The Complexity of Factors of Multivariate Polynomials *

Peter Bürgisser

Dept. of Mathematics and Computer Science
University of Paderborn
D-33095 Paderborn, Germany
buergisser@upb.de

July 3, 2003

Abstract

The existence of string functions, which are not polynomial time computable, but whose graph is checkable in polynomial time, is a basic assumption in cryptography. We prove that in the framework of algebraic complexity, there are no such families of polynomial functions of polynomially bounded degree over fields of characteristic zero. The proof relies on a polynomial upper bound on the approximative complexity of a factor g of a polynomial f in terms of the (approximative) complexity of f and the degree of the factor g. This extends a result by Kaltofen (STOC 1986). The concept of approximative complexity allows to cope with the case that a factor has an exponential multiplicity, by using a perturbation argument. Our result extends to randomized (two-sided error) decision complexity.

1 Introduction

Checking or verifying a solution to a computational problem might be easier than computing a solution. In a certain sense, this is the contents of the famous $P \neq NP$ hypothesis. In [39] Valiant made an attempt to clarify the principal relationship between the complexity of checking and evaluating. In particular, he asked whether any (string) function, for which values can

---

be checked in polynomial time, can also be evaluated in polynomial time. Cryptographers hope that the answer to this question is negative, since it turns out to be intimately connected to the existence of one-way functions. Indeed, the inverse φ of a one-way function is not polynomial time computable, but membership to the graph of φ can be decided in polynomial time. The converse is also known to be true [20, 35] and equivalent to $P \neq UP$.

The goal of this paper is to investigate the relationship between the complexity of computational and decisional tasks in an algebraic framework of computation, a line of research initiated by Lickteig [29, 30]. Unless stated otherwise, k denotes a fixed field of characteristic zero. Are there families of polynomials (φ_n) over k, for which checking the value can be done with a polynomial number of arithmetic operations and tests, but which cannot be evaluated with a polynomial number of arithmetic operations? We do not know the answer to this question. However, we will be able to show that the answer is negative under the restriction that the degree of φ_n grows at most polynomially in n. Actually, our result is slightly weaker in the sense that we know it to be true only for a notion of approximative complexity.

1.1 Decision, Computation, and Factors

We discuss a basic relationship between the complexity of decision and computation in our algebraic framework of computation and raise some natural open questions.

By the *straight-line complexity* $L(g)$ of a multivariate polynomial g over k we understand the minimal number of arithmetic operations sufficient to compute $g(X_1, \ldots, X_n)$ by a straight-line program without divisions from the variables X_i and constants in k. The *decision complexity* $C(g)$ of g is defined as the minimal number of arithmetic operations and tests sufficient for an algebraic computation tree to decide for given points x in k^n whether $g(x) = 0$. If $k = \mathbb{R}$, we allow also $\leq$-tests. Clearly, $C(g) \leq L(g) + 1$ (the 1 accounts for the zero test). We define the *exclusion complexity* $EC(g)$ of g similarly as in [30, 11]

$$EC(g) := \min\{L(f) \mid f \in k[X_1, \ldots, X_n] \setminus \{0\}, \ g \mid f\}.$$

We clearly have $EC(g) \leq L(g)$. For formal definitions, we refer to [10].

In the following, we assume that g is the irreducible generator of a hypersurface in k^n and either $k = \mathbb{R}$ or $\mathbb{C}$. Let f be a nonzero polynomial multiple of g, say $f = g^e h$ with a polynomial h coprime to g and $e \in \mathbb{N}_{>0}$. Then any straight-line program for f can be used to exclude membership

2

to the zeroset of g: $f(x) \neq 0$ implies $g(x) \neq 0$ and the converse is also true, provided $h(x) \neq 0$. Thus we may consider $EC(g)$ as a "generic decision complexity" of g.

The following well-known lemma provides a link between decisional and computational complexity (cf. [12, 3]). The proof is a rather straightforward consequence of the Nullstellensatz.

Lemma 1.1 *Let g be the irreducible generator of a hypersurface in $\mathbb{R}^n$ or $\mathbb{C}^n$. Then $EC(g) \leq C(g)$.*

Over the reals, we need both assumptions that g is irreducible (see the comment on question (2) below) and that the zeroset of g is a hypersurface (take $g = X_1^{2m} + \ldots + X_n^{2m}$ over $\mathbb{R}$). Over the complex numbers, one can relax these assumptions and show that $EC(g) \leq C(g) + r - 1$ if g is squarefree with r irreducible factors. Moreover, we remark that the conclusion of this lemma remains true over any infinite field k if g is the generator of the graph of a polynomial φ, that is, $g = Y - \varphi(X_1, \ldots, X_n)$.

Under the assumption of the lemma we have $EC(g) \leq C(g) \leq L(g) + 1$. Asking about inequalities in the reverse direction, it is natural to raise the following questions:

$$
\begin{align}
L(g) &\leq EC(g)^{O(1)} \text{ ?} \tag{1} \\
L(g) &\leq C(g)^{O(1)} \text{ ?} \tag{2} \\
L(\varphi) &\leq C(\mathrm{graph}(\varphi))^{O(1)} \text{ ?} \tag{3}
\end{align}
$$

Again, g denotes the irreducible generator of a hypersurface in $\mathbb{R}^n$ or $\mathbb{C}^n$ and φ denotes a polynomial. We have the following chain of implications: $(1) \Rightarrow (2) \Rightarrow (3)$.

We believe that all these three questions have negative answers, but we have been unable to prove this for irreducible g. However, the following counterexamples are known for questions (1) and (2) when allowing for reducible polynomials g, and assuming $k = \mathbb{R}$ for question (2).

Referring to question (1), there exist univariate polynomials f having reducible factors g with a complexity exponential in the complexity of f, a fact first discovered by Lipton and Stockmeyer [32]. The simplest known example illustrating this is as follows: Consider $f_n = X^{2^n} - 1 = \prod_{j < 2^n}(X - \zeta^j)$, where $\zeta = \exp(2\pi i / 2^n)$. By repeated squaring we get $L(f_n) \leq n + 1$. On the other hand, one can prove that for almost all $M \subseteq \{0, 1, \ldots, 2^n - 1\}$ the random factor $\prod_{j \in M}(X - \zeta^j)$ has a complexity which is exponential in n, cf. [10, Exercise 9.8]. A similar reasoning can be made over the reals using Chebychev polyomials.

Commenting on question (2), we remark that the answer is negative if we drop the irreducibility assumption and assume $k = \mathbb{R}$. This follows from the following trivial example from [12], which shows that $L(g)$ may be exponentially larger than $C(g)$: Let $g_n \in \mathbb{R}[X]$ have n distinct real roots. Then $C(g_n) \leq \log n$ using binary search, but $L(g_n) \geq n$ if the roots of g are algebraically independent over $\mathbb{Q}$.

We regard to question (3), we note that its truth would imply that there are no "one-way functions" in the algebraic setting of computations with polynomials.

It would be interesting to find out whether the truth of the above questions is equivalent to the collapse of some complexity classes, similarly as P = UP in the bit-model.

1.2 Main Results

The counterexamples discussed above established polynomials g whose degree was exponential in the exclusion or decision complexity of g, respectively. We restrict now our attention to factors g having a degree polynomially bounded in the complexity of f.

The *Factor Conjecture* from [7, Conj. 8.3] states that for polynomials g

$$L(g) \leq (EC(g) + \deg g)^{O(1)}. \tag{4}$$

A partial step towards establishing this conjecture is an older result due to Kaltofen [22], which can be seen as a byproduct of his achievements [23] to factor polynomials given by a straight-line program representations (see also [25]). Kaltofen proved that the complexity of any factor g of f is polynomially bounded in the complexity of f and in the degree and the *multiplicity* of the factor g.

Before stating the precise result, let us fix some notation. For the remainder of this paper, $M(d)$ denotes an upper bound on the complexity for the multiplication of two univariate polynomials of degree d over k, that is, for computing the coefficients of the product polynomial from the coefficients of the given polynomials. It is well-known that $M(d) \leq O(d \log d)$ for $k = \mathbb{R}$ or $\mathbb{C}$, cf. [10]. We will assume that $M(d_1) \leq M(d_2)$ for $d_1 \leq d_2$ and the subadditivity property $M(d_1) + M(d_2) \leq M(d_1 + d_2)$.

Here is the precise statement of Kaltofen's result, which was independently found by the author, compare [7, Thm. 8.14].

Theorem 1.2 *Let $f = g^e h$ with coprime polynomials $g, h \in k[X_1, \dots, X_n]$. Let $d \geq 1$ be the degree of g. We suppose that k is a field of characteristic*

4

zero. Then we have

$$L(g) \leq O(M(d^3 e)(L(f) + d \log e)).$$

Thus our Factor Conjecture claims that the dependence on the multiplicity can be omitted. It is known [22] that this is true in the case $f = g^e$, in which case $L(g) \leq O(M(d)L(f))$ with $d = \deg g$, see Proposition 6.1 in the appendix.

The main result of this paper states that the dependence on the multiplicity can indeed be omitted when switching to an approximative complexity measure. The approximative complexity $\underline{L}(g)$ of a polynomial g is the minimal cost of "approximative straight-line programs" computing approximations of g with any precision required. A formal definition will be given in Section 2.

The precise formulation is as follows:

Theorem 1.3 *Let k be a field of characteristic zero. Assume that n by n matrices over k can be multiplied with $O(n^\gamma)$ arithmetic operations in k. For $g \in k[X_1, \ldots, X_n]$ of degree d we have*

$$\underline{L}(g) \leq O(M(d)M(d^4)EC(g) + d^{2\gamma}M(d)^2).$$

We remark that the "exponent γ of matrix multiplication" may be chosen as $2 \leq \gamma < 2.38$, see [13, 10].

Remark 1.4 There is certainly room for improvement in this bound. In fact, the proof of Theorem 1.3 yields better estimates in the following cases.

1. If g is the generator of the graph of a polynomial function φ, we obtain $\underline{L}(g) \leq O(M(d^2)EC(g))$.

2. We have $\underline{L}(g) \leq O(M(d^4)EC(g) + d^{2\gamma}M(d))$ if g is the irreducible generator of a hypersurface in $\mathbb{R}^n$ or $\mathbb{C}^n$.

An interesting consequence is the following degree bounded version of question (3):

Corollary 1.5 *The approximative complexity $\underline{L}(\varphi)$ of a polynomial φ is polynomially bounded in the decision complexity of the graph of φ and the degree d of φ, namely $\underline{L}(\varphi) \leq O(M(d^2) C(\mathrm{graph}(\varphi)))$. This remains true if we allow randomization with two-sided error.*

5

Coming back to the discussion of one-way functions, we remark that Sturtivant and Zhang [38] obtained the following related result, which excludes the existence of certain one-way functions in the algebraic framework of computation. Let $\psi \colon k^n \to k^n$ be bijective such that ψ as well as ψ^{-1} are polynomial mappings. Then the complexity to evaluate ψ is polynomially bounded in the complexity to evaluate the inverse ψ^{-1} and the maximal degree of the component functions of ψ. Again, it is unknown whether the degree restriction can be omitted.

The paper is organized as follows: In Section 2 we introduce the concept of approximative complexity. Section 3 contains the proof of the main result. We then shortly discuss some applications in Section 4, where we also build in the concept of approximative complexity into Valiant's algebraic P-NP framework [40, 42] (see also [10, 7]). Section 5 is devoted to a more detailed analysis of the concept of approximative complexity. Finally, the appendix contains a proof of Theorem 1.2.

For some other aspects of the issues discussed in this paper see [9].

2   Approximative Complexity

In complexity theory it has proven useful to study "approximative algorithms", which use arithmetic with infinite precision and nevertheless only give us an approximation of the solution to be computed, however with any precision required. This concept was systematically studied in the framework of bilinear complexity (border rank) and there it has turned out to be one of the main keys to the currently best known fast matrix multiplication algorithms [13]. We refer to [10, Chap. 15] and the references there for further information.

Although approximative complexity is a very natural concept, it has been investigated in less detail for computations of polynomials or rational functions. Originally, it had been introduced by Strassen in a topological

way [37]. Griesser [18] generalized most of the known lower bounds for multiplicative complexity to approximative complexity. Lickteig systematically studied the notion of approximative complexity with the goal of proving lower bounds [30]. In Grigoriev and Karpinski [19] the notion of approximative complexity is also employed for proving lower bounds.

It is not known how to meaningfully relate the complexity of trailing coefficients or of factors of a polynomial to the complexity of the polynomial itself. However, by allowing approximative computations, we are able to establish quite satisfactory reductions in these cases. The deeper reason why this is possible seems to be the lower semicontinuity of the approximative complexity, which allows a controlled passage to the limit and can be used in perturbation arguments.

Assume the polynomial f is expanded with respect to Y:

$$f = f_q(X_1, \ldots, X_n)Y^q + f_{q+1}(X_1, \ldots, X_n)Y^{q+1} + \ldots.$$

We do not know whether the complexity of the trailing coefficient f_q can be polynomially bounded in the the complexity of f. However, we can make the following observation. For the moment assume that k is the field of real or complex numbers. We have $\lim_{y \to 0} y^{-q} f(X, y) = f_q(X)$ and $L(f(X, y)) \leq L(f)$ for all $y \in k$. Thus we can approximate f_q with arbitrary precision by polynomials having complexity at most $L(f)$. We will say that f_q has "approximate complexity" at most $L(f)$.

In what follows, we will formalize this in an algebraic way; a topological interpretation will be given later. Throughout the paper, $K := k(\epsilon)$ is a rational function field in the indeterminate ϵ over the field k and R denotes the local subring of K consisting of the rational functions defined at $\epsilon = 0$. We write $F_{\epsilon=0}$ for the image of $F \in R[X]$ under the morphism $R[X] \to k[X]$ induced by $\epsilon \mapsto 0$.

Definition 2.1 Let $f \in k[X_1, \ldots, X_n]$. The *approximative complexity* $\underline{L}(f)$ of the polynomial f is the smallest natural number r such that there exists F in $R[X_1, \ldots, X_n]$ satisfying $F_{\epsilon=0} = f$ and $L(F) \leq r$. Here the complexity L is to be interpreted with respect to the larger field of constants K.

Even though L refers to division-free straight-line programs, divisions may occur implicitly since our model allows the free use of any elements of K as constants (e.g., division by powers of ϵ). In fact, the point is that even though F is defined with respect to the morphism $\epsilon \mapsto 0$, the intermediate results of the computation may not be so! Note that $\underline{L}(f) \leq L(f)$.

We remark that the assumption that any elements of K are free constants is just made for conceptual simplicity. We may as well require to build up the needed elements of K from ϵ, ϵ^{-1} and elements of k. It is easy to see that this would not change our main result (i.e., Theorem 1.3).

Assume that $\underline{L}(f) \leq r$ over $k = \mathbb{R}$, say $F_\epsilon(x) = f(x) + \epsilon R_\epsilon(x)$ and $L(F_\epsilon) \leq r$. Let S be the supremum of $R_\epsilon(x)$ over all $\epsilon \in [0, 1]$ and $x \in \mathbb{R}^n$ with $||x||_\infty \leq 1$. Then we have for such ϵ and x that $|F_\epsilon(x) - f(x)| = \epsilon |R_\epsilon(x)| \leq S\epsilon$. Therefore, for each $\epsilon > 0$ we can compute on input x an approximation to $f(x)$ with absolute error less than $S\epsilon$ with only r arithmetic operations. If we would additionally require in the definition of $\underline{L}$ to build up the needed constants in K from ϵ, ϵ^{-1}, then $\underline{L}(f) \leq r$ would even mean that one can compute an approximation with error less than $S\epsilon$ with only r arithmetic operations on input x *and* ϵ.

Example 2.2 Let us illustrate the notion of approximative complexity with an example. The convex hull of the support $\operatorname{supp} f$ of a polynomial

$$f = \sum_{a \in \operatorname{supp} f} c_a X_1^{a_1} \cdots X_n^{a_n} \qquad (c_a \neq 0)$$

is called the Newton polytope P of f. To a supporting hyperplane H of P we may assign the corresponding initial term polynomial

$$\operatorname{in}_H f := \sum_{a \in H \cap \operatorname{supp} f} c_a X_1^{a_1} \cdots X_n^{a_n}.$$

We claim that

$$\underline{L}(\operatorname{in}_H f) \leq L(f) + n + 1.$$

Indeed, we may obtain $\operatorname{in}_H(f)$ as a "degeneration" of f as follows. Assume that $\langle w, x \rangle - c = 0$ is the equation of H, say $\langle w, x \rangle \geq c$ on P. We can always achieve that $w \in \mathbb{Z}^n$, $c \in \mathbb{Z}$. Then we have

$$F := \epsilon^{-c} f(\epsilon^{w_1} X_1, \ldots, \epsilon^{w_n} X_n) = \sum_{a \in \operatorname{supp} f} c_a \epsilon^{\langle w, a \rangle - c} X_1^{a_1} \cdots X_n^{a_n} = \operatorname{in}_H f + O(\epsilon)$$

using the convenient, intuitive Big-Oh notation. Therefore, $F_{\epsilon=0} = \operatorname{in}_H f$ and $L(F) \leq L(f) + n + 1$ which proves our claim. (Recall that the powers of ϵ are considered as constants.)

The next lemma states some of the basic properties of $\underline{L}$.

Lemma 2.3   1. (Semicontinuity) *If F is defined over R and $f = F_{\epsilon=0}$, then $\underline{L}(f) \le \underline{L}(F)$. Note that the quantity $\underline{L}(F)$ is well-defined for a polynomial F over K (adjoining a further indeterminate to K).*

2. (Elimination of constants) *Let k_1 be a field extension of k of degree at most d and f be a polynomial over k. Then $\underline{L}(f) \le O(M(d)\,\underline{L}_{k_1}(f))$, where $\underline{L}_{k_1}(f)$ denotes the approximative complexity of f interpreted as a polynomial over k_1 (i.e., constants in k_1 may be used freely).*

3. (Transitivity) *The approximative complexity $\underline{L}(f \mid g)$ to compute f from g and the variables is defined in a natural way. We have $\underline{L}(f) \le \underline{L}(f \mid g) + \underline{L}(g)$, and an analogous inequality is true for the computation of several polynomials.*

Proof.   (1) We start with a general observation: Let Φ be a rational function in two variables ϵ, δ. We assume that Φ, viewed as a rational function in δ over $k(\epsilon)$, is defined at $\delta = 0$ with value $\Phi_{\delta=0}$. Moreover, we assume that the rational function $\Phi_{\delta=0}$ is defined at $\epsilon = 0$ with value $\lambda := (\Phi_{\delta=0})_{\epsilon=0}$. Then $\Phi(\epsilon, \epsilon^N)$ is defined at $\epsilon = 0$ with value λ for sufficiently large N. Indeed, if B is the denominator of Φ and $B(\epsilon, 0) = B_m \epsilon^m + O(\epsilon^{m+1})$, $B_m \ne 0$, then it is easy to check that it suffices to take $N > m$.

Let now $\Phi \in k(\epsilon, \delta)[X]$ be such that $\Phi_{\delta=0} = F$ and $L(\Phi) = \underline{L}(F)$. An optimal computation of Φ takes place in a finitely generated subring of $k(\epsilon, \delta)[X]$. The morphism $\delta \mapsto \epsilon^N$ is defined on this subring if N is chosen sufficiently large. Then we have $L(\phi) \le L(\Phi)$ for $\phi := \Phi(\epsilon, \epsilon^N, X)$. If N is chosen sufficiently large, we have $\phi_{\epsilon=0} = F_{\epsilon=0}$ by the observation at the beginning of the proof. This implies the claim.

(2) This follows easily from [7, Prop. 4.1(iii)].

(3) By definition there exists $G \in k(\epsilon)[X]$ such that $G_{\epsilon=0} = g$ and $\underline{L}(g) = L(G)$. Moreover, there exists $F \in k(\epsilon)[X]$ such that $F_{\epsilon=0} = f$ and

$$\underline{L}(f \mid g) = L(F \mid g).$$

Let Γ be an optimal straight-line program computing F from g, variables X_i, and constants in $k(\epsilon)$. We replace ϵ by a new indeterminate δ and denote the element thus corresponding to G by $G(\delta) \in k(\delta)[X]$ (abusing notation). If we replace the input g by $G(\delta)$, then the program Γ, using the same constants in $k(\epsilon)$ as Γ, will compute an element $\Phi \in k(\epsilon, \delta)[X]$. Clearly, $\Phi_{\delta=0} = F$.

Since the computation of Φ takes place in a finitely generated subring of $k(\epsilon, \delta)[X]$, the morphism $\delta \mapsto \epsilon^N$ is defined on this subring if N is chosen

large enough. If we denote the image of Φ under this morphism by ϕ and the image of $G(\delta)$ by $G(\epsilon^N)$, then we have

$$L(\phi \mid G(\epsilon^N)) \ \leq \ L(F \mid g) \ = \ \underline{L}(f \mid g).$$

Moreover, we clearly have $L(G(\epsilon^N)) \leq L(G) = \underline{L}(g)$. By the transitivity of L, we get $L(\phi) \leq \underline{L}(f \mid g) + \underline{L}(g)$. From the observation at the beginning of the proof of part (1) of the lemma, we conclude that $\phi_{\epsilon=0} = f$ for sufficiently large N. This implies the claim. $\qquad\square$

We proceed with a topological interpretation of approximative complexity, which points out the naturality of this notion from a mathematical point of view. It will not be needed for the proof of the main Theorem 1.3.

Assume k to be an algebraically closed field. There is a natural way to put a Zariski topology on the polynomial ring $A_n := k[X_1, \ldots, X_n]$ as a limit of the Zariski topologies on the finite dimensional subspaces $\{f \in A_n \mid \deg f \leq d\}$ for $d \in \mathbb{N}$. If k is the field of complex numbers, we may define the Euclidean topology on A_n in a similar way.

If $f \in A_n$ satisfies $\underline{L}(f) \leq r$, then it easy to see that f lies in the closure (Zariski or Euclidean) of the set $\{f \in A_n \mid L(f) \leq r\}$. Indeed, we have $L(F_{\epsilon=y}) \leq L(F)$ for all but finitely many $y \in k$ and $\lim_{y \to 0} F_{\epsilon=y} = F_{\epsilon=0} = f$. Alder [1] has shown that the converse is true and obtained the following topological characterization of the approximative complexity.

Theorem 2.4 *Let k be algebraically closed. The set $\{f \in A_n \mid \underline{L}(f) \leq r\}$ is the closure of the set $\{f \in A_n \mid L(f) \leq r\}$ for the Zariski topology. If $k = \mathbb{C}$, this is also true for the Euclidean topology.*

This essentially claims that $\underline{L}$ is the largest lower semicontinuous function of f bounded by $L(f)$. The proof of Theorem 5.7 in Section 5.2 implies the above result as a special case. We remark that this theorem can also be easily deduced from [10, Lemma 20.28]. One can show that the above statement is also true over the reals with the Euclidean topology, similar as in Lehmkuhl and Lickteig [26].

3   Approximative Complexity of Factors

We will supply here the proof of our main result Theorem 1.3. The outline of the proof is as follows: Let $f = g^e h$ with coprime g and h and assume $k = \mathbb{C}$. After a suitable coordinate transformation one can interpret the zeroset of the factor g locally as the graph of some analytic function φ. In order to

cope with a possibly large multiplicity e of g, we apply a small perturbation to the polynomial f without affecting its complexity too much. This results in a small perturbation of φ. We compute now the homogeneous parts of the perturbed φ by a Newton iteration up to a certain order. Using efficient polynomial arithmetic, this gives us an upper bound on the approximative complexity of the homogeneous parts of φ up to a predefined order (Proposition 3.4). In the special case, where the factor g is the generator of the graph of a polynomial function, we are already done. This is essentially the contents of Section 3.2.

In a second step, elaborated in Section 3.3, we view the factor g as the minimal polynomial of φ in $Y := X_n$ over the field $k(X_1, \ldots, X_{n-1})$. We show that the Taylor approximations up to order $2d^2$ uniquely determine the factor g and compute the bihomogeneous components of g with respect to the degrees in the X-variables and Y by fast linear algebra.

3.1   Preliminaries

The following result is obtained by a straightforward application of a technique introduced by Strassen [36] for the computation of homogeneous components and avoiding divisions. A proof will be sketched in Section 5.1.

Proposition 3.1 *Assume that $F(X,Y) = \sum_{i,\delta} F_i^{(\delta)} Y^i$ is the bihomogeneous decomposition of the polynomial $F \in k[X_1, \ldots, X_n, Y]$ with respect to the total degree in the X-variables and the degree Y. Thus $F_i^{(\delta)}$ is a homogeneous polynomial in the X-variables of degree δ. Then we have for all $D \geq 1$*

$$L(\{F_i^{(\delta)} \mid i, \delta \leq D\}) \ \leq \ O(M(D^2)L(F))$$

and the same is true if the complexity L is replaced by the approximative complexity $\underline{L}$.

Part (1) of the next lemma follows immediately from the well-known algorithms for the multiplication and division of univariate power series described in [10, §2.4] by interpreting the homogeneous components of a multivariate power series $f \in k[[X_1, \ldots, X_n]]$ as the T-adic coefficients of the transformed series $f(TX_1, \ldots, TX_n)$. Part (2) of this lemma is obtained from part (1) by applying Horner's rule.

Lemma 3.2   (1)   *We can compute the homogeneous parts up to degree D of the product $F \cdot G$ and of the quotient F/G (if $G(0) \neq 0$) of multivariate power series F and G from the homogeneous parts of F and G up to degree D by $M(D)$ arithmetic operations.*

11

(2) *Assume that the multivariate power series $F_0, \ldots, F_D$ and Φ are given by their homogeneous parts up to degree D. Then we can compute from this data the homogeneous parts of $\sum_{i=0}^{D} F_i \Phi^i$ up to degree D by $O(DM(D))$ arithmetic operations.*

We remark that D^2 nonscalar operations are needed for the composition problem (2) in the generic case. For proving this, we assume that we have just one variable and choose for Φ a constant power series: $\Phi = a$. Let $F_i = \sum_j F_{i,j} X^j$. The problem then reduces to the simultaneous evaluation of $\sum_{i \le D} F_{i,j} a^i$ for $j \le D$, a problem known to be of nonscalar complexity $(D+1)^2 - 1$, see [10, Exercise 6.2].

3.2   Approximative Computation of Graph

We need the following lemma.

Lemma 3.3 *Let $g, h \in k[X_1, \ldots, X_n]$ be coprime, g irreducible and $d := \deg g$. Then there is field extension k_1 of degree at most d over k and a point $p \in k_1^n$ such that*

$$g(p) = 0, \ h(p) \neq 0, \ \operatorname{grad} g(p) \neq 0.$$

Moreover, we may assume in this statement that $k_1 = k$ if either $k = \mathbb{C}$ or if $k = \mathbb{R}$ and g is the irreducible generator of a hypersurface in $\mathbb{R}^n$.

Proof.   The claim for $k = \mathbb{C}$ is a straightforward consequence of the Nullstellensatz. In the case $k = \mathbb{R}$ we apply Theorem 4.5.1 in [5], which tells us that $\{x \in \mathbb{R}^n \mid g(x) = 0, \operatorname{grad} g(x) \neq 0\}$ is Zariski dense in the zeroset of g and that the vanishing ideal of this zeroset is generated by g. This implies the claim.

In the general case, we apply a linear coordinate transformation $X_i \mapsto X_i + u_i Y$ $(i < n)$, $Y \mapsto vY$ for suitable $u_i, v \in k$ in order to achieve that g is monic of degree d with respect to the variable $Y := X_n$. From now on we write $X := (X_1, \ldots, X_{n-1})$. Since g is irreducible and g, h are coprime, the resultants $\operatorname{res}_Y(g, \partial_Y g)$ and $\operatorname{res}_Y(g, h)$ in $k[X]$ with respect to the variable Y are not the zero polynomials. We choose a point $\xi \in k^{n-1}$ where these resultants do not vanish. From the properties of the resultant we conclude that the univariate polynomials $\tilde{g} := g(\xi, Y)$ and $\tilde{h} := h(\xi, Y)$ are coprime and that $\tilde{g}$ is squarefree. Let η be a root of $\tilde{g}$ is some extension field k_1 of k of degree at most d. Then $\partial_Y \tilde{g}$ and $\tilde{h}$ do not vanish at η and the point $p = (\xi, \eta)$ satisfies the claim of the lemma. $\square$

We assume now that we are in the situation of Theorem 1.3. Without loss of generality we may assume that g is irreducible (apply Theorem 1.3 to the irreducible factors of g and use the subadditivity and monotonicity of M). From now on we use the notations $Y := X_n$ and $X := (X_1, \ldots, X_{n-1})$.

Let $f = g^e h$, where g and h coprime such that $EC(g) = L(f)$. We choose the field extension k_1 and the point $p = (\xi, \eta) \in k_1^n$ according to Lemma 3.3. To simplify notation, we assume that $k_1 = k$, an assumption which will be eliminated at the end of Section 3.3 at the price of an additional factor $M(d)$ in the complexity bound.

We are now going to transform the polynomials into a special form by suitable linear transformations. By a coordinate shift we can always achieve that $(\xi, \eta) = (0, 0)$. By a substitution $\tilde{g}(X, Y) := g(X_1 + u_1 Y, \ldots, X_{n-1} + u_{n-1} Y, vY)$ we may achieve that the degree of $\tilde{g}$ in Y equals d and that $\partial_Y \tilde{g}(0, 0)$ does not vanish. Indeed, if $g^{(d)}$ denotes the homogeneous component of g of degree d, then the coefficient of Y^d in $\tilde{g}$ equals $\tilde{g}^{(d)}(0, 1) = g^{(d)}(u, v)$. Moreover, $\partial_Y \tilde{g}(0, 0) = u_1 \partial_{X_1} g(0, 0) + \ldots + u_{n-1} \partial_{X_{n-1}} g(0, 0) + v \partial_Y g(0, 0)$. Hence it suffices to choose u, v such that this linear combination does not vanish and such that $g^{(d)}(u, v) \neq 0$. By scaling, we may assume without loss of generality that $\tilde{g}$ is monic with respect to Y. In the following, we will assume that this transformation has already been done, i.e., $\tilde{g} = g$, which results in a complexity increase of f of at most $2n$. Note that $L(f) \geq n$ if all the variables occur in f.

Summarizing, we achieved the following by a suitable choice of a linear transformation:

$$g(0, 0) = 0, \ h(0, 0) \, \partial_Y g(0, 0) \neq 0, \ \deg_Y g = d. \tag{5}$$

The implicit function theorem implies that there exists a unique formal power series $\varphi \in k[[X]]$ such that

$$g(X, \varphi(X)) = 0, \ \varphi(0) = 0. \tag{6}$$

Moreover, this power series can be recursively computed by the following *Newton iteration*: if we put $\varphi_0 = 0$ and define

$$\varphi_{\nu+1} = \varphi_\nu - \frac{g(X, \varphi_\nu)}{\partial_Y g(X, \varphi_\nu)}, \tag{7}$$

then we have quadratic convergence of the φ_ν towards φ, in the sense that $\varphi_\nu \equiv \varphi \mod (X)^{2^\nu}$, where (X) denotes the maximal ideal of $k[[X]]$ (cf. [10, Theorem 2.31]).

It is easy to see that if the partial derivative $\partial_Y f(0,0)$ would not vanish, then the above power series φ could also be recursively computed by the Newton recursion (7) with g replaced by f. However, $\partial_Y f(0,0) = 0$ always vanishes for multiplicities $e > 1$. The key idea is now to enforce the nonvanishing of this partial derivative by a suitable perturbation of the given polynomial f. By doing so, we have to content ourselves with an approximative computation of the factor g.

Based on these ideas, we prove the following assuming the conditions (5):

Proposition 3.4 *The homogeneous parts $\varphi^{(\delta)}$ of φ of degree δ satisfy*

$$\forall D \geq 1: \quad \underline{L}(\varphi^{(1)}, \ldots, \varphi^{(D)}) \; \leq \; O(M(D^2)\underline{L}(f)).$$

Proof. Note that g, viewed as a polynomial in Y over $k(X)$, is the minimal polynomial of φ over $k(X)$. W.l.o.g. we may assume that φ is not a rational function (otherwise $d = 1$, φ would be linear, and the claim obvious).

We define the perturbed polynomial $F(X, Y) := f(X, Y + \epsilon) - f(0, \epsilon)$ over the coefficient ring R. It is clear that $F(0,0) = 0$ and $F_{\epsilon=0} = f$. By a straight-forward calculation we get

$$\partial_Y F(0,0) = (eh\, \partial_Y g + g\, \partial_Y h)(0, \epsilon) \cdot g^{e-1}(0, \epsilon).$$

Assumptions (5) tell us that $g(0, \epsilon) = \lambda\epsilon + O(\epsilon^2)$ with $\lambda \in k^\times$, hence

$$\partial_Y F(0,0) = e\, \lambda^e\, h(0,0)\, \epsilon^{e-1} + O(\epsilon^e)$$

and we conclude that this partial derivative does not vanish (char $k = 0$).

As in the reasoning before, the implicit function theorem implies that there exists a unique formal power series Φ over the field $K = k(\epsilon)$ such that $F(X, \Phi(X)) = 0$, $\Phi(0) = 0$ and this power series can be recursively computed by the Newton iteration

$$\Phi_0 = 0, \;\; \Phi_{\nu+1} = \Phi_\nu - \frac{F(X, \Phi_\nu)}{\partial_Y F(X, \Phi_\nu)} \tag{8}$$

with quadratic convergence: $\Phi_\nu \equiv \Phi \mod (X)^{2^\nu}$.

Claim: Φ_ν is defined over the coefficient ring R for all ν.

We prove this claim by induction on ν, the induction start $\nu = 0$ being clear. So let us assume that Φ_ν is defined over R and set $\psi_\nu := (\Phi_\nu)_{\epsilon=0}$. By applying the morphism $R[[X]] \rightarrow k[[X]], \epsilon \mapsto 0$ we obtain

$$(\partial_Y F(X, \Phi_\nu))_{\epsilon=0} = \partial_Y f(X, \psi_\nu)$$
$$= (eh\, \partial_Y g + g\, \partial_Y h)(X, \psi_\nu) \cdot g^{e-1}(X, \psi_\nu).$$

The first parenthesis maps under the substitution $X \mapsto 0$ to $(eh\,\partial_Y g)(0,0)$, which is nonzero by our assumptions. The second factor $g(X, \psi_\nu)$ can only vanish if $\psi_\nu = \varphi$ since the power series φ is uniquely determined by the conditions (6). In this case, φ would be a rational function, which we have excluded at the beginning of the proof. We have thus shown that $(\partial_Y F(X, \Phi_\nu))_{\epsilon=0}$ nonzero. By equation (8) this implies that $\Phi_{\nu+1}$ is defined over R and proves the claim.

The claim implies that Φ is defined over R. From $F(X, \Phi(X)) = 0$ we get $f(X, (\Phi(X))_{\epsilon=0}) = 0$, hence $g(X, (\Phi(X))_{\epsilon=0}) = 0$, as $h(X, (\Phi(X))_{\epsilon=0}) \neq 0$. We conclude that $(\Phi(X))_{\epsilon=0} = \varphi$. If $\Phi^{(\delta)}$ denotes the homogeneous part of Φ of degree δ, we have $(\Phi_\nu)^{(\delta)} = \Phi^{(\delta)}$ for $\delta < 2^\nu$. This implies for $\delta < 2^\nu$ that

$$((\Phi_\nu)^{(\delta)})_{\epsilon=0} = (\Phi^{(\delta)})_{\epsilon=0} = (\Phi_{\epsilon=0})^{(\delta)} = \varphi^{(\delta)}.$$

As a word of warning, we point out that a certain care in these argumentations is necessary. For instance, Example 3.5 below shows that in general $(\Phi_\nu)_{\epsilon=0} \neq \varphi_\nu$.

We turn now to the algorithmic analysis of the proof. First of all we note that $L(F) \leq L(f)+2$. A moment's thought shows that also $\underline{L}(F) \leq \underline{L}(f)+2$. In order to prove the proposition it is enough to show that

$$L(\Phi_N^{(1)}, \ldots, \Phi_N^{(D)}) \ \leq \ O(M(D^2)\underline{L}(f)), \tag{9}$$

where $N := \lceil \log(D+1) \rceil$. In fact, by the semicontinuity of $\underline{L}$ (Lemma 2.3(1)), we only need to prove this estimate for approximative complexity on the lefthand side.

The following computation deals with polynomials in the X-variables, which are truncated at a certain degree and represented by their homogeneous parts up to this degree. We obtain from Proposition 3.1 for the bihomogeneous decomposition of F that

$$\underline{L}(\{F_i^{(\delta)} \mid i, \delta \leq D\}) \ \leq \ O(M(D^2)\underline{L}(F)). \tag{10}$$

In the following, we assume that we have already computed the bihomogeneous components $F_i^{(\delta)}$ for $i, \delta \leq D$.

Inductively, we suppose that we have computed the homogeneous parts of Φ_ν up to degree 2^ν. The main work of one Newton step (8) consists in the computation of the substituted polynomials $F(X, \Phi_\nu)$ and $\partial_Y F(X, \Phi_\nu)$. By Lemma 3.2 we can compute the homogeneous parts up to degree $2^{\nu+1}$ of $F(X, \Phi_\nu)$ by $O(2^\nu M(2^\nu))$ arithmetic operations. Analogously, we get the homogeneous parts up to degree $2^{\nu+1}$ of $\partial_Y F(X, \Phi_\nu)$ by the same number of

15

arithmetic operations. By a division and a subtraction we obtain from this the homogeneous parts of $\Phi_{\nu+1}$ up to degree $2^{\nu+1}$ using further $O(M(2^\nu))$ arithmetic operations. Altogether, we obtain

$$L(\Phi_N^{(1)}, \ldots, \Phi_N^{(D)} \mid \{F_i^{(\delta)} \mid i, \delta \le D\})$$
$$\le\ O(\sum_{\nu=0}^{N} 2^\nu M(2^\nu))\ \le\ O(DM(D))\ \le\ O(M(D^2)),$$

by the monotonicity and subadditivity of M. The assertion (9) follows from this estimate and equation (10) by the transitivity of approximative complexity (Lemma 2.3(3)). □

Example 3.5 Consider the bivariate polynomial $g := (1 + Y)^2 - 1 - X^2$ and put $f = g^2$, $h = 1$. Then the conditions (5) are satisfied. The first Newton iterate according to (7) satisfies $\varphi_1 = \frac{1}{2}X^2$ and the power series φ defined by (6) has the expansion

$$\varphi = \frac{1}{2}X^2 - \frac{1}{8}X^4 + \cdots.$$

As in the proof of Proposition 3.4 we set $F := f(X, Y + \epsilon) - f(0, \epsilon)$. A straightforward computation (e.g., using a computer algebra system) yields for the first Newton approximation Φ_1 according to (8) that

$$\Phi_1 = -\frac{1}{4}\frac{(-X^2 + 2\epsilon + \epsilon^2)^2 - \epsilon^2(2 + \epsilon)^2}{(-X^2 + 2\epsilon + \epsilon^2)(1 + \epsilon)}.$$

Therefore, $(\Phi_1)_{\epsilon=0} = \frac{1}{4}X^2 \ne \varphi_1$. On the other hand, we note that the expansion of Φ_1 starts as follows

$$\Phi_1 = \frac{1}{2(1 + \epsilon)}X^2 + \frac{1}{4\epsilon(1 + \epsilon)(2 + \epsilon)}X^4 + \cdots$$

and we see that $(\Phi_1^{(2)})_{\epsilon=0} = \frac{1}{2}X^2 = \varphi_1^{(2)}$. Note that the fourth order term of this expansion is not defined for $\epsilon = 0$ even though Φ_1 is defined under this substitution!

3.3 Reconstruction of Minimal Polynomial

Consider the bihomogeneous decomposition $g(X, Y) = \sum_{i,\alpha \le d} g_i^{(\alpha)} Y^i$. Let T be an additional indeterminate and perform the substitution $X_j \mapsto TX_j$.

16

The condition $g(X, \varphi(X)) = 0 \bmod (X)^{D+1}$ then translates to

$$\sum_{i,\alpha \leq d} g_i^{(\alpha)} T^\alpha (\sum_{\delta \leq D} \varphi^{(\delta)} T^\delta)^i \equiv 0 \bmod T^{D+1}$$

for any $D \geq 1$. Moreover, we have $g_d^{(0)} = 1$ and $g_d^{(\alpha)} = 0$ for $0 < \alpha \leq d$, since g is monic of degree d in Y. The next lemma states that these conditions uniquely determine the bihomogeneous components of g if we choose $D \geq d^2$. The proof is based on well-known ideas from the application of the LLL-algorithm to polynomial factoring [28] (see also [17, Lemma 16.20]), adapted from $\mathbb{Z}$ to the setting of a polynomial ring.

Lemma 3.6 *By comparing the coefficients of the powers of the indeterminate T, one can interpret the conditions*

$$\sum_{i,\alpha \leq d} Z_{i,\alpha} T^\alpha (\sum_{\delta \leq 2d^2} \varphi^{(\delta)} T^\delta)^i \equiv 0 \bmod T^{2d^2+1},$$
$$Z_{d,0} = 1, Z_{d,1} = 0, \ldots, Z_{d,d} = 0$$

as a system of linear equations over the field $k(X)$ in the unknowns $Z_{i,\alpha}$. (There are $2d^2 + 1$ equations and $(d+1)^2$ unknowns). This linear system has as the unique solution the bihomogeneous components $Z_{i,\alpha} = g_i^{(\alpha)}$ of g.

Proof. We define the bivariate polynomial $A(T, Y) := \sum_{i,\alpha \leq d} g_i^{(\alpha)} T^\alpha Y^i$ over $k(X)$ and assign to a solution $\zeta_{i,\alpha}$ of the above linear system of equations the bivariate polynomial $B(T, Y) := \sum_{i,\alpha \leq d} \zeta_{i,\alpha} T^\alpha Y^i$. Note that A is an irreducible polynomial in Y over $k(X, T)$ since we assume g to be irreducible and monic with respect to Y. The polynomial $\psi := \sum_{\delta \leq 2d^2} \varphi^{(\delta)} T^\delta$ is an approximative common root of A and B in the sense that

$$A(T, \psi) \equiv 0 \bmod T^{2d^2+1}, \quad B(T, \psi) \equiv 0 \bmod T^{2d^2+1}.$$

The resultant $\mathrm{res}(A, B) \in k(X)[T]$ of A and B with respect to Y satisfies the degree estimate

$$\deg_T \mathrm{res}(A, B)) \leq \deg_Y A \cdot \deg_T B + \deg_T A \cdot \deg_Y B \leq 2d^2,$$

which is easily seen from the description of the resultant as the determinant of the Sylvester matrix (cf. [17, §6.3]). It is well-known that there exist polynomials $u, v \in k(X)[T, Y]$ such that $uA + vB = \mathrm{res}(A, B)$. Substituting the approximative common root ψ for Y in this equation implies that $\mathrm{res}(A, B) \equiv 0 \bmod T^{2d^2+1}$, hence the resultant vanishes. Since A is irreducible, it must be a factor of B over $k(X, T)$. However, we assume A and B both to be monic with respect to Y. This implies that in fact $A = B$ as claimed. $\qquad\square$

The coefficients of the linear system of equations in Lemma 3.6 can be computed from the homogenous components $\varphi^{(\delta)}$, $\delta \leq 2d^2$, with d multiplications of power series given by their coefficients up to degree $2d^2$. This can be done with $O(dM(d^2))$ arithmetic operations (Lemma 3.2).

Assume that n by n matrices can be multiplied with $O(n^\gamma)$ arithmetic operations. Then we can compute from the coefficients of the linear system the unique solution with $O(d^{2\gamma})$ operations (see [10, Chap. 16]). This computation can be interpreted as a straight-line program involving divisions. However, as the bihomogeneous components of g we are seeking for are homogenous of degree at most d, we can apply Strassen's idea of avoiding divisions [36] and transform this straight-line program into one without divisions, which is at most by a factor of $O(M(d))$ longer. Summarizing, we obtain the following:

$$ L(\{g_i^{(\delta)} \mid i, \delta \leq d\} \mid \varphi^{(1)}, \ldots, \varphi^{(2d^2)}) \ \leq \ O(d^{2\gamma}M(d)). \qquad (11) $$

Our main Theorem 1.3 is a consequence of this estimate and Proposition 3.4. In fact, this provides an upper bound on $\underline{L}(g)$ with respect to the field extension k_1 of k of degree at most d considered at the beginning of Section 3.2. To simplify notation, we assumed there that $k_1 = k$. This assumption can now be eliminated at the price of an additional factor $M(d)$ in the complexity bound according to Lemma 2.3(2). As noted in the proof, we may directly take $k_1 = k$ in the cases $k = \mathbb{R}$ or $\mathbb{C}$, so that this additional factor is not necessary in these cases. Moreover, note that if g is the generator of the graph of a polynomial function, we obtain the improved bound stated in Remark 1.4 directly from Proposition 3.4. Summarizing, we have now provided the proof of the main Theorem 1.3 as well as of Remark 1.4.

Remark 3.7 Alternatively, one can compute the bihomogeneous components of g by an analogue of the LLL-algorithm applied to the lattice

$$ \{A \in k(X)[T, Y] \mid \deg_Y A \leq d, A(T, \psi) \equiv 0 \bmod T^{2d^2+1}\} \simeq R + RY + \cdots R^d $$

over the principal ideal domain $R := k(X)[T]$. The complexity bound resulting from this approach is $O(d^4 \cdot d^2)$ operations in R (cf. [15, Thm. 4.8] or [27]). This results in $O(d^6 M(d))$ operations in $k(X)$, including divisions, which is worse than the bound $O(d^{2\gamma})$ of Proposition 11 for Gaussian elimination ($\gamma = 3$).

We think that an improvement upon the bound of inequality (11) is possible by taking account of the structure of the linear system of equations under consideration, based on the ideas of Wiedemann [43, 24]. This reduces

to the question of how fast the matrix underlying the above linear system can be multiplied with a vector. This is an interesting problem in its own right, which will be addressed elsewhere.

4   Applications to Decision Complexity

By combining Theorem 1.3, Remark 1.4, and Lemma 1.1, we obtain the following corollary.

Corollary 4.1 *Let g be the generator of an irreducible hypersurface in $\mathbb{R}^n$ or $\mathbb{C}^n$ of degree d. Assume that n by n matrices can be multiplied with $O(n^\gamma)$ arithmetic operations. Then we have*

$$\underline{L}(g) \ \leq \ O(M(d^4)C(g) + d^{2\gamma}M(d)).$$

We remark that if the hypersurface is the graph of a polynomial function, then we obtain the better bound $\underline{L}(g) \leq O(M(d^2)C(g))$. This implies Corollary 1.5 of the introduction for deterministic decision complexity. The claim about randomized complexity (formalized by randomized algebraic computation trees) then follows easily by the results in [33, 11, 14].

In [40, 42] Valiant had proposed an analogue of the theory of NP-completeness in a framework of algebraic complexity, in connection with his famous hardness result for the permanent [41]. This theory features algebraic complexity classes VP and VNP as well as VNP-completeness results for many families of generating functions of graph properties, the most prominent being the family of permanents. There is rather strong evidence for Valiant's hypothesis VP $\neq$ VNP. In fact, if it were false, then the nonuniform versions of the complexity classes NC and PH would collapse [8]. For a comprehensive presentation of this theory, we refer to [16, 10, 7]. In the following, we assume some basic familiarity with the concepts introduced there.

It is quite natural to incorporate the concept of approximative complexity into Valiant's framework.

Definition 4.2 An *approximatively p-computable family* is a *p*-family (f_n) such that $\underline{L}(f_n)$ is a *p*-bounded function of n. The complexity class $\underline{\text{VP}}$ comprises all such families over a fixed field k.

It is obvious that VP $\subseteq \underline{\text{VP}}$. If the polynomial f is a projection of a polynomial g, then we clearly have $\underline{L}(f) \leq \underline{L}(g)$. Therefore, the complexity

class $\underline{VP}$ is closed under p-projections. We remark that $\underline{VP}$ is also closed under the polynomial oracle reductions introduced in [6].

We know very few about the relationship between the complexity classes VP, $\underline{VP}$, and VNP. We therefore raise the following question:

Problem 4.3 Is the class VP strictly contained $\underline{VP}$?

Intuitively, one would think that $\underline{VP}$ should not differ too much from VP. Commenting on this, we remark that by Lemma 5.6(3), an improvement of Theorem 5.7 of the form $\max\{q, \underline{L}_q(f)\} \leq (\underline{L}(f) + \deg f)^{O(1)}$ would imply that $VP = \underline{VP}$. However, we do not see how to achieve such an improvement.

The class VNP is closed under taking coefficients (cf. [7, §2.3]). This makes it plausible that $\underline{VP}$ is contained in VNP. Nevertheless, this is not clear as the occuring polynomials might have a degree exponential in ϵ.

The hypothesis
$$VNP \not\subseteq \underline{VP} \tag{12}$$
is a strengthening of Valiant's hypothesis, which is equivalent to saying that VNP-complete families are not approximately p-computable.

This hypothesis should be compared with the known work on polynomial time deterministic or randomized approximation algorithms for the permanent of non-negative matrices [31, 2, 21]. Based on the Markov chain approach, Jerrum, Sinclair and Vigoda [21] have recently established a fully-polynomial randomized approximation scheme for computing the permanent of an arbitrary real matrix with non-negative entries. We note that this result does not contradict hypothesis (12), since the above mentioned algorithm works only for matrices with *non-negative* entries, while approximative straight-line programs a fortiori work on all real inputs.

Under the hypothesis $VNP \not\subseteq \underline{VP}$, we can conclude from Corollary 4.1 that checking the values of polynomials forming VNP-complete families is hard, even when we allow randomized algorithms with two-sided error.

Corollary 4.4 *Assume* $VNP \not\subseteq \underline{VP}$ *over a field k of characteristic zero. Then for any* VNP-*complete family* (g_n), *checking the value* $y = g_n(x)$ *cannot be done by deterministic or randomized algebraic computation trees with a polynomial number of arithmetic operations and tests in n.*

Hypothesis (12) implies a separation of complexity classes in the Blum-Shub-Smale model of computation [4]. See [3] for the definition of the classes $P_{\mathbb{R}}$ and $PAR_{\mathbb{R}}$. (For the proof use Corollary 4.1 with the permanent polynomial g.)

Corollary 4.5 *If* $\mathrm{VNP} \not\subseteq \underline{\mathrm{VP}}$ *is true, then we have* $\mathrm{P}_{\mathbb{R}} \neq \mathrm{PAR}_{\mathbb{R}}$ *in the Blum-Shub-Smale model over the reals.*

5 Properties of Approximative Complexity

We perform here a more detailed analysis of the concept of approximative complexity. The results of this section are not needed for understanding the main results of the paper. The field k may here also be of positive characteristic.

5.1 Trailing p-adic Coefficients

We discuss first a result about the complexity to compute the p-adic expansion of a polynomial, which is related to Proposition 3.1 and proved in a similar way.

Let A be a commutative algebra over the field k and $p \in A[Y]$ be a fixed monic polynomial of degree $d \geq 1$. Any polynomial $f \in A[Y]$ has a unique p-adic expansion $f = \sum_{i \geq 0} f_i p^i$, where $f_i = \sum_{\mu < d} f_{i,\mu} Y^\mu \in A[T]$ is of degree strictly less than d. We will write

$$C_D^p(f) := \{f_{i,\mu} \mid i \leq D, \mu < d\} \subseteq A$$

for the set of coefficients of the p-adic coefficients of f up to order D.

Lemma 5.1 *For $f, g \in A[T]$ we have*

$$L(C_D^p(f \cdot g) \mid C_D^p(f, g)) \ \leq \ O(M(Dd)).$$

Proof. Let $f = \sum_{i,\mu} f_{i,\mu} Y^\mu p^i$ and $g = \sum_{j,\nu} g_{j,\nu} Y^\nu p^j$ be the p-adic expansions of f and g. Assume that

$$\sum_{\ell \leq 2D, \lambda < 2d-1} h_{\ell,\lambda} Y^\lambda U^\ell = \left(\sum_{i \leq D, \mu < d} f_{i,\mu} Y^\mu U^i \right) \left(\sum_{j \leq D, \nu < d} g_{j,\nu} Y^\nu U^j \right),$$

where U is a new indeterminate. The coefficients $h_{\ell,\lambda}$ can be computed by bivariate polynomial multiplication with $O(M(Dd))$ operations. Put $h_\ell := \sum_{\lambda < 2d-1} h_{\ell,\lambda} Y^\lambda$.

Suppose that $\sum_{\ell \leq 2D+1} \overline{h}_\ell p^\ell$ is the p-adic expansion of $f \cdot g = \sum_{\ell \leq 2D} h_\ell p^\ell$. It is easy to see that the $\overline{h}_0, \ldots, \overline{h}_{2D+1}$ can be obtained from the $h_0, \ldots, h_{2D}$ by $2D$ divisions with remainder by p. Such a division with remainder can be performed with $M(d)$ arithmetic operations in A (cf. [10, Cor. 2.26]). Therefore, the coefficients of $\overline{h}_0, \ldots, \overline{h}_{2D+1}$ can be obtained from the coefficients of $h_0, \ldots, h_{2D}$ with $O(DM(d))$ arithmetic operations in A. $\square$

The next proposition shows that the computation of the (coefficients of the) p-adic coefficients of a polynomial up to a certain order D is not much harder than the computation of the polynomial.

Proposition 5.2 *For $D \geq 1$ we have*

$$L(C_D^p(f)) \leq O(M(Dd)\, L(f)).$$

Proof. Let $g_1, \ldots, g_r \in A[Y]$ be the sequence of intermediate results of a computation of f. Suppose that $g_\rho = g_i \cdot g_j$, $i, j < \rho$. By Lemma 5.1, we can compute the elements of $C_D^p(g_\rho)$ from the elements of $C_D^p(g_i, g_j)$ using $O(M(Dd))$ arithmetic operations in A. If $g_\rho = g_i \pm g_j$, then we can clearly do this with $O(Dd)$ operations. In this way, we can successively compute the elements in $C_D^p(g_1), \ldots, C_D^p(g_r)$ with the required number of arithmetic operations in A. □

We note that the statement of Proposition 5.2 does also hold for approximative complexity $\underline{L}$. (The proof is obvious.) We remark that Proposition 3.1 of Section 3.1 may be derived from the above Proposition 5.2 by applying to $F(X,Y)$ the substitution $X_i \mapsto Y X_i$, $Y \mapsto p := Y^{d_X+1}$, where $d_X = \deg_X F$, and by taking $A = k[X_1, \ldots, X_n]$. (Of course, it can also be derived directly.)

Our initial motivation for the introduction of approximative complexity was the study of trailing coefficients. We come now back to this issue in a more general setting.

In the following let $A = k[X_1, \ldots, X_n]$ and $p \in A[Y]$ be monic of degree $d \geq 1$. Let $f = \sum_i f_i p^i$ be the p-adic expansion of $f \in A[Y]$. By Proposition 5.2 we know that the complexity of the p-adic coefficient polynomial f_i of Y^i is polynomially bounded in d, i, and $L(f)$. The following proposition essentially going back to Valiant [42] shows that the dependence on the degree i cannot be avoided in general.

Proposition 5.3 *The complexity of the coefficient polynomials in the p-adic expansion of a polynomial f with respect to a polynomial p is not polynomially bounded in $L(f)$ and $\deg p$, unless Valiant's hypothesis is false.*

Proof. We take $p = Y$ and consider the Y-adic expansion of the following polynomial f_n of complexity $L(f_n) \leq O(n^2)$:

$$f_n := \prod_{i=1}^{n} \left(\sum_{j=1}^{n} X_{ij} Y^{2^{j-1}} \right) = \sum_i f_{n,i}(X) Y^i.$$

The coefficient $f_{n,2^n-1}(X)$ equals the sum over all products $X_{1j_1} \cdots X_{nj_n}$ such that $\{j_1, \ldots, j_n\} = \{1, 2, \ldots, n\}$. That is, $f_{n,2^n-1}(X)$ equals the permanent $\mathrm{PER}_n(X)$ of the matrix $[X_{ij}]$. An estimate as claimed in the proposition would imply that $L(\mathrm{PER}_n(X)) \leq n^{O(1)}$, which contradicts Valiant's hypothesis. □

Assume now that the p-adic expansion $f = f_q p^q + f_{q+1} p^{q+1} + \ldots$ starts at order q ($f_q \neq 0$). We call f_q the *trailing coefficient* of f with respect to the base p. By contrast with Proposition 5.3, we can say the following about the approximative complexity of the trailing coefficient in relation to the complexity of f.

Proposition 5.4 *The approximative complexity of the trailing coefficient f_q with respect to p is polynomially bounded in $d = \deg p$ and $\underline{L}(f)$; we have*

$$\underline{L}(f_q) \leq O(M(d)\underline{L}(f)).$$

Proof. By the semicontinuity of $\underline{L}$ (Lemma 2.3(1)) it is sufficient to prove the statement for $L(f)$ on the right-hand side. Let $K = k(\epsilon)$ and R be as usual. We have $f = f_q p^q + u p^{q+1}$ with some $u \in k[X, Y]$, hence $f \equiv \epsilon^q f_q + \epsilon^{q+1} u \mod (p - \epsilon)$. Let $\rho(u) \in k[X, Y]$ denote the remainder of u by division with $p - \epsilon$ (viewed as a polynomial in Y). Then we conclude that $\rho(f) = \epsilon^q(f_q + \epsilon\rho(u))$. From the definition of approximative complexity we obtain

$$\underline{L}(f_q) \leq L(\epsilon^{-q}\rho(f)) \leq 1 + L(\rho(f)).$$

On the other hand, we conclude from Proposition 5.2 that $\rho(f)$ can be computed with $O(M(d)L(f))$ arithmetic operations. This proves the claim.
□

Note that the main reason for us to work with approximative complexity is that we do not know whether a statement similar to Prop. 5.4 does hold for complexity (compare Problem 4.3).

5.2 Further Characterizations

In order to investigate the relationship between $\underline{L}$ and L, it is useful to introduce a variant $\underline{L}_\infty$ of approximative complexity, which differs from $\underline{L}$ at most by a factor of two.

Definition 5.5 The *approximative complexity* $\underline{L}_q(f)$ *of order* $q \in \mathbb{N}$ of a polynomial f in $k[X_1, \ldots, X_n]$ is the smallest natural number r such that

there exists $f' \in k[[\epsilon]][X_1, \ldots, X_n]$ satisfying

$$L(\epsilon^q f + \epsilon^{q+1} f') \leq r,$$

where L refers here to the total (division-free) complexity in the polynomial ring $k[[\epsilon]][X_1, \ldots, X_n]$ with free constants in the ring of formal power series $k[[\epsilon]]$. Moreover, we define the *modified approximative complexity* $\underline{L}_\infty(f) := \min_q \underline{L}_q(f)$.

The following lemma summarizes some of the basic properties of this notion. The field $k((\epsilon))$ of formal Laurent series is defined as the quotient field of $k[[\epsilon]]$.

Lemma 5.6   (1) *We have* $\frac{1}{2}\underline{L}_\infty(f) \leq \underline{L}(f) \leq \underline{L}_\infty(f) + 1$.

(2) *In Definition 5.5 one can equivalently work with the polynomial ring $k[\epsilon]$ instead of with the coefficient ring $k[[\epsilon]]$ of formal power series. In Definition 2.1 one can equivalently work with $R = k[[\epsilon]]$ and $K = k((\epsilon))$.*

(3) *We have* $L(f) \leq O(M(q) \underline{L}_q(f))$.

Proof.   (1) For proving the right-hand estimate of (1), we assume that we have an optimal straight-line program of length $r = \underline{L}_\infty(f) = \underline{L}_q(f)$ computing $\epsilon^q f + \epsilon^{q+1} f'$ in $k[[\epsilon]][X]$. We can execute this straight-line program in $k[[\epsilon]][X]/(\epsilon^{q+1}) \simeq k[\epsilon][X]/(\epsilon^{q+1})$ by applying the canonical projection. By interpreting this computation back in $k[\epsilon][X]$ we obtain that $L(\epsilon^q f + \epsilon^{q+1} f'') \leq r$ for some suitable $f'' \in k[\epsilon][X]$, where L refers here to $k[\epsilon][X]$. We multiply the result with the constant ϵ^{-q} (this is the only computational step leading outside $k[\epsilon][X]$) and conclude that $\underline{L}(f) \leq r+1$.

For proving the left-hand estimate of (1), we use the embedding of $K = k(\epsilon)$ in the field $k((\epsilon))$ of formal Laurent series. This leads to an embedding $K[X] \hookrightarrow k((\epsilon))[X]$. The elements of $k((\epsilon))[X]$ can be written in the form $\epsilon^{-\alpha} \cdot A$ with $\alpha \in \mathbb{N}$ and $A \in k[[\epsilon]][X]$. Note that for $\alpha \geq \beta$

$$\epsilon^{-\alpha} A \pm \epsilon^{-\beta} B = \epsilon^{-\alpha}(A \pm \epsilon^{\alpha-\beta} \cdot B), \quad \epsilon^{-\alpha} A \cdot \epsilon^{-\beta} B = \epsilon^{-\alpha-\beta}(A \cdot B).$$

If we encode an element $\epsilon^{-\alpha} \cdot A$ by the constant $\epsilon^{-\alpha}$ and the polynomial A over the ring $k[[\epsilon]]$, then we can simulate any division-free straight-line computation in $k((\epsilon))[X]$ of length r by a straight-line computation in $k[[\epsilon]][X]$ of length at most $2r$. (The number of nonscalar multiplications even remains the same.) This way, a computation of $F = f + \epsilon f'$ in $k(\epsilon)[X]$ with

$f' \in k[[\epsilon]][X]$ will lead to a computation in $k[[\epsilon]][X]$ of some C such that $\epsilon^{-\gamma} C = F$ for some $\gamma \in \mathbb{N}$, hence $C = \epsilon^{\gamma} f + \epsilon^{\gamma+1} f'$.

(2) This follows from the first part of the proof of part (1).

(3) This is a consequence of part (2) and Proposition 5.2 applied to compute ϵ-adic coefficients. $\qquad\square$

Part (3) of the above lemma provides a polynomial bound on the complexity in terms of the approximative complexity of a certain order of approximation q and this order q. Unfortunately, the best general upper bound on the order q, that we are able to prove, is exponential in the complexity.

Theorem 5.7 *For polynomials f over an algebraically closed field k we have $\underline{L}_q(f) \leq 2\underline{L}(f)$ with $q \leq 4^{\underline{L}(f)}$.*

Proof. We proceed as in Lehmkuhl and Lickteig [26], who proved a similar bound on the order of approximation for border rank (approximative bilinear complexity).

The proof is based on the following geometric description of the set $\{f \in A_n \mid L(f) \leq r\}$. The field k is assumed to be algebraically closed. A straight-line program Γ is a description for a computation of a polynomial from constants $z_1, \ldots, z_m$ and variables $X_1, \ldots, X_n$ (recall that we do not allow divisions). Let $\phi_\Gamma(z)$ denote the polynomial in $A_n := k[X_1, \ldots, X_n]$ computed by Γ from the list of constants $z \in k^m$ and write $\mathcal{C}_\Gamma$ for the image of ϕ_Γ. The map ϕ_Γ can be interpreted as a morphism of affine varieties $k^m \to \{f \in A_n \mid \deg f < 2^{r_*}\}$, where r_* denotes the number of multiplication instructions of Γ. Hence the images $\mathcal{C}_\Gamma$ are irreducible, constructible sets. By Bézout's inequality, the geometric degree of the graph of ϕ_Γ satisfies $\deg \operatorname{graph}(\phi_\Gamma) \leq 2^{r_*}$ (compare [10, §8.3]). We have for fixed r that

$$\{f \in A_n \mid L(f) \leq r\} = \bigcup_\Gamma \mathcal{C}_\Gamma,$$

where the union is over all straight-line programs Γ of length r.

Assume now that f is in the Zariski-closure of the set on the left-hand side. Then we have $f \in \overline{\mathcal{C}_\Gamma}$ for some Γ. (We remark that in the case $k = \mathbb{C}$ the Zariski-closure of constructible sets coincides with the closure with respect to the Euclidean topology (cf. [34, Theorem 2.33]).)

We apply now two results proven in Lehmkuhl and Lickteig [26] to the morphism ϕ_Γ. Proposition 1 of [26] claims that there exists an irreducible curve $C \subseteq k^m$ such that $f \in \overline{\phi_\Gamma(C)}$ and $\deg C \leq \deg \operatorname{graph}(\phi_\Gamma)$. The Corollary to Proposition 3 in [26] states that there exists a point

$\zeta = (\zeta_1, \ldots, \zeta_m) \in k((\epsilon))^m$ such that $F := \phi_\Gamma(\zeta)$ is defined over $k[[\epsilon]]$, satisfies $F_{\epsilon=0} = f$ and such that all formal Laurent series ζ_i have order at least $-\deg C$. We conclude with Lemma 5.6(2) that $L(F) \leq r$ and hence $\underline{L}(f) \leq r$, which proves the nontrivial direction of Theorem 2.4. Moreover, we have shown that there is a straight-line program of length r, which computes F in $k((\epsilon))[X]$ from the X-variables and constants ζ_i having order at least $-\deg C \geq -2^{r_*}$. By a similar reasoning as in the proof of Lemma 5.6(1) we can construct from this a straight-line program of length at most $2r$, which computes in $k[[\epsilon]][X]$ an element of the form $\epsilon^q f + \epsilon^{q+1} f'$ with $q \leq 2^{r_*} \deg C \leq 4^{r_*} \leq 4^{\underline{L}(f)}$. We therefore have $\underline{L}_q(f) \leq 2r$, which finishes the proof of Theorem 5.7. $\qquad\square$

By tracing the proofs of the above results it is straightforward to show the following statement.

Remark 5.8 By counting only nonscalar multiplications, one can introduce the notions $\underline{L}^{ns}, \underline{L}_q^{ns}$ in an analogous way. We then have $\underline{L}^{ns} = \underline{L}_\infty^{ns} = \underline{L}_q^{ns}$ with $q \leq 4^{\underline{L}^{ns}}$.

Finally, we show that the restriction to division-free straight-line programs in the definition of approximative complexity is not a serious one.

Lemma 5.9 *If $\underline{L}'(f)$ denotes the approximative complexity of a multivariate polynomial f of degree d, where divisons are allowed, then the divison-free approximative complexity $\underline{L}(f)$ can be bounded by $\underline{L}(f) \leq O(M(d)\underline{L}'(f))$. Here the ground field k is assumed to be infinite.*

Proof. Formally, $\underline{L}'(f) \leq r$ means there exists $F \in R(X)$ with $F_{\epsilon=0} = f$ and such that the complexity of F in $K(X)$ (allowing divisions) is at most r. We can avoid the divisions using the well-known ideas of [36]. Accordingly, there exists $\xi \in k^n$ such that an optimal computation of F takes places in the local subring $\mathcal{O}_\xi$ of $K(X)$ consisting of the rational functions defined at $X = \xi$. To simplify notion, we assume without loss of generality that $\xi = 0$. Let $F^{(\delta)}$ denote the homogeneous component of F of degree δ. By [10, Theorem 7.1], the division-free complexity L of these components of degree up to d satisfies $L(F^{(0)}, \ldots, F^{(d)}) = O(M(d)r)$. As F is defined over R, also its homogeneous components are defined over R and we have $(F^{(\delta)})_{\epsilon=0} = f^{(\delta)}$. This implies that $\underline{L}(f^{(0)}, \ldots, f^{(d)}) = O(M(d)r)$ as claimed. $\qquad\square$

6 Appendix

We include here the proofs of Theorem 1.2 and Proposition 6.1. Although being essentially the same as the original proofs in [22] (with minor improvements in complexity), we believe that our exposition, integrated in the coherent framework of this paper, will facilitate the reader's understanding of the difficulties encountered in extending these results to Theorem 1.3.

Proof of Theorem 1.2. Let $f = g^e h$ with coprime polynomials $g, h \in k[X_1, \ldots, X_n]$ and $d = \deg g$. As in the proof of Lemma 3.3 we may achieve by a linear coordinate transformation that g is monic of degree d with respect to the variable $Y := X_n$. We put $X := (X_1, \ldots, X_{n-1})$ and $A = k[X]$. Using resultants, we see that there is some point $\xi \in k^{n-1}$ such that the univariate polynomials $g^{(0)} := g(\xi, Y)$ and $h^{(0)} := h(\xi, Y)$ are coprime. By a coordinate translation, we may assume that $\xi = 0$. Note that $g^{(0)}$ is a monic univariate polynomial over k of degree d. We have $u h^{(0)} + v g^{(0)} = 1$ with uniquely determined $u, v \in k[Y]$ such that $\deg u < d$.

The basic idea is to use Hensel lifting in order to successively compute the factorization $f = g^e h$ from $f^{(0)} := (g^{(0)})^e h^{(0)}$. The crux is the choice of the suitable valuation by which to lift. We will lift with respect to the total degree in the X-variables. Consider the decomposition of polynomials in $A[Y]$ into homogeneous parts with respect to the total degree in the X-variables:

$$f = \sum_{\delta \geq 0} f^{(\delta)}, \ g = \sum_{\delta \geq 0} g^{(\delta)}, \ h = \sum_{\delta \geq 0} h^{(\delta)},$$

where $f^{(\delta)}, g^{(\delta)}, h^{(\delta)} \in A[Y]$ are homogeneous of degree δ in the X-variables. This notation is consistent with our earlier introduction of $f^{(0)}, g^{(0)}, h^{(0)}$. We have $\deg_Y g^{(\delta)} < d$ for $\delta > 0$, and $g^{(\delta)} = 0$ for $i > d$.

We are going to derive a formula, which allows to compute $g^{(s+1)}, h^{(s+1)}$ from the homogeneous parts of g and h up to degree s. From $f = g^e h$ we obtain modulo the ideal generated by the monomials of degree $s + 2$ in the X-variables that

$$
\begin{aligned}
f &\equiv \left(\sum_{\delta=0}^{s} g^{(\delta)} + g^{(s+1)} \right)^e \left(\sum_{\delta=0}^{s} h^{(\delta)} + h^{(s+1)} \right) \\
&\equiv \left(\left(\sum_{\delta=0}^{s} g^{(\delta)} \right)^e + e(g^{(0)})^{e-1} g^{(s+1)} \right) \left(\sum_{\delta=0}^{s} h^{(\delta)} + h^{(s+1)} \right) \\
&\equiv \left(\sum_{\delta=0}^{s} g^{(\delta)} \right)^e \left(\sum_{\delta=0}^{s} h^{(\delta)} \right) + e(g^{(0)})^{e-1} g^{(s+1)} h^{(0)} + (g^{(0)})^e h^{(s+1)}.
\end{aligned}
$$

If we write

$$F := \sum_{\delta \geq 0} F^{(\delta)} := \left(\sum_{\delta=0}^{s} g^{(\delta)} \right)^e \left(\sum_{\delta=0}^{s} h^{(\delta)} \right)$$

(omitting the dependence of F on s) and set $\Delta^{(s+1)} := f^{(s+1)} - F^{(s+1)}$, then we obtain

$$Q := \frac{\Delta^{(s+1)}}{(g^{(0)})^{e-1}} = eg^{(s+1)}h^{(0)} + g^{(0)}h^{(s+1)}. \tag{13}$$

Since $\deg_Y g^{(s+1)} < d$, this relation uniquely determines the polynomial $g^{(s+1)}$. On the other hand, we have $Q = uQh^{(0)} + vQg^{(0)}$. It follows that $eg^{(s+1)}$ is the remainder of the division of uQ by $g^{(0)}$, hence $eg^{(s+1)}$ is the $(e-1)$th $g^{(0)}$-adic coefficient of $u\Delta^{(s+1)}$.

We write elements $a \in A[Y]$ in the form

$$a = \sum_{\delta \geq 0} a^{(\delta)} = \sum_{\delta \geq 0, i \geq 0} a_i^{(\delta)} (g^{(0)})^i = \sum_{\delta \geq 0, i \geq 0, j < d} a_{ij}^{(\delta)} Y^j (g^{(0)})^i$$

with *coefficients* $a_{ij}^{(\delta)} \in A$, which are homogeneous polynomials of degree δ in the X-variables. Note that the $a_i^{(\delta)} = \sum_{j<d} a_{ij}^{(\delta)} Y^j$ are the $g^{(0)}$-adic coefficients of $a^{(\delta)}$. The collection of coefficients

$$a_{ij}^{(\delta)} \quad \text{for } 0 \leq \delta \leq d,\ 0 \leq j < d,\ 0 \leq i \leq D$$

will be used to represent the element a approximatively. We will call this an approximation of order D of a.

Assume that $c = a \cdot b$ in $A[Y]$. A straightforward generalization of Lemma 5.1 yields that the coefficients $c_{ij}^{(\delta)}$ of c with $\delta \leq d, j < d, i \leq D$ can be computed from the corresponding coefficients of a and b with $O(M(d^2 D))$ arithmetic operations in A. Using this generalization of Lemma 5.1, we can generalize Proposition 5.2 in an obvious way and obtain that the coefficients $f_{ij}^{(\delta)}$ for $\delta \leq d, j < d, i \leq de$ can be computed with $O(M(d^2 \cdot de)L(f))$ arithmetic operations in A.

For $0 \leq s \leq d$ we define C_s as the set of the coefficients $g_{ij}^{(\delta)}, h_{ij}^{(\delta)}$ for $\delta \leq s, j < d, i \leq (d-s)e$. Note that C_s is a subset of $A = k[X]$. Moreover, $C_0 \subseteq k$, thus the elements of C_0 can be considered as free constants. Note also that $g^{(\delta)} = g_0^{(\delta)}$ for $\delta > 0$ as $\deg_Y g^{(\delta)} < d$.

Inductively, we assume now that we have already computed the elements of C_s for $s < d$. From these data we can compute the coefficients $F_{ij}^{(s+1)}$ of F for $j < d, i \leq (d-s)e$ with $O(M(d^2(d-s)e \log e))$ arithmetic operations using

the above mentioned generalization of Lemma 5.1 ($\log e$ squarings). From this and the coefficients $f_{ij}^{(\delta)}$, $\delta \le d, j < d, i \le de$, we compute the coefficients of $u\Delta^{(s+1)}$ up to order $(d-s)e$ with $O(M(d(d-s)e))$ arithmetic operations. In particular, we have thus computed the coefficients of $eg^{(s+1)}$, since $eg^{(s+1)}$ is the $(e-1)$th $g^{(0)}$-adic coefficient of $u\Delta^{(s+1)}$. From equation (13) we obtain

$$\Delta^{(s+1)} - eg^{(s+1)}(g^{(0)})^{e-1}h^{(0)} = h^{(s+1)}(g^{(0)})^e.$$

We can compute the coefficients of this polynomial up to order $(d-s)e$ with further $O(M(d(d-s)e))$ arithmetic operations. This way, we get the coefficients of $h^{(s+1)}$ up to order $(d-s-1)e$. Note that the order has decreased by e.

Summarizing, the cost of each induction step is $O(M(d^3 e \log e))$ and there are at most d induction steps. The polynomial $g = g^{(0)} + g^{(1)} + \ldots + g^{(d)}$ can be computed with further $O(d^2)$ additions from the coefficients of the $g^{(\delta)} = g_0^{(\delta)}$ for $1 \le \delta \le d$. Altogether, we obtain $L(g) = O(M(d^3 e)(L(f) + d \log e))$ as claimed. $\qquad \square$

Proposition 6.1 *Assume that $f = g^e$ in $k[X_1, \ldots, X_n]$, $d = \deg g \ge 1$ and char $k = 0$. Then $L(g) \le O(M(d)L(f))$.*

Proof. By a coordinate shift, we may assume that $g(0) \ne 0$; without loss of generality $g(0) = 1$. The polynomial $\varphi = g - 1 \in k[X]$ is the unique solution of the equation

$$(1 + \varphi)^e - f = 0, \ \varphi(0) = 0$$

to be solved in the ring of formal power series $k[[X]]$. This power series φ can be recursively computed by the Newton iteration $\varphi_0 = 0$ and

$$\varphi_{\nu+1} = \varphi_\nu - \frac{(1+\varphi_\nu)^e - f}{e(1+\varphi_\nu)^{e-1}} = -\frac{1}{e} + (1 - \frac{1}{e})\varphi_\nu + \frac{1}{e}\frac{f}{(1+\varphi_\nu)^{e-1}}$$

satisfying $\varphi_\nu \equiv \varphi \bmod (X)^{2^\nu}$. (Compare Section 3.2 and [10, Theorem 2.31].)

We first compute the homogeneous parts of f up to degree d with $O(M(d)L(f))$ arithmetic operations by a variant of Proposition 3.1. Using Lemma 3.2, we can compute from this and the homogeneous parts of φ_ν up to degree 2^ν the homogeneous parts of $\varphi_{\nu+1}$ up to degree $2^{\nu+1}$ with $O(M(2^\nu)\log e)$ arithmetic operations ($\log e$ squarings). As $\sum_{\nu=0}^{N} M(2^\nu) \le M(2^{N+1} - 1)$, a total of $O(M(d)(L(f) + \log e))$ arithmetic operations is sufficient. Since $L(f) \ge \log \deg f \ge \log e$, the claim follows. $\qquad \square$

References

[1] A. Alder. *Grenzrang und Grenzkomplexität aus algebraischer und topologischer Sicht.* PhD thesis, Zürich University, 1984.

[2] A.I. Barvinok. Polynomial time algorithms to approximate permanents and mixed discriminants within a simply exponential factor. *Random Structures and Algorithms*, 14:29–61, 1999.

[3] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation.* Springer, 1998.

[4] L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers. *Bull. Amer. Math. Soc.*, 21:1–46, 1989.

[5] J. Bochnak, M. Coste, and M.F. Roy. *Géometrie algébrique réelle*, volume 12 of *Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge.* Springer Verlag, 1987.

[6] P. Bürgisser. On the structure of Valiant's complexity classes. *Discr. Math. Theoret. Comp. Sci.*, 3:73–94, 1999.

[7] P. Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and Computation in Mathematics.* Springer Verlag, 2000.

[8] P. Bürgisser. Cook's versus Valiant's hypothesis. *Theoret. Comp. Sci.*, 235:71–88, 2000.

[9] P. Bürgisser. On implications between P-NP-hypotheses: Decision versus computation in algebraic complexity. In J. Sgall, A. Pultr, and P. Kolman, editors, *Proc. 26th MFCS*, number 2136 in LNCS, pages 3–17. Springer Verlag, 2001.

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

[11] P. Bürgisser, M. Karpinski, and T. Lickteig. On randomized semialgebraic decision complexity. *J. Compl.*, 9:231–251, 1993.

[12] P. Bürgisser, T. Lickteig, and M. Shub. Test complexity of generic polynomials. *J. Compl.*, 8:203–215, 1992.

[13] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *J. Symb. Comp.*, 9:251–280, 1990.

[14] F. Cucker, M. Karpinski, P. Koiran, T. Lickteig, and K. Werther. On real Turing machines that toss coins. In *Proc. 27th ACM STOC, Las Vegas*, pages 335–342, 1995.

[15] J. von zur Gathen. Hensel and Newton methods in valuation rings. *Math. Comp.*, 42:637–661, 1984.

[16] J. von zur Gathen. Feasible arithmetic computations: Valiant's hypothesis. *J. Symb. Comp.*, 4:137–172, 1987.

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

[18] B. Griesser. Lower bounds for the approximative complexity. *Theoret. Comp. Sci.*, 46:329–338, 1986.

[19] D.Yu. Grigoriev and M. Karpinski. Randomized quadratic lower bound for knapsack. In *Proc. 29th ACM STOC*, pages 76–85, 1997.

[20] J. Grollmann and A.L. Selman. Complexity measures for public-key cryptosystems. *SIAM J. Comp.*, 17(2):309–335, 1988.

[21] M.R. Jerrum, A. Sinclair, and E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. *Electronic Colloquium on Computational Complexity*, 2000. Report No. 79.

[22] E. Kaltofen. Single-factor Hensel lifting and its application to the straight-line complexity of certain polynomials. In *Proc. 19th ACM STOC*, pages 443–452, 1986.

[23] E. Kaltofen. Factorization of polynomials given by straight-line programs. In S. Micali, editor, *Randomness and Computation*, pages 375–412. JAI Press, Greenwich CT, 1989.

[24] E. Kaltofen and B.D. Saunders. On Wiedemann's method of solving sparse linear systems. In *Proc. AAECC-9*, number 539 in LNCS, pages 29–38. Springer Verlag, 1991.

[25] E. Kaltofen and B.M. Trager. Computing with polynomials given by black boxes for their evaluations: greatest common divisors, factorization, separation of numerators and denominators. *J. Symb. Comp.*, 9:301–320, 1990.

[26] T. Lehmkuhl and T. Lickteig. On the Order of Approximation in Approximative Triadic Decompositions of Tensors. *Theoret. Comp. Sci.*, 69:1–14, 1989.

[27] A.K. Lenstra. Factoring multivariate polynomials over finite fields. *J. Comp. Syst. Sci.*, 30:235–248, 1985.

[28] A.K. Lenstra, H.W. Lenstra Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[29] T. Lickteig. Testing Polynomials for zero (extended abstract). Internal Report 1988. Universität Tübingen.

[30] T. Lickteig. On semialgebraic decision complexity. Technical Report TR-90-052, Int. Comp. Sc. Inst., Berkeley, 1990. Habilitationsschrift, Universität Tübingen.

[31] N. Linial, A. Samorodnitsky, and A. Wigderson. A deterministic polynomial algorithm for matrix scaling and approximate permanents. In *Proc. 30th ACM STOC*, pages 644–652, 1998.

[32] R.J. Lipton and L.J. Stockmeyer. Evaluation of polynomials with superpreconditioning. *J. Comp. Syst. Sci.*, 16:124–139, 1978.

[33] F. Meyer auf der Heide. Simulating probabilistic by deterministic algebraic computation trees. *Theoret. Comp. Sci.*, 41:325–330, 1985.

[34] D. Mumford. *Algebraic Geometry I: Complex Projective Varieties.* Springer Verlag, 1976.

[35] A.L. Selman. A survey of one-way functions in complexity theory. *Math. Systems Theory*, 25:203–221, 1992.

[36] V. Strassen. Vermeidung von Divisionen. *Crelles J. Reine Angew. Math.*, 264:184–202, 1973.

[37] V. Strassen. Polynomials with rational coefficients which are hard to compute. *SIAM J. Comp.*, 3:128–149, 1974.

[38] C. Sturtivant and Z. L. Zhang. Efficiently inverting bijections given by straight line programs. In *Proc. 31th FOCS*, pages 327–334, 1990.

[39] L.G. Valiant. Relative complexity of checking and evaluating. *Inf. Proc. Letters*, 5:20–23, 1976.

[40] L.G. Valiant. Completeness classes in algebra. In *Proc. 11th ACM STOC*, pages 249–261, 1979.

[41] L.G. Valiant. The complexity of computing the permanent. *Theoret. Comp. Sci.*, 8:189–201, 1979.

[42] L.G. Valiant. Reducibility by algebraic projections. In *Logic and Algorithmic: an International Symposium held in honor of Ernst Specker*, volume 30, pages 365–380. Monogr. No. 30 de l'Enseign. Math., 1982.

[43] D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Trans. Inform. Theory*, 32:54–62, 1985.