Fast computation of approximant bases in canonical form

Claude-Pierre Jeannerod, Vincent Neiger, Gilles Villard

Fast computation of approximant bases in canonical form

Claude-Pierre Jeannerod

Univ Lyon, Inria, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR 5668, F-69007 Lyon, France

Vincent Neiger

Univ. Limoges, CNRS, XLIM, UMR 7252, F-87000 Limoges, France

Gilles Villard

Univ Lyon, CNRS, ENS de Lyon, Inria, Université Claude Bernard Lyon 1, LIP UMR 5668, F-69007 Lyon, France

Abstract

In this article, we design fast algorithms for the computation of approximant bases in shifted Popov normal form. We first recall the algorithm known as PM-Basis, which will be our second fundamental engine after polynomial matrix multiplication: most other fast approximant basis algorithms basically aim at efficiently reducing the input instance to instances for which PM-Basis is fast. Such reductions usually involve partial linearization techniques due to Storjohann, which have the effect of balancing the degrees and dimensions in the manipulated matrices.

Following these ideas, Zhou and Labahn gave two algorithms which are faster than PM-Basis for important cases including Hermite-Padé approximation, yet only for shifts whose values are concentrated around the minimum or the maximum value. The three mentioned algorithms were designed for balanced orders and compute approximant bases that are generally not normalized. Here, we show how they can be modified to return the shifted Popov basis without impact on their cost bound; besides, we extend Zhou and Labahn's algorithms to arbitrary orders.

Furthermore, we give an algorithm which handles arbitrary shifts with one extra logarithmic factor in the cost bound compared to the above algorithms. To the best of our knowledge, this improves upon previously known algorithms for arbitrary shifts, including for particular cases such as Hermite-Padé approximation. This algorithm is based on a recent divide and conquer approach which reduces the general case to the case where information on the output degree is available. As outlined above, we solve the latter case via partial linearizations and PM-Basis.

Keywords: Hermite-Padé approximation; minimal approximant basis; order basis; polynomial matrix; shifted Popov form.

1. Introduction

Let $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{>0}$, and let $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ be a matrix of univariate polynomials over a field $\mathbb{K}$, which represents a matrix of formal power series with the jth column truncated at order d_j. We consider a matrix-type generalization of Hermite-Padé approximation, which consists in computing polynomial row vectors $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ such that

$$\mathbf{pF} = 0 \bmod \mathbf{X^d}, \quad \text{where} \;\; \mathbf{X^d} = \operatorname{diag}(X^{d_1}, \ldots, X^{d_n}). \tag{1}$$

Here, $\mathbf{pF} = 0 \bmod \mathbf{X^d}$ means that $\mathbf{pF} = \mathbf{qX^d}$ for some $\mathbf{q} \in \mathbb{K}[X]^{1\times n}$. The set of all such approximants forms a free $\mathbb{K}[X]$-module of rank m denoted by $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$; its bases are represented as the rows of nonsingular matrices in $\mathbb{K}[X]^{m\times m}$. One is usually interested in bases having minimal row degrees with respect to a *shift* $\mathbf{s} \in \mathbb{Z}^m$, used as column weights.

In this paper, we improve complexity bounds for the computation of such $\mathbf{s}$-*minimal approximant bases*. In addition, our algorithms return a canonical $\mathbf{s}$-minimal basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, called the $\mathbf{s}$-*Popov* basis (Popov, 1972; Beckermann et al., 1999) and defined in Section 2.1. The properties of this basis allow us to compute it faster than $\mathbf{s}$-minimal bases in general (for more insight, see Jeannerod et al., 2016) and also, once obtained, to efficiently perform operations with this basis (see for example Rosenkilde and Storjohann, 2016, Thm. 12).

Our problem is stated in Problem 1; $\mathrm{cdeg}(\mathbf{F})$ denotes the tuple of the n column degrees of the matrix $\mathbf{F}$. Here and hereafter, tuples of integers are always compared componentwise. The assumption that $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$ is harmless: truncating the column j of $\mathbf{F}$ modulo X^{d_j} does not affect the module of approximants.

Problem 1 – *Approximant basis in shifted Popov form*

Input:
- approximation order $\mathbf{d} \in \mathbb{Z}^n_{>0}$,
- matrix $\mathbf{F}$ in $\mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$ componentwise,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: the $\mathbf{s}$-Popov basis $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ of the $\mathbb{K}[X]$-module

$$\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[X]^{1\times m} \mid \mathbf{pF} = 0 \bmod \mathbf{X^d}\}.$$

For estimating the tightness of the cost bounds below, we consider the number of field elements used to represent the input and output of the problem. Representing polynomials in the standard monomial basis, the matrix $\mathbf{F}$ is represented by $m\sigma$ coefficients from $\mathbb{K}$, where

$$\sigma = d_1 + \cdots + d_n = |\mathbf{d}|;$$

here, $|\cdot|$ denotes the sum of a tuple of nonnegative integers. By definition of $\mathbf{s}$-Popov forms, the output basis can be written $\mathbf{P} = \mathbf{X}^{\boldsymbol{\delta}} + \mathbf{A}$ for a matrix $\mathbf{A}$ such that $\mathrm{cdeg}(\mathbf{A}) < \boldsymbol{\delta} = \mathrm{cdeg}(\mathbf{P})$. Importantly, we have $|\boldsymbol{\delta}| \leq \sigma$ (see Lemma 2.2). Thus, $\mathbf{P}$ can be represented by the degrees $\boldsymbol{\delta}$ together with $m|\boldsymbol{\delta}| \leq m\sigma$ coefficients from $\mathbb{K}$ for the columns of $\mathbf{A}$ (not counting those corresponding to identity columns in $\mathbf{P}$). The tuple $\boldsymbol{\delta}$, called the $\mathbf{s}$-*minimal degree of* $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, plays a central role in our algorithms; knowing $\boldsymbol{\delta}$ amounts to knowing the degrees of the columns of the sought basis.

Our cost model estimates the number of arithmetic operations in $\mathbb{K}$ on an algebraic RAM. We consider an exponent ω for matrix multiplication: two matrices in $\mathbb{K}^{m\times m}$ can be multiplied in $O(m^\omega)$ operations in $\mathbb{K}$. *In this paper, all cost bounds are given for $\omega > 2$*; additional logarithmic factors may appear if $\omega = 2$. (Coppersmith and Winograd, 1990; Le Gall, 2014) show that one can take $\omega < 2.373$. We also use a cost function $\mathsf{MM}(\cdot,\cdot)$ for the multiplication of polynomial matrices, defined as follows: for two real numbers $m, d > 0$, $\mathsf{MM}(m, d)$ is such that two matrices of degree at most d in $\mathbb{K}[X]^{\bar{m}\times\bar{m}}$ with $\bar{m} \leq m$ can be multiplied using $\mathsf{MM}(m, d)$ operations in $\mathbb{K}$. Furthermore, we will use $\mathsf{MM}'(m, d) = \sum_{0\leq i\leq \log(d)} 2^i \mathsf{MM}(m, 2^{-i}d)$ from (Storjohann, 2003; Giorgi et al., 2003), which is typically related to divide-and-conquer computations.

We will always give cost bounds in function of $\mathsf{MM}(m, d)$ and $\mathsf{MM}'(m, d)$; the current best known upper bounds on the former quantity can be found in (Cantor and Kaltofen, 1991; Bostan and Schost, 2005; Harvey et al., 2017). The first of these references proves

$$\mathsf{MM}(m,d) \in O(m^\omega d\log(d) + m^2 d\log(d)\log(\log(d)) + m^\omega)$$

for an arbitrary field $\mathbb{K}$, while the last two show better bounds in the case of fields that are either finite or of characteristic zero. For the sake of presentation, we will also give simplified cost bounds for our main results, relying on the following assumption:

$$\mathcal{H}_{\mathsf{MM}}: \; \mathsf{MM}(m,d) + \mathsf{MM}(m,d') \le \mathsf{MM}(m,d+d') \text{ for } m,d,d' > 0 \qquad \textit{(super-linearity)}.$$

We remark that $\mathcal{H}_{\mathsf{MM}}$ implies $\mathsf{MM}'(m,d) \in O(\mathsf{MM}(m,d)\log(d))$.

It is customary to assume $\mathsf{MM}(m,d) \in O(m^\omega \mathsf{M}(d))$ for a cost function $\mathsf{M}(\cdot)$ such that two polynomials in $\mathbb{K}[X]$ of degree at most d can be multiplied in $\mathsf{M}(d)$ operations in $\mathbb{K}$. However this does not always reflect well the actual cost of polynomial matrix multiplication, which tends to have a term in $m^2 d$ with several (sub)logarithmic factors, and a term in $m^\omega d$ with at most one logarithmic factor. In fact, even the above general bound on $\mathsf{MM}(m,d)$ is asymptotically better than $O(m^\omega \mathsf{M}(d))$ if we replace $\mathsf{M}(d)$ by the best known bound.

As a consequence, and since we will be discussing cost bound improvements on the level of logarithmic factors, we will not follow this custom. Instead, and as in (Storjohann, 2003) for example, we will prefer to write our cost bounds with general expressions involving $\mathsf{MM}(m,d)$ and $\mathsf{MM}'(m,d)$, which one can then always replace with context-dependent upper bounds.

Main result. We give an efficient solution to Problem 1 for arbitrary orders and shifts.

Theorem 1.1. *Let $\mathbf{d} \in \mathbb{Z}^n_{>0}$, let $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, and let $\mathbf{s} \in \mathbb{Z}^m$. Then, writing $\sigma = |\mathbf{d}|$ for the sum of the entries of $\mathbf{d}$ and assuming $m \in O(\sigma)$, Problem 1 can be solved in*

$$O\left(\left(\sum_{k=0}^{\lceil \log_2(\sigma/m)\rceil} 2^k \mathsf{MM}'(m, 2^{-k}\sigma/m)\right) + m^{\omega-1}\sigma\log(m)\right)$$

operations in $\mathbb{K}$. Assuming $\mathcal{H}_{\mathsf{MM}}$, this is in $O(\mathsf{MM}(m,\sigma/m)\log(\sigma/m)^2 + m^{\omega-1}\sigma\log(m))$.

Hiding logarithmic factors, this cost bound is $\tilde{O}(m^{\omega-1}\sigma)$, the same as for the multiplication of two $m\times m$ matrices of degree σ/m. As mentioned above, the output basis has average column degree at most σ/m, which is reached generically. Furthermore, there are instances of Problem 1 whose resolution does require at least as many field operations as the multiplication of two matrices in $\mathbb{K}[X]^{m\times m}$ of degree about σ/m (see Section 2.4).

In the case $\sigma \in O(m)$, less common in applications, the current fastest known algorithm for solving Problem 1 uses $O(m\sigma^{\omega-1} + \sigma^\omega \log(\max(\mathbf{d})))$ operations (Jeannerod et al., 2017, Prop. 7.1).

The overall design of our main algorithm is based on (Jeannerod et al., 2016, Algo. 1); we refer to (ibid., Sec. 1.2) for an overview of this approach. In short, we use a divide and conquer strategy which splits the order $\mathbf{d}$ into two parts whose sums are about $\sigma/2$. Two corresponding shifted Popov bases are found recursively and yield the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$, which then helps us to efficiently compute the $\mathbf{s}$-Popov approximant basis.

In fact, (ibid., Algo. 1) solves a more general problem; we refer to (Van Barel and Bultheel, 1992; Beckermann, 1992; Beckermann and Labahn, 1997) for details about and earlier solutions to *matrix rational interpolation* problems. Eq. (1) is indeed a particular case of

$$\mathbf{pF} = 0 \bmod (\mathbf{X}-\mathbf{x})^{\mathbf{d}}, \quad \text{where } (\mathbf{X}-\mathbf{x})^{\mathbf{d}} = \mathrm{diag}([(X-x_j)^{d_j}]_{1\le j\le n}), \tag{2}$$

where these diagonal entries are given by their roots $\mathbf{x}$ and multiplicities $\mathbf{d}$.

For such equations, (Beckermann and Labahn, 2000, Algo. FFFG) returns the $\mathbf{s}$-Popov basis of solutions in $O(m\sigma^2)$ operations (Neiger, 2016, Sec. 6.4). At each step of this iterative algorithm, one normalizes the computed basis to better control its degrees, and thus achieve better efficiency. Indeed, similar algorithms without normalization, such as the one in (Van Barel and Bultheel, 1992), have a cost of $O(m^2\sigma^2)$ operations in general.

The algorithm of (Jeannerod et al., 2016) also addresses Eq. (2). Here, we obtain a faster algorithm in the case $\mathbf{x} = \mathbf{0}$ by improving one of its core components: solving Problem 1 when the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$ is known a priori. Explicitly, the gain here compared to the cost bound in (ibid., Thm. 1.3) is in $\Omega(\log(\sigma))$.

This extra logarithmic factor in (ibid.) has two independent sources. First, it originates from the computation of *residuals*, which are matrix remainders of the form $\mathbf{PF} \bmod (\mathbf{X} - \mathbf{x})^{\mathbf{d}}$; here, with $\mathbf{x} = \mathbf{0}$, these are simply truncated products. Second, it also comes from the strategy for handling unbalanced output degrees, by relying on (Jeannerod et al., 2017, Algo. 2) which uses unbalanced polynomial matrix products and changes of shifts. Here we rather make use of the overlapping linearization from (Storjohann, 2006, Sec. 2), allowing us to reduce more directly to cases solved by (Giorgi et al., 2003, Algo. PM-Basis) using balanced polynomial matrix products.

Balanced orders: obtaining the canonical basis via PM-Basis*.* Let us now consider the case where the n entries of the order $\mathbf{d}$ are roughly the same. More precisely, we assume that

$$\mathcal{H}_{\mathbf{d}}: \quad \max(\mathbf{d}) \in O(\sigma/n) \qquad\qquad \textit{(balanced order)},$$

and we let $d = \max(\mathbf{d})$. We note that any algorithm designed for a uniform order $(d, \ldots, d)$ can straightforwardly be used to deal with any order $\mathbf{d}$ (see Remark 3.3); yet, this might lead to poor performance if the latter order is not balanced.

Under $\mathcal{H}_{\mathbf{d}}$, the divide and conquer algorithm of (Beckermann and Labahn, 1994), improved as in (Giorgi et al., 2003, Algo. PM-Basis), computes an $\mathbf{s}$-minimal approximant basis using $O((1 + n/m)\mathsf{MM}'(m, \sigma/n))$ operations. This is achieved for arbitrary shifts, despite the existence of $\mathbf{s}$-minimal bases with arbitrarily large degree: PM-Basis always returns a basis of degree $\leq d$. It is particularly efficient in the case $n = \Theta(m)$, the cost bound being then in $\tilde{O}(m^{\omega-1}\sigma)$.

Here, we slightly modify PM-Basis so that its output basis reveals the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$. For this, we ensure that, in addition to being $\mathbf{s}$-minimal, this basis exhibits a so-called *pivot entry* on each row; it is then said to be in $\mathbf{s}$-weak Popov form (Mulders and Storjohann, 2003). Computing bases in this form to obtain $\boldsymbol{\delta}$ will be a common thread in all the algorithms we present.

Then, we show that the canonical basis can be obtained by using essentially two successive calls to PM-Basis: the first one to find $\boldsymbol{\delta}$, and the second one to find the basis by using $-\boldsymbol{\delta}$ in place of the shift. The correctness of this approach is detailed in Lemma 2.3.

Theorem 1.2. *Let $\mathbf{d} \in \mathbb{Z}_{>0}^n$, let $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, and let $\mathbf{s} \in \mathbb{Z}^m$. Then,*

- *Problem 1 can be solved in $O((1 + n/m)\mathsf{MM}'(m, d))$ operations in $\mathbb{K}$, where $d = \max(\mathbf{d})$; assuming $\mathcal{H}_{\mathsf{MM}}$, this is in $O((1 + n/m)\mathsf{MM}(m, d)\log(d))$.*
- *If $n > m$ (hence also $\sigma > m$), Problem 1 can be solved in $O(\mathsf{MM}'(m, \sigma/m) + \mathsf{MM}'(m, d))$ operations in $\mathbb{K}$; assuming $\mathcal{H}_{\mathsf{MM}}$, this is in $O(\mathsf{MM}(m, \sigma/m)\log(\sigma/m) + \mathsf{MM}(m, d)\log(d))$.*

When $n > m$, the cost bound in the second item improves upon that in the first item for some unbalanced orders. Take for example $\mathbf{d} = (\sigma/2, 1, \ldots, 1)$ with $n = \sigma/2 + 1 \geq m$: then, $d = \sigma/2$ and the first bound is $O(\frac{\sigma}{m}\mathsf{MM}'(m, \sigma))$ whereas the second bound is only $O(\mathsf{MM}'(m, \sigma))$. This is obtained via an algorithm which reduces the column dimension to $n < m$ (first term in the cost) and then applies PM-Basis on the remaining instance (second term in the cost). The first step is itself done by applying PM-Basis a logarithmic number of times to process all columns whose corresponding order is less than σ/m; there are at least $n - m$ such columns by definition of σ.

To illustrate the involved logarithmic factors, let us consider $m = n + 1 = 2$. The cost bounds in the last theorem become $O(\mathsf{M}(\sigma)\log(\sigma))$, the same as for the related half-gcd algorithm in $\mathbb{K}[X]$ of Knuth (1970); Schönhage (1971); Moenck (1973). Besides, the bound $O(\mathsf{M}(\sigma)\log(\sigma)^3)$ from (Jeannerod et al., 2016) is replaced by $O(\mathsf{M}(\sigma)\log(\sigma)^2)$ in Theorem 1.1. We will see that this remaining extra logarithmic factor compared to the half-gcd comes from two layers of recursion: at each node of the global divide and conquer scheme, there is a call to PM-Basis, which itself is a divide and conquer algorithm performing a polynomial matrix product at each node. To avoid this factor for the general approximation problem considered here is an open question.

Weakly unbalanced shifts, around their minimal or maximum value. In this paragraph, we report cost bounds from (Zhou and Labahn, 2012) which are proved under the following assumptions:

$$\begin{aligned}\mathcal{H}_{\mathsf{M}}:\ &\mathsf{MM}(m,d) \in \Theta(m^{\omega}\mathsf{M}(d)), \mathsf{M}(kd) \in O(k^{\omega-1}\mathsf{M}(d)),\\ &\text{and } \mathsf{M}(d) + \mathsf{M}(d') \le \mathsf{M}(d + d') \text{ for } m, d, d' > 0 \text{ and } k \ge 1.\end{aligned}$$

Note that $\mathcal{H}_{\mathsf{M}}$ implies $\mathcal{H}_{\mathsf{MM}}$. Hereafter, for an integer t and a shift $\mathbf{s} = (s_1, \ldots, s_m)$, we denote by $\mathbf{s} + t$ the shift $(s_1 + t, \ldots, s_m + t)$, and notation such as the inequality $\mathbf{s} \le t$ stands for $\max(\mathbf{s}) \le t$.

The algorithm PM-Basis discussed above is efficient for $n \in \Omega(m)$ and assuming $\mathcal{H}_{\mathbf{d}}$. Yet, when n is small compared to m, this assumption $\mathcal{H}_{\mathbf{d}}$ becomes weaker and so does the bound $d = \max(\mathbf{d})$ controlling the output degree. In the extreme case $n = 1$, $\mathcal{H}_{\mathbf{d}}$ is void since $d \le \sigma = |\mathbf{d}|$ always holds; then, PM-Basis manipulates bases of degree up to $d = \sigma$, and its cost bound is $O\tilde{~}(m^{\omega}\sigma)$. Focusing on the case $n < m$, Zhou and Labahn (2012) noted that both the assumption

$$\mathcal{H}_{\mathbf{s},\mathrm{bal}}:\ \max(\mathbf{s}) - \min(\mathbf{s}) \in O(\sigma/m) \qquad \textit{(balanced shift)}$$

and the weaker assumption

$$\mathcal{H}_{\mathbf{s},\min}:\ |\mathbf{s} - \min(\mathbf{s})| \in O(\sigma) \qquad \textit{(weakly unbalanced shift, around } \min\textit{)}$$

imply that the average row degree of any $\mathbf{s}$-minimal approximant basis is in $O(\sigma/m)$. Then, using the *overlapping linearization* technique from (Storjohann, 2006, Sec. 2) at most $\log(m/n)$ times, they reduced to the case $n = \Theta(m)$ and obtained the cost bound $O(m^{\omega}\mathsf{M}(\sigma/m)\log(\sigma/n)) \subseteq O\tilde{~}(m^{\omega-1}\sigma)$ (Zhou and Labahn, 2012, Sec. 3 to 5), under $\mathcal{H}_{\mathsf{M}}$, $\mathcal{H}_{\mathbf{d}}$, and $\mathcal{H}_{\mathbf{s},\min}$. The partial linearizations are done at a degree δ which is doubled at each iteration, each of them allowing to recover the rows of degree $\le \delta$ of the sought basis. There are many such rows since the average row degree is small by assumption: after the kth iteration, only $O(m/2^k)$ rows remain to be found. An essential property for efficiency is that the found rows can be discarded in the further iterations; this yields a dimension decrease which compensates for the increase of the degree δ.

On the other hand, assuming

$$\mathcal{H}_{\mathbf{s},\max}:\ |\max(\mathbf{s}) - \mathbf{s}| \in O(\sigma) \qquad \textit{(weakly unbalanced shift, around } \max\textit{)}$$

implies roughly that the sought basis has average row degree in $O(\sigma/m)$ up to a small number of columns whose degree is large, and that the shift can be used to guess locations for these columns. Then, Zhou and Labahn (2012, Sec. 6) use $\log(m)$ calls to the *output column linearization* from

(Storjohann, 2006, Sec. 3) in degree δ. At each call, this transformation reduces to the case $\mathcal{H}_{\mathbf{s},\mathrm{bal}}$ and allows one to uncover rows of the sought basis whose degree is at a distance at most δ from the expected one. Again, there must be many such rows under $\mathcal{H}_{\mathbf{s},\max}$, and since the remaining rows have degrees which do not agree well with the shift, they must contain large blocks of zeroes; this leads to decreasing the dimensions while δ is doubled. This approach has the same asymptotic cost as above, still under $\mathcal{H}_{\mathsf{M}}$ and $\mathcal{H}_{\mathbf{d}}$; we summarize this in Fig. 1 (top).

Most often, the approximant bases returned by the algorithms in (Zhou and Labahn, 2012) are not normalized. Here, we show how to modify these algorithms to obtain the $\mathbf{s}$-Popov basis without impacting the cost bound. Furthermore, we generalize them to arbitrary orders; in other words, we remove the assumptions $n < m$ and $\mathcal{H}_{\mathbf{d}}$. Instead of making assumptions on $\mathbf{s}$ such as $\mathcal{H}_{\mathbf{s},\min}$ and $\mathcal{H}_{\mathbf{s},\max}$, we extend the algorithms to arbitrary shifts and give cost bounds parametrized by the quantities $|\mathbf{s} - \min(\mathbf{s})|$ and $|\max(\mathbf{s}) - \mathbf{s}|$ which appear in the latter assumptions and are inherent to the approach. Then, the obtained cost bounds range from $\tilde{O}(m^{\omega-1}\sigma)$ under $\mathcal{H}_{\mathbf{s},\min}$ or $\mathcal{H}_{\mathbf{s},\max}$, thus matching Theorem 1.1 up to logarithmic factors, to $\tilde{O}(m^{\omega}d)$ when the quantities above exceed some threshold, thus matching Theorem 1.2; in the latter case, the algorithms essentially boil down to a single call to PM-Basis. Precisely, we obtain the next result.

Theorem 1.3. *Let* $\mathbf{d} \in \mathbb{Z}^n_{>0}$*, let* $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ *with* $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$*, and let* $\mathbf{s} \in \mathbb{Z}^m$*. Consider the parameters* $\sigma = |\mathbf{d}|$*,* $d = \max(\mathbf{d})$*,* $\xi = \sigma + |\mathbf{s} - \min(\mathbf{s})|$*, and* $\zeta = \sigma + |\max(\mathbf{s}) - \mathbf{s}|$*. Then,*

- *If* $\xi \le md$*, Problem 1 can be solved in* $O(C(\xi, m, d))$ *operations in* $\mathbb{K}$*, where*

$$C(\xi, m, d) = \sum_{k=0}^{\lceil \log_2(d/\lceil \xi/m\rceil)\rceil} \mathsf{MM}'(2^{-k}m, 2^k\lceil \xi/m\rceil) + 2^k\mathsf{MM}(2^{-k}m, 2^k\lceil \xi/m\rceil). \qquad (3)$$

  Assuming $\mathcal{H}_{\mathsf{M}}$*, the latter quantity is in* $O(m^{\omega}\mathsf{M}(\lceil \xi/m\rceil)\log(d))$*.*
- *If* $\zeta \le md$*, Problem 1 can be solved in*

$$O\left(\mathsf{MM}'(\mu, \lceil \sigma/\mu\rceil) + \mathsf{MM}'(\mu, d) + \sum_{k=0}^{\lfloor \log_2(md/\zeta)\rfloor} C(\zeta, 2^{-k}m, d)\right)$$

  operations in $\mathbb{K}$*, for some integer* $\mu \in \mathbb{Z}_{>0}$ *such that* $\mu \le m$ *and* $\mu d < \zeta$*. Assuming* $\mathcal{H}_{\mathsf{M}}$*, this cost bound is in* $O(m^{\omega}\mathsf{M}(\lceil \zeta/m\rceil)\log(d) + \mu^{\omega}\mathsf{M}(\lceil \sigma/\mu\rceil)\log(\lceil \sigma/\mu\rceil))$*.*

As above, consider these cost bounds for $\sigma \ge m$. They can be written $\tilde{O}(m^{\omega-1}\xi)$ and $\tilde{O}(m^{\omega-1}\zeta)$ and they improve upon those in Theorem 1.2 when $\xi \in o(md)$ and when $\zeta \in o(md)$, respectively. Note that $\mathcal{H}_{\mathbf{s},\min}$ and $\mathcal{H}_{\mathbf{s},\max}$ are equivalent to $\xi \in O(\sigma)$ and $\zeta \in O(\sigma)$, respectively; under either of these two assumptions, the corresponding cost bound in the above theorem improves upon that in Theorem 1.1 at the level of logarithmic factors, assuming $\mathcal{H}_{\mathsf{M}}$.

An important example of a shift which satisfies neither $\xi \le md$ nor $\zeta \le md$ is the one which yields the approximant basis in Hermite form; namely, $\mathbf{s} = (\sigma, 2\sigma, \ldots, m\sigma)$ for which we have $\xi = \zeta = \frac{m(m-1)}{2}\sigma \ge \frac{m-1}{2}md$. Then, only the cost in Theorem 1.1 meets the target $\tilde{O}(m^{\omega-1}\sigma)$ in general: Theorem 1.3 is void with such ξ and ζ, while the cost $\tilde{O}(m^{\omega-1}\sigma + m^{\omega}d)$ in Theorem 1.2 has an extra factor md/σ which can be as large as m.

The cost bounds in Theorem 1.3 refine those in (Zhou and Labahn, 2012, Thm. 5.3 and 6.14). Jeannerod et al. (2017) gave an algorithm achieving a cost similar to that in the first item above, in the more general context of Eq. (2) and thus covering the case of arbitrary orders as well; the cost bound above improves upon that given in (ibid., Thm. 1.5) by a logarithmic factor.

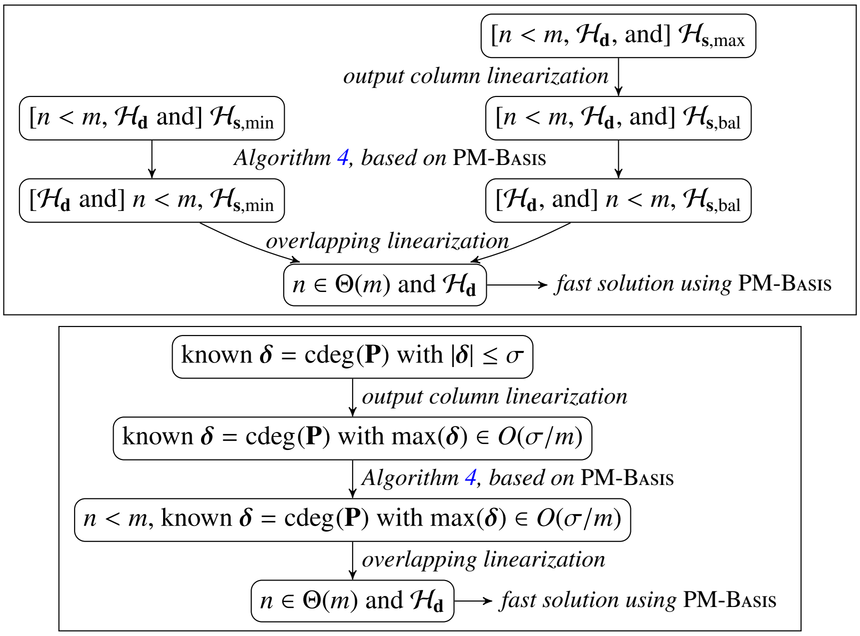

Figure 1: *(Top)* Fast algorithm from (Zhou and Labahn, 2012) assuming either $\mathcal{H}_{\mathbf{s},\min}$ or $\mathcal{H}_{\mathbf{s},\max}$, via a logarithmic number of partial linearizations from (Storjohann, 2006) and calls to PM-Basis. In brackets, assumptions that we have removed in our modified algorithm; we have also inserted the column dimension reduction (Algorithm 4) which is not necessary in (Zhou and Labahn, 2012) where $n < m$ is assumed. *(Bottom)* Fast algorithm when the shifted minimal degree is known, using two partial linearizations from (Storjohann, 2006) and calls to (Giorgi et al., 2003, Algo. PM-Basis).

Known minimal degree. The main new ingredient behind Theorem 1.1 is an efficient algorithm for Problem 1 when the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$ of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is known.

As noted above, knowing $\boldsymbol{\delta}$ leads us to consider the shift $-\boldsymbol{\delta}$ instead of $\mathbf{s}$. This new shift is weakly unbalanced around its maximum value, since $|\boldsymbol{\delta}| \le \sigma$. Inspired by the efficient algorithms of (Zhou and Labahn, 2012) for such shifts, we consider the same overall strategy while exploiting the additional information given by $\boldsymbol{\delta}$ to design a simpler and more efficient algorithm.

To handle the unbalancedness of the output column degrees, (ibid.) uses a logarithmic number of output column linearizations, each of them leading to find some rows of the sought basis. Thanks to the knowledge of $\boldsymbol{\delta}$, we are able to use the same linearization only once, with parameters which directly yield the full basis (Algorithm 5, Step **1**). This transformation builds a new instance for which the new shifted minimal degree $\boldsymbol{\delta}$ is known and balanced: $\max(\boldsymbol{\delta}) \in O(\sigma/m)$.

Then, we use PM-Basis to efficiently reduce to the case $n < m$ (Algorithm 5, Step **2**). This is not done in (ibid.) since $n < m$ holds by assumption in this reference (yet, we do resort to column dimension reduction in our generalized version of this algorithm, see Algorithm 7, Step **1**).

Now, to handle balanced shifts such as the new $-\boldsymbol{\delta}$, (ibid.) uses a logarithmic number of overlapping linearizations. Each of these transformations gives an instance satisfying $n \in \Theta(m)$ and $\mathcal{H}_{\mathbf{d}}$, which can thus be solved efficiently via PM-Basis, thereby uncovering some rows of the output basis. Here, since the output degree is $\max(\boldsymbol{\delta}) \in O(\sigma/m)$, a single call to overlapping linearization (Algorithm 5, Step **3**) yields a new instance which directly gives the full basis; as above, it satisfies $n \in \Theta(m)$ and $\mathcal{H}_{\mathbf{d}}$ and thus can be solved efficiently via PM-Basis.

We summarize our approach in Fig. 1 (bottom diagram). We note that similar ideas were already used in (Gupta and Storjohann, 2011, Sec. 3), in the context of Hermite form computation when the degrees of the diagonal entries are known.

To summarize, we obtain the cost bound $O(\mathsf{MM}'(m, \sigma/m))$ for solving Problem 1 when $\boldsymbol{\delta}$ is known (see Proposition 5.1), without any further assumption. This improves over the algorithm in (Jeannerod et al., 2016, Sec. 4), designed for the same purpose but in the more general context of Eq. (2), in which it is unclear to us how to generalize the overlapping linearization.

Outline of the paper. In Section 2, we present preliminary definitions and properties. Then, in Section 3, we describe the algorithm PM-Basis and prove the first item of Theorem 1.2. We use this algorithm in Section 4 to show how to reduce to $n < m$ efficiently; this implies the second item of Theorem 1.2. Together with partial linearizations that we recall, this allows us to solve Problem 1 when the $\mathbf{s}$-minimal degree is known (Section 5). Then, in Section 6, we give our main algorithm and the proof of Theorem 1.1. Finally, we present generalizations of the algorithms of (Zhou and Labahn, 2012) and we prove Theorem 1.3 in Section 7.

2. Preliminaries

2.1. Minimal bases, Popov bases, and minimal degree

For a shift $\mathbf{s} = (s_j)_j \in \mathbb{Z}^m$, the $\mathbf{s}$-*degree* of $\mathbf{p} = [p_j]_j \in \mathbb{K}[X]^{1\times m}$ is $\max_j(\deg(p_j) + s_j)$, with the convention $\deg(0) = -\infty$. If $\mathbf{p}$ is nonzero, its $\mathbf{s}$-*pivot* is its rightmost entry p_i such that $\deg(p_i) + s_i = \mathrm{rdeg}(\mathbf{p})$; then, i and $\deg(p_i)$ are called the $\mathbf{s}$-pivot *index* and the $\mathbf{s}$-pivot *degree* of $\mathbf{p}$, respectively. The $\mathbf{s}$-*row degree* of a matrix $\mathbf{P} \in \mathbb{K}[X]^{k\times m}$ is $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) = (r_1, \ldots, r_k)$ where r_i is the $\mathbf{s}$-degree of the ith row of $\mathbf{P}$, and the $\mathbf{s}$-*leading matrix* of $\mathbf{P} = [p_{ij}]_{ij}$ is the matrix $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}) \in \mathbb{K}^{k\times m}$ whose entry (i, j) is the coefficient of degree $r_i - s_j$ of p_{ij}. Furthermore, if $\mathbf{P}$ has no zero row, its $\mathbf{s}$-pivot index (resp. degree) is the tuple of the $\mathbf{s}$-pivot indices (resp. degrees) of its rows. The column degree of $\mathbf{P}$ is $\mathrm{cdeg}(\mathbf{P}) = \mathrm{rdeg}_{\mathbf{0}}(\mathbf{P}^\mathsf{T})$, where $\mathbf{P}^\mathsf{T}$ is the transpose of $\mathbf{P}$. We use the following definitions from (Kailath, 1980; Beckermann et al., 1999; Mulders and Storjohann, 2003).

Definition 2.1. *For* $\mathbf{s} \in \mathbb{Z}^m$*, a nonsingular matrix* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *is said to be in*

- $\mathbf{s}$-reduced form *if* $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ *is invertible;*
- $\mathbf{s}$-ordered weak Popov form *if* $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ *is invertible and lower triangular;*
- $\mathbf{s}$-weak Popov form *if it is in* $\mathbf{s}$*-ordered weak Popov form up to row permutation;*
- $\mathbf{s}$-Popov form *if* $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ *is unit lower triangular and* $\mathrm{lm}_{\mathbf{0}}(\mathbf{P}^\mathsf{T})$ *is the identity matrix.*

In particular, the $\mathbf{s}$-pivot degree of a matrix $\mathbf{P}$ in $\mathbf{s}$-ordered weak Popov form is the tuple $\boldsymbol{\delta} \in \mathbb{Z}^m_{\geq 0}$ of the degrees of its diagonal entries, and for $\mathbf{P}$ in $\mathbf{s}$-Popov form we have $\boldsymbol{\delta} = \mathrm{cdeg}(\mathbf{P})$.

For $\mathbf{d} \in \mathbb{Z}^n_{>0}$ and $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$, a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ in $\mathbf{s}$-reduced form is said to be an $\mathbf{s}$-*minimal* basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. We further call $\mathbf{s}$-*minimal degree of* $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ the $\mathbf{s}$-pivot degree of the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, and in fact of any $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ (Jeannerod et al., 2016, Lem. 3.3). The importance of these degrees is highlighted by the next two lemmas.

The first one allows us to control the degrees in the computed bases and can be found in (Van Barel and Bultheel, 1992, Thm. 4.1) in a more general context. The second one follows from (Sarkar and Storjohann, 2011, Lem. 15 and 17) and shows that when the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$ is known, the computations may be performed with the shift $-\boldsymbol{\delta}$.

Lemma 2.2. *Let* $\mathbf{d} \in \mathbb{Z}^n_{>0}$*, let* $\sigma = |\mathbf{d}|$*, and let* $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ *with* $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$*. Then, for any basis* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *of* $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$*, we have* $\deg(\det(\mathbf{P})) \leq \sigma$*. Furthermore, for* $\mathbf{s} \in \mathbb{Z}^m$*, the* $\mathbf{s}$*-minimal degree* $\boldsymbol{\delta} \in \mathbb{Z}^m_{\geq 0}$ *of* $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ *satisfies* $|\boldsymbol{\delta}| \leq \sigma$ *and* $\max(\boldsymbol{\delta}) \leq \max(\mathbf{d})$.

Proof. Let $\mathbf{P}$ be the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Then, $\mathbf{P}$ is in particular $\mathbf{0}$-column reduced, hence $\deg(\det(\mathbf{P})) = |\mathrm{cdeg}(\mathbf{P})| = |\boldsymbol{\delta}|$ (Kailath, 1980, Sec. 6.3.2); and since any basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ has determinant $\lambda \det(\mathbf{P})$ for some nonzero $\lambda \in \mathbb{K}$, it is enough to prove that $|\boldsymbol{\delta}| \le \sigma$.

Since $\mathbf{P}$ has column degree $(\delta_1, \ldots, \delta_m)$, according to (Kailath, 1980, Thm. 6.3.15) the quotient $\mathbb{K}[X]^{1\times m}/\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is isomorphic to $\mathbb{K}[X]/(X^{\delta_1}) \times \cdots \times \mathbb{K}[X]/(X^{\delta_m})$ as a $\mathbb{K}$-vector space, and thus has dimension $|\boldsymbol{\delta}|$. Now, this dimension is at most σ, since $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is the kernel of the morphism $\mathbf{p} \in \mathbb{K}[X]^{1\times m} \mapsto \mathbf{pF} \bmod \mathbf{X}^{\mathbf{d}} \in \mathbb{K}[X]/(X^{d_1}) \times \cdots \times \mathbb{K}[X]/(X^{d_n})$, whose codomain has dimension $|\mathbf{d}| = \sigma$ as a $\mathbb{K}$-vector space.

The matrix $X^{\max(\mathbf{d})}\mathbf{I}_m$ is a left-multiple of $\mathbf{P}$ since $X^{\max(\mathbf{d})}\mathbf{I}_m\mathbf{F} = 0 \bmod \mathbf{X}^{\mathbf{d}}$; thus the inequality $\max(\boldsymbol{\delta}) \le \max(\mathbf{d})$ follows from the predictable degree property (Forney, Jr., 1975). $\square$

Lemma 2.3 (Jeannerod et al. (2016, Lem. 4.1))**.** *Let $\mathbf{s} \in \mathbb{Z}^m$ and let $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ be in $\mathbf{s}$-Popov form with column degree $\boldsymbol{\delta} \in \mathbb{Z}^m_{\ge 0}$. Then $\mathbf{P}$ is also in $-\boldsymbol{\delta}$-Popov form, and we have $\mathrm{rdeg}_{-\boldsymbol{\delta}}(\mathbf{P}) = \mathbf{0}$. In particular, for any matrix $\mathbf{R} \in \mathbb{K}[X]^{m\times m}$ which is unimodularly equivalent to $\mathbf{P}$ and $-\boldsymbol{\delta}$-reduced, $\mathbf{R}$ has column degree $\boldsymbol{\delta}$, and $\mathbf{P} = \mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$.*

Let $\boldsymbol{\delta}$ be the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. This result states that, up to a *constant* transformation, the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is equal to any of its $-\boldsymbol{\delta}$-minimal bases $\mathbf{R}$. Furthermore, $\mathrm{cdeg}(\mathbf{R}) = \boldsymbol{\delta}$ implies that $\mathbf{R}$ has average column degree $|\boldsymbol{\delta}|/m \le \sigma/m$. We have no such control on the column degree of $\mathbf{s}$-minimal bases when $\mathbf{s}$ is not linked to $\boldsymbol{\delta}$, even under assumptions on the shift such as $\mathcal{H}_{\mathbf{s},\max}$, $\mathcal{H}_{\mathbf{s},\min}$, or $\mathcal{H}_{\mathbf{s},\mathrm{bal}}$.

2.2. *Recursive computation of approximant bases*

Here, we state the correctness of the approach which consists in computing a first basis from the input, then a residual instance, then a second basis from the residual, and finally combining both bases by multiplication to obtain the output basis. This scheme is followed for example by the iterative algorithms in (Van Barel and Bultheel, 1991; Beckermann and Labahn, 2000) and by the divide and conquer algorithms in (Beckermann and Labahn, 1994; Giorgi et al., 2003).

In the next lemma, the first and second items focus on minimal bases and extend (Beckermann and Labahn, 1997, Sec. 5.1); the third item gives a similar result for ordered weak Popov bases. The fourth item, from (Jeannerod et al., 2016, Sec. 3), shows how to retrieve the $\mathbf{s}$-minimal degree from two bases in normal form without computing their product.

Lemma 2.4. *Let $\mathcal{M} \subseteq \mathcal{M}_1$ be two $\mathbb{K}[X]$-submodules of $\mathbb{K}[X]^m$ of rank m, and let $\mathbf{P}_1 \in \mathbb{K}[X]^{m\times m}$ be a basis of $\mathcal{M}_1$. Let further $\mathbf{s} \in \mathbb{Z}^m$ and $\mathbf{t} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1)$. Then,*

- *(i) The rank of the module $\mathcal{M}_2 = \{\boldsymbol{\lambda} \in \mathbb{K}[X]^{1\times m} \mid \boldsymbol{\lambda}\mathbf{P}_1 \in \mathcal{M}\}$ is m, and for any basis $\mathbf{P}_2 \in \mathbb{K}[X]^{m\times m}$ of $\mathcal{M}_2$, the product $\mathbf{P}_2\mathbf{P}_1$ is a basis of $\mathcal{M}$.*
- *(ii) If $\mathbf{P}_1$ is $\mathbf{s}$-reduced and $\mathbf{P}_2$ is $\mathbf{t}$-reduced, then $\mathbf{P}_2\mathbf{P}_1$ is $\mathbf{s}$-reduced.*
- *(iii) If $\mathbf{P}_1$ is in $\mathbf{s}$-ordered weak Popov form and $\mathbf{P}_2$ is in $\mathbf{t}$-ordered weak Popov form, then $\mathbf{P}_2\mathbf{P}_1$ is in $\mathbf{s}$-ordered weak Popov form.*
- *(iv) If $\boldsymbol{\delta}_1$ is the $\mathbf{s}$-minimal degree of $\mathcal{M}_1$ and $\boldsymbol{\delta}_2$ is the $\mathbf{t}$-minimal degree of $\mathcal{M}_2$, then the $\mathbf{s}$-minimal degree of $\mathcal{M}$ is $\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2$.*

Proof. (*i*) Let $\mathbf{A} \in \mathbb{K}[X]^{m\times m}$ denote the adjugate of $\mathbf{P}_1$. Then, we have $\mathbf{AP}_1 = \det(\mathbf{P}_1)\mathbf{I}_m$. Thus, $\mathbf{pAP}_1 = \det(\mathbf{P}_1)\mathbf{p} \in \mathcal{M}$ for all $\mathbf{p} \in \mathcal{M}$, and therefore $\mathcal{M}\mathbf{A} \subseteq \mathcal{M}_2$. Now, the nonsingularity of $\mathbf{A}$ ensures that $\mathcal{M}\mathbf{A}$ has rank m; from (Dummit and Foote, 2004, Sec. 12.1, Thm. 4), this implies that $\mathcal{M}_2$ has rank m as well. The matrix $\mathbf{P}_2\mathbf{P}_1$ is nonsingular since $\det(\mathbf{P}_2\mathbf{P}_1) \ne 0$. Now let $\mathbf{p} \in \mathcal{M}$; we want to prove that $\mathbf{p}$ is a $\mathbb{K}[X]$-linear combination of the rows of $\mathbf{P}_2\mathbf{P}_1$. First, $\mathbf{p} \in \mathcal{M}_1$, so there exists $\boldsymbol{\lambda} \in \mathbb{K}[X]^{1\times m}$ such that $\mathbf{p} = \boldsymbol{\lambda}\mathbf{P}_1$. But then $\boldsymbol{\lambda} \in \mathcal{M}_2$, and thus there exists $\boldsymbol{\mu} \in \mathbb{K}[X]^{1\times m}$ such that $\boldsymbol{\lambda} = \boldsymbol{\mu}\mathbf{P}_2$. This yields the combination $\mathbf{p} = \boldsymbol{\mu}\mathbf{P}_2\mathbf{P}_1$.

(*ii*) Let $\mathbf{d} = \mathrm{rdeg}_{\mathbf{t}}(\mathbf{P}_2)$; we have $\mathbf{d} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1)$ by the predictable degree property. Using $\mathbf{X}^{-\mathbf{d}}\mathbf{P}_2\mathbf{P}_1\mathbf{X}^{\mathbf{s}} = \mathbf{X}^{-\mathbf{d}}\mathbf{P}_2\mathbf{X}^{\mathbf{t}}\mathbf{X}^{-\mathbf{t}}\mathbf{P}_1\mathbf{X}^{\mathbf{s}}$, we obtain that $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1) = \mathrm{lm}_{\mathbf{t}}(\mathbf{P}_2)\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_1)$. By assumption, $\mathrm{lm}_{\mathbf{t}}(\mathbf{P}_2)$ and $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_1)$ are invertible, hence $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1)$ is invertible as well; thus $\mathbf{P}_2\mathbf{P}_1$ is $\mathbf{s}$-reduced.

(*iii*) The matrix $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1) = \mathrm{lm}_{\mathbf{t}}(\mathbf{P}_2)\mathrm{lm}_{\mathbf{s}}(\mathbf{P}_1)$ is lower triangular and invertible.

(*iv*) Let $\mathbf{P}_1$ be the $\mathbf{s}$-Popov basis of $\mathcal{M}_1$ and $\mathbf{P}_2$ be the $\mathbf{t}$-Popov basis of $\mathcal{M}_2$. Then, by the items (*i*) and (*iii*) above, $\mathbf{P}_2\mathbf{P}_1$ is a $\mathbf{s}$-ordered weak Popov basis of $\mathcal{M}$. Thus, from (Jeannerod et al., 2016, Lem. 3.3), it is enough to show that the $\mathbf{s}$-pivot degree of $\mathbf{P}_2\mathbf{P}_1$ is $\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2$, that is, $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1) = \mathbf{s}+\boldsymbol{\delta}_1+\boldsymbol{\delta}_2$. This follows from the predictable degree property, since $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_2\mathbf{P}_1) = \mathrm{rdeg}_{\mathbf{t}}(\mathbf{P}_2) = \mathbf{t} + \boldsymbol{\delta}_2 = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1) + \boldsymbol{\delta}_2 = \mathbf{s} + \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2$. ☐

Now, consider the case where the basis $\mathbf{P}_1$ of $\mathcal{M}_1$ already has some rows in $\mathcal{M}$: we show that we may directly store these rows in the basis of $\mathcal{M}$ being computed, and that $\mathbf{P}_2$ can be obtained by focusing only on the rows of $\mathbf{P}_1$ not in $\mathcal{M}$. In the next lemma, we use standard notation for submatrices and subtuples: $\mathbf{P}_{I,*}$, $\mathbf{P}_{*,J}$, $\mathbf{P}_{I,J}$, $\mathbf{s}_I$, where I and J are subsets of $\{1,\ldots,m\}$.

Lemma 2.5. *(Using notation from Lemma 2.4.) Let I be a subset of $\{1,\ldots,m\}$ of cardinality $k \in \{0,\ldots,m\}$ and such that all rows of $\mathbf{P}_1$ with index in I are in $\mathcal{M}$. Let also $I^c = \{1,\ldots,m\} \setminus I$ be the complement of I. Then, the module $\mathcal{M}_3 = \{\boldsymbol{\mu} \in \mathbb{K}[X]^{1\times(m-k)} \mid \boldsymbol{\mu}(\mathbf{P}_1)_{I^c,*} \in \mathcal{M}\}$ has rank $m-k$, and for any basis $\mathbf{P}_3$ of $\mathcal{M}_3$, the matrix $\mathbf{P}_2 \in \mathbb{K}[X]^{m\times m}$ defined by its submatrices*

$$\begin{bmatrix} (\mathbf{P}_2)_{I,I} & (\mathbf{P}_2)_{I,I^c} \\ (\mathbf{P}_2)_{I^c,I} & (\mathbf{P}_2)_{I^c,I^c} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_3 \end{bmatrix}$$

is a basis of $\mathcal{M}_2$. Furthermore, if $\mathbf{P}_1$ and $\mathbf{P}_3$ are in $\mathbf{s}$- and $\mathbf{t}_{I^c}$-ordered weak Popov form, then $\mathbf{P}_2\mathbf{P}_1$ is an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{M}$.

Proof. Let $\boldsymbol{\lambda} \in \mathbb{K}[X]^{1\times m}$, and consider $\boldsymbol{\mu} = \boldsymbol{\lambda}_{*,I^c} \in \mathbb{K}[X]^{1\times(m-k)}$. Then, we have the equivalence $\boldsymbol{\lambda} \in \mathcal{M}_2 \Leftrightarrow \boldsymbol{\mu}(\mathbf{P}_1)_{I^c,*} \in \mathcal{M}$ since the rows of $(\mathbf{P}_1)_{I,*}$ are already in $\mathcal{M}$. Hence $\boldsymbol{\lambda} \in \mathcal{M}_2 \Leftrightarrow \boldsymbol{\mu} \in \mathcal{M}_3$, by definition of $\mathcal{M}_3$. This shows that $\mathcal{M}_3$ has rank $m-k$, and since $\mathbf{P}_3$ is a basis of $\mathcal{M}_3$, we also deduce that $\mathbf{P}_2$ is a basis of $\mathcal{M}_2$.

It is easily verified that if $\mathbf{P}_3$ is in $\mathbf{t}_{I^c}$-ordered weak Popov form, then $\mathbf{P}_2$ is in $\mathbf{t}$-ordered weak Popov form. Hence the conclusion, by the first and third items of Lemma 2.4. ☐

We remark that the left-multiplication by $\mathbf{P}_2$ amounts to simply copying the submatrix $(\mathbf{P}_1)_{I,*}$, and left-multiplying the submatrix $(\mathbf{P}_1)_{I^c,*}$ by $\mathbf{P}_3$.

2.3. Computing residuals

Approximant basis algorithms commonly make use of *residuals*, which are truncated matrix products $\mathbf{PF} \bmod \mathbf{X}^{\mathbf{d}}$. Here, we discuss their efficient computation in two cases: when we control $\deg(\mathbf{P})$, and when we control the average column degree of $\mathbf{P}$.

Lemma 2.6. *Let $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ and $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$. Then,*

- *for $d,\sigma \in \mathbb{Z}_{\geq 0}$ such that $\deg(\mathbf{P}) \leq d$ and $|\mathrm{cdeg}(\mathbf{F})| \leq \sigma$, one can compute $\mathbf{PF}$ using $O\left(\left\lceil \frac{n+\sigma/(d+1)}{m} \right\rceil \mathsf{MM}(m,d)\right)$ operations in $\mathbb{K}$ if $d > 0$ and $O\left(\left\lceil \frac{n+\sigma}{m} \right\rceil m^{\omega}\right)$ operations if $d = 0$;*
- *for $\mathbf{d} \in \mathbb{Z}_{>0}^{n}$ and $\sigma \geq m$ such that $|\mathbf{d}| \leq \sigma$ and $|\mathrm{cdeg}(\mathbf{P})| \leq \sigma$, one can compute $\mathbf{PF} \bmod \mathbf{X}^{\mathbf{d}}$ using $O(\mathsf{MM}(m,\sigma/m))$ operations in $\mathbb{K}$, assuming $n \leq m$.*

Proof. For the first item, we use column partial linearization on $\mathbf{F}$ to transform it into a matrix $\overline{\mathbf{F}}$ with m rows, $n+\sigma/(d+1)$ columns, and degree at most d. Then, we compute $\mathbf{P}\overline{\mathbf{F}}$, and the columns of this product are compressed back to obtain $\mathbf{PF}$. More details can be found for example in the discussion preceding (Jeannerod et al., 2017, Prop. 4.1).

For the second item, using column partial linearization on $\mathbf{P}$ we obtain $\overline{\mathbf{P}} \in \mathbb{K}[X]^{m\times\overline{m}}$ such that $m \le \overline{m} \le 2m$, $\deg(\overline{\mathbf{P}}) \le \lceil\sigma/m\rceil$, and $\mathbf{P} = \overline{\mathbf{P}}\mathbf{C}$ where the form of $\mathbf{C} \in \mathbb{K}[X]^{\overline{m}\times m}$ is as in Eq. (6). Then $\mathbf{PF} \bmod \mathbf{X}^{\mathbf{d}} = \overline{\mathbf{P}}\,\overline{\mathbf{F}} \bmod \mathbf{X}^{\mathbf{d}}$, where $\overline{\mathbf{F}} = \mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}}$ is obtained for free since each row of $\mathbf{C}$ is of the form $[0\cdots 0\; X^{\alpha}\; 0\cdots 0]$ for some $\alpha \in \mathbb{Z}_{\ge 0}$. Now, up to augmenting $\overline{\mathbf{P}}$ with $\overline{m}-m$ zero rows, we can apply the first item to compute $\overline{\mathbf{P}}\,\overline{\mathbf{F}}$. Here we take $d = \lceil\sigma/m\rceil$, implying $\sigma/(d+1) \le m$ and thus $(n+\sigma/(d+1))/\overline{m} \le 2$, since $\overline{m} \ge m \ge n$. Hence, computing $\overline{\mathbf{P}}\,\overline{\mathbf{F}}$ costs $O(\mathsf{MM}(\overline{m}, \lceil\sigma/m\rceil))$ operations, which is within the claimed bound since $\overline{m} \le 2m$ and $\sigma \ge m$. □

2.4. *Computing matrix products via approximant bases*

Consider a constant matrix $\mathbf{F} \in \mathbb{K}^{m\times n}$ and $\mathbf{d} = (1,\ldots,1)$; note that $\sigma = n$. Then, as detailed in Section 3, finding the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is equivalent to computing a left nullspace basis in reduced row echelon form for the matrix $\mathbf{F}$ with rows permuted according to the entries of $\mathbf{s}$. The multiplication of constant matrices can be embedded in such nullspace computations. More generally, any algorithm for Problem 1 can be used to multiply polynomial matrices, following ideas from (Sarkar and Storjohann, 2011).

Lemma 2.7. *Let $\mathcal{P}$ be an algorithm which solves Problem 1. Then, for $\mathbf{A}, \mathbf{B} \in \mathbb{K}[X]^{m\times m}$ of degree at most d, the product $\mathbf{AB}$ can be read off from the output of $\mathcal{P}(\mathbf{d}, \mathbf{F}, \mathbf{0})$, where*

$$\mathbf{d} = (6d+4,\ldots,6d+4) \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} X^{2d+1}\mathbf{I}_m & \mathbf{B} \\ -X^{2d+1}\mathbf{A} & X^{2d+1}\mathbf{I}_m \\ -\mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_m \end{bmatrix} \in \mathbb{K}[X]^{4m\times 2m}.$$

Proof. This follows from the results in (Sarkar and Storjohann, 2011, Sec. 4 and 6), which imply that the $\mathbf{0}$-Popov left kernel basis of $\mathbf{F}$ is

$$\begin{bmatrix} \mathbf{I}_m & \mathbf{0} & X^{2d+1}\mathbf{I}_m & \mathbf{B} \\ \mathbf{A} & \mathbf{I}_m & \mathbf{0} & \mathbf{AB} + X^{2d+1}\mathbf{I}_m \end{bmatrix}$$

and appears as the last $2m$ rows of the $\mathbf{0}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. □

2.5. *Stability of ordered weak Popov forms under some permutations*

When computing a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, it is sometimes useful to permute the rows of $\mathbf{F}$, that is, to consider $\mathcal{A}_{\mathbf{d}}(\boldsymbol{\pi}\mathbf{F})$ for some $m\times m$ permutation matrix $\boldsymbol{\pi}$. Then, it is easily verified that an $\mathbf{s}$-minimal basis $\mathbf{P}$ of $\mathcal{A}_{\mathbf{d}}(\boldsymbol{\pi}\mathbf{F})$ yields an $\mathbf{s}\boldsymbol{\pi}$-minimal basis $\mathbf{P}\boldsymbol{\pi}$ of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. However, the more specific weak Popov forms are not preserved in this process: if $\mathbf{P}$ is in $\mathbf{s}$-weak Popov form, then the column permuted basis $\mathbf{P}\boldsymbol{\pi}$ might for example have all its $\mathbf{s}\boldsymbol{\pi}$-pivot entries in its last column. Still, for specific permutations and when considering a submatrix of $\mathbf{P}\boldsymbol{\pi}$, we have the following result (we remark that it will only be used in Section 7.1).

Lemma 2.8. *Let $1 \le n < m$ and consider a partition $\{1,\ldots,m\} = \{i_1,\ldots,i_n\} \cup \{j_1,\ldots,j_{m-n}\}$ with $(i_k)_k$ and $(j_k)_k$ both strictly increasing. Let further $\boldsymbol{\pi} = (\pi_{i,j})$ be the $m\times m$ permutation matrix such that $\pi_{k,i_k} = 1$ for $1 \le k \le n$ and $\pi_{k+n,j_k} = 1$ for $1 \le k \le m-n$, and let $\mathbf{s} = (s_j) \in \mathbb{Z}^m$. Then,*

- *if a matrix* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *is in* $\mathbf{s}$*-ordered weak Popov form, then the leading principal* $n \times n$ *submatrix of* $\boldsymbol{\pi}\mathbf{P}\boldsymbol{\pi}^{-1}$ *is in* $(s_{i_1}, \ldots, s_{i_n})$*-ordered weak Popov form;*
- *for a tuple* $\mathbf{d} \in \mathbb{Z}_{\geq 0}^{m-n}$ *and matrices* $\mathbf{P} \in \mathbb{K}[X]^{n\times n}$ *and* $\mathbf{Q} \in \mathbb{K}[X]^{n\times(m-n)}$*, if the matrix*

$$\hat{\mathbf{P}} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{X}^{\mathbf{d}} \end{bmatrix} \in \mathbb{K}[X]^{m\times m}$$

is in $\mathbf{s}$*-ordered weak Popov, then* $\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi}$ *is in* $\mathbf{s}\boldsymbol{\pi}$*-ordered weak Popov form.*

Proof. Concerning the first item, let $\mathbf{t} = (s_{i_1}, \ldots, s_{i_n})$ and write $[p_{i,j}]$ for the entries of $\mathbf{P}$. Then, the leading principal $n \times n$ submatrix of $\boldsymbol{\pi}\mathbf{P}\boldsymbol{\pi}^{-1}$ is $[p_{i_k,i_\ell}]_{1\leq k,\ell\leq n}$. Now, $\mathrm{lm}_{\mathbf{t}}([p_{i_k,i_\ell}])$ is the submatrix of $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ formed by its rows and columns indexed by $(i_1, \ldots, i_n)$, and $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ is unit lower triangular since $\mathbf{P}$ is in $\mathbf{s}$-ordered weak Popov form. Since $i_1 < \cdots < i_n$, $\mathrm{lm}_{\mathbf{t}}([p_{i_k,i_\ell}])$ is unit lower triangular as well, and therefore $[p_{i_k,i_\ell}]_{1\leq k,\ell\leq n}$ is in $\mathbf{t}$-ordered weak Popov form.

For the second item, we prove that the $\mathbf{s}\boldsymbol{\pi}$-leading matrix of $\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi}$ is unit lower triangular. For $1 \leq k \leq m - n$, the row j_k of $\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi}$ is $[0 \ \cdots \ 0 \ X^{d_k} \ 0 \ \cdots \ 0]$ with X^{d_k} at index j_k; thus, the row j_k of $\mathrm{lm}_{\mathbf{s}\boldsymbol{\pi}}(\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi})$ is $[0 \ \cdots \ 0 \ 1 \ 0 \cdots \ 0]$ with 1 on the diagonal. It remains to show that, for $1 \leq k \leq n$, the row i_k of $\mathrm{lm}_{\mathbf{s}\boldsymbol{\pi}}(\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi})$ has the form $[* \ \cdots \ * \ 1 \ 0 \ \cdots \ 0]$ with 1 on the diagonal, that is, at index i_k. The row i_k of $\mathrm{lm}_{\mathbf{s}\boldsymbol{\pi}}(\boldsymbol{\pi}^{-1}\hat{\mathbf{P}}\boldsymbol{\pi})$ is the row k of $\mathrm{lm}_{\mathbf{s}}(\hat{\mathbf{P}})\boldsymbol{\pi}$; the latter has the desired form $[* \ \cdots \ * \ 1 \ 0 \ \cdots \ 0]$ with 1 at index i_k, since the row k of $\mathrm{lm}_{\mathbf{s}}(\hat{\mathbf{P}})$ has the form $[* \ \cdots \ * \ 1 \ 0 \ \cdots \ 0]$ with 1 at index k and since $i_1 < \cdots < i_k$. □

3. Algorithm PM-Basis: approximant bases via polynomial matrix multiplication

In this section, we focus on the case of a uniform order, that is, $\mathbf{d} = (d, \ldots, d) \in \mathbb{Z}_{>0}^{n}$ and $\sigma = nd$. For simplicity, we write $\mathcal{A}_d(\mathbf{F})$ to refer to $\mathcal{A}_{(d,\ldots,d)}(\mathbf{F})$. Then, for any shift, (Giorgi et al., 2003, Algo. PM-Basis) computes an $\mathbf{s}$-minimal basis of $\mathcal{A}_d(\mathbf{F})$ using $O((1 + n/m)\mathsf{MM}'(m, d))$ operations; this is in $\tilde{O}(m^{\omega-1}\sigma)$ when $n \in \Omega(m)$.

PM-Basis follows a divide and conquer approach, splitting the instance at order d into two instances at order $d/2$ and combining the recursively obtained bases by polynomial matrix multiplication. The base case ($d = 1$) is solved via fast dense linear algebra over the field $\mathbb{K}$. Here, we describe PM-Basis with a modified base case, ensuring that it returns the normalized basis. As a consequence, the whole algorithm returns an $\mathbf{s}$-ordered weak Popov basis; this has the advantage of directly revealing the $\mathbf{s}$-minimal degree of $\mathcal{A}_d(\mathbf{F})$, a fact used multiple times in this paper.

We now consider the base case: $d = 1$ and $\mathbf{F} \in \mathbb{K}^{m\times n}$ is constant. Then, we will see that the $\mathbf{s}$-Popov basis of $\mathcal{A}_1(\mathbf{F})$ has two sets of rows: rows corresponding to a nullspace basis for $\mathbf{F}$, and elementary rows of the form $[0 \ \cdots \ 0 \ X \ 0 \ \cdots \ 0]$. Algorithm 1 is a modified version of (Giorgi et al., 2003, Algo. M-Basis with $d = 1$), and also a specialization of (Jeannerod et al., 2017, Algo. 9) when the multiplication matrix is zero.

Proposition 3.1. *Algorithm 1 is correct and uses* $O(r^{\omega-2}mn)$ *operations in* $\mathbb{K}$*, where* r *is the rank of* $\mathbf{F}$*.*

Proof. Concerning the cost bound, the LSP decomposition at Step **2** uses $O(r^{\omega-2}mn)$ operations (Storjohann, 2000, Sec. 2.2), and reveals the row rank profile.

For the correctness, we prove the following three properties: all the rows of the output $\mathbf{P} = \pi_{\mathbf{s}}^{-1}\hat{\mathbf{P}}\pi_{\mathbf{s}}$ are in $\mathcal{A}_1(\mathbf{F})$, the rows of $\mathbf{P}$ generate $\mathcal{A}_1(\mathbf{F})$, and $\mathbf{P}$ is in $\mathbf{s}$-Popov form.

Algorithm 1 – M-Basis-1 *(Popov basis at order* $(1,\ldots,1)$*)*

Input:

- constant matrix $\mathbf{F} \in \mathbb{K}^{m\times n}$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: the $\mathbf{s}$-Popov basis of $\mathscr{A}_1(\mathbf{F})$.

1. $\pi_{\mathbf{s}} \leftarrow m\times m$ permutation matrix such that $\pi_{\mathbf{s}}\,[(s_1,1)\ \cdots\ (s_m,m)]^{\mathsf{T}}$ is lexicographically increasing
2. $(\boldsymbol{\rho},\mathbf{L}) \in \mathbb{Z}^r_{>0}\times\mathbb{K}^{m\times m} \leftarrow$ row rank profile of $\pi_{\mathbf{s}}\mathbf{F}$, and L-factor in the LSP decomposition of $\pi_{\mathbf{s}}\mathbf{F}$, where $\mathbf{L}_{*,j}$ is an identity column for $j\notin\boldsymbol{\rho}$
3. $\mathbf{M} \in \mathbb{K}^{m\times m} \leftarrow$ matrix whose ith row is $\mathbf{L}_{i,*}$ with negated off-diagonal entries if $i\notin\boldsymbol{\rho}$, and is the identity row if $i\in\boldsymbol{\rho}$
4. $\hat{\mathbf{P}} \in \mathbb{K}[X]^{m\times m} \leftarrow$ the matrix $\mathbf{X}^{\boldsymbol{\mu}}\mathbf{M}$ with $\boldsymbol{\mu} = (\mu_1,\ldots,\mu_m)$ such that $\mu_i = 1$ if $i\in\boldsymbol{\rho}$, and $\mu_i = 0$ otherwise
5. Return $\pi_{\mathbf{s}}^{-1}\hat{\mathbf{P}}\pi_{\mathbf{s}}$

First, we have that $\mathbf{PF} = 0 \bmod X$ since the rows of $\mathbf{P}$ are either multiples of X or, by definition of $\mathbf{M}$, in the left nullspace of $\mathbf{F}$. Indeed, by property of the LSP decomposition, the rows $\mathbf{L}_{i,*}$ with negated off-diagonal entries for all $i\notin\boldsymbol{\rho}$ form a basis of the left nullspace of $\pi_{\mathbf{s}}\mathbf{F}$.

Second, we show that any $\mathbf{p} \in \mathscr{A}_1(\mathbf{F})$ belongs to the row space of $\mathbf{P}$. Writing $\mathbf{p} = \mathbf{q}X + \mathbf{r}$ with $\mathbf{q} \in \mathbb{K}[X]^{1\times m}$ and $\mathbf{r} \in \mathbb{K}^{1\times m}$, we have the identity $\mathbf{q}X = \mathbf{q}\pi_{\mathbf{s}}^{-1}\mathbf{M}^{-1}\mathbf{X}^{1-\boldsymbol{\mu}}\pi_{\mathbf{s}}\mathbf{P}$. Furthermore, $\mathbf{pF} = \mathbf{rF} = \mathbf{r}\pi_{\mathbf{s}}^{-1}\pi_{\mathbf{s}}\mathbf{F} = 0 \bmod X$, and therefore $\mathbf{r}\pi_{\mathbf{s}}^{-1} = \boldsymbol{\lambda}\mathbf{M}$ for some $\boldsymbol{\lambda} = [\lambda_i]_i \in \mathbb{K}^{1\times m}$ such that $\lambda_i = 0$ if $i\in\boldsymbol{\rho}$. Recalling that $\mu_i = 0$ if $i\notin\boldsymbol{\rho}$, we obtain $\mathbf{r} = \boldsymbol{\lambda}\mathbf{X}^{\boldsymbol{\mu}}\mathbf{M}\pi_{\mathbf{s}} = \boldsymbol{\lambda}\pi_{\mathbf{s}}\mathbf{P}$.

Finally, we prove that $\mathbf{P}$ is in $\mathbf{s}$-Popov form. By construction, $\hat{\mathbf{P}}_{*,j}$ is the jth column of the identity if $j\notin\boldsymbol{\rho}$, while for $j\in\boldsymbol{\rho}$, it has constants everywhere but at position j, where $\hat{p}_{jj} = X$. It follows that $\mathrm{lm}_{\mathbf{0}}(\hat{\mathbf{P}}^{\mathsf{T}}) = \mathbf{I}_m$, and it is then easily checked that $\mathrm{lm}_{\mathbf{0}}(\mathbf{P}^{\mathsf{T}}) = \mathbf{I}_m$.

It remains to prove that $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ is unit lower triangular, or, equivalently, that

$$p_{ii} \text{ is monic and } \begin{cases} \deg(p_{ij}) + s_j \le \deg(p_{ii}) + s_i & \text{if } j \le i, \\ \deg(p_{ij}) + s_j < \deg(p_{ii}) + s_i & \text{if } j > i. \end{cases} \tag{4}$$

where p_{ij} is the entry of $\mathbf{P}$ at (i,j). Writing $[1\ \cdots\ m]\pi_{\mathbf{s}} = [\pi_1\ \cdots\ \pi_m]$, we have $p_{ij} = \hat{p}_{\pi_i\pi_j}$ for all i,j. If $\mathbf{P}_{i,*}$ is nonconstant, then so is $\hat{\mathbf{P}}_{\pi_i,*}$ and thus, by construction, its only nonzero entry is $\hat{p}_{\pi_i\pi_i} = X$. Hence $\mathbf{P}_{i,*} = [0\ \cdots\ 0\ X\ 0\ \cdots\ 0]$ with X at index i, so that Eq. (4) holds.

Let now $\mathbf{P}_{i,*}$ be a constant row. In this case, $\hat{\mathbf{P}}_{\pi_i,*}$ is constant as well and $\hat{p}_{\pi_i\pi_i} = 1$. Consequently, $p_{ii} = 1$ and Eq. (4) is now equivalent to

$$\text{if } (j \le i \text{ and } s_j > s_i) \text{ or } (j > i \text{ and } s_j \ge s_i), \text{ then } p_{ij} = 0.$$

Now, by definition of $\pi_{\mathbf{s}}$, if i and j are such that $s_j > s_i$, or such that $s_j \ge s_i$ and $j > i$, then $\pi_j > \pi_i$. Since $\hat{\mathbf{P}}$ is lower triangular, this implies $\hat{p}_{\pi_i\pi_j} = 0$, that is, $p_{ij} = 0$. □

Now, we recall PM-Basis in Algorithm 2. Note that it computes a basis of degree at most d, although there often exist $\mathbf{s}$-minimal bases with larger degree. As a result, the two bases obtained recursively can be multiplied in $\mathsf{MM}(m,d)$ operations.

Algorithm 2 – PM-BASIS *(Minimal basis for a uniform order)*

Input:

- order $d \in \mathbb{Z}_{>0}$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ of degree less than d,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output:

- an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_d(\mathbf{F})$ of degree at most d.

1. If $d = 1$ then return M-BASIS-1$(\mathbf{F}, \mathbf{s})$

2. Else:

a. $\mathbf{P}_1 \leftarrow$ PM-BASIS$(\lceil d/2 \rceil, \mathbf{F} \bmod X^{\lceil d/2 \rceil}, \mathbf{s})$

b. $\mathbf{G} \leftarrow (X^{-\lceil d/2 \rceil}\mathbf{P}_1\mathbf{F}) \bmod X^{\lfloor d/2 \rfloor}$; $\mathbf{t} \leftarrow \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1)$

c. $\mathbf{P}_2 \leftarrow$ PM-BASIS$(\lfloor d/2 \rfloor, \mathbf{G}, \mathbf{t})$

d. Return $\mathbf{P}_2\mathbf{P}_1$

Proposition 3.2. *Algorithm 2 is correct and uses $O((1 + \frac{n}{m})\mathsf{MM}'(m, d))$ operations in $\mathbb{K}$.*

Proof. From Proposition 3.1, Step **1** computes the $\mathbf{s}$-Popov basis of $\mathcal{A}_1(\mathbf{F})$, which has degree at most 1. Then, it follows by induction that the output has degree at most $d = \lceil d/2 \rceil + \lfloor d/2 \rfloor$, and items (*i*) and (*iii*) of Lemma 2.4 prove the correctness.

For the cost analysis, let us assume that d is a power of 2. From Proposition 3.1, Step **1** uses $O(m^{\omega-1}n)$ operations. The tree of the recursion has d leaves, which altogether account for $O(m^{\omega-1}nd)$ field operations. Note that $m^{\omega-1}nd \in O(\frac{n}{m}\mathsf{MM}'(m, d))$.

Then, there are recursive calls at Steps **2.a** and **2.c**, in dimension m and at order $d/2$. The residual $\mathbf{G}$ at Step **2.b** is obtained from the product $\mathbf{P}_1\mathbf{F}$, where $\mathbf{P}_1$ is an $m \times m$ matrix of degree at most $d/2$, and $\mathbf{F}$ is an $m \times n$ matrix of degree at most d. This product is done in $O(\mathsf{MM}(m, d))$ operations if $n \leq m$, and in $O(\frac{n}{m}\mathsf{MM}(m, d))$ operations if $m \leq n$. The multiplication at Step **2.d** involves two $m \times m$ matrices of degree at most $d/2$, and hence is done in $O(\mathsf{MM}(m, d/2))$ operations in $\mathbb{K}$. The cost bound follows from the definition of the cost function $\mathsf{MM}'(\cdot, \cdot)$. □

Based on Lemma 2.3, we show how to obtain the $\mathbf{s}$-Popov approximant basis using two calls to PM-BASIS (Algorithm 3). This yields an efficient solution to Problem 1 when $n \in \Omega(m)$ and the order is balanced as in $\mathcal{H}_{\mathbf{d}}$, and this proves the first item of Theorem 1.2. Note that here we allow the order to be non-uniform, based on the following remark.

Remark 3.3. Let $\mathbf{d} \in \mathbb{Z}^n_{>0}$ and $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$. Then, for any $\mathbf{d}' \in \mathbb{Z}^n_{>0}$ such that $\mathbf{d}' \geq \mathbf{d}$, we have $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \mathcal{A}_{\mathbf{d}'}(\mathbf{F}\mathbf{X}^{\mathbf{d}'-\mathbf{d}})$. In particular, algorithms for uniform orders can be used to solve the case of arbitrary orders: for $d = \max(\mathbf{d})$ and $\mathbf{G} = \mathbf{F}\mathbf{X}^{(d,\ldots,d)-\mathbf{d}}$, we have $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \mathcal{A}_d(\mathbf{G})$. For example, for a balanced order ($\mathbf{d}$ such that $\mathcal{H}_{\mathbf{d}}$: $d \in O(\sigma/n)$), PM-BASIS uses $O((1 + n/m)\mathsf{MM}'(m, \sigma/n))$ operations, where $\sigma = |\mathbf{d}|$. □

The correctness of Algorithm 3 follows from that of PM-BASIS, and from Lemma 2.3 and Remark 3.3. Besides, the cost bound $O((1 + n/m)\mathsf{MM}'(m, d))$ follows from Proposition 3.2, noting that Step **5** uses $O(m^\omega d) \subseteq O(\mathsf{MM}'(m, d))$ operations since $\deg(\mathbf{R}) \leq d$.

Algorithm 3 – Popov-PM-Basis *(Popov basis via PM-Basis)*

Input:

- order $\mathbf{d} \in \mathbb{Z}_{>0}^n$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

1. $d \leftarrow \max(\mathbf{d})$; $\mathbf{G} \leftarrow \mathbf{F}\mathbf{X}^{(d,\ldots,d)-\mathbf{d}}$
2. $\mathbf{P} \leftarrow$ PM-Basis$(d, \mathbf{G}, \mathbf{s})$
3. $\boldsymbol{\delta} \leftarrow$ the diagonal degrees of $\mathbf{P}$
4. $\mathbf{R} \leftarrow$ PM-Basis$(d, \mathbf{G}, -\boldsymbol{\delta})$
5. Return $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$

4. Reduction to the case $n < m$

Let $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}_{>0}^n$, $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ such that $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, and let $\mathbf{s} \in \mathbb{Z}^m$. In this section we assume $n \geq m$, which also implies $\sigma = d_1 + \cdots + d_n \geq m$, and we present an efficient procedure relying on PM-Basis to reduce to the case $n < m$.

Here is an overview of the reduction, assuming $d_1 \geq \cdots \geq d_n$ for simplicity. The idea is to efficiently compute a basis $\mathbf{P}$ of a truncated instance, namely of $\mathcal{A}_{\mathbf{d}'}(\mathbf{F} \bmod \mathbf{X}^{\mathbf{d}'})$ for the order

$$\mathbf{d}' = (d_m, \ldots, d_m, d_{m+1}, \ldots, d_n) \in \mathbb{Z}_{>0}^n.$$

Then, the residual instance consists of the order $\hat{\mathbf{d}} = \mathbf{d} - \mathbf{d}'$ and the matrix $\hat{\mathbf{F}} = \mathbf{P}\mathbf{F}\mathbf{X}^{-\mathbf{d}'} \bmod \mathbf{X}^{\hat{\mathbf{d}}}$: by Lemma 2.4, for any basis $\hat{\mathbf{P}}$ of $\mathcal{A}_{\hat{\mathbf{d}}}(\hat{\mathbf{F}})$, the product $\hat{\mathbf{P}}\mathbf{P}$ is a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. By construction, the residual matrix $\hat{\mathbf{F}}$ has m rows and less than m nonzero columns.

In Algorithm 4, we detail how to efficiently obtain $\mathbf{P}$ and the residual instance $(\hat{\mathbf{d}}, \hat{\mathbf{F}})$. We now sketch this algorithm, assuming that $d_m, \ldots, d_n$ are powers of 2 for ease of presentation. How to reduce to this case follows from Remark 3.3.

Then, denoting by ℓ the integer such that $d_m = 2^\ell$, we define

$$\nu_i = \mathrm{Card}(\{j \in \{1, \ldots, n\} \mid d_j = 2^i\})$$

for $0 \leq i \leq \ell$, as well as $\nu_{\ell+1} = n - \nu_0 - \cdots - \nu_\ell$. This can be illustrated as follows:

$$\mathbf{d} \;=\; (\;\overbrace{\underbrace{> 2^\ell, \ldots, > 2^\ell}_{\nu_{\ell+1}},\; \underbrace{2^\ell, \ldots, 2^\ell}_{\nu_\ell}}^{m},\; \underbrace{2^{\ell-1}, \ldots, 2^{\ell-1}}_{\nu_{\ell-1}},\; \ldots,\; \underbrace{1, \ldots, 1}_{\nu_0}\;).$$

Furthermore, we let $\mu_i = \nu_{\ell+1} + \nu_\ell + \cdots + \nu_i = \max\{j \mid d_j \geq 2^i\}$. Then, guided by this decomposition of $\mathbf{d}$, we obtain $\mathbf{P}$ in $O(\mathsf{MM}'(m, \sigma/m))$ operations via $\ell + 1$ calls to PM-Basis. This is faster than the straightforward approach consisting in a single call to PM-Basis with order $d = 2^\ell \in O(\sigma/m)$, which uses $O(\frac{n}{m}\mathsf{MM}'(m, \sigma/m))$ operations.

The first call is with $d = 1$ and computes an approximant basis $\mathbf{P}_0$ for all $\mu_0 = n$ columns of $\mathbf{F} \bmod X$. After this, we are left with the residual matrix $\mathbf{G} = X^{-1}\mathbf{P}_0\mathbf{F}$ and the order $(d_1 - 1, \ldots, d_n - 1)$, whose last ν_0 entries are zero. Thus, the second call is with $d = 2^1 - 2^0 = 1$ and for the first $\mu_1 = n - \nu_0$ columns of $\mathbf{G} \bmod X$, giving an approximant basis $\mathbf{P}_1$. Then $\mathbf{P}_1\mathbf{P}_0$ is a basis of $\mathcal{A}_{(2,\ldots,2)}(\mathbf{F})$. Considering the residual $\mathbf{G} = X^{-2}\mathbf{P}_1\mathbf{P}_0\mathbf{F}$, the third call is with $d = 2^2 - 2^1 = 2$ and for the first μ_2 columns of $\mathbf{G} \bmod X^2$, yielding an approximant basis $\mathbf{P}_2$. Thus, $\mathbf{P}_2\mathbf{P}_1\mathbf{P}_0$ is a basis of $\mathcal{A}_{(4,\ldots,4)}(\mathbf{F})$. Continuing this process until reaching the order $(2^\ell, \ldots, 2^\ell)$, we obtain $\mathbf{P} = \mathbf{P}_\ell \cdots \mathbf{P}_0$ and we are left with a residual matrix having at most $\nu_{\ell+1} = \mu_{\ell+1} < m$ nonzero columns.

Algorithm 4 – ReduceColDim *(Reduction to $n < m$ via PM-Basis)*

Input:

- order $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{>0}$ with $d_1 \geq \cdots \geq d_n$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$ and $n \geq m$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output:

- $\hat{\mathbf{d}} = (d_1 - d_m, \ldots, d_\nu - d_m) \in \mathbb{Z}^\nu_{>0}$, where $\nu = \max\{j \mid d_j > d_m\}$,
- $\hat{\mathbf{F}} = X^{-d_m}[(\mathbf{P}\mathbf{F}_{*,1}) \bmod X^{d_1} | \cdots | (\mathbf{P}\mathbf{F}_{*,\nu}) \bmod X^{d_\nu}] \in \mathbb{K}[X]^{m\times\nu}$,
- $\hat{\mathbf{s}} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) \in \mathbb{Z}^m$,
- $\mathbf{P}$ an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}-(\hat{\mathbf{d}},\mathbf{0})}(\mathbf{F})$.

1. $\tilde{d}_j \leftarrow 2^{\lceil \log_2(d_j)\rceil}$ for $m \leq j \leq n$; and $\tilde{d}_j \leftarrow d_j + \tilde{d}_m - d_m$ for $1 \leq j < m$
2. $\tilde{\mathbf{F}} \leftarrow \mathbf{F}\mathbf{X}^{\tilde{\mathbf{d}}-\mathbf{d}}$ where $\tilde{\mathbf{d}} = (\tilde{d}_1, \ldots, \tilde{d}_n)$
3. $\ell \leftarrow \log_2(\tilde{d}_m)$; $\mu_i \leftarrow \max\{j \mid \tilde{d}_j \geq 2^i\}$ for $1 \leq i \leq \ell$; and $\nu \leftarrow \max\{j \mid \tilde{d}_j > 2^\ell\}$
4. $\mathbf{P} \leftarrow$ M-Basis-1($\tilde{\mathbf{F}} \bmod X, \mathbf{s}$)
5. For i from 1 to ℓ:
  a. $\mathbf{G} \leftarrow (X^{-2^{i-1}}\mathbf{P}[\tilde{\mathbf{F}}_{*,1}|\cdots|\tilde{\mathbf{F}}_{*,\mu_i}]) \bmod X^{2^{i-1}}$
  b. $\mathbf{P}_i \leftarrow$ PM-Basis($2^{i-1}, \mathbf{G}, \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})$)
  c. $\mathbf{P} \leftarrow \mathbf{P}_i\mathbf{P}$
6. $\hat{\mathbf{d}} \leftarrow (d_1 - d_m, \ldots, d_\nu - d_m)$; and $\hat{\mathbf{s}} \leftarrow \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})$
7. $\hat{\mathbf{F}} \leftarrow X^{-d_m}[(\mathbf{P}\mathbf{F}_{*,1}) \bmod X^{d_1}|\cdots|(\mathbf{P}\mathbf{F}_{*,\nu}) \bmod X^{d_\nu}]$
8. Return $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}}, \mathbf{P})$

Proposition 4.1. *Algorithm 4 is correct and uses $O(\mathsf{MM}'(m, \sigma/m))$ operations in $\mathbb{K}$, where $\sigma = d_1 + \cdots + d_n$. Furthermore, the output is such that $\hat{\mathbf{F}}$ has m rows and $\nu < m$ columns, $|\hat{\mathbf{d}}| \leq \sigma$, $\deg(\mathbf{P}) \leq 2\sigma/m$, and for any basis $\mathbf{Q}$ of $\mathcal{A}_{\hat{\mathbf{d}}}(\hat{\mathbf{F}})$, then $\mathbf{Q}\mathbf{P}$ is a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.*

Proof. Steps **1** and **2** compute $\tilde{\mathbf{d}}$ and $\tilde{\mathbf{F}}$ such that $(\tilde{d}_i)_{i\geq m}$ are the smallest powers of two larger than or equal to $(d_i)_{i\geq m}$, and $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \mathcal{A}_{\tilde{\mathbf{d}}}(\tilde{\mathbf{F}})$ (see Remark 3.3). Step **3** defines parameters, and Step **4** computes the $\mathbf{s}$-Popov basis $\mathbf{P}$ of $\mathcal{A}_{(1,\ldots,1)}(\tilde{\mathbf{F}})$.

Then, Lemma 2.4 shows that we have the following invariant for the loop at Step **5**: at the end of the iteration i, $\mathbf{P}$ is an $\mathbf{s}$-ordered weak Popov approximant basis for $\tilde{\mathbf{F}}$ at order $(2^i, \ldots, 2^i, \tilde{d}_{\mu_i+1}, \ldots, \tilde{d}_n)$. Thus, after exiting the loop, $\mathbf{P}$ is an $\mathbf{s}$-ordered weak Popov approximant basis for $\tilde{\mathbf{F}}$ at order

$$(2^\ell, \ldots, 2^\ell, \tilde{d}_{\mu_\ell+1}, \ldots, \tilde{d}_n) = (\tilde{d}_m, \ldots, \tilde{d}_m, \tilde{d}_{m+1}, \ldots, \tilde{d}_n) = \tilde{\mathbf{d}} - (\hat{\mathbf{d}}, \mathbf{0}).$$

By choice of $\tilde{\mathbf{F}}$, we obtain that $\mathbf{P}$ is an approximant basis for $\mathbf{F}$ at order

$$\mathbf{d} - (\hat{\mathbf{d}}, \mathbf{0}) = (d_m, \ldots, d_m, d_{m+1}, \ldots, d_n).$$

In particular, it follows from Lemma 2.4 that $\mathbf{QP}$ is a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

Now, concerning the cost bound, Proposition 3.1 states that Step **4** costs $O(m^{\omega-1}n)$ operations, since $n \geq m$. This is within $O(\mathsf{MM}'(m, \sigma/m))$, since we have $m^{\omega-1}n \in O(\mathsf{MM}'(m, n/m))$, with $n \leq \sigma$. The resulting basis $\mathbf{P}$ has degree at most 1.

To obtain the residual at Step **5.a**, we compute $\mathbf{P}[\tilde{\mathbf{F}}_{*,1}|\cdots|\tilde{\mathbf{F}}_{*,\mu_i}] \bmod X^{2^i}$; this is done in $O(\frac{\mu_i}{m}\mathsf{MM}(m, 2^i))$ operations since $\mu_i \geq m$. Then, according to Proposition 3.2, Step **5.b** uses $O(\frac{\mu_i}{m}\mathsf{MM}'(m, 2^{i-1}))$ operations and $\deg(\mathbf{P}_i) \leq 2^{i-1}$. Thus, at Step **5.c** we multiply two $m \times m$ matrices of degree at most 2^{i-1}, which uses $O(\mathsf{MM}(m, 2^i))$ operations.

Altogether, the loop at Step **5** uses $O(\sum_{1\leq i\leq \ell} \frac{\mu_i}{m}\mathsf{MM}'(m, 2^{i-1})) \subseteq O(\mathsf{MM}'(m, \sigma/m))$ operations in $\mathbb{K}$, where we prove the inclusion as follows. By definition of $\mathsf{MM}'(\cdot,\cdot)$,

$$\begin{aligned}\sum_{1\leq i\leq \ell} \frac{\mu_i}{m}\mathsf{MM}'(m, 2^{i-1}) &= \sum_{1\leq i\leq \ell,\ 0\leq k<i} \frac{\mu_i}{m} 2^{i-1-k}\mathsf{MM}(m, 2^k) \\ &\leq 2 \sum_{0\leq k<\ell} 2^{-k}\frac{\sigma}{m}\mathsf{MM}(m, 2^k) \ \leq\ 4\mathsf{MM}'(m, \sigma/m).\end{aligned}$$

Both inequalities are consequences of the construction of $\tilde{\mathbf{d}}$: the first one follows from

$$2\sigma \geq |\tilde{\mathbf{d}}| = \tilde{d}_1 + \cdots + \tilde{d}_\nu + (\mu_\ell - \mu_{\ell+1})2^\ell + \cdots + (\mu_1 - \mu_2)2 + (n - \mu_1) \geq \textstyle\sum_{1\leq i\leq \ell} \mu_i 2^{i-1},$$

while the second one comes from the fact that we have $\ell - 1 \leq \log(\sigma/m)$, since

$$m2^\ell = m\tilde{d}_m \leq \tilde{d}_1 + \cdots + \tilde{d}_m \leq |\tilde{\mathbf{d}}| \leq 2\sigma.$$

Finally, the matrix $\hat{\mathbf{F}}$ at Step **7** is directly obtained from the product $\mathbf{P}[\mathbf{F}_{*,1}|\cdots|\mathbf{F}_{*,\nu}]$. This is computed in $O(\mathsf{MM}(m, \sigma/m))$ operations, according to the first item of Lemma 2.6 with $d = 2\sigma/m$, noting that $(\nu + \sigma/(d+1))/m < 2$ since $\nu < m$. □

As a result, we obtain the second item in Theorem 1.2; we only consider the case $n \geq m$, hence also $\sigma \geq m$, since otherwise the claimed bound follows from that of the first item in the same theorem. We first apply Algorithm 4 to reduce the column dimension in $O(\mathsf{MM}'(m, \sigma/m))$ operations. This gives a first basis, in $\mathbf{s}$-ordered weak Popov form, and a new instance $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}})$. Then we compute a second basis, in $\hat{\mathbf{s}}$-ordered weak Popov form for $\mathcal{A}_{\hat{\mathbf{d}}}(\hat{\mathbf{F}})$, via Algorithm 2; since $\hat{\mathbf{F}}$ has fewer columns than rows by construction, this uses $O(\mathsf{MM}'(m, \max(\mathbf{d})))$ operations.

Multiplying both bases costs $O(\mathsf{MM}(m, \sigma/m + \max(\mathbf{d})))$ and yields an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. To obtain the canonical basis, one would rather deduce the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$ from the two bases (without computing the product), and then either restart the process with the shift $-\boldsymbol{\delta}$ (similarly to Algorithm 3) or call the more general algorithm in the next section.

5. Computing approximant bases when the minimal degree is known

Let $(\mathbf{d}, \mathbf{F}, \mathbf{s})$ be the input of Problem 1, and suppose that the $\mathbf{s}$-minimal degree $\boldsymbol{\delta} \in \mathbb{Z}^m_{\geq 0}$ of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is known. In this context, Lemma 2.3 suggests that we may focus on computing a basis $\mathbf{R}$ of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ which is $-\boldsymbol{\delta}$-minimal; then, the $\mathbf{s}$-Popov basis can be easily retrieved via the constant transformation $\mathrm{lm}_{-\delta}(\mathbf{R})^{-1}\mathbf{R}$. An obstacle towards computing $\mathbf{R}$ efficiently is the possible unbalancedness of $\boldsymbol{\delta} = \mathrm{cdeg}(\mathbf{R})$, which also impacts the shift $-\boldsymbol{\delta}$. As sketched in Section 1 and in Fig. 1 (bottom), we handle this in Algorithm 5 by using the partial linearizations from (Storjohann, 2006) which allow us to compute $\mathbf{R}$ using essentially one call to ReduceColDim and then one call to PM-Basis. We defer the proof of Proposition 5.1 to Section 5.3, and we first present the partial linearizations.

Algorithm 5 – KnownDegAppBasis *(Popov basis for known minimal degree)*

Input:
- order $\mathbf{d} \in \mathbb{Z}^{n}_{>0}$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$,
- shift $\mathbf{s} \in \mathbb{Z}^{m}$,
- the $\mathbf{s}$-minimal degree $\boldsymbol{\delta} \in \mathbb{Z}^{m}_{\geq 0}$ of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

Output: the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

1. `/* Output column linearization ⇒ balanced minimal degree */`
$\delta \leftarrow \lceil |\mathbf{d}|/m \rceil$
$(-\overline{\boldsymbol{\delta}}, \mathbf{C}, (\alpha_i)_{1\leq i\leq m}, \overline{m}) \leftarrow \textsc{ColParLin}(-\boldsymbol{\delta}, \delta, \max(-\boldsymbol{\delta}))$ `// see Section 5.1`

2. `/*` ReduceColDim `⇒ fewer columns than rows */`
permute $\mathbf{d}$ into nonincreasing order, and permute the columns of $\mathbf{F}$ accordingly
$$(\hat{\mathbf{d}}, \hat{\mathbf{F}}, -\hat{\boldsymbol{\delta}}, \mathbf{R}_1) \leftarrow \begin{cases} \textsc{ReduceColDim}(\mathbf{d}, \mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}}, -\overline{\boldsymbol{\delta}}) & \text{if } n \geq \overline{m} \\ (\mathbf{d}, \mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}}, -\overline{\boldsymbol{\delta}}, \mathbf{I}_{\overline{m}}) & \text{if } n < \overline{m} \end{cases}$$
$\nu \leftarrow$ the number of columns of $\hat{\mathbf{F}}$ `//` $\hat{\mathbf{F}} \in \mathbb{K}[X]^{\overline{m}\times\nu}$ `with` $\nu < \overline{m}$

3. `/* Overlapping linearization ⇒ balanced order and dimensions */`
Construct $\mathcal{L}_{\delta}(\hat{\mathbf{d}}) \in \mathbb{Z}^{\overline{m}+\overline{\nu}}_{>0}$ and $\mathcal{L}_{\hat{\mathbf{d}},\delta}(\hat{\mathbf{F}}) \in \mathbb{K}[X]^{(\overline{m}+\overline{\nu})\times(\nu+\overline{\nu})}$ as in Definition 5.5
$\mathbf{t} \leftarrow (-\hat{\boldsymbol{\delta}}, -\delta, \ldots, -\delta) \in \mathbb{Z}^{\overline{m}+\overline{\nu}}_{\leq 0}$

4. `/* Compute approximant basis for linearized instance */`
$\hat{d} \leftarrow \max(\mathcal{L}_{\delta}(\hat{\mathbf{d}}))$; $\boldsymbol{\Delta} \leftarrow (\hat{d}, \ldots, \hat{d}) - \mathcal{L}_{\delta}(\hat{\mathbf{d}})$
$\overline{\mathbf{P}} \leftarrow \textsc{PM-Basis}(\hat{d}, \mathcal{L}_{\hat{\mathbf{d}},\delta}(\hat{\mathbf{F}})\mathbf{X}^{\boldsymbol{\Delta}}, \mathbf{t})$

5. `/* Deduce basis for original instance and normalize */`
$\mathbf{R}_2 \leftarrow$ leading principal $\overline{m} \times \overline{m}$ submatrix of $\overline{\mathbf{P}}$
$\mathbf{R} \leftarrow$ submatrix of $\mathbf{R}_2\mathbf{R}_1\mathbf{C}$ formed by its rows at indices $\alpha_1 + \cdots + \alpha_i$ for $1 \leq i \leq m$
Return $\mathrm{lm}_{-\delta}(\mathbf{R})^{-1}\mathbf{R}$

Proposition 5.1. *Algorithm 5 is correct and uses $O(\mathsf{MM}'(m, \sigma/m))$ operations in $\mathbb{K}$, where we assume that $\sigma = |\mathbf{d}| \in \Omega(m)$.*

5.1. Output column linearization to balance the output degrees

Here, we detail the transformation used in Step **1** of Algorithm 5, for which we closely follow ideas from (Storjohann, 2006, Sec. 3) and (Zhou and Labahn, 2012, Sec. 6). Yet, there are a few differences due to our goal of handling arbitrary orders $\mathbf{d}$ and computing bases in Popov form.

This transformation corresponds to modifying the input matrix $\mathbf{F}$ and the input shift $\mathbf{s}$ so that the computed basis $\overline{\mathbf{P}}$ is a column partial linearization of the sought approximant basis $\mathbf{P}$, the benefit being that $\overline{\mathbf{P}}$ has uniformly small degrees. Like all partial linearizations, this increases the matrix dimensions, m in this case. This transformation is thus mostly useful when we are able to predict which columns of $\mathbf{P}$ may have large degree: then, we only perform partial linearization for the columns that require it, and m is typically at most doubled. If the prediction was not completely accurate, this will only yield a subset of the rows of $\mathbf{P}$ (see Section 7.2).

When the shifted minimal degree is known, it directly gives the column degree of the sought basis $\mathbf{P}$. Thanks to this information, the original transformation of Storjohann (2006, Sec. 3) allows us to reduce to the case where the output has degree in $O(\sigma/m)$, and yet to retrieve the full Popov approximant basis $\mathbf{P}$. This has already been stated in (Jeannerod et al., 2016, Lem. 4.2) in a more general context; for the purpose of this section, the latter result would be sufficient.

Still, in Section 7.2 we will deal with situations where the $\mathbf{s}$-minimal degree is not available a priori, but where assumptions on the shift allow us to guess the locations of large degree columns. Hence we present, in the next lemma, the details of a more general transformation similar to that in (Zhou and Labahn, 2012, Sec. 6) but for arbitrary orders $\mathbf{d}$; in Lemma 5.4, we apply it to the specific case where the minimal degree is known. For more insight into this transformation, we refer the reader to the latter reference as well as to (Storjohann, 2006, Sec. 3).

From the next lemma we derive a procedure ColParLin which, on input $(\mathbf{s}, \delta, t)$, returns the partial linearization objects $(\overline{\mathbf{s}}, \mathbf{C}, (\alpha_i)_{1\le i\le m}, \overline{m})$. It is used in Algorithms 5 and 8. The parameter δ is a degree for partial linearization: roughly, columns of degree more than δ will be split into several columns of degree less than δ, or shift entries that are more than δ will be split into several shift entries that are less than δ. On the other hand, the parameter t has an impact on the degree threshold beyond which we can recover the approximants for the original instance from those for the partially linearized instance, as stated in Lemma 5.3.

Lemma 5.2. *Let $\mathbf{s} \in \mathbb{Z}^m$ and consider two parameters $\delta \in \mathbb{Z}_{>0}$ and $t \in \mathbb{Z}$ for partial linearization.*

Define the shift $\mathbf{t} = (t_1, \ldots, t_m) = \mathbf{s} - \max(\mathbf{s}) + t \in \mathbb{Z}^m_{\le t}$, and for each $i \in \{1, \ldots, m\}$ write $-t_i = (\alpha_i - 1)\delta + \beta_i$ with $\alpha_i = \lceil -t_i/\delta \rceil$ and $1 \le \beta_i \le \delta$ if $t_i < 0$, and with $\alpha_i = 1$ and $\beta_i = -t_i$ if $t_i \ge 0$. Let $\overline{m} = \alpha_1 + \cdots + \alpha_m$, and define the shift $\overline{\mathbf{s}} \in \mathbb{Z}^{\overline{m}}_{\le 0}$ as

$$\overline{\mathbf{s}} = (\underbrace{-\delta, \ldots, -\delta, -\beta_1}_{\alpha_1}, \ldots, \underbrace{-\delta, \ldots, -\delta, -\beta_m}_{\alpha_m}). \tag{5}$$

We have $-\delta \le \overline{\mathbf{s}} \le \max(t, -1)$ and $m \le \overline{m}$, and if $t \ge 0$ then $\overline{m} \le m + |\max(\mathbf{s}) - \mathbf{s}|/\delta$.

Define also the compression-expansion matrix $\mathbf{C} \in \mathbb{K}[X]^{\overline{m}\times m}$ as the transpose of

$$\mathbf{C}^\mathsf{T} = \begin{bmatrix} 1 & X^\delta & \cdots & X^{(\alpha_1-1)\delta} & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & X^\delta & \cdots & X^{(\alpha_m-1)\delta} \end{bmatrix}. \tag{6}$$

Then, for each $i \in \{1, \ldots, m\}$,

- *If a vector $\overline{\mathbf{p}} \in \mathbb{K}[X]^{1\times\overline{m}}$ has $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$ and $\overline{\mathbf{s}}$-pivot degree $\overline{\gamma}$, then $\overline{\mathbf{p}}\mathbf{C}$ has $\mathbf{s}$-pivot index i and $\mathbf{s}$-pivot degree $\overline{\gamma} + (\alpha_i - 1)\delta = \overline{\gamma} - t_i - \beta_i$.*
- *If a vector $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ has $\mathbf{s}$-pivot index i and $\mathbf{s}$-pivot degree $\gamma \ge -t_i$, then $\mathbf{p} = \overline{\mathbf{p}}\mathbf{C}$ for some $\overline{\mathbf{p}} \in \mathbb{K}[X]^{1\times\overline{m}}$ which has $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$ and $\overline{\mathbf{s}}$-pivot degree $\gamma + t_i + \beta_i$.*

Proof. Since $\alpha_i \ge 1$ for $1 \le i \le m$, we have $m \le \overline{m}$. Besides, the bound on $\overline{\mathbf{s}}$ follows from $\min(-t, 1) = \min(-\mathbf{t}, 1) \le \beta_i \le \delta$, which holds by definition. Now, if $t \ge 0$, for all i we have $\alpha_i \le 1 + (t - t_i)/\delta$ since $t \ge t_i$, hence the upper bound on $\overline{m}$.

Let $\overline{\mathbf{p}}$ be as in the first item, and let $\mathbf{p} = \overline{\mathbf{p}}\mathbf{C}$. We write $\overline{\mathbf{p}} = [\overline{p}_j]_{1 \le j \le \overline{m}}$, $\mathbf{p} = [p_j]_{1 \le j \le m}$, and $\overline{\mathbf{s}} = [\overline{s}_j]_{1 \le j \le \overline{m}}$. Our assumption on the $\overline{\mathbf{s}}$-pivot of $\overline{\mathbf{p}}$ implies that $\deg(\overline{p}_j) \le \overline{\gamma} - \beta_i - \overline{s}_j$ holds for $1 \le j \le \overline{m}$, with equality if $j = \alpha_1 + \cdots + \alpha_i$ (in which case $\overline{s}_j = -\beta_i$) and strict inequality if $j > \alpha_1 + \cdots + \alpha_i$. By construction, $p_j = \sum_{1 \le k \le \alpha_j} \overline{p}_{\alpha_1 + \cdots + \alpha_{j-1} + k} X^{(k-1)\delta}$ holds for $1 \le j \le m$, hence

$$\begin{aligned}\deg(p_j) &\le \max_{1 \le k \le \alpha_j} \left(\overline{\gamma} - \beta_i - \overline{s}_{\alpha_1 + \cdots + \alpha_{j-1} + k} + (k-1)\delta\right) \\ &= \overline{\gamma} - \beta_i + \beta_j + (\alpha_j - 1)\delta = \overline{\gamma} - \beta_i - t_j,\end{aligned}$$

with equality if $j = i$ and strict inequality if $j > i$. Thus, $\mathbf{p}$ has $\mathbf{t}$-pivot index i and $\mathbf{t}$-pivot degree $\overline{\gamma} - \beta_i - t_i$; its $\mathbf{s}$-pivot index and degree are the same since $\mathbf{s}$ and $\mathbf{t}$ only differ by a constant.

Let $\mathbf{p}$ be as in the second item, and write $\mathbf{p} = [p_j]_{1 \le j \le m}$. We define $\overline{\mathbf{p}} = [\overline{p}_k]_{1 \le k \le \overline{m}} \in \mathbb{K}[X]^{1 \times \overline{m}}$ as the (unique) vector such that $\mathbf{p} = \overline{\mathbf{p}}\mathbf{C}$ and $\deg(\overline{p}_k) < \delta$ if $k \notin \{\alpha_1 + \cdots + \alpha_j, 1 \le j \le m\}$. Thus, the entry $\overline{p}_{\alpha_1 + \cdots + \alpha_j}$ is the nonnegative degree part of $X^{-(\alpha_j - 1)\delta} p_j$. In particular, for $j = i$, since by assumption $\deg(p_i) = \gamma \ge \max(-t_i, 0) \ge (\alpha_i - 1)\delta$, we obtain that $\overline{p}_{\alpha_1 + \cdots + \alpha_i}$ has degree exactly $\deg(p_i) - (\alpha_i - 1)\delta = \gamma + t_i + \beta_i$, which we denote by $\overline{\gamma}$. Then, our assumption on the $\mathbf{s}$-pivot index and degree of $\mathbf{p}$, which are the same as its $\mathbf{t}$-pivot index and degree, implies that

$$\begin{aligned}\deg(\overline{p}_{\alpha_1 + \cdots + \alpha_j}) \le \deg(p_j) - (\alpha_j - 1)\delta &\le \gamma + t_i - t_j - (\alpha_j - 1)\delta \\ &= \overline{\gamma} - \beta_i + \beta_j = \overline{\gamma} + \overline{s}_{\alpha_1 + \cdots + \alpha_i} - \overline{s}_{\alpha_1 + \cdots + \alpha_j},\end{aligned}$$

where the second inequality is strict if $j > i$. Furthermore, for $k \notin \{\alpha_1 + \cdots + \alpha_j, 1 \le j \le m\}$, the requirement $\deg(\overline{p}_k) < \delta = -\overline{s}_k$ implies that $\deg(\overline{p}_k) + \overline{s}_k < 0 \le \gamma + t_i = \overline{\gamma} - \beta_i = \overline{\gamma} + \overline{s}_{\alpha_1 + \cdots + \alpha_i}$. Thus, $\overline{\mathbf{p}}$ has $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$ and $\overline{\mathbf{s}}$-pivot degree $\overline{\gamma}$. ☐

Lemma 5.3. *Let $\mathbf{d} \in \mathbb{Z}^n_{>0}$, let $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, and let $\mathbf{s} \in \mathbb{Z}^m$. Let $\delta \in \mathbb{Z}_{>0}$ and $t \in \mathbb{Z}$. Below, we use notation from the construction in Lemma 5.2 on input $(\mathbf{s}, \delta, t)$, and in particular, $(\overline{\mathbf{s}}, \mathbf{C}, (\alpha_i)_{1 \le i \le m}, \overline{m}) = \textsc{ColParLin}(\mathbf{s}, \delta, t)$. Then, we have $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})\mathbf{C}$.*

Let $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m) \in \mathbb{Z}^m_{\ge 0}$ be the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, let $\overline{\mathbf{P}} \in \mathbb{K}[X]^{\overline{m} \times \overline{m}}$ be an $\overline{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$, and let $i \in \{1, \ldots, m\}$. If $\delta_i \ge -t_i$, the approximant $\overline{\mathbf{P}}_{\alpha_1 + \cdots + \alpha_i, *}\mathbf{C} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$ has $\mathbf{s}$-pivot index i and $\mathbf{s}$-pivot degree δ_i. Furthermore, if $\overline{\mathbf{P}}_{\alpha_1 + \cdots + \alpha_i, *}$ has $\overline{\mathbf{s}}$-pivot degree larger than β_i (or, equivalently, $\mathrm{rdeg}_{\overline{\mathbf{s}}}(\overline{\mathbf{P}}_{\alpha_1 + \cdots + \alpha_i, *}) > 0$), then $\delta_i > -t_i$.

Proof. The inclusion $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) \supseteq \mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})\mathbf{C}$ is obvious: any $\overline{\mathbf{p}} \in \mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$ satisfies $\overline{\mathbf{p}}\mathbf{CF} = 0 \bmod \mathbf{X}^{\mathbf{d}}$ by definition, hence $\overline{\mathbf{p}}\mathbf{C} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Conversely, from any $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$ one can construct $\overline{\mathbf{p}}$ such that $\mathbf{p} = \overline{\mathbf{p}}\mathbf{C}$, since $\mathbf{C}$ contains $\mathbf{I}_m$ as a submatrix; then $\overline{\mathbf{p}}\mathbf{CF} = \mathbf{pF} = 0 \bmod \mathbf{X}^{\mathbf{d}}$, hence $\overline{\mathbf{p}} \in \mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$ and therefore $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})\mathbf{C}$.

Now, let $\overline{\mathbf{p}} = \overline{\mathbf{P}}_{\alpha_1 + \cdots + \alpha_i, *}$. The above paragraph shows $\overline{\mathbf{p}}\mathbf{C} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Since $\overline{\mathbf{P}}$ is in $\overline{\mathbf{s}}$-ordered weak Popov form, $\overline{\mathbf{p}}$ has $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$; let $\overline{\gamma}$ be the $\overline{\mathbf{s}}$-pivot degree of $\overline{\mathbf{p}}$.

From the first item in Lemma 5.2, we obtain that $\overline{\mathbf{p}}\mathbf{C}$ has $\mathbf{s}$-pivot index i and $\mathbf{s}$-pivot degree $\overline{\gamma} - t_i - \beta_i$; this must be at least δ_i by minimality of $\boldsymbol{\delta}$. On the other hand, the second item implies that there exists an approximant in $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$ which has $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$ and $\overline{\mathbf{s}}$-pivot degree $\delta_i + t_i + \beta_i$; this must be at least $\overline{\gamma}$ by minimality of $\overline{\mathbf{P}}$. Thus, we have $\overline{\gamma} - t_i - \beta_i = \delta_i$.

To prove our last claim, we assume that $\overline{\gamma} > \beta_i$, and we show that $\delta_i \le -t_i$ leads to a contradiction. Indeed, in this case there exists $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$ with $\mathbf{s}$-pivot index i and $\mathbf{s}$-pivot degree $\gamma = -t_i$. Then, the second item in Lemma 5.2 proves the existence of an approximant in $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$ with $\overline{\mathbf{s}}$-pivot index $\alpha_1 + \cdots + \alpha_i$ and $\overline{\mathbf{s}}$-pivot degree $\gamma + t_i + \beta_i = \beta_i < \overline{\gamma}$, which is impossible by minimality of $\overline{\gamma}$. ☐

We now specialize this result to the case where the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is known.

Lemma 5.4. *Let $\mathbf{d} \in \mathbb{Z}^{n}_{>0}$, let $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, let $\mathbf{s} \in \mathbb{Z}^{m}$, and let $\boldsymbol{\delta} \in \mathbb{Z}^{m}_{\geq 0}$ be the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Choose parameters $\delta \geq \lceil |\boldsymbol{\delta}|/m \rceil$ and $t = \max(-\boldsymbol{\delta})$. We use notation from Lemma 5.2 on input $(-\boldsymbol{\delta}, \delta, t)$; in particular, $(\bar{\mathbf{s}}, \mathbf{C}, (\alpha_i)_{1\leq i\leq m}, \overline{m}) = \textsc{ColParLin}(-\boldsymbol{\delta}, \delta, t)$.*

Then, we have $m \leq \overline{m} \leq 2m$, $-\delta \leq \bar{\mathbf{s}} \leq 0$, and $\bar{\mathbf{s}} = -\bar{\boldsymbol{\delta}}$ where $\bar{\boldsymbol{\delta}}$ is the $\bar{\mathbf{s}}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$. Let $\overline{\mathbf{P}} \in \mathbb{K}[X]^{\overline{m}\times\overline{m}}$ be an $\bar{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$ and $\mathbf{R} \in \mathbb{K}[X]^{m\times m}$ be the submatrix of $\overline{\mathbf{P}}\mathbf{C}$ formed by its rows at indices $\{\alpha_1+\cdots+\alpha_i, 1 \leq i \leq m\}$. Then, $\mathbf{R}$ is a $-\boldsymbol{\delta}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ and therefore, as a consequence of Lemma 2.3, $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$ is the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

Proof. The lower bound on $\overline{m}$ follows directly from Lemma 5.2, and so do the bounds on $\bar{\mathbf{s}}$ since $\max(t, -1) \leq 0$. By choice of t, we have $\mathbf{t} = -\boldsymbol{\delta}$, whose entries are nonpositive. Thus, for each $i \in \{1, \ldots, m\}$, we have $\alpha_i = 1$ if $\delta_i = t_i = 0$ and $\alpha_i = \lceil \delta_i/\delta \rceil$ otherwise; in both cases, $\alpha_i \leq 1+\delta_i/\delta$. Summing these inequalities, we obtain $\overline{m} \leq m + |\boldsymbol{\delta}|/\delta \leq 2m$ by choice of δ. Furthermore, since $-\mathbf{t} \leq \boldsymbol{\delta}$ entry-wise, Lemma 5.3 shows that $\mathbf{R}$ is a $-\boldsymbol{\delta}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

Our claim on $\bar{\boldsymbol{\delta}}$ can be showed using the minimality of $\boldsymbol{\delta}$ and the arguments used for proving the two items of Lemma 5.2; details can be found in the proof of (Jeannerod et al., 2016, Lem. 4.2) which contains an explicit description of the $\bar{\mathbf{s}}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}})$. □

5.2. *Overlapping linearization to balance orders and dimensions*

Now, we study Step **3** of Algorithm 5: assuming that the shifted minimal degree is known, balanced (Step **1**), and that $n < m$ (Step **2**), we reduce to an instance which is solved efficiently by PM-Basis. Namely, we use the *overlapping linearization* of Storjohann (2006, Sec. 2) to further transform the instance of Problem 1 into one with a balanced order and $n \in \Theta(m)$. In the latter reference, as well as in (Zhou and Labahn, 2012, Sec. 3), this linearization has been considered in the case of a uniform order $\mathbf{d} = (d, \ldots, d)$. Here, we extend the construction to arbitrary orders, and we show how it can be used in our specific situation where the $\mathbf{s}$-minimal degree is known.

We first give an overview of the construction and of its properties. Let $\mathbf{d} \in \mathbb{Z}^{n}_{>0}$ and $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, and choose a positive integer δ. Then, we build an order $\mathcal{L}_{\delta}(\mathbf{d}) \in \mathbb{Z}^{n+\bar{n}}_{>0}$ and a matrix $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) \in \mathbb{K}[X]^{(m+\bar{n})\times(n+\bar{n})}$ such that

- the largest entry of the order $\mathcal{L}_{\delta}(\mathbf{d})$ is at most 2δ,
- the increase in dimension is $\bar{n} < \sigma/\delta$, where $\sigma = |\mathbf{d}|$,
- approximants $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$ of degree at most δ correspond to approximants $[\mathbf{p} \quad \mathbf{q}] \in \mathcal{A}_{\mathcal{L}_{\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$ for some $\mathbf{q}$ of degree less than $\mathrm{rdeg}(\mathbf{p})$.

The last item, detailed in Lemma 5.6, gives a link between the original approximation instance and the one obtained after linearization. This implies that a minimal basis for the original instance can be retrieved from a minimal basis for the transformed instance, assuming we choose δ as an upper bound on the degree of the former basis; this approach is detailed in Lemma 5.7.

The first two items are direct consequences of the construction, given in Definition 5.5. They specify the dimensions of the transformed instance. In the context of Algorithm 5, the output column linearization of Section 5.1 has already been applied, which ensures that we are seeking a basis of degree about σ/m, and hence that δ can be chosen to be about σ/m. Then, the new order is balanced and the dimension increase is only about m: the transformed instance can be solved efficiently using a single call of PM-Basis. More details about Step **3** of Algorithm 5 can be found in Section 5.3.

Note that, in general, the $\mathbf{s}$-Popov approximant basis may have degree up to σ, in which case one would choose $\delta \geq \sigma$ in the above approach: this would not lead to any improvement since the entries of the order have not been decreased by the linearization. Still, in some contexts it is known that the sought basis has rows of small degree: using a small parameter δ will not yield the whole basis but does give the small degree part of the basis (see Lemma 5.6). This was one of the key properties mentioned in the original design of this linearization in (Storjohann, 2006), and used in (Zhou and Labahn, 2012) to handle shifts that are weakly unbalanced around their minimum value (see also Section 7.1).

Let us now present the construction of $\mathcal{L}_\delta(\mathbf{d})$ and $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$.

Definition 5.5. *Let* $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}_{>0}^n$, *let* $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ *with* $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, *and let* $\delta \in \mathbb{Z}_{>0}$. *Then, for* $1 \leq i \leq n$, *let* $d_i = \alpha_i\delta + \beta_i$ *with* $\alpha_i = \left\lceil \frac{d_i}{\delta} - 1 \right\rceil$ *and* $1 \leq \beta_i \leq \delta$.

Let also $\overline{n} = \max(\alpha_1 - 1, 0) + \cdots + \max(\alpha_n - 1, 0)$, *and define*

$$\mathcal{L}_\delta(\mathbf{d}) = (\overline{d}_1, \ldots, \overline{d}_n) \in \mathbb{Z}_{>0}^{n+\overline{n}},$$

where $\overline{d}_i = (2\delta, \ldots, 2\delta, \delta + \beta_i) \in \mathbb{Z}_{>0}^{\alpha_i}$ *if* $\alpha_i > 1$ *and* $\overline{d}_i = d_i$ *otherwise. Considering the ith column of* $\mathbf{F}$, *we write its* X^δ*-adic representation as*

$$\mathbf{F}_{*,i} = \mathbf{F}_{*,i}^{(0)} + \mathbf{F}_{*,i}^{(1)}X^\delta + \cdots + \mathbf{F}_{*,i}^{(\alpha_i)}X^{\alpha_i\delta}$$
$$\text{where } \mathrm{cdeg}([\mathbf{F}_{*,i}^{(0)} \;\; \mathbf{F}_{*,i}^{(1)} \;\; \cdots \;\; \mathbf{F}_{*,i}^{(\alpha_i)}]) < (\delta, \ldots, \delta, \beta_i).$$

If $\alpha_i > 1$, *we define*

$$\overline{\mathbf{F}}_{*,i} = \begin{bmatrix}\mathbf{F}_{*,i}^{(0)} + \mathbf{F}_{*,i}^{(1)}X^\delta & \mathbf{F}_{*,i}^{(1)} + \mathbf{F}_{*,i}^{(2)}X^\delta & \cdots & \mathbf{F}_{*,i}^{(\alpha_i-1)} + \mathbf{F}_{*,i}^{(\alpha_i)}X^\delta\end{bmatrix} \in \mathbb{K}[X]^{m\times\alpha_i}$$

and $\mathbf{E}_i = [\mathbf{0} \;\; \mathbf{I}_{\alpha_i-1}] \in \mathbb{K}[X]^{(\alpha_i-1)\times\alpha_i}$, *and otherwise we let* $\overline{\mathbf{F}}_{*,i} = \mathbf{F}_{*,i}$ *and* $\mathbf{E}_i \in \mathbb{K}[X]^{0\times 1}$. *Then,*

$$\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) = \begin{bmatrix}\overline{\mathbf{F}}_{*,1} & \overline{\mathbf{F}}_{*,2} & \cdots & \overline{\mathbf{F}}_{*,n} \\ \mathbf{E}_1 & & & \\ & \mathbf{E}_2 & & \\ & & \ddots & \\ & & & \mathbf{E}_n\end{bmatrix} \in \mathbb{K}[X]^{(m+\overline{n})\times(n+\overline{n})}$$

is called the overlapping linearization *of* $\mathbf{F}$ *with respect to* $\mathbf{d}$ *and* δ.

The next lemma gives a correspondence between the approximants of degree bounded by δ in $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ and in $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$. It uses notation from Definition 5.5.

Lemma 5.6. *Let* $\mathbf{d} \in \mathbb{Z}_{>0}^n$, *let* $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ *with* $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, *and let* $\delta \in \mathbb{Z}_{>0}$. *Then,*

- *If* $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ *is in* $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$, *then there exists a unique* $\mathbf{q} \in \mathbb{K}[X]^{1\times\overline{n}}$ *such that* $[\mathbf{p} \;\; \mathbf{q}] \in \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$, $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p})$, *and* $\mathrm{cdeg}(\mathbf{q}) < \mathcal{L}_\delta(\mathbf{d})\mathbf{E}^\mathsf{T}$ *where* $\mathbf{E} = \mathrm{diag}(\mathbf{E}_1, \ldots, \mathbf{E}_n)$. *Explicitly, it is defined as* $\mathbf{q} = -\mathbf{p}[\overline{\mathbf{F}}_{*,1} \;\; \cdots \;\; \overline{\mathbf{F}}_{*,n}]\mathbf{E}^\mathsf{T} \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{E}^\mathsf{T}}$.
- *If* $[\mathbf{p} \;\; \mathbf{q}] \in \mathbb{K}[X]^{1\times(m+\overline{n})}$ *is in* $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$ *and such that* $\mathrm{rdeg}(\mathbf{q}) < \delta$ *and* $\mathrm{rdeg}(\mathbf{p}) \leq \delta$, *then* $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$; *in particular,* $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p})$.

Proof. Concerning the first item, we first consider $i \in \{1, \ldots, n\}$ such that $\alpha_i \in \{0, 1\}$. Then, we have $\overline{\mathbf{F}}_{*,i} = \mathbf{F}_{*,i}$, $\overline{d}_i = d_i$, and $\mathbf{E}_i \in \mathbb{K}[X]^{0\times 1}$. Defining $\mathbf{q}_i$ as an empty matrix in $\mathbb{K}[X]^{1\times 0}$, the identity $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{d_i}$ can be rewritten as $\mathbf{p}\overline{\mathbf{F}}_{*,i} + \mathbf{q}_i\mathbf{E}_i = 0 \bmod \mathbf{X}^{\overline{d}_i}$.

Now, for i such that $\alpha_i > 1$, we define $\mathbf{q}_i = [q_{1,i} \ \cdots \ q_{\alpha_i-1,i}] \in \mathbb{K}[X]^{1\times(\alpha_i-1)}$ as

$$\begin{cases} q_{j,i} = X^{-j\delta}\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \cdots + \mathbf{F}^{(j-1)}_{*,i}X^{(j-1)\delta}) \bmod X^{2\delta}, \ \text{ for } 1 \le j < \alpha_i - 1, \\ q_{\alpha_i-1,i} = X^{-(\alpha_i-1)\delta}\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \cdots + \mathbf{F}^{(\alpha_i-2)}_{*,i}X^{(\alpha_i-2)\delta}) \bmod X^{\delta+\beta_i}. \end{cases} \tag{7}$$

These are polynomials since $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{d_i}$, and $\operatorname{rdeg}(\mathbf{q}_i) < \operatorname{rdeg}(\mathbf{p})$ holds since by construction $\operatorname{cdeg}(\mathbf{F}^{(k)}_{*,i}) < \delta$ for all k. For $j < \alpha_i - 1$, $\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \cdots + \mathbf{F}^{(j+1)}_{*,i}X^{(j+1)\delta}) = 0 \bmod X^{(j+2)\delta}$ becomes $q_{j,i}X^{j\delta} + \mathbf{p}(\mathbf{F}^{(j)}_{*,i}X^{j\delta} + \mathbf{F}^{(j+1)}_{*,i}X^{(j+1)\delta}) = 0 \bmod X^{(j+2)\delta}$, hence $\mathbf{p}(\mathbf{F}^{(j)}_{*,i} + \mathbf{F}^{(j+1)}_{*,i}X^{\delta}) + q_{j,i} = 0 \bmod X^{2\delta}$. Similarly, we obtain $\mathbf{p}(\mathbf{F}^{(\alpha_i-1)}_{*,i} + \mathbf{F}^{(\alpha_i)}_{*,i}X^{\delta}) + q_{\alpha_i-1,i} = 0 \bmod X^{\delta+\beta_i}$. In short, we have

$$\begin{bmatrix} \mathbf{p} & \mathbf{q}_i \end{bmatrix} \begin{bmatrix} \overline{\mathbf{F}}_{*,i} \\ \mathbf{E}_i \end{bmatrix} = 0 \bmod \mathbf{X}^{\overline{d_i}}, \quad \text{where } \overline{d_i} = (2\delta, \ldots, 2\delta, \delta + \beta_i). \tag{8}$$

Thus, by construction of $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$ and $\mathcal{L}_\delta(\mathbf{d})$, we have $[\mathbf{p} \ \mathbf{q}_1 \ \cdots \ \mathbf{q}_n] \in \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$. Besides, we have proved the degree bound for $[\mathbf{q}_1 \ \cdots \ \mathbf{q}_n]$; the explicit formula follows from Eq. (8), since the latter gives $\mathbf{q}_i = \mathbf{q}_i\mathbf{E}_i\mathbf{E}_i^\mathsf{T} = -\mathbf{p}\overline{\mathbf{F}}_{*,i}\mathbf{E}_i^\mathsf{T} \bmod \mathbf{X}^{\overline{d}_i\mathbf{E}_i^\mathsf{T}}$.

Now, we prove the second item. We write $\mathbf{q} = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_n]$ with $\mathbf{q}_i \in \mathbb{K}[X]^{1\times 0}$ if $\alpha_i \in \{0,1\}$ and $\mathbf{q}_i = [q_{1,i}, \ldots, q_{\alpha_i-1,i}] \in \mathbb{K}[X]^{1\times(\alpha_i-1)}$ if $\alpha_i > 1$. Let $i \in \{1,\ldots,n\}$. If $\alpha_i \in \{0,1\}$, then we have $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{d_i}$. If $\alpha_i > 1$, then the identity in Eq. (8) holds and yields

$$\begin{aligned} &\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \mathbf{F}^{(1)}_{*,i}X^{\delta}) = 0 \bmod X^{2\delta}, \\ &\mathbf{p}(\mathbf{F}^{(j)}_{*,i} + \mathbf{F}^{(j+1)}_{*,i}X^{\delta}) = -q_{j,i} \bmod X^{2\delta} \quad \text{for } 1 \le j \le \alpha_i - 2, \\ &\mathbf{p}(\mathbf{F}^{(\alpha_i-1)}_{*,i} + \mathbf{F}^{(\alpha_i)}_{*,i}X^{\delta}) = -q_{\alpha_i-1,i} \bmod X^{\delta+\beta_i}, \end{aligned}$$

where $\mathbf{q}_i = [q_{1,i}, \ldots, q_{\alpha_i-1,i}]$. The first identity and the second one for $j = 1$ imply that

$$\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \mathbf{F}^{(1)}_{*,i}X^{\delta} + \mathbf{F}^{(2)}_{*,i}X^{2\delta}) = \mathbf{p}\mathbf{F}^{(0)}_{*,i} - q_{1,i}X^{\delta} = 0 \bmod X^{2\delta};$$

using the bounds $\operatorname{rdeg}(\mathbf{q}) < \delta$ and $\operatorname{rdeg}(\mathbf{p}) \le \delta$ we obtain $q_{1,i} = X^{-\delta}\mathbf{p}\mathbf{F}^{(0)}_{*,i}$ and $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{3\delta}$. Then the same arguments with the above identity for $j = 2$, we obtain $q_{2,i} = X^{-2\delta}\mathbf{p}(\mathbf{F}^{(0)}_{*,i} + \mathbf{F}^{(1)}_{*,i}X^{\delta})$ and $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{4\delta}$. Continuing this process, we eventually obtain $\mathbf{p}\mathbf{F}_{*,i} = 0 \bmod X^{d_i}$. ☐

We now show that the $\mathbf{s}$-Popov basis $\mathbf{P}$ of $\mathcal{A}_\mathbf{d}(\mathbf{F})$ can be deduced from one for the transformed problem, as long as δ is chosen to be at least $\deg(\mathbf{P})$.

Lemma 5.7. *Let $\mathbf{d} \in \mathbb{Z}^n_{>0}$, let $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\operatorname{cdeg}(\mathbf{F}) < \mathbf{d}$, let $\mathbf{s} \in \mathbb{Z}^m$, let $\boldsymbol{\delta} \in \mathbb{Z}^m_{\ge 0}$ be the $\mathbf{s}$-minimal degree of $\mathcal{A}_\mathbf{d}(\mathbf{F})$, and let $\delta \in \mathbb{Z}_{>0}$ be such that $\delta \ge \max(\boldsymbol{\delta})$. Let $\overline{\mathbf{P}}$ be a $(-\boldsymbol{\delta}, -\delta, \ldots, -\delta)$-ordered weak Popov basis of $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$. Then, the leading principal submatrix $\mathbf{R} \in \mathbb{K}[X]^{m\times m}$ of $\overline{\mathbf{P}}$ is a $-\boldsymbol{\delta}$-ordered weak Popov basis of $\mathcal{A}_\mathbf{d}(\mathbf{F})$ and therefore, as a consequence of Lemma 2.3, $\operatorname{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$ is the $\mathbf{s}$-Popov basis of $\mathcal{A}_\mathbf{d}(\mathbf{F})$.*

Proof. In this proof, we use the notation $\mathbf{t} = (-\boldsymbol{\delta}, -\delta, \ldots, -\delta) \in \mathbb{Z}^{m+\overline{n}}$.

Let $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ be a $-\boldsymbol{\delta}$-ordered weak Popov basis of $\mathcal{A}_\mathbf{d}(\mathbf{F})$. Then, we have $\operatorname{rdeg}_{-\boldsymbol{\delta}}(\mathbf{P}) = \mathbf{0}$ according to Lemma 2.3, hence in particular all the rows of $\mathbf{P}$ have degree at most δ. The first item of Lemma 5.6 implies that there exists a matrix $\mathbf{Q} \in \mathbb{K}[X]^{m\times\overline{n}}$ such that all the rows of $[\mathbf{P} \ \ \mathbf{Q}]$ are in $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$ and $\operatorname{rdeg}(\mathbf{Q}) < \operatorname{rdeg}(\mathbf{P})$. Then, by choice of $\mathbf{t}$, we have $\operatorname{lm}_\mathbf{t}([\mathbf{P} \ \ \mathbf{Q}]) =$

$[\mathrm{lm}_{-\delta}(\mathbf{P}) \quad \mathbf{0}]$, with $\mathrm{lm}_{-\delta}(\mathbf{P})$ lower triangular by assumption. Thus $[\mathbf{P} \quad \mathbf{Q}]$ is in $\mathbf{t}$-ordered weak Popov form with all $\mathbf{t}$-pivots in $\mathbf{P}$.

Now, let us write

$$\overline{\mathbf{P}} = \begin{bmatrix} \mathbf{R} & \overline{\mathbf{P}}_{12} \\ \overline{\mathbf{P}}_{21} & \overline{\mathbf{P}}_{22} \end{bmatrix} \text{ with } \mathbf{R} \in \mathbb{K}[X]^{m \times m} \text{ and } \overline{\mathbf{P}}_{22} \in \mathbb{K}[X]^{\overline{n} \times \overline{n}}.$$

Since the $\mathbf{t}$-pivots of $[\mathbf{R} \;\; \overline{\mathbf{P}}_{12}]$ are on the diagonal of $\mathbf{R}$, by minimality of $\overline{\mathbf{P}}$ we obtain $\mathrm{rdeg}_{-\boldsymbol{\delta}}(\mathbf{R}) = \mathrm{rdeg}_{\mathbf{t}}([\mathbf{R} \;\; \overline{\mathbf{P}}_{12}]) \leq \mathrm{rdeg}_{\mathbf{t}}([\mathbf{P} \;\; \mathbf{Q}]) = \mathbf{0}$. Thus $\deg(\mathbf{R}) \leq \max(\boldsymbol{\delta}) \leq \delta$ and $\deg(\overline{\mathbf{P}}_{12}) < \delta$, and the second item of Lemma 5.6 applied to the rows of $[\mathbf{R} \;\; \overline{\mathbf{P}}_{12}]$ shows that each row of $\mathbf{R}$ is in $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Since $\mathbf{R}$ is in $-\boldsymbol{\delta}$-ordered weak Popov form, this gives $\mathrm{rdeg}_{-\boldsymbol{\delta}}(\mathbf{R}) \geq \mathrm{rdeg}_{-\boldsymbol{\delta}}(\mathbf{P}) = \mathbf{0}$ by minimality of $\mathbf{P}$. Thus, we have $\mathrm{rdeg}_{-\boldsymbol{\delta}}(\mathbf{R}) = \mathbf{0}$ and $\mathbf{R}$ is a $-\boldsymbol{\delta}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. $\square$

5.3. Proof of Proposition 5.1

We first give some properties of the manipulated quantities to verify that the assumptions of the lemmas and corollary referred to in the next paragraph are indeed satisfied. In what follows, we let $\overline{\mathbf{F}} = \mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}}$. First, we have $|\boldsymbol{\delta}| \leq \sigma = |\mathbf{d}|$ by Lemma 2.2, hence $\delta = \lceil \sigma/m \rceil \geq \lceil |\boldsymbol{\delta}|/m \rceil$ and thus we can apply Lemma 5.4; it ensures that the tuple $\overline{\boldsymbol{\delta}}$ computed at Step **1** is the $-\overline{\boldsymbol{\delta}}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\overline{\mathbf{F}})$ and satisfies $-\overline{\boldsymbol{\delta}} \geq -\delta$, that is, $\max(\overline{\boldsymbol{\delta}}) \leq \delta$. Besides, since $\mathbf{R}_1$ is in $-\overline{\boldsymbol{\delta}}$-ordered weak Popov form, it has $-\overline{\boldsymbol{\delta}}$-pivot degree $\mathrm{rdeg}_{-\overline{\boldsymbol{\delta}}}(\mathbf{R}_1) + \overline{\boldsymbol{\delta}} = -\hat{\boldsymbol{\delta}} + \overline{\boldsymbol{\delta}}$, by definition of $\hat{\boldsymbol{\delta}}$ at Step **2**. Thus, by the fourth item of Lemma 2.4 and by Proposition 4.1, $\hat{\boldsymbol{\delta}}$ is the $-\hat{\boldsymbol{\delta}}$-minimal degree of $\mathcal{A}_{\hat{\mathbf{d}}}(\hat{\mathbf{F}})$. This further implies $\hat{\boldsymbol{\delta}} \leq \overline{\boldsymbol{\delta}}$, and therefore $\max(\hat{\boldsymbol{\delta}}) \leq \max(\overline{\boldsymbol{\delta}}) \leq \delta$.

By Remark 3.3, Step **4** computes a $\mathbf{t}$-ordered weak Popov basis $\overline{\mathbf{P}}$ of $\mathcal{A}_{\mathcal{L}_\delta(\hat{\mathbf{d}})}(\mathcal{L}_{\hat{\mathbf{d}},\delta}(\hat{\mathbf{F}}))$. Then, Lemma 5.7 applied to $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, -\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\delta}}, \delta)$ shows that $\mathbf{R}_2$ is a $-\hat{\boldsymbol{\delta}}$-ordered weak Popov basis of $\mathcal{A}_{\hat{\mathbf{d}}}(\hat{\mathbf{F}})$. Then, Proposition 4.1 implies that $\mathbf{R}_2\mathbf{R}_1$ is a basis of $\mathcal{A}_{\mathbf{d}}(\overline{\mathbf{F}})$ and the third item of Lemma 2.4 shows that it is in $-\overline{\boldsymbol{\delta}}$-ordered weak Popov form, since $-\hat{\boldsymbol{\delta}} = \mathrm{rdeg}_{-\overline{\boldsymbol{\delta}}}(\mathbf{R}_1)$. It then follows from Lemma 5.4 applied to $(\mathbf{d}, \mathbf{F}, \mathbf{s}, \boldsymbol{\delta})$ that $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$ is the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

Concerning the cost, Steps **1** and **3** use no field operation. At Step **2**, obtaining the matrix $\mathbf{CF} \bmod \mathbf{X}^{\mathbf{d}}$ involves no field operation given the form of $\mathbf{C}$, but only at most $\overline{m}\sigma$ read/write of field elements, where $\overline{m} \leq 2m$ according to Lemma 5.4. Then Proposition 4.1 indicates that Step **2** uses $O(\mathsf{MM}'(m, \sigma/m))$ operations, which is within the announced bound.

From $\overline{\nu} \leq |\hat{\mathbf{d}}|/\delta$ by Definition 5.5 and $|\hat{\mathbf{d}}| \leq \sigma$ by Proposition 4.1, we get $\overline{\nu} \leq \sigma/\lceil \sigma/m \rceil \leq m$. Thus, $\mathcal{L}_{\hat{\mathbf{d}},\delta}(\overline{\mathbf{F}})$ has $\overline{m} + \overline{\nu} \leq 3m$ rows and $\nu + \overline{\nu} < \overline{m} + \overline{\nu} \leq 3m$ columns. Besides, by construction of $\mathcal{L}_\delta(\hat{\mathbf{d}})$ we have $\hat{d} \leq 2\delta = 2\lceil \sigma/m \rceil$, hence $\hat{d} \in O(\sigma/m)$. Note that we can discard the ceiling since we have assumed $\sigma \in \Omega(m)$. Then, according to Proposition 3.2, the call to PM-Basis at Step **4** uses $O(\mathsf{MM}'(\overline{m} + \overline{\nu}, \hat{d})) \subseteq O(\mathsf{MM}'(m, \sigma/m))$ operations.

Now, $\deg(\mathbf{R}_1) \leq 2\sigma/m$ by Proposition 4.1. We have seen that $\mathbf{R}_2$ has $-\hat{\boldsymbol{\delta}}$-pivot degree $\hat{\boldsymbol{\delta}}$, which implies $\mathrm{cdeg}(\mathbf{R}_2) = \hat{\boldsymbol{\delta}}$ by Lemma 2.3. Thus $\deg(\mathbf{R}_2) = \max(\hat{\boldsymbol{\delta}}) \leq \lceil \sigma/m \rceil$, which gives $\deg(\mathbf{R}_2) \in O(\sigma/m)$ (remark that here only the case $\sigma \geq m$ is relevant, since otherwise $n \leq \sigma < m \leq \overline{m}$ and then $\mathbf{R}_1 = \mathbf{I}_{\overline{m}}$). Thus, computing $\mathbf{R}_2\mathbf{R}_1$ uses $O(\mathsf{MM}(m, \sigma/m))$ operations. Then, given the shape of $\mathbf{C}$, obtaining $\mathbf{R}$ from $\mathbf{R}_2\mathbf{R}_1$ uses $O(m\overline{m}\sigma/m) \subseteq O(m\sigma)$ additions in $\mathbb{K}$.

Finally, the computation of $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}$ at Step **5** uses $O(m^\omega)$ operations. Since $\mathrm{cdeg}(\mathbf{R}) = \boldsymbol{\delta}$ by Lemma 2.3 and $|\boldsymbol{\delta}| \leq \sigma$ by Lemma 2.2, applying the first item of Lemma 2.6 with $d = 0$ shows that the product $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$ costs $O(\lceil (m+\sigma)/m \rceil m^\omega)$ operations. Since $\sigma \in \Omega(m)$ this bound is in $O(m^{\omega-1}\sigma)$, which itself is in $O(\mathsf{MM}'(m, \sigma/m))$.

6. Computing approximant bases for arbitrary shifts

We now describe our algorithm for solving the general case of Problem 1 (Algorithm 6), and we prove that it is correct and admits the cost bound announced in Theorem 1.1.

Algorithm 6 – PopovAppBasis *(Shifted Popov approximant basis)*

Input:

- order $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{>0}$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

1. If $\sigma = d_1 + \cdots + d_n \le m$: `// Base case`

  a. For i from 1 to n:

   (i) $\mathbf{E}_i \leftarrow \begin{bmatrix} \mathbf{f}_i^{(0)} & \mathbf{f}_i^{(1)} & \cdots & \mathbf{f}_i^{(d_i-1)} \end{bmatrix} \in \mathbb{K}^{m \times d_i}$ where $\mathbf{F}_{*,i} = \sum_{0 \le k < d_i} \mathbf{f}_i^{(k)} X^k$

   (ii) $\mathbf{Z}_i \leftarrow \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \in \mathbb{K}^{d_i \times d_i}$

  b. $\mathbf{E} \leftarrow \begin{bmatrix} \mathbf{E}_1 & \cdots & \mathbf{E}_n \end{bmatrix} \in \mathbb{K}^{m \times \sigma}$; $\mathbf{Z} \leftarrow \mathrm{diag}(\mathbf{Z}_1, \ldots, \mathbf{Z}_n) \in \mathbb{K}^{\sigma \times \sigma}$

  c. Return LinearizationInterpolationBasis$(\mathbf{E}, \mathbf{Z}, \mathbf{s}, \max(\mathbf{d}))$
   `// (Jeannerod et al., 2017, Algo. 9)`

2. Else if $n \ge m$: `// Entered at most once at initial call`

  a. permute $\mathbf{d}$ into nonincreasing order, and the columns of $\mathbf{F}$ accordingly

  b. $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}}, \mathbf{P}_1) \leftarrow$ ReduceColDim$(\mathbf{d}, \mathbf{F}, \mathbf{s})$

  c. $\mathbf{P}_2 \leftarrow$ PopovAppBasis$(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}})$

  d. $\boldsymbol{\delta}_1 \leftarrow$ diagonal degrees of $\mathbf{P}_1$; $\boldsymbol{\delta}_2 \leftarrow$ diagonal degrees of $\mathbf{P}_2$

  e. Return KnownDegAppBasis$(\mathbf{d}, \mathbf{F}, \mathbf{s}, \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$

3. Else: `// Divide and conquer`

  a. $1 \le i_0 \le n$ and $1 \le d \le d_{i_0}$ such that $d_1 + \cdots + d_{i_0-1} + d = \lfloor \sigma/2 \rfloor$

  b. $\mathbf{f}_{i_0,1} \leftarrow \mathbf{F}_{*,i_0} \bmod X^d$; $\mathbf{f}_{i_0,2} \leftarrow X^{-d}(\mathbf{F}_{*,i_0} - \mathbf{f}_{i_0,1})$

  c. $\mathbf{d}_1 \leftarrow (d_1, \ldots, d_{i_0-1}, d)$; $\mathbf{F}_1 \leftarrow [\mathbf{F}_{*,1}|\cdots|\mathbf{F}_{*,i_0-1}|\mathbf{f}_{i_0,1}]$

  d. $\mathbf{d}_2 \leftarrow (d_{i_0} - d, d_{i_0+1}, \ldots, d_n)$; $\mathbf{F}_2 \leftarrow [\mathbf{f}_{i_0,2}|\mathbf{F}_{*,i_0+1}|\cdots|\mathbf{F}_{*,n}]$

  e. $\mathbf{P}_1 \leftarrow$ PopovAppBasis$(\mathbf{d}_1, \mathbf{F}_1, \mathbf{s})$; $\boldsymbol{\delta}_1 \leftarrow$ diagonal degrees of $\mathbf{P}_1$

  f. $\mathbf{G} \leftarrow \mathbf{P}_1 \mathbf{F}_2 \bmod \mathbf{X}^{\mathbf{d}_2}$ `// using partial linearization`

  g. $\mathbf{P}_2 \leftarrow$ PopovAppBasis$(\mathbf{d}_2, \mathbf{G}, \mathbf{s} + \boldsymbol{\delta}_1)$; $\boldsymbol{\delta}_2 \leftarrow$ diagonal degrees of $\mathbf{P}_2$

  h. Return KnownDegAppBasis$(\mathbf{d}, \mathbf{F}, \mathbf{s}, \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$

Proof of Theorem 1.1. Concerning the base case of the recursion at Step **1**, (Jeannerod et al., 2017, Prop. 7.1) shows that it correctly computes the $\mathbf{s}$-Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ using $O(m^\omega \log(m))$ operations. When the algorithm is called on an instance with $\sigma > m$, Step **1** is performed less than $2\sigma/m$ times in the whole computation, thus leading to a total contribution of $O(m^{\omega-1}\sigma \log(m))$ operations in the cost bound.

Let us now study Step **3**, where $\sigma > m$ and $n < m$. The instance $(\mathbf{d}, \mathbf{F})$ is first split into two instances $(\mathbf{d}_1, \mathbf{F}_1)$ and $(\mathbf{d}_2, \mathbf{F}_2)$ such that $|\mathbf{d}_1| = \lfloor \sigma/2 \rfloor$ and $|\mathbf{d}_2| = \lceil \sigma/2 \rceil$, and with $\mathrm{cdeg}(\mathbf{F}_1) < \mathbf{d}_1$ and $\mathrm{cdeg}(\mathbf{F}_2) < \mathbf{d}_2$. Furthermore, since $n < m$, the column dimensions of both $\mathbf{F}_1$ and $\mathbf{F}_2$ are less than their row dimension, so that the recursive calls at Steps **3.e** and **3.g** will not lead to entering Step **2**. We note that when $d = d_{i_0}$ the first entry of $\mathbf{d}_2$ is zero; then, one can discard this entry and the corresponding zero column of $\mathbf{F}_2$.

At Step **3.f**, the residual $\mathbf{G}$ is computed in $O(\mathsf{MM}(m, \sigma/m))$ operations according to the second item of Lemma 2.6. Indeed, we have $\sigma > m > n$, $|\mathrm{cdeg}(\mathbf{P}_1)| \le \lfloor \sigma/2 \rfloor \le \sigma$ by Lemma 2.2, and $|\mathbf{d}_2| = \lceil \sigma/2 \rceil \le \sigma$ by construction.

Let us define the shift $\mathbf{t} \in \mathbb{Z}^m$ as $\mathbf{t} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1) = \mathbf{s} + \boldsymbol{\delta}_1$. Suppose that the recursive calls correctly compute the $\mathbf{s}$- and $\mathbf{t}$-Popov bases $\mathbf{P}_1$ and $\mathbf{P}_2$ of $\mathcal{A}_{\mathbf{d}_1}(\mathbf{F}_1)$ and $\mathcal{A}_{\mathbf{d}_2}(\mathbf{G})$. Then, the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is $\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2$ according to the fourth item of Lemma 2.4. Thus, by Proposition 5.1, Step **3.h** computes the sought approximant basis in $O(\mathsf{MM}'(m, \sigma/m))$ operations.

The recursive calls (Steps **3.e** and **3.g**) are with the same dimension m and half the total order $\sigma/2$, hence the cost bound in the case $n < m$.

Step **2** deals with the case $n \ge m$, and starts by calling Algorithm 4 to efficiently reduce to $n < m$. According to the above discussion, Step **2** may only be entered once, at the initial call to the algorithm. The correctness and cost bound in the case $n \ge m$ then follow from Proposition 4.1 and from the arguments used above concerning Step **3**. □

7. Computing approximant bases for weakly unbalanced shifts

In this section, we describe approximant basis algorithms which are efficient when the shift is weakly unbalanced around its minimum value (Section 7.1) or around its maximum value (Section 7.2). We recall these notions from Section 1. In the first case, this means that $\mathbf{s}$ satisfies the assumption $\mathcal{H}_{\mathbf{s},\min}$, that is, $|\mathbf{s} - \min(\mathbf{s})| \in O(\sigma)$ with $\sigma = |\mathbf{d}|$. In particular, a balanced shift (that is, one which satisfies $\mathcal{H}_{\mathbf{s},\mathrm{bal}}$: $\max(\mathbf{s}) - \min(\mathbf{s}) \in O(\sigma/m)$) also satisfies $\mathcal{H}_{\mathbf{s},\min}$. In the second case, this means that $\mathbf{s}$ satisfies $\mathcal{H}_{\mathbf{s},\max}$: $|\max(\mathbf{s}) - \mathbf{s}| \in O(\sigma)$.

For shifts satisfying $\mathcal{H}_{\mathbf{s},\min}$, any $\mathbf{s}$-minimal approximant basis $\mathbf{P}$ has small average row degree δ, which means that the overlapping linearization of Section 5.2 at degree δ will efficiently recover a large number of the rows of $\mathbf{P}$ (all those of degree $\le \delta$). Then, Zhou and Labahn (2012) show how the computed rows allow us to discard a corresponding large number of rows and columns in the overlapping linearization at degree 2δ, making it efficient to recover the rows of $\mathbf{P}$ of degree $\le 2\delta$. This process is continued until all rows are obtained.

In Section 7.1, we present a generalization of (Zhou and Labahn, 2012, Algo. 1) which supports arbitrary orders and returns the basis in $\mathbf{s}$-ordered weak Popov form. We do not assume that $\mathbf{s}$ satisfies $\mathcal{H}_{\mathbf{s},\min}$, but we describe the algorithm and a complexity analysis using the parameter $|\mathbf{s} - \min(\mathbf{s})|$ (see Proposition 7.3). Besides, we observe that this generalized algorithm remains efficient: it has the same cost bound as in (ibid., Thm. 5.3) if we assume $\mathcal{H}_{\mathbf{s},\min}$.

For shifts satisfying $\mathcal{H}_{\mathbf{s},\max}$, an $\mathbf{s}$-minimal approximant basis $\mathbf{P}$ may have both large average row degree and large average column degree. Nevertheless, under this assumption, the size of $\mathbf{P}$ remains in $O(m\sigma)$, and we can guess the location of the columns of $\mathbf{P}$ which may have uniformly large degrees: they correspond to the smallest entries of the shift. For example, with $\mathbf{s} = (-\sigma, 0, \ldots, 0)$, only the first column of $\mathbf{P}$ may have all its entries of degree close to σ. Based on this, (ibid., Algo. 2) uses output column linearization to balance the degrees according to this guessed column degree profile of $\mathbf{P}$. This is similar to the output column linearization of

Algorithm 5, except that here we have no guarantee that the guessed column degree is the actual column degree of $\mathbf{P}$. As a result, the linearization will be called a logarithmic number of times, until all rows of $\mathbf{P}$ are revealed. The efficiency of each step depends on the quantity $|\max(\mathbf{s}) - \mathbf{s}|$, which is assumed small in $\mathcal{H}_{\mathbf{s},\max}$.

In Section 7.2, we present a generalization of (ibid., Algo. 2) which supports arbitrary orders and returns a basis in $\mathbf{s}$-ordered weak Popov form. We do not assume that $\mathbf{s}$ satisfies $\mathcal{H}_{\mathbf{s},\max}$ but the cost bound is parametrized by $|\max(\mathbf{s}) - \mathbf{s}|$ (see Proposition 7.4). As above, this generalized algorithm is efficient: it has the same cost bound as in (ibid., Thm. 6.14) if we assume $\mathcal{H}_{\mathbf{s},\max}$.

Before going into detail, we remark that the first item (resp. second item) of Theorem 1.3 follows as a corollary of Proposition 7.3 (resp. Proposition 7.4), although these propositions only prove that we can compute an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ within the claimed cost bound. Indeed, such a basis reveals the $\mathbf{s}$-minimal degree of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ and therefore it only remains to call Algorithm 5, which also fits within the claimed cost bound, to obtain the $\mathbf{s}$-Popov basis.

7.1. Weakly unbalanced shift around its minimum value

Here we consider $\mathbf{s}$-minimal approximant bases for shifts such that $|\mathbf{s} - \min(\mathbf{s})|$ is small. We extend the approach of (Zhou and Labahn, 2012, Sec. 3 to 5) to work with an arbitrary order, and we seek a basis in $\mathbf{s}$-ordered weak Popov form. In this approach, one computes approximants for overlapping linearizations of $(\mathbf{d}, \mathbf{F})$, for a linearization degree parameter δ which is doubled iteratively until a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is obtained. The correctness is based on the next result, which shows how to use the knowledge of a basis of $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$ to find a basis of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$ (see Definition 5.5 for the overlapping linearization giving the matrix $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$ and the order $\mathcal{L}_\delta(\mathbf{d})$).

Hereafter, for $m \in \mathbb{Z}_{>0}$, we write $\mathbf{J}_m$ for the $m \times (\lceil m/2 \rceil - 1)$ matrix whose column k is the column $2k$ of $\mathbf{I}_m$, and $\mathbf{J}_m^{\mathrm{c}}$ for the $m \times (\lfloor m/2 \rfloor + 1)$ submatrix of $\mathbf{I}_m$ formed by the remaining columns. We stress that if m is even, the last column of $\mathbf{I}_m$ does not appear in $\mathbf{J}_m$ but in $\mathbf{J}_m^{\mathrm{c}}$. In particular, $\mathbf{J}_1$ and $\mathbf{J}_2$ are the empty 1×0 and 2×0 matrices, while $\mathbf{J}_2^{\mathrm{c}} = \mathbf{I}_2$. Besides, in what follows $\mathbf{J}_m$ and $\mathbf{J}_m^{\mathrm{c}}$ refer to the 0×0 matrix when $m \in \{-1, 0\}$, and we use the notation $\mathbf{0}_{m\times ?}$ or $\mathbf{0}_{?\times n}$ for the zero matrix when the row dimension m or the column dimension n is not clear from the context.

Lemma 7.1. *Let $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}_{>0}^n$, let $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$, let $\mathbf{s} \in \mathbb{Z}^m$, and let $\delta \in \mathbb{Z}_{>0}$. As in Definition 5.5, let $\alpha_i = \lceil \frac{d_i}{\delta} - 1 \rceil$ for $1 \le i \le n$ and $\overline{n} = \sum_{1\le i\le n} \max(\alpha_i - 1, 0)$. Then, consider the overlapping linearization $\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) \in \mathbb{K}[X]^{(m+\overline{n}_2)\times(n+\overline{n}_2)}$, with*

$$\overline{n}_2 = \sum_{1\le i\le n} \max\left(\left\lceil \frac{d_i}{2\delta} - 1 \right\rceil - 1, 0\right) = \sum_{1\le i\le n} \max(\lfloor \alpha_i/2 \rfloor - 1, 0).$$

We augment this matrix with $\overline{n} - \overline{n}_2$ zero rows in order to define

$$\check{\mathbf{F}}_2 = \mathrm{diag}(\mathbf{I}_m, \mathbf{J}_{\alpha_1 - 1}, \ldots, \mathbf{J}_{\alpha_n - 1})\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) = \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) \\ \mathbf{0} \end{bmatrix} \in \mathbb{K}[X]^{(m+\overline{n})\times(n+\overline{n}_2)},$$

where $\boldsymbol{\pi}$ is the inverse of the permutation matrix

$$\boldsymbol{\pi}^{-1} = \begin{bmatrix} \mathbf{I}_m & & & & & & \\ & \mathbf{J}_{\alpha_1-1} & & & \mathbf{J}_{\alpha_1-1}^{\mathrm{c}} & & \\ & & \ddots & & & \ddots & \\ & & & \mathbf{J}_{\alpha_n-1} & & & \mathbf{J}_{\alpha_n-1}^{\mathrm{c}} \end{bmatrix} \in \mathbb{K}^{(m+\overline{n})\times(m+\overline{n})}.$$

Now define a matrix $\mathbf{S}$ which, through right-multiplication, selects a given set of $n+\overline{n}_2$ columns from any matrix with $n+\overline{n}$ columns, and a matrix $\mathbf{S}^{\mathrm{c}}$ which selects the $\overline{n}-\overline{n}_2$ remaining columns: $\mathbf{S} = \mathrm{diag}(\mathbf{S}_1, \ldots, \mathbf{S}_n) \in \mathbb{K}^{(n+\overline{n})\times(n+\overline{n}_2)}$ *and* $\mathbf{S}^{\mathrm{c}} = \mathrm{diag}(\mathbf{S}^{\mathrm{c}}_1, \ldots, \mathbf{S}^{\mathrm{c}}_n) \in \mathbb{K}^{(n+\overline{n})\times(\overline{n}-\overline{n}_2)}$ *with, for $1 \le i \le n$,*

$$\mathbf{S}_i = \begin{bmatrix} 1 & \\ & \mathbf{J}_{\alpha_i-1} \end{bmatrix} \in \mathbb{K}^{\max(\alpha_i,1)\times\max(\lfloor\alpha_i/2\rfloor,1)} \quad \textit{and} \quad \mathbf{S}^{\mathrm{c}}_i = \begin{bmatrix} \mathbf{0}_{1\times ?} \\ \mathbf{J}^{\mathrm{c}}_{\alpha_i-1} \end{bmatrix} \in \mathbb{K}^{\max(\alpha_i,1)\times(\max(\alpha_i,1)-\max(\lfloor\alpha_i/2\rfloor,1))}.$$

By construction, we have $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S} = \check{\mathbf{F}}_2 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}}$ and $0 \le \mathcal{L}_{2\delta}(\mathbf{d}) - \mathcal{L}_\delta(\mathbf{d})\mathbf{S} \le 2\delta$.

Let us define the order $\check{\mathbf{d}} = (\mathcal{L}_\delta(\mathbf{d}), \mathcal{L}_{2\delta}(\mathbf{d})) \in \mathbb{Z}^{2n+\overline{n}+\overline{n}_2}_{>0}$, the shifts $\check{\mathbf{s}} = (\mathbf{s} - \min(\mathbf{s}), \mathbf{0}) \in \mathbb{Z}^{m+\overline{n}}$ and $\overline{\mathbf{s}} = (\mathbf{s} - \min(\mathbf{s}), \mathbf{0}) \in \mathbb{Z}^{m+\overline{n}_2}$, and the matrix $\check{\mathbf{F}} = [\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) \;\; \check{\mathbf{F}}_2] \in \mathbb{K}[X]^{(m+\overline{n})\times(2n+\overline{n}+\overline{n}_2)}$. Then,

- *For any $\overline{\mathbf{s}}$-ordered weak Popov basis $\mathbf{P} \in \mathbb{K}[X]^{(m+\overline{n}_2)\times(m+\overline{n}_2)}$ of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$, the matrix*

$$\boldsymbol{\pi}^{-1} \begin{bmatrix} \mathbf{P} & -\mathbf{P}_\ell \, \overline{\mathbf{F}} \, \mathbf{S}^{\mathrm{c}} \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \\ \mathbf{0} & \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \end{bmatrix} \boldsymbol{\pi} \in \mathbb{K}[X]^{(m+\overline{n})\times(m+\overline{n})} \tag{9}$$

  is an $\check{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$, where $\mathbf{P}_\ell \in \mathbb{K}[X]^{(m+\overline{n}_2)\times m}$ is the submatrix of $\mathbf{P}$ formed by its leftmost m columns and $\overline{\mathbf{F}} \in \mathbb{K}[X]^{m\times(n+\overline{n})}$ is the submatrix of $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$ formed by its top m rows.
- *For any $\check{\mathbf{s}}$-ordered weak Popov basis $\check{\mathbf{P}} \in \mathbb{K}[X]^{(m+\overline{n})\times(m+\overline{n})}$ of $\mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$, the leading principal $(m+\overline{n}_2)\times(m+\overline{n}_2)$ submatrix of $\boldsymbol{\pi}\check{\mathbf{P}}\boldsymbol{\pi}^{-1}$ is an $\overline{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$.*
- *For any vectors $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ and $\mathbf{q} \in \mathbb{K}[X]^{1\times\overline{n}}$ such that $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p}) \le \delta$ and $[\mathbf{p} \;\; \mathbf{q}] \in \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$, we have $[\mathbf{p} \;\; \mathbf{q}] \in \mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$.*

Proof. (First item.) We define $\mathbf{Q} = -\mathbf{P}_\ell \, \overline{\mathbf{F}} \, \mathbf{S}^{\mathrm{c}} \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \in \mathbb{K}[X]^{(m+\overline{n}_2)\times(\overline{n}-\overline{n}_2)}$ and we denote by $\mathbf{B}$ the matrix in Eq. (9). Then, we start by showing that all rows of $\mathbf{B}$ are in $\mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$, that is, $\mathbf{B}\check{\mathbf{F}}_2 = 0 \bmod \mathbf{X}^{\mathcal{L}_{2\delta}(\mathbf{d})}$ and $\mathbf{B}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})}$. First, we have

$$\mathbf{B}\check{\mathbf{F}}_2 = \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \end{bmatrix} \boldsymbol{\pi}\boldsymbol{\pi}^{-1} \begin{bmatrix} \mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) \\ \mathbf{0} \end{bmatrix} = \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathbf{P}\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) \\ \mathbf{0} \end{bmatrix} = 0 \bmod \mathbf{X}^{\mathcal{L}_{2\delta}(\mathbf{d})}$$

by assumption on $\mathbf{P}$. Since $\mathcal{L}_{2\delta}(\mathbf{d}) \ge \mathcal{L}_\delta(\mathbf{d})\mathbf{S}$, this also gives $\mathbf{B}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S} = \mathbf{B}\check{\mathbf{F}}_2 = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}}$ and thus it remains to show that $\mathbf{B}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}$. By construction, the last $\overline{n}$ rows of $\boldsymbol{\pi}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}}$ are formed by $\overline{n}_2$ zero rows followed by the identity matrix:

$$\begin{bmatrix} \mathbf{0}_{?\times m} & \mathbf{I}_{\overline{n}} \end{bmatrix} \boldsymbol{\pi}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = \begin{bmatrix} (\mathbf{J}_{\alpha_1-1})^\mathsf{T} & & \\ & \ddots & \\ & & (\mathbf{J}_{\alpha_n-1})^\mathsf{T} \\ (\mathbf{J}^{\mathrm{c}}_{\alpha_1-1})^\mathsf{T} & & \\ & \ddots & \\ & & (\mathbf{J}^{\mathrm{c}}_{\alpha_n-1})^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{?\times 1} \;\; \mathbf{I}_{\alpha_1-1} & & \\ & \ddots & \\ & & \mathbf{0}_{?\times 1} \;\; \mathbf{I}_{\alpha_n-1} \end{bmatrix} \begin{bmatrix} \begin{matrix} \mathbf{0}_{1\times ?} \\ \mathbf{J}^{\mathrm{c}}_{\alpha_1-1} \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} \mathbf{0}_{1\times ?} \\ \mathbf{J}^{\mathrm{c}}_{\alpha_n-1} \end{matrix} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{\overline{n}_2\times ?} \\ \mathbf{I}_{\overline{n}-\overline{n}_2} \end{bmatrix}. \tag{10}$$

As a consequence, we have

$$\mathbf{B}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{0} & \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \end{bmatrix} \boldsymbol{\pi}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathbf{P}_\ell \, \overline{\mathbf{F}} \, \mathbf{S}^{\mathrm{c}} + \mathbf{Q} \\ \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}} \end{bmatrix} = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}.$$

Now, we prove that any $\check{\mathbf{p}} \in \mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$ is a combination of the rows of $\mathbf{B}$. Write $\check{\mathbf{p}} = [\mathbf{p} \;\; \mathbf{q}]\boldsymbol{\pi}$ with $\mathbf{p} \in \mathbb{K}[X]^{1\times(m+\overline{n}_2)}$ and $\mathbf{q} \in \mathbb{K}[X]^{1\times(\overline{n}-\overline{n}_2)}$. Then, $\check{\mathbf{p}} \in \mathcal{A}_{\check{\mathbf{d}}}(\check{\mathbf{F}})$ implies first $\mathbf{p} \in \mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$, hence $\mathbf{p} = \boldsymbol{\lambda}\mathbf{P}$ for some $\boldsymbol{\lambda} \in \mathbb{K}[X]^{1\times(m+\bar{n}_2)}$, and second $\boldsymbol{\lambda}\mathbf{P}_\ell\overline{\mathbf{F}}\mathbf{S}^{\mathrm{c}} + \mathbf{q} = [\mathbf{p} \;\; \mathbf{q}]\boldsymbol{\pi}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}$, hence $\mathbf{q} = \boldsymbol{\lambda}\mathbf{Q} + \boldsymbol{\mu}\mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}$ for some $\boldsymbol{\mu} \in \mathbb{K}[X]^{1\times(\bar{n}-\bar{n}_2)}$. Thus, $\check{\mathbf{p}} = [\boldsymbol{\lambda} \;\; \boldsymbol{\mu}]\boldsymbol{\pi}\mathbf{B}$.

It remains to prove that $\boldsymbol{\pi}\mathbf{B}\boldsymbol{\pi}^{-1}$ is in $\check{\mathbf{s}}$-ordered weak Popov form; then, the second item of Lemma 2.8 shows that $\mathbf{B}$ is also in $\check{\mathbf{s}}$-ordered weak Popov form (note that $\check{\mathbf{s}}\boldsymbol{\pi} = \check{\mathbf{s}}$). Since the bottom-right block of $\boldsymbol{\pi}\mathbf{B}\boldsymbol{\pi}^{-1}$ is a diagonal matrix and the top-left block is already in $\bar{\mathbf{s}}$-ordered weak Popov form, where $\check{\mathbf{s}} = (\bar{\mathbf{s}}, \mathbf{0})$, it is enough to show that $\mathrm{rdeg}(\mathbf{Q}) < \mathrm{rdeg}_{\bar{\mathbf{s}}}(\mathbf{P})$. Since $\bar{\mathbf{s}} \geq 0$, we have $\mathrm{rdeg}(\mathbf{P}) \leq \mathrm{rdeg}_{\bar{\mathbf{s}}}(\mathbf{P})$ and thus it is enough to show that $\mathrm{rdeg}(\mathbf{Q}) < \mathrm{rdeg}(\mathbf{P})$. Consider a row $[\mathbf{p} \;\; \mathbf{q}]$ of $[\mathbf{P} \;\; \mathbf{Q}]$. If $\mathrm{rdeg}(\mathbf{p}) \geq 2\delta$, then $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}_{\bar{\mathbf{s}}}(\mathbf{p})$ follows since by construction we have $\mathrm{rdeg}(\mathbf{q}) < \max(\mathcal{L}_\delta(\mathbf{d})) \leq 2\delta$. If $\mathrm{rdeg}(\mathbf{p}) < 2\delta$, since $\mathbf{p}$ is in $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$, the second item of Lemma 5.6 (with parameter 2δ) shows that the m leftmost entries of $\mathbf{p}$ are in $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$; then, the first item of the same lemma (with parameter δ) gives in particular $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p})$.

(Second item.) The first item implies that $\check{\mathbf{P}} = \mathbf{U}\mathbf{B}$ for some unimodular matrix $\mathbf{U}$. Let $\mathbf{U}_0$ and $\mathbf{P}_0$ denote the leading principal $(m + \bar{n}_2) \times (m + \bar{n}_2)$ submatrices of $\boldsymbol{\pi}\mathbf{U}\boldsymbol{\pi}^{-1}$ and $\boldsymbol{\pi}\check{\mathbf{P}}\boldsymbol{\pi}^{-1}$. The first item of Lemma 2.8 shows that $\mathbf{P}_0$ is in $\bar{\mathbf{s}}$-ordered weak Popov form. Besides, the identity $\boldsymbol{\pi}\check{\mathbf{P}}\boldsymbol{\pi}^{-1} = \boldsymbol{\pi}\mathbf{U}\boldsymbol{\pi}^{-1}\boldsymbol{\pi}\mathbf{B}\boldsymbol{\pi}^{-1}$ and the triangular shape of $\boldsymbol{\pi}\mathbf{B}\boldsymbol{\pi}^{-1}$ yield $\mathbf{P}_0 = \mathbf{U}_0\mathbf{P}$. Furthermore, $\boldsymbol{\pi}\check{\mathbf{P}}\boldsymbol{\pi}^{-1}$ and $\boldsymbol{\pi}\mathbf{B}\boldsymbol{\pi}^{-1}$ being $\check{\mathbf{s}}$-ordered weak Popov bases of the same module, they have the same $\check{\mathbf{s}}$-minimal degree (see Section 2.1), and thus the same $\check{\mathbf{s}}$-row degree. This implies that their leading principal submatrices $\mathbf{P}_0$ and $\mathbf{P}$ have the same $\bar{\mathbf{s}}$-row degree, hence

$$\deg(\det(\mathbf{U}_0)) = \deg(\det(\mathbf{P}_0)) - \deg(\det(\mathbf{P})) = |\mathrm{rdeg}_{\bar{\mathbf{s}}}(\mathbf{P}_0)| - |\mathrm{rdeg}_{\bar{\mathbf{s}}}(\mathbf{P})| = 0.$$

This means that $\mathbf{U}_0$ is unimodular, and therefore $\mathbf{P}_0$ is a basis of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$.

(Third item.) We want to prove that $[\mathbf{p} \;\; \mathbf{q}] \in \mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\check{\mathbf{F}}_2)$. The second item of Lemma 5.6 implies that $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$, while its first item gives the uniqueness of $\mathbf{q}$: if $\mathbf{r} \in \mathbb{K}[X]^{1\times\bar{n}}$ is such that $\mathrm{rdeg}(\mathbf{r}) < \mathrm{rdeg}(\mathbf{p})$ and $[\mathbf{p} \;\; \mathbf{r}] \in \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$, then $\mathbf{r} = \mathbf{q}$. (Note that here the constraint $\mathrm{cdeg}(\mathbf{r}) < \mathcal{L}_\delta(\mathbf{d})\mathbf{E}^{\mathsf{T}}$ from Lemma 5.6 is implied by $\mathrm{rdeg}(\mathbf{r}) < \delta < \min(\mathcal{L}_\delta(\mathbf{d})\mathbf{E}^{\mathsf{T}})$.)

Lemma 5.6 gives $\mathbf{q}_2 \in \mathbb{K}[X]^{1\times\bar{n}_2}$ such that $\mathrm{rdeg}(\mathbf{q}_2) < \mathrm{rdeg}(\mathbf{p})$ and $[\mathbf{p} \;\; \mathbf{q}_2] \in \mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$. Then, define $\mathbf{q}_3 = -\mathbf{p}\overline{\mathbf{F}}\mathbf{S}^{\mathrm{c}} \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}$, which is a subvector of $\mathbf{q} = -\mathbf{p}\overline{\mathbf{F}}\mathbf{E}^{\mathsf{T}} \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{E}^{\mathsf{T}}}$ since $\mathbf{S}^{\mathrm{c}}$ selects a subset of the columns selected by $\mathbf{E}^{\mathsf{T}}$. Let further $\mathbf{r} = [\mathbf{q}_2 \;\; \mathbf{q}_3][\mathbf{0} \;\; \mathbf{I}_{\bar{n}}]\boldsymbol{\pi} \in \mathbb{K}[X]^{1\times\bar{n}}$; by construction, we have $\mathrm{rdeg}(\mathbf{r}) < \mathrm{rdeg}(\mathbf{p})$. We are going to show that $[\mathbf{p} \;\; \mathbf{r}] \in \mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\check{\mathbf{F}}_2)$ and $[\mathbf{p} \;\; \mathbf{r}] \in \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$: the latter point implies $\mathbf{r} = \mathbf{q}$ by the mentioned uniqueness, and then the former point gives $[\mathbf{p} \;\; \mathbf{q}] \in \mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\check{\mathbf{F}}_2)$, thus concluding the proof.

Noticing that $[\mathbf{p} \;\; \mathbf{r}] = [\mathbf{p} \;\; \mathbf{q}_2 \;\; \mathbf{q}_3]\boldsymbol{\pi}$, the first point follows by construction of $\check{\mathbf{F}}_2$:

$$[\mathbf{p} \;\; \mathbf{r}]\check{\mathbf{F}}_2 = [\mathbf{p} \;\; \mathbf{q}_2 \;\; \mathbf{q}_3]\boldsymbol{\pi}\boldsymbol{\pi}^{-1}\begin{bmatrix}\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) \\ \mathbf{0}\end{bmatrix} = [\mathbf{p} \;\; \mathbf{q}_2]\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}) = 0 \bmod \mathbf{X}^{\mathcal{L}_{2\delta}(\mathbf{d})}.$$

Furthermore, since $\mathcal{L}_{2\delta}(\mathbf{d}) \geq \mathcal{L}_\delta(\mathbf{d})\mathbf{S}$ we can consider the same identity modulo $\mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}}$. Using $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S} = \check{\mathbf{F}}_2 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}}$, this directly yields $[\mathbf{p} \;\; \mathbf{r}]\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S} = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}}$. For the second point, it remains to show $[\mathbf{p} \;\; \mathbf{r}]\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = 0 \bmod \mathbf{X}^{\mathcal{L}_\delta(\mathbf{d})\mathbf{S}^{\mathrm{c}}}$. This follows from the definition of $\mathbf{q}_3$ since Eq. (10) gives $[\mathbf{p} \;\; \mathbf{r}]\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = [\mathbf{p} \;\; \mathbf{q}_2 \;\; \mathbf{q}_3]\boldsymbol{\pi}\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})\mathbf{S}^{\mathrm{c}} = \mathbf{p}\overline{\mathbf{F}}\mathbf{S}^{\mathrm{c}} + \mathbf{q}_3$. □

We remark that working with matrices in ordered weak Popov form allows us to directly locate the submatrix that contains the sought basis, and thus to avoid resorting to computations of row rank profiles as was done for example in (Zhou and Labahn, 2012, Thm. 3.15 and Algo. 1).

The second item in this lemma implies that, knowing a basis of $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$, we can obtain a basis of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$ via the classical approach of computing a residual, a second approximant basis, and the product of the two bases. Furthermore, the third item shows that rows of degree less than δ in the first basis are already in $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$. Thus, they can be discarded when computing the second basis (see Lemma 2.5); this is a key property for the efficiency of Algorithm 7. The next result formalizes these remarks, using notation from Lemma 7.1.

Corollary 7.2. *Let $\mathbf{P} \in \mathbb{K}[X]^{(m+\overline{n})\times(m+\overline{n})}$ be an $\check{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$, let $I \subseteq \{1,\ldots,m+\overline{n}\}$ be the set of indices i of the rows $\mathbf{P}_{i,*} = [\mathbf{p} \;\; \mathbf{q}]$ such that $\operatorname{rdeg}(\mathbf{q}) < \operatorname{rdeg}(\mathbf{p}) \leq \delta$, where $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ and $\mathbf{q} \in \mathbb{K}[X]^{1\times\overline{n}}$. Let further $I^c = \{1,\ldots,m+\overline{n}\} \setminus I$ be the complement of I and let i denote the cardinality of I. We have $I \subseteq \{1,\ldots,m\}$.*

Now, consider the tuples $\boldsymbol{\mu} = \mathcal{L}_\delta(\mathbf{d})\mathbf{S}$ and $\boldsymbol{\nu} = \mathcal{L}_{2\delta}(\mathbf{d})$ both in $\mathbb{Z}_{>0}^{n+\overline{n}_2}$, as well as the residual $\mathbf{G} = \mathbf{P}_{I^c,*}\check{\mathbf{F}}_2\mathbf{X}^{-\boldsymbol{\mu}} \bmod \mathbf{X}^{\boldsymbol{\nu}-\boldsymbol{\mu}} \in \mathbb{K}[X]^{(m+\overline{n}-i)\times(n+\overline{n}_2)}$ and a basis $\mathbf{P}_2 \in \mathbb{K}[X]^{(m+\overline{n}-i)\times(m+\overline{n}-i)}$ of $\mathcal{A}_{\boldsymbol{\nu}-\boldsymbol{\mu}}(\mathbf{G})$ in $\operatorname{rdeg}_{\check{\mathbf{s}}}(\mathbf{P}_{I^c,*})$-ordered weak Popov form. Modify $\mathbf{P}$ by left-multiplying its submatrix $\mathbf{P}_{I^c,*}$ by $\mathbf{P}_2$, that is, perform the operation $\mathbf{P}_{I^c,*} \leftarrow \mathbf{P}_2\mathbf{P}_{I^c,*}$. Then, the leading principal $(m+\overline{n}_2)\times(m+\overline{n}_2)$ submatrix of $\boldsymbol{\pi}\mathbf{P}\boldsymbol{\pi}^{-1}$ is an $\overline{\mathbf{s}}$-ordered weak Popov basis of $\mathcal{A}_{\mathcal{L}_{2\delta}(\mathbf{d})}(\mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F}))$.

Proof. The fact that $I \subseteq \{1,\ldots,m\}$ follows by definition of the $\check{\mathbf{s}}$-ordered weak Popov form. Indeed, since $\check{\mathbf{s}} = (\mathbf{s}-\min(\mathbf{s}), \mathbf{0})$, such a row $[\mathbf{p} \;\; \mathbf{q}]$ with $\operatorname{rdeg}(\mathbf{q}) < \operatorname{rdeg}(\mathbf{p}) \leq \operatorname{rdeg}_{\mathbf{s}-\min(\mathbf{s})}(\mathbf{p})$ must have its $\check{\mathbf{s}}$-pivot entry in $\mathbf{p}$, or in other words, its $\check{\mathbf{s}}$-pivot index in $\{1,\ldots,m\}$. Since the $\check{\mathbf{s}}$-pivot entries are on the diagonal, $[\mathbf{p} \;\; \mathbf{q}]$ must be one of the first m rows of $\mathbf{P}$.

The other claims follow directly from Lemmas 7.1 and 2.5. □

This suggests an algorithm which computes approximant bases iteratively for the overlapping linearized problems with a linearization parameter δ which is doubled at each step. When the parameter reaches $\delta > \max(\mathbf{d})$, we actually have $\mathcal{L}_\delta(\mathbf{d}) = \mathbf{d}$ and $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) = \mathbf{F}$, and therefore the computed basis is a basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F}) = \mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$. In what follows, let $\sigma = |\mathbf{d}|$.

In this process, the number of columns of the approximant instances steadily decreases. On the first hand, the number of columns $\overline{n}$ added by the overlapping linearization is roughly halved when δ is doubled. On the other hand, only the $\leq 2\sigma/\delta$ columns of $\mathbf{F}$ with corresponding order $d_i \geq \delta/2$ need to be considered in the iteration with linearization parameter δ, since all the others have been fully processed already (see the proof of Proposition 7.3 for more details).

Furthermore, the corollary above indicates that if at some iteration one of the computed approximants in $\mathcal{A}_{\mathcal{L}_\delta(\mathbf{d})}(\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}))$ has degree less than δ, then it can be stored as a row of the sought basis and can be discarded in the computation of the residual and of the second basis. In the process outlined above, this allows us to decrease the row dimension each time such a small degree approximant has been found.

Yet, there remains an obstacle towards efficiency: if the output basis has no row of small degree, there will be no such row dimension decrease before the very last few iterations. In this case, some iterations may ask us to solve instances with roughly the same dimensions and degrees as the original instance $(\mathbf{d},\mathbf{F})$; then, this approach is not faster than a direct call to PM-Basis.

Nevertheless, there are many shifts for which this worst-case scenario cannot occur, since the sum of the row degree of an $\mathbf{s}$-minimal basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$ is at most $\xi = \sigma + |\mathbf{s}-\min(\mathbf{s})|$ (Van Barel and Bultheel, 1992, Thm. 4.1). Thus, this $\mathbf{s}$-minimal basis has at most ξ/δ rows of degree $\geq \delta$; this is especially beneficial when ξ is small, that is, for shifts that are weakly unbalanced around their minimum value (assumption $\mathcal{H}_{\mathbf{s},\min}$). For example, for the uniform shift, a $\mathbf{0}$-minimal basis has at most $m/2^i$ rows of degree $\geq 2^i\lceil\sigma/m\rceil$, which means that in our process at least $m - m/2^i$ rows can be discarded when δ has reached $2^i\lceil\sigma/m\rceil$.

Proposition 7.3. *Algorithm 7 is correct. Let $\sigma = |\mathbf{d}|$, let $\xi = \sigma + |\mathbf{s}-\min(\mathbf{s})|$, and let $d = \max(\mathbf{d})$. If $\xi \leq md$, then Algorithm 7 uses $C(\xi,m,d)$ operations in $\mathbb{K}$, where $C(\cdot)$ is defined as in Eq. (3). If $\xi > md$, it uses $O(\mathsf{MM}'(m,\lceil\sigma/m\rceil) + \mathsf{MM}'(m,d))$ operations in $\mathbb{K}$.*

Algorithm 7 – $\textsc{ShiftAroundMinAppBasis}$ *(Minimal basis for small* $|\mathbf{s} - \min(\mathbf{s})|$*)*

Input:

- order $\mathbf{d} \in \mathbb{Z}_{>0}^{n}$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m \times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$,
- shift $\mathbf{s} \in \mathbb{Z}^{m}$.

Output: an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

1. If $n \geq m$:

  a. permute $\mathbf{d}$ into nonincreasing order, and the columns of $\mathbf{F}$ accordingly
  b. $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}}, \mathbf{P}_1) \leftarrow \textsc{ReduceColDim}(\mathbf{d}, \mathbf{F}, \mathbf{s})$
  c. $\mathbf{P}_2 \leftarrow \textsc{ShiftAroundMinAppBasis}(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}})$
  d. $\boldsymbol{\delta}_1 \leftarrow$ diagonal degrees of $\mathbf{P}_1$; $\boldsymbol{\delta}_2 \leftarrow$ diagonal degrees of $\mathbf{P}_2$
  e. Return $\textsc{KnownDegAppBasis}(\mathbf{d}, \mathbf{F}, \mathbf{s}, \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$

2. Else:

  a. $\delta \leftarrow \lceil (|\mathbf{d}| + |\mathbf{s} - \min(\mathbf{s})|)/m \rceil$
  Construct $\mathcal{L}_{\delta}(\mathbf{d}) \in \mathbb{Z}_{>0}^{m+\overline{n}}$ and $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) \in \mathbb{K}[X]^{(m+\overline{n}) \times (n+\overline{n})}$ as in Definition 5.5
  $\mathbf{P} \leftarrow \textsc{PM-Basis}(2\delta, \mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) \mathbf{X}^{2\delta - \mathcal{L}_{\delta}(\mathbf{d})}, (\mathbf{s} - \min(\mathbf{s}), \mathbf{0}))$
  $I \leftarrow \{i \in \{1, \ldots, m + \overline{n}\} \mid \mathbf{P}_{i,*} = [\mathbf{p} \;\; \mathbf{q}]$ is such that $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p}) \leq \delta\}$,
  where $\mathbf{p} \in \mathbb{K}[X]^{1 \times m}$ and $\mathbf{q} \in \mathbb{K}[X]^{1 \times \overline{n}}$ `// for these rows,` $\mathbf{p} \in \mathcal{A}_{\mathbf{d}}(\mathbf{F})$

  b. While $\mathrm{Card}(I) < m$: `//` $I \subseteq \{1, \ldots, m\}$ `holds, cf. Corollary 7.2`

    (i) Construct matrices $\boldsymbol{\pi} \in \mathbb{K}^{(m+\overline{n}) \times (m+\overline{n})}$ and $\mathbf{S} \in \mathbb{K}^{(n+\overline{n}) \times (n+\overline{n}_2)}$ as in Lemma 7.1,
    tuples $\boldsymbol{\mu} \leftarrow \mathcal{L}_{\delta}(\mathbf{d}) \mathbf{S}$ and $\boldsymbol{\nu} \leftarrow \mathcal{L}_{2\delta}(\mathbf{d})$ both in $\mathbb{Z}_{>0}^{n+\overline{n}_2}$, and sets
    $J \leftarrow \{j \in \{1, \ldots, n + \overline{n}_2\} \mid \nu_j - \mu_j > 0\}$ and $I^{c} \leftarrow \{1, \ldots, m + \overline{n}\} \setminus I$
    (ii) $\mathbf{G} \leftarrow \mathbf{P}_{I^{c},*} \boldsymbol{\pi}^{-1} \begin{bmatrix} \mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F})_{*,J} \\ \mathbf{0} \end{bmatrix} \mathbf{X}^{-\boldsymbol{\mu}_J} \bmod \mathbf{X}^{\boldsymbol{\nu}_J - \boldsymbol{\mu}_J}$
    (iii) $\mathbf{P}_2 \leftarrow \textsc{PM-Basis}(2\delta, \mathbf{G}\mathbf{X}^{2\delta - \boldsymbol{\nu}_J + \boldsymbol{\mu}_J}, \mathrm{rdeg}_{(\mathbf{s} - \min(\mathbf{s}), \mathbf{0})}(\mathbf{P}_{I^{c},*}))$
    (iv) $\mathbf{P}_{I^{c},*} \leftarrow \mathbf{P}_2 \mathbf{P}_{I^{c},*}$ `// this modifies` $\mathbf{P}$
    (v) $\mathbf{P} \leftarrow$ leading principal $(n + \overline{n}_2) \times (n + \overline{n}_2)$ submatrix of $\boldsymbol{\pi} \mathbf{P} \boldsymbol{\pi}^{-1}$
    $\delta \leftarrow 2\delta$; $\overline{n} \leftarrow \overline{n}_2$; $I \leftarrow I \cup \{i \in I^{c} \mid \mathbf{P}_{i,*} = [\mathbf{p} \;\; \mathbf{q}]$ is such that $\mathrm{rdeg}(\mathbf{q}) < \mathrm{rdeg}(\mathbf{p}) \leq \delta\}$, where $\mathbf{p} \in \mathbb{K}[X]^{1 \times m}$ and $\mathbf{q} \in \mathbb{K}[X]^{1 \times \overline{n}}$

  c. Return $\mathbf{P}$

Proof. The correctness of Step **1** follows from Lemma 2.4 and Proposition 4.1. Concerning Step **2**, we first note that if $\lceil \xi/m \rceil > d$, then $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) = \mathbf{F}$ and $\mathcal{L}_{\delta}(\mathbf{d}) = \mathbf{d}$ and therefore the call to PM-Basis at Step **2.a** computes a whole **s**-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$. Then, the loop at Step **2.b** is not entered, and Step **2** uses $O(\mathsf{MM}'(m,d))$ operations according to Proposition 3.2.

On the other hand, if $\lceil \xi/m \rceil \leq d$, the correctness of Step **2** follows from Corollary 7.2, noticing that the loop terminates after at most $1 + \lfloor \log_2(d/\lceil \xi/m \rceil) \rfloor$ iterations since δ is doubled at each iteration, and as mentioned above $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F}) = \mathbf{F}$ and $\mathcal{L}_{\delta}(\mathbf{d}) = \mathbf{d}$ for $\delta > d$. Furthermore, in this algorithm we use the set J to explicitly filter out columns for which the correct order has already been reached, thus for which the residual columns are zero. This was not done in Corollary 7.2 which focused on correctness, yet here it makes it easier to describe column dimensions in the following cost analysis.

Concerning Step **2**, we place ourselves at the beginning of an iteration, and we start by describing the dimensions and the degrees of the matrices involved in the computations. Then,

- $\mathbf{P}_{I^c,*}$ has dimensions $\mathrm{Card}(I^c) \times (m + \overline{n})$ and degree $< 2\delta$;
- $\boldsymbol{\pi}^{-1} \begin{bmatrix} \mathcal{L}_{\mathbf{d},2\delta}(\mathbf{F})_{*,J} \\ \mathbf{0} \end{bmatrix}$ has dimensions $(m + \overline{n}) \times \mathrm{Card}(J)$ and degree $< \max(\mathcal{L}_{2\delta}(\mathbf{d})) \leq 4\delta$;
- $\mathbf{G}$ has dimensions $\mathrm{Card}(I^c) \times \mathrm{Card}(J)$ and degree $< \max(\boldsymbol{\nu} - \boldsymbol{\mu}) \leq 2\delta$;
- $\mathbf{P}_2$ has dimensions $\mathrm{Card}(I^c) \times \mathrm{Card}(I^c)$ and degree $< 2\delta$.

As above, $\overline{n}$ is such that $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$ has dimensions $(m + \overline{n}) \times (n + \overline{n})$, and $\overline{n} < \sigma/\delta$ where $\sigma = |\mathbf{d}|$.

Besides, as a consequence of (Van Barel and Bultheel, 1992, Thm. 4.1), the sum of the degrees of the rows of the sought basis is at most ξ, and thus this basis has more than $m - \xi/\delta$ rows of degree $\leq \delta$; Lemma 5.6 shows that the set $I \subseteq \{1, \ldots, m\}$ precisely contains the indices of the latter rows. Thus, $\mathrm{Card}(I) > m - \xi/\delta$, and $\mathrm{Card}(I^c) = m + \overline{n} - \mathrm{Card}(I) < \overline{n} + \xi/\delta \leq 2\xi/\delta$.

Furthermore, note that the entries of $\mathcal{L}_{\delta}(\mathbf{d})\mathbf{S}$ and $\mathcal{L}_{2\delta}(\mathbf{d})$ which coincide are exactly those corresponding to columns with order $d_i \leq 2\delta$ (or, equivalently, $\alpha_i = 1$): these are columns $\mathbf{F}_{*,i}$ which appear as such in $\mathcal{L}_{\mathbf{d},\delta}(\mathbf{F})$ and also in $\mathcal{L}_{\mathbf{d},2^k\delta}(\mathbf{F})$ for all subsequent iterations. Indeed, if $d_i > 2\delta$, the corresponding entries in $\mathcal{L}_{\delta}(\mathbf{d})\mathbf{S}$ are at most 2δ and cannot coincide with those in $\mathcal{L}_{2\delta}(\mathbf{d})$ which are at least $2\delta + 1$. As a result, $\mathrm{Card}(J)$ is the sum of the number $\overline{n}_2$ of columns added by the overlapping linearization with degree parameter 2δ, and of the number of indices $i \in \{1, \ldots, n\}$ such that $d_i > 2\delta$; both numbers are less than $\sigma/(2\delta)$. Thus, $\mathrm{Card}(J) < \sigma/\delta$.

Now, let $\delta_0 = \lceil \xi/m \rceil$ be the initial value of δ. Then, at the beginning of the k-th iteration of the loop (the first one being for $k = 1$), we have $\delta = 2^{k-1}\delta_0$ and the dimensions satisfy $m + \overline{n} < 2m$, $\mathrm{Card}(I^c) < 2\xi/\delta = 2^{2-k}\xi/\delta_0 \leq 2^{2-k}m$, and $\mathrm{Card}(J) < \sigma/\delta = 2^{1-k}\sigma/\delta_0 \leq 2^{1-k}\xi/\delta_0 \leq 2^{1-k}m$.

Then, both matrix multiplications at Steps **2.b.**(ii) and **2.b.**(iv) use $O(2^{k-1}\mathsf{MM}(2^{1-k}m, 2^{k-1}\delta_0))$ operations. Besides, the call to PM-Basis at Step **2.b.**(iii) uses $O(\mathsf{MM}'(2^{1-k}m, 2^{k-1}\delta_0))$ operations according to Proposition 3.2, while the call at Step **2.a** uses $O(\mathsf{MM}'(m, \delta_0))$ operations. Summing these terms over all iterations gives the cost bound announced in the statement, since as explained above the loop terminates before or when k reaches $1 + \lfloor \log_2(d/\lceil \xi/m \rceil) \rfloor$.

Now, independently from assumptions on $\lceil \xi/m \rceil$, Steps **1.b** and **1.e** both use $O(\mathsf{MM}'(m, \sigma/m))$ operations according to Propositions 4.1 and 5.1; here $\lceil \sigma/m \rceil \in \Theta(\sigma/m)$ since $\sigma \geq n \geq m$. Besides, the former proposition and the specification of ReduceColDim ensure that:

- $\deg(\mathbf{P}_1) \leq 2\sigma/m$, hence $\mathbf{s} \leq \hat{\mathbf{s}} \leq \mathbf{s} + 2\sigma/m$ since $\hat{\mathbf{s}} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1)$;
- $|\hat{\mathbf{d}}| \leq \sigma$, hence $\sigma + |\hat{\mathbf{s}} - \min(\hat{\mathbf{s}})| \leq \xi + 2\sigma \leq 3\xi$;
- $\hat{\mathbf{F}}$ has fewer columns than rows, hence the call at Step **1.c** will enter Step **2**.

Then, the cost bounds given above hold for Step **1.c**: if $\lceil \xi/m \rceil \leq d$ this step is thus the bottleneck of Step **1**, and if $\lceil \xi/m \rceil > d$ we obtain the claimed bound $O(\mathsf{MM}'(m, \sigma/m) + \mathsf{MM}'(m, d))$. □

We remark that it would also be correct, instead of Steps **1.d** and **1.e**, to directly compute and return the product $\mathbf{P}_2\mathbf{P}_1$; this uses $O(\mathsf{MM}(m, \lceil \xi/m \rceil))$ operations and thus does not impact the cost bound if $\xi \in O(\sigma)$. In addition, for input instances with $\sigma \ll m$, one may rather rely on linear algebra over $\mathbb{K}$ instead of the above algorithm (see Steps **1.a**, **1.b**, and **1.c** of Algorithm 6).

We now show the upper bound on $C(\xi, m, d)$ given in Theorem 1.3, for the case $\xi \leq md$. Under the assumption $\mathcal{H}_{\mathsf{M}}$, we obtain

$$\begin{aligned}
&\mathsf{MM}'(2^{-k}m, 2^k\lceil \xi/m \rceil) + 2^k\mathsf{MM}(2^{-k}m, 2^k\lceil \xi/m \rceil) \\
&\quad \in O\Big((2^{-k}m)^\omega \mathsf{M}(2^k\lceil \xi/m \rceil)\log(2^k\lceil \xi/m \rceil) + 2^k(2^{-k}m)^\omega \mathsf{M}(2^k\lceil \xi/m \rceil)\Big) \\
&\quad \subseteq O\Big(m^\omega \mathsf{M}(\lceil \xi/m \rceil)(2^{-k}(k + \log(\lceil \xi/m \rceil)) + 1)\Big),
\end{aligned}$$

since $\mathcal{H}_{\mathsf{M}}$ implies in particular $\mathsf{M}(2^k\lceil \xi/m \rceil) \in O(2^{(\omega-1)k}\mathsf{M}(\lceil \xi/m \rceil))$. Since $\sum_{k\geq 0} k2^{-k}$ is the constant 2, summing over $0 \leq k \leq 1 + \log(d/\lceil \xi/m \rceil)$ gives the sought bound

$$C(\xi, m, d) \in O(m^\omega \mathsf{M}(\lceil \xi/m \rceil)(\log(\lceil \xi/m \rceil) + \log(d/\lceil \xi/m \rceil))) = O(m^\omega \mathsf{M}(\lceil \xi/m \rceil)\log(d)),$$

valid under $\mathcal{H}_{\mathsf{M}}$ and for an arbitrary order and shift.

We remark that the latter bound is precisely the one which was obtained (Zhou and Labahn, 2012, Thm. 5.3), under the additional assumptions that $\xi \in O(\sigma)$ and that $\mathbf{d} = (d, \ldots, d) \in \mathbb{Z}^n_{>0}$ with $n \leq m \leq \sigma = nd$; in that case the bound can be written $O(m^\omega \mathsf{M}(nd/m)\log(d))$.

7.2. Weakly unbalanced shift around its maximum value

Here, we will only sketch the correctness and cost bound of the algorithm, and refer to (Zhou and Labahn, 2012, Sec. 6) for more details and examples. Indeed, it can be noticed that the output column linearization does not modify the order $\mathbf{d}$ and does not depend on it. As a result, generalizing (ibid., Algo. 2) to the case of arbitrary orders was mostly done in Section 5.1 where the definition and properties of the output column linearization were presented.

In Algorithm 8, we interrupt the iterative use of output column linearization as soon as it becomes more efficient to directly resort to PM-Basis (Step **4**). We remark that, while this may seem to differ from (ibid., Algo. 2), it is in fact mentioned in the proof of (ibid., Thm. 6.14) that the algorithm should behave like this to avoid weakening its efficiency.

We recall that $C(\cdot)$ was defined in Eq. (3).

Proposition 7.4. *Algorithm 8 is correct. Let $\sigma = |\mathbf{d}|$, let $\zeta = \sigma + |\max(\mathbf{s}) - \mathbf{s}|$, and let $d = \max(\mathbf{d})$. If $\zeta > md$, Algorithm 8 uses $O(\mathsf{MM}'(m, \lceil \sigma/m \rceil) + \mathsf{MM}'(m, d))$ operations in $\mathbb{K}$. If $\zeta \leq md$, it uses*

$$O\left(\mathsf{MM}'(\mu, \lceil \sigma/\mu \rceil) + \mathsf{MM}'(\mu, d) + \sum_{k=0}^{\lfloor \log_2(md/\zeta) \rfloor} C(\zeta, 2^{-k}m, d)\right)$$

*operations in $\mathbb{K}$, where $C(\cdot)$ is defined as in Eq. (3) and μ is the cardinality of the set I after Step **3** has been performed; it is such that $\mu < \zeta/d$.*

Proof. First, if $\zeta > md$, the loop at Step **3** is not entered, and at Step **4** we have $I = \{1, \ldots, m\}$; in particular, $\mathbf{P}_{I,*} = \mathbf{P}$ and Step **4.f** simply amounts to $\mathbf{P} \leftarrow \overline{\mathbf{P}}$. In this case, the correctness and cost bound follow from Propositions 4.1 and 3.2, the fourth item of Lemma 2.4, and Proposition 5.1.

From now on, suppose $\zeta \leq md$. The same results prove the correctness of Step **4** while Lemma 5.3 proves that of Step **3**, using in addition (Zhou and Labahn, 2012, Thm. 6.11) to show

Algorithm 8 – ShiftAroundMaxAppBasis *(Minimal basis for small* $|\max(\mathbf{s}) - \mathbf{s}|$*)*

Input:

- order $\mathbf{d} \in \mathbb{Z}^n_{>0}$,
- matrix $\mathbf{F} \in \mathbb{K}[X]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d}$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: an $\mathbf{s}$-ordered weak Popov basis of $\mathcal{A}_{\mathbf{d}}(\mathbf{F})$.

1. $\mathbf{P} \leftarrow$ empty matrix in $\mathbb{K}[X]^{0\times m}$

2. $I \leftarrow \{1,\dots,m\}$ `// indices of rows still to be found`

3. While $\sigma + |\max(\mathbf{s}) - \mathbf{s}| \le \mathrm{Card}(I)d$:

 a. $\delta \leftarrow 1 + 2\lfloor |\max(\mathbf{s}) - \mathbf{s}|/\mathrm{Card}(I)\rfloor$

 b. $(\overline{\mathbf{s}}, \mathbf{C}, (\alpha_i)_{1\le i\le m}, \overline{m}) \leftarrow$ ColParLin$(\mathbf{s}_I, \delta, \delta)$ `// see Section 5.1`

 c. $\overline{\mathbf{P}} \leftarrow$ ShiftAroundMinAppBasis$(\mathbf{d}, \mathbf{C}\mathbf{F}_{I,*} \bmod \mathbf{X^d}, \overline{\mathbf{s}})$

 d. $\mathbf{E} \in \mathbb{K}^{m\times m} \leftarrow \mathrm{diag}(e_1,\dots,e_m)$ with $e_i = 1$ if $i \in I$ and $e_i = 0$ otherwise

 e. For $i \in I$ such that $s_i \ge \max(\mathbf{s}) - \delta$ or $\mathrm{rdeg}_{\overline{\mathbf{s}}}(\overline{\mathbf{P}}_{\alpha_1+\cdots+\alpha_i,*}) > 0$:

  $\mathbf{P}_{i,*} \leftarrow \overline{\mathbf{P}}_{\alpha_1+\cdots+\alpha_i,*}\mathbf{C}\mathbf{E}$; $I \leftarrow I \setminus \{i\}$

4. If $I \ne \emptyset$: `// compute remaining rows via PM-Basis`

 a. permute $\mathbf{d}$ into nonincreasing order, and the columns of $\mathbf{F}_{I,*}$ accordingly

 b. $(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}}, \mathbf{P}_1) \leftarrow$ ReduceColDim$(\mathbf{d}, \mathbf{F}_{I,*}, \mathbf{s}_I)$

 c. $\mathbf{P}_2 \leftarrow$ PM-Basis$(\hat{\mathbf{d}}, \hat{\mathbf{F}}, \hat{\mathbf{s}})$

 d. $\boldsymbol{\delta}_1 \leftarrow$ diagonal degrees of $\mathbf{P}_1$; $\boldsymbol{\delta}_2 \leftarrow$ diagonal degrees of $\mathbf{P}_2$

 e. $\overline{\mathbf{P}} \leftarrow$ KnownDegAppBasis$(\mathbf{d}, \mathbf{F}_{I,*}, \mathbf{s}_I, \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$

 f. $\mathbf{P}_{I,*} \leftarrow \overline{\mathbf{P}}\,\mathrm{diag}(e_1,\dots,e_m)$, where $e_i = 1$ if $i \in I$ and $e_i = 0$ otherwise

5. Return $\mathbf{P}$

that we may discard the rows of $\mathbf{F}$ with index not in I (Steps **3.b** and **4.b**) and fill corresponding columns of $\mathbf{P}$ with zeroes (multiplication by $\mathbf{E}$ in Step **3.e** and by the diagonal in **4.f**).

Furthermore, the above propositions show that Step **4** uses $O(\mathsf{MM}'(\mu, \lceil\sigma/\mu\rceil) + \mathsf{MM}'(\mu, d))$ operations; since the loop at Step **3** has exited, we have $\zeta > \mu d$.

Concerning Step **3**, the main point is that the cardinality of I is at least halved at the end of each iteration of the While loop. Indeed, let $c > 0$ be the cardinality of I at the beginning of an iteration; hence $\delta > 2|\max(\mathbf{s}) - \mathbf{s}|/c$. Then, at the end of the iteration, we have that I is contained in $\{i \in \{1,\dots,m\} \mid s_i < \max(\mathbf{s}) - \delta\}$ which has cardinality at most $|\max(\mathbf{s}) - \mathbf{s}|/\delta$. Thus, we obtain $\mathrm{Card}(I) \le |\max(\mathbf{s}) - \mathbf{s}|/\delta < c/2$.

As a consequence, the worst case in terms of cost occurs when $\mathrm{Card}(I)$ is divided by only slightly more than 2 at each iteration. Then, this cardinality is about $2^{-k}m$ at the end of the kth iteration of the While loop. This iteration then uses $C(\zeta, 2^{-k}m, d)$ operations in $\mathbb{K}$; this follows from the bounds on $\overline{m}$ and $\overline{\mathbf{s}}$ in Lemma 5.2 and from the cost of Step **3.c** given in Proposition 7.3. We remark that the condition $\zeta \le \mathrm{Card}(I)d$ of the loop precisely ensures that we are in the case "$\xi \le md$" of the latter proposition. □

To conclude this section, we derive the upper bound given in the second item of Theorem 1.3 under the assumption $\mathcal{H}_{\mathsf{M}}$. We first remark that we have $\lceil 2^k\zeta/m\rceil \le 2^k\lceil\zeta/m\rceil$, since $\lceil\lceil\alpha r\rceil/\alpha\rceil = \lceil r\rceil$

holds for any real number r and any positive integer α. Besides, the assumption $\mathcal{H}_{\mathsf{M}}$ implies that $\mathsf{M}(2^k\lceil\zeta/m\rceil) \in O(2^{(\omega-1)k}\mathsf{M}(\lceil\zeta/m\rceil))$. Then, the first item in Theorem 1.3 yields

$$C(\zeta, 2^{-k}m, d) \in O\Big((2^{-k}m)^\omega \mathsf{M}(\lceil 2^k\zeta/m\rceil)\log(d)\Big) \subseteq O\Big(2^{-k}m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d)\Big),$$

from which we obtain

$$\sum_{k=0}^{\lfloor \log_2(md/\zeta)\rfloor} C(\zeta, 2^{-k}m, d) \ \in O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d)).$$

Now, using $d \le \frac{m}{\mu}\lceil\frac{\mu}{m}d\rceil \le \frac{m}{\mu}\lceil\frac{\zeta}{m}\rceil$ and the assumption $\mathcal{H}_{\mathsf{M}}$ leads to $\mathsf{M}(d) \in O((m/\mu)^{\omega-1}\mathsf{M}(\lceil\zeta/m\rceil))$, and therefore we also have

$$\mathsf{MM}'(\mu, d) \in O\Big(m^{\omega-1}\mu\mathsf{M}(\lceil\zeta/m\rceil)\log(d)\Big) \subseteq O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d)).$$

This completes the proof of the upper bound in the second item of Theorem 1.3, since we have $\mathsf{MM}'(\mu, \lceil\sigma/\mu\rceil) \in O(\mu^\omega \mathsf{M}(\lceil\sigma/\mu\rceil)\log(\lceil\sigma/\mu\rceil))$.

One can simplify the latter bound slightly further, in order to facilitate the comparison with (Zhou and Labahn, 2012, Thm. 6.14). Indeed, we have $\lceil\frac{\sigma}{\mu}\rceil \in O(\frac{m}{\mu}\lceil\frac{\sigma}{m}\rceil)$ since $m \ge \mu$. Then, using

$$\mathsf{M}(\lceil\sigma/\mu\rceil)\log(\lceil\sigma/\mu\rceil) \in O((m/\mu)^{\omega-1}\mathsf{M}(\lceil\sigma/m\rceil)\log(\lceil\sigma/m\rceil)),$$

which is a minor strengthening of the assumption $\mathcal{H}_{\mathsf{M}}$, the last bound in Theorem 1.3 becomes:

$$\begin{aligned} &O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d) + \mu^\omega \mathsf{M}(\lceil\sigma/\mu\rceil)\log(\lceil\sigma/\mu\rceil)) \\ \subseteq\ &O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d) + m^\omega \mathsf{M}(\lceil\sigma/m\rceil)\log(\lceil\sigma/m\rceil)) \\ \subseteq\ &O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d\lceil\sigma/m\rceil)). \end{aligned}$$

Finally, we remark that if $n \le m$, then we have $\sigma \le md$ and therefore this upper bound becomes $O(m^\omega \mathsf{M}(\lceil\zeta/m\rceil)\log(d))$. This matches the bound in (Zhou and Labahn, 2012, Thm. 6.14), where $n \le m$ is assumed. We further note that in the specific case considered in this reference (the order $\mathbf{d}$ is uniform and $n \le m$), the algorithm stops as soon as the row and column dimensions become roughly equal, and therefore it does not need to rely on column dimension reduction; thus, in this case, the term $\mathsf{MM}'(\mu, \lceil\sigma/\mu\rceil)$ can be removed from the above cost bounds.

Acknowledgement

The authors want to thank Éric Schost for his useful comments. The research leading to these results was partly done while Vincent Neiger was affiliated with the Department of Applied Mathematics and Computer Science of the Technical University of Denmark, with funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 609405 (COFUNDPostdocDTU).

References

Beckermann, B., 1992. A reliable method for computing M-Padé approximants on arbitrary staircases. J. Comput. Appl. Math. 40 (1), 19–42.
URL https://doi.org/10.1016/0377-0427(92)90039-Z

Beckermann, B., Labahn, G., 1994. A uniform approach for the fast computation of matrix-type Padé approximants. SIAM J. Matrix Anal. Appl. 15 (3), 804–823.
URL https://doi.org/10.1137/S0895479892230031
Beckermann, B., Labahn, G., 1997. Recursiveness in matrix rational interpolation problems. J. Comput. Appl. Math. 77, 5–34.
URL https://doi.org/10.1016/S0377-0427(96)00120-3
Beckermann, B., Labahn, G., 2000. Fraction-free computation of matrix rational interpolants and matrix gcds. SIAM J. Matrix Anal. Appl. 22 (1), 114–144.
URL https://doi.org/10.1137/S0895479897326912
Beckermann, B., Labahn, G., Villard, G., 1999. Shifted normal forms of polynomial matrices. In: ISSAC'99. ACM, pp. 189–196.
URL https://doi.org/10.1145/309831.309929
Bostan, A., Schost, É., 2005. Polynomial evaluation and interpolation on special sets of points. J. Complexity 21 (4), 420–446.
URL https://doi.org/10.1016/j.jco.2004.09.009
Cantor, D. G., Kaltofen, E., 1991. On fast multiplication of polynomials over arbitrary algebras. Acta Inform. 28 (7), 693–701.
URL https://doi.org/10.1007/BF01178683
Coppersmith, D., Winograd, S., 1990. Matrix multiplication via arithmetic progressions. J. Symbolic Comput. 9 (3), 251–280.
URL https://doi.org/10.1016/S0747-7171(08)80013-2
Dummit, D. S., Foote, R. M., 2004. Abstract Algebra. John Wiley & Sons.
Forney, Jr., G. D., 1975. Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems. SIAM Journal on Control 13 (3), 493–520.
URL https://doi.org/10.1137/0313029
Giorgi, P., Jeannerod, C.-P., Villard, G., 2003. On the complexity of polynomial matrix computations. In: ISSAC'03. ACM, pp. 135–142.
URL https://doi.org/10.1145/860854.860889
Gupta, S., Storjohann, A., 2011. Computing Hermite forms of polynomial matrices. In: ISSAC'11. ACM, pp. 155–162.
URL https://doi.org/10.1145/1993886.1993913
Harvey, D., van der Hoeven, J., Lecerf, G., 2017. Faster polynomial multiplication over finite fields. J. ACM 63 (6), 52:1–52:23.
URL http://doi.acm.org/10.1145/3005344
Jeannerod, C.-P., Neiger, V., Schost, E., Villard, G., 2016. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts. In: ISSAC'16. ACM, pp. 295–302.
URL https://doi.org/10.1145/2930889.2930928
Jeannerod, C.-P., Neiger, V., Schost, E., Villard, G., 2017. Computing minimal interpolation bases. J. Symbolic Comput. 83, 272–314.
URL https://doi.org/10.1016/j.jsc.2016.11.015
Kailath, T., 1980. Linear Systems. Prentice-Hall.
Knuth, D. E., 1970. The analysis of algorithms. In: Congrès int. Math., Nice, France. Vol. 3. pp. 269–274.
URL http://www.mathunion.org/ICM/ICM1970.3/Main/icm1970.3.0269.0274.ocr.pdf
Le Gall, F., 2014. Powers of tensors and fast matrix multiplication. In: ISSAC'14. ACM, pp. 296–303.
URL https://doi.org/10.1145/2608628.2608664
Moenck, R. T., 1973. Fast computation of GCDs. In: Proc. 5th ACM Symp. Theory Comp. pp. 142–151.
URL https://doi.org/10.1145/800125.804045
Mulders, T., Storjohann, A., 2003. On lattice reduction for polynomial matrices. J. Symbolic Comput. 35, 377–401.
URL https://doi.org/10.1016/S0747-7171(02)00139-6
Neiger, V., 2016. Fast computation of shifted Popov forms of polynomial matrices via systems of modular polynomial equations. In: ISSAC'16. ACM, pp. 365–372.
URL https://doi.org/10.1145/2930889.2930936
Popov, V. M., 1972. Invariant description of linear, time-invariant controllable systems. SIAM Journal on Control 10 (2), 252–264.
URL https://doi.org/10.1137/0310020
Rosenkilde, J., Storjohann, A., 2016. Algorithms for simultaneous Padé approximations. In: ISSAC'16. ACM, New York, NY, USA, pp. 405–412.
URL https://doi.org/10.1145/2930889.2930933
Sarkar, S., Storjohann, A., 2011. Normalization of row reduced matrices. In: ISSAC'11. ACM, pp. 297–304.
URL https://doi.org/10.1145/1993886.1993931

Schönhage, A., 1971. Schnelle Berechnung von Kettenbruchentwicklungen. Acta Inform. 1, 139–144, in German.
URL https://doi.org/10.1007/BF00289520

Storjohann, A., 2000. Algorithms for matrix canonical forms. Ph.D. thesis, Swiss Federal Institute of Technology – ETH.
URL https://doi.org/10.3929/ethz-a-004141007

Storjohann, A., 2003. High-order lifting and integrality certification. J. Symbolic Comput. 36 (3-4), 613–648.
URL https://doi.org/10.1016/S0747-7171(03)00097-X

Storjohann, A., 2006. Notes on computing minimal approximant bases. In: Challenges in Symbolic Computation Software. Dagstuhl Seminar Proceedings. pp. 1–6.
URL http://drops.dagstuhl.de/opus/volltexte/2006/776

Van Barel, M., Bultheel, A., 1991. The computation of non-perfect Padé-Hermite approximants. Numer. Algorithms 1 (3), 285–304.
URL https://doi.org/10.1007/BF02142327

Van Barel, M., Bultheel, A., 1992. A general module theoretic framework for vector M-Padé and matrix rational interpolation. Numer. Algorithms 3, 451–462.
URL https://doi.org/10.1007/BF02141952

Zhou, W., Labahn, G., 2012. Efficient algorithms for order basis computation. J. Symbolic Comput. 47 (7), 793–819.
URL https://doi.org/10.1016/j.jsc.2011.12.009