Fast Computation of Minimal Interpolation Bases in Popov Form for Arbitrary Shifts

Claude-Pierre Jeannerod
Inria, Université de Lyon
Laboratoire LIP (CNRS, Inria, ENSL, UCBL)
claude-pierre.jeannerod@ens-lyon.fr

Vincent Neiger
ENS de Lyon, Université de Lyon
Laboratoire LIP (CNRS, Inria, ENSL, UCBL)
vincent.neiger@ens-lyon.fr

Éric Schost
University of Waterloo
David R. Cheriton School of Computer Science
eschost@uwaterloo.ca

Gilles Villard
CNRS, Université de Lyon
Laboratoire LIP (CNRS, Inria, ENSL, UCBL)
gilles.villard@ens-lyon.fr

ABSTRACT

We compute minimal bases of solutions for a general interpolation problem, which encompasses Hermite-Padé approximation and constrained multivariate interpolation, and has applications in coding theory and security.

This problem asks to find univariate polynomial relations between m vectors of size σ; these relations should have small degree with respect to an input degree shift. For an arbitrary shift, we propose an algorithm for the computation of an interpolation basis in shifted Popov normal form with a cost of $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$ field operations, where ω is the exponent of matrix multiplication and the notation $\mathcal{O}\tilde{}(\cdot)$ indicates that logarithmic terms are omitted.

Earlier works, in the case of Hermite-Padé approximation [34] and in the general interpolation case [18], compute non-normalized bases. Since for arbitrary shifts such bases may have size $\Theta(m^2\sigma)$, the cost bound $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$ was feasible only with restrictive assumptions on the shift that ensure small output sizes. The question of handling arbitrary shifts with the same complexity bound was left open.

To obtain the target cost for any shift, we strengthen the properties of the output bases, and of those obtained during the course of the algorithm: all the bases are computed in shifted Popov form, whose size is always $\mathcal{O}(m\sigma)$. Then, we design a divide-and-conquer scheme. We recursively reduce the initial interpolation problem to sub-problems with more convenient shifts by first computing information on the degrees of the intermediate bases.

Keywords

M-Padé approximation; Hermite-Padé approximation; order basis; polynomial matrix; shifted Popov form.

ISSAC'16, July 19–22, 2016, Waterloo, ON, Canada

ISBN 978-1-4503-4380-0/16/07…$15.00

DOI: http://dx.doi.org/10.1145/2930889.2930928

1. INTRODUCTION

1.1 Problem and main result

We focus on the following interpolation problem from [31, 2]. For a field $\mathbb{K}$ and some positive integer σ, we have as input m vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ in $\mathbb{K}^{1\times\sigma}$, seen as the rows of a matrix $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$. We also have a *multiplication matrix* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ which specifies the multiplication of vectors $\mathbf{e} \in \mathbb{K}^{1\times\sigma}$ by polynomials $p \in \mathbb{K}[X]$ as $p \cdot \mathbf{e} = \mathbf{e}\, p(\mathbf{J})$. Then, we want to find $\mathbb{K}[X]$-linear relations between these vectors, that is, some $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{K}[X]^m$ such that $\mathbf{p} \cdot \mathbf{E} = p_1 \cdot \mathbf{e}_1 + \cdots + p_m \cdot \mathbf{e}_m = 0$. Such a $\mathbf{p}$ is called an *interpolant for* $(\mathbf{E}, \mathbf{J})$.

Hereafter, the matrix $\mathbf{J}$ is in Jordan canonical form: this assumption is satisfied in many interesting applications, as explained below. The notion of interpolant we consider is directly related to the one introduced in [31, 2]. Suppose that $\mathbf{J}$ has n Jordan blocks of dimensions $\sigma_1 \times \sigma_1, \ldots, \sigma_n \times \sigma_n$ and with respective eigenvalues $x_1, \ldots, x_n$; in particular, $\sigma = \sigma_1 + \cdots + \sigma_n$. Then, one may identify $\mathbb{K}^\sigma$ with

$$\mathfrak{F} = \mathbb{K}[X]/(X^{\sigma_1}) \times \cdots \times \mathbb{K}[X]/(X^{\sigma_n}),$$

by mapping any $\mathbf{f} = (f_1, \ldots, f_n)$ in $\mathfrak{F}$ to the vector $\mathbf{e} \in \mathbb{K}^\sigma$ made from the concatenation of the coefficient vectors of $f_1, \ldots, f_n$. Over $\mathfrak{F}$, the $\mathbb{K}[X]$-module structure on $\mathbb{K}^\sigma$ given by $p \cdot \mathbf{e} = \mathbf{e}\, p(\mathbf{J})$ becomes

$$p \cdot \mathbf{f} = (p(X + x_1)f_1 \bmod X^{\sigma_1}, \ldots, p(X + x_n)f_n \bmod X^{\sigma_n}).$$

Now, if $(\mathbf{e}_1, \ldots, \mathbf{e}_m) \in \mathbb{K}^{m\times\sigma}$ is associated to $(\mathbf{f}_1, \ldots, \mathbf{f}_m) \in \mathfrak{F}^m$, with $\mathbf{f}_i = (f_{i,1}, \ldots, f_{i,n})$ and $f_{i,j}$ in $\mathbb{K}[X]/(X^{\sigma_j})$ for all i, j, the relation $p_1 \cdot \mathbf{e}_1 + \cdots + p_m \cdot \mathbf{e}_m = 0$ means that for all j in $\{1, \ldots, n\}$, we have

$$p_1(X + x_j)f_{1,j} + \cdots + p_m(X + x_j)f_{m,j} = 0 \bmod X^{\sigma_j};$$

applying a translation by $-x_j$, this is equivalent to

$$p_1 f_{1,j}(X - x_j) + \cdots + p_m f_{m,j}(X - x_j) = 0 \bmod (X - x_j)^{\sigma_j}.$$

Thus, in terms of vector M-Padé approximation as in [31, 2], $(p_1, \ldots, p_m)$ is an interpolant for $(\mathbf{f}_1, \ldots, \mathbf{f}_m)$, $x_1, \ldots, x_n$, and $\sigma_1, \ldots, \sigma_n$.

The set of all interpolants for $(\mathbf{E}, \mathbf{J})$ is a free $\mathbb{K}[X]$-module of rank m. We are interested in computing a basis of this module, represented as a matrix in $\mathbb{K}[X]^{m\times m}$ and called an *interpolation basis for* $(\mathbf{E}, \mathbf{J})$. Its rows are interpolants for $(\mathbf{E}, \mathbf{J})$, and any interpolant for $(\mathbf{E}, \mathbf{J})$ can be written as a unique $\mathbb{K}[X]$-linear combination of its rows.

Besides, we look for interpolants that have some type of minimal degree. Following [31, 34], for a nonzero $\mathbf{p} = [p_1, \ldots, p_m] \in \mathbb{K}[X]^{1\times m}$ and a *shift* $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{Z}^m$, we define the $\mathbf{s}$-*degree* of $\mathbf{p}$ as $\max_{1\leqslant j\leqslant m}(\deg(p_j)+s_j)$. Up to a change of sign, this notion of $\mathbf{s}$-degree is equivalent to the one in [3] and to the notion of *defect* from [1, Definition 3.1].

Then, the $\mathbf{s}$-*row degree* of a matrix $\mathbf{P} \in \mathbb{K}[X]^{k\times m}$ of rank k is the tuple $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) = (d_1, \ldots, d_k) \in \mathbb{Z}^k$ with d_i the $\mathbf{s}$-degree of the i-th row of $\mathbf{P}$. The $\mathbf{s}$-*leading matrix* of $\mathbf{P} = [p_{ij}]_{i,j}$ is the matrix in $\mathbb{K}^{k\times m}$ whose entry (i, j) is the coefficient of degree $d_i - s_j$ of p_{ij}. Then, $\mathbf{P}$ is $\mathbf{s}$-*reduced* if its $\mathbf{s}$-leading matrix has rank k; see [3].

Our aim is to compute an $\mathbf{s}$-*minimal* interpolation basis for $(\mathbf{E}, \mathbf{J})$, that is, one which is $\mathbf{s}$-reduced: equivalently, it is an interpolation basis whose $\mathbf{s}$-row degree, once written in nondecreasing order, is lexicographically minimal. This corresponds to Problem 1 below. In particular, an interpolant of minimal degree can be read off from an $\mathbf{s}$-minimal interpolation basis for the *uniform* shift $\mathbf{s} = \mathbf{0}$.

Problem 1 (Minimal interpolation basis).
Input:

- *the base field* $\mathbb{K}$,
- *the dimensions* m *and* σ,
- *a matrix* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$,
- *a Jordan matrix* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$,
- *a shift* $\mathbf{s} \in \mathbb{Z}^m$.

Output: *an* $\mathbf{s}$-*minimal interpolation basis for* $(\mathbf{E}, \mathbf{J})$.

A well-known particular case of this problem is Hermite-Padé approximation, that is, the computation of *order bases* (or σ-bases, or minimal approximant bases), where $\mathbf{J}$ has only eigenvalue 0. Previous work on this case includes [1, 14, 30, 34] with algorithms focusing on $\mathbf{J}$ with n blocks of identical size σ/n. For a shift $\mathbf{s} \in \mathbb{N}^m$ with nonnegative entries, we write $|\mathbf{s}|$ for the sum of its entries. Then, in this context, the cost bound $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$ has been obtained under each of the following assumptions:

(H_1) $\max(\mathbf{s}) - \min(\mathbf{s}) \in \mathcal{O}(\sigma/m)$ in [34, Theorem 5.3] and more generally $|\mathbf{s} - \min(\mathbf{s})| \in \mathcal{O}(\sigma)$ in [33, Section 4.1];

(H_2) $|\max(\mathbf{s}) - \mathbf{s}| \in \mathcal{O}(\sigma)$ in [34, Theorem 6.14].

These assumptions imply in particular that any $\mathbf{s}$-minimal basis has size in $\mathcal{O}(m\sigma)$, where by *size* we mean the number of field elements used to represent the matrix.

An interesting example of a shift not covered by (H_1) or (H_2) is $\mathbf{h} = (0, \sigma, 2\sigma, \ldots, (m-1)\sigma)$ which is related to the Hermite form [3, Lemma 2.6]. In general, as detailed in Appendix A, one may assume without loss of generality that $\min(\mathbf{s}) = 0$, $\max(\mathbf{s}) \in \mathcal{O}(m\sigma)$, and $|\mathbf{s}| \in \mathcal{O}(m^2\sigma)$.

There are also applications of Problem 1 to multivariate interpolation, where $\mathbf{J}$ is not nilpotent anymore, and for which we have neither (H_1) nor (H_2), as we will see in Subsection 1.3. It was left as an open problem in [34, Section 7] to obtain algorithms with cost bound $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$ for such matrices $\mathbf{J}$ and for arbitrary shifts. In this paper, we solve this open problem.

An immediate challenge is that for an arbitrary shift $\mathbf{s}$, the size of an $\mathbf{s}$-minimal interpolation basis may be beyond our target cost: we show this in Appendix B with an example of Hermite-Padé approximation. Our answer is to compute a basis in $\mathbf{s}$-*Popov form*: among its many interesting features, it can be represented using at most $m(\sigma+1)$ elements from $\mathbb{K}$, and it is canonical: for every nonsingular $\mathbf{A} \in \mathbb{K}[X]^{m\times m}$ and $\mathbf{s} \in \mathbb{Z}^m$, there is a unique matrix $\mathbf{P}$ in $\mathbf{s}$-Popov form which is left-unimodularly equivalent to $\mathbf{A}$. We use the definition from [2, Section 7], phrased using the notion of pivot [19, Section 6.7.2].

Definition 1.1 (Pivot of a row). *Let* $\mathbf{p} = [p_j]_j \in \mathbb{K}[X]^{1\times m}$ *be a nonzero row vector and let* $\mathbf{s} \in \mathbb{Z}^m$. *The* $\mathbf{s}$-pivot index *of* $\mathbf{p}$ *is the largest index* $j \in \{1, \ldots, m\}$ *such that* $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{p}) = \deg(p_j) + s_j$*; then,* p_j *and* $\deg(p_j)$ *are called the* $\mathbf{s}$-pivot entry *and the* $\mathbf{s}$-pivot degree *of* $\mathbf{p}$.

Definition 1.2 (Popov form). *Let* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *be nonsingular and let* $\mathbf{s} \in \mathbb{Z}^m$. *Then,* $\mathbf{P}$ *is said to be in* $\mathbf{s}$-Popov form *if its* $\mathbf{s}$*-pivot entries are monic and on its diagonal, and if in each column of* $\mathbf{P}$ *the nonpivot entries have degree less than the pivot entry.*

We call $\mathbf{s}$-*Popov interpolation basis for* $(\mathbf{E}, \mathbf{J})$ the unique interpolation basis for $(\mathbf{E}, \mathbf{J})$ which is in $\mathbf{s}$-Popov form; in particular, it is an $\mathbf{s}$-minimal one. For small values of σ, namely $\sigma \in \mathcal{O}(m)$, we gave in [18, Section 7] an algorithm which computes the $\mathbf{s}$-Popov interpolation basis in $\mathcal{O}\tilde{}(\sigma^{\omega-1}m)$ operations for an arbitrary $\mathbf{s}$ [18, Theorem 1.4]. Hence, in what follows, we focus on the case $m \in \mathcal{O}(\sigma)$.

We use the convenient assumption that $\mathbf{J}$ is given to us as a list of eigenvalues and block sizes:

$$\mathbf{J} = ((x_1, \sigma_{1,1}), \ldots, (x_1, \sigma_{1,r_1}), \ldots, (x_t, \sigma_{t,1}), \ldots, (x_t, \sigma_{t,r_t})),$$

for some pairwise distinct eigenvalues $x_1, \ldots, x_t$, with $r_1 \geqslant \cdots \geqslant r_t$ and $\sigma_{i,1} \geqslant \cdots \geqslant \sigma_{i,r_i}$ for all i; we say that this representation is *standard*.

Theorem 1.3. *Assuming that* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ *is a Jordan matrix given by a standard representation, there is a deterministic algorithm which solves Problem 1 using*

$$\begin{aligned} &\mathcal{O}(m^{\omega-1}\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)^2) && \textit{if } \omega > 2, \\ &\mathcal{O}(m\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)^2\log(m)^3) && \textit{if } \omega = 2 \end{aligned}$$

operations in $\mathbb{K}$ *and returns the* $\mathbf{s}$*-Popov interpolation basis for* $(\mathbf{E}, \mathbf{J})$.

In this result, $\mathsf{M}(\cdot)$ is such that polynomials of degree at most d in $\mathbb{K}[X]$ can be multiplied using $\mathsf{M}(d)$ operations in $\mathbb{K}$, and $\mathsf{M}(\cdot)$ satisfies the super-linearity properties of [13, Chapter 8]. It follows from [8] that $\mathsf{M}(d)$ can be taken in $\mathcal{O}(d\log(d)\log(\log(d)))$. The exponent ω is so that we can multiply $m \times m$ matrices in $\mathcal{O}(m^\omega)$ ring operations on any ring, the best known bound being $\omega < 2.38$ [11, 22].

Compared to our work in [18], our algorithm here has two key new features:

- it supports arbitrary shifts with a cost $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$;
- it computes the basis in $\mathbf{s}$-Popov form.

To the best of our knowledge, no algorithm for Problem 1 with cost $\mathcal{O}\tilde{}(m^{\omega-1}\sigma)$ was known previously for *arbitrary* shifts, even for the specific case of order basis computation.

If $\mathbf{J}$ is given as an arbitrary list $((x_1, \sigma_1), \ldots, (x_n, \sigma_n))$, we can reorder it (and permute the columns of $\mathbf{E}$ accordingly) to obtain an equivalent standard representation in time $\mathcal{O}(\mathsf{M}(\sigma)\log(\sigma)^3)$ [5, Proposition 12]; if $\mathbb{K}$ is equipped with an order, and if we assume that comparisons take unit time, this can of course be done in time $\mathcal{O}(\sigma\log(\sigma))$.

1.2 Overview of our approach

Several previous algorithms for order basis computation, such as those in [1, 14], follow a divide-and-conquer scheme inspired by the Knuth-Schönhage-Moenck algorithm [20, 29, 23] This paper builds on our previous work in [18], where we extended this recursive approach to more general interpolation problems. However, the main algorithm in [18] does not handle an arbitrary shift $\mathbf{s}$ with a satisfactory complexity; here, we use it as a black box, after showing how to reduce the problem to a new one with suitable shift.

Let $\mathbf{E}$, $\mathbf{J}$, and $\mathbf{s}$ be our input, and write $\mathbf{J}^{(1)}$ and $\mathbf{J}^{(2)}$ for the $\sigma/2\times\sigma/2$ leading and trailing principal submatrices of $\mathbf{J}$. First, compute an $\mathbf{s}$-minimal interpolation basis $\mathbf{P}^{(1)}$ for $\mathbf{J}^{(1)}$ and the first $\sigma/2$ columns of $\mathbf{E}$; then, compute the last $\sigma/2$ columns $\mathbf{E}^{(2)}$ of the *residual* $\mathbf{P}^{(1)}\cdot\mathbf{E}$; then, compute a $\mathbf{t}$-minimal interpolation basis $\mathbf{P}^{(2)}$ for $(\mathbf{E}^{(2)},\mathbf{J}^{(2)})$ with $\mathbf{t}=\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}^{(1)})$; finally, return the matrix product $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$.

This approach allows to solve Problem 1 using $\mathcal{O}\tilde{~}(m^\omega\sigma)$ operations in $\mathbb{K}$. In the case of Hermite-Padé approximation, this is the divide-and-conquer algorithm in [1]. Besides, an $\mathbf{s}$-minimal basis computed by this method has degree at most σ and thus size in $\mathcal{O}(m^2\sigma)$, and there are indeed instances of Problem 1 for which this size reaches $\Theta(m^2\sigma)$. In Appendix B, we show such an instance for the algorithm in [1], in the case of Hermite-Padé approximation.

It is known that the average degree of the rows of any $\mathbf{s}$-minimal interpolation basis is at most $(\sigma+\xi)/m$, where $\xi=|\mathbf{s}-\min(\mathbf{s})|$ [31, Theorem 4.1]. In [18], focusing on the case where ξ is small compared to σ, and preserving such a property in recursive calls via changes of shifts, we obtained the cost bound

$$\mathcal{O}(m^{\omega-1}\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)+m^{\omega-1}\mathsf{M}(\xi)\log(\xi/m)) \quad (1)$$

to solve Problem 1; this cost is for $\omega>2$, and a similar one holds for $\omega=2$, both being in $\mathcal{O}\tilde{~}(m^{\omega-1}(\sigma+\xi))$. The fundamental reason for this kind of improvement over $\mathcal{O}\tilde{~}(m^\omega\sigma)$, already seen with [34], is that one controls the average row degree of the bases $\mathbf{P}^{(2)}$ and $\mathbf{P}^{(1)}$, *and* of their product $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$.

This result is $\mathcal{O}\tilde{~}(m^{\omega-1}\sigma)$ for ξ in $\mathcal{O}(\sigma)$. The main difficulty to extend it to any shift $\mathbf{s}$ is to control the size of the computed bases: the Hermite-Padé example pointed out above corresponds to $\xi=\Theta(m\sigma)$ and leads to an output of size $\Theta(m^2\sigma)$ for the algorithm of [18] as well.

The key ingredient to control this size is to work with bases in $\mathbf{s}$-Popov form: for any $\mathbf{s}$, the $\mathbf{s}$-Popov interpolation basis $\mathbf{P}$ for $(\mathbf{E},\mathbf{J})$ has average *column* degree at most σ/m and size at most $m(\sigma+1)$, as detailed in Section 2.

Now, suppose that we have computed recursively the bases $\mathbf{P}^{(2)}$ and $\mathbf{P}^{(1)}$ in $\mathbf{s}$- and $\mathbf{t}$-Popov form; we want to output the $\mathbf{s}$-Popov form $\mathbf{P}$ of $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$. In general, this product is not normalized and may have size $\Theta(m^2\sigma)$: its computation is beyond our target cost. Thus, one main idea is that we will *not* rely on polynomial matrix multiplication to combine the bases obtained recursively; instead, we use a minimal interpolation basis computation for a shift that has good properties as explained below.

An important remark is that if we know *a priori* the column degree $\boldsymbol{\delta}$ of $\mathbf{P}$, then the problem becomes easier. This idea was already used in algorithms for the Hermite form $\mathbf{H}$ of a polynomial matrix [15, 33], which first compute the column degree $\boldsymbol{\delta}$ of $\mathbf{H}$, and then obtain $\mathbf{H}$ as a submatrix of some minimal nullspace basis for a shift involving $-\boldsymbol{\delta}$.

In Section 4, we study the problem of computing the $\mathbf{s}$-Popov interpolation basis $\mathbf{P}$ for $(\mathbf{E},\mathbf{J})$ having its column degree $\boldsymbol{\delta}$ as an additional input. We show that this reduces to the computation of a $\mathbf{d}$-minimal interpolation basis $\mathbf{R}$ with the specific shift $\mathbf{d}=-\boldsymbol{\delta}$. The properties of this shift $\mathbf{d}$ allow us first to compute $\mathbf{R}$ in $\mathcal{O}\tilde{~}(m^{\omega-1}\sigma)$ operations using the partial linearization framework from [30, Section 3] and the minimal interpolation basis algorithm in [18, Section 3], and second to easily retrieve $\mathbf{P}$ from $\mathbf{R}$.

Still, in general we do not know $\boldsymbol{\delta}$. We will thus compute it, relying on a variation of the divide-and-conquer strategy at the beginning of this subsection. We stop the recursion as soon as $\sigma\leqslant m$, in which case we do not need $\boldsymbol{\delta}$ to achieve efficiency: the algorithm from [18, Section 7] computes the $\mathbf{s}$-Popov interpolation basis in $\mathcal{O}\tilde{~}(\sigma^{\omega-1}m)$ operations for any $\mathbf{s}$ [18, Theorem 1.4]. Then, we show in Section 3 that from $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ computed recursively *in shifted Popov form*, we can obtain $\boldsymbol{\delta}$ for free. Finally, instead of considering $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$, we use the knowledge of $\boldsymbol{\delta}$ to compute the basis $\mathbf{P}$ from scratch as explained in the previous paragraph.

This summarizes our main algorithm, which is presented in Section 2.

1.3 Previous work and applications

As a particular case of Problem 1, when all the eigenvalues of $\mathbf{J}$ are zero, we obtain the following complexity result about *order basis* computation [34, Definition 2.2].

THEOREM 1.4. *Let $m,n\in\mathbb{Z}_{>0}$, let $(\sigma_1,\ldots,\sigma_n)\in\mathbb{Z}^n_{>0}$, let $\mathbf{s}\in\mathbb{Z}^m$, and let $\mathbf{F}\in\mathbb{K}[X]^{m\times n}$ with its j-th column $\mathbf{F}_{*,j}$ of degree less than σ_j. The unique basis $\mathbf{P}\in\mathbb{K}[X]^{m\times m}$ in $\mathbf{s}$-Popov form of the $\mathbb{K}[X]$-module of approximants*

$$\{\mathbf{p}\in\mathbb{K}[X]^{1\times m} \mid \mathbf{p}\mathbf{F}_{*,j}=0 \bmod X^{\sigma_j} \text{ for each } j\}$$

can be computed deterministically using

$$\mathcal{O}(m^{\omega-1}\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)^2) \qquad \text{if } \omega>2,$$
$$\mathcal{O}(m\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)^2\log(m)^3) \qquad \text{if } \omega=2$$

operations in $\mathbb{K}$, where $\sigma=\sigma_1+\cdots+\sigma_n$.

Previous work on this problem includes [1, 14, 30, 34, 18], mostly with identical orders $\sigma_1=\cdots=\sigma_n$; an interesting particular case is Hermite-Padé approximation with $n=1$. To simplify matters, for all our comparisons, we consider $\omega>2$. For order basis computation with $\sigma_1=\cdots=\sigma_n$ and $n\leqslant m$, the cost bound $\mathcal{O}(m^\omega\mathsf{M}(\sigma/m)\log(\sigma/n))$ was achieved in [34] under either of the assumptions (H_1) and (H_2) on the shift. Still, the corresponding algorithm returns a basis $\mathbf{P}$ which is only $\mathbf{s}$-reduced, and because both the shift $\mathbf{s}$ and the degrees in $\mathbf{P}$ may be unbalanced, one cannot directly rely on the fastest known normalization algorithm [28] to compute the $\mathbf{s}$-Popov form of $\mathbf{P}$ within the target cost.

Another application of Problem 1 is a multivariate interpolation problem that arises for example in the first step of algorithms for the list-decoding of Parvaresh-Vardy codes [26] and of folded Reed-Solomon codes [16], as well as in robust Private Information Retrieval [12]. The bivariate case corresponds to the interpolation steps of Kötter and Vardy's soft-decoding [21] and Guruswami and Sudan's list-decoding [17] algorithms for Reed-Solomon codes.

Given a set of points in $\mathbb{K}^{r+1}$ and associated multiplicities, this problem asks to find a multivariate polynomial $Q(X, Y_1, \ldots, Y_r)$ such that: (a) Q has prescribed exponents for the Y variables, so that the problem can be linearized with respect to Y, leaving us with a linear algebra problem over $\mathbb{K}[X]$; (b) Q vanishes at all the given points with their multiplicities, inducing a structure of $\mathbb{K}[X]$-module on the set of solutions; (c) Q has some type of minimal weighted degree, which can be seen as the minimality of the shifted degree of the vector over $\mathbb{K}[X]$ that represents Q.

Following the coding theory context [17, 26], given a point $(x, y) \in \mathbb{K} \times \mathbb{K}^r$ and a set of exponents $\mu \subset \mathbb{N}^{r+1}$, we say that the polynomial $Q(X, Y) \in \mathbb{K}[X, Y_1, \ldots, Y_r]$ *vanishes at* (x, y) *with multiplicity support* μ if the shifted polynomial $Q(X + x, Y + y)$ has no monomial with exponent in μ. We will only consider supports that are *stable under division*, meaning that if $(\gamma_0, \gamma_1, \ldots, \gamma_r)$ is in μ, then any $(\gamma_0', \gamma_1', \ldots, \gamma_r')$ with $\gamma_j' \leqslant \gamma_j$ for all j is also in μ.

Now, given a set of exponents $\Gamma \subset \mathbb{N}^r$, we represent $Q(X, Y) = \sum_{\gamma \in \Gamma} p_\gamma Y^\gamma$ as the row $\mathbf{p} = [p_\gamma]_{\gamma \in \Gamma} \in \mathbb{K}[X]^{1 \times m}$ where m is the cardinality of Γ. Again, we assume that the exponent set Γ is stable under division; then, the set of solutions is a free $\mathbb{K}[X]$-module of rank m. In the mentioned applications, we typically have $\Gamma = \{(\gamma_1, \ldots, \gamma_r) \in \mathbb{N}^r \mid \gamma_1 + \cdots + \gamma_r \leqslant \ell\}$ for an integer ℓ called the *list-size parameter*.

Besides, we are given some weights $\mathbf{w} = (w_1, \ldots, w_r) \in \mathbb{N}^r$ on the variables $Y = Y_1, \ldots, Y_r$, and we are looking for $Q(X, Y)$ which has minimal $\mathbf{w}$-weighted degree, which is the degree in X of the polynomial

$$Q(X, X^{w_1} Y_1, \ldots, X^{w_r} Y_r) = \sum_{\gamma \in \Gamma} p_\gamma X^{\gamma_1 w_1 + \cdots + \gamma_r w_r} Y_1^{\gamma_1} \cdots Y_r^{\gamma_r}.$$

This is exactly requiring that the $\mathbf{s}$-degree of $\mathbf{p} = [p_\gamma]_\gamma$ be minimal, for $\mathbf{s} = [\gamma_1 w_1 + \cdots + \gamma_r w_r]_\gamma$. We note that it is sometimes important, for example in [12], to return a whole $\mathbf{s}$-minimal interpolation basis and not only one interpolant of small $\mathbf{s}$-degree.

> PROBLEM 2 (MULTIVARIATE INTERPOLATION).
> Input:
>
> - *number of Y variables $r > 0$,*
> - *set $\Gamma \subset \mathbb{N}^r$ of cardinality m, stable under division,*
> - *pairwise distinct points $\{(x_k, y_k) \in \mathbb{K} \times \mathbb{K}^r\}_{1 \leqslant k \leqslant p}$,*
> - *supports $\{\mu_k \subset \mathbb{N}^{r+1}\}_{1 \leqslant k \leqslant p}$, stable under division,*
> - *a shift $\mathbf{s} \in \mathbb{Z}^m$.*
>
> Output: *a matrix $\mathbf{P} \in \mathbb{K}[X]^{m \times m}$ such that*
>
> - *the rows of $\mathbf{P}$ form a basis of the $\mathbb{K}[X]$-module*
> $$\left\{ \mathbf{p} = [p_\gamma]_{\gamma \in \Gamma} \in \mathbb{K}[X]^{1 \times m} \;\middle|\; \sum_{\gamma \in \Gamma} p_\gamma(X) Y^\gamma \text{ vanishes at } (x_k, y_k) \text{ with support } \mu_k \text{ for } 1 \leqslant k \leqslant p \right\},$$
> - *$\mathbf{P}$ is $\mathbf{s}$-reduced.*

For more details about the reduction from Problem 2 to Problem 1, explaining how to build the input matrices $(\mathbf{E}, \mathbf{J})$ with $\mathbf{J}$ a Jordan matrix in standard representation, we refer the reader to [18, Subsection 2.4]. In particular, the dimension σ is the sum of the cardinalities of the multiplicity supports. In the mentioned applications to coding theory, we have $m = \binom{r+\ell}{r}$ where ℓ is the list-size parameter; and σ is the so-called *cost* in the soft-decoding context [21, Section III], that is, the number of linear equations when linearizing the problem over $\mathbb{K}$. As a consequence of Theorem 1.3, we obtain the following complexity result.

THEOREM 1.5. *Let $\sigma = \sum_{1 \leqslant k \leqslant p} \# \mu_k$. There is a deterministic algorithm which solves Problem 2 using*

$$\mathcal{O}(m^{\omega-1} \mathsf{M}(\sigma) \log(\sigma) \log(\sigma/m)^2) \qquad \text{if } \omega > 2,$$
$$\mathcal{O}(m \mathsf{M}(\sigma) \log(\sigma) \log(\sigma/m)^2 \log(m)^3) \qquad \text{if } \omega = 2$$

operations in $\mathbb{K}$, and returns the unique basis of solutions which is in $\mathbf{s}$-Popov form.

Under the assumption that the x_k are pairwise distinct, the cost bound $\mathcal{O}(m^{\omega-1} \mathsf{M}(\sigma) \log(\sigma)^2)$ was achieved for an arbitrary shift using fast structured linear algebra [9, Theorems 1 and 2], following work by [25, 27, 32]. However, the corresponding algorithm is randomized and returns only one interpolant of small $\mathbf{s}$-degree. For a broader overview of previous work on this problem, we refer the reader to the introductive sections of [4, 9] and to [18, Section 2].

The term $\mathcal{O}(m^{\omega-1} \mathsf{M}(\xi) \log(\xi/m))$ reported in (1) for the cost of the algorithm of [18] can be neglected if $\xi \in \mathcal{O}(\sigma)$; this is for instance satisfied in the context of bivariate interpolation for soft- or list-decoding of Reed-Solomon codes [18, Sections 2.5 and 2.6]. However, we do not have this bound on ξ in the list-decoding of Parvaresh-Vardy codes and folded Reed-Solomon codes and in Private Information Retrieval. Thus, in these cases our algorithm achieves the best known cost bound, improving upon [7, 6, 10, 12, 18].

2. FAST POPOV INTERPOLATION BASIS

In this section, we present our main result, Algorithm 1. It relies on three subroutines; two of them are from [18], while the third is a key new ingredient, detailed in Section 4.

- LINEARIZATIONMIB [18, Algorithm 9] solves the base case $\sigma \leqslant m$ using linear algebra over $\mathbb{K}$. The inputs are $\mathbf{E}$, $\mathbf{J}$, $\mathbf{s}$, as well as an integer for which we can take the first power of two greater than or equal to σ.
- COMPUTERESIDUALS [18, Algorithm 5] (with an additional pre-processing detailed at the end of Section 4) computes the residual $\mathbf{P}^{(1)} \cdot \mathbf{E}$ from the first basis $\mathbf{P}^{(1)}$ obtained recursively.
- KNOWNMINDEGMIB, detailed in Section 4, computes the $\mathbf{s}$-Popov interpolation basis when one knows *a priori* the $\mathbf{s}$-minimal degree of $(\mathbf{E}, \mathbf{J})$ (see below).

In what follows, by *$\mathbf{s}$-minimal degree of* $(\mathbf{E}, \mathbf{J})$ we mean the tuple of degrees of the diagonal entries of the $\mathbf{s}$-Popov interpolation basis $\mathbf{P}$ for $(\mathbf{E}, \mathbf{J})$. Because $\mathbf{P}$ is in $\mathbf{s}$-Popov form, this is also the column degree of $\mathbf{P}$, and the sum of these degrees is $\deg(\det(\mathbf{P}))$. As a consequence, using Theorem 4.1 in [31] (or following the lines of [19] and [2]) we obtain the following lemma, which implies in particular that the size of $\mathbf{P}$ is at most $m(\sigma + 1)$.

LEMMA 2.1. *Let* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$, $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$, $\mathbf{s} \in \mathbb{Z}^m$, *and let* $(\delta_1, \ldots, \delta_m)$ *be the* $\mathbf{s}$*-minimal degree of* $(\mathbf{E}, \mathbf{J})$. *Then, we have* $\delta_1 + \cdots + \delta_m \leqslant \sigma$.

ALGORITHM 1. POPOVMIB

Input:

- *a matrix* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$,
- *a Jordan matrix* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ *in standard representation,*
- *a shift* $\mathbf{s} \in \mathbb{Z}^m$.

Output:

- *the* $\mathbf{s}$*-Popov interpolation basis* $\mathbf{P}$ *for* $(\mathbf{E}, \mathbf{J})$,
- *the* $\mathbf{s}$*-minimal degree* $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m)$ *of* $(\mathbf{E}, \mathbf{J})$.

1. *If* $\sigma \leqslant m$, *return* LINEARIZATIONMIB$(\mathbf{E}, \mathbf{J}, \mathbf{s}, 2^{\lceil \log_2(\sigma)\rceil})$

2. *Else*

  a. $\mathbf{E}^{(1)} \leftarrow$ *first* $\lceil \sigma/2 \rceil$ *columns of* $\mathbf{E}$

  b. $(\mathbf{P}^{(1)}, \boldsymbol{\delta}^{(1)}) \leftarrow$ POPOVMIB$(\mathbf{E}^{(1)}, \mathbf{J}^{(1)}, \mathbf{s})$

  c. $\mathbf{E}^{(2)} \leftarrow$ *last* $\lfloor \sigma/2 \rfloor$ *columns of* $\mathbf{P}^{(1)} \cdot \mathbf{E} =$ COMPUTERESIDUALS$(\mathbf{J}, \mathbf{P}^{(1)}, \mathbf{E})$

  d. $(\mathbf{P}^{(2)}, \boldsymbol{\delta}^{(2)}) \leftarrow$ POPOVMIB$(\mathbf{E}^{(2)}, \mathbf{J}^{(2)}, \mathbf{s} + \boldsymbol{\delta}^{(1)})$

  e. $\mathbf{P} \leftarrow$ KNOWNMINDEGMIB$(\mathbf{E}, \mathbf{J}, \mathbf{s}, \boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)})$

  f. *Return* $(\mathbf{P}, \boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)})$

Taking for granted the results in the next sections, we now prove our main theorem.

PROOF OF THEOREM 1.3. For the case $\sigma \leqslant m$, the correctness and the cost bound of Algorithm 1 both follow from [18, Theorem 1.4]: it uses $\mathcal{O}(\sigma^{\omega-1}m + \sigma^\omega \log(\sigma))$ operations (with an extra $\log(\sigma)$ factor if $\omega = 2$).

Now, we consider the case $\sigma > m$. Using the notation in the algorithm, assume that $\mathbf{P}^{(1)}$ is the $\mathbf{s}$-Popov interpolation basis for $(\mathbf{E}^{(1)}, \mathbf{J}^{(1)})$, and $\mathbf{P}^{(2)}$ is the $\mathbf{t}$-Popov interpolation basis for $(\mathbf{E}^{(2)}, \mathbf{J}^{(2)})$, where $\mathbf{t} = \mathbf{s} + \boldsymbol{\delta}^{(1)} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}^{(1)})$, and $\boldsymbol{\delta}^{(1)}$ and $\boldsymbol{\delta}^{(2)}$ are the $\mathbf{s}$- and $\mathbf{s} + \boldsymbol{\delta}^{(1)}$-minimal degrees of $(\mathbf{E}^{(1)}, \mathbf{J}^{(1)})$ and $(\mathbf{E}^{(2)}, \mathbf{J}^{(2)})$, respectively.

We claim that $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$ is $\mathbf{s}$-reduced: this will be proved in Lemma 3.2. Let us then prove that $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$ is an interpolation basis for $(\mathbf{E}, \mathbf{J})$. Let $\mathbf{p} \in \mathbb{K}[X]^{1\times m}$ be an interpolant for $(\mathbf{E}, \mathbf{J})$. Since $\mathbf{J}$ is upper triangular, $\mathbf{p}$ is in particular an interpolant for $(\mathbf{E}^{(1)}, \mathbf{J}^{(1)})$, so there exists $\mathbf{v} \in \mathbb{K}[X]^{1\times m}$ such that $\mathbf{p} = \mathbf{v}\mathbf{P}^{(1)}$. Besides, we have $\mathbf{P}^{(1)} \cdot \mathbf{E} = [0|\mathbf{E}^{(2)}]$, so that $0 = \mathbf{p} \cdot \mathbf{E} = \mathbf{v}\mathbf{P}^{(1)} \cdot \mathbf{E} = [0|\mathbf{v} \cdot \mathbf{E}^{(2)}]$, and thus $\mathbf{v} \cdot \mathbf{E}^{(2)} = 0$. Then, there exists $\mathbf{w} \in \mathbb{K}[X]^{1\times m}$ such that $\mathbf{v} = \mathbf{w}\mathbf{P}^{(2)}$, which gives $\mathbf{p} = \mathbf{w}\mathbf{P}^{(2)}\mathbf{P}^{(1)}$.

In particular, the $\mathbf{s}$-Popov interpolation basis for $(\mathbf{E}, \mathbf{J})$ is the $\mathbf{s}$-Popov form of $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$. Thus, Lemma 3.2 combined with Lemma 3.3 will show that the $\mathbf{s}$-minimal degree of $(\mathbf{E}, \mathbf{J})$ is $\boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)}$. As a result, Proposition 4.3 states that Step **2.e** correctly computes the $\mathbf{s}$-Popov interpolation basis for $(\mathbf{E}, \mathbf{J})$.

Concerning the cost bound, the recursion stops when $\sigma \leqslant m$, and thus the algorithm uses $\mathcal{O}(m^\omega \log(m))$ operations (with an extra $\log(m)$ factor if $\omega = 2$). The depth of the recursion is $\mathcal{O}(\log(\sigma/m))$; we have two recursive calls in dimensions $m \times \sigma/2$, and two calls to subroutines with cost bounds given in Corollary 4.5 and Proposition 4.3, respectively. The conclusion follows from the super-linearity properties of $\mathsf{M}(\cdot)$. $\square$

3. OBTAINING THE MINIMAL DEGREE FROM RECURSIVE CALLS

In this section, we show that the $\mathbf{s}$-minimal degree of $(\mathbf{E}, \mathbf{J})$ can be deduced for free from two bases computed recursively as in Algorithm 1. To do this, we actually prove a slightly more general result about the degrees of the $\mathbf{s}$-pivot entries of so-called weak Popov matrix forms [24].

DEFINITION 3.1 (WEAK POPOV FORM, PIVOT DEGREE). *Let* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *be nonsingular and let* $\mathbf{s} \in \mathbb{Z}^m$. *Then,* $\mathbf{P}$ *is said to be in* $\mathbf{s}$-weak Popov form *if the* $\mathbf{s}$*-pivot indices of its rows are pairwise distinct;* $\mathbf{P}$ *is said to be in* $\mathbf{s}$-diagonal weak Popov form *if its* $\mathbf{s}$*-pivot entries are on its diagonal.*

If $\mathbf{P}$ *is in* $\mathbf{s}$*-weak Popov form, the* $\mathbf{s}$-pivot degree *of* $\mathbf{P}$ *is the tuple* $(\delta_1, \ldots, \delta_m)$ *where for* $j \in \{1, \ldots, m\}$, δ_j *is the* $\mathbf{s}$*-pivot degree of the row of* $\mathbf{P}$ *which has* $\mathbf{s}$*-pivot index* j.

We recall from Section 1 that for $\mathbf{P} \in \mathbb{K}[X]^{k\times m}$, its $\mathbf{s}$-leading matrix $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}) \in \mathbb{K}^{k\times m}$ is formed by the coefficients of degree 0 of $\mathbf{X}^{-\mathbf{d}}\mathbf{P}\mathbf{X}^{\mathbf{s}}$, where $\mathbf{d} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})$ and $\mathbf{X}^{\mathbf{s}}$ stands for the diagonal matrix with entries $X^{s_1}, \ldots, X^{s_m}$. Then, a nonsingular $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ is in $\mathbf{s}$-diagonal weak Popov form with $\mathbf{s}$-pivot degree $\boldsymbol{\delta}$ if and only if $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ is lower triangular and invertible and $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) = \mathbf{s} + \boldsymbol{\delta}$.

For example, at all stages of the algorithms in [31, 1, 18] for Problem 1 (as well as [14] if avoiding row permutations at the base case of the recursion), the computed bases are in shifted diagonal weak Popov form. This is due to the compatibility of this form with matrix multiplication, as stated in the next lemma.

LEMMA 3.2. *Let* $\mathbf{s} \in \mathbb{Z}^m$, $\mathbf{P}^{(1)} \in \mathbb{K}[X]^{m\times m}$ *in* $\mathbf{s}$*-diagonal weak Popov form with* $\mathbf{s}$*-pivot degree* $\boldsymbol{\delta}^{(1)}$, $\mathbf{t} = \mathbf{s} + \boldsymbol{\delta}^{(1)} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}^{(1)})$, *and* $\mathbf{P}^{(2)} \in \mathbb{K}[X]^{m\times m}$ *in* $\mathbf{t}$*-diagonal weak Popov form with* $\mathbf{t}$*-pivot degree* $\boldsymbol{\delta}^{(2)}$. *Then,* $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$ *is in* $\mathbf{s}$*-diagonal weak Popov form with* $\mathbf{s}$*-pivot degree* $\boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)}$.

PROOF. By the predictable-degree property [19, Theorem 6.3-13] we have $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}^{(2)}\mathbf{P}^{(1)}) = \mathrm{rdeg}_{\mathbf{t}}(\mathbf{P}^{(2)}) = \mathbf{t} + \boldsymbol{\delta}^{(2)} = \mathbf{s} + \boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)}$. The result follows since $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}^{(2)}\mathbf{P}^{(1)}) = \mathrm{lm}_{\mathbf{t}}(\mathbf{P}^{(2)})\mathrm{lm}_{\mathbf{s}}(\mathbf{P}^{(1)})$ is lower triangular and invertible. $\square$

For matrices in $\mathbf{s}$-Popov form, the $\mathbf{s}$-pivot degree coincides with the column degree: in particular, the $\mathbf{s}$-minimal degree of $(\mathbf{E}, \mathbf{J})$ is the $\mathbf{s}$-pivot degree of the $\mathbf{s}$-Popov interpolation basis for $(\mathbf{E}, \mathbf{J})$. With the notation of Algorithm 1, the previous lemma proves that the $\mathbf{s}$-pivot degree of $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$ is $\boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)}$. In the rest of this section, we prove that the $\mathbf{s}$-Popov form of $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$ has the same $\mathbf{s}$-pivot degree as $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$. Consequently, the $\mathbf{s}$-minimal degree of $(\mathbf{E}, \mathbf{J})$ is $\boldsymbol{\delta}^{(1)} + \boldsymbol{\delta}^{(2)}$ and thus can be found from $\mathbf{P}^{(2)}$ and $\mathbf{P}^{(1)}$ without computing their product.

It is known that left-unimodularly equivalent $\mathbf{s}$-reduced matrices have the same $\mathbf{s}$-row degree up to permutation [19, Lemma 6.3-14]. Here, we prove that the $\mathbf{s}$-pivot degree is invariant among left-unimodularly equivalent matrices in $\mathbf{s}$-weak Popov form.

Lemma 3.3. *Let* $\mathbf{s} \in \mathbb{Z}^m$ *and let* $\mathbf{P}$ *and* $\mathbf{Q}$ *in* $\mathbb{K}[X]^{m\times m}$ *be two left-unimodularly equivalent nonsingular polynomial matrices in* $\mathbf{s}$*-weak Popov form. Then* $\mathbf{P}$ *and* $\mathbf{Q}$ *have the same* $\mathbf{s}$*-pivot degree.*

Proof. Since row permutations preserve both the $\mathbf{s}$-pivot degrees and left-unimodular equivalence, we can assume that $\mathbf{P}$ and $\mathbf{Q}$ are in $\mathbf{s}$-diagonal weak Popov form. The $\mathbf{s}$-pivot degrees of $\mathbf{P}$ and $\mathbf{Q}$ are then $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) - \mathbf{s}$ and $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{Q}) - \mathbf{s}$, and it remains to check that $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{Q})$.

For any nonsingular $\mathbf{W} \in \mathbb{K}[X]^{m\times m}$ in $\mathbf{s}$-weak Popov form we have $|\mathrm{rdeg}_{\mathbf{s}}(\mathbf{W})| = \deg(\det(\mathbf{W})) + |\mathbf{s}|$ [19, Section 6.3.2]. Thus, if $\mathbf{W}$ is furthermore comprised entirely of rows in the $\mathbb{K}[X]$-row space of $\mathbf{P}$ (that is, $\mathbf{W}$ is a left multiple of $\mathbf{P}$) then we must have $|\mathrm{rdeg}_{\mathbf{s}}(\mathbf{W})| \geqslant |\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})|$.

To arrive at a contradiction, suppose there exists a row index i such that the $\mathbf{s}$-degree of $\mathbf{P}_{i,*}$ differs from that of $\mathbf{Q}_{i,*}$ and without loss of generality assume that the s-degree of $\mathbf{Q}_{i,*}$ is strictly less than that of $\mathbf{P}_{i,*}$. Then the matrix $\mathbf{W}$ obtained from $\mathbf{P}$ by replacing the i-th row of $\mathbf{P}$ with $\mathbf{Q}_{i,*}$ is in $\mathbf{s}$-diagonal weak Popov form. This is a contradiction, since $|\mathrm{rdeg}_{\mathbf{s}}(\mathbf{W})| < |\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})|$ and $\mathbf{Q}_{i,*}$ is in the $\mathbb{K}[X]$-row space of $\mathbf{P}$ for $\mathbf{Q}$ is left-unimodularly equivalent to $\mathbf{P}$. □

In particular, any nonsingular matrix in $\mathbf{s}$-weak Popov form has the same $\mathbf{s}$-pivot degree as its $\mathbf{s}$-Popov form, which proves our point about the $\mathbf{s}$-Popov form of $\mathbf{P}^{(2)}\mathbf{P}^{(1)}$.

4. COMPUTING INTERPOLATION BASES WITH KNOWN MINIMAL DEGREE

In this section, we propose an efficient algorithm for computing the $\mathbf{s}$-Popov interpolation basis $\mathbf{P}$ for $(\mathbf{E},\mathbf{J})$ when the $\mathbf{s}$-minimal degree $\boldsymbol{\delta}$ of $(\mathbf{E},\mathbf{J})$ is known *a priori*.

First, we show that the shift $\mathbf{d} = -\boldsymbol{\delta}$ leads to the same $\mathbf{d}$-Popov interpolation basis $\mathbf{P}$ as the initial shift $\mathbf{s}$. Then, we prove that $\mathbf{P}$ can be easily recovered from any interpolation basis which is simply $\mathbf{d}$-reduced. The following lemma extends [28, Lemmas 15 and 17] to the case of any shift $\mathbf{s}$.

Lemma 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* $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_m)$*. Then* $\mathbf{P}$ *is also in* $\mathbf{d}$*-Popov form for* $\mathbf{d} = (-\delta_1,\ldots,-\delta_m)$*, and we have* $\mathrm{rdeg}_{\mathbf{d}}(\mathbf{P}) = (0,\ldots,0)$*. In particular, for any matrix* $\mathbf{R} \in \mathbb{K}[X]^{m\times m}$ *which is unimodularly equivalent to* $\mathbf{P}$ *and* $\mathbf{d}$*-reduced,* $\mathbf{R}$ *has column degree* $\boldsymbol{\delta}$*, and* $\mathbf{P} = \mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R}$.

Proof. Let us denote $\mathbf{P} = [p_{ij}]_{i,j}$, and let $i \in \{1,\ldots,m\}$. Since $\mathbf{P}$ is in $\mathbf{s}$-Popov form, it is enough to prove that the $\mathbf{d}$-pivot entries of the rows of $\mathbf{P}$ are on its diagonal. We have $\deg(p_{ij}) < \deg(p_{jj}) = \delta_j$ for all $j \neq i$, and $\deg(p_{ii}) = \delta_i$. Then, the i-th row of $\mathbf{P}$ has $\mathbf{d}$-pivot index i and $\mathbf{d}$-degree 0. Thus $\mathbf{P}$ is in $\mathbf{d}$-Popov form with $\mathbf{d}$-row degree $(0,\ldots,0)$.

Now, let $\mathbf{R}$ be a $\mathbf{d}$-reduced matrix left-unimodularly equivalent to $\mathbf{P}$. Then, $\mathrm{rdeg}_{\mathbf{d}}(\mathbf{R}) = \mathrm{rdeg}_{\mathbf{d}}(\mathbf{P}) = (0,\ldots,0)$, so that we can write $\mathbf{R} = \mathrm{lm}_{\mathbf{d}}(\mathbf{R})\mathbf{X}^{\boldsymbol{\delta}} + \mathbf{Q}$ with the j-th column of $\mathbf{Q}$ of degree less than δ_j. In particular, since $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})$ is invertible, the column degree of $\mathbf{R}$ is $\boldsymbol{\delta}$. Besides, we obtain $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R} = \mathbf{X}^{\boldsymbol{\delta}} + \mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{Q}$, and the j-th column of $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{Q}$ has degree less than δ_j. Thus $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R}$ is in $\mathbf{d}$-Popov form and unimodularly equivalent to $\mathbf{P}$, hence equal to $\mathbf{P}$. □

In particular, if $\boldsymbol{\delta}$ is the $\mathbf{s}$-minimal degree of $(\mathbf{E},\mathbf{J})$ and $\mathbf{d} = -\boldsymbol{\delta}$, any $\mathbf{d}$-minimal interpolation basis $\mathbf{R}$ for $(\mathbf{E},\mathbf{J})$ has size at most $m^2 + m|\boldsymbol{\delta}|$, which for $\sigma \geqslant m$ is in $\mathcal{O}(m\sigma)$. Still, the algorithm in [18] cannot directly be used to compute such an $\mathbf{R}$ efficiently, because $|\mathbf{d} - \min(\mathbf{d})|$ can be as large as $\Theta(m\sigma)$, for example when $\boldsymbol{\delta} = (\sigma,0,\ldots,0)$; in this case, this algorithm uses $\mathcal{O}\tilde{~}(m^\omega\sigma)$ operations.

By Lemma 2.1, however, $\mathbf{d} = -\boldsymbol{\delta}$ satisfies $|\max(\mathbf{d}) - \mathbf{d}| \leqslant \sigma$. For this type of unbalanced shift, a solution in $\mathcal{O}\tilde{~}(m^{\omega-1}\sigma)$ already exists in the particular case of order basis computation [34, Section 6], building upon the partial linearization technique in [30, Section 3]. Here, we adopt a similar approach, taking advantage of the *a priori* knowledge of the column degree of the output matrix.

Lemma 4.2. *Let* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$*,* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$*, and* $\mathbf{s} \in \mathbb{Z}^m$*, and let* $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_m)$ *denote the* $\mathbf{s}$*-minimal degree of* $(\mathbf{E},\mathbf{J})$*.*

Then, let $\delta = \lceil\sigma/m\rceil \geqslant 1$*, and for* $i \in \{1,\ldots,m\}$ *write* $\delta_i = (\alpha_i - 1)\delta + \beta_i$ *with* $\alpha_i = \lceil\delta_i/\delta\rceil$ *and* $1 \leqslant \beta_i \leqslant \delta$ *if* $\delta_i > 0$*, and with* $\alpha_i = 1$ *and* $\beta_i = 0$ *if* $\delta_i = 0$*. Let* $\overline{m} = \alpha_1 + \cdots + \alpha_m$*. Then, define* $\overline{\boldsymbol{\delta}} \in \mathbb{N}^{\overline{m}}$ *as*

$$\overline{\boldsymbol{\delta}} = (\underbrace{\delta,\ldots,\delta,\beta_1}_{\alpha_1},\ldots,\underbrace{\delta,\ldots,\delta,\beta_m}_{\alpha_m}) \qquad (2)$$

and the expansion-compression matrix $\mathcal{E} \in \mathbb{K}[X]^{\overline{m}\times m}$ *as*

$$\mathcal{E} = \begin{bmatrix} 1 & & \\ X^{\delta} & & \\ \vdots & & \\ X^{(\alpha_1-1)\delta} & & \\ & \ddots & \\ & & 1 \\ & & X^{\delta} \\ & & \vdots \\ & & X^{(\alpha_m-1)\delta} \end{bmatrix}. \qquad (3)$$

Let further $\mathbf{d} = -\overline{\boldsymbol{\delta}} \in \mathbb{Z}^{\overline{m}}$ *and* $\mathbf{R} \in \mathbb{K}[X]^{\overline{m}\times\overline{m}}$ *be a* $\mathbf{d}$*-minimal interpolation basis for* $(\mathcal{E}\cdot\mathbf{E},\mathbf{J})$*. Then, the* $\mathbf{s}$*-Popov interpolation basis for* $(\mathbf{E},\mathbf{J})$ *is the submatrix of* $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R}\mathcal{E}$ *formed by its rows at indices* $\alpha_1 + \cdots + \alpha_i$ *for* $1 \leqslant i \leqslant m$*.*

Proof. Let $\mathbf{P}$ denote the $\mathbf{s}$-Popov interpolation basis for $(\mathbf{E},\mathbf{J})$; $\mathbf{P}$ has column degree $\boldsymbol{\delta}$. First, we partially linearize the columns of $\mathbf{P}$ in degree δ to obtain $\widetilde{\mathbf{P}} \in \mathbb{K}[X]^{m\times\overline{m}}$; more precisely, $\widetilde{\mathbf{P}}$ is a matrix of degree at most δ and such that $\mathbf{P} = \widetilde{\mathbf{P}}\mathcal{E}$. Then, we define $\overline{\mathbf{P}} \in \mathbb{K}[X]^{\overline{m}\times\overline{m}}$ as follows:

- for $1 \leqslant i \leqslant m$, the row $\alpha_1 + \cdots + \alpha_i$ of $\overline{\mathbf{P}}$ is the row i of $\widetilde{\mathbf{P}}$;
- for $0 \leqslant i \leqslant m-1$ and $1 \leqslant j \leqslant \alpha_{i+1} - 1$, the row $\alpha_1 + \cdots + \alpha_i + j$ of $\overline{\mathbf{P}}$ is the row $[0,\cdots,0,X^{\delta},-1,0,\cdots,0] \in \mathbb{K}[X]^{1\times\overline{m}}$ with the entry X^{δ} at column index $\alpha_1 + \cdots + \alpha_i + j$.

Since $\mathbf{P}$ is in $\mathbf{s}$-Popov form with column degree $\boldsymbol{\delta}$, it is in $-\boldsymbol{\delta}$-Popov form by Lemma 4.1. Then, one can check that $\overline{\mathbf{P}}$ is in $\mathbf{d}$-Popov form and has $\mathbf{d}$-row degree $(0,\ldots,0)$.

By construction, every row of $\overline{\mathbf{P}}$ is an interpolant for $(\mathcal{E}\cdot\mathbf{E},\mathbf{J})$. In particular, since $\mathbf{R}$ is an interpolation basis for $(\mathcal{E}\cdot\mathbf{E},\mathbf{J})$, there is a matrix $\mathbf{U} \in \mathbb{K}[X]^{\overline{m}\times\overline{m}}$ such that $\overline{\mathbf{P}} = \mathbf{U}\mathbf{R}$. Besides, there exists no interpolant $\overline{\mathbf{p}} \in \mathbb{K}[X]^{1\times\overline{m}}$ for $(\mathcal{E}\cdot\mathbf{E},\mathbf{J})$ which has $\mathbf{d}$-degree less than 0: otherwise, $\overline{\mathbf{p}}\mathcal{E}$ would be an interpolant for $(\mathbf{E},\mathbf{J})$, and it is easily checked that it would have $-\boldsymbol{\delta}$-degree less than 0, which is impossible.

Thus every row of $\mathbf{R}$ has $\mathbf{d}$-degree at least 0, and the predictable degree property [19, Theorem 6.3.13] shows that $\mathbf{U}$ is a constant matrix, and therefore unimodular. Then, $\overline{\mathbf{P}}$ is an interpolation basis for $(\mathcal{E} \cdot \mathbf{E}, \mathbf{J})$, and since it is in $\mathbf{d}$-Popov form, by Lemma 4.1 we obtain that $\overline{\mathbf{P}} = \mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R}$. The conclusion follows. □

Then, it remains to prove that such a basis $\mathbf{R}$ can be computed efficiently using the algorithm MinimalInterpolationBasis in [18]; this leads to Algorithm 2.

Algorithm 2. KnownMinDegMIB

Input:

- *a matrix* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$ *with* $\sigma \geqslant m > 0$,
- *a Jordan matrix* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ *in standard representation,*
- *a shift* $\mathbf{s} \in \mathbb{Z}^m$,
- $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m) \in \mathbb{N}^m$ *the* $\mathbf{s}$*-minimal degree of* $(\mathbf{E}, \mathbf{J})$.

Output: *the* $\mathbf{s}$*-Popov interpolation basis* $\mathbf{P}$ *for* $(\mathbf{E}, \mathbf{J})$.

1. $\delta \leftarrow \lceil \sigma/m \rceil$, $\alpha_i \leftarrow \max(1, \lceil \delta_i/\delta \rceil)$ *for* $1 \leqslant i \leqslant m$, $\overline{m} \leftarrow \alpha_1 + \cdots + \alpha_m$

2. *Let* $\overline{\boldsymbol{\delta}} \in \mathbb{N}^{\overline{m}}$ *as in* (2) *and* $\mathbf{d} \leftarrow -\overline{\boldsymbol{\delta}} \in \mathbb{N}^{\overline{m}}$

3. *Let* $\mathcal{E} \in \mathbb{K}[X]^{\overline{m}\times m}$ *as in* (3) *and* $\overline{\mathbf{E}} \leftarrow \mathcal{E} \cdot \mathbf{E}$

4. $\mathbf{R} \leftarrow$ MinimalInterpolationBasis$(\overline{\mathbf{E}}, \mathbf{J}, \mathbf{d} + (\delta, \ldots, \delta))$

5. $\overline{\mathbf{P}} \leftarrow \mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}\mathbf{R}$

6. *Return the submatrix of* $\overline{\mathbf{P}}\mathcal{E}$ *formed by the rows at indices* $\alpha_1 + \cdots + \alpha_i$ *for* $1 \leqslant i \leqslant m$

Proposition 4.3. *Assuming that* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ *is a Jordan matrix given by a standard representation, and assuming we have the* $\mathbf{s}$*-minimal degree of* $(\mathbf{E}, \mathbf{J})$ *as an additional input, there is a deterministic algorithm* KnownMinDegMIB *which solves Problem 1 using*

$$\mathcal{O}(m^{\omega-1}\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)) \qquad \text{if } \omega > 2,$$
$$\mathcal{O}(m\mathsf{M}(\sigma)\log(\sigma)\log(\sigma/m)\log(m)^3) \qquad \text{if } \omega = 2$$

operations in $\mathbb{K}$.

Proof. We focus on the case $\sigma \geqslant m$; otherwise, a better cost bound can be achieved even without knowing $\boldsymbol{\delta}$ [18, Theorem 1.4]. The correctness of Algorithm 2 follows from Lemma 4.2. We remark that it uses $\mathbf{d} + (\delta, \ldots, \delta)$ rather than $\mathbf{d}$ because the minimal interpolation basis algorithm in [18] requires the input shift to have non-negative entries. Since adding a constant to every entry of $\mathbf{d}$ does not change the notion of $\mathbf{d}$-reducedness, the basis $\mathbf{R}$ obtained at Step **4** is a $\mathbf{d}$-minimal interpolation basis for $(\mathbf{E}, \mathbf{J})$.

Concerning the cost bound, we will show that it is dominated by the time spent in Step **4**. First, we prove that $|\mathbf{d} - \min(\mathbf{d})| \in \mathcal{O}(\sigma)$, so that the cost of Step **4** follows from [18, Theorem 1.5]. We have $\alpha_i \leqslant 1 + \delta_i/\delta \leqslant 1 + m\delta_i/\sigma$ for all i. Thus, $\overline{m} = \alpha_1 + \cdots + \alpha_m \leqslant m + \sum_{1\leqslant i\leqslant m} m\delta_i/\sigma \leqslant 2m$ thanks to Lemma 2.1. Then, since all entries of $\mathbf{d}$ are in $\{-\delta, \ldots, 0\}$, we obtain $|\mathbf{d} - \min(\mathbf{d})| \leqslant \overline{m}\delta \leqslant 2m(1 + \sigma/m) \leqslant 4\sigma$.

Step **3** can be done in $\mathcal{O}(m\mathsf{M}(\sigma)\log(\sigma))$ operations according to Lemma 4.4 below.

Lemma 4.1 proves that the sum of the column degrees of $\mathbf{R}$ is $|\overline{\boldsymbol{\delta}}| = |\boldsymbol{\delta}| \leqslant \sigma$. Then, the product in Step **5** can be done in $\mathcal{O}(m^{\omega-1}\sigma)$ operations, by first linearizing the columns of $\mathbf{R}$ into a $\overline{m} \times \overline{m} + |\overline{\boldsymbol{\delta}}|$ matrix over $\mathbb{K}$, then left-multiplying this matrix by $\mathrm{lm}_{\mathbf{d}}(\mathbf{R})^{-1}$ (itself computed using $\mathcal{O}(m^{\omega})$ operations), and finally performing the inverse linearization.

Because of the degrees in $\overline{\mathbf{P}}$ and the definition of $\mathcal{E}$, the output in Step **6** can be formed without using any arithmetic operation. □

The efficient computation of $\mathcal{E} \cdot \mathbf{E}$ can be done with the algorithm for computing residuals in [18, Section 6].

Lemma 4.4. *The product* $\mathcal{E} \cdot \mathbf{E}$ *at Step 3 of Algorithm 2 can be computed using* $\mathcal{O}(m\mathsf{M}(\sigma)\log(\sigma))$ *operations in* $\mathbb{K}$.

Proof. The product $\mathcal{E} \cdot \mathbf{E}$ has $\overline{m}$ rows, with $\overline{m} \leqslant 2m$ as above. Besides, by definition of $\mathcal{E}$, each row of $\mathcal{E} \cdot \mathbf{E}$ is a product of the form $X^{i\delta} \cdot \mathbf{E}_{j,*}$, where $0 \leqslant i \leqslant m$, $1 \leqslant j \leqslant m$, and $\mathbf{E}_{j,*}$ denotes the row j of $\mathbf{E}$. In particular, $i\delta \leqslant 2\sigma$: then, according to [18, Proposition 6.1], each of these $\overline{m}$ products can be performed using $\mathcal{O}(\mathsf{M}(\sigma)\log(\sigma))$ operations in $\mathbb{K}$. □

This lemma and the partial linearization technique can also be used to compute the residual at Step **2.c** of Algorithm 1, that is, a product of the form $\mathbf{P} \cdot \mathbf{E}$ with the sum of the column degrees of $\mathbf{P}$ bounded by σ. First, we expand the high-degree columns of $\mathbf{P}$ to obtain $\overline{\mathbf{P}} \in \mathbb{K}[X]^{m\times\overline{m}}$ of degree less than $\lceil \sigma/m \rceil$ such that $\mathbf{P} = \overline{\mathbf{P}}\mathcal{E}$; then, we compute $\overline{\mathbf{E}} = \mathcal{E} \cdot \mathbf{E}$; and finally we rely on the algorithm in [18, Proposition 6.1] to compute $\overline{\mathbf{P}} \cdot \overline{\mathbf{E}} = \mathbf{P} \cdot \mathbf{E}$ efficiently.

Corollary 4.5. *Let* $\mathbf{E} \in \mathbb{K}^{m\times\sigma}$ *with* $\sigma \geqslant m$, *and let* $\mathbf{J} \in \mathbb{K}^{\sigma\times\sigma}$ *be a Jordan matrix given by a standard representation. Let* $\mathbf{P} \in \mathbb{K}[X]^{m\times m}$ *with column degree* $(\delta_1, \ldots, \delta_m)$ *such that* $\delta_1 + \cdots + \delta_m \leqslant \sigma$. *Then, the product* $\mathbf{P} \cdot \mathbf{E}$ *can be computed using* $\mathcal{O}(m^{\omega-1}\mathsf{M}(\sigma)\log(\sigma))$ *operations in* $\mathbb{K}$.

Acknowledgments. We thank B. Beckermann and G. Labahn for their valuable comments, as well as an anonymous referee for suggesting a shorter proof of Lemma 3.3. C.-P. Jeannerod and G. Villard were partly supported by the ANR project HPAC (ANR 11 BS02 013). V. Neiger was supported by the international mobility grants Explo'ra Doc from *Région Rhône-Alpes*, *PALSE*, and *Mitacs Globalink - Inria*. É. Schost was supported by NSERC.

5. REFERENCES

[1] B. Beckermann and G. Labahn. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM J. Matrix Anal. Appl.*, 15(3):804–823, 1994.

[2] B. Beckermann and G. Labahn. Fraction-free computation of matrix rational interpolants and matrix gcds. *SIAM J. Matrix Anal. Appl.*, 22(1):114–144, 2000.

[3] B. Beckermann, G. Labahn, and G. Villard. Normal forms for general polynomial matrices. *J. Symbolic Comput.*, 41(6):708–737, 2006.

[4] P. Beelen and K. Brander. Key equations for list decoding of Reed-Solomon codes and how to solve them. *J. Symbolic Comput.*, 45(7):773–786, 2010.

[5] A. Bostan, C.-P. Jeannerod, and É. Schost. Solving structured linear systems with large displacement rank. *Theor. Comput. Sci.*, 407(1-3):155–181, 2008.

[6] K. Brander. *Interpolation and List Decoding of Algebraic Codes*. PhD thesis, Technical University of Denmark, 2010.

[7] P. Busse. *Multivariate List Decoding of Evaluation Codes with a Gröbner Basis Perspective.* PhD thesis, University of Kentucky, 2008.

[8] D. G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Inform.*, 28(7):693–701, 1991.

[9] M. Chowdhury, C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. Faster algorithms for multivariate interpolation with multiplicities and simultaneous polynomial approximations. *IEEE Trans. Inf. Theory*, 61(5):2370–2387, 2015.

[10] H. Cohn and N. Heninger. Approximate common divisors via lattices. In *Tenth Algorithmic Number Theory Symposium*, pages 271–293. Mathematical Sciences Publishers (MSP), 2012-2013.

[11] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990.

[12] C. Devet, I. Goldberg, and N. Heninger. Optimally robust private information retrieval. In *USENIX Security 12*, pages 269–283. USENIX, 2012.

[13] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra (third edition).* Cambridge University Press, 2013.

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

[15] S. Gupta and A. Storjohann. Computing Hermite forms of polynomial matrices. In *ISSAC'11*, pages 155–162. ACM, 2011.

[16] V. Guruswami and A. Rudra. Explicit codes achieving list decoding capacity: Error-correction with optimal redundancy. *IEEE Trans. Inf. Theory*, 54(1):135–150, 2008.

[17] V. Guruswami and M. Sudan. Improved decoding of Reed-Solomon and algebraic-geometry codes. *IEEE Trans. Inf. Theory*, 45(6):1757–1767, 1999.

[18] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. Computing minimal interpolation bases. HAL Open archive - https://hal.inria.fr/hal-01241781, 2015.

[19] T. Kailath. *Linear Systems.* Prentice-Hall, 1980.

[20] D. E. Knuth. The analysis of algorithms. In *Congrès int. Math., Nice, France*, volume 3, pages 269–274, 1970.

[21] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. *IEEE Trans. Inf. Theory*, 49(11):2809–2825, 2003.

[22] F. Le Gall. Powers of tensors and fast matrix multiplication. In *ISSAC'14*, pages 296–303. ACM, 2014.

[23] R. T. Moenck. Fast computation of GCDs. In *Proc. 5th ACM Symp. Theory Comp.*, pages 142–151, 1973.

[24] T. Mulders and A. Storjohann. On lattice reduction for polynomial matrices. *J. Symbolic Comput.*, 35:377–401, 2003.

[25] V. Olshevsky and M. A. Shokrollahi. A displacement approach to efficient decoding of algebraic-geometric codes. In *STOC'99*, pages 235–244. ACM, 1999.

[26] F. Parvaresh and A. Vardy. Correcting errors beyond the Guruswami-Sudan radius in polynomial time. In *FOCS'05*, pages 285–294. IEEE, 2005.

[27] R. M. Roth and G. Ruckenstein. Efficient decoding of Reed-Solomon codes beyond half the minimum distance. *IEEE Trans. Inf. Theory*, 46(1):246–257, 2000.

[28] S. Sarkar and A. Storjohann. Normalization of row reduced matrices. In *ISSAC'11*, pages 297–304. ACM, 2011.

[29] A. Schönhage. Schnelle Berechnung von Kettenbruchentwicklungen. *Acta Inform.*, 1:139–144, 1971.

[30] A. Storjohann. Notes on computing minimal approximant bases. In *Dagstuhl Seminar Proceedings*, 2006.

[31] M. Van Barel and A. Bultheel. A general module theoretic framework for vector M-Padé and matrix rational interpolation. *Numer. Algorithms*, 3:451–462, 1992.

[32] A. Zeh, C. Gentner, and D. Augot. An interpolation procedure for list decoding Reed-Solomon codes based on generalized key equations. *IEEE Trans. Inf. Theory*, 57(9):5946–5959, 2011.

[33] W. Zhou. *Fast Order Basis and Kernel Basis Computation and Related Problems.* PhD thesis, University of Waterloo, 2012.

[34] W. Zhou and G. Labahn. Efficient algorithms for order basis computation. *J. Symbolic Comput.*, 47(7):793–819, 2012.

APPENDIX

A. REDUCING THE ENTRIES OF THE SHIFT

Let $\mathbf{A} \in \mathbb{K}[X]^{m\times m}$ be nonsingular, let $\mathbf{s} \in \mathbb{Z}^m$, and consider $\sigma \in \mathbb{N}$ such that $\sigma > \deg(\det(\mathbf{A}))$. Here, we show how to construct a shift $\mathbf{t} \in \mathbb{N}^m$ such that

- the $\mathbf{s}$-Popov form $\mathbf{P}$ of $\mathbf{A}$ is also in $\mathbf{t}$-Popov form;
- $\min(\mathbf{t}) = 0$, $\max(\mathbf{t}) \leqslant (m-1)\sigma$, and $|\mathbf{t}| \leqslant m^2\sigma/2$.

We write $\hat{\mathbf{s}} = (s_{\pi(1)}, \ldots, s_{\pi(m)})$ where π is a permutation of $\{1, \ldots, m\}$ such that $\hat{\mathbf{s}}$ is non-decreasing. Then, we define $\hat{\mathbf{t}} = (\hat{t}_1, \ldots, \hat{t}_m)$ by $\hat{t}_1 = 0$ and, for $2 \leqslant i \leqslant m$,

$$\hat{t}_i - \hat{t}_{i-1} = \begin{cases} \sigma & \text{if } \hat{s}_i - \hat{s}_{i-1} \geqslant \sigma, \\ \hat{s}_i - \hat{s}_{i-1} & \text{otherwise.} \end{cases}$$

Let $\mathbf{t} = (\hat{t}_{\pi^{-1}(1)}, \ldots, \hat{t}_{\pi^{-1}(m)})$. Since the diagonal entries of $\mathbf{P}$ have degree at most $\deg(\det(\mathbf{A})) < \sigma$, we obtain that $\mathbf{P}$ is in $\mathbf{t}$-diagonal weak Popov form and thus in $\mathbf{t}$-Popov form.

B. EXAMPLE OF ORDER BASIS WITH SIZE BEYOND OUR TARGET COST

We focus on a Hermite-Padé approximation problem with input $\mathbf{F}$ of dimensions $2m \times 1$ as below, order σ with $\sigma \geqslant m$, and shift $\mathbf{s} = (0, \ldots, 0, \sigma, \ldots, \sigma) \in \mathbb{N}^{2m}$ with m entries 0 and m entries σ.

Let f be a polynomial in X with nonzero constant coefficient, and let $f_1, \ldots, f_m$ be generic polynomials in X of degree less than σ. Then, we consider the following input with all entries truncated modulo X^σ:

$$\mathbf{F} = [f, f{+}Xf, X(f{+}Xf), \cdots, X^{m-2}(f{+}Xf), f_1, \cdots, f_m]^\mathsf{T}.$$

After m steps, the iterative algorithm in [1] has computed an $\mathbf{s}$-minimal basis $\mathbf{P}^{(m)}$ of approximants for $\mathbf{F}$ and order m, which is such that $\mathbf{t} = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}^{(m)}) = (1, \ldots, 1, \sigma, \ldots, \sigma)$ and $\mathbf{P}^{(m)}\mathbf{F} = [0, \cdots, 0, X^m f, X^m g_1, \cdots, X^m g_m]^\mathsf{T} \bmod X^\sigma$, for some polynomials $g_1, \ldots, g_m$.

Now we finish the process up to order σ. Since the coefficient of degree m of $X^m f$ is nonzero and because of the specific shift $\mathbf{t}$, the obtained $\mathbf{s}$-minimal basis $\mathbf{P}$ of approximants for $\mathbf{F}$ has degree profile

$$\mathbf{P} = \begin{bmatrix} [1] & [0] & & & & & \\ \vdots & \ddots & \ddots & & & & \\ [1] & \cdots & [1] & [0] & & & \\ [d+1] & \cdots & [d+1] & [d+1] & & & \\ [d] & \cdots & [d] & [d] & [0] & & \\ \vdots & \cdots & \vdots & \vdots & & \ddots & \\ [d] & \cdots & [d] & [d] & & & [0] \end{bmatrix},$$

where $d = \sigma - m$, $[i]$ denotes an entry of degree i, the entries left blank correspond to the zero polynomial, and the entries $[d{+}1]$ are on the m-th row. In particular, $\mathbf{P}$ has size $\Theta(m^2\sigma)$.