Asymptotically fast polynomial matrix algorithms for multivariable systems

Claude-Pierre Jeannerod, Gilles Villard

Asymptotically Fast Polynomial Matrix Algorithms for Multivariable Systems

Claude-Pierre Jeannerod,
Gilles Villard

September 2005

Asymptotically Fast Polynomial Matrix Algorithms for Multivariable Systems

Claude-Pierre Jeannerod,
Gilles Villard

September 2005

Abstract

We present the asymptotically fastest known algorithms for some basic problems on univariate polynomial matrices: *rank, nullspace, determinant, generic inverse, reduced form* [8, 9, 16, 17]. We show that they essentially can be reduced to two computer algebra techniques, *minimal basis computations* and *matrix fraction expansion/reconstruction*, and to polynomial matrix multiplication. Such reductions eventually imply that all these problems can be solved in about the same amount of time as polynomial matrix multiplication (up to logarithmic factors and the size of the output).

Keywords: Linear algebra, polynomial matrix, matrix rank, matrix determinant, nullspace basis, matrix inversion, matrix reduced form.

Résumé

Cet article présente les algorithmes actuellement les plus rapides asymptotiquement pour effectuer certaines opérations de base sur les matrices polynomiales : *calcul du rang, d'une base du noyau, du déterminant, de l'inverse générique, d'une forme réduite* [8, 9, 16, 17]. On montre que ces problèmes se ramènent essentiellement à deux techniques, *le calcul de bases minimales* et *le développement et la reconstruction de fractions de matrices polynomiales*, ainsi qu'au produit de matrices polynomiales. Ces réductions impliquent pour ces problèmes l'existence d'algorithmes de résolution dont le coût est de l'ordre de celui du produit de matrices polynomiales (à la taille de la sortie et aux facteurs logarithmiques près).

Mots-clés: Algèbre linéaire, matrice polynomiale, rang, déterminant, base du noyau, inversion, forme réduite.

1 Introduction

We aim at drawing attention to today's asymptotically fastest known algorithms for computing with polynomial matrices. In particular, we shall focus on the following problems: compute the *rank*, a right or *left nullspace*, the *determinant*, the *inverse* and a column- or *row-reduced form* of a given polynomial matrix. Polynomial matrices are quite common in the analysis of multivariable linear systems and Kailath's treatise *Linear Systems* [10] is a good illustration of this.

Recently, algorithms have been designed [8, 9, 16, 17] that allow to compute solutions to these problems in essentially the same amount of time as when multiplying two polynomial matrices together. More precisely, given a field K—for example the complex numbers, the rationals or a finite field—and given a polynomial matrix $A \in \mathsf{K}[x]^{n \times n}$ whose entries have degree in x bounded by d, these algorithms allow to compute rank A, ker A, det A and to row-reduce A in $O\tilde{\ }(n^\omega d)$ operations in K, and to compute A^{-1} when A is generic in $O\tilde{\ }(n^3 d)$ operations in K. Here, $O\tilde{\ }(n^\omega d)$ is the best known asymptotic bound for multiplying two matrices in $\mathsf{K}[x]^{n \times n}$ of degree d [5, 3], where $2 \leq \omega < 2.376$ is the exponent of matrix multiplication over K [4, Chapter 15]. Using schoolbook matrix multiplication, we have $\omega = 3$ and the bound $O\tilde{\ }(n^\omega d)$ becomes $O\tilde{\ }(n^3 d)$. Furthermore, the soft-O notation $O\tilde{\ }$ simply indicates some missing logarithmic factors of the form $\alpha(\log n)^\beta(\log d)^\gamma$ for three positive real numbers α, β, γ. By achieving the complexity estimate $O\tilde{\ }(n^\omega d)$, these algorithms improve upon all the complexity estimates that were known previously.

In this paper, evidence is given that the key tools for such improvements are:

- Minimal bases of $\mathsf{K}[x]$-modules;

- Expansion/reconstruction of polynomial matrix fractions.

The former has the same flavour as in [6] while for the fractions we heavily rely on the concepts in [10, Chapter 6]. Two kinds of minimal bases, namely *approximant bases* and *nullspace bases*, are studied in Section 2. There we will see that such bases are small enough to be computed fast, that is, in $O\tilde{\ }(n^\omega d)$ operations in K. Polynomial matrix fractions are matrices $F \in \mathsf{K}(x)^{n \times n}$, where $\mathsf{K}(x)$ is the field of rational functions over K. By expansion of F, we thus mean a power series expansion $F = \sum_{i=0}^\infty F_i x^i \in \mathsf{K}[[x]]^{n \times n}$ and by reconstruction of F we mean a left or right quotient of polynomial matrices like $F = A^{-1}B$ or $F = BA^{-1}$. It turns out that all we need is truncated expansions and reconstructed quotients of rather low degree, both of which can be computed fast as seen in Section 3. The key idea here is that an *approximant* of sufficiently high order—with respect to the input problem—may lead to an *exact* solution over $\mathsf{K}[x]$. This is well-known in computer algebra, at least for scalar rational functions [7, §5.7], but as far we know the extension to the matrix case is more recent [8, 9, 16, 17].

Minimal bases and matrix fractions are interesting not only because they can be computed fast, but also—and, perhaps, mainly—because computing a minimal basis and expanding/reconstructing a matrix fraction are problems to which we can reduce all other problems like *rank*, *left nullspace*, *determinant*, *generic inverse* and *row-reduced form*. The goal of Section 4 is precisely to show this: there the above problems are thus seen as applications of the techniques studied in Sections 2 and 3.

If we assume given an $O\tilde{\ }(n^\omega d)$ algorithm for multiplying two n by n polynomial matrices of degree d, combining the reductions of Section 4 with the cost estimates of Sections 2

and 3 then yields $O\tilde{}(n^\omega d)$ solutions to all our problems under consideration. Of course, we could have introduced a cost function $\mathsf{MM}(n, d)$ for polynomial matrix multiplication and derived more precise complexity estimates for each of the problems, in terms of (functions of) $\mathsf{MM}(n, d)$. However, we prefer for this paper to stick to the more readable $O\tilde{}(n^\omega d)$ bound, which already gives a good sense of the link with polynomial matrix multiplication.

A first task remaining would be to relax the regularity assumptions made for *inversion* (the input should be generic and of dimensions a power of two, see Section 4.1) and for *row-reduction* (the input should be non-singular, see Section 4.3). But even these "generic" situations are enough for our purpose here of showing how to rely on minimal bases and matrix fraction expansions/reconstructions.

Also, recently, other problems on polynomial matrices than those treated in this paper have been shown to have about the same complexity as polynomial matrix multiplication. An example is the problem of computing the *Smith normal form* and thus also the determinant, whose solution in [16] gives us Theorem 3.1. However—and this is the second task remaining—, the list of problems that can be solved in about the same number of operations as for polynomial matrix multiplication still has to be augmented. The question is particularly interesting for the problem of computing the *characteristic polynomial* and the *Frobenius normal form*, for which the best known solutions [11, 12] have cost $O\tilde{}(n^{2.7}d)$ still greater than $O\tilde{}(n^\omega d)$.

Notation and basic reminders. Here and hereafter log denotes the logarithm in base two and I_n the n by n identity matrix. For a matrix A over $\mathsf{K}[x]$, we denote its value at $x = 0$ by $A(0)$. For $d \in \mathbb{N}$ and a matrix F over $\mathsf{K}[[x]]$, $F \equiv 0 \mod x^d$ means that each entry of F is a multiple of x^d, and $F \mod x^d$ means that we truncate F into a polynomial matrix where only powers in x strictly less than d appear. By *size* of a polynomial matrix over $\mathsf{K}[x]$ we mean the number of elements of K that are necessary to represent it. For example, $M \in \mathsf{K}[x]^{n \times m}$ of degree d has size at most $nm(d + 1) = O(nmd)$. A polynomial matrix is said to be *non-singular* when it is square and when its determinant is a non identically zero polynomial. Two matrices $A, R \in \mathsf{K}[x]^{n \times n}$ are *unimodularly left equivalent* when there exists $U \in \mathsf{K}[x]^{n \times n}$ such that $\det U$ is a non-zero constant—that is, U is *unimodular*— and when $UA = R$.

2 Minimal approximant bases and minimal nullspace bases

Our solutions for solving a class of polynomial matrix problems in about the same number of operations in K as for multiplying two polynomial matrices will fundamentally rely on computing minimal bases of $\mathsf{K}[x]$-modules. The target complexity estimate $O\tilde{}(n^\omega d)$ is reached since the bases we use are small, with size $O(n^2 d)$ is most cases, and may be computed fast (see Theorem 2.2 below).

Definition 2.1 *Let $\mathcal{M}$ be a $\mathsf{K}[x]$-submodule of $\mathsf{K}[x]^n$ of dimension $\mathcal{D}$. A basis $N_1, \ldots, N_{\mathcal{D}} \in \mathsf{K}[x]^n$ of $\mathcal{M}$ with degrees $\delta_1 \leq \cdots \leq \delta_{\mathcal{D}}$ is called a* minimal basis *if any other basis of $\mathcal{M}$ with degrees $d_1 \leq \cdots \leq d_{\mathcal{D}}$ satisfies $d_i \geq \delta_i$ for $1 \leq i \leq \mathcal{D}$. The degrees δ_i are called the* minimal indices *of $\mathcal{M}$.*

In applications to multivariable systems, this definition follows the study of minimal polynomial bases of vector spaces in [6]. The two important examples of such bases that we use

in this paper are minimal approximant bases and minimal nullspace bases. The approximant bases are defined from a power series matrix F over $\mathsf{K}[[x]]$, the nullspace bases are computed as special approximant bases from a polynomial matrix $F = A$ over $\mathsf{K}[x]$.

2.1 Minimal approximant bases

Given a formal power series $F \in \mathsf{K}[[x]]^{n \times n}$ and an order $d \in \mathbb{N}$, we take for $\mathcal{M}$ the set of all approximants for F of order d:

$$\mathcal{M} = \{v \in \mathsf{K}[x]^{1 \times n} \ : \ vF \equiv 0 \mod x^d\}.$$

The minimal bases of $\mathcal{M}$ are called *minimal approximant bases for F of order d*. Since $\mathcal{M}$ has dimension n, such bases form non-singular $n \times n$ polynomial matrices. These polynomial matrices further have degree up to d and their size is thus of the order of $n^2 d$.

Theorem 2.2 [8]. *Let $F \in \mathsf{K}[[x]]^{n \times n}$ and $d \in \mathbb{N}$. A minimal approximant basis for F of order d can be computed in $O\tilde{\ }(n^\omega d)$ operations in K.*

Our notion of minimal approximant bases is directly inspired by [1] with some adaptations for fully reflecting the polynomial matrix point of view. The cost estimate of Theorem 2.2 is a matrix polynomial generalization of the recursive Knuth/Schönhage half-gcd algorithm for scalar polynomials [13, 15] (see also [7, §11.1]), that takes into account fast polynomial matrix multiplication.

For a matrix A over $\mathsf{K}[x]$, we denote by d_i its ith *row degree*, that is, the highest degree of all the entries of the ith row of A. The *row leading matrix* of A is the constant matrix whose ith row consists of the coefficients of x^{d_i} in the ith row of A. We recall from [10, §6.3.2] that a full row rank A is *row-reduced* when its row leading matrix also has full rank. As a consequence of their minimality, minimal approximant bases have the following properties, which will be used in Section 2.2 when specializing approximants for power series matrices to approximants for polynomial matrices.

Property 2.3 *Let N be a minimal approximant basis for F of order d. Then,*

  I. *N is row-reduced;*

  II. *If $v \in \mathcal{M}$ has degree at most d, then there is a unique $u \in \mathsf{K}[x]^{1 \times n}$ such that $v = uN$. Furthermore, N has at least one row of degree at most d.*

Property I above is a consequence of the minimality of the basis [10, Theorem 6.5-10]. Property II is the fact that the rows of N form a basis, together with the predictable degree property [10, Theorem 6.3-13].

2.2 Minimal nullspace bases

Given a polynomial matrix $A \in \mathsf{K}[x]^{n \times n}$ of rank r, we now take

$$\mathcal{M} = \{v \in \mathsf{K}[x]^{1 \times n} \ : \ vA = 0\}.$$

This is a $\mathsf{K}[x]$-submodule of $\mathsf{K}[x]^n$ of dimension $n - r$. Its bases are called *minimal nullspace bases for A* and form full rank $(n - r) \times n$ polynomial matrices. The minimal indices $\delta_1 \leq$

$\cdots \leq \delta_{n-r}$ (see Definition 2.1) are called the (left) *Kronecker indices* of A [10, §6.5.4]. For any given degree threshold δ, we further define

$$\kappa = \max\{1 \leq i \leq n - r \ : \ \delta_i \leq \delta\}.$$

A corresponding family of κ linearly independent vectors of degrees $\delta_1, \cdots, \delta_\kappa$ is a family of *minimal nullspace vectors* of degree at most δ. The theorem below says that if $F = A$ is a polynomial matrix then any minimal approximant basis for A of sufficiently high order actually contains a family of minimal nullspace vectors for A.

Theorem 2.4 *Let $A \in \mathsf{K}[x]^{n \times n}$ be of degree d. Let N be a minimal approximant basis for A of order $\delta + d + 1$. Then exactly κ rows of N have degree at most δ; these rows are in the (left) nullspace of A and their degrees are the Kronecker indices $\delta_1, \ldots, \delta_\kappa$.*

Proof. A row v of N of degree bounded by δ satisfies $vA \equiv 0 \mod x^{\delta+d+1}$, and using $\deg vA \leq \delta + d$, $vA = 0$. Let k be the number of such $v's$ in N, from the definition of κ and since N is non-singular, $k \leq \kappa$. We now verify that $k \geq \kappa$. We consider κ linearly independent vectors v_i of degrees δ_i in the nullspace of A. From Property 2.3 we have $v_1 = u_1 M$ and deduce that one row of N has degree bounded by δ_1. Now, if N has $i - 1$ rows of degrees bounded by $\delta_1, \ldots, \delta_{i-1}$, then the same reasoning with v_i as for v_1 shows that N has a row of degree bounded by δ_i, linearly independent with respect to the first $i-1$ chosen ones. It follows that $k \geq \kappa$ rows of N have degrees bounded by $\delta_1, \ldots, \delta_\kappa$, and are in the nullspace of A. Hence $k = \kappa$, and we conclude using Definition 2.1 and the minimality of the δ_i's. ∎

For some applications, a *shifted degree* may be introduced (see [2] and the references therein), and some aspects of Theorem 2.4 may be generalized accordingly (see [2, Theorem 4.2] or [17, Lemma 6.3]).

Notice that if the Kronecker indices of A are all bounded by d then an *entire* minimal nullspace basis for A can already be computed fast: by Theorem 2.4, it suffices to compute a minimal approximant basis for A of order $2d + 1$ and, by Theorem 2.2, this computation can be done in time $\tilde{O}(n^\omega d)$.

However, in the general case of unbalanced degrees, computing a nullspace basis fast is much less immediate and the method we shall give in Section 4.4 relies on the complexity result given below in Theorem 2.5. The cost given here is the one of a randomized algorithm of the Las Vegas kind—always correct, probably fast. The algorithm outputs correct minimal vectors in time $\tilde{O}(n^\omega d)$ with good probability, say greater than $1/2$, otherwise returns *failure* (a correct result will be obtained after repetition).

Theorem 2.5 [17]. *Let $A \in \mathsf{K}[x]^{(n+m) \times n}$ with $m \leq n$ be of full column rank and degree bounded by d. If $\delta \in \mathbb{N}$ satisfies*

$$\delta m = O(nd), \tag{1}$$

then a family of minimal nullspace vectors of degree at most δ can be computed by a randomized Las Vegas (certified) algorithm in $\tilde{O}(n^\omega d)$ operations in K.

Note that the cost estimate $\tilde{O}(n^\omega d)$ relies on the compromise (1) between the minimal nullspace vector degree bound δ and the row dimension of matrix A. For example, when $m = 1$ one can compute a nullspace vector of degree as large as $O(nd)$, whereas when $m = n$ one may compute up to n nullspace vectors of degree $O(d)$. Random values are introduced

essentially through a random compression matrix $P \in \mathsf{K}[x]^{n \times m}$ that allows to compute minimal vectors more efficiently using the matrix $AP \in \mathsf{K}[x]^{(n+m) \times m}$ rather than directly from $A \in \mathsf{K}[x]^{(n+m) \times m}$ (see [17, Proposition 5.4]).

3   Matrix fraction expansion and reconstruction

Matrix fraction expansion and *reconstruction* will be key tools especially for the row reduction and the nullspace problems. Fraction reconstruction is a useful tool in computer algebra (*e.g.* see [7, §5.7] for scalar polynomials), that is directly connected to *coprime factorization* (see below, and [10, Chapter 6] or [14] and the references therein).

For a polynomial matrix A that is non-singular at $x = 0$ and a polynomial matrix B, the techniques of [16, Proposition 17] reduce the computation of parts of the power series expansion

$$A^{-1}B = \sum_{i=0}^{\infty} F_i x^i$$

to polynomial matrix multiplication. By parts of the expansion, we mean a given number of consecutive matrix coefficients F_i. This is summarized in the following theorem.

Theorem 3.1 [16]. *Let $A \in \mathsf{K}[x]^{n \times n}$ with $A(0)$ non-singular, and $B \in \mathsf{K}[x]^{n \times m}$. Assume that A and B have degree bounded by d and let $h \in \mathbb{N}$ be such that $h = O(nd)$. If $\delta \in \mathbb{N}$ satisfies*

$$\delta m = O(nd), \tag{2}$$

then the δ coefficients $F_h, F_{h+1}, \ldots, F_{h+\delta-1} \in \mathsf{K}^{n \times m}$ of the expansion of $A^{-1}B$ at $x = 0$ can be computed in $\tilde{O}(n^\omega d)$ operations in K.

Similarly to Theorem 2.5, the cost estimate $\tilde{O}(n^\omega d)$ relies on the compromise (2) between approximation order δ and the column dimension of matrix B. For instance, for a vector $B = b \in \mathsf{K}[x]^{n \times 1}$ and $h = 0$, one can expand $A^{-1}b$ up to order $O(nd)$, whereas with $B = I_n$ and $h = 0$, one gets the expansion of A^{-1} up to order $O(d)$. In Section 4.3, we shall use this result with $B = I_n$ and $h = (n-1)d + 1$ in order to get a high-order slice of length $O(nd)$ of the expansion of A^{-1}.

Notice also that the regularity assumption $\det A(0) \neq 0$ in Theorem 3.1 is not restrictive. Indeed, it can be satisfied with high probability using random shifts, thus yielding randomized algorithms for any $A(0)$. Typically, with a randomly chosen $x_0 \in \mathsf{K}$, we shift x in the input like $x \leftarrow x + x_0$ to get a regular input at zero and, at the end of the computation, we shift x back like $x \leftarrow x - x_0$ to recover the result (see [16, 8, 17]).

A rational matrix $H \in \mathsf{K}(x)^{n \times m}$ is *strictly proper* if $\lim_{x \to \infty} H(x) = 0 \in \mathsf{K}^{n \times m}$. In most applications, difficulties arise when $A^{-1} \in \mathsf{K}(x)^{n \times n}$ is *not* strictly proper. However, one can define another fraction that is always strictly proper and shares some invariants with A^{-1}. Before seeing this, we first need to recall some facts about *greatest common divisors* of two polynomial matrices.

Definition 3.2 *A (left) matrix gcd of $A \in \mathsf{K}[x]^{n \times n}$ and $B \in \mathsf{K}[x]^{n \times m}$ is any full column rank polynomial matrix G such that $[G \quad 0]U = [A \quad B]$ with U unimodular over $\mathsf{K}[x]$.*

Definition 3.2 is for instance from [10, Lemma 6.3-3]. If $[A \ B]$ has full row rank then all the gcd's of A and B are non-singular and equivalent with respect to multiplication on the right by any unimodular matrix in $\mathsf{K}[x]^{n \times n}$ (see [10, Lemma 6.3-4]). A non-singular $A \in \mathsf{K}[x]^{n \times n}$ is said to be (left) *coprime* with $B \in \mathsf{K}[x]^{n \times m}$ if any gcd of A and B is unimodular (*the* gcd may be chosen as being the identity matrix I_n). Similar definitions hold for rights gcd's and right coprimeness.

Theorem 3.3 [8]. *Let $A \in \mathsf{K}[x]^{n \times n}$ of degree bounded by d, with $A(0)$ non-singular. For $A^{-1} = \sum_{i=0}^{\infty} F_i x^i$ and $h > (n-1)d$, let $H \in \mathsf{K}(x)^{n \times n}$ be given by $H = \sum_{i=0}^{\infty} F_{h+i} x^i$. Then $H = A^{-1}(AH) = (HA)A^{-1}$ is strictly proper, and AH and HA are polynomial matrices that are respectively left and right coprime with A.*

Proof. Let $B = AH$. By definition of H we have $I_n = A(A^{-1} \mod x^h) + x^h B$ which in [16] is (17) on the left with B and T respectively set to I_n and A. It follows that B is a polynomial matrix. On the other hand, $H = A^{-1}B$ is strictly proper because $A^{-1}B = x^{-h}A^{-1} - x^{-h}(A^{-1} \mod x^h)$ and $h > (n-1)d \geq \deg A^*$ where A^* is the adjoint matrix of A. For establishing coprimeness we use

$$[A \ \ x^h B] \begin{bmatrix} I_n & (A^{-1} \mod x^h) \\ 0 & I_n \end{bmatrix} \begin{bmatrix} 0 & I_n \\ I_n & -A \end{bmatrix} = [I_n \ \ 0], \tag{3}$$

and the fact that if G is a left gcd of A and B it satisfies

$$[G \ \ 0]U = [A \ \ B] \tag{4}$$

with U unimodular. Identities (3) and (4) give that their exists a polynomial matrix V such that $[G \ \ 0]V = [I_n \ \ 0]$, hence a polynomial matrix W such that $GW = I_n$. Since G is a polynomial matrix this implies that G is unimodular, and A and B are left coprime. With $B = HA$, one could show similary right coprimeness. ∎

For our application in Section 4.3, we will need only the first, say δ, coefficients of the expansion of H as in Theorem 3.3. These coefficients thus correspond to a slice of order h and length δ of the expansion of A^{-1} and, to recover them, we shall use Theorem 3.1 with $B = I_n$.

Matrix power series expansion will be used in conjunction with matrix *(irreducible) fraction reconstruction* or, equivalently, *(coprime) factorization*. We show below that minimal approximant bases are appropriate tools for solving these problems.

Definition 3.4 *A (left)* factorization *of degree δ of a rational matrix $H \in \mathsf{K}(x)^{n \times n}$ is a representation $H = V^{-1}U$ with U and V two polynomial matrices of degree bounded by δ. This factorization is said to be* coprime *when U and V are (left) coprime.*

A similar definition holds on the right. Hence, given $H \in \mathsf{K}(x)^{n \times n}$, the reconstruction or factorization problem is to recover U and V over $\mathsf{K}[x]$ such that $V^{-1}U = H$. If H is defined at $x = 0$ and given by its formal expansion $F \in \mathsf{K}[[x]]^{n \times n}$, this problem reduces to computing a suitable $[U \ V] \in \mathsf{K}[x]^{n \times 2n}$ such that

$$\begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} -I_n \\ F \end{bmatrix} = 0.$$

Theorem 3.5 *Let $H \in \mathsf{K}(x)^{n \times n}$ be strictly proper, with expansion $F \in \mathsf{K}[[x]]^{n \times n}$ at $x = 0$. Assume that H admits a right factorization of degree δ_R and a left factorization of degree δ_L. Let $N \in \mathsf{K}[x]^{2n \times 2n}$ be a minimal approximant basis for $[-I_n \quad F^T]^T$ of order $\delta_L + \delta_R + 1$. Then exactly n rows of N have degree bounded by δ_L; these rows form a matrix $[U \quad V] \in \mathsf{K}[x]^{n \times 2n}$ such that $V^{-1}U$ is a left coprime factorization of H, with V row-reduced.*

Proof. Let BA^{-1} be a right factorization of H of degree δ_R and $T^{-1}S$ be a left factorization of H degree δ_L. Since $[-I_n \quad F^T]^T A = [-A^T \quad B^T]^T$, N is also a minimal approximant basis of the latter matrix whose rank is n. Using $[S \quad T][-A^T \quad B^T]^T = 0$, with the threshold $\delta = \delta_L$ we have $\kappa = n$. (See before Theorem 2.4 for a definition of κ.) Hence, applying Theorem 2.4 to $[-A^T \quad B^T]^T$ (augmented on the right with n zero columns) with $\delta = \delta_L$ and $d = \delta_R$, we know that exactly n rows of N have degree bounded by δ_L and are in the nullspace of $[-A^T \quad B^T]^T$. We denote the corresponding matrix by $[U \quad V]$. The matrix V is non-singular, for otherwise there would be a non-zero vector v such that $vV = 0$. This would imply $vVB = vUA = 0$, hence either $vU = 0$ or $wA = 0$ for $w = vU \neq 0$, and would contradict either that $\mathrm{rank}[U \quad V] = n$ or that A is non-singular. Therefore, $V^{-1}U$ is a left factorization of H.

This factorization must further be left coprime. Indeed, non-coprimeness would imply that U and V have a non-trivial left gcd, that is, there exists a polynomial matrix G such that $U = GU'$, $V = GU'$ and $\deg(\det G) > 0$. Then $[GU' \quad GV']$ would be a submatrix of the minimal approximant basis, which would contradict its irreducibility in [10, Theorem 6.5-10] by considering a zero of $\det G$. In addition, the fact that $[U \quad V]$ as a submatrix of N is row-reduced (see Property 2.3), implies that V is row-reduced. Indeed, since $H = V^{-1}U$ is strictly proper, the row degrees of U are strictly smaller than those of V [10, Lemma 6.3-10], and the row leading matrix of $[U \quad V]$ has the form $[0 \quad L]$ where L is the row leading matrix of V, which is then non-singular. ∎

As an immediate consequence of Theorem 3.5 and Theorem 2.2, coprime factorizations can be computed fast when the input matrix fractions admit left and right factorizations of degree $O(d)$. This corollary, given below, will be applied in Section 4.3 to the particular matrix fraction of Theorem 3.3.

Corollary 3.6 *Let $H \in \mathsf{K}(x)^{n \times n}$ be as in Theorem 3.5 with $\delta_L = O(d)$ and $\delta_R = O(d)$. Given the first $\delta_L + \delta_R + 1$ coefficients of the expansion of H at $x = 0$, one can compute a left coprime factorization of H in $O^\sim(n^\omega d)$ operations in K.*

4 Applications

In this section, we show how the techniques presented in Sections 2 and 3 can be used to solve the following problems asymptotically fast:

- $\mathsf{Inv}_{n,d}$: given a non-singular $A \in \mathsf{K}[x]^{n \times n}$ of degree d, compute A^{-1}.

- $\mathsf{Det}_{n,d}$: given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, compute $\det A$.

- $\mathsf{RowRed}_{n,d}$: given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, compute a row-reduced form of A.

- $\mathsf{Nullspace}_{n,d}$: given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, compute the rank r of A and a full rank $N \in \mathsf{K}[x]^{(n-r) \times n}$ such that $NA = 0$.

- $\mathsf{Factor}_{n,d}$: given a right factorization of degree d of $H \in \mathsf{K}(x)^{n \times n}$, compute a left factorization of H.

Our approach here is to reduce each of the above five problems to (collections of) the problems below, for which $\tilde{O}(n^\omega d)$ solutions are known:

- $\mathsf{MatMul}_{n,d}$: given $A, B \in \mathsf{K}[x]^{n \times n}$ of degree d, compute the product AB.

  $\hookrightarrow$ *for solutions in time $\tilde{O}(n^\omega d)$ see [5], [3].*

- $\mathsf{PartialNullSpace}_{m,\delta}$: given $\delta = O(nd/m)$ with n, d fixed, and given $A \in \mathsf{K}[x]^{(n+m) \times n}$ of degree d, compute the minimal nullspace vectors of A of degree at most δ.

  $\hookrightarrow$ *solved in time $\tilde{O}(n^\omega d)$ by Theorem 2.5.*

- $\mathsf{MatFracExp}_{m,\delta}$: given $\delta = O(nd/m)$ with n, d, h fixed such that $h = O(nd)$, and given $A \in \mathsf{K}[x]^{n \times n}, B \in \mathsf{K}[x]^{n \times m}$ of degree d with $A(0)$ non-singular, compute the δ coefficients $F_h, F_{h+1}, \ldots, F_{h+\delta-1}$ of the expansion of $A^{-1}B$ at $x = 0$.

  $\hookrightarrow$ *solved in time $\tilde{O}(n^\omega d)$ by Theorem 3.1.*

- $\mathsf{MatFracRec}_{n,d}$: given $\delta_L, \delta_R = O(d)$ and the first $\delta_L + \delta_R + 1$ coefficients of the expansion at $x = 0$ of $H \in \mathsf{K}(x)^{n \times n}$ as in Theorem 3.5, compute a left coprime factorization of H with row-reduced denominator.

  $\hookrightarrow$ *solved in time $\tilde{O}(n^\omega d)$ by Corollary 3.6.*

Assuming that n is a power of two and given a problem $\mathsf{P}_{n,d}$ or $\mathsf{P}_{m,\delta}$ such as any of those just introduced, we define the collections of problems we shall rely on as

$$\mathsf{P}^*_{n,d} := \big\{ \text{solve } O(2^i) \text{ problems } \mathsf{P}_{n/2^i, 2^i d} \big\}_{0 \leq i < \log n}. \tag{5}$$

Such collections can be solved at about the same cost as polynomial matrix multiplication, as shown below. Here subscripts n, d and m, δ should be added to P and P^* depending on the underlying problem.

Lemma 4.1 *For all $\mathsf{P} \in \{\mathsf{MatMul}, \mathsf{PartialNullSpace}, \mathsf{MatFracExp}, \mathsf{MatFracRec}\}$, one can solve P^* in $\tilde{O}(n^\omega d)$ operations in K.*

Proof. This an immediate consequence of (5) and of the bound $\tilde{O}(n^\omega d)$ on the cost of each of these four problems. ∎

4.1 Polynomial matrix inversion ($\mathsf{Inv}_{n,d}$)

Given $A \in \mathsf{K}[x]^{n \times n}$ non-singular of degree d, the problem is to compute $A^{-1} \in \mathsf{K}(x)^{n \times n}$.

Assuming that A is generic and that n is a power of two, we recall from [9] how $\mathsf{Inv}_{n,d}$ reduces to $\mathsf{PartialNullSpace}^*_{n,d}$ plus some polynomial matrix multiplications. The algorithm in [9, p.75] essentially consists in computing in $\log n$ steps a non-singular matrix $U \in \mathsf{K}[x]^{n \times n}$ and a diagonal matrix $B \in \mathsf{K}[x]^{n \times n}$ such that

$$UA = B. \tag{6}$$

The inverse of A is then recovered as $A^{-1} = B^{-1}U$. The first step is as follows. Let $A = [A_L \ A_R]$ where $A_L, A_R \in \mathsf{K}[x]^{n \times n/2}$ and let $\underline{N}, \overline{N} \in \mathsf{K}[x]^{n/2 \times n}$ be minimal nullspace bases for, respectively, A_L, A_R. This gives the first block-elimination step towards the diagonalization of A:

$$A = \begin{bmatrix} A_L & A_R \end{bmatrix} \quad \rightarrow \quad NA = \begin{bmatrix} \overline{N} \\ \underline{N} \end{bmatrix} \begin{bmatrix} A_L & A_R \end{bmatrix} = \begin{bmatrix} \overline{N}A_L & \\ & \underline{N}A_R \end{bmatrix}. \tag{7}$$

When A is generic of degree d, it turns out that all the minimal indices of both $\underline{N}$ and $\overline{N}$ are equal to d [9, Fact 1] and that $\overline{N}A_L$ and $\underline{N}A_R$ are $n/2 \times n/2$ polynomial matrices of degree exactly $2d$ on which we iterate.

We show in [9] that the property "dimension $\times$ degree $= nd$" generically carries from one iteration to the other: at step i, starting from 2^{i-1} blocks of dimensions $(n/2^{i-1}) \times (n/2^{i-1})$ and degree $2^{i-1}d$, we compute 2^{i-1} pairs $(\underline{N}_i^{(j)}, \overline{N}_i^{(j)})$ of minimal nullspace bases of dimensions $(n/2^i) \times (n/2^{i-1})$ and whose minimal indices are all equal to $2^{i-1}d$. Let $(U, B) = (I_n, A)$ before the first step. Step i also requires to update the matrix transform as $U \leftarrow \operatorname{diag}[N_i^{(j)}]_j \times U$ and the right hand side as $B \leftarrow \operatorname{diag}(N_i^{(j)})_j \times B$. Because of the special block-structure of the polynomial matrices involved, it can be shown that these updates reduce to solving $O(2^{2i})$ problems $\mathsf{MatMul}_{n/2^{i-1}, 2^{i-1}d}$.

Overall, the $\log n$ block-diagonalization steps thus reduce to $\mathsf{PartialNullSpace}_{n,d}^*$ and to

$$\left\{ \text{solve } O(2^{2i}) \text{ problems } \mathsf{MatMul}_{n/2^i, 2^i d} \right\}_{0 \le i < \log n}. \tag{8}$$

By Lemma 4.1 and (8), we therefore obtain a solution to $\mathsf{Inv}_{n,d}$ in $O^\sim(n^3 d)$ operations in K.

Since by Cramer's rule each entry of A^{-1} has the form $p/(\det A)$ where $p \in \mathsf{K}[x]$ may have degree at large as $(n-1)d$, the size of A^{-1} is of the order of $n^3 d$. The above inversion algorithm, defined for A generic and n a power of two, is therefore nearly optimal.

4.2 Determinant computation ($\mathsf{Det}_{n,d}$)

Given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, the problem is to compute $\det A \in \mathsf{K}[x]$.

We assume here that A is generic with n is a power of two, and we use the inversion algorithm of Section 4.1. It has been shown in [8] that the diagonal entries of the diagonal matrix B in (6) are constant multiples of $\det A$. Since $\det A(0)$ is generically non-zero, we have

$$\det A = \frac{\det A(0)}{b_{i,i}(0)} b_{i,i} \quad \text{for all } 1 \le i \le n.$$

The problem $\mathsf{Det}_{n,d}$ thus reduces essentially to computing the determinant of the constant matrix $A(0)$ and to the computation of, say, $b_{1,1}$. It is well-known that over K computing the determinant reduces to matrix multiplication [4, Section 16.4] (that is, $\mathsf{Det}_{n,0}$ reduces to $\mathsf{MatMul}_{n,0}$ using our notations). Concerning $b_{1,1}$, we perform $\log n$ steps as for inversion but, since $b_{1,1}$ is the upper-left corner of B, we use instead of (7) the simpler step

$$A = \begin{bmatrix} A_L & A_R \end{bmatrix} \quad \rightarrow \quad \overline{N}A_L. \tag{9}$$

As in (7), $\overline{N}$ is a minimal nullspace basis for A_R. Step i now consists in computing a single minimal nullspace basis of dimensions $(n/2^i) \times (n/2^{i-1})$ and minimal indices $2^{i-1}d$, and then in multiplying this basis with the left half of an $n/2^i$ by $n/2^i$ block of degree $2^{i-1}d$, as in (9).

Hence, computing $b_{1,1}$ by performing these $\log n$ steps reduces to solving $\mathsf{PartialNullSpace}^*_{n,d}$ and $\mathsf{MatMul}^*_{n,d}$. By Lemma 4.1, this gives a solution to $\mathsf{Det}_{n,d}$ in $O\tilde{\ }(n^\omega d)$ operations in K.

Notice that when A is not generic or when n is not a power of two, a Las Vegas $O\tilde{\ }(n^\omega d)$ solution to $\mathsf{Det}_{n,d}$ can be obtained using the Smith normal form algorithm in [16].

4.3   Row reduction ($\mathsf{RowRed}_{n,d}$)

Given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, the problem is to compute $R \in \mathsf{K}[x]^{n \times n}$ that is row-reduced and unimodularly left equivalent to A.

We assume here that $A(0)$ is non-singular. Recall from Section 2.1 and [10, §6.3.2] that $R = A$ is a row-reduced form of A when R is row-reduced and $R = UA$ for some unimodular polynomial matrix U. The solution in [8] works by expansion/reconstruction of the matrix fraction H as in Theorem 3.3 with $h = (n-1)d + 1$.

First, we expand H up to order $2d + 1$. This is done by solving $\mathsf{MatFracExp}_{n,2d+1}$ once, taking $B = I_n$ and $h = (n-1)d+1 = O(nd)$. From Theorem 3.3 we know that H is a strictly proper matrix fraction which admits left and right factorizations $A^{-1}(AH)$ and $(HA)A^{-1}$. Strict properness further implies that the degrees of both AH and HA must be less than the degree of A [10, Lemma 6.3-10], and are thus bounded by d as well. Therefore, these left and right factorizations of H are factorizations of degree d and, using Theorem 3.5, we can reconstruct H from its expansion up to order $2d + 1$ as $H = R^{-1}S$. This reconstruction corresponds to solving problem $\mathsf{MatFracRec}_{n,d}$ once. On one hand, we know by Theorem 3.5 that R is row-reduced. On the other hand, $A^{-1}(AH)$ and $R^{-1}S$ are coprime factorizations of the same fraction, which implies that there exists a unimodular U such that $UA = R$ [10, Theorem 6.5-4]. It follows that R is indeed a row-reduced form of A. By Lemma 4.1, this reduction to $\mathsf{MatFracExp}_{n,2d+1}$ and $\mathsf{MatFracRec}_{n,d}$ gives a solution to $\mathsf{RowRed}_{n,d}$ in $O\tilde{\ }(n^\omega d)$ operations in K.

4.4   Small nullspace computation ($\mathsf{Nullspace}_{n,d}$)

Given $A \in \mathsf{K}[x]^{n \times n}$ of degree d, the problem is to compute the rank r of A, and $N \in \mathsf{K}[x]^{(n-r) \times n}$ of rank $n - r$ such that $NA = 0$.

As already seen, a solution in the restrictive (*e.g.* generic) case when all minimal vectors have degrees in $O(d)$ is provided by a solution to $\mathsf{PartialNullSpace}_{n,d}$. In the general case the row degrees in a nullspace basis of A may be unbalanced, they range between 0 and nd [17, Theorem 3.3]. Previously known methods, whose cost is essentially driven by the highest Kronecker index, do not seem to allow the target complexity estimate $O\tilde{\ }(n^\omega d)$ (see for instance [17, Section 2]).

Our solution in [17] first reduces the general nullspace problem to the full column rank case via randomization. This consists in evaluating the rank r of A at a random $x = x_0$, then in compressing A to a full column rank matrix. We also derive a particular strategy when $n \gg r$. Consequently, for a simplified explanation here, we now assume that A has full column rank n and dimensions $(n + m) \times n$ with $m = O(n)$.

The algorithm then works in i steps with $1 \leq i \leq \log n$. At step i we compute a set of about $m/2^i$ nullspace vectors of degrees less that $\delta = 2^i d$. These vectors are obtained from $\log n$ solutions to $\mathsf{PartialNullSpace}_{m,\delta}$ for nullspace vectors of bounded degree $\delta = 2^i d$, and

involving matrices of decreasing dimensions $n + m/2^i$. Hence we essentially have a reduction to $\mathsf{PartialNullSpace}^*_{m,\delta}$. We may point out that the proof of Theorem 2.5 for the cost of the partial nullspace itself relies on solutions to $\mathsf{MatFracExp}_{m,\delta}$, and $\mathsf{MatFracRec}_{m,\delta}$. Nullspace vectors are computed using a matrix fraction expansion /reconstruction scheme.

The appropriate instances for $\mathsf{PartialNullSpace}_{m/2^i,2^i d}$, $1 \le i \le \log n$, are built as submatrices of the input matrix A. Our choices for these submatrices ensure the linear independency of the successive computed sets of nullspace vectors. The algorithm hence outputs a union of a logarithmic number of sets of linearly independent nullspace vectors. Each set, corresponding to an instance of $\mathsf{PartialNullSpace}_{m/2^i,2^i d}$, is a family of minimal vectors for a submatrix of A. The minimality is not preserved in general with respect to A, however we prove that small degree vectors are obtained [17, Proposition 7.1].

This reduction of $\mathsf{NullSpace}_{n,d}$ to $\mathsf{PartialNullSpace}^*_{m,\delta}$ and to $\mathsf{MatMul}^*_{n,d}$ for additional matrix multiplications establishes that a solution matrix N such that $NA = 0$ can be computed in $O\tilde{\ }(n^\omega d)$ operations in K by a randomized Las Vegas (certified) algorithm.

4.5 Factorization ($\mathsf{Factor}_{n,d}$)

Given a right factorization BA^{-1} of degree d of $H \in \mathsf{K}(x)^{n \times n}$, the problem is to compute polynomial matrices U and V such that $V^{-1}U = H$.

Corollary 3.6, together with the expansion of $H = BA^{-1}$, provides a solution to $\mathsf{FracMatRec}_{n,d}$ if H admits factorizations of degree d on both sides. The solution of the general case, we mean for an arbitrary left side factorization, induces several difficulties for dealing with unbalanced row degrees. These difficulties are bypassed using the techniques of Section 4.4.

By considering the polynomial matrix $[-A^T \; B^T]$ and solving $\mathsf{Nullspace}_{2n,d}$ we get U and V such that

$$[U \; V] \begin{bmatrix} -A \\ B \end{bmatrix} = 0.$$

Arguments similar to those used in the proof of Theorem 3.5 lead to the fact that V is non-singular. Hence a solution $V^{-1}U$ to the factorization problem is computed in $O\tilde{\ }(n^\omega d)$ operations in K. Note that since a solution to $\mathsf{Nullspace}_{2n,d}$ may not be minimal, the factorization $V^{-1}U$ may not be coprime.

References

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

[2] B. Beckermann, G. Labahn, and G. Villard. Normal forms for general polynomial matrices. *Journal of Symbolic Computation*. To appear.

[3] A. Bostan and É. Schost. Polynomial evaluation and interpolation on special sets of points. *Journal of Complexity*, 21(4):420–446, 2005.

[4] B. Bürgisser, C. Clausen, and M.A. Shokrollahi. *Algebraic Complexity Theory*, volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1997.

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

[6] G.D. Forney. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM J. Control*, 13:493–520, 1975.

[7] J. von zur Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, 1999.

[8] P. Giorgi, C.-P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *Proc. International Symposium on Symbolic and Algebraic Computation, Philadelphia, Pennsylvania, USA*, pages 135–142. ACM Press, August 2003.

[9] C.-P. Jeannerod and G. Villard. Essentially optimal computation of the inverse of generic polynomial matrices. *Journal of Complexity*, 21(1):72–86, 2005.

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

[11] E. Kaltofen. On computing determinants without divisions. In *International Symposium on Symbolic and Algebraic Computation, Berkeley, California USA*, pages 342–349. ACM Press, July 1992.

[12] E. Kaltofen and G. Villard. On the complexity of computing determinants. *Computational Complexity*, 13:91–130, 2004.

[13] D.E. Knuth. The analysis of algorithms. In *Proc. International Congress of Mathematicians, Nice, France*, volume 3, pages 269–274, 1970.

[14] C. Oară and A. Varga. Minimal degree coprime factorization of rational matrices. *SIAM J. Matrix Anal. Appl.*, 21:245–278, 1999.

[15] A. Schönhage. Schnelle Berechnung von Kettenbruchenwicklungen. *Acta Informatica*, 1:139–144, 1971.

[16] A. Storjohann. High-order lifting and integrality certification. *Journal of Symbolic Computation*, 36(3-4):613–648, 2003. Special issue International Symposium on Symbolic and Algebraic Computation (ISSAC'2002). Guest editors: M. Giusti & L. M. Pardo.

[17] A. Storjohann and G. Villard. Computing the rank and a small nullspace basis of a polynomial matrix. In *Proc. International Symposium on Symbolic and Algebraic Computation, Beijing, China*. ACM Press, July 2005.