HAL
open science

Normal Forms for General Polynomial Matrices

Bernd Beckermann, George Labahn, Gilles Villard

To cite this version:

Bernd Beckermann, George Labahn, Gilles Villard. Normal Forms for General Polynomial Matrices. [Research Report] LIP RR-2002-1, Laboratoire de l'informatique du parallélisme. 2002, 2+31p. hal-02101933

HAL Id: hal-02101933

https://hal-lara.archives-ouvertes.fr/hal-02101933

Submitted on 17 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon
Unité Mixte de Recherche CNRS-INRIA-ENS LYON n° 5668

Normal Forms for General Polynomial Matrices

Bernd Beckermann, George Labahn and Gilles Villard

Janvier 2002

Research Report N° 2002-1

École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.37
Télécopieur : +33(0)4.72.72.80.80
Adresse électronique : lip@ens-lyon.fr

Normal Forms for General Polynomial Matrices

Bernd Beckermann, George Labahn
and Gilles Villard

Janvier 2002

Abstract

We present an algorithm for the computation of a *shifted Popov Normal Form* of a rectangular polynomial matrix. For specific input shifts, we obtain methods for computing the *matrix greatest common divisor* of two matrix polynomials (in normal form) or such polynomial normal form computation as the classical *Popov form* and the *Hermite Normal Form*. The method is done by embedding the problem of computing shifted forms into one of computing matrix rational approximants. This has the advantage of allowing for fraction-free computations over integral domains such as $\mathbf{Z}[z]$ or $\mathsf{K}[z_1, \ldots, z_n][z]$.
In the case of rectangular matrix input, the corresponding multipliers for the shifted forms are not unique. We use the concept of minimal matrix approximants to introduce a notion of *minimal mutipliers* and show how such multipliers are computed by our methods.

Keywords: Popov Normal Form, Hermite Normal Form, Matrix Gcd, Exact Arithmetic, Fraction-Free Algorithm.

Résumé

Nous présentons un algorithme pour le calcul de la *forme normale de Popov à décalage* d'une matrice polynomiale rectangulaire. Pour des entrées particulières l'algorithme fournit des méthodes pour le calcul du *pgcd de deux polynômes matriciels* ou pour le calcul de *formes normales* telles que celles de *Popov* ou d'*Hermite*. Notre solution consiste à réduire le calcul d'une forme à décalage au calcul d'un approximant rationnel matriciel. Cette solution a notamment l'avantage de permettre une approche sans fraction sur des anneaux comme $\mathbf{Z}[z]$ ou $\mathsf{K}[z_1, \ldots, z_n][z]$.
Pour une matrice rectangulaire générale, les matrices de transformations associées ne sont pas uniques. Nous utilisons la notion d'approximants minimaux pour introduire et calculer des *transformations minimales*.

Mots-clés: Forme normale de Popov, forme normale d'Hermite, pgcd de matrices, arithmétique exacte, algorithme sans fraction.

NORMAL FORMS FOR GENERAL POLYNOMIAL MATRICES [1]

Bernhard Beckermann

Laboratoire d'Analyse Numérique et d'Optimisation
Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France
bbecker@ano.univ-lille1.fr http://ano.univ-lille1.fr/bb/bb_page.html

George Labahn

Department of Computer Science
University of Waterloo, Waterloo, Ontario, Canada
glabahn@scg.uwaterloo.ca http://www.scg.uwaterloo.ca/~glabahn

Gilles Villard

Cnrs, Laboratoire Lip, École Normale Supérieure de Lyon
46, Allée d'Italie, 69364 Lyon Cedex 07, France
Gilles.Villard@ens-lyon.fr http://www.ens-lyon.fr/~gvillard

Abstract

We present an algorithm for the computation of a *shifted Popov Normal Form* of a rectangular polynomial matrix. For specific input shifts, we obtain methods for computing the *matrix greatest common divisor* of two matrix polynomials (in normal form) or such polynomial normal form computation as the classical *Popov form* and the *Hermite Normal Form*. The method is done by embedding the problem of computing shifted forms into one of computing matrix rational approximants. This has the advantage of allowing for fraction-free computations over integral domains such as $\mathbf{Z}[z]$ or $\mathsf{K}[z_1, \ldots, z_n][z]$.

In the case of rectangular matrix input, the corresponding multipliers for the shifted forms are not unique. We use the concept of minimal matrix approximants to introduce a notion of *minimal mutipliers* and show how such multipliers are computed by our methods.

Keywords. Popov Normal Form, Hermite Normal Form, Matrix Gcd, Exact Arithmetic, Fraction-Free Algorithm.

[1] Also publication Ano, Université de Lille.

This material is based on work supported in part by the Centre National de la Recherche Scientifique, Actions Incitatives No 5929 et Stic LinBox 2001 (Villard).

1 Introduction

Two polynomial matrices $\mathbf{A}(z)$ and $\mathbf{B}(z)$ in $\mathsf{K}[z]^{m\times n}$, K a field, are column equivalent if there exists a unimodular matrix $\mathbf{U}(z) \in \mathsf{K}[z]^{n\times n}$ 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. There exists a number of normal forms for such an equivalence problem, the best known being the *Hermite normal form*, initially discovered by Hermite [16] for the domain of integers [22, 25]. This is an upper triangular matrix that has the added constraints that the diagonals have the largest degrees in each row. While a triangular form has many obvious advantages for such operations as solving linear equations it also has certain disadvantages. In particular, the column degrees of a Hermite normal form can increase. For such reasons the *Popov normal form* [26] of a polynomial matrix can give a better form for many problems. This is a form that specifies certain normalizations of matrix leading row coefficients and which has the property that columns degrees are minimal over all column equivalent matrices.

A related problem is the computation of *matrix greatest common divisors*. For two matrix polynomials $\mathbf{A}(z)$ and $\mathbf{B}(z)$, both with the same row dimension, a left matrix Gcd is a matrix polynomial $\mathbf{C}(z)$ satisfying $\mathbf{A}(z) = \mathbf{C}(z)\cdot\hat{\mathbf{A}}(z)$ and $\mathbf{B}(z) = \mathbf{C}(z)\cdot\hat{\mathbf{B}}(z)$ and where $\hat{\mathbf{A}}(z)$ and $\hat{\mathbf{B}}(z)$ have only unimodular left divisors. The Matrix Gcd plays an important role in such areas as linear systems theory [18], minimal partial realizations and other application areas. Normal forms plays two important roles in the Matrix Gcd problem. In the first place Matrix Gcd's are only unique up to multiplication on the right by unimodular polynomial matrices. In order to specify a single answer one asks that the Gcd be in a specific normal form. In addition, matrix Gcd's are usually computed by converting the rectangular matrix polynomial $[\mathbf{A}(z), \mathbf{B}(z)]$ into a normal form $[\mathbf{0}, \mathbf{C}(z)]$ where $\mathbf{C}(z)$ is precisely the Matrix Gcd in normal form.

Shifted normal forms. The solution of a number of normal form problems involving matrix polynomials – particularly the forms mentioned above and others that we will introduce as examples in the paper – may be unified by the notion of *shifted form* [6]. Matrix normal forms such as Hermite or Popov have certain degree structures in their requirements. They are typically computed by reducing degrees in certain rows or columns in such a way that the degree requirements will eventually be met. An input shift associated to an input matrix $\mathbf{A}(z)$ is a vector $\vec{a}$ of integers that may be seen as weights attributed to the rows of the matrix (see Definition 2.3). These weights govern the order in which operations are performed during the algorithm and thus allow – by different ways of chosing them – one to use the same process for the computation of different forms including for example the Popov and the Hermite forms. One can illustrate this using the two latter forms as an example. The column degrees of

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

may be reduced by unimodular column transformations to obtain

$$\mathbf{T}(z) = \begin{bmatrix} z & -1 \\ 1 & z-1 \end{bmatrix}$$

the Popov form of $\mathbf{A}(z)$. With the shift $\vec{a} = [0, -1]$ one will give preference to the degrees – and

hence to the elimination – in the second row over those in the first row. This leads to

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

which is the Hermite normal form of $\mathbf{A}(z)$ (see [6, Lemma 2.4] and the definition in Section 2). Additional examples of the use of shifted normal forms are also included later in this paper. For example, it can happen that certain preferences of a whole set of rows of a rectangular matrix will provide a *right inverse computation* for such a matrix (cf. Example 2.4).

Shifted Popov forms were introduced in [5] as a natural and convenient normal form for describing the properties of *Mahler systems*. These systems are used as basic building blocks for recursively computing solutions to module bases for matrix rational approximation and matrix rational interpolation problems (see also [4] in the case of matrix Padé systems and section 4.2). Shifts also appear naturally in the context of computing normal forms of matrices over $\mathbf{Z}$. Schrijver has shown that integer weights on the rows of a matrix $\mathbf{A}$ in $\mathbf{Z}^{n\times n}$ leads to a lattice whose reduced basis gives the Hermite form of $\mathbf{A}$ [28, p74]; a similar approach is found in [15, §6] where the powers of a positive integer γ give an appropriate shift. For matrix polynomials we develop a more complete study.

Computing shifted forms. The primary aim in this paper is to give a new algorithm for computing a shifted Popov form of an arbitrary rank polynomial matrix. In the case where the input is square and *nonsingular* or rectangular and of *full column rank*, an algorithm for determining a shifted Popov form has been given in [6]. We will extend the methods to deal with the general case of singular matrices, in particular for those not having full column rank. The Matrix Gcd problem is an important example where one requires normal forms for arbitrary rank matrices. The reader may refer to Examples 2.5 and 3.2. We will also see *coprime matrix rational functions* with Example 3.6. The new algorithm solves two difficulties. The first one concerns the fact that – unlike in the full column rank case – the multipliers $\mathbf{U}(z)$ are not unique. We will define and compute *minimal multipliers*. The second difficulty is the intermediate expression swell with respect to polynomial coefficients. This will be solved by proposing a *fraction-free* method.

Our methods view the equation $\mathbf{A}(z) \cdot \mathbf{U}(z) = \mathbf{T}(z)$ as a kernel computation

$$[\mathbf{A}(z), -\mathbf{I}] \cdot \begin{bmatrix} \mathbf{U}(z) \\ \mathbf{T}(z) \end{bmatrix} = 0 \tag{1}$$

and the normal form problem as one of computing a special shifted form of a basis for the kernel of $[\mathbf{A}(z), -\mathbf{I}]$. This special form provides both the normal form $\mathbf{T}(z)$ along with a unimodular multiplier $\mathbf{U}(z)$ having certain minimal degree properties. The minimality of degrees overcomes the problem that the unimodular multiplier is not unique. Starting from (1), the same idea of computing column reduced forms using minimal polynomial basis has been used by Beelen *et al.* [7]. We generalize the approach by using non-constant shifts and by computing different forms.

For various purposes, the shift technique may also be applied to the unimodular multipliers. In addition to the input shift $\vec{a}$ associated to $\mathbf{A}(z)$ we introduce a second shift $\vec{b}$ associated to $\mathbf{U}(z)$. Specific choices of $\vec{b}$ result in unimodular matrices $\mathbf{U}(z)$ having special properties that are useful for *nullspace* and *polynomial system solving* problems. For example, two bases for the nullspace of

$$\mathbf{A}(z) = \left[\begin{array}{ccc|c} z^3 + 4\,z^2 - 2 & z^2 + 2\,z - 1 & z^3 + z^2 & z^2 - 1 \\ z^4 + z^3 + 4\,z^2 + 2\,z & z^3 + z^2 + z & 2\,z^3 + 2\,z^2 & z^2 + z \end{array}\right]$$

are given by

$$\mathbf{U}(z) = \begin{bmatrix} z+1 & 0 \\ -z^2-z & -z-1 \\ -1 & z+1 \\ -2-z & -z^2+1-z \end{bmatrix} \text{ and } \mathbf{U}'(z) = \begin{bmatrix} -z^3-2\,z^2+1 & z^2-1 \\ z^4+2\,z^3-z^2-4\,z-2 & -z^3+2\,z+1 \\ 2\,z^2+4\,z+1 & -2\,z \\ 0 & \underline{1} \end{bmatrix}.$$

These two bases are submatrices of multipliers for shifted Popov forms of $\mathbf{A}(z)$ (see the detailed study at section 3). The first basis $\mathbf{U}(z)$, computed with no shift, has the smallest possible degrees. The second one, $\mathbf{U}'(z)$, is computed with a shift $\vec{b} = [0, 0, 0, -3]$ which forces a preference of elimination in its last row. By definition, the second column $\mathbf{U}'_{\mathbf{2}}(z)$ of $\mathbf{U}'(z)$ satisfies $\mathbf{A}(z)\mathbf{U}'_{\mathbf{2}}(z) = 0$ and since the last entry has been made constant by the shift, one may solve the Diophantine linear system given by $\mathbf{A}(z)$ and its last column. This is not directly possible from $\mathbf{U}(z)$.

The *shifted multipliers* are also useful for deriving degree bounds and analyzing the cost of the algorithm. Bounds for the minimal unimodular multiplier are obtained in terms of the input parameters and the invariants to the problem – the *shifted minimal degrees* and the *shifted Kronecker indices*. The invariants are themselves bounded in terms of the input parameters. The bounds that we obtain for the minimal unimodular multiplier are interesting in their own right. Indeed, as it will be shown at Example 5.4 they can be used for instance to determine bounds for the cofactor polynomials in the extended Gcd problem for $n \geq 2$ scalar polynomials, that is degree bounds for the $u_k(z)$ in

$$a_1(z) \cdot u_1(z) + \cdots + a_n(z) \cdot u_n(z) = Gcd\,(a_1(z), \ldots, a_n(z)).$$

Following (1), the computation of a shifted form for the kernel of $[\mathbf{A}(z), -\mathbf{I}]$ is done using the σ-basis algorithm of Beckermann and Labahn [5]. This algorithm computes *all* solutions to a rational approximation problem, that of vector Hermite-Padé approximants to a certain order. The basis for this approximation problem is in shifted Popov form and includes the desired kernel for high enough orders. The algorithm has the advantage that the computations are fraction-free for integral domains, a significant advantage when the input is parameterized, for example when the input entries are from $\mathbf{Z}[z]$ or $\mathsf{K}[a_1, \ldots, a_n][z]$, classical domains for computer algebra systems. To our knowledge, a specific fraction-free method for computing Hermite or (shifted) Popov forms for general matrix polynomials has not been previously given.

Algorithms and complexities to compute the Popov form or column reduced forms over $\mathsf{K}[z]$ with K an abstract field have been studied in [34] (see also the references therein) or in [24]. Many algorithms have been proposed to compute the Hermite form over $\mathsf{K}[z]$, the reader may refer to [30] for an overview of the domain. For concrete coefficient domains like $\mathbf{Z}$, expression swell on the coefficient level leads in general to a severe breakdown of the methods performance. The case of matrices over $\mathbf{Z}[z]$ has only been considered for the Hermite form using chinese remaindering in [29, Chap. 4 & 6]. Our idea to introduce fraction-free techniques to handle the complexity of coefficients for general matrix polynomial computations (shifted forms) is a natural solution.

Organization of the paper. Section 2 gives the basic definitions of our normal forms along with proofs of existence and uniqueness. Section 3 discusses shifted minimal polynomial bases and shifted minimal multipliers. Section 4 gives the main algorithm for the computation of the shifted forms. We first show in 4.1 how they may be computed as minimal polynomial bases and

then how to compute these bases as approximant order bases in 4.2. The termination of the algorithm is specified using invariants of the problems (minimal degrees and Kronecker indices). These invariants are studied in Section 5 where the concern is to bound the degrees of the unimodular multipliers. From these bounds, Section 6 gives the cost of the algorithm. The last section includes a conclusion along with a discussion of future research directions.

Notations. Except for the presentation and the cost study of the fraction-free technique where the domain of the entries will be precised, we will work with constant and polynomial matrices over K and $\mathsf{K}[z]$ for an abstract commutative field K. Given a 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 $\mathbf{A}(z)$, where 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).

For any multi-index $\vec{a}$ (that is, a vector of integers) we denote by $|\vec{a}|$ the sum of its components, $\max[\vec{a}]$ and $\min[\vec{a}]$ its maximum and minimum components, $\operatorname{perm}[\vec{a}]$ a permutation of its components and $\vec{a}_I$ the subvector given by the indices in I. The multi-index $\vec{e}$ is the vector $(1, \ldots, 1)$ of appropriate size. Three main types of multi-index are involved. The shifts will be indicated by latin letters (*e.g.* $\vec{a}$, $\vec{b}$), we will use greek letters (*e.g.* $\vec{\alpha}$, $\vec{\beta}$) for the shifted column degrees and add a "*" to the latters (*e.g.* $\vec{\alpha}^*$, $\vec{\beta}^*$) for the actual (non-shifted) column degrees (see Definition 2.3). The multi-index given by the column degrees of matrix $\mathbf{A}(z)$ is denoted by $\operatorname{cdeg}\mathbf{A}(z)$ and the row degree is denoted by $\operatorname{rdeg}\mathbf{A}(z)$.

2 Preliminaries

This section gives some of the basic definitions required for the remainder of the paper. We give examples of matrix normal forms and provide motivation for the concept of shifted normal forms for both square and rectangular matrix polynomials. Information for normal forms of matrix polynomials in the full column case have been handled in a previous paper [6].

The best known normal form for matrix polynomials is the Hermite normal form. This is a triangular matrix with the additional normalization properties than the diagonal polynomials are monic and that the degrees of the offdiagonal entries are less than the degrees of the diagonal entry in the same row. For example, the matrix

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

is in Hermite normal form. From a computational point of view, the Hermite normal form has the disadvantage that it does not minimize or even necessarily reduce the column degrees. A second well known matrix normal form was introduced by Popov [26]. Called the *Popov normal form* or the *polynomial-echelon form* [18], this form requires normalization properties similar to those of the Hermite form: the leading (by rows) coefficient matrix have to be normalized to the identity. Specifically we have, in the case of nonsingular square matrices:

Definition 2.1 *An $m \times m$ non-singular matrix polynomial $\mathbf{T}(z) \in \mathsf{K}[z]^{m \times m}$ is in Popov form (with column degree $\vec{\alpha}^*$) if there exists a multi-index $\vec{\alpha}^*$ such that $\mathbf{T}(z)$ satisfies the degree constraints*

$$\mathbf{T}(z) \cdot z^{-\vec{\alpha}^*} = \mathbf{T}' + \mathcal{O}(z^{-1})_{z \to \infty}, \quad \mathbf{T}' \in \mathsf{K}^{m \times m} \text{ being upper triangular.} \tag{3}$$

$$z^{-\vec{\alpha}^*} \cdot \mathbf{T}(z) = \mathbf{I}_m + \mathcal{O}(z^{-1})_{z \to \infty}. \tag{4}$$

If only condition (3) holds with $\mathbf{T}' \in \mathsf{K}^{m\times m}$ *being simply nonsingular then the matrix polynomial is said to be column reduced.* □

Notice that in the first part of the definition the matrix $\mathbf{T}'$ in (3) is necessarily nonsingular because of (4). Up to a (unique) permutation of columns, we obtain the classical Popov normal form [18, §6.7.2, p. 481]. When the form is used in a matrix fraction description or as a minimal polynomial basis then the degree $\vec{\alpha}^*$ is usually referred to as the vector of *controllability* or *Kronecker indices*. It is known [18, 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 resulting polynomial matrix is unique. As an example, the Popov normal form for (2) is given by

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

with $\vec{\alpha}^* = [3,2,2]$ and

$$\mathbf{T}' = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since this matrix is column-reduced it gives the minimal column degrees of all matrices column equivalent to (2) (up to a permutation of the columns). That is, if $\mathbf{B}(z)$ is any other matrix column equivalent to the above then its column degrees are at least $(3,2,2)$ [18, §6.5.4, p. 456] (see Section 3).

Our example matrix (2) is not in Popov normal form - indeed it is not even in column reduced form. However, we can make a shift, that is give "weights" to the degrees of the rows by multiplication on the left by $z^{(-3,-1,0)}$ to make it a column reduced matrix. This has led in [6] to the extension of the notion of Popov forms to the notion of $\vec{a}$-Popov for full column rank matrix polynomials. The matrix (2) is said to be in $(3,1,0)$-Popov form with shifted degrees $(4,2,1)$. In order to include more general applications such as normal forms for Matrix Gcd's and for minimal nullspace bases, we will need to consider the more general case of rectangular matrix polynomials having an arbitrary column rank. The normalization of leading matrices is now ensured by column echelon matrices:

Definition 2.2 (Column echelon matrices) *A scalar matrix* $\mathbf{T}' \in \mathsf{K}^{m\times n}$ *of rank* r *is in upper echelon form with pivot set* $I = (i_1, ..., i_r)$ *if* $1 \le i_1 < i_2 < ... < i_r \le m$, $\mathbf{T}'_{i,j} = 0$ *for* $i > i_j$, *and* $\mathbf{T}'_{i_j,j} \neq 0$, $j = 1, ..., r$. □

Notice that we may transform any scalar matrix by column operations to upper echelon form and that 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. The general definition of shifted forms is then:

Definition 2.3 (Shifted-Popov Normal Form) *A matrix polynomial* $\mathbf{T}(z) \in \mathsf{K}[z]^{m\times n}$ *is called* $\vec{a}$-column reduced *with* $\vec{a}$*-degree* $\vec{\alpha}$ *(or shifted degree) if it may be partitioned as*

$$\mathbf{T}(z) = [\mathbf{0}, \quad \mathbf{T}(z)_{*,J_c}] \quad \textit{where } J = (1,2,...,n-r),\ J_c = (n-r+1,...,n), \tag{5}$$

and if its right hand part is such that there exists a full column rank scalar matrix $\mathbf{T}'$ *with pivot set* I *satisfying*

$$z^{-\vec{a}} \cdot \mathbf{T}(z)_{*,J_c} \cdot z^{\vec{a}_I - \vec{\alpha}} = \mathbf{T}' + \mathcal{O}(z^{-1})_{z\to\infty}. \tag{6}$$

If it satisfies the additional normalization degree and leading coefficient constraint

$$z^{-\vec{\alpha}} \cdot \mathbf{T}(z)_{I,J_c} = \mathbf{I}_r + \mathcal{O}(z^{-1})_{z\to\infty}. \tag{7}$$

then $\mathbf{T}(z)$ *is said to be in* $\vec{a}$–Popov normal form. *We define the multi-index* $\vec{\alpha}^*$ *by*

$$\vec{\alpha}^* = \operatorname{cdeg} z^{-\vec{a}} \cdot \mathbf{T}(z)_{*,J_c}$$

with the convention that though both the $\vec{a}$*-degree* $\vec{\alpha}$ *and the column degree* $\vec{\alpha}^*$ *have the same sizes, we index them in different ways –* $\vec{\alpha}$ *is indexed by the pivot rows* I *while* $\vec{\alpha}^*$ *is indexed by the columns* J_c. □

In the square, nonsingular case the pivot set is $(1,\ldots,m)$ and may be dropped from our notation. For the above forms, r equals the rank of $\mathbf{T}(z)$, $\mathbf{T}(z)_{*,J_c}$ and of $\mathbf{T}'$. Since $I = (i_1,\ldots,i_r)$ is the pivot set of $\mathbf{T}'$, the $r \times r$ sub-matrices $\mathbf{T}(z)_{I,J_c}$ and $\mathbf{T}'_{I,J_c}$ are invertible and for any other list $I' = (i'_1,..,i'_l,i'_{l+1},..,i'_r)$ with $i'_l > i_l$ for some l, $\mathbf{T}'_{I',J_c}$ is singular. By extension we will call the *pivot set* of an $\vec{a}$-column reduced matrix, the pivot set of an associated matrix $\mathbf{T}'$ in (6). Notice that Definition 2.3 may be used directly for Laurent matrix polynomials and that $\mathbf{A}(z)$ is $\vec{a}$-column reduced (respectively, in $\vec{a}$-Popov form) if and only if $z^{-\vec{a}} \cdot \mathbf{A}(z)$ is column reduced (respectively, in Popov form).

Example 2.4 *Consider the matrix*

$$\mathbf{A}(z) = \begin{bmatrix} 3\,z-6 & -3\,z & 6 \\ -3\,z+3 & 3\,z & -3 \\ 2\,z+3 & -2 & -2\,z-1 \\ z & -1 & -z+1 \end{bmatrix}$$

Its Popov form $\mathbf{T}^0(z)$ *and its* $(2,2,0,0)$*-Popov form* $\mathbf{T}(z)$ *provide two different bases for its dimension* 2 *column space:*

$$\mathbf{T}^0(z) = \begin{bmatrix} 0 & -z & -6 \\ 0 & \underline{z} & 3 \\ 0 & -\frac{2}{3} & 2\,z+1 \\ 0 & -\frac{1}{3} & \underline{z}-1 \end{bmatrix}, \quad \mathbf{T}(z) = \begin{bmatrix} 0 & -z^2+z-2 & 2z^2+z+4 \\ 0 & z^2-z+1 & -2z^2-z-2 \\ 0 & \underline{1} & 0 \\ 0 & 0 & \underline{1} \end{bmatrix}$$

with pivot sets $I^0 = (2,4)$ *and* $I = (3,4)$ *(pointed out by the underlined terms). The shifted degrees are* $\vec{\alpha}^0 = [1,1]$ *and* $\vec{\alpha} = [0,0]$. *A transformation matrix* $\mathbf{U}(z)$ *for* $\mathbf{T}(z)$ *will give a right inverse for* $\mathbf{A}(z)_{I,*}$ *since it satisfies* $\mathbf{A}(z)_{I,*}\mathbf{U}(z)_{*,(2,3)} = \mathbf{I}_2$. □

Example 2.5 *Consider the problem of computing the left Matrix Gcd of the two matrix polynomials*

$$\mathbf{P}(z) = \begin{bmatrix} -z^3+4z^2+z+1 & z-1 \\ -z^2+7z+4 & z+2 \end{bmatrix} \quad \textit{and} \quad \mathbf{Q}(z) = \begin{bmatrix} 2z^2+2z-2 & -z^2 \\ z^2+6z+6 & -2z \end{bmatrix}.$$

One determines a left Matrix Gcd by forming the augmented matrix $\mathbf{A}(z) = [\mathbf{P}(z), \mathbf{Q}(z)]$ *and then making this column equivalent to a matrix of the form* $[\mathbf{0}, \mathbf{C}(z)]$ *with* $\mathbf{C}(z)$ *a left Matrix Gcd. Popov and shifted Popov forms for* $\mathbf{A}(z)$ *gives Matrix Gcd's in normal form. For example, the Popov form for* $\mathbf{A}(z)$ *is given by*

$$\begin{bmatrix} 0 & 0 & \underline{z} & -1 \\ 0 & 0 & 2 & \underline{z} \end{bmatrix}$$

and so the last two columns give a Matrix Gcd in Popov form. On the other hand the $(2,0)$*-Popov form for* $\mathbf{A}(z)$ *is*

$$\begin{bmatrix} 0 & 0 & \underline{z}^2+2 & \frac{1}{2}z \\ 0 & 0 & 0 & \underline{1} \end{bmatrix}$$

and its last two columns give a Matrix Gcd in Hermite form (see [6, Lemma 2.4]). □

Using Definition 2.3, we can now handle the Hermite normal form more precisely. A matrix $\mathbf{T}(z)$ in $\mathsf{K}[z]^{m\times n}$ is in Hermite normal form (see [22, 25]) if it may be partitioned following (5) as

$$\mathbf{T}(z) = [\mathbf{0}, \quad \mathbf{T}(z)_{*,J_c}]$$

where $\mathbf{T}(z)_{*,J_c}$: (i) has full column rank r; (ii) is in upper echelon form with pivot set $I = (i_1, i_2, \ldots, i_r)$; (iii) satisfies the normalization constraint

$$z^{-\vec{\alpha}} \cdot \mathbf{T}(z)_{I,J_c} = \mathbf{I}_r + \mathcal{O}(z^{-1})_{z\to\infty}, \tag{8}$$

that is, the pivot entries $\mathbf{T}(z)_{i_j,n-r+j}$ are monic and have a degree $\vec{\alpha}_{i_j}$ strictly larger than the entries in the same row $\mathbf{T}(z)_{i_j,*}$. As mentioned already before, for an appropriate shift the Hermite form may be viewed as a shifted-Popov form:

Lemma 2.6 *If* $\mathbf{T}(z)$ *is in* $\vec{a}$*-Popov normal form with pivot set* $I = (i_1, i_2, \ldots, i_r)$ *and shifted degree* $\vec{\alpha}$ *such that*

$$\vec{a}_i - \vec{a}_l \geq \vec{\alpha}_i, \text{ for } i \in I \text{ and } l > i \tag{9}$$

then $\mathbf{T}(z)$ *is in Hermite normal form.*

Proof: Since $\mathbf{T}(z)$ is in $\vec{a}$-Popov normal form with shifted degree $\vec{\alpha}$, conditions (i) and (iii) for the Hermite form are clearly true. In addition, from identities (6) and (7), for $l > i_j$ we get $-\vec{a}_l + \deg \mathbf{T}(z)_{l,n-r+j} + \vec{a}_{i_j} - \vec{\alpha}_{i_j} < 0$. Thus if the shift satisfies (9) then $\mathbf{T}(z)_{l,n-r+j} = 0$ for $l > i_j$ and hence $\mathbf{T}(z)_{*,J_c}$ is in upper echelon form. □

A practical *a priori* shift $\vec{a}$ will depend on bounds for the degree $\vec{\alpha}$. This will be considered at section 6 using Theorem 5.1 below.

The existence and uniqueness of shifted Popov forms for rectangular matrix polynomials having full column rank can be found in [6, Theorem 3.5]. In the general case we have the following.

Theorem 2.7 *Any matrix* $\mathbf{A}(z) \in \mathsf{K}[z]^{m\times n}$ *is column equivalent to a unique matrix* $\mathbf{T}(z)$ *in* $\vec{a}$*-Popov form.*

Proof: As noted in [18, p. 375-376] one can always do elementary column operations so that the matrix polynomial $\mathbf{A}(z)$ is of the form $[\mathbf{0}, \mathbf{B}(z)]$ where the first $n-r$ columns are 0 and $\mathbf{B}(z)$ has full column rank r. The existence of an $\vec{a}$-Popov form therefore follows from the full column rank case. To show uniqueness, suppose $[\mathbf{0}, \mathbf{T}(z)]$ and $[\mathbf{0}, \mathbf{T}'(z)]$ are two $\vec{a}$-Popov forms for $\mathbf{A}(z)$. Then there exists a unimodular matrix polynomial $\mathbf{V}(z)$ such that

$$[\mathbf{0}, \mathbf{T}(z)] = [\mathbf{0}, \mathbf{T}'(z)] \cdot \mathbf{V}(z). \tag{10}$$

If $\mathbf{V}'(z)$ denotes the bottom right hand $r \times r$ submatrix of $\mathbf{V}(z)$ then (10) implies that

$$\mathbf{T}(z) = \mathbf{T}'(z) \cdot \mathbf{V}'(z). \tag{11}$$

Similarly, one can find a second $r \times r$ matrix polynomial $\mathbf{W}'(z)$ such that

$$\mathbf{T}'(z) = \mathbf{T}(z) \cdot \mathbf{W}'(z). \tag{12}$$

Therefore,

$$\mathbf{T}(z) = \mathbf{T}(z) \cdot \mathbf{W}'(z) \cdot \mathbf{V}'(z)$$

implying that $\mathbf{W}'(z) \cdot \mathbf{V}'(z) = \mathbf{I}_r$ since $\mathbf{T}(z)$ is of full column rank. Thus $\mathbf{V}'(z)$ and $\mathbf{W}'(z)$ are unimodular and $\mathbf{T}(z)$ and $\mathbf{T}'(z)$ are two full column rank equivalent $\vec{a}$-Popov forms. By [6, Theorem 3.5] they must be equal. □

3 Minimal Multipliers

In the case of full-column rank matrices, both an $\vec{a}$-Popov form $\mathbf{T}(z)$ and its unimodular multiplier $\mathbf{U}(z)$ are unique. In the case of singular input Theorem 2.7 implies that the shifted form is also unique. However, the same cannot be said for the associated unimodular multiplier. Indeed one simply needs to look at the definition to see that one can take multiples of any of the first $n-r$ columns of the multiplier and add them to the last r columns without having any effect on the associated shifted form. In this section we look for a multiplier that has a type of *shifted minimal* property. This type of reduction (without a shift) has already been done in the context of linear systems theory [35] (see Remark 3.7). We will show in a later section that the addition of a shift gives certain useful properties for our multipliers.

If $\mathbf{A}(z) \cdot \mathbf{U}(z) = [\mathbf{0}\ ,\ \mathbf{T}(z)_{*,J_c}]$ with $\mathbf{U}(z) = [\mathbf{U}(z)_{*,J},\ \mathbf{U}(z)_{*,J_c}]$ unimodular and $\mathbf{T}(z)_{*,J_c}$ of full column rank, then $\mathbf{A}(z) \cdot \mathbf{U}_{*,J}(z) = \mathbf{0}$. Since $\mathbf{U}(z)$ is unimodular, $\mathbf{U}(z)_{*,J}$ has full column rank and therefore forms a polynomial basis for the kernel of $\mathbf{A}(z)$, that is, a basis for the kernel as a module over $\mathsf{K}[z]$. The minimality of the columns J will thus be naturally captured using the well-known concept of *minimal polynomial basis* (see [13] or [18, §6.5.4]) which we extend here to include shifted bases.

Definition 3.1 (Shifted Minimal Polynomial Bases) *Let* $\mathbf{A}(z) \in \mathsf{K}[z]^{m\times n}$ *be of rank* r *and* $\mathbf{B}(z) \in \mathsf{K}[z]^{n\times(n-r)}$ *with* $\mathbf{A}(z) \cdot \mathbf{B}(z) = \mathbf{0}$. *If* $\mathbf{B}(z)$ *is* $\vec{b}$*-column reduced then* $\mathbf{B}(z)$ *is a* $\vec{b}$-Minimal Polynomial Basis ($\vec{b}$-MPB) *for the nullspace of* $\mathbf{A}(z)$. *If* $\mathbf{B}(z)$ *is also in* $\vec{b}$*-Popov form then* $\mathbf{B}(z)$ *is a* $\vec{b}$-Popov Minimal Polynomial Basis ($\vec{b}$-Popov MPB). □

If $\vec{b} = 0$ then Definition 3.1 gives the classical definition of a Minimal Polynomial Basis (MPB) [13]. Such bases are called *minimal* since if a MPB for the nullspace of $\mathbf{A}(z)$ has column degree $\vec{\beta}$ (with components in increasing order), then the degree $\vec{\beta}'$ (with components in increasing order) of any other basis satisfy $\vec{\beta}'_l \geq \vec{\beta}_l$, $1 \leq l \leq n-r$ [18, §6.5.4, p. 456]. Clearly, the same property holds for shifted MPB. The shifted degrees $\vec{\beta}_l$ may be called the *$\vec{b}$-right minimal* or *$\vec{b}$-right Kronecker* indices of $\mathbf{A}(z)$. The existence and the uniqueness of a shifted MPB in Popov form follows from Theorem 2.7.

Example 3.2 *Consider*

$$\mathbf{A}(z) = \begin{bmatrix} -z^3+4\,z^2+z+1 & z-1 & 2\,z^2+2\,z-2 & -z^2 \\ -z^2+7\,z+4 & z+2 & z^2+6\,z+6 & -2\,z \end{bmatrix},$$

the augmented matrix used in the Matrix Gcd problem of Example 2.5. The Popov MPB and $[0,-3,0,0]$-Popov MBP for the nullspace of $\mathbf{A}(z)$ have the same pivot set $K=(2,4)$:

$$\mathbf{U}^0(z)_{*,J} = \begin{bmatrix} -1 & -1 \\ \underline{z}^2-7 & -2\,z-7 \\ -z+3 & 3 \\ -1 & \underline{z} \end{bmatrix}, \quad \mathbf{U}(z)_{*,J} = \begin{bmatrix} -\frac{2}{21}\,z+\frac{1}{7} & -z^2-2\,z \\ \underline{1} & 0 \\ \frac{2}{21}\,z-\frac{3}{7} & z^2-z \\ \frac{2}{21}\,z^2-\frac{1}{3}\,z-\frac{4}{21} & \underline{z}^3-9\,z-7 \end{bmatrix}.$$

One may use the first column of $\mathbf{U}(z)_{*,J}$ with pivot 1 to write the second column of $\mathbf{A}(z)$ as a polynomial combination of the other columns. This combination is not seen directly from $\mathbf{U}^0(z)_{*,J}$. □

As described above, the minimal bases provide a normalization for the columns J of the multipliers $\mathbf{U}(z)$. If $\mathbf{U}(z)_{*,J}$ is a MPB with pivot set K, then the remaining columns J_c may be normalized by reducing their row degrees with respect to $\mathbf{U}(z)_{K,J}$. This leads to the notion of $(\vec{a},\vec{b})$-minimal multipliers:

Theorem 3.3 *Let $\mathbf{A}(z) \in \mathsf{K}[z]^{m\times n}$ of rank r, and $\vec{a}$, $\vec{b}$ be multi–indices of length m and n, respectively. Let $\mathbf{U}(z)$ be a unimodular matrix such that $\mathbf{A}(z)\cdot\mathbf{U}(z) = \mathbf{T}(z)$ with $\mathbf{T}(z)$ the unique $\vec{a}$–Popov normal form.*

(i) *A unimodular multiplier $\mathbf{U}(z)$ is unique up to multiplication on the right by matrices of the form*

$$\mathbf{W}(z) = \begin{bmatrix} \mathbf{W}(z)_{J,J} & \mathbf{W}(z)_{J,J_c} \\ \mathbf{0} & \mathbf{I}_r \end{bmatrix}, \quad \mathbf{W}(z)_{J,J} \in \mathsf{K}[z]^{(n-r)\times(n-r)} \text{ being unimodular.}$$

(ii) *There exists a unique multiplier $\mathbf{U}(z)$ verifying*

$$[\mathbf{U}(z)_{K,J}]^{-1}\cdot\mathbf{U}(z)_{K,J_c} = \mathcal{O}(z^{-1})_{z\to\infty}, \tag{13}$$

with $\mathbf{U}(z)_{*,J}$ being a $\vec{b}$-Popov MPB for the nullspace of $\mathbf{A}(z)$.

(iii) *Under all multipliers mentioned in (i), the sum of the degrees of the $\vec{b}$–column degrees of the unique multiplier $\mathbf{U}(z)$ of (ii) is minimal.*

We will refer to the unique multiplier $\mathbf{U}(z)$ satisfying (ii) as the $(\vec{a},\vec{b})$*–minimal multiplier* or, when $(\vec{a},\vec{b}) = (0,0)$, as the *minimal multiplier*.

Before proving Theorem 3.3 we give two lemmas, one which shows a useful property of column reduced matrices ([18, Theorem 6.3-13, p. 387]) and a second lemma which describes the division/remainder properties of polynomial matrices. These results will be used for the proof of the Theorem.

Lemma 3.4 (Predictable-Degree Property) *Let $\mathbf{B}(z)$ be a full column rank and $\vec{b}$-column reduced matrix polynomial with $\vec{\beta}^* = \operatorname{cdeg} z^{-\vec{b}} \cdot \mathbf{B}(z)$. If $\mathbf{P}(z)$ and $\mathbf{C}(z)$ are two matrix polynomials such that $\mathbf{B}(z) \cdot \mathbf{P}(z) = \mathbf{C}(z)$ with $\vec{\delta}^* = \operatorname{cdeg} z^{-\vec{b}} \cdot \mathbf{C}(z)$ then $\deg \mathbf{P}(z)_{i,j} \leq \vec{\delta}^*_j - \vec{\beta}^*_i$.* □

Lemma 3.5 (Matrix polynomials division) *Let $\mathbf{B}(z)$ be a nonsingular $m \times m$ matrix polynomial with $\vec{\beta}^* = \operatorname{cdeg} z^{-\vec{b}} \cdot \mathbf{B}(z)$. For any $m \times n$ matrix polynomial $\mathbf{A}(z)$ with $\vec{\delta}^* = \operatorname{cdeg} z^{-\vec{b}} \cdot \mathbf{A}(z)$ there exist unique matrix polynomials $\mathbf{Q}(z) \in \mathsf{K}[z]^{m\times n}$ and $\mathbf{R}(z) \in \mathsf{K}[z]^{m\times n}$ such that*

$$\begin{aligned} \mathbf{A}(z) &= \mathbf{B}(z)\mathbf{Q}(z) + \mathbf{R}(z), \\ \mathbf{B}(z)^{-1} \cdot \mathbf{R}(z) &= \mathcal{O}(z^{-1})_{z\to\infty}. \end{aligned} \tag{14}$$

If $\mathbf{B}(z)$ is $\vec{b}$-column reduced then $\deg \mathbf{Q}(z)_{i,j} \leq \vec{\delta}^*_j - \vec{\beta}^*_i$, for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Proof: The first statement is Theorem 6.3-15 in [18, p. 387]. For the second statement, the matrix quotient $\mathbf{Q}(z)$ is the polynomial part of $\mathbf{B}(z)^{-1}\mathbf{A}(z) = \operatorname{adj}(\mathbf{B}(z))\mathbf{A}(z)/(\det \mathbf{B}(z))$ where $\operatorname{adj}(\mathbf{B}(z))$ denotes the adjoint of $\mathbf{B}(z)$. Since $\mathbf{B}(z)\operatorname{adj}(\mathbf{B}(z)) = \operatorname{diag}(\det \mathbf{B}(z), \ldots, \det \mathbf{B}(z))$, by Lemma 3.4 we get that $\deg \operatorname{adj}(\mathbf{B}(z))_{i,j} \leq d - \vec{b}_j - \vec{\beta}^*_i$ with $d = \deg \det \mathbf{B}(z)$. It follows that $\deg(\operatorname{adj}(\mathbf{B}(z))\mathbf{A}(z))_{i,j} \leq (d - \vec{b}_l - \vec{\beta}^*_i) + (\vec{b}_l + \vec{\delta}^*_j) = d + \vec{\delta}^*_j - \vec{\beta}^*_i$ where the index l is from the matrix product. The quotient by $\det \mathbf{B}(z)$ then leads to $\deg \mathbf{Q}(z)_{i,j} \leq \vec{\delta}^*_j - \vec{\beta}^*_i$. □

Proof of Theorem 3.3: For statement (i), if $\mathbf{U}^{(1)}(z)$ and $\mathbf{U}^{(2)}(z)$ are two multipliers for the $\vec{a}$-Popov form, then $\mathbf{U}^{(1)}(z)_{*,J}$ and $\mathbf{U}^{(2)}(z)_{*,J}$ are two bases for the nullspace of $\mathbf{A}(z)$ and thus for the same $\mathsf{K}[z]$-module. Consequently there exists a unimodular multiplier $\mathbf{W}(z)_{J,J}$ which makes these matrices column equivalent. By the uniqueness of $\mathbf{T}(z)_{*,J_c}$, the columns of $\mathbf{U}^{(2)}(z)_{*,J_c} - \mathbf{U}^{(1)}(z)_{*,J_c}$ are in the nullspace of $\mathbf{A}(z)$ and there exists a matrix $\mathbf{W}(z)_{J,J_c}$ such that $(\mathbf{U}^{(2)}(z)_{*,J_c} - \mathbf{U}^{(1)}(z)_{*,J_c}) = \mathbf{U}^{(1)}(z)_{*,J}\mathbf{W}(z)_{J,J_c}$ or $\mathbf{U}^{(2)}(z)_{*,J_c} = \mathbf{U}^{(1)}(z)_{*,J_c} + \mathbf{U}^{(1)}(z)_{*,J}\mathbf{W}(z)_{J,J_c}$. This gives the general form of the multipliers as announced in (i).

For (ii), assume now that $\mathbf{U}(z)_{*,J}$ is the unique $\vec{b}$-Popov MPB for the nullspace, say with pivot set K, so that by definition $\mathbf{U}(z)_{K,J}$ is invertible. Given any multiplier $\mathbf{U}^{(0)}(z)$ we may thus divide $\mathbf{U}^{(0)}(z)_{K,J_c}$ on the left by $\mathbf{U}(z)_{K,J}$:

$$\mathbf{U}^{(0)}(z)_{K,J_c} = \mathbf{U}(z)_{K,J}\mathbf{W}(z)_{J,J_c} + \mathbf{U}(z)_{K,J_c}$$

and (13) is identity (14) of Lemma 3.5. Since in addition the matrix remainder $\mathbf{U}(z)_{K,J_c}$ is the unique matrix such that (13) is satisfied, using the generic form of a multiplier given at (i) and taking

$$\mathbf{U}(z)_{*,J_c} = \mathbf{U}^{(0)}(z)_{*,J_c} - \mathbf{U}(z)_{*,J}\mathbf{W}(z)_{J,J_c} \tag{15}$$

shows that the $(\vec{a},\vec{b})$-minimal multiplier $\mathbf{U}(z)$ is well-defined and unique. This proves (ii).

It remains to conclude the proof of (iii). Let $\mathbf{U}^{(0)}(z)$ be a second unimodular multiplier. From the general form of the multipliers, the sum of the degrees in the columns J and in the column J_c can be minimized independently. Since the degrees in the columns J are minimized by choosing a MPB, we only have to look at what happens in the columns J_c. Thus we need to show that the degree sum of $z^{-\vec{b}}\cdot\mathbf{U}^{(0)}(z)_{*,J_c}$ is at least the degree sum of $z^{-\vec{b}}\cdot\mathbf{U}(z)_{*,J_c}$. Let $\vec{\beta}^*$, $\vec{\delta}^*$, $\vec{\gamma}^*$ be the column degrees of $z^{-\vec{b}}\cdot\mathbf{U}(z)_{*,J}$, of $z^{-\vec{b}_K}\cdot\mathbf{U}^{(0)}(z)_{K,J_c}$ and of $z^{-\vec{b}_{K_c}}\cdot\mathbf{U}^{(0)}(z)_{K_c,J_c}$, respectively. The degree sum for the columns J_c of $\mathbf{U}^{(0)}(z)$ is $\sigma_{\min}=\Sigma_j\max\{\vec{\delta}^*_j,\vec{\gamma}^*_j\}$. By Lemma 3.5, we have a matrix quotient $\mathbf{W}(z)_{J,J_c}$ such that (15) is satisfied and where $\deg\mathbf{W}(z)_{i,j}\le\vec{\delta}^*_j-\vec{\beta}^*_i$. Therefore, by (15), after the division we have, for $1\le i\le m$ and $j\in J_c$:

$$\deg\left(z^{-\vec{b}}\cdot\mathbf{U}(z)\right)_{i,j}\le\max\{\max\{\vec{\delta}^*_j,\vec{\gamma}^*_j\},\vec{\delta}^*_j\}=\max\{\vec{\delta}^*_j,\vec{\gamma}^*_j\}.$$

This shows that the degree sum for the columns of J_c is not increased by the normalizing division and gives (iii). □

Example 3.6 (Coprime Rational Matrix Functions) *Suppose we are given a left coprime proper matrix rational function $\mathbf{R}(z)=\mathbf{D}(z)^{-1}\cdot\mathbf{N}(z)$ with $\mathbf{D}(z)$ square of size $p\times p$ in Popov form and $\mathbf{N}(z)$ of size $p\times q$. Then it is well known that there exists a right coprime matrix rational function* – i.e. *of the form $\mathbf{Q}(z)\cdot\mathbf{P}(z)^{-1}$ – for $\mathbf{R}(z)$. This is done as follows. Let*

$$\mathbf{A}(z)=[\mathbf{D}(z),\mathbf{N}(z)]$$

a matrix of size $m\times n$ with $m=p$ and $n=p+q$. Let $\mathbf{U}(z)$ be the unique minimal multiplier (with $\vec{a}=\vec{b}=0$) such that

$$\mathbf{A}(z)\cdot\mathbf{U}(z)=[\mathbf{0},\mathbf{I}_m],\tag{16}$$

the Popov form for $\mathbf{A}(z)$. In this case the pivots are given by $I=(1,\ldots,m)$ and, since $\mathbf{R}(z)$ is proper (degree constraints), by $K=(n-m+1,\ldots,m)$. Indentity (16) leads to:

$$\begin{array}{rcl}\mathbf{D}(z)\cdot\mathbf{U}(z)_{K_c,J}+\mathbf{N}(z)\cdot\mathbf{U}(z)_{K,J} &=& \mathbf{0}\\ \mathbf{D}(z)\cdot\mathbf{U}(z)_{K_c,J_c}+\mathbf{N}(z)\cdot\mathbf{U}(z)_{K,J_c} &=& \mathbf{I}_m,\end{array}$$

where $J=(1,...,n-m)$. In this case the matrix fraction $-\mathbf{U}(z)_{K_c,J}\cdot\mathbf{U}(z)^{-1}_{K,J}$ gives the right coprime proper rational function with $\mathbf{U}(z)_{K,J}$ in Popov form. □

Remark 3.7 *With $\vec{\beta}$ the $\vec{b}$-Kronecker indices of $\mathbf{A}(z)$, that is, the $\vec{b}$-degree of the MPB $\mathbf{U}(z)_{*,J}$ and $\vec{\beta}^*=\operatorname{cdeg}z^{-\vec{b}}\cdot\mathbf{U}(z)_{*,J}$, using the index convention of Definition 2.3, we have the following degree bounds for a minimal multiplier:*

$$\deg\mathbf{U}(z)_{k,j}\le\begin{cases}\min(\vec{\beta}_k,\vec{b}_k+\vec{\beta}^*_j), & k\in K,\ j\in J,\\ \vec{b}_k+\vec{\beta}^*_j, & k\in K_c, j\in J,\\ \vec{\beta}_k-1, & k\in K,\ j\in J_c\end{cases}.\tag{17}$$

The two first bounds are from the fact that $\mathbf{U}(z)_{K,J}$ is in $\vec{b}$-Popov form and from the definition of $\vec{\beta}^*$. The last one is deduced from the reduction identity (13) and has been given in [35, Theorem 2] in the case $\vec{a}=\vec{b}=\vec{0}$. □

Example 3.8 *Consider again the matrix* $\mathbf{A}(z)$ *of Example 3.2. In this case the minimal multiplier for the Popov form satisfies*

$$\mathbf{A}(z)\cdot\mathbf{U}^0(z)=\mathbf{A}(z)\cdot\begin{bmatrix}-1 & -1 & 0 & 0\\ \underline{z}^2-7 & -2\,z-7 & -z-2 & z+3\\ -z+3 & 3 & 1 & -1\\ -1 & \underline{z} & 1 & -1\end{bmatrix}=\begin{bmatrix}0 & 0 & z & -1\\ 0 & 0 & 2 & z\end{bmatrix}=\mathbf{T}(z).$$

On the other hand the $(\vec{0},\vec{b})$*-minimal multiplier for* $\vec{b}=[0,-3,0,0]$ *is given by*

$$\mathbf{U}(z)=\begin{bmatrix}-\frac{2}{21}z+\frac{1}{7} & -z^2-2\,z & \frac{1}{7}z+\frac{2}{7} & -\frac{1}{21}z-\frac{3}{7}\\ \underline{1} & 0 & 0 & 0\\ \frac{2}{21}\,z-\frac{3}{7} & z^2-z & -\frac{1}{7}z+\frac{1}{7} & \frac{1}{21}z+\frac{2}{7}\\ \frac{2}{21}z^2-\frac{1}{3}-\frac{4}{21} & \underline{z}^3-9\,z-7 & -\frac{1}{7}z^2+\frac{9}{7} & \frac{1}{21}z^2+\frac{1}{3}z-\frac{23}{21}\end{bmatrix}.$$

The Kronecker and the $\vec{b}$*-Kronecker indices are* $(2,\ 1)$ *and* $(0,\ 3)$ *respectively, with pivot set* $K=(2,4)$ *in both cases. They impose strict degree bounds for the entries of* $\mathbf{U}(z)_{K,*}$*, here for the non-pivot entries in the second and the last rows.* □

4 Computing Popov Normal Forms

As mentioned in the introduction, the shifted Popov form of $\mathbf{A}(z)$ together with its minimal multiplier is going to be computed by determining a minimal polynomial basis for the kernel of $[\mathbf{A}(z),-\mathbf{I}_m]$ – see identity (1) – considered as a module over $\mathsf{K}[z]$. We first show in subsection 4.1 that the entire normal form problem may actually be stated as a special MPB computation. Subsection 4.2 will then be concerned by the algorithm itself using known techniques for computing MPB. Most of the results here are given in terms of unknown degree bounds for the entries of $\mathbf{U}(z)$. These bounds will be estimated in terms of the input parameters by Theorem 5.1 in the next section and leads to simplified interpretations of what follows when we will study the complexity in Section 6.

4.1 Computing Popov Forms as Minimal Polynomial Bases

The approach used for computing Popov Forms as Minimal Polynomial Bases has already been roughly described in [6, Section 5] for the case of full column rank matrices $\mathbf{A}(z)$. In this subsection we will give more details and extend our considerations to the case of general matrices $\mathbf{A}(z)$. Consider the stacked matrix polynomial

$$\mathbf{S}(z)=\begin{bmatrix}\mathbf{U}(z)\\ \mathbf{T}(z)\end{bmatrix}\in\mathsf{K}^{(m+n)\times n}. \tag{18}$$

Notice that the columns of $\mathbf{S}(z)$ are elements of the kernel of $[\mathbf{A}(z),-\mathbf{I}_m]$ if and only if $\mathbf{A}(z)\mathbf{U}(z)=\mathbf{T}(z)$. In addition, they form a basis of the kernel if and only if $\mathbf{U}(z)$ is unimodular. Conversely, it is not the case that any MPB of the kernel of $[\mathbf{A}(z),-\mathbf{I}_m]$ will give the $\mathbf{U}(z)$ and $\mathbf{T}(z)$ that is desired [7]. By generalizing the work of [7] with the introduction of shifts and the full rank case

in [6, Theorem 5.2], we show that we can recover the shifted normal form by imposing a certain degree structure on the MPB that is computed. This structure is itself one of a shifted normal form.

Theorem 4.1 (Popov forms via MPB) *Let $\mathbf{A}(z)$ be a matrix polynomial and $\vec{a}, \vec{b}$ multi-indices. We denote by $\mathbf{T}(z)$ the $\vec{a}$–Popov form of $\mathbf{A}(z)$ with pivot set I and $\vec{a}$-degree $\vec{\alpha}$ and by $\vec{\alpha}^*$ the column degree of $z^{-\vec{a}} \cdot \mathbf{T}(z)_{*,J_c}$. Let $\mathbf{U}(z)$ be the associated $(\vec{a}, \vec{b})$–minimal multiplier with pivot set K and $\vec{b}$-degree $\vec{\beta}$. Assume that $\vec{\tau}^*$ is some upper bound for the column degree of $z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c}$.*

Let $\mathcal{N}_0 = \max[\vec{\tau}^* - \vec{\alpha}^*]$. Then for any integer $N \geq \mathcal{N}_0$, the matrix polynomial $\mathbf{S}(z)$ of (18) with $\mathbf{T}(z)$ and $\mathbf{U}(z)$ as above is the unique MPB of the kernel of $[\mathbf{A}(z), -\mathbf{I}_m]$ in $\vec{n}(N)$-Popov form where $\vec{n}(N) = (\vec{b} + N \cdot \vec{e}, \vec{a})$. As a shifted MPB, its pivot set is (K, I) and is $\vec{n}(N)$-degree is $\vec{\mu} = (\vec{\beta}, \vec{\alpha})$.

Proof: We may first show that $\mathbf{S}(z)$ obtained in (18) using the shifted Popov form of $\mathbf{A}(z)$ and its multiplier is indeed a $\vec{n}(N)$-Popov MPB for any $N \geq \mathcal{N}_0$, with the pivot sets and the shifted degrees as indicated above. By construction $\mathbf{S}(z)$ is a basis for the kernel, it is thus sufficient to prove the shifted Popov form properties of Definition 2.3. Since $\mathbf{T}(z)_{*,J} = \mathbf{0}$ and since $\mathbf{T}(z)_{*,J_c}$ and $\mathbf{U}(z)_{*,J}$ are in shifted Popov form and thus both satisfy (7), we have

$$z^{-\vec{\mu}} \cdot \mathbf{S}(z)_{(K,I),*} = \begin{bmatrix} z^{-\vec{\beta}} \cdot \mathbf{U}(z)_{K,J} & z^{-\vec{\beta}} \cdot \mathbf{U}(z)_{K,J_c} \\ \mathbf{0} & z^{-\vec{\alpha}} \cdot \mathbf{T}(z)_{I,J_c} \end{bmatrix},$$

with $z^{-\vec{\beta}} \cdot \mathbf{U}(z)_{K,J} = \mathbf{I}_{n-r} + \mathcal{O}(z^{-1})$ and $z^{-\vec{\alpha}} \cdot \mathbf{T}(z)_{I,J_c} = \mathbf{I}_r + \mathcal{O}(z^{-1})$. By the division (13) we also have

$$z^{-\vec{\beta}} \cdot \mathbf{U}(z)_{K,J_c} = z^{-\vec{\beta}} \cdot \mathbf{U}(z)_{K,J} \cdot \left[[\mathbf{U}(z)_{K,J}]^{-1} \cdot \mathbf{U}(z)_{K,J_c}\right] = \mathcal{O}(z^{-1}).$$

Thus $z^{-\vec{\mu}} \cdot \mathbf{S}(z)_{(K,I),*} = \mathbf{I}_n + \mathcal{O}(z^{-1})$ as required for the row degrees of $\mathbf{S}(z)$. For the column degrees we have

$$z^{-\vec{n}(N)} \cdot \mathbf{S}(z) \cdot z^{-\vec{\mu}+\vec{n}(N)_{(K,I)}} = \begin{bmatrix} z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J} \cdot z^{-\vec{\beta}+\vec{b}_K} & z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c} \cdot z^{-\vec{\alpha}^*-N\cdot\vec{e}} \\ \mathbf{0} & z^{-\vec{a}} \cdot \mathbf{T}(z)_{*,J_c} \cdot z^{-\vec{\alpha}^*} \end{bmatrix},$$

with

$$z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c} \cdot z^{-\vec{\alpha}^*-N\cdot\vec{e}} = \left[z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c} \cdot z^{-\vec{\tau}^*}\right] \cdot z^{\vec{\tau}^*-\vec{\alpha}^*-N\cdot\vec{e}} = \mathcal{O}(1)$$

by the definition of τ^* and the assumption that $N \geq \mathcal{N}_0$. Since, in addition, $U(z)_{*,J}$ and $T(z)_{*,J_c}$ both satisfy the column degree property (6), $\mathbf{S}(z)$ itself satisfies the property and is therefore in shifted Popov form as required for the first part of the proof. Since two $\vec{n}(N)$-Popov MPB – as bases of the $\mathsf{K}[z]$-module $\ker[\mathbf{A}(z), -\mathbf{I}_m]$ – must be equivalent, the uniqueness follows from the uniqueness of shifted Popov forms in Theorem 2.7. □

The theorem states that if the shift is large enough (when compared to the column degrees of $\mathbf{U}(z)$), then preference is given to the last m rows of $\mathbf{S}(z)$. These rows are forced to be in shifted Popov form and so lead to the unique $\mathbf{T}(z)$ in the MPB. As noticed previously, in order to exploit

the statement and in particular, in order to obtain an explicit value for $\mathcal{N}_0$ (which is needed for the algorithm), we rely on an upper bound $\vec{\tau}^*$ for the column degree of $z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c}$. In the case of square nonsingular $\mathbf{A}(z)$ ($r = m = n$ and $K_c = J_c = (1, ..., n)$) or more generally full column rank matrices $\mathbf{A}(z)$ ($r = n$, $K_c = (1, ..., n) = J_c = (1, ..., n)$), such bounds have been obtained in [6, Theorem 4.1]. The case of general matrices $\mathbf{A}(z)$ is considerably more involved and will be discussed in Section 5.

4.2 Computing Minimal Polynomial Bases as Matrix Rational Approximants

Several algorithms exist for computing Minimal Polynomial Bases (in normal form) of matrix polynomial kernels. Most of them are based on matrix pencil normal forms – see for example [7, 8] and the references therein. Another approach could be to compute a nullspace basis, for example as in [30, Chap 5], and from there compute a minimal basis by column reduction and normalization. Here, in order to take shifts into account and in order to do fraction-free computation, we will follow the idea of [27, Chap. 4] as applied by [6] in the full column rank case. We will use Hermite-Padé approximation by adapting the methods FPHPS and SPHPS of [2] (which have been developed for fixed cost arithmetic) and their fraction-free generalization FFFG [5].

A key point of the latter algorithms is the notion of order (see *e.g.* [3] or [5, Lemma 2.8]): given some $m \times s$ matrix polynomial $\mathbf{F}(z)$ and a multi-index $\vec{\sigma}$ of length m, we say that a vector $\mathbf{Q}(z) \in \mathsf{K}[z]^s$ has order $\vec{\sigma}$ if $z^{-\vec{\sigma}} \cdot \mathbf{F}(z) \cdot \mathbf{Q}(z) = \mathcal{O}(1)_{z \to 0}$. If we compared this to (1) and keeping the notation of (18), we see that we can take $\mathbf{F}(z) = [\mathbf{A}(z), -\mathbf{I}_m]$ and consider that the columns of $\mathbf{S}(z)$ have order infinity. Based on this remark, Theorem 4.2 below will state that the columns of $\mathbf{S}(z)$ can be computed as order vectors for $\vec{\sigma}$ large enough with respect to degree bounds on $\mathbf{U}(z)$ and $\mathbf{T}(z)$. In what follows we will always take $\mathbf{F}(z) = [\mathbf{A}(z), -\mathbf{I}_m]$ though many of the properties mentioned are also true more generally for a matrix of formal power series at zero.

From Theorem 4.2 we obtain the required order σ for an approximation problem. We then give the main algorithm for computation up to such an order (and hence for computing our minimal multiplier and normal form). The set of all polynomials of order $\vec{\sigma}$ forms a $\mathsf{K}[z]$-submodule of the module $\mathsf{K}[z]^s$. This module contains all the elements of the kernel of $\mathbf{F}(z)$. The basic idea of our method is to construct successively for increasing order vectors some $s \times s$ matrix polynomial $\mathbf{M}(z)$ with columns forming a basis for the module of order $\vec{\sigma}$. From a certain order on, n of the columns of $\mathbf{M}(z)$ will also form a basis of the kernel of $\mathbf{F}(z)$. We will check these additional properties by counting zero columns in the *residual polynomial*

$$\mathbf{R}(z) = z^{-\vec{\sigma}} \cdot \mathbf{F}(z) \cdot \mathbf{M}(z). \tag{19}$$

According to Theorem 4.1, we require not only a basis of the kernel of $\mathbf{F}(z)$ but a $\vec{n}(N)$-Popov MPB. Therefore we also need to impose degree contraints on the matrix $\mathbf{M}(z)$. These constraints will actually be ensured by algorithm FFFG where all intermediate bases are *Mahler systems* of type $\vec{n}$, that is, order bases which are in $\vec{n}$-Popov form for any input multi-index $\vec{n}$ [5, Theorem 7.2 and Theorem 7.3]. Before stating explicitly the algorithm, let us thus show that order bases lead to shifted Popov forms and multipliers for sufficiently large order $\vec{\sigma}$:

Theorem 4.2 (MPB via FFFG) *Let $\mathbf{A}(z)$ be an $m \times n$ matrix polynomial. As in Theorem 4.1 set $\mathcal{N}_0 = \max[\vec{\tau}^* - \vec{\alpha}^*]$, $\vec{n}(N) = (\vec{b} + N \cdot \vec{e}, \vec{a})$ with $\vec{a}, \vec{b}$ the input shifts, and $\vec{\alpha}, \vec{\alpha}^*, \vec{\beta}, \beta^*, \vec{\tau}^*$ the corresponding degrees in $\mathbf{T}(z)$ and $\mathbf{U}(z)$. We assume in addition that the multi-index $\vec{\gamma}^*$ is an upper bound for the row degree of $\mathbf{A}(z) \cdot z^{\vec{b}}$, that is, $z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}} = \mathcal{O}(1)$.*

If $\mathbf{F}(z) = [\mathbf{A}(z), -\mathbf{I}_m]$ *and* $\mathbf{M}(z)$ *is a Mahler system of type* $\vec{n}(N)$ *with order vector* $\vec{\sigma}$, $\vec{n}(N)$-*degree* $\vec{\mu}$ *and residual polynomial* $\mathbf{R}(z)$ *as in (19), then for any integer*

$$N \geq \max\{\mathcal{N}_0, \max[\vec{a} - \vec{\gamma}^*]\}, \tag{20}$$

$\mathbf{M}(z)$ *and* $\mathbf{R}(z)$ *satisfy:*

(a) *As soon as* $\vec{\sigma}$ *is greater than*

$$\vec{\sigma}_0 = \vec{\gamma}^* + \Big(1 + \max\{N + \max[\vec{\alpha}^*], \max[\vec{\beta}^*]\}\Big)\vec{e} \tag{21}$$

there exists a list $L = (1, ..., s)$ *of* n *elements with* $\mathbf{R}(z)_{*,L} = \mathbf{0}$.

(b) *If there exists a list* $L \subset (1, ..., s)$ *of* n *elements with* $\mathbf{R}(z)_{*,L} = \mathbf{0}$, *then the last* m *rows of* $\mathbf{M}(z)_{*,L}$ *give the* $\vec{a}$-*Popov form of* $\mathbf{A}(z)$, *the first* n *rows of* $\mathbf{M}(z)_{*,L}$ *give the corresponding* $(\vec{a}, \vec{b})$-*minimal multiplier, and* $\vec{\mu}_L = (\vec{\beta}, \vec{\alpha})$.

Proof: Let us start by proving (b). Suppose that $\mathbf{R}(z)_{*,L} = \mathbf{0}$ with $\#L = n$. Then, by (19), the columns of $\mathbf{M}(z)_{*,L}$ are elements of the kernel of $\mathbf{F}(z)$. A basis of the kernel is given by the matrix $\mathbf{S}(z)$ build up with the shifted Popov form and the minimal multiplier of $\mathbf{A}(z)$ as described in (18). Hence there exists a unique $n \times n$ matrix polynomial $\mathbf{Q}(z)$ with

$$\mathbf{M}(z)_{*,L} = \mathbf{S}(z) \cdot \mathbf{Q}(z).$$

On the other hand, the columns of $\mathbf{S}(z)$ are of order $\vec{\sigma}$ and thus may be uniquely represented as polynomial linear combinations of the columns of $\mathbf{M}(z)$. Thus there exists a matrix polynomial $\mathbf{P}(z)$ with

$$\mathbf{S}(z) = \mathbf{M}(z) \cdot \mathbf{P}(z). \tag{22}$$

Combining these two identities gives $\mathbf{M}(z)_{*,L} = \mathbf{M}(z) \cdot \mathbf{P}(z) \cdot \mathbf{Q}(z)$. Since the columns of $\mathbf{M}(z)$ are linearly independent, we find that $\mathbf{P}(z)_{L_c,*} = \mathbf{0}$ and $\mathbf{P}(z)_{L,*} = \mathbf{Q}(z)^{-1}$ is unimodular. Hence $\mathbf{S}(z)$ and $\mathbf{M}(z)_{*,L}$ are column equivalent matrices. They are both in $\vec{n}(N)$-Popov form: $\mathbf{S}(z)$ as seen in the proof of Theorem 4.1 since $N \geq \mathcal{N}_0$ and $\mathbf{M}(z)_{*,L}$ as a subset of columns of the Mahler system $\mathbf{M}(z)$ which is in $\vec{n}(N)$-Popov form. It follows that $\mathbf{S}(z) = \mathbf{M}(z)_{*,L}$ by uniqueness of the normal form, showing part (b).

A proof of (a) is slightly more involved. Let $\vec{\delta}^*$ be defined by

$$\vec{\delta}^* = \mathrm{cdeg}\,(z^{-\vec{n}(N)} \cdot \mathbf{S}(z)) = (\vec{\beta}^* - N\vec{e}, \vec{\alpha}^*).$$

Using the Predictable-Degree Property stated in Lemma 3.4 we may deduce for the unique $\mathbf{P}(z)$ satisfying (22) that

$$\deg \mathbf{P}(z)_{i,j} \leq \vec{\delta}^*_j - [\vec{\mu} - \vec{n}(N)]_i, \;\; 1 \leq i \leq m+n, 1 \leq j \leq n. \tag{23}$$

When the order $\vec{\sigma}$ is increased, the degrees in $\mathbf{M}(z)$ also are increased. We show that this forces some rows of $\mathbf{P}(z)$ and consequently some columns of the residual to be null. We proceed by showing that otherwise, inequality (23) would be impossible. From the definition (19) of the residual

$$\mathbf{R}(z) = z^{-\vec{\sigma}} \cdot \mathbf{F}(z) \cdot \mathbf{M}(z)$$

we can write:

$$\begin{aligned} z^{\vec{\sigma}-\vec{\gamma}^*} \cdot \mathbf{R}(z) \cdot z^{\vec{n}(N)-\vec{\mu}} &= z^{-\vec{\gamma}^*} \cdot \mathbf{F}(z) \cdot \mathbf{M}(z) \cdot z^{\vec{n}(N)-\vec{\mu}} \\ &= z^{-\vec{\gamma}^*} \cdot [\mathbf{A}(z), -\mathbf{I}_m] \cdot \mathbf{M}(z) \cdot z^{\vec{n}(N)-\vec{\mu}} \\ &= [z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}+N\cdot\vec{e}}, -z^{-\vec{\gamma}^*+\vec{a}}] \cdot [z^{-\vec{n}(N)} \cdot \mathbf{M}(z) \cdot z^{\vec{n}(N)-\vec{\mu}}]. \end{aligned}$$

Since $N \geq \max[\vec{a} - \vec{\gamma}^*]$ and since $\vec{\gamma}^*$ is defined by $z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}} = \mathcal{O}(1)$ we will have that $[z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}+N\cdot\vec{e}}, -z^{-\vec{\gamma}^*+\vec{a}}] = \mathcal{O}(z^N)$. Since in addition $\vec{\mu}$, the shifted degree of $\mathbf{M}(z)$, gives $[z^{-\vec{n}(N)} \cdot \mathbf{M}(z) \cdot z^{\vec{n}(N)-\vec{\mu}}] = \mathcal{O}(1)$, the residual satisfies:

$$z^{\vec{\sigma}-\vec{\gamma}^*-N\vec{e}} \cdot \mathbf{R}(z) \cdot z^{\vec{n}(N)-\vec{\mu}} = \mathcal{O}(z^0)_{z\to\infty}. \tag{24}$$

From [3, Lemma 2.8] it is known that, with $\mathbf{F}(0)$, also $\mathbf{R}(0)$ and thus $\mathbf{R}(z)$ all have full row rank m for all order $\vec{\sigma}$. Therefore we may find a nonsingular square submatrix $\mathbf{R}(z)_{*,L_c}$ and more precisely some bijective map $\rho : \{1, ..., m\} \to L_c$ such that

$$\mathbf{R}(z)_{j,\rho(j)} \neq 0, \quad j = 1, 2, ..., m.$$

Together with (24) this leads to:

$$[\vec{\sigma} - \vec{\gamma}^* - N\vec{e}]_j + [\vec{n}(N) - \vec{\mu}]_{\rho(j)} \leq 0, \quad j = 1, 2, ..., m.$$

Replacing $\vec{\sigma}$ by its lower bound

$$\begin{aligned} \vec{\sigma}_0 &= \vec{\gamma}^* + \Big(1 + \max\big\{N + \max[\vec{\alpha}^*], \max[\vec{\beta}^*]\big\}\Big)\vec{e} \\ &= \vec{\gamma}^* + \Big(1 + N + \max[\vec{\delta}^*]\Big)\vec{e} \end{aligned}$$

we get

$$1 + \max[\vec{\delta}^*] + [\vec{n}(N) - \vec{\mu}]_i \leq 0, \quad i \in L_c$$

or

$$\vec{\delta}^*_j - [\vec{\mu} - \vec{n}(N)]_i \leq -1, \ i \in L_c, 1 \leq j \leq n. \tag{25}$$

Comparing (25) and (23), we may conclude that

$$\mathbf{P}(z)_{L_c,*} = \mathbf{0}, \quad \mathbf{S}(z) = \mathbf{M}(z)_{*,L} \cdot \mathbf{P}(z)_{L,*}. \tag{26}$$

Multiplying this last identity on the left by $[\mathbf{A}(z), -\mathbf{I}_m]$ leads to

$$\mathbf{0} = z^{-\vec{\sigma}} \cdot [\mathbf{A}(z), -\mathbf{I}_m] \cdot \mathbf{S}(z) = \mathbf{R}(z)_{*,L} \cdot \mathbf{P}(z)_{L,*}.$$

On the other hand, $\mathbf{S}(z)$ is a polynomial basis and thus of full column rank. From (26) it follows that $\mathbf{P}(z)_{L,*}$ is invertible implying that the matrix polynomial $\mathbf{R}(z)_{*,L}$ is identically zero, as claimed by (a). □

Algorithm SPF given on page 18 computes the shifted normal form and the associated minimal multiplier. As described previously, they are obtained in a stack matrix $\mathbf{S}(z)$ as a submatrix of an order basis $\mathbf{M}(z)$ once n columns of the residual matrix have been zeroed (see the *stopping criterion* $\#L = n$). Concerning the invariants and the main iteration, the algorithm is essentially

Algorithm SPF – Shifted Popov Form via FFFG.

INPUT: A matrix polynomial $\mathbf{A}(z) \in \mathsf{D}[z]^{m \times n}$ of degree d, $s = m + n$,
a multi-index $\vec{a}$ of length m (default $\vec{a} = 0$),
a multi-index $\vec{b}$ of length n (default $\vec{b} = -\mathrm{cdeg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z))$.

INVARIANTS: For (increasing) order vectors $\vec{\sigma}$,
Mahler system $\mathbf{M}(z)$ of size $s \times s$ in $\vec{n}$–Popov form with shifted degree $\vec{\mu}$,
its columns form an order basis for $[\mathbf{A}(z), -\mathbf{I}_m]$ and order $\vec{\sigma}$,
g is the corresponding multigradient,
$\mathbf{R}(z)$ is the corresponding residual polynomial,
L is the set of indices of zero columns of the residual polynomial

INITIALIZATION: $\mathbf{M}(z) \leftarrow \mathbf{I}_s$, $g \leftarrow 1$, $\mathbf{R}(z) \leftarrow [\mathbf{A}(z), -\mathbf{I}_m]$, $L \leftarrow \{\ \}$, $\vec{\sigma} \leftarrow \vec{0}$
$N \leftarrow d \cdot \min\{m, n\} + \max[\vec{a}] - \min[\vec{b}]$, $\vec{n} \leftarrow (\vec{b} + N \cdot \vec{e}, \vec{a})$,
$\vec{\gamma}^* \leftarrow \mathrm{rdeg}\,(\mathbf{A}(z) \cdot z^{\vec{b}})$, $\vec{\mu} \leftarrow \vec{0}$

ITERATIVE STEP:
Find j such that $\vec{\gamma}^*_j - \vec{\sigma}_j = \max[\vec{\gamma}^* - \vec{\sigma}]$
Define for $\ell = 1, ..., s$: $r_\ell \leftarrow \mathbf{R}(0)_{j,\ell}$
Define set $\Lambda = \{\ell \in \{1, .., s\} : r_\ell \neq 0\}$ ($\neq \{\ \}$)
Define pivot $\pi = \min\Big\{\ell \in \Lambda : \vec{n}_\ell - \vec{\mu}_\ell = \max_{k \in \Lambda}\{\vec{n}_k - \vec{\mu}_k\}\Big\}$
Define leading coefficients $p_\ell \leftarrow \mathit{coefficient}(\mathbf{M}(z)_{\ell,\pi}, z^{\vec{\mu}_\ell - 1})$, $\ell \neq \pi$

Check stopping criterion:
add all indices $\ell \notin L \cup \Lambda$ with $\mathbf{R}_{*,\ell}(z) = \mathbf{0}$ to L
STOP ITERATION if $\#L = n$.

Increase order for $\ell = 1, .., m$, $\ell \neq \pi$:
$\mathbf{M}(z)_{*,\ell} \leftarrow [\mathbf{M}(z)_{*,\ell} \cdot r_\pi - \mathbf{M}(z)_{*,\pi} \cdot r_\ell]/g$
$\mathbf{R}(z)_{*,\ell} \leftarrow [\mathbf{R}(z)_{*,\ell} \cdot r_\pi - \mathbf{R}(z)_{*,\pi} \cdot r_\ell]/g$

Increase order for $\ell = \pi$ and adjust degree constraints:
$\mathbf{M}(z)_{*,\pi} \leftarrow [z \cdot \mathbf{M}(z)_{*,\pi} \cdot r_\pi - \sum_{\ell \neq \pi} \mathbf{M}(z)_{*,\ell} \cdot p_\ell]/g$
$\mathbf{R}(z)_{*,\pi} \leftarrow [z \cdot \mathbf{R}(z)_{*,\pi} \cdot r_\pi - \sum_{\ell \neq \pi} \mathbf{R}(z)_{*,\ell} \cdot p_\ell]/g$

Adjust residual in row j:
$\mathbf{R}(z)_{j,*} \leftarrow [\mathbf{R}(z)_{j,*}/z]$

Update multigradient: $g \leftarrow r_\pi$
Update order vector: $\vec{\sigma} \leftarrow \vec{\sigma} + \vec{e}_j$
Update shifted degree vector: $\vec{\mu} \leftarrow \vec{\mu} + \vec{e}_\pi$

FINAL STEP AND OUTPUT: If L is the increasing list $(\ell_1, ..., \ell_n)$:
Rank r of $\mathbf{A}(z)$: unique index with $\ell_{n-r} \leq n < \ell_{n-r+1}$ (with the convention $\ell_0 = 0$)
$g \cdot \mathbf{T}(z) = \mathbf{M}(z)_{(n+1,\ldots,n+m),L}$ being the $\vec{a}$–Popov form of $\mathbf{A}(z)$
$g \cdot \mathbf{U}(z) = \mathbf{M}(z)_{(1,\ldots,n),L}$ being the corresponding $(\vec{a}, \vec{b})$–minimal multiplier
Pivot sets: $K \leftarrow (\ell_1, ..., \ell_{n-r})$, $I \leftarrow (\ell_{n-r+1} - n, ..., \ell_n - n)$
Shifted degrees: $(\vec{\beta}, \vec{\alpha}) \leftarrow \vec{\mu}_L$.

the algorithm FFFG of [5, Section 7] for fraction-free order bases computation. A slight difference is in the computation of the residual polynomials which is made more explicit here. For a proof of correctness of FFFG the reader will refer to [5, Theorem 7.2]. Some further properties, in particular the link to an underlying system of linear equations has been investigated in [5, Theorem 7.3]. The complexity study will be given in Section 6.

For D an integral domain, Algorithm SPF takes as input an $m \times n$ matrix $\mathbf{A}(z)$ with entries of degrees less than d in $\mathsf{D}[z]$. Since FFFG is fraction-free, all the divisions are exact in D. The claim of Theorem 4.2 relies on two quantities: the input shift must satisfy $N \geq \max\{\mathcal{N}_0, \max[\vec{a} - \vec{\gamma}^*]\}$ and the order $\vec{\sigma}$ must be greater than $\vec{\sigma}_0$. For the choice of the input shift we use Corollary 5.9 of Section 5. The corresponding worst-case bound (see the initialization of N) works for any $\mathbf{A}(z)$, $\vec{a}$ and $\vec{b}$. Finer bounds could be derived from the forthcoming Theorem 5.1 if additional properties are available on the matrix and the shifts. Default shift values may be proposed (see the input data): for example, the choices $\vec{a} = 0$ and $\vec{b} = -\mathrm{cdeg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z))$ lead to the simplification $\vec{\gamma}^* = \vec{a} = 0$. Concerning the order, an *a priori* bound for $\vec{\sigma}_0$ is actually not needed for the algorithm and will be used only for the complexity estimates. The stopping criterion ensures that the algorithm will automatically stop when a sufficient order is reached. At this point, the columns $(l_1, \ldots, l_n)$ are null. In $\mathbf{S}(z)$, the pivot indices corresponding to K are lower than n and those corresponding to I are greater than n. The same is thus true in $\mathbf{M}(z)$ which is itself, by construction (Algorithm FFFG), in shifted Popov form. Identity (7) implies that the pivot entries are diagonal in $\mathbf{M}(z)$, thus K and I can also be found from the column indices. The algorithm thus finds the rank of $\mathbf{A}(z)$ ($\#I$) and I from the l_i's greater than n; it finds the set K from the l_i's smaller than n (see the final step and the output data).

Notice that as usual for fraction-free algorithms, the algorithm only outputs g times the "correct" answers $\mathbf{T}(z)$ and $\mathbf{U}(z)$ for a scalar $g \in \mathsf{D}$. Here, correct means with $\mathbf{T}(z)$ having a normalized leading matrix following (7). Indeed, generically the coefficients of the normal form are not elements of the initial integral domain but only of its quotient field and g is a multiple of the denominators. In [5, Definition 4.1], the authors give a characterization of the *multigradient* g as a determinant of a striped Krylov matrix (here, a striped Sylvester matrix) which seems to indicate that in general the factor g cannot made be smaller (though of course for particular examples a better choice of g might be suitable).

Example 4.3 *If we keep the matrix* $\mathbf{A}(z) \in \mathbf{Z}[z]^{m \times n}$ *of Examples 3.2 and 3.8, Algorithm* SPF *with shifts* $\vec{a} = [0, 0]$ *and* $\vec{b} = [0, -3, 0, 0]$ *constructs*

$$\mathbf{S}(z) = g \cdot \begin{bmatrix} \mathbf{U}(z) \\ \mathbf{T}(z) \end{bmatrix} = \begin{bmatrix} -2\,z+3 & -21z^2-42\,z & 3\,z+6 & -z-9 \\ \underline{21} & 0 & 0 & 0 \\ 2\,z-9 & 21\,z^2-21\,z & -3\,z+3 & z+6 \\ 2z^2-7z-4 & \underline{21}\,z^3-189z-147 & -3\,z^2+27 & z^2+7\,z-23 \\ 0 & 0 & \underline{21}\,z & -21 \\ 0 & 0 & 42 & \underline{21}\,z \end{bmatrix}.$$

Notice that $\mathbf{S}(z)$ *is not in* $(\vec{b}, \vec{a})$*-Popov form since the leading degrees in the two last columns are not in the last two rows* – i.e. *the part corresponding to* $\mathbf{T}(z)$. *Here it would have been sufficient to consider* $N \geq 1$*: with* $\vec{n}(1) = (\vec{b} + \vec{e}, \vec{a})$, $\mathbf{S}(z)$ *is in* $\vec{n}(1)$*-Popov form with pivot indices* $L = (l_1, l_2, l_3, l_4) = (2, 4, 5, 6)$. *From there it is seen that* $K = (2, 4)$ *and* $I = (5 - n, 6 - n) = (1, 2)$.

If we compare with Example 3.8, we see that SPF *has computed the multiple* $g \cdot \mathbf{T}(z) = 21\mathbf{T}(z)$ *of the normal form. However, since this factor appears naturally in the denominators of* $\mathbf{U}(z)$ *in the normalized case, we may consider that no spurious factor has been introduced.* □

The fraction–free computation of shifted normal forms via FFFG has been implemented using the computer algebra system MAPLE and can be obtained from the authors.

5 Degree Bounds for Minimal Multipliers

The algorithm described in the previous section requires that we know in advance degree bounds for the $(\vec{a}, \vec{b})$–minimal unimodular multiplier $\mathbf{U}(z)$ needed for transforming a given matrix polynomial $\mathbf{A}(z)$ into $\vec{a}$–Popov form $\mathbf{T}(z)$. The aim of this section is to give such degree bounds, generalizing those already determined in [6] for the case of full-column rank matrices. Theorem 5.1 gives estimates in terms of the degrees in $\mathbf{A}(z)$, of the shifts and of the invariants $\vec{\alpha}$ and $\vec{\beta}$ of the problem. These invariants are generally unknown and the estimates are simply worst-case bounds. Nevertheless Example 5.6 will show that there are cases where the bounds are tight.

We do our degree bounds in two steps. In part (a) of Theorem 5.1, we first formulate our bounds in terms of the input parameters $\mathbf{A}(z), \vec{a}, \vec{b}$, and of the invariants to our problem, these being $\mathbf{T}(z)_{*,J_c}, \vec{\alpha}, \vec{\beta}\,(\vec{\alpha}^*, \vec{\beta}^*)$ along with the pivot sets I, K. Note that the degree bounds for the MPB part $\mathbf{U}(z)_{*,J}$ and for $\mathbf{U}(z)_{K,J_c}$ follow immediately from (17) in Remark 3.7 with only degree bounds for $\mathbf{U}(z)_{K_c,J_c}$ remaining unknown at this stage. The aim of the second step, in parts (c), (d) and (e), is to estimate the invariants in terms of the input parameters only.

Part (b) is a precise characterization of the pivot sets I and K. It generalizes the following well known relations for full column rank matrices. In the case of square matrix polynomials, if $\mathbf{A}(z)$ is column reduced and if $\mathbf{T}(z)$ is its Popov normal form with column degree vector $\vec{\alpha}^*$ then we have the invariant

$$|\vec{\alpha}^*| = \deg\det \mathbf{T}(z) = \deg\det \mathbf{A}(z). \tag{27}$$

It is also not difficult to see how one obtains degree bounds for the multiplier in this case. Indeed, one writes

$$\mathbf{A}(z) = \mathbf{T}(z) \cdot \mathbf{V}(z)$$

where $\mathbf{V}(z) = \mathbf{U}(z)^{-1}$ and uses the Predictable Degree Property to obtain bounds for $\mathbf{V}(z)$ in terms of the column degrees of $\mathbf{A}(z)$ and $\mathbf{T}(z)$. Bounds for $\mathbf{U}(z)$ are then determined by making use of Cramers rule for the adjoint of $\mathbf{V}(z)$.

In the rectangular case a relation corresponding to (27) and degree bounds for the multiplier requires some classical tools from linear system theory. For a matrix polynomial $\mathbf{A}(z)$ of rank r, we define the *Minor degree* – denoted by Minor-deg $\mathbf{A}(z)$ – as the maximum of the degrees of the determinants of $r \times r$ submatrices of $\mathbf{A}(z)$ (see [18, Eq. (34), p. 454]). For a matrix polynomial, this degree is the polar content at infinity and is thus equal to the sum of the polar contents at all poles which is the *MacMillan degree* [19] – denoted by MM-deg $\mathbf{A}(z)$.

If $\mathbf{A}(z)$ has full column rank then it is well-known that

$$\text{Minor-deg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z)) \leq |\text{cdeg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z))|,$$

with equality if and only if $\mathbf{A}(z)$ is $\vec{a}$–column reduced [18, § 6.3.2, p. 384]. In the latter case, denoting by $I = (i_1, ..., i_r)$ the pivot set of the leading coefficient matrix, and $\vec{\alpha}^* = \text{cdeg}\,(z^{-\vec{a}}\mathbf{A}(z))$,

we also have that

$$\text{Minor-deg}\,(z^{-\vec{a}}\mathbf{A}(z)) = \deg\det\,(z^{-\vec{a}_I}\mathbf{A}(z)_{I,*}) = |\vec{\alpha}^*| > \deg\det\,(z^{-\vec{a}_{I'}}\mathbf{A}(z)_{I',*}) \tag{28}$$

for any list of the form $I' = (i'_1, .., i'_\ell, i'_{\ell+1}, .., i'_r)$ for some ℓ, where $i'_\ell > i_\ell$ and $i'_k \geq i_k$.

Theorem 5.1 (Degree bounds for Multiplier) *Let $\mathbf{A}(z) \cdot \mathbf{U}(z) = \mathbf{T}(z)$ with $\mathbf{T}(z)_{*,J} = \mathbf{0}$ and $\mathbf{T}(z)_{*,J_c}$ $\vec{a}$–column reduced with pivot set I and $\vec{a}$-degree $\vec{\alpha}$. Assume that $\mathbf{U}(z)$ is unimodular with $\mathbf{U}(z)_{*,J}$ $\vec{b}$–column reduced with pivot set K, $\vec{b}$-degree $\vec{\beta}$, and satisfying $[\mathbf{U}(z)_{K,J}]^{-1} \cdot \mathbf{U}(z)_{K,J_c} = \mathcal{O}(z^{-1})_{z\to\infty}$. Set $\vec{\gamma}^* = \text{rdeg}\,(\mathbf{A}(z) \cdot z^{\vec{b}})$, $\vec{\alpha}^* = \text{cdeg}\,(z^{-\vec{a}}\mathbf{T}(z)_{*,J_c}) = \vec{\alpha} - \vec{a}_I$, $\vec{\beta}^* = \text{cdeg}\,(z^{-\vec{b}}\mathbf{U}(z)_{*,J}) = \vec{\beta} - \vec{b}_K$ and define $\Delta^{\vec{a},\vec{b}} = |\vec{\gamma}^*_I| - |\vec{b}_{K_c}| - |\vec{\alpha}| - |\vec{\beta}|$. Then the following are true.*

(a) (Degree bounds for multiplier) *For $j \in J_c$, $k \in K_c$ we have the degree bounds*

$$\deg z^{-\vec{b}_k} \cdot \mathbf{U}(z)_{k,j} \leq \max(\vec{\alpha}^*_j + \max[\vec{a}_I - \vec{\gamma}^*_I] + \Delta^{\vec{a},\vec{b}},\quad \max[\vec{\beta}^*] - 1). \tag{29}$$

(b) (Pivot sets and minor degree) *The set K_c consists of the smallest column indices such that*

$$\text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}}) = \text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z)_{*,K_c} \cdot z^{\vec{b}_{K_c}}).$$

The set I consists of the largest row indices such that

$$\text{Minor-deg}\,(z^{-\vec{a}} \cdot \mathbf{A}(z) \cdot z^{\vec{b}}) = \deg\det\,(z^{-\vec{a}_I} \cdot \mathbf{A}(z)_{I,K_c} \cdot z^{\vec{b}_{K_c}}),$$

with $\deg\det\,(\mathbf{A}(z)_{I,K_c}) = |\vec{\alpha}| + |\vec{\beta}|$.

(c) (Bounds for $\Delta^{\vec{a},\vec{b}}$) *We have*

$$0 \leq -\text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z)_{*,K_c} \cdot z^{\vec{b}_{K_c}}) \leq \Delta^{\vec{a},\vec{b}} \leq |\vec{\gamma}^*_I| - |\vec{b}_{K_c}|,$$

and $\Delta^{\vec{a},\vec{b}} = 0$ if and only if $\mathbf{A}(z)_{*,K_c}$ is $\vec{\gamma}^*$–column reduced, with pivot set I and $\vec{b}_{K_c} = \text{cdeg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z)_{*,K_c})$.

(d) (Bounds for $\vec{\alpha}$) *$\mathbf{T}(z)_{I,J_c}$ is a left maximal factor of $\mathbf{A}(z)_{I,*}$, and thus $|\vec{\alpha}|$ equals the degree of the determinant of a left maximal factor of $\mathbf{A}(z)_{I,*}$. Furthermore we also have, $\vec{\alpha}^* \leq \max[\vec{\gamma}^*_I - \vec{a}_I] \cdot \vec{e} - \text{perm}\,[\vec{b}_{K_c}]$ and by definition, $\vec{\alpha} \geq \vec{0}$.*

(e) (Bounds for the $\vec{b}$-Kronecker indices $\vec{\beta}$) *We have $|\vec{\beta}| \leq |\vec{\gamma}^*_I| - |\vec{b}_{K_c}| - |\vec{\alpha}|$ and by definition, $\vec{\beta} \geq \vec{0}$.*

We remark that the statement of the theorem simplifies for matrices having full row rank (where $I = (1, ..., m)$), and also for matrices having full column rank. In the latter case, $K_c = J_c = (1, ..., n)$, $J = K = \emptyset$ and all terms involving $\vec{\beta}$ have to be dropped (c.f. [6]). The quantity $\Delta^{\vec{a},\vec{b}}$ can be thought of as a measure of the distance that our input matrix is from being $\vec{\gamma}^*$–column reduced (see the discussion in [6, Section 4]).

We now give several examples to illustrate the theorem. Its proof will be given subsequently.

Example 5.2 *Let us use Theorem 5.1 to study the Kronecker indices of*

$$\mathbf{A}(z) = \begin{bmatrix} z^5 - 2\,z^2 - 3 & z^2 + 2\,z - 5 & z^3 - z - 2 & z^2 + z - 4 \\ z^6 + z^3 - z^2 + z - 1 & 2\,z^2 + z - 1 & z^4 - z^2 + z - 1 & z^2 + z - 1 \end{bmatrix}.$$

With the shift $\vec{a} = [0, -3]$ one computes the Hermite form of $\mathbf{A}(z)$ with a determinantal degree satisfying $|\vec{\alpha}| = 4$. The degrees of the entries of the associated minimal multipliers with respective shifts $\vec{b} = [0, 0, 0, 0]$ and $\vec{b}' = [-5, 0, -3, 0]$ are:

$$\deg \mathbf{U}(z) = \begin{bmatrix} 0 & -\infty & -\infty & -\infty \\ 0 & 2 & 2 & 1 \\ \underline{2} & 0 & 0 & -\infty \\ 0 & \underline{2} & 1 & 1 \end{bmatrix}, \ \deg \mathbf{U}'(z) = \begin{bmatrix} \underline{0} & -\infty & -\infty & -\infty \\ 4 & 2 & 2 & 1 \\ -\infty & \underline{0} & -\infty & -\infty \\ 4 & 2 & 2 & 1 \end{bmatrix}.$$

For $\vec{b} = [0, 0, 0, 0]$, the pivot set K is $(3, 4)$ with $\vec{\beta} = [2, 2]$. One can see that the choice $\vec{b}' = [-5, 0, -3, 0]$ forces another pivot set $K = (1, 3)$ and leads to $\vec{\beta} = [0, 0]$. In both cases one can check statement (b) of Theorem 5.1:

$$\begin{cases} \vec{b} & = [0, 0, 0, 0], \quad K_c = (1, 2), \quad \deg\det \mathbf{A}(z)_{*,(1,2)} = 4 + 4 = 8, \\ \vec{b}' & = [-5, 0, -3, 0], \quad K_c = (2, 4), \quad \deg\det \mathbf{A}(z)_{*,(2,4)} = 4 + 0 = 4. \end{cases}$$

This makes explicit the relation between the Kronecker indices and the selection of a particular submatrix of $\mathbf{A}(z)$. When $\vec{b} = [0, 0, 0, 0]$, the bound (e) on the Kronecker indices is $|\vec{\beta}| \leq (5 + 6) - (0 + 0) - 4 = 7$ which is pessimistic compared to the actual value $|\vec{\beta}| = 4$. The value $\Delta^{\vec{a},\vec{b}} = 7 - 4 = 3$ shows that the prediction is limited by the structure of $\mathbf{A}(z)_{*,(1,2)}$. When $\vec{b} = [-5, 0, -3, 0]$ the estimation $|\vec{\beta}| \leq (2 + 2) - (0 + 0) - 4 = 0$ gives exactly $|\vec{\beta}| = 0$: the good predictions takes advantage of the fact that $\mathbf{A}(z)_{*,(2,4)}$ is column reduced. □

Example 5.3 *As in Example 3.6 assume that we have a left coprime proper matrix rational function $\mathbf{R}(z) = \mathbf{D}(z)^{-1} \cdot \mathbf{N}(z)$ with $\mathbf{D}(z)$ square of size $p \times p$ in Popov form and $\mathbf{N}(z)$ of size $p \times q$. Let $\mathbf{U}(z)$ be the unique minimal multiplier giving the Popov form:*

$$[\mathbf{D}(z), \mathbf{N}(z)] \cdot \mathbf{U}(z) = [\mathbf{0}, \mathbf{I}_m],$$

Then Theorem 5.1 gives the pivot sets and degree bounds on $\mathbf{U}(z)$ with $\vec{a} = \vec{b} = \vec{0}$. In this case $\vec{\gamma}^*$ is the row degree of $\mathbf{D}(z)$ ($\mathbf{R}(z)$ is proper), $\Delta^{\vec{a},\vec{b}} = \vec{0}$ since $\mathbf{D}(z)$ is $\vec{\gamma}^*$–column reduced ($\mathbf{D}(z)$ is in Popov form) and $\vec{\alpha} = \vec{0}$. The degree bound in part (a) then gives

$$\deg \mathbf{U}(z)_{k,j} \leq \max[\vec{\beta}] - 1 \text{ for } k \in K_c, j \in J_c.$$

This in turn is bounded by $|\vec{\beta}| - 1 \leq |\vec{\gamma}^*| - 1$ by part (e). □

In the following example, we illustrate the use of Theorem 5.1 to give bounds for the size of cofactors for the Gcd of $n \geq 2$ scalar polynomials. To our knowledge, such a bound has not been given before.

Example 5.4 (Gcd of several scalar polynomials) *Let $\mathbf{A}(z)$ be a row vector of n polynomials $[a_1(z), a_2(z), \ldots a_n(z)] \in \mathsf{K}[z]^n$ with degrees $\vec{d} = [d_1, d_2, \ldots d_n]$, where without loss of generality we assume that $d_1 = \min_j d_j$ and $d_n = \max_j d_j$. We are interested in degree bounds for "small" multipliers $u_k(z)$ in the diophantine equation*

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

We will derive these degree bounds from the last column of a shifted minimal multiplier according to Theorem 5.1(a),(b),(c). In fact we show that there exist multipliers verifying

$$\deg u_1(z) \leq d_n - \delta - 1, \quad \sum_{\substack{k=2 \\ u_k \neq 0}}^{n} (1 + \deg u_k(z)) \leq d_1 - \delta, \tag{30}$$

where $\delta = \deg Gcd\,(a_1(z), \ldots, a_n(z))$. Notice that our bound includes the classical one for $n = 2$ (cf. [14]). Also, a straight forward generalization of the integer bound of [15] to the polynomial case would lead to the weaker estimate $\deg u_k(z) \leq d_n - 1$ for all k.

In order to show (30), we choose $\vec{a} = \vec{0}$ and $\vec{b} = -\vec{d}$ and we take as the $u_k(z)$ the entries of the last column of a $(\vec{0}, \vec{b})$–minimal multiplier. Degree bounds for the $u_k(z)$ are then determined by using Theorem 5.1 along with Remark 3.7. The row degree of $\mathbf{A}(z)z^{\vec{b}}$ is $\vec{\gamma}^* = \vec{0}$ and in this case $I = (1)$, $J = (1, ..., n-1)$, and $\vec{\alpha} = [\delta]$. Furthermore from our choice of $\vec{b}$ we determine that $K_c = (1)$ from part (b) while since $\mathbf{A}(z) \cdot z^{\vec{b}}$ is column reduced we know from part (c) that $\Delta^{\vec{a},\vec{b}} = 0$ and $|\vec{\beta}| = d_1 - \delta$. The degree bounds from part (a) then implies that

$$\begin{aligned} \deg\ u_1(z) &= \max(-d_1 + \delta, -d_1 + \max[\vec{\beta} + \vec{d}_K] - 1) \\ &= -d_1 + \max[\vec{\beta} + \vec{d}_K] - 1 \\ &\leq -d_1 + |\vec{\beta}| + d_n - 1 \\ &= d_n - \delta - 1. \end{aligned}$$

Bounds for the degrees of the remaining $u_k(z)$, for $2 \leq k \leq n$ are given in Remark 3.7. This allows us to bound the sum of these degrees by

$$\sum_{\substack{k=2 \\ u_k \neq 0}}^{n} (1 + \deg u_k(z)) \leq \sum_{k=2}^{n} \vec{\beta}_k \leq |\vec{\beta}| = d_1 - \delta$$

giving us our desired degree bounds. □

In terms of the total degree of our input matrix polynomial Theorem 5.1 gives the following upper bounds.

Corollary 5.5 *Let $d = \deg \mathbf{A}(z)$ where $\mathbf{A}(z)$ is of size $m \times n$, of rank r and $\vec{a} = \vec{b} = \vec{0}$. Then in the singular case, we have the bounds*

$$\deg \mathbf{U}(z) \leq r \cdot d, \ \deg \mathbf{T}(z) \leq d,$$

while in the nonsingular case we have

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

Proof: The degree bounds follow directly from part (a) of Theorem 5.1 and Remark 3.7. Indeed, for $j \in J_c$, $k \in K_c$ we have from Theorem 5.1 the degree bounds

$$\deg \mathbf{U}(z)_{k,j} \leq \max((r-1)\cdot d - |\vec{\beta}|, \max[\vec{\beta}] - 1)$$

with $0 \leq |\vec{\beta}| \leq r \cdot d$. The remaining entries are bounded using Remark 3.7. □

The following example shows that the bounds from Corollary 5.5 can indeed be tight.

Example 5.6 *Let $q_1, \ldots, q_n$ be polynomials of degree d and set $Q_L = \prod_{j=1, j\notin L}^{n} q_j$. Assume the q_i are chosen so that $Q_{(1)}, \ldots, Q_{(n)}$ are coprime (for example the q_j could have distinct sets of zeros). Let u_i, $i = 1, \ldots, n$, be polynomials such that $\sum_{i=1}^{n} u_j \cdot Q_{(j)} = 1$ and satisfying $\deg u_i < d$ (this is possible since otherwise we can replace each u_i by its remainder after division by q_i). Then the matrix polynomial*

$$\mathbf{A}(z) = \begin{bmatrix} u_1 & u_2 & \ldots & \ldots & u_n \\ q_1 & -q_2 & 0 & & 0 \\ q_1 & 0 & -q_3 & & 0 \\ \vdots & & & \ddots & \vdots \\ & & & & 0 \\ q_1 & 0 & \ldots & 0 & -q_n \end{bmatrix}$$

is unimodular of degree d. Indeed if $\mathbf{U}(z)$ is the matrix given by

$$\mathbf{U}(z)_{k,j} = \begin{cases} Q_{(k)}, & \text{if } k = 1, \ldots, n, j = 1 \\ \frac{u_j \cdot Q_{(j)} - 1}{q_j}, & \text{if } k = j \neq 1 \\ u_j \cdot Q_{(k,j)}, & \text{otherwise} \end{cases}$$

then it is a simple exercise to show that $\mathbf{U}(z)$ is the inverse of $\mathbf{A}(z)$. In this case $\mathbf{U}(z)$ is the unimodular multiplier giving the Popov form $\mathbf{I}_n$ for $\mathbf{A}(z)$ and has degree $(n-1)\cdot d$. If we remove the first row of $\mathbf{A}(z)$ then it is easy to check that $\mathbf{U}(z)$ becomes the minimal multiplier with $\max[\vec{\beta}] = |\vec{\beta}| = r \cdot d = \deg \mathbf{U}(z)$. Notice that the high degree is not just in the kernel part but occurs also in the rest of $\mathbf{U}(z)$. □

For a proof of Theorem 5.1, we will need the following

Lemma 5.7 *Let $\mathbf{U}(z)$ be unimodular with inverse $\mathbf{V}(z)$, and K', J some index lists, with $|K'| = |J|$, and complements K'_c, J_c. Then $\mathbf{V}(z)_{J_c, K'_c}$ is invertible if and only if $\mathbf{U}(z)_{K',J}$ is invertible and in this case*

$$\mathbf{U}(z)_{K'_c, J_c} = \mathbf{V}(z)^{-1}_{J_c, K'_c} + \mathbf{U}(z)_{K'_c, J}[\mathbf{U}(z)_{K',J}]^{-1}\mathbf{U}(z)_{K', J_c}, \tag{31}$$

and

$$\deg\det \mathbf{U}(z)_{K',J} = \deg\det \mathbf{V}(z)_{J_c, K'_c}. \tag{32}$$

Proof: Since $\mathbf{U}(z)$ is invertible, the first statement follows from well-known Schur complement techniques. It remains to show (32) in the case of invertible $\mathbf{U}(z)_{K',J}$. From $\mathbf{V}(z)\cdot\mathbf{U}(z) = \mathbf{I}$ we have

$$\mathbf{U}(z)_{K'_c, J} \cdot [\mathbf{U}(z)_{K',J}]^{-1} = -[\mathbf{V}(z)_{J_c, K'_c}]^{-1} \cdot \mathbf{V}(z)_{J_c, K'}. \tag{33}$$

Both matrix fraction descriptions in (33) are coprime since they result from columns of unimodular matrices [18, Subsection 6.3.1, p. 380]. This implies that we have two partial minimal realizations [18, Theorem 6.5-1, p. 439] of the same rational function and so the denominators must satisfy (32). □

Proof of Theorem 5.1: We start by proving part (b). Let $\mathbf{V}(z) = \mathbf{U}(z)^{-1}$. Taking into account the full rank decomposition

$$\mathbf{A}(z) = \mathbf{T}(z) \cdot \mathbf{V}(z) = \mathbf{T}(z)_{*,J_c} \cdot \mathbf{V}(z)_{J_c,*},$$

we have

$$\text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z) \cdot z^{\vec{b}}) = \text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{T}(z)_{*,J_c}) + \text{Minor-deg}\,(\mathbf{V}(z)_{J_c,*} \cdot z^{\vec{b}}).$$

The right hand term can be computed using Lemma 5.7 and (28)

$$\begin{aligned}
\text{Minor-deg}\,(\mathbf{V}(z)_{J_c,*} \cdot z^{\vec{b}}) &= \max_{|K'_c|=|J_c|} \deg\det\,(\mathbf{V}(z)_{J_c,K'_c} \cdot z^{\vec{b}_{K'_c}}) \\
&= |\vec{b}| + \max_{|K'|=|J|} \deg\det\,(z^{-\vec{b}_{K'}} \cdot \mathbf{U}(z)_{K',J}) \\
&= |\vec{b}| + \deg\det\,(z^{-\vec{b}_K} \cdot \mathbf{U}(z)_{K,J}) \\
&= \deg\det\,(\mathbf{V}(z)_{J_c,K_c} \cdot z^{\vec{b}_{K_c}}) = |\vec{b}_{K_c}| + |\vec{\beta}|
\end{aligned}$$

from which the first part of (b) is readily determined. A similar argument shows that

$$\text{Minor-deg}\,(z^{-\vec{a}} \cdot \mathbf{T}(z)_{*,J_c}) = \deg\det\,(z^{-\vec{a}_I} \cdot \mathbf{T}(z)_{I,J_c}) = |\vec{\alpha}| - |\vec{a}_I|$$

which along with $\mathbf{A}(z)_{I,K_c} = \mathbf{T}(z)_{I,J_c} \cdot \mathbf{V}(z)_{J_c,K_c}$ gives the second part of (b).

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

$$0 \geq \text{Minor-deg}\,(z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z)_{*,K_c} \cdot z^{\vec{b}_{K_c}}) \geq \text{Minor-deg}\,(z^{-\vec{\gamma}_I^*} \cdot \mathbf{A}(z)_{I,K_c} \cdot z^{\vec{b}_{K_c}}).$$

The latter quantity is equal to $-\Delta^{\vec{a},\vec{b}}$ by its definition along with making use of the fact that $\deg\det\,(\mathbf{A}(z)_{I,K_c}) = |\vec{\alpha}| + |\vec{\beta}|$. It remains to discuss the case $\Delta^{\vec{a},\vec{b}} = 0$. By (28) and the above inequalities, $\Delta^{\vec{a},\vec{b}} = 0$ is equivalent to the facts that $\vec{b}_{K_c}$ is the column degree both of $z^{-\vec{\gamma}_I^*} \cdot \mathbf{A}(z)_{I,K_c}$ and $z^{-\vec{\gamma}^*} \cdot \mathbf{A}(z)_{*,K_c}$, and that both matrices are column reduced, as claimed in part (c).

We now turn our attention to a proof of parts (d) and (e). Since $\mathbf{A}(z)_{I,*} \cdot \mathbf{U}(z) = [\mathbf{0}\ ,\ \mathbf{T}(z)_{I,J_c}]$ with square $\mathbf{T}(z)_{I,J_c}$ and unimodular $\mathbf{U}(z)$, the first sentence follows, for example from [18, Lemma 6.3-3, p. 377]. The upper bound for $|\vec{\beta}|$ is implied by part (c). In order to show the upper bound for $\vec{\alpha}^* = \vec{\alpha} - \vec{a}_I$ (with index set J_c), recall that $\mathbf{A}(z)_{I,K_c} = \mathbf{T}(z)_{I,J_c}\mathbf{V}(z)_{J_c,K_c}$ is nonsingular. Consequently, if we set $\vec{\nu}^* = \text{cdeg}\, z^{-\vec{a}_I} \cdot \mathbf{A}(z)_{I,K_c}$ then the Predictable Degree Property, Lemma 3.4, gives $\deg \mathbf{V}(z)_{j,k} \leq \vec{\nu}_k^* - \vec{\alpha}_j^*$, $j \in J_c$, $k \in K_c$. Since $z^{-\vec{\gamma}_I^*} \cdot \mathbf{A}(z)_{I,*} \cdot z^{\vec{b}} = \mathcal{O}(1)_{z\to\infty}$, by bounding $\vec{\nu}^*$, we also have that $\deg \mathbf{V}(z)_{j,k} \leq \max[\vec{\gamma}_I^* - \vec{a}_I] - \vec{b}_k - \vec{\alpha}_j^*$ for all $j \in J_c$, $k \in K_c$. Our bounds now follows by taking into account that $\mathbf{V}(z)_{J_c,K_c}$ is nonsingular, since then for each j in J_c there must exist at least one k in K_c such that $\deg \mathbf{V}(z)_{j,k} \geq 0$.

Finally we prove part (a). From Lemma 5.7 we need only to consider degree constraints for both terms on the right in equation (31). From Remark 3.7 we know that $\vec{b} + \max[\vec{\beta}^*]\vec{e}$ bounds the row degree of $\mathbf{U}(z)_{*,J}$. Since $[\mathbf{U}(z)_{K,J}]^{-1} \cdot \mathbf{U}(z)_{K,J_c} = \mathcal{O}(z^{-1})_{z\to\infty}$ the second term is then bounded by observing that

$$z^{-\vec{b}_{K_c} - \max[\vec{\beta}^*]\vec{e}} \mathbf{U}(z)_{K_c,J} [\mathbf{U}(z)_{K,J}]^{-1} \mathbf{U}(z)_{K,J_c} = \mathcal{O}(z^{-1})_{z\to\infty}.$$

For the first term on the right of equation (31) we use $[\mathbf{V}(z)_{J_c,K_c}]^{-1} = [\mathbf{A}(z)_{I,K_c}]^{-1} \cdot \mathbf{T}(z)_{I,J_c}$ and so

$$z^{-\vec{b}_{K_c}} [\mathbf{V}(z)_{J_c,K_c}]^{-1} z^{-\vec{\alpha}^*} = \underbrace{\frac{1}{\det \mathbf{A}(z)_{I,K_c}}}_{\mathcal{O}(z^{-|\vec{\alpha}|-|\vec{\beta}|})_{z\to\infty}} z^{-\vec{b}_{K_c}} \operatorname{adj}(\mathbf{A}(z)_{I,K_c}) z^{\vec{a}_I} \underbrace{z^{-\vec{a}_I} \mathbf{T}(z)_{I,J_c} z^{-\vec{\alpha}^*}}_{\mathcal{O}(1)_{z\to\infty}}. \tag{34}$$

Since we know that $z^{-\vec{\gamma}_I^*} \mathbf{A}(z)_{I,K_c} z^{\vec{b}_{K_c}} = \mathcal{O}(1)_{z\to\infty}$ we can obtain degree bounds for the adjoint in the above equation by making use of Cramer's formula. This gives

$$\begin{aligned} \deg \operatorname{adj}(\mathbf{A}(z)_{I,K_c})_{i,k} &\le |\vec{\gamma}_I^*| - |\vec{b}_{K_c}| + \vec{b}_i - \vec{\gamma}_k^* \\ &\le |\vec{\gamma}_I^*| - |\vec{b}_{K_c}| + \vec{b}_i - \vec{a}_k + \max[\vec{a}_I - \vec{\gamma}_I^*] \end{aligned}$$

for all i in I and k in K_c. Therefore the middle term of equation (34) has order $\mathcal{O}(z^{|\vec{\gamma}_I^*| - |\vec{b}_{K_c}| + \max[\vec{a}_I - \vec{\gamma}_I^*]})_{z\to\infty}$ and so we have the desired bounds for part (a). □

Remark 5.8 *Note that there are other possibilities for degree bounds for the first part of equation (31) and hence (from the above proof of Theorem 5.1 (a)) for degree bounds for* $\mathbf{U}(z)_{K_c,J_c}$. *For example,* $\vec{\gamma}^*$ *may not be the actual row degrees of* $\mathbf{A}(z) \cdot z^{\vec{b}}$ *but only an upper bound, the proof for part (a) also hold in this latter case. In addition, if* $d = \deg \mathbf{A}(z)$ *is the total degree of* $\mathbf{A}(z)$ *then the middle term of equation (34) is of order* $\mathcal{O}(z^{(r-1)d + \max[\vec{a}] - \min[\vec{b}]})_{z\to\infty}$ *where* r *is the rank of* $\mathbf{A}(z)$. *In this case the bounds in Theorem 5.1 (a) could be replaced by:*

$$\deg z^{\vec{b}_k} \cdot \mathbf{U}(z)_{k,j} \le \max(\vec{\alpha}_j^* + (r-1)d + \max[\vec{a}] - \min[\vec{b}] - |\vec{\alpha}| - |\vec{\beta}|, \quad \max[\vec{\beta}^*] - 1) \tag{35}$$

for all $j \in J_c$ *and* $k \in K_c$. □

Using Theorem 5.1 and Remark 5.8 allows us to deduce a simplified estimation of a degree bound on $\mathbf{U}(z)_{*,J_c}$ and of the shift threshold defined in Theorem 4.2. These estimations have already been used for the presentation of Algorithm SPF at the end of Section 4.

Corollary 5.9 *Let* $\mathbf{A}(z) \in \mathsf{K}[z]^{m\times n}$ *with* $d = \deg \mathbf{A}(z)$. *Then, for any* $\vec{a}$ *and* $\vec{b}$, *in Theorem 4.2 we may choose*

$$N = \min\{m,n\}d + \max[\vec{a}] - \min[\vec{b}].$$

For the corresponding worst case order vector $\vec{\sigma}_0$ *we obtain*

$$|\vec{\sigma}_0| \le m \cdot \Big(2(\min\{m,n\}+1)d + \max[\vec{a}] - \min[\vec{a}] + \max[\vec{b}] - \min[\vec{b}]\Big).$$

Proof: For the first part of the statement we need to find an $N \geq \max\{\max[\vec{\tau}^* - \vec{\alpha}^*], \max[\vec{a} - \vec{\gamma}^*]\}$ where we recall that $\vec{\tau}^*$ is an upper bound for the column degree of $z^{-\vec{b}} \cdot \mathbf{U}(z)_{*,J_c}$. Let r be the rank of $\mathbf{A}(z)$. Clearly $\max[\vec{a} - \vec{\gamma}^*] \leq d + \max[\vec{a}] - \min[\vec{b}]$. In addition, the second term in equation (35) of Remark 5.8 can be estimated using Theorem 5.1(b). In this case

$$\max[\vec{\beta}^*] \leq |\vec{\beta}| - \min[\vec{b}] \leq \deg\det\left(\mathbf{A}(z)_{I,K_c}\right) - \min[\vec{b}] \leq rd - \min[\vec{b}]$$

which implies that

$$\deg \mathbf{U}(z)_{k,j} \leq \max(\vec{b}_k + \vec{\alpha}_j^* + (r-1)d + \max[\vec{a}] - \min[\vec{b}] - |\vec{\alpha}| - |\vec{\beta}|, \quad \vec{b}_k + rd - \min[\vec{b}] - 1)$$

for all $j \in J_c$ and $k \in K_c$. Consequently, we obtain $rd + \max[\vec{a}] - \min[\vec{b}]$ as an estimate for $\max[\vec{\tau}^* - \vec{\alpha}^*]$. Replacing the rank r by the larger quantity $\min\{m, n\}$ completes the proof of the first statement.

We now turn to the proof of the second statement. With N as before, we have to estimate the order vector

$$\vec{\sigma}_0 = \vec{\gamma}^* + \Big(1 + \max\{N + \max[\vec{\alpha}^*], \max[\vec{\beta}^*]\}\Big)\vec{e}$$

given in Theorem 4.2. From the above reasoning we have

$$N + \max[\vec{\alpha}^*] \geq \max[\vec{\tau}^*] \geq \max[\vec{\beta}^*] - 1$$

and hence

$$\max\{N + \max[\vec{\alpha}^*], \max[\vec{\beta}^*]\} \leq N + 1 + \max[\vec{\alpha}^*] \leq N + 1 + \max[\vec{\alpha}] - \min[\vec{a}].$$

Substituting the explicit value for N and using the rough bounds $|\vec{\gamma}^*| \leq (\max[\vec{b}] + d)m$ and $\max[\vec{\alpha}] \leq |\vec{\alpha}| \leq \deg\det \mathbf{A}(z)_{I,K_c} \leq r \cdot d$ leads to the claimed upper bound for $|\vec{\sigma}_0|$. □

6 Cost of the Algorithm

A worst case bound for the cost of Algorithm SPF will depend on the size of the input matrix (dimensions, bit sizes and degrees) and on the input shifts. We consider an input matrix $\mathbf{A}(z) \in \mathsf{D}[z]^{m \times n}$ with D an integral domain. We assume that D is such that for any two elements a and b in D, the elementary operations (addition, product, exact division) are using $O(\text{size}(a) \cdot \text{size}(b))$ bit operations – with a standard arithmetic – or $O^{\sim}(\text{size}(a) + \text{size}(b))$ bit operations – with a fast arithmetic (based on FFT for instance). Here, $O^{\sim}(p(n))$ denotes $p(n)^{1+o(1)}$. The function *size* is such that the result of an operation has a size of $O(\text{size}(a) + \text{size}(b))$ bits.

To reach a Mahler system with order $\vec{\sigma}$, we know from [5, Theorem 6.2] that SPF has cost $O((m+n)|\vec{\sigma}|^2)$ operations in D, on elements of size bounded by $O(|\vec{\sigma}| \log ||\mathbf{A}||)$. By $\log ||\mathbf{A}||$ we denote the length of the entries in $\mathbf{A}$. In terms of bit operations the cost is thus $O((m+n)|\vec{\sigma}|^4 \log^2 ||\mathbf{A}||)$ using a standard arithmetic or $O^{\sim}((m+n)|\vec{\sigma}|^3 \log ||\mathbf{A}||)$ using a fast arithmetic. With the order threshold $\vec{\sigma}_0$ defined by (21), the bound of Corollary 5.9 on $\vec{\sigma}_0$ and taking $\vec{a} = \vec{b} = 0$ we get:

Corollary 6.1 *Let $\mathbf{A} \in \mathsf{D}[z]^{m \times n}$ be of degree d. The Popov normal form of $\mathbf{A}$ and a corresponding minimal multipliers can be computed by the fraction-free algorithm* SPF *using $O^{\sim}((m+n)(md \min\{m,n\})^3 \log ||\mathbf{A}||)$ bit operations (with a fast arithmetic).* □

The worst-case value N proposed in Corollary 5.9 is very easy to compute and applies for any set of data. However, it could be improved for certain classes of matrices and hence lead to smaller complexities in special cases. We should also mention that the cost of the algorithm SPF may very well depend on the choice of the shift parameter N, even for values over the threshold (20), as becoming clear from the following example.

Example 6.2 *Let $\mathbf{A}(z)$ be a square matrix in $\mathsf{K}[z]^{n\times n}$ with $d = \deg \mathbf{A}(z)$, $\mathbf{A}(0)$ being invertible and $\vec{a} = \vec{b} = \vec{0}$* i.e. *we compute the (unshifted) Popov form of $\mathbf{A}(z)$. For any choice of $N \geq 0$ we thus consider Mahler systems of type $\vec{n} = (N\vec{e}, \vec{0})$. It is not difficult to show that, after $k \cdot n$ steps of the algorithm* SPF *, $0 \leq k \leq N$, we obtain*

$$\vec{\sigma} = k \cdot \vec{e}, \quad \vec{\mu} = (k \cdot \vec{e}, \vec{0}), \quad \mathbf{M}(z) = \pm[\det \mathbf{A}(0)]^k \cdot \begin{bmatrix} z^k \mathbf{I}_n & \mathbf{B}(z)^{(k)} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} \tag{36}$$

where $\mathbf{B}(z)^{(k)}$, of total degree $k-1$, is the partial sum of the inverse power series $\mathbf{A}(z)^{-1}$. In particular, one may check that by construction, for any $0 \leq k \leq N$, $\mathbf{M}(z)$ in (36) is a Mahler system of type $\vec{n}$ with order vector $k \cdot \vec{e}$. We thus see that to compute the Popov form of $\mathbf{A}(z)$, SPF *implicitely proceeds by first computing an approximation of order N of $\mathbf{A}(z)^{-1}$. This may be compared to the method of [34, Lemma 2].* □

If $\mathbf{A}(z)$ is of degree d, from Theorem 5.1 (b), we know that $0 \leq \vec{\alpha}_i \leq |\vec{\alpha}| \leq \deg\det \mathbf{A}(z)_{I,K_c} \leq d \min\{m,n\}$. Hence, the shift $\vec{a} = (id \min\{m,n\})_{i=(m-1),\ldots,0}$ satisfies the sufficient condition (9) for computing the Hermite form. With the corresponding bound for $|\vec{\sigma}_0|$ in Corollary 5.9 we obtain:

Corollary 6.3 *Let $\mathbf{A} \in \mathsf{D}[z]^{m\times n}$ be of degree d. The Hermite normal form of $\mathbf{A}$ and a corresponding minimal multipliers can be computed by the fraction-free algorithm* SPF *using $\tilde{O}((m+n)(m^2 d \min\{m,n\})^3 \log||\mathbf{A}||)$ bit operations (with a fast arithmetic).* □

Concerning the computation of small multipliers for the gcd of n polynomials, one can use the notations and the results of Example 5.4. The input polynomials have minimum degree d_1 and maximum degree d_n. Their gcd has degree δ. In Theorem 4.2 this gives $\mathcal{N}_0 = d_1 + d_n - 2\delta - 1$ and the order threshold $\sigma_0 = d_1 + d_n - \delta + 1$, hence:

Corollary 6.4 *Multipliers that verify the degree bounds (30) for the gcd of n polynomials can be computed by the fraction-free algorithm* SPF *using $\tilde{O}(n(d_1 + d_n - \delta)^3 \log||\mathbf{A}||)$ bit operations (with a fast arithmetic).* □

7 Conclusion

In this paper we have presented an algorithm for the computation of a shifted Popov Normal Form of an arbitrary rank rectangular polynomial matrix. For specific input shifts, our approach gives methods for computing matrix normal forms (such as Hermite and Popov) and the matrix greatest common divisor of two matrix polynomials (in normal form). The method used is to embed the problem of computing shifted forms into one of computing matrix rational approximants.

In the case of matrix normal forms, our methods compute both the form and a unimodular matrix that describes the elementary operations used to obtain the form. In the case of rectangular matrix input, the corresponding multipliers for the shifted forms are not unique. We use the concept of minimal matrix approximants to introduce a notion of minimal mutipliers and show how such multipliers are computed by our methods.

The proposed method has the advantage that in the case of exact arithmetic domains all computations can be done using fraction-free arithmetic. This ensures that the problem of intermediate expression swell is minimized for such computations. To our knowledge we know of no fraction-free methods that handle all the normal forms and matrix greatest common divisor problems covered in this paper.

There are other methods that can also be used to reduce intermediate expression swell in exact arithmetic computations. In particular, modular methods can be used for such cases. These methods reduce a single computation to a number of similar computations in simpler domains and then reconstruct the result using a chinese remaindering technique. Since modular methods are typically an order of magnitude faster than fraction-free methods we plan on investigating such methods in the future. We remark that our present paper already contributes to such a method since modular methods require that we reconstruct an object in the original domain and know when to stop. Since our computations are all done in the original domain, as opposed to moving to a quotient domain, our results are of interest for such problems.

Our method embeds our problem into a rational approximation problem and then relies on a variation of the method of [5] for computing a solution to the problem. The concern with this method is that it is not a reduction process and, as such, does not recognize early the case where we quickly convert to a normal form. We plan on investigating reduction methods for computing shifted normal forms. Since we are interested in exact arithmetic domains along with fixed cost domains we will look for methods which are fraction-free. We also hope to address similar problems with respect to algorithms for the computation of matrices of linear difference and differential operators. Popov forms for such noncommutative domains are interesting for their use in finding series and closed form solutions of systems of difference and differential equations.

References

[1] B. Beckermann and G. Labahn, A uniform approach for Hermite Padé and simultaneous Padé Approximants and their matrix generalizations, *Numerical Algorithms* **3** (1992) 45-54.

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

[3] B. Beckermann and G. Labahn, Recursiveness in Matrix Rational Interpolation Problems, *J. Comput. Appl. Math.* **77** (1997) 5-34.

[4] B. Beckermann, S. Cabay and G. Labahn, Fraction–free Computation of Matrix Padé Systems, *Proceeding of International Symposium on Symbolic and Algebraic Computation, ISSAC'97*, Maui, ACM Press, (1997) 125-132.

[5] B. Beckermann and G. Labahn, Fraction-free Computation of Matrix Gcd's and Rational Interpolants, *SIAM J. Matrix Anal. Appl.*, **22**:1 (2000) 114-144.

[6] B. Beckermann, G. Labahn and G. Villard, Shifted Normal Forms of Polynomial Matrices, *Proceeding of International Symposium on Symbolic and Algebraic Computation, ISSAC'99*, Vancouver, ACM Press, (1999) 189-196.

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

[8] Th.G.J. Beelen and P.M. Van Dooren, An improved algorithm for the computation of Kronecker's canonical form of a singular pencil, *Linear Algebra and its Applications* **105** (1988) 9–65.

[9] V.A. Belyĭ, V.B. Khazanov and V.N. Kublanovskaya, Spectral problems for matrix pencils. Methods and algorithms III, *Soviet J. Numer. Anal. Math. Model.* **4**:1 (1989) 19-51.

[10] R.R. Bitmead, S.Y. Kung, B.D.O. Anderson and T. Kailath, Greatest Common Divisors via Generalized Sylvester and Bezout Matrices, *IEEE Trans. Automat. Contr.* **AC-23** (1978) 1043-1046.

[11] A. Bultheel and M. Van Barel, A matrix Euclidean Algorithm and the Matrix Minimal Padé Approximation Problem, in: *Continued Fractions and Padé Approximants*, C. Brezinski, ed., Elsevier, North-Holland (1990) 11-51.

[12] P.D. Domich, R. Kannan, R. and L.E. Trotter Jr., Hermite normal form computation using modulo determinant arithmetic, *Mathematics of Operations Research*, **12**:1, (1987) 50-59.

[13] G.D. Forney, Jr., Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems, *SIAM J. Control and Optimization*, **13**, (1975) 493-520.

[14] K.O. Geddes, S.R. Czapor and G. Labahn, *Algorithms for Computer Algebra*, Kluwer, Boston, MA (1992).

[15] G. Havas, B.S. Majewski and K.R. Matthews, Extended gcd and Hermite normal form algorithms via lattice basis reduction, *Experimental Mathematics*, **7**:2, (1998) 125-135.

[16] C. Hermite, Sur l'introduction des variables continues dans la théorie des nombres, *J. Reine Angew. Math.*, **41**, (1851) 191-216.

[17] C.S. Iliopoulos, Worst-case complexity bounds on algorithms for computing the canonical structure of finite abelian groups and the Hermite and Smith normal forms of an integer matrix, *SIAM J. Comput.*, **18**:4, (1989) 658-669.

[18] T. Kailath, *Linear systems*, Prentice–Hall (1980).

[19] R.E. Kalman, Irreducible realizations and the degree of a rational matrix, *SIAM J. Appl. Math.* **13**, (1965) 520-544.

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

[21] S.Y. Kung, T. Kaliath and M. Morf, A generalized resultant matrix for polynomial matrices, in: *Proceeding of IEEE Conf. on Decision and Control*, Florida (1976) 892-895.

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

[23] K. Mahler, Perfect systems, *Compos. Math.* **19** (1968) 95-166.

[24] T. Mulders and A. Storjohann, *On lattice reduction for polynomial matrices*, Tech. Rep. 356, Department of Computer Science, ETH Zurich Switzerland, dec. 2000.

[25] M. Newman, *Integral matrices*, Academic Press, New-York (1972).

[26] V. M. Popov, Some Properties of Control Systems with Irreducible Matrix Transfer Functions, in *Lecture Notes in Mathematics*, Vol. **144**, Springer, Berlin, (1969) 169-180.

[27] M.P. Quéré-Stuchlik, *Algorithmique des faisceaux linéaires de matrices, application à la théorie des systèmes linéaires et à la résolution d'équations algébro-différentielles*, Thèse de Doctorat, Université de Paris VI, France (1997).

[28] A. Schrijver, *Theory of linear and integer programming*, Wiley-Interscience series in Discrete Mathematics (1986).

[29] A. Storjohann, *Computation of Hermite and Smith normal forms of matrices*, Master's Thesis, University of Waterloo, Canada (1994).

[30] A. Storjohann, *Algorithms for Matrix Canonical Form*, PhD Thesis, Eidgenössiche Technische Hochshule ETH, Zürich, Switzerland (2000).

[31] A. Storjohann and G. Labahn, Preconditioning of Rectangular Polynomial Matrices for efficient Hermite Normal Form computation, *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC'95*, Montreal, Canada, ACM Press, (1995) 119–125.

[32] A. Storjohann and G. Labahn, Asymptotically fast computation of Hermite normal forms of integer matrices, *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC'96*, Zürich, ACM Press, (1996) 259-266.

[33] M. Van Barel and A. Bultheel, A general module theoretic framework for vector M-Padé and matrix rational interpolation, *Numerical Algorithms* **3** (1992) 451-462.

[34] G. Villard, Computing Popov and Hermite forms of polynomial matrices, *Proceedings of the International Symposium on Symbolic and Algebraic Computation, ISSAC'96*, Zürich, ACM Press (1996) 250-258.

[35] W.A. Wolovich and P.J. Antsaklis, The Canonical Diophantine Equations with Applications, *SIAM J. Control and Optimization*, **22** (1984) 777-787.

[36] W.M. Wonham, *Linear multivariable control*, Springer (1974).