{T}' \in \mathbb{Q}^{m \times m} \text{ nonsingular.} \tag{1}$$

If this condition holds with $\mathbf{T}'$ nonsingular and upper triangular then $\mathbf{T}(z)$ is said to be in $\vec{a}$ quasi Popov form. *When it satisfies the additional normalization degree and leading coefficient constraint*

$$z^{-\vec{\alpha}} \cdot \mathbf{T}(z) = \mathbf{I}_m + \mathcal{O}(z^{-1})_{z \to \infty}. \tag{2}$$

then $T(z)$ is said to be in $\vec{a}$ Popov normal form. □

Remark 2.2 *Properties (1) and (2) are invariant under adding a constant to all components of $\vec{a}$. This allows us, in particular, to extend our definitions to* Laurent matrix polynomials *where needed.*

Up to a (unique) permutation of columns, Definition 2.1 gives the classical Popov normal form [12, Subsection 6.7.2, p.481] in the case $\vec{a} = \vec{0}$. When the form is used in a matrix fraction description or as a minimal polynomial basis, the degree $\vec{\alpha}$ are referred to as the vector of controllability *or* Kronecker indices.

For any two column equivalent and $\vec{a}$-column reduced matrices, the corresponding vectors $\vec{\alpha} - \vec{a}$ as in (1) coincide up to permutation [12, Lemma 6.3-14, p.388]. Also, for any matrix polynomial $\mathbf{A}(z)$ column equivalent to $\mathbf{T}(z)$ as in (1),

$$|\vec{\alpha}| = \deg \det \mathbf{T}(z) = \deg \det \mathbf{A}(z). \tag{3}$$

is an invariant. □

It is known [12, § 6.7.2, p.484] that any square nonsingular matrix polynomial may be transformed to Popov normal form by multiplication on the right by a unimodular matrix polynomial, and that the form is unique. A similar statement is also true for an $\vec{a}$-Popov form.

Lemma 2.3 *For a nonsingular matrix polynomial $\mathbf{A}(z)$ and a multi-index $\vec{a}$, set $\mathbf{A}(z)^{\#} = z^{-\vec{a}} \cdot \mathbf{A}(z)$. Let $\mathbf{T}(z)^{\#} = \mathbf{A}(z)^{\#} \cdot \mathbf{U}(z)$ be the $\vec{0}$–Popov form with degree $\vec{\alpha}^{\#}$ of this resulting Laurent matrix polynomial $\mathbf{A}(z)^{\#}$. Then $\mathbf{T}(z) = z^{\vec{a}} \cdot \mathbf{T}(z)^{\#} = \mathbf{A}(z) \cdot \mathbf{U}(z)$ is an $\vec{a}$–Popov form of $\mathbf{A}(z)$ with degree $\vec{\alpha} = \vec{\alpha}^{\#} + \vec{a}$. Conversely, for any $\vec{a}$–Popov form $\mathbf{T}(z)$ of $\mathbf{A}(z)$ we have $\mathbf{T}(z)^{\#} = z^{-\vec{a}} \cdot \mathbf{T}(z)$.* □

Lemma 2.3 says that it is possible to consider only $\vec{0}$–Popov forms. However, the introduction of an additional parameter $\vec{a}$ is convenient for a number of reasons. It appears naturally in the context of the approximation problems studied in [4], and was the primary purpose for introducing this form. Our vector shift can also be used to simplify the Matrix Euclidean algorithm of [8]. Indeed their six reduction steps can be viewed as moving from one shifted Popov form to a new shifted Popov form with a reduced shift. In the case of column-reduction a shift (in this case a scalar shift of certain components) was used as a tool in the algorithm of [7] for constructing a column reduced polynomial matrix for a matrix polynomial of full column rank.

A vector shift is also very useful in that it allows one to describe a number of other important matrix normal forms. For example, triangular forms and the Hermite normal form [15, §22, p.32] are obtained with shifts as follows.

Lemma 2.4 *Let $\mathbf{A}(z)$ be a matrix polynomial and $\mathbf{T}(z)$ be one of its $\vec{a}$-quasi Popov forms with $\vec{a}$-degree $\vec{\alpha}$. If $\vec{a}$ and $\vec{\alpha}$ satisfy*

$$\vec{a}_j - \vec{a}_i \geq \vec{\alpha}_j, \text{ for } i > j, \tag{4}$$

then $\mathbf{T}(z)$ is upper triangular. Furthermore the $\vec{a}$-Popov form of $\mathbf{A}(z)$ coincides with its Hermite normal form.

Proof. Let us show that $\mathbf{T}(z)_{i,j} = 0$ for $i > j$. By Definition 2.1, $\mathbf{T}'$ is upper triangular thus for $i > j$, $-\vec{a}_i + \deg \mathbf{T}(z)_{i,j} + \vec{a}_j - \vec{\alpha}_j < 0$ and if (4) is true then

$\deg \mathbf{T}(z)_{i,j} < 0$ and $\mathbf{T}(z)$ is upper triangular. Also condition (2) implies that the degree of a diagonal entry is strictly larger than the degree of the other entries of this row and the $\vec{\alpha}$-Popov form is the Hermite normal form. □

The condition $\mathbf{T}'$ upper triangular in Definition 2.1 plays a minor role in the entire triangularization of $\mathbf{A}(z)$. Indeed the upper triangularization may be ensured using only a column reduced form with a slightly different shift. If $\mathbf{T}(z)$ is $\vec{\alpha}$-column reduced with $\vec{\alpha}_j - \vec{\alpha}_i > \vec{\alpha}_j$, for $i > j$, then $\mathbf{T}(z)$ is upper triangular.

Even if for any given matrix $\mathbf{A}(z)$, the degrees $\vec{\alpha}_j$ are not known in advance, one may always use Lemma 2.4 and shifted forms to compute triangular forms. Indeed, the degrees are bounded by (3) thus choosing $w \geq \deg \det \mathbf{A}(z)$ and as shift vector $\vec{\alpha} = [(m-1)w, \ldots, w, 0]$, condition (4) is satisfied.

Example 2.5 *Let $\mathbf{A}(z)$ be the following $\vec{0}$-column reduced matrix:*

$$\mathbf{A}(z) = \begin{bmatrix} z^2-1 & z+1 & 2 \\ 0 & -z & z-2 \\ z+2 & 0 & z+1 \end{bmatrix}$$

From the sum of the column degrees one may take $w = 4$ as an upper bound on the degree of the determinant. If $\vec{\alpha} = [8, 4, 0]$ then an $\vec{\alpha}$-quasi Popov form of $\mathbf{A}(z)$ is

$$\begin{bmatrix} z^4+z^2-z+4 & -z^3+4z+5 & z^4-4z^3+4z^2+15z \\ 0 & -4 & -8z \\ 0 & 0 & 8 \end{bmatrix},$$

which in fact is $[4, 1, 0]$-reduced with degree $\vec{\alpha} = [4, 0, 0]$. By Lemma 2.4, with the same shifts, the $\vec{\alpha}$-Popov form of $\mathbf{A}(z)$:

$$\begin{bmatrix} z^4-4z^2-z+4 & \frac{1}{4}z^3-z-\frac{5}{4} & \frac{1}{2}z^3+z^2+z-\frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is also its Hermite normal form. □

This application of shifted forms of matrix polynomial, is very similar to a well known application of integer lattice basis reduction. Given an integer matrix $\mathbf{A}$, it is shown in [18, p 74] that the reduction of a well chosen lattice deduced from $\mathbf{A}$, gives the integer Hermite form.

3 Shifted Forms of Full Column Rank Matrices

In order to include in our framework such applications as the extended gcd problem (cf. Example 3.4), determination of matrix structures [21] (cf. Example 3.5) and normalization of module bases, we extend the results of the previous section to full column rank polynomial matrices. We introduce shifts in the classical treatment for column reduced forms [17, 12, p481] and, as done in [13] in the Hermite case, we ensure the normalization of leading matrices by considering column echelon forms.

Given a matrix polynomial $\mathbf{A}(z)$, we denote its elements by $\mathbf{A}(z)_{i,j}$. Furthermore, given lists I, J of increasing row / column indices of $\mathbf{A}(z)$, we denote by $\mathbf{A}(z)_{I,J}$ the corresponding submatrix of $\mathbf{A}(z)$. Also, for $\mathbf{A}(z)_{I,*}$ (and $\mathbf{A}(z)_{*,J}$) we just extract rows with index in I (and columns with index in J, respectively).

Definition 3.1 (Column echelon matrices) *A full column rank scalar matrix $\mathbf{T}' \in \mathbb{Q}^{m \times n}$ is in upper echelon form with pivot set $I = (i_1, \ldots, i_n)$ if $1 \leq i_1 < i_2 < \ldots < i_n \leq m$, $\mathbf{T}'_{i,j} = 0$ for $i > i_j$, and $\mathbf{T}'_{i_j,j} \neq 0$, $j = 1, \ldots, n$.* □

Any full column rank scalar matrix may be transformed by column operations to upper echelon form and the corresponding pivot set is unique. The (row) pivot set of any matrix is thus well defined as the pivot set of its (column) upper echelon forms. In addition we have:

Lemma 3.2 *The pivot set I is maximal for $\det \mathbf{T}'_{I,*}$ being nonzero, that is, if $I' = (i'_1, \ldots, i'_\ell, i'_{\ell+1}, \ldots, i'_n)$ with $i'_\ell > i_\ell$ and $i'_k \geq i_k$, $\ell + 1 \leq k \leq n$, then $\det \mathbf{T}'_{I',*} = 0$.*

Proof. By Definition 3.1, the rows of indices $i'_\ell, \ldots, i'_n$ cannot be linearly independent since their first ℓ entries are zero. □

Definition 2.1 is generalized for non-square matrices.

Definition 3.3 *A matrix polynomial $\mathbf{T}(z) \in \mathbb{Q}^{m \times n}[z]$ is called* $\vec{\alpha}$-column reduced *with degree $\vec{\alpha}$ if there exists a full column rank scalar matrix $\mathbf{T}'$ with pivot set I satisfying*

$$z^{-\vec{\alpha}} \cdot \mathbf{T}(z) \cdot z^{\vec{\alpha}_I - \vec{\alpha}} = \mathbf{T}' + \mathcal{O}(z^{-1})_{z \to \infty}. \tag{5}$$

If $\mathbf{T}'$ is in upper echelon form then $\mathbf{T}(z)$ is said to be in $\vec{\alpha}$-quasi Popov form. *If it satisfies the additional normalization degree and leading coefficient constraint*

$$z^{-\vec{\alpha}} \cdot \mathbf{T}(z)_{I,*} = \mathbf{I}_n + \mathcal{O}(z^{-1})_{z \to \infty}. \tag{6}$$

then $\mathbf{T}(z)$ is said to be in $\vec{\alpha}$–Popov normal form. □

By extension of Definition 3.1, we will refer to the set $I = (i_1, \ldots, i_n)$ of Definition 3.3 as the pivot set of an $\vec{\alpha}$-column reduced matrix $\mathbf{T}(z)$. If $\mathbf{T}(z)$ in addition is in $\vec{\alpha}$-quasi Popov form then its entry $\mathbf{T}(z)_{i_j,j}$ – referred to as the jth pivot element – has precise degree $\vec{\alpha}_j$. Notice that the statement of Lemma 2.3 also holds in the rectangular case. In particular, $\mathbf{A}(z)$ is $\vec{\alpha}$-column reduced if and only if $z^{-\vec{\alpha}} \cdot \mathbf{A}(z)$ is $\vec{0}$-column reduced. Non normal shifted forms have been defined in [20] with a slightly different definition of the $\vec{\alpha}$-degree.

The next example gives another illustration of the link between shifted forms of polynomial matrices and the integer case. We use the fact that a shift vector with only one nonzero entry may be used to "select" a given row in a matrix. As shown in [11, Algorithm 2] for integers, this is useful in solving the extended gcd problem.

Example 3.4 *To compute the gcd $c(z)$ of $a(z) = z^2 - 4z + 3$ and $b(z) = z^3 - 7z^2 + 14z - 8$ together with a corresponding multiplier $\mathbf{U}(z)$ such that $[a(z)\ b(z)] \cdot \mathbf{U}(z) = [c(z)\ 0]$, one may compute a shifted form equivalent to*

$$\mathbf{A}(z) = \begin{bmatrix} z^2-4z+3 & z^3-7z^2+14z-8 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This matrix is built using the identity matrix, to keep track of the column operations that are performed and that will give $\mathbf{U}(z)$. A $\vec{0}$-column reduced form of $\mathbf{A}(z)$ is

$$\begin{bmatrix} z^2-4z+3 & -z+1 \\ 1 & -z+3 \\ 0 & 1 \end{bmatrix}$$

but does not provide the target result. One could verify here that a $[-2,0,0]$-reduced form:

$$\begin{bmatrix} -z+1 & 0 \\ -z+3 & -z^2+6z-8 \\ 1 & z-3 \end{bmatrix}.$$

leads to $c(z)$ and to a possible $\mathbf{U}(z)$. □

The generalization of Remark 2.2 and in particular of identity (3) relies on classical tools from linear system theory. For a matrix polynomial $\mathbf{A}(z)$ of full column rank we define the *Minor degree* - denoted by Minor-deg $\mathbf{A}(z)$ - as the maximum of the degrees of the determinants of $n \times n$ submatrices of $\mathbf{A}(z)$ (see [12, Eq. (34), p.454]). This can be naturally extended to Laurent matrix polynomials. Setting $\vec{\alpha}^\star = \text{cdeg}\,(z^{-\vec{a}}\mathbf{A}(z))$ where *cdeg* denotes the unshifted column degree, it is well-known that

$$|\vec{\alpha}^\star| \geq \text{Minor-deg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z)),$$

with equality iff $\mathbf{A}(z)$ is $\vec{a}$–column reduced [12, § 6.3.2, p.384]. In this latter case, with pivot set I, Lemma 3.2 provides the additional information

$$\begin{aligned} \text{Minor-deg}\,(z^{-\vec{a}}\mathbf{A}(z)) &= \deg\det\,(z^{-\vec{a}_I}\mathbf{A}(z)_{I,\star}) \\ &= |\vec{\alpha}^\star| \\ &> \deg\det\,(z^{-\vec{a}_{I'}}\mathbf{A}(z)_{I',\star}) \end{aligned} \tag{7}$$

where I' is any list as in the lemma.

Example 3.5 *The $\vec{0}$-Popov normal of*

$$\mathbf{A}(z) = \begin{bmatrix} z^2 & -z^2+1 & z^4-z^2+z \\ z-1 & -z+2 & z^3-2z^2 \\ z & 0 & 1 \\ 1 & -1 & z^2 \\ z+1 & -z-2 & z^3+3z^2 \end{bmatrix}$$

has pivot set $I=(1,3,5)$ and vector degree $\vec{\alpha}=[2,1,2]$:

$$\mathbf{T}(z) = \begin{bmatrix} \underline{z^2}-1 & 1 & z \\ z-2 & 1 & 0 \\ 0 & \underline{z} & 1 \\ 1 & 0 & 0 \\ z+2 & -1 & \underline{z^2} \end{bmatrix}.$$

The form reveals that $\mathbf{A}(z)$ has Minor degree $|\vec{\alpha}|=5$. Here, the degrees are also called the minimal indices *of the module generated by the columns of $\mathbf{A}(z)$ (see also Definition 5.1) [9]. With $\vec{a}=[0,0,0,0,3]$, the shifted form*

$$\mathbf{T}^{\vec{a}}(z) = \begin{bmatrix} \underline{z} & 0 & 1 \\ 0 & \underline{z}-1 & 1 \\ 1 & 0 & \underline{z} \\ 0 & 1 & 0 \\ z^2 & -z^3+z+1 & -1 \end{bmatrix}$$

has pivot set $I=(1,2,3)$ and vector degree $\vec{\alpha}=[1,1,1]$. Now $|\vec{\alpha}|=3$ is the Minor degree of the first 4 rows of $\mathbf{A}(z)$. □

Theorem 3.7 below shows that any polynomial matrix can be transformed to a unique $\vec{a}$-Popov form by multiplication on the right with some unimodular matrix. We first state a lemma which gives a useful property of column reduced matrices (see [12, Theorem 6.3-13, p387]).

Lemma 3.6 (Predictable-Degree Property)
Let $\mathbf{B}(z)$ with $\text{cdeg}\,z^{-\vec{b}}\cdot\mathbf{B}(z)=\vec{\beta}^\star$, be a $\vec{b}$-column reduced matrix polynomial. If $\mathbf{P}(z)$ and $\mathbf{C}(z)$ are two matrix polynomials such that $\mathbf{B}(z)\mathbf{P}(z)=\mathbf{C}(z)$ with $\text{cdeg}\,z^{-\vec{b}}\cdot\mathbf{C}(z)=\vec{\gamma}^\star$ then $\deg\mathbf{P}(z)_{i,j}\leq\vec{\gamma}^\star_j-\vec{\beta}^\star_i$. □

Theorem 3.7 ($\vec{a}$-Popov form)
Any matrix $\mathbf{A}(z)\in\mathbb{Q}^{m\times n}[z]$ of full column rank is equivalent to a unique matrix $\mathbf{T}(z)$ in $\vec{a}$–Popov normal form.

Proof: The existence of a column reduced form $\mathbf{B}(z)$ equivalent to $\mathbf{A}(z)$ is a classical fact, see [12, p 386] for instance. The same applies directly to shifted forms. Following (5) let $\mathbf{B}'$ be the corresponding leading scalar matrix. We denote by $\vec{\beta}^\star$ the unshifted degree:

$$\begin{aligned} z^{-\vec{a}}\cdot\mathbf{A}(z)\cdot\mathbf{U}(z)\cdot z^{-\vec{\beta}^\star} &= z^{-\vec{a}}\cdot\mathbf{B}(z)\cdot z^{-\vec{\beta}^\star} \\ &= \mathbf{B}'+\mathcal{O}(z^{-1})_{z\to\infty} \end{aligned}$$

and let I be the pivot set of $\mathbf{B}'$. If $\mathbf{T}(z)_{I,\star}=\mathbf{B}(z)_{I,*}\mathbf{V}(z)$ is in $\vec{a}_I$-Popov form (Definition 2.1 in the nonsingular case) then $\mathbf{T}(z)=\mathbf{A}(z)\cdot\mathbf{U}(z)\cdot\mathbf{V}(z)$ is in $\vec{a}$-Popov form. Indeed, by Remark 2.2, the components of the unshifted degree $\vec{\alpha}^\star$ of $z^{-\vec{a}_I}\cdot\mathbf{T}(z)_{I,*}$ are given by a permutation of the components of $\vec{\beta}^\star$: $\vec{\alpha}^\star=\text{perm}\,(\vec{\beta}^\star)$. Lemma 3.6 implies that $\deg\,[\mathbf{V}(z)]_{i,j}\leq\vec{\alpha}^\star_j-\vec{\beta}^\star_i$. Thus all the entries in the j-th column of $z^{-\vec{a}}\cdot\mathbf{T}(z)$ have degrees bounded by $\vec{\alpha}^\star_j$. Using the second identity of (7) we have Minor-deg $(z^{-\vec{a}}\mathbf{T}(z))=$ Minor-deg $(z^{-\vec{a}}\mathbf{B}(z))$ thus the two matrices must have the same pivot set I, otherwise the inequality in (7) would be violated. For the same reason, it follows that below a pivot entry in $z^{-\vec{a}}\mathbf{T}(z)$ the degrees must be strictly less than $\vec{\alpha}^\star_j$, this establishes the existence of the form.

By (7), as used above, any two $\vec{a}$-column reduced matrices that are unimodular column equivalent have the same pivot set. From the uniqueness of the form in the nonsingular case, if $\mathbf{T}^{(1)}(z)$ and $\mathbf{T}^{(2)}(z)$ are two $\vec{a}$-Popov forms of $\mathbf{A}(z)$ then $\mathbf{T}^{(1)}(z)_{I,*}=\mathbf{T}^{(2)}(z)_{I,*}$. This implies that the two corresponding transformation matrices are the same and that $\mathbf{T}^{(1)}(z)=\mathbf{T}^{(2)}(z)$. □

We have seen in the preceding proof that, for any two column equivalent and $\vec{a}$-column reduced matrices, the pivot sets coincide. Nevertheless, by (7), the pivot set will in general depend on $\vec{a}$. For instance, generalizing Lemma 2.4 to the full rank case we may define Hermite shifts $\vec{a}$ by

$$\vec{a}_{i_j}-\vec{a}_i\geq\vec{\alpha}_j,\ \text{for } i>i_j.$$

In such cases, I will consist of the largest row indices such that $\mathbf{A}(z)_{I,\star}$ is nonsingular.

4 Degree bounds for multipliers

In this section we give degree bounds for the unimodular multiplier $\mathbf{U}(z)$ used to transform a given matrix polynomial $\mathbf{A}(z)$ into an $\vec{a}$–Popov form (or some similar form) $\mathbf{T}(z)$. The

bounds will be used in the next section to embed the shifted Popov problem into one of computing a certain basis of the kernel of an associated matrix polynomial, a computation that can be efficiently done in computer algebra systems.

Our degree estimates will be given in terms of a free parameter $\vec{c}$ which may be chosen in order to reflect particular properties of the input $\mathbf{A}(z)$ (e.g., in the case where $\mathbf{A}(z)$ is $\vec{c}$–column reduced). In the first part of the next theorem our estimates are formulated in terms of the inputs $\mathbf{A}(z), \vec{a}, \vec{c}$, and of the invariants $\mathbf{T}(z), \vec{\alpha}, I$. The aim of the second part is to estimate the invariants in terms of the input.

Theorem 4.1 (Degree bounds for Multiplier) *Let $\mathbf{A}(z) \cdot \mathbf{U}(z) = \mathbf{T}(z)$ with $\mathbf{U}(z)$ unimodular and $\mathbf{T}(z)$ being $\vec{a}$–column reduced with pivot set I and degree $\vec{\alpha}$. Furthermore, for some arbitrary multi-index $\vec{c}$, let* [1]

$$\vec{\gamma}^{*\vec{c}} = \operatorname{cdeg}(z^{-\vec{c}} \cdot \mathbf{A}(z)), \quad \vec{\tau}^{*\vec{c}} = \operatorname{cdeg}(z^{-\vec{c}} \cdot \mathbf{T}(z)), \qquad (8)$$

and define $\Delta^{\vec{c}} := |\vec{\gamma}^{*\vec{c}}| + |\vec{c}_I| - |\vec{\alpha}|$. Then

$$\deg \mathbf{U}(z)_{k,j} \le \vec{\tau}_j^{*\vec{c}} - \vec{\gamma}_k^{*\vec{c}} + \Delta^{\vec{c}}, \quad j,k = 1,\ldots,n. \qquad (9)$$

Also

$$0 \le -\text{Minor-deg}(z^{-\vec{c}} \cdot \mathbf{A}(z) \cdot z^{-\vec{\gamma}^{*\vec{c}}}) \le \Delta^{\vec{c}} \le |\vec{\gamma}^{*\vec{c}}| + |\vec{c}_I|, \qquad (10)$$

and $\Delta^{\vec{c}} = 0$ iff $\mathbf{A}(z)$ is $\vec{c}$–column reduced. In addition,

$$\vec{\tau}^{*\vec{c}} \le \vec{\alpha} - \vec{a}_I + \vec{c} \cdot \max[\vec{a} - \vec{c}], \qquad (11)$$

$$\vec{\tau}^{*\vec{a}} = \vec{\alpha} - \vec{a}_I \le \operatorname{perm}(\vec{\gamma}^{*\vec{a}}). \qquad (12)$$

Before giving a proof of Theorem 4.1 we give some examples which illustrate the sharpness of (9) for different choices of $\vec{c}$. In the examples $\mathbf{A}(z)$ will always be nonsingular so $m = n$ and $I = (1,\ldots,m)$. Also, we will just compare the choices $\vec{c} = \vec{a} \ge \vec{0}$ and $\vec{c} = \vec{0}$. If $\vec{\gamma}^{*\vec{0}} = \operatorname{cdeg} \mathbf{A}(z)$ and $\vec{\gamma}^{*\vec{a}} = \operatorname{cdeg}(z^{-\vec{a}}\mathbf{A}(z))$ are of the same magnitude, then $\Delta^{\vec{0}} \approx \Delta^{\vec{a}} - |\vec{a}| \le \Delta^{\vec{a}} - \max[\vec{a}]$, and from (9) and (11) we see that the choice $\vec{c} = \vec{0}$ leads to tighter bounds (compare with $\mathbf{A}^{(3)}(z)$ below). If the shift is relevant for the degree bounds, typically when $\Delta^{\vec{a}} \le \Delta^{\vec{0}}$ (see matrix $\mathbf{A}^{(2)}(z)$), then the choice $\vec{c} = \vec{a}$ will be more appropriate. However, the optimal choice of the parameter $\vec{c}$ remains an open problem.

Example 4.2 *Given the shift $\vec{a} = [0,6,9]$, we have $\vec{\gamma}^{*\vec{0}} = \vec{\gamma}^{*\vec{a}} + \vec{a} = [8,3,1]$ (and thus $\Delta^{\vec{a}} = \Delta^{\vec{0}}$) for the matrix*

$$\mathbf{A}^{(1)}(z) = \begin{bmatrix} 2z^8 & 0 & 0 \\ z & z^3+z & 0 \\ z+1 & z+1 & 2z \end{bmatrix}$$

which is $\vec{a}$-column reduced. The transformation matrix for the $\vec{a}$-Popov form is constant, and the two estimates for $\vec{c} \in \{\vec{0}, \vec{a}\}$ give the exact degree 0 for most of its coefficients.

For the shift $\vec{a} = [7,2,0]$ on the matrix,

$$\mathbf{A}^{(2)}(z) = \begin{bmatrix} 2z^{10}+z & z^8-1 & z^{11}+1 \\ z^5+1 & 2z^3-1 & z^4+1 \\ -z^2 & 2 & 2z \end{bmatrix}$$

one gets $\vec{\gamma}^{*\vec{a}} = \vec{a} + [10,3,4]$ and $\vec{\gamma}^{*\vec{0}} = [10,8,11]$. The $\vec{a}$-Popov form of $\mathbf{A}^{(2)}(z)$ has the degree vector $\vec{\alpha} = [10,3,3]$. With $\Delta^{\vec{a}} = 1$ and $\Delta^{\vec{0}} = 13 > \Delta^{\vec{a}}$, the bound for $\vec{c} = \vec{a}$ is precise and gives the exact maximum degree 3 for the entries of the transformation, when the other one is pessimistic and gives a maximum degree 15.

The $[10,3,0]$-Popov form of

$$\mathbf{A}^{(3)}(z) = \begin{bmatrix} 2z+1 & 2z^5+z^4+2z^3+2z & 2z^2+2 \\ 2z+1 & z^5+z^3+2z^2+1 & z^2+1 \\ z+1 & z^5+2z^4+2z^2+z & 2z^2+z \end{bmatrix}$$

is in upper triangular Hermite form with $\vec{\alpha} = [8,0,0]$. In this case $\Delta^{\vec{a}} = 21 > \Delta^{\vec{0}} = 0$. The entries of the transformation $\mathbf{U}(z)$ have degree at most 7 and all the degrees are exactly predicted by the choice $\vec{c} = \vec{0}$. On the other hand the bound for $\vec{c} = \vec{a}$ gives a maximum degree 12. □

Proof of Theorem 4.1: For the remainder of the proof we fix $\vec{c}$ and use the shorter notation $\vec{\gamma} = \vec{\gamma}^{*\vec{c}}$, $\vec{\tau} = \vec{\tau}^{*\vec{c}}$. For a proof of (9), we recall that $\deg \det \mathbf{A}(z)_{I,*} = |\vec{\alpha}|$. Since $\mathbf{A}(z)_{i,j} \le \vec{c}_i - \vec{\gamma}_j$, we get

$$z^{\vec{\gamma}} \cdot (\operatorname{adj} \mathbf{A}(z)_{I,*}) \cdot z^{\vec{c}_I} = \mathcal{O}(z^{|\vec{\gamma}|+|\vec{c}_I|}).$$

Therefore, by applying (8) and Cramer's rule, the quantity $z^{\vec{\gamma}} \cdot \mathbf{U}(z) \cdot z^{-\vec{\tau}}$ reduces to

$$\frac{z^{\vec{\gamma}} \cdot \operatorname{adj} \mathbf{A}(z)_{I,*} \cdot z^{\vec{c}_I}}{\det \mathbf{A}(z)_{I,*}} \cdot z^{-\vec{c}_I} \cdot \mathbf{T}(z)_{I,*} \cdot z^{-\vec{\tau}} = \mathcal{O}(z^{\Delta^{\vec{c}}}),$$

giving equation (9).

In order to show (10), recall first that $z^{-\vec{c}} \cdot \mathbf{A}(z) \cdot z^{-\vec{\gamma}}$ is a polynomial in $1/z$ by the definition of $\vec{\gamma}$. Consequently,

$$\begin{aligned} 0 &\ge \text{Minor-deg}(z^{-\vec{c}} \cdot \mathbf{A}(z) \cdot z^{-\vec{\gamma}}) \\ &\ge \text{Minor-deg}(z^{-\vec{c}_I} \cdot \mathbf{A}(z)_{I,*} \cdot z^{-\vec{\gamma}}), \end{aligned}$$

the latter quantity being equal to $-\Delta^{\vec{c}}$. It remains to discuss the case $\Delta^{\vec{c}} = 0$. By (7) and the above inequalities, $\Delta^{\vec{c}} = 0$ is equivalent to the facts that $\vec{\gamma}$ is the column degree both of $z^{-\vec{c}_I} \cdot \mathbf{A}(z)_{I,*}$ and $z^{-\vec{c}} \cdot \mathbf{A}(z)$, and that both matrices are column reduced, as claimed in (10).

Assertion (11) and the first part of (12) follow immediately from the definition of $\vec{\tau}$. In order to show the final estimate for $\vec{c} = \vec{a}$, we introduce the unimodular matrix $\mathbf{V}(z) = \mathbf{U}(z)^{-1}$. Since $z^{-\vec{a}}\mathbf{A}(z) = z^{-\vec{a}}\mathbf{T}(z) \cdot \mathbf{V}(z)$ and $z^{-\vec{a}}\mathbf{T}(z)$ is $\vec{0}$–column reduced, it follows from the Predictable-Degree Property (Lemma 3.6) and (8) that $z^{\vec{\tau}} \cdot \mathbf{V}(z) \cdot z^{-\vec{\gamma}} = \mathcal{O}(z^0)$. Since $\mathbf{V}(z)$ is nonsingular, we find some permutation ρ such that $\mathbf{V}(z)_{j,\rho(j)} \ne 0$, leading to (12). □

Remark 4.3 *Recall that $\vec{\alpha}_j$ is the (a priori unknown) degree of a pivot element of the normal form, in particular it is nonnegative. Thus, in terms of the input data, combining (9) with (11) gives the weaker degree bound*

$$\deg \mathbf{U}(z)_{k,j} \le |\vec{c}_I| - \vec{a}_{i_j} + |\vec{\gamma}^{*\vec{c}}| - \vec{\gamma}_k^{*\vec{c}} + \max[\vec{a} - \vec{c}].$$

In particular, in terms of $d = \deg \mathbf{A}(z)$, we obtain in the unshifted Popov case ($\vec{a} = \vec{c} = \vec{0}$)

$$\deg \mathbf{U}(z) \le (n-1) \cdot d, \quad \deg \mathbf{T}(z) \le d.$$

[1] As seen in the proof below, estimates (9) and (12) remain valid if instead of (8) we only assume that $\vec{\gamma}^{*\vec{c}} \ge \operatorname{cdeg}(z^{-\vec{c}_I} \cdot \mathbf{A}(z)_{I,*})$ and $\vec{\tau}^{*\vec{c}} \ge \operatorname{cdeg}(z^{-\vec{c}_I} \cdot \mathbf{T}(z)_{I,*})$.

In contrast, in the square Hermite case ($m = n$, $\vec{c} = \vec{0}$, $\vec{a}_j = \sum_{k=j}^{m} \vec{\alpha}_k$, $j = 1, ..., m$ as in (4), and thus $\vec{\tau}_j^{*\vec{c}} \leq |\vec{\alpha}|$ by (11)) we have

$$\deg \mathbf{U}(z) \leq (n-1) \cdot d,$$
$$\deg \mathbf{T}(z) \leq |\vec{\alpha}| = \deg \det \mathbf{A}(z) \leq n \cdot d.$$

□

5 Computing Popov Forms via Minimal Polynomial Bases

In this section we show that shifted Popov forms $\mathbf{T}(z)$ of $\mathbf{A}(z)$ together with their multiplier $\mathbf{U}(z)$ can be obtained by a particular polynomial basis for the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$ (considered as a module over $\mathbb{Q}[z]$). It makes intuitive sense to look at the kernel since the columns of the $(m+n) \times n$ matrix polynomial

$$\mathbf{S}(z) = \begin{bmatrix} \mathbf{U}(z) \\ \mathbf{T}(z) \end{bmatrix} \tag{13}$$

lie in the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$. These columns form a basis iff $\mathbf{T}(z) = \mathbf{A}(z) \cdot \mathbf{U}(z)$ and $\mathbf{U}(z)$ is unimodular. Finally, as observed for example in [19], we may use the algorithms FPHPS and SPHPS of [3] to compute column reduced polynomial bases of matrix polynomials, otherwise known as minimal polynomial bases (MPB). In our case we wish to use a shifted version of a basis of the kernel to compute shifted Popov forms.

Definition 5.1 (Shifted Minimal Polynomial Bases)

Let $\mathbf{C}(z) \in \mathbb{Q}^{m \times n}[z]$ be of rank r and $\mathbf{B}(z) \in \mathbb{Q}^{n \times (n-r)}[z]$ irreducible (of full rank for all finite values of z in the complex plane), with $\mathbf{C}(z) \cdot \mathbf{B}(z) = 0$. If in addition $\mathbf{B}(z)$ is $\vec{b}$-column reduced then it is called a $\vec{b}$-Minimal Polynomial Basis ($\vec{b}$-MPB) *of the $\mathbb{Q}[z]$-module kernel of $\mathbf{C}(z)$.* □

The existence and the uniqueness of a $\vec{b}$-MPB in $\vec{b}$-Popov form follows from Theorem 3.7. If $\vec{b} = 0$ then Definition 5.1 gives the classical definition of a Minimal Polynomial Basis (MPB) [9]. Such bases are called *minimal* since if a MPB for the kernel of $\mathbf{A}(z)$ has degree $\vec{\beta}$ then any other basis has degree $\vec{\beta}' \geq \mathrm{perm}(\vec{\beta})$ [12, §6.5.4, p456].

Theorem 5.2 (Popov forms via MPB)
Let $\mathbf{A}(z)$ be an $m \times n$ matrix polynomial of full column rank, and $\vec{a}, \vec{c}$ multi-indices. Furthermore, let $\vec{\gamma}^{*\vec{c}} \geq \mathrm{cdeg}(z^{-\vec{c}}\mathbf{A}(z))$, and write $\vec{n}(N) = (N \cdot \vec{e} - \vec{\gamma}^{*\vec{c}}, \vec{a})$ for any integer N.

The matrix polynomial $\mathbf{S}(z)$ is of the form (13) with $\mathbf{T}(z)$ the $\vec{a}$–Popov form of $\mathbf{A}(z)$ and $\mathbf{U}(z)$ its corresponding multiplier if and only if, for some

$$N \geq N_1 := \Delta^{\vec{c}} + \max[\vec{a} - \vec{c}],$$

$\mathbf{S}(z)$ is a MPB of the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$ in $\vec{n}(N)$–Popov form.

In this case, the latter property is true for all $N \geq N_1$. Furthermore, $\mathbf{S}(z)$ and $\mathbf{T}(z)$ have the same pivots, and the same shifted column degree.

Proof: For the first implication it is sufficient to show that the stacked matrix $\mathbf{S}(z)$ constructed with help of the Popov form $\mathbf{T}(z)$ and the multiplier $\mathbf{U}(z)$ is $\vec{n}(N)$–column reduced for all $N \geq N_1$, with the properties as specified in the last part of Theorem 5.2. Indeed, denoting by $\vec{a}$ the $\vec{a}$ column degree of $\mathbf{T}(z)$ and by I the corresponding pivot set, we get

$$z^{-\vec{n}(N)} \cdot \mathbf{S}(z) \cdot z^{-\vec{\alpha}+\vec{a}_I} = \begin{bmatrix} z^{\vec{\gamma}^{*\vec{c}}} \cdot \mathbf{U}(z) \cdot z^{-\vec{\alpha}+\vec{a}_I - N \cdot \vec{e}} \\ z^{-\vec{a}} \cdot \mathbf{T}(z) \cdot z^{-\vec{\alpha}+\vec{a}_I} \end{bmatrix},$$

and it only remains to show that

$$z^{\vec{\gamma}^{*\vec{c}}} \cdot \mathbf{U}(z) \cdot z^{-\vec{\alpha}+\vec{a}_I - N \cdot \vec{e}} = \mathcal{O}(z^0)_{z \to \infty}$$

for all $N \geq N_1$. This latter statement follows however from (9) and (11).

In order to show the other implication, let $\mathbf{S}(z)$ (and $\widetilde{\mathbf{S}}(z)$) be as described in the first part of the assertion (and of the second part, respectively). Then the columns of both matrices form bases of the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$, and therefore there exists some unimodular $\mathbf{W}(z)$ such that $\widetilde{\mathbf{S}}(z) = \mathbf{S}(z) \cdot \mathbf{W}(z)$. On the other hand, both $\mathbf{S}(z)$ and $\widetilde{\mathbf{S}}(z)$ are of full column rank and in $\vec{n}(N)$–Popov form for the specified value of N, and thus $\mathbf{S}(z) = \widetilde{\mathbf{S}}(z)$ by the uniqueness of $\vec{n}(N)$–Popov forms. □

Theorem 5.2 implies that one can compute shifted forms by computing shifted forms of bases matrices for an associated kernel. That this is useful is shown by Theorem 5.3 below which implies that one can use the Mahler basis algorithm FFFG of [4] in order to compute an MPB $\mathbf{S}(z)$ in shifted Popov form. In addition, FFFG only uses fraction-free arithmetic, an advantage if the entries in the original matrix are polynomials having coefficients from an integral domain (for example a matrix with entries from $\mathbf{Z}[z]$ or $\mathbb{Q}[a_1, \ldots, a_k][z]$).

To be more precise, let us briefly recall the definition of order for vector Hermite–Padé approximation (see, e.g., [5] or [4, Example 2.3]): given an $m \times s$ matrix $\mathbf{F}(z)$ of polynomials (or formal power series) and a multi-index $\vec{\sigma}$ of length m, a vector $\mathbf{P} \in \mathbb{Q}^s[z]$ is said to have order $\vec{\sigma}$ if $z^{-\vec{\sigma}} \cdot \mathbf{F}(z) \cdot \mathbf{P}(z) = \mathcal{O}(1)_{z \to 0}$. Obviously, the set of polynomial vectors having a given order is a submodule of $\mathbb{Q}^s[z]$, containing the kernel of $\mathbf{F}(z)$. The method FFFG mentioned above determines $s \times s$ bases of such order modules in shifted Popov form, by successively increasing one (arbitrary) component of the order vector by one unit. For sufficiently large order vectors and suitable shifts, we may recover our desired MPB in shifted Popov form as a part of such an order basis.

Theorem 5.3 *Let $\mathbf{A}(z), \vec{a}, \vec{c}, \vec{\gamma}^{*\vec{c}}, \vec{n}$ be as in Theorem 5.2, and define $\vec{n} := (N \cdot \vec{e} - \vec{\gamma}^{*\vec{c}}, \vec{a})$, where $N \geq N_1$.*

Apply algorithm FFFG to the matrix $[\mathbf{A}(z), -\mathbf{I}_m]$ along the offdiagonal path induced by $\vec{n}$ and increasing order vectors $\vec{\sigma} \geq \vec{0}$ in order to obtain fraction-free polynomial bases $\mathbf{M}(z)$ of $\vec{n}$–column degree $\vec{\nu}$ along with polynomial *residuals*

$$\mathbf{R}(z) = z^{-\vec{\sigma}} \cdot [\mathbf{A}(z), -\mathbf{I}_m] \cdot \mathbf{M}(z) \in \mathbb{Q}^{m \times (n+m)}[z],$$

and stop the algorithm if a $m \times n$ submatrix $\mathbf{R}(z)_{*,L}$ of $\mathbf{R}(z)$ is zero.

Then there exists some scalar $c \neq 0$ such that the last m rows of $\mathbf{M}(z)_{*,L}$ give c times the $\vec{a}$–Popov form of $\mathbf{A}(z)$ with

$\vec{a}$-column degree $\vec{\alpha} = \vec{\nu}_L$ and pivot set $(\ell_1 - n, ..., \ell_n - n)$, where $L = (\ell_1, ..., \ell_n)$, and the first n rows of $\mathbf{M}(z)_{*,L}$ give c times the corresponding unimodular multiplier.

The algorithm will terminate at the latest by the order vector

$$\vec{\sigma} \geq \vec{\sigma}^f := \vec{c} + (N + 1 + \max[\vec{\gamma}^{*\cdot\vec{a}}])\vec{e}.$$

Proof: Notice that, for any order vector $\vec{\sigma}$, $\mathbf{R}(0)$ is of full rank m [5, Lemma 2.8], and $\mathbf{M}(z)/c$ is a nonsingular matrix in $\vec{n}$–Popov form for some scalar $c \neq 0$ [4, Theorem 7.2], with its degree vector denoted by $\vec{\nu}$. Let $\mathbf{S}(z)$ be as in the first part of Theorem 5.2. Then the columns of $\mathbf{S}(z)$ have order $\vec{\sigma}$, and from [4, Theorem 7.3(a)] we know that there exists a unique matrix polynomial $\mathbf{P}(z)$ such that

$$\mathbf{S}(z) = \mathbf{M}(z)\mathbf{P}(z), \quad \text{and} \quad z^{\vec{\nu}-\vec{n}} \cdot \mathbf{P}(z) \cdot z^{-\vec{\alpha}+\vec{a}_I} = \mathcal{O}(1)_{z\to\infty}. \tag{14}$$

If the index set L is as above, then the columns of $\mathbf{M}(z)_{*,L}$ are elements of the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$. From the basis property of $\mathbf{S}(z)$ we may conclude that there exists a unique matrix polynomial $\mathbf{Q}(z)$ such that $\mathbf{M}(z)_{*,L} = \mathbf{S}(z)\mathbf{Q}(z)$. In combination with (14) we obtain $\mathbf{M}(z)_{*,L} = \mathbf{M}(z)\mathbf{P}(z)\mathbf{Q}(z)$. Consequently, $\mathbf{P}(z)_{L_c,*} = \mathbf{0}$, and $\mathbf{P}(z)_{L,*} = \mathbf{Q}(z)^{-1}$ is unimodular. Uniqueness of shifted Popov forms gives us the desired result $\mathbf{S}(z) = \mathbf{M}(z)_{*,L}/c$.

In order to prove the last part, notice that

$$\begin{aligned} z^{\vec{\sigma}-\vec{c}} \cdot \mathbf{R}(z) \cdot z^{\vec{n}-\vec{\nu}} &= z^{-\vec{a}} \cdot [\mathbf{A}(z), -\mathbf{I}_m] \cdot \mathbf{M}(z) \cdot z^{\vec{n}-\vec{\nu}} \\ &= [z^{-\vec{c}} \cdot \mathbf{A}(z) \cdot z^{-\vec{\gamma}^{*\vec{c}}}, -\mathbf{I}_s] \\ &\quad \cdot \begin{bmatrix} z^N \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & z^{\vec{a}-\vec{c}} \end{bmatrix} \cdot z^{-\vec{n}} \cdot \mathbf{M}(z) \cdot z^{\vec{n}-\vec{\nu}}. \end{aligned}$$

Since $N - \max[\vec{a} - \vec{c}] \geq \Delta^{\vec{c}} \geq 0$ by assumption on N and relation (10) of Theorem 4.1, it follows that $z^{\vec{\sigma}-\vec{c}} \cdot \mathbf{R}(z) \cdot z^{\vec{n}-\vec{\nu}-N\vec{e}} = \mathcal{O}(z^0)_{z\to\infty}$. We now may choose an index list L with $L_c := (1, ..., m+n) \setminus L$ such that the square matrix $\mathbf{R}(z)_{*,L_c}$ is invertible. Consequently, there exists a bijective map $\rho : \{1, ..., m\} \to L_c$ with $\mathbf{R}(z)_{j,\rho(j)} \neq 0$, $j = 1, .., m$, and thus

$$\operatorname{perm}(\vec{\sigma} - \vec{c} - N\vec{e}) \leq (\vec{\nu} - \vec{n})_{L_c}$$

where we recall that perm denotes a permutation of the components. If now $\vec{\sigma} \geq \sigma^f = \vec{c} + N' \cdot \vec{e}$, then $\min[(\vec{\nu} - \vec{n})_{L_c}] \geq N' - N \geq \max[\vec{\gamma}^{*\vec{a}}] + 1 \geq \max[\vec{\alpha} - \vec{a}_I] + 1$ by assumption on N' and (12). It follows that $\mathbf{P}(z)_{L_c,*} = \mathbf{0}$ in (14). On the other hand, with $\mathbf{S}(z)$, $\mathbf{P}(z)$ must also have full column rank, and thus $\mathbf{P}(z)_{L,*}$ is invertible. Multiplying the left-hand equation in (14) on the left with $[\mathbf{A}(z), -\mathbf{I}_m]$, we obtain $\mathbf{R}(z)_{*,L} \cdot \mathbf{P}(z)_{L,*} = \mathbf{0}$ and hence $\mathbf{R}(z)_{*,L} = \mathbf{0}$, as required for the final part of Theorem 5.3. □

The fraction free computation of shifted Popov forms via FFFG has been implemented in the computer algebra system MAPLE and can be obtained from the authors web sites or via email.

Example 5.4 *Let $\mathbf{A}(z)$ be as in Example 2.5 with $\vec{a} = \vec{0}$ and $\vec{c} = cdeg\ \mathbf{A}(z)$. Then applying the FFFG algorithm along the path $[6, 10, 6, 8, 4, 0]$ results in a Mahler system that produces an integer multiple of a $[8, 4, 0]$-Popov normal form $\mathbf{A}(z)$ along with a unimodular multiplier in 18 iterations. The result is 4 times the true answer because it works with fraction-free arithmetic. The unimodular multiplier in this case is given by*

$$\begin{bmatrix} 4z^2 + 4z & z + 1 & -2z + 2 \\ -4z^2 + 16 & -z & 2z - 4 \\ -4z^2 - 8z & -z - 2 & 2z \end{bmatrix}$$

again 4 times the true answer. □

We can give a worst case complexity for this approach in terms of $d = \deg \mathbf{A}(z)$. As in the discussion at the end of section 3 we will restrict ourselves to the choices $\vec{c} = 0$ and $\vec{a} \geq 0$, $\max[\vec{a}] \leq |\vec{\alpha}|$, which includes the classical Popov and Hermite normal forms. Then $\vec{\gamma}^{*\vec{a}} \leq \vec{\gamma}^{\vec{0}} \leq d \cdot \vec{e}$, and $\Delta^{\vec{0}} + \max[\vec{a} - \vec{c}] \leq |\vec{\gamma}^{*\vec{0}}| \leq n \cdot d$. Consequently, a practical choice for N would be $N = n \cdot d$, leading to $|\vec{\sigma}^f| = m \cdot n \cdot d + m \cdot (d + 1)$.

From [5, Lemma 2.8] and [4, Section 4] we know that the module $K := |\vec{\nu}|$ of the degree vector of the final basis in FFFG coincides with the module of the final order vector $|\vec{\sigma}| \leq |\vec{\sigma}^f|$. It is shown in [3, 4] that the complexity of FFFG to reach this order vector is bounded by $\mathcal{O}((m + n) \cdot K^2)$ in floating point arithmetic, and by $\mathcal{O}((m + n) \cdot \kappa^2 \cdot K^4)$ in exact arithmetic, where κ is an upper bound for the size (in bits) of the coefficients of $\mathbf{A}(z)$ (remember that the corresponding complexities for solving the underlying systems of linear equations by Gaussian elimination are obtained by replacing $m + n$ above by K). Under the above assumptions, we obtain rough worst case bounds in terms of the degree $d = \deg A(z)$, namely $\mathcal{O}(m^3 \cdot n^2 \cdot d^2)$ floating point operations, and $\mathcal{O}(\kappa^2 \cdot m^5 \cdot n^4 \cdot d^4)$ bit operations. There are no other fraction-free algorithms for Popov or Hermite that can be used for comparison with our approach.

6 Conclusions

In this paper we have studied reduced and normal forms of matrix polynomials by looking at so called *shifted* forms of matrix polynomials. These forms include column-reduced, triangular, Hermite normal and Popov normal forms along with their shifted counterparts. We have determined degree bounds for a unimodular matrix which transforms the input matrix polynomial into an equivalent matrix in the desired form. The degree bounds allow one to embed a shifted Popov normal problem into a problem of determining a minimal polynomial basis in shifted Popov form for an associated stacked matrix polynomial. These shifted minimal polynomial bases can in turn be computed in a fraction-free way via the Mahler system algorithm of [4].

As mentioned in the introduction the results in this paper can be viewed as a first step in a program to obtain efficient symbolic methods for computing matrix normal forms of arbitrary matrix polynomials. The full row rank rectangular case is important for computing matrix polynomial GCDs in normal form. The singular case gives information on minimal polynomial bases for the kernel of the matrix polynomial. The case of singular or rectangular matrices is considerably more complex because one no longer has a unique unimodular multiplier. In such cases one needs to determine a unimodular multiplier with minimal degree properties [6]. Degree bounds for this case are also more difficult. At the same time degree bounds for such multipliers in this case also lead to interesting degree bounds for important classes

of problems of interest in computer algebra. For example, from [6] we have:

Theorem 6.1 (GCD of several scalar polynomials) *Let* $\mathbf{A}(z) = [a_1(z), a_2(z), \ldots a_n(z)] \in \mathbb{Q}^{1\times n}[z]$ *with degrees* $\vec{\gamma} = [\gamma_1, \gamma_2, \ldots \gamma_n]$ *and* $d(z) = GCD(a_1(z), \ldots, a_n(z))$ *with degree* δ. *Assume (without loss of generality) that* $\gamma_1 = \min_j \gamma_j$ *and* $\gamma_n = \max_j \gamma_j$. *Then there are "small" multipliers* $u_k(z)$ *for the diophantine equation*

$$a_1(z)\cdot u_1(z) + a_2(z)\cdot u_2(z) + \ldots + a_n(z)\cdot u_n(z) = d(z),$$

which satisfy

$$\deg u_1(z) \le \gamma_n - \delta - 1, \quad \sum_{\substack{k=2 \\ u_k \neq 0}}^{n} (1 + \deg u_k(z)) \le \gamma_1 - \delta. \quad (15)$$

Notice that these bounds include the classical one for $n = 2$ (cf. [10]). Also, a straight forward generalization of the integer bound of [11] to the polynomial case would lead to the weaker estimate $\deg u_k(z) \le \gamma_n - 1$ for all k.

There are a number of interesting problems that still remain to be solved. We have shown that it is possible to solve the shifted Popov form problem via some fraction free algorithm by noting that it is embedded inside an order basis computation (or an MPB computation). The major problem with using such an approach to compute our form is that this method is not really a reduction procedure. In particular it does not recognize when a matrix polynomial $\mathbf{A}(z)$ is in shifted Popov form until the final step of the computation. We are interested in obtaining a fraction-free algorithm which computes minimal polynomial bases (and hence our normal form) in a reduction procedure. We expect that this may be done by determining an associated linear system along with determinental representations as in [4] and then making use of modified Schur complements as done in [1].

References

[1] Beckermann, B., Cabay, S., and Labahn, G. Fraction-free computation of matrix padé systems. In *Proceedings of ISSAC 97* (Maui, USA, 21–23 July 1997), W. Küchlin, Ed., ACM Press, pp. 125–132.

[2] Beckermann, B., and Labahn, G. A uniform approach for hermite padé and simultaneous padé approximants and their matrix generalizations. *Numerical Algorithms 3* (1992), 45–54.

[3] Beckermann, B., and Labahn, G. A uniform approach for the fast, reliable computation of matrix-type padé approximants. *SIAM J. Matrix Anal. Appl. 15* (1994), 804 823.

[4] Beckermann, B., and Labahn, G. Fraction-free computation of matrix gcd's and rational interpolants. *Submitted to SIAM J. Matrix Anal. Appl.* (1997), 45pgs.

[5] Beckermann, B., and Labahn, G. Recursiveness in matrix rational interpolation problems. *J. Comput. Appl. Math. 77* (1997), 5–34.

[6] Beckermann, B., Labahn, G., and Villard, G. Shifted normal forms of general polynomial matrices. manuscript. 1999.

[7] Beelen, T., van der Hurk, G., and Praagman, C. A new method for computing a column reduced polynomial matrix. *Systems and Control Letters 10* (1988), 217–224.

[8] Bultheel, A., and Van Barel, M. A matrix euclidean algorithm and the matrix minimal padé approximation problem. *Continued Fractions and Padé Approximants* (1990).

[9] G.D. Forney, J. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM J. Control 13* (1975), 493 520.

[10] Geddes, K., Czapor, S., and Labahn, G. *Algorithms for Computer Algebra.* Kluwer, Boston, MA, 1992.

[11] Havas, G., Majewski, B., and Matthews, K. Extended gcd and hermite normal form algorithms via lattice basis reduction. *Experimental Mathematics 7(2)* (1998), 125–135.

[12] Kailath, T. *Linear Systems.* Prentice–Hall, 1980.

[13] Kaltofen, E., Krishnamoorthy, M., and Saunders, B. Parallel algorithms for matrix normal forms. *Linear Algebra and its Applications 136* (1990), 189–208.

[14] Kannan, R. Solving systems of linear equations over polynomials. *Theoretical Computer Science 39* (1985), 69 88.

[15] MacDuffee, C. *The Theory of Matrices.* Chelsea, New-York, 1956.

[16] Newman, M. *Integral Matrices.* Academic Press, New-York, 1972.

[17] Popov, V. Some properties of control systems with irreducible matrix transfer functions. *Lecture Notes in Mathematics 144* (1969), 169–180.

[18] Schrijver, A. *Theory of Linear and Integer Programming.* Wiley-Interscience series in Discrete Mathematics, 1986.

[19] Stuchlik-Quéré, M. How to compute minimal bases using padé approximants. Rapport de recherche lip6 1997/035, Laboratoire d'Informatique de Paris 6, France, 1997.

[20] Van Barel, M., and Bultheel, A. A generalized minimal partial realization problem. *Linear Algebra and its Applications 254* (1997), 527–551.

[21] Verghese, G., and Kailath, T. Rational matrix structure. *IEEE Trans. Automat. Control 26* (1981), 434 438.

[22] Villard, G. Computing popov and hermite forms of polynomial matrices. In *Proceedings of ISSAC 96* (Zurich, Switzerland, 24–26 July 1996), Y. Lakshman, Ed., ACM Press, pp. 250–258.