# Efficient Algorithms for Order Bases Computation

Wei Zhou and George Labahn

*Cheriton School of Computer Science*
*University of Waterloo,*
*Waterloo, Ontario, Canada*

**Abstract**

In this paper we present two algorithms for the computation of a shifted order basis of an $m \times n$ matrix of power series over a field $\mathbb{K}$ with $m \leq n$. For a given order $\sigma$ and balanced shift $\vec{s}$ the first algorithm determines an order basis with a cost of $O^{\sim}(n^{\omega}\lceil m\sigma/n \rceil)$ field operations in $\mathbb{K}$, where $\omega$ is the exponent of matrix multiplication. Here an input shift is balanced when $\max(\vec{s}) - \min(\vec{s}) \in O(m\sigma/n)$. This extends earlier work of Storjohann which only determines a subset of an order basis that is within a specified degree bound $\delta$ using $O^{\sim}(n^{\omega}\delta)$ field operations for $\delta \geq \lceil m\sigma/n \rceil$.

While the first algorithm addresses the case when the column degrees of a complete order basis are unbalanced given a balanced input shift, it is not efficient in the case when an unbalanced shift results in the row degrees also becoming unbalanced. We present a second algorithm which balances the high degree rows and computes an order basis also using $O^{\sim}(n^{\omega}\lceil m\sigma/n \rceil)$ field operations in the case that the shift is unbalanced but satisfies the condition $\sum_{i=1}^{n}(\max(\vec{s}) - \vec{s}_i) \leq m\sigma$. This condition essentially allows us to locate those high degree rows that need to be balanced. This extends the earlier work by the authors from ISSAC'09.

## 1. Introduction

Let $\mathbf{F} \in \mathbb{K}[[x]]^{m \times n}$ be a matrix of power series over a field $\mathbb{K}$ with $m \leq n$. Given a nonnegative integer $\sigma$, we say a vector $\mathbf{p} \in \mathbb{K}[x]^{n \times 1}$ of polynomials gives an *order* $\sigma$ approximation of $\mathbf{F}$, or $\mathbf{p}$ has order $(\mathbf{F}, \sigma)$, if

$$\mathbf{F} \cdot \mathbf{p} \equiv \mathbf{0} \mod x^{\sigma},$$

that is, the first $\sigma$ terms of $\mathbf{F} \cdot \mathbf{p}$ are zero. Historically such problems date back to their use in Hermite's proof of the transcendence of $e$ in 1873. In 1893 Padé, a student of Hermite, formalized the concepts introduced by Hermite and defined what is now known

*Email address:* {w2zhou,glabahn}@uwaterloo.ca (Wei Zhou and George Labahn).

as Hermite-Padé approximants (where $m = 1$), Padé approximants (where $m = 1, n = 2$) and simultaneous Padé approximants (where $\mathbf{F}$ has a special structure). Such rational approximations also specified degree constraints on the polynomials $\mathbf{p}$ and had their order conditions related to these degree constraints. Additional order problems include vector and matrix versions of rational approximation, partial realizations of matrix sequences and vector rational reconstruction just to name a few (cf. the references in Beckermann and Labahn (1997)). As an example, the factorization of differential operators algorithm of Van Hoeij (1997) makes use of vector Hermite-Padé approximation to reconstruct differential factorizations over rational functions from factorizations of differential operators over power series domains.

The set of all such order $(\mathbf{F}, \sigma)$ approximations forms a module over $\mathbb{K}[x]$. An *order basis* - or minimal approximant basis or $\sigma$-basis - is a basis of this module having a type of minimal degree property (called reduced order basis in (Beckermann and Labahn, 1997)). The minimal degree property parameterizes solutions to an order problem by the degrees of the columns of the order basis. In the case of rational approximation, order bases can be viewed as a natural generalization of the Padé table of a power series (Baker and Graves-Morris, 1996) since they are able to describe *all* solutions to such problems given particular degree bounds. They can even be used to show the well known block structure of the Padé and related Rational Interpolation tables (Beckermann and Labahn, 1997). Order bases are used in such diverse applications as the inversion of structured matrices (Labahn, 1992), normal forms of matrix polynomials (Beckermann et al., 1999, 2006), and other important problems in matrix polynomial arithmetic including matrix inversion, determinant and nullspace computation (Giorgi et al., 2003; Storjohann and Villard, 2005). In our case we also allow the minimal degree property to include a shift $\vec{s}$. Such a shift is important, for example, for matrix normal form problems (Beckermann et al., 1999, 2006).

In this paper we focus on the efficient computation of order basis. Algorithms for fast computation of order basis include that of Beckermann and Labahn (1994) which converts the matrix problem into a vector problem of higher order (which they called the Power Hermite-Padé problem). Their divide and conquer algorithm has complexity of $O^{\sim}(n^2 m\sigma + nm^2\sigma)$ field operations. As usual, the soft-$O$ notation $O^{\sim}$ is simply Big-$O$ with polylogarithmic factors $(\log(nm\sigma))^{O(1)}$ ommited. By working more directly on the input $m \times n$ input matrix, Giorgi et al. (2003) give a divide and conquer method with cost $O^{\sim}(n^{\omega}\sigma)$ arithmetic operations. Their method is very efficient if $m$ is close to the size of $n$ but can be improved if $m$ is small.

In a novel construction, Storjohann (2006) effectively reverses the approach of Beckermann and Labahn. Namely, rather than convert a high dimension matrix order problem into a lower dimension vector problem of higher order, Storjohann converts a low dimension problem to a high dimension problem with lower order. For example, computing an order basis for a $1 \times n$ vector input $\mathbf{f}$ and order $\sigma$ can be converted to a problem of order basis computation with an $O(n) \times O(n)$ input matrix and an order $O(\lceil \sigma/n \rceil)$. Combining this conversion with the method of Giorgi et al. can then be used effectively for problems with small row dimensions to achieve a cost of $O^{\sim}(n^{\omega} \lceil m\sigma/n \rceil)$.

However, while order bases of the original problem can have degree up to $\sigma$, the nature of Storjohann's conversion limits the degree of an order basis of the converted problem to $O(\lceil m\sigma/n \rceil)$ in order to be computationally efficient. In other words, this approach does not in general compute a complete order basis. Rather, in order to achieve efficiency,

it only computes a partial order basis containing basis elements with degrees within $O(\lceil m\sigma/n \rceil)$, referred to by Storjohann as a *minbasis*. Fast methods for computing a minbasis are particularly useful for certain problems, for example, in the case of inversion of structured block matrices where one needs only precisely the minbasis (Labahn, 1992). However, in other applications, such as those arising in matrix polynomial arithmetic, one needs a complete basis which specifies all solutions of a given order, not just those within a particular degree bound (cf. Beckermann and Labahn (1997)).

In this paper we present two algorithms which compute an entire order basis with a cost of $O^{\sim}(n^{\omega}\lceil m\sigma/n \rceil)$ field operations. This work extends the previous results first reported in Zhou and Labahn (2009). The two algorithms differ depending on the nature of the degree shift required for the reduced order basis. In the first case we use a transformation that can be considered as an extension of Storjohann's transformation. This new transformation provides a way to extend the results from one transformed problem to another transformed problem of a higher degree. This enables us to use an idea from the null space basis algorithm found in (Storjohann and Villard, 2005) in order to achieve efficient computation. At each iteration, basis elements within a specified degree bound are computed via a Storjohann transformed problem. Then the partial result is used to simplify the next Storjohann transformed problem of a higher degree, allowing basis elements within a higher degree bound to be computed efficiently. This is repeated until all basis elements are computed.

In order to compute an order basis efficiently, the first algorithm requires that the degree shifts are balanced. In the case where the shift is not balanced, the row degrees of the basis can also become unbalanced in addition to the unbalanced column degrees. We give a second algorithm that balances the high degree rows and uses $O^{\sim}(n^{\omega}\lceil m\sigma/n \rceil)$ field operations when the shift $\vec{s}$ is unbalanced but satisfies the condition $\sum_{i=1}^{n}(\max(\vec{s})-\vec{s}_i) \leq m\sigma$. This condition essentially allows us to locate the high degree unbalanced rows that need to be balanced. The algorithm converts a problem of unbalanced shift to one with balanced shift, based on a second idea from (Storjohann, 2006). Then the first algorithm is used to efficiently compute the elements of an order basis whose shifted degrees exceed a specified parameter. The problem is then reduced to one where we remove the computed elements. This results in a new problem with smaller dimension and higher degree. The same process is repeated again on this new problem in order to compute the elements with the next highest shifted degrees.

The remaining paper is structured as follows. Basic definitions and properties of order bases are given in the next section. Section 3 provides an extension to Storjohann's transformation to allow higher degree basis elements to be computed. Based on this new transformation, Section 4 establishes a link between two Storjohann transformed problems of different degrees, from which an recursive method and then an iterative algorithm are derived. The time complexity is analyzed in the next section. After this, Section 6 describes an algorithm which handles problems with a type of unbalanced shift. This is followed by a conclusion along with a description for topics for future research.

## 2. Preliminaries

The computational cost in this paper is analysed by bounding the number of arithmetic operations (additions, subtractions, multiplications, and divisions) in the coefficient field $\mathbb{K}$ on an algebraic random access machine. We use $\mathrm{MM}(n, d)$ to denote the cost of

multiplying two polynomial matrices with dimension $n$ and degree $d$, and $\mathrm{M}(n)$ to denote the cost of multiplying two polynomials with degree $d$. We define a cost function $\bar{\mathrm{M}}(d) = d \log d \log\log d$, then $\bar{\mathrm{M}}(ab) \in O\left(\bar{\mathrm{M}}(a)\bar{\mathrm{M}}(b)\right)$ and $\bar{\mathrm{M}}(t) \in O(n^{\omega-1})$. We take $\mathrm{MM}(n,d) \in O\left(n^{\omega}\,\mathrm{M}(d)\right) \subset O(n^{\omega}\,\mathrm{M}'(d))$, where the multiplication exponent $\omega$ 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.

In the remaining of this section, we provide some of the background needed in order to understand the basic concepts and tools needed for order basis computation. This includes basic definitions and a look at the size of the input and the output for computing such bases. The challenges of balancing input and handling unbalanced output are discussed along with the techniques which we plan to use to overcome the difficulties. We review the construction by Storjohann (2006) which transforms the inputs to those having dimensions and degree balance better suited for fast computation and discuss an idea from Storjohann and Villard (2005) for handling the case where the output degree is unbalanced.

### *2.1. Order Basis*

Let $\mathbb{K}$ be a field, $\mathbf{F} \in \mathbb{K}\left[[x]\right]^{m\times n}$ a matrix of power series and $\vec{\sigma} = [\sigma_1, \ldots, \sigma_m]$ a vector of non-negative integers.

**Definition 2.1.** A vector of polynomials $\mathbf{p} \in \mathbb{K}\left[x\right]^{n\times 1}$ has *order* $(\mathbf{F}, \vec{\sigma})$ (or *order* $\vec{\sigma}$ with respect to $\mathbf{F}$) if $\mathbf{F} \cdot \mathbf{p} \equiv \mathbf{0} \mod x^{\vec{\sigma}}$, that is,

$$\mathbf{F} \cdot \mathbf{p} = x^{\vec{\sigma}}\mathbf{r} = \begin{bmatrix} x^{\sigma_1} & & \\ & \ddots & \\ & & x^{\sigma_m} \end{bmatrix} \mathbf{r}$$

for some $\mathbf{r} \in \mathbb{K}\left[[x]\right]^{m\times 1}$. If $\vec{\sigma} = [\sigma, \ldots, \sigma]$ is uniform, then we say that $\mathbf{p}$ has order $(\mathbf{F}, \sigma)$. The set of all order $(\mathbf{F}, \vec{\sigma})$ vectors is a $\mathbb{K}\left[x\right]$-module denoted by $\langle(\mathbf{F}, \vec{\sigma})\rangle$.

An order basis for $\mathbf{F}$ and $\vec{\sigma}$ is simply a basis for the module $\langle(\mathbf{F}, \vec{\sigma})\rangle$. In this paper we compute those order bases having a type of minimality degree condition (also referred to as a reduced order basis in (Beckermann and Labahn, 1997)). 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).

The shifted column degree of a column polynomial vector $\mathbf{p}$ with shift $\vec{s} = [s_1, \ldots, s_n] \in \mathbb{Z}^n$ is given by

$$\deg_{\vec{s}} \mathbf{p} = \max_{1\leq i\leq n} [\deg p^{(i)} + s_i] = \deg(x^{\vec{s}} \cdot \mathbf{p}).$$

We call this the *$\vec{s}$-column degree*, or simply the *$\vec{s}$-degree* of $\mathbf{p}$. A shifted column degree defined this way is equivalent to the notion of *defect* commonly used in the literature. Our definition of $\vec{s}$-degree is also equivalent to the notion of $\mathbf{H}$-degree from (Beckermann and Labahn, 1997) for $\mathbf{H} = x^{\vec{s}}$. 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 $\mathbf{P}$ is a $\vec{s}$-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. Similarly, $\mathbf{P}$ is in $\vec{s}$-Popov form if $x^{\vec{s}} \cdot \mathbf{P}$ is in Popov form (Beckermann et al., 1999, 2006).

An *order basis* (Beckermann and Labahn, 1994, 1997) $\mathbf{P}$ of $\mathbf{F}$ with order $\vec{\sigma}$ and shift $\vec{s}$, or simply an $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis, is a basis for the module $\langle(\mathbf{F}, \vec{\sigma})\rangle$ having minimal $\vec{s}$-column degrees. If $\vec{\sigma} = [\sigma, \ldots, \sigma]$ are constant vectors then we simply write $(\mathbf{F}, \sigma, \vec{s})$-basis. The precise definition of an $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis is as follows.

**Definition 2.2.** A polynomial matrix $\mathbf{P}$ is an order basis of $\mathbf{F}$ of order $\sigma$ and shift $\vec{s}$, denoted by $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis, if the following properties hold:

(1) $\mathbf{P}$ is a nonsigular matrix of dimension $n$.
(2) $\mathbf{P}$ is $\vec{s}$-column reduced.
(3) $\mathbf{P}$ has order $(\mathbf{F}, \vec{\sigma})$ (or equivalently, each column of $\mathbf{P}$ is in $\langle(\mathbf{F}, \vec{\sigma})\rangle$).
(4) Any $\mathbf{q} \in \langle(\mathbf{F}, \vec{\sigma})\rangle$ can be expressed as a linear combination of the columns of $\mathbf{P}$, given by $\mathbf{P}^{-1}\mathbf{q}$.

Although we allow different orders for each row in this definition, we focus on order basis computation problems having uniform order. However special cases of non-uniform order problems are still needed in our analysis. We also assume $m \leq n$ for simplicity. The case of $m > n$ can be transformed to the case of $m \leq n$ by compression (Storjohann and Villard, 2005). We further assume, without any loss of generality, that $n/m$ and $\sigma$ are powers of two. This can be achieved by padding zero rows to the input matrix and multiplying it by some power of $x$.

From (Beckermann and Labahn, 1997) we have the following lemma.

**Lemma 2.3.** *The following are equivalent for a polynomial matrix* $\mathbf{P}$*:*

*(1)* $\mathbf{P}$ *is a* $(\mathbf{F}, \vec{\sigma}, \vec{s})$*-basis.*
*(2)* $\mathbf{P}$ *is comprised of a set of* $n$ *minimal* $\vec{s}$*-degree polynomial vectors that are linearly independent and each having order* $(\mathbf{F}, \vec{\sigma})$*.*
*(3)* $\mathbf{P}$ *does not contain a zero column, has order* $(\mathbf{F}, \vec{\sigma})$*, is* $\vec{s}$*-column reduced, and any* $\mathbf{q} \in \langle(\mathbf{F}, \vec{\sigma})\rangle$ *can be expressed as a linear combination of the columns of* $\mathbf{P}$*.*

In some cases an entire order basis is unnecessary and instead one looks for a minimal basis that generates only the elements of $\langle(\mathbf{F}, \vec{\sigma})\rangle$ with $\vec{s}$-degrees bounded by a given $\delta$. Such a minimal basis is a partial $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis comprised of elements of a $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis with $\vec{s}$-degrees bounded by $\delta$. This is called a *minbasis* in Storjohann (2006).

**Definition 2.4.** Let $\langle(\mathbf{F}, \vec{\sigma}, \vec{s})\rangle_\delta \subset \langle(\mathbf{F}, \vec{\sigma})\rangle$ denote the set of order $(\mathbf{F}, \vec{\sigma})$ polynomial vectors with $\vec{s}$-degree bounded by $\delta$. A $(\mathbf{F}, \vec{\sigma}, \vec{s})_\delta$-basis is a polynomial matrix $\mathbf{P}$ not containing a zero column and satisfying:

(1) $\mathbf{P}$ has order $(\mathbf{F}, \vec{\sigma})$.
(2) Any element of $\langle(\mathbf{F}, \vec{\sigma}, \vec{s})\rangle_\delta$ can be expressed as a linear combination of the columns of $\mathbf{P}$.
(3) $\mathbf{P}$ is $\vec{s}$-column reduced.

A $(\mathbf{F}, \vec{\sigma}, \vec{s})_\delta$-basis is, in general, not square unless $\delta$ is large enough to contain all $n$ basis elements in which case it is a complete $(\mathbf{F}, \vec{\sigma}, \vec{s})$-basis.

*2.2. Balancing Input with Storjohann's Transformation*

For computing a $(\mathbf{F}, \sigma, \vec{s})$-basis with input matrix $\mathbf{F} \in \mathbb{K}[[x]]^{m\times n}$, shift $\vec{s}$ and order $\sigma$ one can view $\mathbf{F}$ as a polynomial matrix with degree $\sigma - 1$, as higher order terms are not needed in the computation. As such the total input size of an order basis problem is $mn\sigma$ coefficients. One can apply the method of Giorgi et al. (2003) directly, which gives a cost of

$$\begin{aligned}\sum_{i=0}^{\log\sigma} 2^i \operatorname{MM}(n, 2^{-i}\sigma) &= \sum_{i=0}^{\log\sigma} 2^{-i}\sigma \operatorname{MM}(n, 2^i)\\ &\subset O\left(\sum_{i=0}^{\log\sigma} 2^{-i} n^\omega \sigma 2^i \log 2^i \log\log 2^i\right)\\ &= O\left(n^\omega \sigma \sum_{i=0}^{\log\sigma} i \log i\right)\\ &\subset O\left(n^\omega \sigma \sum_{i=0}^{\log\sigma} \log\sigma \log\log\sigma\right)\\ &= O\left(n^\omega \sigma \log^2\sigma \log\log\sigma\right) = O(n^\omega \bar{\mathrm{M}}(\sigma)\log\sigma),\end{aligned}$$

close to the cost of multiplying two matrices with dimension $n$ and degree $\sigma$. Note that this cost is independent of the degree shift. This is very efficient if $m \in \Theta(n)$. However, for small $m$, say $m = 1$ as in Hermite Padé approximation, the total input size is only $n\sigma$ coefficients. Matrix multiplication cannot be used effectively on a such vector input.

Storjohann (2006) provides a novel way to transform an order basis problem with small row dimension to a problem with higher row dimension and possibly lower degree to take advantage of Giorgi et al. (2003)'s algorithm. We provide a quick overview of a slightly modified version of Storjohann's method. Our small modification allows a nonuniform degree shift for the input and provides a slightly simpler degree shift, degree, and order for the transformed problem. The proof of its correctness is provided in Section 3. In order to compute a $(\mathbf{F}, \sigma, \vec{s})$-basis, assuming without loss of generality that $\min(\vec{s}) = 0$, we first write

$$\mathbf{F} = \mathbf{F}_0 + \mathbf{F}_1 x^\delta + \mathbf{F}_2 x^{2\delta} + \cdots + \mathbf{F}_l x^{l\delta},$$

with $\deg \mathbf{F}_i < \delta$ for a positive integer $\delta$, and where we assume (again without loss of generality) that $\sigma = (l+1)\delta$. Set

$$\bar{\mathbf{F}} = \left[\begin{array}{c|ccccc} \mathbf{F}_0 + \mathbf{F}_1 x^\delta & \mathbf{0}_m & \mathbf{0}_m & \cdots & & \mathbf{0}_m \\ \hline \mathbf{F}_1 + \mathbf{F}_2 x^\delta & \mathbf{I}_m & \mathbf{0}_m & & & \\ \mathbf{F}_2 + \mathbf{F}_3 x^\delta & \mathbf{0}_m & \mathbf{I}_m & & & \\ \vdots & & & \ddots & & \\ \mathbf{F}_{l-1} + \mathbf{F}_l x^\delta & & & & & \mathbf{I}_m \end{array}\right]_{ml\times(n+m(l-1))}.$$

On the left side of $\bar{\mathbf{F}}$, each block $\mathbf{F}_i + \mathbf{F}_{i+1}x^\delta$ has dimension $m \times n$. On the right side, there are $l \times (l-1)$ blocks of $\mathbf{0}_m$'s or $\mathbf{I}_m$'s each having dimension $m \times m$. The overall dimension of $\bar{\mathbf{F}}$ is $ml \times (n + m(l-1))$. Set $\vec{s}' = [\vec{s}, 0, \ldots, 0]$ ($\vec{s}$ followed by $m(l-1)$ 0's).

A $(\bar{\mathbf{F}}, 2\delta, \vec{s'})$-basis can then be computed by the method of Giorgi et al. with a cost of $O^{\sim}(n^{\omega}\delta)$ for $\delta \geq \lceil m\sigma/n \rceil$. This transformation of Storjohann can be viewed as a partial linearization of the original problem, where $\bar{\mathbf{F}}$ is analogous to the coefficient matrix of $\mathbf{F}$. Note that $\bar{\mathbf{F}}$ has $l$ block rows each containing $m$ rows. We continue to use each block row to represent $m$ rows for the remainder of the paper.

Clearly a $(\bar{\mathbf{F}}, 2\delta, \vec{s'})$-basis $\bar{\mathbf{P}}$ of the transformed problem is not a $(\mathbf{F}, \sigma, \vec{s})$-basis of the original problem, as $\bar{\mathbf{P}}$ has a higher dimension and lower degree. However, the first $n$ rows of the $(\bar{\mathbf{F}}, 2\delta, \vec{s'})_{\delta-1}$-basis contained in $\bar{\mathbf{P}}$ is a $(\mathbf{F}, \sigma, \vec{s})_{\delta-1}$-basis.

Note that there is no need to set the degree parameter $\delta$ to less than $\lceil m\sigma/n \rceil$, as this produces fewer basis elements without a better cost. The lowest cost is achieved when $\bar{\mathbf{F}}$ is close to square so matrix multiplication can be used most effectively. This requires the number of block rows $l$ of $\bar{\mathbf{F}}$ to be close to $n/m$, which requires $\delta = \Theta(\lceil m\sigma/n \rceil)$. Recall that $mn\sigma$ is the total size of the original $m \times n$ input matrix $\mathbf{F}$, hence $d = mn\sigma/n^2 = m\sigma/n$ is the average degree of each entry of $\mathbf{F}$ if the $m$ rows of $\mathbf{F}$ are spread out over $n$ rows. Choosing $\delta = \Theta(\lceil d \rceil)$, the cost of computing a $(\bar{\mathbf{F}}, 2\delta, \vec{s'})$-basis is then $O^{\sim}(n^{\omega} \lceil d \rceil) = O^{\sim}(n^{\omega} \lceil m\sigma/n \rceil)$. The ceiling function here is used to take care of the case of $m\sigma < n$. For the remainder of the paper, we assume that $m\sigma \geq n$ in order to avoid the need for the ceiling function and so simplify the presentation. Together with the assumption that $\sigma$ and $n/m$ are both powers of two, $m\sigma/n$ is then always a positive integer in this paper.

**Example 2.5.** Let $\mathbb{K} = \mathbb{Z}_2$, $\sigma = 8$, $\delta = 2$ and

$$\mathbf{F} = [x+x^2+x^3+x^4+x^5+x^6,\ 1+x+x^5+x^6+x^7,\ 1+x^2+x^4+x^5+x^6+x^7,\ 1+x+x^3+x^7]$$

a vector of size $1 \times 4$. Then

$$\bar{\mathbf{F}} = \left[\begin{array}{cccc|cc} x+x^2+x^3 & 1+x & 1+x^2 & 1+x+x^2 & 0 & 0 \\ \hline 1+x+x^2+x^3 & x^3 & 1+x^2+x^3 & x & 1 & 0 \\ 1+x+x^2 & x+x^2+x^3 & 1+x+x^2+x^3 & x^3 & 0 & 1 \end{array}\right]_{3\times 6}$$

and a $\left(\bar{\mathbf{F}}, 4, \vec{0}\right)$-basis is given by

$$\bar{\mathbf{P}} = \left[\begin{array}{cc|cccc} 1 & x & 1 & x^2+x^3 & 0 & x+x^2+x^3 \\ 0 & 1 & 0 & x^2 & x^2+x^3 & 0 \\ 1 & 1+x & x+x^2 & x^2 & x^2 & x^2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 & x^2 & x+x^2+x^3 \\ 0 & 1 & 1+x^2 & 0 & x^2 & x+x^2 \end{array}\right].$$

The first two columns of $\bar{\mathbf{P}}$ have degree less than 2, hence its top left $4 \times 2$ submatrix is

a $\left(\mathbf{F}, 8, \vec{0}\right)_1$-basis. This is a low degree part of the $(\mathbf{F}, 8, \vec{0})$-basis

$$\mathbf{P} = \begin{bmatrix} 1 & x & 1 & x^2 \\ 0 & 1 & x^2 + x^3 & 0 \\ 1 & 1+x & x & x^3 + x^4 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Note that if $\delta$ is set to $\sigma/2 = 4$, then the transformed problem is the same as the original problem.

### *2.3. Unbalanced Output*

Storjohann's transformation can be used to efficiently compute a $(\mathbf{F}, \sigma, \vec{s})_{\delta-1}$-basis if the degree parameter $\delta$ is close to the average degree $d = m\sigma/n$. However, if $\delta$ is large, say $\delta = \Theta(\sigma)$, or if we want to compute a complete $(\mathbf{F}, \sigma, \vec{s})$-basis, then the current analysis for the computation still gives the cost estimate of $O^{\sim}(n^{\omega}\sigma)$.

The underlying difficulty with computing a complete order basis is that the basis can have degree up to $\sigma$. As the output of this problem has dimension $n \times n$ and degree up to $\Theta(\sigma)$, this may seem to suggest $O^{\sim}(n^{\omega}\sigma)$ is about the best that can be done. However, the total size of the output, that is, the total number of coefficients of all $n^2$ polynomial entries can still be bounded by $O(mn\sigma)$, the same as the size of the input. This gives some hope for a more efficient method.

**Lemma 2.6.** *Let $\vec{t}$ be the $\vec{s}$-column degrees of a $(\mathbf{F}, \sigma, \vec{s})$-basis. Then $\sum_i \left(\vec{t}_i - \vec{s}_i\right) \leq m\sigma$. In addition, the total size of any $(\mathbf{F}, \sigma, \vec{s})$-basis in $\vec{s}$-Popov form is bounded by $nm\sigma$.*

**Proof.** This can be shown by considering the sizes of the pivots in the iterative order basis computation given in (Beckermann and Labahn, 1994; Giorgi et al., 2003). □

Let us now look at the average column degree of the output. In the first part of this paper, we assumed, without loss of generality, that $\min(\vec{s}) = 0$ so $\deg \mathbf{q} \leq \deg_{\vec{s}} \mathbf{q}$ for any $\mathbf{q} \in \mathbb{K}[x]^n$. The situation is simpler if the shift $\vec{s}$ is uniform since then $\sum_i \vec{t}_i \leq m\sigma$ by Lemma 2.6 and the average column degree is therefore bounded by $d = m\sigma/n$. In the first part of this paper, we consider a slightly more general case, when the shift $\vec{s}$ is *balanced*, which is defined as follows.

**Definition 2.7.** A shift $\vec{s}$ is balanced if $\max \vec{s} - \min \vec{s} \in O(d) = O(m\sigma/n)$.

By assuming $\min \vec{s} = 0$, $\vec{s}$ is balanced if $\max \vec{s} \in O(d)$. In this case, Lemma 2.6 implies $\sum_i \left(\vec{t}_i\right) \leq m\sigma + \sum_i (\vec{s}_i) \in O(m\sigma + nd) = O(m\sigma)$. Hence the average column degree of the output basis remains $O(d)$.

The fact that a $(\mathbf{F}, \sigma, \vec{s})$-basis can have degree up to $\sigma$ while its average column degree is $O(m\sigma/n)$ implies that an order basis can have quite unbalanced column degrees, especially if $m$ is small. A similar problem with unbalanced output is encountered in null space basis computation. Storjohann and Villard (2005) deal with this in the following way.

Let $d$ be the average column degree of the output. Set the degree parameter $\delta$ to twice that of $d$. This allows one to compute at least half the columns of a basis (since the number of columns with degree at least $\delta$ must be at most a half of the total number of columns). One can then simplify the problem, so that the computed basis elements are completely removed from the problem. This reduces the dimension of the problem by at least a factor of 2. One then doubles the degree bound $\delta$ in order to have at least 3/4 of the basis elements computed. Repeating this, at iteration $i$, at most $1/2^i$ of the basis elements are remaining. Therefore, no more than $\log n$ iterations are needed to compute all basis elements.

## 3. Extending Storjohann's Transformation

In this section, we introduce a transformation that can be viewed as an extension of Storjohann's transformation which allows for computation of a full, rather than partial, order basis. More generally (as discussed in the next section) this transformation provides a link between two Storjohann transformed problems constructed using different degree parameters. For easier understanding, we first focus on a particular case of this transformation in Subsection 3.1 and then generalize this in Subsection 3.2.

### *3.1. A Particular Case*

Consider the problem of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis. We assume $\sigma = 4\delta$ for a positive integer $\delta$ and write the input matrix polynomial as $\mathbf{F} = \mathbf{F}_0 + \mathbf{F}_1 x^\delta + \mathbf{F}_2 x^{2\delta} + \mathbf{F}_3 x^{3\delta}$ with $\deg \mathbf{F}_i < \delta$. In the following, we show that computing a $(\mathbf{F}, \sigma, \vec{s})$-basis can be done by computing a $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis where

$$\mathbf{F}' = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{F}'_{21} & \mathbf{F}'_{22} \end{bmatrix} = \left[\begin{array}{c|cc} \mathbf{F}_0 + \mathbf{F}_1 x^\delta + \mathbf{F}_2 x^{2\delta} + \mathbf{F}_3 x^{3\delta} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{F}_1 + \mathbf{F}_2 x^\delta & \mathbf{I}_m & \mathbf{0} \\ \mathbf{F}_2 + \mathbf{F}_3 x^\delta & \mathbf{0} & \mathbf{I}_m \end{array}\right] \tag{3.1}$$

with order $\vec{\omega} = [4\delta, \ldots, 4\delta, 2\delta, \ldots, 2\delta]$ (with $m$ $4\delta$'s and $2m$ $2\delta$'s) and degree shift $\vec{s'} = [\vec{s}, e, \ldots, e]$ (with $2m$ $e$'s), where $e$ is an integer less than or equal to 1. We set $e$ to 0 in this paper for simplicity [1] .

We first look at the correspondence bettween the elements of $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_\tau$ and the elements of $\langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle_\tau$ in Lemma 3.1 to Lemma 3.5. The correspondence between $(\mathbf{F}, \sigma, \vec{s})$-bases and $(\mathbf{F}', \vec{\omega}, \vec{s'})$-bases is then considered in Corollary 3.7 to Theorem 3.10.

Let

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_n \\ x^{-\delta}\mathbf{F}_0 \\ x^{-2\delta}\left(\mathbf{F}_0 + \mathbf{F}_1 x^\delta\right) \end{bmatrix}.$$

**Lemma 3.1.** *If* $\mathbf{q} \in \langle(\mathbf{F}, \sigma)\rangle$, *then* $\mathbf{Bq} \in \langle(\mathbf{F}', \vec{\omega})\rangle$.

[1] Storjohann used $e = 1$ in (Storjohann, 2006). All results in this section still hold for any other $e \leq 1$.

**Proof.** The lemma follows from

$$\mathbf{F}'\mathbf{B}\mathbf{q} = \begin{bmatrix} \mathbf{F}_0 + \mathbf{F}_1 x^{\delta} + \mathbf{F}_2 x^{2\delta} + \mathbf{F}_3 x^{3\delta} \\ \mathbf{F}_0 x^{-\delta} + \mathbf{F}_1 + \mathbf{F}_2 x^{\delta} \\ \mathbf{F}_0 x^{-2\delta} + \mathbf{F}_1 x^{-\delta} + \mathbf{F}_2 + \mathbf{F}_3 x^{\delta} \end{bmatrix} \mathbf{q} \equiv \mathbf{0} \mod x^{\vec{\omega}}.$$

Note that the bottom rows of $\mathbf{B}$ may not be polynomials. However, $\mathbf{B}\mathbf{q}$ is a polynomial vector since $\mathbf{q} \in \langle(\mathbf{F}, \sigma)\rangle$ implies $\mathbf{q} \in \langle(\mathbf{F}_0, \delta)\rangle$ and $\mathbf{q} \in \left\langle \left(\mathbf{F}_0 + \mathbf{F}_1 x^{\delta}, 2\delta\right)\right\rangle$. □

The following lemma shows that the condition $e \leq 1$ forces $\deg_{\vec{s'}} \mathbf{B}\mathbf{q}$ to be determined by $\mathbf{q}$.

**Lemma 3.2.** *If $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\tau}$ for any degree bound $\tau \in \mathbb{Z}$, then $\deg_{\vec{s'}} \mathbf{B}\mathbf{q} = \deg_{\vec{s}} \mathbf{q}$.*

**Proof.** By assumption $s_i \geq 0$, so $\deg \mathbf{q} \leq \deg_{\vec{s}} \mathbf{q}$. Now consider the degree of the bottom $2m$ entries, $\mathbf{q}_2, \mathbf{q}_3$, of

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{q}_2 \\ \mathbf{q}_3 \end{bmatrix} = \mathbf{B}\mathbf{q} = \begin{bmatrix} \mathbf{q} \\ x^{-\delta}\mathbf{F}_0 \cdot \mathbf{q} \\ x^{-2\delta}\left(\mathbf{F}_0 + \mathbf{F}_1 x^{\delta}\right) \cdot \mathbf{q} \end{bmatrix}.$$

Our goal is to show $\deg_{\vec{e}} \left[\mathbf{q}_2^T, \mathbf{q}_3^T\right]^T \leq \deg_{\vec{s}} \mathbf{q}$. Note that

$$\deg \mathbf{q}_2 = \deg\left(\mathbf{F}_0 \mathbf{q}/x^{\delta}\right) \leq \deg \mathbf{q} + \delta - 1 - \delta \leq \deg_{\vec{s}} \mathbf{q} - 1,$$

and similarly $\deg \mathbf{q}_3 \leq \deg_{\vec{s}} \mathbf{q} - 1$. Therefore

$$\deg_{\vec{e}} \begin{bmatrix} \mathbf{q}_2 \\ \mathbf{q}_3 \end{bmatrix} = \deg \begin{bmatrix} \mathbf{q}_2 \\ \mathbf{q}_3 \end{bmatrix} + e \leq \deg_{\vec{s}} \mathbf{q} - 1 + e \leq \deg_{\vec{s}} \mathbf{q}.$$

□

**Corollary 3.3.** *If $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\tau}$ for any degree bound $\tau \in \mathbb{Z}$ , then $\mathbf{B}\mathbf{q} \in \langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle_{\tau}$.*

**Corollary 3.4.** *Let $\bar{\mathbf{S}}_{\tau}$ be a $(\mathbf{F}', \vec{\omega}, \vec{s'})_{\tau}$-basis and $\mathbf{S}_{\tau}$ be the top $n$ rows of $\bar{\mathbf{S}}_{\tau}$ for any bound $\tau \in \mathbb{Z}$. Then any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\tau}$ is a linear combination of the columns of $\mathbf{S}_{\tau}$.*

**Proof.** By Corollary 3.3, $\mathbf{B}\mathbf{q} \in \langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle_{\tau}$, and so is a linear combination of columns of $\bar{\mathbf{S}}_{\tau}$. That is, there exists a polynomial vector $\mathbf{u}$ such that $\mathbf{B}\mathbf{q} = \bar{\mathbf{S}}_{\tau}\mathbf{u}$. This remains true if we restrict the equation to the top $n$ rows, that is, $\mathbf{q} = [\mathbf{I}_n, \mathbf{0}]\,\mathbf{B}\mathbf{q} = [\mathbf{I}_n, \mathbf{0}]\,\bar{\mathbf{S}}_{\tau}\mathbf{u} = \mathbf{S}_{\tau}\mathbf{u}$. □

**Lemma 3.5.** *Let $\bar{\mathbf{q}} \in \langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle_{\tau}$ for any degree bound $\tau \in \mathbb{Z}$, and $\mathbf{q}_1$ the first $n$ entries of $\bar{\mathbf{q}}$. Then $\mathbf{q}_1 \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\tau}$.*

**Proof.** The top rows of

$$\mathbf{F}'\mathbf{q} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{F}'_{21} & \mathbf{F}'_{22} \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{F}\mathbf{q}_1 \\ \mathbf{F}'_{21}\mathbf{q}_1 + \mathbf{F}'_{22}\mathbf{q}_2 \end{bmatrix} \equiv \mathbf{0} \mod x^{\vec{\omega}}$$

give $\mathbf{F}\mathbf{q}_1 \equiv \mathbf{0} \mod x^{\sigma}$. □

The next lemma shows a $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis can be constructed from a $(\mathbf{F}, \sigma, \vec{s})$-basis. This well-formed $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis restricts the elements of $\langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle$ to a simple form shown in Corollary 3.7. This in turn helps to establish a close correspondence between a $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis and a $(\mathbf{F}, \sigma, \vec{s})$-basis in Lemma 3.8, Lemma 3.9, and Theorem 3.10.

**Lemma 3.6.** *If* $\mathbf{P}$ *is a* $(\mathbf{F}, \sigma, \vec{s})$*-basis, then*

$$\bar{\mathbf{T}} = \left[\begin{array}{c|c} \mathbf{BP} & \begin{matrix} \mathbf{0}_{n\times 2m} \\ x^{2\delta}\mathbf{I}_{2m} \end{matrix} \end{array}\right] = \left[\begin{array}{c|cc} \mathbf{P} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times m} \\ \hline x^{-\delta}\mathbf{F}_0 \cdot \mathbf{P} & x^{2\delta}\mathbf{I}_m & \mathbf{0}_m \\ x^{-2\delta}\left(\mathbf{F}_0 + \mathbf{F}_1 x^{\delta}\right) \cdot \mathbf{P} & \mathbf{0}_m & x^{2\delta}\mathbf{I}_m \end{array}\right]$$

*is a* $(\mathbf{F}', \vec{\omega}, \vec{s'})$*-basis.*

**Proof.** By Lemma 3.1, $\bar{\mathbf{T}}$ has order $(\mathbf{F}', \vec{\omega})$ and is $\vec{s'}$-column reduced since $\mathbf{P}$ dominates the $\vec{s'}$-degrees of $\bar{\mathbf{T}}$ on the left side by Lemma 3.2. It remains to show that any $\bar{\mathbf{q}} \in \langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle$ is a linear combination of the columns of $\bar{\mathbf{T}}$.

Let $\mathbf{q}$ be the top $n$ entries of $\bar{\mathbf{q}}$. Then by Lemma 3.5, $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle$, hence is a linear combination of the columns of $\mathbf{P}$, that is $\mathbf{q} = \mathbf{P}\mathbf{u}$ with $\mathbf{u} = \mathbf{P}^{-1}\mathbf{q} \in \mathbb{K}[x]^{n\times 1}$. Subtracting the contribution of $\mathbf{P}$ from $\bar{\mathbf{q}}$, we get

$$\mathbf{q}' = \bar{\mathbf{q}} - \mathbf{BPu} = \bar{\mathbf{q}} - \mathbf{Bq} = \begin{bmatrix} \mathbf{0} \\ \mathbf{v} \end{bmatrix},$$

which is still in $\langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle$, that is,

$$\mathbf{F}'\mathbf{q}' = \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_{2m}\mathbf{v} \end{bmatrix} \equiv \mathbf{0} \mod x^{\vec{\omega}}.$$

This forces $\mathbf{v}$ to be a linear combination of the columns of $x^{2\delta}\mathbf{I}_{2m}$, the bottom right submatrix of $\bar{\mathbf{T}}$. Now $\bar{\mathbf{q}} = \bar{\mathbf{T}}\left[\mathbf{u}^T, \mathbf{v}^T\right]^T$ as required. □

**Corollary 3.7.** *Let* $\tau \in \mathbb{Z}$ *be any degree bound and* $\mathbf{P}_\tau \in \mathbb{K}[x]^{n\times t}$ *be a* $(\mathbf{F}, \sigma, \vec{s})_\tau$*-basis. If* $\bar{\mathbf{q}} \in \langle(\mathbf{F}', \vec{\omega}, \vec{s'})\rangle_\tau$ *and* $\mathbf{q}$ *is the top* $n$ *entries of* $\bar{\mathbf{q}}$*, then* $\bar{\mathbf{q}}$ *must have the form*

$$\bar{\mathbf{q}} = \mathbf{BP}_\tau\mathbf{u} + x^{2\delta}\begin{bmatrix} \mathbf{0} \\ \mathbf{v} \end{bmatrix} = \mathbf{Bq} + x^{2\delta}\begin{bmatrix} \mathbf{0} \\ \mathbf{v} \end{bmatrix}$$

*for some polynomial vector* $\mathbf{u} \in \mathbb{K}[x]^{t\times 1}$ *and* $\mathbf{v} \in \mathbb{K}[x]^{2m\times 1}$. *In particular, if* $\deg_{\vec{s'}} \bar{\mathbf{q}} < 2\delta$, *then* $\bar{\mathbf{q}} = \mathbf{B}\mathbf{P}_\tau\mathbf{u} = \mathbf{B}\mathbf{q}$.

**Proof.** This follows directly from Lemma 3.6 with $\vec{s'}$-degrees restricted to $\tau$. □

**Lemma 3.8.** *If* $\bar{\mathbf{S}}^{(1)}$ *is a* $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})_{2\delta-1}$*-basis, then the matrix* $\mathbf{S}^{(1)}$ *consisting of its first* $n$ *rows is a* $(\mathbf{F}, \sigma, \vec{s})_{2\delta-1}$*-basis.*

**Proof.** By Lemma 3.5, $\mathbf{S}^{(1)}$ has order $(\mathbf{F}, \sigma)$. By Corollary 3.4, any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta-1}$ is a linear combination of $\mathbf{S}^{(1)}$. It remains to show that $\mathbf{S}^{(1)}$ is $\vec{s}$-column reduced.

By Corollary 3.7, $\bar{\mathbf{S}}^{(1)} = \mathbf{B}\mathbf{S}^{(1)}$, and by Lemma 3.5, the columns of $\mathbf{S}^{(1)}$ are in $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta-1}$. Thus, by Lemma 3.2, $\mathbf{S}^{(1)}$ determines the $\vec{s'}$-column degrees of $\mathbf{S}^{(1)}$. Therefore, $\bar{\mathbf{S}}^{(1)}$ being $\vec{s'}$-column reduced implies that $\mathbf{S}^{(1)}$ is $\vec{s}$-column reduced. □

**Lemma 3.9.** *Let* $\bar{\mathbf{S}}^{(12)} = [\bar{\mathbf{S}}^{(1)}, \bar{\mathbf{S}}^{(2)}]$ *be a* $(\mathbf{F}', \vec{\omega}, \vec{s'})_{2\delta}$*-basis, with* $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(1)} \leq 2\delta - 1$ *and* $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(2)} = 2\delta$, *and* $\mathbf{S}^{(12)}, \mathbf{S}^{(1)}, \mathbf{S}^{(2)}$ *the first* $n$ *rows of* $\bar{\mathbf{S}}^{(12)}, \bar{\mathbf{S}}^{(1)}, \bar{\mathbf{S}}^{(2)}$, *respectively. Let* $I$ *be the column rank profile (the lexicographically smallest sequence of column indices that indicates a full column rank submatrix) of* $\mathbf{S}^{(12)}$. *Then the submatrix* $\mathbf{S}^{(12)}_I$ *comprised of the columns of* $\mathbf{S}^{(12)}$ *indexed by* $I$ *is a* $(\mathbf{F}, \sigma, \vec{s})_{2\delta}$*-basis.*

**Proof.** Consider doing $\vec{s}$-column reduction on $\mathbf{S}^{(12)}$. From Lemma 3.8, we know that $\mathbf{S}^{(1)}$ is a $(\mathbf{F}, \sigma, \vec{s})_{2\delta-1}$-basis. Therefore, only $\mathbf{S}^{(2)}$ may be $\vec{s}$-reduced. If a column $\mathbf{c}$ of $\mathbf{S}^{(2)}$ can be further $\vec{s}$-reduced, then it becomes an element of $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta-1}$, which is generated by $\mathbf{S}^{(1)}$. Thus $\mathbf{c}$ must be reduced to zero by $\mathbf{S}^{(1)}$. The only nonzero columns of $\mathbf{S}^{(12)}$ remaining after $\vec{s}$-column reduction are therefore the columns that cannot be $\vec{s}$-reduced. Hence $\mathbf{S}^{(12)}$ $\vec{s}$-reduces to $\mathbf{S}^{(12)}_I$. In addition, $\mathbf{S}^{(12)}_I$ has order $(\mathbf{F}, \sigma)$ as $\mathbf{S}^{(12)}$ has order $(\mathbf{F}, \sigma)$ by Lemma 3.5. From Corollary 3.4 any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta}$ is a linear combination of $\mathbf{S}^{(12)}$ and hence is also a linear combination of $\mathbf{S}^{(12)}_I$. □

To extract $\mathbf{S}^{(12)}_I$ from $\mathbf{S}^{(12)}$, note that doing $\vec{s}$-column reduction on $\mathbf{S}^{(12)}$ is equivalent to the more familiar problem of doing column reduction on $x^{\vec{s}}\mathbf{S}^{(12)}$. As $\mathbf{S}^{(12)}$ $\vec{s}$-column reduces to $\mathbf{S}^{(12)}_I$, this corresponds to determining the column rank profile of the *leading column coefficient matrix* of $x^{\vec{s}}\mathbf{S}^{(12)}$. Recall that the leading column coefficient matrix of a matrix $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_k]$ used for column reduction is

$$\begin{aligned}\operatorname{lcoeff}(\mathbf{A}) &= [\operatorname{lcoeff}(\mathbf{a}_1), \ldots, \operatorname{lcoeff}(\mathbf{a}_k)] \\ &= [\operatorname{coeff}(\mathbf{a}_1, \deg(\mathbf{a}_1)), \ldots, \operatorname{coeff}(\mathbf{a}_k, \deg(\mathbf{a}_k))].\end{aligned}$$

The column rank profile of $\operatorname{lcoeff}(x^{\vec{s}}\mathbf{S}^{(12)})$ can be determined by (the transposed version of) LSP factorization (Ibarra et al., 1982), which factorizes $\operatorname{lcoeff}(x^{\vec{s}}\mathbf{S}^{(12)}) = PSU$ as the product of a permutation matrix $P$, a matrix $S$ with its nonzero columns forming a lower triangular submatrix, and an upper triangular matrix $U$ with 1's on the diagonal. The indices, $I$, of the nonzero columns of $S$ then give $\mathbf{S}^{(12)}_I$ in $\mathbf{S}^{(12)}$.

**Theorem 3.10.** *Let* $\bar{\mathbf{S}} = [\bar{\mathbf{S}}^{(12)}, \bar{\mathbf{S}}^{(3)}]$ *be a* $(\mathbf{F}', \vec{\omega}, \vec{s'})$*-basis, with* $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(12)} \leq 2\delta$ *and* $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(3)} \geq 2\delta + 1$*, and* $\mathbf{S}, \mathbf{S}^{(12)}, \mathbf{S}^{(3)}$ *the first* $n$ *rows of* $\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(12)}, \bar{\mathbf{S}}^{(3)}$*, respectively. If* $I$ *is the column rank profile of* $\mathbf{S}^{(12)}$*, then the submatrix* $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$ *of* $\mathbf{S}$ *is a* $(\mathbf{F}, \sigma, \vec{s})$*-basis.*

**Proof.** By Lemma 3.5, $\mathbf{S}$ has order $(\mathbf{F}, \sigma)$, and so $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$ also has order $(\mathbf{F}, \sigma)$. By Corollary 3.4, any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle$ is a linear combination of the columns of $\mathbf{S}$, and so $\mathbf{q}$ is also a linear combination of the columns of $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$. It only remains to show that $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$ is $\vec{s}$-column reduced.

Let $\mathbf{P}$ be a $(\mathbf{F}, \sigma, \vec{s})$-basis and $\bar{\mathbf{T}}$ be the $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis constructed from $\mathbf{P}$ as in Lemma 3.6. Let $\bar{\mathbf{T}}^{(3)}$ be the columns of $\bar{\mathbf{T}}$ with $\vec{s'}$-degrees greater than $2\delta$, and $\mathbf{P}^{(3)}$ be the columns of $\mathbf{P}$ with $\vec{s}$-degrees greater than $2\delta$. Assume without loss of generality that $\mathbf{S}$, $\mathbf{P}$, and $\bar{\mathbf{T}}$ have their columns sorted according to their $\vec{s}$-degrees and $\vec{s'}$-degrees, respectively. Then $\deg_{\vec{s}} \mathbf{S}^{(3)} \leq \deg_{\vec{s'}} \bar{\mathbf{S}}^{(3)} = \deg_{\vec{s'}} \bar{\mathbf{T}}^{(3)} = \deg_{\vec{s}} \mathbf{P}^{(3)}$. Combining this with the $\vec{s}$-minimality of $\mathbf{S}_I^{(12)}$ from Lemma 3.9, it follows that $\deg_{\vec{s}} [\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}] \leq \deg_{\vec{s}} \mathbf{P}$. This combined with the fact that $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$ still generates $\langle(\mathbf{F}, \sigma, \vec{s})\rangle$ implies that $\deg_{\vec{s}} [\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}] = \deg_{\vec{s}} \mathbf{P}$. Therefore, $[\mathbf{S}_I^{(12)}, \mathbf{S}^{(3)}]$ is a $(\mathbf{F}, \sigma, \vec{s})$-basis. □

**Corollary 3.11.** *Let* $\bar{\mathbf{S}}$ *be a* $(\mathbf{F}', \vec{\omega}, \vec{s'})$*-basis with its columns sorted in an increasing order of their* $\vec{s'}$ *degrees, and* $\mathbf{S}$ *the first* $n$ *rows of* $\bar{\mathbf{S}}$*. If* $J$ *is the column rank profile of* $\operatorname{lcoeff}(x^{\vec{s}}\mathbf{S})$*, then the submatrix* $\mathbf{S}_J$ *of* $\mathbf{S}$ *indexed by* $J$ *is a* $(\mathbf{F}, \sigma, \vec{s})$*-basis.*

**Proof.** This follows directly from Theorem 3.10. □

This rank profile $J$ can be determined by LSP factorization on $\operatorname{lcoeff}(x^{\vec{s}} \cdot \mathbf{S}^{(12)})$.

**Example 3.12.** For the problem in Example 2.5, $\check{\mathbf{F}}$ is given by

$$\begin{bmatrix} x+x^2+x^3+x^4+x^5+x^6 & 1+x+x^5+x^6+x^7 & 1+x^2+x^6+x^7 & 1+x+x^3+x^7 & 0 & 0 \\ 1+x+x^2+x^3 & x^3 & 1+x^2+x^3 & x & 1 & 0 \\ 1+x+x^2 & x+x^2+x^3 & 1+x+x^2+x^3 & x^3 & 0 & 1 \end{bmatrix},$$

and a $\left(\mathbf{F}', [8,4,4], \vec{0}\right)$-basis is given as

$$\left[\begin{array}{cccc|cc} 1 & x & 1 & x^2 & x^2+x^4 & 1+x^2+x^3+x^4 \\ 0 & 1 & x^2+x^3 & 0 & x^3 & 0 \\ 1 & 1+x & x & x^3+x^4 & 0 & x+x^2+x^3 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1+x^2 & x^2 & x^2+x^3 & 1+x^2+x^3+x^4 \\ 0 & 1 & 1 & x^2+x^4 & x^2+x^3 & 1+x^3 \end{array}\right].$$

Column reduction on the top 4 rows gives the top left $4 \times 4$ submatrix, which is a $(\mathbf{F}, 8, \vec{0})$-basis.

The following two lemmas verify Storjohann's result in the case of degree parameter $\delta = \sigma/4$. More specifically, we show that the matrix of the top $n$ rows of a $(\bar{\mathbf{F}}, 2\delta, \vec{s'})_{\delta-1}$-basis is a $(\mathbf{F}, \sigma, \vec{s})_{\delta-1}$-basis, with the transformed input matrix

$$\bar{\mathbf{F}} = \left[\begin{array}{c|cc} \mathbf{F}_0 + \mathbf{F}_1 x^\delta & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{F}_1 + \mathbf{F}_2 x^\delta & \mathbf{I}_m & \mathbf{0} \\ \mathbf{F}_2 + \mathbf{F}_3 x^\delta & \mathbf{0} & \mathbf{I}_m \end{array}\right] \equiv \mathbf{F}' \mod x^{2\delta}. \tag{3.2}$$

**Lemma 3.13.** *If $\bar{\mathbf{q}} \in \langle(\bar{\mathbf{F}}, 2\delta, \vec{s'})\rangle_{\delta-1}$ and $\mathbf{q}$ denotes the first $n$ entries of $\bar{\mathbf{q}}$, then $\bar{\mathbf{q}}$ must have the form*

$$\bar{\mathbf{q}} = \mathbf{B}\mathbf{q} = \begin{bmatrix} \mathbf{q} \\ x^{-\delta}\mathbf{F}_0 \cdot \mathbf{q} \\ x^{-2\delta}\left(\mathbf{F}_0 + \mathbf{F}_1 x^\delta\right) \cdot \mathbf{q} \end{bmatrix}$$

*and $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\delta-1}$.*

**Proof.** Let $\mathbf{q}, \mathbf{q}_2, \mathbf{q}_3$ consist of the top $n$ entries, middle $m$ entries, and bottom $m$ entries, respectively, of $\bar{\mathbf{q}}$ so that

$$\bar{\mathbf{F}}\bar{\mathbf{q}} \equiv \begin{bmatrix} \mathbf{F}_0\mathbf{q} + x^\delta \mathbf{F}_1\mathbf{q} \\ \mathbf{q}_2 + \mathbf{F}_1\mathbf{q} + x^\delta \mathbf{F}_2\mathbf{q} \\ \mathbf{q}_3 + \mathbf{F}_2\mathbf{q} + x^\delta \mathbf{F}_3\mathbf{q} \end{bmatrix} \equiv \mathbf{0} \mod x^{2\delta}. \tag{3.3}$$

From the first and second block rows, we get $\mathbf{F}_0\mathbf{q} + x^\delta\mathbf{F}_1\mathbf{q} \equiv \mathbf{0} \mod x^{2\delta}$ and $\mathbf{q}_2 + \mathbf{F}_1\mathbf{q} \equiv \mathbf{0} \mod x^\delta$, which implies

$$\mathbf{F}_0\mathbf{q} \equiv x^\delta \mathbf{q}_2 \mod x^{2\delta}. \tag{3.4}$$

Similarly, from the second and third rows, we get $\mathbf{q}_2 + \mathbf{F}_1\mathbf{q} + x^\delta\mathbf{F}_2\mathbf{q} \equiv \mathbf{0} \mod x^{2\delta}$ and $\mathbf{q}_3 + \mathbf{F}_2\mathbf{q} \equiv \mathbf{0} \mod x^\delta$, which implies $\mathbf{q}_2 + \mathbf{F}_1\mathbf{q} \equiv x^\delta\mathbf{q}_3 \mod x^{2\delta}$.

Since $\deg \mathbf{q} \le \deg_{\vec{s}} \mathbf{q} = \delta - 1$, we have $\deg \mathbf{F}_0\mathbf{q} \le 2\delta - 2$, hence from (3.4) $\deg \mathbf{q}_2 \le \delta - 2$ and $\mathbf{q}_2 x^\delta = \mathbf{F}_0\mathbf{q}$. Similarly, $\deg \mathbf{q}_3 \le \delta - 2$ and $\mathbf{q}_3 x^{2\delta} = \mathbf{q}_2 x^\delta + \mathbf{F}_1\mathbf{q}x^\delta = \mathbf{F}_0\mathbf{q} + \mathbf{F}_1\mathbf{q}x^\delta$. Substituting this to $\mathbf{F}\mathbf{q} = (\mathbf{F}_0\mathbf{q} + \mathbf{F}_1\mathbf{q}x^\delta) + (\mathbf{F}_2\mathbf{q}x^{2\delta} + \mathbf{F}_3\mathbf{q}x^{3\delta})$, we get $\mathbf{F}\mathbf{q} = \mathbf{q}_3 x^{2\delta} + (\mathbf{F}_2\mathbf{q}x^{2\delta} + \mathbf{F}_3\mathbf{q}x^{3\delta}) \equiv \mathbf{0} \mod x^{4\delta}$ using the bottom block row of (3.3). □

**Lemma 3.14.** *If $\bar{\mathbf{S}}_{\delta-1}$ is a $(\bar{\mathbf{F}}, 2\delta, \vec{s'})_{\delta-1}$-basis, then the matrix of its first $n$ rows, $\mathbf{S}_{\delta-1}$, is a $(\mathbf{F}, \sigma, \vec{s})_{\delta-1}$-basis.*

**Proof.** By Lemma 3.13, $\mathbf{S}_{\delta-1}$ has order $(\mathbf{F}, \sigma)$. Following Lemmas 3.1 and 3.2 and Corollaries 3.3 and 3.4 (replacing $\vec{\omega}$ by $2\delta$), we conclude that any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\delta-1}$ is a linear combination of the columns of $\mathbf{S}_{\delta-1}$. In addition, since $\bar{\mathbf{S}}_{\delta-1} = \mathbf{B}\mathbf{S}_{\delta-1}$ by Lemma 3.13, and the columns of $\mathbf{S}_{\delta-1}$ are in $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\delta-1}$, it follows from Lemma 3.2 that $\mathbf{S}_{\delta-1}$ determines the $\vec{s'}$-column degrees of $\bar{\mathbf{S}}_{\delta-1}$. Hence $\bar{\mathbf{S}}_{\delta-1}$ $\vec{s'}$-column reduced implies that $\mathbf{S}_{\delta-1}$ is $\vec{s}$-column reduced. □

*3.2. More General Results*

Let us now consider an immediate extension of the results in the previous subsection. Suppose that instead of a $(\mathbf{F}, \sigma, \vec{s})$-basis we now want to compute a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis with a Storjohann transformed input matrix

$$\bar{\mathbf{F}}^{(i)} = \left[\begin{array}{c|c} \mathbf{F}_0 + \mathbf{F}_1 x^{\delta^{(i)}} & \mathbf{0}_m \ \cdots \ \cdots \ \mathbf{0}_m \\ \hline \mathbf{F}_1 + \mathbf{F}_2 x^{\delta^{(i)}} & \mathbf{I}_m \\ \mathbf{F}_2 + \mathbf{F}_3 x^{\delta^{(i)}} & \quad \mathbf{I}_m \\ \vdots & \quad\quad \ddots \\ \mathbf{F}_{l^{(i)}-1} + \mathbf{F}_{l^{(i)}} x^{\delta^{(i)}} & \quad\quad\quad \mathbf{I}_m \end{array}\right]_{ml^{(i)} \times (n+m(l^{(i)}-1))}$$

made with degree parameter $\delta^{(i)} = 2^i d$ for some integer $i$ between 2 and $\log(\sigma/d) - 1$, and a shift $\vec{s}^{(i)} = [\vec{s}, 0, \ldots, 0]$ (with $m(l^{(i)}-1)$ 0's), where $l^{(i)} = \sigma/\delta^{(i)} - 1$ is the number of block rows [2]. To apply a transformation analogous to (3.1), we write each $\mathbf{F}_j = \mathbf{F}_{j0} + \mathbf{F}_{j1}\delta^{(i-1)}$ and set

$$\mathbf{F}'^{(i)} = \left[\begin{array}{l|c} \mathbf{F}_{00} + \mathbf{F}_{01} x^{\delta^{(i-1)}} + \mathbf{F}_{10} x^{2\delta^{(i-1)}} + \mathbf{F}_{11} x^{3\delta^{(i-1)}} & \mathbf{0} \\ \hline \mathbf{F}_{01} + \mathbf{F}_{10} x^{\delta^{(i-1)}} & \\ \mathbf{F}_{10} + \mathbf{F}_{11} x^{\delta^{(i-1)}} + \mathbf{F}_{20} x^{2\delta^{(i-1)}} + \mathbf{F}_{21} x^{3\delta^{(i-1)}} & \\ \mathbf{F}_{11} + \mathbf{F}_{20} x^{\delta^{(i-1)}} & \\ \vdots & \mathbf{I} \\ \mathbf{F}_{(l^{(i)}-1)0} + \mathbf{F}_{(l^{(i)}-1)1} x^{\delta^{(i-1)}} + \mathbf{F}_{l^{(i)}0} x^{2\delta^{(i-1)}} + \mathbf{F}_{l^{(i)}1} x^{3\delta^{(i-1)}} & \\ \mathbf{F}_{(l^{(i)}-1)1} + \mathbf{F}_{l^{(i)}0} x^{\delta^{(i-1)}} & \\ \mathbf{F}_{l^{(i)}0} + \mathbf{F}_{l^{(i)}1} x^{\delta^{(i-1)}} & \end{array}\right], \qquad (3.5)$$

and $\vec{\omega}^{(i)} = \left[\left[[2\delta^{(i)}]^m, [\delta^{(i)}]^m\right]^{l^{(i)}}, [\delta^{(i)}]^m\right]$, where $[\circ]^k$ represents $\circ$ repeated $k$ times. The order entries $2\delta^{(i)}$, $\delta^{(i)}$ in $\vec{\omega}^{(i)}$ correspond to the degree $2\delta^{(i)} - 1$, degree $\delta^{(i)} - 1$ rows in $\mathbf{F}'^{(i)}$ respectively. Let

$$\mathbf{E}^{(i)} = \left[\begin{array}{c||c|c|c|c||c} \mathbf{I}_n & & & & & \mathbf{0}_{n\times m} \ \mathbf{0}_{n\times m} \\ \hline\hline & \mathbf{0}_m \ \mathbf{I}_m & & & & \\ \hline & & \mathbf{0}_m \ \mathbf{I}_m & & & \\ \hline & & & \ddots \ \ddots & & \\ \hline & & & & \mathbf{0}_m \ \mathbf{I}_m & \end{array}\right]$$

[2] Recall that $d = m\sigma/n$ is the average degree of the input matrix $\mathbf{F}$ if we treat $\mathbf{F}$ as a square $n \times n$ matrix. Also, $i$ starts at 2 because $i = 1$ is our base case in the computation of an order basis, which may become more clear in the next section. The base case can be computed efficiently using the method of Giorgi et al. (2003) directly and does not require the transformation discussed in this section.

with $l^{(i)}-1$ blocks of $[\mathbf{0}_m, \mathbf{I}_m]$ and hence an overall dimension of $(n+m(l^{(i)}-1)) \times (n+m(l^{(i-1)}-1))$. Thus $\mathbf{E}^{(i)}\mathbf{M}$ picks out from $\mathbf{M}$ the first $n$ rows and the even block rows from the remaining rows except the last block row for a matrix $\mathbf{M}$ with $n+m(l^{(i-1)}-1)$ rows. In particular, if $i = \log(n/m)-1$, then $(\mathbf{F}'^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)}) = (\mathbf{F}', \vec{\omega}, \vec{s'})$, which for $d = m\sigma/n$ gives the problem considered earlier in Subsection 3.1, and $\mathbf{E}^{(i)} = [\mathbf{I}_n, \mathbf{0}_{n\times m}, \mathbf{0}_{n\times m}]$ is used to select the top $n$ rows of a $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis for a $(\mathbf{F}, \sigma, \vec{s})$-basis to be extracted.

We can now state the analog of Corollary 3.11:

**Theorem 3.15.** *Let $\mathbf{S}'^{(i)}$ be a $(\mathbf{F}'^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis with its columns sorted in an increasing order of their $\vec{s}^{(i-1)}$ degrees. Let $\hat{\mathbf{S}}^{(i)} = \mathbf{E}^{(i)}\mathbf{S}'^{(i)}$. Let $J$ be the column rank profile of $\operatorname{lcoeff}(x^{\vec{s}^{(i)}}\hat{\mathbf{S}}^{(i)})$. Then $\hat{\mathbf{S}}^{(i)}_J$ is a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis.*

**Proof.** One can follow the same arguments used before from Lemma 3.1 to Corollary 3.11. Alternatively, this can be derived from Corollary 3.11 by noticing the redundant block rows that can be disregarded after applying transformation (3.1) directly to the input matrix $\bar{\mathbf{F}}^{(i)}$. □

Lemma 3.14 can also be extended in the same way to capture Storjohann's transformation with more general degree parameters:

**Lemma 3.16.** *If $\bar{\mathbf{P}}_1^{(i-1)}$ is a $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})_{\delta^{(i-1)}-1}$-basis, then $\mathbf{E}^{(i)}\bar{\mathbf{P}}_1^{(i-1)}$ is a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})_{\delta^{(i-1)}-1}$-basis and the matrix of the top $n$ rows of $\bar{\mathbf{P}}_1^{(i-1)}$ is a $(\mathbf{F}, \sigma, \vec{s})_{\delta^{(i-1)}-1}$-basis.*

**Proof.** Again, this can be justified as done in Lemma 3.14. Alternatively, one can apply Storjohann's transformation with degree parameter $\delta^{(i-1)}$ to $\bar{\mathbf{F}}^{(i)}$ as in (3.2). The lemma then follows from Lemma 3.14 after noticing the redundant block rows that can be disregarded. □

Notice that if $i = \log(n/m) - 1$, then Theorem 3.15 and Lemma 3.16 specialize to Corollary 3.11 and Lemma 3.14.

## 4. Computation of Order Bases

In this section, we establish a link between two different Storjohann transformed problems by dividing the transformed problem from the previous section into two subproblems and then simplifying the second subproblem. This leads to a recursive method for computing order bases. We also present an equivalent, iterative method for computing order bases. The iterative approach is usually more efficient in practice, as it uses just $O(1)$ iterations in the generic case.

*4.1. Dividing into Subproblems*

In Section 3 we have shown that the problem of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis can be converted to the problem of computing a $(\mathbf{F}', \vec{\omega}, \vec{s'})$-basis and, more generally, that the computation of a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis, a Storjohann transformed problem with degree parameter $\delta^{(i)}$, can be converted to the problem of computing a $(\mathbf{F}'^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis. We now consider dividing the new converted problem into two subproblems.

The first subproblem is to compute a $(\mathbf{F}'^{(i)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis or equivalently a $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis $\bar{\mathbf{P}}^{(i-1)}$, a Storjohann transformed problem with degree parameter $\delta^{(i-1)}$. The second subproblem is computing a $(\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis $\bar{\mathbf{Q}}^{(i)}$ using the residual $\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}$ from the first subproblem along with a degree shift $\vec{t}^{(i-1)} = \deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}^{(i-1)}$. From Theorem 5.1 in (Beckermann and Labahn, 1997) we then know that the product $\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$ is a $(\mathbf{F}'^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis and $\deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)} = \deg_{\vec{t}^{(i-1)}} \bar{\mathbf{Q}}^{(i)}$.

**Example 4.1.** Let us continue with Example 2.5 and Example 3.12 in order to compute a $\left(\mathbf{F}, 8, \vec{0}\right)$-basis (or equivalently a $(\bar{\mathbf{F}}^{(2)}, 8, \vec{0})$-basis). This can be determined by computing a $(\mathbf{F}'^{(2)}, [8,4,4], \vec{0})$-basis as shown in Example 3.12 where we have $\mathbf{F}'^{(2)} = \mathbf{F}'$. Computing a $(\mathbf{F}'^{(2)}, [8,4,4], \vec{0})$-basis can be divided into two subproblems. The first subproblem is computing a $(\bar{\mathbf{F}}^{(1)}, 4, \vec{0})$-basis $\bar{\mathbf{P}}^{(1)}$, the Storjohann partial linearized problem in Example 2.5. The residual

$$\mathbf{F}'^{(2)}\bar{\mathbf{P}}^{(1)} = \begin{bmatrix} 0 & x^8 & x^6+x^9 & x^4+x^6+x^9 & x^6+x^8+x^9+x^{10} & x^5+x^8 \\ 0 & 0 & x^5 & x^4+x^6 & x^4+x^6 & x^5+x^6 \\ 0 & x^4 & x^5 & x^5 & x^4+x^5+x^6 & x^4 \end{bmatrix}$$

is then used as the input matrix for the second subproblem. The shift for the second subproblem $\vec{t}^{(1)} = [0,1,2,3,3,3]$ is the list of column degrees of $\bar{\mathbf{P}}^{(1)}$ and so the second subproblem is to compute a $(\mathbf{F}'^{(2)}\bar{\mathbf{P}}^{(1)}, [8,4,4], [0,1,2,3,3,3])$-basis, which is

$$\bar{\mathbf{Q}}^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & x^2 & x & 1 \\ 0 & 0 & 0 & 0 & x & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x \end{bmatrix}. \tag{4.1}$$

Then $\bar{\mathbf{P}}^{(1)}\bar{\mathbf{Q}}^{(2)}$ gives the $(\mathbf{F}'^{(2)}, [8,4,4], \vec{0})$-basis shown in Example 3.12.

We now show that the dimension of the second subproblem can be significantly reduced. First, the row dimension can be reduced by over a half. Let $\hat{\mathbf{P}}^{(i-1)} = \mathbf{E}^{(i)}\bar{\mathbf{P}}^{(i-1)}$.

**Lemma 4.2.** *A $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)}, \vec{t}^{(i-1)})$-basis is a $(\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis.*

**Proof.** This follows because $\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}$ is a submatrix of $\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}$ after removing rows which already have the correct order $2\delta^{(i-1)}$. □

The column dimension of the second subproblem can be reduced by disregarding the $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})_{\delta^{(i-1)}-1}$-basis which has already been computed. More specifically, after sorting the columns of $\bar{\mathbf{P}}^{(i-1)}$ in an increasing order of their $\vec{s}^{(i-1)}$-degrees, let $[\bar{\mathbf{P}}_1^{(i-1)}, \bar{\mathbf{P}}_2^{(i-1)}] = \bar{\mathbf{P}}^{(i-1)}$ be such that $\deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}_1^{(i-1)} \leq \delta^{(i-1)} - 1$ and $\deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}_2^{(i-1)} \geq \delta^{(i-1)}$. Then $\hat{\mathbf{P}}_1^{(i-1)} = \mathbf{E}^{(i)}\bar{\mathbf{P}}_1^{(i-1)}$ is a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})_{\delta^{(i-1)}-1}$-basis by Lemma 3.16. In the second subproblem, the remaining basis elements of a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis can then be computed without $\bar{\mathbf{P}}_1^{(i-1)}$.

Let $\hat{\mathbf{P}}_2^{(i-1)} = \mathbf{E}^{(i)}\bar{\mathbf{P}}_2^{(i-1)}$, $\vec{b}^{(i-1)} = \deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}_2^{(i-1)}$, $\bar{\mathbf{Q}}_2^{(i)}$ be a $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}_2^{(i-1)}, 2\delta^{(i)}, \vec{b}^{(i-1)})$-basis (or equivalently a $(\mathbf{F}'^{(i)}\bar{\mathbf{P}}_2^{(i-1)}, \vec{\omega}^{(i)}, \vec{b}^{(i-1)})$-basis), and $k^{(i-1)}$ be the column dimension of $\bar{\mathbf{P}}_1^{(i-1)}$. We then have the following result.

**Lemma 4.3.** *The matrix*

$$\bar{\mathbf{Q}}^{(i)} = \begin{bmatrix} \mathbf{I}_{k^{(i-1)}} & \\ & \bar{\mathbf{Q}}_2^{(i)} \end{bmatrix}$$

*is a* $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)}, \vec{t}^{(i-1)})$*-basis (equivalently a* $(\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$*-basis).*

**Proof.** First note that $\bar{\mathbf{Q}}^{(i)}$ has order $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)})$ as

$$\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)} = [\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}_1^{(i-1)}, \bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}_2^{(i-1)}\bar{\mathbf{Q}}_2^{(i)}] \equiv 0 \mod x^{2\delta^{(i)}}.$$

In addition, $\bar{\mathbf{Q}}^{(i)}$ has minimal $\vec{t}^{(i-1)}$ degrees as $\bar{\mathbf{Q}}_2^{(i)}$ is $\vec{b}$-minimal. Hence, by Lemma 2.3, $\bar{\mathbf{Q}}^{(i)}$ is a $(\bar{\mathbf{F}}^{(i)} \cdot \hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)}, \vec{t}^{(i-1)})$-basis. □

Lemma 4.3 immediately leads to the following.

**Lemma 4.4.** *Let* $\hat{\mathbf{S}} = [\hat{\mathbf{P}}_1^{(i-1)}, \hat{\mathbf{P}}_2^{(i-1)}\bar{\mathbf{Q}}_2^{(i)}]$*, and let* $I$ *be the column rank profile of* $\mathrm{lcoeff}(x^{\vec{s}^{(i)}}\hat{\mathbf{S}})$*. Then* $\hat{\mathbf{S}}_I$ *is a* $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$*-basis.*

**Proof.** From Lemma 4.3, $\bar{\mathbf{Q}}^{(i)}$ is a $(\mathbf{F}'^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis and hence $\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$ is a $(\mathbf{F}'^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis. Since $[\hat{\mathbf{P}}_1^{(i-1)}, \hat{\mathbf{P}}_2^{(i-1)}\bar{\mathbf{Q}}_2^{(i)}] = \mathbf{E}^{(i)}\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$, the result follows from Theorem 3.15. □

**Example 4.5.** Continuing with Example 2.5, Example 3.12, and Example 4.1, notice that in the computation of the second subproblem, instead of using $\mathbf{F}'^{(2)}$, $\bar{\mathbf{P}}^{(1)}$, $\bar{\mathbf{Q}}^{(2)}$, and $\bar{\mathbf{P}}^{(1)}\bar{\mathbf{Q}}^{(2)}$, the previous lemmas show that we can just use their submatrices, $\bar{\mathbf{F}}^{(2)}$ the top left $1 \times 4$ submatrix of $\mathbf{F}'^{(2)}$, $\hat{\mathbf{P}}_2^{(1)}$ the top right $4 \times 4$ submatrix of $\bar{\mathbf{P}}^{(1)}$, $\bar{\mathbf{Q}}_2^{(2)}$ the bottom right $4 \times 4$ submatrix of $\bar{\mathbf{Q}}^{(2)}$, and $\hat{\mathbf{P}}_2^{(1)}\bar{\mathbf{Q}}_2^{(2)}$ the top right $4 \times 4$ submatrix of $\bar{\mathbf{P}}^{(1)}\bar{\mathbf{Q}}^{(2)}$of lower dimensions.

Lemma 4.4 gives us a way of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis. We can set $i$ to $\log(n/m) - 1$ so that $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)}) = (\mathbf{F}, \sigma, \vec{s})$, and compute a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis. By Lemma 4.4, this can be divided into two subproblems. The first produces $[\hat{\mathbf{P}}_1^{(i-1)}, \hat{\mathbf{P}}_2^{(i-1)}] = \hat{\mathbf{P}}^{(i-1)} = \mathbf{E}^{(i)}\bar{\mathbf{P}}^{(i-1)}$ from computing a $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis $\bar{\mathbf{P}}^{(i-1)}$. The second subproblem then computes a $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}_2^{(i-1)}, 2\delta^{(i)}, \vec{b}^{(i-1)})$-basis $\bar{\mathbf{Q}}_2^{(i)}$. Note the first subproblem of computing a $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis can again be divided into two subproblems just as before. This can be repeated recursively until we reach the base case with degree parameter $\delta^{(1)} = 2d$. The total number of recursion levels is therefore $\log(n/m) - 1$.

Notice that the transformed matrix $\mathbf{F}'^{(i)}$ is not used explicitly in the computation, even though it is crucial for deriving our results.

### *4.2. The Iterative View*

In this subsection we present our algorithm, which uses an iterative version of the computation discussed above. The iterative version is usually more efficient in practice, considering that the generic case has balanced output that can be computed with just one iteration, whereas the recursive method has to go through $\log(n/m) - 1$ levels of recursion.

Algorithm 1 uses a subroutine OrderBasis, the algorithm from Giorgi et al. (2003), for computing order bases with balanced input. Specifically, $[\mathbf{Q}, \vec{a}] = \text{OrderBasis}(\mathbf{G}, \sigma, \vec{b})$ computes a $(\mathbf{G}, \sigma, \vec{b})$-basis and also returns its $\vec{b}$-column degrees $\vec{a}$. The other subroutine StorjohannTransform is the transformation described in Subsection 2.2.

Algorithm 1 proceeds as follows. In the first iteration, which is the base case of the recursive approach, we set the degree parameter $\delta^{(1)}$ to be twice the average degree $d$ and apply Storjohann's transformation to produce a new input matrix $\bar{\mathbf{F}}^{(1)}$, which has $l^{(1)}$ block rows. Then a $(\bar{\mathbf{F}}^{(1)}, 2\delta^{(1)}, \vec{s}^{(1)})$-basis $\bar{\mathbf{P}}^{(1)}$ is computed. Note this is in fact the first subproblem of computing a $(\bar{\mathbf{F}}^{(2)}, 2\delta^{(2)}, \vec{s}^{(2)})$-basis, which is another Storjohann transformed problem and also the problem of the second iteration. At the second iteration, we work on a new Storjohann transformed problem with the degree doubled and the number of block rows $l^{(2)} = (l^{(1)} - 1)/2$ reduced by over a half. The column dimension is reduced by using the result from the previous iteration. More specifically, we know that the basis $\bar{\mathbf{P}}^{(1)}$ already provides a $(\bar{\mathbf{F}}^{(2)}, 2\delta^{(2)}, \vec{s}^{(2)})_{\delta^{(1)}-1}$-basis $\hat{\mathbf{P}}_1^{(1)}$, which can be disregarded in the remaining computation. The remaining work in the second iteration is to compute a $(\bar{\mathbf{F}}^{(2)}\hat{\mathbf{P}}_2^{(1)}, 2\delta^{(2)}, \vec{b}^{(1)})$-basis $\bar{\mathbf{Q}}^{(2)}$, where $\vec{b}^{(1)} = \deg_{\vec{s}^{(1)}} \bar{\mathbf{P}}_2^{(1)}$, and then to combine it with the result from the previous iteration to form a matrix $[\hat{\mathbf{P}}_1^{(1)}, \hat{\mathbf{P}}_2^{(1)}\bar{\mathbf{Q}}^{(2)}]$ in order to extract a $(\bar{\mathbf{F}}^{(2)}, 2\delta^{(2)}, \vec{s}^{(2)})$-basis $\bar{\mathbf{P}}^{(2)}$.

With a $(\bar{\mathbf{F}}^{(2)}, 2\delta^{(2)}, \vec{s}^{(2)})$-basis computed, we can repeat the same process to use it for computing a $(\bar{\mathbf{F}}^{(3)}, 2\delta^{(3)}, \vec{s}^{(3)})$-basis. Continue, using the computed $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis to compute a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis, until all $n$ elements of a $(\mathbf{F}, \sigma, \vec{s})$-basis have been determined.

## 5. Computational Complexity

In this section, we analyze the computational complexity of Algorithm 1.

**Lemma 5.1.** *Algorithm 1 computes a $(\mathbf{F}, \sigma, \vec{s})$-basis in no more than $\log(n/m) - 1$ iterations.*

**Algorithm 1** FastBasis $(\mathbf{F}, \sigma, \vec{s})$

**Input:** $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$, $\sigma \in \mathbb{Z}_{\geq 0}$,$\vec{s} \in \mathbb{Z}^n$ satisfying $n \geq m$, $n/m$ and $\sigma$ are powers of 2 and $\min(\vec{s}) = 0$
**Output:** a $(\mathbf{F}, \sigma, \vec{s})$-basis $\mathbf{P} \in K[x]^{n \times n}$ and $\deg_{\vec{s}} \mathbf{P}$

1: **if** $2m \geq n$ **then return** OrderBasis $(\mathbf{F}, \sigma, \vec{s})$;
2: $i := 1$; $d := m\sigma/n$; $\delta^{(1)} := 2d$;
3: $\bar{\mathbf{F}}^{(1)} :=$ StorjohannTransform$(\mathbf{F}, \delta^{(1)})$;
4: $l^{(1)} :=$ rowDimension$(\bar{\mathbf{F}}^{(1)})/m$;
5: $\vec{b}^{(0)} := [\vec{s}, 0, \ldots, 0]$; // $m(l_1 - 1)$ 0's
6: $[\bar{\mathbf{P}}^{(1)}, \vec{a}^{(1)}] :=$ OrderBasis$(\bar{\mathbf{F}}^{(1)}, 2\delta^{(1)}, \vec{b}^{(0)})$;
7: Sort the columns of $\bar{\mathbf{P}}^{(i)}$ and $\vec{a}^{(i)}$ by the shifted column degrees $\vec{a}^{(i)} = \deg_{\vec{b}} \bar{\mathbf{P}}^{(i)}$ in increasing order;
8: $\vec{t}^{(i)} := \vec{a}^{(i)}$;
9: $k^{(i)} :=$ number of entries of $\vec{a}^{(i)}$ less than $\delta^{(i)}$;
10: $[\bar{\mathbf{P}}_1^{(i)}, \bar{\mathbf{P}}_2^{(i)}] := \bar{\mathbf{P}}^{(i)}$ with $\bar{\mathbf{P}}_1^{(i)} \in K[x]^{n \times k^{(i)}}$;
11: **while** columnDimension$(\bar{\mathbf{P}}_1^{(i)}) < n$ **do**
12: $i := i + 1$; $\delta^{(i)} := 2\delta^{(i-1)}$; $l^{(i)} := (l^{(i-1)} - 1)/2$;
13: $\bar{\mathbf{F}}^{(i)} :=$ StorjohannTransform$(\mathbf{F}, \delta^{(i)})$;
14: $\hat{\mathbf{P}}_2^{(i-1)} := \mathbf{E}^{(i)} \bar{\mathbf{P}}_2^{(i-1)}$;
15: $\mathbf{G}^{(i)} := \bar{\mathbf{F}}^{(i)} \hat{\mathbf{P}}_2^{(i-1)}$;
16: $\vec{b}^{(i-1)} := \vec{t}^{(i-1)}[k^{(i-1)} + 1 \ldots n + m(l^{(i-1)} - 1)]$;
// w:=v[k..l] means that w receives a slice of v whose indices range from k to l
17: $[\mathbf{Q}^{(i)}, \vec{a}^{(i)}] :=$ OrderBasis$(\mathbf{G}^{(i)}, 2\delta^{(i)}, \vec{b}^{(i-1)})$;
18: Sort the columns of $\mathbf{Q}^{(i)}$ and $\vec{a}^{(i)}$ by $\vec{a}^{(i)} = \deg_{\vec{b}^{(i-1)}} \mathbf{Q}^{(i)}$ in increasing order;
19: $\check{\mathbf{P}}^{(i)} := \hat{\mathbf{P}}_2^{(i-1)} \mathbf{Q}^{(i)}$;
20: $J :=$ the column rank profile of lcoeff$(x^{[\vec{s},0,\ldots,0]}[\mathbf{E}^{(i)} \bar{\mathbf{P}}_1^{(i-1)}, \check{\mathbf{P}}^{(i)}])$;
21: $\bar{\mathbf{P}}^{(i)} := [\mathbf{E}^{(i)} \bar{\mathbf{P}}_1^{(i-1)}, \check{\mathbf{P}}^{(i)}]_J$,
22: $\vec{t}^{(i)} := \deg_{[\vec{s},0,\ldots,0]} \bar{\mathbf{P}}^{(i)}$;
23: $k^{(i)} :=$ number of entries of $\vec{t}^{(i)}$ less than $\delta^{(i)}$;
24: $[\bar{\mathbf{P}}_1^{(i)}, \bar{\mathbf{P}}_2^{(i)}] := \bar{\mathbf{P}}^{(i)}$ with $\bar{\mathbf{P}}_1^{(i)} \in K[x]^{n \times k^{(i)}}$ ;
25: **end while**
26: **return** the top $n$ rows of $\bar{\mathbf{P}}_1^{(i)}$, $\vec{t}^{(i)}[1..n]$;

**Proof.** Each iteration $i$ computes a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis. At iteration $i^* = \log(n/m) - 1$, the degree parameter is $\sigma/2$ and $(\bar{\mathbf{F}}^{(i^*)}, 2\delta^{(i^*)}, \vec{s}^{(i^*)}) = (\mathbf{F}, \sigma, \vec{s})$. □

**Lemma 5.2.** *If the shift $\vec{s} = [0, \ldots, 0]$, then a $(\mathbf{F}, \sigma, \vec{s})_{\delta^{(i)}-1}$-basis (or equivalently a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})_{\delta^{(i)}-1}$-basis) computed at iteration $i$ has at least $n - n/2^i$ elements, and hence at most $n/2^i$ elements remain to be computed. If the shift $\vec{s}$ is balanced, that is, $\max \vec{s} \in O(d)$ assuming $\min \vec{s} = 0$, then the number $n^{(i)}$ of remaining basis elements at iteration $i$ is $O(n/2^i)$.*

**Proof.** The uniform case follows from the idea of Storjohann and Villard (2005) on null space basis computation discussed in Subsection 2.3. For the balanced case, the average column degree is bounded by $cd = cm\sigma/n$ for some constant $c$. The first iteration $\lambda$ such that $\delta^{(\lambda)}$ reaches $cd$ is therefore a constant. That is, $\delta^{(\lambda)} = 2^\lambda d \geq cd > \delta^{(\lambda-1)}$ and hence $\lambda = \lceil \log c \rceil$. By the same argument as in the uniform case, the number of remaining basis elements $n^{(i)} \leq n/2^{i-\lambda} = 2^\lambda(n/2^i) \in O(n/2^i)$ at iteration $i \geq \lambda$. For iterations $i < \lambda$, certainly $n^{(i)} \leq n < 2^\lambda(n/2^i) \in O(n/2^i)$. □

**Theorem 5.3.** *If the shift $\vec{s}$ is balanced with $\min(\vec{s}) = 0$, then Algorithm 1 computes a $(\mathbf{F}, \sigma, \vec{s})$-basis with a cost of $O\left(n^\omega \bar{\mathrm{M}}(d)\log\sigma\right) = O\left(n^\omega d \log d \log\log d \log\sigma\right)) \subset O^\sim\left(n^\omega d\right)$ field operations.*

**Proof.** The computational cost depends on the degree, the row dimension, and the column dimension of the problem at each iteration. The degree parameter $\delta^{(i)}$ is $2^i d$ at iteration $i$. The number of block rows $l^{(i)}$ is $\sigma/\delta^{(i)} - 1$, which is less than $\sigma/(2^i d) = n/(2^i m)$ at iteration $i$. The row dimension is therefore less than $n/2^i$ at iteration $i$.

The column dimension of interest at iteration $i$ is the column dimension of $\hat{\mathbf{P}}_2^{(i-1)}$ (equivalently the column dimension of $\bar{\mathbf{P}}_2^{(i-1)}$), which is the sum of two components, $n^{(i-1)} + (l^{(i-1)} - 1)m$. The first component $n^{(i-1)} \in O(n/2^i)$ by Lemma 5.2. The second component $(l^{(i-1)} - 1)m < n/2^{i-1} - m < n/2^{i-1}$ comes from the size of the identity matrix added in Storjohann's transformation. Therefore, the overall column dimension of the problem at iteration $i$ is $O(n/2^i)$.

At each iteration, the four most expensive operations are the multiplications at line 15 and line 19, the order basis computation at line 17, and extracting the basis at line 20.

The matrices $\bar{\mathbf{F}}^{(i)}$ and $\hat{\mathbf{P}}_2^{(i-1)}$ have degree $O(2^i d)$ and dimensions $O(n/2^i) \times O(n)$ and $O(n) \times O(n/2^i)$. The multiplication cost is therefore $2^i\,\mathrm{MM}(n/2^i, 2^i d)$ field operations, which is bounded by

$$\begin{aligned} 2^i\,\mathrm{MM}(n/2^i, 2^i d) &\in O\left(2^i \left(n/2^i\right)^\omega \bar{\mathrm{M}}(2^i d)\right) \\ &\subset O\left(n^\omega \left(2^i\right)^{1-\omega} \bar{\mathrm{M}}\left(2^i\right) \bar{\mathrm{M}}(d)\right) \qquad (5.1) \\ &\subset O\left(n^\omega \left(2^i\right)^{1-\omega} \left(2^i\right)^{\omega-1} \bar{\mathrm{M}}(d)\right) \qquad (5.2) \\ &\subset O\left(n^\omega \bar{\mathrm{M}}(d)\right). \end{aligned}$$

Equation (5.1) follows from $\bar{\mathrm{M}}(ab) \in O\left(\bar{\mathrm{M}}(a)\bar{\mathrm{M}}(b)\right)$ while equation (5.2) follows from $\bar{\mathrm{M}}(t) \in O(t^{\omega-1})$.

The matrices $\hat{\mathbf{P}}_2^{(i-1)}$ and $\bar{\mathbf{Q}}^{(i)}$ of the second multiplication have the same degree $O(2^i d)$ and dimensions $O(n) \times O(n/2^i)$ and $O(n/2^i) \times O(n/2^i)$ and can also be multiplied with a cost of $O\left(n^\omega \bar{\mathrm{M}}(d)\right)$ field operations. The total cost of the multiplications over $O(\log(n/m))$ iterations is therefore $O\left(n^\omega \bar{\mathrm{M}}(d)\log(n/m)\right)$.

The input matrix $\mathbf{G}^{(i)} = \bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}_2^{(i-1)}$ of the order basis computation problem at iteration $i$ has dimension $O(n/2^i) \times O(n/2^i)$ and the order of the problem is $2\delta^{(i)} \in O(2^i d)$.

Thus, the cost of the order basis computation at iteration $i$ is $O\left(\left(n/2^i\right)^\omega \bar{\mathrm{M}}\left(2^i d\right) \log\left(2^i d\right)\right)$. The total cost over $O(\log(n/m))$ iterations is bounded by

$$\begin{aligned}
&O\left(\sum_{i=1}^{\infty}\left(\left(n/2^i\right)^\omega \bar{\mathrm{M}}\left(2^i d\right)\log\left(2^i d\right)\right)\right)\\
\subset &O\left(\sum_{i=1}^{\infty}\left(\left(n/2^i\right)^\omega \bar{\mathrm{M}}\left(2^i\right)\log\left(2^i\right)\bar{\mathrm{M}}\left(d\right)\log\left(d\right)\right)\right)\\
\subset &O\left(\sum_{i=1}^{\infty}\left(n^\omega\left(2^i\right)^{-\omega}\left(2^i\right)^{\omega-1}\bar{\mathrm{M}}\left(d\right)\log\left(d\right)\right)\right)\\
\subset &O\left(n^\omega\bar{\mathrm{M}}\left(d\right)\log\left(d\right)\sum_{i=1}^{\infty}\left(2^{-i}\right)\right)\\
\subset &O\left(n^\omega\bar{\mathrm{M}}\left(d\right)\log\left(d\right)\right).
\end{aligned}$$

Finally, extracting an order basis by LSP factorization costs $O\left(n^\omega\right)$, which is dominated by the other costs. Combining the above gives

$$O\left(n^\omega\bar{\mathrm{M}}\left(d\right)\log(n/m) + n^\omega\bar{\mathrm{M}}\left(d\right)\log d\right) = O\left(n^\omega\bar{\mathrm{M}}\left(d\right)\log\sigma\right))$$

as the total cost of the algorithm. □

## 6. Unbalanced Shifts

Section 5 shows that Algorithm 1 can efficiently compute a $(\mathbf{F}, \sigma, \vec{s})$-basis when the shift $\vec{s}$ is balanced. When the $\vec{s}$ is unbalanced (something important for example in normal form computation (Beckermann et al., 1999, 2006)), then Algorithm 1 still returns a correct answer but may be less efficient. The possible inefficiency results because there may not be enough partial results from the intermediate subproblems to sufficiently reduce the column dimension of the subsequent subproblem. This is clear from the fact that the column degrees of the output can be much larger and no longer sum up to $O\left(m\sigma\right)$ as in the balanced shift case. The shifted $\vec{s}$-column degrees, however, still behave well. In particular, the total $\vec{s}$-degree increase is still bounded by $m\sigma$ as stated in Lemma 2.6, while the shifted degree of any column can also increase by up to $\sigma$. Recall that Lemma 2.6 states that for any shift $\vec{s}$, there exists a $(\mathbf{F}, \sigma, \vec{s})$-basis still having a total size bounded by $nm\sigma$ which gives hope for efficient computation.

In this section, we describe an algorithm for an important special case of unbalanced shift – when the input shift $\vec{s}$ satisfies the condition:

$$\sum_{i=1}^{n}(\max(\vec{s}) - \vec{s}_i) \leq m\sigma.$$

For simpler presentation, we use the equivalent condition

$$\vec{s} \leq 0 \text{ and } \sum_i -\vec{s}_i \leq m\sigma, \tag{6.1}$$

which can always be obtained from the previous condition by using $\vec{s} - \max\vec{s}$ as the new shift. Note that translating every entry of the shift by the same constant does not change the problem.

In the balanced shift case, a central problem is to find a way to handle unbalanced column degrees of the output order basis. In this section, the unbalanced shift makes row degrees of the output also unbalanced, which is a major problem that needs to be resolved. Here we note a second transformation by Storjohann (2006) which converts the input in such a way that each high degree row of the output becomes multiple rows of lower degrees. We refer to this as Storjohann's second transformation to distinguish it from that described in Subsection 2.2. The transformed problem can then be computed efficiently using Algorithm 1. After the computation, rows can then be combined appropriately to form a basis of the original problem. The method is computationally efficient.

Unfortunately, the bases computed this way are not minimal and hence do not in general produce our reduced order bases. In the following, we describe a transformation that incorporates Storjohann's second transformation and guarantees the minimality of some columns of the output, hence providing a partial order basis. We can then work on the remaining columns iteratively as done in the balanced shift case to compute a full order basis.

Condition (6.1) essentially allows us to locate the potential high degree rows that need to be balanced. In more general cases, we may not know in advance which are the high degree rows that need to be balanced, so our approach given in this section does not work directly. This suggests that one possible future direction to pursue is to find an effective way to estimate the row degree of the result pivot entries. Such an estimate may allow us to apply the method given in this section efficiently for general unbalanced shifts. One example of a case not covered by Condition (6.1) is when the shift $\vec{s} = [0, -n\sigma, -2n\sigma, \ldots, -(n-1)n\sigma]$. This shift makes the resulting order basis close to Hermite normal form but with possibly higher degree non-pivot entries.

### *6.1. Transform to Balanced Shifts*

We now describe the transformation for balancing the high degree rows of the resulting basis. Consider the problem of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis, where the input shift $\vec{s}$ satisfies the conditions (6.1). Let $\alpha, \beta \in \mathbb{Z}_{>0}$ be two parameters. For each shift entry $s_i$ in $\vec{s}$ with $-s_i > \alpha + \beta$, let

$$r_i = \operatorname{rem}(-s_i - \alpha - 1, \beta) + 1$$

be the remainder when $-s_i - \alpha$ is divided by $\beta$, and where $r_i = \beta$ in the case where the remainder is 0, and set

$$q_i = \begin{cases} 1 & \text{if } -s_i \leq \alpha + \beta \\ 1 + (-s_i - \alpha - r_i)/\beta & \text{otherwise} \end{cases}$$

Then, for each $q_i > 1$, we expand the corresponding $i$th column $\mathbf{f}_i$ of $\mathbf{F}$ and shift $s_i$ to

$$\tilde{\mathbf{F}}^{(i)} = \left[ \mathbf{f}_i,\ x^{r_i}\mathbf{f}_i,\ x^{r_i+\beta}\mathbf{f}_i,\ \ldots\ ,\ x^{r_i+(q_i-2)\beta}\mathbf{f}_i \right], \quad \tilde{s}_i = [-\alpha-\beta,\ \ldots\ , -\alpha-\beta]$$

with $q_i$ entries in each case. When $q_i = 1$, the corresponding shift entry and input column remain the same, that is, $\tilde{s}_i = s_i$, and $\tilde{\mathbf{F}}^{(i)} = \mathbf{f}_i$. Then for the transformed problem, the new shift becomes $\bar{s} = [\tilde{s}_1, \ldots, \tilde{s}_{\bar{n}}] \in \mathbb{Z}^{\bar{n}}_{\leq 0}$, and the new input matrix becomes

$\bar{\mathbf{F}} = [\tilde{\mathbf{F}}^{(1)}, \ldots, \tilde{\mathbf{F}}^{(n)}] \in \mathbb{K}[x]^{m \times \bar{n}}$, with the new column dimension $\bar{n}$ satisfies $\bar{n} = \sum_{i=1}^{n} q_i$. Note that every entry of the new shift $\bar{s}$ is an integer from $-\alpha - \beta$ to 0. Let

$$\mathbf{E} = \left[\begin{array}{c|c|c} 1 \; x^{r_1} \; x^{r_1+\beta} \; \cdots \; x^{r_1+(q_1-2)\beta} & & \\ \hline & \ddots & \\ & \quad \ddots & \\ \hline & & 1 \; x^{r_n} \; x^{r_n+\beta} \; \cdots \; x^{r_n+(q_n-2)\beta} \end{array}\right]_{n \times \bar{n}}.$$

Then $\bar{\mathbf{F}} = \mathbf{F}\mathbf{E}$. Storjohann's second transformation is determined by setting $\alpha = -1$, a value not allowed in our transformation (we show later in Theorem 6.10 that this value is not useful in our case). One can verify that the new dimension

$$\bar{n} = \sum_{i=1}^{n} q_i \leq n + \sum_{i=1}^{n} -s_i/\beta \leq m\sigma/\beta + n.$$

Thus by setting $\beta \in \Theta(m\sigma/n) = \Theta(d)$, we can make $\bar{n} \in \Theta(n)$. Furthermore, by also setting $\alpha \in \Theta(d)$, we have a balanced shift problem since

$$\max \bar{s} - \min \bar{s} \leq -\min \bar{s} \leq \alpha + \beta \in \Theta(d).$$

Hence Algorithm 1 can compute a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis with cost $O^{\sim}(n^{\omega} d)$ in this case.

With a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis $\bar{\mathbf{P}} \in \mathbb{K}[x]^{\bar{n} \times \bar{n}}$ computed, let us now consider $\mathbf{E}\bar{\mathbf{P}} \in \mathbb{K}[x]^{n \times \bar{n}}$. While it is easy to see that $\mathbf{E}\bar{\mathbf{P}}$ has order $(\mathbf{F}, \sigma)$ since $\mathbf{F}\mathbf{E}\bar{\mathbf{P}} = \bar{\mathbf{F}}\bar{\mathbf{P}} \equiv 0 \mod x^{\sigma}$, in general it is not a minimal basis (in fact, $\mathbf{E}\bar{\mathbf{P}}$ is not even square). However, our transformation does guarantee that the highest degree columns of $\mathbf{E}\bar{\mathbf{P}}$ having $\vec{s}$-degrees exceed $-\alpha$ are minimal. That is, the columns of $\mathbf{E}\bar{\mathbf{P}}$ whose $\vec{s}$-degrees exceed $-\alpha$ are exactly the columns of a $(\mathbf{F}, \sigma, \vec{s})$-basis whose $\vec{s}$-degrees exceed $-\alpha$. We have therefore correctly computed a partial $(\mathbf{F}, \sigma, \vec{s})$-basis.

**Example 6.1.** Let us use the same input as in Example 2.5, but with shift $\vec{s} = [0, -3, -5, -6]$, and parameters $\alpha = \beta = 1$. Then we get the transformed input

$$\begin{aligned}\bar{\mathbf{F}} = [&x + x^2 + x^3 + x^4 + x^5 + x^6,\ 1 + x + x^5 + x^6 + x^7,\ x + x^2 + x^6 + x^7 + x^8,\\ &1 + x^2 + x^4 + x^5 + x^6 + x^7,\ x + x^3 + x^5 + x^6 + x^7 + x^8,\ x^2 + x^4 + x^6 + x^7 + x^8 + x^9,\\ &x^3 + x^5 + x^7 + x^8 + x^9 + x^{10},\ 1 + x + x^3 + x^7,\ x + x^2 + x^4 + x^8,\\ &x^2 + x^3 + x^5 + x^9,\ x^3 + x^4 + x^6 + x^{10},\ x^4 + x^5 + x^7 + x^{11}]\end{aligned}$$

having 12 components, and $\bar{s} = [0, -2, -2, -2, -2, -2, -2, -2, -2, -2, -2, -2]$. In this case $r_1 = r_2 = r_3 = r_4 = 1$, $q_1 = 1$, $q_2 = 2$, $q_3 = 4$, $q_4 = 5$ and the transformation matrix is

$$\mathbf{E} = \left[\begin{array}{c|cc|cccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & x & x^2 & x^3 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x & x^2 & x^3 & x^4 \end{array}\right].$$

Using the earlier algorithm for balanced shift, we compute a $(\bar{\mathbf{F}}, 8, \bar{s})$-basis

$$\bar{\mathbf{P}} = \left[\begin{array}{cccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hline
x & 1 & 0 & 0 & 1 & 0 & x & 0 & 0 & 0 & x & 0 \\
0 & 0 & 1 & 0 & 0 & x & 1+x & x & x & x & 1 & 0 \\
\hline
x & 1 & 0 & 1 & 1+x & 1 & x & 0 & 0 & 0 & 0 & 1 \\
x & 0 & 1 & 1 & 1+x & 1+x & 1 & x & x & 0 & 0 & 0 \\
x & 0 & 0 & 1 & 1+x & 1+x & 1 & x & 0 & 1 & 0 & 0 \\
x & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 1 & x & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & x & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]$$

with $\bar{s}$-degrees $[-1,-2,-2,-2,-1,-1,-1,-1,-1,-1,-1,0]$. Only the last column has $\bar{s}$-degree exceeding $-\alpha = -1$ and so is the only column guaranteed to give a correct $(\mathbf{F}, 8, \vec{s})$-basis element. Comparing

$$\mathbf{E}\bar{\mathbf{P}} = \left[\begin{array}{cccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
x & 1 & x & 0 & 1 & x^2 & x^2 & x^2 & x^2 & x^2 & 0 & 0 \\
x+x^2+x^3+x^4 & 1 & x & 1+x+x^2+x^3 & 1 & 1+x+x^3 & x^2 & x^2 & x^2 & x^2 & 0 & 1 \\
0 & x & x^2 & 1+x^3+x^4 & x & 1+x^4 & x^3 & x^3 & x^3 & x^3 & 0 & 1
\end{array}\right]$$

to a $(\mathbf{F}, 8, \vec{s})$-basis

$$\mathbf{P} = \left[\begin{array}{cccc}
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
1 & x^2+x^3+x^4 & 1+x+x^2+x^3 & 1 \\
x & x^2 & 1+x^3+x^4 & 1
\end{array}\right]$$

with $\vec{s}$-degrees $[-3,-1,-2,\ 0\ ]$, we see that the last column of $\mathbf{E}\bar{\mathbf{P}}$ is a element of a $(\mathbf{F}, 8, \vec{s})$-basis.

If we set $\alpha = 2, \beta = 1$, then the new transformed problem gives

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

with $\bar{s}$-degrees $[-3, -1, -2, -2, -2, -2, -2, -2, \ 0\ ]$. In this case the second column also has $\bar{s}$-degree exceeding $-\alpha = -2$, and so it is guaranteed to produce another element of a $(\mathbf{F}, 8, \vec{s})$-basis. Computing

$$\mathbf{E}\bar{\mathbf{P}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & x & 1+x & x & x & x & 0 \\ 1 & x^2+x^3+x^4 & 1+x+x^2+x^3 & x & 1+x & x & x & x & 1 \\ x & x^2 & 1+x^3+x^4 & x^2 & x+x^2 & x^2 & x^2 & x^2 & 1 \end{bmatrix},$$

we notice the second column is indeed an element of a $(\mathbf{F}, 8, \vec{s})$-basis.

### *6.2. Correspondence Between the Original Problem and the Transformed Problem*

We now work towards establishing the correspondence between the high degree columns of a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis whose $\bar{s}$-degrees exceed $-\alpha$ and those of a $(\mathbf{F}, \sigma, \vec{s})$-basis whose $\vec{s}$-degrees exceed $-\alpha$. A useful link is provided by the following a matrix .

Set

$$\mathbf{A}_i = \begin{bmatrix} x^{r_i} & & & \\ -1 & x^{\beta} & & \\ & -1 & \ddots & \\ & & \ddots & x^{\beta} \\ & & & -1 \end{bmatrix}_{q_i \times (q_i - 1)} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & & \\ & \ddots & \\ & & \mathbf{A}_n \end{bmatrix}_{\bar{n} \times (\bar{n} - n)}.$$

If $q_i = 1$, $\mathbf{A}_i$ has dimension $1 \times 0$, which just adds a zero row and no column in $\mathbf{A}$.

We now show that for any $\bar{\mathbf{w}} \in \left\langle \left(\bar{\mathbf{F}}, \sigma, \bar{s}\right) \right\rangle$, $\bar{\mathbf{w}}$ can be transformed by $\mathbf{A}$ to one of the two forms that correspond to the original problem and transformed problem. This is made more precise in the following lemma. We then use unimodular equivalence of these two forms to show the equivalence between the high degree part of the result from the transformed problem and that of the original problem.

**Lemma 6.2.** *Let*

$$\bar{\mathbf{w}} = \begin{bmatrix} \bar{\mathbf{w}}_1 \\ \vdots \\ \bar{\mathbf{w}}_n \end{bmatrix} \in \langle (\bar{\mathbf{F}}, \sigma, \bar{s}) \rangle \text{ with } \bar{\mathbf{w}}_i = \begin{bmatrix} \bar{w}_{i,0} \\ \vdots \\ \bar{w}_{i,q_i-1} \end{bmatrix}_{q_i \times 1}.$$

*Then there exists a vector* $\mathbf{u} \in \mathbb{K}[x]^{(\bar{n}-n)\times 1}$ *such that* $\bar{\mathbf{w}} + \mathbf{A}\mathbf{u}$ *has one of the following two forms.*

***(a)*** *The first form is*

$$\mathbf{w}^{[1]} = \begin{bmatrix} \mathbf{w}_1^{[1]} \\ \vdots \\ \mathbf{w}_n^{[1]} \end{bmatrix} \text{ with } \mathbf{w}_i^{[1]} = \begin{bmatrix} w_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{q_i \times 1},$$

*where* $w_i = \bar{w}_{i,0} + \bar{w}_{i,1}x^{r_i} + \bar{w}_{i,2}x^{r_i+\beta} + \cdots + \bar{w}_{i,q_i-1}x^{r_i+(q_i-2)\beta}$.

***(b)*** *The second form is*

$$\mathbf{w}^{[2]} = \begin{bmatrix} \mathbf{w}_1^{[2]} \\ \vdots \\ \mathbf{w}_n^{[2]} \end{bmatrix} \text{ with } \mathbf{w}_i^{[2]} = \begin{bmatrix} w_{i,0} \\ \vdots \\ w_{i,q_i-1} \end{bmatrix},$$

*where* $\deg w_{i,j} < r_i \le \beta$ *when* $j = 0$ *and* $\deg w_{i,j} < \beta$ *when* $j \in \{1, \ldots, q_i - 2\}$. *There is no degree restriction on* $w_{i,q_i-1}$.

**Proof.** The first form is obtained by setting

$$\mathbf{u}^{[1]} = \begin{bmatrix} \mathbf{u}_1^{[1]} \\ \vdots \\ \mathbf{u}_n^{[1]} \end{bmatrix} \text{ with } \mathbf{u}_i^{[1]} = \begin{bmatrix} \bar{w}_{i,1} + \bar{w}_{i,2}x^{\beta} + \bar{w}_{i,3}x^{2\beta} + \cdots + \bar{w}_{i,q_i-1}x^{(q_i-2)\beta} \\ \bar{w}_{i,2} + \bar{w}_{i,3}x^{\beta} + \cdots + \bar{w}_{i,q_i-1}x^{(q_i-3)\beta} \\ \vdots \\ \bar{w}_{i,q_i-1} \end{bmatrix}.$$

Then $\bar{\mathbf{w}} + \mathbf{A}\mathbf{u}^{[1]}$ gives the first form. Note that $\mathbf{u}_i^{[1]}$ is empty if $q_i = 1$ and $\bar{\mathbf{w}}_i = \mathbf{w}_i^{[1]} = [\bar{w}_{i,0}]$ is not changed by the transformation.

The second form can be obtained based on the first form. Let

$$t_{i,j} = \begin{cases} r_i & \text{if } j = 0 \\ \beta & \text{if } j \in \{1, \ldots, q_i - 2\} \end{cases}$$

and write $w_i$ from the first form as

$$w_i = w_{i,0} + w_{i,1}x^{r_i} + w_{i,2}x^{r_i+\beta} + \cdots + w_{i,q_i-1}x^{r_i+(q_i-2)\beta} \tag{6.2}$$

with $\deg w_{i,j} < t_{i,j}$ for $j < q_i - 1$. Note that in general $w_{i,j} \neq \bar{w}_{i,j}$, as $\deg \bar{w}_{i,j}$ may not be less than $t_{i,j}$. Now set

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} \text{ with } \mathbf{v}_i = \begin{bmatrix} w_{i,1} + w_{i,2}x^{\beta} + w_{i,3}x^{2\beta} + \cdots + w_{i,q_i-1}x^{(q_i-2)\beta} \\ w_{i,2} + w_{i,3}x^{\beta} + \cdots + w_{i,q_i-1}x^{(q_i-3)\beta} \\ \vdots \\ w_{i,q_i-1} \end{bmatrix}$$

and $\mathbf{u}^{[2]} = \mathbf{u}^{[1]} - \mathbf{v}$, which comes from the unimodular transformation

$$[\bar{\mathbf{w}}, \mathbf{A}] \left[\begin{array}{c|c} 1 & \\ \hline \mathbf{u}^{[1]} & \mathbf{I} \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline -\mathbf{v} & \mathbf{I} \end{array}\right] = [\bar{\mathbf{w}}, \mathbf{A}] \left[\begin{array}{c|c} 1 & \\ \hline \mathbf{u}^{[1]} - \mathbf{v} & \mathbf{I} \end{array}\right].$$

Then $\mathbf{w}^{[2]} = \bar{\mathbf{w}} + \mathbf{A}\mathbf{u}^{[2]}$ is in the second form. Again note that $\mathbf{v}_i$ and $\mathbf{u}_i^{[2]}$ are empty if $q_i = 1$ and $\mathbf{w}_i^{[2]} = \bar{\mathbf{w}}_i = [\bar{w}_{i,0}]$. □

**Lemma 6.3.** *Let* $\bar{\mathbf{w}} \in \langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$ *and* $\mathbf{w}^{[2]}$ *be in the second form. If* $\deg_{\bar{s}} \mathbf{E}\bar{\mathbf{w}} > -\alpha$ *or* $\deg_{\bar{s}} \mathbf{w}^{[2]} > -\alpha$, *then* $\deg_{\bar{s}} \mathbf{E}\bar{\mathbf{w}} = \deg_{\bar{s}} \mathbf{w}^{[2]}$.

**Proof.** Consider the $i$th entry $w_i$ of $\mathbf{E}\bar{\mathbf{w}}$ and the entries $\mathbf{w}_i^{[2]} = [w_{i,0}, \ldots, w_{i,q_i-1}]^T$ in $\mathbf{w}^{[2]}$. If $q_i = 1$, then $w_i = w_{i,0}$ and the corresponding shifts satisfies $s_i = \bar{s}_{\ell(i)}$, where $\ell(i) = \sum_{k=1}^{i} q_k$. Hence $\deg w_i + s_i = \deg w_{i,0} + \bar{s}_{\ell(i)}$. Thus we only need to consider the case where $q_i > 1$. Write $w_i$ as in Equation (6.2). Note that $\deg w_{i,q_i-1} = \deg w_i - r_i - \beta\,(q_i - 2)$ and hence $\deg w_{i,q_i-1} - \alpha - \beta = \deg w_i - r_i - \alpha - \beta\,(q_i - 1)$, that is, $\deg w_{i,q_i-1} + \bar{s}_{\ell(i)} = \deg w_i + s_i$. It follows that

$$\begin{aligned} \deg_{\bar{s}} \mathbf{E}\bar{\mathbf{w}} &= \max_i(\deg w_i + s_i) = \max_i(\deg w_{i,q_i-1} + \bar{s}_{\ell(i)}) \\ &\leq \max_{i,j}\left(\deg w_{i,j} + \bar{s}_{\ell(i-1)+j+1}\right) = \deg_{\bar{s}} \mathbf{w}^{[2]}. \end{aligned}$$

The only possible indices $j$ where the inequality can be strict occur when $j < q_i - 1$. But $\deg w_{i,j} < \beta$ for all $j < q_i - 1$, which implies $\deg w_{i,j} + \bar{s}_{\ell(i-1)+j+1} = \deg w_{i,j} - \alpha - \beta < -\alpha$, and so it follows that the entries at these indices $j$ do not contribute to $\deg_{\bar{s}} \mathbf{w}^{[2]}$ when $\deg_{\bar{s}} \mathbf{w}^{[2]} > -\alpha$ or $\deg_{\bar{s}} \mathbf{E}\bar{\mathbf{w}} = \max_i(\deg w_{i,q_i-1} + \bar{s}_{\ell(i)}) > -\alpha$. In other words, if one of them exceeds $-\alpha$, then $\deg_{\bar{s}} \mathbf{w}^{[2]}$ and $\deg_{\bar{s}} \mathbf{E}\bar{\mathbf{w}}$ are determined only by entries at indices $j = q_i - 1$, but the equality always holds for these entries. □

**Remark 6.4.** Notice that the first form $\mathbf{w}^{[1]}$ of $\bar{\mathbf{w}}$ has nonzero entries only at indices $I = [1, q_1+1, \ldots, \sum_{k=1}^{n-1} q_k+1]$. Let $\mathbf{B}$ be a $\bar{n}\times n$ matrix with 1's at position $(\sum_{k=1}^{n-1} q_k+1, i)$ and 0's everywhere else. Then the first form satisfies $\mathbf{w}^{[1]} = \mathbf{B}\mathbf{E}\bar{\mathbf{w}}$. Hence Lemma 6.3 provides the degree correspondence between the degrees of the first form $\mathbf{B}\mathbf{E}\bar{\mathbf{w}}$, which is just $\mathbf{E}\bar{\mathbf{w}}$ with zero rows added, and the second form $\bar{\mathbf{w}}^{[2]}$ of $\bar{\mathbf{w}}$.

**Corollary 6.5.** *Let* $\bar{\mathbf{w}} \in \langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$ *and* $\mathbf{w}^{[2]}$ *be its second form. Then* $\deg_{\vec{s}} \mathbf{E}\bar{\mathbf{w}} > -\alpha$ *if and only if* $\deg_{\bar{s}} \mathbf{w}^{[2]} > -\alpha$.

**Proof.** The proof follows directly from Lemma 6.3. □

**Lemma 6.6.** *Let* $\bar{\mathbf{w}} \in \langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$. *Then* $\deg_{\vec{s}} \mathbf{E}\bar{\mathbf{w}} \le \deg_{\bar{s}} \bar{\mathbf{w}}$.

**Proof.** As in Lemma 6.3, consider the $i$th entry $w_i$ of $\mathbf{E}\bar{\mathbf{w}}$ and the corresponding entries $\bar{\mathbf{w}}_i = [\bar{w}_{i,0}, \ldots, \bar{w}_{i,q_i-1}]^T$ in $\bar{\mathbf{w}}$. If $q_i = 1$, then $\deg w_i + s_i = \deg w_{i,0} + \bar{s}_{\ell(i)}$ as before. Thus we just need to consider the case $q_i > 1$, where the shifts for $\bar{\mathbf{w}}_i$ are $-\alpha - \beta$. Since $w_i = \bar{w}_{i,0} + \bar{w}_{i,1}x^{r_i} + \bar{w}_{i,2}x^{r_i+\beta} + \cdots + \bar{w}_{i,q_i-1}x^{r_i+(q_i-2)\beta}$, we get

$$\deg w_i = \max\{\deg \bar{w}_{i,0}, \deg \bar{w}_{i,1} + r_i, \deg \bar{w}_{i,2} + r_i + \beta, \ldots, \deg \bar{w}_{i,q_i-2} + r_i + (q_i-2)\beta\}.$$

Then

$$\begin{aligned}
\deg w_i + s_i &= \deg w_i - r_i - \alpha - \beta(q_i - 1) \\
&= \max\{\deg \bar{w}_{i,0} - r_i - \alpha - \beta(q_i-1),\ \deg \bar{w}_{i,1} - \alpha - \beta(q_i-1),\ \ldots, \\
&\qquad \ldots, \deg \bar{w}_{i,q_i-2} - \alpha - \beta\} \\
&\le \max\{\deg \bar{w}_{i,0} - \alpha - \beta, \deg \bar{w}_{i,1} - \alpha - \beta, \ldots, \deg \bar{w}_{i,q_i-2} - \alpha - \beta\},
\end{aligned}$$

and so $\deg_{\vec{s}} \mathbf{E}\bar{\mathbf{w}} \le \deg_{\bar{s}} \bar{\mathbf{w}}$. □

**Corollary 6.7.** *Let* $\bar{\mathbf{P}} = [\bar{\mathbf{P}}_1, \bar{\mathbf{P}}_2]$ *be a* $(\bar{\mathbf{F}}, \sigma, \bar{s})$*-basis, where* $\deg_{\bar{s}} \bar{\mathbf{P}}_1 \le -\alpha$ *and* $\deg_{\bar{s}} \bar{\mathbf{P}}_2 > -\alpha$. *Let* $\bar{\mathbf{P}}_2^{[2]}$ *be the second form of* $\bar{\mathbf{P}}_2$. *Then* $\deg_{\bar{s}} \bar{\mathbf{P}}_2 = \deg_{\bar{s}} \bar{\mathbf{P}}_2^{[2]} = \deg_{\vec{s}} \mathbf{E}\bar{\mathbf{P}}_2$. *Hence* $[\bar{\mathbf{P}}_1, \bar{\mathbf{P}}_2^{[2]}]$ *is also a* $(\bar{\mathbf{F}}, \sigma, \bar{s})$*-basis.*

**Proof.** Since any column $\bar{\mathbf{p}}$ of $\bar{\mathbf{P}}_2$ satisfies $\deg_{\bar{s}} \bar{\mathbf{p}} > -\alpha$, from Lemma 6.3 and Lemma 6.6, we get

$$\deg_{\bar{s}} \bar{\mathbf{p}}^{[2]} = \deg_{\vec{s}} \mathbf{E}\bar{\mathbf{p}} \le \deg_{\bar{s}} \bar{\mathbf{p}}.$$

The inequality is in fact an equality, since otherwise, $\bar{\mathbf{p}}$ in $\bar{\mathbf{P}}$ can be replaced by $\bar{\mathbf{p}}^{[2]}$ to get a basis of lower degree, contradicting the minimality of $\bar{\mathbf{P}}$. Note that $\bar{\mathbf{P}}$ with its column $\bar{\mathbf{p}}$ replaced by $\bar{\mathbf{p}}^{[2]}$ remains to be a $(\bar{\mathbf{F}}, \sigma, \bar{s})$-basis, since $\bar{\mathbf{p}}^{[2]} = \bar{\mathbf{p}} + \mathbf{A}\mathbf{u}$ involves column operations with only columns in $\bar{\mathbf{P}}_1$ as $\mathbf{A}$ has $\bar{s}$-degrees bounded by $-\alpha$ and hence is generated by $\bar{\mathbf{P}}_1$. □

**Lemma 6.8.** *If* $\mathbf{P}$ *is a* $(\mathbf{F}, \sigma, \vec{s})$*-basis, then* $[\mathbf{BP}, \mathbf{A}]$ *is a basis for* $\langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$.

**Proof.** Any $\bar{\mathbf{w}} \in \langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$ can be transformed by $\mathbf{A}$ to the first form

$$\mathbf{w}^{[1]} = \bar{\mathbf{w}} + \mathbf{A}\mathbf{u}^{[1]} = \mathbf{BE}\bar{\mathbf{w}},$$

where $\mathbf{E}\bar{\mathbf{w}} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle$ is generated by $\mathbf{P}$. That is,

$$\bar{\mathbf{w}} = \mathbf{w}^{[1]} - \mathbf{A}\mathbf{u}^{[1]} = \mathbf{BE}\bar{\mathbf{w}} - \mathbf{A}\mathbf{u}^{[1]} = \mathbf{BP}\mathbf{v} - \mathbf{A}\mathbf{u}^{[1]} = [\mathbf{BP}, \mathbf{A}]\,[\mathbf{v}, -\mathbf{u}^{[1]}]^T.$$

One can also see that the columns of $\mathbf{A}$ and the columns of $\mathbf{BP}$ are linearly independent, as each zero row of $\mathbf{BP}$ has a $-1$ from a column of $\mathbf{A}$. □

**Lemma 6.9.** *If $\bar{\mathbf{P}}$ is a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis, then $\mathbf{E}\bar{\mathbf{P}}$ generates $\langle(\mathbf{F}, \sigma, \vec{s})\rangle$. That is, for any $\mathbf{w} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle$, there is an $\mathbf{u} \in \mathbb{K}[x]^{\bar{n}\times 1}$ such that $\mathbf{w} = \mathbf{E}\bar{\mathbf{P}}\mathbf{u}$.*

**Proof.** For any $(\mathbf{F}, \sigma, \vec{s})$-basis $\mathbf{P}$, the columns of $\mathbf{BP}$ are in $\langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$ generated by $\bar{\mathbf{P}}$, that is, $\mathbf{BP} = \bar{\mathbf{P}}\mathbf{U}$ for some $\mathbf{U} \in \mathbb{K}[x]^{\bar{n}\times n}$. Hence $\mathbf{EBP} = \mathbf{P}$ is generated by $\mathbf{E}\bar{\mathbf{P}}$. That is, $\mathbf{P} = \mathbf{E}\bar{\mathbf{P}}\mathbf{U}$. Then any $\mathbf{w} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle$, which satisfies $\mathbf{w} = \mathbf{P}\mathbf{v}$ for some $\mathbf{v} \in \mathbb{K}[x]^{n\times 1}$, satisfies $\mathbf{w} = \mathbf{E}\bar{\mathbf{P}}\mathbf{U}\mathbf{v}$. □

We are now ready to prove the main result on the correspondence between a high degree part of a basis of the transformed problem and that of the original problem.

**Theorem 6.10.** *Let $\bar{\mathbf{P}} = [\bar{\mathbf{P}}_1, \bar{\mathbf{P}}_2]$ be a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis, where $\deg_{\bar{s}} \bar{\mathbf{P}}_1 \leq -\alpha$ and $\deg_{\bar{s}} \bar{\mathbf{P}}_2 > -\alpha$. Then $\mathbf{E}\bar{\mathbf{P}}_2$ is the matrix of the columns of a $(\mathbf{F}, \sigma, \vec{s})$-basis whose $\vec{s}$-degrees exceed $-\alpha$.*

**Proof.** We want to show that $[\mathbf{P}_1, \mathbf{E}\bar{\mathbf{P}}_2]$ is a $(\mathbf{F}, \sigma, \vec{s})$-basis for any $(\mathbf{F}, \sigma, \vec{s})_{-\alpha}$-basis $\mathbf{P}_1$. First, $\mathbf{E}\bar{\mathbf{P}}$ has order $(\mathbf{F}, \sigma)$ since $\bar{\mathbf{F}}\bar{\mathbf{P}} = \mathbf{F}\mathbf{E}\bar{\mathbf{P}}$ and $\bar{\mathbf{P}}$ has order $\left(\bar{\mathbf{F}}, \sigma\right)$. Also, since $\mathbf{E}\bar{\mathbf{P}}$ generates $\langle(\mathbf{F}, \sigma, \vec{s})\rangle$ by Lemma 6.9, and from Corollary 6.5 $\mathbf{E}\bar{\mathbf{P}}_1$ has $\vec{s}$-degree bounded by $-\alpha$ hence is generated by $\mathbf{P}_1$, it follows that $\left[\mathbf{P}_1, \mathbf{E}\bar{\mathbf{P}}_2\right]$ generates $\langle(\mathbf{F}, \sigma, \vec{s})\rangle$.

It only remains to show that the $\vec{s}$-degrees of $\mathbf{E}\bar{\mathbf{P}}_2$ are minimal. Suppose not, then $[\mathbf{P}_1, \mathbf{E}\bar{\mathbf{P}}_2]$ can be reduced to $[\mathbf{P}_1, \tilde{\mathbf{P}}_2]$ where $\tilde{\mathbf{P}}_2$ has a column having lower $\vec{s}$-degree than that of the corresponding column in $\mathbf{E}\bar{\mathbf{P}}_2$. That is, assuming the columns of $\tilde{\mathbf{P}}_2$ and $\mathbf{E}\bar{\mathbf{P}}_2$ are in non-decreasing $\vec{s}$-degrees order, then we can find the first index $i$ where the $\vec{s}$-degree of $i$th column of $\tilde{\mathbf{P}}_2$ is lower than the $\vec{s}$-degree of the $i$th column of $\mathbf{E}\bar{\mathbf{P}}_2$. It follows that $[\mathbf{BP}_1, \mathbf{BE}\bar{\mathbf{P}}_2]$ can be reduced to $[\mathbf{BP}_1, \mathbf{B}\tilde{\mathbf{P}}_2]$ and $[\mathbf{BP}_1, \mathbf{BE}\bar{\mathbf{P}}_2, \mathbf{A}]$ can be reduced to $[\mathbf{BP}_1, \mathbf{B}\tilde{\mathbf{P}}_2, \mathbf{A}]$. Since $[\mathbf{BP}_1, \mathbf{B}\tilde{\mathbf{P}}_2, \mathbf{A}]$ generates $\langle(\bar{\mathbf{F}}, \sigma, \bar{s})\rangle$ by Lemma 6.8, it can be reduced to $\bar{\mathbf{P}} = [\bar{\mathbf{P}}_1, \bar{\mathbf{P}}_2]$. But it can also be reduced to $[\bar{\mathbf{P}}_1, \tilde{\mathbf{P}}_2^{[2]}, \mathbf{A}]$ with $\tilde{\mathbf{P}}_2^{[2]}$ the second form of $\mathbf{B}\tilde{\mathbf{P}}_2$, and to $[\bar{\mathbf{P}}_1, \tilde{\mathbf{P}}_2^{[2]}]$ as the columns of $\mathbf{A}$ are generated by the $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)_{-\alpha}$-basis $\bar{\mathbf{P}}_1$.

In order to reach a contradiction we just need to show that $\tilde{\mathbf{P}}_2^{[2]}$ has a column with $\bar{s}$-degree less than that of the corresponding column in $\bar{\mathbf{P}}_2$. Let $\tilde{\mathbf{w}}$ be the first column of $\tilde{\mathbf{P}}_2$ with $\vec{s}$-degree less than that of the corresponding column $\mathbf{w}$ in $\mathbf{E}\bar{\mathbf{P}}_2$ and let $\bar{\mathbf{w}}$ be the corresponding column in $\bar{\mathbf{P}}_2$. By Corollary 6.7 $\deg_{\vec{s}} \mathbf{w} = \deg_{\bar{s}} \bar{\mathbf{w}}$. Let $\tilde{\mathbf{w}}^{[2]}$ be the second form of $\mathbf{B}\tilde{\mathbf{w}}$, which is a column in $\tilde{\mathbf{P}}_2^{[2]}$ corresponding to the column $\bar{\mathbf{w}}$ in $\bar{\mathbf{P}}_2$. We know that either $\deg_{\bar{s}} \tilde{\mathbf{w}}^{[2]} \leq -\alpha$ or $\deg_{\bar{s}} \tilde{\mathbf{w}}^{[2]} = \deg_{\vec{s}} \tilde{\mathbf{w}}$ by Lemma 6.3, as $\mathbf{E}\tilde{\mathbf{w}}^{[2]} = \mathbf{E}(\mathbf{B}\tilde{\mathbf{w}} + \mathbf{A}\mathbf{u}) = \tilde{\mathbf{w}}$. In either case, $\deg_{\bar{s}} \tilde{\mathbf{w}}^{[2]} < \deg_{\bar{s}} \bar{\mathbf{w}}$, as $\deg_{\bar{s}} \bar{\mathbf{w}}$ is greater than both $-\alpha$ and $\deg_{\vec{s}} \tilde{\mathbf{w}}$. Hence we have $[\bar{\mathbf{P}}_1, \tilde{\mathbf{P}}_2^{[2]}]$ is another $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis with lower $\bar{s}$-degrees than $\bar{\mathbf{P}}$, contradicting with the minimality of $\bar{\mathbf{P}}$. □

### *6.3. Achieving Efficient Computation*

Theorem 6.10 essentially tells us that a high degree part of a $(\mathbf{F}, \sigma, \vec{s})$-basis can be determined by computing a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis, something we know can be done efficiently. Notice the parallel between the situation here and in the earlier balanced shift case, where the transformed problem also allows us to compute a partial $(\mathbf{F}, \sigma, \vec{s})$-basis, albeit a low degree part, in each iteration.

After a $\left(\bar{\mathbf{F}}, \sigma, \bar{s}\right)$-basis, or equivalently a high degree part of a $(\mathbf{F}, \sigma, \vec{s})$-basis, is computed, for the remaining problem of computing the remaining basis elements, we can in fact reduce the dimension of the input $\mathbf{F}$ by removing some of its columns corresponding to the high shift entries.

**Theorem 6.11.** *Suppose without loss of generality that the entries of $\vec{s}$ are in non-decreasing order. Let $I$ be the index set containing the indices of entries $s_i$ in $\vec{s}$ such that $s_i \leq -\alpha$. Let $\mathbf{F}_I$ be the columns of $\mathbf{F}$ indexed by $I$. Then a $(\mathbf{F}_I, \sigma, \vec{s})_{-\alpha}$-basis $\mathbf{P}_1$ gives a $(\mathbf{F}, \sigma, \vec{s})_{-\alpha}$-basis $\left[\mathbf{P}_1^T, \mathbf{0}\right]^T$.*

**Proof.** For any $\mathbf{p} \in \mathbb{K}[x]^{n\times 1}$ and $\deg_{\vec{s}} \mathbf{p} \leq -\alpha$, note that if the $i$th entry of the shift satisfies $s_i \leq -\alpha$, then the corresponding entry $p_i$ of $\mathbf{p}$ is zero. Otherwise, if $p_i \neq 0$ then the $\vec{s}$-degree of $\mathbf{p}$ is at least $s_i > -\alpha$, contradicting the assumption that the $\vec{s}$-degree of $\mathbf{p}$ is lower than or equal to $-\alpha$. □

Thus, these zero entries do not need to be considered in the remaining problem of computing a $(\mathbf{F}, \sigma, \vec{s})_{-\alpha}$-basis. As such the corresponding columns from the input matrix $\mathbf{F}$ can be removed.

**Example 6.12.** Let us return to Example 6.1. When the parameters $\alpha = \beta = 1$, after computing an element of a $(\mathbf{F}, 8, \vec{s})$-basis with $\vec{s}$-degree 0 that exceeds $-\alpha = -1$, the first row of any $(\mathbf{F}, \sigma, \vec{s})_{-1}$-basis must be zero by Theorem 6.11 (since the first entry of $\vec{s} = [0, -3, -5, -6]$ is $0 > -\alpha$). This is illustrated by the $(\mathbf{F}, 8, \vec{s})$-basis $\mathbf{P}$ given in Example 6.1. This implies that the first column of $\mathbf{F}$ is not needed in the subsequent computation of the remaining basis elements.

**Corollary 6.13.** *If the shift $\vec{s}$ satisfies condition* (6.1) *and $c$ is a constant greater than or equal to* 1*, then a $(\mathbf{F}, \sigma, \vec{s})_{-cd}$-basis has at most $n/c$ basis elements.*

**Proof.** Since $d = m\sigma/n \geq -\sum_{i=1}^{n} s_i/n$ under condition (6.1), there cannot be more than $n/c$ entries of $\vec{s}$ less than or equal to $-cd$. By Theorem 6.11, the only possible nonzero rows of a $(\mathbf{F}, \sigma, \vec{s})_{-cd}$-basis are the ones corresponding to (with the same indices as) the shift entries that are less than or equal to $-cd$. Hence there cannot be more than $n/c$ nonzero rows and at most $n/c$ columns, as the columns are linearly independent. □

We now have a situation similar to that found in the balanced shift case. Namely, for each iteration we transform the problem using appropriate parameters $\alpha$ and $\beta$ to efficiently compute the basis elements with degrees greater than $-\alpha$. Then we can remove columns from the input matrix $\mathbf{F}$ corresponding to the shift entries that are greater than $-\alpha$. We can then repeat the same process again, with a larger $\alpha$ and $\beta$, in order to compute more basis elements.

**Theorem 6.14.** *If the shift $\vec{s}$ satisfies condition* (6.1)*, then a $(\mathbf{F}, \sigma, \vec{s})$-basis can be computed with cost $O\left(n^{\omega}\bar{\mathrm{M}}(d)\log\sigma\right) = O(n^{\omega} d \log d \log\log d \log\sigma) \subset O^{\sim}(n^{\omega}d)$.*

**Proof.** We give the following constructive proof. Initially, we set transformation parameters $\alpha_1 = \beta_1 = 2d$ with $d = m\sigma/n \geq -\sum_{i=1}^{n} s_i/n$. Algorithm 1 works efficiently on the transformed problem as the shift $\bar{s}^{(1)}$ is balanced and the dimension of $\bar{\mathbf{F}}_1$ remains $O(n)$. By Theorem 6.10 this gives the basis elements of $(\mathbf{F}, \sigma, \vec{s})$-basis with $\vec{s}$-degree exceeding $-\alpha_1 = -2d$. By Corollary 6.13, the number of basis elements remaining to be computed is at most $n/2$, hence the number of elements correctly computed is at least $n/2$. By Theorem 6.11, this also allows us to remove at least half of the columns from the input $\mathbf{F}$ and correspondingly at least half of the rows from the output for the remaining problem. Thus the new input matrix $\mathbf{F}_2$ has a new column dimension $n_2 \leq n/2$ and the corresponding shift $\vec{s}^{(2)}$ has $n_2$ entries. The average degree of the new problem is $d_2 = m\sigma/n_2$.

For the second iteration, we set $\alpha_2$ and $\beta_2$ to $2d_2$. Since

$$\alpha_2 = 2m\sigma/n_2 \geq -2\sum_{i=1}^{n} s_i/n_2 \geq -2\sum_{i=1}^{n_2} s_i^{(2)}/n_2,$$

this allows us to reduce the dimension $n_3$ of $\mathbf{F}_3$ to at most $n_2/2$ after finishing computing a $\left(\bar{\mathbf{F}}_2, \sigma, \bar{s}^{(2)}\right)_{-\alpha_1}$-basis. Again, this can be done using Algorithm 1 with a cost of $O\left(n_2^{\omega}\bar{\mathrm{M}}(d_2)\log\sigma\right)$ as the shift $\bar{a}_2$ is balanced and the dimension of $\bar{\mathbf{F}}_2$ is $O(n_2)$. Repeating this process, at iteration $i$, we set $\alpha_i = \beta_i = 2d_i = 2m\sigma/n_i$. The transformed problem has a balanced shift $\bar{a}_i$ and column dimension $O(n_i)$. So a $\left(\bar{\mathbf{F}}_i, \sigma, \bar{s}^{(i)}\right)_{-\alpha_{i-1}}$-basis can be computed with a cost of

$$O\left(n_i^{\omega}\bar{\mathrm{M}}(d_i)\log\sigma\right) \subset O\left(\left(2^{-i}n\right)^{\omega}\bar{\mathrm{M}}\left(2^i d\right)\log\sigma\right) \subset O\left(2^{-i}n^{\omega}\bar{\mathrm{M}}(d)\log\sigma\right).$$

Since

$$\alpha_i = 2m\sigma/n_i \geq -2\sum_{i=1}^{n} s_i/n_i \geq -2\sum_{i=1}^{n_i} s_i^{(i)}/n_i,$$

the column dimension $n_{i+1}$ of the next problem can again be reduced by a half. After iteration $i$, at most $n/2^i$ $(\mathbf{F}, \sigma, \vec{s})$-basis elements remain to be computed. We can stop this process when the column dimension $n_i$ of the input matrix $\mathbf{F}_i$ reaches the row dimension $m$, as an order basis can be efficiently computed in such case. Therefore, a complete $(\mathbf{F}, \sigma, \vec{s})$-basis can be computed in at most $\log(n/m)$ iterations, so the overall cost is

$$O\left(\sum_{i=1}^{\log(n/m)}\left(2^{-i}n^{\omega}\bar{\mathrm{M}}(d)\log\sigma\right)\right) = O\left(n^{\omega}\bar{\mathrm{M}}(d)\log\sigma\sum_{i=1}^{\log(n/m)}2^{-i}\right) \subset O\left(n^{\omega}\bar{\mathrm{M}}(d)\log\sigma\right)$$

field operations. □

Finally, we remark that when the condition (6.1) is relaxed to $\sum_{i=1}^{n} -s_i \in O(m\sigma)$, so that $\sum_{i=1}^{n} -s_i \leq cm\sigma$ for a constant $c$, we can still compute a $(\mathbf{F}, \sigma, \vec{s})$-basis with the same complexity, by setting $\alpha_i = \beta_i = 2cm\sigma/n_i$ at each iteration $i$ and following the same procedure as above. The cost at each iteration $i$ remains $O^{\sim}(n^{\omega}d)$, and the entire computation still uses at most $\log(n/m)$ iterations.

**Algorithm 2** UnbalancedFastBasis $(\mathbf{F}, \sigma, \vec{s})$

Input: $\mathbf{F} \in K[x]^{m \times n}$, $\sigma \in \mathbb{Z}_{\geq 0}$, $\vec{s}$ satisfies condition (6.1).
Output: $\mathbf{P} \in K[x]^{n \times n}$, an $(\mathbf{F}, \sigma, \vec{s})$-basis.
Uses:
(a) TransformUnbalanced : converts an unbalanced shift problem to a balanced one using the transformation described in Section 6. Returns transformed input matrix, transformed shift, and transformation matrix.
(b) FastBasis : computes order basis with balanced shift.

$i := 1$; $\mathbf{P} = [\,]$;
$\mathbf{F}^{(i)} := \mathbf{F}$, $\vec{s}^{(i)} := \vec{s}$;
**while** columnDimension$(\mathbf{P}) \neq n$ **do**
  $d_i = \lceil m\sigma / \text{columnDimension}(\mathbf{F}^{(i)}) \rceil$;
  $\alpha_i := \beta_i := 2d_i$;
  $\left[\bar{\mathbf{F}}^{(i)}, \bar{s}^{(i)}, \mathbf{E}\right] :=$ TransformUnbalanced $\left(\mathbf{F}^{(i)}, \vec{s}^{(i)}, \alpha_i, \beta_i\right)$;
  $\bar{\mathbf{P}}^{(i)} :=$ FastBasis $\left(\bar{\mathbf{F}}^{(i)}, \sigma, \bar{s}^{(i)}\right)$;
  Set $\mathbf{P}^{(i)}$ to be the columns of $\mathbf{E}\bar{\mathbf{P}}^{(i)}$ with $\bar{s}_i$-column degrees in $(-\alpha_i, -\alpha_{i-1}]$;
  $\mathbf{P} := \left[\mathbf{P}^{(i)}, \mathbf{P}\right]$;
  Set $I$ as the set of indices $i$ satisfying $s_i \leq -\alpha_i$;
  $\mathbf{F}^{(i+1)} := \mathbf{F}^{(i)}_I$, $\vec{s}^{(i+1)} := \vec{s}^{(i)}_I$;
  $i := i + 1$;
**end while**
return $\mathbf{P}$ ;

## 7. Future Research

The algorithms in this paper give fast procedures for efficiently computing a large class of order basis problems, including those without shift, those with a balanced shift or with a restricted unbalanced shift. However a number of problems remain to be solved. In particular, the efficient computation of order basis with a general unbalanced shift remains an open problem. In addition, order bases are closely related to many other problems in polynomial matrix computation, for example nullspace basis and matrix normal forms. We are interested in seeing how our tools can be used to solve these problems more efficiently. Our work assumes that we are working with polynomials and power series represented in standard bases. We would like to obtain efficient methods for computation of order bases represented in arbitrary bases, particularly those associated to interpolation bases. Finally, the constructions used in this paper assume exact arithmetic where coefficient growth is not an issue. We are interested in determining how our tools can be used with methods such as fraction-free or modular construction of order bases, particularly combining the constructions found in (Beckermann and Labahn, 2000).

**Acknowledgements** We would like to thank Arne Storjohann and an anonymous referee for their valuable comments.

## References

Baker, G., Graves-Morris, P., 1996. Padé Approximants, 2nd edition. Cambridge.

Beckermann, B., Labahn, G., 1994. A uniform approach for the fast computation of matrix-type Padé approximants. SIAM Journal on Matrix Analysis and Applications 15 (3), 804–823.
URL http://www.cs.uwaterloo.ca/ glabahn/publications.html

Beckermann, B., Labahn, G., 1997. Recursiveness in matrix rational interpolation problems. Journal of Computational and Applied Math 5-34.
URL http://www.cs.uwaterloo.ca/ glabahn/Papers/recursive.pdf

Beckermann, B., Labahn, G., 2000. Fraction-free computation of matrix rational interpolants and matrix GCDs. SIAM Journal on Matrix Analysis and Applications 22 (1), 114–144.
URL http://www.cs.uwaterloo.ca/ glabahn/publications.html

Beckermann, B., Labahn, G., Villard, G., 1999. Shifted normal forms of polynomial matrices. In: Proceedings of the International Symposium on Symbolic and Algebraic Computation. ISSAC'99. pp. 189–196.
URL http://perso.ens-lyon.fr/gilles.villard/BIBLIOGRAPHIE/biblio.html

Beckermann, B., Labahn, G., Villard, G., 2006. Normal forms for general polynomial matrices. Journal of Symbolic Computation 41 (6), 708–737.
URL http://perso.ens-lyon.fr/gilles.villard/BIBLIOGRAPHIE/biblio.html

Giorgi, P., Jeannerod, C.-P., Villard, G., 2003. On the complexity of polynomial matrix computations. In: Proceedings of the International Symposium on Symbolic and Algebraic Computation, Philadelphia, Pennsylvania, USA. ACM Press, pp. 135–142.
URL http://perso.ens-lyon.fr/gilles.villard/BIBLIOGRAPHIE/biblio.html

Ibarra, O., Moran, S., Hui, R., 1982. A generalization of the fast LUP matrix decomposition algorithm and applications. J. Algorithms 3 (1), 45–56.

Labahn, G., 1992. Inversion components for block Hankel-like matrices. Linear Algebra and Its Applications 177, 7–48.

Storjohann, A., 2006. Notes on computing minimal approximant bases. In: Challenges in Symbolic Computation Software. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

Storjohann, A., Villard, G., 2005. 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. pp. 309–316.
URL http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0505030

Van Hoeij, M., November 1997. Factorization of differential operators with rational functions coefficients. Journal of Symbolic Computation 24, 537–561.
URL http://portal.acm.org/citation.cfm?id=271276.271278

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

Zhou, W., Labahn, G., 2009. Efficient computation of order bases. In: Proceedings of the International Symposium on Symbolic and Algebraic Computation. ISSAC'09. ACM, pp. 375–382.
URL http://doi.acm.org/10.1145/1576702.1576753