Improved Dense Multivariate Polynomial Factorization Algorithms

Grégoire Lecerf

Laboratoire de Mathématiques, UMR 8100 CNRS
Université de Versailles Saint-Quentin-en-Yvelines
45 avenue des États-Unis, 78035 Versailles, France

Abstract

We present new deterministic and probabilistic algorithms that reduce the factorization of dense polynomials from several to one variable. The deterministic algorithm runs in sub-quadratic time in the dense size of the input polynomial, and the probabilistic algorithm is softly optimal when the number of variables is at least three. We also investigate the reduction from several to two variables and improve the quantitative version of Bertini's irreducibility theorem.

Key words: Polynomial factorization, Hensel lifting, Bertini's irreducibility theorem.

Introduction

The factorization of multivariate polynomials is a classical problem in computer algebra, which intervenes in many fields of application. So far no softly optimal algorithm is known. In this article we propose new faster methods for reducing this factorization to one or two variables.

Let $\mathbb{K}$ be a commutative field. Throughout this article F denotes a polynomial in $\mathbb{K}[z_1, \ldots, z_n, y]$, of total degree $d := \deg(F)$ such that the following hypothesis holds:

Hypothesis (C) $\mathbb{K}$ has characteristic 0 or at least $d(d-1)+1$.

We are interested in the complexity of computing the irreducible factors $F_1, \ldots, F_r$ of F. We use the *dense representation* for the polynomials, which means that a polynomial of

⋆ This work was supported in part by the French Research Agency (ANR Gecko).
Email address: Gregoire.Lecerf@math.uvsq.fr (Grégoire Lecerf).
URL: http://www.math.uvsq.fr/~lecerf (Grégoire Lecerf).

Preliminary version of January 23, 2005. Revised version: 18 January 2007.

total degree d is stored as the vector of its coefficients in the basis of the monomials of degree at most d. We shall often use the quantity

$$N_{d,n} := \binom{d+n}{n}$$

to represent the number of monomials in n variables of degree at most d. In particular the size of F equals $N_{d,n+1}$. Under Hypothesis (C) it is always possible to suppose that F is squarefree. Thus, up to a linear change of variables, we can assume that the following hypothesis holds, without loss of generality:

$$\text{Hypothesis (H)} \quad \begin{cases} (i)\ F \text{ is monic in } y \text{ and } \deg_y(F) = d, \\ (ii)\ \operatorname{Res}\left(F(0,\ldots,0,y), \frac{\partial F}{\partial y}(0,\ldots,0,y)\right) \neq 0, \end{cases}$$

where $\deg_y(F)$ represents the partial degree of F in the variable y. Here $\operatorname{Res}(A,B)$ denotes the resultant of two univariate polynomials A and B. Under the latter hypothesis, we apply the *lifting and recombination* technique, popularized by Zassenhaus (1969), in order to compute the factorization of F. This technique can be made very efficient for bivariate polynomials, as demonstrated by Bostan et al. (2004) and Lecerf (2006). One of the main goals of this article is to generalize the results of Lecerf (2006) to several variables. Mistakes in the two latter references are corrected in Appendix A.

Main Results

The first section of this article is devoted to the deterministic and probabilistic reductions to one variable. We use and generalize the lifting and recombination algorithm of Lecerf (2006). More precisely, the deterministic reduction algorithm is presented in Section 1.2 and proceeds as follows:

(1) Factor the univariate polynomial $F(0,\ldots,0,y)$.
(2) *Lift* the resulting factors in order to obtain the irreducible factorization of F in $\mathbb{K}[[z_1,\ldots,z_n]][y]$ to precision $(z_1,\ldots,z_n)^{2d}$, where $\mathbb{K}[[z_1,\ldots,z_n]]$ represents the power series algebra in n variables. Here $(z_1,\ldots,z_n)^{2d}$ represents the $2d$th power of the maximal ideal $(z_1,\ldots,z_n)$.
(3) Solve a linear system in order to determine how the lifted factors *recombine* into the true factors.

For fixed n, we show that the costs of the second and third steps are sub-quadratic with respect to the dense size of F.

The probabilistic reduction algorithm is given in Section 1.3. It starts with the reduction to two variables by means of substituting random linear forms in the new variable x for the z_i. We say that $(\alpha_1,\ldots,\alpha_n) \in \mathbb{K}^n$ is a *Bertinian good point* for F if $F_i(\alpha_1 x,\ldots,\alpha_n x, y)$ is irreducible for all i. In other words, the irreducible factors of F are in one-to-one correspondence with those of $H(x,y) := F(\alpha_1 x,\ldots,\alpha_n x, y) \in \mathbb{K}[x,y]$. From a practical point of view, the knowledge of a Bertinian good point naturally gives rise to the following algorithm:

(1) Factor $H(x,y)$.
(2) Lift the resulting factors in order to recover $F_1,\ldots,F_r$.

Of course the factorization of $H(x,y)$ can be handled by any algorithm (probabilistic or not), but we show that the use of the probabilistic recombination algorithm of Appendix A.1 leads to a softly optimal reduction to one variable as soon as $n \geq 2$.

In practice, the point $(\alpha_1, \ldots, \alpha_n)$ is chosen with coordinates in a finite subset of $\mathbb{K}$. In order to estimate the probability of success of this probabilistic reduction, we need to upper bound the density of Bertinian bad points. This is the purpose of Section 2.1, where we provide a nearly optimal bound.

Lastly, Section 2.2 is devoted to a new quantitative Bertini irreducibility theorem. There we do not work under Hypothesis (H) anymore. In order to avoid confusion we consider a polynomial P of degree $d \geq 1$ in the variables $v_1, \ldots, v_n$ over $\mathbb{K}$. For any points $(\alpha_1, \ldots, \alpha_n)$, $(\beta_1, \ldots, \beta_n)$ and $(\gamma_1, \ldots, \gamma_n)$ in $\mathbb{K}^n$, we define the bivariate polynomial $P_{\alpha,\beta,\gamma}$ in the variables x and y by:

$$P_{\alpha,\beta,\gamma} := P(\alpha_1 x + \beta_1 y + \gamma_1, \ldots, \alpha_n x + \beta_n y + \gamma_n). \tag{1}$$

According to the classical Bertini irreducibility theorem (e.g. Shafarevich, 1994, Chapter II, Section 6.1) and if P is irreducible, then there exists a proper Zariski open subset of $(\mathbb{K}^n)^3$ such that $P_{\alpha,\beta,\gamma}$ is irreducible for any triple $(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n), (\gamma_1, \ldots, \gamma_n)$ in this subset. We say that, for any irreducible factor Q of P, such a triple in $(\mathbb{K}^n)^3$ is a *Bertinian good point* for P if $Q(\alpha_1 x + \beta_1 y + \gamma_1, \ldots, \alpha_n x + \beta_n y + \gamma_n)$ is irreducible with the same total degree of Q. In other words, the irreducible factors of P are in one-to-one correspondence with those of $P_{\alpha,\beta,\gamma}$. The complementary set of Bertinian good points is written $\mathfrak{B}(P)$ and is called the set of *Bertinian bad points*.

For algorithmic purposes, the entries of $(\alpha_1, \ldots, \alpha_n)$, $(\beta_1, \ldots, \beta_n)$ and $(\gamma_1, \ldots, \gamma_n)$ must be taken in a finite subset S of $\mathbb{K}$, so that we are naturally interested in upper bounding the number of Bertinian bad points in $(S^n)^3$. We refer to such a bound as a *quantitative Bertini theorem*. The density of Bertinian bad points with entries in a non-empty finite subset S of $\mathbb{K}$ is defined by:

$$\mathfrak{B}(P, S) := \frac{|\mathfrak{B}(P) \cap (S^n)^3|}{|S|^{3n}},$$

where $|S|$ represents the cardinality of S. At the end of this article we show that $\mathfrak{B}(P, S) \leq 3d^2/|S|$ (see Corollary 8), which improves the previously known bounds, under Hypothesis (C).

Working under Hypothesis (H) is interesting from several points of view. Most polynomials satisfy this hypothesis, so that the substitution of $\alpha_i x$ for z_i is more efficient than the one of (1). In particular, the former preserves the sparsity whereas the latter does not. Hypothesis (H) is naturally satisfied in the *geometric resolution algorithm* for solving algebraic systems (see Lecerf, 2003). Thus, for any equidimensional algebraic closed set encoded by a *lifting fiber*, our Corollary 7 of Section 2.1 can be directly applied in order to bound the density of associated *lifting curves* which preserve the irreducible decomposition.

Related Works

Works on polynomial factorization are too numerous to be all cited here. Several aspects are treated in the following references: Kaltofen (1982a), Zippel (1993), Schinzel (2000), von zur Gathen and Gerhard (2003), and Chèze and Galligo (2005). Historical surveys can be found in: Kaltofen (1990, 1992, 1995, 2003), and Gao (2003). The first polynomial time multivariate factorization algorithm is due to Kaltofen (1982b,c, 1985c). Then Chistov, von zur Gathen, Grigoriev, Kaltofen and A. K. Lenstra contributed to this suject. An important breakthrough has been accomplished by Gao (2003) who designed

a quadratic time probabilistic reduction from two to one variable for the first time. Then, Bostan et al. (2004) and Lecerf (2006) proposed faster reductions: a deterministic one with sub-quadratic cost and a probabilistic one with a cost in $\tilde{\mathcal{O}}(d^3)$ (see the errata in Appendix A).

Reduction to One Variable

The present work is closely connected to previous results of Heintz and Sieveking (1981), von zur Gathen (1985), and Kaltofen (1985a,c, 1995). Compared to Kaltofen's methods, our main gain is essentially due to using a precision linear in d during the lifting stage instead of a quadratic precision. Briefly speaking, Kaltofen's quadratic precision (1985c) comes from using algebraic approximant algorithms, whereas our linear precision is proved in (Lecerf, 2006) thanks to Ruppert's ideas (1986; 1999). Ruppert's original idea relies on considering the first algebraic de Rham cohomology group of $\mathbb{K}[z, y, 1/F(z, y)]$ (here $n = 1$ and we let $z = z_1$): if $\mathbb{K}$ is algebraically closed and has characteristic 0, then

$$\left(\frac{\hat{F}_i \frac{\partial F_i}{\partial z}}{F} dz + \frac{\hat{F}_i \frac{\partial F_i}{\partial y}}{F} dy \right)_{i \in \{1,\ldots,r\}}$$

is basis of this group, where $\hat{F}_i := \frac{F}{F_i}$ (see Ruppert, 1986, Satz 2). In consequence, this group can be obtained by searching for closed differential 1-forms with denominators F and numerators of degrees at most $d - 1$. As shown by Gao (2003), this computation boils down to linear algebra and is still valid for sufficiently large positive characteristics.

It is worth to mention special cases for which specific methods exist. Over finite fields, Kaltofen (1987) and Gao et al. (2004) have shown how to test the irreducibility and even to count the number of factors in a deterministic way. When $\mathbb{K} = \mathbb{Q}$, Kaltofen (1985c) has given a specific deterministic reduction from several to two variables. Searching for the factors in the algebraic closure of $\mathbb{K}$ is called the *absolute factorization*: the absolute factorization can be computed by a polynomial time deterministic algorithm with operations in $\mathbb{K}$ alone. Advanced results can be found in: Kaltofen (1995), Gao (2003), Chèze (2004), Chèze and Galligo (2005), Chèze and Lecerf (2005). Finally, concerning other polynomial representations, such as straight-line program, circuit, black box and sparse representations, the reader can consult: von zur Gathen (1985), von zur Gathen and Kaltofen (1985a,b), Kaltofen (1989), Kaltofen and Trager (1990).

Reduction to Two Variables

What we call "Bertini's theorem" in this paper is a particular but central case of more general theorems such as in (Shafarevich, 1994, Chapter II, Section 6.1). We refer the reader to Kleiman's survey (1998) on Bertini's life and mathematical work, and to Jouanolou's book (1983) for an extensive mathematical treatment. As pointed out by Kaltofen (1995), the particular case of Bertini's theorem that only concerns the reduction of the factorization problem from several to two variables goes back at least to Hilbert (1892, p. 117). This is the reason why some authors say "Hilbert's theorem" instead of "Bertini's theorem".

Bertini's theorem was popularized in complexity theory by Heintz and Sieveking (1981), and Kaltofen (1982b). A few years later, Bertini's theorem became a cornerstone of many factorization or reduction techniques including: Kaltofen (1985a,b,c,d), von zur Gathen (1985), von zur Gathen and Kaltofen (1985b). For any characteristic and under Hypothesis (H), von zur Gathen (1985) showed that the set of Bertinian bad

points for F is included in a proper hypersurface of degree at most 9^{d^2}. This bound is to be compared to the one of our Theorem 6 of Section 2.1. When $\mathbb{K}$ is the field of complex numbers, Bajaj et al. (1993) obtained the bound $\mathfrak{B}(P,S) \leq (d^4 - 2d^3 + d^2 + d + 1)/|S|$ by following Mumford's proof (1995, Theorem 4.17) of Bertini's theorem. This proof starts with reducing to Hypothesis (H). For any perfect field $\mathbb{K}$, Kaltofen (1995) proved that $\mathfrak{B}(P,S) \leq 2d^4/|S|$ by using his factorization algorithm. If $\mathbb{K}$ has characteristic 0 or larger than $2d^2$, Gao (2003) proved the sharper bound $\mathfrak{B}(P,S) \leq 2d^3/|S|$. He made use of his factorization algorithm adapted from Ruppert's theorems (1986; 1999). Recently, Chèze has pointed out (2004, Chapter 1) that the latter bound can even be refined to $\mathfrak{B}(P,S) \leq d(d^2-1)/|S|$ by using directly (Ruppert, 1986, Satz C). A nice presentation of Ruppert's results is made in Schinzel's book (2000, Chapter V).

Complexity Model

For our complexity analysis, we use the *computation tree* model (see Bürgisser et al., 1997, Chapter 4) from the *total complexity* point of view. Roughly speaking, this means that complexity estimates charge a constant cost for each arithmetic operation ($+$, $-$, $\times$, $\div$) and the equality test. Yet all the constants in the base fields (or rings) of the trees are thought to be freely at our disposal. Univariate factorization algorithms fall outside this model. Therefore, for convenience, we enlarge the model with a univariate factorization algorithm. We use the classical $\mathcal{O}$ and $\tilde{\mathcal{O}}$ (read "*soft Oh*") notation in the neighborhood of infinity as defined in (von zur Gathen and Gerhard, 2003, Chapter 25.7). Informally speaking, "soft Oh"s are used for readability in order to hide logarithmic factors in complexity estimates.

For each integer d, we assume that we are given a computation tree that computes the products of two polynomials of degree at most d with at most $\mathsf{M}(d)$ operations, independently of the base ring. As in (von zur Gathen and Gerhard, 2003, Chapter 8.3), for any positive integers d_1 and d_2, we assume that M satisfies: $\mathsf{M}(d_1 d_2) \leq d_1^2 \mathsf{M}(d_2)$ and $\mathsf{M}(d_1)/d_1 \leq \mathsf{M}(d_2)/d_2$ if $d_1 \leq d_2$. In particular, this implies the *super-additivity* of M, that is $\mathsf{M}(d_1) + \mathsf{M}(d_2) \leq \mathsf{M}(d_1 + d_2)$. We recall that the resultant and the extended greatest common divisor of two univariate polynomials of degree at most d over $\mathbb{K}$ can be computed with $\mathcal{O}(\mathsf{M}(d)\log(d))$ operations in $\mathbb{K}$ (von zur Gathen and Gerhard, 2003, Chapter 11). Series are thought to be represented by dense vectors of their coefficients in the usual monomial basis. We assume that, for each d and n, we are given a computation tree that computes the product of two power series over $\mathbb{K}$ in n variables, truncated in total degree d, and that performs at most $\mathsf{S}(d,n)$ operations in $\mathbb{K}$. In addition, we assume that S is super-additive with respect to d, that is: $\mathsf{S}(d_1,n) + \mathsf{S}(d_2,n) \leq \mathsf{S}(d_1+d_2,n)$ for any positive integers d_1 and d_2. If $\mathbb{K}$ has characteristic 0 then the algorithm presented by Lecerf and Schost (2003) allows us to take $\mathsf{S}(d,n) \in \tilde{\mathcal{O}}(N_{d-1,n})$, which is softly optimal with respect to the dense size of the series to be multiplied. For any characteristic and any truncation, we expect that softly optimal algorithms may exist: recent advances in this direction have been made by van der Hoeven (2004, 2005) and Schost (2005).

Lastly, we assume that, for each n, we are given a computation tree that computes the product of two $n \times n$ matrices over $\mathbb{K}$ with at most $\mathcal{O}(n^\omega)$ field operations, for a fixed constant ω. We require that $2 < \omega \leq 3$ in order to use (Storjohann, 2000, Theorem 2.10) later. In contrast to polynomials, we only deal with matrices over $\mathbb{K}$.

1. Reduction to One Variable

We carry on with the notation of the introduction: F denotes a polynomial of degree d which satisfies Hypotheses (C) and (H). Recall that the irreducible factors of F are denoted by $F_1, \ldots, F_r$. Without loss of generality we assume that $F_1, \ldots, F_r$ are monic in y. Let $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ denote the monic irreducible factors of F in $\mathbb{K}[[z_1, \ldots, z_n]][y]$. Under Hypothesis (H), and because of the Hensel lemma, $\mathfrak{F}_i$ remains irreducible when substituting $0, \ldots, 0$ for $z_1, \ldots, z_n$, for all $i \in \{1, \ldots, s\}$. To each $i \in \{1, \ldots, r\}$, we associate the vector $\mu_i \in \{0,1\}^s$, defined by

$$F_i = \prod_{j=1}^{s} \mathfrak{F}_j^{\mu_{i,j}}. \tag{2}$$

Since the μ_i have entries in $\{0,1\}$ and have pairwise disjoint supports, we can assume that they form a reduced echelon basis, without loss of generality.

1.1. Theoretical Reduction to Two Variables

We introduce the set of auxiliary variables $a_1, \ldots, a_n$ and the polynomial

$$G := F(a_1 x, \ldots, a_n x, y) \in \mathbb{K}_a[x, y], \text{ where } \mathbb{K}_a := \mathbb{K}(a_1, \ldots, a_n),$$

on which we are going to apply the deterministic reduction algorithm of Lecerf (2006). The polynomial G is monic in y when seen in $\mathbb{K}[a_1, \ldots, a_n, x][y]$, thus its irreducible factors in $\mathbb{K}_a[x, y]$ are in one-to-one correspondence to those of F. In other words, the irreducible factors of G are the $G_i(x, y) := F_i(a_1 x, \ldots, a_n x, y)$, for $i \in \{1, \ldots, r\}$. It is straightforward to check that Hypothesis (H) implies:

$$(\mathrm{H}_a) \quad \begin{cases} (i) \ \deg_y(G) = \deg(G) = d, \\ (ii) \ \mathrm{Res}\left(G(0,y), \frac{\partial G}{\partial y}(0,y)\right) \neq 0. \end{cases}$$

We introduce the irreducible factors $\mathfrak{G}_1, \ldots, \mathfrak{G}_s$ of G in $\mathbb{K}_a[[x]][y]$, which are related to the $\mathfrak{F}_i$ by $\mathfrak{G}_i(x, y) = \mathfrak{F}_i(a_1 x, \ldots, a_n x, y)$, for all $i \in \{1, \ldots, s\}$. It follows that $\mathfrak{G}_i$ belongs to $\mathbb{K}[a_1, \ldots, a_n][[x]][y]$. Furthermore the coefficient of $x^j y^k$ in $\mathfrak{G}_i$ is either 0 or homogeneous of degree j. As a direct consequence of (2), we observe that μ_i satisfies and is uniquely determined by $G_i = \prod_{j=1}^{s} \mathfrak{G}_j^{\mu_{i,j}}$. Lastly, we introduce

$$\hat{\mathfrak{F}}_i := \prod_{j=1, j\neq i}^{s} \mathfrak{F}_j \quad \text{and} \quad \hat{\mathfrak{G}}_i := \prod_{j=1, j\neq i}^{s} \mathfrak{G}_j, \text{ for all } i \in \{1, \ldots, s\}.$$

1.2. Deterministic Reduction Algorithm

We are now ready to apply the deterministic recombination algorithm of (Lecerf, 2006, Section 3) to G, that makes use of the following linear system over $\mathbb{K}_a$ in the unknowns $(\ell_1, \ldots, \ell_s)$:

$$D_{a,\sigma} \begin{cases} \displaystyle\sum_{i=1}^{s} \ell_i \, \mathrm{coeff}\left(\hat{\mathfrak{G}}_i \frac{\partial \mathfrak{G}_i}{\partial y}, x^j y^k\right) = 0, \ k \leq d-1, \ d \leq j+k \leq \sigma - 1, \\ \displaystyle\sum_{i=1}^{s} \ell_i \, \mathrm{coeff}\left(\hat{\mathfrak{G}}_i \frac{\partial \mathfrak{G}_i}{\partial x}, x^j y^k\right) = 0, \ k \leq d-1, \ j \leq \sigma - 2, \ d \leq j+k \leq \sigma - 1, \end{cases}$$

where coeff$(G, x^i y^j)$ represents the coefficient of the monomial $x^j y^k$ in G, and where σ denotes a positive integer. Since (H) implies (H_a), the combination of (Lecerf, 2006, Theorem 1 and Lemma 4) implies:

Lemma 1 *Under Hypotheses (C) and (H), for any $\sigma \geq 2d$, the reduced echelon solution basis of $D_{a,\sigma}$ is $\mu_1, \ldots, \mu_r$.*

Let us recall here that the condition $\sigma \geq 2d$ in Lemma 1 is a consequence of Ruppert's irreducibility test (1986; 1999). If we applied this lemma directly, we would be led to solve $D_{a,\sigma}$ over $\mathbb{K}_a$, which is very expensive. In the next lemma, we show that this computation can be avoided by means of solving the linear system over $\mathbb{K}$ instead of $\mathbb{K}_a$. Moreover we show that the resolution can be performed over any subfield $\mathbb{E}$ of $\mathbb{K}$. In particular and when possible, the use of the prime field of $\mathbb{K}$ is expected to yield a practical speed-up in the resolution.

Lemma 2 *Let $\mathbb{E}$ be a subfield of $\mathbb{K}$. For any $\sigma \geq 2d$, the reduced echelon basis of the restriction to $\mathbb{E}^s$ of the solution set of $D_{a,\sigma}$ is $\mu_1, \ldots, \mu_r$.*

Proof. Since the entries of the μ_i are in $\{0, 1\}$, one has $\mu_i \in \mathbb{E}^s$, hence μ_i is a solution of $D_{a,\sigma}$ over $\mathbb{E}$. Let $(\ell_1, \ldots, \ell_s) \in \mathbb{E}^s$ denote a solution of $D_{a,\sigma}$. According to the previous lemma, there exists $(\gamma_1, \ldots, \gamma_r) \in \mathbb{K}_a^r$ such that $(\ell_1, \ldots, \ell_s) = \gamma_1\mu_1 + \cdots + \gamma_r\mu_r$. Since $\mu_1, \ldots, \mu_r$ form a reduced echelon basis it follows that $\gamma_i \in \mathbb{E}$, for all $i \in \{1, \ldots, r\}$. □

For the sake of efficiency, we wish to avoid the substitution of the $a_i x$ for the z_i, and to obtain a direct generalization of (Lecerf, 2006, Section 3). For this purpose, we introduce the differential operator

$$\theta := z_1 \frac{\partial}{\partial z_1} + \cdots + z_n \frac{\partial}{\partial z_n},$$

which is easy to compute by means of the following formula:

$$\theta(z_1^{j_1} \cdots z_n^{j_n}) = (j_1 + \cdots + j_n) z_1^{j_1} \cdots z_n^{j_n}.$$

We write $\bar{j} := j_1 + \cdots + j_n$, and consider the new linear system defined by:

$$D_\sigma \begin{cases} \displaystyle\sum_{i=1}^{s} \ell_i \,\mathrm{coeff}\left(\hat{\mathfrak{F}}_i \frac{\partial \mathfrak{F}_i}{\partial y}, z_1^{j_1} \cdots z_n^{j_n} y^k\right) = 0,\ k \leq d-1,\ d \leq \bar{j} + k \leq \sigma - 1, \\ \displaystyle\sum_{i=1}^{s} \ell_i \,\mathrm{coeff}\left(\hat{\mathfrak{F}}_i \theta \mathfrak{F}_i, z_1^{j_1} \cdots z_n^{j_n} y^k\right) = 0,\ k \leq d-1,\ \bar{j} \leq \sigma - 1,\ d+1 \leq \bar{j} + k \leq \sigma. \end{cases}$$

Lemma 3 *Under Hypotheses (C) and (H), and for any $\sigma \geq 2d$, the reduced echelon solution basis of D_σ is $\mu_1, \ldots, \mu_r$.*

Proof. By Lemma 2, it remains to verify that the solutions of D_σ coincide with the solutions of $D_{a,\sigma}$ in $\mathbb{K}^s$. From $\hat{\mathfrak{G}}_i \dfrac{\partial \mathfrak{G}_i}{\partial y} = \left(\hat{\mathfrak{F}}_i \dfrac{\partial \mathfrak{F}_i}{\partial y}\right)(a_1 x, \ldots, a_n x, y)$ we deduce that coeff $\left(\hat{\mathfrak{G}}_i \dfrac{\partial \mathfrak{G}_i}{\partial y}, x^j y^k\right)$ equals the homogeneous component of degree j of the coefficient of

y^k in $(\hat{\mathfrak{F}}_i \frac{\partial \mathfrak{F}_i}{\partial y})(a_1, \ldots, a_n, y)$ seen in $\mathbb{K}[[a_1, \ldots, a_n]][y]$. Thus the first subsets of equations of $D_{a,\sigma}$ and D_σ coincide over $\mathbb{K}^s$. On the other hand, a basic calculation gives:

$$\begin{aligned} x\frac{\partial \mathfrak{G}_i}{\partial x} &= x\frac{\partial}{\partial x}\Big(\mathfrak{F}_i(a_1 x, \ldots, a_n x, y)\Big) \\ &= \sum_{j=1}^{n} x a_j \frac{\partial \mathfrak{F}_i}{\partial z_j}(a_1 x, \ldots, a_n x, y) = (\theta \mathfrak{F}_i)(a_1 x, \ldots, a_n x, y), \end{aligned}$$

from which we deduce that $\text{coeff}\left(\hat{\mathfrak{G}}_i \frac{\partial \mathfrak{G}_i}{\partial x}, x^j y^k\right) = \text{coeff}\left(\hat{\mathfrak{G}}_i x \frac{\partial \mathfrak{G}_i}{\partial x}, x^{j+1} y^k\right)$ equals the homogeneous component of degree $j+1$ of the coefficient of y^k in $(\hat{\mathfrak{F}}_i \theta \mathfrak{F}_i)(a_1, \ldots, a_n, y)$ seen in $\mathbb{K}[[a_1, \ldots, a_n]][y]$. Finally, the second subsets of equations of $D_{a,\sigma}$ and D_σ also coincide over $\mathbb{K}^s$. $\square$

Based on Lemma 3, the factorization algorithm proceeds as follows:

Algorithm 1 Deterministic factorization algorithm.
Input: *F of total degree d satisfying Hypotheses (C) and (H).*
Output: *the irreducible factors $F_1, \ldots, F_r$ of F.*

(1) Compute $\mathfrak{F}_1(0, \ldots, 0, y), \ldots, \mathfrak{F}_s(0, \ldots, 0, y)$ as the irreducible factors of the univariate polynomial $F(0, \ldots, 0, y)$.

(2) Lifting step. *Call a fast multi-factor Hensel lifting algorithm in order to obtain $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to precision $(z_1, \ldots, z_n)^\sigma$ with $\sigma := 2d$.*

(3) Recombination step.

- *(a) For each $i \in \{1, \ldots, s\}$ compute $\hat{\mathfrak{F}}_i$ as the quotient of F by $\mathfrak{F}_i$ to precision $(z_1, \ldots, z_n)^\sigma$ in $\mathbb{K}[[z_1, \ldots, z_n]][y]$.*
- *(b) Compute $(\hat{\mathfrak{F}}_1 \frac{\partial \mathfrak{F}_1}{\partial y}, \ldots, \hat{\mathfrak{F}}_s \frac{\partial \mathfrak{F}_s}{\partial y})$ to precision $(z_1, \ldots, z_n)^\sigma$.*
- *(c) Compute $(\hat{\mathfrak{F}}_1 \theta \mathfrak{F}_1, \ldots, \hat{\mathfrak{F}}_s \theta \mathfrak{F}_s)$ to precision $(z_1, \ldots, z_n)^\sigma$.*
- *(d) Build the linear system D_σ and compute its reduced echelon solution basis $\mu_1, \ldots, \mu_r$.*
- *(e) If $r = 1$ then return F. Otherwise, for each i in $\{1, \ldots, r\}$, compute F_i as $\prod_{j=1}^{s} \mathfrak{F}_j^{\mu_{i,j}}$ to precision $(z_1, \ldots, z_n)^{\deg(F_i)+1}$, and return $F_1, \ldots, F_r$.*

Proposition 4 *Algorithm 1 is correct and performs one factorization of a univariate polynomial of degree d over $\mathbb{K}$ plus a number of operations in $\mathbb{K}$ belonging to:*

$$\mathcal{O}\Big(s\mathsf{S}(\sigma, n)\mathsf{M}(d) + dN_{\sigma-1,n} s^{\omega-1}\Big). \tag{3}$$

Proof. The correctness follows from Lemma 3. It remains to analyze the costs of steps (2) and (3). A fast multi-factor Hensel lifting algorithm is given in (von zur Gathen and Gerhard, 2003, Algorithm 15.17) for I-adic topologies when I is a principal ideal. Here $I := (z_1, \ldots, z_n)$ is not principal but this algorithm still applies. From the complexity point of view, we must take care to perform the last step of the lifting to precision σ and not to the next power of 2 of σ. Subject to this slight modification and thanks to the super-additivity of S, the cost of step (2) follows *mutatis mutandis* from (von zur Gathen and Gerhard, 2003, part (ii) of Theorem 15.18): it belongs to $\mathcal{O}((\mathsf{S}(\sigma, n) + \log(d))\mathsf{M}(d)\log(s)) \subseteq$

$\mathcal{O}(\mathsf{S}(\sigma,n)\mathsf{M}(d)\log(s))$. A slightly faster (by a constant factor) lifting algorithm is described in (Bostan et al., 2004, Section 3) but the same modifications are necessary to deal with multivariate power series.

The total cost of steps (3a), (3b) and (3c) clearly belongs to $\mathcal{O}(s\mathsf{S}(\sigma,n)\mathsf{M}(d))$. The construction of D_σ is negligible. Since D_σ has s unknowns and less than $2dN_{\sigma-1,n}$ equations, the cost of step (3d) belongs to $\mathcal{O}(dN_{\sigma-1,n}s^{\omega-1})$ by (Storjohann, 2000, Theorem 2.10). The computation of F_i can benefit of the sub-product tree technique of (von zur Gathen and Gerhard, 2003, Algorithm 10.3). Thus, by (von zur Gathen and Gerhard, 2003, Lemma 10.3) we deduce that each F_i can be computed in time $\mathcal{O}(\mathsf{S}(\deg(F_i)+1,n)\mathsf{M}(\deg(F_i))\log(s_i)))$, where $s_i := \sum_{j=1}^{s}\mu_{i,j}$ represents the number of lifted factors involved in F_i. Thanks to the super-additivities, the cost of step (3e) drops to $\mathcal{O}(\mathsf{S}(d,n)\mathsf{M}(d)\log(s))$. □

Since a power series in n variables to precision σ has dense size $N_{\sigma-1,n}$, one necessarily has $\mathsf{S}(\sigma,n) \geq N_{\sigma-1,n}$. By using a softly optimal polynomial multiplication, that is $\mathsf{M}(d)$ in $\tilde{\mathcal{O}}(d)$, and using the assumption $\omega > 2$, we deduce that cost (3) drops to $\mathcal{O}\Big(\mathsf{S}(\sigma,n)d^\omega\Big)$. Furthermore, when softly optimal series multiplication is available, that is $\mathsf{S}(\sigma,n) \in \tilde{\mathcal{O}}(N_{2d-1,n})$, this cost drops further to $\tilde{\mathcal{O}}(2^n d^{\omega-1}N_{d,n+1})$, by using

$$\frac{N_{2d-1,n}}{N_{d,n}} = \frac{(2d-1+n)\cdots(2d)}{(d+n)\cdots(d+1)} \leq 2^n,$$

and

$$dN_{d,n} \in \mathcal{O}\Big(N_{d,n+1}\log(N_{d,n+1})\Big) \text{ (Lecerf and Schost, 2003, Lemma 3).} \tag{4}$$

In general, when using the naive series multiplication, $S(\sigma,n)$ is quadratic in $N_{2d-1,n}$. Thus, combining

$$\frac{N_{2d-1,n}}{N_{d,n}^2} = \frac{n(2d-1+n)}{(d+n)^2}\cdots\frac{2d}{(d+1)^2} \leq 1$$

and inequality (4), we deduce that cost (3) belongs to $\tilde{\mathcal{O}}(N_{d,n+1}^4)$: the cost of this reduction is polynomial (in the size of F).

Let us now consider that n is fixed. We use the notation $\mathcal{O}_d$ to specify that the $\mathcal{O}$ concerns the only parameter d. In this setting, softly optimal series multiplication is always possible. Precisely, we can take $\mathsf{S}(\sigma,n) \in \mathcal{O}(\mathsf{M}(\sigma)^n) \subseteq \tilde{\mathcal{O}}_d(\sigma^n) \subseteq \tilde{\mathcal{O}}_d(d^n)$, thus cost (3) drops to $\tilde{\mathcal{O}}_d\Big(d^{n+\omega}\Big)$. Since the dense size $N_{d,n+1}$ of F is greater than $d^{n+1}/(n+1)!$, and since ω is at most 3, we can say that the cost of the deterministic reduction algorithm is sub-quadratic.

1.3. Probabilistic Reduction Algorithm

We could adapt the probabilistic algorithm of Appendix A.1 to several variables by using the reduction to G as in the previous subsection. Roughly speaking, we would only gain a factor of d in the size of linear system to be solved. The natural probabilistic strategy actually consists of factoring $H(x,y) = F(\alpha_1 x,\ldots,\alpha_n x,y) \in \mathbb{K}[x,y]$ for a Bertinian good point $(\alpha_1,\ldots,\alpha_n) \in \mathbb{K}^n$, as presented in the introduction. The detailed algorithm depends on parameters $u_2,\ldots,u_m$, with $m := 2d^2-1$.

Algorithm 2 Probabilistic factorization algorithm.

Input: *F of total degree d satisfying Hypotheses (C) and (H), $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, and $(u_2, \ldots, u_m) \in \mathbb{K}^{m-1}$.*

Output: *the irreducible factors $F_1, \ldots, F_r$ of F.*

(1) Compute $\mathfrak{H}_1(0, y), \ldots, \mathfrak{H}_s(0, y)$ as the irreducible factors of the univariate polynomial $H(0, y)$.

(2) Lifting step. *Call a fast multi-factor Hensel lifting algorithm in order to obtain the irreducible factors $\mathfrak{H}_1, \ldots, \mathfrak{H}_s$ of $H(x, y)$ in $\mathbb{K}[[x]]/(x^\sigma)[y]$, where $\sigma := 2d$.*

(3) Recombination step.

(a) Compute $\mu_1, \ldots, \mu_r$ by means of Algorithm 3 of the appendix called with input $\mathfrak{H}_1, \ldots, \mathfrak{H}_s$, and $(u_2, \ldots, u_m)$.

(b) Verify that $\mu_1, \ldots, \mu_r$ give the irreducible factorization of H by means of Algorithm 4 of the appendix. If $r = 1$ then return F.

(c) Obtain $F_i(0, \ldots, 0, y)$ as $\prod_{j=1}^{s} \mathfrak{H}_j(0, y)^{\mu_{i,j}}$, for each i in $\{1, \ldots, r\}$.

(d) Let $\bar{d} := \max(\deg(F_i(0, \ldots, 0, y)) \mid i \in \{1, \ldots, r\}) \leq d - 1$. Call a fast multi-factor Hensel lifting algorithm up to precision $(z_1, \ldots, z_n)^{\bar{d}+1}$ in order to recover all the F_i, and return $F_1, \ldots, F_r$.

Proposition 5 *Assume that Hypotheses (C) and (H) hold and that $(\alpha_1, \ldots, \alpha_n)$ is a Bertinian good point for F. There exists a nonzero polynomial $\mathcal{P} \in \mathbb{K}[z_2, \ldots, z_m]$ of total degree at most s such that Algorithm 2 is correct whenever $\mathcal{P}(u_2, \ldots, u_m) \neq 0$. Algorithm 2 performs one factorization of a univariate polynomial of degree d over $\mathbb{K}$ plus a number of operations in $\mathbb{K}$ belonging to*

$$\mathcal{O}\Big(d(\mathsf{M}(d)^2 + \mathsf{M}(d^2)) + \mathsf{S}(d, n)\mathsf{M}(d)\log(d)\Big). \tag{5}$$

Proof. The correctness mainly follows from Propositions 10 and 11 of the appendix. In step (2) we can directly use (von zur Gathen and Gerhard, 2003, Algorithm 15.17), which performs $\mathcal{O}(\mathsf{M}(\sigma)\mathsf{M}(d)\log(s))$ operations, by (von zur Gathen and Gerhard, 2003, part ii of Theorem 15.18). By Proposition 10, the cost of step (3a) is in $\mathcal{O}(s(\mathsf{M}(d^2) + \mathsf{M}(d)^2))$. By Proposition 11, step (3b) takes $\mathcal{O}(M(d)^2\log(d))$ operations. The computation of $F_i(0, \ldots, 0, y)$ can be done by means of (von zur Gathen and Gerhard, 2003, Algorithm 10.3). Thus, by (von zur Gathen and Gerhard, 2003, Lemma 10.3), each $F_i(0, \ldots, 0, y)$ can be computed in time $\mathcal{O}(\mathsf{M}(\deg(F_i(0, \ldots, 0, y)))\log(s)))$. Thanks to the super-additivity of M, step (3c) only takes $\mathcal{O}(\mathsf{M}(d)\log(s))$ operations. The cost of the last step has already been discussed in the proof of Proposition 4: it belongs to $\mathcal{O}((\mathsf{S}(\bar{d}+1, n) + \log(d))\mathsf{M}(d)\log(d))$. □

When softly optimal polynomial and series multiplications are available, that is when $\mathsf{M}(d) \in \tilde{\mathcal{O}}(d)$ and $\mathsf{S}(d, n) \in \tilde{\mathcal{O}}(N_{d-1,n})$, cost (5) drops to:

$$\tilde{\mathcal{O}}(d^3 + dN_{d-1,n}) \subseteq \tilde{\mathcal{O}}(dN_{d,n} + d^3) \subseteq \tilde{\mathcal{O}}(N_{d,n+1} + d^3),$$

where the latter inclusion uses (4). If $n \geq 2$ then d^3 belongs to $\mathcal{O}(N_{d,n+1})$, hence this reduction algorithm is softly optimal.

2. Reduction to Two Variables

We start with a sharp estimate for the density of the Bertinian bad points $(\alpha_1, \ldots, \alpha_n)$ for F. Then we deduce our new quantitative Bertini theorem.

2.1. Under Hypothesis (H)

The set of Bertinian bad points $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ of F is denoted by $\mathfrak{B}_H(F)$. Let us start with an example that will provide us with the lower bounds stated in Theorem 6 and Corollary 7 below.

Example. Let $n \geq 2, \mathbb{K} := \mathbb{C}, F := y^d + z_1^{d-1} y - z_2^{d-1} - 1$. The Stepanov-Schmidt criterion implies that $F(0, z_2, 0, \ldots, 0, y) = y^d - z_2^{d-1} - 1$ is irreducible, thus F is irreducible (for this criterion and recent generalizations see Gao, 2001). Let S denote the set of roots of $z^{d(d-1)} - 1$. For any $(\alpha_1, \ldots, \alpha_n) \in S^n$, the polynomial $y - (\alpha_2/\alpha_1)^{d-1}$ divides $H = y^d - 1 + x^{d-1}(\alpha_1^{d-1} y - \alpha_2^{d-1})$. Therefore, all the points of S^n are Bertinian bad points for F, whence

$$\frac{|\mathfrak{B}_H(F) \cap S^n|}{|S|^n} = d(d-1)/|S| = 1.$$

By the classical Schwartz-Zippel lemma (Zippel, 1979; Schwartz, 1980): a nonzero polynomial $\mathcal{A}$ in n variables can not have more than $\deg(\mathcal{A})|S|^{n-1}$ roots in S^n. We deduce that there exists no polynomial $\mathcal{A}$ of degree at most $d(d-1)-1$ that vanishes on $\mathfrak{B}_H(F)$.

Now we deal with the upper bound:

Theorem 6 *Under Hypotheses (C) and (H), there exists a polynomial in $\mathbb{K}[a_1, \ldots, a_n] \setminus \{0\}$ of total degree at most $(d-1)(2d-1)$ that vanishes on $\mathfrak{B}_H(F)$. In addition we have:*

$$\max\Big(\min\big(\deg(\mathcal{A}) \mid \mathcal{A}(\mathfrak{B}_H(F)) = 0 \big) \mid F \text{ satisfies (C) and (H)} \Big) \geq d(d-1),$$

where $\mathcal{A}$ is taken over all the nonzero polynomials in $\mathbb{K}[a_1, \ldots, a_n]$.

Proof. Since the map $\delta : d \mapsto (d-1)(2d-1)$ satisfies $\delta(d_1) + \delta(d_2) \leq \delta(d_1 + d_2)$, for any positive integers d_1 and d_2, we can assume that F is irreducible.

In the rest of the proof we let $\sigma := 2d$. For any $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$, we introduce the following linear system:

$$D_{\alpha,\sigma} \begin{cases} \displaystyle\sum_{i=1}^{s} \ell_i \operatorname{coeff}\left(\hat{\mathfrak{H}}_i \frac{\partial \mathfrak{H}_i}{\partial y}, x^j y^k \right) = 0, \ k \leq d-1, \ d \leq j+k \leq \sigma - 1, \\ \displaystyle\sum_{i=1}^{s} \ell_i \operatorname{coeff}\left(\hat{\mathfrak{H}}_i \frac{\partial \mathfrak{H}_i}{\partial x}, x^j y^k \right) = 0, \ k \leq d-1, \ j \leq \sigma - 2, \ d \leq j+k \leq \sigma - 1, \end{cases}$$

where $\mathfrak{H}_1(x,y), \ldots, \mathfrak{H}_s(x,y)$ represent the monic irreducible factors of $H(x,y)$ in $\mathbb{K}[[x]][y]$, and $\hat{\mathfrak{H}}_i := H/\mathfrak{H}_i$, for all $i \in \{1, \ldots, s\}$. By (Lecerf, 2006, Theorem 1 and Lemma 4), the rank of the solution space of $D_{\alpha,\sigma}$ equals the number of irreducible factors of H.

From Lemma 1, we know that $D_{a,\sigma}$ has rank $s-1$. Therefore there exists a nonzero minor $\mathcal{A} \in \mathbb{K}[a_1, \ldots, a_n]$ of size $s-1$ in $D_{a,\sigma}$. Now remark that $D_{\alpha,\sigma}$ coincides with the specialization of $D_{a,\sigma}$ at $a_1 = \alpha_1, \ldots, a_n = \alpha_n$. Therefore, if $\mathcal{A}(\alpha_1, \ldots, \alpha_n) \neq 0$ then $D_{\alpha,\sigma}$ has rank $s-1$, hence H is irreducible.

For any $j \in \{0, \ldots, \sigma - 1\}$, the coefficient $\operatorname{coeff}\left(\hat{\mathfrak{G}}_i(x,y) \dfrac{\partial \mathfrak{G}_i}{\partial y}(x,y), x^j y^k \right)$ is a polynomial of degree at most j. For any $j \in \{0, \ldots, \sigma - 2\}$, $\operatorname{coeff}\left(\hat{\mathfrak{G}}_i(x,y) \dfrac{\partial \mathfrak{G}_i}{\partial x}(x,y), x^j y^k \right)$ is a polynomial of degree at most $j+1$ (see the proof of Lemma 3). It follows that $\mathcal{A}$ is a polynomial of total degree at most $(s-1)(2d-1)$. □

In terms of counting Bertinian bad points, we deduce the following corollary thanks to the Schwartz-Zippel lemma mentioned above.

Corollary 7 *Under Hypotheses (C) and (H), for any finite non-empty subset S of $\mathbb{K}$, we have*

$$\frac{|\mathfrak{B}_H(F) \cap S^n|}{|S|^n} \leq (d-1)(2d-1)/|S|.$$

This bound is asymptotically sharp up to a constant factor:

$$\max\left(\frac{|\mathfrak{B}_H(F) \cap S^n|}{|S|^{n-1}} \mid S \subseteq \mathbb{K} \text{ and } F \text{ satisfies (C) and (H)}\right) \geq d(d-1).$$

Roughly speaking, this corollary asserts that it is necessary and sufficient to take $|S| \gg d^2$ in order to pick up a Bertinian good point at random in S^n with a high probability of success.

2.2. Quantitative Bertini Theorem

Now we are ready to deduce our quantitative Bertini theorem. We keep on using the notation of the introduction: P denotes a polynomial in $\mathbb{K}[v_1, \ldots, v_n]$ of total degree $d \geq 1$. If P is squarefree then the variables can be changed in order to recover Hypothesis (H) for a suitable polynomial F. This is the main idea for proving:

Corollary 8 *Under Hypothesis (C), for any $P \in \mathbb{K}[v_1, \ldots, v_n]$ of total degree d and for any non-empty finite subset S of $\mathbb{K}$, we have $\mathfrak{B}(P, S) \leq (3d(d-1)+1)/|S|$.*

Proof. Since the map $\delta : d \mapsto 3d(d-1)+1$ satisfies $\delta(d_1) + \delta(d_2) \leq \delta(d_1 + d_2)$, for any positive integers d_1 and d_2, we can assume that P is irreducible. Let $w_1, \ldots, w_n$, $z_1, \ldots, z_n$ be new sets of variables. For any $(\beta_1, \ldots, \beta_n)$ and $(\gamma_1, \ldots, \gamma_n)$, we define:

$$\begin{aligned} P_\beta &:= P(w_1 + \beta_1 y, \ldots, w_n + \beta_n y) \in \mathbb{K}[w_1, \ldots, w_n, y], \\ P_{\beta,\gamma} &:= P_\beta(z_1 + \gamma_1, \ldots, z_n + \gamma_n, y) \in \mathbb{K}[z_1, \ldots, z_n, y]. \end{aligned}$$

Let $\mathcal{B} \in \mathbb{K}[b_1, \ldots, b_n]$ represent the homogeneous component of P of highest degree d. It is straightforward to verify that if $(\beta_1, \ldots, \beta_n)$ is not a zero of $\mathcal{B}$ then P_β is monic in y.

For any $(\beta_1, \ldots, \beta_n) \in \mathbb{K}^n$ such that $\mathcal{B}(\beta_1, \ldots, \beta_n) \neq 0$, we introduce the discriminant $\mathcal{C}_\beta \in \mathbb{K}[c_1, \ldots, c_n]$ of P_β with respect to y. Since P_β is squarefree, Hypothesis (C) implies that $\mathcal{C}_\beta$ is a nonzero polynomial of degree at most $d(d-1)$. For any $(\gamma_1, \ldots, \gamma_n) \in \mathbb{K}^n$ such that $\mathcal{C}_\beta(\gamma_1, \ldots, \gamma_n) \neq 0$, the polynomial $F := P_{\beta,\gamma}$ satisfies Hypothesis (H). Therefore, Theorem 6 ensures the existence of a nonzero polynomial $\mathcal{A}_{\beta,\gamma} \in \mathbb{K}[a_1, \ldots, a_n]$ of degree at most $(d-1)(2d-1)$ satisfying the following property: for any $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ such that $\mathcal{A}_{\beta,\gamma}(\alpha_1, \ldots, \alpha_n) \neq 0$, the polynomial $H = P_{\beta,\gamma}(\alpha_1 x, \ldots, \alpha_n x, y)$ is irreducible. This way we obtain:

$$\begin{aligned} \mathfrak{B}(P) \subseteq & \mathbb{K}^n \times \{(\beta_1, \ldots, \beta_n) \mid \mathcal{B}(\beta_1, \ldots, \beta_n) = 0\} \times \mathbb{K}^n \\ & \cup \mathbb{K}^n \times \{(\beta_1, \ldots, \beta_n), (\gamma_1, \ldots, \gamma_n) \mid \mathcal{B}(\beta_1, \ldots, \beta_n) \neq 0,\ \mathcal{C}_\beta(\gamma_1, \ldots, \gamma_n) = 0\} \\ & \cup \{(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n), (\gamma_1, \ldots, \gamma_n) \mid \mathcal{B}(\beta_1, \ldots, \beta_n) \neq 0, \\ & \qquad \mathcal{C}_\beta(\gamma_1, \ldots, \gamma_n) \neq 0,\ \mathcal{A}_{\beta,\gamma}(\alpha_1, \ldots, \alpha_n) = 0\}. \end{aligned}$$

Finally, by using the Schwartz-Zippel lemma with $\mathcal{B}$, $\mathcal{C}_b$ and $\mathcal{A}_{b,c}$, it follows that:

$$|\mathfrak{B}(P) \cap (S^n)^3| \leq d|S|^{3n-1} + d(d-1)|S|^{3n-1} + (d-1)(2d-1)|S|^{3n-1},$$

which yields the claimed bound. □

Acknowledgements

This article had been accepted for presentation at the MEGA 2005 conference. It had then been accepted to the related special issue of Journal of Symbolic Computation before I found the mistakes pointed in Appendix A. I am grateful to the anonymous referees involved in these processes for their useful comments.

A. Errata for "Complexity Issues in Bivariate Polynomial Factorization" and "Sharp Precision in Hensel Lifting for Bivariate Polynomial Factorization"

In (Bostan et al., 2004; Lecerf, 2006) we presented deterministic and probabilistic recombination algorithms for the factorization of dense bivariate polynomials. It turns out that the analyzes of the probability of success of the probabilistic algorithms are wrong. In this appendix, we explain what is wrong and what can be fixed. We follow the notation of (Lecerf, 2006). Recall that we are interested in computing the irreducible factorization of a polynomial F of total degree d in two variables x and y over a commutative field $\mathbb{K}$, under the assumption that the characteristic of $\mathbb{K}$ is zero or at least $d(d-1)+1$.

What is Wrong

In (Bostan et al., 2004, Corollary 2) and (Lecerf, 2006, Proposition 3) it was claimed that the recombination problem could be solved with $\mathcal{O}(d^\omega)$ arithmetic operations in $\mathbb{K}$ in average (here ω denotes a feasible matrix multiplication exponent). This result is wrong. The error appears in (Bostan et al., 2004, Lemma 1), and is repeated in (Lecerf, 2006, Lemma 5). More precisely, the error is at the end of the proof of (Bostan et al., 2004, Lemma 1): there it is said that the restriction to $\mathbb{K}$ of the solution set of a $\mathbb{K}(x)$-linear system $\mathcal{S}$ can be obtained by means of the only resolution over $\mathbb{K}$ of specializations of $\mathcal{S}$ at only two suitable values for x in $\mathbb{K}$, which is in general wrong. This error implies that the probability of success of the probabilistic recombination algorithms presented in (Bostan et al., 2004, Section 2.2) and (Lecerf, 2006, Section 3) is erroneous. The other results of (Bostan et al., 2004; Lecerf, 2006) are not affected by this error. In the previous version of the present paper, that was accepted at the MEGA 2005 conference, only the constants in the upper bounds on the density of Bertinian bad points of Section 2 suffered from this error.

What can be Fixed

We will only focus on fixing the statements of (Lecerf, 2006). Similar corrections for (Bostan et al., 2004) follow *mutatis mutandis*. In the first subsection we present a new probabilistic algorithm with a cost in $\tilde{\mathcal{O}}(d^3)$ in average. In the second subsection we correct the wrong probabilistic algorithm of (Lecerf, 2006) so that is always return a correct answer, and refer to it as the *heuristic algorithm*. This heuristic gives the best performances in practice, as observed in (Bostan et al., 2004, Section 2.4), and we leave the question of its probability of success to future work.

A.1. The Corrected Probabilistic Recombination Algorithm

In this subsection we present a new probabilistic recombination algorithm with an average cost in $\tilde{\mathcal{O}}(d^3)$. We start with recalling a classical preconditioning technique, due to Kaltofen and Saunders (1991, Theorem 2), for solving overdetermined linear systems faster. We briefly recall the proof for convenience. For other possible strategies, we refer the reader to Chen et al. (2002).

Lemma 9 *Let A be a $m \times s$ matrix over $\mathbb{K}$ of rank $s - r$, and let U be the following upper triangular $s \times m$ Toepliz matrix with entries in $\mathbb{K}$:*

$$U := \begin{pmatrix} 1 & u_2 & u_3 & \cdots & u_{m-1} & u_m \\ & 1 & u_2 & u_3 & \cdots & u_{m-1} \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & \dots & u_{m-s+1} \end{pmatrix}.$$

There exists a nonzero polynomial $\mathcal{P} \in \mathbb{K}[z_2, \dots, z_m]$ of total degree at most s such that the matrix UA has rank $s - r$ whenever $\mathcal{P}(u_2, \dots, u_m) \neq 0$.

Proof. The case when $m \leq s$ is immediate (we can take $\mathcal{P} = 1$). Let us now assume that $m > s$. Let $B := UA$, $t := s - r$, $I := \{1, \dots, t\}$, and $J := \{j_1, \dots, j_t\}$ be such that the columns of A indexed by $j_1, \dots, j_t$ are linearly independent. Let us assume that U has generic entries $z_2, \dots, z_m$ replacing $u_2, \dots, u_m$, and let us take

$$\mathcal{P} := \sum_K U_{I,K} A_{K,J},$$

where the sum is taken over all the subsets K of $\{1, \dots, m\}$ with cardinality t, and where $U_{I,K}$ represents the determinant of the submatrix of U composed of the rows indexed by I and columns indexed by K. The polynomials $U_{I,K}$ are $\mathbb{K}$-linearly independent (see the proof of Kaltofen and Saunders, 1991, Theorem 2). Since there exists K such that $A_{K,J} \neq 0$, we have that $\mathcal{P} \neq 0$. The conclusion follows from the classical Cauchy-Binet formula that provides us with $\mathcal{P} = B_{I,J}$. □

From now let $\sigma := 2d$, and let A denote the $m \times s$ matrix associated to D_σ, where $m := 2d^2 - 1$. With a lucky matrix U, the computation of the kernel of A reduces to computing the kernel of UA, which has size $s \times s$. Since U is Toepliz, the product UA can be obtained efficiently. This is the main idea in the following algorithm:

Algorithm 3 Probabilistic recombination algorithm.
Input: $\mathfrak{F}_1, \dots, \mathfrak{F}_s$ *to precision* (x^σ), *and* $(u_2, \dots, u_m) \in \mathbb{K}^{m-1}$.
Output: $\mu_1, \dots, \mu_r$.

(1) For each $i \in \{1, \dots, s\}$, compute $\hat{\mathfrak{F}}_i$ as the quotient of F by $\mathfrak{F}_i$ to precision (x^σ).
(2) Compute $(\hat{\mathfrak{F}}_1 \frac{\partial \mathfrak{F}_1}{\partial y}, \dots, \hat{\mathfrak{F}}_s \frac{\partial \mathfrak{F}_s}{\partial y})$ to precision (x^σ).
(3) Compute $(\hat{\mathfrak{F}}_1 \frac{\partial \mathfrak{F}_1}{\partial x}, \dots, \hat{\mathfrak{F}}_s \frac{\partial \mathfrak{F}_s}{\partial x})$ to precision $(x^{\sigma-1})$.
(4) Build the matrix A, and compute $B = UA$.
(5) Return the reduced echelon basis of the kernel of B.

Proposition 10 *Under Hypothesis (H), for any F, there exists a nonzero polynomial $\mathcal{P} \in \mathbb{K}[z_2, \ldots, z_m]$ of total degree at most s such that Algorithm 3 returns a correct answer whenever $\mathcal{P}(u_2, \ldots, u_m) \neq 0$. The cost of Algorithm 3 belongs to $\mathcal{O}(s(\mathsf{M}(d)^2 + \mathsf{M}(d^2)))$, or to $\tilde{\mathcal{O}}(d^3)$.*

Proof. The correctness follows from (Lecerf, 2006, Theorem 1 and Lemma 4) and Lemma 9. The cost analysis of steps (1) to (3) is the same as in (Lecerf, 2006, Proposition 1). In step (4), it is classical that the product UA costs $\mathcal{O}(s\mathsf{M}(d^2))$. The final kernel computation is in $\mathcal{O}(s^\omega)$ by (Storjohann, 2000, Theorem 2.10). □

Let S be a finite subset of $\mathbb{K}$ of cardinality $|S|$ and assume that $u_2, \ldots, u_m$ are uniformly taken at random in S. By the classical Schwartz-Zippel lemma (Zippel, 1979; Schwartz, 1980) and Lemma 9, we obtain that the probability of getting a zero of $\mathcal{P}$ is at most $s/|S|$. Since the output of Algorithm 3 can be verified in softly optimal time (see Algorithm 4 below), we can thus deduce a recombination algorithm that always returns a correct answer with an average cost in $\tilde{\mathcal{O}}(d^3)$.

A.2. Heuristic Recombination Algorithm

The heuristic recombination algorithm we are to present can be seen as a variant of the deterministic one (of Lecerf, 2006, Section 3). We solve the over-determined linear system D_σ progressively: D_σ is split into d subsystems of sizes $\mathcal{O}(d) \times s$. Each subsystem can be built efficiently and independently of the others. This way we can compute the intersection of their solution set in sequence and stop the resolution when the softly optimal early exit criterion given below is satisfied. In practice we observe that only a few subsystems are necessary, and that this approach is faster than the one of the previous subsection.

Algorithm 4 Early exit criterion.

Input: *$\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to precision (x^d), and a reduced echelon basis $\nu_1, \ldots, \nu_t$ such that $\langle \mu_1, \ldots, \mu_r \rangle \subseteq \langle \nu_1, \ldots, \nu_t \rangle$.*

Output: *"true" if $(\mu_1, \ldots, \mu_r) = (\nu_1, \ldots, \nu_t)$, and "false" otherwise.*

(1) If $t = 1$ then return "true".

(2) If the entries of the ν_i are not in $\{0, 1\}$ the return "false".

(3) If the supports of $\nu_1, \ldots, \nu_t$ do not form a partition of $\{1, \ldots, s\}$ of size t then return "false".

(4) For each $i \in \{1, \ldots, t\}$, let $d_i := \sum_{j=1}^{s} \nu_{i,j}$, and let $\tilde{F}_i \in \mathbb{K}[x, y]_{d_i}$ be computed as the truncation of $\prod_{j=1}^{s} \mathfrak{F}_j^{\nu_{i,j}}$ modulo $(x, y)^{d_i+1}$.

(5) If $\prod_{i=1}^{t} \tilde{F}_j = F$ then return "true" else return "false".

Proposition 11 *Under Hypothesis (H), Algorithm 4 is correct and takes $\mathcal{O}(\mathsf{M}(d)^2 \log(d))$ operations in $\mathbb{K}$.*

Proof. It is clear that the "false" answer is always correct. If "true" comes from the first step, this means that $r = t = 1$ (that is F is irreducible). If "true" is returned by the last step then each $\tilde{F}_j$ is a product of some irreducible factors of F, whence $\langle \nu_1, \ldots, \nu_t \rangle \subseteq \langle \mu_1, \ldots, \mu_r \rangle$.

Steps (1) to (3) take $\mathcal{O}(st)$ operations in $\mathbb{K}$. In step (4) each $\tilde{F}_i$ can be computed by the sub-product tree technique with $\mathcal{O}(\mathsf{M}(d_i)^2 \log(d_i))$ operations, by (von zur Gathen and

Gerhard, 2003, Lemma 10.3). The total cost of this step thus belongs to $\mathcal{O}(\mathsf{M}(d)^2 \log(d))$. Similarly step (5) also takes the $\mathcal{O}(\mathsf{M}(d)^2 \log(d))$. $\square$

Let $\tau := 2d + 1$. Here we require precision (x^τ) for the lifted factors $\mathfrak{F}_1, \ldots, \mathfrak{F}_s$. For any $u \in \mathbb{K}$, we introduce the following linear system P_τ^u:

$$P_\tau^u \begin{cases} \sum_{i=1}^s \ell_i \operatorname{coeff}\left(\hat{\mathfrak{F}}_i(x, ux)\frac{\partial \mathfrak{F}_i}{\partial x}(x, ux), x^j\right) = 0, \ d \le j \le \tau - 2, \\ \sum_{i=1}^s \ell_i \operatorname{coeff}\left(\hat{\mathfrak{F}}_i(x, ux)\frac{\partial \mathfrak{F}_i}{\partial y}(x, ux), x^j\right) = 0, \ d \le j \le \tau - 2. \end{cases}$$

The heuristic recombination algorithm proceeds as follows:

Algorithm 5 Heuristic recombination algorithm.

Input: *$\mathfrak{F}_1, \ldots, \mathfrak{F}_s$ to precision (x^τ), and $\{u_1, \ldots, u_d\} \subseteq \mathbb{K}^d$.*

Output: *$\mu_1, \ldots, \mu_r$.*

(1) Initialize t with s, and $\nu_1, \ldots, \nu_t$ with the canonical basis of $\mathbb{K}^s$.

(2) For u in $\{u_1, \ldots, u_d\}$ do:

- *(a) For each $i \in \{1, \ldots, s\}$, compute $f_i := \mathfrak{F}_i(x, ux)$, $g_i := \frac{\partial \mathfrak{F}_i}{\partial y}(x, ux)$, and $h_i := \frac{\partial \mathfrak{F}_i}{\partial x}(x, ux)$ to precision $(x^{\tau-1})$.*
- *(b) Let $A_1 := 1$, $B_s := 1$. For each i from 2 to s, compute $A_i := A_{i-1} f_{i-1}$, $B_{s-i+1} := B_{s-i+2} f_{s-i+2}$ to precision $(x^{\tau-1})$.*
- *(c) For each $i \in \{1, \ldots, s\}$, compute $\hat{\mathfrak{F}}_i(x, ux)\frac{\partial \mathfrak{F}_i}{\partial y}(x, ux)$ as $g_i A_i B_i$ to precision $(x^{\tau-1})$ and $\hat{\mathfrak{F}}_i(x, ux)\frac{\partial \mathfrak{F}_i}{\partial x}(x, ux)$ as $h_i A_i B_i$ to precision $(x^{\tau-1})$ (remark that $A_i B_i = \prod_{j=1, j\neq i}^s f_j$).*
- *(d) Update $\nu_1, \ldots, \nu_t$ with the reduced echelon basis of the restriction to $\langle \nu_1, \ldots, \nu_t \rangle$ of the solutions of P_τ^u.*
- *(e) If Algorithm 4 returns "true" then return $\nu_1, \ldots, \nu_t$.*

Proposition 12 *Under Hypothesis (H), if $u_1, \ldots, u_d$ are pairwise distinct then Algorithm 5 is correct. Each step of the main loop (2) takes $\mathcal{O}(ds^{\omega-1} + \mathsf{M}(d)^2 \log(d))$ operations in $\mathbb{K}$.*

Proof. When u is seen as a transcendental parameter over $\mathbb{K}$ then the solution set over $\mathbb{K}$ of P_τ^u coincides with the one of D_σ (this comes from Lecerf, 2006, equation (7)). Since P_τ^u involves polynomials in u of degree at most $d - 1$, the latter solution set coincides with the common solutions of $P_\tau^{u_1}, \ldots, P_\tau^{u_d}$, whenever $u_1, \ldots, u_d$ are pairwise distinct.

The cost of steps (2a) to (2c) belongs to $\mathcal{O}(d\tau + \mathsf{M}(\tau)s)$ (see the proof of Lecerf, 2006, Proposition 2). The computations in step (2d) can be done as follows. Let M denote the matrix of P_τ^u, let N denote the matrix whose columns are $\nu_1, \ldots, \nu_t$, and let $\tilde{N}$ be a matrix whose columns are a basis of the kernel of MN. Then the columns of $N\tilde{N}$ generate the restriction to $\langle \nu_1, \ldots, \nu_t \rangle$ of the solutions of P_τ^u. This way, and thanks to (Storjohann, 2000, Theorem 2.10), step (2d) costs $\mathcal{O}(\tau s^{\omega-1})$. The cost of step (2e) comes from Proposition 11. We finally deduce that each step of the main loop costs $\mathcal{O}(\mathsf{M}(d)^2 \log(d) + \tau s^{\omega-1})$. $\square$

On all the examples we have tested, the early exit happens after only one or two steps of the main loop. Therefore the interesting question is the following: what is the average cost of Algorithm 5 when taking $u_1, \ldots, u_d$ uniformly at random in a given finite subset of $\mathbb{K}$?

References

Bajaj, C., Canny, J., Garrity, T., Warren, J., 1993. Factoring rational polynomials over the complex numbers. SIAM J. Comput. 22 (2), 318–331.

Bostan, A., Lecerf, G., Salvy, B., Schost, É., Wiebelt, B., 2004. Complexity issues in bivariate polynomial factorization. In: Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation. ACM Press, pp. 42–49.

Bürgisser, P., Clausen, M., Shokrollahi, M. A., 1997. Algebraic complexity theory. Springer-Verlag.

Chen, L., Eberly, W., Kaltofen, E., Saunders, B. D., Turner, W. J., Villard, G., 2002. Efficient matrix preconditioners for black box linear algebra. Linear Algebra Appl. 343/344, 119–146, special issue on structured and infinite systems of linear equations.

Chèze, G., 2004. Des méthodes symboliques-numériques et exactes pour la factorisation absolue des polynômes en deux variables. Ph.D. thesis, Université de Nice-Sophia Antipolis (France).

Chèze, G., Galligo, A., 2005. Four lectures on polynomial absolute factorization. In: Dickenstein, A., Emiris, I. Z. (Eds.), Solving polynomial equations: foundations, algorithms, and applications. Vol. 14 of Algorithms Comput. Math. Springer-Verlag, pp. 339–392.

Chèze, G., Lecerf, G., 2005. Lifting and recombination techniques for absolute factorization, manuscript.

Gao, S., 2001. Absolute irreducibility of polynomials via Newton polytopes. J. Algebra 237 (2), 501–520.

Gao, S., 2003. Factoring multivariate polynomials via partial differential equations. Math. Comp. 72, 801–822.

Gao, S., Kaltofen, E., Lauder., A., 2004. Deterministic distinct degree factorization for polynomials over finite fields. J. Symbolic Comput. 38 (6), 1461–1470.

von zur Gathen, J., 1985. Irreducibility of multivariate polynomials. J. Comput. System Sci. 31 (2), 225–264.

von zur Gathen, J., Gerhard, J., 2003. Modern computer algebra, 2nd Edition. Cambridge University Press.

von zur Gathen, J., Kaltofen, E., 1985a. Factoring sparse multivariate polynomials. J. Comput. System Sci. 31, 265–287.

von zur Gathen, J., Kaltofen, E., 1985b. Factorization of multivariate polynomials over finite fields. Math. Comp. 45 (171), 251–261.

Heintz, J., Sieveking, M., 1981. Absolute primality of polynomials is decidable in random polynomial time in the number of variables. In: Automata, languages and programming (Akko, 1981). Vol. 115 of Lecture Notes in Comput. Sci. Springer-Verlag, pp. 16–28.

Hilbert, D., 1892. Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. J. reine angew. Math. 110.

van der Hoeven, J., 2004. The truncated Fourier transform and applications. In: Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation. ACM Press, pp. 290–296.

van der Hoeven, J., 2005. Notes on the truncated Fourier transform, manuscript.

Jouanolou, J.-P., 1983. Théorèmes de Bertini et applications. Vol. 42 of Progress in Mathematics. Birkhäuser.

Kaltofen, E., 1982a. Polynomial factorization. In: Buchberger, B., Collins, G., Loos, R. (Eds.), Computer algebra. Springer-Verlag, pp. 95–113.

Kaltofen, E., 1982b. A polynomial reduction from multivariate to bivariate integral polynomial factorization. In: Proceedings of the 14th Symposium on Theory of Computing. ACM, pp. 261–266.

Kaltofen, E., 1982c. A polynomial-time reduction from bivariate to univariate integral polynomial factorization. In: Proceedings of the 23rd Symposium on Foundations of Computer Science. IEEE, pp. 57–64.

Kaltofen, E., 1985a. Effective Hilbert irreducibility. Inform. and Control 66 (3), 123–137.

Kaltofen, E., 1985b. Fast parallel absolute irreducibility testing. J. Symbolic Comput. 1 (1), 57–67.

Kaltofen, E., 1985c. Polynomial-time reductions from multivariate to bi- and univariate integral polynomial factorization. SIAM J. Comput. 14 (2), 469–489.

Kaltofen, E., 1985d. Sparse Hensel lifting. In: Proceedings of EUROCAL '85, Vol. 2 (Linz, 1985). Vol. 204 of Lecture Notes in Comput. Sci. Springer-Verlag, pp. 4–17.

Kaltofen, E., 1987. Deterministic irreducibility testing of polynomials over large finite fields. J. Symbolic Comput. 4 (1), 77–82.

Kaltofen, E., 1989. Factorization of polynomials given by straight-line programs. In: Micali, S. (Ed.), Randomness and Computation. Vol. 5 of Advances in Computing Research. JAI Press Inc., pp. 375–412.

Kaltofen, E., 1990. Polynomial factorization 1982–1986. In: Computers in mathematics (Stanford, CA, 1986). Vol. 125 of Lecture Notes in Pure and Appl. Math. Dekker, pp. 285–309.

Kaltofen, E., 1992. Polynomial factorization 1987–1991. In: LATIN '92 (São Paulo, 1992). Vol. 583 of Lecture Notes in Comput. Sci. Springer-Verlag, pp. 294–313.

Kaltofen, E., 1995. Effective Noether irreducibility forms and applications. J. Comput. System Sci. 50 (2), 274–295.

Kaltofen, E., 2003. Polynomial factorization: a success story. In: Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation. ACM Press, pp. 3–4.

Kaltofen, E., Saunders, B. D., 1991. On Wiedemann's method of solving sparse linear systems. In: Mattson, H. F., Mora, T., Rao, T. R. N. (Eds.), Proceedings of AAECC-9. Vol. 539 of Lect. Notes Comput. Sci. Springer-Verlag, pp. 29–38.

Kaltofen, E., Trager, B., 1990. Computing with polynomials given by black boxes for their evaluations: Greatest common divisors, factorization, separation of numerators and denominators. J. Symbolic Comput. 9 (3), 301–320.

Kleiman, S. L., 1998. Bertini and his two fundamental theorems. Rend. Circ. Mat. Palermo (2) Suppl. (55), 9–37, studies in the history of modern mathematics, III.

Lecerf, G., 2003. Computing the equidimensional decomposition of an algebraic closed set by means of lifting fibers. J. Complexity 19 (4), 564–596.

Lecerf, G., 2006. Sharp precision in Hensel lifting for bivariate polynomial factorization. Math. Comp. 75, 921–933.

Lecerf, G., Schost, É., 2003. Fast multivariate power series multiplication in characteristic zero. SADIO Electronic Journal on Informatics and Operations Research 5 (1), 1–10.

Mumford, D., 1995. Algebraic geometry. I Complex projective varieties. Classics in Mathematics. Springer-Verlag, reprint of the 1976 edition.

Ruppert, W. M., 1986. Reduzibilität ebener Kurven. J. Reine Angew. Math. 369, 167–191.

Ruppert, W. M., 1999. Reducibility of polynomials $f(x, y)$ modulo p. J. Number Theory 77 (1), 62–70.

Schinzel, A., 2000. Polynomials with special regard to reducibility. Vol. 77 of Encyclopedia of Mathematics and its Applications. Cambridge University Press.

Schost, É., 2005. Multivariate power series multiplication. In: Proceedings of the 2005 international symposium on Symbolic and algebraic computation. ACM Press, pp. 293–300.

Schwartz, J. T., 1980. Fast probabilistic algorithms for verification of polynomial identities. Journal of the ACM 27 (4), 701–717.

Shafarevich, I. R., 1994. Basic algebraic geometry. 1 Varieties in projective space, 2nd Edition. Springer-Verlag.

Storjohann, A., 2000. Algorithms for matrix canonical forms. Ph.D. thesis, ETH, Zürich (Switzerland).

Zassenhaus, H., 1969. On Hensel factorization I. J. Number Theory 1 (1), 291–311.

Zippel, R., 1979. Probabilistic algorithms for sparse polynomials. In: Proceedings of EUROSAM '79. No. 72 in Lecture Notes in Comput. Sci. Springer-Verlag, pp. 216–226.

Zippel, R., 1993. Effective Polynomial Computation. Kluwer Academic Publishers.