Bivariate polynomial reduction and elimination ideal over finite fields

Gilles Villard

Bivariate polynomial reduction and elimination ideal over finite fields

Gilles Villard

CNRS, Univ. Lyon, ENS de Lyon, Inria, UCBL - LIP UMR 5668, Lyon, France

Abstract

Given two polynomials a and b in $\mathbb{F}_q[x, y]$ which have no non-trivial common divisors, we prove that a generator of the elimination ideal $\langle a, b\rangle \cap \mathbb{F}_q[x]$ can be computed in quasi-linear time. To achieve this, we propose a randomized algorithm of the Monte Carlo type which requires $(de \log q)^{1+o(1)}$ bit operations, where d and e bound the input degrees in x and in y respectively.

The same complexity estimate applies to the computation of the largest degree invariant factor of the Sylvester matrix associated with a and b (with respect to either x or y), and of the resultant of a and b if they are sufficiently generic, in particular such that the Sylvester matrix has a unique non-trivial invariant factor.

Our approach is to exploit reductions to problems of minimal polynomials in quotient algebras of the form $\mathbb{F}_q[x, y]/\langle a, b\rangle$. By proposing a new method based on structured polynomial matrix division for computing with the elements of the quotient, we succeed in improving the best-known complexity bounds.

Keywords: Complexity, algorithm, computer algebra, elimination ideal, resultant, polynomial structured matrix

1. Introduction

Given two coprime polynomials $a, b \in \mathbb{K}[x, y]$, where $\mathbb{K}$ is a commutative field, we consider the problem of computing the monic generator μ of the elimination ideal $\langle a, b\rangle \cap \mathbb{K}[x]$. The problem is related to that of bivariate lexicographic Gröbner bases. The polynomial μ is in fact the unique element in $\mathbb{K}[x]$ of a Gröbner basis of the ideal $\langle a, b\rangle$ for the lexicographic order $y > x$ (Cox et al., 2007, Chap. 3). The question is also close to that of the resultant (Lazard, 1985; Cox and D'Andrea, 2023). The resultant $\text{Res}_y(a, b)$ of a and b with respect to y is the determinant of the associated Sylvester matrix S_y over $\mathbb{K}[x]$ (von zur Gathen and Gerhard, 1999, Chap. 6). We have that μ is the last (highest degree) invariant factor of S_y in particular when the polynomial system $a = b = 0$ has no roots at infinity with respect to y (the y-leading coefficients of a and b are coprime). Here we mean invariant factors of S_y in the sense of the diagonal entries of its Smith normal form over $\mathbb{K}[x]$ (Jacobson, 2009, Thm. 3.8). When a and b are sufficiently generic, the Smith form of S_y has a unique non-trivial invariant factor, and up to a non-zero element in $\mathbb{K}$ the generator of the elimination ideal is the resultant (Cox et al., 2005, Chap. 3, Sec. 5-6). (Genericity is considered in the Zariski sense: a property is generic if it holds except on a hypersurface of the parameter space.)

To our knowledge, computing a generator of the elimination ideal or the resultant in quasi-linear time with respect to the input/output size is still beyond our reach in the general case and for an arbitrary field $\mathbb{K}$ (see Section 1.6). The generator μ has degree in $O(d^2)$ and can be computed using $\tilde{O}(d^3)$ arithmetic operations in $\mathbb{K}$ if a and b have total degree at most d (Lebreton et al., 2013). (The soft-O notation $\tilde{O}(c)$ is shorthand for a bound of the form $O(c \log^k c)$ for some positive k.) In the rest of the text we prefer to work with the degrees individually: we let d and e be the respective bounds for the input degrees of a and b in x and in y. The resultant $\mathrm{Res}_y(a, b)$ has degree in $O(de)$ and can be computed using $\tilde{O}(de^2)$ arithmetic operations (von zur Gathen and Gerhard, 1999, Chap. 11).

In this paper we consider the special case of a finite field $\mathbb{K} = \mathbb{F}_q$ with q elements, and a and b of x-degree at most d and y-degree at most e in $\mathbb{F}_q[x, y]$, with no non-trivial common divisors. Our main result is a randomized Monte Carlo algorithm with computes the monic generator of the elimination ideal using a quasi-linear number of $(de \log q)^{1+o(1)}$ bit operations (Theorem 6.2). This extends what was established in (Villard, 2023) which was limited to the situation where $a = b = 0$ has no roots at infinity. (Much of the material in this paper is shared with (Villard, 2023).) The same complexity bound applies to the computation of the last invariant factor of the Sylvester matrix (Corollary 6.3). Even in the case of a finite field, we do not know how to compute the resultant in quasi-linear time except under appropriate genericity assumptions (Section 7).

The complexity bound we establish has already been obtained in special cases. Our approach over finite fields is inspired by and goes further than the major steps taken with: the change of order algorithm of Poteaux and Schost (2013) for triangular sets and radical ideals ; the algorithm of van der Hoeven and Lecerf (2021a), which computes the resultant of generic polynomials with respect to the total degree. In the bivariate case, these two works are the first to provide solutions in quasi-linear expected time in the input/output size (Poteaux and Schost (2013) treat general multivariate cases). They are part of the same long line of research reducing elimination problems to linear algebra (Lazard, 1981; Cox et al., 2005, Chap. 2, Sec. 4 & Chap. 3, Sec. 6), and especially to the computation of minimal polynomials in quotient algebras (Lazard, 1992; Shoup, 1999).

1.1. *Minimal polynomials*

Let $I = \langle a, b \rangle$ be the (zero-dimensional) ideal generated by a and b in $\mathbb{K}[x, y]$, and consider the associated quotient algebra $\mathbb{A} = \mathbb{K}[x, y]/I$. The monic generator of the elimination ideal $I \cap \mathbb{K}[x]$ is the minimal polynomial μ of the multiplication by x in $\mathbb{A}$. We also recall in Section 2 that the last invariant factor of the Sylvester matrix S_y can be computed as a minimal polynomial under the condition that there are no roots at infinity (Lazard, 1985). By slightly modifying the input polynomials we will ensure that this condition is satisfied.

The two problems we are considering can therefore be instantiated as minimal polynomial computations. For efficiency, the minimal polynomial problem itself is reduced to a power projection problem (Kaltofen, 2000, Sec. 6). (A more complete list of references is given later in Section 1.7.) Given a linear form ℓ in the dual of $\mathbb{A}$ over $\mathbb{K}$, the minimal polynomial of x in $\mathbb{A}$ is computed as the one of the linearly generated sequence $\{\ell(x^i \bmod I)\}_{i\geq 0}$ over $\mathbb{K}$. The use of a random linear form preserves the recursion which is sought in $\mathbb{A}$ (Wiedemann, 1986) (Section 6). As observed by Shoup (1994), the power projection problem is dual to the modular composition problem (Brent and Kung, 1978). Finally, we rely on the approach of Kedlaya and Umans (2011) to solve the latter problems in quasi-linear time over finite fields. As we shall now see, this is made possible by a new algorithm we propose for arithmetic operations modulo the ideal.

1.2. Bivariate polynomial reduction

One of the bottlenecks in the above strategy is to perform arithmetic operations in $\mathbb{A}$ (van der Hoeven, 2017), even if only to compute the multiplication of two polynomials or the powers of x that need to be projected modulo the ideal I. Herein lies an important aspect of our contribution. In (Poteaux and Schost, 2013), the special case of triangular sets is considered. That is, in our context, when either a or b is univariate. On the other hand, the generic resultant algorithm of van der Hoeven and Lecerf (2021a) relies on Gröbner bases techniques, and on the normal form algorithm modulo I of van der Hoeven and Larrieu (2019).

In Section 3, we use polynomial matrix division instead (Kailath, 1980, Sec. 6.3). Considering a polynomial f in $\mathbb{K}[x,y]$ as a vector with entries in $\mathbb{K}[x]$, we reduce its x-degree using division by the polynomial Sylvester matrix S_y; note also that we may need to construct a Sylvester matrix from multiples of a and b if the dimensions do not match. By definition of the Sylvester matrix, the remainder of this division gives a new polynomial in the coset $f+I$. By a similar division after swapping the roles of x and y (using the Sylvester matrix S_x with respect to x), this leads to a normal form algorithm modulo I, up to a regularity assumption related to roots at infinity (Lemma 2.4 and Proposition 3.4). This algorithm is algebraic and deterministic for arbitrary fields. If f has x-degree at most δ and y-degree at most η, then it uses $\tilde{O}((d+\delta)(e+\eta))$ arithmetic operations (Proposition 3.4). A Sylvester matrix is a Toeplitz-like matrix (Bini and Pan, 1994). Our cost bound is based on fast structured matrix arithmetic which is discussed in Section 3.1. In particular, the normal form algorithm allows multiplication in $\mathbb{K}[x,y]/\langle a,b\rangle$ using $\tilde{O}(de)$ operations when the leading coefficients of a and b are sufficiently generic (Lemma 2.4). In the case of the total degree, for generic polynomials a,b with $\deg a \geq \deg b$, the algorithm of van der Hoeven and Larrieu (2019) costs $\tilde{O}((\deg a)(\deg b))$, after precomputing a concise Gröbner basis representation of the ideal using $\tilde{O}((\deg a)^2)$ operations. Thus the two algorithms are complementary in their assumptions (Section 3.3).

1.3. Extension of Kedlaya and Umans' techniques for the power projections

Once the normal form algorithm is available, i.e. the arithmetic operations in $\mathbb{A}$, it is possible to develop the general strategy of Shoup (1994, 1999) for the computation of modular power projections, coupled by duality with the algorithm of Kedlaya and Umans for modular composition Kedlaya and Umans (2011) (this latter reference treats the case of a univariate ideal I in $\mathbb{F}_q[x]$). This is what was generalized by both Poteaux and Schost (2013) and van der Hoeven and Lecerf (2021a), with respective forms of the ideal I that we have seen above. We proceed in the same way, in Section 4, integrating the new division algorithm into this overall process:

(i) For a polynomial $f \in \mathbb{F}_q[x]$ of degree $O(de)$ we compute f modulo I (Corollary 4.2). This operation is considered as modular composition according to $f(g(x)) \bmod I$ with $g = x$. Modular composition relies on multivariate multipoint evaluation (Kedlaya and Umans, 2011, Thm. 3.1);

(ii) The transposition principle (Bürgisser et al., 1997, Thm. 13.20) is used to obtain appropriate power projections (Section 3.4 and Corollary 4.3).

Since the degree of the resultant of a and b with respect to y (or x) is at most $2de$, it is sufficient to be able to compute (i) f modulo I for $\deg f < 4de$, and (ii) $\{\ell(x^i \bmod I)\}_{0 \le i < 4de}$, to derive the minimal polynomial of the sequence as desired in Section 1.1.

We establish in Section 4 that (i) and (ii) can be performed within our target cost bound over a finite field using bit operations.

1.4. Random conditioning of the Sylvester matrices

In order to exploit what we have seen above for general polynomials a and b, we need to address some regularity issues. The normal form algorithm modulo I we propose is based on matrix polynomial division. It is effective only if the leading (matrix) coefficient of the Sylvester matrices S_x and S_y with respect to x and y are invertible (see Lemma 2.4 and Proposition 3.4). If the input polynomials a and b result in S_x and S_y with singular leading coefficients, then we construct two new polynomials a' and b' which allow us to circumvent the difficulty. To do this, we use rudimentary transformations in Section 5 based on polynomial shifts and reversals. These transformations do not change much the structure of the ideal. The target polynomials for the ideal I are efficiently recovered from those computed with $\langle a', b' \rangle$. Our extension of the work in (Villard, 2023) for computing the elimination ideal for arbitrary a and b is obtained by also considering the minimal polynomial of an appropriate power κ of y with respect to the multiplication by x (Lemma 5.1). Required power projections of the form $\ell(y^\kappa x^i \bmod I)$ are computed in the same way as above for the projections of the powers of x (compare Corollary 4.3 and Corollary 4.4).

1.5. Elimination ideal $\langle a, b \rangle \cap \mathbb{F}_q[x]$, last invariant factor and resultant.

The ideas and results presented so far allow us to establish in Section 6 the announced complexity bounds for the generator of the elimination ideal (Theorem 6.2) and the last invariant factor of the Sylvester matrix (Corollary 6.3).

For general a and b we only compute a specific factor of the resultant, which is the last invariant factor of the Sylvester matrix. We consider in Section 7 some favorable cases in which the resultant can be obtained within the same complexity bound. In particular, the resultant is computed when the associated Sylvester matrix has a unique non-trivial invariant factor.

1.6. Comparison to previous work

Over an abstract field, we refer to (Lecerf, 2019; Villard, 2018; Pernet et al., 2024) and to the pointers found there for the problem of the bivariate resultant. Improvements have been achieved over the bound $O(de^2)$ for generic polynomials, but no general quasi-linear bounds seem to be known. The same applies to the modular composition problem which, as we saw in Sections 1.1 and 1.3, is involved in today's general approaches for the elimination ideal (Neiger et al., 2023).

Over a finite field, which is the case we are looking at, the approach of van der Hoeven and Lecerf (2021a) allows a quasi-linear bit cost for the resultant; for generic polynomials with respect to the total degree, and any $\epsilon > 0$, this leads to the complexity bound $O(((\deg a)(\deg b)\log q)^{1+\epsilon}) + \tilde{O}((\deg a)^2 \log q)$ when $\deg a \geq \deg b$ (van der Hoeven and Lecerf, 2021a, Thm. 1). Our algorithm covers this case, in particular. Genericity ensures that there is a unique non-trivial invariant factor (see Section 7), and we obtain a comparable asymptotic bound. By considering degree conditions on the variables individually we treat a larger class of problems with weaker assumptions. For polynomials of x-degree d and y-degree e, we compute the resultant in quasi-linear time when the Sylvester matrix S_y has a unique non-trivial invariant factor.

On the other hand, our approach is designed to compute a generator of the elimination ideal for arbitrary a and b. We are not aware of any previous method whose cost would be quasi-linear over finite fields under the same assumptions. The complexity of this problem is related to that of the resultant and bivariate lexicographic Gröbner bases (Dahan, 2022). In particular, for a and b of total degree at most d, we arrive at the bound $d^{2+o(1)} \log(q)^{1+o(1)}$, while previous estimates are $\tilde{O}(d^3 \log q)$ (Lebreton et al., 2013).

1.7. Bibliographical notes

We give here some additional references from which the results we use largely inherit. The adaptation of numerical matrix methods to the finite field setting began with the solution of sparse linear systems (Coppersmith et al., 1986; Wiedemann, 1986). These methods lead to projections of the powers of the involved matrix in order to compute its minimal polynomial, as evidenced by Wiedemann's approach (Wiedemann, 1986).

The connection is to be made with the use of power projections for computing minimal polynomials in quotient algebras, using the trace map in (Thiong Ly, 1989; Rifà and Borrell, 1991) and general projections in (Shoup, 1994; Kaltofen and Shoup, 1998; Shoup, 1999). (We do indeed have a multiplication endomorphism in the quotient.)

The duality between the power projection problem and the modular composition problem is observed by Shoup (1994).

In the context of solving polynomial systems, for which the literature is vast, we may refer to the use of the trace map in (Alonso et al., 1996; Rouillier, 1999; Gonzalez-Vega et al., 1999), or of arbitrary linear forms in (Bostan et al., 2003b). Structured matrices and duality are applied to multivariate polynomial problems in (Mourrain and Pan, 2000). Multivariate power projections are considered in (Shoup, 1999; Kaltofen, 2000), especially for minimal polynomials, and are exploited for the computation of special resultants in (Bostan et al., 2006), and for the change of order of variables for triangular sets in (Pascal and Schost, 2006). The connection between the change of ordering and linear algebra is also fruitful with the use of power projections of a multiplication matrix in (Faugère and Mou, 2011; Faugère et al., 2014), and in order to take advantage of sparsity (Faugère and Mou, 2017), which brings us back to Wiedemann's algorithm.

The Sylvester matrix S_y (or S_x) is a polynomial matrix which we manipulate as such (Lazard, 1985). This can be thought of as working in a $\mathbb{K}[y]$-module rather than in a $\mathbb{K}$-vector space in order to represent the quotient algebra $\mathbb{A}$ (Jacobson, 2009, Sec. 3.10), and implement the operations on its elements. This direction has been taken in particular (Berthomieu et al., 2022b) for a change of ordering of Gröbner bases algorithm, and in (Schost and St-Pierre, 2023b) for a p-adic Gröbner basis algorithm.

Regarding implicitly represented matrices, an open problem is to compute the characteristic polynomial in essentially the same time as for the minimal polynomial (Kaltofen, 2000, Sec. 3). This is especially true for sparse or structured matrices. The question of whether the bivariate resultant can be computed in essentially the same time as for the last invariant factor of the Sylvester matrix seems to be similar to Kaltofen's open problem.

1.8. Model of computation and notations

The normal form algorithm and its transpose for polynomials in $\mathbb{K}[x, y]$ modulo the ideal $\langle a, b\rangle$ are presented using an algebraic model (Section 3), and work e.g. with computation trees (Bürgisser et al., 1997, Sec. 4.4). Complexity bounds correspond to numbers of arithmetic operations performed in $\mathbb{K}$. The application of Kedlaya and Uman's techniques in Section 4, and thus the computation of the generator of the elimination ideal and of the last invariant factor in Section 6, is based on a RAM bit complexity model. We assume that arithmetic operations in $\mathbb{F}_q$ can be done in time $\tilde{O}(\log q)$, and that the RAM can produce a random element uniformly distributed in $\mathbb{F}_q$ at the same cost.

Throughout the paper we consider two polynomials $a, b \in \mathbb{K}[x, y]$, of degrees d_a and d_b in x, and e_a and e_b in y, respectively. We use the notations $d = \max\{d_a, d_b\}$ and $e = \max\{e_a, e_b\}$. The corresponding Sylvester matrices with respect to x and y are $S_x \in \mathbb{K}[y]^{n_x\times n_x}$ and $S_y \in \mathbb{K}[x]^{n_y\times n_y}$, with dimensions $n_x = d_a + d_b$ and $n_y = e_a + e_b$. The resultants $\mathrm{Res}_x(a, b) \in \mathbb{K}[y]$ and $\mathrm{Res}_y(a, b) \in \mathbb{K}[x]$, of a and b with respect to x and y, are the respective determinants of S_x and S_y (von zur Gathen and Gerhard, 1999, Chap. 6). We assume that a and b have no non-trivial common divisors, so both S_x and S_y are non-singular. We concentrate on computations in relation to $\langle a, b\rangle \cap \mathbb{K}[x]$ and $\mathrm{Res}_y(a, b) = \det S_y$ (the conclusions would be unchanged in relation to $\langle a, b\rangle \cap \mathbb{K}[y]$ and $\mathrm{Res}_x(a, b)$).

We use expressions such as "x-degree" or "y-leading coefficient" to indicate the variable in question, and use $\deg_x$ and $\deg_y$ in formulas when bivariate polynomials are involved. For example, subscripts in $\mathbb{K}[x, y]_{<(d,n_y)}$ indicate degree bounds in x and y, and $\mathbb{K}[x]_d$ is the set of polynomials of degree d. We are often led to manipulate reversals of polynomials. For $k \geq 0$, we define the reversal of a polynomial $f \in \mathbb{K}[x]$ with respect to k as $\mathrm{rev}_k(f) = x^k f(1/x)$; by default, if k is not specified, the reversal is taken with respect to the degree of the polynomial. This is generalized to polynomial matrices, which are considered as matrix polynomials, i.e. with matrix coefficients. The polynomials in $\mathbb{K}[x, y]$ are identified with the (column) vectors of their coefficients, using dimensions that are clear from the context. For example, given $f = f_0(x) + f_1(x)y + \ldots + f_d(x)y^d$ and $n \geq d+1$, $v_y(f) \in \mathbb{K}[x]^n$ denotes the vector $[0 \ \ldots \ 0 \ f_d \ \ldots \ f_0]^{\mathsf{T}}$.

2. Bivariate ideals and resultant

Given two coprime polynomials a and b in $\mathbb{K}[x, y]$, let $I = \langle a, b \rangle$ be the (zero-dimensional) ideal generated by a and b in $\mathbb{K}[x, y]$, and $\mathbb{A} = \mathbb{K}[x, y]/I$ be the associated quotient algebra. In this section we give the basic notions and results we need concerning the connections between the elimination ideal $I \cap \mathbb{K}[x]$, the resultant $\operatorname{Res}_y(a, b)$ and some invariant properties of the Sylvester matrix S_y with respect to y. The elimination ideal is treated with its specific generator which is the minimal polynomial of the multiplication by x in $\mathbb{A}$ (Cox et al., 2005, Chap. 2, Sec. 4).

As univariate polynomial matrix, the Sylvester matrix S_y is unimodularly equivalent to a matrix $\operatorname{diag}(s_1, \ldots, s_n) \in \mathbb{K}[x]^{n_y \times n_y}$ in Smith normal form, where s_n is the highest degree invariant factor. We are not always able to compute the resultant within the cost objective. However, we are able to compute the last invariant factor s_n (Corollary 6.3). Using the structure theory of finitely generated modules, this last invariant factor can be seen as the minimal polynomial of a linear transformation in a finite dimensional $\mathbb{K}$-vector space (Jacobson, 2009, Sec. 3.10). Such a formalism has occasionally been used to efficiently compute general matrix normal forms Villard (1997); Storjohann (2000).

For Sylvester matrices, and in the broader context of solving polynomial systems, this is related to the use of a multiplication map on a quotient algebra Lazard (1981). We rely on the following results, which are direct consequences of a theorem of Lazard.

Lemma 2.1 (Lazard, 1985, Thm. 4)**.** *The last invariant factor of S_y is a multiple of the minimal polynomial of the multiplication by x in $\mathbb{A}$. Both polynomials coincide if the y-leading coefficients of a and b are coprime in $\mathbb{K}[x]$.*

Proof. The last invariant factor is a multiple of the last diagonal entry $h \in \mathbb{K}[x]$ of the Hermite form, where the latter is lower triangular and obtained by unimodular column transformations. The polynomial h is in $I \cap \mathbb{K}[x]$ (combinations of columns of S_y are seen as combinations of a and b), which gives the first assertion. When the leading coefficients are coprime, the divisibility property (i) in (Lazard, 1985, Thm. 4) shows that the Hermite form of S_y can be brought to Smith form using unimodular (row) transformations, without modifying the diagonal. From (ii) in (Lazard, 1985, Thm. 4), the last invariant factor is therefore an element of a reduced Gröbner basis of I, and as polynomial in $\mathbb{K}[x]$ it generates the elimination ideal $I \cap \mathbb{K}[x]$. □

The condition on the leading coefficients of a and b in Lemma 2.1 is the fact that the system $a = b = 0$ has no roots at infinity with respect to y. In general, the resultant and the last invariant factor may have terms coming from both the affine variety and the behavior at infinity (Cox et al., 2005, Chap. 3).

The resultant can be deduced from Lemma 2.1 when the Smith form of S_y has a unique non-trivial invariant factor and there are no roots at infinity. This corresponds to situations in which the ideal I has a shape basis (Gianni et al., 1988; Becker et al., 1994).

Lemma 2.2 (Consequence of (Lazard, 1985, Thm. 4))**.** *The y-leading coefficients of a and b are coprime in $\mathbb{K}[x]$ and there exist two polynomials $\mu, \lambda \in \mathbb{K}[x]$ such that $I = \langle \mu(x), y - \lambda(x) \rangle$ if and only if, up to a non-zero element in $\mathbb{K}$, the resultant $\operatorname{Res}_y(a, b)$ is the minimal polynomial μ of the multiplication by x in $\mathbb{A}$.*

Proof. From (Lazard, 1985, Thm. 4), under the hypothesis $I = \langle \mu(x), y - \lambda(x) \rangle$ and using the coprimeness, we know that the Hermite form of S_y has a unique non-trivial diagonal entry, which is μ. Therefore, the latter is also the determinant of S_y, up to the normalization to a monic polynomial in the Hermite form.

Conversely, the last element $h \in \mathbb{K}[x]$ of the diagonal of the Hermite form of S_y is in I. Hence h must be a multiple of μ, and of the resultant by assumption. It follows that $h = \operatorname{Res}_y(a, b) = \mu$, and all the other diagonal entries of the Hermite form are equal to 1. This proves that the y-leading coefficients of a and b are coprime since otherwise the first diagonal entry of the Hermite form would be a non-constant polynomial in $\mathbb{K}[x]$. Item (ii) (Lazard, 1985, Thm. 4) allows to conclude. □

About the links between the resultant and the associated ideal, the reader may especially refer to (Cox and D'Andrea, 2023, Thm. 1.3), which is a general multivariate version of Lemma 2.2.

Example 2.3. *The Sylvester matrix may have a unique non-trivial invariant factor (that our algorithm will compute) even though there are roots at infinity. With $\mathbb{K} = \mathbb{F}_2$, take $a = (x + 1)\, y + x^2$ and $(x + 1)\, y^2 + y$. The Hermite normal form of S_y is*

$$S_y U = \begin{bmatrix} x+1 & 0 & 0 \\ 1 & x+1 & 0 \\ 0 & x^2 & x^2\,(x+1) \end{bmatrix},$$

where U is unimodular in $\mathbb{K}[x]^{3 \times 3}$. None of the arguments used for Lemmas 2.1 and 2.2 apply: the Hermite form cannot be brought to Smith form using unimodular row operations without modifying the diagonal (as used in the proof of Lemma 2.1), and the form is not either trivial (proof of Lemma 2.2). We have $I = \langle x^2, y \rangle$, hence $I \cap \mathbb{K}[x] = \langle x^2 \rangle$, and the last invariant factor of S_y is $\operatorname{Res}_y(a, b) = x^2\,(x+1)^3$. □

As mentioned in Sections 1.2 and 1.4, our reduction algorithm modulo I relies on regularity assumptions on the Sylvester matrices with respect to both x and y. For efficient polynomial matrix division (Section 3.2), we need the leading matrix coefficients to be invertible. This can be specified in terms of the column reducedness of polynomial matrices (Kailath, 1980, Sec. 6.3, p.384), and related to the absence of roots at infinity. Let S be a matrix in $\mathbb{K}[x]^{n \times n}$ whose column j has degree d_j. The (column) leading (matrix) coefficient of S is the matrix in $\mathbb{K}^{n \times n}$ whose entry (i, j) is the coefficient of degree d_j of the entry (i, j) of S. We manipulate non-singular univariate polynomial matrices, and say that such a matrix is column reduced if its leading coefficient is invertible.

Lemma 2.4. *S_x is column reduced if and only if, the y-leading coefficients of a and b are coprime and at least one of latter polynomials in $\mathbb{K}[x]$ has maximal degree d_a or d_b, respectively.*

Proof. Let $s, t \in \mathbb{K}[x]$ be the y-leading coefficients of a, b, with respective degrees d_s and d_t. The columns of the leading coefficient of S_x are given by the vectors in $\mathbb{K}^{d_a+d_b}$ associated with

$$x^{d_b-1}s, x^{d_b-2}s, \ldots, s, x^{d_a-1}t, x^{d_a-2}t, \ldots, t.$$

If S_x is column reduced then the first row of its leading matrix is non-zero and either $d_s = d_a$ or $d_t = d_b$. Let's say that $d_s = d_a$ (up to a column permutation). The leading coefficient of S_x is therefore given by

$$x^{d_b-1}s, x^{d_b-2}s, \ldots, x^{d_t}s, x^{d_t-1}s, x^{d_t-2}s, \ldots, s, x^{d_s-1}t, x^{d_s-2}t, \ldots, t, \tag{1}$$

and we see that its rank ensures the non-singularity of the Sylvester matrix associated with s and t since the latter is given by

$$x^{d_t-1}s, x^{d_t-2}s, \ldots, s, x^{d_s-1}t, x^{d_s-2}t, \ldots, t. \tag{2}$$

Conversely, from the independence of the vectors in Eq. (2) we deduce the independence of the vectors in Eq. (1), and hence the column reducedness of S_x. □

The work of Lazard (1985) is central to our approach to the bivariate problems we are interested in. Extensions of (Lazard, 1985, Thm. 4) are provided by (Schost and St-Pierre, 2023b, Prop. 1) without any assumptions about roots at infinity, and by (Berthomieu et al., 2022b, Thm. 4.1) in a more general multivariate context. For a study of differences between the resultant and the generator of the elimination ideal the reader may refer to (Mantzaflaris et al., 2016).

3. Bivariate polynomial division

In this section we propose a normal form algorithm for bivariate polynomials modulo the ideal $I = \langle a, b\rangle$. The algorithm is based on matrix polynomial division. Bivariate polynomials in $\mathbb{K}[x, y]$ are treated as univariate polynomial vectors alternately over $\mathbb{K}[x]$ and $\mathbb{K}[y]$, dividing such a vector by S_y or S_x is actually equivalent to reducing the associated polynomial modulo the ideal (Proposition 3.4).

Sylvester matrices are Toeplitz-like matrices. We first recall in Section 3.1 how operations can be performed on matrices of this class taking into account their structure (Bini and Pan, 1994; Pan, 2001). We then study the division with remainder of a polynomial vector by S_y or S_x in Section 3.2. To define a normal form and perform the division efficiently, we rely on a regularity assumption on the leading coefficient matrices: we assume that S_x and S_y are column reduced. This assumption is ultimately harmless for computing the minimal polynomial of x and the last invariant factor of the Sylvester matrix (Section 5). In Section 3.3 we present the normal form algorithm. We keep the same notation as before for the degrees of a and b, and the dimensions of the matrices; in particular, d is the maximum degree in x and S_y is $n_y \times n_y$. Given a polynomial $f \in \mathbb{K}[x, y]$, we show how to compute a unique polynomial $\hat{f} \in \mathbb{K}[x, y]_{<(d,n_y)}$, that we denote by $\hat{f} = f \text{ rem } I$, such that $f - \hat{f} \in I$ (Proposition 3.4). Uniqueness is ensured using a properness property provided by the polynomial matrix division. The construction is a $\mathbb{K}$-linear map sending f to $\hat{f}$ whose y-coefficients are given by the entries of a vector $v_y(\hat{f}) \in \mathbb{K}[x]^{n_y}_{<d}$ such that $S_y^{-1} v_y(\hat{f})$ is strictly proper (each entry has its numerator degree less than its denominator degree), see Eq. (3). This allows us to represent the elements in $\mathbb{A}$ by normal forms. The transpose algorithm, which computes the corresponding power projections, is derived in Section 3.4. With $\deg_x a = d_a$, $\deg_x b = d_b$, $\deg_y a = e_a$, and $\deg_y b = e_b$, the quotient algebra $\mathbb{A}$ has at most the dimension $d_a e_b + d_b e_a$. To represent its elements, the quotient is embedded in the space $\mathbb{K}[x]^{n_y}_{<d}$ of dimension $dn_y = \max\{d_a, d_b\}(e_a + e_b)$ which can therefore be slightly larger than the dimension of $\mathbb{A}$ (Example 3.2).

3.1. Structured matrix arithmetic

The normal form algorithm exploits the fact that Sylvester matrices are structured. The class of structure we are dealing with is that of Toeplitz-like polynomial matrices, which are commonly treated using the notion of displacement rank (Kailath et al., 1979). The notion allows for a concise matrix representation through which matrix arithmetic can be efficiently implemented (Bini and Pan, 1994; Pan, 2001).

Given by the polynomials a and b, S_x and S_y are represented using $O(de)$ elements of $\mathbb{K}$. The division algorithm requires the solution of associated linear systems and uses matrix inversion with truncated power series entries. We consider that polynomial Sylvester matrices and their inverses are represented using their concise Toeplitz-like representations (Bini and Pan, 1994). This is achieved, for example, by extending the ΣLU form defined over fields by Kaltofen (1994), to polynomials or truncated power series (Pernet et al., 2024, Sec. 3).

Multiplication of an $n \times n$ polynomial Sylvester matrix of degree d by a polynomial vector of degree at most l over $\mathbb{K}[x]$, can be done using $\tilde{O}(n(d+l))$ arithmetic operations in $\mathbb{K}$ (Bini and Pan, 1994). This cost bound is valid for the same type of multiplication using instead the inverse of the matrix modulo x^l, if it exists.

If $T \in \mathbb{K}^{n \times n}$ is a non-singular Sylvester matrix and $v \in \mathbb{K}^n$, then the linear system $T^{-1}v$ can be solved using $\tilde{O}(n)$ arithmetic operations. This is obtained by combining an inversion formula for the Sylvester matrix Labahn (1992), and matrix Padé approximation Beckermann and Labahn (1994) (see also (Bini and Pan, 1994, Chap. 2, Sec. 9) and (Villard, 2018, Sec. 5)). The declination of this is used in Section 3.2 over truncated power series modulo x^l. Let $S \in \mathbb{K}[x]^{n \times n}_d$ be a polynomial Sylvester matrix such that $\det S(0) \neq 0$, and consider a vector $v \in \mathbb{K}[x]^n$ of degree at most l. The system $S^{-1}v$ can be solved modulo x^l using $\tilde{O}(n(d + l))$ arithmetic operations. This can be done by x-adic lifting using that $S(0)$ is nonsingular (Moenck and Carter, 1979; Dixon, 1982). We refer to the description of the method in (Dixon, 1982), carried over to the case $\mathbb{K}[x]$. The cost is essentially that of $O(d+l)$ multiplications of $S(0)^{-1}$ by a vector over $\mathbb{K}$. From (Villard, 2018, Prop. 5.1), the matrix inverse modulo x^l can itself be computed (with concise representation) within the same cost bound.

3.2. Matrix and bivariate polynomial division

Consider S in $\mathbb{K}[x]^{n\times n}$, non-singular of degree d. For any vector $v \in \mathbb{K}[x]^n$, we know from (Kailath, 1980, Thm. 6.3-15, p. 389) that there exist unique $w, \hat{v} \in \mathbb{K}[x]^n$ such that

$$v = Sw + \hat{v}, \tag{3}$$

and $S^{-1}\hat{v}$ is strictly proper. We mean that each entry of the vector has its numerator degree less than its denominator degree.

From (Kailath, 1980, Thm. 6.3-10, p. 383) we further have that the polynomial remainder vector $\hat{v}$ has degree less than d; note however that uniqueness is ensured by properness and not by the latter degree property (Example 3.2). The following will be applied to both S_x and S_y, hence we take a general notation S for the statement. We propose a structured matrix polynomial adaptation of the Cook-Sieveking-Kung algorithm for (scalar) polynomial division with remainder, for which the reader may refer to (von zur Gathen and Gerhard, 1999, Sec. 9.1).

Lemma 3.1. *Let $S \in \mathbb{K}[x]^{n\times n}$ be a column reduced Sylvester matrix of degree d. Consider a vector $v \in \mathbb{K}[x]^n$ of degree at most l. The unique remainder $\hat{v}$ of the division of v by S as in Eq. (3) can be computed using $\tilde{O}(n(d+l))$ arithmetic operations in $\mathbb{K}$.*

Proof. Consider that S is associated with two polynomials $a, b \in K[x,y]$, such that $S = S_y$, with e_1 columns of degree d_1 and e_2 columns of degree d_2. Up to row and column permutations we assume that $d = \max\{d_1, d_2\} = d_1$.

We first treat the case $d = d_1 = d_2$. All the columns of S have the same degree, hence since S is column reduced it is also row reduced (use the definition given before Lemma 2.4, on the rows).

If $l < d$, then we take $\hat{v} = v$. From (Kailath, 1980, Thm. 6.3-11, p. 385), by row reducedness, we know that $S^{-1}\hat{v}$ is strictly proper. If $l \geq d$, the polynomial division can be performed by reformulating (von zur Gathen and Gerhard, 1999, Sec. 9.1, Eq. (2)) on matrices. Since S has non-singular leading matrix, by the predictable degree property (Kailath, 1980, Thm. 6.3-13, p. 387) we know that the quotient vector w has degree $\deg v - d$, hence at most $l - d$. Using reversals of matrix polynomials, Eq. (3) can be rewritten as $\mathrm{rev}_l(v) = \mathrm{rev}_d(S)\cdot\mathrm{rev}_{l-d}(w) + x^{l-d+1}\mathrm{rev}_{d-1}(\hat{v})$, hence we have

$$\mathrm{rev}_{l-d}(w) \equiv \mathrm{rev}_d(S)^{-1}\mathrm{rev}_l(v) \bmod x^{l-d+1}. \tag{4}$$

Remark that by reducedness assumption the coefficient matrix of degree 0 of $\mathrm{rev}_d(S)$ is non-singular, thus the latter matrix in invertible modulo x^{l-d+1}. As soon as $w' = \mathrm{rev}_{l-d}(w)$ hence $w = \mathrm{rev}_{l-d}(w')$ are known, then $\hat{v}$ can be deduced using $\hat{v} = v - Sw$. We know that $S^{-1}\hat{v}$ is strictly proper using reducedness, with the argument we saw earlier. Using fast structured matrix arithmetic (Section 3.1), $\mathrm{rev}_{l-d}(w)$ is computed from Eq. (4) and $\hat{v}$ is obtained within the claim cost bound.

When $d = d_1 > d_2$, first we balance the columns degrees. With $\delta = d_1 - d_2 > 0$, take $D = \mathrm{diag}(x^\delta, \ldots, x^\delta, 1, \ldots, 1)$, with e_1 entries x^δ. The matrix $T = SD^{-1}$ has all its column degrees equal to d_2. Here and below the degree of a rational function is the difference between the degrees of the numerator and the denominator. Column and row reducedness are extended accordingly.

If $l < d_2$ we let $v' = v$, otherwise we can compute a polynomial vector w' of degree at most $l - d_2$ and $v' = v - Tw'$ of degree less than d_2 such that $T^{-1}v'$ is strictly proper. This is done using Eq. (3) after multiplying everything by x^δ so as to be reduced to a division with polynomial matrices, in time $\tilde{O}(n(d+l))$. This is similar to the $d_1 = d_2$ case above since T is column reduced. Then, taking the quotient of the first e_1 entries of w' by x^δ, we write $w' = Dw + z$, where z is of degree less than δ and such that only its first e_1 entries may be non-zero. The vector w remains of degree at most $l - d_2$, and we obtain $\hat{v}$ in time $\tilde{O}(n(d+l))$ as $\hat{v} = v - Sw = v - SD^{-1}(w' - z) = v' + SD^{-1}z$. To complete the proof we check that $S^{-1}\hat{v}$ is strictly proper. This vector is $S^{-1}\hat{v} = S^{-1}v' + D^{-1}z = D^{-1}T^{-1}v' + D^{-1}z$. By construction, $T^{-1}v'$ is strictly proper, so it is the same for $D^{-1}T^{-1}v'$; z has degree at most $\delta - 1$ for its first e_1 entries (the others are zero), hence $D^{-1}z$ is strictly proper. □

3.3. Normal form modulo the bivariate ideal

Given a polynomial $f \in \mathbb{K}[x,y]$ whose y-degree is less than the dimension n_y of S_y, we can apply Lemma 3.1 to the vector $v_y(f) \in \mathbb{K}[x]^{n_y}$ of the coefficients of f. Equation (3) becomes $v_y(f) = S_y w + v_y(\hat{f})$ on vectors, and by definition of the Sylvester matrix we have $\hat{f} = f - ua - vb \in f + I$ for some $u, v \in \mathbb{K}[x,y]$, with $\hat{f}$ of x-degree less than d. We show with Proposition 3.4 that, thanks to the uniqueness of the remainder, this allows us to define a normal form modulo $\langle a, b\rangle$. The general y-degree case for f is treated using a preliminary division by S_x (whose entries are in $\mathbb{K}[y]$) in order to reduce the degree in y. The whole construction gives a $\mathbb{K}$-linear map

$$\begin{array}{rcl} \varphi : \mathbb{K}[x,y] & \to & \mathbb{K}[x,y]_{<(d,n_y)} \\ f & \mapsto & \hat{f} = f \text{ rem } I \end{array} \tag{5}$$

such that $f - \varphi(f) \in I$, and $\varphi(g) = 0$ if $g \in I$. The map φ is suitable to represent the elements in $\mathbb{A}$ by normal forms.

Example 3.2. *For $\mathbb{K} = \mathbb{Q}$, consider $a = x^2y + y$ and $b = xy^2 + x$, with $d = d_a = 2$ and $n_y = e_a + e_b = 1 + 2 = 3$. If $f = x$ then both f and $f - b = -xy^2$ are in $\mathbb{K}[x,y]_{<(2,3)}$, hence the map φ might not be surjective. The division as in Eq. (3) leads to*

$$v_y(f) = S_y w + v_y(\hat{f}) = \begin{bmatrix} x^2+1 & 0 & x \\ 0 & x^2+1 & 0 \\ 0 & 0 & x \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -x \\ 0 \\ 0 \end{bmatrix},$$

and $\hat{f} = f - b$ since $S_y^{-1}v_y(\hat{f})$ is strictly proper, whereas $S_y^{-1}v_y(f)$ is not. Note that the quotient algebra $\mathbb{K}[x,y]/\langle a,b\rangle$ has dimension 5, which is smaller than the dimension of $\mathbb{K}[x,y]_{<(2,3)}$. □

If $\eta = \deg_y f \geq n_y$ and $\delta = \deg_x f$ is less than the dimension n_x of S_x we can go straight to the division with S_x. Otherwise, as we now see with Lemma 3.3, we first extend S_x to a larger suitable Sylvester matrix T_x of dimension $\delta + 1$. By linearization, we assign to f a vector $v_x(f) \in \mathbb{K}[y]^{\delta+1}$ of y-degree η. Then using division by T_x, whose y-degree is the degree e of S_x, we can compute f' of y-degree less than $e < n_y$, such that $f - f' \in I$.

Lemma 3.3. *Assume that the Sylvester matrix S_x associated with a and b with respect to x is column reduced. Consider $f \in \mathbb{K}[x, y]$ of x-degree at most δ and y-degree at most $\eta \geq n_y$. Using $\tilde{O}((n_x + \delta)\eta)$ arithmetic operations in $\mathbb{K}$ we can compute a polynomial $f' \in \mathbb{K}[x, y]$, of x-degree at most $\max\{n_x - 1, \delta\}$ and y-degree less than $e < n_y$, such that $f - f' \in I$.*

Proof. If δ is less than n_x, we take $T_x = S_x$. Otherwise, let $m = \delta - n_x + 1$, and denote the y-leading coefficients of a, b by $s, t \in \mathbb{K}[x]$. Since S_x is column reduced, we know from Lemma 2.4 that $\gcd(s, t) = 1$. Either s or t is not divisible by x, let us assume that it is s, and take for T_x over $\mathbb{K}[y]$, the Sylvester matrix associated with a and $x^m b$ with respect to x. The Sylvester matrix associated with s and $x^m t$ is non-singular, hence T_x is column reduced by Lemma 2.4 again: if either $\deg s = d_a$ or $\deg t = d_b$, then either $\deg s = d_a$ or $\deg t + m = d_b + m$.

This matrix T_x has dimension $\max\{n_x, \delta + 1\}$, and degree $e = \max\{e_a, e_b\}$ in y. The remainder of the division of $v_x(f)$ by T_x gives f' such that $f - f' \in I$, its y-degree is less than that of T_x, and its x-degree is less than the dimension of T_x. The cost bound is from Lemma 3.1, with a matrix of dimension $n = \max\{n_x, \delta + 1\}$ and degree e, and a vector of degree $l = \eta \geq e$. □

Lemma 3.3 allows to first reduce the y-degree, the reduction of the degree in x now also ensures the normal form.

Proposition 3.4. *Assume that the Sylvester matrices S_x and S_y associated with a and b are column reduced, and consider $f \in \mathbb{K}[x, y]$. The $\mathbb{K}$-linear map φ in Eq. (5) is well defined by choosing for $\hat{f}$ the unique polynomial in $\mathbb{K}[x, y]_{<(d, n_y)}$ such that $f - \hat{f} \in I$, and $S_y^{-1} v_y(\hat{f})$ is strictly proper. If f has x-degree at most δ and y-degree at most η, then this normal form for $f + I$ in $\mathbb{A}$ can be computed using $\tilde{O}((d + \delta)(e + \eta))$ arithmetic operations in $\mathbb{K}$.*

Proof. We show the existence of such an $\hat{f}$ for every f, then show that $\hat{g} = 0$ if $g \in I$. After division by S_x using Lemma 3.3 we have $f' \in \mathbb{K}[x, y]$ of y-degree less than $e < n_y = e_a + e_b$ such that $f - f' \in I$. Then by Lemma 3.1, that is by division by S_y, we obtain $\hat{f} \in \mathbb{K}[x, y]_{<(d, n_y)}$ such that $f' - \hat{f} \in I$, hence $f - \hat{f} \in I$. By construction, $S_y^{-1} v_y(\hat{f})$ is strictly proper. For $g \in I$, this first leads to some g' of y-degree less than $e < n_y$. Since S_y is column reduced, we know from Lemma 2.4 that the y-leading coefficients of a and b are relatively prime, hence using (Lazard, 1985, Lem. 7) there exist polynomials $r, s \in \mathbb{K}[x, y]$ such that $g' - ra - sb = 0$, $\deg_y r < e_b$, and $\deg_y s < e_a$. From uniqueness it follows that we must have $\hat{g} = 0$ because this value is appropriate using the above identity. The map $\varphi(f) = \hat{f}$ is well defined and provides a normal form. For f_1, f_2 in the coset $f + I$ we indeed have $\varphi(f_1 - f_2) = 0$ hence $\varphi(f_1) = \varphi(f_2)$ by $\mathbb{K}$-linearity of the divisions. From Lemma 3.3, the first division by S_x costs $\tilde{O}((d+\delta)(e+\eta))$, where we use that S_x has dimension $n_x \le 2d$ and degree e. This leads to the next division of a vector of degree at most $\max\{n_x-1,\delta\}$ by S_y, whose dimension is $n_y < 2e$ and degree d. Using Lemma 3.1 this adds $\tilde{O}(e(d+\delta))$ operations. □

Example 3.5. *We continue with $a = x^2y+y$ and $b = xy^2+x$ as in Example 3.2; S_x and S_y have dimension $n_x = n_y = 3$. For $f = y^3+x^3y^2+1$, we first reduce the degree in y using S_x. Since $\delta = \deg_x f \ge n_x$, we cannot directly use S_x which is 3×3. Following the proof of Lemma 3.3 we increase the dimension and consider $T_x \in \mathbb{K}[y]^{4\times 4}$, the Sylvester matrix with respect to x associated with a and xb. The first division is therefore:*

$$v_x(f) = T_x w_1 + v_x(f') = \begin{bmatrix} y & 0 & y^2+1 & 0 \\ 0 & y & 0 & y^2+1 \\ y & 0 & 0 & 0 \\ 0 & y & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y^2 \\ 1 \\ -y \end{bmatrix} + \begin{bmatrix} -1 \\ y \\ 0 \\ 1 \end{bmatrix}.$$

The new polynomial is $f' = -x^3+yx^2+1$, its degree in y is $1 < n_y$. So the division by $S_y \in \mathbb{K}[x]^{3\times 3}$ in order to reduce the degree in x is now possible:

$$v_y(f') = S_y w_2 + v_y(\hat{f}) = \begin{bmatrix} x^2+1 & 0 & x \\ 0 & x^2+1 & 0 \\ 0 & 0 & x \end{bmatrix} \begin{bmatrix} x \\ 1 \\ -x^2 \end{bmatrix} + \begin{bmatrix} -x \\ -1 \\ 1 \end{bmatrix},$$

and we obtain the normal form $\hat{f} = -xy^2-y+1 \in \mathbb{K}[x,y]_{<(2,3)}$. □

The assumptions of Proposition 3.4 are central to be able to reduce the degree in x and also ensure the normal form. The following example describes a situation with the existence of roots at infinity with respect to y.

Example 3.6. *With $\mathbb{K} = \mathbb{F}_7$, take $a = (x+3)y+x^2+5x+5$ and $b = (x+3)(x+4)y+x^2+4x+2$. The minimal polynomial of x in the quotient algebra is $x+2$ but cannot be obtained by combinations of a and b of y-degree less than $e_a = e_b = 1$, i.e. using combinations of the columns of S_y. The vector $[0 \;\; x+2]^\intercal$ is its own remainder of the division by S_y, hence $x+2$ is not reduced to zero while being in the ideal. In this case however, the Smith normal form of S_y has a unique non-trivial invariant factor. Thanks to the random conditioning of Section 5 and the use of certain powers of y (Lemma 5.1), we correctly compute the minimal polynomial and the resultant (see Sections 6 and 7).* □

Since the multiplication in $\mathbb{K}[x,y]$ can be computed in quasi-linear time (von zur Gathen and Gerhard, 1999, Sec. 8.4), Proposition 3.4 allows multiplication in $\mathbb{K}[x,y]/\langle a,b\rangle$ using $\tilde{O}(de)$ arithmetic operations. This is true as long as both Sylvester matrices are column reduced. From Lemma 2.4 this means that the x-leading (or y-) coefficients of a and b are coprime, and one of them has maximal degree d_a or d_b (respectively e_a or e_b). In a complementary situation, that is with a sufficiently generic ideal $\langle a,b\rangle$ for the graded lexicographic order and using the total degree, a quasi-linear complexity was already obtained by van der Hoeven and Larrieu (2019) for the multiplication in such a quotient.

Following the dichotomic scheme of van der Hoeven and Larrieu, 2018, an efficient reduction algorithm modulo a bivariate lexicographic Gröbner basis is given in (Schost and St-Pierre, 2023a). Finally, although it retains certain assumptions about the ideal, we should also mention the multiplication bound $\tilde{O}((de)^{1.5})$ of (Hyun et al., 2019, Sec. 4.5).

3.4. Power projections: transposed normal form

Using Shoup's general approach for the computation of minimal polynomials in a quotient algebra, we rely in particular on the fact that the power projection problem is the transpose of the modular composition problem (Shoup, 1994; Kaltofen, 2000, Sec. 6). The normal form algorithm of Proposition 3.4 treats a special case of modular composition since f mod I can be seen as $f(g,h)$ mod I for $g = x$ and $h = y$ (see Section 4). Certain power projections can therefore already be derived by transposition from what we have done so far, as we explain in this section. This is used at the core of the general algorithm in Section 4 for the computation of a larger number of $O(de)$ projections efficiently for $\mathbb{K} = \mathbb{F}_q$.

Consider the restriction $\varphi_{\delta,\eta}$ of φ given by Eq. (5) to the $\mathbb{K}$-vector space $\mathcal{U} = \mathbb{K}[x,y]_{\le(\delta,\eta)}$, and denote $\mathbb{K}[x,y]_{<(d,n_y)}$ as a $\mathbb{K}$-vector space by $\mathcal{V}$. We introduce the dual spaces $\widehat{\mathcal{U}}$ and $\widehat{\mathcal{V}}$ of the $\mathbb{K}$-linear forms on $\mathcal{U}$ and $\mathcal{V}$, respectively. The transpose of $\varphi_{\delta,\eta}$ is the $\mathbb{K}$-linear map

$$\begin{aligned} \varphi_{\delta,\eta}^{\mathsf{T}} : \widehat{\mathcal{V}} &\to \widehat{\mathcal{U}} \\ \ell &\mapsto \ell \circ \varphi_{\delta,\eta}. \end{aligned} \tag{6}$$

We view the polynomials in $\mathcal{U}$ as vectors on the basis $\mathcal{B} = \{1, x, \ldots, x^\delta, y, xy, \ldots, x^\delta y^\eta\}$. The linear forms in $\widehat{\mathcal{U}}$ on the dual basis of $\mathcal{B}$ are represented by vectors in $\mathbb{K}^{(\delta+1)(\eta+1)}$. In accordance with the definition of the Sylvester matrix S_y, the elements in $\mathcal{V}$ and $\widehat{\mathcal{V}}$ are viewed in $\mathbb{K}^{dn_y}$ from the basis $\{y^{n_y-1}, y^{n_y-1}x, \ldots, y^{n_y-1}x^{d-1}, y^{n_y-2}, y^{n_y-2}x, \ldots, x^{d-1}\}$ of $\mathcal{V}$. From Eq. (6), the entries of $\varphi_{\delta,\eta}^{\mathsf{T}}(\ell)$ are the bivariate power projections

$$(\ell \circ \varphi_{\delta,\eta})(x^i y^j) \quad \text{for } 0 \le i \le \delta \text{ and } 0 \le j \le \eta. \tag{7}$$

We compute these projections by applying the transposition principle (Bordewijk, 1957, Fiduccia, 1973; Shoup, 1994). The principle states that if a $\mathbb{K}$-linear map $\phi : \mathcal{E}_1 \to \mathcal{E}_2$ can be computed by a linear straight-line program of length l, then the transpose map can be computed by a program of length $l + \dim \mathcal{E}_2$ (l if ϕ is an isomorphism) (Bürgisser et al., 1997, Thm. 13.20). We use a common strategy to implement the principle and refer the reader in particular to Bostan et al. (2003a). Proposition 3.7 follows directly from Proposition 3.4, e.g. by imitating the results of (Bostan et al., 2006, Sec. 4) in $\mathbb{K}[x,y]/\langle f(x), g(y)\rangle$, or from Pascal and Schost (2006) and Poteaux and Schost (2013) modulo triangular sets. The change only concerns the way in which the ideal is represented.

Proposition 3.7. *Assume that the Sylvester matrices S_x and S_y associated with a and b are column reduced. Given two integers $\delta, \eta \ge 0$ and $\ell \in \widehat{\mathcal{V}}$ one can compute $\varphi_{\delta,\eta}^{\mathsf{T}}(\ell)$ using $\tilde{O}((d+\delta)(e+\eta))$ arithmetic operations in $\mathbb{K}$.*

Proof. The claim on $\varphi_{\delta,\eta}^{\mathsf{T}}$ will follow from the application of the transposition principle to the algorithm of Proposition 3.4 for $\varphi_{\delta,\eta}$, with portions written as $\mathbb{K}$-linear straight-line programs.

The computation of $\varphi_{\delta,\eta}$ reduces to two applications of Lemma 3.1, hence it is sufficient to study the transposition of the matrix division with remainder algorithm. Regarding this algorithm, observe that if the matrix inverses such as in Eq. (4) are precomputed, then afterwards only $\mathbb{K}$-linear forms in the entries of the input vector are involved. Moreover, these linear forms can be computed by $\mathbb{K}$-linear straight-line computations. The division algorithm of Lemma 3.1 can therefore be viewed as follows. The inverse of the reversed Sylvester matrix in Eq. (4) is precomputed over truncated power series in time as stated in Lemma 3.1, using the complexity bounds in Section 3.1. Then the linear operations involving the input vector, including structured matrix times vector products (Bini and Pan, 1994), are performed within the same cost bound. The transposed division algorithm follows: it uses the precomputed inverse as a parameter, and it is obtained from the transposition principle applied to the linear straight-line remaining portions. For the transposed division, this leads to the same complexity bound as stated in Lemma 3.1, and for the transpose $\varphi_{\delta,\eta}^{\mathsf{T}}$, to the bound as in Proposition 3.4. □

We have used the transposition principle directly. However, the transpose algorithm could be stated explicitly as was done for the univariate case in (Bostan et al., 2003a)— using duality with linear recurrence sequence extension (Shoup, 1991), and for multivariate triangular sets in (Pascal and Schost, 2006) and (Poteaux and Schost, 2013).

4. Application of Kedlaya & Umans' techniques

Minimal polynomials using the multiplication by x in $\mathbb{A}$ involve projections with $O(de)$ power of x (Section 6). Proposition 3.7 is therefore not sufficient to achieve quasi-linear complexity: with $\delta \in O(de)$ it only gives the cost bound $O(de^2)$. This now leads us to apply the techniques of Kedlaya and Umans, 2011, and their extensions in (Poteaux and Schost, 2013; van der Hoeven and Lecerf, 2021a), for efficient modular composition over a finite field (Corollary 4.2) and power projection by transposition (Corollaries 4.3 and 4.4).

Given three polynomials $f, g, h \in \mathbb{K}[x]$ with $\deg(f) < n$ and $\deg(g) < n$ where $n = \deg(h)$, the problem of modular composition is to compute $f(g) \bmod h$ (Brent and Kung, 1978). (The problem is more fundamentally stated over a ring.) In this case, we benefit from the fact that for such polynomials the division with remainder can be computed using $\tilde{O}(n)$ arithmetic operations (von zur Gathen and Gerhard, 1999, Sec. 9.1). One of the difficulties in the bivariate case is to be able to start from an analogous point, we mean from an efficient division with remainder modulo I. Once this is achieved, the approach of Kedlaya and Umans (2011) can be followed for both modular composition and power projection. This is what has been accomplished in Poteaux and Schost (2013) (multivariate case) and van der Hoeven and Lecerf (2021a) (special case $g = x$), with corresponding forms of the ideal I that we have already mentioned. We proceed in the same way, integrating the new division (normal form) algorithm into the whole process. We therefore do not repeat all the details for the proof of Theorem 4.1 and its corollaries, and refer the reader to the stem papers. As for Proposition 3.7, our change is in the way the ideal is represented, leading to a new modular bivariate projection algorithm in Corollary 4.3.

The first main ingredient is to reduce the problem of division (of modular composition) to divisions with smaller input degrees (Proposition 3.4), and to a problem of multipoint evaluation (Kedlaya and Umans, 2011, Pb. 2.1). Theorem 4.1 shows that the problem of computing the normal form of $f \in \mathbb{K}[x]_{<\delta}$ modulo I can be reduced, for $2 \leq d_\epsilon < \delta$, to normal forms of polynomials of x- and y-degree less than $d_\epsilon d \log \delta$ and $d_\epsilon e \log \delta$, respectively, and to multipoint evaluation. Remember the notation $d = \max\{\deg_x a, \deg_x b\}$ and $e = \max\{\deg_y a, \deg_y b\}$. The Sylvester matrix S_y is $n_y \times n_y$ over $\mathbb{K}$.

Theorem 4.1 ((Kedlaya and Umans, 2011, Thm. 3.1), generalized in (Poteaux and Schost, 2013; van der Hoeven and Lecerf, 2021a))**.** *Consider $f \in \mathbb{K}[x]$ of degree less than δ, and an arbitrary integer $2 \leq d_\epsilon < \delta$. Assume that the Sylvester matrices S_x and S_y associated with a and b are column reduced, and $|\mathbb{K}| > l(d_\epsilon - 1)\max\{d-1, n_y - 1\}$ where $l = \lceil \log_{d_\epsilon}(\delta) \rceil$. If $\delta = O(de)$ then $f(x)$ rem I can be computed using $\tilde{O}(d_\epsilon^2 de)$ arithmetic operations in $\mathbb{K}$, plus one multivariate multipoint evaluation of a polynomial with l variables over $\mathbb{K}$ and individual degrees less than d_ϵ, at $O(l^2 d_\epsilon^2 de)$ points in $\mathbb{K}^l$.*

Proof. The following six steps are those of the proof of (Kedlaya and Umans, 2011, Thm. 3.1).

1. We first appeal to the inverse Kronecker substitution (Kedlaya and Umans, 2011, Dfn. 2.3), in order to map f to a polynomial with l variables and degree less than d_ϵ in each variable. This $\mathbb{K}$-linear map from $\mathbb{K}[x]_{<\delta}$ to $\mathbb{K}[z_0, \ldots, z_{l-1}]_{<(d_\epsilon,\ldots,d_\epsilon)}$ is defined as follow. For $0 \leq k < \delta$, the monomial x^k is sent to $z_0^{k_0} z_1^{k_1} \ldots z_{l-1}^{k_{l-1}}$, where $k_0, k_1, \ldots, k_{l-1}$ are the coefficients of the expansion of k in base d_ϵ. This is extended linearly to $\mathbb{K}[x]_{<\delta}$, and f is mapped in this way to a polynomial $\phi \in \mathbb{K}[z_0, \ldots, z_{l-1}]_{<(d_\epsilon,\ldots,d_\epsilon)}$. The map is injective on $\mathbb{K}[x]_{<\delta}$ and is computed in linear time using monomial bases.
2. Then we compute the polynomials $\chi_i = x^{d_\epsilon^i}$ rem I in $\mathbb{K}[x, y]_{<(d,n_y)}$, for $i = 0, \ldots, l-1$. This corresponds to l exponentiations by d_ϵ modulo I. By successive bivariate multiplications (von zur Gathen and Gerhard, 1999, Sec. 8.4), each followed by a reduction modulo the ideal, this can be done in time $\tilde{O}(de)$ from Proposition 3.4.
   A key property is that $f(x)$ rem $I = \phi(\chi_0, \ldots, \chi_{l-1})$ rem I. This leads to the idea of computing $\phi(\chi_0, \ldots, \chi_{l-1})$ by evaluation-interpolation first, and then do the reduction modulo the ideal. We have that the degree of $\phi(\chi_0, \ldots, \chi_{l-1})$ is at most $\delta' = l(d_\epsilon - 1)(d-1)$ in x, and $\eta' = l(d_\epsilon - 1)(n_y - 1)$ in y.
3. We choose subsets K_1 and K_2 of $\mathbb{K}$ or cardinalities $\delta' + 1$ and $\eta' + 1$, respectively. By multipoint bivariate evaluation, we compute all values $\mu_{i,j,k} = \chi_i(\lambda_j, \lambda_k) \in \mathbb{K}$ for $i = 0, \ldots, l-1$ and $(\lambda_j, \lambda_k) \in K_1 \times K_2$. Using univariate evaluation (von zur Gathen and Gerhard, 1999, Sec. 10.1), variable by variable, this can be done using $\tilde{O}(l\delta'\eta')$ hence $\tilde{O}(d_\epsilon^2 de)$ operations ($n_y \leq 2e$).
4. This is followed by all the evaluations $\phi(\mu_{1,j,k}, \ldots, \mu_{l,j,k})$, which is multipoint evaluation of a polynomial with l variables, with individual degrees less than d_ϵ, at $(\delta' + 1)(\eta' + 1)$ points in $\mathbb{K}^l$.

5. From there, $\phi(\chi_0, \dots, \chi_{l-1})$ is recovered using bivariate interpolation from its values just obtained at $K_1 \times K_2$, this uses $\tilde{O}(d_\epsilon^2 de)$ operations in a way similar to multipoint bivariate evaluation above.
6. Finally, $f(x) \operatorname{rem} I = \phi(\chi_0, \dots, \chi_{l-1}) \operatorname{rem} I$. We know from Proposition 3.4 that this costs $\tilde{O}(\delta' \eta')$ operations, which is $\tilde{O}(d_\epsilon^2 de)$. □

In line with (Kedlaya and Umans, 2011, Thm 7.1) and (Poteaux and Schost, 2013; van der Hoeven and Lecerf, 2021a), thanks to fast multipoint evaluation (Kedlaya and Umans, 2011, Cor. 4.5), the cost of the reduction of a univariate polynomial modulo the ideal can then be bounded.

Corollary 4.2. *Let $\mathbb{K}$ be a finite field $\mathbb{F}_q$. Assume that the Sylvester matrices S_x and S_y associated with a and b are column reduced, and consider $f \in \mathbb{F}_q[x]$ of degree less than $\delta = 4de$. For every constant $\epsilon > 0$, if $q \geq \delta^{1+\epsilon}$, then $f \operatorname{rem} I$ can be computed using $O((de \log q)^{1+\epsilon})$ bit operations.*

Proof. Depending on ϵ, we choose a large enough constant integer c for $d_\epsilon = \lceil \delta^{1/c} \rceil$ to be sufficiently small compared to de. We have in particular, $d_\epsilon < \delta^\epsilon$, and $l = \lceil \log_{d_\epsilon}(\delta) \rceil \leq c$. For δ large enough this leads to $l(d_\epsilon - 1)\max\{d-1, n_y - 1\} \leq q$ and therefore we can apply Theorem 4.1. We know that $f \operatorname{rem} I$ can be computed using $\tilde{O}(d_\epsilon^2 de)$ operations in $\mathbb{F}_q$, which is $O(de)^{1+\epsilon}$, plus the cost of the associated multipoint evaluation. Then we use the fact that for every constant $\gamma > 0$, there is an algorithm for evaluating a polynomial in $\mathbb{F}_q[z_0, \dots, z_{l-1}]_{<(d_\epsilon, \dots, d_\epsilon)}$ at n points in $\mathbb{F}_q^l$ using $(d_\epsilon^l + n)^{1+\gamma} \log(q)^{1+o(1)}$ bit operations, when the individual degrees d_ϵ are sufficiently large, and the number of variables l is at most $d_\epsilon^{o(1)}$ (Kedlaya and Umans, 2011, Cor. 4.5). Considering the evaluation parameters in Theorem 4.1 and $l \leq c$, for every $\gamma > 0$, the cost of the multipoint evaluation then is $O((\delta + d_\epsilon^2 de)^{1+\gamma} \log(q)^{1+o(1)})$, which allows to obtain the claimed complexity bound. □

As said before, our presentation is simplified compared to that of Kedlaya and Umans (2011). The dependence in q, in complexity bounds analogous to those in Corollary 4.2, is made explicit using polylogarithmic functions in Poteaux and Schost (2013). According to (Kedlaya and Umans, 2011, Rem. p. 1790), ϵ can be chosen to be a subconstant function of the other parameters in the complexity bounds. In this respect the study of van der Hoeven and Lecerf (2021a) uses an explicit function of slow increase for the number of variables of the multipoint evaluation problem. Sharper bounds and improved algorithms can be found in (van der Hoeven and Lecerf, 2020; Bhargava et al., 2022) for multipoint evaluation, and (van der Hoeven and Lecerf, 2021b) for multivariate modular composition over finite fields. Since we are also relying on a solution to the dual problem, and since it is not covered in the latter references, we remain based on (Kedlaya and Umans, 2011).

Similar to what we did in Section 3.4 we now transpose the algorithm of Corollary 4.2. The non-algebraic portions of the algorithm involved in multipoint evaluation are treated by means of (Kedlaya and Umans, 2011, Thm. 7.6). Our argument is that of (Kedlaya and Umans, 2011, Thm 7.7), (Poteaux and Schost, 2013, Thm. 3.3) and (van der Hoeven and Lecerf, 2021a, Prop. 1) for modular power projection. We use the φ notation of Eq. (5) for the normal form map.

Corollary 4.3. *Let $\mathbb{K}$ be a finite field $\mathbb{F}_q$. Assume that the Sylvester matrices S_x and S_y associated with a and b are column reduced. Let ℓ be a linear form in the dual of $\mathbb{F}_q[x,y]_{<(d,n_y)}$. For every constant $\epsilon > 0$, if $q \geq \delta^{1+\epsilon}$ with $\delta = 4de$, then $(\ell \circ \varphi)(x^i)$ for $0 \leq i < \delta$ can be computed using $O((de \log q)^{1+\epsilon})$ bit operations.*

Proof. From Eq. (7), we have to compute $\varphi_{\delta-1,0}^{\mathsf{T}}(\ell)$. The claim follows from the transposition principle (Section 3.4) applied to the successive algebraic steps of the normal form algorithm of Corollary 4.2, in reverse order. The non-algebraic portions of the algorithm involved in multipoint evaluation are treated by means of (Kedlaya and Umans, 2011, Thm. 7.6). The steps of the algorithm are given in the proof of Theorem 4.1 (Kedlaya and Umans, 2011, Thm. 3.1). The four of them that depend on the input f have to be considered, these are Steps 1, 4, 5, and 6, that we see as $\mathbb{F}_q$-linear maps. The last step 6 is reduction modulo I, the transpose is obtained from Proposition 3.7. Step 5 is bivariate interpolation, computed by interpolating in x then in y. This is transposed using two transposed univariate interpolation (Kaltofen and Lakshman, 1988; Bostan et al., 2003a). Step 4 is multivariate evaluation using (Kedlaya and Umans, 2011, Cor. 4.5). The transpose is given by (Kedlaya and Umans, 2011, Thm. 7.6) when the ambient dimension is equal to the number of evaluation points, i.e. the linear map can be represented by a square matrix. The general case in which we are, with a larger number of evaluation points, is treated as in the proof of (Kedlaya and Umans, 2011, Thm. 7.7) using several instances of the square case with a cost that fits the claimed bound. Finally, the transpose of the inverse Kronecker substitution at Step 1 is a projection that takes linear time. In view of the transposition principle and of (Kedlaya and Umans, 2011, Thm. 7.6), the algorithm obtained from those transpositions computes the power projections using $O((de)^{1+\epsilon} \log(q)^{1+o(1)})$ bit operations, as in Corollary 4.2. □

For the computation of the minimal polynomial in the presence of roots at infinity (Section 6) we will further need the projections $\{(\ell \circ \varphi)(y^\kappa x^i)\}_{0 \leq i < \delta}$ with $\kappa \in O(\delta)$. Let ρ_κ be the $\mathbb{K}$-linear map which sends $\hat{f} \in \mathbb{K}[x,y]_{<(d,n_y)}$ to $y^\kappa \hat{f}$ rem I. The entries of $(\rho_\kappa \circ \varphi_{\delta-1,0})^{\mathsf{T}}(\ell)$ are $(\ell \circ \varphi)(y^\kappa x^i)$ for $0 \leq i < \delta$. The following is thus a consequence of Corollary 4.3 by introducing ρ_κ^{T}.

Corollary 4.4. *Under the assumptions of Corollary 4.3 and with $\kappa \in O(de)$, the projections $(\ell \circ \varphi)(y^\kappa x^i)$ for $0 \leq i < \delta$ can be computed using $O((de \log q)^{1+\epsilon})$ bit operations.*

Proof. Corollary 4.3 provides us with the means to compute $\varphi_{\delta-1,0}^{\mathsf{T}}(\ell)$. We thus have to compute $\rho_\kappa^{\mathsf{T}}(\ell)$ and apply $(\rho_\kappa \circ \varphi_{\delta-1,0})^{\mathsf{T}} = \varphi_{\delta-1,0}^{\mathsf{T}} \circ \rho_\kappa^{\mathsf{T}}$. We have seen in Section 3.3 that the multiplication in $\mathbb{K}[x,y]/\langle a,b \rangle$ can be implemented using $\tilde{O}(de)$ operations, hence with binary powering the same bound holds for computing $\rho_\kappa(\hat{f}) = (y^\kappa \text{ rem } I) \cdot \hat{f} \text{ rem } I$. To conclude the proof, we apply the transposition principle and obtain the transpose of the latter product in arithmetic time $\tilde{O}(de)$. Alternatively, explicit multivariate polynomial products are detailed in (Bostan et al., 2003b; Pascal and Schost, 2006), which could be combined with Proposition 3.7. □

5. Reversals and random shifts

To use the power projections of Corollary 4.4 to compute the minimal polynomial of the multiplication by x in $\mathbb{A}$ and the last invariant factor of the Smith normal form of S_y, we have to deal with column reducedness issues.

If the input polynomials a and b result in S_x and S_y with singular leading coefficients, then we construct two new polynomials a' and b' which allow us to circumvent the difficulty. We take advantage of the fact that the ideal $\langle a, b\rangle$ is zero-dimensional and construct a' and b' using polynomial shifts and reversals. These transformations are rudimentary and change the structure of the ideal only slightly; the target information for $\langle a, b\rangle$ is efficiently recovered from that computed with a' and b'.

We first modify S_x in Section 5.1. After a random shift with respect to y the constant term of the new Sylvester matrix is not singular, then reversed polynomials ensure column reducedness of a corresponding Sylvester matrix S'_x (Lemma 5.4). The Smith normal form of the associated matrices S_y and S'_y remains unchanged. Therefore, if $a = b = 0$ has no roots at infinity, the minimal polynomial and the last invariant factor can be computed together by Lemma 2.1. If there are roots at infinity, the last invariant factor is preserved, but the minimal polynomials of x modulo $\langle a, b\rangle$ and $\langle a', b'\rangle$ are different (Example 5.3). The minimal polynomial modulo $\langle a, b\rangle$ is obtained by using an appropriate power of y modulo $\langle a', b'\rangle$ (Lemma 5.1).

Similarly, in Section 5.2 we shift and reverse polynomials with respect to x, and we end up with final Sylvester matrices with respect to x and y both column reduced (Proposition 5.6).

5.1. Conditioning of S_x

We recall the notation $\deg_x a = d_a$, $\deg_x b = d_b$, $d = \max\{d_a, d_b\}$, and $\deg_y a = e_a$, $\deg_y b = e_b$, $e = \max\{e_a, e_b\}$. For coprime polynomials a, b and any integer $l \geq 0$, we define the minimal polynomial μ_l of y^l with respect to the multiplication by x to be the monic polynomial of smallest degree in $\mathbb{K}[x]$ such that $y^l \mu_l(x) \in I$. Any $p \in \mathbb{K}[x]$ such that $y^l p(x) \in \langle a, b\rangle$ is a multiple of $\mu_l(x)$.

Lemma 5.1 (Reversals w.r.t y and minimal polynomials)**.** *Consider two coprime polynomials $a, b \in \mathbb{K}[x, y]_{\leq(d,e)}$, and the reversed polynomials $a' = y^{e_a} a(1/y), b' = y^{e_b} b(1/y)$. Assume that a' and b' also have y-degree e_a and e_b, and have coprime y-leading coefficients. If $\kappa \geq 2de$ then the minimal polynomial of y^κ w.r.t to the multiplication by x in $\mathbb{K}[x, y]/\langle a', b'\rangle$ is the minimal polynomial μ of the multiplication by x in $\mathbb{K}[x, y]/\langle a, b\rangle$.*

Lemma 5.1 means that we work with the saturation $\bar{I} = \langle a', b'\rangle : y^\infty$ of $\langle a', b'\rangle$ with respect to y (Cox et al., 2007, Chap. 4, Sec. 4), and that μ is a generator of $\bar{I} \cap \mathbb{K}[x]$.

We first need the following in order to bound the degree of cofactors when $a = b = 0$ has no roots at infinity.

Lemma 5.2 (Lazard, 1985, Lem. 7)**.** *Consider $a, b \in \mathbb{K}[x, y]_{\leq(d,e)}$ with coprime y-leading coefficients. If for some $l \geq e_a + e_b$, $p \in \langle a, b\rangle$ has y-degree less than l, then one can choose $r, s \in \mathbb{K}[x, y]$ with $\deg_y r < l - e_a$ and $\deg_y s < l - e_b$ such that $p = ra + sb$.*

Proof. Take $p = ra + sb$ with $\deg_y r \geq l - e_a$ or $\deg_y s \geq l - e_b$. Then, ra and sb have identical y-degrees and there y-leading coefficients, say $\mathrm{lc}(ra)$ and $\mathrm{lc}(sb)$, cancel. Since the y-leading coefficients $\mathrm{lc}(a)$ and $\mathrm{lc}(b)$ of a and b are coprime, there exists $w \in \mathbb{K}[x]$ such that $\mathrm{lc}(r) = w\mathrm{lc}(b)$ and $\mathrm{lc}(s) = -w\mathrm{lc}(a)$. The polynomials $r_1 = r - wby^{\deg_y r - e_b}$ and $s_1 = s + way^{\deg_y s - e_a}$ have y-degree at most $\deg_y r - 1$ and $\deg_y s - 1$, respectively. In addition, using that $\deg_y r + e_a = \deg_y s + e_b$, we obtain $r_1 a + s_1 b = p - waby^{\deg_y r - e_b} + waby^{\deg_y s - e_a} = p$. Following Lazard, this gives the result by induction. □

Example 5.3. *(Example 3.6 cont., $\mathbb{K} = \mathbb{F}_7$). The y-leading coefficients of a and b have a non-trivial gcd and Lemma 5.2 do not apply for the minimal polynomial $\mu = x + 2 \in \langle a, b\rangle$. Following Lemma 5.1, take $a' = x+3+(x^2+5x+5)y$ and $b' = (x+3)(x+4)+(x^2+4x+2)y$, and work modulo $\langle a', b'\rangle$. The new minimal polynomial of x is $(x+2)(x+3)^3$ and the minimal polynomial of y^3 w.r.t. x is $x + 2$.* □

Proof of Lemma 5.1. We note first that for $\kappa \geq 0$, the minimal polynomial μ'_κ of y^κ with respect to the multiplication by x in $\mathbb{K}[x,y]/\langle a', b'\rangle$ is a multiple of μ. Then we show that μ is a multiple of μ'_κ if κ is large enough, and we conclude by showing that the latter is true as soon as the degree of $\langle a', b'\rangle$ is reached.

If $y^\kappa p(x) \equiv 0 \bmod \langle a', b'\rangle$ with $\kappa \geq e_a + e_b - 1$, then we can write $y^\kappa p(x) = ra' + sb'$ with $r, s \in \mathbb{K}[x,y]$ whose degrees are bounded as in the conclusion of Lemma 5.2 ($l = \kappa + 1$). This leads to

$$(y^\kappa (1/y)^\kappa)\, p(x) = \left(y^{l-e_a-1} r(1/y)\right)\left(y^{e_a} a'(1/y)\right) + \left(y^{l-e_b-1} s(1/y)\right)\left(y^{e_b} b'(1/y)\right),$$

which is also $p(x) = r'a + s'b$ for some $r', s' \in \mathbb{K}[x,y]$. The minimal polynomial μ'_κ is thus a multiple of μ for $\kappa \geq e_a + e_b - 1$, hence for any $\kappa \geq 0$ ($\mu'_{\kappa-1}$ is a multiple of μ'_κ).

Conversely, since $\mu \in \langle a, b\rangle$, there exist $u, v \in \mathbb{K}[x,y]$ such that $\mu = ua + vb$. Using polynomial reversals with respect to y, it follows that $y^\kappa \mu(x) \in \langle a', b'\rangle$ for some $\kappa \geq 0$. From what we have seen so far, this implies that μ is a generator of $\bar{I} \cap \mathbb{K}[x]$ where $\bar{I} = \langle a', b'\rangle : y^\infty$ is the saturation of $\langle a', b'\rangle$ with respect to y.

Now, for $\kappa \geq 0$, consider the quotient $Q_\kappa = \langle a', b'\rangle : y^\kappa$. The ideals $\{Q_k\}_{k\geq 0}$ form an ascending chain; as soon as $Q_{\kappa_0} = Q_{\kappa_0+1}$ for some κ_0, we know that $Q_{\kappa_0} = \bar{I}$ (Cox et al., 2007, Chap. 4, Sec. 4, Exe. 8). Before this happens, say for $0 < \kappa < \kappa_0$, since Q_κ is strictly contained in $Q_{\kappa+1}$, the quotient algebras $\mathbb{K}[x,y]/Q_\kappa$ have decreasing dimensions as k increases. (See e.g. (Berthomieu et al., 2022a, Lem. 3.4) and (Cox et al., 2007, Chap. 5, Sec. 3, Prop. 4).) Since the dimension of Q_0 is at most $2de$, we know that $Q_\kappa = \bar{I}$ for $\kappa \geq 2de$. This allows us to conclude the proof since $Q_\kappa \cap \mathbb{K}[x] = \bar{I} \cap \mathbb{K}[x]$ for $\kappa \geq 2de$ gives that the minimal polynomial of y^κ with respect to x is μ for $\kappa \geq 2de$. □

According to the notation $\mathbb{A} = \mathbb{K}[x,y]/\langle a, b\rangle$, we now denote $\mathbb{K}[x,y]/\langle a', b'\rangle$ by $\mathbb{A}'$.

Lemma 5.4 (Conditioning of S_x)**.** *Given $\alpha \in \mathbb{K}$ not a root of $\mathrm{Res}_x(a, b)$ (the ideal is zero-dimensional), in arithmetic time $\tilde{O}(de)$ we can compute two polynomials a' and b' with degrees as those of a and b, such that:*

- *the new Sylvester matrix S'_x is column reduced;*

- *for κ as in Lemma 5.1, the minimal polynomial of y^κ w.r.t to the multiplication by x in $\mathbb{A}'$ is the minimal polynomial of the multiplication by x in $\mathbb{A}$;*
- *the Smith normal form of S_y' is that of S_y.*

Proof. Consider $a^{(1)}(x,y) = a(x, y+\alpha)$ and $b^{(1)}(x,y) = b(x, y+\alpha)$. The new Sylvester matrix $S_x^{(1)}$ with respect to x has a non-singular constant term since $(\mathrm{Res}_x(a,b))(\alpha) \neq 0$. The minimal polynomials of the multiplication by x in $\mathbb{A}$ and in $\mathbb{K}[x,y]/\langle a^{(1)}, b^{(1)}\rangle$ are identical. The Smith normal form of $S_y^{(1)}$ is equal to the Smith normal form of S_y. Indeed, let $Q_{\alpha,k} \in \mathbb{K}^{k\times k}$ be the matrix of the endomorphism that shifts a polynomial of degree less than k by α; $Q_{\alpha,k}$ is lower triangular with unit diagonal. We have

$$S_y^{(1)} = Q_{\alpha, e_a+e_b}\, S_y\, \mathrm{diag}(Q_{\alpha,e_b}^{-1}, Q_{\alpha,e_a}^{-1}), \tag{8}$$

hence $S_y^{(1)}$ and S_y are unimodularly equivalent.

Then we consider the reversed polynomials a' and b' of $a^{(1)}$ and $b^{(1)}$ with respect to y, using the respective degrees e_a and e_b. Note that a' and b' must keep the same y-degrees, otherwise $S_x^{(1)}$ could not have a non-singular constant coefficient. For the same reason, the new matrix S_x' associated with a' and b' is column reduced, which proves the first claim.

For the second claim we can apply Lemma 5.1 with $a^{(1)}, b^{(1)}$ since a' and b' have appropriate degrees and their leading coefficients are coprime by Lemma 2.4. As noted previously, the minimal polynomial of x in $\mathbb{A}$ is obtained.

The last claim follows from the fact that the Smith form with respect to y is unchanged:

$$S_y' = J_{e_a+e_b}\, S_y^{(1)}\, \mathrm{diag}(J_{e_b}, J_{e_a}), \tag{9}$$

where J_k is the reversal matrix of dimension k. The cost is dominated by the one of at most $2(d+1)$ shifts of polynomials of degree at most e in $\mathbb{K}[y]$, see e.g. (Bini and Pan, 1994)[Chap. 1, Pb. 3.5]. □

Lemma 5.4 preserves the Smith normal form of S_y but not necessarily its Hermite form. From Lemma 2.4, $a' = b' = 0$ has no roots at infinity with respect to y, so the last invariant factor of S_y is the minimal polynomial of the multiplication by x in the new quotient algebra (Lemma 2.1). The latter may have changed, with an additional factor coming from possible roots at infinity for $a = b = 0$. So Lemma 5.4 is useful for computing the minimal polynomial of x if $a = b = 0$ has roots at infinity (second claim). Otherwise, with Lemma 2.1, we can simply rely on the last invariant factor computation (third claim).

5.2. Conditioning of both S_x and S_y

We now do the same kind of manipulation as in Section 5.1 for the column reducedness of S_y and need a preliminary observation about reversed polynomial matrices. The reversal by columns of a matrix polynomial is the matrix whose entries are reversed with respect to the degree of their column.

Lemma 5.5 (Reversed Smith normal form)**.** *The last invariant factor of the reversal of $A \in \mathbb{K}[x]^{n\times n}$ by columns is the reversal of the last invariant factor of A made monic and multiplied by some power of x.*

Proof. Let X in $\mathbb{K}[x]^{n\times n}$ with a determinant which is a power of x be such that the reversal R of A by columns is $A(1/x)X$. Let S_A be the Smith normal form of A, with unimodular matrices U and V such that $AV = US_A$. We have

$$RX^{-1}V(1/x) = U(1/x)S_A(1/x). \tag{10}$$

Let S_A^* be the diagonal matrix whose entries are the reversals of the diagonal entries of S, made monic by division by their leading coefficients. By multiplying Eq. (10) by an appropriate power of x, we obtain

$$RW_1 = W_2S_A^*$$

for two matrices W_1 and W_2 in $\mathbb{K}[x]^{n\times n}$ whose determinants are powers of x. Now let S_R^* be the diagonal matrix whose diagonal entries are the invariant factors of R divided by the largest power of x they contain. Using similar manipulations as above we get

$$S_R^*W_3 = W_4S_A^*$$

for two matrices W_3 and W_4 in $\mathbb{K}[x]^{n\times n}$ whose determinants are also powers of x. By the multiplicativity of the Smith normal form (Newman, 1972, Chap. II, Thm. 2.15), noting that S_A^* and S_R^* are themselves in Smith normal form, we arrive at $S_R^* = S_A^*$. The claim follows since the last invariant factor of R is the one of S_R^* multiplied by some power of x, and the last invariant factor of S_A^* is the reversal of the last invariant factor of A divided by its leading coefficient. □

Proposition 5.6 (Conditioning of S_x and S_y)**.** *Given $\alpha, \beta \in \mathbb{K}$ not roots of* $\mathrm{Res}_x(a,b)$ *and* $\mathrm{Res}_y(a,b) \in \mathbb{K}[x]$, *respectively (the ideal is zero-dimensional), in arithmetic time $\tilde{O}(de)$ we can compute two polynomials a' and b' with degrees as those of a and b such that:*

- *the new Sylvester matrices S'_x and S'_y are column reduced;*
- *for κ as in Lemma 5.1, the minimal polynomial of the multiplication by x in $\mathbb{A}$ can be deduced from the one of y^κ w.r.t to the multiplication by x in $\mathbb{A}'$, using $\tilde{O}(de)$ additional arithmetic operations;*
- *the last invariant factor of S_y can be deduced from that of S'_y using $\tilde{O}(de)$ additional arithmetic operations.*

Proof. For the purpose of the proof we use the notation a_0 and b_0 for the initial polynomials (a and b in the statement), and let now a and b denote the polynomials after an application of Lemma 5.4. We can thus assume that a and b are such that S_x is column reduced, without modifying the Smith form of the Sylvester matrix and the resultant with respect to y. We denote the last invariant factor of S_y by $\sigma \in \mathbb{K}[x]$. For κ as in Lemma 5.1, we can also assume that the minimal polynomial $\mu \in \mathbb{K}[x]$ of y^κ w.r.t to the multiplication by x in $\mathbb{K}[x,y]/\langle a,b\rangle$ is that of x in $\mathbb{K}[x,y]/\langle a_0,b_0\rangle$.

We use arguments similar to those used in the proof of Lemma 5.4. First we take $a^{(1)}(x,y) = a(x+\beta,y)$ and $b^{(1)}(x,y) = b(x+\beta,y)$. The new Sylvester matrix $S_y^{(1)}$ with respect to y has a non-singular constant term since $(\mathrm{Res}_y(a,b))(\beta) \neq 0$. The last invariant factor of $S_y^{(1)}$ is $\sigma_\beta = \sigma(x+\beta)$ and satisfies $\sigma_\beta(0) \neq 0$. The new minimal polynomial of y^κ with respect to the multiplication by x is $\mu_\beta = \mu(x+\beta)$. We also have $\mu_\beta(0) \neq 0$ by Lemma 2.1 since μ_β divides the minimal polynomial of x as well as σ_β.

Then we consider the reversed polynomials a' and b' of $a^{(1)}$ and $b^{(1)}$ with respect to x, using the respective degrees d_a and d_b. Since $S_y^{(1)}$ has a non-singular constant term, a' and b' retain the same x-degrees and the new matrix S'_y associated with a' and b' is column reduced.

We now prove the claims with a' and b'. We have just seen for the column reducedness of the y-Sylvester matrix. With respect to x, the Sylvester matrix is column reduced after the initial application of Lemma 5.4. Using Eqs. (8) and (9) from the proof of the latter lemma, now with β, $S_x^{(1)}$, and S'_x, we deduce that S'_x remains column reduced and we have proved the first claim.

Let μ'_β be the reversal of μ_β with respect to x, and $x^{l_1}\lambda'$ be the minimal polynomial of y^κ with respect to x in $\mathbb{A}'$ for some $l_1 \geq 0$ and $\lambda' \in \mathbb{K}[x]$ such that $\lambda'(0) \neq 0$. For the second claim we need to show how to obtain μ from $x^{l_1}\lambda'$. Writing that $y^\kappa \mu_\beta = ua + vb$ with $u, v \in \mathbb{K}[x,y]$, and reverting the polynomials, we deduce that $y^\kappa x^{l_2} \mu'_\beta$ is in $\langle a', b' \rangle$ for some $l_2 \geq 0$. So μ'_β is a multiple of λ'. Conversely, $y^\kappa x^{l_1} \lambda' \in \langle a', b' \rangle$ gives that λ' is a multiple of μ'_β, hence $\lambda' = c\mu'_\beta$ for a non-zero $c \in \mathbb{K}$. Since $\lambda'(0) \neq 0$, the reversal of $x^{l_1}\lambda' = cx^{l_1}\mu'_\beta$ is $c\mu_\beta$. Using a shift of $-\beta$ and making the polynomial monic gives us μ.

Finally, we compute the last invariant factor σ of S_y from that of S'_y. The Sylvester matrix S'_y is the reversal by columns of $S_y^{(1)}$ (entries reversed with respect to the degree of their column). Let σ'_β the reversal of σ_β. From Lemma 5.5 we deduce that the last invariant factor of S'_y is $cx^l\sigma'_\beta$ for some integer $l \geq 0$, and a non-zero $c \in \mathbb{K}$. Since $\sigma_\beta(0) \neq 0$, the reversal of $cx^l\sigma'_\beta$ is $c\sigma_\beta$. The invariant factor σ is derived as done above for μ.

In addition to the cost in Lemma 5.4, we essentially have to perform at most $2(e+1)$ shifts of polynomials of degree at most d in $\mathbb{K}[x]$ for having a' and b'; final shifts of polynomials of degree $O(de)$ give μ and σ (Bini and Pan, 1994)[Chap. 1, Pb. 3.5]. □

6. Elimination ideal and invariant factor computation

Recall that φ is the $\mathbb{K}$-linear map that defines the normal form in $\mathbb{A}$ (Proposition 3.4). For a random linear form ℓ in the dual of $\mathbb{F}_q[x,y]_{<(d,n_y)}$ and an integer $\kappa \geq 0$, the minimal polynomial μ_κ of y^κ with respect to the multiplication by x in $\mathbb{K}[x,y]/\langle a,b \rangle$ is also the one of the linearly generated sequence $(\ell \circ \varphi)(y^\kappa x^i)_{i \geq 0}$ with high probability (Shoup, 1995, Sec. 4; Kaltofen and Shoup, 1998, Lem. 6). In essence, this minimal polynomial approach is a transcription of that of Wiedemann, 1986, with multiplication matrices rather than sparse ones (Shoup, 1994). This allows us in this section to bound the complexity of the minimal polynomial problem and of the last invariant factor problem from the power projection complexity bound obtained earlier (Corollary 4.4). Since the minimal polynomials we consider have degree at most $2de$, they can be computed in fact from the first $4de$ terms of the power projection sequences.

Lemma 6.1. *Consider two polynomials $a, b \in \mathbb{F}_q[x,y]_{\leq(d,e)}$ and assume that the associated Sylvester matrices S_x and S_y are column reduced. For every constant $\epsilon > 0$, there exists a randomized Monte Carlo algorithm which computes the minimal polynomial of y^κ with respect to the multiplication by x in $\mathbb{F}_q[x,y]/\langle a,b\rangle$ using $O((de\log q)^{1+\epsilon})$ bit operations if $q \geq \delta^{1+\epsilon}$ with $\delta = 4de$ and $\kappa \in O(de)$. The algorithm returns a divisor of the minimal polynomial, to which it is equal with probability at least $1 - 2de/q \geq 1/2$.*

Proof. The modular power projections as in Corollary 4.4 are computed for a random linear map ℓ. The sequence $\{(\ell\circ\varphi)(y^\kappa x^i)\}_{i\geq 0}$ is linearly generated; its minimal polynomial $\tilde{\mu}_\kappa$ is a divisor of the minimal polynomial μ_κ of y^κ with respect to the multiplication by x in $\mathbb{A}$. Since $\deg \mu_\kappa \leq 2de$, $\tilde{\mu}_\kappa$ can be computed using $\tilde{O}(de)$ additional operations in $\mathbb{K}$ from the $4de$ first terms of the sequence (von zur Gathen and Gerhard, 1999, Algo 12.9). We can conclude by proving that $\tilde{\mu}_\kappa = \mu_\kappa$ with high probability. Following the construction of φ in Eq. (5), one can define the multiplication map $\psi : \mathbb{F}_q[x,y]_{<(d,n_y)} \to \mathbb{F}_q[x,y]_{<(d,n_y)}; f \mapsto xf \text{ rem } I$. For an appropriate basis of $\mathbb{F}_q[x,y]_{<(d,n_y)}$ as a $\mathbb{F}_q$-vector space, we consider that ψ is represented by a $(dn_y)\times(dn_y)$ matrix M over $\mathbb{F}_q$ and that $y^\kappa \bmod I$ is represented by the vector $u = \varphi(y^\kappa) \in \mathbb{F}_q^{dn_y}$. According to what we have seen in Section 3.4, we also represent linear forms in the dual of $\mathbb{F}_q[x,y]_{<(d,n_y)}$ by vectors in $\mathbb{F}_q^{dn_y}$. With this, μ_κ is the minimal polynomial of u with respect to M. Hence for a random linear form ℓ represented by $v \in \mathbb{F}_q^{dn_y}$, the minimal polynomial of the linearly generated sequence $\{(\ell\circ\varphi)(y^\kappa x^i)\}_{i\geq 0} = \{v^\intercal M^i u\}_{k\geq 0}$ is μ_κ with probability at least $1 - \deg\mu_\kappa/q$ (Kaltofen and Pan, 1991, Lem. 2; Kaltofen and Saunders, 1991, Lem. 1). □

Lemma 6.1 is only valid if the Sylvester matrices involved are column reduced. From the random shifts and reversals seen in Section 5, we now compute the minimal polynomial of the multiplication by x or y for arbitrary coprime a and b.

Theorem 6.2. *Consider two coprime polynomials a, b in $\mathbb{F}_q[x,y]_{\leq(d,e)}$. There exists a randomized Monte Carlo algorithm which computes the minimal polynomial of the multiplication by either x or y in $\mathbb{K}[x,y]/\langle a,b\rangle$, using $(de\log q)^{1+o(1)}$ bit operations. The algorithm either returns the target minimal polynomial, and this with probability at least $1/2$, one of its divisors, or "failure".*

Proof. Given $\epsilon > 0$, when $q \geq (12de)^{1+\epsilon}$, we choose random α and β in $\mathbb{F}_q$, then check whether S'_x and S'_y as in Proposition 5.6 are column reduced. This is performed using $\tilde{O}(d+e)$ operations, see Lemma 2.4 and e.g. (von zur Gathen and Gerhard, 1999, Thm. 11.10). Since $\mathrm{Res}_y(a,b) \in \mathbb{F}_q[x]$ and $\mathrm{Res}_x(a,b) \in \mathbb{F}_q[y]$ have degree at most $2de$, the probability of success is at least $1-4de/q$. If the Sylvester matrices are column reduced, from Lemma 6.1 and with κ as in Lemma 5.1, we compute the minimal polynomial of y^κ (or x^κ) w.r.t the multiplication by x (or y) in the quotient algebra associated with S'_y and S'_x. The probability of success is at least $1 - 2de/q$. From Proposition 5.6 again, we finally derive the minimal polynomial of x (or y) in $\mathbb{A}$. If q is too small, we construct an extension field of $\mathbb{F}_q$ with cardinality at least $(12de)^{1+\epsilon}$, that is of degree $O(\log(de))$. This can be done using an expected number of $\tilde{O}((\log(de)^2 + \log(de)\log(q))$ bit operations Shoup (1994) (see also Couveignes and Lercier (2013) and (von zur Gathen and Gerhard, 1999, Sec. 14.9) in this regard). We then work in this extension, the costs induced are logarithmic factors which do not change our target cost bound, and the probability of success can be adjusted. According to Kedlaya and Umans, 2011, Rem. p. 1790 and the comments in Section 4, ϵ can be chosen to be a subconstant function of the other parameters in the complexity bound of Lemma 6.1. □

The following is proved in the same way as for Theorem 6.2, using the third rather than the second assertion in Proposition 5.6.

Corollary 6.3. *Consider two coprime polynomials a, b in $\mathbb{F}_q[x,y]_{\le(d,e)}$. There exists a randomized Monte Carlo algorithm which computes the last invariant factor of the Sylvester matrix associated with a and b with respect to either x or y, using $(de\log q)^{1+o(1)}$ bit operations. The algorithm either returns the target invariant factor, and this with probability at least $1/2$, one of its divisors, or "failure".*

Proof. We do not repeat all the proof of Theorem 6.2 but simply point out that if the modified Sylvester matrices S'_x and S'_y are column reduced, from Lemma 6.1, we compute the minimal polynomial of the multiplication by x (or y) in the quotient algebra associated with S'_y and S'_x. Then Lemma 2.1 tells us that we have actually computed the last invariant factor of S'_y (or S'_x), and from Proposition 5.6 we obtain the last invariant factor of S_y (or S_x). □

7. Resultant

For general a and b we only compute a specific factor of the resultant, which is the last invariant factor of the Sylvester matrix.

When the system $a = b = 0$ has no roots at infinity with respect to y, Lemma 2.2 indicates that if, moreover, the ideal has a shape basis $I = \langle \mu(x), y - \lambda(x)\rangle$ (Gianni et al., 1988; Becker et al., 1994), then the resultant is known from the minimal polynomial. Note that the extra non-zero constant in Lemma 2.2 can be computed at the cost of $\tilde{O}(de)$ operations in $\mathbb{F}_q$ using evaluation in x. We see that this leads to a weaker genericity assumption than in (van der Hoeven and Lecerf, 2021a), where the total degree is used. Suppose that the ideal $\langle a, b\rangle$ is in generic position for the lexicographic order $y > x$, so that $\mathrm{Res}_y(a,b) = c\mu$ with $c \neq 0 \in \mathbb{F}_q$. In this case we can compute the resultant in quasi-linear time without the use of an additional condition with respect to the graded reverse lexicographic order (van der Hoeven and Lecerf, 2021a).

This also allows us to deal with more general situations than that of the total degree: we obtain the resultant in all cases where $\mathrm{Res}_y(a,b) = c\mu$. This condition is sufficient but not necessary (Example 3.6), the resultant is computed when S_y has a unique non-trivial invariant factor. The latter property can be formalized in the Zariski sense, e.g. by relying on ideals in general position without roots at infinity (Cox et al., 2005, Chap. 3, Sec.5). More precisely, there exists a non-zero polynomial Φ in $2(d+1)(e+1)$ variables over $\mathbb{K}$, such that the Smith form of S_y has a unique non-trivial invariant factor if Φ does not vanish at the coefficients of a and b. The generic resultant algorithm becomes of the Las Vegas type when the degree of the resultant is known in advance, especially when the Sylvester matrix S_y is column reduced. In the latter case the degree of the resultant is actually the sum of the column degrees of S_y (Kailath, 1980, Eq. (24), p. 385).

Whether the resultant can be computed in quasi-linear time for arbitrary a and b is an interesting question.

References

Alonso, M.E., Becker, E., Roy, M.F., Wörmann, T., 1996. Zeros, multiplicities, and idempotents for zero-dimensional systems, in: Algorithms in Algebraic Geometry and Applications. PM vol. 143, Birkhaüser, pp. 1–15. doi:10.1007/978-3-0348-9104-2_1.

Becker, E., Mora, T., Marinari, M.G., Traverso, C., 1994. The shape of the Shape Lemma, in: ISSAC'94: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 129–133. doi:10.1145/190347.190382.

Beckermann, B., Labahn, G., 1994. A Uniform Approach for the Fast Computation of Matrix-Type Padé Approximants. SIAM J. Matrix Analysis and Applications 15, 804–823. doi:10.1137/S0895479892230031.

Berthomieu, J., Eder, C., Safey El Din, M., 2022a. New efficient algorithms for computing Gröbner bases of saturation ideals (F4SAT) and colon ideals (Sparse-FGLM-colon). arXiv:2202.13387. doi:10.48550/arXiv.2202.13387.

Berthomieu, J., Neiger, V., Safey El Din, M., 2022b. Faster Change of Order Algorithm for Gröbner Bases under Shape and Stability Assumptions, in: ISSAC'22: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 409–418. doi:10.1145/3476446.3535484.

Bhargava, V., Ghosh, S., Guo, Z., Kumar, M., Umans, C., 2022. Fast Multivariate Multipoint Evaluation Over All Finite Fields, in: IEEE 63rd Annual Symposium on Foundations of Computer Science, IEEE. pp. 221–232. doi:10.1109/FOCS54457.2022.00028.

Bini, D., Pan, V.Y., 1994. Polynomial and Matrix Computations. Birkhäuser. doi:10.1007/978-1-4612-0265-3.

Bordewijk, J.L., 1957. Inter-reciprocity applied to electrical networks. Appl. Sci. Res. B 6, 1–74. doi:10.1007/BF02920362. (Ph. D. thesis, Technische Hogeschool Delft, 1956).

Bostan, A., Flajolet, P., Salvy, B., Schost, E., 2006. Fast computation of special resultants. J. Symb. Comput. 41, 1–29. doi:10.1016/j.jsc.2005.07.001.

Bostan, A., Lecerf, G., Schost, É., 2003a. Tellegen's principle into practice, in: ISSAC'03: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 37–44. doi:10.1145/860854.860870.

Bostan, A., Salvy, B., Schost, É., 2003b. Fast Algorithms for Zero-Dimensional Polynomial Systems using Duality. Appl. Algebr. Eng. Comm. 14, 239–272. doi:10.1007/s00200-003-0133-5.

Brent, R.P., Kung, H.T., 1978. Fast algorithms for manipulating formal power series. J. ACM 25, 581–595. doi:10.1145/322092.322099.

Bürgisser, P., Clausen, M., Shokrollahi, M.A., 1997. Algebraic complexity theory. volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer. doi:10.1007/978-3-662-03338-8.

Coppersmith, D., Odlzyko, A.M., Schroeppel, R., 1986. Discrete logarithms in GF(p). Algorithmica 1, 1–15. doi:10.1007/BF01840433.

Couveignes, J.M., Lercier, R., 2013. Fast construction of irreducible polynomials over finite fields. Isr. J. Math. 194, 77–105. doi:10.1007/s11856-012-0070-8.

Cox, D.A., D'Andrea, C., 2023. Subresultants and the Shape Lemma. Math. Comput. doi:10.1090/mcom/3840.

Cox, D.A., Little, J., O'Shea, D., 2005. Using Algebraic Geometry. Springer. doi:10.1007/b138611. Second Edition.

Cox, D.A., Little, J., O'Shea, D., 2007. Ideals, varieties, and algorithms. Springer. doi:10.1007/978-3-319-16721-3. Third Edition.

Dahan, X., 2022. Lexicographic Gröbner bases of bivariate polynomials modulo a univariate one. J. Symb. Comput. 110, 24–65. doi:10.1016/j.jsc.2021.10.001.

Dixon, J.D., 1982. Exact solution of linear equations using p-adic expansions. Numer. Math. 40, 137–141. doi:10.1007/BF01459082.

Faugère, J.C., Gaudry, P., Huot, L., Renault, G., 2014. Sub-cubic change of ordering for Gröbner basis: a probabilistic approach, in: ISSAC'14: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 170–177. doi:10.1145/2608628.2608669.

Faugère, J.C., Mou, C., 2011. Fast algorithm for change of ordering of zero-dimensional Gröbner bases with sparse multiplication matrices, in: Proc. ISSAC. ACM Press, pp. 115–122. doi:10.1145/1993886.1993908.

Faugère, J.C., Mou, C., 2017. Sparse FGLM algorithms. J. Symb. Comput. 80, 538–569. doi:10.1016/j.jsc.2016.07.025.

Fiduccia, C.M., 1973. On the algebraic complexity of matrix multiplication. Ph.D. thesis. Brown University. URL: https://bruknow.library.brown.edu/permalink/01BU_INST/9mvq88/alma991027307479706966.

von zur Gathen, J., Gerhard, J., 1999. Modern Computer Algebra. Cambridge University Press. doi:10.1017/CBO9781139856065. third edition 2013.

Gianni, P., Trager, B., Zacharias, G., 1988. Gröbner bases and primary decomposition of polynomial ideals. J. Symb. Computation 6, 149–167. doi:10.1016/S0747-7171(88)80040-3.

Gonzalez-Vega, L., Rouillier, F., Roy, M.F., 1999. Symbolic Recipes for Polynomial System Solving, in: Algorithms and Computation in Mathematics, Some Tapas of Computer Algebra. Springer, pp. 34–65. doi:10.1007/978-3-662-03891-8_2.

van der Hoeven, J., 2017. On the Complexity of Multivariate Polynomial Division, in: ACA 2015: Applications of Computer Algebra, Springer, PROMS 198. pp. 447–458. doi:10.1007/978-3-319-56932-1_28.

van der Hoeven, J., Larrieu, R., 2018. Fast reduction of bivariate polynomials with respect to sufficiently regular Gröbner bases, in: ISSAC'18: Proceedings of the International Symposium on Symbolic and Algebraic Computation. ACM Press, pp. 199–206. doi:10.1145/3208976.3209003.

van der Hoeven, J., Larrieu, R., 2019. Fast Gröbner basis computation and polynomial reduction for generic bivariate ideals. Appl. Algebr. Eng. Comm. 30, 509–539. doi:10.1007/s00200-019-00389-9.

van der Hoeven, J., Lecerf, G., 2020. Fast multivariate multi-point evaluation revisited. J. of Complexity 56. doi:10.1016/j.jco.2019.04.001.

van der Hoeven, J., Lecerf, G., 2021a. Fast computation of generic bivariate resultants. J. of Complexity 62. doi:10.1016/j.jco.2020.101499.

van der Hoeven, J., Lecerf, G., 2021b. On the Complexity Exponent of Polynomial System Solving. Found. Comput. Math. 21, 1–57. doi:10.1007/s10208-020-09453-0.

Hyun, S.G., Melczer, S., Schost, É., St-Pierre, C., 2019. Change of basis for m-primary ideals in one and two variables, in: ISSAC'19: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 227–234. doi:10.1145/3326229.3326268.

Jacobson, N., 2009. Basic Algebra I. Dover Publications Inc. URL: https://store.doverpublications.com/0486471896.html. Second Edition W.H. Freeman 1985.

Kailath, T., 1980. Linear Systems. Prentice-Hall.

Kailath, T., Kung, S.Y., Morf, M., 1979. Displacement ranks of matrices and linear equations. J. Math. Anal. Appl. 68, 395–407. doi:10.1016/0022-247X(79)90124-0.

Kaltofen, E., 1994. Asymptotically fast solution of Toeplitz-like singular linear systems, in: ISSAC'94: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 297–304. doi:10.1145/190347.190431.

Kaltofen, E., 2000. Challenges of Symbolic Computation: My Favorite Open Problems. J. Symbolic Computation 29, 891–919. doi:10.1006/jsco.2000.0370.

Kaltofen, E., Lakshman, Y., 1988. Improved Sparse Multivariate Polynomial Interpolation Algorithms, in: Proc. ISSAC, Springer, LNCS 358. pp. 467–474. doi:10.1007/3-540-51084-2_44.

Kaltofen, E., Pan, V.Y., 1991. Processor efficient parallel solution of linear systems over an abstract field, in: SPAA '91: Proceedings of the third annual ACM symposium on Parallel algorithms and architectures, ACM. pp. 180–191. doi:10.1145/113379.113396.

Kaltofen, E., Saunders, D., 1991. On Wiedemann's method of solving sparse linear systems, in: AAECC 1991: Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Springer Verlag. pp. 29–38. doi:10.1007/3-540-54522-0_93.

Kaltofen, E., Shoup, V., 1998. Subquadratic-time factoring of polynomials over finite fields. Mathematics of Computation 67, 1179–1197. doi:10.1090/S0025-5718-98-00944-2.

Kedlaya, K.S., Umans, C., 2011. Fast Polynomial Factorization and Modular Composition. SIAM J. on Computing 40, 1767–1802. doi:10.1137/08073408X.

Labahn, G., 1992. Inversion components of block Hankel-like matrices. Linear Algebra Appl. 177, 7–48. doi:10.1016/0024-3795(92)90316-3.

Lazard, D., 1981. Résolution des systèmes d'équations algébriques. Theor. Comput. Sci. 15, 77–110. doi:10.1016/0304-3975(81)90064-5.

Lazard, D., 1985. Ideal Bases and Primary Decomposition: Case of Two Variables. J. Symb. Comput. 1, 261–270. doi:10.1016/S0747-7171(85)80035-3.

Lazard, D., 1992. Solving zero-dimensional algebraic systems. J. Symb. Comput. 13, 117–131. doi:10.1016/S0747-7171(08)80086-7.

Lebreton, R., Mehrabi, E., Schost, É., 2013. On the complexity of solving bivariate systems: the case of non-singular solutions, in: ISSAC'13: Proceedings of the International Symposium on Symbolic and Algebraic Computation, pp. 251–258. doi:10.1145/2465506.2465950.

Lecerf, G., 2019. On the complexity of the Lickteig-Roy subresultant algorithm. J. Symb. Comput. 92, 243–268. doi:10.1016/j.jsc.2018.04.017.

Mantzaflaris, A., Rahkooy, H., Zafeirakopoulos, Z., 2016. Efficient computation of dual space and directional multiplicity of an isolated point. Comput. Aided Geom. Des. 47, 114–129. doi:10.1016/J.CAGD.2016.05.002.

Moenck, R.T., Carter, J.H., 1979. Approximate algorithms to derive exact solutions to systems of linear equations, in: Proc. EUROSAM, pp. 65–73. doi:10.1007/3-540-09519-5_60.

Mourrain, B., Pan, V.Y., 2000. Multivariate Polynomials, Duality, and Structured matrices. J. of Complexity 16, 110–180. doi:10.1006/jcom.1999.0530.

Neiger, V., Salvy, B., É. Schost, Villard, G., 2023. Faster Modular Composition. J. ACM URL: https://arxiv.org/abs/2110.08354. To appear.

Newman, M., 1972. Integral Matrices. Academic Press. First edition.

Pan, V.Y., 2001. Structured Matrices and Polynomials: Unified Superfast Algorithms. Springer. doi:10.1007/978-1-4612-0129-8.

Pascal, C., Schost, É., 2006. Change of order for bivariate triangular sets, in: ISSAC'06: Proceedings of the International Symposium on Symbolic and Algebraic Computation. ACM Press, pp. 277–284. doi:10.1145/1145768.1145814.

Pernet, C., Signargout, H., Villard, G., 2024. High-order lifting for polynomial Sylvester matrices. J. of Complexity 80. doi:10.1016/j.jco.2023.101803.

Poteaux, A., Schost, É., 2013. Modular Composition Modulo Triangular Sets and Applications. Comput. Complex. 22, 463–516. doi:10.1007/s00037-013-0063-y.

Rifà, J., Borrell, J., 1991. Improving the time complexity of the computation of irreducible and primitive polynomials in finite fields, in: AAECC 1991: Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, pp. 352–359. doi:10.1007/3-540-54522-0_123.

Rouillier, F., 1999. Solving Zero-Dimensional Systems Through the Rational Univariate Representation. Appl. Algebra Eng. Commun. Comput. 9, 433–461. doi:10.1007/s002000050114.

Schost, É., St-Pierre, C., 2023a. Newton iteration for lexicographic Gröbner bases in two variables. arXiv:2302.03766. doi:10.48550/arxiv.2302.03766.

Schost, É., St-Pierre, C., 2023b. p-adic algorithm for bivariate Gröbner bases, in: ISSAC'23: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 508–516. doi:10.1145/3597066.3597086.

Shoup, V., 1991. A Fast Deterministic Algorithm for Factoring Polynomials over Finite Fields of Small Characteristic, in: ISSAC'91: Proceedings of the International Symposium on Symbolic and Algebraic Computation, pp. 14–21. doi:10.1145/120694.120697.

Shoup, V., 1994. Fast Construction of Irreducible Polynomials over Finite Fields. J. Symb. Comput. 17, 371–391. doi:10.1006/jsco.1994.1025.

Shoup, V., 1995. A New Polynomial Factorization Algorithm and its Implementation. J. Symb. Comput. 20, 363–397. doi:10.1006/jsco.1995.1055.

Shoup, V., 1999. Efficient computation of minimal polynomials in algebraic extensions of finite fields, in: ISSAC'99: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 53–58. doi:10.1145/309831.309859.

Storjohann, A., 2000. Algorithms for Matrix Canonical Forms. Ph.D. thesis. Institut für Wissenschaftliches Rechnen, ETH-Zentrum, Zurich, Switzerland. URL: https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/145127/eth-24018-02.pdf.

Thiong Ly, J.A., 1989. Note for computing the minimun polynomial of elements in large finite fields, in: Coding Theory 1988: Coding Theory and Applications, pp. 185–192. doi:10.1007/BFb0019856.

Villard, G., 1997. Fast Parallel Algorithms for Matrix Reduction to Normal Forms. Appl. Algebra Eng. Commun. Comput. 8, 511–537. doi:10.1007/s002000050089.

Villard, G., 2018. On Computing the Resultant of Generic Bivariate Polynomials, in: ISSAC'18: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 391–398. doi:10.1145/3208976.3209020.

Villard, G., 2023. Elimination ideal and bivariate resultant over finite fields, in: ISSAC'23: Proceedings of the International Symposium on Symbolic and Algebraic Computation, ACM Press. pp. 526–534. doi:10.1145/3597066.3597100.

Wiedemann, D., 1986. Solving sparse linear equations over finite fields. IEEE Trans. Information Theory 32, 54–62. doi:10.1109/TIT.1986.1057137.