Deterministic computation of the characteristic polynomial in the time of matrix multiplication

Vincent Neiger, Clément Pernet

Deterministic computation of the characteristic polynomial in the time of matrix multiplication

Vincent Neiger

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

Clément Pernet

Université Grenoble Alpes, Laboratoire Jean Kuntzmann, CNRS, UMR 5224
700 avenue centrale, IMAG - CS 40700, 38058 Grenoble cedex 9, France

Abstract

This paper describes an algorithm which computes the characteristic polynomial of a matrix over a field within the same asymptotic complexity, up to constant factors, as the multiplication of two square matrices. Previously, this was only achieved by resorting to genericity assumptions or randomization techniques, while the best known complexity bound with a general deterministic algorithm was obtained by Keller-Gehrig in 1985 and involves logarithmic factors. Our algorithm computes more generally the determinant of a univariate polynomial matrix in reduced form, and relies on new subroutines for transforming shifted reduced matrices into shifted weak Popov matrices, and shifted weak Popov matrices into shifted Popov matrices.

Keywords: Characteristic polynomial, polynomial matrices, determinant, fast linear algebra.

1. Introduction

The last five decades witnessed a constant effort towards computational reductions of linear algebra problems to matrix multiplication. It has been showed that most classical problems are not harder than multiplying two square matrices, such as matrix inversion, LU decomposition, nullspace basis computation, linear system solving, rank and determinant computations, etc. [7] [25] [8, Chap. 16]. In this context, one major challenge stands out: designing a similar reduction to matrix multiplication for the computation of characteristic polynomials and related objects such as minimal polynomials and Frobenius forms. For the characteristic polynomial, significant progress was achieved by Keller-Gehrig [31], and more recently by Pernet and Storjohann [39] who solved the problem if one allows randomization. This paper closes the problem by providing a deterministic algorithm with the same asymptotic complexity as matrix multiplication.

The characteristic polynomial of a square matrix over a field $\mathbb{K}$, say $\mathbf{M} \in \mathbb{K}^{m\times m}$, is defined as $\det(x\mathbf{I}_m - \mathbf{M})$. Specific algorithms exist for sparse or structured matrices; here we consider the classical, dense case. In this paper the complexity of an algorithm is measured as an upper bound on its arithmetic cost, that is, the number of basic field operations it uses to compute the output.

Theorem 1.1. *Let $\mathbb{K}$ be a field. Using a subroutine which multiplies two matrices in $\mathbb{K}^{m\times m}$ in $O(m^\omega)$ field operations for some $\omega > 2$, the characteristic polynomial of a matrix in $\mathbb{K}^{m\times m}$ can be computed deterministically in $O(m^\omega)$ field operations.*

Outline. The rest of this introduction gives more details about our framework for complexity bounds (Section 1.1), summarizes previous work (Section 1.2), describes our contribution on polynomial matrix determinant computation (Section 1.3), gives an overview of our approach and of new tools that we designed to avoid logarithmic factors (Sections 1.4 and 1.5), and finally lists a few perspectives (Section 1.6). Section 2 introduces the notation, main definitions, and basic properties used in this paper. Then Section 3 presents the main algorithm of this paper along with a detailed complexity analysis. This algorithm uses two new technical tools described in Sections 4 and 5: the transformation of reduced forms into weak Popov forms and of weak Popov forms into Popov forms, in the case of shifted forms.

1.1. Framework for complexity bounds

In this paper, $\mathbb{K}$ is any field and we seek upper bounds on the complexity of algorithms which operate on objects such as matrices and polynomials over $\mathbb{K}$. We consider the arithmetic cost of these algorithms, i.e. the number of basic operations in $\mathbb{K}$ that are used to compute the output from some input of a given size. The basic operations are addition, subtraction, multiplication, and inversion in the field, as well as testing whether a given field element is zero.

As already highlighted in Theorem 1.1, in this paper we fix any $2 < \omega \leq 3$ as well as any algorithm which multiplies matrices in $\mathbb{K}^{m\times m}$ using $O(m^\omega)$ operations in $\mathbb{K}$: this algorithm is assumed to be the one used as a black box for all matrix multiplications arising in the algorithms we design. The current best known cost bounds ensure that any $\omega > 2.373$ is feasible [33]. In practice, one often considers a cubic algorithm with $\omega = 3$ or Strassen's algorithm with $\omega = \log_2(7)$ [47]. Our results hold with the only assumption that $2 < \omega \leq 3$.

In the computer algebra literature, this setting is classical and often implicit; we still emphasize it because here, and more generally when one studies the logarithmic factors in the cost bound of some algorithm, this clarification of how the underlying matrix multiplications are performed is of the utmost importance. Indeed, if one were allowed to use any matrix multiplication subroutine, then the question of logarithmic factors becomes void: for any exponent ω known to be feasible at the time of writing, it is known that $\omega - \varepsilon$ is feasible as well for a sufficiently small $\varepsilon > 0$; then one might rather rely on this faster subroutine, and apply Keller-Gehrig's algorithm to obtain the characteristic polynomial in $O(m^{\omega-\varepsilon}\log(m))$ operations in $\mathbb{K}$, which is in $O(m^\omega)$.

Similarly, we consider a nondecreasing function $d \mapsto \mathsf{M}(d)$ and an algorithm which multiplies two polynomials in $\mathbb{K}[x]$ of degree at most d using at most $\mathsf{M}(d)$ operations in $\mathbb{K}$; our algorithms rely on this subroutine for polynomial multiplication. Here d is any nonnegative real number; it will often be a fraction D/m of positive integers; we assume that $\mathsf{M}(d) = 1$ for $0 \leq d < 1$, so that $\mathsf{M}(d) \geq 1$ for all $d \geq 0$. To help derive complexity upper bounds, we also consider the following assumptions $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_\omega$.

$\mathcal{H}_{\mathrm{sl}}$: $2\,\mathsf{M}(d) \leq \mathsf{M}(2d)$ for all $d \geq 1$ (superlinearity).
$\mathcal{H}_{\mathrm{sm}}$: $\mathsf{M}(d_1 d_2) \leq \mathsf{M}(d_1)\,\mathsf{M}(d_2)$ for all $d_1, d_2 \geq 0$ (submultiplicativity).
$\mathcal{H}_\omega$: $\mathsf{M}(d) \in O(d^{\omega-1-\epsilon})$ for some $\epsilon > 0$.

The first assumption is customary, see e.g. [19, Sec. 8.3]; note that it implies $\mathsf{M}(d) \geq d$ for all $d \geq 1$. The second and last assumptions are commonly made in complexity analyses for divide and conquer algorithms on polynomial matrices [45, 23]: we refer to [45, Sec. 2] for further comments on these assumptions. They are satisfied by the cost bounds of polynomial multiplication algorithms such as the quasi-linear algorithm of Cantor and Kaltofen [9] and, for suitable fields $\mathbb{K}$, the quasi-linear algorithm of Harvey and van der Hoeven and Lecerf [24], and most of Toom-Cook subquadratic algorithms [50, 10]. For the latter only $\mathcal{H}_\omega$ might not

be satisfied, depending on ω and on the number of points used. Note that with the current estimates having $\omega > 2.373$, an order 5 Toom-Cook multiplication (requiring a field with at least 9 points) has exponent $\log(9)/\log(5) \approx 1.365 < \omega - 1$; thus for such exponents ω all Toom-Cook algorithms of order 5 or more satisfy all the above assumptions.

Following [45, 23], we also define a function $d \mapsto \mathsf{M}'(d)$ related to the cost of divide and conquer methods such as the half-gcd algorithm: $\mathsf{M}'(d) = \sum_{0 \le i \le \lceil \log_2(d) \rceil} 2^i \, \mathsf{M}(2^{-i}d)$ for $d \ge 1$, and $\mathsf{M}'(d) = 1$ for $0 \le d \le 1$. By definition one has $\mathsf{M}'(d) \ge \mathsf{M}(d) \ge 1$ for all $d \ge 0$, and the identity $\mathsf{M}'(2d) = 2\,\mathsf{M}'(d) + \mathsf{M}(2d)$ for $d \ge 1$ ensures that $\mathsf{M}'(d)$ is superlinear: $2\,\mathsf{M}'(d) \le \mathsf{M}'(2d)$ for all $d \ge 1$. Assuming $\mathcal{H}_{\mathrm{sl}}$ yields the asymptotic bound $\mathsf{M}'(d) \in O(\mathsf{M}(d)\log(d))$ where the $\log(d)$ factor only occurs if a quasi-linear polynomial multiplication is used; in particular, $\mathcal{H}_{\mathrm{sl}}$ and $\mathcal{H}_{\omega}$ imply $\mathsf{M}'(d) \in O(d^{\omega-1-\varepsilon})$ for some $\varepsilon > 0$. Furthermore, if one assumes $\mathcal{H}_{\mathrm{sl}}$ and $\mathcal{H}_{\mathrm{sm}}$, then $\mathsf{M}'(\cdot)$ is submultiplicative as well: $\mathsf{M}'(d_1 d_2) \le \mathsf{M}'(d_1)\,\mathsf{M}'(d_2)$ for all $d_1, d_2 \ge 0$.

In what follows we assume that two polynomial matrices in $\mathbb{K}[x]^{m\times m}$ of degree at most $d \ge 0$ can be multiplied in $O(m^\omega \, \mathsf{M}(d))$ operations in $\mathbb{K}$. This is a very mild assumption: it holds as soon as $\mathsf{M}(d)$ corresponds to one of the above-mentioned polynomial multiplication algorithms, and it also holds if the chosen matrix multiplication algorithm defining ω supports matrices over a commutative ring using only the operations $\{+, -, \times\}$ (so that one can use it to multiply $m \times m$ matrices over $\mathbb{K}[x]/(x^{2d+1})$). Note still that this bound $O(m^\omega \, \mathsf{M}(d))$ is slightly worse than the best known ones [9, 24]; for example, Cantor and Kaltofen's algorithm performs polynomial matrix multiplication in $O(m^\omega d \log(d) + m^2 d \log(d)\log(\log(d)))$ field operations, which is finer than the bound $O(m^\omega \, \mathsf{M}(d))$ with $\mathsf{M}(d) = \Theta(d\log(d)\log(\log(d)))$ in that case. This simplification is frequent in the polynomial matrix literature, and it is made here for the sake of presentation, to improve the clarity of our main complexity results and of the analyses that lead to them.

1.2. Previous work

Previous algorithms based on linear algebra over $\mathbb{K}$ for computing the characteristic polynomial of $\mathbf{M} \in \mathbb{K}^{m\times m}$ mainly fall in three types of methods.

Traces of powers: combining the traces of the first n powers of the input matrix using the Newton identities reveals the coefficients of the characteristic polynomial. Known as the Faddeev-LeVerrier algorithm, it was introduced in [34], refined and rediscovered in [43, 16, 18], and used in [11] to prove that the problem is in the $\mathcal{NC}^2$ parallel complexity class.

Determinant expansion formula: introduced in [41] and improved in [6], this approach does not involve division, and is therefore well suited for computing over integral domains. Later developments in this field include [1, 30], the latter reaching the best known cost bound of $O(m^{2.6973}\log(m)^c)$ ring operations using a deterministic algorithm, for some constant $c > 0$.

Krylov methods: based on sequences of iterates of vectors under the application of the matrix: $(\mathbf{v}, \mathbf{M}\mathbf{v}, \mathbf{M}^2\mathbf{v}, \ldots)$. These methods rely on the fact that the first linear dependency between these iterates defines a polynomial which divides the characteristic polynomial. Some algorithms construct the Krylov basis explicitly [31, 20, 14], while others can be interpreted as an implicit Krylov iteration with structured vectors [12, 39].

Methods based on traces of powers use $O(m^4)$ or $O(m^{\omega+1})$ field operations, and are mostly competitive for their parallel complexity. Methods based on determinant expansions use $O(m^4)$ or $O(m^{\omega+1})$ field operations and are relevant for division-free algorithms. Lastly, the Krylov methods run in $O(m^3)$ [12, 14] or $O(m^\omega \log m)$ [31] field operations with deterministic algorithms, or in $O(m^\omega)$ field operations with the Las Vegas randomized algorithm in [39].

Note that the characteristic polynomial of $\mathbf{M}$ cannot be computed faster than the determinant of $\mathbf{M}$, since the latter is the constant coefficient of the former. Furthermore, under the model of computation trees, the determinant of $m \times m$ matrices cannot be computed faster than the product of two $m \times m$ matrices [8, Sec. 16.4], a consequence of Baur and Strassen's theorem [2].

Another type of characteristic polynomial algorithms is based on operations on matrices over $\mathbb{K}[x]$, called polynomial matrices in what follows. Indeed the characteristic polynomial may be obtained by calling a determinant algorithm on the characteristic matrix $x\mathbf{I}_m - \mathbf{M}$, which is in $\mathbb{K}[x]^{m\times m}$. Existing algorithms, which accept any matrix in $\mathbb{K}[x]^{m\times m}$ of degree d as input, include

- the evaluation-interpolation method, which costs $O(m^{\omega+1}d + m^3\,\mathsf{M}'(d))$ field operations, requires that the field $\mathbb{K}$ is large enough, and mainly relies on the computation of about md determinants of matrices in $\mathbb{K}^{m\times m}$;
- the algorithm of Mulders and Storjohann [36] based on weak Popov form computation, which uses $O(m^3d^2)$ field operations;
- retrieving the determinant as the product of the diagonal entries of the Smith form, itself computed by a Las Vegas randomized algorithm in $O(m^\omega\,\mathsf{M}'(d)\log(m)^2)$ field operations [45, Prop. 41], assuming $\mathsf{M}'(d) \in O(d^{\omega-1})$;
- the algorithm based on unimodular triangularization in [32], which is deterministic and uses $O(m^\omega d\log(d)^a\log(m)^b)$ field operations for some constants $a, b \in \mathbb{Z}_{>0}$.

In the last two items the cost bound is, up to logarithmic factors, the same as the cost of multiplying matrices $\mathbb{K}[x]^{m\times m}$ of degree d by relying on both fast linear algebra over $\mathbb{K}$ and fast arithmetic in $\mathbb{K}[x]$, as showed in [9]. The last two of these cost bounds do involve factors logarithmic in m, whereas the first two have an exponent on m which exceeds ω.

In summary, the fastest characteristic polynomial algorithms either are randomized or have a cost a logarithmic factor away from the lower bound. This paper, with Theorem 1.1, bridges this gap by proposing the first deterministic algorithm with cost $O(m^\omega)$.

1.3. A more general result: determinant of reduced polynomial matrices

Our algorithm falls within the category of polynomial matrix determinant computation. Yet unlike the above-listed approaches ours is tailored to a specific family of polynomial matrices, which contains the characteristic matrix $x\mathbf{I}_m - \mathbf{M}$: the family of *row reduced matrices* [52, 29]. Restricting to such matrices provides us with good control of the degrees in computations; as a typical example, it is easy to predict the degree of a vector-matrix product $\mathbf{v}(x\mathbf{I}_m - \mathbf{M})$ by observing the degrees in $\mathbf{v}$, without actually computing the product. As we explain below, this degree control allows us to avoid searches of degree profiles, which would add logarithmic terms to the cost. Although the characteristic matrix has other properties besides row reducedness (it has degree 1, and is in Popov form [40] hence column reduced), we do not exploit them.

When appropriate, the average row degree D/m, where D is the sum of the degrees of the rows of the matrix, is chosen as a measure of the input degree which refines the matrix degree d used above. This gives cost bounds more sensitive to the input degrees and also, most importantly, leverages the fact that even if the algorithm starts from a matrix with uniform degrees such as $x\mathbf{I}_m - \mathbf{M}$, it may end up handling matrices with unbalanced row degrees in the process.

Theorem 1.2. *Assuming $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_\omega$ (hence in particular $\omega > 2$), there is an algorithm which takes as input a row reduced matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times m}$ and computes its determinant using*

$$O\left(m^\omega\,\mathsf{M}'(D/m)\right) \subseteq O\left(m^\omega\,\mathsf{M}'(\deg(\mathbf{A}))\right)$$

operations in $\mathbb{K}$, where $D = \deg(\det(\mathbf{A}))$ is equal to the sum of the degrees of the rows of $\mathbf{A}$.

The fact that $\deg(\det(\mathbf{A}))$ is the sum of row degrees is a consequence of row reducedness [29], and the cost bound inclusion follows from $\deg(\det(\mathbf{A})) \leq m \deg(\mathbf{A})$. Taking $\mathbf{A} = x\mathbf{I}_m - \mathbf{M}$ for $\mathbf{M} \in \mathbb{K}^{m\times m}$, Theorem 1.1 is a direct corollary of Theorem 1.2. The only assumption needed in Theorem 1.1 is $\omega > 2$, since it implies the existence of a polynomial multiplication algorithm such that $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_{\omega}$ hold, such as Cantor and Kaltofen's algorithm [9].

Previous polynomial matrix determinant algorithms with costs of the order of $m^{\omega} \deg(\mathbf{A})$, up to logarithmic factors, have been listed above: a randomized one from [45], and a deterministic one from [32]. To our knowledge, this paper gives the first description of an algorithm achieving such a cost involving no factor logarithmic in m. Our approach partially follows the algorithm of [32], but also substantially differs from it in a way that allows us to benefit from the reducedness of $\mathbf{A}$. The cost bound $O(m^{\omega}\,\mathsf{M}'(\deg(\mathbf{A})))$ has been obtained before in [21, Sec. 4.2.2] in the particular case of a "sufficiently generic[1]" matrix $\mathbf{A}$. In that case, both the algorithm of [32] and the one here coincide and become the algorithm of [21, Sec. 4.2.2]; when $\mathbf{A}$ is the characteristic matrix $x\mathbf{I}_m - \mathbf{M}$, this also relates to the fast algorithm in [31, Sec. 6] for a generic $\mathbf{M}$.

1.4. Approach, and existing tools

For the sake of presentation, suppose m is a power of 2. Writing $\mathbf{A} = \left[\begin{smallmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{A}_3 & \mathbf{A}_4\end{smallmatrix}\right]$ with the $\mathbf{A}_i$'s of dimensions $(m/2)\times(m/2)$, the algorithm of [32] is based on the block triangularization

$$\begin{bmatrix} * & * \\ \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$$

where the entries "$*$" are not computed, $\mathbf{B} = \mathbf{K}_1\mathbf{A}_2 + \mathbf{K}_2\mathbf{A}_4$, and $\mathbf{R}$ and $[\mathbf{K}_1 \;\; \mathbf{K}_2]$ are computed from $\left[\begin{smallmatrix}\mathbf{A}_1\\ \mathbf{A}_3\end{smallmatrix}\right]$ as a *row basis* and a *kernel basis*, respectively (see Section 2.2 for definitions). Then the leftmost matrix in the above identity is unimodular [32, Lem. 3.1] and thus, up to a constant factor, $\det(\mathbf{A})$ can be computed recursively as $\det(\mathbf{R})\det(\mathbf{B})$.

A first observation is that neither the kernel basis computation nor the matrix multiplication giving $\mathbf{B}$ is an obstacle towards a cost which is free of $\log(m)$. (The fastest known method for multiplying matrices with unbalanced degrees, such as in $\mathbf{B} = \mathbf{K}_1\mathbf{A}_2 + \mathbf{K}_2\mathbf{A}_4$, splits the computation into $O(\log(m))$ multiplications of smaller matrices with balanced degrees [58, Sec. 3.6], suggesting that its cost may involve a $\log(m)$ factor.) Indeed we show that, under the above assumptions on $\mathsf{M}(\cdot)$, the cost of these operations is in $O(m^{\omega}\,\mathsf{M}(D/m))$ and $O(m^{\omega}\,\mathsf{M}'(D/m))$, thus only involving factors logarithmic in D/m. In previous work, cost bounds either hide logarithmic factors [58] or they are derived without assuming $\mathcal{H}_{\mathrm{sm}}$ and have the form $O(m^{\omega-1}\,\mathsf{M}(D))$ and $O(m^{\omega-1}\,\mathsf{M}'(D))$ [27], thus resulting in factors logarithmic in D. Proving this observation is straightforward from the analyses in [58, 27] (see Section 2.5). This is a first key towards our main result: the characteristic matrix has $D = m$, and $O(m^{\omega}\,\mathsf{M}(1))$ is the same as $O(m^{\omega})$ whereas $O(m^{\omega-1}\,\mathsf{M}(m))$ involves factors logarithmic in m.

However, the computation of the row basis $\mathbf{R}$ remains an obstacle which prevents the algorithm of [32] from being a candidate for Theorem 1.2. Indeed, among the row basis algorithms we are aware of, only one has a cost bound which fits into our target up to logarithmic factors: the one of [56]. It relies on three kernel bases computations, and while one of them is similar to the computation of $[\mathbf{K}_1 \;\; \mathbf{K}_2]$ and is handled via the algorithm of [58], the two others have different constraints on the input and were the subject of a specific algorithm described in [56, Sec. 4].

[1] Precisely, if the upper triangular, row-wise Hermite normal form of $\mathbf{A}$ has diagonal entries $(1,\ldots,1,\lambda\det(\mathbf{A}))$, for some $\lambda \in \mathbb{K}\setminus\{0\}$ making $\lambda\det(\mathbf{A})$ monic.

In this reference, cost bounds were given without showing logarithmic factors; our own analysis revealed the presence of a factor logarithmic in m. The algorithm has a loop over $\Theta(\log(m))$ iterations, each of them calling [55, Algo. 2] for minimal approximant bases with unbalanced input. This approximant basis algorithm may spend a logarithmic number of iterations for finding some degree profile of the output basis, in a way reminiscent of Keller-Gehrig's algorithm in [31, Sec. 5] which finds the lengths of Krylov sequences (the link between the two situations becomes more explicit for approximant bases at small orders, see [27, Sec. 7]).

Our attempts at accelerating the row basis algorithm of [56] having not succeeded, the algorithm in this paper follows an approach which is more direct at first: remove the obstacle. Instead of computing a row basis $\mathbf{R}$ and relying on the identity $\det(\mathbf{A}) = \det(\mathbf{R})\det(\mathbf{B})$ (up to a constant), keep the first block row of $\mathbf{A}$:

$$\begin{bmatrix} \mathbf{I}_{m/2} & \mathbf{0} \\ \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \tag{1}$$

and rely on the identity $\det(\mathbf{A}) = \det(\mathbf{A}_1)\det(\mathbf{B})/\det(\mathbf{K}_2)$. The nonsingularity of $\mathbf{A}_1$ and $\mathbf{K}_2$ is easily ensured thanks to the assumption that $\mathbf{A}$ is reduced, as discussed in Section 1.5.

This leads to an unusual recursion scheme: we are not aware of a similar scheme being used in the literature on computational linear algebra. The algorithm uses three recursive calls with $(m/2) \times (m/2)$ matrices whose determinant has degree at most $D/2$ for two of them and at most D for the third; our complexity analysis in Section 3.3 shows that such a recursion gives the cost in Theorem 1.2. Precisely, if $\deg(\det(\mathbf{A}_1)) \leq D/2$ then degree properties of minimal kernel bases imply that $\deg(\det(\mathbf{K}_2)) \leq D/2$, yielding the two calls in half the degree; otherwise the algorithm uses inexpensive row and column operations on $\mathbf{A}$ to reduce to the case $\deg(\det(\mathbf{A}_1)) \leq D/2$.

Although this approach removes the obstacle of row basis computation which arises in [32], it adds a requirement: all recursive calls must take input matrices that are reduced. In the next section we discuss how to ensure the reducedness of $\mathbf{A}_1$ and $\mathbf{B}$ thanks to a straightforward generalization of [42, Sec. 3], and we describe a new algorithm which handles the more involved case of $\mathbf{K}_2$.

1.5. New tools, and ensuring reduced form in recursive calls

When outlining the approach of our determinant algorithm via the identity in Eq. (1), we implicitly assumed that the matrices used as input in recursive calls, i.e. $\mathbf{A}_1$ and $\mathbf{K}_2$ and $\mathbf{B}$, do satisfy the input requirement of row reducedness: this is not necessarily the case, even if starting from a reduced matrix $\mathbf{A}$.

Concerning $\mathbf{A}_1$, one may locate such a reduced submatrix of $\mathbf{A}$ and then permute rows and columns of $\mathbf{A}$ (which only affects the sign of $\det(\mathbf{A})$) to make this submatrix become the leading principal submatrix $\mathbf{A}_1$. This is a classical operation on reduced matrices which suggests using a form slightly stronger than reduced form called *weak Popov form* [36] (see Section 2.4). Assuming that $\mathbf{A}$ has this form ensures that its leading principal submatrix $\mathbf{A}_1$ has it as well. This assumption is acceptable in terms of complexity since one can transform a reduced $\mathbf{A}$ into a weak Popov $\mathbf{P}$ by means of fast linear algebra in a cost negligible compared to our target [42, Sec. 3]; note that $\mathbf{A}$ and $\mathbf{P}$ have the same determinant up to an easily found constant (see Algorithm 1).

Next, the cases of $\mathbf{K}_2$ and $\mathbf{B}$ are strongly linked. First, we will not discuss $\mathbf{K}_2$ but the whole kernel basis $\begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix}$. The fastest known algorithm for computing such a basis is that of [58], and for best efficiency it outputs a matrix in *shifted* reduced form, which is a generalization of reducedness involving degree weights given by a tuple $\boldsymbol{s} \in \mathbb{Z}^m$ called a *shift* (see Sections 2.3

and 2.4 for definitions); the non-shifted case is for $\boldsymbol{s} = \mathbf{0}$. As in the determinant algorithm of [32], here the shift for $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ is taken as the list of row degrees of $\mathbf{A}$, denoted by $\boldsymbol{s} = \mathrm{rdeg}(\mathbf{A})$; for the characteristic matrix one has $\boldsymbol{s} = (1, \ldots, 1)$ but non-uniform shifts may arise in recursive calls. We want $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ to be not only $\boldsymbol{s}$-reduced, but in $\boldsymbol{s}$-weak Popov form: a direct consequence is that $\mathbf{B}$ is in weak Popov form, and is thus suitable input for a recursive call.

To obtain $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ we use the kernel basis algorithm of [58] and transform its output into $\boldsymbol{s}$-weak Popov form. A minor issue is that the fastest known algorithm for such transformations was written in [42, Sec. 3] for non-shifted forms; yet it easily extends to shifted forms as we show in Section 4, obtaining the next result.

Theorem 1.3. *There is an algorithm* ReducedToWeakPopov *which takes as input a matrix* $\mathbf{A} \in \mathbb{K}[x]^{m \times n}$ *with* $m \leq n$ *and a shift* $\boldsymbol{s} \in \mathbb{Z}^n$ *such that* $\mathbf{A}$ *is in* $\boldsymbol{s}$*-reduced form, and returns an* $\boldsymbol{s}$*-weak Popov form of* $\mathbf{A}$ *using* $O(m^{\omega-2}nD + m^{\omega-1}n)$ *operations in* $\mathbb{K}$*, where* $D = |\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})| - m \cdot \min(\boldsymbol{s})$.

Here, following usual notation recalled in Section 2.1, $|\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})|$ is the sum of the $\boldsymbol{s}$-degrees of the rows of $\mathbf{A}$. This result extends [42, Thm. 13] since for $\boldsymbol{s} = \mathbf{0}$ the quantity D is the sum of the row degrees of $\mathbf{A}$ and in particular $D \leq m \deg(\mathbf{A})$, leading to the cost bound $O(m^{\omega-1} n \deg(\mathbf{A}))$.

To summarize, at this stage we have outlined how to ensure, without exceeding our target cost bound, that $\mathbf{A}_1$ and $\mathbf{B}$ are valid input for recursive calls, i.e. are in weak Popov form. Having $\det(\mathbf{A}_1)$ and $\det(\mathbf{B})$, it remains to find $\det(\mathbf{K}_2)$ and then the sought $\det(\mathbf{A})$ follows. We noted that, to ensure the form of $\mathbf{B}$ but also for efficiency reasons, the kernel basis $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ is computed in $\boldsymbol{s}$-weak Popov form for the shift $\boldsymbol{s} = \mathrm{rdeg}(\mathbf{A})$. This causes the main difficulty related to our modification of the determinant algorithm of [32]: $\mathbf{K}_2$ is not valid input for a recursive call since it is in $\boldsymbol{\nu}$-weak Popov form for some shift $\boldsymbol{\nu}$, a subtuple of $\boldsymbol{s}$ which is possibly nonzero.

A first idea is to extend our approach to the shifted case, allowing recursive calls with such a $\boldsymbol{\nu}$-reduced matrix: this is straightforward but gives an inefficient algorithm. Indeed, along the recursion the shift drifts away from its initial value and becomes arbitrarily large and unbalanced with respect to the degrees of the input matrices of recursive calls. For example, as mentioned above the sum of row degrees of the initial non-shifted $m \times m$ matrix $\mathbf{A}$ is $D = \deg(\det(\mathbf{A}))$, whereas for the $\boldsymbol{\nu}$-shifted $(m/2) \times (m/2)$ matrix $\mathbf{K}_2$ we only have the same bound D instead of one related to $\deg(\det(\mathbf{K}_2))$ itself, which is known to be at most $D/2$ in our algorithm. This gap, here between D and $D/2$, will only grow as the algorithm goes down the tree of recursive calls, meaning that degrees in matrices handled recursively are not sufficiently well controlled.

Another idea is to compute a $\mathbf{0}$-reduced matrix which has the same determinant as $\mathbf{K}_2$. Finding a $\mathbf{0}$-reduced form of $\mathbf{K}_2$ within our target cost seems to be a difficult problem. The best known algorithms for general $\mathbf{0}$-reduction involve $\log(m)$ factors, either explicitly [23] or implicitly [38] (in the latter approach one starts by using the above-discussed triangularization procedure of [32] which we are modifying here to avoid $\log(m)$ factors). More specific algorithms exploit the form of $\mathbf{K}_2$, interpreting the problem as a change of shift from $\boldsymbol{\nu}$ to $\mathbf{0}$; yet at the time of writing efficient changes of shifts have only been achieved when the target shift is larger than the origin shift [27, Sec. 5], a fact that offers degree control for the transformation between the two matrices. Another possibility is to compute the so-called $\boldsymbol{\nu}$-Popov form $\mathbf{P}$ of $\mathbf{K}_2$, since its transpose $\mathbf{P}^{\mathsf{T}}$ is $\mathbf{0}$-reduced by definition (see Section 2.4), and $\det(\mathbf{P}^{\mathsf{T}}) = \det(\mathbf{P})$ is $\det(\mathbf{K}_2)$ up to a constant. However this suffers from the same issue, as computing $\mathbf{P}$ is essentially the same as changing the shift $\boldsymbol{\nu}$ into the nonpositive shift $-\boldsymbol{\delta}$, where $\boldsymbol{\delta}$ is the list of diagonal degrees of $\mathbf{K}_2$ [42, 26].

To circumvent these issues, we use the property that the transpose $\mathbf{K}_2^{\mathsf{T}}$ of a $\boldsymbol{\nu}$-reduced matrix is in $-\boldsymbol{d}$-reduced form where $\boldsymbol{d} = \mathrm{rdeg}_{\boldsymbol{\nu}}(\mathbf{K}_2)$. This fact naturally comes up here since $\det(\mathbf{K}_2) =$

$\det(\mathbf{K}_2^\intercal)$, but seems otherwise rarely exploited in polynomial matrix algorithms: in fact we did not find a previous occurrence of it apart from related degree considerations in [56, Lem. 2.2].

Transposing the above two approaches using $\mathbf{K}_2^\intercal$ instead of $\mathbf{K}_2$, we observe that computing a $\mathbf{0}$-reduced form of $\mathbf{K}_2^\intercal$ is a change of shift from $-\boldsymbol{d}$ to $\mathbf{0}$, and computing the $-\boldsymbol{d}$-Popov form $\mathbf{P}$ of $\mathbf{K}_2^\intercal$ is essentially a change of shift from $-\boldsymbol{d}$ to $-\boldsymbol{\delta}$. In both cases the target shift is larger than the origin shift, implying that the kernel-based change of shift of [27, Sec. 5] involves matrices of well-controlled degrees. Still, this is not enough to make this change of shift efficient as such, the difficulty being now that the average row degree of $\mathbf{K}_2^\intercal$ may not be small: only its average column degree, which corresponds to the average row degree of $\mathbf{K}_2$, is controlled.

Our solution uses the second approach, computing the $-\boldsymbol{d}$-Popov form $\mathbf{P}$, because it offers the a priori knowledge that the column degrees of $\mathbf{P}$ are exactly $\boldsymbol{\delta}$. We exploit this degree knowledge to carry out partial linearization techniques, originally designed for approximant bases [46, 28], which we extend here to kernel bases. These techniques allow us to reduce our problem to a kernel basis computation where the matrix entries have uniformly small degrees, implying that it can be efficiently handled via the minimal approximant basis algorithm PM-Basis from [21]. The next result summarizes the new algorithmic tool developed in Section 5 for finding $\mathbf{P}$.

Theorem 1.4. *Let $\boldsymbol{s} \in \mathbb{Z}^m$, let $\mathbf{A} \in \mathbb{K}[x]^{m\times m}$ be in $-\boldsymbol{s}$-weak Popov form, let $\boldsymbol{\delta} \in \mathbb{Z}^m_{\geq 0}$ be the $-\boldsymbol{s}$-pivot degree of $\mathbf{A}$, and assume that $\boldsymbol{s} \geq \boldsymbol{\delta}$. There is an algorithm* WeakPopovToPopov *which takes as input $(\mathbf{A}, \boldsymbol{s})$ and computes the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$ by*

- *performing* PM-Basis *at order less than $|\boldsymbol{s}|/m + 4$ on an input matrix of row dimension at most $6m$ and column dimension at most $3m$,*
- *multiplying the inverse of a matrix in $\mathbb{K}^{m\times m}$ by a matrix in $\mathbb{K}[x]^{m\times m}$ of column degree $\boldsymbol{\delta}$,*
- *and performing $O(m^2)$ extra operations in $\mathbb{K}$.*

Thus, computing the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$ can be done in $O(m^\omega\, \mathsf{M}'(|\boldsymbol{s}|/m))$ operations in $\mathbb{K}$.

This theorem is a generalization of [42, Sec. 4] to shifted forms, for shifts $-\boldsymbol{s}$ that satisfy the assumption $\boldsymbol{s} \geq \boldsymbol{\delta}$. Indeed, if $\mathbf{A}$ is $\mathbf{0}$-weak Popov, then one recovers [42, Thm. 20] by taking $\boldsymbol{s} = (\deg(\mathbf{A}), \ldots, \deg(\mathbf{A}))$ in the above theorem. For comparison, the naive generalization of [42, Sec. 4] to shifted forms runs in $O(m^\omega\, \mathsf{M}'(\max(\boldsymbol{s})))$, which exceeds our target complexity as soon as $\max(\boldsymbol{s}) \gg |\boldsymbol{s}|/m$. Hence the use of partial linearization techniques, which were not needed in the non-shifted case featuring $\max(\boldsymbol{s}) = |\boldsymbol{s}|/m = \deg(\mathbf{A})$.

As mentioned above, our Algorithm WeakPopovToPopov is based on the computation of a kernel basis with a priori knowledge of the degree profile of the output. This kernel problem is very close to the one handled in [56, Sec. 4], except that in this reference one only has upper bounds on the output degrees, implying a certain number—possibly logarithmic in m—of calls to PM-Basis to recover the output and its actual degrees. In the same spirit but in the context of approximant bases, [28, Sec. 5] uses partial linearization techniques to reduce an arbitrary input with known output degrees to essentially one call to PM-Basis, whereas [55, Algo. 2] assumes weaker output degree information and makes a potentially logarithmic number of calls to PM-Basis.

1.6. Perspectives

We plan to implement our characteristic polynomial algorithm in the LinBox ecosystem [48, 49]. First prototype experiments suggest that, for large finite fields, it could be competitive with the existing fastest-known implementation, based on the randomized algorithm of [39]. The native support for small fields of our algorithm should outperform the algorithm of [39]

which requires expensive field extensions. Another perspective stems from the remark that our algorithm resorts to fast polynomial multiplication (see assumption $\mathcal{H}_\omega$), while previous ones did not [31, 39]: we woud like to understand whether the same cost can be achieved by a purely linear algebraic approach. Finally, perhaps the most challenging problem related to characteristic polynomial computation is to compute Frobenius forms deterministically in the time of matrix multiplication, the current best known complexity bound being $O(m^\omega \log(m) \log(\log(m)))$ [44]; and more generally computing Smith forms of polynomial matrices with a cost free of factors logarithmic in the matrix dimension.

2. Preliminaries on polynomial matrices

In this section we present the notation as well as basic definitions and properties that will be used throughout the paper.

2.1. Notation

Tuples of integers will often be manipulated entry-wise. In particular, for tuples $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Z}^n$ of the same length n, we write $\boldsymbol{s} + \boldsymbol{t}$ for their entry-wise sum, and the inequality $\boldsymbol{s} \leq \boldsymbol{t}$ means that each entry in $\boldsymbol{s}$ is less than or equal to the corresponding entry in $\boldsymbol{t}$. The concatenation of tuples is denoted by $(\boldsymbol{s}, \boldsymbol{t})$. We write $|\boldsymbol{t}|$ for the sum of the entries of $\boldsymbol{t}$. The tuple of zeros is denoted by $\mathbf{0} = (0, \ldots, 0)$; its length is understood from the context.

For an $m \times n$ matrix $\mathbf{A}$ over some ring, we write $\mathbf{A}_{i,j}$ for its entry at index (i, j). We extend this to submatrices: given sets $I \subseteq \{1, \ldots, m\}$ and $J \subseteq \{1, \ldots, n\}$ of row and column indices, we write $\mathbf{A}_{I,J}$ for the submatrix of $\mathbf{A}$ formed by its entries indexed by $I \times J$. Besides, $\mathbf{A}_{I,*}$ stands for the submatrix of $\mathbf{A}$ formed by its rows with index in I, and we use the similar notation $\mathbf{A}_{*,J}$. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}^\mathsf{T}$. The identity matrix of size n is denoted by $\mathbf{I}_n$, while the $n \times n$ matrix with 1 on the antidiagonal and 0 elsewhere is denoted by $\mathbf{J}_n$. In particular, when writing $\boldsymbol{s}\mathbf{J}_n$ for a tuple $\boldsymbol{s} = (s_1, \ldots, s_n) \in \mathbb{Z}^n$, we mean the reversed tuple $\boldsymbol{s}\mathbf{J}_n = (s_n, \ldots, s_1)$.

Now consider $\mathbf{A}$ with polynomial entries, i.e. $\mathbf{A} \in \mathbb{K}[x]^{m \times n}$. The *degree* of $\mathbf{A}$ is denoted by $\deg(\mathbf{A})$ and is the largest of the degrees of its entries, or $-\infty$ if $\mathbf{A} = \mathbf{0}$. The *row degree* of $\mathbf{A}$ is the tuple $\operatorname{rdeg}(\mathbf{A}) \in (\mathbb{Z}_{\geq 0} \cup \{-\infty\})^m$ whose ith entry is $\max_{1 \leq j \leq n}(\deg(\mathbf{A}_{i,j}))$. More generally, for a tuple $\boldsymbol{s} = (s_1, \ldots, s_n) \in \mathbb{Z}^n$, the *$\boldsymbol{s}$-row degree* of $\mathbf{A}$ is the tuple $\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A}) \in (\mathbb{Z} \cup \{-\infty\})^m$ whose ith entry is $\max_{1 \leq j \leq n}(\deg(\mathbf{A}_{i,j}) + s_j)$. In this context, the tuple $\boldsymbol{s}$ is commonly called a *(degree) shift* [4]. The *(shifted) column degree* of $\mathbf{A}$ is defined similarly.

We write $\mathbf{X}^{\boldsymbol{s}}$ for the $n \times n$ diagonal matrix $\operatorname{diag}(x^{s_1}, \ldots, x^{s_n})$ which is over the ring $\mathbb{K}[x, x^{-1}]$ of Laurent polynomials over $\mathbb{K}$. Note that, hereafter, Laurent polynomials will only arise in proofs and explanations, more specifically in considerations about shifted degrees: they never arise in algorithms, which for the sake of clarity only involve polynomials in $\mathbb{K}[x]$. The usefulness of this matrix $\mathbf{X}^{\boldsymbol{s}}$ will become clear in the definition of leading matrices in the next subsection.

The next lemma gives a link between shifted row degrees and shifted column degrees. We will mostly use the following particular case of it: the column degree of $\mathbf{A}$ is at most $\boldsymbol{d} \in \mathbb{Z}^n_{\geq 0}$ (entry-wise) if and only if the $-\boldsymbol{d}$-row degree of $\mathbf{A}$ is nonpositive.

Lemma 2.1 ([56, Lemma 2.2])**.** *Let $\mathbf{A}$ be a matrix in $\mathbb{K}[x]^{m \times n}$, $\boldsymbol{d}$ be a tuple in $\mathbb{Z}^n$, and $\boldsymbol{t}$ be a tuple in $\mathbb{Z}^m$. Then,* $\operatorname{cdeg}_{\boldsymbol{t}}(\mathbf{A}) \leq \boldsymbol{d}$ *if and only if* $\operatorname{rdeg}_{-\boldsymbol{d}}(\mathbf{A}) \leq -\boldsymbol{t}$.

2.2. Bases of modules, kernel bases and approximant bases

We recall that any $\mathbb{K}[x]$-submodule $\mathcal{M}$ of $\mathbb{K}[x]^{1\times n}$ is free, and admits a basis formed by r elements of $\mathbb{K}[x]^{1\times n}$, where $r \leq n$ is called the rank of $\mathcal{M}$ [see e.g. 15]. Such a basis can thus be represented as an $r \times n$ matrix $\mathbf{B}$ over $\mathbb{K}[x]$ whose rows are the basis elements; this basis matrix $\mathbf{B}$ has rank r.

For a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$, its *row space* is the $\mathbb{K}[x]$-submodule $\{\mathbf{pA}, \mathbf{p} \in \mathbb{K}[x]^{1\times m}\}$ of $\mathbb{K}[x]^{1\times n}$, that is, the set of all $\mathbb{K}[x]$-linear combinations of its rows. If $\mathbf{B} \in \mathbb{K}[x]^{r\times n}$ is a basis of this row space, then $\mathbf{B}$ is said to be a *row basis* of $\mathbf{A}$; in particular, r is the rank of $\mathbf{B}$ and of $\mathbf{A}$.

The *left kernel* of $\mathbf{A}$, denoted by $\mathcal{K}(\mathbf{A})$, is the $\mathbb{K}[x]$-module $\{\mathbf{p} \in \mathbb{K}[x]^{1\times m} \mid \mathbf{pA} = \mathbf{0}\}$. A matrix $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$ is a *left kernel basis of* $\mathbf{A}$ if its rows form a basis of $\mathcal{K}(\mathbf{A})$, in which case $k = m - r$. Similarly, a *right* kernel basis of $\mathbf{A}$ is a matrix $\mathbf{K} \in \mathbb{K}[x]^{n\times(n-r)}$ whose *columns* form a basis of the right kernel of $\mathbf{A}$.

Given positive integers $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n) \in \mathbb{Z}^n_{>0}$ and a matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, the set of *approximants for* $\mathbf{F}$ *at order* $\boldsymbol{\gamma}$ [see e.g. 51, 3] is the $\mathbb{K}[x]$-submodule of $\mathbb{K}[x]^{1\times m}$ defined as

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

The identity $\mathbf{pF} = \mathbf{0} \bmod \mathbf{X}^{\boldsymbol{\gamma}}$ means that $\mathbf{pF}_{*,j} = 0 \bmod x^{\gamma_j}$ for $1 \leq j \leq n$. Since all m rows of the matrix $x^{\max(\boldsymbol{\gamma})}\mathbf{I}_m$ are in $\mathcal{A}_{\boldsymbol{\gamma}}(\mathbf{F})$, this module has rank m.

2.3. Leading matrices and reduced forms of polynomial matrices

We will often compute with polynomial matrices that have a special form, called the *(shifted) reduced form*. It corresponds to a type of minimality of the degrees of such matrices, and also provides good control of these degrees during computations as illustrated by the *predictable degree property* [17] [29, Thm. 6.3-13] which we recall below. In this section, we introduce the notion of *row* reducedness; to avoid confusion, we will not use the similar notion of column reducedness in this paper, and thus all further mentions of reducedness refer to row reducedness.

For shifted reduced forms, we follow the definitions in [4, 5]. Let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ and $\boldsymbol{s} \in \mathbb{Z}^n$, and let $\boldsymbol{t} = (t_1, \ldots, t_m) = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})$. Then, the $\boldsymbol{s}$-leading matrix of $\mathbf{A}$ is the matrix $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A}) \in \mathbb{K}^{m\times n}$ whose entry (i, j) is the coefficient of degree $t_i - s_j$ of the entry (i, j) of $\mathbf{A}$, or 0 if $t_i = -\infty$. Equivalently, $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ is the coefficient of degree zero of $\mathbf{X}^{-\boldsymbol{t}}\mathbf{A}\mathbf{X}^{\boldsymbol{s}}$, whose entries are in $\mathbb{K}[x^{-1}]$. The matrix $\mathbf{A}$ is said to be in $\boldsymbol{s}$-reduced form if its $\boldsymbol{s}$-leading matrix has full row rank. In particular, a matrix in $\boldsymbol{s}$-reduced form must have full row rank.

For a matrix $\mathbf{M} \in \mathbb{K}[x]^{k\times m}$, we have $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{MA}) \leq \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{M})$ and this is an equality when no cancellation of leading terms occurs in this left-multiplication. The predictable degree property states that $\mathbf{A}$ is $\boldsymbol{s}$-reduced if and only if $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{MA}) = \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{M})$ holds for any $\mathbf{M} \in \mathbb{K}[x]^{k\times m}$. Here is a useful consequence of this characterization.

Lemma 2.2. *Let $\boldsymbol{s} \in \mathbb{Z}^n$, let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$, and let $\boldsymbol{t} = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})$. If $\mathbf{A}$ is $\boldsymbol{s}$-reduced, then the identity $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{MA}) = \mathrm{lm}_{\boldsymbol{t}}(\mathbf{M})\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ holds for any $\mathbf{M} \in \mathbb{K}[x]^{k\times m}$.*

Proof. Let $\boldsymbol{d} = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{MA})$. By definition, $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{MA})$ is the coefficient of degree 0 of the matrix $\mathbf{X}^{-\boldsymbol{d}}\mathbf{MA}\mathbf{X}^{\boldsymbol{s}} = \mathbf{X}^{-\boldsymbol{d}}\mathbf{M}\mathbf{X}^{\boldsymbol{t}}\mathbf{X}^{-\boldsymbol{t}}\mathbf{A}\mathbf{X}^{\boldsymbol{s}}$, whose entries are in $\mathbb{K}[x^{-1}]$. Besides, since $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A}) = \boldsymbol{t}$ and since the predictable degree property gives $\mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{M}) = \boldsymbol{d}$, the matrices $\mathbf{X}^{-\boldsymbol{d}}\mathbf{M}\mathbf{X}^{\boldsymbol{t}}$ and $\mathbf{X}^{-\boldsymbol{t}}\mathbf{A}\mathbf{X}^{\boldsymbol{s}}$ are over $\mathbb{K}[x^{-1}]$ and their coefficients of degree 0 are $\mathrm{lm}_{\boldsymbol{t}}(\mathbf{M})$ and $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$, respectively. □

Another characterization of matrices in $\boldsymbol{s}$-reduced form is that they have minimal $\boldsymbol{s}$-row degree among all matrices which represent the same $\mathbb{K}[x]$-module [53, Def. 2.13]; in this paper, we will use the following consequence of this minimality.

Lemma 2.3. *Let $\mathcal{M}$ be a submodule of $\mathbb{K}[x]^{1\times n}$ of rank m, let $\boldsymbol{s} \in \mathbb{Z}^n$, and let $\boldsymbol{t} \in \mathbb{Z}^m$ be the $\boldsymbol{s}$-row degree of some $\boldsymbol{s}$-reduced basis of $\mathcal{M}$. Without loss of generality, assume that $\boldsymbol{t}$ is nondecreasing. Let $\mathbf{B} \in \mathbb{K}[x]^{m\times n}$ be a matrix of rank m whose rows are in $\mathcal{M}$, and let $\boldsymbol{d} \in \mathbb{Z}^m$ be its $\boldsymbol{s}$-row degree sorted in nondecreasing order. If $\boldsymbol{d} \le \boldsymbol{t}$, then $\mathbf{B}$ is an $\boldsymbol{s}$-reduced basis of $\mathcal{M}$, and $\boldsymbol{d} = \boldsymbol{t}$.*

Proof. Up to permuting the rows of $\mathbf{B}$, we assume that $\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{B}) = \boldsymbol{d}$ without loss of generality. Let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ be an $\boldsymbol{s}$-reduced basis of $\mathcal{M}$ such that $\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A}) = \boldsymbol{t}$. Since the rows of $\mathbf{B}$ are in $\mathcal{M}$, there exists a matrix $\mathbf{U} \in \mathbb{K}[x]^{m\times m}$ such that $\mathbf{B} = \mathbf{U}\mathbf{A}$; and $\mathbf{U}$ is nonsingular since $\mathbf{B}$ and $\mathbf{A}$ have rank m. Since $\mathbf{A}$ is $\boldsymbol{s}$-reduced, the predictable degree property applies, ensuring that

$$\boldsymbol{d} = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{B}) = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{U}\mathbf{A}) = \operatorname{rdeg}_{\boldsymbol{t}}(\mathbf{U}).$$

This means that $\deg(\mathbf{U}_{i,j}) \le d_i - t_j$ for all $1 \le i, j \le m$.

Now, assume by contradiction that $\boldsymbol{d} = \boldsymbol{t}$ does not hold. Thus, $d_k < t_k$ for some $1 \le k \le m$. Then, for $i \le k$ and $j \ge k$ we have $d_i \le d_k < t_k \le t_j$, hence $\deg(\mathbf{U}_{i,j}) < 0$. Thus, the submatrix $\mathbf{U}_{\{1,\ldots,k\},\{k,\ldots,m\}}$ is zero, which implies that $\mathbf{U}$ is singular; this is a contradiction, hence $\boldsymbol{d} = \boldsymbol{t}$.

Since $\boldsymbol{t}$ is nondecreasing, the inequality $\deg(\mathbf{U}_{i,j}) \le t_i - t_j$ implies that $\mathbf{U}$ is a block lower triangular matrix whose diagonal blocks have degree 0; hence these blocks are invertible matrices over $\mathbb{K}$, and $\mathbf{U}$ is unimodular [see 42, Lemma 6 for similar degree considerations, starting from stronger assumptions on $\mathbf{A}$ and $\mathbf{B}$]. Thus, $\mathbf{B}$ is a basis of $\mathcal{M}$.

Furthermore, it is easily observed that $\operatorname{lm}_{\boldsymbol{d}}(\mathbf{U}) \in \mathbb{K}^{m\times m}$ is block lower triangular with the same invertible diagonal blocks as $\mathbf{U}$; hence $\operatorname{lm}_{\boldsymbol{d}}(\mathbf{U})$ is invertible. On the other hand, Lemma 2.2 states that $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{B}) = \operatorname{lm}_{\boldsymbol{d}}(\mathbf{U})\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$. Thus $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{B})$ has rank $m = \operatorname{rank}(\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A}))$, and $\mathbf{B}$ is $\boldsymbol{s}$-reduced. □

2.4. Pivots and weak Popov forms of polynomial matrices

For algorithmic purposes, it is often convenient to work with reduced forms that satisfy some additional requirements, called *weak Popov forms*. These are intrinsically related to the notion of *pivot* of a polynomial matrix.

For a nonzero vector $\mathbf{p} \in \mathbb{K}[x]^{1\times n}$ and a shift $\boldsymbol{s} \in \mathbb{Z}^n$, the $\boldsymbol{s}$-pivot of $\mathbf{p}$ is its rightmost entry p_j such that $\deg(p_j) + s_j = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{p})$ [4, 36]; it corresponds to the rightmost nonzero entry of $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{p})$. The index $j = \pi$ and the degree $\deg(p_\pi) = \delta$ of this entry are called the $\boldsymbol{s}$-pivot index and $\boldsymbol{s}$-pivot degree, respectively. For brevity, in this paper the pair (π, δ) is called the *$\boldsymbol{s}$-pivot profile* of $\mathbf{p}$. By convention, the zero vector in $\mathbb{K}[x]^{1\times n}$ has $\boldsymbol{s}$-pivot index 0 and $\boldsymbol{s}$-pivot degree $-\infty$. These notions are extended to matrices $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ by forming row-wise lists. For example, the $\boldsymbol{s}$-pivot index of $\mathbf{A}$ is $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m) \in \mathbb{Z}^m_{>0}$ where π_i is the $\boldsymbol{s}$-pivot index of the row $\mathbf{A}_{i,*}$. The $\boldsymbol{s}$-pivot degree $\boldsymbol{\delta}$ and the $\boldsymbol{s}$-pivot profile $(\pi_i, \delta_i)_{1\le i\le m}$ of $\mathbf{A}$ are defined similarly.

Then, $\mathbf{A}$ is said to be in $\boldsymbol{s}$-weak Popov form if it has no zero row and $\boldsymbol{\pi}$ is strictly increasing; and $\mathbf{A}$ is said to be in $\boldsymbol{s}$-unordered weak Popov form if it is in $\boldsymbol{s}$-weak Popov form up to row permutation, i.e. the entries of $\boldsymbol{\pi}$ are pairwise distinct. Furthermore, a matrix is in $\boldsymbol{s}$-Popov form if it is in $\boldsymbol{s}$-weak Popov form, its $\boldsymbol{s}$-pivots are monic, and each of these $\boldsymbol{s}$-pivots has degree strictly larger than the other entries in the same column. For a given $\mathbb{K}[x]$-submodule $\mathcal{M}$ of $\mathbb{K}[x]^{1\times n}$, there is a unique basis of $\mathcal{M}$ which is in $\boldsymbol{s}$-Popov form [4].

For a given matrix $\mathbf{B}$, the matrix $\mathbf{A}$ is said to be an $\boldsymbol{s}$-reduced (resp. $\boldsymbol{s}$-weak Popov, $\boldsymbol{s}$-Popov) form of $\mathbf{B}$ if $\mathbf{A}$ is a row basis of $\mathbf{B}$ and $\mathbf{A}$ is in $\boldsymbol{s}$-reduced (resp. $\boldsymbol{s}$-weak Popov, $\boldsymbol{s}$-Popov) form.

Like for $\boldsymbol{s}$-reducedness, the property of a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ to be in $\boldsymbol{s}$-weak Popov form depends only on its $\boldsymbol{s}$-leading matrix $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A}) \in \mathbb{K}^{m\times n}$, namely on the fact that it has a staircase shape. Indeed, $\mathbf{A}$ is in $\boldsymbol{s}$-weak (resp. $\boldsymbol{s}$-unordered weak) Popov form if and only if $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$ has no

zero row and $\mathbf{J}_m \mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n$ is in row echelon form (resp. in row echelon form up to row permutation); this was used as a definition by Beckermann et al. [4, 5]. In particular, for any constant matrix $\mathbf{C} \in \mathbb{K}^{m\times n}$, we have $\mathrm{lm}_{\mathbf{0}}(\mathbf{C}) = \mathbf{C}$ and therefore $\mathbf{C}$ is in $\mathbf{0}$-weak (resp. $\mathbf{0}$-unordered weak) Popov form if and only if it has no zero row and $\mathbf{J}_m\mathbf{C}\mathbf{J}_n$ is in row echelon form (resp. in row echelon form up to row permutation). Taking $\mathbf{C} = \mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$, the next lemma follows.

Lemma 2.4. *Let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ and let $\boldsymbol{s} \in \mathbb{Z}^n$. Then, $\mathbf{A}$ is in $\boldsymbol{s}$-weak (resp. $\boldsymbol{s}$-unordered weak) Popov form if and only if $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ is in $\mathbf{0}$-weak (resp. $\mathbf{0}$-unordered weak) Popov form.*

Furthermore, if $\mathbf{A}$ is in $\boldsymbol{s}$-weak Popov form and $(j_1, \ldots, j_m)$ is the list of indices of pivot columns in the row echelon form $\mathbf{J}_m\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n$ (in other words, this list is the column rank profile of that matrix), then the $\boldsymbol{s}$-pivot index of $\mathbf{A}$ is equal to $(n+1-j_m, \ldots, n+1-j_1)$. This leads to the following lemma which states that the $\boldsymbol{s}$-pivot profile is an invariant of left-unimodularly equivalent $\boldsymbol{s}$-weak Popov forms [29, 4, 5], generalizing the fact that for matrices over $\mathbb{K}$ the set of indices of pivot columns is an invariant of left-equivalent row echelon forms.

Lemma 2.5. *Let $\boldsymbol{s} \in \mathbb{Z}^n$ and let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ be in $\boldsymbol{s}$-unordered weak Popov form with $\boldsymbol{s}$-pivot profile $(\pi_i, \delta_i)_{1\le i\le m}$. Then, the $\boldsymbol{s}$-pivot profile of the $\boldsymbol{s}$-Popov form of $\mathbf{A}$ is $(\pi_{\sigma(i)}, \delta_{\sigma(i)})_{1\le i\le m}$, where $\sigma : \{1, \ldots, m\} \to \{1, \ldots, m\}$ is the permutation such that $(\pi_{\sigma(i)})_{1\le i\le m}$ is strictly increasing.*

Proof. Without loss of generality we assume that $\mathbf{A}$ is in $\boldsymbol{s}$-weak Popov form, implying also $\sigma(i) = i$ for $1 \le i \le m$. Let $\mathbf{P} \in \mathbb{K}[x]^{m\times n}$ be the $\boldsymbol{s}$-Popov form of $\mathbf{A}$: we want to prove that $\mathbf{A}$ and $\mathbf{P}$ have the same $\boldsymbol{s}$-pivot index and the same $\boldsymbol{s}$-pivot degree. Let $\mathbf{U}$ be the unimodular matrix such that $\mathbf{P} = \mathbf{U}\mathbf{A}$; then Lemma 2.2 yields $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{P}) = \mathrm{lm}_{\boldsymbol{t}}(\mathbf{U})\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$, where $\boldsymbol{t} = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})$. Since both $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{P})$ and $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ have full row rank, $\mathrm{lm}_{\boldsymbol{t}}(\mathbf{U}) \in \mathbb{K}^{m\times m}$ is invertible. Then

$$\mathbf{J}_m\mathrm{lm}_{\boldsymbol{s}}(\mathbf{P})\mathbf{J}_n = \mathbf{J}_m\mathrm{lm}_{\boldsymbol{t}}(\mathbf{U})\mathbf{J}_m\mathbf{J}_m\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n$$

holds, and thus the row echelon forms $\mathbf{J}_m\mathrm{lm}_{\boldsymbol{s}}(\mathbf{P})\mathbf{J}_n$ and $\mathbf{J}_m\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n$ have the same pivot columns since $\mathbf{J}_m\mathrm{lm}_{\boldsymbol{t}}(\mathbf{U})\mathbf{J}_m \in \mathbb{K}^{m\times m}$ is invertible. It follows from the discussion preceding this lemma that $\mathbf{P}$ has the same $\boldsymbol{s}$-pivot index as $\mathbf{A}$.

As a consequence, $\mathbf{P}$ has the same $\boldsymbol{s}$-pivot degree as $\mathbf{A}$ if and only if $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{P}) = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A})$. Suppose by contradiction that there exists an index i such that $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{P}_{i,*}) < \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A}_{i,*})$. Then, build the matrix $\mathbf{B} \in \mathbb{K}[x]^{m\times n}$ which is equal to $\mathbf{A}$ except for its ith row which is replaced by $\mathbf{P}_{i,*}$. By construction, $\mathbf{B}$ has rank m (since it is in $\boldsymbol{s}$-weak Popov form) and its rows are in the row space of $\mathbf{A}$. Writing $\boldsymbol{d}$ for the tuple $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{B})$ sorted in nondecreasing order, and $\boldsymbol{u}$ for the tuple $\boldsymbol{t}$ sorted in nondecreasing order, we have $\boldsymbol{d} \le \boldsymbol{u}$ and $\boldsymbol{d} \ne \boldsymbol{u}$, which contradicts Lemma 2.3. Hence there is no such index i, and since this proof by contradiction is symmetric in $\mathbf{A}$ and $\mathbf{P}$, there is no index i such that $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A}_{i,*}) < \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{P}_{i,*})$ either. Thus $\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A}) = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{P})$. □

We will also use the following folklore fact, which is a corollary of Lemma 2.2, and has often been used in algorithms for approximant bases or kernel bases in order to preserve the reducedness of matrices during the computation.

Lemma 2.6. *Let $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ and $\mathbf{B} \in \mathbb{K}[x]^{k\times m}$, and let $\boldsymbol{s} \in \mathbb{Z}^n$ and $\boldsymbol{t} = \mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{A}) \in \mathbb{Z}^m$. Then,*

- *if $\mathbf{A}$ is $\boldsymbol{s}$-reduced and $\mathbf{B}$ is $\boldsymbol{t}$-reduced, then $\mathbf{B}\mathbf{A}$ is $\boldsymbol{s}$-reduced;*
- *if $\mathbf{A}$ is in $\boldsymbol{s}$-weak Popov form and $\mathbf{B}$ is in $\boldsymbol{t}$-weak Popov form, then $\mathbf{B}\mathbf{A}$ is in $\boldsymbol{s}$-weak Popov form.*

Proof. Since $\mathbf{A}$ is $\boldsymbol{s}$-reduced, Lemma 2.2 states that $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{BA}) = \mathbf{ML}$ where $\mathbf{L} = \mathrm{lm}_{\boldsymbol{s}}(\mathbf{A}) \in \mathbb{K}^{m\times n}$ and $\mathbf{M} = \mathrm{lm}_{\boldsymbol{t}}(\mathbf{B}) \in \mathbb{K}^{k\times m}$. The first item then follows from the fact that if $\mathbf{M}$ has rank k and $\mathbf{L}$ has rank m, then $\mathbf{ML}$ has rank k. Similarly, the second item reduces to prove that, assuming $\mathbf{M}$ and $\mathbf{L}$ are in row echelon form with full row rank, then $\mathbf{ML}$ is also in row echelon form. Let $(a_1, \ldots, a_k)$ (resp. $(b_1, \ldots, b_m)$) be the pivot indices of $\mathbf{M}$ (resp. $\mathbf{L}$). Then the ith row of $\mathbf{ML}$ is a nonzero multiple of row a_i of $\mathbf{L}$ combined with multiples of rows of $\mathbf{L}$ of index greater than a_i. Consequently, the pivot indices of the rows of $\mathbf{ML}$ are $b_{a_1} < \cdots < b_{a_k}$, which proves that $\mathbf{ML}$ is in row echelon form. □

Finally, under assumptions that generalize the situation encountered in our determinant algorithm below, we show that the pivot entries of a kernel basis $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ are located in its rightmost columns, that is, in $\mathbf{K}_2$.

Lemma 2.7. *Let $\boldsymbol{t} \in \mathbb{Z}^n$, let $\mathbf{F} \in \mathbb{K}[x]^{n\times n}$ be in $\boldsymbol{t}$-weak Popov form, and let $\boldsymbol{u} = \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{F})$. Let $\mathbf{G} \in \mathbb{K}[x]^{m\times n}$ and $\boldsymbol{v} \in \mathbb{Z}^m$ be such that $\boldsymbol{v} \geq \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{G})$, and let $\mathbf{K} = [\mathbf{K}_1 \ \ \mathbf{K}_2] \in \mathbb{K}[x]^{m\times(m+n)}$ be a $(\boldsymbol{u}, \boldsymbol{v})$-weak Popov basis of $\mathcal{K}([\begin{smallmatrix}\mathbf{F}\\ \mathbf{G}\end{smallmatrix}])$, where $\mathbf{K}_1$ and $\mathbf{K}_2$ have m and n columns, respectively. Then, the $(\boldsymbol{u}, \boldsymbol{v})$-pivot entries of $\mathbf{K}$ are all located in $\mathbf{K}_2$; in particular, $\mathbf{K}_2$ is in $\boldsymbol{v}$-weak Popov form.*

Proof. Since the $(\boldsymbol{u}, \boldsymbol{v})$-pivot entry of a row is the rightmost entry of that row which reaches the $(\boldsymbol{u}, \boldsymbol{v})$-row degree, it is enough to prove that $\mathrm{rdeg}_{\boldsymbol{v}}(\mathbf{K}_2) \geq \mathrm{rdeg}_{\boldsymbol{u}}(\mathbf{K}_1)$. First, from $\boldsymbol{v} \geq \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{G})$, we obtain $\mathrm{rdeg}_{\boldsymbol{v}}(\mathbf{K}_2) \geq \mathrm{rdeg}_{\mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{G})}(\mathbf{K}_2)$. Now, by definition, $\mathrm{rdeg}_{\mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{G})}(\mathbf{K}_2) \geq \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{K}_2\mathbf{G})$. Since the rows of $\mathbf{K}$ are in $\mathcal{K}([\begin{smallmatrix}\mathbf{F}\\ \mathbf{G}\end{smallmatrix}])$, we have $\mathbf{K}_2\mathbf{G} = -\mathbf{K}_1\mathbf{F}$, hence $\mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{K}_2\mathbf{G}) = \mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{K}_1\mathbf{F})$. Since $\mathbf{F}$ is $\boldsymbol{t}$-reduced, we can apply the predictable degree property: $\mathrm{rdeg}_{\boldsymbol{t}}(\mathbf{K}_1\mathbf{F}) = \mathrm{rdeg}_{\boldsymbol{u}}(\mathbf{K}_1)$. This proves the sought inequality. For the last point, note that the $(\boldsymbol{u}, \boldsymbol{v})$-pivot entries of $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ located in $\mathbf{K}_2$ correspond to $\boldsymbol{v}$-pivot entries in $\mathbf{K}_2$. Thus, since $[\mathbf{K}_1 \ \ \mathbf{K}_2]$ is in $(\boldsymbol{u}, \boldsymbol{v})$-weak Popov form with all $(\boldsymbol{u}, \boldsymbol{v})$-pivot entries in $\mathbf{K}_2$, it follows that the $\boldsymbol{v}$-pivot index of $\mathbf{K}_2$ is increasing. □

2.5. Basic subroutines and their complexity

To conclude these preliminaries, we recall known fast algorithms for three polynomial matrix subroutines used in our determinant algorithm: multiplication with unbalanced degrees, minimal approximant bases, and minimal kernel bases; we give the corresponding complexity estimates adapted to our context and in particular using our framework stated in Section 1.1.

Unbalanced multiplication. Polynomial matrix algorithms often involve multiplication with matrix operands whose entries have degrees that may be unbalanced but still satisfy properties that can be exploited to perform the multiplication efficiently. Here we will encounter products of reduced matrices with degree properties similar to those discussed in [58, Sec. 3.6], where an efficient approach for computing such products was given.

Lemma 2.8. *There is an algorithm* UnbalancedMultiplication *which takes as input a matrix $\mathbf{B} \in \mathbb{K}[x]^{k\times m}$ with $k \leq m$, a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$ with $n \leq m$, and an integer D greater than or equal to both the sum of the positive entries of $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})$ and that of $\mathrm{rdeg}_{\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})}(\mathbf{B})$, and returns the product $\mathbf{BA}$ using $O(m^\omega\, \mathsf{M}(D/m))$ operations in $\mathbb{K}$, assuming $\mathcal{H}_{\mathrm{sm}}$ and $\mathcal{H}_\omega$.*

Proof. Zhou et al. [58, Sec. 3.6] gave such an algorithm, yet with a cost analysis which hides logarithmic factors; because these factors are our main concern here we will rely on the version in [27, Sec. 4]. In this reference, Algorithm UnbalancedMultiplication was described for square matrices. One could adapt it to the case of rectangular $\mathbf{A}$ and $\mathbf{B}$ as in the statement above.

However, for the sake of conciseness and with no impact on the asymptotic cost bound, we consider the more basic approach of forming the square $m \times m$ matrices $\mathbf{D} = \left[\begin{smallmatrix}\mathbf{B}\\ \mathbf{0}\end{smallmatrix}\right]$ and $\mathbf{C} = [\mathbf{A} \;\; \mathbf{0}]$, computing $\mathbf{DC}$ using the above-cited algorithm, and retrieving $\mathbf{BA}$ from it. Now, by construction, both the sum of the positive entries of $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{C})$ and that of $\mathrm{rdeg}_{\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})}(\mathbf{D})$ are at most D, hence [27, Prop. 4.1] applies: defining $\bar{m}$ and d as the smallest powers of 2 greater than or equal to m and D/m, it states that the computation of $\mathbf{DC}$ costs $O(\sum_{0\le i\le \log_2(\bar{m})} 2^i (2^{-i}\bar{m})^\omega \,\mathsf{M}(2^i d))$ operations in $\mathbb{K}$. Using $\mathcal{H}_{\mathrm{sm}}$ and $\mathcal{H}_\omega$, which ensure respectively that $\mathsf{M}(2^i d) \le \mathsf{M}(2^i)\,\mathsf{M}(d)$ and $\mathsf{M}(2^i) \in O(2^{i(\omega-1-\varepsilon)})$ for some $\varepsilon > 0$, we obtain that this bound is in $O(\bar{m}^\omega \,\mathsf{M}(d) \sum_{0\le i\le \log_2(\bar{m})} 2^{-i\varepsilon}) \subseteq O(\bar{m}^\omega \,\mathsf{M}(d))$. This is in $O(m^\omega \,\mathsf{M}(D/m))$, since $\bar{m} \in \Theta(m)$ and $d \in \Theta(1 + D/m)$. □

Minimal approximant basis. The second basic tool we will use is approximant bases computation; for this, we will use the algorithm PM-Basis, originally described in [21]. Precisely, we rely on the slightly modified version presented in [28] which ensures that the computed basis is in shifted weak Popov form.

Lemma 2.9. *There is an algorithm* PM-Basis *which takes as input a tuple $\boldsymbol{\gamma} \in \mathbb{Z}^n_{>0}$, a matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ with $\mathrm{cdeg}(\mathbf{F}) < \boldsymbol{\gamma}$, and a shift $\boldsymbol{s} \in \mathbb{Z}^m$, and returns a basis of $\mathcal{A}_{\boldsymbol{\gamma}}(\mathbf{F})$ in $\boldsymbol{s}$-weak Popov form using $O((m+n)m^{\omega-1}\,\mathsf{M}'(\max(\boldsymbol{\gamma})))$ operations in $\mathbb{K}$.*

Proof. The algorithm is [28, Algo. 2]; to accommodate non-uniform order $\boldsymbol{\gamma}$, it is called with input order $\Gamma = \max(\boldsymbol{\gamma})$ and input matrix $\mathbf{F}\mathbf{X}^{(\Gamma,\dots,\Gamma)-\boldsymbol{\gamma}}$ as explained in [28, Rmk. 3.3]. According to [28, Prop. 3.2], this costs $O((1+\frac{n}{m}) \sum_{0\le i\le \lceil \log_2(\Gamma)\rceil} 2^i m^\omega \,\mathsf{M}(2^{-i}\Gamma))$ operations in $\mathbb{K}$, which is precisely the claimed bound by definition of $\mathsf{M}'(\cdot)$. □

Minimal kernel basis. We will make use of the algorithm of Zhou et al. [58], which itself relies on unbalanced products and approximant bases, and returns a kernel basis in shifted reduced form efficiently for input matrices with small average row degree.

Lemma 2.10. *There is an algorithm* KernelBasis *which takes as input a full column rank matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ with $m \ge n$ and $m \in O(n)$, and a shift $\boldsymbol{s} \in \mathbb{Z}^m_{\ge 0}$ such that $\boldsymbol{s} \ge \mathrm{rdeg}_{\mathbf{0}}(\mathbf{F})$, and returns a basis of $\mathcal{K}(\mathbf{F})$ in $\boldsymbol{s}$-reduced form using $O(m^\omega\,\mathsf{M}'(D/m))$ operations in $\mathbb{K}$, assuming $\mathcal{H}_{\mathrm{sl}}, \mathcal{H}_{\mathrm{sm}}, \mathcal{H}_\omega$. Here D is the sum of the entries of $\boldsymbol{s}$, and the sum of the $\boldsymbol{s}$-row degree of this basis is at most D.*

Proof. The algorithm of Zhou et al. [58] computes an $\boldsymbol{s}$-reduced basis $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$ of $\mathcal{K}(\mathbf{F})$; precisely, this reference is about computing a basis of the right kernel in column reduced form, yet this naturally translates into left kernels and row reduced forms by taking suitable transposes. Furthermore, the last claim in the lemma follows from [58, Thm. 3.4], which states that any such basis $\mathbf{K}$ is such that $|\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{K})| \le |\boldsymbol{s}| = D$. For the complexity, we rely on the analysis in [27, Prop. B.1] which shows that, defining $\bar{m}$ and d as the smallest powers of 2 greater than or equal to m and D/m, this computation costs

$$O\left(\sum_{j=0}^{\log_2(\bar{m})} 2^j \left(\sum_{i=0}^{\log_2(2^{-j}\bar{m})} 2^i (2^{-i-j}\bar{m})^\omega \,\mathsf{M}\big(2^{i+j}d\big) + \sum_{i=0}^{\log_2(2^j d)} 2^i (2^{-j}\bar{m})^\omega \,\mathsf{M}\big(2^{j-i}d\big)\right)\right)$$

operations in $\mathbb{K}$. Now the same analysis as in the proof of Lemma 2.8 shows that, assuming $\mathcal{H}_{\mathrm{sm}}$ and $\mathcal{H}_\omega$, the first inner sum is in $O((2^{-j}\bar{m})^\omega\,\mathsf{M}(2^j d))$, and by definition of $\mathsf{M}'(\cdot)$ the second inner sum is in $O((2^{-j}\bar{m})^\omega\,\mathsf{M}'(2^j d))$. Thus the total cost is in $O(\sum_{0\le j\le \log_2(\bar{m})} 2^j (2^{-j}\bar{m})^\omega\,\mathsf{M}'(2^j d))$, which is in $O(\bar{m}^\omega\,\mathsf{M}'(d)\sum_{0\le j\le \log_2(\bar{m})} 2^{j(1-\omega)}\,\mathsf{M}'(2^j))$ since $\mathcal{H}_{\mathrm{sl}}$ and $\mathcal{H}_{\mathrm{sm}}$ ensure that $\mathsf{M}'(\cdot)$ is sub-multiplicative. Similarly to the proof of Lemma 2.8, this bound is in $O(m^\omega\,\mathsf{M}'(D/m))$ thanks to $\mathcal{H}_\omega$. □

3. Determinant algorithm for reduced matrices

In this section, we present the main algorithm in this paper, which computes the determinant of a matrix in reduced form using the subroutines listed in Section 2.5 as well as the algorithms ReducedToWeakPopov and WeakPopovToPopov from Theorems 1.3 and 1.4. Taking for granted the proof of these theorems in Sections 4 and 5, here we prove the correctness of our determinant algorithm in Section 3.2 and analyse its complexity in Section 3.3, thus proving Theorem 1.2.

3.1. Two properties of determinants of reduced matrices

Leading coefficient of the determinant. All bases of a given submodule of $\mathbb{K}[x]^{1\times n}$ of rank n have the same determinant up to a constant factor, i.e. up to multiplication by an element of $\mathbb{K}\setminus\{0\}$. Many algorithms operating on polynomial matrices such as PM-Basis and KernelBasis compute such bases, so that their use in a determinant algorithm typically leads to obtaining the sought determinant up to a constant factor; then finding the actual determinant requires to efficiently recover this constant [see e.g. 32, Sec. 4]. Since in this paper we seek determinants of matrices in reduced form, this issue is easily handled using the next result.

Lemma 3.1. *Let $\boldsymbol{s}\in\mathbb{Z}^n$ and $\mathbf{A}\in\mathbb{K}[x]^{n\times n}$. If $\mathbf{A}$ is in $\boldsymbol{s}$-reduced form, the leading coefficient of* $\det(\mathbf{A})$ *is* $\det(\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A}))$. *In particular, if $\mathbf{A}$ is in $\boldsymbol{s}$-weak Popov form, the leading coefficient of* $\det(\mathbf{A})$ *is the product of the leading coefficients of the diagonal entries of $\mathbf{A}$.*

Proof. The second claim is a direct consequence of the first, since for $\mathbf{A}$ in $\boldsymbol{s}$-weak Popov form, $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ is lower triangular with diagonal entries equal to the leading coefficients of the diagonal entries of $\mathbf{A}\mathbf{X}^{\boldsymbol{s}}$, which are the leading coefficients of the diagonal entries of $\mathbf{A}$. For the first claim in the case $\boldsymbol{s}=\mathbf{0}$, we refer to [29, Sec. 6.3.2], and in particular Eq. (23) therein. Now, for an arbitrary $\boldsymbol{s}$ and $\mathbf{A}$ in $\boldsymbol{s}$-reduced form, we consider the nonnegative shift $\boldsymbol{t}=\boldsymbol{s}-(\min(\boldsymbol{s}),\dots,\min(\boldsymbol{s}))$ and observe that $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})=\mathrm{lm}_{\mathbf{0}}(\mathbf{A}\mathbf{X}^{\boldsymbol{t}})$, hence $\mathbf{A}\mathbf{X}^{\boldsymbol{t}}$ is $\mathbf{0}$-reduced, and thus the leading coefficient of $\det(\mathbf{A}\mathbf{X}^{\boldsymbol{t}})=\det(\mathbf{A})\det(\mathbf{X}^{\boldsymbol{t}})$ (which is the same as that of $\det(\mathbf{A})$) is equal to $\det(\mathrm{lm}_{\mathbf{0}}(\mathbf{A}\mathbf{X}^{\boldsymbol{t}}))=\det(\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A}))$. □

From shifted to non-shifted. In Section 1.5, we explained that one step of our algorithm consists in finding $\det(\mathbf{K})$ for a matrix $\mathbf{K}$ in $\boldsymbol{v}$-weak Popov form, and that it achieves this by computing the $-\boldsymbol{d}$-Popov form of $\mathbf{K}^{\mathsf{T}}$, which is already in $-\boldsymbol{d}$-weak Popov form. The next lemma substantiates this; note that in Section 1.5 we had left out the reversal matrix $\mathbf{J}_n$ for the sake of exposition.

Lemma 3.2. *Let $\boldsymbol{v}\in\mathbb{Z}^n$, let $\mathbf{K}\in\mathbb{K}[x]^{n\times n}$, and let $\boldsymbol{d}=\mathrm{rdeg}_{\boldsymbol{v}}(\mathbf{K})$.*

(a) if $\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})$ has no zero column, then $\mathrm{lm}_{-\boldsymbol{d}}(\mathbf{K}^{\mathsf{T}})=\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})^{\mathsf{T}}$ and $\mathrm{rdeg}_{-\boldsymbol{d}}(\mathbf{K}^{\mathsf{T}})=-\boldsymbol{v}$;

(b) if $\mathbf{K}$ is in $\boldsymbol{v}$-reduced form, then $\mathbf{K}^{\mathsf{T}}$ is in $-\boldsymbol{d}$-reduced form;

(c) if $\mathbf{K}$ is in $\boldsymbol{v}$-weak Popov form, then $\mathbf{J}_n\mathbf{K}^{\mathsf{T}}\mathbf{J}_n$ is in $-\boldsymbol{d}\mathbf{J}_n$-weak Popov form;

(d) if furthermore $\mathbf{P}$ is the $-\boldsymbol{d}\mathbf{J}_n$-Popov form of $\mathbf{J}_n\mathbf{K}^{\mathsf{T}}\mathbf{J}_n$, then $\mathbf{P}^{\mathsf{T}}$ is in $\mathbf{0}$-weak Popov form and $\det(\mathbf{K})=\det(\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K}))\det(\mathbf{P}^{\mathsf{T}})$.

Proof. By definition, $\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})^{\mathsf{T}}$ is the coefficient of degree 0 of $(\mathbf{X}^{-\boldsymbol{d}}\mathbf{K}\mathbf{X}^{\boldsymbol{v}})^{\mathsf{T}}=\mathbf{X}^{\boldsymbol{v}}\mathbf{K}^{\mathsf{T}}\mathbf{X}^{-\boldsymbol{d}}$, which is a matrix over $\mathbb{K}[x^{-1}]$. The assumption on $\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})$ implies that this coefficient of degree 0 has no zero row. It follows that $\mathrm{rdeg}_{-\boldsymbol{d}}(\mathbf{K}^{\mathsf{T}})=-\boldsymbol{v}$ and that this coefficient of degree 0 is $\mathrm{lm}_{-\boldsymbol{d}}(\mathbf{K}^{\mathsf{T}})$. Item (b) follows from Item (a) by definition of shifted reduced forms.

From now on, we assume that $\mathbf{K}$ is in $\boldsymbol{v}$-weak Popov form. Then $\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})$ is invertible and lower triangular, and in particular $\mathrm{lm}_{-\boldsymbol{d}}(\mathbf{K}^{\mathsf{T}})=\mathrm{lm}_{\boldsymbol{v}}(\mathbf{K})^{\mathsf{T}}$. Since $\mathbf{J}_n$ is a permutation matrix, we

obtain $\mathrm{lm}_{-\boldsymbol{d}\mathbf{J}_n}(\mathbf{J}_n\mathbf{K}^\mathsf{T}\mathbf{J}_n) = \mathbf{J}_n\mathrm{lm}_{-\boldsymbol{d}}(\mathbf{K}^\mathsf{T})\mathbf{J}_n = \mathbf{J}_n\mathrm{lm}_{\boldsymbol{\nu}}(\mathbf{K})^\mathsf{T}\mathbf{J}_n$, which is invertible and lower triangular. Hence $\mathbf{J}_n\mathbf{K}^\mathsf{T}\mathbf{J}_n$ is in $-\boldsymbol{d}\mathbf{J}_n$-weak Popov form.

For Item (d), since $\mathbf{P}$ is $n \times n$ and in $-\boldsymbol{d}\mathbf{J}_n$-Popov form, we have $\mathrm{lm}_{\mathbf{0}}(\mathbf{P}^\mathsf{T}) = \mathbf{I}_n$, hence $\mathbf{P}^\mathsf{T}$ is in $\mathbf{0}$-weak Popov form. Furthermore, since $\mathbf{J}_n\mathbf{K}^\mathsf{T}\mathbf{J}_n$ is unimodularly equivalent to $\mathbf{P}$, its determinant is $\det(\mathbf{K}) = \det(\mathbf{J}_n\mathbf{K}^\mathsf{T}\mathbf{J}_n) = \lambda\det(\mathbf{P})$ for some $\lambda \in \mathbb{K} \setminus \{0\}$. Applying Lemma 3.1 to $\mathbf{P}$ shows that $\det(\mathbf{P})$ is monic, hence λ is the leading coefficient of $\det(\mathbf{K})$; applying the same lemma to $\mathbf{K}$ yields $\lambda = \det(\mathrm{lm}_{\boldsymbol{\nu}}(\mathbf{K}))$. □

3.2. Algorithm and correctness

Our main determinant algorithm is DeterminantOfWeakPopov (Algorithm 2), which takes as input a matrix in $\mathbf{0}$-weak Popov form and computes its determinant using recursive calls on matrices of smaller dimension. Then, we compute the determinant of a $\mathbf{0}$-reduced matrix by first calling ReducedToWeakPopov to find a $\mathbf{0}$-weak Popov matrix which has the same determinant up to a nonzero constant, and then calling the previous algorithm on that matrix. This is detailed in Algorithm 1.

Algorithm 1 DeterminantOfReduced($\mathbf{A}$)

Input: a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times m}$ in $\mathbf{0}$-reduced form.
Output: the determinant of $\mathbf{A}$.

1: $\mathbf{P} \in \mathbb{K}[x]^{m\times m} \leftarrow$ ReducedToWeakPopov($\mathbf{A}, \mathbf{0}$)
2: $\Delta \leftarrow$ DeterminantOfWeakPopov($\mathbf{P}$); $\ell_\Delta \in \mathbb{K} \setminus \{0\} \leftarrow$ leading coefficient of Δ
3: **return** $\det(\mathrm{lm}_{\mathbf{0}}(\mathbf{A}))\,\Delta/\ell_\Delta$

The correctness of Algorithm 1 is obvious: according to Theorem 1.3 and Proposition 3.3, $\mathbf{P}$ is a $\mathbf{0}$-weak Popov form of $\mathbf{A}$, and Δ is the determinant of $\mathbf{P}$ up to multiplication by some element of $\mathbb{K} \setminus \{0\}$. Thus $\det(\mathbf{A}) = \ell\Delta$ for some $\ell \in \mathbb{K} \setminus \{0\}$, and Lemma 3.1 yields $\ell = \det(\mathrm{lm}_{\mathbf{0}}(\mathbf{A}))/\ell_\Delta$.

Concerning the cost bound, Theorem 1.3 states that the first step uses $O(m^\omega(1 + D/m))$ operations in $\mathbb{K}$, where $D = |\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})|$; since $\mathbf{A}$ is $\mathbf{0}$-reduced, this is $D = \deg(\det(\mathbf{A}))$ [29, Sec. 6.3.2]. In the last step, the determinant computation costs $O(m^\omega)$ operations, while scaling Δ by a constant costs $O(D)$ operations. The second step uses $O(m^\omega\,\mathsf{M}'(D/m))$ operations according to Proposition 3.3, under the assumptions $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_\omega$.

We now describe the main algorithm of this paper (Algorithm 2) and focus on its correctness. We also mention cost bounds for all steps of the algorithm that are not recursive calls, but we defer the core of the complexity analysis to Section 3.3.

Proposition 3.3. *Algorithm 2 is correct, and assuming that $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_\omega$ hold (hence in particular $\omega > 2$), it uses $O(m^\omega\,\mathsf{M}'(D/m))$ operations in $\mathbb{K}$.*

Proof of correctness. The fact that $\mathbf{A}$ is in $\mathbf{0}$-weak Popov form has two consequences on the tuple $\boldsymbol{s}$ computed at Line 1: first, it is the $\mathbf{0}$-pivot degree of $\mathbf{A}$ (i.e. its diagonal degrees), and second, the sum $D = |\boldsymbol{s}|$ is equal to the degree of the determinant of $\mathbf{A}$ [29, Sec. 6.3.2].

The main base case of the recursion is when $m = 1$ and is handled at Line 2; it uses no operation in $\mathbb{K}$. We use a second base case at Line 3: if $D = 0$, then $\mathbf{A}$ is an $m \times m$ matrix over $\mathbb{K}$. Since it is in $\mathbf{0}$-weak Popov, it is invertible and lower triangular, hence $\det(\mathbf{A})$ is the product of its diagonal entries, which is computed in $O(m)$ multiplications in $\mathbb{K}$. This base case is not necessary for obtaining the correctness and the cost bound in Proposition 3.3; still, not using it would incur a cost of $O(m^\omega)$ operations in the case $D = 0$.

Algorithm 2 DeterminantOfWeakPopov($\mathbf{A}$)

Input: a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times m}$ in $\mathbf{0}$-weak Popov form.
Output: the determinant of $\mathbf{A}$, up to multiplication by some element of $\mathbb{K} \setminus \{0\}$.

1: $\boldsymbol{s} = (s_1, \ldots, s_m) \in \mathbb{Z}^m_{\geq 0} \leftarrow \mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})$; $D \leftarrow s_1 + \cdots + s_m$ ▷ *D = degree of* $\det(\mathbf{A})$
2: **if** $m = 1$ **then return** the polynomial f such that $\mathbf{A} = [f]$ ▷ *base case: 1×1 matrix*
3: **if** $D = 0$ **then return** the product of diagonal entries of $\mathbf{A}$ ▷ *base case: matrix over $\mathbb{K}$*
4: **if** $D < m$ **then** ▷ *handle constant rows to reduce to dimension $\leq D$*
5: $\quad \mathbf{B} \leftarrow \mathbf{A}\,\mathrm{lm}_0(\mathbf{A})^{-1}$ from which rows and columns with indices in $\{i \mid s_i = 0\}$ are removed
6: $\quad$ **return** DeterminantOfWeakPopov($\mathbf{B}$)
7: **if** $s_1 + \cdots + s_{\lfloor m/2 \rfloor} > D/2$ **then return** DeterminantOfWeakPopov($\mathbf{J}_m \mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}\mathbf{J}_m$)
8: Write $\mathbf{A} = \left[\begin{smallmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{smallmatrix}\right]$, with $\mathbf{A}_1$ of size $\lfloor m/2 \rfloor \times \lfloor m/2 \rfloor$ and $\mathbf{A}_4$ of size $\lceil m/2 \rceil \times \lceil m/2 \rceil$
9: $\mathbf{K} \in \mathbb{K}[x]^{\lceil m/2 \rceil \times m} \leftarrow$ KernelBasis($\left[\begin{smallmatrix} \mathbf{A}_1 \\ \mathbf{A}_3 \end{smallmatrix}\right], \boldsymbol{s}$)
10: $[\mathbf{K}_1 \;\; \mathbf{K}_2] \in \mathbb{K}[x]^{\lceil m/2 \rceil \times m} \leftarrow$ ReducedToWeakPopov($\mathbf{K}, \boldsymbol{s}$), where $\mathbf{K}_2$ is $\lceil m/2 \rceil \times \lceil m/2 \rceil$
11: $\mathbf{B} \leftarrow$ UnbalancedMultiplication($[\mathbf{K}_1 \;\; \mathbf{K}_2], \left[\begin{smallmatrix} \mathbf{A}_2 \\ \mathbf{A}_4 \end{smallmatrix}\right], D$) ▷ $\mathbf{B} = \mathbf{K}_1\mathbf{A}_2 + \mathbf{K}_2\mathbf{A}_4$
12: $\Delta_1 \leftarrow$ DeterminantOfWeakPopov($\mathbf{B}$) ▷ *first recursive call*
13: $\Delta_2 \leftarrow$ DeterminantOfWeakPopov($\mathbf{A}_1$) ▷ *second recursive call*
14: $\mathbf{P} \leftarrow$ WeakPopovToPopov($\mathbf{J}_{\lceil m/2 \rceil}\mathbf{K}_2^{\mathsf{T}}\mathbf{J}_{\lceil m/2 \rceil}, \mathrm{rdeg}_{(s_{\lfloor m/2 \rfloor+1}, \ldots, s_m)}(\mathbf{K}_2)\mathbf{J}_{\lceil m/2 \rceil}$)
15: $\Delta_3 \leftarrow$ DeterminantOfWeakPopov($\mathbf{P}^{\mathsf{T}}$) ▷ *third recursive call*
16: **return** $\Delta_1\Delta_2/\Delta_3$

For the recursion, we proceed inductively: we assume that the algorithm correctly computes the determinant for all $\mathbf{0}$-weak Popov matrices of dimension less than m, and based on this we show that it is also correct for any $\mathbf{0}$-weak Popov matrix $\mathbf{A}$ of dimension m.

Case 1: $D < m$. Then $\mathbf{A}$ has at least one constant row; using linear algebra we reduce to the case of a matrix $\mathbf{B}$ of dimension at most D with all rows of degree at least 1. Since $\boldsymbol{s} = \mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})$, we can write $\mathbf{A} = \mathbf{X}^{\boldsymbol{s}}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A}) + \mathbf{R}$ for a matrix $\mathbf{R} \in \mathbb{K}[x]^{m\times m}$ such that $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{R}) < \boldsymbol{s}$. Since $\mathbf{A}$ is $\mathbf{0}$-reduced, $\mathrm{lm}_{\mathbf{0}}(\mathbf{A})$ is invertible and $\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1} = \mathbf{X}^{\boldsymbol{s}} + \mathbf{R}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$ with $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{R}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}) < \boldsymbol{s}$, which implies $\mathrm{lm}_{\mathbf{0}}(\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}) = \mathbf{I}_m$. In particular, $\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$ is in $\mathbf{0}$-weak Popov form, and for each i such that the row $\mathbf{A}_{i,*}$ is constant, i.e. $s_i = 0$, the ith row of $\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$ is the ith row of the identity matrix. Therefore the matrix $\mathbf{B}$ at Line 5 is in $\mathbf{0}$-weak Popov form, has the same determinant as $\mathbf{A}$ up to a constant, and has dimension $\#\{i \mid s_i \neq 0\} \leq D$. Hence the correctness in this case. In terms of complexity, computing $\mathbf{B}$ essentially amounts to computing the product $\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$, which is done by row-wise expanding $\mathbf{A}$ into a $(m + D) \times m$ matrix over $\mathbb{K}$, right-multiplying by $\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$, and compressing the result back into a polynomial matrix: this costs $O(m^\omega(1 + D/m)) \subseteq O(m^\omega)$ operations.

Case 2: $s_1 + \cdots + s_{\lfloor m/2 \rfloor} > D/2$. Then we modify the input $\mathbf{A}$ so as to reduce to *Case 3*. As we have seen above, $\mathrm{lm}_{\mathbf{0}}(\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}) = \mathbf{I}_m$. We now reverse the diagonal entries by reversing the order of rows and columns: let $\mathbf{B} = \mathbf{J}_m\mathbf{A}\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}\mathbf{J}_m$. Then $\mathrm{lm}_{\mathbf{0}}(\mathbf{B}) = \mathbf{J}_m\mathrm{lm}_{\mathbf{0}}(\mathbf{A}\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1})\mathbf{J}_m = \mathbf{I}_m$, hence $\mathbf{B}$ is in $\mathbf{0}$-weak Popov form: Line 7 calls the algorithm on this matrix to obtain $\det(\mathbf{B})$ up to a constant, and this yields $\det(\mathbf{A})$ since it is equal to $\det(\mathrm{lm}_{\mathbf{0}}(\mathbf{A}))\det(\mathbf{B})$. To conclude the proof of correctness in that case (assuming correctness in *Case 3*), it remains to observe that $\mathbf{B}$ has the same matrix dimension m as $\mathbf{A}$, and that the matrix $\mathbf{B}$ has degrees such that calling the algorithm with input $\mathbf{B}$ does not enter *Case 2* but *Case 3*. Indeed, we have $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{B}) = \boldsymbol{s}\mathbf{J}_m$, hence the sum of the first $\lfloor m/2 \rfloor$ entries of the tuple $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{B})$ is $s_m + \cdots + s_{\lceil m/2 \rceil+1} = D - (s_1 + \cdots + s_{\lceil m/2 \rceil})$, which is at most $D/2$ by assumption. In terms of complexity, the main step is to compute the

product $\mathbf{A}\,\mathrm{lm}_{\mathbf{0}}(\mathbf{A})^{-1}$, which costs $O(m^\omega(1+D/m))$ operations as we have seen above; this is in $O(m^\omega\,\mathsf{M}'(D/m))$.

Case 3: $s_1+\cdots+s_{\lfloor m/2\rfloor}\le D/2$. Then, Line 7 performs no action, and Line 8 defines submatrices of $\mathbf{A}$. By construction, $\left[\begin{smallmatrix}\mathbf{A}_1\\ \mathbf{A}_3\end{smallmatrix}\right]$ has full column rank and $\boldsymbol{s}\ge \mathrm{rdeg}_{\mathbf{0}}(\left[\begin{smallmatrix}\mathbf{A}_1\\ \mathbf{A}_3\end{smallmatrix}\right])$ holds. Thus, according to Lemma 2.10, Line 9 uses $O(m^\omega\,\mathsf{M}'(D/m))$ operations to compute an $\boldsymbol{s}$-reduced basis $\mathbf{K}$ of $\mathcal{K}(\left[\begin{smallmatrix}\mathbf{A}_1\\ \mathbf{A}_3\end{smallmatrix}\right])$, with $|\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{K})|\le D$. Then, Theorem 1.3 states that Line 10 transforms $\mathbf{K}$ into an $\boldsymbol{s}$-weak Popov basis $[\mathbf{K}_1\ \ \mathbf{K}_2]$ of this kernel at a cost of $O(m^\omega(1+D/m))$ operations, since $|\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{K})|\le D$ and $\min(\boldsymbol{s})\ge 0$. Since all $\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have the same $\boldsymbol{s}$-row degree up to permutation, $|\mathrm{rdeg}_{\boldsymbol{s}}([\mathbf{K}_1\ \ \mathbf{K}_2])|\le D$ holds, hence the assumptions of Lemma 2.8 are satisfied and Line 11 uses $O(m^\omega\,\mathsf{M}(D/m))$ operations to compute $\mathbf{B}=\mathbf{K}_1\mathbf{A}_2+\mathbf{K}_2\mathbf{A}_4$.

The important observation at this stage is the identity

$$\begin{bmatrix}\mathbf{I}_{\lceil m/2\rceil} & \mathbf{0}\\ \mathbf{K}_1 & \mathbf{K}_2\end{bmatrix}\begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{A}_3 & \mathbf{A}_4\end{bmatrix}=\begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{0} & \mathbf{B}\end{bmatrix} \tag{2}$$

which, provided that $\mathbf{K}_2$ is nonsingular, implies $\det(\mathbf{A})=\det(\mathbf{B})\det(\mathbf{A}_1)/\det(\mathbf{K}_2)$. We are going to show that this is the formula used in Line 16 to compute $\det(\mathbf{A})$.

First, $\mathbf{A}_1$ has dimension less than m and, being a principal submatrix of the $\mathbf{0}$-weak Popov matrix $\mathbf{A}$, it is also in $\mathbf{0}$-weak Popov form. Hence the recursive call at Line 13 is sound and Δ_2 is equal to $\det(\mathbf{A}_1)$ up to a constant.

Since $\mathbf{A}$ is in $\mathbf{0}$-weak Popov form and $\left[\begin{smallmatrix}\mathbf{I}_{\lceil m/2\rceil} & \mathbf{0}\\ \mathbf{K}_1 & \mathbf{K}_2\end{smallmatrix}\right]$ is in $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})$-weak Popov form, their product $\left[\begin{smallmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{0} & \mathbf{B}\end{smallmatrix}\right]$ is in $\mathbf{0}$-weak Popov form; see Lemma 2.6, or note that $\mathrm{lm}_{\mathbf{0}}(\left[\begin{smallmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{0} & \mathbf{B}\end{smallmatrix}\right])$ is invertible and lower triangular according to Lemma 2.2. It follows that $\mathbf{B}$ is in $\mathbf{0}$-weak Popov form and has dimension less than m: Line 12 recursively computes Δ_1, equal to $\det(\mathbf{B})$ up to a constant.

It remains to prove that Δ_3 computed at Lines 14 and 15 is equal to $\det(\mathbf{K}_2)$ up to a constant. Let $\boldsymbol{\nu}=\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A}_4)=(s_{\lfloor m/2\rfloor+1},\ldots,s_m)$ be the shift used at Line 14, and let $\boldsymbol{d}=\mathrm{rdeg}_{\boldsymbol{\nu}}(\mathbf{K}_2)=\mathrm{rdeg}_{\boldsymbol{s}}([\mathbf{K}_1\ \ \mathbf{K}_2])$. Applying Lemma 2.7 (with $\mathbf{F}=\mathbf{A}_1$, $\mathbf{G}=\mathbf{A}_3$, $\boldsymbol{t}=\mathbf{0}$, and $\boldsymbol{\nu}$ as above) shows that $[\mathbf{K}_1\ \ \mathbf{K}_2]$ has all its $\boldsymbol{s}$-pivot entries in $\mathbf{K}_2$, and in particular $\mathbf{K}_2$ is in $\boldsymbol{\nu}$-weak Popov form. Let $\boldsymbol{\delta}\in\mathbb{Z}^n_{\ge 0}$ be the $\boldsymbol{\nu}$-pivot degree of $\mathbf{K}_2$, where $n=\lceil m/2\rceil$, and note that $\boldsymbol{d}=\boldsymbol{\delta}+\boldsymbol{\nu}\ge\boldsymbol{\delta}$ since $\boldsymbol{\nu}\ge\mathbf{0}$. Then, Lemma 3.2 states that $\mathbf{J}_n\mathbf{K}_2^{\mathsf{T}}\mathbf{J}_n$ is in $-\boldsymbol{d}\mathbf{J}_n$-weak Popov form; its $-\boldsymbol{d}\mathbf{J}_n$-pivot degree is the list of degrees of its diagonal entries, that is, $\boldsymbol{\delta}\mathbf{J}_n$. Since $\boldsymbol{d}\mathbf{J}_n\ge\boldsymbol{\delta}\mathbf{J}_n$, we can apply Theorem 1.4, which implies that Line 14 computes the $-\boldsymbol{d}\mathbf{J}_n$-Popov form $\mathbf{P}$ of $\mathbf{J}_n\mathbf{K}_2^{\mathsf{T}}\mathbf{J}_n$ using $O(m^\omega\,\mathsf{M}'(|\boldsymbol{d}|/m))$ operations; as we have seen above, $|\boldsymbol{d}|=|\mathrm{rdeg}_{\boldsymbol{s}}([\mathbf{K}_1\ \ \mathbf{K}_2])|\le D$. Then, from the last item of Lemma 3.2, $\mathbf{P}^{\mathsf{T}}$ is in $\mathbf{0}$-weak Popov form and $\det(\mathbf{K})=\det(\mathrm{lm}_{\boldsymbol{\nu}}(\mathbf{K}))\det(\mathbf{P}^{\mathsf{T}})$, hence Line 15 correctly computes $\det(\mathbf{K})$ up to a constant. □

To conclude this presentation of our determinant algorithm, we note that it would be beneficial, in a practical implementation, to add an early exit. Precisely, just after computing $\det(\mathbf{B})$ at Line 12, one could perform the following action before (possibly) proceeding to the next steps:

12b: **if** $\deg(\Delta_1)=D$ **then return** Δ_1 ▷ *early exit*

Indeed, recall that Δ_1 is $\det(\mathbf{B})$ up to a constant; furthermore we claim that

- for a generic $\mathbf{A}$, we have $\deg(\Delta_1)=D$,
- if $\deg(\Delta_1)=D$ (i.e. $\deg(\det(\mathbf{B}))=D$), then $\det(\mathbf{B})$ is $\det(\mathbf{A})$ up to a constant.

It follows that for a generic matrix $\mathbf{A}$, then Δ_1 is $\det(\mathbf{A})$ up to a constant, hence the correctness of this early exit (see also [21, Sec. 4.2.2] for similar considerations). To prove the above claim,

first note that since $[\mathbf{K}_1 \;\; \mathbf{K}_2]$ is a kernel basis, it has unimodular column bases [22, Lem. 2.2], and thus it can be completed into a unimodular matrix $\mathbf{U} = \left[\begin{smallmatrix}\mathbf{U}_1 & \mathbf{U}_2\\ \mathbf{K}_1 & \mathbf{K}_2\end{smallmatrix}\right] \in \mathbb{K}[x]^{m\times m}$ [57, Lem. 2.10]. Therefore

$$\mathbf{UA} = \begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\\ \mathbf{K}_1 & \mathbf{K}_2\end{bmatrix}\begin{bmatrix}\mathbf{A}_1 & \mathbf{A}_2\\ \mathbf{A}_3 & \mathbf{A}_4\end{bmatrix} = \begin{bmatrix}\mathbf{B}_1 & \mathbf{B}_2\\ \mathbf{0} & \mathbf{B}\end{bmatrix}$$

where $[\mathbf{B}_1 \;\; \mathbf{B}_2] = [\mathbf{U}_1 \;\; \mathbf{U}_2]\mathbf{A}$. Since $\det(\mathbf{U})$ is in $\mathbb{K}\setminus\{0\}$, $\det(\mathbf{A})$ is $\det(\mathbf{B}_1)\det(\mathbf{B})$ up to a constant. For the second item of the claim, $\deg(\Delta_1) = D$ implies $\deg(\det(\mathbf{B})) = D = \deg(\det(\mathbf{A}))$, hence $\det(\mathbf{B}_1)$ is in $\mathbb{K}\setminus\{0\}$. The first item follows from the fact that $\mathbf{B}_1$ is a row basis of $\left[\begin{smallmatrix}\mathbf{A}_1\\ \mathbf{A}_3\end{smallmatrix}\right]$ [56, Lem. 3.1]; since the latter matrix has more rows than columns, if $\mathbf{A}_1$ and $\mathbf{A}_3$ have generic entries, then such a row basis $\mathbf{B}_1$ is unimodular which means $\det(\mathbf{B}_1) \in \mathbb{K}\setminus\{0\}$ and thus $\deg(\Delta_1) = D$.

3.3. Complexity analysis

We have seen above that all computations in Algorithm 2 other than recursive calls have an arithmetic cost in $O(m^\omega \mathsf{M}'(D/m))$ operations in $\mathbb{K}$; here, we complete the proof of the cost bound in Proposition 5.14. In this section, we use the assumptions $\mathcal{H}_{\mathrm{sl}}$, $\mathcal{H}_{\mathrm{sm}}$, and $\mathcal{H}_\omega$ as well as their consequences stated in Section 1.1.

Let $C(m, D)$ denote the arithmetic cost of Algorithm 2; recall that D is the degree of the determinant of the input, which is also the sum of its row degrees. First consider the case $m \le D$. If $s_1 + \cdots + s_{\lfloor m/2\rfloor} > D/2$, the reduction to the case $s_1 + \cdots + s_{\lfloor m/2\rfloor} \le D/2$ with the same m and D performed at Line 7 costs $O(m^\omega(1 + D/m))$. Once we are in the latter case, there are three recursive calls with input matrices having the following dimensions and degrees:

- At Line 12, the matrix $\mathbf{B}$ is $\lceil m/2\rceil \times \lceil m/2\rceil$, and applying the predictable degree property on Eq. (2) gives in particular $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{B}) = \mathrm{rdeg}_s([\mathbf{K}_1 \;\; \mathbf{K}_2])$, hence $|\mathrm{rdeg}_{\mathbf{0}}(\mathbf{B})| \le D$.
- At Line 13, the matrix $\mathbf{A}_1$ is $\lfloor m/2\rfloor \times \lfloor m/2\rfloor$ and the sum of its row degrees is $s_1 + \cdots + s_{\lfloor m/2\rfloor}$, which is at most $D/2$ by assumption.
- At Line 15, the matrix $\mathbf{P}^\mathsf{T}$ is $\lceil m/2\rceil \times \lceil m/2\rceil$ and its $\mathbf{0}$-pivot degree is $\mathrm{rdeg}_{\mathbf{0}}(\mathbf{P}^\mathsf{T}) = \boldsymbol{\delta}\mathbf{J}_m$. Recall indeed that this is the list of diagonal degrees of $\mathbf{P}^\mathsf{T}$, which is the same as that of $\mathbf{P}$, and thus the same as that of $\mathbf{J}_n\mathbf{K}_2^\mathsf{T}\mathbf{J}_n$ according to Lemma 2.5. Now, from $|\boldsymbol{\delta} + \boldsymbol{\nu}| = |\boldsymbol{\delta}| + |\boldsymbol{\nu}| \le D$ and the assumption $|\boldsymbol{\nu}| = s_{\lfloor m/2\rfloor+1} + \ldots + s_m > D/2$, we obtain $|\mathrm{rdeg}_{\mathbf{0}}(\mathbf{P}^\mathsf{T})| = |\boldsymbol{\delta}| \le D/2$.

We assume without loss of generality that m is a power of 2. If it is not, a given input matrix can be padded with zeros, and ones on the main diagonal, so as to form a square matrix with dimension the next power of two and the same determinant. According to the three items above, the cost bound then satisfies:

$$C(m, D) \le 2C(m/2, \lfloor D/2\rfloor) + C(m/2, D) + O\left(m^\omega \mathsf{M}'(D/m)\right).$$

Letting the $O(\cdot)$ term aside, we illustrate this recurrence relation in Fig. 1.

Let $\mu = \log_2(m)$ and let K be the constant of the $O(\cdot)$ term above. Recalling that $m \le D$ and $\lfloor\lfloor D/2^j\rfloor/2\rfloor = \lfloor D/2^{j+1}\rfloor$, unrolling this recurrence to the ith recursion level for $0 \le i \le \mu$ yields

$$\begin{aligned} C(m, D) &\le \sum_{j=0}^{i} a_{i,j} C\left(\frac{m}{2^i}, \left\lfloor\frac{D}{2^j}\right\rfloor\right) + K\sum_{k=0}^{i-1}\sum_{j=0}^{k} a_{k,j}\left(\frac{m}{2^k}\right)^\omega \mathsf{M}'\left(\frac{D/2^j}{m/2^k}\right)\\ &\le \sum_{j=0}^{i} a_{i,j} C\left(\frac{m}{2^i}, \left\lfloor\frac{D}{2^j}\right\rfloor\right) + Km^\omega \mathsf{M}'\left(\frac{D}{m}\right)\left(\sum_{k=0}^{i-1}\sum_{j=0}^{k} a_{k,j} 2^{-k\omega}\mathsf{M}'\left(2^{k-j}\right)\right), \end{aligned}$$

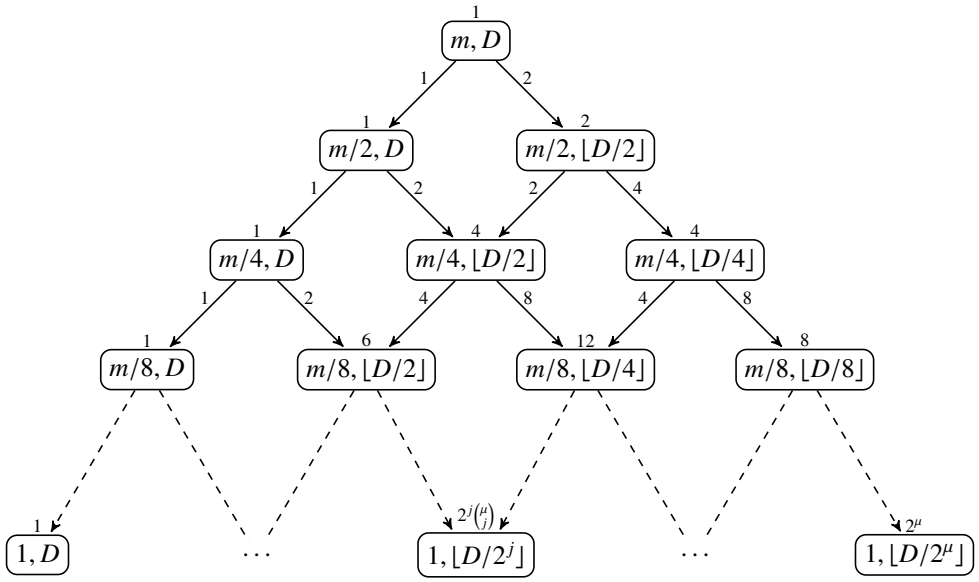


Figure 1: Directed acyclic graph of recursive calls, of depth $\mu = \log_2(m)$. Each boxed node shows the matrix dimensions and the determinantal degree of a recursive call. Beginning with one call in dimension and determinantal degree (m, D), for a given node the number above it indicates the number of times a recursive call corresponding to this node is made, and the numbers of recursive sub-calls this node generates are indicated on both arrows starting from this node.

where the last inequality comes from the submultiplicativity of $\mathsf{M}'(\cdot)$ and the coefficients $a_{i,j}$ satisfy

$$\begin{cases} a_{i,0} &= 1, \\ a_{i,i} &= 2^i, \\ a_{i,j} &= a_{i-1,j} + 2a_{i-1,j-1} \text{ for } 0 < j < i. \end{cases}$$

In Fig. 1, one can observe the similarity between Pascal's triangle and the number of calls with parameters $(m/2^i, D/2^j)$. This translates as a connection between the $a_{i,j}$'s and the binomial coefficients: one can prove by induction that $a_{i,j} = 2^j\binom{i}{j}$.

Now, by assumption $\mathsf{M}'(d) \in O(d^{\omega-1-\varepsilon})$ for some $\varepsilon > 0$; for such an $\varepsilon > 0$, let $\tilde{K}$ be a constant such that $\mathsf{M}'(d) \leq \tilde{K} d^{\omega-1-\varepsilon}$ for all $d \geq 0$. Then

$$\sum_{k=0}^{i-1}\sum_{j=0}^{k} a_{k,j}2^{-k\omega}\,\mathsf{M}'\left(2^{k-j}\right) \leq \tilde{K}\sum_{k=0}^{i-1} 2^{-(1+\varepsilon)k}\sum_{j=0}^{k}\frac{a_{k,j}}{2^{j(\omega-1-\varepsilon)}} \leq \tilde{K}\sum_{k=0}^{i-1} 2^{-(1+\varepsilon)k}\sum_{j=0}^{k}\binom{k}{j} = \tilde{K}\sum_{k=0}^{i-1} 2^{-\varepsilon k}$$

and, defining the constant $\hat{K} = K\tilde{K}\sum_{k=0}^{+\infty} 2^{-\varepsilon k}$, for $i = \mu$ we obtain

$$C(m, D) \leq \sum_{j=0}^{\mu} a_{\mu,j} C\left(1, \left\lfloor \frac{D}{2^j} \right\rfloor\right) + \hat{K} m^{\omega}\,\mathsf{M}'\left(\frac{D}{m}\right).$$

As we have seen above, parameters $(1, d)$ for any $d \in \mathbb{Z}_{\geq 0}$ correspond to base cases with a 1×1 matrix, and they incur no arithmetic cost (one might want to consider them to use $O(1)$ operations

each; then the total cost of these base cases is bounded asymptotically by $\sum_{j=0}^{\mu} a_{\mu,j} \leq m^{\log_2(3)})$. Thus we obtain $C(m, D) = O(m^\omega \mathsf{M}'(D/m))$, under the assumption $m \leq D$.

For the case $D < m$ handled at Line 4, Section 3.2 showed that $C(m, D) = C(\hat{m}, D) + O(m^\omega)$ where $\hat{m}$ is the number of non-constant rows of the input matrix. Since $\hat{m} \leq D$, our proof above shows that $C(\hat{m}, D)$ is in $O(\hat{m}^\omega \mathsf{M}'(D/\hat{m}))$. Now our assumptions on $\mathsf{M}(\cdot)$ imply in particular that $\mathsf{M}'(\cdot)$ is subquadratic, hence this bound is in $O(\hat{m}^\omega (D/\hat{m})^2) = O(\hat{m}^{\omega-2} D^2) \subseteq O(m^\omega)$. This concludes the complexity analysis.

4. Shifted forms: from reduced to weak Popov

This section proves Theorem 1.3 by generalizing the approach of [42, Sec. 2 and 3], which focuses on the non-shifted case $\boldsymbol{s} = \mathbf{0}$. It first uses Gaussian elimination on the $\mathbf{0}$-leading matrix of $\mathbf{A}$ to find a unimodular matrix $\mathbf{U}$ such that $\mathbf{UA}$ is in $\mathbf{0}$-weak Popov form, and then exploits the specific form of $\mathbf{U}$ to compute $\mathbf{UA}$ efficiently. Here we extend this approach to arbitrary shifts and show how to take into account the possible unbalancedness of the row degree of $\mathbf{A}$.

First, we generalize [42, Lem. 8] from $\boldsymbol{s} = \mathbf{0}$ to an arbitrary $\boldsymbol{s}$, by describing how $\mathbf{U}$ can be obtained by computing a $\mathbf{0}$-weak Popov form of the $\boldsymbol{s}$-leading matrix of $\mathbf{A}$.

Lemma 4.1. *Let $\boldsymbol{s} \in \mathbb{Z}^n$ and let $\mathbf{A} \in \mathbb{K}[x]^{m \times n}$ be $\boldsymbol{s}$-reduced with $\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A})$ nondecreasing. There exists an invertible lower triangular $\mathbf{T} \in \mathbb{K}^{m \times m}$ which can be computed in $O(m^{\omega-1} n)$ operations in $\mathbb{K}$ and is such that $\mathbf{T} \operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$ is in $\mathbf{0}$-unordered weak Popov form. For any such matrix $\mathbf{T}$,*

- *$\mathbf{U} = \mathbf{X}^{\boldsymbol{t}} \mathbf{T} \mathbf{X}^{-\boldsymbol{t}}$ has polynomial entries and is unimodular, where $\boldsymbol{t} = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A})$,*
- *$\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{UA}) = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A})$ and $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{UA}) = \mathbf{T} \operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$,*
- *$\mathbf{UA}$ is in $\boldsymbol{s}$-unordered weak Popov form.*

Proof. Consider the matrix $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n$ and its generalized Bruhat decomposition $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})\mathbf{J}_n = \mathbf{CPR}$ as defined in [35]: $\mathbf{C} \in \mathbb{K}^{m \times m}$ is in column echelon form, $\mathbf{R} \in \mathbb{K}^{m \times n}$ is in row echelon form, and $\mathbf{P} \in \mathbb{K}^{m \times m}$ is a permutation matrix. Therefore $\mathbf{J}_m \mathbf{R} \mathbf{J}_n$ is in $\mathbf{0}$-weak Popov form (see the paragraph before Lemma 2.4 in Section 2.4), and $\mathbf{PRJ}_n$ is in $\mathbf{0}$-unordered weak Popov form. Since $\operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$ has full row rank, $\mathbf{C}$ is lower triangular and invertible, hence $\mathbf{PRJ}_n = \mathbf{C}^{-1} \operatorname{lm}_{\boldsymbol{s}}(\mathbf{A})$ which proves the existence of $\mathbf{T} = \mathbf{C}^{-1}$. Computing the decomposition costs $O(m^{\omega-1} n)$ operations [13, Cor. 25] while inverting $\mathbf{C}$ costs $O(m^\omega)$ operations. Alternatively, [42, Sec. 3] shows how to compute $\mathbf{T}$ within the same cost bound using an LUP decomposition with a modified pivoting strategy.

For any such matrix $\mathbf{T}$, write $\mathbf{T} = (T_{ij})_{ij}$ and $\boldsymbol{t} = (t_i)_i$. Then the entry (i, j) of $\mathbf{U}$ is $T_{ij} x^{t_i - t_j}$. Thus, $\mathbf{U}$ is lower triangular in $\mathbb{K}(x)^{m \times m}$ with diagonal entries in $\mathbb{K} \setminus \{0\}$, and since $\boldsymbol{t} = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A})$ is nondecreasing, $\mathbf{U}$ is a unimodular matrix in $\mathbb{K}[x]^{m \times m}$.

Now, consider the nonnegative shift $\boldsymbol{u} = \boldsymbol{t} - (\min(\boldsymbol{t}), \ldots, \min(\boldsymbol{t})) \in \mathbb{Z}^m_{\geq 0}$, and note that $\mathbf{U} = \mathbf{X}^{\boldsymbol{u}} \mathbf{T} \mathbf{X}^{-\boldsymbol{u}}$. (The introduction of $\boldsymbol{u}$ is to circumvent the fact that we have not defined the row degree of a matrix over the Laurent polynomials, a notion which would be needed if we used $\mathbf{X}^{\boldsymbol{t}}$ rather than $\mathbf{X}^{\boldsymbol{u}}$ in Eq. (4).) Since $\mathbf{A}$ is in $\boldsymbol{s}$-reduced form, the predictable degree property yields

$$\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{UA}) = \operatorname{rdeg}_{\boldsymbol{t}}(\mathbf{U}) = \operatorname{rdeg}_{\boldsymbol{u}}(\mathbf{U}) + (\min(\boldsymbol{t}), \ldots, \min(\boldsymbol{t})). \tag{3}$$

On the other hand,

$$\operatorname{rdeg}_{\boldsymbol{u}}(\mathbf{U}) = \operatorname{rdeg}_{\mathbf{0}}(\mathbf{U}\mathbf{X}^{\boldsymbol{u}}) = \operatorname{rdeg}_{\mathbf{0}}(\mathbf{X}^{\boldsymbol{u}} \mathbf{T}) = \boldsymbol{u} = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A}) - (\min(\boldsymbol{t}), \ldots, \min(\boldsymbol{t})), \tag{4}$$

where the third equality follows from the fact that $\mathbf{T}$ is a constant matrix with no zero row. Then, from Eqs. (3) and (4), we obtain $\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{UA}) = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A}) = \boldsymbol{t}$.

Then, $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{UA})$ is formed by the coefficients of nonnegative degree of

$$\mathbf{X}^{-\mathrm{rdeg}_{\boldsymbol{s}}(\mathbf{UA})}\,\mathbf{UAX}^{\boldsymbol{s}} \;=\; \mathbf{X}^{-\boldsymbol{t}}\,\mathbf{UAX}^{\boldsymbol{s}} \;=\; \mathbf{TX}^{-\boldsymbol{t}}\,\mathbf{AX}^{\boldsymbol{s}}\,.$$

Since $\mathbf{T}$ is constant and $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$ is formed by the coefficients of nonnegative degree of $\mathbf{X}^{-\boldsymbol{t}}\,\mathbf{AX}^{\boldsymbol{s}}$, we obtain that $\mathrm{lm}_{\boldsymbol{s}}(\mathbf{UA}) = \mathbf{T}\,\mathrm{lm}_{\boldsymbol{s}}(\mathbf{A})$. The third item then directly follows from Lemma 2.4. □

Knowing $\mathbf{T}$, and therefore $\mathbf{U}$, the remaining difficulty is to efficiently compute $\mathbf{UA}$. For this, in the case $\boldsymbol{s} = \mathbf{0}$, the approach in [42] has a cost bound which involves the maximum degree $d = \deg(\mathbf{A}) = \max(\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A}))$ [42, Thm. 13], and uses the following steps:

- first compute $x^d\mathbf{X}^{-\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})}\,\mathbf{UA} = \mathbf{T}(\mathbf{X}^{(d,\ldots,d)-\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})}\,\mathbf{A})$;
- then scale the rows via the left multiplication by $x^{-d}\mathbf{X}^{\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})}$.

While the scaling does not use arithmetic operations, the first step asks to multiply the constant matrix $\mathbf{T}$ by an $m \times n$ polynomial matrix of degree d: this costs $O(m^{\omega-1}nd)$ operations in $\mathbb{K}$. Such a cost would not allow us to reach our target bound for the computation of characteristic polynomials, since d may be too large, namely when $\mathbf{A}$ has unbalanced row degrees.

Below, we refine the approach to better conform with the row degrees of $\mathbf{A}$. This leads to an improvement of the above cost bound to $O(m^{\omega-1}n(1 + D/m))$ where $D = |\mathrm{rdeg}_{\mathbf{0}}(\mathbf{A})|$, thus involving the average row degree of $\mathbf{A}$ instead of its maximum degree d. For this, we follow the strategy of splitting the rows of $\mathbf{A}$ into subsets having degrees in prescribed intervals; when these intervals get further away from the average row degree of $\mathbf{A}$, the corresponding subset contains a smaller number of rows. Earlier works using a similar strategy such as [54], [58, Sec. 2.6], and [22, Sec. 4], are not directly applicable to the problem here. Furthermore, we exploit the fact that $\mathbf{A}$ has nondecreasing row degree and that $\mathbf{T}$ has a lower triangular shape to avoid logarithmic factors in the cost bound. This results in Algorithm 3.

Algorithm 3 ReducedToWeakPopov($\mathbf{A}, \boldsymbol{s}$)

Input: a matrix $\mathbf{A} \in \mathbb{K}[x]^{m\times n}$, a shift $\boldsymbol{s} \in \mathbb{Z}^n$ such that $\mathbf{A}$ is in $\boldsymbol{s}$-reduced form.
Output: an $\boldsymbol{s}$-weak Popov form of $\mathbf{A}$.

▷ *Step 1: Ensure nonnegative shift and nondecreasing $\boldsymbol{s}$-row degree* ◁
$\hat{\boldsymbol{s}} \in \mathbb{Z}^n_{\geq 0} \leftarrow \boldsymbol{s} - (\min(\boldsymbol{s}), \ldots, \min(\boldsymbol{s}))$
$(\mathbf{B}, \boldsymbol{t}) \in \mathbb{K}[x]^{m\times n} \times \mathbb{Z}^m_{\geq 0} \leftarrow$ matrix $\mathbf{B}$ obtained from $\mathbf{A}$ by row permutation such that the tuple $\boldsymbol{t} = \mathrm{rdeg}_{\hat{\boldsymbol{s}}}(\mathbf{B})$ is nondecreasing
▷ *Step 2: Compute the factor $\mathbf{T}$ in the unimodular transformation $\mathbf{U} = \mathbf{X}^{\boldsymbol{t}}\,\mathbf{TX}^{-\boldsymbol{t}}$* ◁
$\mathbf{L} \in \mathbb{K}^{m\times n} \leftarrow \mathrm{lm}_{\hat{\boldsymbol{s}}}(\mathbf{B})$, that is, the entry (i, j) of $\mathbf{L}$ is the coefficient of degree $t_i - s_j$ of the entry (i, j) of $\mathbf{B}$, where $\hat{\boldsymbol{s}} = (s_1, \ldots, s_n)$ and $\boldsymbol{t} = (t_1, \ldots, t_m)$
$\mathbf{T} \in \mathbb{K}^{m\times m} \leftarrow$ invertible lower triangular matrix such that $\mathbf{TL}$ is in $\mathbf{0}$-unordered weak Popov form ▷ *can be computed via a generalized Bruhat decomposition, see Lemma 4.1*
▷ *Step 3: Compute the product $\mathbf{P} = \mathbf{UB}$* ◁
$D \leftarrow t_1 + \cdots + t_m$; $K \leftarrow \lfloor \log_2(mt_m/D) \rfloor + 1$ ▷ $K = \min\{k \in \mathbb{Z}_{>0} \mid t_m < 2^k D/m\}$
$i_0 \leftarrow 0$; $i_k \leftarrow \max\{i \mid t_i < 2^k D/m\}$ for $1 \leq k \leq K$ ▷ $0 = i_0 < i_1 \leq \cdots \leq i_{K-1} < i_K = m$
$\mathbf{P} \leftarrow$ zero matrix in $\mathbb{K}[x]^{m\times n}$
for k from 1 to K **do**
    $\mathcal{R} \leftarrow \{i_{k-1} + 1, \ldots, m\}$; $C \leftarrow \{i_{k-1} + 1, \ldots, i_k\}$; $\theta \leftarrow t_{i_k} = \max(\boldsymbol{t}_C)$
    $\mathbf{P}_{\mathcal{R},*} \leftarrow \mathbf{P}_{\mathcal{R},*} + \mathbf{X}^{\boldsymbol{t}_{\mathcal{R}}-(\theta,\ldots,\theta)}\,\mathbf{T}_{\mathcal{R},C}\,\mathbf{X}^{(\theta,\ldots,\theta)-\boldsymbol{t}_C}\,\mathbf{B}_{C,*}$
return the row permutation of $\mathbf{P}$ which has increasing $\hat{\boldsymbol{s}}$-pivot index

Proof of Theorem 1.3. The first step builds a nonnegative shift $\hat{\boldsymbol{s}}$ which only differs from $\boldsymbol{s}$ by an additive constant, and builds a matrix $\mathbf{B}$ which is a row permutation of $\mathbf{A}$; hence any $\hat{\boldsymbol{s}}$-weak Popov form of $\mathbf{B}$ is an $\boldsymbol{s}$-weak Popov form of $\mathbf{A}$. Since the $\hat{\boldsymbol{s}}$-row degree $\boldsymbol{t}$ of $\mathbf{B}$ is nondecreasing, the construction of $\mathbf{T}$ at Step 2 ensures that the matrix $\mathbf{U} = \mathbf{X}^{\boldsymbol{t}}\mathbf{T}\mathbf{X}^{-\boldsymbol{t}} \in \mathbb{K}[x]^{m\times m}$ is unimodular and such that $\mathbf{UB}$ is in $\hat{\boldsymbol{s}}$-unordered weak Popov form, according to Lemma 4.1. Thus, for the correctness, it remains to prove that the matrix $\mathbf{P}$ computed at Step 3 is $\mathbf{P} = \mathbf{UB}$.

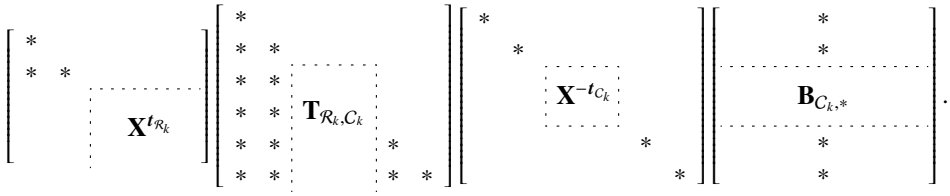
Using notation from the algorithm, define $C_k = \{i_{k-1}+1, \ldots, i_k\}$ for $1 \le k \le K$. The C_k's are pairwise disjoint and such that $\{1, \ldots, m\} = C_1 \cup \cdots \cup C_K$. Then, slicing the columns of $\mathbf{T}$ according to these sets of column indices, we obtain

$$\mathbf{UB} = \mathbf{X}^{\boldsymbol{t}}\mathbf{T}\mathbf{X}^{-\boldsymbol{t}}\mathbf{B} = \sum_{1\le k\le K} \mathbf{X}^{\boldsymbol{t}}\mathbf{T}_{*,C_k}\mathbf{X}^{-\boldsymbol{t}_{C_k}}\mathbf{B}_{C_k,*}.$$

Furthermore, since $\mathbf{T}$ is lower triangular, all rows of $\mathbf{T}_{*,C_k}$ with index not in $\mathcal{R}_k = \{i_{k-1}+1, \ldots, m\}$ are zero. Hence, more precisely,

$$\mathbf{UB} = \sum_{1\le k\le K} \begin{bmatrix} \mathbf{0}_{i_{k-1}\times n} \\ \mathbf{X}^{\boldsymbol{t}_{\mathcal{R}_k}}\ \mathbf{T}_{\mathcal{R}_k,C_k}\ \mathbf{X}^{-\boldsymbol{t}_{C_k}}\ \mathbf{B}_{C_k,*} \end{bmatrix}.$$

This formula corresponds to the slicing of the product $\mathbf{P} = \mathbf{X}^{\boldsymbol{t}}\mathbf{T}\mathbf{X}^{-\boldsymbol{t}}\mathbf{B}$ in blocks as follows:



The **for** loop at Step 3 computes $\mathbf{P}$ by following this formula, hence $\mathbf{P} = \mathbf{UB}$. (Note indeed that the scaling by $(\theta, \ldots, \theta)$ in the algorithm can be simplified away and is just there to avoid computing with Laurent polynomials.)

Concerning the complexity bound, we first note that the quantity D defined in the algorithm is the same as that in Theorem 1.3; indeed, $\operatorname{rdeg}_{\hat{\boldsymbol{s}}}(\mathbf{B}) = \operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{B}) - (\min(\boldsymbol{s}), \ldots, \min(\boldsymbol{s}))$, hence

$$D = t_1 + \cdots + t_m = |\operatorname{rdeg}_{\hat{\boldsymbol{s}}}(\mathbf{B})| = |\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{B})| - m\cdot\min(\boldsymbol{s}) = |\operatorname{rdeg}_{\boldsymbol{s}}(\mathbf{A})| - m\cdot\min(\boldsymbol{s}).$$

By Lemma 4.1, Step 2 uses $O(m^{\omega-1}n)$ operations. As for Step 3, its cost follows from bounds on the cardinalities of C_k and $\mathcal{R}_k$. Precisely, we have $C_k \subseteq \mathcal{R}_k$, and $\min(\mathcal{R}_k) > i_{k-1}$ implies that each entry of the subtuple $\boldsymbol{t}_{\mathcal{R}_k}$ is at least $2^{k-1}D/m$. Hence $\#\mathcal{R}_k \cdot 2^{k-1}D/m \le \sum_{i\in\mathcal{R}_k} t_i \le t_1 + \cdots + t_m = D$, and $\#C_k \le \#\mathcal{R}_k \le m/2^{k-1}$.

Then, in the product $\mathbf{X}^{\boldsymbol{t}_{\mathcal{R}_k}-(\theta,\ldots,\theta)}\ \mathbf{T}_{\mathcal{R}_k,C_k}\ \mathbf{X}^{(\theta,\ldots,\theta)-\boldsymbol{t}_{C_k}}\ \mathbf{B}_{C_k,*}$, the left multiplication by $\mathbf{X}^{\boldsymbol{t}_{\mathcal{R}_k}-(\theta,\ldots,\theta)}$ does not use arithmetic operations; the matrix $\mathbf{T}_{\mathcal{R}_k,C_k}$ is over $\mathbb{K}$ and has at most $m/2^{k-1}$ rows and at most $m/2^{k-1}$ columns; and the matrix $\mathbf{X}^{(\theta,\ldots,\theta)-\boldsymbol{t}_{C_k}}\ \mathbf{B}_{C_k,*}$ is over $\mathbb{K}[x]$ and has n columns and at most $m/2^{k-1}$ rows. Furthermore, the latter matrix has degree at most θ: indeed, $\hat{\boldsymbol{s}} \ge \mathbf{0}$ implies that $\boldsymbol{t} = \operatorname{rdeg}_{\hat{\boldsymbol{s}}}(\mathbf{B}) \ge \operatorname{rdeg}_{\mathbf{0}}(\mathbf{B})$, hence in particular $\boldsymbol{t}_{C_k} \ge \operatorname{rdeg}_{\mathbf{0}}(\mathbf{B}_{C_k,*})$. Recall that $\theta = t_{i_k} \le 2^k D/m$ holds, by definition of θ and i_k. From these bounds on the dimensions and the degrees of the involved matrices, and using the fact that $m/2^{k-1} \le m \le n$, it follows that computing the product

$\mathbf{X}^{t_{\mathcal{R}_k}-(\theta,\ldots,\theta)}$ $(\mathbf{T}_{\mathcal{R}_k,C_k}$ $(\mathbf{X}^{(\theta,\ldots,\theta)-t_{C_k}}$ $\mathbf{B}_{C_k,*}))$ uses $O((m/2^{k-1})^{\omega-1}n(\theta+1)) \subseteq O((m/2^{k-1})^{\omega-1}n(2^kD/m+1))$ operations in $\mathbb{K}$. Thus, since $\omega > 2$, the cost of Step 3 is

$$O\left(\sum_{1\le k\le K}\left(\frac{m}{2^{k-1}}\right)^{\omega-1} n\left(\frac{2^kD}{m}+1\right)\right) \subseteq O\left(m^{\omega-2}nD\left(\sum_{1\le k\le K}2^{k(2-\omega)}\right) + m^{\omega-1}n\left(\sum_{1\le k\le K}2^{k(1-\omega)}\right)\right)$$
$$\subseteq O\left(m^{\omega-2}nD + m^{\omega-1}n\right).$$ □

5. Shifted forms: from weak Popov to Popov

This section culminates in Section 5.5 with the description of Algorithm WeakPopovToPopov and a proof of Theorem 1.4. Based on [42, Lem. 14] (which extends to the shifted case), the result in Theorem 1.4 can easily be used to solve the same problem in the rectangular case with an $m\times n$ matrix $\mathbf{A}$; while this is carried out in Algorithm WeakPopovToPopov, it is out of the main scope of this paper and thus for conciseness we only give a cost analysis in the case $m = n$.

Our approach is to obtain the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$ from a shifted reduced kernel basis of some matrix $\mathbf{F}$ built from $\mathbf{A}$. This fact is substantiated in Section 5.1, which also proves that we have precise a priori information on the degrees of this sought kernel basis.

A folklore method for kernel basis computation is to find an approximant basis at an order sufficiently large so that it contains a kernel basis as a submatrix. More precisely, assuming we know a list of bounds $\boldsymbol{s}$ such that there exists a basis of $\mathcal{K}(\mathbf{F})$ with column degree bounded by $\boldsymbol{s}$, the following algorithm computes such a kernel basis which is furthermore $-\boldsymbol{s}$-reduced:

- $\boldsymbol{\gamma} \leftarrow \mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F}) + 1$;
- $\mathbf{M} \leftarrow$ basis of $\mathcal{A}_{\boldsymbol{\gamma}}(\mathbf{F})$ in $-\boldsymbol{s}$-reduced form;
- **return** the submatrix of $\mathbf{M}$ formed by its rows with nonpositive $-\boldsymbol{s}$-degree.

The idea is that any row $\mathbf{p}$ of $\mathbf{M}$ is such that $\mathbf{pF} = \mathbf{0} \bmod \mathbf{X}^{\boldsymbol{\gamma}}$, and if it further satisfies $\mathrm{cdeg}(\mathbf{p}) \le \boldsymbol{s}$ then $\mathrm{cdeg}(\mathbf{pF}) \le \mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F}) < \boldsymbol{\gamma}$, so that $\mathbf{pF} = \mathbf{0}$ holds. Here the complexity mainly depends on $|\boldsymbol{s}|$ and $|\mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F})|$, quantities that are often large in which case the algorithm of Zhou et al. [58] is more efficient. Nevertheless there are cases, such as the one arising in this section, where both sums are controlled, and elaborating over this approach leads to an efficient algorithm.

In order to propose Algorithm KnownDegreeKernelBasis in Section 5.4, efficiently computing the kernel basis using essentially a single call to PM-Basis, we transform the input into one with a balanced shift and a balanced order. Here and in what follows, for a nonnegative tuple $\boldsymbol{t} \in \mathbb{Z}^n_{\ge 0}$, we say that $\boldsymbol{t}$ is *balanced* if $\max(\boldsymbol{t}) \in O(|\boldsymbol{t}|/n)$, meaning that the maximum entry in $\boldsymbol{t}$ is not much larger than the average of all entries of $\boldsymbol{t}$. Section 5.2 deals with the shifts by describing a transformation of the input inspired from [46, Sec. 3], allowing us to reduce to the case of a shift $\boldsymbol{s}$ whose entries are balanced. Section 5.3 deals with balancing the order $\boldsymbol{\gamma}$ by performing the overlapping partial linearization of [46, Sec. 2].

For the latter transformation, as noted above, we assume that there exists a basis of the considered kernel whose column degree is bounded by $\boldsymbol{s}$, or equivalently that $-\boldsymbol{s}$-reduced kernel bases have nonpositive $-\boldsymbol{s}$-row degree. On the other hand, for the first transformation we must ensure that $-\boldsymbol{s}$-reduced kernel bases have nonnegative $-\boldsymbol{s}$-row degree. Thus our algorithm works under the requirement that $-\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have $-\boldsymbol{s}$-row degree exactly $\mathbf{0}$, hence its name. This is a restriction compared to [56, Algo. 1] which only assumes that $-\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have nonpositive $-\boldsymbol{s}$-row degree and has to perform several approximant basis computations to

recover the whole kernel basis, as outlined in the introduction. In short, we have managed to exploit the fact that we have better a priori knowledge of degrees in kernel bases than in the context of [56], leading to a faster kernel computation which, when used in our determinant algorithm, brings no logarithmic factor in the complexity.

5.1. Normalization via kernel basis computation

We normalize the matrix $\mathbf{A}$ into its $-\boldsymbol{s}$-Popov form $\mathbf{P}$ using a kernel basis computation, an approach already used in a context similar to ours in the non-shifted case in [42, Lem. 19]. Roughly, this stems from the fact that the identity $\mathbf{UA} = \mathbf{P}$ with a unimodular $\mathbf{U}$ can be rewritten as

$$\begin{bmatrix}\mathbf{U} & \mathbf{P}\end{bmatrix}\begin{bmatrix}\mathbf{A}\\ -\mathbf{I}_m\end{bmatrix} = \mathbf{0};$$

and that, for a well-chosen shift, one retrieves $[\mathbf{U}\ \ \mathbf{P}]$ as a shifted reduced kernel basis. The next statement gives a choice of shift suited to our situation, and describes the degree profile of such kernel bases. The focus on the shift $-\boldsymbol{\delta}$ comes from the fact that any $-\boldsymbol{\delta}$-reduced form $\mathbf{R}$ of $\mathbf{A}$ is only a constant transformation away from being the $-\boldsymbol{s}$-Popov form $\mathbf{P}$ of $\mathbf{A}$ [see 26, Lem. 4.1].

Lemma 5.1. *Let $\boldsymbol{s} \in \mathbb{Z}^m$, let $\mathbf{A} \in \mathbb{K}[x]^{m\times m}$ be in $-\boldsymbol{s}$-weak Popov form, let $\boldsymbol{\delta} \in \mathbb{Z}^m_{\geq 0}$ be the $-\boldsymbol{s}$-pivot degree of $\mathbf{A}$, and assume that $\boldsymbol{s} \geq \boldsymbol{\delta}$. Let $\mathbf{R}$ be a $-\boldsymbol{\delta}$-weak Popov form of $\mathbf{A}$ and let $\mathbf{U}$ be the unimodular matrix such that $\mathbf{UA} = \mathbf{R}$. Let further $\boldsymbol{d} = (\boldsymbol{s}-\boldsymbol{\delta}, \boldsymbol{\delta}) \in \mathbb{Z}^{2m}$ and $\mathbf{F} = \left[\begin{smallmatrix}\mathbf{A}\\ -\mathbf{I}_m\end{smallmatrix}\right] \in \mathbb{K}[x]^{2m\times m}$. Then,*

- *the $-\boldsymbol{d}$-pivot profile of $[\mathbf{U}\ \ \mathbf{R}]$ is $(m + j, \delta_j)_{1\leq j\leq m}$,*
- *$[\mathbf{U}\ \ \mathbf{R}]$ is in $-\boldsymbol{d}$-weak Popov form with $\operatorname{rdeg}_{-\boldsymbol{d}}([\mathbf{U}\ \ \mathbf{R}]) = \mathbf{0}$ and $\operatorname{cdeg}([\mathbf{U}\ \ \mathbf{R}]) \leq \boldsymbol{d}$,*
- *$[\mathbf{U}\ \ \mathbf{R}]$ is a basis of $\mathcal{K}(\mathbf{F})$.*

Proof. First, we prove that $\mathbf{R}$ has $-\boldsymbol{\delta}$-pivot degree $\boldsymbol{\delta}$; note that this implies $\operatorname{rdeg}_{-\boldsymbol{\delta}}(\mathbf{R}) = \mathbf{0}$ and $\operatorname{cdeg}(\mathbf{R}) = \boldsymbol{\delta}$ since $\mathbf{R}$ is in $-\boldsymbol{\delta}$-weak Popov form. Lemma 2.5 shows that the $-\boldsymbol{s}$-Popov form $\mathbf{P}$ of $\mathbf{A}$ has the same $-\boldsymbol{s}$-pivot degree as $\mathbf{A}$, that is, $\boldsymbol{\delta}$. Hence, by definition of Popov forms, $\operatorname{cdeg}(\mathbf{P}) = \boldsymbol{\delta}$. Then, [26, Lem. 4.1] states that $\mathbf{P}$ is also in $-\boldsymbol{\delta}$-Popov form. Since $\mathbf{R}$ is a $-\boldsymbol{\delta}$-weak Popov form of $\mathbf{P}$, by Lemma 2.5 it has the same $-\boldsymbol{\delta}$-pivot degree as $\mathbf{P}$, that is, $\boldsymbol{\delta}$.

Now, by the predictable degree property and since $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{A}) = -\boldsymbol{s} + \boldsymbol{\delta}$,

$$\operatorname{rdeg}_{-\boldsymbol{s}+\boldsymbol{\delta}}(\mathbf{U}) = \operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{UA}) = \operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{R}) \leq \operatorname{rdeg}_{-\boldsymbol{\delta}}(\mathbf{R}) = \mathbf{0},$$

where the inequality holds because $-\boldsymbol{s} \leq -\boldsymbol{\delta}$. Thus, by choice of $\boldsymbol{d}$, the $-\boldsymbol{d}$-pivot entries of $[\mathbf{U}\ \ \mathbf{R}]$ are the $-\boldsymbol{\delta}$-pivot entries of its submatrix $\mathbf{R}$; this proves the first item.

Then, the second item follows: the matrix $[\mathbf{U}\ \ \mathbf{R}]$ is in $-\boldsymbol{d}$-weak Popov form since its $-\boldsymbol{d}$-pivot index is increasing; its $-\boldsymbol{d}$-row degree is equal to the $-\boldsymbol{\delta}$-row degree of $\mathbf{R}$ which is $\mathbf{0}$; and $\operatorname{rdeg}_{-\boldsymbol{d}}([\mathbf{U}\ \ \mathbf{R}]) = \mathbf{0}$ implies $\operatorname{cdeg}([\mathbf{U}\ \ \mathbf{R}]) \leq \boldsymbol{d}$ by Lemma 2.1.

Let $[\mathbf{K}_1\ \ \mathbf{K}_2] \in \mathbb{K}[x]^{m\times 2m}$ be a basis $\mathcal{K}(\mathbf{F})$ (it has $m = 2m - m$ rows since $\mathbf{F}$ is $2m \times m$ and has rank m). Since $\mathbf{UA} = \mathbf{R}$, the rows of $[\mathbf{U}\ \ \mathbf{R}]$ are in this kernel. As a result, there exists a matrix $\mathbf{V} \in \mathbb{K}[x]^{m\times m}$ such that $[\mathbf{U}\ \ \mathbf{R}] = \mathbf{V}[\mathbf{K}_1\ \ \mathbf{K}_2]$. In particular, $\mathbf{U} = \mathbf{VK}_1$, and since $\mathbf{U}$ is unimodular, this implies that $\mathbf{V}$ is unimodular as well. Thus, the basis $[\mathbf{K}_1\ \ \mathbf{K}_2]$ is unimodularly equivalent to $[\mathbf{U}\ \ \mathbf{R}]$, and the latter matrix is also a basis of $\mathcal{K}(\mathbf{F})$. □

One may note similarities with [27, Lem. 5.1], which is about changing the shift of reduced forms via kernel basis computations. Here, we consider $\mathbf{A}$ in $-\boldsymbol{s}$-reduced form and are interested

in its $-\boldsymbol{\delta}$-reduced forms $\mathbf{R}$: we change $-\boldsymbol{s}$ into $-\boldsymbol{\delta} = -\boldsymbol{s} + (\boldsymbol{s} - \boldsymbol{\delta})$, with a nonnegative difference $\boldsymbol{s} - \boldsymbol{\delta} \geq \mathbf{0}$. Still, the above lemma does not follow from [27] since here the origin shift $-\boldsymbol{s}$ is nonpositive and thus we cannot directly incorporate it in the matrix $\mathbf{F}$ by considering $\mathbf{A}\mathbf{X}^{-\boldsymbol{s}}$.

The next corollary uses notation from Lemma 5.1 and shows how to obtain the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$ via a $-\boldsymbol{d}$-reduced basis of the kernel of $\mathbf{F}$. Then Sections 5.2 to 5.4 focus on the efficient computation of such a kernel basis.

Corollary 5.2. *Let* $[\hat{\mathbf{U}} \;\; \hat{\mathbf{R}}] \in \mathbb{K}[x]^{m\times 2m}$ *be a* $-\boldsymbol{d}$*-reduced basis of* $\mathcal{K}(\mathbf{F})$*. Then,* $\hat{\mathbf{R}}$ *is a* $-\boldsymbol{\delta}$*-reduced form of* $\mathbf{A}$*, and* $\mathbf{P} = (\mathrm{lm}_{-\boldsymbol{\delta}}(\hat{\mathbf{R}}))^{-1}\hat{\mathbf{R}}$ *is the* $-\boldsymbol{s}$*-Popov form of* $\mathbf{A}$*.*

Proof. It suffices to prove that $\hat{\mathbf{R}}$ is a $-\boldsymbol{\delta}$-reduced form of $\mathbf{A}$; then, the conclusion follows from [26, Lem. 4.1]. Both matrices $[\hat{\mathbf{U}} \;\; \hat{\mathbf{R}}]$ and $[\mathbf{U} \;\; \mathbf{R}]$ are $-\boldsymbol{d}$-reduced and thus have the same $-\boldsymbol{d}$-row degree up to permutation. Lemma 5.1 yields $\mathrm{rdeg}_{-\boldsymbol{d}}([\mathbf{U} \;\; \mathbf{R}]) = \mathbf{0}$, hence $\mathrm{rdeg}_{-\boldsymbol{d}}([\hat{\mathbf{U}} \;\; \hat{\mathbf{R}}]) = \mathbf{0}$. In particular, since $-\boldsymbol{d} = (\boldsymbol{\delta} - \boldsymbol{s}, -\boldsymbol{\delta})$, we have $\mathrm{rdeg}_{-\boldsymbol{\delta}}(\hat{\mathbf{R}}) \leq \mathbf{0}$.

We conclude by applying Lemma 2.3 to the row space of $\mathbf{A}$, which has rank m, and which has a basis $\mathbf{R}$ in $-\boldsymbol{\delta}$-reduced form with $-\boldsymbol{\delta}$-row degree $\mathbf{0}$. From $\hat{\mathbf{U}}\mathbf{A} = \hat{\mathbf{R}}$, we obtain that the rows of $\hat{\mathbf{R}}$ are in this row space. This identity also implies that $\hat{\mathbf{R}}$ has rank m, otherwise there would exist a nonzero vector in the left kernel of $\hat{\mathbf{R}}$, which would also be in the left kernel of $\hat{\mathbf{U}}$ since $\mathbf{A}$ is nonsingular, hence it would be in the left kernel of the full row rank matrix $[\hat{\mathbf{U}} \;\; \hat{\mathbf{R}}]$. The assumptions of the lemma are satisfied, and thus $\hat{\mathbf{R}}$ is a $-\boldsymbol{\delta}$-reduced form of $\mathbf{A}$. □

5.2. Reducing to the case of balanced pivot degree

Here, we show that the output column partial linearization, used previously in algorithms for approximant bases and generalizations of them [46, Sec. 3], [55, 26], can be applied to kernel computations when the sought kernel basis has nonnegative shifted row degree. The main effect of this transformation is to make the shift and output degrees more balanced, while preserving most other properties of $(\boldsymbol{s}, \mathbf{F})$. The transformation itself, defined below, is essentially a column-wise x^{δ}-adic expansion for a well-chosen integer parameter δ.

Definition 5.3. *Let* $\boldsymbol{s} = (s_1, \ldots, s_m) \in \mathbb{Z}^m_{\geq 0}$*, and let* $\delta \in \mathbb{Z}_{>0}$*. For* $1 \leq j \leq m$*, write* $s_j = (\alpha_j - 1)\delta + \beta_j$ *with* $\alpha_j = \lceil s_j/\delta \rceil$ *and* $1 \leq \beta_j \leq \delta$ *if* $s_j \neq 0$*, and* $\alpha_j = 1$ *and* $\beta_j = 0$ *if* $s_j = 0$*. Define* $\overline{m} = \alpha_1 + \cdots + \alpha_m$*, and the expansion-compression matrix* $\mathbf{E} \in \mathbb{K}[x]^{\overline{m}\times m}$ *as*

$$\mathbf{E} = \begin{bmatrix} 1 & & & \\ x^{\delta} & & & \\ \vdots & & & \\ x^{(\alpha_1-1)\delta} & & & \\ & \ddots & & \\ & & 1 & \\ & & x^{\delta} & \\ & & \vdots & \\ & & x^{(\alpha_m-1)\delta} & \end{bmatrix}.$$

Define also

$$\overline{\boldsymbol{s}} = (\underbrace{\delta, \ldots, \delta, \beta_1}_{\alpha_1}, \;\; \ldots \;\; , \underbrace{\delta, \ldots, \delta, \beta_m}_{\alpha_m}) \in \mathbb{Z}^{\overline{m}}_{\geq 0}.$$

In this context, for a matrix $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$*, we define the* column partial linearization of $\mathbf{K}$ *as the unique matrix* $\overline{\mathbf{K}} \in \mathbb{K}[x]^{k\times\overline{m}}$ *such that* $\mathbf{K} = \overline{\mathbf{K}}\mathbf{E}$ *and all the columns of* $\overline{\mathbf{K}}$ *whose index is not in* $\{\alpha_1 + \cdots + \alpha_j, 1 \leq j \leq m\}$ *have degree less than* δ*.*

We use the notation in this definition in all of Section 5.2. More explicitly, $\overline{\mathbf{K}} = [\overline{\mathbf{K}}_1 \ \cdots \ \overline{\mathbf{K}}_m]$ where $\overline{\mathbf{K}}_j \in \mathbb{K}[x]^{k\times\alpha_j}$ is the unique matrix such that

$$\mathbf{K}_{*,j} = \overline{\mathbf{K}}_j \begin{bmatrix} 1 \\ x^\delta \\ \vdots \\ x^{(\alpha_j-1)\delta} \end{bmatrix}$$

and the first $\alpha_j - 1$ columns of $\overline{\mathbf{K}}_j$ have degree less than δ.

This construction originates from [46, Sec. 3], where it was designed for approximant basis computations, in order to make the shift more balanced at the cost of a small increase of the dimension; this is stated in Lemma 5.4. As mentioned above, several slightly different versions of this column partial linearization have been given in the literature: each version requires some minor adaptations of the original construction to match the context. Here, in order to benefit from properties proved in [28, Sec. 5.1], we follow the construction in [28, Lem. 5.2]: one can check that Definition 5.3 is a specialization of the construction in that reference, for the shift $-\boldsymbol{s} \in \mathbb{Z}^m_{\leq 0}$ and taking the second parameter to be $t = \max(-\boldsymbol{s})$ so that $\boldsymbol{t} = -\boldsymbol{s} - \max(-\boldsymbol{s}) + t = -\boldsymbol{s}$.

Lemma 5.4. *The entries of $\overline{\boldsymbol{s}}$ are in $\{0, 1, \ldots, \delta\}$. Furthermore, if $\delta \geq |\boldsymbol{s}|/m$, then $m \leq \overline{m} \leq 2m$.*

Proof. The first remark is obvious. For the second one, note that $1 \leq \alpha_j \leq 1 + s_j/\delta$ holds by construction, for $1 \leq j \leq m$. Hence $m \leq \overline{m} \leq m + |\boldsymbol{s}|/\delta \leq 2m$. □

Importantly, this column partial linearization behaves well with respect to shifted row degrees and shifted reduced forms.

Lemma 5.5. *Let $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$ with $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$, and let $\overline{\mathbf{K}} \in \mathbb{K}[x]^{k\times\overline{m}}$ be its column partial linearization. Then, row degrees are preserved: $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = \operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K})$. Furthermore, if $\mathbf{K}$ is in $-\boldsymbol{s}$-weak Popov form, then $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form.*

Proof. We show that this follows from the second item in [28, Lem. 5.2]; as noted above, the shift $\boldsymbol{t} = (t_1, \ldots, t_m)$ in that reference corresponds to $-\boldsymbol{s} = (-s_1, \ldots, -s_m)$ here. Let $(\pi_i, \delta_i)_{1\leq i\leq k}$ denote the $-\boldsymbol{s}$-pivot profile of $\mathbf{K}$. By definition of the $-\boldsymbol{s}$-pivot index and degree, $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) = (\delta_i - s_{\pi_i})_{1\leq i\leq k}$, hence our assumption $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$ means that $\delta_i \geq s_{\pi_i}$ for $1 \leq i \leq k$. Thus, we can apply [28, Lem. 5.2] to each row of $\mathbf{K}$, from which we conclude that $\overline{\mathbf{K}}$ has $-\overline{\boldsymbol{s}}$-pivot profile $(\alpha_1 + \cdots + \alpha_{\pi_i}, \delta_i - s_{\pi_i} + \beta_{\pi_i})_{1\leq i\leq k}$.

First, since the entry of $-\overline{\boldsymbol{s}}$ at index $\alpha_1 + \cdots + \alpha_{\pi_i}$ is $-\beta_{\pi_i}$, this implies that

$$\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = (\delta_i - s_{\pi_i} + \beta_{\pi_i} - \beta_{\pi_i})_{1\leq i\leq k} = (\delta_i - s_{\pi_i})_{1\leq i\leq k} = \operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}).$$

Furthermore, if $\mathbf{K}$ is in $-\boldsymbol{s}$-weak Popov form, then $(\pi_i)_{1\leq i\leq k}$ is increasing, hence $(\alpha_1+\cdots+\alpha_{\pi_i})_{1\leq i\leq k}$ is increasing as well and $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form. □

Remark 5.6. *One may further note the following properties:*

- *Writing $\operatorname{lm}_{-\boldsymbol{s}}(\mathbf{K}) = [\mathbf{L}_{*,1} \ \cdots \ \mathbf{L}_{*,m}] \in \mathbb{K}^{k\times m}$, we have*

$$\operatorname{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = [\underbrace{\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,1}}_{\alpha_1} \ \cdots \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,m}}_{\alpha_m}] \in \mathbb{K}^{k\times\overline{m}}.$$

- *If* $\mathbf{K}$ *is in* $-\boldsymbol{s}$*-Popov form, then* $\overline{\mathbf{K}}$ *is in* $-\overline{\boldsymbol{s}}$*-Popov form.*

These properties are not used here, but for reference we provide a proof in Appendix A.

We will also need properties for the converse operation, going from some matrix $\mathbf{P} \in \mathbb{K}[x]^{m \times \overline{m}}$ to its compression $\mathbf{PE}$.

Lemma 5.7. *Let* $\mathbf{P} \in \mathbb{K}[x]^{k \times \overline{m}}$ *have* $-\overline{\boldsymbol{s}}$*-pivot profile* $(\pi_i, \delta_i)_{1 \le i \le k}$ *and assume* $\pi_i = \alpha_1 + \cdots + \alpha_{j_i}$ *for some* $j_i \in \mathbb{Z}_{>0}$*, for* $1 \le i \le k$*. Then,* $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{PE}) = \operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P})$*, and* $\mathbf{PE}$ *has* $-\boldsymbol{s}$*-pivot profile*

$$(j_i, \delta_i + (\alpha_{j_i} - 1)\delta)_{1 \le i \le k} = (j_i, \delta_i + s_{j_i} - \beta_{j_i})_{1 \le i \le k}.$$

If $\mathbf{P}$ *is in* $-\overline{\boldsymbol{s}}$*-weak Popov form, then* $\mathbf{PE}$ *is in* $-\boldsymbol{s}$*-weak Popov form.*

Proof. The $-\boldsymbol{s}$-pivot profile of $\mathbf{PE}$ is directly obtained by applying the first item in [28, Lem. 5.2] to each row of $\mathbf{P}$. From this $-\boldsymbol{s}$-pivot profile, we get

$$\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{PE}) = (\delta_i + s_{j_i} - \beta_{j_i} - s_{j_i})_{1 \le i \le k} = (\delta_i - \beta_{j_i})_{1 \le i \le k}.$$

On the other hand, since the entry of $-\overline{\boldsymbol{s}}$ at index $\alpha_1 + \cdots + \alpha_{j_i}$ is $-\beta_{j_i}$, the $-\overline{\boldsymbol{s}}$-pivot profile of $\mathbf{P}$ yields $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P}) = (\delta_i - \beta_{j_i})_{1 \le i \le k}$. Thus, $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{PE}) = \operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P})$. Now, if $\mathbf{P}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form, then $(\alpha_1 + \cdots + \alpha_{j_i})_{1 \le i \le k}$ is increasing, which implies that $(j_i)_{1 \le i \le k}$ is increasing. As a result, $\mathbf{PE}$ is in $-\boldsymbol{s}$-weak Popov form, since $(j_i)_{1 \le i \le k}$ is the $-\boldsymbol{s}$-pivot index of $\mathbf{PE}$. □

Our approach for computing the kernel of $\mathbf{F}$ is based on the fact that if $\mathbf{K}$ is a basis of $\mathcal{K}(\mathbf{F})$ and $\overline{\mathbf{K}}$ is its column partial linearization, then $\mathbf{KF} = \overline{\mathbf{K}}\mathbf{EF} = \mathbf{0}$. This identity shows that the kernel $\mathcal{K}(\mathbf{EF})$ contains the rows of $\overline{\mathbf{K}}$, so we may hope to recover $\overline{\mathbf{K}}$, and thus $\mathbf{K} = \overline{\mathbf{K}}\mathbf{E}$, from a basis of $\mathcal{K}(\mathbf{EF})$; the main advantage is that the latter basis is computed with the balanced shift $-\overline{\boldsymbol{s}}$. Note that a basis of $\mathcal{K}(\mathbf{EF})$ does not straightforwardly yield $\overline{\mathbf{K}}$, at least because this kernel also contains $\mathcal{K}(\mathbf{E})$. In Lemma 5.8 we exhibit a basis $\mathbf{S}$ for $\mathcal{K}(\mathbf{E})$, and then in Lemma 5.9 we show that $\mathcal{K}(\mathbf{EF})$ is generated by $\overline{\mathbf{K}}$ and $\mathbf{S}$. We also give properties which allow us, from a basis of $\mathcal{K}(\mathbf{EF})$, to easily recover a basis $\mathbf{K}$ of $\mathcal{K}(\mathbf{F})$ which has the sought form (see Lemma 5.11).

Lemma 5.8. *The matrix* $\mathbf{S} = \operatorname{diag}(\mathbf{S}_1, \ldots, \mathbf{S}_m) \in \mathbb{K}[x]^{(\overline{m}-m) \times \overline{m}}$*, where*

$$\mathbf{S}_j = \begin{bmatrix} x^\delta & -1 & & \\ & \ddots & \ddots & \\ & & x^\delta & -1 \end{bmatrix} \in \mathbb{K}[x]^{(\alpha_j - 1) \times \alpha_j}$$

for $1 \le j \le m$*, is the* $\mathbf{0}$*-Popov basis of the kernel* $\mathcal{K}(\mathbf{E})$*. Furthermore,* $\mathbf{S}$ *is also in* $-\overline{\boldsymbol{s}}$*-Popov form, it has* $-\overline{\boldsymbol{s}}$*-row degree* $\mathbf{0}$*, and its* $-\overline{\boldsymbol{s}}$*-pivot profile is* $(\alpha_1 + \cdots + \alpha_{j-1} + i, \delta)_{1 \le i < \alpha_j, 1 \le j \le m}$*.*

Proof. By construction, $\mathbf{SE} = \mathbf{0}$ and $\mathbf{S}$ is in $\mathbf{0}$-Popov form. Besides, $\mathbf{S}$ has rank $\overline{m} - m$, which is the rank of $\mathcal{K}(\mathbf{E})$ since $\mathbf{E}$ has rank m. Now, observe that there is no nonzero vector of degree less than δ in the left kernel of the vector $[1 \quad x^\delta \quad \cdots \quad x^{(\alpha_j - 1)\delta}]^\mathsf{T}$, and thus there is no nonzero vector of degree less than δ in $\mathcal{K}(\mathbf{E})$. It follows that $\mathbf{S}$ is a basis of $\mathcal{K}(\mathbf{E})$. Indeed, if $\mathbf{K} \in \mathbb{K}[x]^{(\overline{m}-m) \times \overline{m}}$ is a basis of $\mathcal{K}(\mathbf{E})$ in $\mathbf{0}$-reduced form, then $\operatorname{rdeg}(\mathbf{K}) \ge (\delta, \ldots, \delta)$. Since $\mathbf{SE} = \mathbf{0}$, we have $\mathbf{S} = \mathbf{UK}$ for some nonsingular $\mathbf{U} \in \mathbb{K}[x]^{(\overline{m}-m) \times (\overline{m}-m)}$. By the predictable degree property,

$$(\delta, \ldots, \delta) = \operatorname{rdeg}(\mathbf{S}) = \operatorname{rdeg}(\mathbf{UK}) = \operatorname{rdeg}_{\operatorname{rdeg}(\mathbf{K})}(\mathbf{U}) \ge \operatorname{rdeg}(\mathbf{U}) + (\delta, \ldots, \delta),$$

hence $\mathbf{U}$ is constant. Thus $\mathbf{U}$ is unimodular, and $\mathbf{S}$ is a basis of $\mathcal{K}(\mathbf{E})$.

Now consider j such that $\alpha_j > 1$, and write $\boldsymbol{t} = (-\delta, \dots, -\delta, -\beta_j) \in \mathbb{Z}^{\alpha_j}$. Then, the definition of α_j and β_j, notably the fact that $-\beta_j < 0$, ensures that $\mathbf{S}_j$ is in $\boldsymbol{t}$-Popov form with $\boldsymbol{t}$-pivot degree $(\delta, \dots, \delta)$ and $\boldsymbol{t}$-pivot index $(1, \dots, \alpha_j - 1)$. The conclusion about $\mathbf{S}$ follows. □

Lemma 5.9. *Let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$. The following properties hold.*

(i) Column degrees are preserved: $\operatorname{cdeg}_{\overline{\boldsymbol{s}}}(\mathbf{EF}) = \operatorname{cdeg}_{\boldsymbol{s}}(\mathbf{F})$.

(ii) The kernels of $\mathbf{F}$ and $\mathbf{EF}$ are related by $\mathcal{K}(\mathbf{EF})\mathbf{E} = \mathcal{K}(\mathbf{F})$.

(iii) Let $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$ be a basis of $\mathcal{K}(\mathbf{F})$ (hence $k = m - \operatorname{rank}(\mathbf{F})$), and let $\overline{\mathbf{K}} \in \mathbb{K}[x]^{k\times\overline{m}}$ be any matrix such that $\mathbf{K} = \overline{\mathbf{K}}\mathbf{E}$. Then,

$$\mathbf{B} = \begin{bmatrix} \overline{\mathbf{K}} \\ \mathbf{S} \end{bmatrix} \in \mathbb{K}[x]^{(k+\overline{m}-m)\times\overline{m}} \tag{5}$$

is a basis of $\mathcal{K}(\mathbf{EF})$, where $\mathbf{S}$ is the matrix defined in Lemma 5.8.

(iv) Assume that $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$ and that $\overline{\mathbf{K}}$ is the column partial linearization of $\mathbf{K}$. Then, $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{B}) = (\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}), \mathbf{0})$, and if $\mathbf{K}$ is in $-\boldsymbol{s}$-weak Popov form, then $\mathbf{B}$ is in $-\overline{\boldsymbol{s}}$-unordered weak Popov form.

Proof. (i) By definition, $\operatorname{cdeg}_{\overline{\boldsymbol{s}}}(\mathbf{EF}) = \operatorname{cdeg}(\mathbf{X}^{\overline{\boldsymbol{s}}}\mathbf{EF})$ and $\operatorname{cdeg}_{\boldsymbol{s}}(\mathbf{F}) = \operatorname{cdeg}(\mathbf{X}^{\boldsymbol{s}}\mathbf{F})$. Since the matrix

$$\mathbf{X}^{\overline{\boldsymbol{s}}}\mathbf{E} = \begin{bmatrix} X^{\delta} & & & \\ \vdots & & & \\ X^{(\alpha_1-1)\delta} & & & \\ X^{s_1} & & & \\ & \ddots & & \\ & & X^{\delta} & \\ & & \vdots & \\ & & X^{(\alpha_m-1)\delta} & \\ & & X^{s_m} & \end{bmatrix}$$

has column degree $\boldsymbol{s}$ and contains $\mathbf{X}^{\boldsymbol{s}}$ as a subset of its rows, we get $\operatorname{cdeg}(\mathbf{X}^{\overline{\boldsymbol{s}}}\mathbf{EF}) = \operatorname{cdeg}(\mathbf{X}^{\boldsymbol{s}}\mathbf{F})$.

(ii) For any vector $\mathbf{p} \in \mathcal{K}(\mathbf{EF})$, we have $\mathbf{pEF} = \mathbf{0}$, which means $\mathbf{pE} \in \mathcal{K}(\mathbf{F})$. Conversely, from any $\mathbf{p} \in \mathcal{K}(\mathbf{F})$ we can construct $\mathbf{q} \in \mathbb{K}[x]^{1\times\overline{m}}$ such that $\mathbf{qE} = \mathbf{p}$ since $\mathbf{E}$ has the identity as a subset of its rows; then, $\mathbf{qEF} = \mathbf{pF} = \mathbf{0}$, which means $\mathbf{q} \in \mathcal{K}(\mathbf{EF})$.

(iii) We recall that, by Lemma 5.8, $\mathbf{S}$ is a basis of $\mathcal{K}(\mathbf{E})$. The rows of $\mathbf{B}$ are in $\mathcal{K}(\mathbf{EF})$, since $\mathbf{SEF} = \mathbf{0}$ and $\overline{\mathbf{K}}\mathbf{EF} = \mathbf{KF} = \mathbf{0}$. Now, we want to prove that any $\mathbf{u} \in \mathcal{K}(\mathbf{EF})$ is a $\mathbb{K}[x]$-linear combination of the rows of $\mathbf{B}$. By Item (ii), $\mathbf{uE} \in \mathcal{K}(\mathbf{F})$; thus $\mathbf{uE} = \mathbf{vK} = \mathbf{v}\overline{\mathbf{K}}\mathbf{E}$ for some $\mathbf{v} \in \mathbb{K}[x]^{1\times k}$. Therefore $\mathbf{u} - \mathbf{v}\overline{\mathbf{K}} \in \mathcal{K}(\mathbf{E})$, and it follows that $\mathbf{u} - \mathbf{v}\overline{\mathbf{K}} = \mathbf{wS}$ for some $\mathbf{w} \in \mathbb{K}[x]^{1\times\overline{m}}$. This yields $\mathbf{u} = [\mathbf{v} \;\; \mathbf{w}]\mathbf{B}$.

(iv) Since $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$, we can apply Lemma 5.5, which yields $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) = \operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}})$, hence $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{B}) = (\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}), \mathbf{0})$. Now, if $\mathbf{K}$ is in $-\boldsymbol{s}$-weak Popov form, then Lemma 5.5 shows that $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form and that all entries of its $-\overline{\boldsymbol{s}}$-pivot index are in $\{\alpha_1 + \cdots + \alpha_i, 1 \leq i \leq m\}$. Besides, by Lemma 5.8, $\mathbf{S}$ is in $-\overline{\boldsymbol{s}}$-Popov form with a $-\overline{\boldsymbol{s}}$-pivot index which is disjoint from $\{\alpha_1 + \cdots + \alpha_i, 1 \leq i \leq m\}$. Hence $\mathbf{B}$ is in $-\overline{\boldsymbol{s}}$-unordered weak Popov form. □

Remark 5.10. *Similarly to Item (iv), one may observe that if $\mathbf{K}$ is in $-\boldsymbol{s}$-reduced form, then $\mathbf{B}$ is in $-\overline{\boldsymbol{s}}$-reduced form; and that if $\mathbf{K}$ is in $-\boldsymbol{s}$-Popov form, then $\mathbf{B}$ is in $-\overline{\boldsymbol{s}}$-Popov form up to row permutation. These points will not be used here; for reference a proof is given in Appendix A.*

Finally, we combine the above results to show that one can compute a basis of $\mathcal{K}(\mathbf{F})$ by computing a $-\overline{\boldsymbol{s}}$-weak Popov basis of $\mathcal{K}(\mathbf{EF})$ and taking a submatrix of it.

Lemma 5.11. *Let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ and let $r = \overline{m} - \operatorname{rank}(\mathbf{F})$ be the rank of $\mathcal{K}(\mathbf{EF})$. Assume that $-\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have nonnegative $-\boldsymbol{s}$-row degree. Let $\mathbf{Q} \in \mathbb{K}[x]^{r\times\overline{m}}$ be a $-\overline{\boldsymbol{s}}$-weak Popov basis of $\mathcal{K}(\mathbf{EF})$. Let $\mathbf{P} \in \mathbb{K}[x]^{k\times\overline{m}}$ be the submatrix of the rows of $\mathbf{Q}$ whose $-\overline{\boldsymbol{s}}$-pivot index is in $\{\alpha_1 + \cdots + \alpha_j, 1 \le j \le m\}$. Then, $\mathbf{PE}$ is a $-\boldsymbol{s}$-weak Popov basis of $\mathcal{K}(\mathbf{F})$.*

Proof. We first prove that the number of rows of $\mathbf{PE}$ is the rank of $\mathcal{K}(\mathbf{F})$, that is, $k = m - \operatorname{rank}(\mathbf{F})$. Indeed, by Item (iv) of Lemma 5.9, the $-\overline{\boldsymbol{s}}$-pivot index of the $-\overline{\boldsymbol{s}}$-Popov basis of $\mathcal{K}(\mathbf{EF})$ contains the $-\overline{\boldsymbol{s}}$-pivot index of $\mathbf{S}$. By Lemma 5.8, the latter is the tuple formed by the integers in the set $\{1,\dots,\overline{m}\} \setminus \{\alpha_1 + \cdots + \alpha_j, 1 \le j \le m\}$ sorted in increasing order. Since $\mathbf{P}$ is the submatrix of the rows of $\mathbf{Q}$ whose $-\overline{\boldsymbol{s}}$-pivot index is not in this set, $\mathbf{P}$ has $k = r - (\overline{m} - m) = m - \operatorname{rank}(\mathbf{F})$ rows.

Now, by construction, $\mathbf{P}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form and its $-\overline{\boldsymbol{s}}$-pivot index has entries in $\{\alpha_1 + \cdots + \alpha_i, 1 \le i \le m\}$. Thus we can apply Lemma 5.7; it ensures that $\mathbf{PE}$ is in $-\boldsymbol{s}$-weak Popov form and that $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{PE}) = \operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P})$.

It remains to prove that $\mathbf{PE}$ is a basis of $\mathcal{K}(\mathbf{F})$. Let $\mathbf{K} \in \mathbb{K}^{k\times m}$ be the $-\boldsymbol{s}$-Popov basis of $\mathcal{K}(\mathbf{F})$, and let $\overline{\mathbf{K}} \in \mathbb{K}^{k\times\overline{m}}$ be its column partial linearization. Let $\boldsymbol{d} = \operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K})$, which has nonnegative entries by assumption. Then, according to Lemma 5.5, $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form, and $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = \boldsymbol{d}$. Then, we define $\mathbf{B} \in \mathbb{K}^{(k+\overline{m}-m)\times\overline{m}}$ as in Eq. (5); by Item (iv) of Lemma 5.9, the matrix $\mathbf{B}$ is a $-\overline{\boldsymbol{s}}$-unordered weak Popov basis of $\mathcal{K}(\mathbf{EF})$.

Then, Lemma 2.5 shows that $\mathbf{Q}$ has the same $-\overline{\boldsymbol{s}}$-pivot profile as the row permutation of $\mathbf{B}$ which is in $-\overline{\boldsymbol{s}}$-weak Popov form. Since $\mathbf{P}$ (resp. $\overline{\mathbf{K}}$) is the submatrix of the rows of $\mathbf{Q}$ (resp. $\mathbf{B}$) whose $-\overline{\boldsymbol{s}}$-pivot index is in $\{\alpha_1 + \cdots + \alpha_j, 1 \le j \le m\}$, and since both $\mathbf{P}$ and $\overline{\mathbf{K}}$ are in $-\overline{\boldsymbol{s}}$-weak Popov form, it follows that $\mathbf{P}$ and $\overline{\mathbf{K}}$ have the same $-\overline{\boldsymbol{s}}$-pivot profile. In particular, we have $\operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P}) = \operatorname{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = \boldsymbol{d}$, from which we get $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{PE}) = \boldsymbol{d}$. According to Lemma 2.3, since the rows of $\mathbf{PE}$ are in $\mathcal{K}(\mathbf{F})$, this implies that $\mathbf{PE}$ is a basis of $\mathcal{K}(\mathbf{F})$. □

5.3. Reducing to the case of a balanced order

Now we apply the overlapping partial linearization of [46, Sec. 2], more precisely the version in [28, Sec. 5.2] which supports arbitrary $\boldsymbol{\gamma}$ as showed in the definition below that we recall for completeness. In the next lemma, we will show that the sought kernel basis can be retrieved as a submatrix of an approximant basis for the transformed problem.

Definition 5.12 ([46, 28])**.** *Let $\boldsymbol{\gamma} = (\gamma_1,\dots,\gamma_n) \in \mathbb{Z}^n_{>0}$, let $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ with $\operatorname{cdeg}(\mathbf{F}) < \boldsymbol{\gamma}$, and let $\mu \in \mathbb{Z}_{>0}$. Then, for $1 \le i \le n$, let $\gamma_i = \alpha_i\mu + \beta_i$ with $\alpha_i = \left\lceil \frac{\gamma_i}{\mu} - 1 \right\rceil$ and $1 \le \beta_i \le \mu$. Let also $\overline{n} = \max(\alpha_1 - 1, 0) + \cdots + \max(\alpha_n - 1, 0)$, and define*

$$\mathcal{L}_\mu(\boldsymbol{\gamma}) = (\overline{\gamma}_1,\dots,\overline{\gamma}_n) \in \mathbb{Z}^{n+\overline{n}}_{>0},$$

where $\overline{\gamma}_i = (2\mu,\dots,2\mu,\mu+\beta_i) \in \mathbb{Z}^{\alpha_i}_{>0}$ if $\alpha_i > 1$ and $\overline{\gamma}_i = \gamma_i$ otherwise. Considering the ith column of $\mathbf{F}$, we write its x^μ-adic representation as

$$\mathbf{F}_{*,i} = \mathbf{F}^{(0)}_{*,i} + \mathbf{F}^{(1)}_{*,i}x^\mu + \cdots + \mathbf{F}^{(\alpha_i)}_{*,i}x^{\alpha_i\mu}$$
$$\text{where } \operatorname{cdeg}([\mathbf{F}^{(0)}_{*,i} \ \ \mathbf{F}^{(1)}_{*,i} \ \ \cdots \ \ \mathbf{F}^{(\alpha_i)}_{*,i}]) < (\mu,\dots,\mu,\beta_i).$$

If $\alpha_i > 1$, we define

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

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

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

is called the overlapping linearization *of* $\mathbf{F}$ *with respect to* $\boldsymbol{\gamma}$ *and* μ*.*

Lemma 5.13. *Let* $\mathbf{F} \in \mathbb{K}[x]^{m \times n}$*, and let* $\boldsymbol{s} \in \mathbb{Z}^m_{\geq 0}$ *be such that there exists a basis of* $\mathcal{K}(\mathbf{F})$ *with nonpositive* $-\boldsymbol{s}$*-row degree. Let* $\mu \in \mathbb{Z}_{>0}$ *with* $\mu > \max(\boldsymbol{s})$*,* $\boldsymbol{t} = (\boldsymbol{s}, \mu - 1, \ldots, \mu - 1) \in \mathbb{Z}^{m+\overline{n}}_{\geq 0}$*, and* $\boldsymbol{\gamma} \in \mathbb{Z}^n_{>0}$ *with* $\boldsymbol{\gamma} \geq \operatorname{cdeg}_{\boldsymbol{s}}(\mathbf{F}) + 1$*. Let* $\mathbf{M} \in \mathbb{K}[x]^{(m+\overline{n}) \times (m+\overline{n})}$ *be a* $-\boldsymbol{t}$*-reduced basis of* $\mathcal{A}_{\mathcal{L}_\mu(\boldsymbol{\gamma})}(\mathcal{L}_{\boldsymbol{\gamma},\mu}(\mathbf{F}))$*. Then, exactly* $k = m - \operatorname{rank}(\mathbf{F})$ *rows of* $\mathbf{M}$ *have nonpositive* $-\boldsymbol{t}$*-degree, and the first* m *columns of these rows form a matrix* $\mathbf{K} \in \mathbb{K}[x]^{k \times m}$ *which is a* $-\boldsymbol{s}$*-reduced basis of* $\mathcal{K}(\mathbf{F})$*. Besides, if* $\mathbf{M}$ *is in* $-\boldsymbol{t}$*-weak Popov form, then* $\mathbf{K}$ *is in* $-\boldsymbol{s}$*-weak Popov form.*

Proof. Let $[\mathbf{K} \;\; \mathbf{Q}]$ be the submatrix of $\mathbf{M}$ formed by its rows of nonpositive $-\boldsymbol{t}$-degree, where $\mathbf{K} \in \mathbb{K}[x]^{k \times m}$ and $\mathbf{Q} \in \mathbb{K}[x]^{k \times \overline{n}}$; we have $0 \leq k \leq m + \overline{n}$. By choice of $\boldsymbol{t}$, from $\operatorname{rdeg}_{-\boldsymbol{t}}([\mathbf{K} \;\; \mathbf{Q}]) \leq \mathbf{0}$ we get $\deg(\mathbf{Q}) < \mu$ and $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \leq \mathbf{0}$.

In particular, $\deg(\mathbf{K}) \leq \max(\boldsymbol{s}) < \mu$: applying the second item in [28, Lem. 5.6] to each row of $[\mathbf{K} \;\; \mathbf{Q}]$ shows that $\operatorname{rdeg}(\mathbf{Q}) < \operatorname{rdeg}(\mathbf{K})$ and that the rows of $\mathbf{K}$ are in $\mathcal{A}_{\boldsymbol{\gamma}}(\mathbf{F})$, that is, $\mathbf{K}\mathbf{F} = \mathbf{0} \bmod \mathbf{X}^{\boldsymbol{\gamma}}$. On the other hand, $\operatorname{cdeg}(\mathbf{K}) \leq \boldsymbol{s}$ implies that $\operatorname{cdeg}(\mathbf{K}\mathbf{F}) \leq \operatorname{cdeg}_{\boldsymbol{s}}(\mathbf{F}) < \boldsymbol{\gamma}$, hence $\mathbf{K}\mathbf{F} = \mathbf{0}$, i.e. the rows of $\mathbf{K}$ are in $\mathcal{K}(\mathbf{F})$. This implies that the rank of $\mathbf{K}$ is at most the rank of the module $\mathcal{K}(\mathbf{F})$, i.e. $\operatorname{rank}(\mathbf{K}) \leq m - r$ where $r = \operatorname{rank}(\mathbf{F})$.

Furthermore, from $\operatorname{rdeg}(\mathbf{Q}) < \operatorname{rdeg}(\mathbf{K})$ and $\max(\boldsymbol{s}) < \mu$ we obtain

$$\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \operatorname{rdeg}(\mathbf{K}) - \max(\boldsymbol{s}) > \operatorname{rdeg}(\mathbf{Q}) - \mu + 1 = \operatorname{rdeg}_{(-\mu+1,\ldots,-\mu+1)}(\mathbf{Q}),$$

hence by choice of $\boldsymbol{t}$ we have $\operatorname{lm}_{-\boldsymbol{t}}([\mathbf{K} \;\; \mathbf{Q}]) = [\operatorname{lm}_{-\boldsymbol{s}}(\mathbf{K}) \;\; \mathbf{0}]$. Since this matrix is a subset of the rows of the nonsingular matrix $\operatorname{lm}_{-\boldsymbol{t}}(\mathbf{M})$, it has full row rank, and thus $\operatorname{lm}_{-\boldsymbol{s}}(\mathbf{K})$ has full row rank. This shows that $\mathbf{K}$ is in $-\boldsymbol{s}$-reduced form, and that $k = \operatorname{rank}(\mathbf{K})$. If $\mathbf{M}$ is furthermore in $-\boldsymbol{t}$-weak Popov form, then $[\mathbf{K} \;\; \mathbf{Q}]$ is in $-\boldsymbol{t}$-weak Popov form as well; then, the identity $\operatorname{lm}_{-\boldsymbol{t}}([\mathbf{K} \;\; \mathbf{Q}]) = [\operatorname{lm}_{-\boldsymbol{s}}(\mathbf{K}) \;\; \mathbf{0}]$ shows that the $-\boldsymbol{t}$-pivot entries of $[\mathbf{K} \;\; \mathbf{Q}]$ are all located in $\mathbf{K}$, hence $\mathbf{K}$ is in $-\boldsymbol{s}$-weak Popov form.

It remains to prove that $k = m - r$ and that the rows of $\mathbf{K}$ generate $\mathcal{K}(\mathbf{F})$.

By assumption, there exists a basis $\mathbf{P} \in \mathbb{K}[x]^{(m-r) \times m}$ of $\mathcal{K}(\mathbf{F})$ such that $\operatorname{cdeg}(\mathbf{P}) \leq \boldsymbol{s}$. In particular, the rows of $\mathbf{P}$ are in $\mathcal{A}_{\boldsymbol{\gamma}}(\mathbf{F})$, and applying the first item of [28, Lem. 5.6] to each of these rows shows that there exists a matrix $\mathbf{R} \in \mathbb{K}[x]^{(m-r) \times \overline{n}}$ with $\operatorname{rdeg}(\mathbf{R}) < \operatorname{rdeg}(\mathbf{P})$ and such that the rows of $[\mathbf{P} \;\; \mathbf{R}]$ are in $\mathcal{A}_{\mathcal{L}_\mu(\boldsymbol{\gamma})}(\mathcal{L}_{\boldsymbol{\gamma},\mu}(\mathbf{F}))$. Thus, $[\mathbf{P} \;\; \mathbf{R}]$ is a left multiple of $\mathbf{M}$.

A key remark now is that $\operatorname{rdeg}_{-\boldsymbol{t}}([\mathbf{P} \;\; \mathbf{R}]) \leq \mathbf{0}$. Indeed, by Lemma 2.1 $\operatorname{rdeg}_{-\boldsymbol{s}}(\mathbf{P}) \leq \mathbf{0}$ follows from $\operatorname{cdeg}(\mathbf{P}) \leq \boldsymbol{s}$, and we have

$$\deg(\mathbf{R}) < \deg(\mathbf{P}) = \max(\operatorname{cdeg}(\mathbf{P})) \leq \max(\boldsymbol{s}) < \mu.$$

Thus, since $\mathbf{M}$ is $-\boldsymbol{t}$-reduced, the predictable degree property ensures that $[\mathbf{P} \;\; \mathbf{R}]$ is a left multiple of $\mathbf{M}$ which does not involve the rows of $\mathbf{M}$ of positive $-\boldsymbol{t}$-degree, i.e. a left multiple of $[\mathbf{K} \;\; \mathbf{Q}]$. In particular, $\mathbf{P}$ is a left multiple of $\mathbf{K}$: there exists a matrix $\mathbf{V} \in \mathbb{K}[x]^{(m-r) \times k}$ such that $\mathbf{V}\mathbf{K} = \mathbf{P}$.

Since $\mathbf{P}$ has rank $m - r$, we obtain $k \geq m - r$, hence $k = m - r$. On the other hand, since the rows of $\mathbf{K}$ are in $\mathcal{K}(\mathbf{F})$, there exists a matrix $\mathbf{W} \in \mathbb{K}[x]^{k\times k}$ such that $\mathbf{WP} = \mathbf{K}$. It follows that $\mathbf{P} = \mathbf{VWP}$, and since $\mathbf{P}$ has full row rank this implies $\mathbf{VW} = \mathbf{I}_k$. This means that $\mathbf{P}$ and $\mathbf{K}$ are left unimodularly equivalent, hence $\mathbf{K}$ is a basis of $\mathcal{K}(\mathbf{F})$, which concludes the proof. □

5.4. Computing kernel bases with known pivot degree

After applying the transformations presented in Sections 5.2 and 5.3, we are left with the computation of an approximant basis for a balanced order $\mathcal{L}_\mu(\boldsymbol{\gamma})$ and a balanced shift $-\overline{\boldsymbol{s}}$: this is done efficiently by PM-Basis, designed in [21] as an improvement of [3, Algo. SPHPS]. Here, we use the version in [28, Algo. 2] which ensures that the output basis is in $-\overline{\boldsymbol{s}}$-weak Popov form.

Algorithm 4 KnownDegreeKernelBasis($\mathbf{F}, \boldsymbol{s}$)

Input: a matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$, and a nonnegative shift $\boldsymbol{s} \in \mathbb{Z}^m_{\geq 0}$.
Requirement: $-\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have $-\boldsymbol{s}$-row degree $\mathbf{0}$.
Output: a $-\boldsymbol{s}$-weak Popov basis of $\mathcal{K}(\mathbf{F})$.

1: ▷ *Step 1: Output column partial linearization* ◁
2: $\delta \leftarrow \lceil D/m \rceil \in \mathbb{Z}_{>0}$, where $D = \max(|\boldsymbol{s}|, |\mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F})|, 1)$
3: Apply Definition 5.3 to $(\boldsymbol{s}, \delta)$ to obtain the parameters and expansion-compression matrix: $(\alpha_1, \ldots, \alpha_m) \in \mathbb{Z}^m_{>0}$, $\overline{m} \in \mathbb{Z}_{>0}$, $\overline{\boldsymbol{s}} \in \mathbb{Z}^{\overline{m}}_{\geq 0}$, $\mathbf{E} \in \mathbb{K}[x]^{\overline{m}\times m}$
4: ▷ *Step 2: Overlapping partial linearization* ◁
5: $\boldsymbol{\gamma} \leftarrow \mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F}) + 1 \in \mathbb{Z}^n_{>0}$ ▷ *order for approximation, equal to* $\mathrm{cdeg}_{\overline{\boldsymbol{s}}}(\mathbf{EF}) + 1$
6: Apply Definition 5.12 to $(\boldsymbol{\gamma}, \mathbf{EF}, \delta + 1)$ to obtain the order $\mathcal{L}_{\delta+1}(\boldsymbol{\gamma}) \in \mathbb{Z}^{n+\overline{n}}_{>0}$ and the matrix $\mathcal{L}_{\boldsymbol{\gamma},\delta+1}(\mathbf{EF}) \in \mathbb{K}[x]^{(\overline{m}+\overline{n})\times(n+\overline{n})}$
7: ▷ *Step 3: Compute $-\boldsymbol{t}$-weak Popov basis of $\mathcal{A}_{\mathcal{L}_{\delta+1}(\boldsymbol{\gamma})}(\mathcal{L}_{\boldsymbol{\gamma},\delta+1}(\mathbf{EF}))$* ◁
8: $\boldsymbol{t} \leftarrow (\overline{\boldsymbol{s}}, \delta, \ldots, \delta) \in \mathbb{Z}^{\overline{m}+\overline{n}}_{\geq 0}$
9: $\Gamma \leftarrow \max(\mathcal{L}_{\delta+1}(\boldsymbol{\gamma}))$; $\mathbf{G} = \mathcal{L}_{\boldsymbol{\gamma},\delta+1}(\mathbf{EF})\mathbf{X}^{(\Gamma,\ldots,\Gamma)-\mathcal{L}_{\delta+1}(\boldsymbol{\gamma})}$ ▷ *use uniform order* $(\Gamma, \ldots, \Gamma)$
10: $\mathbf{M} \in \mathbb{K}[x]^{(\overline{m}+\overline{n})\times(\overline{m}+\overline{n})} \leftarrow$ PM-Basis$(\Gamma, \mathbf{G}, -\boldsymbol{t})$
11: ▷ *Step 4: Deduce first $-\overline{\boldsymbol{s}}$-weak Popov basis of $\mathcal{K}(\mathbf{EF})$, then $-\boldsymbol{s}$-weak Popov basis of $\mathcal{K}(\mathbf{F})$* ◁
12: $\mathbf{Q} \in \mathbb{K}[x]^{\overline{k}\times\overline{m}} \leftarrow$ first $\overline{m}$ columns of the rows of $\mathbf{M}$ which have nonpositive $-\boldsymbol{t}$-degree
13: $\mathbf{P} \in \mathbb{K}[x]^{k\times\overline{m}} \leftarrow$ the rows of $\mathbf{Q}$ whose $-\overline{\boldsymbol{s}}$-pivot index is in $\{\alpha_1 + \cdots + \alpha_j, 1 \leq j \leq m\}$
14: **return PE**

Proposition 5.14. *Algorithm 4 is correct. Let $D = \max(|\boldsymbol{s}|, |\mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F})|, 1)$. Then, assuming $m \geq n$ and using notation from the algorithm, its cost is bounded by the sum of:*

- *the cost of performing* PM-Basis *at order at most $2\lceil D/m \rceil + 2$ on an input matrix of row dimension $\overline{m} + \overline{n} \leq 3m$ and column dimension $n + \overline{n} \leq 2m$;*
- *$O(m^2)$ extra operations in $\mathbb{K}$.*

Thus, Algorithm 4 uses $O(m^\omega\, \mathsf{M}'(D/m))$ operations in $\mathbb{K}$.

Proof. Using the assumption that $-\boldsymbol{s}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have $-\boldsymbol{s}$-row degree $\mathbf{0}$ along with Item (iv) of Lemma 5.9 shows that the $-\overline{\boldsymbol{s}}$-reduced bases of $\mathcal{K}(\mathbf{EF})$ have $-\overline{\boldsymbol{s}}$-row degree $\mathbf{0}$. Thus, we can apply Lemma 5.13 to $(\mathbf{EF}, \overline{\boldsymbol{s}}, \mu, \boldsymbol{\gamma})$ with $\mu = \delta + 1 > \max(\overline{\boldsymbol{s}})$ and $\boldsymbol{\gamma} = \mathrm{cdeg}_{\overline{\boldsymbol{s}}}(\mathbf{EF}) + 1$, which is $\boldsymbol{\gamma} = \mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F}) + 1$ according to Item (i) of Lemma 5.9. Note that $\mathbf{M}$ is a $-\boldsymbol{t}$-weak Popov basis of $\mathcal{A}_{(\Gamma,\ldots,\Gamma)}(\mathbf{G}) = \mathcal{A}_{\mathcal{L}_{\delta+1}(\boldsymbol{\gamma})}(\mathcal{L}_{\boldsymbol{\gamma},\delta+1}(\mathbf{EF}))$ (see e.g. [28, Rmk. 3.3] for this approach to make the order uniform). Then, Lemma 5.13 states that the matrix $\mathbf{Q}$ at Line 12 has $\overline{k} = \overline{m} - \mathrm{rank}(\mathbf{EF}) =$

$\overline{m}$ − rank(**F**) rows and is a $-\overline{\boldsymbol{s}}$-weak Popov basis of $\mathcal{K}(\mathbf{EF})$. Then, by Lemma 5.11, **PE** is a $-\boldsymbol{s}$-weak Popov basis of $\mathcal{K}(\mathbf{F})$, hence the correctness.

For the cost analysis, we assume $m \geq n$, and we start by summarizing bounds on the dimensions and degrees at play. Lemma 5.4 yields $m \leq \overline{m} \leq 2m$, while Item (i) of Lemma 5.9 ensures $\mathrm{cdeg}_{\overline{\boldsymbol{s}}}(\mathbf{EF}) = \mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F})$. Each entry of $\mathcal{L}_{\delta+1}(\boldsymbol{\gamma})$ is at most $2(\delta+1) = 2\lceil D/m \rceil + 2$, by construction. Writing $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)$, by Definition 5.12 we have

$$\overline{n} = \max\left(\left\lceil \frac{\gamma_1}{\delta+1} - 1 \right\rceil, 0\right) + \cdots + \max\left(\left\lceil \frac{\gamma_n}{\delta+1} - 1 \right\rceil, 0\right) \leq \frac{\gamma_1}{\delta+1} + \cdots + \frac{\gamma_n}{\delta+1} = \frac{|\boldsymbol{\gamma}|}{\delta+1}.$$

Since $|\boldsymbol{\gamma}| = |\mathrm{cdeg}_{\boldsymbol{s}}(\mathbf{F})| + n \leq D + m$, and since $\delta + 1 \geq (D+m)/m$, it follows that $\overline{n} \leq m$. Besides, $\mathrm{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{P}) \leq \mathbf{0}$ by construction, so that $\mathrm{cdeg}(\mathbf{P}) \leq \overline{\boldsymbol{s}}$ by Lemma 2.1.

The only steps that involve operations in $\mathbb{K}$ are the call to PM-Basis at Line 10 and the multiplication **PE** at Line 14. The construction of **E** and the inequality $\mathrm{cdeg}(\mathbf{P}) \leq \overline{\boldsymbol{s}}$ imply that the product **PE** mainly involves concatenating vectors of coefficients; concerning operations in $\mathbb{K}$, there are $\overline{m} - m$ columns of **P** for which we may add the constant term of that column to the term of degree δ of the previous column. Therefore Line 14 has cost bound $O(\overline{m}k)$; since $\overline{m} \leq 2m$ and $k \leq m$ this is in $O(m^2)$. At Line 10, we call the approximant basis subroutine PM-Basis discussed in Section 2.5; since $\Gamma = \max(\mathcal{L}_{\delta+1}(\boldsymbol{\gamma}))$ is at most $2\lceil D/m \rceil + 2 \in O(1 + D/m)$, Lemma 2.9 states that this call uses

$$O\left((\overline{m} + \overline{n} + n + \overline{n})(\overline{m} + \overline{n})^{\omega-1}\, \mathsf{M}'(D/m)\right)$$

operations in $\mathbb{K}$. Since $n + \overline{n} \leq \overline{m} + \overline{n} \leq 3m$, this yields the claimed cost bound. □

5.5. Proof of Theorem 1.4

Algorithm 5 WeakPopovToPopov($\mathbf{A}, \boldsymbol{s}$)

Input: a matrix $\mathbf{A} \in \mathbb{K}[x]^{m \times n}$, a shift $\boldsymbol{s} \in \mathbb{Z}^n$ such that $\mathbf{A}$ is in $-\boldsymbol{s}$-weak Popov form.
Requirement: $\boldsymbol{s}_{\boldsymbol{\pi}} \geq \boldsymbol{\delta}$, where $(\boldsymbol{\pi}, \boldsymbol{\delta})$ is the $-\boldsymbol{s}$-pivot profile of $\mathbf{A}$.
Output: the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$.

1: ▷ *Step 1: Find unimodular transformation and $-\boldsymbol{\delta}$-reduced form of $\mathbf{A}_{*,\boldsymbol{\pi}}$* ◁
2: $(\boldsymbol{\pi}, \boldsymbol{\delta}) \leftarrow$ the $-\boldsymbol{s}$-pivot profile of $\mathbf{A}$
3: $[\mathbf{U} \;\; \mathbf{R}] \in \mathbb{K}[x]^{m \times 2m} \leftarrow$ KnownDegreeKernelBasis$\left(\begin{bmatrix} \mathbf{A}_{*,\boldsymbol{\pi}} \\ -\mathbf{I}_m \end{bmatrix}, (\boldsymbol{s}_{\boldsymbol{\pi}} - \boldsymbol{\delta}, \boldsymbol{\delta})\right)$
4: ▷ *Step 2: Deduce $-\boldsymbol{s}$-Popov form of $\mathbf{A}$* ◁
5: $\mathbf{P} \leftarrow$ zero matrix in $\mathbb{K}[x]^{m \times n}$
6: $\mathbf{P}_{*,\boldsymbol{\pi}} \leftarrow \mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$
7: $\mathbf{P}_{*,\{1,\ldots,n\}\setminus\boldsymbol{\pi}} \leftarrow \mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{U}\,\mathbf{A}_{*,\{1,\ldots,n\}\setminus\boldsymbol{\pi}}$
8: **return P**

For proving Theorem 1.4, we describe Algorithm WeakPopovToPopov (Algorithm 5) and we focus on the square case, $m = n$. Then, by definition of the $-\boldsymbol{s}$-weak Popov form, $\boldsymbol{\delta}$ is the tuple of degrees of the diagonal entries of $\mathbf{A}$, and $\boldsymbol{\pi} = (1, \ldots, m)$. Furthermore, in this case, $\boldsymbol{s}_{\boldsymbol{\pi}} = \boldsymbol{s}$, $\mathbf{A}_{*,\boldsymbol{\pi}} = \mathbf{A}$, and $\mathbf{P}_{*,\boldsymbol{\pi}} = \mathbf{P}$; in particular, we can discard the step at Line 7 since the submatrices it involves are empty.

First note that the shift $\boldsymbol{d} = (\boldsymbol{s}_{\boldsymbol{\pi}} - \boldsymbol{\delta}, \boldsymbol{\delta}) = (\boldsymbol{s} - \boldsymbol{\delta}, \boldsymbol{\delta})$ used at Line 3 is nonnegative. Besides, Lemma 5.1 shows that $-\boldsymbol{d}$-reduced bases of $\mathcal{K}(\mathbf{F})$ have $-\boldsymbol{d}$-row degree $\mathbf{0}$. Thus the requirements

of Algorithm KnownDegreeKernelBasis are met, and Proposition 5.14 shows that the matrix $[\mathbf{U} \quad \mathbf{R}]$ computed at Line 3 is a $-\boldsymbol{d}$-weak Popov basis of $\mathcal{K}(\mathbf{F})$. Then, Corollary 5.2 shows that the matrix $\mathbf{P} = \mathbf{P}_{*,\boldsymbol{\pi}} = \mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{R}$ computed at Line 6 is the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$. This proves that Algorithm 5 is correct.

Concerning the cost bound, we focus on the case $|\boldsymbol{s}| > 0$. Indeed, since $\boldsymbol{s} \geq \boldsymbol{\delta} \geq \mathbf{0}$, if $|\boldsymbol{s}| = 0$, then $\boldsymbol{s} = \boldsymbol{\delta} = \mathbf{0}$. In this case, no computation needs to be done: the $-\boldsymbol{s}$-Popov form of $\mathbf{A}$ is $\mathbf{I}_m$, the unique matrix in Popov form whose pivot degree is $\mathbf{0}$.

The cost of Line 6 is that of multiplying $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1} \in \mathbb{K}^{m\times m}$ by $\mathbf{R} \in \mathbb{K}[x]^{m\times m}$, which has column degree $\boldsymbol{\delta}$ as explained in Section 5.1; this computation corresponds to the second item in Theorem 1.4. This multiplication can be done by first performing a column linearization of $\mathbf{R}$ into a $m \times (m + |\boldsymbol{\delta}|)$ matrix $\overline{\mathbf{R}}$ over $\mathbb{K}$, computing $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\overline{\mathbf{R}}$, and finally compressing the result back into a polynomial matrix. This uses $O(m^\omega(1 + |\boldsymbol{\delta}|/m))$ operations in $\mathbb{K}$.

Concerning Line 3, we rely on Proposition 5.14. Here, the matrix we give as input to Algorithm KnownDegreeKernelBasis has dimensions $2m \times m$, hence the dimensions in Proposition 5.14 satisfy $\overline{m} \leq 4m$ and $\overline{n} \leq 2m$ (the latter bound comes from the proof of that proposition). Then, Proposition 5.14 states that Line 3 costs:

- $O(m^2)$ operations in $\mathbb{K}$, which is the third item in Theorem 1.4,
- one call to PM-Basis on a matrix of row dimension $\overline{m} + \overline{n} \leq 6m$, column dimension $m + \overline{n} \leq 3m$, and at order at most $2\lceil D/(2m)\rceil + 2$, where $D = \max(|\boldsymbol{d}|, |\mathrm{cdeg}_{\boldsymbol{d}}(\mathbf{F})|, 1)$ and $\mathbf{F}$ is the input matrix $[\mathbf{A}^\mathsf{T} \quad -\mathbf{I}_m]^\mathsf{T}$.

Besides, Proposition 5.14 also implies that Line 3 uses $O(m^\omega\,\mathsf{M}'(D/m))$ operations in $\mathbb{K}$.

We are going to prove that $D = |\boldsymbol{s}|$, which concludes the proof. Indeed, the previous paragraph then directly gives the overall cost bound $O(m^\omega\,\mathsf{M}'(|\boldsymbol{s}|/m))$ in Theorem 1.4, and using

$$2\left\lceil \frac{D}{2m} \right\rceil + 2 < 2\left(\frac{D}{2m} + 1\right) + 2 = |\boldsymbol{s}|/m + 4,$$

the previous paragraph also gives the first item in that theorem.

To observe that $D = |\boldsymbol{s}|$, we first use the definition of $\boldsymbol{d}$ to obtain $|\boldsymbol{d}| = |\boldsymbol{s} - \boldsymbol{\delta}| + |\boldsymbol{\delta}| = |\boldsymbol{s}|$. Since $|\boldsymbol{s}| \geq 1$, this gives $D = \max(|\boldsymbol{s}|, |\mathrm{cdeg}_{\boldsymbol{d}}(\mathbf{F})|)$. Now, $\mathrm{cdeg}_{\boldsymbol{d}}(\mathbf{F})$ is the entry-wise maximum of $\mathrm{cdeg}_{\boldsymbol{s}-\boldsymbol{\delta}}(\mathbf{A})$ and $\mathrm{cdeg}_{\boldsymbol{\delta}}(-\mathbf{I}_m) = \boldsymbol{\delta}$. Since $\mathbf{A}$ is in $-\boldsymbol{s}$-row reduced form with $\mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{A}) = -\boldsymbol{s} + \boldsymbol{\delta}$, Lemma 3.2 ensures that $\mathbf{A}^\mathsf{T}$ is in $\boldsymbol{s} - \boldsymbol{\delta}$-reduced form with $\mathrm{cdeg}_{\boldsymbol{s}-\boldsymbol{\delta}}(\mathbf{A}) = \mathrm{rdeg}_{\boldsymbol{s}-\boldsymbol{\delta}}(\mathbf{A}^\mathsf{T}) = \boldsymbol{s}$. Then, since $\boldsymbol{s} \geq \boldsymbol{\delta}$ we obtain $\mathrm{cdeg}_{\boldsymbol{d}}(\mathbf{F}) = \boldsymbol{s}$, and thus $D = |\boldsymbol{s}|$. This concludes the proof of Theorem 1.4.

As for the rectangular case $m < n$, the correctness of Algorithm 5 follows from the above proof in the square case, which shows that the algorithm correctly computes the $-\boldsymbol{s}_{\boldsymbol{\pi}}$-Popov form $\mathbf{P}_{*,\boldsymbol{\pi}}$ of $\mathbf{A}_{*,\boldsymbol{\pi}}$, and from [37, Lem. 5.1] concerning the computation of $\mathbf{P}_{*,\{1,\ldots,n\}\setminus\boldsymbol{\pi}}$ and the fact that $\mathrm{lm}_{-\boldsymbol{\delta}}(\mathbf{R})^{-1}\mathbf{U}$ is the unimodular matrix which transforms $\mathbf{A}$ into $\mathbf{P}$. The cost bound can be derived using the degree bounds on rectangular shifted Popov forms given in [5, 37] (we do not detail this here since this would add technical material unrelated to the main results of this article).

Appendix A.

In this appendix, we give proofs for Remarks 5.6 and 5.10. We use notation from Section 5.2: a shift $\boldsymbol{s} \in \mathbb{Z}^m_{\geq 0}$, an integer $\delta \in \mathbb{Z}_{>0}$, and a matrix $\mathbf{F} \in \mathbb{K}[x]^{m\times n}$ are given; $\mathbf{K} \in \mathbb{K}[x]^{k\times m}$ is a basis of $\mathcal{K}(\mathbf{F})$; $(\alpha_j, \beta_j)_{1\leq j\leq m}$ and $\overline{\boldsymbol{s}} \in \mathbb{Z}^{\overline{m}}_{\geq 0}$ and $\overline{\mathbf{K}} \in \mathbb{K}[x]^{k\times\overline{m}}$ are as described in Definition 5.3. Following the context of Remarks 5.6 and 5.10, we assume $\mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$, hence in particular $\mathrm{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = \mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{K})$ (see Lemma 5.5). We start with the claims in Remark 5.6.

Lemma A.1. *Writing* $\mathrm{lm}_{-\boldsymbol{s}}(\mathbf{K}) = [\mathbf{L}_{*,1} \ \cdots \ \mathbf{L}_{*,m}] \in \mathbb{K}^{k\times m}$*, we have*

$$\mathrm{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = [\underbrace{\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,1}}_{\alpha_1} \ \cdots \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,m}}_{\alpha_m}] \in \mathbb{K}^{k\times\overline{m}}.$$

If $\mathbf{K}$ *is in* $-\boldsymbol{s}$*-Popov form, then* $\overline{\mathbf{K}}$ *is in* $-\overline{\boldsymbol{s}}$*-Popov form.*

Proof. For given $i \in \{1,\ldots,k\}$ and $j \in \{1,\ldots,m\}$, we rely on the definition of a leading matrix (see Section 2.3): the entry (i,j) of $\mathrm{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}})$ is the coefficient of degree $d_i + \overline{\boldsymbol{s}}_j$ of $\overline{\mathbf{K}}_{i,j}$, where $d_i = \mathrm{rdeg}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}_{i,*}) = \mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{K}_{i,*}) \geq 0$. First consider the case $j \notin \{\alpha_1 + \cdots + \alpha_\pi, 1 \leq \pi \leq m\}$. Then the jth entry of $\overline{\boldsymbol{s}}$ is $\overline{\boldsymbol{s}}_j = \delta$, and by definition of the output column partial linearization $\overline{\mathbf{K}}_{i,j}$ has degree less than δ, hence its coefficient of degree $d_i + \delta \geq \delta$ must be zero. This proves that all columns of $\mathrm{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}})$ with index not in $\{\alpha_1 + \cdots + \alpha_\pi, 1 \leq \pi \leq m\}$ are indeed zero. It remains to prove that in the case $j = \alpha_1 + \cdots + \alpha_\pi$ for some $1 \leq \pi \leq m$, then the entry (i,j) of $\mathrm{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}})$ is equal to $\mathbf{L}_{i,\pi}$. This holds, since in this case we have $\overline{\boldsymbol{s}}_j = \beta_\pi$, and by construction of $\overline{\mathbf{K}}$ the coefficient of degree $d_i + \beta_\pi$ of $\overline{\mathbf{K}}_{i,j}$ is equal to the coefficient of degree $d_i + (\alpha_\pi - 1)\delta + \beta_\pi = d_i + s_\pi$ of $\mathbf{K}_{i,\pi}$, which itself is equal to $\mathbf{L}_{i,\pi}$ by definition of a leading matrix.

Now assume $\mathbf{K}$ is in $-\boldsymbol{s}$-Popov form. We have showed in Lemma 5.5 that $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-weak Popov form and that the row $\overline{\mathbf{K}}_{i,*}$ has its $-\overline{\boldsymbol{s}}$-pivot at index $\alpha_1 + \cdots + \alpha_{\pi_i}$ where π_i is the $-\boldsymbol{s}$-pivot index of $\mathbf{K}_{i,*}$. By construction, the column $\overline{\mathbf{K}}_{*,\alpha_1+\cdots+\alpha_{\pi_i}}$ is the part of nonnegative degree of the column $x^{-(\alpha_{\pi_i}-1)\delta}\mathbf{K}_{*,\pi_i} = x^{-s_{\pi_i}+\beta_i}\mathbf{K}_{*,\pi_i}$. It follows first that the $-\overline{\boldsymbol{s}}$-pivot entry $\overline{\mathbf{K}}_{i,\alpha_1+\cdots+\alpha_{\pi_i}}$ is monic since it is a high degree part of the (monic) $-\boldsymbol{s}$-pivot entry $\mathbf{K}_{i,\pi_i}$; and second that $\overline{\mathbf{K}}_{i,\alpha_1+\cdots+\alpha_{\pi_i}}$ has degree strictly larger than all other entries in the column $\overline{\mathbf{K}}_{*,\alpha_1+\cdots+\alpha_{\pi_i}}$ since $\mathbf{K}_{i,\pi_i}$ has degree strictly larger than all other entries in the column $\mathbf{K}_{*,\pi_i}$. Hence $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-Popov form. □

Now we prove the claims in Remark 5.10 concerning variants of Item (iv) of Lemma 5.9, using further notation from Section 5.2: $\mathbf{S}$ is the basis of $\mathcal{K}(\mathbf{E})$ described in Lemma 5.8 and $\mathbf{B} = \left[\begin{smallmatrix}\overline{\mathbf{K}}\\ \mathbf{S}\end{smallmatrix}\right] \in \mathbb{K}[x]^{(k+\overline{m}-m)\times\overline{m}}$ is the basis of $\mathcal{K}(\mathbf{EF})$ given in Eq. (5). As above we assume $\mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{K}) \geq \mathbf{0}$, and therefore $\mathrm{rdeg}_{-\overline{\boldsymbol{s}}}(\mathbf{B}) = (\mathrm{rdeg}_{-\boldsymbol{s}}(\mathbf{K}), \mathbf{0})$ as stated in Item (iv) of Lemma 5.9.

Lemma A.2. *With the above notation and assumptions,*

- *If* $\mathbf{K}$ *is in* $-\boldsymbol{s}$*-reduced form, then* $\mathbf{B}$ *is in* $-\overline{\boldsymbol{s}}$*-reduced form.*
- *If* $\mathbf{K}$ *is in* $-\boldsymbol{s}$*-Popov form, then* $\mathbf{B}$ *is in* $-\overline{\boldsymbol{s}}$*-Popov form up to row permutation.*

Proof. Suppose first that $\mathbf{K}$ is in $-\boldsymbol{s}$-reduced form. We apply Lemma A.1:

$$\mathrm{lm}_{-\overline{\boldsymbol{s}}}(\overline{\mathbf{K}}) = [\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,1} \ \cdots \ \mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{L}_{*,m}]$$

where $\mathrm{lm}_{-\boldsymbol{s}}(\mathbf{K}) = [\mathbf{L}_{*,1} \ \cdots \ \mathbf{L}_{*,m}] \in \mathbb{K}^{k\times m}$. The latter matrix has full rank, since $\mathbf{K}$ is in $-\boldsymbol{s}$-reduced form. On the other hand, for matrices such as the diagonal blocks of $\mathbf{S}$ we have

$$\mathrm{lm}_{(-\delta,\ldots,-\delta,-\beta)}\left(\begin{bmatrix} x^\delta & -1 & & \\ & \ddots & \ddots & \\ & & x^\delta & -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}$$

for any integer $\beta \geq 1$. As a result,

$$\mathrm{lm}_{-\overline{s}}(\mathbf{B}) = \begin{bmatrix} \mathrm{lm}_{-\overline{s}}(\overline{\mathbf{K}}) \\ \mathrm{lm}_{-\overline{s}}(\mathbf{S}) \end{bmatrix} = \left[\begin{array}{ccc|c|ccc} & & & \multicolumn{1}{c|}{\mathbf{L}_{*,1}} & \cdots & & \multicolumn{1}{c}{\mathbf{L}_{*,m}} \\ \multicolumn{3}{c|}{} & & & & \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ \hline & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{array}\right]$$

has full rank as well, which means that $\mathbf{B}$ is in $-\overline{s}$-reduced form.

Now, assume further that $\mathbf{K}$ is in $-\boldsymbol{s}$-Popov form. Then Lemma A.1 states that $\overline{\mathbf{K}}$ is in $-\overline{\boldsymbol{s}}$-Popov form, and the above description of $\mathrm{lm}_{-\overline{s}}(\mathbf{B})$ shows that $\overline{\mathbf{K}}$ and $\mathbf{S}$ have disjoint $-\overline{s}$-pivot indices, and that all their $-\overline{s}$-pivots are monic. Hence $\mathbf{B}$ is in $-\overline{s}$-unordered weak Popov form with monic $-\overline{s}$-pivots. It remains to show that each of these $-\overline{s}$-pivots has degree strictly larger than the other entries in the same column. Each row of the submatrix $\mathbf{S}$ of $\mathbf{B}$ has $-\overline{s}$-pivot entry x^{δ} and the other entries in the same column of $\mathbf{B}$ are either -1, which has degree $0 < \delta$, or are in a column of $\overline{\mathbf{K}}$ whose index is not in $\{\alpha_1 + \cdots + \alpha_j, 1 \leq j \leq m\}$, which has degree less than δ by definition of the column partial linearization. The jth row of the submatrix $\overline{\mathbf{K}}$ of $\mathbf{B}$ has $-\overline{s}$-pivot index $\alpha_1 + \cdots + \alpha_j$ and strictly positive $-\overline{s}$-pivot degree since $\mathrm{rdeg}_{-\overline{s}}(\overline{\mathbf{K}}) \geq \mathbf{0}$ and the entry of $-\overline{s}$ at index $\alpha_1 + \cdots + \alpha_j$ is $-\beta_j < 0$. This concludes the proof since the column $\mathbf{S}_{*,\alpha_1+\cdots+\alpha_j}$ has degree 0, and since $\overline{\mathbf{K}}$ is in $-\overline{s}$-Popov form. □

References

[1] Abdeljaoued, J., Malaschonok, G.I., 2001. Efficient algorithms for computing the characteristic polynomial in a domain. Journal of Pure and Applied Algebra 156, 127–145. doi:10.1016/S0022-4049(99)00158-9.

[2] Baur, W., Strassen, V., 1983. The complexity of partial derivatives. Theoretical computer science 22, 317–330. doi:10.1016/0304-3975(83)90110-X.

[3] Beckermann, B., Labahn, G., 1994. A uniform approach for the fast computation of matrix-type Padé approximants. SIAM J. Matrix Anal. Appl. 15, 804–823. doi:10.1137/S0895479892230031.

[4] Beckermann, B., Labahn, G., Villard, G., 1999. Shifted normal forms of polynomial matrices, in: ISSAC'99, ACM. pp. 189–196. doi:10.1145/309831.309929.

[5] Beckermann, B., Labahn, G., Villard, G., 2006. Normal forms for general polynomial matrices. J. Symbolic Comput. 41, 708–737. doi:10.1016/j.jsc.2006.02.001.

[6] Berkowitz, S.J., 1984. On computing the determinant in small parallel time using a small number of processors. Information Processing Letters 18, 147–150. doi:10.1016/0020-0190(84)90018-8.

[7] Bunch, J.R., Hopcroft, J.E., 1974. Triangular factorization and inversion by fast matrix multiplication. Mathematics of Computation 28, 231–236. doi:10.2307/2005828.

[8] Bürgisser, P., Clausen, M., Shokrollahi, A., 1997. Algebraic Complexity Theory. 1st ed., Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-662-03338-8.

[9] Cantor, D.G., Kaltofen, E., 1991. On fast multiplication of polynomials over arbitrary algebras. Acta Inform. 28, 693–701. doi:10.1007/BF01178683.

[10] Cook, S.A., 1966. On the minimum computation time of functions. Ph.D. thesis.

[11] Csanky, L., 1975. Fast parallel matrix inversion algorithms, in: 16th Annual Symposium on Foundations of Computer Science (sfcs 1975), pp. 11–12. doi:10.1109/SFCS.1975.14.

[12] Danilevskij, A.M., 1937. The numerical solution of the secular equation. Matem. Sbornik 44, 169–171. In Russian.

[13] Dumas, J.G., Pernet, C., Sultan, Z., 2017. Fast computation of the rank profile matrix and the generalized Bruhat decomposition. J. Symbolic Comput. 83, 187–210. doi:10.1016/j.jsc.2016.11.011.
[14] Dumas, J.G., Pernet, C., Wan, Z., 2005. Efficient computation of the characteristic polynomial, in: ISSAC'05, ACM. pp. 140–147. doi:10.1145/1073884.1073905.
[15] Dummit, D.S., Foote, R.M., 2004. Abstract Algebra. John Wiley & Sons.
[16] Faddeev, D., Sominskii, I., 1949. Collected Problems in Higher Algebra, Problem n°979.
[17] Forney, Jr., G.D., 1975. Minimal Bases of Rational Vector Spaces, with Applications to Multivariable Linear Systems. SIAM Journal on Control 13, 493–520. doi:10.1137/0313029.
[18] Frame, J., 1949. A simple recurrent formula for inverting a matrix (abstract). Bull. of Amer. Math. Soc. 55, 1045.
[19] Gathen, J.v.z., Gerhard, J., 2013. Modern Computer Algebra (third edition). Cambridge University Press. doi:10.1017/CBO9781139856065.
[20] Giesbrecht, M., 1995. Nearly optimal algorithms for canonical matrix forms. SIAM Journal on Computing 24, 948–969. doi:10.1137/S0097539793252687.
[21] Giorgi, P., Jeannerod, C.P., Villard, G., 2003. On the complexity of polynomial matrix computations, in: ISSAC'03, ACM. pp. 135–142. doi:10.1145/860854.860889.
[22] Giorgi, P., Neiger, V., 2018. Certification of minimal approximant bases, in: ISSAC'18, ACM. pp. 167–174. doi:10.1145/3208976.3208991.
[23] Gupta, S., Sarkar, S., Storjohann, A., Valeriote, J., 2012. Triangular x-basis decompositions and derandomization of linear algebra algorithms over $K[x]$. J. Symbolic Comput. 47, 422–453. doi:10.1016/j.jsc.2011.09.006.
[24] Harvey, D., Van Der Hoeven, J., Lecerf, G., 2017. Faster polynomial multiplication over finite fields. J. ACM 63. doi:10.1145/3005344.
[25] Ibarra, O.H., Moran, S., Hui, R., 1982. A generalization of the fast LUP matrix decomposition algorithm and applications. Journal of Algorithms 3, 45–56. doi:10.1016/0196-6774(82)90007-4.
[26] Jeannerod, C.P., Neiger, V., Schost, E., Villard, G., 2016. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts, in: ISSAC'16, ACM. pp. 295–302. doi:10.1145/2930889.2930928.
[27] Jeannerod, C.P., Neiger, V., Schost, E., Villard, G., 2017. Computing minimal interpolation bases. J. Symbolic Comput. 83, 272–314. doi:10.1016/j.jsc.2016.11.015.
[28] Jeannerod, C.P., Neiger, V., Villard, G., 2020. Fast computation of approximant bases in canonical form. J. Symbolic Comput. 98, 192–224. doi:10.1016/j.jsc.2019.07.011.
[29] Kailath, T., 1980. Linear Systems. Prentice-Hall.
[30] Kaltofen, E., Villard, G., 2005. On the complexity of computing determinants. Computational Complexity 13, 91–130. doi:10.1007/s00037-004-0185-3.
[31] Keller-Gehrig, W., 1985. Fast algorithms for the characteristic polynomial. Theoretical Computer Science 36, 309–317. doi:10.1016/0304-3975(85)90049-0.
[32] Labahn, G., Neiger, V., Zhou, W., 2017. Fast, deterministic computation of the Hermite normal form and determinant of a polynomial matrix. J. Complexity 42, 44–71. doi:10.1016/j.jco.2017.03.003.
[33] Le Gall, F., 2014. Powers of tensors and fast matrix multiplication, in: ISSAC'14, ACM. pp. 296–303. doi:10.1145/2608628.2608664.
[34] Le Verrier, U., 1840. Sur les variations séculaires des éléments elliptiques des sept plantètes principales. Journal des Mathématiques Pures et Appliquées 5, 220–254.
[35] Manthey, W., Helmke, U., 2007. Bruhat canonical form for linear systems. Linear Algebra Appl. 425, 261–282. doi:10.1016/j.laa.2007.01.022.
[36] Mulders, T., Storjohann, A., 2003. On lattice reduction for polynomial matrices. J. Symbolic Comput. 35, 377–401. doi:10.1016/S0747-7171(02)00139-6.
[37] Neiger, V., Rosenkilde, J., Solomatov, G., 2018. Computing Popov and Hermite Forms of Rectangular Polynomial Matrices, in: ISSAC'18, ACM. pp. 295–302. doi:10.1145/3208976.3208988.
[38] Neiger, V., Vu, T.X., 2017. Computing canonical bases of modules of univariate relations, in: ISSAC'17, ACM. pp. 357–364. doi:10.1145/3087604.3087656.
[39] Pernet, C., Storjohann, A., 2007. Faster Algorithms for the Characteristic Polynomial, in: ISSAC'07, ACM. pp. 307–314. doi:10.1145/1277548.1277590.
[40] Popov, V.M., 1972. Invariant description of linear, time-invariant controllable systems. SIAM Journal on Control 10, 252–264. doi:10.1137/0310020.
[41] Samuelson, P.A., 1942. A method of determining explicitly the coefficients of the characteristic equation. Annals of Mathematical Statistics 13, 424–429.
[42] Sarkar, S., Storjohann, A., 2011. Normalization of row reduced matrices, in: ISSAC'11, ACM. pp. 297–304. doi:10.1145/1993886.1993931.
[43] Souriau, J.M., 1948. Une méthode pour la décomposition spectrale et l'inversion des matrices. Comptes-Rendus de l'Académie des Sciences 227, 1010–1011.
[44] Storjohann, A., 2001. Deterministic computation of the frobenius form, in: Proceedings 42nd IEEE Symposium

on Foundations of Computer Science, pp. 368–377. doi:10.1109/SFCS.2001.959911.
[45] Storjohann, A., 2003. High-order lifting and integrality certification. J. Symbolic Comput. 36, 613–648. doi:10.1016/S0747-7171(03)00097-X.
[46] Storjohann, A., 2006. Notes on computing minimal approximant bases, in: Challenges in Symbolic Computation Software. URL: http://drops.dagstuhl.de/opus/volltexte/2006/776.
[47] Strassen, V., 1969. Gaussian elimination is not optimal. Numer. Math. 13, 354–356. doi:10.1007/BF02165411.
[48] The FFLAS-FFPACK Group, 2019. FFLAS-FFPACK: Finite Field Linear Algebra Subroutines / Package, version 2.4.3. http://github.com/linbox-team/fflas-ffpack.
[49] The LinBox Group, 2019. Linbox: Linear algebra over black-box matrices, version 1.6.3. https://github.com/linbox-team/linbox/.
[50] Toom, A.L., 1963. The complexity of a scheme of functional elements realizing the multiplication of integers. Soviet Mathematics Doklady 3, 714–716.
[51] Van Barel, M., Bultheel, A., 1992. A general module theoretic framework for vector M-Padé and matrix rational interpolation. Numer. Algorithms 3, 451–462. doi:10.1007/BF02141952.
[52] Wolovich, W.A., 1974. Linear Multivariable Systems. volume 11 of *Applied Mathematical Sciences*. Springer-Verlag New-York. doi:10.1007/978-1-4612-6392-0.
[53] Zhou, W., 2012. Fast Order Basis and Kernel Basis Computation and Related Problems. Ph.D. thesis. University of Waterloo. URL: http://hdl.handle.net/10012/7326.
[54] Zhou, W., Labahn, G., 2009. Efficient computation of order bases, in: ISSAC'09, ACM. pp. 375–382. doi:10.1145/1576702.1576753.
[55] Zhou, W., Labahn, G., 2012. Efficient algorithms for order basis computation. J. Symbolic Comput. 47, 793–819. doi:10.1016/j.jsc.2011.12.009.
[56] Zhou, W., Labahn, G., 2013. Computing column bases of polynomial matrices, in: ISSAC'13, ACM. pp. 379–386. doi:10.1145/2465506.2465947.
[57] Zhou, W., Labahn, G., 2014. Unimodular completion of polynomial matrices, in: ISSAC'14, ACM. pp. 413–420. doi:10.1145/2608628.2608640.
[58] Zhou, W., Labahn, G., Storjohann, A., 2012. Computing minimal nullspace bases, in: ISSAC'12, ACM. pp. 366–373. doi:10.1145/2442829.2442881.