Efficient Computation of Order Bases

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

ABSTRACT

In this paper we give an efficient algorithm for computation of order basis of a matrix of power series. For a problem with an $m \times n$ input matrix over a field $\mathbb{K}$, $m \leq n$, and order σ, our algorithm uses $O^{\sim}(\mathrm{MM}(n, \lceil m\sigma/n \rceil)) \subset O^{\sim}(n^{\omega} \lceil m\sigma/n \rceil)$ field operations in $\mathbb{K}$, where the soft-O notation $O^{\sim}$ is Big-O with log factors omitted and $\mathrm{MM}(n, d)$ denotes the cost of multiplying two polynomial matrices with dimension n and degree d. The algorithm extends earlier work of Storjohann, whose method can be used to find a subset of an order basis that is within a specified degree bound δ using $O^{\sim}(\mathrm{MM}(n, \delta))$ field operations for $\delta \geq \lceil m\sigma/n \rceil$.

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

General Terms: Algorithms, Theory

Keywords: Order basis, Complexity

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 non-negative integer σ, we say a vector $\mathbf{p} \in \mathbb{K}[x]^{n \times 1}$ of polynomials gives an *order* σ 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 σ terms of $\mathbf{F} \cdot \mathbf{p}$ are zero. Examples of such problems include Padé approximation, Hermite-Padé, Simultaneous Padé, partial realizations of matrix sequences and vector rational reconstruction just to name a few.

The set of all such order $(\mathbf{F}, \sigma)$ approximations form a module over $\mathbb{K}[x]$. An *order basis* - or minimal approximant basis or σ-basis - is a basis of this module having a type of minimal degree property. In the case of rational approximation order bases can be viewed as a natural generalization of the Padé table of a power series [1] since they are able to describe *all* solutions to such problems given particular degree bounds [3]. Order bases are used in such diverse applications as inversion of structured matrices [8], normal forms of matrix polynomials [5, 4], and other important problems in matrix polynomial arithmetic including matrix inversion, determinant and nullspace computation [6, 10].

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 [2] 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 log factors omitted. By working more directly on the input $m \times n$ input matrix, Giorgi, Jeannerod and Villard [6] give a divide and conquer method with cost $O(\mathrm{MM}(n, \sigma) \log \sigma) \subset O^{\sim}(\mathrm{MM}(n, \sigma))$ arithmetic operations, where $\mathrm{MM}(n, \sigma) \in O^{\sim}(n^{\omega}\sigma)$ denotes the cost of multiplying two polynomial matrices with dimension n and degree σ. Their method is nearly optimal if m is close to the size of n but can be improved if m is small.

In a novel construction, Storjohann [9] 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 high 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 σ 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 $O^{\sim}(\mathrm{MM}(n, \lceil m\sigma/n \rceil))$.

However, while order bases of the original problem can have degree up to σ, the nature of Storjohann's conversion limits the degree of 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 be computationally efficient, it 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 [8]. However, in other applications such as those arising in matrix polynomial arithmetic one needs complete basis which specifies all solutions of a given order, not just those within a particular degree bound.

ISSAC'09, July 28–31, 2009, Seoul, Republic of Korea.

In this paper we give an algorithm which computes an entire order basis with a cost of $O(\mathrm{MM}(n, \lceil m\sigma/n \rceil) \log \sigma) \subset O^{\sim}(\mathrm{MM}(n, \lceil m\sigma/n \rceil))$ field operations. 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 Storjohann's transformed problem to another Storjohann's transformed problem of a higher degree, enabling us to use an idea from Storjohann and Villard's null space basis computation [10] 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.

The remaining paper is structured as follows. Some 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. This is followed in the next section by the complexity analysis and then a concluding section which includes topics for future research.

2. PRELIMINARIES

In this section, we provide some of the background needed in order to understand the basic concepts and tools needed for order basis computation. We give the basic definitions and look at the size of the input and the output for computing such bases. We point out the challenges of balancing input and handling unbalanced output. We review construction by Storjohann [9] which transforms the inputs to those having dimensions and degree balance better suited for fast computation. We then discuss an idea from Storjohann and Villard [10] for handling unbalanced output.

2.1 Order Basis

Let $\mathbb{K}$ be a field, $\mathbf{F} \in \mathbb{K}[[x]]^{m \times n}$ 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* $\mathbf{p} \in \mathbb{K}[x]^{n \times 1}$ *has* order $(\mathbf{F}, \vec{\sigma})$ *(or* order $\vec{\sigma}$ *with respect to* $\mathbf{F}$*) if* $\mathbf{F} \cdot \mathbf{p} \equiv \mathbf{0} \bmod x^{\vec{\sigma}}$*, that is,*

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

for some $\mathbf{r} \in \mathbb{K}[[x]]^{m \times 1}$*. 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 module over* $\mathbb{K}[x]$ *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 [3]). While minimality is often given in terms of the degrees alone it is sometimes important to consider this in terms of shifted degrees [5].

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

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

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. It is also equivalent to the notion of $\mathbf{H}$-degree from [3] 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 lowered by column operations. Note that a matrix $\mathbf{P}$ is $\vec{s}$-column reduced if and only if $x^{\vec{s}} \cdot \mathbf{P}$ is column reduced.

DEFINITION 2.2. *An* order basis *[2, 3]* $\mathbf{P}$ *of* $\mathbf{F}$ *with order* $\vec{\sigma}$ *and shift* $\vec{s}$*, or simply a* $(\mathbf{F}, \vec{\sigma}, \vec{s})$*-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.*

Although we allow different orders for each row in this definition, we focus on order basis computation problem with uniform order. However special cases of non-uniform order problems are still needed in our analysis. We also assume $m \le n$ for simplicity. The case of $m > n$ can be transformed to the case of $m \le n$ by compression [10]. We further assume without loss of generality that n/m and σ are powers of two, which can be achieved by padding zero rows to the input matrix and multiplying it by some power of x.

From [3] we have the following.

LEMMA 2.3. *An* $(\mathbf{F}, \vec{\sigma}, \vec{s})$*-basis* $\mathbf{P}$ *satisfies the following properties:*

(a) $\mathbf{P}$ *consists of* n *linearly independent columns, hence* $\mathbf{P}$ *is a square* $n \times n$ *matrix.*

(b) $\mathbf{P}$ *is* $\vec{s}$*-column reduced.*

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

(d) 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* $\mathbf{P}^{-1}\mathbf{q}$*.*

The following also comes from [3]:

LEMMA 2.4. *The following are equivalent for a polynomial matrix* $\mathbf{P}$*:*

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

(b) $\mathbf{P}$ *is comprised of a set of* n *minimal* $\vec{s}$*-degree polynomial vectors that are linearly independent and each having order* $(\mathbf{F}, \vec{\sigma})$*.*

(c) $\mathbf{P}$ *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 δ. Such 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 δ. This is called a *minbasis* by Storjohann in [9].

DEFINITION 2.5. *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 vector with $\vec{s}$-degree bounded by δ. A $(\mathbf{F},\vec{\sigma},\vec{s})_\delta$-basis is a polynomial matrix $\mathbf{P}$ not containing a zero column and satisfying:*

(a) $\mathbf{P}$ has order $(\mathbf{F},\vec{\sigma})$.

(b) Any element of $\langle(\mathbf{F},\vec{\sigma},\vec{s})\rangle_\delta$ can be expressed as a linear combination of the columns of $\mathbf{P}$.

(c) $\mathbf{P}$ is $\vec{s}$-column reduced.

A $(\mathbf{F},\vec{\sigma},\vec{s})_\delta$-basis is, in general, not square unless δ is large enough to contain all n basis elements making it 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 σ 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 [6] directly, which gives a cost of $O(\mathrm{MM}(n,\sigma)\log\sigma) \subset O^\sim(\mathrm{MM}(n,\sigma))$ field operations, close to the cost of multiplying two matrices with dimension n and degree σ. (Note that this cost is independent of the degree shift.) This is optimal within a log factor 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 [9] provides a novel way to transform an order basis problem with small row dimension to a problem with higher row dimension and lower degree to take advantage of Giorgi et al's algorithm. We now provide a quick overview of a slightly modified version of Storjohann's method (The small modification allows nonuniform degree shift for the input, and provides 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 $\deg \mathbf{F}_i \le \delta-1$ for a positive integer δ, and where we assume (without loss of generality) that $\sigma=(l+1)\delta$. Set

$$\bar{\mathbf{F}} = \left[\begin{array}{c|c} \mathbf{F}_0+\mathbf{F}_1x^\delta & \mathbf{0}_{n,m(l-1)} \\ \hline \mathbf{F}_1+\mathbf{F}_2x^\delta & \\ \mathbf{F}_2+\mathbf{F}_3x^\delta & \mathbf{I}_{m(l-1)} \\ \vdots & \\ \mathbf{F}_{l-1}+\mathbf{F}_lx^\delta & \end{array}\right]$$

and $\vec{s'}=[\vec{s},0,\ldots,0]$ ($\vec{s}$ followed by $m(l-1)$ 0's). A $(\bar{\mathbf{F}},2\delta,\vec{s'})$-basis can then be computed by Giorgi et al's method with a cost of $O^\sim(\mathrm{MM}(n,\delta))$ for $\delta \ge \lceil m\sigma/n\rceil$. Note that $\bar{\mathbf{F}}$ has l block rows each containing m rows. We continue to use a block row to represent m rows for the remaining 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 top 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 δ 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 if we treat $\mathbf{F}$ as square. Choosing $\delta=\Theta(\lceil d\rceil)$, the cost of computing a $(\bar{\mathbf{F}},2\delta,\vec{s'})$-basis is then $O^\sim(\mathrm{MM}(n,d)) = O^\sim(\mathrm{MM}(n,\lceil m\sigma/n\rceil))$. The ceiling function here is used to take care of the case of $m\sigma<n$. For the remaining of the paper, we assume that $m\sigma \ge n$ to avoid this case for simplicity. Together with the assumption σ and n/m are both powers of two, $m\sigma/n$ is then always a positive integer in this paper.

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 δ is close to the average degree $d=m\sigma/n$. However, if δ is large, say $\delta=\Theta(\sigma)$, or if we want to compute a complete $(\mathbf{F},\sigma,\vec{s})$-basis, then the computation still cost $O^\sim(\mathrm{MM}(n,\sigma))$.

The underlying difficulty with computing a complete order basis is that the basis can have degree up to σ. As the output of this problem has dimension $n\times n$ and degree up to $\Theta(\sigma)$, this may seem to suggest $O^\sim(\mathrm{MM}(n,\sigma))$ is about the best that can be done. However, the total size of the output is still $O(mn\sigma)$, 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 $\sum_i(\vec{t}_i-\vec{s}_i) \le m\sigma$ and $\max_i(\vec{t}_i-\vec{s}_i)\le\sigma$. In addition, the total size of any $(\mathbf{F},\sigma,\vec{s})$-basis is bounded by $nm\sigma$.*

PROOF. This can be shown by considering the sizes of the pivots in the iterative order basis computation [2, 6]. □

As a result, the average degree of the entries of the output matrix is also bounded by $d=m\sigma/n$.

Let us now look at the average column degree of the output. We assume without loss of generality that $\min(\vec{s})=0$ so $\deg\mathbf{q} \le \deg_{\vec{s}}\mathbf{q}$ for any $\mathbf{q}\in\mathbb{K}^n[x]$. The situation is simpler if the shift $\vec{s}$ is uniform, then $\sum_i \vec{t}_i \le m\sigma$ by Lemma 2.6, and the average column degree is therefore bounded by $d=m\sigma/n$. In this paper, we consider a slightly more general case – when the shift $\vec{s}$ is *balanced,* that is, when $\max\vec{s}\in O(d)$. In this case, Lemma 2.6 implies $\sum_i(\vec{t}_i) \le m\sigma+\sum_i(\vec{s}_i)\in O(m\sigma+nd)=O(m\sigma)$, 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 σ while its average column degree is $O(m\sigma/n)$ implies that an order basis can have quite unbalanced column degrees, especially if m small. A similar problem with unbalanced output is encountered in null space basis computation. Storjohann and Villard in [10] deal with this in the following way.

Let d be the average column degree of the output. Set the degree parameter δ to twice of d which allows one to compute at least half the columns of a basis (since the number of columns with degree at least δ must be at most a half of the total number of columns). One then simplify the problem so 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 δ 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.

In this paper, we discuss a way to compute order basis involving the ideas of [10] and [9].

3. EXTENDING STORJOHANN'S TRANSFORMATION

In this section, we discuss a transformation that can be viewed as an extension of Storjohann's transformation for a full order basis to be computed. 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, which extends immediately to the more general results, as discussed towards the end of this section.

Consider the problem of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis. We assume $\sigma = 4\delta$ for a positive integer δ and write the input matrix $\mathbf{F}$ as $\mathbf{F} = \mathbf{F}_0 + \mathbf{F}_1 x^\delta + \mathbf{F}_2 x^{2\delta} + \mathbf{F}_3 x^{3\delta}$ with $\deg \mathbf{F}_i \leq \delta - 1$. In the following, we show computing a $(\mathbf{F}, \sigma, \vec{s})$-basis can be done by computing a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis of $\check{\mathbf{F}} =$

$$\left[\begin{array}{cc} \mathbf{F} & \mathbf{0} \\ \check{\mathbf{F}}_{21} & \check{\mathbf{F}}_{22} \end{array}\right] = \left[\begin{array}{c|cc} \mathbf{F}_0 + \mathbf{F}_1 x^\delta + \mathbf{F}_2 x^{2\delta} + \mathbf{F}_3 x^{3\delta} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{F}_1 + \mathbf{F}_2 x^\delta & \mathbf{I}_m & \mathbf{0} \\ \mathbf{F}_2 + \mathbf{F}_3 x^\delta & \mathbf{0} & \mathbf{I}_m \end{array}\right] \tag{1}$$

with order $\vec{\omega} = [4\delta, \ldots, 4\delta, 2\delta, \ldots, 2\delta]$ (with m 4δ's and $2m$ 2δ'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 (Storjohann used $e = 1$ in [9]). In fact, it is quite easy to construct a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis from a $(\mathbf{F}, \sigma, \vec{s})$-basis, as we show later in Lemma 3.6. It requires more work to extract a $(\mathbf{F}, \sigma, \vec{s})$-basis from a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis, which is addressed eventually in Corollary 3.11.

Let

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_n \\ \left(1/x^\delta\right) \mathbf{F}_0 \\ \left(1/x^{2\delta}\right) \left(\mathbf{F}_0 + \mathbf{F}_1 x^\delta\right) \end{bmatrix}.$$

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

PROOF. The lemma follows from $\check{\mathbf{F}}\mathbf{Bq} =$

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

Note that the bottom rows of $\mathbf{B}$ may not be polynomials, but $\mathbf{Bq}$ is a polynomial vector as $\mathbf{q} \in \langle(\mathbf{F}, \sigma)\rangle$ 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 $\deg_{\vec{s'}} \mathbf{Bq}$ 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 $\deg_{\vec{s'}} \mathbf{Bq} = \deg_{\vec{s}} \mathbf{q}$.*

PROOF. First from our assumption $s_i \geq 0$ note that $\deg \mathbf{q} \leq \deg_{\vec{s}} \mathbf{q}$. Now consider the degree of the bottom $2m$ entries $\mathbf{q}_1, \mathbf{q}_2$ of

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{q}_2 \\ \mathbf{q}_3 \end{bmatrix} = \mathbf{Bq} = \begin{bmatrix} \mathbf{q} \\ \left(1/x^\delta\right) \mathbf{F}_0 \cdot \mathbf{q} \\ \left(1/x^{2\delta}\right) \left(\mathbf{F}_0 + \mathbf{F}_1 x^\delta\right) \cdot \mathbf{q} \end{bmatrix}.$$

Our goal is to show $\deg_{\vec{e}} \left[\mathbf{q}_2^T, \mathbf{q}_3^T\right]^T \leq \deg_{\vec{s}} \mathbf{q}$. Since $\deg \mathbf{q}_2 = \deg\left(\mathbf{F}_0 \mathbf{q}/x^\delta\right) \leq \deg \mathbf{q} + \delta - 1 - \delta \leq \deg_{\vec{s}} \mathbf{q} - 1$, and similarly $\deg \mathbf{q}_3 \leq \deg_{\vec{s}} \mathbf{q} - 1$, it follows that $\deg_{\vec{e}} \left[\mathbf{q}_2^T, \mathbf{q}_3^T\right]^T = \deg \left[\mathbf{q}_2^T, \mathbf{q}_3^T\right]^T + e \leq \deg_{\vec{s}} \mathbf{q} - 1 + e \leq \deg_{\vec{s}} \mathbf{q}$. □

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

PROOF. This follows from Lemma 3.1 and Lemma 3.2. □

COROLLARY 3.4. *Let $\bar{\mathbf{S}}_\tau$ be a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})_\tau$-basis and $\mathbf{S}_\tau$ the top n rows of $\bar{\mathbf{S}}_\tau$ for any degree 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, $\mathbf{Bq} \in \langle(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})\rangle_\tau$, hence is a linear combination of columns of $\bar{\mathbf{S}}_\tau$. I.e., there exists a polynomial vector $\mathbf{u}$ such that $\mathbf{Bq} = \bar{\mathbf{S}}_\tau \mathbf{u}$. This remains true if we restrict the equation to the top n rows, that is, $\mathbf{q} = [\mathbf{I}_n, \mathbf{0}]\, \mathbf{Bq} = [\mathbf{I}_n, \mathbf{0}]\, \bar{\mathbf{S}}_\tau \mathbf{u} = \mathbf{S}_\tau \mathbf{u}$. □

LEMMA 3.5. *Let $\bar{\mathbf{q}} \in \langle(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})\rangle_\tau$ 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 $\check{\mathbf{F}}\mathbf{q} =$

$$\left[\begin{array}{cc} \mathbf{F} & \mathbf{0} \\ \check{\mathbf{F}}_{21} & \check{\mathbf{F}}_{22} \end{array}\right] \left[\begin{array}{c} \mathbf{q}_1 \\ \mathbf{q}_2 \end{array}\right] = \begin{bmatrix} \mathbf{F}\mathbf{q}_1 \\ \check{\mathbf{F}}_{21}\mathbf{q}_1 + \check{\mathbf{F}}_{22}\mathbf{q}_2 \end{bmatrix} \equiv \mathbf{0} \mod x^{\vec{\omega}}$$

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

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

$$\begin{aligned} \bar{\mathbf{T}} &= \left[\mathbf{BP} \middle| \begin{array}{c} \mathbf{0}_{n,2m} \\ x^{2\delta}\mathbf{I}_{2m} \end{array}\right] \\ &= \left[\begin{array}{c|cc} \mathbf{P} & \mathbf{0}_{n,m} & \mathbf{0}_{n,m} \\ \hline \left(1/x^\delta\right) \mathbf{F}_0 \cdot \mathbf{P} & x^{2\delta}\mathbf{I}_m & \mathbf{0}_m \\ \left(1/x^{2\delta}\right) \left(\mathbf{F}_0 + \mathbf{F}_1 x^\delta\right) \cdot \mathbf{P} & \mathbf{0}_m & x^{2\delta}\mathbf{I}_m \end{array}\right] \end{aligned}$$

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

PROOF. $\bar{\mathbf{T}}$ has order $(\check{\mathbf{F}}, \vec{\omega})$ by Lemma 3.1 and $\bar{\mathbf{T}}$ is $\vec{s'}$-column reduced as $\mathbf{P}$ dominates the $\vec{s'}$-degrees of $\bar{\mathbf{T}}$ on the left side by Lemma 3.2. It remains to show that any $\bar{\mathbf{q}} \in \langle(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})\rangle$ 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 $\mathbf{P}$. I.e., $\mathbf{q} = \mathbf{Pu}$ 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

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

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

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

which requires $\mathbf{v}$ to have order $(\mathbf{I}_{2m}, 2\delta)$. 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. Let $\mathbf{P}_\tau \in \mathbb{K}[x]^{n \times k}$ be a $(\mathbf{F}, \sigma, \vec{s})_\tau$-basis. Let $\bar{\mathbf{q}} \in \langle(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})\rangle_\tau$ and $\mathbf{q}$ be the top n entries of $\bar{\mathbf{q}}$. Then $\bar{\mathbf{q}}$ must have the form $\bar{\mathbf{q}} = \mathbf{BP}_\tau\mathbf{u} + x^{2\delta}[\mathbf{0}, \mathbf{v}^T]^T = \mathbf{Bq} + x^{2\delta}[\mathbf{0}, \mathbf{v}^T]^T$ for some polynomial vector $\mathbf{u} \in \mathbb{K}[x]^{k \times 1}$ and $\mathbf{v} \in \mathbb{K}[x]^{2m \times 1}$. In particular, if $\deg_{\vec{s'}} \bar{\mathbf{q}} < 2\delta$, then $\bar{\mathbf{q}} = \mathbf{BP}_\tau\mathbf{u} = \mathbf{Bq}$.*

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

LEMMA 3.8. *If $\bar{\mathbf{S}}^{(1)}$ is a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})_{2\delta-1}$-basis, then its first n rows $\mathbf{S}^{(1)}$ 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 $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta-1}$, hence by Lemma 3.2 $\mathbf{S}^{(1)}$ determines the $\vec{s'}$-column degrees of $\mathbf{S}^{(1)}$. Therefore, the $\vec{s'}$-column reducedness of $\bar{\mathbf{S}}^{(1)}$ 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 $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})_{2\delta}$-basis, with $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(2)} = 2\delta$ and $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(1)} \leq 2\delta - 1$. Let $\mathbf{S}^{(12)}, \mathbf{S}^{(1)}$, $\mathbf{S}^{(2)}$ be 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 lowest column indices that indicate a full column rank submatrix) of $\mathbf{S}^{(12)}$, which contains all columns of $\mathbf{S}^{(1)}$ by Lemma 3.8. Then the submatrix $\mathbf{S}_I^{(12)}$ 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 already 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 $\mathbf{s}$ 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 $\mathbf{s}$ 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 $\mathbf{S}_I^{(12)}$. In addition, $\mathbf{S}_I^{(12)}$ has order $(\mathbf{F}, \sigma)$ as $\mathbf{S}^{(12)}$ has order $(\mathbf{F}, \sigma)$ by Lemma 3.5, and by Corollary 3.4 any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{2\delta}$ is a linear combination of $\mathbf{S}^{(12)}$ hence also a linear combination of $\mathbf{S}_I^{(12)}$. □

To extract $\mathbf{S}_I^{(12)}$ from $\mathbf{S}^{(12)}$, recall 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 $\mathbf{S}_I^{(12)}$, 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

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

The column rank profile of $\text{lcoeff}(x^{\vec{s}}\mathbf{S}^{(12)})$ can be determined by (the transposed version of) LSP factorization[7], which factorizes $\text{lcoeff}(x^{\vec{s}}\mathbf{S}^{(12)}) = PSU$ 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. Then the indices I of the nonzero columns of S indicate $\mathbf{S}_I^{(12)}$ in $\mathbf{S}^{(12)}$.

THEOREM 3.10. *Let $\bar{\mathbf{S}} = [\bar{\mathbf{S}}^{(12)}, \bar{\mathbf{S}}^{(3)}]$ be a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis, with $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(12)} \leq 2\delta$ and $\deg_{\vec{s'}} \bar{\mathbf{S}}^{(3)} \geq 2\delta+1$. Let $\mathbf{S}, \mathbf{S}^{(12)}, \mathbf{S}^{(3)}$ be the first n rows of $\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(12)}, \bar{\mathbf{S}}^{(3)}$ respectively. Let I be 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 columns $\mathbf{S}$, 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 $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-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 2δ, and $\mathbf{P}^{(3)}$ be the columns of $\mathbf{P}$ with $\vec{s}$-degrees greater than 2δ. 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 we have $\deg_{\vec{s}} \mathbf{S}^{(3)} \leq \deg_{\vec{s'}} \bar{\mathbf{S}}^{(3)} = \deg_{\vec{s'}} \bar{\mathbf{T}}^{(3)} = \deg_{\vec{s}} \mathbf{P}^{(3)}$. 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 $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis with its columns sorted in an increasing order of their $\vec{s'}$ degrees, and $\mathbf{S}$ the first n rows of $\bar{\mathbf{S}}$. Let J be the column rank profile of $\text{lcoeff}(x^{\vec{s}}\mathbf{S})$. 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 the Theorem 3.10. □

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

The following two lemmas verify Storjohann's result in the case of degree parameter $\delta = \sigma/4$. More specifically, we show that 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 Storjohann transformed input matrix

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

LEMMA 3.12. *If $\bar{\mathbf{q}} \in \langle(\bar{\mathbf{F}}, 2\delta, \vec{s'})\rangle_{\delta-1}$ and $\mathbf{q}$ is the first n entries of $\bar{\mathbf{q}}$, then $\bar{\mathbf{q}}$ must have the form*

$$\bar{\mathbf{q}} = \mathbf{B}\mathbf{q} = \left[\begin{array}{c} \mathbf{q} \\ (1/x^\delta)\,\mathbf{F}_0 \cdot \mathbf{q} \\ (1/x^{2\delta})\,(\mathbf{F}_0 + \mathbf{F}_1 x^\delta) \cdot \mathbf{q} \end{array}\right]$$

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

PROOF. Let $\mathbf{q}, \mathbf{q}_2, \mathbf{q}_3$ be be the top n entries, middle m entries, and bottom m entries, respectively, of $\bar{\mathbf{q}}$ so

$$\bar{\mathbf{F}}\bar{\mathbf{q}} \equiv \left[\begin{array}{c} \mathbf{F}_0\mathbf{q} + x^\delta \mathbf{F}_1 \mathbf{q} \\ \mathbf{q}_2 + \mathbf{F}_1\mathbf{q} + x^\delta \mathbf{F}_2 \mathbf{q} \\ \mathbf{q}_3 + \mathbf{F}_2\mathbf{q} + x^\delta \mathbf{F}_3 \mathbf{q} \end{array}\right] \equiv \mathbf{0} \mod x^{2\delta}. \quad (3)$$

From the first and the 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

$$\mathbf{F}_0\mathbf{q} \equiv x^\delta \mathbf{q}_2 \mod x^{2\delta}. \quad (4)$$

Similarly, from the second row and the third row, 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 $\deg \mathbf{q} \leq \deg_{\vec{s}} \mathbf{q} = \delta - 1$, we have $\deg \mathbf{F}_0\mathbf{q} \leq 2\delta - 2$, hence from (4) $\deg \mathbf{q}_2 \leq \delta - 2$ and $\mathbf{q}_2 x^\delta = \mathbf{F}_0\mathbf{q}$. Similarly, $\deg \mathbf{q}_3 \leq \delta - 2$ 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). □

LEMMA 3.13. *If $\bar{\mathbf{S}}_{\delta-1}$ is a $(\bar{\mathbf{F}}, 2\delta, \vec{s'})_{\delta-1}$-basis, then its first n rows $\mathbf{S}_{\delta-1}$ is a $(\mathbf{F}, \sigma, \vec{s})_{\delta-1}$-basis.*

PROOF. By Lemma 3.12, $\mathbf{S}_{\delta-1}$ has order $(\mathbf{F}, \sigma)$. Following Lemma 3.1, Lemma 3.2, Corollary 3.3, Corollary 3.4 but replace $\vec{\omega}$ by 2δ, we can conclude that any $\mathbf{q} \in \langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\delta-1}$ is a linear combination of $\mathbf{S}_{\delta-1}$. In addition, since $\bar{\mathbf{S}}_{\delta-1} = \mathbf{B}\mathbf{S}_{\delta-1}$ by Lemma 3.12 and $\mathbf{S}_{\delta-1}$ are in $\langle(\mathbf{F}, \sigma, \vec{s})\rangle_{\delta-1}$, it follows that $\mathbf{S}_{\delta-1}$ determines the $\vec{s'}$-column degrees of $\bar{\mathbf{S}}_{\delta-1}$ by Lemma 3.2, hence the $\vec{s'}$-column reducedness of $\bar{\mathbf{S}}_{\delta-1}$ implies that $\mathbf{S}_{\delta-1}$ is $\vec{s}$-column reduced. □

Let us now consider an immediate extension of the above results. 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

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

constructed 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. To apply a transformation analogous to (1), we write each $\mathbf{F}_j = \mathbf{F}_{j0} + \mathbf{F}_{j1}\delta^{(i-1)}$ and set $\check{\mathbf{F}}^{(i)} =$

$$\left[\begin{array}{l|l} \mathbf{F}_{00} + \mathbf{F}_{01}x^{\delta^{(i-1)}} + \mathbf{F}_{10}x^{2\delta^{(i-1)}} + \mathbf{F}_{11}x^{3\delta^{(i-1)}} & \mathbf{0} \\ \hline \mathbf{F}_{01} + \mathbf{F}_{10}x^{\delta^{(i-1)}} & \\ \mathbf{F}_{10} + \mathbf{F}_{11}x^{\delta^{(i-1)}} + \mathbf{F}_{20}x^{2\delta^{(i-1)}} + \mathbf{F}_{21}x^{3\delta^{(i-1)}} & \\ \mathbf{F}_{11} + \mathbf{F}_{20}x^{\delta^{(i-1)}} & \\ \vdots & \mathbf{I} \\ \mathbf{F}_{(l^{(i)}-1)0} + \mathbf{F}_{(l^{(i)}-1)1}x^{\delta^{(i-1)}} + \mathbf{F}_{l^{(i)}}x^{\delta^{(i)}} & \\ \mathbf{F}_{(l^{(i)}-1)1} + \mathbf{F}_{l^{(i)}0}x^{\delta^{(i-1)}} & \\ \mathbf{F}_{l^{(i)}0} + \mathbf{F}_{l^{(i)}1}x^{\delta^{(i-1)}} & \end{array}\right], \tag{5}$$

and $\vec{\omega}^{(i)} = \left[\left[[2\delta^{(i)}]^m, [\delta^{(i)}]^m\right]^{l^{(i)}}, [\delta^{(i)}]^m\right]$, where $[\circ]^k$ represent $\circ$ 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 $\check{\mathbf{F}}^{(i)}$ respectively. Let $\mathbf{E}^{(i)} =$

$$\left[\begin{array}{c|cc|cc|cc|cc} \mathbf{I}_n & & & & & & & \mathbf{0}_{n,m} & \mathbf{0}_{n,m} \\ \hline & \mathbf{0}_m & \mathbf{I}_m & & & & & & \\ \hline & & & \mathbf{0}_m & \mathbf{I}_m & & & & \\ \hline & & & & & \ddots & \ddots & & \\ \hline & & & & & & & \mathbf{0}_m & \mathbf{I}_m \end{array}\right]$$

with $l^{(i)}-1$ blocks of $[\mathbf{0}_m, \mathbf{I}_m]$ and hence an overall dimension of $(n + m(l^{(i)} - 1)) \times (n + m(l^{(i-1)} - 1))$, so $\mathbf{E}^{(i)}\mathbf{M}$ 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 $(\check{\mathbf{F}}^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)}) = (\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$, which gives the problem considered earlier in this section, and $\mathbf{E}^{(i)} = [\mathbf{I}_n, \mathbf{0}_{n,m}, \mathbf{0}_{n,m}]$ is used to pick up the top n rows of a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis for a $(\mathbf{F}, \sigma, \vec{s})$-basis to be extracted.

We can now state the analog of Corollary 3.11:

THEOREM 3.14. *Let $\check{\mathbf{S}}^{(i)}$ be a $(\check{\mathbf{F}}^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis with its columns sorted in an increasing order of their $\vec{s}^{(i-1)}$ degrees. Let $\hat{\mathbf{S}}^{(i)} = \mathbf{E}^{(i)}\check{\mathbf{S}}^{(i)}$. Let J be the column rank profile of $\mathrm{lcoeff}(x^{\vec{s}^{(i)}}\hat{\mathbf{S}}^{(i)})$. 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 (1) directly to the input matrix $\bar{\mathbf{F}}^{(i)}$. □

Lemma 3.13 can also be extended in the same way to capture Storjohann's transformation with more general degree parameters:

LEMMA 3.15. *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 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 in the same way as Lemma 3.13. Alternatively, one can apply Storjohann's transformation with degree parameter $\delta^{(i-1)}$ to $\bar{\mathbf{F}}^{(i)}$ as in (2). The lemma then follows from Lemma 3.13 after noticing the redundant block rows that can be disregarded. □

Notice that if $i = \log(n/m) - 1$, then Theorem 3.14 and Lemma 3.15 specialize to Corollary 3.11 and Lemma 3.13.

4. COMPUTATION OF ORDER BASES

In this section, we first establish a link between two different Storjohann transformed problems, by dividing the transformed problem from the previous section to two subproblems and then simplifying the second subproblem. This immediately leads to a recursive way of computing order bases. We then describe our algorithm in the equivalent iterative way.

4.1 Dividing to Subproblems

Section 3 shows that the problem of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis can be converted to the problem of computing a $(\check{\mathbf{F}}, \vec{\omega}, \vec{s'})$-basis and more generally, computing 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 a $(\check{\mathbf{F}}^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis computation problem. We now consider dividing the new converted problem to two subproblems.

The first subproblem is computing a $(\check{\mathbf{F}}^{(i)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis or equivalently a $(\bar{\mathbf{F}}^{(i-1)}, 2\delta^{(i-1)}, \vec{s}^{(i-1)})$-basis $\bar{\mathbf{P}}^{(i-1)}$, which is another Storjohann transformed problem with degree parameter $\delta^{(i-1)}$. The second subproblem is computing a $(\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis $\bar{\mathbf{Q}}^{(i)}$ using the residual $\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}$ 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 [3] we then know that the product $\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$ is a $(\check{\mathbf{F}}^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis and $\deg_{\vec{s}^{(i-1)}} \bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)} = \deg_{\vec{t}^{(i-1)}} \bar{\mathbf{Q}}^{(i)}$.

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 4.1. *A $(\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)}, \vec{t}^{(i-1)})$-basis is a $(\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis.*

PROOF. This is because $\bar{\mathbf{F}}^{(i)}\hat{\mathbf{P}}^{(i-1)}$ is a submatrix of $\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}$ after removing rows already having 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 already 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 3.15. 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 $(\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}_2^{(i-1)}, \vec{\omega}^{(i)}, \vec{b}^{(i-1)})$-basis), and $k^{(i-1)}$ be the column dimension of $\bar{\mathbf{P}}_1^{(i-1)}$. Then

LEMMA 4.2. *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 $(\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-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 $\bar{\mathbf{Q}}^{(i)}$ is a $(\bar{\mathbf{F}}^{(i)} \cdot \hat{\mathbf{P}}^{(i-1)}, 2\delta^{(i)}, \vec{t}^{(i-1)})$-basis by Lemma 2.4. □

Lemma 4.2 immediately leads to the following.

LEMMA 4.3. *Let $\hat{\mathbf{S}} = [\hat{\mathbf{P}}_1^{(i-1)}, \hat{\mathbf{P}}_2^{(i-1)}\bar{\mathbf{Q}}_2^{(i)}]$, I 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 4.2, $\bar{\mathbf{Q}}^{(i)}$ is a $(\check{\mathbf{F}}^{(i)}\bar{\mathbf{P}}^{(i-1)}, \vec{\omega}^{(i)}, \vec{t}^{(i-1)})$-basis, hence $\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$ is a $(\check{\mathbf{F}}^{(i)}, \vec{\omega}^{(i)}, \vec{s}^{(i-1)})$-basis. Now Theorem 3.14 applies as $[\hat{\mathbf{P}}_1^{(i-1)}, \hat{\mathbf{P}}_2^{(i-1)}\bar{\mathbf{Q}}_2^{(i)}] = \mathbf{E}^{(i)}\bar{\mathbf{P}}^{(i-1)}\bar{\mathbf{Q}}^{(i)}$. □

Lemma 4.3 gives us a way of computing a $(\mathbf{F}, \sigma, \vec{s})$-basis. We can simply set i to $\log(n/m)-1$ so $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)}) = (\mathbf{F}, \sigma, \vec{s})$, and compute a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis. By Lemma 4.3, this can be divided to 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 $(\bar{\mathbf{F}}^{(i-1)}, \delta^{(i-1)}, \vec{s}^{(i-1)})$-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 $(\bar{\mathbf{F}}^{(i-1)}, \delta^{(i-1)}, \vec{s}^{(i-1)})$-basis can again be divided to two subproblems just like before. This can be repeated recursively until we reach the base case with degree parameter $\delta^{(1)} = 2d$. The total number of levels of recursion is therefore $\log(n/m) - 1$.

Notice that the transformed matrix $\check{\mathbf{F}}^{(i)}$ 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.

Algorithm 1 uses a subroutine OrderBasis, the algorithm from Giorgi et al., for computing order bases with balanced input. Specifically, $[\mathbf{Q}, \vec{a}] = \mathrm{OrderBasis}(\mathbf{G}, \sigma, \vec{b})$ 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 Section 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 of 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. Now 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 a $\bar{\mathbf{P}}^{(1)}$ 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. Now 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 $\vec{b}^{(1)} = \deg_{\vec{s}^{(1)}} \hat{\mathbf{P}}_2^{(1)}$, 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

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

PROOF. Each iteration i computes a $(\bar{\mathbf{F}}^{(i)}, 2\delta^{(i)}, \vec{s}^{(i)})$-basis. At iteration $\log(n/m) - 1$, 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, hence at most $n/2^i$ elements remain to be computed. If the shift $\vec{s}$ is balanced, the number of remaining basis elements $n^{(i)}$ at iteration i is $O(n/2^i)$.*

PROOF. The uniform case follows from the idea of Storjohann and Villard [10] on null space basis computation discussed in Section 2.3. For the balanced case, the average column degree is bounded by $cd = cm\sigma/n$ for some constant c. The first iteration λ that $\delta^{(\lambda)}$ reaches cd is therefore a constant. I.e., $\delta^{(\lambda)} = 2^\lambda d \geq cd > \delta^{(\lambda-1)}$, 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 $O(\mathrm{MM}(n, d)\log\sigma) \subset O^\sim(\mathrm{MM}(n, d))$ field operations.*

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 σ 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: $\vec{b}^{(i-1)} := \vec{t}^{(i-1)}[k^{(i-1)} + 1 \ldots n + m(l^{(i-1)} - 1)]$;
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. 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 19, the order basis computation at line 17, and extracting 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 MM$(n/2^i, 2^i d)$ field operations, which is in $O(\mathrm{MM}(n, d))$. 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 $O(\mathrm{MM}(n, d))$ field operations. The total cost of the multiplications over $O(\log(n/m))$ iterations is therefore $O(\mathrm{MM}(n, d) \log(n/m))$.

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 $2\delta^{(i)} \in O(2^i d)$. Therefore, the cost of the order basis computation at iteration i is $O(\mathrm{MM}(n/2^i, 2^i d) \log(2^i d))$. The total cost over $O(\log(n/m))$ iterations is then $O(\sum \mathrm{MM}(n/2^i, 2^i d) \log 2^i d)$ $= O(\mathrm{MM}(n, d) \log d)$.

Extracting an order basis by LSP factorization cost $O(n^\omega)$, which is dominated by other costs.

Combining the above gives $O(\mathrm{MM}(n, d)(\log n/m + \log d))$ $= O(\mathrm{MM}(n, d) \log \sigma)$ as the total cost of the algorithm. □

6. FUTURE RESEARCH

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. Order bases are also closely related to many other problems in polynomial matrix computation, for example matrix normal forms. We are interested in seeing how our tools can be used to solve these problems more efficiently. We are primarily interested in the computation of shifted normal forms [5] and in particular those related to shifted Popov forms.

7. REFERENCES

[1] G. Baker and P. Graves-Morris. *Padé Approximants, 2nd edition*. Cambridge, 1996.

[2] B. Beckermann and G. Labahn. A uniform approach for fast computation of matrix-type Padé approximants. *SIAM J. Matrix Analysis and its Applications*, 15(3):804–823, July 1994.

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

[4] B. Beckermann, G. Labahn, and G. Villard. Shifted normal forms of polynomial matrices. In *Proceedings of ISSAC'99*, 1999.

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

[6] P. Giorgi, C.P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *Proceedings of ISSAC'03*, pages 135–142. ACM Press, August 2003.

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

[8] G. Labahn. Inversion components for block Hankel-like matrices. *Linear Algebra and its Applications*, 177:7–48, 1992.

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

[10] A. Storjohann and G. Villard. Computing the rank and a small nullspace basis of a polynomial matrix. In *Proceedings of ISSAC'05*, page 309, 2005.