Computing Minimal Nullspace Bases

Wei Zhou, George Labahn, and Arne Storjohann
Cheriton School of Computer Science
University of Waterloo,
Waterloo, Ontario, Canada
{w2zhou,glabahn,astorjoh}@uwaterloo.ca

ABSTRACT

In this paper we present a deterministic algorithm for the computation of a minimal nullspace basis of an $m \times n$ input matrix of univariate polynomials over a field $\mathbb{K}$ with $m \leq n$. This algorithm computes a minimal nullspace basis of a degree d input matrix with a cost of $O^{\sim}(n^{\omega}\lceil md/n \rceil)$ field operations in $\mathbb{K}$. Here the soft-O notation is Big-O with log factors removed while ω is the exponent of matrix multiplication. The same algorithm also works in the more general situation on computing a shifted minimal nullspace basis, with a given degree shift $\vec{s} \in \mathbb{Z}^n_{\geq 0}$ whose entries bound the corresponding column degrees of the input matrix. In this case if ρ is the sum of the m largest entries of $\vec{s}$, then a $\vec{s}$-minimal right nullspace basis can be computed with a cost of $O^{\sim}(n^{\omega}\rho/m)$ field operations.

Categories and Subject Descriptors: I.1.2 [Symbolic and Algebraic Manipulation]: Algorithms; F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems

General Terms: Algorithms, Theory

Keywords: Nullspace basis, Complexity

1. INTRODUCTION

Let $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$ be a matrix of polynomials over a field $\mathbb{K}$ with rank $r \leq m \leq n$. The set

$$\{\mathbf{p} \in \mathbb{K}[x]^n \mid \mathbf{F}\mathbf{p} = 0\},$$

is a (right) nullspace of $\mathbf{F}$, which is also a $\mathbb{K}[x]$-module. It can be generated by a basis – a nullspace basis of $\mathbf{F}$, that can be represented as a matrix in $\mathbb{K}[x]^{n \times (n-r)}$, with the columns being the basis elements.

Nullspaces of polynomial matrices appear in a large number of applications, being first used as an algebraic formalism in the area of control theory (Kucera, 1979). For example, in linear system theory if a system is respresented by a transfer function given in terms of a left coprime matrix fraction decomposition $\mathbf{T} = \mathbf{D}_\ell^{-1}\mathbf{N}_\ell$, with $\mathbf{D}_\ell$ and $\mathbf{N}_\ell$ polynomial matrices, then one often wants to find a right coprime matrix fraction representation $\mathbf{T} = \mathbf{N}_r\mathbf{D}_r^{-1}$ with $\mathbf{D}_r$ and $\mathbf{N}_r$ polynomial matrices of appropriate dimensions (Kailath, 1980). This is equivalent to the nullspace basis computation

$$[\mathbf{D}_\ell \quad -\mathbf{N}_\ell] \begin{bmatrix} \mathbf{N}_r \\ \mathbf{D}_r \end{bmatrix} = 0. \tag{1}$$

Solving and determining fundamental properties of the basic matrix equation $\mathbf{AZ} = \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ have polynomial elements can be determined by finding a complete description (that is, a basis) of the nullspace of $[\mathbf{A}, -\mathbf{B}]$. Other examples of the use of nullspaces and their bases include fault diagnostics (Frisk, 2001), column reduction of matrix polynomials, matrix inverse and determinant computations (Beelen et al., 1988; Jeannerod and Villard, 2006, 2005).

In most applications one is interested in finding a *minimal nullspace basis* of $\mathbf{F}$ in $\mathbb{K}[x]^n$ (Forney, 1975). A nullspace basis $\mathbf{N}$ of $\mathbf{F}$ is said to be minimal if it has the minimal possible column degrees among all right nullspace bases. This is also often referred to as a *minimal polynomial basis*. Examples where minimality are needed include the right coprime matrix factorization problem and the problem of column reducing a polynomial matrix. As an example, finding a basis for the nullspace corresponding to the right matrix fraction problem (1) finds a matrix fraction while a minimal nullspace basis finds such a fraction in reduced form having a minimal column degree denominator (needed for example in minimal partial realization problems). In some cases, for example when using nullspace bases for column reduction, as in (Beelen et al., 1988), or for normal form computation, as in (Beckermann et al., 1999, 2006), one is interested in shifting the importance of the degrees of some of the rows of a basis via a vector. If $\vec{s} = [s_1, \ldots, s_n] \in \mathbb{Z}^n$ then the shifted $\vec{s}$-column degree of a column vector of polynomials $\mathbf{p} = [p_1, \ldots, p_n]^T$ is

$$\deg_{\vec{s}} \mathbf{p} = \max_i \{\deg(p_i) + s_i\}.$$

The $\vec{s}$-column degree specializes to the column degree when $\vec{s} = 0$. A nullspace basis $\mathbf{N}$ is said to be $\vec{s}$-minimal if it has the minimal possible $\vec{s}$-column degrees among all nullspace bases, or equivalently, the column degrees of

$$x^{\vec{s}}\mathbf{N} = \mathrm{diag}(x^{s_1}, \ldots, x^{s_n}) \cdot \mathbf{N}$$

are the minimal possible among all nullspace bases of $\mathbf{F}$.

In this paper we are interested in fast computation of minimal nullspace bases and shifted minimal nullspace bases in exact environments. Historically computation of a minimal

ISSAC 2012, July 22–25, 2012, Grenoble, France.

nullspace basis has made use of either matrix pencil or resultant methods (often called linearized approaches). Matrix pencil methods convert a nullspace basis computation problem to one of larger matrix size but having polynomial degree one. In this case a minimal nullspace basis is determined from the computation of the Kronecker canonical form, with efficient algorithms given by (Beelen and Dooren, 1988; Misra et al., 1994; Oara and Dooren, 1997). The cost of these algorithms is $O(m^2nd^3)$. Resultant methods convert the nullspace basis computation of the matrix polynomial $\mathbf{F}$ into a block Toeplitz kernel problem with much higher dimension with the resulting complexity again being high. In (Storjohann and Villard, 2005) the authors give an algorithm for computing a nullspace basis with a cost of $O^{\sim}\left(nmr^{\omega-1}d\right)$ where $O^{\sim}$ is the same as Big-O but without log factors and where ω is the power of fast matrix multiplication. However, their algorithm is randomized Las Vegas and, in addition, the bases they compute are not minimal in general.

In this paper we present a new, deterministic algorithm for computing a minimal nullspace basis with a complexity cost of $O^{\sim}\left(n^{\omega}\lceil md/n\rceil\right)$ field operations in $\mathbb{K}$. This cost reduces to $O^{\sim}\left(n^{\omega-1}md\right)$ when $md \in \Omega(n)$, that is, when md is asymptotically bounded below by a constant factor of n, as in the case of $md \geq n$. The same algorithm can also compute a $\vec{s}$-minimal nullspace basis of $\mathbf{F}$ with a cost of $O^{\sim}(n^{\omega}\rho/m)$ if the entries of $\vec{s}$ bound the corresponding column degrees of $\mathbf{F}$, where ρ is the sum of the m largest entries of $\vec{s}$. The computational cost in this paper is analyzed by bounding the number of arithmetic operations in the coefficient field $\mathbb{K}$ on an algebraic random access machine. We assume the cost of multiplying two polynomial matrices with dimension n and degree d is $O^{\sim}(n^{\omega}d)$ field operations, where the multiplication exponent ω is assumed to satisfy $2 < \omega \leq 3$. We refer to the book by von zur Gathen and Gerhard (2003) for more details and reference about the cost of polynomial multiplication and matrix multiplication.

Our method uses three complementary techniques: reduce the column dimension by using an order basis computation to compute a partial nullspace basis, reduce the row dimension by first computing a nullspace basis of a subset of the rows, and finally, controlling the degrees by maintaining a bound throughout the computation. One key component of the algorithm, the computation of order basis (also known as minimal approximant basis or σ-basis) (Beckermann and Labahn, 1994) can be done efficiently using the algorithms from Giorgi et al. (2003) and Zhou and Labahn (2009, 2012), allowing us to efficiently reduce the column dimension. The problem can then be separated to two subproblems of smaller row dimensions, which can then be handled recursively.

The remainder of this paper is structured as follows. Basic definitions and properties of order bases and nullspace bases are given in the next section. The details of our nullspace basis computation and a formal statement of the algorithm can be found in Section 3. A complexity analysis of the algorithm is provided in the following section. The paper ends with a conclusion and topics for future research.

2. PRELIMINARIES

In this section, we provide some of the background needed in order to understand the basic concepts and tools needed for nullspace basis computation. We also provide a brief introduction to order basis, a key ingredient in our algorithm.

2.1 Nullspace Basis

Let $\mathbb{K}$ be a field. Given a polynomial matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, we are interested in computing a minimal (right) nullspace basis of $\mathbf{F}$, or more generally, a shifted minimal nullspace basis of $\mathbf{F}$. While minimality is often given in terms of the degrees alone it is sometimes important to consider this in terms of shifted degrees (Beckermann et al., 2006) as given in the introduction.

A shifted column degree (called the *$\vec{s}$-column degree*, or simply the *$\vec{s}$-degree*) is equivalent to the notion of *defect* commonly used in the literature. As in the uniform shift case, we say a matrix is *$\vec{s}$-column reduced* or *$\vec{s}$-reduced* if its $\vec{s}$-degrees cannot be decreased by unimodular column operations. More precisely, if a matrix $\mathbf{P}$ is $\vec{s}$-column reduced and $[d_1,\ldots,d_n]$ are the $\vec{s}$-degrees of columns of $\mathbf{P}$ sorted in nondecreasing order, then $[d_1,\ldots,d_n]$ is lexicographically minimal among all matrices right equivalent to $\mathbf{P}$. Note that a matrix $\mathbf{P}$ is $\vec{s}$-column reduced if and only if $x^{\vec{s}}\cdot\mathbf{P}$ is column reduced. A $\vec{s}$-minimal (right) nullspace basis of a given polynomial matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ is then defined as follows.

Definition 2.1. *Given $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, a polynomial matrix $\mathbf{N} \in \mathbb{K}[x]^{n\times *}$ is a $\vec{s}$-minimal (right) nullspace basis of $\mathbf{F}$ if the following properties hold:*

1. *$\mathbf{N}$ is full-rank and $\vec{s}$-column reduced (equivalently, the leading column coefficient matrix of $x^{\vec{s}}\mathbf{N}$ is full-rank).*
2. *$\mathbf{N}$ satisfies $\mathbf{F}\cdot\mathbf{N} = 0$.*
3. *Any $\mathbf{q} \in \mathbb{K}[x]^n$ satisfying $\mathbf{Fq} = 0$ can be expressed as a linear combination of the columns of $\mathbf{N}$, that is, there exists some polynomial vector $\mathbf{p}$ such that $\mathbf{q} = \mathbf{Np}$.*

Note that a $\vec{s}$-column reduced nullspace basis of $\mathbf{F}$ has the minimal $\vec{s}$-column degrees among all nullspace bases of $\mathbf{F}$.

2.2 Order Basis

Order basis computation is a key tool we use in our nullspace basis computation, which is not surprising considering the close relationship between order basis and nullspace basis, as we will see from the definition of order basis. First, let us look at the *order* in order basis.

Let $\mathbb{K}$ be a field, $\mathbf{F} \in \mathbb{K}[[x]]^{m\times n}$ a matrix of power series and σ a non-negative integer.

Definition 2.2. *A vector of polynomials $\mathbf{p} \in \mathbb{K}[x]^{n\times 1}$ has* order $(\mathbf{F},\sigma)$ *(or* order σ *with respect to $\mathbf{F}$) if $\mathbf{F}\cdot\mathbf{p} \equiv \mathbf{0}$* mod x^{σ}*, that is,*

$$\mathbf{F}\cdot\mathbf{p} = x^{\sigma}\mathbf{r}$$

for some $\mathbf{r} \in \mathbb{K}[[x]]^{m\times 1}$. The set of all order $(\mathbf{F},\sigma)$ vectors is a $\mathbb{K}[x]$-module denoted by $\langle(\mathbf{F},\sigma)\rangle$.

An *order basis* $\mathbf{P}$ of $\mathbf{F}$ with order σ and shift $\vec{s}$, or simply a $(\mathbf{F},\sigma,\vec{s})$-basis, is a polynomial matrix whose columns form a basis for the module $\langle(\mathbf{F},\sigma)\rangle$ having minimal $\vec{s}$-column degrees (Beckermann and Labahn, 1994, 1997). Again, note that a $\vec{s}$-column reduced basis of $\langle(\mathbf{F},\sigma)\rangle$ has the minimal $\vec{s}$-column degrees among all bases of $\langle(\mathbf{F},\sigma)\rangle$.

Definition 2.3. *A polynomial matrix $\mathbf{P}$ is an order basis of $\mathbf{F}$ of order σ and shift $\vec{s}$, denoted by $(\mathbf{F},\sigma,\vec{s})$-basis, if the following properties hold:*

1. $\mathbf{P}$ *is nonsingular and* $\vec{s}$*-column reduced.*

2. $\mathbf{P}$ *has order* $(\mathbf{F}, \sigma)$ *(or equivalently, each column of* $\mathbf{P}$ *is in* $\langle(\mathbf{F}, \sigma)\rangle$*).*

3. *Any* $\mathbf{q} \in \langle(\mathbf{F}, \sigma)\rangle$ *can be expressed as a linear combination of the columns of* $\mathbf{P}$*, given by* $\mathbf{P}^{-1}\mathbf{q}$.

Note that the definition of order can be easily extended to having a different order for each row of $\mathbf{F} \cdot \mathbf{P}$ as in (Zhou and Labahn, 2009). However, a single uniform order is sufficient for our discussion of minimal nullspace basis computation in this paper.

2.3 Computing Nullspace Bases via Order Bases

Minimal nullspace bases can be directly computed via order basis computation. Indeed if the order σ of a $(\mathbf{F}, \sigma, \vec{s})$-basis $\mathbf{P}$ is high enough, then $\mathbf{P}$ contains a $\vec{s}$-minimal nullspace basis $\mathbf{N}$, as we will see later in Lemma 3.3. However, this approach may require the order σ to be quite high. For example, if $\mathbf{F}$ has degree d and $\vec{s}$ is uniform, then its minimal nullspace bases can have degree up to md. In that case the order σ would need to be set to $d + md$ in the order basis computation in order to fully compute a minimal nullspace basis. The fastest method of computing such a $(\mathbf{F}, d + md)$-basis would cost $O^{\sim}\left(n^{\omega}\left\lceil m^2 d/n\right\rceil\right)$ using the algorithm from (Zhou and Labahn, 2009).

We can see from this last cost that there is room for improvement when m is large. For example, in the worst case when $m \in \Theta(n)$ this cost would be $O^{\sim}\left(n^{\omega+1}d\right)$. Here $m \in \Theta(n)$ means that $m \in O(n)$ and $n \in O(m)$, that is, m is asymptotically close to n and within a constant factor of n. This points to a root cause for the inefficiency in this approach. Namely, when m is large, the computed nullspace basis, with a column dimension usually $n - m$, is a small subset of the order basis computed. Hence considerable effort is put in the computation of order basis elements that are not part of a nullspace basis. A key to reducing the cost is therefore to reduce such computation of unneeded order basis elements, which is achieved in our algorithm by only using order basis computation to compute partial nullspace bases of low degrees.

3. NULLSPACE BASIS COMPUTATION

In this section, we describe a new, efficient algorithm for computing a shifted minimal nullspace basis. The algorithm uses two computation processes recursively. The first process, described in Subsection 3.2, uses an order basis computation to compute a subset of nullspace basis elements of lower degree, and results in a new problem of lower column dimension. The second process, described in Subsection 3.4, reduces the row dimension of the problem by computing a nullspace basis of a submatrix formed by a subset of the rows of the input matrix.

We require that the entries of the shift $\vec{s}$ to be non-negative and bound the corresponding column degrees of $\mathbf{F}$. For example, we can set $\vec{s}$ to be the list of the column degrees of $\mathbf{F}$, or we can simply set each entry of $\vec{s}$ to be the maximum column degree of $\mathbf{F}$. This is a very useful condition as it helps us to keep track of and bound the shifted degrees throughout the nullspace basis computation, as we will see in Subsection 3.1.

For simplicity, we will also assume without loss of generality that the columns of $\mathbf{F}$ and the corresponding entries of $\vec{s} = [s_1, \ldots, s_n]$ are arranged so that the entries of $\vec{s}$ are in increasing order.

Let $\rho = \sum_{n-m+1}^{n} s_i$ be the sum of m largest entries of $\vec{s}$, and $s = \rho/m$ be their average. The algorithm we present in this section computes a $\vec{s}$-minimal nullspace basis $\mathbf{N}$ with a cost of $O^{\sim}(n^{\omega}s)$ field operations. For a uniform shift $\vec{s} = [s, \ldots, s]$, we improve this later to $O^{\sim}\left(n^{\omega}\left\lceil ms/n\right\rceil\right)$.

3.1 Bounds based on the shift

A key requirement for efficient computation is making sure that the intermediate computations do not blow up in size. We will see that this requirement is satisfied by the existence of a bound, $\xi = \sum \vec{s} = \sum_{i=1}^{n} s_i$, on the sum of all entries of the input shift of all subproblems throughout the computation. Here, and in the rest of this paper, we use the summation notation $\sum$ without index to denote the summation over all elements of the list.

First, we have the following, easily proved, bound on the column degrees of the product of $\mathbf{F}$ with another matrix.

LEMMA 3.1. *Let* $\vec{s}$ *be a shift whose entries bound the corresponding column degrees of* $\mathbf{F}$*. Then for any polynomial matrix* $\mathbf{A}$*, the column degrees of* $\mathbf{FA}$ *are bounded by the corresponding* $\vec{s}$*-column degrees of* $\mathbf{A}$.

The following lemma gives a bound on the $\vec{s}$-column degrees of $(\mathbf{F}, \sigma, \vec{s})$-bases.

LEMMA 3.2. *The sum of the* $\vec{s}$*-column degrees of a* $(\mathbf{F}, \sigma, \vec{s})$*-basis* $\mathbf{P}$ *is at most* $\xi + r\sigma$*, where* r *is the rank of* $\mathbf{F}$.

PROOF. The sum of the $\vec{s}$-column degrees is ξ at order 0, since the identity matrix is a $(\mathbf{F}, 0, \vec{s})$-basis. This sum increases by at most r for each order increase, as can be seen from the iterative computation of order bases in (Beckermann and Labahn, 1994; Giorgi et al., 2003). □

The following lemma shows that any $(\mathbf{F}, \sigma, \vec{s})$-basis contains a partial $\vec{s}$-minimal nullspace basis of $\mathbf{F}$, and as a result, any $(\mathbf{F}, \sigma, \vec{s})$-basis with high enough σ contains a $\vec{s}$-minimal nullspace basis of $\mathbf{F}$.

LEMMA 3.3. *Let* $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2]$ *be any* $(\mathbf{F}, \sigma, \vec{s})$*-basis and* $\mathbf{N} = [\mathbf{N}_1, \mathbf{N}_2]$ *be any* $\vec{s}$*-minimal nullspace basis of* $\mathbf{F}$*, where* $\mathbf{P}_1$ *and* $\mathbf{N}_1$ *contain all columns from* $\mathbf{P}$ *and* $\mathbf{N}$*, respectively, whose* $\vec{s}$*-column degrees are less than* σ*. Then* $[\mathbf{P}_1, \mathbf{N}_2]$ *is a* $\vec{s}$*-minimal nullspace basis of* $\mathbf{F}$*, and* $[\mathbf{N}_1, \mathbf{P}_2]$ *is a* $(\mathbf{F}, \sigma, \vec{s})$*-basis.*

PROOF. From Lemma 3.1, any column $\mathbf{p}$ of $\mathbf{P}_1$ satisfies $\deg \mathbf{Fp} \leq \deg_{\vec{s}} \mathbf{p} < \sigma$. Combining this with the fact that $\mathbf{Fp} \equiv 0 \mod x^{\sigma}$ we get $\mathbf{Fp} = 0$. Thus $\mathbf{P}_1$ is generated by $\mathbf{N}_1$, that is, $\mathbf{P}_1 = \mathbf{N}_1\mathbf{U}$ for some polynomial matrix $\mathbf{U}$. On the other hand, $\mathbf{N}_1$ has order $(\mathbf{F}, \sigma)$ and therefore satisfies $\mathbf{N}_1 = \mathbf{P}_1\mathbf{V}$ for some polynomial matrix $\mathbf{V}$. We now have $\mathbf{P}_1 = \mathbf{P}_1\mathbf{VU}$ and $\mathbf{N}_1 = \mathbf{N}_1\mathbf{UV}$, implying both $\mathbf{U}$ and $\mathbf{V}$ are unimodular. The result then follows from the unimodular equivalence of $\mathbf{P}_1$ and $\mathbf{N}_1$ and the fact that they are $\vec{s}$-column reduced. □

We can now provide a simple bound on the $\vec{s}$-minimal nullspace basis of $\mathbf{F}$.

THEOREM 3.4. *Suppose* $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$ *and* $\vec{s} \in \mathbb{Z}_{\geq 0}^{n}$ *is a shift with entries bounding the corresponding column degrees of* $\mathbf{F}$*. Then the sum of the* $\vec{s}$*-column degrees of any* $\vec{s}$*-minimal nullspace basis of* $\mathbf{F}$ *is bounded by* $\xi = \sum \vec{s}$.

PROOF. Let $\mathbf{P}$ be a $(\mathbf{F}, \sigma, \vec{s})$-basis with high enough order σ so that $\mathbf{P} = \left[\mathbf{N}, \bar{\mathbf{N}}\right]$ contains a complete nullspace basis, $\mathbf{N}$, of $\mathbf{F}$. By Lemma 3.3 we just need σ to be greater than the $\vec{s}$-column degree of a $\vec{s}$-minimal nullspace basis of $\mathbf{F}$. Let r be the column dimension of $\bar{\mathbf{N}}$. Note that this is the same as the rank of $\mathbf{F}$. By Lemma 3.2 the sum of the $\vec{s}$-column degrees of $\mathbf{P}$ is at most $\xi + r\sigma$. By Lemma 3.1 the sum of the $\vec{s}$-column degrees of $\bar{\mathbf{N}}$ is greater than or equal to the sum of the column degrees of $\mathbf{F} \cdot \bar{\mathbf{N}}$, which is at least $r\sigma$, since every column of $\mathbf{F}\bar{\mathbf{N}}$ is nonzero and has order σ. So the sum of the $\vec{s}$-column degrees of $\mathbf{N}$ is bounded by $\xi + r\sigma - r\sigma = \xi$. □

3.2 Reducing the column dimension via order basis computation

In this subsection we look at how an order basis computation can be used to reduce the column dimension of our problem. While order basis computations were also used in (Storjohann and Villard, 2005) to reduce the column dimensions of their problems, here order basis computations are used in a more comprehensive way. In particular, Theorem 3.9 given later in this section, allows us to maintain the minimality of the bases with the use of the shifted degrees and the residuals.

We begin by computing a $(\mathbf{F}, 3s, \vec{s})$-basis $\mathbf{P}$, which can be done with a cost of $O^{\sim}(n^{\omega}s)$ using the algorithm from Giorgi et al. (2003). Note that if $\vec{s}$ is balanced, then we can compute this with a cost of $O^{\sim}(n^{\omega}\lceil \rho/n \rceil)$ using the algorithm from Zhou and Labahn (2009). We will show that at most $\frac{3m}{2}$ columns of $\mathbf{P}$ are not elements of a nullspace basis of $\mathbf{F}$.

REMARK 3.5. *Note that it is not essential to choose* $3s$ *for the order. The order can be set to* ℓs *for any constant* $\ell > 1$. *A smaller* ℓ *means less work to compute a* $(\mathbf{F}, \ell s, \vec{s})$*-basis, but also results in fewer nullspace basis elements computed and leaves more work for computing the remaining basis elements. On the other hand, a larger* ℓ *means more work is needed for order basis computation, but leaves less remaining work. It may be possible to better balance these computations with a better choice of* ℓ. *However, as we will see later, the resulting complexity given in this paper would remain the same for any* $\ell > 1$ *as long as we use the big O notation and do not care about the constant factors in the cost.*

THEOREM 3.6. *Let* $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2]$ *be a* $(\mathbf{F}, \sigma, \vec{s})$*-basis with* $\sigma > s$ *and* $\mathbf{P}_1$ *containing all columns* $\mathbf{n}$ *of* $\mathbf{P}$ *satisfying* $\mathbf{Fn} = 0$. *Then for* $\ell = \sigma/s$ *the column dimension* κ *of* $\mathbf{P}_2$ *is bounded by* $\frac{\ell m}{(\ell-1)}$.

PROOF. Any column $\mathbf{p}$ of $\mathbf{P}_2$ has order σ but also satisfies $\mathbf{Fp} \neq 0$. Thus the degree of $\mathbf{Fp}$ must be at least σ and, by Lemma 3.1, $\mathbf{p}$ must have $\vec{s}$-column degree at least σ. It follows that the sum of the $\vec{s}$-column degrees of the columns of $\mathbf{P}_2$ must satisfy $\sum \deg_{\vec{s}} \mathbf{P}_2 \geq \kappa\sigma$. From Lemma 3.2 we know that the sum of the $\vec{s}$-column degrees of the columns of $\mathbf{P}$ satisfies $\sum \deg_{\vec{s}} \mathbf{P} \leq \sum \vec{s} + m\sigma$, and hence the sum of $\vec{s}$-column degrees of the columns of $\mathbf{P}_1$ must satisfy

$$\sum \deg_{\vec{s}} \mathbf{P}_1 = \sum \deg_{\vec{s}} \mathbf{P} - \sum \deg_{\vec{s}} \mathbf{P}_2 \leq \sum \vec{s} + m\sigma - \kappa\sigma.$$

On the other hand, the lowest possible value of $\sum \deg_{\vec{s}} \mathbf{P}_1$ is $\sum_{i=1}^{n-\kappa} s_i$, the sum of the $n-\kappa$ smallest entries of $\vec{s}$ (which occurs when $\mathbf{P}_1 = [\mathbf{I}, 0]^T$). It follows that

$$\sum \vec{s} + m\sigma - \kappa\sigma \geq \sum_{i=1}^{n-\kappa} s_i,$$

or, after rearrangement,

$$m\sigma \geq \kappa\sigma - \left(\sum \vec{s} - \sum_{i=1}^{n-\kappa} s_i \right).$$

Combining this with the fact that for $\kappa \geq m$ the average of the κ largest entries of $\vec{s}$ is no more than the average of the m largest entries of $\vec{s}$, that is,

$$\left(\sum \vec{s} - \sum_{i=1}^{n-\kappa} s_i \right) / \kappa \leq s, \text{ or } \sum \vec{s} - \sum_{i=1}^{n-\kappa} s_i \leq \kappa s,$$

we get $m\sigma \geq \kappa\sigma - \kappa s$, which gives $\kappa \leq m\sigma/(\sigma - s)$ for $\sigma > s$. Substituting in $\sigma = \ell s$, we get $\kappa \leq \frac{\ell m}{(\ell-1)}$ as required. □

Let $[\mathbf{P}_1, \mathbf{P}_2] = \mathbf{P}$ with $\mathbf{P}_1$ consisting of the nullspace basis elements computed. Then the residual $\mathbf{FP} = [\mathbf{0}, \mathbf{FP}_2]$ can be used to compute the remaining nullspace basis elements. Before showing this can be correctly done, let us first make sure that the matrix multiplication $\mathbf{FP}_2$ can be done efficiently, which may not be obvious since $\mathbf{F}$, $\mathbf{P}_2$, and their product $\mathbf{FP}_2$ can all have degrees up to $\Theta(\xi)$. But we do have the sum of the column degrees of $\mathbf{F}$, that of $\mathbf{FP}_2$, and the sum of the $\vec{s}$-column degrees of $\mathbf{P}_2$ all bounded by $O(\xi)$, which means their total size are not too big but their column degrees can be quite unbalanced. We will encounter this type of multiplication again multiple times, for computing residuals and combining results. In fact, almost all of the matrices in our nullspace basis computation can have such unbalanced degrees. To efficiently multiply these matrices, we provide the following theorem, whose proof we defer until the end of this section.

THEOREM 3.7. *Let* $\mathbf{A} \in \mathbb{K}[x]^{m \times n}$, $\vec{s}$ *a shift with entries bounding the column degrees of* $\mathbf{A}$ *and* ξ, *a bound on the sum of the entries of* $\vec{s}$. *Let* $\mathbf{B} \in \mathbb{K}[x]^{n \times k}$ *with* $k \in O(m)$ *and the sum* θ *of its* $\vec{s}$*-column degrees satisfying* $\theta \in O(\xi)$. *Then we can multiply* $\mathbf{A}$ *and* $\mathbf{B}$ *with a cost of* $O^{\sim}(nm^{\omega-2}\xi)$.

With Theorem 3.7, we can now do the multiplication $\mathbf{FP}_2$ efficiently.

COROLLARY 3.8. *The multiplication of* $\mathbf{F}$ *and* $\mathbf{P}_2$ *can be done with a cost of* $O^{\sim}\left(nm^{\omega-2}\xi\right)$.

PROOF. Since $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2]$ is a $(\mathbf{F}, 3s, \vec{s})$-basis, we have from Lemma 3.2 that the sum of the $\vec{s}$-column degrees of $\mathbf{P}_2$ satisfies $\sum \deg_{\vec{s}} \mathbf{P}_2 \leq 3sm + \xi \leq 4\xi$. Hence Theorem 3.7 applies. □

It remains to show that the residual $\mathbf{FP}_2$ can be used to compute the remaining nullspace basis elements.

THEOREM 3.9. *Let* $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2]$ *be a* $(\mathbf{F}, \sigma, \vec{s})$*-basis such that* $\mathbf{P}_1$ *consists of all the nullspace basis elements of* $\mathbf{F}$ *in* $\mathbf{P}$. *Let* $\vec{b} = [\vec{b}_1, \vec{b}_2]$ *be the* $\vec{s}$*-column degrees of* $\mathbf{P}$, *where* $\vec{b}_1, \vec{b}_2$ *are the* $\vec{s}$*-column degrees of* $\mathbf{P}_1$, $\mathbf{P}_2$ *respectively. Let* $\mathbf{Q}$ *be a* $\vec{b}_2$*-minimal nullspace basis of* $\mathbf{FP}_2$ *with* $\vec{b}_2$*-column degrees* $\vec{b}_2'$. *Then* $[\mathbf{P}_1, \mathbf{P}_2\mathbf{Q}]$ *is a* $\vec{s}$*-minimal nullspace basis of* $\mathbf{F}$ *with* $\vec{s}$*-column degrees* $[\vec{b}_1, \vec{b}_2']$.

PROOF. Let $\mathbf{Q}' = \mathrm{diag}([I, \mathbf{Q}])$, where the dimension of the identity matrix I matches that of $\mathbf{P}_1$. Then $\mathbf{Q}'$ is a $\vec{b}$-minimal nullspace basis of $\mathbf{FP}$ since $\mathbf{FPQ}' = [\mathbf{FP}_1, \mathbf{FP}_2\mathbf{Q}] = 0$. It follows that $\mathbf{PQ}' = [\mathbf{P}_1, \mathbf{P}_2\mathbf{Q}]$ is a nullspace basis of $\mathbf{F}$. We now show that $\mathbf{PQ}'$ is $\vec{s}$-column reduced and has

$\vec{s}$-column degrees $[\vec{b}_1, \vec{b}'_2]$, or equivalently, $x^{\vec{s}}\mathbf{P}\mathbf{Q}'$ is column reduced and has column degrees $[\vec{b}_1, \vec{b}'_2]$. Notice that $x^{\vec{s}}\mathbf{P}$ has column degrees $[\vec{b}_1, \vec{b}_2]$ and a full rank leading column coefficient matrix P. Hence $x^{\vec{s}}\mathbf{P}x^{-[\vec{b}_1,\vec{b}_2]}$ has column degrees $[0, \ldots 0]$. (If one is concerned about the entries not being polynomials, one can simply multiply the matrix by x^{ξ} to shift the degrees up.) Similarly, $x^{\vec{b}_2}\mathbf{Q}x^{-\vec{b}'_2}$ has column degrees $[0, \ldots, 0]$, and so $x^{[\vec{b}_1,\vec{b}_2]}\mathbf{Q}'x^{-[\vec{b}_1,\vec{b}'_2]}$ also has column degrees $[0, \ldots, 0]$ and a full rank leading column coefficient matrix Q'. Putting these together, we see that $x^{\vec{s}}\mathbf{P}x^{-[\vec{b}_1,\vec{b}_2]}x^{[\vec{b}_1,\vec{b}_2]}\mathbf{Q}'x^{-[\vec{b}_1,\vec{b}'_2]} = x^{\vec{s}}\mathbf{P}\mathbf{Q}'x^{-[\vec{b}_1,\vec{b}'_2]}$ has column degrees $[0, \ldots, 0]$ and a full rank leading column coefficient matrix PQ'. It follows that $x^{\vec{s}}\mathbf{P}\mathbf{Q}'$ has column degrees $[\vec{b}_1, \vec{b}'_2]$, or equivalently, the $\vec{s}$-column degrees of $\mathbf{P}\mathbf{Q}'$ is $[\vec{b}_1, \vec{b}'_2]$.

It remains to show that any $\mathbf{n}$ satisfying $\mathbf{F}\mathbf{n} = 0$ must be a linear combination of the columns of $\mathbf{P}\mathbf{Q}'$. Since $\mathbf{n} \in \langle(\mathbf{F}, \sigma)\rangle$, it is generated by the $(\mathbf{F}, \sigma)$-basis $\mathbf{P}$, that is, $\mathbf{n} = \mathbf{P}\mathbf{a}$ with $\mathbf{a} = \mathbf{P}^{-1}\mathbf{n} \in \mathbb{K}[x]^n$. Also, $\mathbf{F}\mathbf{n} = 0$ implies $\mathbf{F}\mathbf{P}\mathbf{a} = 0$, hence $\mathbf{a} = \mathbf{Q}'\mathbf{b}$ for some vector $\mathbf{b}$ as $\mathbf{Q}'$ is a nullspace basis of $\mathbf{F}\mathbf{P}$. We now have $\mathbf{n} = \mathbf{P}\mathbf{Q}'\mathbf{b}$ as required. □

EXAMPLE 3.10. *Let us look at an example of computing nullspace basis using Theorem 3.9. Let* $\mathbf{F}$ *be given by*

$$\begin{bmatrix} x+x^2+x^3 & 1+x & 0 & 1+x \\ 1+x^2+x^3 & x+x^2+x^3 & x+x^2 & x^3 \end{bmatrix} \in \mathbb{Z}_2[x]^{2\times 4}.$$

Let $\sigma = 3$, $\vec{s} = [3,3,3,3]$. *We first compute a* $(\mathbf{F}, \sigma, \vec{s})$*-basis*

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & x^2 & x \\ 1 & 0 & 0 & x^2 \\ 1 & x^2 & x+x^2 & 1+x \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

with the $\vec{s}$*-column degrees* $\vec{b} = [3,5,5,5]$ *and the residual*

$$\mathbf{F}\mathbf{P} = \begin{bmatrix} 0 & 0 & x^3+x^4+x^5 & x^4 \\ 0 & x^3+x^4 & x^5 & x^3+x^5 \end{bmatrix}.$$

Thus $\mathbf{P}_1 = [0,1,1,1]^T$, *with* $\vec{s}$*-column degree* 3, *is the only nullspace basis element computed. Let* $\mathbf{P}_2$ *contain the remaining columns of* $\mathbf{P}$ *and* $\vec{b}_2 = [5,5,5]$ *be its* $\vec{s}$*-column degrees. Next we compute a* $\vec{b}_2$*-minimal nullspace basis of* $\mathbf{F}\mathbf{P}_2$

$$\mathbf{Q} = [1+x+x^4,\ x+x^2,\ 1+x^3]^T$$

which has $\vec{b}_2$*-column degree 9. Then*

$$[\mathbf{P}_1, \mathbf{P}_2\mathbf{Q}] = \begin{bmatrix} 0 & x+x^3 \\ 1 & x^2+x^5 \\ 1 & 1+x+x^6 \\ 1 & 0 \end{bmatrix}$$

is a complete $\vec{s}$*-minimal nullspace basis of* $\mathbf{F}$ *with* $\vec{s}$*-column degrees* $[3,9]$.

Theorem 3.9 shows that the remaining $\vec{s}$-minimal nullspace basis elements $\mathbf{P}_2\mathbf{Q}$ can be correctly computed from the residual $\mathbf{F}\mathbf{P}_2$. Before discussing the computation of a $\vec{b}_2$-minimal nullspace basis $\mathbf{Q}$ of $\mathbf{F}\mathbf{P}_2$, let us first note that the multiplication $\mathbf{P}_2\mathbf{Q}$ can be done efficiently, which again follows from Theorem 3.7.

LEMMA 3.11. *The multiplication of* $\mathbf{P}_2$ *and* $\mathbf{Q}$ *can be done with a cost of* $O^{\sim}\left(nm^{\omega-2}\xi\right)$.

PROOF. Note that the dimension of $\mathbf{P}_2$ is $n \times O(m)$ from Theorem 3.6 and the dimension of $\mathbf{Q}$ is $O(m) \times O(m)$. The column degrees of $\mathbf{P}_2$ are bounded by the $\vec{s}$-column degrees $\vec{b}_2$ of $\mathbf{P}_2$ since $\vec{s}$ is non-negative. Also recall that $\sum \vec{b}_2 \leq 4\xi$ from the proof of Corollary 3.8. By Lemma 3.1 the column degrees of $\mathbf{F}\mathbf{P}_2$ are bounded by the $\vec{s}$-column degrees $\vec{b}_2$ of $\mathbf{P}_2$. By Theorem 3.4, the sum of the $\vec{b}_2$-column degrees of $\mathbf{Q}$ is also bounded by $\sum \vec{b}_2 \leq 4\xi$. Now if we separate $\mathbf{P}_2$ to n/m blocks rows each with no more than m rows, Theorem 3.7 can be used to multiply each block row with $\mathbf{Q}$. Each multiplication involves matrices of dimension $O(m) \times O(m)$. In addition, both the sum of the column degrees of $\mathbf{P}_2$ and the sum of the $\vec{b}_2$-column degrees of $\mathbf{Q}$ are bounded by 4ξ. So each multiplication costs $O^{\sim}(m^{\omega-1}\xi)$. Hence doing this for all n/m block rows costs $O^{\sim}\left(nm^{\omega-2}\xi\right)$. □

3.3 Reducing the degrees

Our next task is computing a $\vec{b}_2$-minimal nullspace basis of the residual $\mathbf{F}\mathbf{P}_2$. It is useful to note that the lower degree terms of $\mathbf{F}\mathbf{P}_2$ are zero since it has order σ. Hence we can use $\mathbf{G} = \mathbf{F}\mathbf{P}_2/x^{\sigma}$ instead to compute the remaining basis elements. In the following, we show that just like the original input matrix $\mathbf{F}$, this new input matrix $\mathbf{G}$ has column degrees bounded by the corresponding entries of $\vec{s}$.

LEMMA 3.12. *If an* $(\mathbf{F}, \sigma, \vec{s})$*-basis has columns arranged in increasing* $\vec{s}$*-column degrees with* $\vec{s}$*-column degrees* $\vec{b}$*, then the entries of* $\vec{b} - [\sigma, \ldots, \sigma] = [b_1 - \sigma, \ldots, b_n - \sigma]$ *are bounded component-wise by* $\vec{s}$.

PROOF. A $(\mathbf{F}, 0, \vec{s})$-basis of order 0 has $\vec{s}$-column degrees given by $\vec{s}$. For each order increase, any column of the basis has its $\vec{s}$-column degree increases by at most one, which occurs when its order is increased by multiplying the column by x. Hence at order σ, the $\vec{s}$-column degree increase for each column is at most σ. □

COROLLARY 3.13. *The column degrees of* $\mathbf{F}\mathbf{P}/x^{\sigma}$ *are bounded component-wise by* $\vec{s}$.

PROOF. From Lemma 3.1, the column degrees of $\mathbf{F}\mathbf{P}$ are bounded component-wise by $\vec{b}$, the $\vec{s}$-column degrees of $\mathbf{P}$. Hence the column degrees of $\mathbf{F}\mathbf{P}/x^{\sigma}$ are bounded component-wise by $\vec{b} - [\sigma, \ldots, \sigma]$. The result then follows from Lemma 3.12. □

From Corollary 3.13, the column degrees of $\mathbf{F}\mathbf{P}_2/x^{\sigma}$ are bounded by the entries of the corresponding subset $\vec{t}$ of $\vec{b} - [\sigma, \ldots, \sigma]$, which is in turn bounded by the entries of the corresponding subset of $\vec{s}$.

EXAMPLE 3.14. *From Example 3.10, note that instead of using the residual*

$$\mathbf{F}\mathbf{P}_2 = \begin{bmatrix} 0 & x^3+x^4+x^5 & x^4 \\ x^3+x^4 & x^5 & x^3+x^5 \end{bmatrix}$$

to compute a $[5,5,5]$*-minimal nullspace basis of* $\mathbf{F}$, *we can instead use*

$$\mathbf{G} = \mathbf{F}\mathbf{P}_2/x^3 = \begin{bmatrix} 0 & 1+x+x^2 & x \\ 1+x & x^2 & 1+x^2 \end{bmatrix}$$

to compute a $[2,2,2]$*-minimal nullspace basis of* $\mathbf{G}$. *The column degrees of* $\mathbf{G}$ *are bounded by the new shift* $[2,2,2]$, *which is in turn bounded by the corresponding entries* $[3,3,3]$ *of* $\vec{s}$.

At this point, using Theorem 3.9 and Corollary 3.13, the problem is reduced to computing a $\vec{t}$-minimal nullspace basis of $\mathbf{G} = \mathbf{F}\mathbf{P}_2/x^{3s}$, which still has row dimension m. But its column dimension is now bounded by $3m/2$. Also notice that as in the original problem, the column degrees of the new input matrix $\mathbf{G}$ are bounded by the corresponding entries of the new shift $\vec{t}$. In addition, as the new shift $\vec{t}$ is bounded component-wise by a subset of the old shift $\vec{s}$, the new problem is no more difficult than the original problem.

3.4 Reducing the row dimension

We now turn to the new problem of computing a $\vec{t}$-minimal nullspace basis of $\mathbf{G}$. Let

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix}$$

with $\mathbf{G}_1$ having $\lfloor m/2 \rfloor$ rows and $\mathbf{G}_2$ having $\lceil m/2 \rceil$ rows. If we compute a $\vec{t}$-minimal nullspace basis $\mathbf{N}_1$ of $\mathbf{G}_1$, where $\mathbf{N}_1$ has $\vec{t}$-column degrees $\vec{u}$, then compute a $\vec{u}$-minimal nullspace basis $\mathbf{N}_2$ of $\mathbf{G}_2\mathbf{N}_1$, then the next theorem shows that $\mathbf{N}_1\mathbf{N}_2$ is a $\vec{t}$-minimal nullspace basis of $\mathbf{G}$.

THEOREM 3.15. *Let $\mathbf{G} = \left[\mathbf{G}_1^T, \mathbf{G}_2^T\right]^T \in \mathbb{K}[x]^{m\times n}$ and $\vec{t}$ a shift vector. If $\mathbf{N}_1$ is a $\vec{t}$-minimal nullspace basis of $\mathbf{G}_1$ with $\vec{t}$-column degrees $\vec{u}$, and $\mathbf{N}_2$ is a $\vec{u}$-minimal nullspace basis of $\mathbf{G}_2\mathbf{N}_1$ with $\vec{u}$-column degrees $\vec{v}$, then $\mathbf{N}_1\mathbf{N}_2$ is a $\vec{t}$-minimal nullspace basis of $\mathbf{G}$ with $\vec{t}$-column degrees $\vec{v}$.*

PROOF. The proof is very similar to the proof of Theorem 3.9. It is clear that $\mathbf{G}\mathbf{N}_1\mathbf{N}_2 = 0$ hence $\mathbf{N}_1\mathbf{N}_2$ is a nullspace basis of $\mathbf{G}$. We now show that $\mathbf{N}_1\mathbf{N}_2$ is $\vec{t}$-column reduced and has $\vec{t}$-column degrees $\vec{v}$, or equivalently, $x^{\vec{t}}\mathbf{N}_1\mathbf{N}_2$ is column reduced. Notice that $x^{\vec{t}}\mathbf{N}_1$ has column degrees $\vec{u}$ and a full rank leading column coefficient matrix N_1. Hence $x^{\vec{t}}\mathbf{N}_1 x^{-\vec{u}}$ has column degrees $[0,\ldots,0]$. Again, if one is concerned about the entries not being polynomials, one can simply multiply the matrix by x^{ξ} to shift the degrees up. Similarly, $x^{\vec{u}}\mathbf{N}_2 x^{\vec{v}}$ has column degrees $[0,\ldots,0]$ and a full rank leading column coefficient matrix N_2. Putting them together, $x^{\vec{t}}\mathbf{N}_1 x^{-\vec{u}} x^{\vec{u}}\mathbf{N}_2 x^{-\vec{v}} = x^{\vec{t}}\mathbf{N}_1\mathbf{N}_2 x^{-\vec{v}}$ has column degrees $[0,\ldots,0]$ and a full rank leading column coefficient matrix N_1N_2. It follows that $x^{\vec{t}}\mathbf{N}_1\mathbf{N}_2$ has column degrees $\vec{v}$, or equivalently, the $\vec{t}$-column degrees of $\mathbf{N}_1\mathbf{N}_2$ is $\vec{v}$.

It remains to show that any $\mathbf{n}$ satisfying $\mathbf{G}\mathbf{n} = 0$ must be a linear combination of the columns of $\mathbf{N}_1\mathbf{N}_2$. First notice that $\mathbf{n} = \mathbf{N}_1\mathbf{a}$ for some polynomial vector $\mathbf{a}$ since $\mathbf{N}_1$ is a nullspace basis of $\mathbf{G}_1$. Also, $\mathbf{G}\mathbf{n} = 0$ implies that $\mathbf{G}_2\mathbf{N}_1\mathbf{a} = 0$, hence $\mathbf{a} = \mathbf{N}_2\mathbf{b}$ for some vector $\mathbf{b}$ as $\mathbf{N}_2$ is a nullspace basis of $\mathbf{G}_2\mathbf{N}_1$. We now have $\mathbf{n} = \mathbf{N}_1\mathbf{N}_2\mathbf{b}$ as required. □

EXAMPLE 3.16. *Let us compute a $\vec{t}$-minimal nullspace basis of*

$$\mathbf{G} = \begin{bmatrix} 0 & 1+x+x^2 & x \\ 1+x & x^2 & 1+x^2 \end{bmatrix}$$

from Example 3.14, where $\vec{t} = [2,2,2]$. Then

$$\mathbf{G}_1 = \begin{bmatrix} 0 & 1+x+x^2 & x \end{bmatrix} \text{ and } \mathbf{G}_2 = \begin{bmatrix} 1+x & x^2 & 1+x^2 \end{bmatrix}.$$

We first compute a $\vec{t}$-minimal nullspace basis $\mathbf{N}_1$ of $\mathbf{G}_1$:

$$\mathbf{N}_1 = \begin{bmatrix} 1 & 0 \\ 0 & x \\ 0 & 1+x+x^2 \end{bmatrix}$$

with its $\vec{t}$-column degrees $\vec{u} = [2,4]$. Next, we compute a $\vec{u}$-minimal nullspace basis $\mathbf{N}_2$ of $\mathbf{G}_2\mathbf{N}_1 = \begin{bmatrix} 1+x & 1+x+x^4 \end{bmatrix}$:

$$\mathbf{N}_2 = [1+x+x^4,\ 1+x]^T.$$

Then

$$\mathbf{N}_1\mathbf{N}_2 = [1+x+x^4,\ x+x^2,\ 1+x^3]^T$$

is a $\vec{t}$-minimal nullspace basis of $\mathbf{G}$.

While Theorem 3.9 allows us to compute nullspace bases by columns, which then reduces the column dimensions, Theorem 3.15 shows that that the nullspace bases can also be computed by rows, which then reduces the row dimensions. Again, we need to check that these computations can be done efficiently. In the following, Lemma 3.17 and Lemma 3.18 show that the multiplication $\mathbf{G}_2\mathbf{N}_1$ and the multiplication $\mathbf{N}_1\mathbf{N}_2$ can be done efficiently, which are again consequences of Theorem 3.7.

LEMMA 3.17. *The multiplication of $\mathbf{G}_2$ and $\mathbf{N}_1$ can be done with a cost of $O^{\sim}(m^{\omega-1}\xi)$.*

PROOF. Theorem 3.7 applies directly here. □

LEMMA 3.18. *The multiplication of $\mathbf{N}_1$ and $\mathbf{N}_2$ can be done with a cost of $O^{\sim}(m^{\omega-1}\xi)$.*

PROOF. Theorem 3.7 applies because the sum of the column degrees of $\mathbf{N}_1$ is bounded by the sum of the $\vec{t}$-column degrees of $\mathbf{N}_1$, which is $\sum \vec{u} \le \xi$, and by Theorem 3.4 the sum of $\vec{u}$-column degrees of $\mathbf{N}_2$ is also bounded by ξ. □

3.5 Recursive computation

The computation of $\mathbf{N}_1$ and $\mathbf{N}_2$ is identical to the original problem, only the dimension has decreased. For computing $\mathbf{N}_1$, the dimension of the input matrix $\mathbf{G}_1$ is bounded by $\lfloor m/2 \rfloor \times (3m/2)$. For computing $\mathbf{N}_2$, the dimension of input matrix $\mathbf{G}_2\mathbf{N}_1$ is bounded by $\lceil m/2 \rceil \times (3m/2)$. The column degrees of $\mathbf{G}_1$ are bounded by the entries of $\vec{t}$, with $\sum \vec{t} \le \xi$. Similarly, the column degrees of $\mathbf{G}_2\mathbf{N}_1$ are bounded by the entries of $\vec{u}$, with $\sum \vec{u} \le \xi$. Hence, the same computation process can be repeated on these two smaller problems. This gives a recursive algorithm, shown in Algorithm 1.

Before analyzing the computational complexity of Algorithm 1 in the following section, we provide a proof of Theorem 3.7, which is needed to efficiently multiply matrices with unbalanced degrees in the algorithm.

3.6 Proof of Theorem 3.7

In this subsection we give a proof of Theorem 3.7.

PROOF. Recall that $\vec{s}$ is a shift with entries ordered in terms of increasing values and ξ is a bound on the sum of the entries of $\vec{s}$. We wish to determine the cost of multiplying the two polynomials matrices $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ and $\mathbf{B} \in \mathbb{K}[x]^{n\times k}$ where $\mathbf{A}$ has column degrees bounded by $\vec{s}$ and where $k \in O(m)$ and the sum θ of its $\vec{s}$-column degrees satisfies $\theta \in O(\xi)$. The goal is to show that these polynomial matrices can be multiplied with a cost of $O^{\sim}(nm^{\omega-2}\xi)$.

For simplicity we assume m is a power of 2, something which can be achieved by appending zero rows to $\mathbf{F}$. We divide the matrix $\mathbf{B}$ into $\log m$ column blocks according to the $\vec{s}$-column degrees of its columns. Let

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(\log m)} & \mathbf{B}^{(\log m-1)} & \cdots & \mathbf{B}^{(2)} & \mathbf{B}^{(1)} \end{bmatrix},$$

Algorithm 1 MinimalNullspaceBasis $(\mathbf{F}, \vec{s})$: Compute a $\vec{s}$-Minimal Nullspace Basis

Input: $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$, $\vec{s} = [s_1, \ldots, s_n] \in \mathbb{Z}^n$ with entries arranged in non-decreasing order and bounding the corresponding column degrees of $\mathbf{F}$.

Output: A $\vec{s}$-minimal nullspace basis of $\mathbf{F}$ and its $\vec{s}$-column degrees $\vec{s}'$.

1: $\xi := \sum_{i=1}^{n} s_i$; $\rho := \sum_{i=n-m+1}^{n} s_i$; $s := \rho/m$;
2: $\left[\mathbf{P}, \vec{b}\right] :=$ orderBasis $(\mathbf{F}, 3s, \vec{s})$, a $(\mathbf{F}, 3s, \vec{s})$-basis with the columns of $\mathbf{P}$ and the entries of its $\vec{s}$-column degrees $\vec{b}$ arranged so that the entries of $\vec{b}$ are in non-decreasing order;
3: $[\mathbf{P}_1, \mathbf{P}_2] := \mathbf{P}$ where $\mathbf{P}_1$ consists of all columns $\mathbf{p}$ of $\mathbf{P}$ satisfying $\mathbf{Fp} = 0$;
4: **if** $m = 1$ **then**
5: **return** $\mathbf{P}_1, \deg_{\vec{s}} \mathbf{P}_1$
6: **else**
7: $\vec{t} := \deg_{\vec{s}} \mathbf{P}_2 - [3s, 3s, \ldots, 3s]$;
8: $\mathbf{G} := \mathbf{F}\mathbf{P}_2 / x^{3s}$;
9: $\left[\mathbf{G}_1^T, \mathbf{G}_2^T\right]^T := \mathbf{G}$, with $\mathbf{G}_1$ having $\lfloor m/2 \rfloor$ rows and $\mathbf{G}_2$ having $\lceil m/2 \rceil$ rows;
10: $[\mathbf{N}_1, \vec{u}] :=$ MinimalNullspaceBasis $\left(\mathbf{G}_1, \vec{t}\right)$;
11: $[\mathbf{N}_2, \vec{v}] :=$ MinimalNullspaceBasis $\left(\mathbf{G}_2\mathbf{N}_1, \vec{u}\right)$;
12: $\mathbf{Q} := \mathbf{N}_1\mathbf{N}_2$;
13: **return** $[\mathbf{P}_1, \mathbf{P}_2\mathbf{Q}], [\deg_{\vec{s}} \mathbf{P}_1, \vec{v}]$
14: **end if**

with $\mathbf{B}^{(\log m)}$, $\mathbf{B}^{(\log m-1)}$, $\mathbf{B}^{(\log m-2)}$, ... , $\mathbf{B}^{(2)}$, $\mathbf{B}^{(1)}$ having $\vec{s}$-column degrees in the range $[0, 2\xi/m]$, $(2\xi/m, 4\xi/m]$, $(4\xi/m, 8\xi/m]$, ...,$(\xi/4, \xi/2]$, $(\xi/2, \theta]$, respectively. We will multiply $\mathbf{A}$ with each $\mathbf{B}^{(i)}$ separately.

We also divide the matrix $\mathbf{A}$ into $\log m$ column blocks and each matrix $\mathbf{B}^{(i)}$ into $\log m$ row blocks according to the size of the corresponding entries in $\vec{s}$. Set

$$\begin{aligned}
\vec{s} &= \begin{bmatrix} \vec{s}_{\log m} & \vec{s}_{\log m-1} & \cdots & \vec{s}_1 \end{bmatrix} \\
\mathbf{A} &= \begin{bmatrix} \mathbf{A}_{\log m} & \mathbf{A}_{\log m-1} & \cdots & \mathbf{A}_1 \end{bmatrix} \\
\mathbf{B} &= \begin{bmatrix} \mathbf{B}^{(\log m)} & \mathbf{B}^{(\log m-1)} & \cdots & \mathbf{B}^{(1)} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{B}^{(\log m)}_{\log m} & \mathbf{B}^{(\log m-1)}_{\log m} & \cdots & \mathbf{B}^{(1)}_{\log m} \\ \vdots & & & \vdots \\ \mathbf{B}^{(\log m)}_{1} & \mathbf{B}^{(\log m-1)}_{1} & \cdots & \mathbf{B}^{(1)}_{1} \end{bmatrix}
\end{aligned}$$

with $\vec{s}_{\log m}, \vec{s}_{\log m-1}, \ldots, \vec{s}_1$ having entries in the range $[0, 2\xi/m]$, $(2\xi/m, 4\xi/m]$, $(4\xi/m, 8\xi/m]$, ..., $(\xi/2, \xi]$ respectively. Also the column dimension of $\mathbf{A}_j$ and the row dimension of $\mathbf{B}^{(i)}_j$ match that of $\vec{s}_j$ for j from 1 to $\log m$.

Notice that $\mathbf{B}^{(i)}_{(j)}$ for $i > j$ must be zero. Otherwise, as $\vec{s}_j > \xi/2^j \geq \xi/2^{i-1}$, the $\vec{s}$-column degree of $\mathbf{B}^{(i)}$ would exceed $\xi/2^{i-1}$, a contradiction since by definition the $\vec{s}$-column degree of $\mathbf{B}^{(i)}$ is bounded by $\xi/2^{i-1}$ when $i > 1$. So $\mathbf{B}$ in fact has a block triangular shape

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}^{(\log m)}_{\log m} & \mathbf{B}^{(\log m-1)}_{\log m} & \cdots & \mathbf{B}^{(1)}_{\log m} \\ & \mathbf{B}^{(\log m-1)}_{\log m-1} & & \vdots \\ & & \ddots & \\ & & & \mathbf{B}^{(1)}_{1} \end{bmatrix}$$

(while remembering that the blocks have varying sizes).

First consider the multiplication

$$\mathbf{A}\mathbf{B}^{(1)} = \begin{bmatrix} \mathbf{A}_{\log m} & \cdots & \mathbf{A}_1 \end{bmatrix} \begin{bmatrix} \mathbf{B}^{(1)}_{\log m} & \cdots & \mathbf{B}^{(1)}_{1} \end{bmatrix}^T.$$

Note that there are $O(1)$ columns in $\mathbf{B}^{(1)}$ since $\theta \in O(\xi)$. We do this in $\log m$ steps. At step j for j from 1 to $\log m$ we multiply $\mathbf{A}_j$ and $\mathbf{B}^{(1)}_j$. The column dimension of $\mathbf{A}_j$, which is the same as the row dimension of $\mathbf{B}^{(1)}_j$, is less than 2^j. The degree of $\mathbf{B}^{(1)}_j$ is $O(\xi)$. To use fast multiplication, we expand $\mathbf{B}^{(1)}_j$ to a matrix $\bar{\mathbf{B}}^{(1)}_j$ with degree less than $\delta \in \Theta(\xi/2^j)$ and column dimension $q \in O(2^j)$ as follows. Write

$$\mathbf{B}^{(1)}_j = \mathbf{B}^{(1)}_{j,0} + \mathbf{B}^{(1)}_{j,1}x^{\delta} + \cdots + \mathbf{B}^{(1)}_{j,q-1}x^{\delta(q-1)} = \sum_{k=0}^{q-1} \mathbf{B}^{(1)}_{j,k} x^{\delta k}$$

with each $\mathbf{B}^{(1)}_{j,k}$ having degree less than δ. Set

$$\bar{\mathbf{B}}^{(1)}_j = \left[\mathbf{B}^{(1)}_{j,0}, \mathbf{B}^{(1)}_{j,1}, \ldots, \mathbf{B}^{(1)}_{j,q-1}\right].$$

We can then multiply $\mathbf{A}_j$, which has dimension $m \times O(2^j)$ for $j < \log m$, and $\bar{\mathbf{B}}^{(1)}_j$, which has dimension $O(2^j) \times O(2^j)$ for $j < \log m$, with a cost of

$$\begin{aligned}
O^{\sim}\left((m/2^j)\left(2^j\right)^{\omega}\xi/2^j\right) &= O^{\sim}\left(\left(2^j\right)^{\omega-2} m\xi\right). \\
&\subset O^{\sim}\left(m^{\omega-1}\xi\right) \subset O^{\sim}(nm^{\omega-2}\xi)
\end{aligned}$$

For $j = \log m$, $\mathbf{A}_j$ has dimension $m \times O(n)$, $\bar{\mathbf{B}}^{(1)}_j$ has dimension $O(n) \times O(m)$, and their degrees are $O(\xi/m)$. Hence they can be multiplied with a cost of $O^{\sim}((n/m)m^{\omega}(\xi/m)) = O^{\sim}\left(nm^{\omega-2}\xi\right)$. Adding up the columns of $\mathbf{A}_j\bar{\mathbf{B}}^{(1)}_j$ gives $\mathbf{A}_j\mathbf{B}^{(1)}_j$ and costs $O(m\xi)$. Therefore, multiplying $\mathbf{A}$ and $\mathbf{B}^{(1)}$ over $\log(m)$ steps costs $O^{\sim}\left(nm^{\omega-2}\xi\right)$.

Next we multiply $\mathbf{A}$ with $\mathbf{B}^{(2)}$. We proceed in the same way as before, but notice that $\mathbf{A}_1\mathbf{B}^{(2)}_1$ is no longer needed since $\mathbf{B}^{(2)}_1 = 0$. Multiplying $\mathbf{A}$ and $\mathbf{B}^{(2)}$ also costs $O^{\sim}\left(nm^{\omega-2}\xi\right)$.

Continuing to doing this, gives a costs of $O^{\sim}\left(nm^{\omega-2}\xi\right)$ to multiply $\mathbf{A}$ with the columns $\mathbf{B}^{(i)}$ for i from 1 to $\log m$. As before, we recall that $\mathbf{B}^{(i)}_{(j)} = 0$ for $j > i$. The overall cost for i from 1 to $\log m$ is therefore $O^{\sim}\left(nm^{\omega-2}\xi\right)$ to multiply $\mathbf{A}$ and $\mathbf{B}$. □

4. COMPUTATIONAL COMPLEXITY

For the cost analysis we first consider the case where the column dimension n is not much bigger than the row dimension m.

THEOREM 4.1. *If $n \in O(m)$, then the cost of Algorithm 1 is $O^{\sim}\left(m^{\omega-1}\xi\right) = O^{\sim}\left(m^{\omega-1}\rho\right)$ field operations.*

PROOF. We may assume m is a power of 2, which can be achieved by appending zero rows to $\mathbf{F}$. Note that $\rho \in \Theta(\xi)$ when $n \in O(m)$. Then the order basis computation at line 2 costs $O^{\sim}(n^{\omega}s) = O^{\sim}\left(m^{\omega-1}\rho\right)$. The multiplications at line 8 and line 13 cost $O^{\sim}\left(nm^{\omega-2}\xi\right) = O^{\sim}\left(m^{\omega-1}\xi\right)$. The remaining operations including multiplications at line 11 and line 12 cost $O^{\sim}\left(m^{\omega-1}\xi\right)$. Let $g(m, \xi)$ be the computational cost of the original problem. Then we have the recurrence relation

$$g(m, \xi) \in O^{\sim}(m^{\omega-1}\xi) + g(m/2, \xi) + g(m/2, \xi),$$

with the base case $g(1,\xi) \in O^{\sim}(\xi)$, the cost of just an order basis computation at $m = 1$. This gives $g(m,\xi) \in O^{\sim}(m^{\omega-1}\xi)$ field operations as the cost of the algorithm. □

We now consider the general case where the column dimension n can be much bigger than the row dimension m.

THEOREM 4.2. *Algorithm 1 costs $O^{\sim}(n^{\omega}s)$ field operations in general.*

PROOF. The order basis computation at line 2 costs $O^{\sim}(n^{\omega}s)$ in general, which dominates the cost of other operations. The problem is then reduced to one where we have column dimension $O(m)$, which is handled by Theorem 4.1 with a cost of $O^{\sim}\left(m^{\omega-1}\xi\right) \in O^{\sim}(n^{\omega}s)$. □

When we have the important special case where the shift $\vec{s} = [s,\ldots,s]$ is uniform then Algorithm 1 has a lower cost. Indeed we notice that the order basis computation at line 2 costs $O^{\sim}(n^{\omega}\lceil ms/n\rceil)$ using the algorithm from Zhou and Labahn (2009). In addition, the multiplication of $\mathbf{F}$ and $\mathbf{P}_2$ at line 8 and the multiplication of $\mathbf{P}_2$ and $\mathbf{Q}$ at line 13 both cost $O^{\sim}\left(nm^{\omega-1}s\right)$ as shown in Lemma 4.3 and Lemma 4.4.

LEMMA 4.3. *If the degree of $\mathbf{F}$ is bounded by s, then the multiplication of $\mathbf{F}$ and $\mathbf{P}_2$ at line 8 costs $O^{\sim}\left(nm^{\omega-1}s\right)$.*

PROOF. Since $\mathbf{P}_2$ is a part of a $(\mathbf{F}, 3s, \vec{s})$-basis, its degree is bounded by $3s$. It has dimension $n \times O(m)$ from Theorem 3.6. Multiplying $\mathbf{F}$ and $\mathbf{P}_2$ therefore costs $(n/m)O^{\sim}(m^{\omega}s) = O^{\sim}\left(nm^{\omega-1}s\right)$. □

LEMMA 4.4. *If $\mathbf{F}$ has degree s, then the multiplication of $\mathbf{P}_2$ and $\mathbf{Q}$ at line 13 costs $O^{\sim}\left(nm^{\omega-1}s\right)$.*

PROOF. First note that the dimension of $\mathbf{Q}$ is $O(m) \times O(m)$ since it is a $\vec{t}$-minimal nullspace basis of $\mathbf{G} = \mathbf{F}\mathbf{P}_2/x^{3s}$, which has dimension $m \times O(m)$. In addition, by Theorem 3.4, the sum of the $\vec{t}$-column degrees of $\mathbf{Q}$ is bounded by $\sum \vec{t}$, which is bounded by $O(ms)$ since $\vec{t}$ has $O(m)$ entries all bounded by s.

Now Theorem 3.7 and its proof still work. The current situation is even simpler as we do not need to subdivide the columns of $\mathbf{P}_2$, which has degree bounded by $3s$ and dimension $n \times O(m)$. We just need to separate the columns of $\mathbf{Q}$ to $O(\log m)$ groups with degree ranges $[0,2s]$, $(2s,4s]$, $(4s,8s]$, $\ldots$, and multiply $\mathbf{P}_2$ with each group in the same way as in Theorem 3.7, with each of these $O(\log m)$ multiplications costs $(n/m)O^{\sim}(m^{\omega}s) = O^{\sim}\left(nm^{\omega-1}s\right)$. □

THEOREM 4.5. *If $\vec{s} = [s,\ldots,s]$ is uniform, then Algorithm 1 costs $O^{\sim}(n^{\omega}\lceil ms/n\rceil)$.*

PROOF. After the initial order basis computation, which costs $O^{\sim}(n^{\omega}\lceil ms/n\rceil)$, and the multiplication of $\mathbf{F}$ and $\mathbf{P}_2$, which costs $O^{\sim}\left(nm^{\omega-1}s\right)$ from Lemma 4.3, the column dimension is reduced to $O(m)$, allowing Theorem 4.1 to apply for computing a $\vec{t}$-minimal nullspace basis of $\mathbf{F}\mathbf{P}_2/x^{3s}$. Hence the remaining work costs $O^{\sim}(m^{\omega}s)$. The overall cost is therefore dominated by the cost $O^{\sim}(n^{\omega}\lceil ms/n\rceil)$ of the initial order basis computation. □

COROLLARY 4.6. *If the input matrix $\mathbf{F}$ has degree d, then a minimal nullspace basis of $\mathbf{F}$ can be computed with a cost of $O^{\sim}(n^{\omega}\lceil md/n\rceil)$.*

PROOF. We can just set the shift $\vec{s}$ to $[d,\ldots,d]$ and apply Theorem 4.5. □

5. CONCLUSION

In this paper we have presented a fast, deterministic procedure for computing a minimal nullspace basis of a polynomial matrix. For a number of extensions and applications of this work, and for other related results, we refer the readers to the upcoming PhD thesis of the first author.

References

B. Beckermann and G. Labahn. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM Journal on Matrix Analysis and Applications*, 15(3):804–823, 1994.

B. Beckermann and G. Labahn. Recursiveness in matrix rational interpolation problems. *Journal of Computational and Applied Math*, 5-34, 1997.

B. Beckermann, G. Labahn, and G. Villard. Shifted normal forms of polynomial matrices. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC'99, pages 189–196, 1999.

B. Beckermann, G. Labahn, and G. Villard. Normal forms for general polynomial matrices. *Journal of Symbolic Computation*, 41(6): 708–737, 2006.

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*, 105:9–65, 1988.

Th. G. J. Beelen, G. J. van den Hurk, and C. Praagman. A new method for computing a column reduced polynomial matrix. *System Control Letters*, 10(4):217–224, 1988.

G.D. Forney. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal of Control*, 13:493–520, 1975.

E Frisk. *Residual Generation for Fault Diagnostics*. PhD thesis, Linköping, University, Sweden, 2001.

P. Giorgi, C.-P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation, Philadelphia, Pennsylvania, USA*, pages 135–142. ACM Press, 2003.

C. P. Jeannerod and G. Villard. Essentially optimal computation of the inverse of generic polynomial matrices. *Journal of Complexity*, 21(1):72–86, 2005.

C.-P. Jeannerod and G. Villard. Asymptotically fast polynomial matrix algorithms for multivariable systems. *Int. J. Control*, 79(11): 1359– 1367, 2006.

T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

V Kucera. *Discrete Linear Control : The Polynomial Equation Approach*. John Wiley and Sons, 1979.

P. Misra, P. Van Dooren, and A. Varga. Computation of structural invariants of generalized state-space systems. *Automatica*, 30:1921–1936, 1994.

C. Oara and P. Van Dooren. An improved algorithm for the computation of structural invariants of a system pencil and related geometric aspects. *Systems and Control Letters*, 30:38–48, 1997.

A. Storjohann and G. Villard. Computing the rank and a small nullspace basis of a polynomial matrix. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC'05, pages 309–316, 2005.

J. von zur Gathen and J Gerhard. *Modern Computer Algebra*. Cambridge University Press, 2nd edition, 2003.

W. Zhou and G. Labahn. Efficient computation of order bases. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC'09, pages 375–382. ACM, 2009.

W. Zhou and G. Labahn. Efficient algorithms for order basis computation. *Journal of Symbolic Computation*, 47:793–819, 2012.