Computing Canonical Bases of Modules of Univariate Relations

Vincent Neiger, Thi Xuan Vu

Computing Canonical Bases of Modules of Univariate Relations

Vincent Neiger
Technical University of Denmark
Kgs. Lyngby, Denmark
vinn@dtu.dk

Vu Thi Xuan
ENS de Lyon, LIP (CNRS, Inria, ENSL, UCBL)
Lyon, France
thi.vu@ens-lyon.fr

ABSTRACT

We study the computation of canonical bases of sets of univariate relations $(p_1, \ldots, p_m) \in \mathbb{K}[x]^m$ such that $p_1 f_1 + \cdots + p_m f_m = 0$; here, the input elements $f_1, \ldots, f_m$ are from a quotient $\mathbb{K}[x]^n/\mathcal{M}$, where $\mathcal{M}$ is a $\mathbb{K}[x]$-module of rank n given by a basis $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ in Hermite form. We exploit the triangular shape of $\mathbf{M}$ to generalize a divide-and-conquer approach which originates from fast minimal approximant basis algorithms. Besides recent techniques for this approach, we rely on high-order lifting to perform fast modular products of polynomial matrices of the form $\mathbf{PF} \bmod \mathbf{M}$.

Our algorithm uses $O^{\sim}(m^{\omega-1}D + n^{\omega}D/m)$ operations in $\mathbb{K}$, where $D = \deg(\det(\mathbf{M}))$ is the $\mathbb{K}$-vector space dimension of $\mathbb{K}[x]^n/\mathcal{M}$, $O^{\sim}(\cdot)$ indicates that logarithmic factors are omitted, and ω is the exponent of matrix multiplication. This had previously only been achieved for a diagonal matrix $\mathbf{M}$. Furthermore, our algorithm can be used to compute the shifted Popov form of a nonsingular matrix within the same cost bound, up to logarithmic factors, as the previously fastest known algorithm, which is randomized.

KEYWORDS

Polynomial matrix; shifted Popov form; division with remainder; univariate equations; syzygy module.

1 INTRODUCTION

In what follows, $\mathbb{K}$ is a field, $\mathbb{K}[x]$ denotes the set of univariate polynomials in x over $\mathbb{K}$, and $\mathbb{K}[x]^{m\times n}$ denotes the set of $m \times n$ (univariate) polynomial matrices.

Univariate relations. Let us consider a (free) $\mathbb{K}[x]$-submodule $\mathcal{M} \subseteq \mathbb{K}[x]^n$ of rank n, specified by one of its bases, represented as the rows of a nonsingular matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$. Besides, let some elements $f_1, \ldots, f_m \in \mathbb{K}[x]^n/\mathcal{M}$ be represented as a matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$. Then, the kernel of the module morphism

$$\begin{array}{rccc} \varphi_{\mathcal{M},f}: & \mathbb{K}[x]^m & \to & \mathbb{K}[x]^n/\mathcal{M} \\ & (p_1, \ldots, p_m) & \mapsto & p_1 f_1 + \cdots + p_m f_m \end{array}$$

consists of relations between the f_i's, and is known as a *syzygy module* [10]. From the matrix viewpoint above, we write it as

$$\mathcal{R}(\mathbf{M}, \mathbf{F}) = \{\mathbf{p} \in \mathbb{K}[x]^{1\times m} \mid \mathbf{pF} = \mathbf{0} \bmod \mathbf{M}\},$$

where the notation $\mathbf{A} = \mathbf{0} \bmod \mathbf{M}$ stands for "$\mathbf{A} = \mathbf{QM}$ for some $\mathbf{Q}$", which means that the rows of $\mathbf{A}$ are in the module $\mathcal{M}$. Hereafter, the elements of $\mathcal{R}(\mathbf{M}, \mathbf{F})$ are called *relations of* $\mathcal{R}(\mathbf{M}, \mathbf{F})$.

Examples of such relations are the following.

- *Hermite-Padé approximants* are relations for $n = 1$ and $\mathcal{M} = x^D\mathbb{K}[x]$. That is, given polynomials $f_1, \ldots, f_m$, the corresponding approximants are all $(p_1, \ldots, p_m) \in \mathbb{K}[x]^m$ such that $p_1 f_1 + \cdots + p_m f_m = 0 \bmod x^D$. Fast algorithms for finding such approximants include [3, 15, 19, 31, 37].
- *Multipoint Padé approximants:* the fast computation of relations when $\mathcal{M}$ is a product of ideals, corresponding to a diagonal basis $\mathbf{M} = \mathrm{diag}(M_1, \ldots, M_n)$, was studied in [2, 4, 19, 20, 26, 32]. Many of these references focus on $M_1, \ldots, M_n$ which split over $\mathbb{K}$ with known roots and multiplicities; then, relations are known as multipoint Padé approximants [1], or also *interpolants* [4, 20]. In this case, a relation can be thought of as a solution to a linear system over $\mathbb{K}[x]$ in which the jth equation is modulo M_j.

Canonical bases. Since $\det(\mathbf{M})\mathbb{K}[x]^m \subseteq \mathcal{R}(\mathbf{M}, \mathbf{F}) \subseteq \mathbb{K}[x]^m$, the module $\mathcal{R}(\mathbf{M}, \mathbf{F})$ is free of rank m [8, Sec. 12.1, Thm. 4]. Hence, any of its bases can be represented as the rows of a nonsingular matrix in $\mathbb{K}[x]^{m\times m}$, which we call a *relation basis for* $\mathcal{R}(\mathbf{M}, \mathbf{F})$.

Here, we are interested in computing relation bases in *shifted Popov form* [5, 27]. Such bases are canonical in terms of the module $\mathcal{R}(\mathbf{M}, \mathbf{F})$ and of a *shift*, the latter being a tuple $\mathbf{s} \in \mathbb{Z}^n$ used as column weights in the notion of degree for row vectors. Furthermore, the degrees in shifted Popov bases are well controlled, which helps to compute them faster than less constrained types of bases (see [19] and [25, Sec. 1.2.2]) and then, once obtained, to exploit them for other purposes (see for example [28, Thm. 12]). Having a shifted Popov basis of a submodule $\mathcal{M} \subseteq \mathbb{K}[x]^n$ is particularly useful for efficient computations in the quotient $\mathbb{K}[x]^n/\mathcal{M}$ (see Section 3).

In fact, shifted Popov bases coincide with Gröbner bases for $\mathbb{K}[x]$-submodules of $\mathbb{K}[x]^n$ [9, Chap. 15], for a term-over-position monomial order weighted by the entries of the shift. For more details about this link, we refer to [24, Chap. 6] and [25, Chap. 1].

For a shift $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbb{Z}^n$, the $\mathbf{s}$-*degree* of a row vector $\mathbf{p} = [p_1, \ldots, p_n] \in \mathbb{K}[x]^{1\times n}$ is $\max_{1\leqslant j\leqslant n}(\deg(p_j) + s_j)$; the $\mathbf{s}$-*row degree* of a matrix $\mathbf{P} \in \mathbb{K}[x]^{m\times n}$ is $\mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}) = (d_1, \ldots, d_m)$ with d_i the $\mathbf{s}$-degree of the ith row of $\mathbf{P}$. Then, the $\mathbf{s}$-*leading matrix* of $\mathbf{P} = [p_{i,j}]_{ij}$ is the matrix $\mathrm{lm}_{\mathbf{s}}(\mathbf{P}) \in \mathbb{K}^{m\times n}$ whose entry (i, j) is the coefficient of degree $d_i - s_j$ of $p_{i,j}$. Similarly, the list of column degrees of a matrix $\mathbf{P}$ is denoted by $\mathrm{cdeg}(\mathbf{P})$.

Definition 1.1 ([5, 21]). 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

- $\mathbf{s}$-*reduced form* if $\mathrm{lm}_{\mathbf{s}}(\mathbf{P})$ is invertible;
- $\mathbf{s}$-*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.

ISSAC '17, Kaiserslautern, Germany

Problem 1: Relation basis
Input:
- nonsingular matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$,
- matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output:
- the $\mathbf{s}$-Popov relation basis $\mathbf{P} \in \mathbb{K}[x]^{m\times m}$ for $\mathcal{R}(\mathbf{M}, \mathbf{F})$.

Hereafter, when we introduce a matrix by saying that it is reduced, it is understood that it is nonsingular. Similar forms can be defined for modules generated by the columns of a matrix rather than by its rows; in the context of polynomial matrix division with remainder, we will use the notion of $\mathbf{P}$ in *column reduced* form, meaning that $\mathrm{lm}_{\mathbf{0}}(\mathbf{P}^{\mathsf{T}})$ is invertible. In particular, we remark that any matrix in shifted Popov form is also column reduced.

Considering relation bases $\mathbf{P}$ for $\mathcal{R}(\mathbf{M}, \mathbf{F})$ in shifted Popov form offers a strong control over the degrees of their entries. As shifted (row) reduced bases, they satisfy the *predictable degree property* [12], which is at the core of the correctness of a divide-and-conquer approach behind most algorithms for the two specific situations described above, for example [3, 15, 16, 20]. Furthermore, as column reduced matrices they have small average column degree, which is central in the efficiency of fast algorithms for non-uniform shifts [19, 26]. Indeed, we will see in Corollary 2.4 that

$$|\mathrm{cdeg}(\mathbf{P})| = \deg(\det(\mathbf{P})) \leqslant \deg(\det(\mathbf{M})),$$

where $|\cdot|$ denotes the sum of the entries of a tuple.

Below, triangular canonical bases will play an important role. A matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ is in *Hermite form* if $\mathbf{M}$ is upper triangular and $\mathrm{lm}_{\mathbf{0}}(\mathbf{M}^{\mathsf{T}})$ is the identity matrix; or, equivalently, if $\mathbf{M}$ is in $(dn, d(n-1), \ldots, d)$-Popov form for any $d \geqslant \deg(\det(\mathbf{M}))$.

Relations modulo Hermite forms. Our main focus is on the case where $\mathbf{M}$ is in Hermite form and $\mathbf{F}$ is already reduced modulo $\mathbf{M}$. In this article, all comparisons of tuples are componentwise.

Theorem 1.2. *If $\mathbf{M}$ is in Hermite form and $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M})$, there is a deterministic algorithm which solves Problem 1 using*

$$\tilde{O}\left(m^{\omega-1}D + n^{\omega}D/m\right)$$

operations in $\mathbb{K}$, where $D = \deg(\det(\mathbf{M})) = |\mathrm{cdeg}(\mathbf{M})|$.

Here, the exponent ω is so that we can multiply $m \times m$ matrices over $\mathbb{K}$ in $O(m^{\omega})$ operations in $\mathbb{K}$, the best known bound being $\omega < 2.38$ [7, 23]. The notation $O^{\tilde{}}(\cdot)$ means that we have omitted the logarithmic factors in the asymptotic bound.

To put this cost bound in perspective, we note that the representation of the input $\mathbf{F}$ and $\mathbf{M}$ requires at most $(m+n)D$ field elements, while that of the output basis uses at most mD elements. In many applications we have $n \in O(m)$, in which case the cost bound becomes $O^{\tilde{}}(m^{\omega-1}D)$, which is satisfactory.

To the best of our knowledge, previous algorithms with a comparable cost bound focus on the case of a diagonal matrix $\mathbf{M}$.

The case of minimal approximant bases $\mathbf{M} = x^d\mathbf{I}_n$ has concentrated a lot of attention. A first algorithm with cost quasi-linear in d was given [3]. It was then improved in [15, 30, 37], obtaining the cost bound $O^{\tilde{}}(m^{\omega-1}nd) = O^{\tilde{}}(m^{\omega-1}D)$ under assumptions on the dimensions m and n or on the shift.

In [20], the divide-and-conquer approach of [3] was carried over and made efficient in the more general case $\mathbf{M} = \mathrm{diag}(M_1, \ldots, M_n)$, where the polynomials M_i split over $\mathbb{K}$ with known linear factors. This approach was then augmented in [19] with a strategy focusing on degree information to efficiently compute the shifted Popov bases for arbitrary shifts, achieving the cost bound $O^{\tilde{}}(m^{\omega-1}D)$.

Then, the case of a diagonal matrix $\mathbf{M}$, with no assumption on the diagonal entries, was solved within $O^{\tilde{}}(m^{\omega-1}D + n^{\omega}D/m)$ [26]. The main new ingredient developed in [26] was an efficient algorithm for the case $n = 1$, that is, when solving a single linear equation modulo a polynomial; we will also make use of this algorithm here.

In this paper we obtain the same cost bound as [26] for any matrix $\mathbf{M}$ in Hermite form. For a more detailed comparison with earlier algorithms focusing on diagonal matrices $\mathbf{M}$, we refer the reader to [26, Sec. 1.2] and in particular Table 2 therein.

Our algorithm essentially follows the approach of [26]. In particular, it uses the algorithm developed there for $n = 1$. However, working modulo Hermite forms instead of diagonal matrices makes the computation of *residuals* much more involved. The residual is a modular product $\mathbf{PF}$ mod $\mathbf{M}$ which is computed after the first recursive call and is to be used as an input replacing $\mathbf{F}$ for the second recursive call. When $\mathbf{M}$ is diagonal, its computation boils down to the multiplication of $\mathbf{P}$ and $\mathbf{F}$, although care has to be taken to account for their possibly unbalanced column degrees. However, when $\mathbf{M}$ is triangular, computing $\mathbf{PF}$ mod $\mathbf{M}$ becomes a much greater challenge: we want to compute a matrix remainder instead of simply taking polynomial remainders for each column separately. We handle this, while still taking unbalanced degrees into account, by resorting to high-order lifting [29].

Shifted Popov forms of matrices. A specific instance of Problem 1 yields the following problem: given a shift $\mathbf{s} \in \mathbb{Z}^n$ and a nonsingular matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$, compute the $\mathbf{s}$-Popov form of $\mathbf{M}$. Indeed, the latter is the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{I}_n)$ (see Lemma 2.7).

To compute this relation basis efficiently, we start by computing the Hermite form $\mathbf{H}$ of $\mathbf{M}$, which can be done deterministically in $O^{\tilde{}}(n^{\omega}\lceil D_{\mathbf{M}}/n\rceil)$ operations [22]. Here, $D_{\mathbf{M}}$ is the *generic determinant bound* [17]; writing $\mathbf{M} = [a_{ij}]$, it is defined as

$$D_{\mathbf{M}} = \max_{\pi\in S_n} \textstyle\sum_{1\leqslant i\leqslant n} \max(0, \deg(a_{i,\pi_i}))$$

where S_n is the set of permutations of $\{1, \ldots, n\}$. In particular, $D_{\mathbf{M}}/n$ is bounded from above by both the average of the degrees of the columns of $\mathbf{M}$ and that of its rows. For more details about this quantity, we refer to [17, Sec. 6] and [22, Sec. 2.3].

Since the rows of $\mathbf{H}$ generate the same module as $\mathbf{M}$, we have $\mathcal{R}(\mathbf{M}, \mathbf{I}_n) = \mathcal{R}(\mathbf{H}, \mathbf{I}_n)$ (see Lemma 2.5). Then, applying our algorithm for relations modulo $\mathbf{H}$ has a cost of $O^{\tilde{}}(n^{\omega-1}\deg(\det(\mathbf{H})))$ operations, according to Theorem 1.2. This yields the next result.

Theorem 1.3. *Given a shift $\mathbf{s} \in \mathbb{Z}^n$ and a nonsingular matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$, there is a deterministic algorithm which computes the $\mathbf{s}$-Popov form of $\mathbf{M}$ using*

$$\tilde{O}\left(n^{\omega}\lceil D_{\mathbf{M}}/n\rceil\right) \subseteq \tilde{O}\left(n^{\omega}\deg(\mathbf{M})\right)$$

operations in $\mathbb{K}$.

A similar cost bound was obtained in [26], yet with a randomized algorithm. The latter follows the approach of [18] for computing Hermite forms, whose first step determines the Smith form $\mathbf{S}$ of $\mathbf{M}$

along with a matrix $\mathbf{F}$ such that the sought matrix is the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{S}, \mathbf{F})$, with $\mathbf{S}$ being therefore a diagonal matrix. Here, relying on the deterministic computation of the Hermite form of $\mathbf{M}$, our algorithm for relation bases modulo Hermite forms allows us to circumvent the computation of $\mathbf{S}$, for which the currently fastest known algorithm is Las Vegas randomized [29]. For a more detailed comparison with earlier row reduction and Popov forms algorithms, we refer to [26, Sec. 1.1] and Table 1 therein.

General relation bases. To solve the general case of Problem 1, one can proceed as follows:

- find the Hermite form $\mathbf{H}$ of $\mathbf{M}$, using [22, Algo. 1 and 3];
- reduce $\mathbf{F}$ modulo $\mathbf{H}$, for example using Algorithm 1;
- apply Algorithm 5 for relations modulo a Hermite form.

Outline. We first give basic properties about matrix division and relation bases (Section 2). We then focus on the fast computation of residuals (Section 3). After that, we discuss three situations which have already been solved efficiently in the literature (Section 4): when $n = 1$, when information on the output degrees is available, and when $D \leqslant m$. Finally, we present our algorithm for relations modulo Hermite forms (Section 5).

2 PRELIMINARIES ON POLYNOMIAL MATRIX DIVISION AND MODULES OF RELATIONS

Division with remainder. Polynomial matrix division is a central notion in this paper, since we aim at solving equations modulo $\mathbf{M}$.

Theorem 2.1 ([13, IV.§2],[21, Thm. 6.3-15]). *For any $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ and any column reduced $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$, there exist unique matrices $\mathbf{Q}, \mathbf{R} \in \mathbb{K}[x]^{m\times n}$ such that $\mathbf{F} = \mathbf{QM} + \mathbf{R}$ and $\operatorname{cdeg}(\mathbf{R}) < \operatorname{cdeg}(\mathbf{M})$.*

Hereafter, we write $\operatorname{Quo}(\mathbf{F}, \mathbf{M})$ and $\operatorname{Rem}(\mathbf{F}, \mathbf{M})$ for the quotient $\mathbf{Q}$ and the remainder $\mathbf{R}$. We have the following properties.

Lemma 2.2. *We have $\operatorname{Rem}(\mathbf{P}\operatorname{Rem}(\mathbf{F}, \mathbf{M}), \mathbf{M}) = \operatorname{Rem}(\mathbf{PF}, \mathbf{M})$ and $\operatorname{Rem}\left(\begin{bmatrix}\mathbf{F}\\ \mathbf{G}\end{bmatrix}, \mathbf{M}\right) = \begin{bmatrix}\operatorname{Rem}(\mathbf{F}, \mathbf{M})\\ \operatorname{Rem}(\mathbf{G}, \mathbf{M})\end{bmatrix}$ for any $\mathbf{F} \in \mathbb{K}[x]^{m\times n}, \mathbf{G} \in \mathbb{K}[x]^{*\times n}$, $\mathbf{P} \in \mathbb{K}[x]^{*\times m}$ and any column reduced $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$.*

Degree control for relation bases. We first relate the vector space dimension of quotients and the degree of determinant of bases.

Lemma 2.3. *Let $\mathcal{M}$ be a $\mathbb{K}[x]$-submodule of $\mathbb{K}[x]^n$ of rank n. Then, the dimension of $\mathbb{K}[x]^n/\mathcal{M}$ as a $\mathbb{K}$-vector space is $\deg(\det(\mathbf{M}))$, for any matrix $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ whose rows form a basis of $\mathcal{M}$.*

Proof. Since the degree of the determinant is the same for all bases of $\mathcal{M}$, we may assume that $\mathbf{M}$ is column reduced. Then, Theorem 2.1 implies that there is a $\mathbb{K}$-vector space isomorphism $\mathbb{K}[x]^n/\mathcal{M} \cong \mathbb{K}[x]/\langle x^{d_1}\rangle \times \cdots \times \mathbb{K}[x]/\langle x^{d_n}\rangle$, where $(d_1, \ldots, d_n) = \operatorname{cdeg}(\mathbf{M})$. Thus, the dimension of $\mathbb{K}[x]^n/\mathcal{M}$ is $d_1 + \cdots + d_n$, which is equal to $\deg(\det(\mathbf{M}))$ according to [21, Sec. 6.3.2]. □

This allows us to bound the sum of column degrees of any column reduced relation basis; for example, a shifted Popov relation basis.

Corollary 2.4. *Let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ be nonsingular. Then, any relation basis $\mathbf{P} \in \mathbb{K}[x]^{m\times m}$ for $\mathcal{R}(\mathbf{M}, \mathbf{F})$ is such that $\deg(\det(\mathbf{P})) \leqslant \deg(\det(\mathbf{M}))$. In particular, if $\mathbf{P}$ is column reduced, then $|\operatorname{cdeg}(\mathbf{P})| \leqslant \deg(\det(\mathbf{M}))$.*

Proof. Let $\mathcal{M}$ be the row space of $\mathbf{M}$. By definition, $\mathcal{R}(\mathbf{M}, \mathbf{F})$ is the kernel of $\varphi_{\mathcal{M},\mathbf{F}}$ (see Section 1), hence $\mathbb{K}[x]^m/\mathcal{R}(\mathbf{M}, \mathbf{F})$ is isomorphic to a submodule of $\mathbb{K}[x]^n/\mathcal{M}$. Since, by Lemma 2.3, the dimensions of $\mathbb{K}[x]^m/\mathcal{R}(\mathbf{M}, \mathbf{F})$ and $\mathbb{K}[x]^m/\mathcal{M}$ are $\deg(\det(\mathbf{P}))$ and $\deg(\det(\mathbf{M}))$, we obtain $\deg(\det(\mathbf{P})) \leqslant \deg(\det(\mathbf{M}))$. □

Properties of relation bases. We now formalize the facts that $\mathcal{R}(\mathbf{M}, \mathbf{F})$ is not changed if $\mathbf{M}$ is replaced by another basis of the module generated by its rows; or if $\mathbf{F}$ and $\mathbf{M}$ are right-multiplied by the same nonsingular matrix; or yet if $\mathbf{F}$ is considered modulo $\mathbf{M}$.

Lemma 2.5. *Let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ be nonsingular. Then, for any nonsingular $\mathbf{A} \in \mathbb{K}[x]^{n\times n}$, any matrix $\mathbf{B} \in \mathbb{K}[x]^{m\times n}$, and any unimodular $\mathbf{U} \in \mathbb{K}[x]^{m\times m}$, we have*

$$\mathcal{R}(\mathbf{M}, \mathbf{F}) = \mathcal{R}(\mathbf{UM}, \mathbf{F}) = \mathcal{R}(\mathbf{MA}, \mathbf{FA}) = \mathcal{R}(\mathbf{M}, \mathbf{F} + \mathbf{BM}).$$

A first consequence is that we may discard identity columns in $\mathbf{M}$.

Corollary 2.6. *Let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ be nonsingular. Suppose that $\mathbf{M}$ has at least $k \in \mathbb{Z}_{>0}$ identity columns, and that the corresponding columns of $\mathbf{F}$ are zero. Then, let π_1, π_2 be $n \times n$ permutation matrices such that*

$$\pi_1 \mathbf{M} \pi_2 = \begin{bmatrix}\mathbf{I}_k & \mathbf{B}\\ \mathbf{0} & \mathbf{N}\end{bmatrix} \quad \text{and} \quad \mathbf{F}\pi_2 = \begin{bmatrix}\mathbf{0} & \mathbf{G}\end{bmatrix},$$

where $\mathbf{G} \in \mathbb{K}[x]^{m\times(n-k)}$. Then, $\mathcal{R}(\mathbf{M}, \mathbf{F}) = \mathcal{R}(\mathbf{N}, \mathbf{G})$.

Another consequence concerns the transformation of a matrix into shifted Popov form. Indeed, Lemma 2.5 together with the next lemma imply in particular that the $\mathbf{s}$-Popov form of $\mathbf{M}$ is the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{H}, \mathbf{I}_n)$, where $\mathbf{H}$ is the Hermite form of $\mathbf{M}$.

Lemma 2.7. *Let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ be nonsingular. Then, $\mathbf{M}$ is a relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{I}_n)$. It follows that the $\mathbf{s}$-Popov form of $\mathbf{M}$ is the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{I}_m)$, for any $\mathbf{s} \in \mathbb{Z}^n$.*

Proof. Let $\mathbf{P} \in \mathbb{K}[x]^{n\times n}$ be a relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{I}_n)$. Then, $\mathbf{PI}_n = \mathbf{QM}$ for some $\mathbf{Q} \in \mathbb{K}[x]^{n\times n}$; since the rows of $\mathbf{M}$ belong to $\mathcal{R}(\mathbf{M}, \mathbf{I}_n)$, we also have $\mathbf{M} = \mathbf{RP}$ for some $\mathbf{R} \in \mathbb{K}[x]^{n\times n}$. Since $\mathbf{P}$ is nonsingular, $\mathbf{P} = \mathbf{QRP}$ implies that $\mathbf{QR} = \mathbf{I}_n$, and therefore $\mathbf{R}$ is unimodular. Thus, $\mathbf{M} = \mathbf{RP}$ is a relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{I}_n)$. □

Divide and conquer approach. Here we give properties in the case of a block triangular matrix $\mathbf{M}$. They imply, if $\mathbf{M}$ is in Hermite form, that Problem 1 can be solved recursively by splitting the instance in dimension n into two instances in dimension $n/2$.

Lemma 2.8. *Let $\mathbf{M}_1 \in \mathbb{K}[x]^{n_1\times n_1}$, $\mathbf{M}_2 \in \mathbb{K}[x]^{n_2\times n_2}$, and $\mathbf{A} \in \mathbb{K}[x]^{n_1\times n_2}$ be such that $\mathbf{M} = \left[\begin{smallmatrix}\mathbf{M}_1 & \mathbf{A}\\ \mathbf{0} & \mathbf{M}_2\end{smallmatrix}\right]$ is column reduced. For any $\mathbf{F}_1 \in \mathbb{K}[x]^{m\times n_1}$ and $\mathbf{F}_2 \in \mathbb{K}[x]^{m\times n_2}$, we have $\operatorname{Rem}([\mathbf{F}_1 \ \ \mathbf{F}_2], \mathbf{M}) = [\operatorname{Rem}(\mathbf{F}_1, \mathbf{M}_1) \ \ \operatorname{Rem}(\mathbf{F}_2 - \operatorname{Quo}(\mathbf{F}_1, \mathbf{M}_1)\mathbf{A}, \mathbf{M}_2)]$.*

Proof. Writing $[\mathbf{F}_1 \ \ \mathbf{F}_2] = [\mathbf{Q}_1 \ \ \mathbf{Q}_2]\mathbf{M} + [\mathbf{R}_1 \ \ \mathbf{R}_2]$ where $\operatorname{cdeg}([\mathbf{R}_1 \ \ \mathbf{R}_2]) < \operatorname{cdeg}(\mathbf{M})$, we obtain $\mathbf{F}_1 = \mathbf{Q}_1\mathbf{M}_1 + \mathbf{R}_1$ as well as $\operatorname{cdeg}(\mathbf{R}_1) < \operatorname{cdeg}(\mathbf{M}_1)$, and therefore $\mathbf{R}_1 = \operatorname{Rem}(\mathbf{F}_1, \mathbf{M}_1)$ and $\mathbf{Q}_1 = \operatorname{Quo}(\mathbf{F}_1, \mathbf{M}_1)$. The result follows from $\mathbf{F}_2 = \mathbf{Q}_1\mathbf{A} + \mathbf{Q}_2\mathbf{M}_2 + \mathbf{R}_2$. □

Theorem 2.9. *Let $\mathbf{M} = \left[\begin{smallmatrix}\mathbf{M}_1 & \ast\\ \mathbf{0} & \mathbf{M}_2\end{smallmatrix}\right]$ be column reduced, where $\mathbf{M}_1 \in \mathbb{K}[x]^{n_1\times n_1}$ and $\mathbf{M}_2 \in \mathbb{K}[x]^{n_2\times n_2}$, and let $\mathbf{F}_1 \in \mathbb{K}[x]^{m\times n_1}$ and $\mathbf{F}_2 \in \mathbb{K}[x]^{m\times n_2}$. If $\mathbf{P}_1$ is a basis for $\mathcal{R}(\mathbf{M}_1, \mathbf{F}_1)$, then $\operatorname{Rem}(\mathbf{P}_1[\mathbf{F}_1 \ \ \mathbf{F}_2], \mathbf{M})$ has the form $[\mathbf{0} \ \ \mathbf{G}]$ for some $\mathbf{G} \in \mathbb{K}[x]^{m\times n_2}$; if furthermore $\mathbf{P}_2$ is a basis for $\mathcal{R}(\mathbf{M}_2, \mathbf{G})$, then $\mathbf{P}_2\mathbf{P}_1$ is a basis for $\mathcal{R}(\mathbf{M}, [\mathbf{F}_1 \ \ \mathbf{F}_2])$.*

Proof. It follows from Lemma 2.8 that the first n_1 columns of $\mathrm{Rem}(\mathbf{P}_1[\mathbf{F}_1 \quad \mathbf{F}_2], \mathbf{M})$ are $\mathrm{Rem}(\mathbf{P}_1\mathbf{F}_1, \mathbf{M}_1)$, which is zero, and that $\mathrm{Rem}([\mathbf{0} \quad \mathbf{G}], \mathbf{M}) = [\mathbf{0} \quad \mathrm{Rem}(\mathbf{G}, \mathbf{M}_2)]$. Then, the first identity in Lemma 2.2 implies both that $\mathcal{R}(\mathbf{M}, [\mathbf{0} \quad \mathbf{G}]) = \mathcal{R}(\mathbf{M}_2, \mathbf{G})$ and that the rows of $\mathbf{P}_2\mathbf{P}_1$ are in $\mathcal{R}(\mathbf{M}, [\mathbf{F}_1 \quad \mathbf{F}_2])$. Now let $\mathbf{p} \in \mathcal{R}(\mathbf{M}, [\mathbf{F}_1 \quad \mathbf{F}_2])$. Lemma 2.8 implies that $\mathbf{p} \in \mathcal{R}(\mathbf{M}_1, \mathbf{F}_1)$, hence $\mathbf{p} = \boldsymbol{\lambda}\mathbf{P}_1$ for some $\boldsymbol{\lambda}$. Then, the first identity in Lemma 2.2 shows that $\mathbf{0} = \mathrm{Rem}(\boldsymbol{\lambda}\mathbf{P}_1[\mathbf{F}_1 \quad \mathbf{F}_2], \mathbf{M}) = \mathrm{Rem}(\boldsymbol{\lambda}[\mathbf{0} \quad \mathbf{G}], \mathbf{M})$, and therefore $\boldsymbol{\lambda} \in \mathcal{R}(\mathbf{M}_2, \mathbf{G})$. Thus $\boldsymbol{\lambda} = \boldsymbol{\mu}\mathbf{P}_2$ for some $\boldsymbol{\mu}$, and $\mathbf{p} = \boldsymbol{\mu}\mathbf{P}_2\mathbf{P}_1$. □

3 COMPUTING MODULAR PRODUCTS

In this section, we aim at designing a fast algorithm for the modular products that arise in our relation basis algorithm.

3.1 Fast division with remainder

For univariate polynomials, fast Euclidean division can be achieved by first computing the reversed quotient via Newton iteration, and then deducing the remainder [14, Chap. 9]. This directly translates into the context of polynomial matrices, as was noted for example in the proof of [15, Lem. 3.4] or in [36, Chap. 10].

In the latter reference, it is showed how to efficiently compute remainders $\mathrm{Rem}(\mathcal{E}, \mathbf{M})$ for a matrix $\mathcal{E}$ as in Eq. (1) below; this is not general enough for our purpose. Algorithms for the general case have been studied [6, 11, 33–35], but we are not aware of any that achieves the speed we desire. Thus, as a preliminary to the computation of residuals in Section 3.2, we now detail this extension of fast polynomial division to fast polynomial matrix division.

As mentioned above, we will start by computing the quotient. The degrees of its entries are controlled thanks to the reducedness of the divisor, which ensures that no high-degree cancellation can occur when multiplying the quotient and the divisor.

Lemma 3.1. *Let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$, $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and $\delta \in \mathbb{Z}_{>0}$ be such that $\mathbf{M}$ is column reduced and $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M}) + (\delta, \ldots, \delta)$. Then, $\deg(\mathrm{Quo}(\mathbf{F}, \mathbf{M})) < \delta$.*

Proof. First, $\mathrm{lm}_{\mathbf{0}}(\mathbf{M}^{\mathsf{T}})^{\mathsf{T}} = \mathrm{lm}_{-\mathbf{d}}(\mathbf{M})$ where $\mathbf{d} = \mathrm{cdeg}(\mathbf{M}) \in \mathbb{Z}^n_{\geqslant 0}$: the $\mathbf{0}$-column leading matrix of $\mathbf{M}$ is equal to its $-\mathbf{d}$-row leading matrix. Since $\mathbf{M}$ is $\mathbf{0}$-column reduced, it is also $-\mathbf{d}$-row reduced.

Thus, by the predictable degree property [21, Thm. 6.3-13] and since since $\mathrm{rdeg}_{-\mathbf{d}}(\mathbf{M}) = \mathbf{0}$, we have $\mathrm{rdeg}_{-\mathbf{d}}(\mathbf{QM}) = \mathrm{rdeg}_{\mathbf{0}}(\mathbf{Q})$. Here, we write $\mathbf{Q} = \mathrm{Quo}(\mathbf{F}, \mathbf{M})$ and $\mathbf{R} = \mathrm{Rem}(\mathbf{F}, \mathbf{M})$.

Now, our assumption $\mathrm{cdeg}(\mathbf{F}) < \mathbf{d} + (\delta, \ldots, \delta)$ and the fact that $\mathrm{cdeg}(\mathbf{R}) < \mathbf{d}$ imply that $\mathrm{cdeg}(\mathbf{F} - \mathbf{R}) < \mathbf{d} + (d, \ldots, d)$, and thus $\mathrm{rdeg}_{-\mathbf{d}}(\mathbf{F} - \mathbf{R}) < (\delta, \ldots, \delta)$. Since $\mathbf{F} - \mathbf{R} = \mathbf{QM}$, from the previous paragraph we obtain $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{Q}) < (\delta, \ldots, \delta)$, hence $\deg(\mathbf{Q}) < \delta$. □

Corollary 3.2. *Let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ and $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ be such that $\mathbf{M}$ is column reduced and $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M})$, and let $\mathbf{P} \in \mathbb{K}[x]^{k\times m}$. Then, $\mathrm{rdeg}(\mathrm{Quo}(\mathbf{PF}, \mathbf{M})) < \mathrm{rdeg}(\mathbf{P})$.*

Proof. For the case $k = 1$, the inequality follows from Lemma 3.1 since $\mathrm{cdeg}(\mathbf{PF}) \leqslant (\delta, \ldots, \delta) + \mathrm{cdeg}(\mathbf{F}) < (\delta, \ldots, \delta) + \mathrm{cdeg}(\mathbf{M})$, where $\delta = \deg(\mathbf{P})$. Then, the general case $k \in \mathbb{Z}_{>0}$ follows by considering separately each row of $\mathbf{P}$. □

Going back to the division $\mathbf{F} = \mathbf{QM} + \mathbf{R}$, to obtain the reversed quotient we will right-multiply the reversed $\mathbf{F}$ by an expansion of the inverse of the reversed $\mathbf{M}$. This operation is performed efficiently by means of high-order lifting; we will use the next result.

Lemma 3.3. *Let $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ with $\mathbf{M}(0)$ nonsingular, and let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$. Then, defining $d = \lceil |\mathrm{cdeg}(\mathbf{M})|/n \rceil$, the truncated x-adic expansion $\mathbf{FM}^{-1} \bmod x^{kd}$ can be computed deterministically using $O^{\sim}(\lceil mk/n \rceil n^{\omega} d)$ operations in $\mathbb{K}$.*

Proof. This is a minor extension of [29, Prop. 15], incorporating the average column degree of the matrix $\mathbf{M}$ instead of the largest degree of its entries. This can be done by means of partial column linearization [17, Sec. 6], as follows. One first expands the high-degree columns of $\mathbf{M}$ and inserts elementary rows to obtain a matrix $\overline{\mathbf{M}} \in \mathbb{K}[x]^{\overline{n}\times\overline{n}}$ such that $n \leqslant \overline{n} < 2n$, $\deg(\overline{\mathbf{M}}) \leqslant d$, and $\mathbf{M}^{-1}$ is the $n \times n$ principal leading submatrix of $\overline{\mathbf{M}}^{-1}$ [17, Thm. 10 and Cor. 2]. Then, defining $\overline{\mathbf{F}} = [\mathbf{F} \quad \mathbf{0}] \in \mathbb{K}[x]^{m\times\overline{n}}$, we have that $\mathbf{FM}^{-1}$ is the submatrix of $\overline{\mathbf{F}}\,\overline{\mathbf{M}}^{-1}$ formed by its first n columns. Thus, the sought truncated expansion is obtained by computing $\overline{\mathbf{F}}\,\overline{\mathbf{M}}^{-1} \bmod x^{kd}$, which is done efficiently by [29, Alg. 4] with the choice $X = x^d$; this is valid since this polynomial is coprime to $\det(\overline{\mathbf{M}}) = \det(\mathbf{M})$ and its degree is at least the degree of $\overline{\mathbf{M}}$. □

Algorithm 1: PM-QuoRem
Input:
- $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ column reduced,
- $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$,
- $\delta \in \mathbb{Z}_{>0}$ such that $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M}) + (\delta, \ldots, \delta)$.

Output: the quotient $\mathrm{Quo}(\mathbf{F}, \mathbf{M})$, the remainder $\mathrm{Rem}(\mathbf{F}, \mathbf{M})$.

1. `/* reverse order of coefficients */`
   $(d_1, \ldots, d_n) \leftarrow \mathrm{cdeg}(\mathbf{M})$
   $\mathbf{M}_{\mathrm{rev}} = \mathbf{M}(x^{-1}) \,\mathrm{diag}(x^{d_1}, \ldots, x^{d_n})$
   $\mathbf{F}_{\mathrm{rev}} = \mathbf{F}(x^{-1}) \,\mathrm{diag}(x^{\delta+d_1-1}, \ldots, x^{\delta+d_n-1})$
2. `/* compute quotient via expansion */`
   $\mathbf{Q}_{\mathrm{rev}} \leftarrow \mathbf{F}_{\mathrm{rev}}\mathbf{M}_{\mathrm{rev}}^{-1} \bmod x^{\delta}$
   $\mathbf{Q} \leftarrow x^{\delta-1}\mathbf{Q}_{\mathrm{rev}}(x^{-1})$
3. *Return* $(\mathbf{Q}, \mathbf{F} - \mathbf{QM})$

Proposition 3.4. *Algorithm 1 is correct. Assuming that both $m\delta$ and n are in $O(D)$, where $D = |\mathrm{cdeg}(\mathbf{M})|$, this algorithm uses $O^{\sim}(\lceil m/n \rceil n^{\omega-1} D)$ operations in $\mathbb{K}$.*

Proof. Let $\mathbf{Q} = \mathrm{Quo}(\mathbf{F}, \mathbf{M})$, $\mathbf{R} = \mathrm{Rem}(\mathbf{F}, \mathbf{M})$, and $(d_1, \ldots, d_n) = \mathrm{cdeg}(\mathbf{M})$. We have the bounds $\mathrm{cdeg}(\mathbf{F}) < (\delta + d_1, \ldots, \delta + d_n)$, $\mathrm{cdeg}(\mathbf{R}) < (d_1, \ldots, d_n)$, and Lemma 3.1 gives $\deg(\mathbf{Q}) < \delta$. Thus, we can define the reversals of these polynomial matrices as

$$\begin{aligned}
\mathbf{M}_{\mathrm{rev}} &= \mathbf{M}(x^{-1}) \,\mathrm{diag}(x^{d_1}, \ldots, x^{d_n}),\\
\mathbf{F}_{\mathrm{rev}} &= \mathbf{F}(x^{-1}) \,\mathrm{diag}(x^{\delta+d_1-1}, \ldots, x^{\delta+d_n-1}),\\
\mathbf{Q}_{\mathrm{rev}} &= x^{\delta-1}\mathbf{Q}(x^{-1}),\\
\mathbf{R}_{\mathrm{rev}} &= \mathbf{R}(x^{-1}) \,\mathrm{diag}(x^{d_1-1}, \ldots, x^{d_n-1}),
\end{aligned}$$

for which the same degree bounds hold. Then, right-multiplying both sides of the identity $\mathbf{F}(x^{-1}) = \mathbf{Q}(x^{-1})\mathbf{M}(x^{-1}) + \mathbf{R}(x^{-1})$ by $\mathrm{diag}(x^{\delta+d_1-1}, \ldots, x^{\delta+d_n-1})$, we obtain $\mathbf{F}_{\mathrm{rev}} = \mathbf{Q}_{\mathrm{rev}}\mathbf{M}_{\mathrm{rev}} + x^{\delta}\mathbf{R}_{\mathrm{rev}}$.

Now, note that the constant term $\mathbf{M}_{\mathrm{rev}}(0) \in \mathbb{K}^{n\times n}$ is equal to the column leading matrix of $\mathbf{M}$, which is invertible since $\mathbf{M}$ is column

reduced, hence $\mathbf{M}_{\text{rev}}$ is invertible (over the fractions). Thus, since $\deg(\mathbf{Q}_{\text{rev}}) < \delta$, this reversed quotient matrix can be determined as the truncated expansion $\mathbf{Q}_{\text{rev}} = \mathbf{F}_{\text{rev}}\mathbf{M}_{\text{rev}}^{-1} \bmod x^\delta$. This proves the correctness of the algorithm.

Concerning the cost bound, Step **2** uses $O^{\tilde{}}(\lceil (m\delta)/(nd)\rceil n^\omega d)$ operations according to Lemma 3.3, where $d = \lceil D/n\rceil$. We have by assumption $d \in \Theta(D/n)$ as well as $m\delta/(nd) \in O(1)$, so that this cost bound is in $O^{\tilde{}}(n^{\omega-1}D)$.

In Step **3**, we multiply the $m \times n$ matrix $\mathbf{Q}$ of degree less than δ with the $n \times n$ matrix $\mathbf{M}$ such that $|\text{cdeg}(\mathbf{M})| = D$. First consider the case $m \leqslant n$. To perform this product efficiently, we expand the rows of $\mathbf{Q}$ so as to obtain a $O(n) \times n$ matrix $\overline{\mathbf{Q}}$ of degree in $O(\lceil m\delta/n\rceil)$ and such that $\mathbf{QM}$ is easily retrieved from $\overline{\mathbf{Q}}\mathbf{M}$ (see Section 3.2 for more details about how such row expansions are carried out). Thus, this product is done in $O^{\tilde{}}(n^{\omega-1}D)$, since $\lceil m\delta/n\rceil \in O(D/n)$. On the other hand, if $m > n$, we have $\delta \in O(D/m) \subseteq O(D/n)$. Then, we can compute the product $\mathbf{QM}$ via $\lceil m/n\rceil$ products of $n \times n$ matrices of degree $O(D/n)$, which cost each $O^{\tilde{}}(n^{\omega-1}D)$ operations; hence the total cost $O^{\tilde{}}(mn^{\omega-2}D)$ when $m > n$. □

3.2 Fast residual computation

Here, we focus on performing modular products $\text{Rem}(\mathbf{PF}, \mathbf{M})$, where $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ and $\mathbf{P} \in \mathbb{K}[x]^{m\times m}$ are such that $\text{cdeg}(\mathbf{F}) < \text{cdeg}(\mathbf{M})$ and $|\text{cdeg}(\mathbf{P})| \leqslant |\text{cdeg}(\mathbf{M})|$, and $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ is column reduced. The difficulty in designing a fast algorithm for this operation comes from the non-uniformity of $\text{cdeg}(\mathbf{P})$: in particular, the product $\mathbf{PF}$ cannot be computed within the target cost bound.

To start with, we use the same strategy as in [19, 26]: we make the column degrees of $\mathbf{P}$ uniform, at the price of introducing another, simpler matrix $\mathcal{E}$ for which we want to compute $\text{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$.

Let $(\delta_1, \ldots, \delta_m) = \text{cdeg}(\mathbf{P})$, $\delta = \lceil(\delta_1 + \cdots + \delta_m)/m\rceil \geqslant 1$, 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 $\delta_i = 0$. Then, let $\overline{m} = \alpha_1 + \cdots + \alpha_m$, and define $\mathcal{E} \in \mathbb{K}[x]^{\overline{m}\times m}$ as the transpose of

$$\mathcal{E}^{\mathsf{T}} = \begin{bmatrix} 1 & x^\delta & \cdots & x^{(\alpha_1-1)\delta} & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & x^\delta & \cdots & x^{(\alpha_m-1)\delta} \end{bmatrix}. \quad (1)$$

Define also the expanded column degrees $\overline{\boldsymbol{\delta}} \in \mathbb{Z}_{\geqslant 0}^{\overline{m}}$ as

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

Then, we expand the columns of $\mathbf{P}$ by considering $\overline{\mathbf{P}} \in \mathbb{K}[x]^{m\times\overline{m}}$ such that $\mathbf{P} = \overline{\mathbf{P}}\mathcal{E}$ and $\deg(\overline{\mathbf{P}}) \leqslant \delta$. (Note that $\overline{\mathbf{P}}$ can be made unique by specifying more constraints on $\text{cdeg}(\overline{\mathbf{P}})$.) The aim of this construction is that the dimension is at most doubled while the degree of the expanded matrix becomes the average column degree of $\mathbf{P}$. Precisely, $m \leqslant \overline{m} < 2m$ and $\max(\overline{\boldsymbol{\delta}}) = \delta = \lceil|\text{cdeg}(\mathbf{P})|/m\rceil$.

Now, we have $\text{Rem}(\mathbf{PF}, \mathbf{M}) = \text{Rem}(\overline{\mathbf{P}}\mathcal{E}\mathbf{F}, \mathbf{M}) = \text{Rem}(\overline{\mathbf{P}}\,\overline{\mathbf{F}}, \mathbf{M})$ by Lemma 2.2, where $\overline{\mathbf{F}} = \text{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$. Thus, $\text{Rem}(\mathbf{PF}, \mathbf{M})$ can be obtained by computing first $\overline{\mathbf{F}}$ and then $\text{Rem}(\overline{\mathbf{P}}\,\overline{\mathbf{F}}, \mathbf{M})$. For the latter, since $\overline{\mathbf{P}}$ has small degree, one can compute the product and then perform the division (Steps **3** and **4** of Algorithm 3). Step **2** of Algorithm 3 efficiently computes $\overline{\mathbf{F}}$, relying on Algorithm 2.

Algorithm 2: RemOfShifts
Input:
- $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ column reduced,
- $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ such that $\text{cdeg}(\mathbf{F}) < \text{cdeg}(\mathbf{M})$,
- $\delta \in \mathbb{Z}_{>0}$ and $k \in \mathbb{Z}_{\geqslant 0}$.

Output: the list of remainders $(\text{Rem}(x^{r\delta}\mathbf{F}, \mathbf{M}))_{0\leqslant r<2^k}$.

1. *If* $k = 0$ *then Return* $\mathbf{F}$
2. *Else*
a. $(*, \mathbf{G}) \leftarrow$ PM-QuoRem$(\mathbf{M}, x^{2^{k-1}\delta}\mathbf{F}, 2^{k-1}\delta)$
b. $\left(\begin{bmatrix}\mathbf{R}_{r0}\\ \mathbf{R}_{r1}\end{bmatrix}\right)_{0\leqslant r<2^{k-1}} \leftarrow$ RemOfShifts$\left(\mathbf{M}, \begin{bmatrix}\mathbf{F}\\ \mathbf{G}\end{bmatrix}, \delta, k-1\right)$
c. *Return* $(\mathbf{R}_{r0})_{0\leqslant r<2^{k-1}} \cup (\mathbf{R}_{r1})_{0\leqslant r<2^{k-1}}$

Proposition 3.5. *Algorithm 2 is correct. Assuming that both* $2^k m\delta$ *and* n *are in* $O(D)$*, where* $D = |\text{cdeg}(\mathbf{M})|$*, this algorithm uses* $O^{\tilde{}}((2^k mn^{\omega-2} + kn^{\omega-1})D)$ *operations in* $\mathbb{K}$.

Proof. The correctness is a consequence of the two properties in Lemma 2.2. Now, if $2^k m\delta$ and n are in $O(D)$, the assumptions in Proposition 3.4 about the input parameters for PM-QuoRem are always satisfied in recursive calls, since the row dimension m is doubled while the exponent $2^k\delta$ is halved. From the same proposition, we deduce the cost bound $O^{\tilde{}}((\sum_{0\leqslant r\leqslant k-1}\lceil 2^r m/n\rceil)n^{\omega-1}D)$. □

Algorithm 3: Residual
Input:
- $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ column reduced,
- $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ such that $\text{cdeg}(\mathbf{F}) < \text{cdeg}(\mathbf{M})$,
- $\mathbf{P} \in \mathbb{K}[x]^{m\times m}$.

Output: the remainder $\text{Rem}(\mathbf{PF}, \mathbf{M})$.

1. `/* expand high-degree columns of P */`
$(\delta_i)_{1\leqslant i\leqslant m} \leftarrow \text{cdeg}(\mathbf{P})$
$\delta \leftarrow \lceil(\delta_1 + \cdots + \delta_m)/m\rceil$
$\alpha_i \leftarrow \max(1, \lceil\delta_i/\delta\rceil)$ for $1 \leqslant i \leqslant m$
$\overline{m} \leftarrow \alpha_1 + \cdots + \alpha_m$
$\overline{\mathbf{P}} \in \mathbb{K}[x]^{\overline{m}\times m} \leftarrow$ matrix such that $\mathbf{P} = \overline{\mathbf{P}}\mathcal{E}$ and $\deg(\overline{\mathbf{P}}) \leqslant \delta$ for $\mathcal{E}$ as in Eq. (1)
2. `/* compute` $\overline{\mathbf{F}} = \text{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$ `*/`
For $1 \leqslant i \leqslant m$ such that $\alpha_i = 1$ *do*
$\overline{\mathbf{F}}_i \in \mathbb{K}[x]^{\alpha_i\times n} \leftarrow$ row i of $\mathbf{F}$
For $1 \leqslant k \leqslant \lceil\log_2(\max_i(\alpha_i))\rceil$ *do*
$(i_1, \ldots, i_\ell) \leftarrow \{i \in \{1, \ldots, m\} \mid 2^{k-1} < \alpha_i \leqslant 2^k\}$
$\mathbf{G} \leftarrow$ submatrix of $\mathbf{F}$ formed by its rows $i_1, \ldots, i_\ell$
$(\mathbf{R}_r)_{0\leqslant r<2^k} \leftarrow$ RemOfShifts$(\mathbf{M}, \mathbf{G}, \delta, k)$
For $1 \leqslant j \leqslant \ell$ *do*
$\overline{\mathbf{F}}_{i_j} \in \mathbb{K}[x]^{\alpha_{i_j}\times n} \leftarrow$ stack the rows j of $(\mathbf{R}_r)_{0\leqslant r<\alpha_{i_j}}$
$\overline{\mathbf{F}} \leftarrow \begin{bmatrix}\overline{\mathbf{F}}_1^{\mathsf{T}} & \cdots & \overline{\mathbf{F}}_m^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}} \in \mathbb{K}[x]^{\overline{m}\times n}$
3. `/* left-multiply by the expanded P */`
$\mathbf{G} \leftarrow \overline{\mathbf{P}}\,\overline{\mathbf{F}}$
4. `/* complete the remainder computation */`
$(*, \mathbf{R}) \leftarrow$ PM-QuoRem$(\mathbf{M}, \mathbf{G}, \delta)$
Return $\mathbf{R}$

Proposition 3.6. *Algorithm 3 is correct. Assuming that all of* $|\mathrm{cdeg}(\mathbf{P})|$, m, *and* n *are in* $O(D)$, *where* $D = |\mathrm{cdeg}(\mathbf{M})|$, *this algorithm uses* $O\tilde{}((m^{\omega-1} + n^{\omega-1})D)$ *operations in* $\mathbb{K}$.

Proof. Let us consider $\mathcal{E} \in \mathbb{K}[x]^{\overline{m}\times m}$ defined as in Eq. (1) from the parameters δ and $\alpha_1, \ldots, \alpha_m$ in Step **1**. We claim that the matrix $\overline{\mathbf{F}}$ computed at Step **2** is equal to $\mathrm{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$. Then, having $\mathrm{cdeg}(\overline{\mathbf{P}}\,\overline{\mathbf{F}}) < \mathrm{cdeg}(\mathbf{M}) + (\delta, \ldots, \delta)$, the correctness of PM-QuoRem implies $\mathbf{R} = \mathrm{Rem}(\overline{\mathbf{P}}\,\overline{\mathbf{F}}, \mathbf{M})$, which is $\mathrm{Rem}(\mathbf{PF}, \mathbf{M})$ by Lemma 2.2.

To prove our claim, it is enough to show that, for $1 \leqslant i \leqslant m$, the ith block $\overline{\mathbf{F}}_i$ of $\overline{\mathbf{F}}$ is the matrix formed by stacking the remainders involving the row i of $\mathbf{F}$, that is, $(\mathrm{Rem}(x^{r\delta}\mathbf{F}_{i,*}, \mathbf{M}))_{0\leqslant r<\alpha_i}$. This is clear from the first *For* loop if $\alpha_i = 1$. Otherwise, let $k \in \mathbb{Z}_{>0}$ be such that $2^{k-1} < \alpha_i \leqslant 2^k$. Then, at the kth iteration of the second loop, we have $i_j = i$ for some $1 \leqslant j \leqslant \ell$. Thus, the correctness of RemOfShifts implies that, for $0 \leqslant r < 2^k$, the row j of $\mathbf{R}_r$ is $\mathrm{Rem}(x^{r\delta}\mathbf{G}_{j,*}, \mathbf{M}) = \mathrm{Rem}(x^{r\delta}\mathbf{F}_{i,*}, \mathbf{M})$. Since $2^k \geqslant \alpha_i$, this contains the wanted remainders and the claim follows.

Let us show the cost bound, assuming that $|\mathrm{cdeg}(\mathbf{P})|$, m, and n are in $O(D)$. Note that this implies $m\delta \in O(D)$.

We first study the cost of the iteration k of the second loop of Step **2**. We have that $2^{k-1}\ell \leqslant \alpha_1 + \cdots + \alpha_m = \overline{m} \leqslant 2m$, the row dimension of $\mathbf{G}$ is ℓ, and $k \leqslant \lceil \log(\max_i(\alpha_i)) \rceil \in O(\log(m))$. Thus, the call to RemOfShifts costs $O\tilde{}((mn^{\omega-2} + n^{\omega-1})D)$ operations according to Proposition 3.5, and the same cost bound holds for the whole Step **2**. Concerning Step **4**, the cost bound $O\tilde{}(\lceil m/n \rceil n^{\omega-1}D)$ follows directly from Proposition 3.4.

The product at Step **3** involves the $m\times\overline{m}$ matrix $\overline{\mathbf{P}}$ whose degree is at most δ and the $\overline{m}\times n$ matrix $\overline{\mathbf{F}}$ such that $\mathrm{cdeg}(\overline{\mathbf{F}}) < \mathrm{cdeg}(\mathbf{M})$; we recall that $\overline{m} \leqslant 2m$. If $n \geqslant m$, we expand the columns of $\overline{\mathbf{F}}$ similarly to how $\overline{\mathbf{P}}$ was obtained from $\mathbf{P}$: this yields a $\overline{m} \times (\leqslant 2n)$ matrix of degree at most $\lceil D/n \rceil$, whose left-multiplication by $\overline{\mathbf{P}}$ directly yields $\overline{\mathbf{P}}\,\overline{\mathbf{F}}$ by compressing back the columns. Thus, this product is done in $O\tilde{}(m^{\omega-2}nD)$ operations since both δ and D/n are in $O(D/m)$ when $n \geqslant m$. If $m \geqslant n$, we do a similar column expansion of $\overline{\mathbf{F}}$, yet into a matrix with $O(m)$ columns and degree $O(D/m)$; thus, the product can be performed in $O\tilde{}(m^{\omega-1}D)$ operations in this case. □

4 FAST ALGORITHMS IN SPECIFIC CASES

Here, we discuss fast solutions to specific instances of Problem 1. This will be important ingredients of our main algorithm for relations modulo Hermite forms (Algorithm 5).

4.1 When the input module is an ideal

We first focus on Problem 1 when $n = 1$; this is one of the two base cases of the recursion in Algorithm 5 (Step **2**). In this case, the input matrix $\mathbf{M}$ is a nonzero polynomial $M \in \mathbb{K}[x]$. In other words, the input module is the ideal (M) of $\mathbb{K}[x]$, and we are looking for the $\mathbf{s}$-Popov basis for the set of relations between m elements of $\mathbb{K}[x]/(M)$. A fast algorithm for this task was given in [26, Sec. 2.2]; precisely, the following result is achieved by running [26, Alg. 2] on input $\mathbf{M}, \mathbf{F}, \mathbf{s}, 2D$.

Proposition 4.1. *Assuming* $n = 1$ *and* $\deg(\mathbf{F}) < D = \deg(\mathbf{M})$, *there is an algorithm which solves Problem 1 using* $O\tilde{}(m^{\omega-1}D)$ *operations in* $\mathbb{K}$.

4.2 When the s-minimal degree is known

Now, we consider Problem 1 with an additional input: the $\mathbf{s}$-minimal degree of $\mathcal{R}(\mathbf{M}, \mathbf{F})$, which is the column degree of its $\mathbf{s}$-Popov basis. This is motivated by a technique from [19] and used in Algorithm 5 to control the degrees of all the bases computed in the process. Namely, we find this $\mathbf{s}$-minimal degree recursively, and then we compute the $\mathbf{s}$-Popov relation basis using this knowledge.

The same question was tackled in [18, Sec. 3] and [26, Sec. 2.1] for a diagonal matrix $\mathbf{M}$. Here, we extend this to the case of a column reduced $\mathbf{M}$, relying in particular on the fast computation of $\mathrm{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$ designed in Section 3.2. We first extend [26, Lem. 2.1] to this more general setting (Lemma 4.2), and then we give the slightly modified version of [26, Alg. 1] (Algorithm 4).

Lemma 4.2. *Let* $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ *be column reduced, let* $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ *be such that* $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M})$, *let* $\mathbf{s} \in \mathbb{Z}^m$. *Furthermore, let* $\mathbf{P} \in \mathbb{K}[x]^{m\times m}$, *and let* $\mathbf{w} \in \mathbb{Z}^n$ *be such that* $\max(\mathbf{w}) \leqslant \min(\mathbf{s})$. *Then,* $\mathbf{P}$ *is the* $\mathbf{s}$*-Popov relation basis for* $\mathcal{R}(\mathbf{M}, \mathbf{F})$ *if and only if* $[\mathbf{P}\ \ \mathbf{Q}]$ *is the* $\mathbf{u}$*-Popov kernel basis of* $[\mathbf{F}^\mathsf{T}\ \ \mathbf{M}]^\mathsf{T}$ *for some* $\mathbf{Q} \in \mathbb{K}[x]^{m\times n}$ *and* $\mathbf{u} = (\mathbf{s}, \mathbf{w}) \in \mathbb{Z}^{m+n}$. *In this case,* $\deg(\mathbf{Q}) < \deg(\mathbf{P})$ *and* $[\mathbf{P}\ \ \mathbf{Q}]$ *has* $\mathbf{u}$*-pivot index* $(1, 2, \ldots, m)$.

Proof. Let $\mathbf{N} = [\mathbf{F}^\mathsf{T}\ \ \mathbf{M}]^\mathsf{T}$. It is easily verified that $\mathbf{P}$ is a relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{F})$ if and only if there is some $\mathbf{Q} \in \mathbb{K}[x]^{m\times n}$ such that $[\mathbf{P}\ \ \mathbf{Q}]$ is a kernel basis of $\mathbf{N}$.

Then, for any matrix $[\mathbf{P}\ \ \mathbf{Q}] \in \mathbb{K}[x]^{m\times(m+n)}$ in the kernel of $\mathbf{N}$, we have $\mathbf{PF} = -\mathbf{QM}$ and therefore Corollary 3.2 shows that $\mathrm{rdeg}(\mathbf{Q}) < \mathrm{rdeg}(\mathbf{P})$; since $\max(\mathbf{w}) \leqslant \min(\mathbf{s})$, this implies $\mathrm{rdeg}_{\mathbf{w}}(\mathbf{Q}) < \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P})$. Thus, we have $\mathrm{lm}_{\mathbf{u}}([\mathbf{P}\ \ \mathbf{Q}]) = [\mathrm{lm}_{\mathbf{s}}(\mathbf{P})\ \ \mathbf{0}]$, and therefore $\mathbf{P}$ is in $\mathbf{s}$-Popov form if and only if $[\mathbf{P}\ \ \mathbf{Q}]$ is in $\mathbf{u}$-Popov form with $\mathbf{u}$-pivot index $(1, \ldots, m)$. □

Algorithm 4: KnownDegreeRelations
Input:

- $\mathbf{M} \in \mathbb{K}[x]^{n\times n}$ column reduced,
- $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ such that $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M})$,
- $\mathbf{s} \in \mathbb{Z}^m$,
- $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m)$ the $\mathbf{s}$-minimal degree of $\mathcal{R}(\mathbf{M}, \mathbf{F})$.

Output: the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{M}, \mathbf{F})$.

1. `/* define partial linearization parameters */`
$\delta \leftarrow \lceil (\delta_1 + \cdots + \delta_m)/m \rceil$,
$\alpha_i \leftarrow \max(1, \lceil \delta_i/\delta \rceil)$ for $1 \leqslant i \leqslant m$,
$\overline{m} \leftarrow \alpha_1 + \cdots + \alpha_m$,
$\overline{\boldsymbol{\delta}} \leftarrow$ tuple as in Eq. (2)

2. `/* for` $\mathcal{E}$ `as in Eq. (1), compute` $\overline{\mathbf{F}} = \mathrm{Rem}(\mathcal{E}\mathbf{F}, \mathbf{M})$ `*/`
$\overline{\mathbf{F}} \leftarrow$ follow Step **2** of Algorithm 3 (Residual)

3. `/* compute the kernel basis */`
$\mathbf{u} \leftarrow (-\overline{\boldsymbol{\delta}}, -\delta, \ldots, -\delta) \in \mathbb{Z}^{\overline{m}+n}$
$\boldsymbol{\tau} \leftarrow (\mathrm{cdeg}(\mathbf{M}_{*,j}) + \delta + 1)_{1\leqslant j\leqslant n}$
$\overline{\mathbf{P}} \leftarrow \mathbf{u}$-Popov approximant basis for $\begin{bmatrix}\overline{\mathbf{F}}\\ \mathbf{M}\end{bmatrix}$ and orders $\boldsymbol{\tau}$

4. `/* retrieve the relation basis */`
$\mathbf{P} \leftarrow$ the principal $\overline{m}\times\overline{m}$ submatrix of $\overline{\mathbf{P}}$
Return the submatrix of $\mathbf{P}\mathcal{E}$ formed by the rows at indices $\alpha_1 + \cdots + \alpha_i$ for $1 \leqslant i \leqslant m$

Proposition 4.3. *Algorithm 4 is correct, and assuming that m and n are in $O(D)$, where $D = |\mathrm{cdeg}(\mathbf{M})|$, it uses $O^{\sim}(m^{\omega-1}D + n^{\omega}D/m)$ operations in $\mathbb{K}$.*

Proof. The correctness follows from the material in [26, Sec. 2.1] and [19, Sec. 4]. Concerning the cost bound, we first note that we have $\delta_1+\cdots+\delta_m \leqslant D$ according to Corollary 2.4. Thus, the cost analysis in Proposition 3.6 shows that Step **2** uses $O^{\sim}((mn^{\omega-2}+n^{\omega-1})D)$ operations. [19, Thm. 1.4] states that the approximant basis computation at Step **3** uses $O^{\sim}((m+n)^{\omega-1}(1+n/m)D)$ operations, since the row dimension of the input matrix is $\overline{m}+n \leqslant 2m+n$ and the sum of the orders is $|\boldsymbol{\tau}| = |\mathrm{cdeg}(\mathbf{M})| + n(\delta+1) \leqslant (1+n/m)D$. □

4.3 Solution based on fast linear algebra

Here, we detail how previous work can be used to handle a base case of the recursion in Algorithm 5 (Step **1**): when the vector space dimension $\deg(\det(\mathbf{M}))$ of the input module is small compared to the number m of input elements. Then, we rely on an interpretation of Problem 1 as a question of dense linear algebra over $\mathbb{K}$, which is solved efficiently by [20, Alg. 9]. This yields the following result.

Proposition 4.4. *Assuming that $\mathbf{M}$ is in shifted Popov form, and that $\mathrm{cdeg}(\mathbf{F}) < \mathrm{cdeg}(\mathbf{M})$, there is an algorithm which solves Problem 1 using $O^{\sim}(D^{\omega}\lceil m/D\rceil)$ operations in $\mathbb{K}$, where $D = \deg(\det(\mathbf{M}))$.*

This cost bound is $O^{\sim}(D^{\omega-1}m) \subseteq O^{\sim}(m^{\omega-1}D)$ when $D \in O(m)$. To see why relying on fast linear algebra is sufficient to obtain a fast algorithm when $D \in O(m)$, we note that this implies that the average column degree of the $\mathbf{s}$-Popov relation basis $\mathbf{P}$ is

$$|\mathrm{cdeg}(\mathbf{P})|/m = \deg(\det(\mathbf{P}))/m \leqslant D/m \in O(1).$$

For example, if $D \leqslant m$, most entries in this basis have degree 0: we are essentially dealing with matrices over $\mathbb{K}$. On the other hand, when $m \in O(D)$, this approach based on linear algebra uses $O^{\sim}(D^{\omega})$ operations, which largely exceeds our target cost.

We now describe how to translate our problem into the $\mathbb{K}$-linear algebra framework in [20]. Let $\mathcal{M}$ denote the row space of $\mathbf{M}$; we assume that $\mathbf{M}$ has no identity column. In order to compute in the quotient $\mathbb{K}[x]^n/\mathcal{M}$, which has finite dimension D, it is customary to make use of the *multiplication matrix* of x with respect to a given monomial basis. Here, since the basis $\mathbf{M}$ of $\mathcal{M}$ is in shifted Popov form with column degree $(d_1,\ldots,d_n) \in \mathbb{Z}^n_{>0}$, Lemma 2.3 suggests to use the monomial basis

$$\{(x^i,0,\ldots,0), 0 \leqslant i < d_1\} \cup \cdots \cup \{(0,\ldots,0,x^i), 0 \leqslant i < d_n\}.$$

Above, we have represented an element in $\mathbb{K}[x]^n/\mathcal{M}$ by a polynomial vector $\mathbf{f} \in \mathbb{K}[x]^{1\times n}$ such that $\mathrm{cdeg}(\mathbf{f}) < (d_1,\ldots,d_n)$. In the linear algebra viewpoint, we rather represent it by a constant vector $\mathbf{e} \in \mathbb{K}^{1\times D}$, which is formed by the concatenations of the coefficient vectors of the entries of $\mathbf{f}$. Applying this to each row of the input matrix $\mathbf{F}$ yields a constant matrix $\mathbf{E} \in \mathbb{K}^{m\times D}$, which is another representation of the same m elements in the quotient.

Besides, the multiplication matrix $\mathbf{X} \in \mathbb{K}^{D\times D}$ is the matrix such that $\mathbf{eX} \in \mathbb{K}^{1\times D}$ corresponds to the remainder in the division of $x\mathbf{f}$ by $\mathbf{M}$. Since the basis $\mathbf{M}$ is in shifted Popov form, the computation of $\mathbf{X}$ is straightforward. Indeed, writing $\mathbf{M} = \mathrm{diag}(x^{d_1},\ldots,x^{d_n}) - \mathbf{A}$ where $\mathbf{A} \in \mathbb{K}[x]^{n\times n}$ is such that $\mathrm{cdeg}(\mathbf{A}) < (d_1,\ldots,d_n)$, then

- the row $d_1+\cdots+d_{i-1}+j$ of $\mathbf{X}$ is the unit vector with 1 at index $d_1+\cdots+d_{i-1}+j+1$, for $1 \leqslant j < d_i$ and $1 \leqslant i \leqslant n$,
- the row $d_1+\cdots+d_i$ of $\mathbf{X}$ is the concatenation of the coefficient vectors of the row i of $\mathbf{A}$, for $1 \leqslant i \leqslant n$.

That is, writing $\mathbf{A} = [a_{ij}]_{1\leqslant i,j\leqslant n}$ and denoting by $\{a_{ij}^{(k)}, 0 \leqslant k < d_j\}$ the coefficients of a_{ij}, the multiplication matrix $\mathbf{X} \in \mathbb{K}^{D\times D}$ is

$$\begin{bmatrix}
 & 1 & & & & & & & \\
 & & \ddots & & & & & & \\
 & & & 1 & & & & & \\
a_{11}^{(0)} & a_{11}^{(1)} & \cdots & a_{11}^{(d_1-1)} & \cdots & a_{1n}^{(0)} & a_{1n}^{(1)} & \cdots & a_{1n}^{(d_n-1)} \\
 & & & & \ddots & & & & \\
 & & & & & & 1 & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & 1 \\
a_{n1}^{(0)} & a_{n1}^{(1)} & \cdots & a_{n1}^{(d_1-1)} & \cdots & a_{nn}^{(0)} & a_{nn}^{(1)} & \cdots & a_{nn}^{(d_n-1)}
\end{bmatrix}.$$

5 RELATIONS MODULO HERMITE FORMS

In this section, we give a fast algorithm for solving Problem 1 when $\mathbf{M}$ is in Hermite form; this matrix is denoted by $\mathbf{H}$ in what follows. The cost bound is given under the assumption that $\mathbf{H}$ has no identity column; how to reduce to this case by discarding columns of $\mathbf{H}$ and $\mathbf{F}$ was discussed in Corollary 2.6. We recall that Steps **1**, **2**, and **3.i** have been discussed in Section 4.

Proposition 5.1. *Algorithm 5 is correct and, assuming the entries $\mathrm{cdeg}(\mathbf{H})$ are positive, it uses $O^{\sim}(m^{\omega-1}D + n^{\omega}D/m)$ operations in $\mathbb{K}$, where $D = |\mathrm{cdeg}(\mathbf{H})| = \deg(\det(\mathbf{H}))$.*

Proof. Following the recursion in the algorithm, our proof is by induction on n, with two base cases (Steps **1** and **2**).

The correctness and the cost bound for Step **1** follows from the discussion in Section 4.3, as summarized in Proposition 4.4. From Section 4.1, Step **2** correctly computes the $\mathbf{s}$-Popov relation basis and uses $O^{\sim}(m^{\omega-1}D)$ operations in $\mathbb{K}$.

Now, we focus on the correctness of Step **3**, assuming that the two recursive calls at Steps **3.d** and **3.g** correctly compute the shifted Popov relation bases. Since KnownDegreeRelations is correct, it is enough to prove that the $\mathbf{s}$-minimal degree of $\mathcal{R}(\mathbf{H},\mathbf{F})$ is $\boldsymbol{\delta}_1+\boldsymbol{\delta}_2$; for this, we will show that $\mathbf{P}_2\mathbf{P}_1$ is a relation basis for $\mathcal{R}(\mathbf{H},\mathbf{F})$ whose $\mathbf{s}$-Popov form has column degree $\boldsymbol{\delta}_1+\boldsymbol{\delta}_2$.

From Theorem 2.9, $\mathbf{P}_2\mathbf{P}_1$ is a relation basis for $\mathcal{R}(\mathbf{H},\mathbf{F})$. Furthermore, the fact that the $\mathbf{s}$-Popov form of $\mathbf{P}_2\mathbf{P}_1$ has column degree $\boldsymbol{\delta}_1+\boldsymbol{\delta}_2$ follows from [19, Sec. 3], since $\mathbf{P}_1$ is in $\mathbf{s}$-Popov form and $\mathbf{P}_2$ is in $\mathbf{t}$-Popov form, where $\mathbf{t} = \mathbf{s}+\boldsymbol{\delta}_1 = \mathrm{rdeg}_{\mathbf{s}}(\mathbf{P}_1)$.

Concerning the cost of Step **3**, we remark that $m < D$, that $n \leqslant D$ is ensured by $\mathrm{cdeg}(\mathbf{H}) > \mathbf{0}$, and that $\boldsymbol{\delta}_1+\boldsymbol{\delta}_2 = \deg(\det(\mathbf{P}_2\mathbf{P}_1)) \leqslant D$ according to Corollary 2.4. Furthermore, there are two recursive calls with dimension about $n/2$, and with $\mathbf{H}_1$ and $\mathbf{H}_2$ that are in Hermite form and have determinant degrees $D_1 = \deg(\det(\mathbf{H}_1))$ and $D_2 = \deg(\det(\mathbf{H}_2))$ such that $D = D_1+D_2$. Besides, the entries of both $\mathrm{cdeg}(\mathbf{H}_1)$ and $\mathrm{cdeg}(\mathbf{H}_2)$ are all positive.

In particular, the assumptions on the parameters in Propositions 3.6 and 4.3, concerning the computation of the residual at

Step **3.f** and of the relation basis when the degrees are known at Step **3.i**, are satisfied. Thus, these steps use $O^{\sim}((m^{\omega-1}+n^{\omega-1})D)$ and $O^{\sim}(m^{\omega-1}D+n^{\omega}D/m)$ operations, respectively. The announced cost bound follows. □

Algorithm 5: RelationsModHermite
Input:

- matrix $\mathbf{H} \in \mathbb{K}[x]^{n\times n}$ in Hermite form,
- matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ such that $\operatorname{cdeg}(\mathbf{F}) < \operatorname{cdeg}(\mathbf{H})$,
- shift $\mathbf{s} \in \mathbb{Z}^m$.

Output: the $\mathbf{s}$-Popov relation basis for $\mathcal{R}(\mathbf{H},\mathbf{F})$.

1. *If* $D = |\operatorname{cdeg}(\mathbf{H})| \leqslant m$:
   a. build $\mathbf{X} \in \mathbb{K}^{D\times D}$ from $\mathbf{H}$ as in Section 4.3
   b. build $\mathbf{E} \in \mathbb{K}^{m\times D}$ from $\mathbf{F}$ as in Section 4.3
   c. $\mathbf{P} \leftarrow$ [20, Alg. 9] on input $(\mathbf{E}, \mathbf{X}, \mathbf{s}, 2^{\lceil \log_2(D)\rceil})$
   d. *Return* $\mathbf{P}$
2. *Else if* $n = 1$ *then*
   a. $\mathbf{P} \leftarrow$ [26, Alg. 2] on input $(\mathbf{H}, \mathbf{F}, \mathbf{s}, 2D)$
   b. *Return* $\mathbf{P}$
3. *Else*:
   a. $n_1 \leftarrow \lfloor n/2 \rfloor$; $n_2 \leftarrow \lceil n/2 \rceil$
   b. $\mathbf{H}_1$ and $\mathbf{H}_2 \leftarrow$ the $n_1 \times n_1$ leading and $n_2 \times n_2$ trailing principal submatrices of $\mathbf{H}$
   c. $\mathbf{F}_1 \leftarrow$ first n_1 columns of $\mathbf{F}$
   d. $\mathbf{P}_1 \leftarrow$ RelationsModHermite$(\mathbf{H}_1, \mathbf{F}_1, \mathbf{s})$
   e. $\boldsymbol{\delta}_1 \leftarrow$ diagonal degrees of $\mathbf{P}_1$
   f. $\mathbf{G} \leftarrow$ last n_2 columns of Residual$(\mathbf{H}, \mathbf{P}_1, \mathbf{F})$
   g. $\mathbf{P}_2 \leftarrow$ RelationsModHermite$(\mathbf{H}_2, \mathbf{G}, \mathbf{s}+\boldsymbol{\delta}_1)$
   h. $\boldsymbol{\delta}_2 \leftarrow$ diagonal degrees of $\mathbf{P}_2$
   i. *Return* KnownDegreeRelations$(\mathbf{H}, \mathbf{F}, \mathbf{s}, \boldsymbol{\delta}_1+\boldsymbol{\delta}_2)$

ACKNOWLEDGMENTS

The authors thank Claude-Pierre Jeannerod for interesting discussions, Arne Storjohann for his helpful comments on high-order lifting, and the reviewers whose remarks helped to prepare the final version of this paper. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 609405 (COFUNDPostdocDTU). Vu Thi Xuan acknowledges financial support provided by the scholarship *Explora Doc* from *Région Rhône-Alpes, France*, and by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program *Investissements d'Avenir* (ANR-11-IDEX-0007) operated by the French National Research Agency.

REFERENCES

[1] G. A. Baker and P. R. Graves-Morris. 1996. *Padé Approximants.* Cambridge University Press.
[2] B. Beckermann. 1992. A reliable method for computing M-Padé approximants on arbitrary staircases. *J. Comput. Appl. Math.* 40, 1 (1992), 19–42.
[3] B. Beckermann and G. Labahn. 1994. A Uniform Approach for the Fast Computation of Matrix-Type Padé Approximants. *SIAM J. Matrix Anal. Appl.* 15, 3 (July 1994), 804–823.
[4] B. Beckermann and G. Labahn. 1997. Recursiveness in matrix rational interpolation problems. *J. Comput. Appl. Math.* 77, 1–2 (1997), 5–34.
[5] B. Beckermann, G. Labahn, and G. Villard. 1999. Shifted Normal Forms of Polynomial Matrices. In *ISSAC'99.* ACM, 189–196.
[6] B. Codenotti and G. Lotti. 1989. A fast algorithm for the division of two polynomial matrices. *IEEE Trans. Automat. Control* 34, 4 (Apr 1989), 446–448.
[7] D. Coppersmith and S. Winograd. 1990. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.* 9, 3 (1990), 251–280.
[8] D. S. Dummit and R. M. Foote. 2004. *Abstract Algebra.* John Wiley & Sons.
[9] D. Eisenbud. 1995. *Commutative Algebra: with a View Toward Algebraic Geometry.* Springer-Verlag, New York.
[10] D. Eisenbud. 2005. *The Geometry of Syzygies.* Springer-Verlag, New York.
[11] P. Favati and G. Lotti. 1991. Parallel algorithms for matrix polynomial division. *Computers and Mathematics with Applications* 22, 7 (1991), 37–42.
[12] G. D. Forney, Jr. 1975. Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems. *SIAM Journal on Control* 13, 3 (1975), 493–520.
[13] F. R. Gantmacher. 1959. *The Theory of Matrices.* Chelsea.
[14] J. von zur Gathen and J. Gerhard. 2013. *Modern Computer Algebra (third edition).* Cambridge University Press. i–xiii, 1–795 pages.
[15] P. Giorgi, C.-P. Jeannerod, and G. Villard. 2003. On the complexity of polynomial matrix computations. In *ISSAC'03.* ACM, 135–142.
[16] P. Giorgi and R. Lebreton. 2014. Online Order Basis Algorithm and Its Impact on the Block Wiedemann Algorithm. In *ISSAC'14.* ACM, 202–209.
[17] S. Gupta, S. Sarkar, A. Storjohann, and J. Valeriote. 2012. Triangular x-basis decompositions and derandomization of linear algebra algorithms over $K[x]$. *J. Symbolic Comput.* 47, 4 (2012), 422–453.
[18] S. Gupta and A. Storjohann. 2011. Computing Hermite Forms of Polynomial Matrices. In *ISSAC'11.* ACM, 155–162.
[19] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. 2016. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts. In *ISSAC'16.* ACM, 295–302.
[20] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. 2017. Computing minimal interpolation bases. *J. Symbolic Comput.* 83 (2017), 272–314.
[21] T. Kailath. 1980. *Linear Systems.* Prentice-Hall.
[22] G. Labahn, V. Neiger, and W. Zhou. 2017. Fast, deterministic computation of the Hermite normal form and determinant of a polynomial matrix. *J. Complexity* (in press) (2017).
[23] F. Le Gall. 2014. Powers of Tensors and Fast Matrix Multiplication. In *ISSAC'14.* ACM, 296–303.
[24] J. Middeke. 2011. *A computational view on normal forms of matrices of Ore polynomials.* Ph.D. Dissertation. Research Institute for Symbolic Computation (RISC). http://www.risc.jku.at/publications/download/risc_4377/diss.pdf
[25] V. Neiger. 2016. *Bases of relations in one or several variables: fast algorithms and applications.* Ph.D. Dissertation. École Normale Supérieure de Lyon. https://tel.archives-ouvertes.fr/tel-01431413
[26] V. Neiger. 2016. Fast computation of shifted Popov forms of polynomial matrices via systems of modular polynomial equations. In *ISSAC'16.* ACM, 365–372.
[27] V. M. Popov. 1972. Invariant Description of Linear, Time-Invariant Controllable Systems. *SIAM Journal on Control* 10, 2 (1972), 252–264.
[28] J. Rosenkilde and A. Storjohann. 2016. Algorithms for Simultaneous Padé Approximations. In *ISSAC'16.* ACM, 405–412.
[29] A. Storjohann. 2003. High-order lifting and integrality certification. *J. Symbolic Comput.* 36, 3-4 (2003), 613–648.
[30] A. Storjohann. 2006. Notes on computing minimal approximant bases. In *Challenges in Symbolic Computation Software (Dagstuhl Seminar Proceedings).* http://drops.dagstuhl.de/opus/volltexte/2006/776
[31] M. Van Barel and A. Bultheel. 1991. The computation of non-perfect Padé-Hermite approximants. *Numer. Algorithms* 1, 3 (1991), 285–304.
[32] M. Van Barel and A. Bultheel. 1992. A general module theoretic framework for vector M-Padé and matrix rational interpolation. *Numer. Algorithms* 3 (1992), 451–462.
[33] Qing-Guo Wang and Chun-Hui Zhou. 1986. An efficient division algorithm for polynomial matrices. *IEEE Trans. Automat. Control* 31, 2 (Feb 1986), 165–166.
[34] W. Wolovich. 1984. A division algorithm for polynomial matrices. *IEEE Trans. Automat. Control* 29, 7 (Jul 1984), 656–658.
[35] Shou-Yuan Zhang and Chi-Tsong Chen. 1983. An algorithm for the division of two polynomial matrices. *IEEE Trans. Automat. Control* 28, 2 (Feb 1983), 238–240.
[36] W. Zhou. 2012. *Fast Order Basis and Kernel Basis Computation and Related Problems.* Ph.D. Dissertation. University of Waterloo.
[37] W. Zhou and G. Labahn. 2012. Efficient Algorithms for Order Basis Computation. *J. Symbolic Comput.* 47, 7 (2012), 793–819.