A Lower Bound for Randomized Algebraic Decision Trees

Dima Grigoriev* Marek Karpinski† Friedhelm Meyer auf der Heide ‡
Roman Smolensky §

Abstract

We extend the lower bounds on the depth of algebraic decision trees to the case of *randomized* algebraic decision trees (with two-sided error) for languages being finite unions of hyperplanes and the intersections of halfspaces. As an application, among other things, we derive, for the first time, an $\Omega(n^2)$ *randomized* lower bound for the *Knapsack Problem* which was previously only known for deterministic algebraic decision trees. It is worth noting that for the languages being finite unions of hyperplanes our proof method yields also a new elementary technique for deterministic algebraic decision trees without making use of Milnor's bound on Betti number of algebraic varieties.

*Dept. of Computer Science and Mathematics, Penn State University, University Park. Email: dima@cse.psu.edu

†Dept. of Computer Science, University of Bonn, 53117 Bonn. Email: marek@cs.uni-bonn.de

‡Heinz Nixdorf Institute and Computer Science Department, University of Paderborn, 33098 Paderborn. Email: fmadh@uni-paderborn.de

§Dept. of Computer Science, University of Bonn, 53117 Bonn. Roman Smolensky has died on 19 October, 1995 in New York. This paper is also dedicated to the memory of Roman by the other co-authors. Email: roman@cs.uni-bonn.de

1 Introduction

Starting with [MT82], [S83], [M85a] and [M85b] there has been a continued effort in the last decade to understand an intrinsic power of randomization in algebraic decision trees (see also [BKL93], [GK93], [GK94] for some more recent results). Several algebraic and topological methods which were introduced in proving lower bounds for deterministic algebraic decision trees (cf. [SY82], [B83], [BLY92], [GKV95], [Y94]), with the exception of [BKL93], and [GK93], were not yielding lower bounds for the case of randomized decision trees. In [M85a] a lower bound has been stated on the depth or randomized *linear* decision trees (randomized algebraic decision trees of degree 1) for the case of languages being finte unions of hyperplanes (a gap in the proof of the Main Lemma of [M85a] for the generic case was closed in [GK94]). Our paper provides the first lower bounds on the depth of randomized algebraic decision trees in the case of the languages being finite unions of hyperplanes as well as intersections of halfspaces. In this case we provide a new method for proving lower bounds also for deterministic algebraic decision trees without making use of Milnor's bound and Betti numbers of algebraic varieties. As an application we derive randomized lower bounds for a number of concrete problems, among others, *Knapsack* ($\Omega(n^2)$ lower bound), and the *Element Distinctness* ($\Omega(n \log n)$ lower bound).

The paper is organized as follows. Section 2 introduces the notation of randomized algebraic decision and computation trees. Section 3 overviews the known results in the area. Section 4 summarizes our results and applies them for the number of concrete problems. Section 5 gives an outline of the lower bound proof, and Sections 6–7 gives the proof of the Main Theorem.

2 Deterministic and Randomized Decision Trees.

An *algebraic decision tree* of degree d, a *d-DT* for inputs $(x_1, \ldots, x_n) \in \mathbb{R}^n$ is a rooted ternary tree. Its root and inner nodes are labelled by polynomials from $\mathbb{R}[X_1, \ldots, X_n]$ of degree at most d, its leaves are accepting or rejecting. The computation of the *d-DT* on input $(x_1, \ldots, x_n) \in \mathbb{R}^n$ consists of traversing the tree from the root to a leaf, always choosing the left/middle/right branch of a node labelled with polynomial g dependent on whether $g(x_1, \ldots, x_n)$ is smaller/equal/greater than 0.

The inputs $(x_1, \ldots, x_n) \in \mathbb{R}^n$ arriving at accepting leaves form the set $S \subseteq \mathbb{R}^n$ *recognized* by the *d-DT*.

A *randomized algebraic decision tree* of degree d, a *d-RDT*, is a finite collection of *d-DT*s T_α. Such a *d-RDT* recognizes $S \subseteq \mathbb{R}^n$, if, for each $x \in \mathbb{R}^n$, at least a fraction of $1 - \gamma$ of the T_α's classify x correctly w. r. t. S, for some $\gamma \in (0, \frac{1}{2})$, called the *error probability.*

The depth of T is the maximum depth of the T_α's. In case of $d = 1$ we talk about deterministic or randomized *linear decision trees*, *LDT*s or *RLDT*s. In case we do not restrict the degree of the polynomials but charge for each arithmetic operation needed to compute them we talk about deterministic and randomized *algebraic computation trees*, *CT*s and *RCT*s (for details see [B83]) with the similar notations for depth (or time).

Note that neither the choice of the error probability from $(0, \frac{1}{2})$, nor the complexity measure "maximum depth" are significant. Choosing any constant error probability smaller than $\frac{1}{2}$ and replacing maximum depth by maximum over all expected path lengths, maximum taken over all inputs ("worst case expected time"), only changes the complexity of a set S by a constant factor, see e. g. [M85c]. Also, we note here without a proof that the restriction that an *RDT* consists of a *finite* collection of *DT*s can be weakened: $\alpha \in \mathbb{N}$, e. g., works as well.

3 Known Results.

The most important results in connection to this research are the variants of the component counting lower bound for deterministic computations: Let $L \subseteq \mathbb{R}^n$ have q connected components. Then each *LDT* for L has depth $\Omega(\log(q))$ [DL75], each *d-DT* for L has depth $\Omega(\frac{\log(q)}{\log(d)} - n)$ (can be deduced from [B83]), each *CT* for L has depth $\Omega(\log(q) - n)$ [B83].

The last two results heavily depend on Milnor's bound on Betti numbers for real algebraic varieties, thus use deep results from algebraic topology.

In order to apply the component counting lower bound one has to count the number of connected components of interesting languages.

Consider, e. g., the Integer Programming Language $L_{n,k} = \{x \in \mathbb{R}^n, \exists a \in \{0, \ldots, k\}^n : xa =$

1} [M85a, M85b]. For each $k \geq 1$, the family $\{L_{n,k}, n \in \mathbb{N}\}$, restricted to integer inputs, is NP-complete, for $k = 1$ this is the famous *Knapsack Problem.*

As shown in [DL78] for $k = 1$ and in [M85b] for arbitrary k, $\mathbb{R}^n - L_{n,k}$ has $(k+1)^{\Omega(n^2)}$ many connected components, yielding lower bounds $\Omega(n^2 \log(k+1))$ in the above models. In [M84] it is shown that all these NP-complete problems have polynomial depth LDTs, for their n-dimensional restrictions.

A further important example was the Element Distinctness Problem, with the connected components bound $n!$, and therefor lower bound $\Omega(n \log n)$.

As far as randomized DTs are concerned much less is known. In [M85a] it is shown that deterministic and randomized LDTs, d-DTs, and CTs, resp., are polynomially related. A randomized lower bound is shown in [M85b] that extends the lower bounds for e. g. the problems mentioned above to randomized LDTs. (A gap in that proof for the generic case was closed in [GK95].)

In [BKL93] it is shown that there are benefits if randomization is used in CTs: Consider the language $\{(x, y) \in \mathbb{R}^{2n} : y$ is permutation of $x\} \subseteq \mathbb{R}^{2n}$. As this language consists of $n!$ n-dimensional linear subspaces of $\mathbb{R}^{2n}$, a restriction to an n-dimensional affine subspace in general position turns it into a set of $n!$ isolated points. Thus its deterministic complexity is $\Omega(n \log n)$ on the above deterministic models. On the other hand, as noted in [BKL93], RCTs need time $O(n)$ only.

4 New Results.

Consider $S = \bigcup_{i=1}^m H_i$ or $S^+ = \bigcap_{i=1}^m H_i^+$, where the H_i's are hyperplanes, and the H_i^+'s are halfspaces. S is often called a *linear arrangement*, S^+ is a *polyhedron.* A k-face L of S is a k-dimensional subspace defined by intersecting $n-k$ of the H_i's. If L is k-dimensional on the boundary of S^+, it is also a k-face of S^+.

We prove the following lower bound.

Main Theorem: *Let $H_1, \ldots, H_m$ be hyperplanes in $\mathbb{R}^n$, $S = \bigcup_{i=1}^m H_i, S^+ = \bigcap_{i=1}^m H_i^+$. If S or S^+ has $m^{\Omega(n-k)}$ k-faces and $m = (n-k)^{\Omega(n-k)}$ for some $k \in \{1, \ldots, n\}$, then each d-RDT for S or S^+ has depth $\Omega((n-k)\log(m))$, even if $d = m^\delta$ for sufficiently small $\delta > 0$.*

Thus, in order to get lower bounds, we need S or S^+ to have $m^{\Omega(n)}$ k-faces for some k. This is true e. g. for all problems mentioned in "previous results", thus all deterministic results mentioned there can be turned into randomized ones, in particular, we get the $\Omega(n^2)$ lower bound for the knapsack problem for d-RDTs, even if $d = m^\delta$ for sufficiently small $\delta > 0$.

Main Theorem yields directly the following concrete applications:

Main Corollary Fix degree $d = m^\delta$ for certain small $\delta > 0$. Then:

(a) Lower bound for the depth of a d-RDT recognizing the *Knapsack Problem* is $\Omega(n^2)$.

(b) Lower bound for the depth of a d-RDT recognizing *Element Distinctness Problem* is $\Omega(n \log n)$.

We note also an interesting application of our method towards the finite set $\{(x, y) \in \mathbb{R}^{2n} : y$ is a permutation of $x\} \subseteq \mathbb{R}^{2n}$ with the deterministic complexity $\Omega(n \log n)$, and with RCTs complexity $O(n)$ [BKL93]. Our method applies also for this set yielding $\Omega(n \log n)$ lower bound for d-RDTs.

5 Outline of the Lower Bound Proof

Fix affinely independent hyperplanes $H_i = \{x \in \mathbb{R}^n, a_i x = b_i\}$ for $i = 1, \ldots, n$. Let $v \in \mathbb{R}^n$ be such that $\bigcap_{i=1}^n H_i = \{v\}$.

Let A denote the $n \times n$-matrix whose rows are $a_1, \ldots, a_n$. For a polynomial $f \in \mathbb{R}[X_1, \ldots, X_n]$ we consider its expansion with origin v and coordinates $a_1, \ldots, a_n$; $f^{(v;H_1,\ldots,H_n)}(Y_1, \ldots, Y_n) := f(v + A(Y_1, \ldots, Y_n))$. Denote for brevity $g = f^{(v;H_1,\ldots,H_n)}$ and define the leading term $lm(g)$ as follows: First take the terms of g with the least degree in Y_n, then among them with the least degree in Y_{n-1} and so on, till Y_1. One could describe $lm(g)$ by means of infinitesimals (cf., e. g., [GV88]).

Namely for a real closed field $\mathbf{F}$ (see e. g. [L65]) we say that an element ε transcendental over F is an infinitesimal (with respect to $\mathbf{F}$) if $0 < \varepsilon < a$ for any element $0 < a \in \mathbf{F}$. This uniquely induces the order on the field $F(\varepsilon)$ of rational functions and further on the real closure $\widetilde{\mathbf{F}(\varepsilon)}$ (see [L65]). Now let $\varepsilon_1 > \ldots > \varepsilon_n > 0$ be the elements such that $\varepsilon_{\ell+1}$ is infinitesimal with respect to the real closed field $\widetilde{\mathbb{R}(\varepsilon)}$ for $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_\ell)$, $0 \le \ell < n$. Then the sign $sgn(g(\varepsilon_1, \ldots, \varepsilon_n)) = sgn(lm(g)(\varepsilon_1, \ldots, \varepsilon_n))$ and this property uniquely determines the term $lm(g)$. Actually, one could stick in the arguing below with the real numbers $1 = \varepsilon_0^{(0)} > \varepsilon_1^{(0)} > \ldots > \varepsilon_n^{(0)} > 0$ instead of $\varepsilon_1, \ldots, \varepsilon_n$ where $\varepsilon_{\ell+1}^{(0)}$ is "considerably smaller" than $\varepsilon_\ell^{(0)}$, $0 \le l \le n-1$. But then one should specify, what does it mean "considerably smaller", and it is more convenient to use infinitesimals.

Now fix a family of polynomials $f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, x_n]$. By $\mathrm{Var}^{(v,H_1,\ldots,H_n)}(f_1, \ldots, f_s)$ we denote the number of variables among $Y_1, \ldots, Y_n$ appearing in the leading terms $lm(f_1^{(v,H_1,\ldots,H_n)}), \ldots, lm(f_s^{(v,H_1,\ldots,H_n)})$. For a d-DT T, $\mathrm{Var}^{(v;H_1,\ldots,H_n)}(T)$ denotes the maximum of all $\mathrm{Var}^{(v;H_1,\ldots,H_n)}(f_1, \ldots, f_s)$, maximum taken over all $f_1, \ldots, f_s$ appearing as testing polynomials on a path in T. We extend the above definition to the case of less hyperplanes $H_1, \ldots, H_{n-k}$ for some $1 \le k \le n-1$. Then $L = \bigcap_{i=1}^{n-k} H_i$ is a k-dimensional affine subspace of $\mathbb{R}^n$. For an arbitrary $v \in L$ we take an $(n-k)$-dimensional subspace U orthogonal to L, with $\{v\} = L \cap U$. We define $\mathrm{Var}_k^{(v;H_1,\ldots,H_{n-k})}(T)$ as above, for the polynomials $f_1, \ldots, f_s$ restricted to U. The following two lemmas imply the lower bound from our Main Theorem. The following chapters contain their proofs.

Lemma 1 *Let T be a d-RDT (or an RCT) recognizing $L = \cup_{i=1}^{n-k} H_i$ or $L^+ = \cap_{i=1}^{n-k} H_i^+$ with error probability $\gamma < \frac{1}{2}$. Then $Var^{(v;H_1,\ldots,H_{n-k})}(T_\alpha) \ge (1-2\gamma)^2 \cdot (n-k)$ for a fraction of $\frac{1-2\gamma}{2-2\gamma}$ of all T_α's.*

Let us denote $\mathbb{R}_+^n = \{(x_1, \ldots, x_n) : x_i \ge 0,\ 1 \le i \le n\}$ and $\mathbb{R}_0^n = (\mathbb{R} \setminus \{0\})^n$. Lemma 1 entails two direct corollaries for both RDTs and RCTs, which give an interesting geometric interpretation of the depth bounds of Lemma 1.

Corollary 1. Any RCT which recognizes $\mathbb{R}_+^n$ or $\mathbb{R}_0^n$ must have the depth greater than or equal to $\frac{1}{2}(1-2\gamma)^2 n$.

Corollary 2. Any d-RDT which recognizes $\mathbb{R}_+^n$ or $\mathbb{R}_0^n$ must have the depth greater or equal to $\frac{1}{d}(1-2\gamma)^2 n$.

Let T' be an d-DT, and $S = \cup_{i=1}^{m} H_i$ or $S = \cap_{i=1}^{m} H_i^+$ for hyperplanes $H_1, \ldots, H_m \in \mathbb{R}^n$.

For a k-face L of S let $1 \leq i_1 < \ldots < i_{n-k} \leq m$ be the lexicographically smallest sequence of $(n-k)$ indices such that $L = H_{i_1} \cap \ldots \cap H_{i_{n-k}}$. Let v_L belong to L but to no lower-dimensional face of S or S^+. We abbreviate $\mathrm{Var}_k^{(v_L, H_{i_1}, \ldots, H_{i_{n-k}})}(T')$ by $\mathrm{Var}^{(v_L)}(T')$.

Lemma 2 *Assume that, for some $c > 0$, there are at least M k-faces L of S with $Var_k^{(v_L)}(T') \geq c(n-k)$. Then the depth t of T' fulfils $M \leq 3^t \cdot m^{(1-c)(n-k)} \cdot (td)^{c(n-k)}$.*

Using these lemmas it is easy to conclude the Main Theorem:
Consider d-RDT for S or S^+ with error probability $\gamma < \frac{1}{2}$. S, S^+ have N many k-faces. Lemma 1 and elementary counting implies that there is α such that T_α fulfils: $\mathrm{Var}_k^{\{v_L\}}(T_\alpha) \geq (1-2\gamma)^2(n-k)$ for $(\frac{1-2\gamma}{2-2\gamma}) \cdot N$ many k-faces L.

Thus Lemma 2 implies the desired $\Omega((n-k)\log(m))$ lower bound, if N is large as demanded in the Main Theorem.

6 Proof of Lemma 1

First observe that it is sufficient to prove the lemma for $k = 0$ and under the assumption that $v = 0$ and the H_i's are defined by $\{x \in \mathbb{R}^n, x_i = 0\}$, in other words the expansion $(y_1, \ldots, y_n) \vdash v + A(y_1, \ldots, y_n)$ is the identity.

Now let the d-RDT (or the RCT) recognize $\bigcap_{i=1}^{n} H_i^+$ with error probability $\gamma < \frac{1}{2}$.

Consider the points $E = (\varepsilon_1, \ldots, \varepsilon_n)$ and $E_i^{(+)} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, -\varepsilon_i, \varepsilon_{i+1}, \ldots, \varepsilon_n), i = 1, \ldots, n$. Easy counting yields that there is a fraction of $(1-2\gamma)/(2-2\gamma)$ of the T_α's that classify E and at least $(1-2\gamma)^2 n$ many E_i's correctly. Take one such T_α and some i_o such that T_α classifies $E_{i_o}^+$ correctly.

Denote by $f_1, \ldots, f_s$ the testing polynomials along the path in T_α followed by input E. We claim that X_{i_0} occurs in one of the leading terms $lm(f_1), \ldots, lm(f_s)$. Indeed, otherwise $sgn(f_\ell(E_{i_0}^{(+)})) = sgn(lm(f_\ell(E_{i_0}^{(+)}))) = sgn(lm(f_\ell(E))) = sgn(f_\ell(E))$, $1 \leq \ell \leq s$, therefore $E_{i_0}^{(+)}$ satisfies all the tests along the same path as E, hence the output for $E_{i_0}^{(+)}$ would be "yes", which contradicts to the choice of i_0. This implies Lemma 1 for $\bigcap_{i=1}^{n} H_i^+$.

In case of T recognizing $\bigcup_{i=1}^{n} H_i$ consider the points $E_i^{(0)} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, 0, \varepsilon_{i+1}, \ldots, \varepsilon_n)$, $1 \leq i \leq n$ and argue as above, replacing $E_i^{(+)}$ by $E_i^{(0)}$, $1 \leq i \leq n$.

7 Proof of Lemma 2

To every k-face L defined by an intersection $H_{i_1} \bigcap \ldots \bigcap H_{i_{n-k}}, i_1 < \ldots < i_{n-k}$, see above, with $\mathrm{Var}^{(v_L)}(T') \geq c(n-k)$, we correspond a path in T' with the testing polynomials $f_1, \ldots, f_s$ for which $\mathrm{Var}^{(v_L)}(T') = \mathrm{Var}^{(v_L)}(f_1, \ldots, f_s)$.

By a flag of L we mean the sequence of embedded planes

$$H_{i_{n-k}} \supset H_{i_{n-k}} \bigcap H_{i_{n-k-1}} \supset H_{i_{n-k}} \bigcap H_{i_{n-k-1}} \bigcap H_{i_{n-k-2}} \supset \ldots \supset \bigcap H_{i_{n-k}} \bigcap \ldots \bigcap H_{i_1}$$

where $i_1 < \ldots < i_{n-k}$ were yielded above. Our purpose is to label some of these planes in an appropriate way. As a result , a labeled flag would be attached to L. Morever, for a fixed path in T' with the testing polynomials $f_1, \ldots, f_s$ we organize the labeled flags attached to all k-faces L which correspond to this path as a regular tree $\mathcal{T} = \mathcal{T}(f_1, \ldots, f_s)$ with all the paths of the same length $n-k$.

We construct the tree $\mathcal{T}$ and thereby the labeled flags by induction on the level . The base of induction. Take L which corresponds to the fixed path (we utilize the introduced above notations for the coordinates in a neighbourhood of v_L). If Y_{n-k} (or in other words, hyperplane $H_{i_{n-k}}$) divides one of $f_1, \dots, f_s$ we construct a vertex, being a son of the root of the tree $\mathcal{T}$, mark it with the hyperplane $H_{i_{n-k}}$ and label. If Y_{n-k} does not divide any of $f_1, \dots, f_s$, we do not label this vertex of $\mathcal{T}$. To complete the construction of the first level of $\mathcal{T}$, we represent the polynomial $f_j = \tilde{f}_j Y_{n-k}^{m_j} \mathcal{L}_{H_{r_1}}^{m_{j,1}} \dots \mathcal{L}_{H_{r_p}}^{m_{j,p}}$, $1 \leq j \leq s$ as a product for maximal possible $m_j, m_{j,1}, \dots, m_{j,p}$ where $i_{n-k} < r_1 < \dots < r_p$ and $\mathcal{L}_{H_{r_1}}, \dots, \mathcal{L}_{H_{r_p}}$ are all linear polynomials determining hyperplanes $H_{r_1}, \dots, H_{r_p}$ which divide f_j with the indices $r_1, \dots, r_p$ greater than i_{n-k}. We assign to the constructed vertex the polynomials $f_j^{(1)}(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-1}) = \widetilde{f}_j(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-1}, 0)$, $1 \leq j \leq s$. One could view the polynomial $f_j^{(1)}$ as being defined on the hyperplane $H_{i_{n-k}}$.

Observe that the linear polynomials $\mathcal{L}_{H_{r_1}} \dots \mathcal{L}_{H_{r_p}}$ do not vanish on L (due to the choice of i_{n-k}) and therefore these linear polynomials do not vanish at v_L, hence the expansion in the coordinates $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$ of $\mathcal{L}_{H_{r_l}}$, $1 \leq l \leq p$ contains nonzero constant term which is thereby its leading term, thus $lm^{(v_L)}(f_j)$ coincides with $lm^{(v_L)}(\widetilde{f}_j Y_{n-k}^{m_j})$ up to a constant factor. Furthermore, $lm^{(v_L)}(\widetilde{f}_j Y_{n-k}^{m_j}) = lm^{(v_L)}(\widetilde{f}_j) Y_{n-k}^{m_j} = lm^{(v_L)}(f_j^{(1)}) Y_{n-k}^{m_j}, 1 \leq j \leq s$, and so the leading term of the new polynomial $f_j^{(1)}$ up to a constant factor is obtained from the leading term of the former polynomial f_j by dividing on $Y_{n-k}^{m_j}$, $1 \leq j \leq s$. We refer to this property as the maintenance of the leading term. In particular, if the vertex of $\mathcal{T}$ under consideration is not labeled, the leading term of all the polynomials change only up to constant factors. If Y_{n-k} occurs in one of $lm^{(v_L)}(f_j)$, $1 \leq j \leq s$ then the vertex is labeled.

Notice that all the k-faces with the same first hyperplane $H_{i_{n-k}}$ in their flags, correspond to the constructed vertex (marked with $H_{i_{n-k}}$). Remark that the polynomials $f_j^{(1)}$, $1 \leq j \leq s$ do not depend on a particular k-face, but still we expand them in the coordinates which depend on L (so, v_L).

Now suppose by induction that $\ell < n$ levels of the tree $\mathcal{T}$ are already constructed. Consider any vertex w of $\mathcal{T}$ at ℓ-th level. To the vertex w leads to path (partially labeled), whose vertices are marked successively by the beginning elements of a flag

$H_{i_{n-k}} \supset H_{i_{n-k}} \bigcap H_{i_{n-k-1}} \supset \dots \supset H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-\ell+1}}$.

Finally, the polynomials $f_j^{(\ell)}, 1 \leq j \leq s$ are assigned to the vertex w. One could look at $f_j^{(\ell)}, 1 \leq j \leq s$ as a polynomial restricted on $(n-\ell)$-dimension plane $H = H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-\ell+1}}$.

If this is the beginning of the flag of a k-face L (we still consider L to keep the notations), then we can regard $f_j^{(\ell)}(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-\ell}), 1 \leq j \leq s$ as the polynomials in the fixed coordinates in the neighbourhood of v_L. As above we construct a new vertex of $\mathcal{T}$ of the level $(\ell+1)$, being a son in $\mathcal{T}$ of the vertex under consideration, and mark it with the $(n-\ell-1)$-dimensional plane $H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-\ell+1}} \bigcap H_{i_{n-k-\ell}} = H \bigcap H_{i_{n-k-\ell}}$.

Represent $f_j^{(\ell)} = \tilde{f}_j^{(\ell)} Y_{n-k-\ell}^{q_j} \mathcal{L}_{H \bigcap H_{t_1}}^{q_{j,1}} \dots \mathcal{L}_{H \bigcap H_{t_\pi}}^{q_{j,\pi}}, 1 \leq j \leq s$ for the maximal possible $q_j, q_{j,1}, \dots, q_{j,\pi}$

where $i_{n-k-\ell} < t_1 < \ldots < t_\pi$ and $\mathcal{L}_{H \bigcap H_{t_1}}, \ldots, \mathcal{L}_{H \bigcap H_{t_\pi}}$ are all the linear polynomials in the plane H determining hyperplanes $H \bigcap H_{t_1}, \ldots, H \bigcap H_{t_\pi}$ (in H) which divide $f_j^{(\ell)}$ with the indices $t_1, \ldots, t_\pi$ greater than $i_{n-k-\ell}$. We assign to the constructed vertex the polynomials $f_j^{(\ell+1)} = \tilde{f}_j^{(\ell)}\left(Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-\ell-1}, 0\right), 1 \leq j \leq s$. One could view the polynomial $f_j^{(\ell+1)}$ as being defined on the plane $H \bigcap H_{i_{n-k-\ell}}$.

If $q_j \geq 1$ for at least one $1 \leq j \leq s$ then we label the constructed vertex. As in the base of the induction we observe that the linear polynomials $\mathcal{L}_{H \bigcap H_{t_1}}, \ldots, \mathcal{L}_{H \bigcap H_{t_\pi}}$ do not vanish on L (due to the choice of i_{n-k-l}) and therefore these linear polynomials do not vanish at v_L, hence the expansion in the coordinates $Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-l}$ of $\mathcal{L}_{H \bigcap H_{t_\theta}}$, $1 \leq \theta \leq \pi$ contains nonzero constant term which is thereby its leading term (with respect to the coordinates $Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-\ell}$). Thus, $lm^{(v_L)} f_j^{(\ell)}$ coincides with $lm^{(v_L)}\left(\widetilde{f}_j^{(l)} Y_{n-k-l}^{q_j}\right)$ up to a constant factor. Furthermore, $lm^{(v_L)}\left(\tilde{f}_j^{(\ell)} Y_{n-k-\ell}^{q_j}\right) = lm^{(v_L)}\left(\tilde{f}_j^{(\ell)}\right) \cdot Y_{n-k-\ell}^{q_j} = lm^{(v_L)}\left(f_j^{(\ell+1)}\right) Y_{n-k-\ell}^{q_j}, 1 \leq j \leq s$. So, the leading term of the new polynomial $f_j^{(\ell+1)}$ up to a constant factor is obtained from the leading term of the former polynomial $f_j^{(\ell)}$ by dividing on $Y_{n-k-\ell}^{q_j}, 1 \leq j \leq s$. Thus, we have ascertained the maintenance property of the leading terms (see the base of induction). Also the vertex is labeled if and only if $Y_{n-k-\ell}$ occurs in one of $lm^{(v_L)}\left(f_j^{(\ell)}\right), 1 \leq j \leq s$.

This completes the inductive construction of $\mathcal{T}$. Observe that to each path in $\mathcal{T}$ corresponds exactly one k-face represented by a flag marked on the path. Vice versa, by the construction of $\mathcal{T}$ every k-face L which corresponds to the fixed path of d - DT T' with the testing polynomials $f_1, \ldots, f_s$, appears in some leaf of $\mathcal{T}$.

Now let us estimate the number of leaves in $\mathcal{T}$. By the assumption of the lemma and due to the property of the maintenance of the leading terms on each path of $\mathcal{T}$ at least $c(n-k)$ vertices are labeled. Observe that in the inductive step of the described construction of $\mathcal{T}$ the constructed vertex (being a son of the vertex w of the level ℓ; we utilize the introduced above notations) which corresponds to the hyperplane $H \bigcap H_{i_{n-k-\ell}}$ (in H) is labeled if and only if the linear polynomial $\mathcal{L}_{H \bigcap H_{i_{n-k-\ell}}}$ divides the product $\prod_{1 \leq j \leq s} f_j^{(\ell)}$. Let $u_1 < \ldots < u_p$ be all the indices such that $\mathcal{L}_{H \bigcap H_{u_q}}$ divides the product $\prod_{1 \leq j \leq s} f_j^{(\ell)}, 1 \leq q \leq p$. By the observed above each labeled son of the vertex w is marked with some $H_{u_{q_0}}, 1 \leq q_0 \leq p$. Since in the construction of $f_j^{(\ell+1)}, 1 \leq j \leq s$ we divided by $\mathcal{L}_{H \bigcap H_{u_q}}$ for all $q > q_0$, we conclude that the degree $\deg\left(\prod_{1 \leq j \leq s} f_j^{(\ell+1)}\right) \leq \deg\left(\prod_{1 \leq j \leq s} f_j^{(\ell)}\right) - (p - q_0 + 1)$. Notice that the polynomials $f_j^{(\ell+1)}, 1 \leq j \leq s$ depend actually on the particular son of the vertex w, although we do not reflect this in the notations.

Besides the labeled sons, any vertex in $\mathcal{T}$ could have at most m unlabeled sons (in fact, each unlabeled son is marked with some H_u with $u < i_{n-k-\ell+1}$, so there are less than m sons in general, but we stick with a rough bound m which suffices).

To estimate the number of leaves in $\mathcal{T}$ denote by $M(R, Q, D)$ the maximal possible number of leaves in a regular tree (actually, we could stick with subtrees of $\mathcal{T}$, so they are partially labeled) with the length of any path equal to R, with at most Q unlabeled vertices on any path and with a polynomial of degree less or equal to D assigned to any vertex (in $\mathcal{T}$ we assign the polynomial $\prod_{1\leq j\leq s} f_j^{(\ell)}$ to the vertex w, see the construction). Assume $w\cdot\ell\cdot o\cdot g\cdot$ that $Q\leq R$ (if $Q>R$ then set $M(R,Q,D)=0$). Considering such a tree and its subtrees with the roots being the sons of the root of the tree we get the following inductive inequality $M(R,Q,D)\leq m\cdot M(R-1,Q-1,D)+\sum_{1\leq p\leq D} M(R-1,Q,D-p)$ (provided that $R>Q$, when $R=Q$ we have $M(Q,Q,D)\leq m\cdot M(Q-1,Q-1,D)$ where the first item in the right side relates the unlabeled sons of the root and the second item relates the labeled sons (see the bound on $\deg\left(\prod_{1\leq j\leq s} f_j^{(\ell+1)}\right)$). ¿From this inequality we get a bound (by induction on R) :

$$M(R,Q,D)\leq m^Q\frac{D^{R-Q}}{(R-Q)!}\binom{R}{Q}.$$

Indeed, the right side of the inequality by inductive hypothesis does not exceed (provided that $R>Q$, when $R=Q$ we have $M(Q,Q,D)\leq m^Q$ by induction on Q)

$$\begin{aligned} & m^Q\frac{D^{R-Q}}{(R-Q)!}\binom{R-1}{Q-1} \\ +\sum_{0\leq p\leq D-1} & m^Q\frac{p^{R-Q-1}}{(R-Q-1)!}\binom{R-1}{Q} \\ & \leq m^Q\left(\frac{D^{R-Q}}{(R-Q)!}\binom{R-1}{Q-1}\right. \\ & \left.+\binom{R-1}{Q}\frac{1}{(R-Q-1)!}\frac{D^{R-Q}}{R-Q}\right) \\ & = m^Q\frac{D^{R-Q}}{(R-Q)!}\binom{R}{Q}\end{aligned}$$

which was to be shown.

Substituting now

$$\begin{aligned} R=n-k, Q=(n-k)(1-c), \\ D=deg(\prod_{1\leq j\leq s} f_j)\leq sd,\end{aligned}$$

we obtain a bound

$$\begin{aligned} m^{(n-k)(1-c)}\frac{(sd)^{c(n-k)}}{(c(n-k))!}2^{n-k} \\ \leq m^{(n-k)(1-c)}(sd)^{c(n-k)}\end{aligned}$$

for the number of leaves in $\mathcal{T}$.

So far, we've considered one path of the d-decision tree T' (with the testing polynomials $f_1,\ldots,f_s$ along this path).

Denote by t the depth of T' (thus, T' has at most 3^t paths). Since each k-face corresponds to a certain path of T' (see the beginning of the proof of the lemma), we conclude that

$$M\leq 3^t m^{(n-k)(1-c)}(td)^{c(n-k)},$$

which proves Lemma 2.

8 Applications

There is a number of applications of our method for concrete problems (see for examples, e. g., [B83] or [M85a]). We shall discuss the full list of concrete problems for which our method applies, and the corresponding randomized lower bounds, in the final version of this paper. Here we give only in Main Corollary, Section 4, the applications for the *Knapsack*, and the *Element Distinctness* Problems with the randomized bounds $\Omega(n^2)$ and $\Omega(n\log n)$, respectively.

9 Conclusion and Open Problems

We have proven that the known counting lower bounds for DTs carry over to RDTs for sets being finite unions of hyperplanes and intersections of halfspaces. Two important questions remain open:

- Does our lower bound for RDTs hold also for sets of other structure, e. g. finite languages?

  Using the method of Example 2 in [BKL93] on polynomial zero-tests we can construct a finite set of $n!$ points (permutations) in $\mathbb{R}^n$, for which an RDT with degree n (cf. also the restriction on M in Theorem 2) needs a constant time. For *Randomized Computation Trees* (RCTs) the above algorithm needs depth $\mathrm{O}(n)$ and Ben-Or's ([B83]) lower bound $\Omega(n \log n)$ holds for deterministic CTs. Our lower bound does not give nontrivial bounds for RDTs of degree m for this problem.

- Is there some analog of our Main Theorem also possible for randomized computation trees (RCTs) ?

References

[B83] M. Ben-Or, Lower Bounds for Algebraic Computation Trees, Proc. 15th ACM STOC (1983), pp. 80–86.

[BLY92] A. Björner, L. Lovasz and A. Yao, Linear Decision Trees: Volume Estimates and Topological Bounds, Proc. 24th ACM STOC (1992), pp. 170–177.

[BKL93] P. Buergisser, M. Karpinski, T. Lickteig, On Randomized Algebraic Test Complexity, J. of Complexity **9** (1993), pp. 231-251.

[DL78] D.P. Dobkin, R.J. Lipton, A Lower Bound of $\frac{1}{2}n^2$ on Linear Search Programms for the Knapsack Problem, J.Compt.Syst.Sci. **16** (1978), pp. 413–417.

[GK93] D. Grigoriev, M. Karpinski, Lower Bounds on Complexity of Testing Membership to a Polygon for Algebraic and Randomized Computation Trees, Technical Report TR-93-042, International Computer Science Institute, Berkeley, 1993.

[GK94] D. Grigoriev, M. Karpinski, Lower Bound for Randomized Linear Decision Tree Recognizing a Union of Hyperplanes in a Generic Position , Research Report No. 85114-CS, University of Bonn, 1994.

[GKV95] D. Grigoriev, M. Karpinski, N. Vorobjov, Improved Lower Bound on Testing Membership to a Polyhedron by Algebraic Decision Trees, Proc. 36th IEEE FOCS (1995), pp. 258-265.

[GV88] D. Grigoriev, N. Vorobjov, Solving Systems of Polynomial Inequalities in Subexponential Time, Journal of Symbolic Comp. **5** (1988), pp. 37–64.

[L65] S. Lang, Algebra, Addison–Wesley, New York, 1965.

[M84] F. Meyer auf der Heide, A Polynomial Linear Search Algorithm for the n-Dimensional Knapsack Problem, J. ACM **31** (1984), pp. 668–676.

[M85a] F. Meyer auf der Heide, Nondeterministic versus Probabilistic Linear Search Algorithms, Proc. IEEE FOCS (1985), pp. 65–73.

[M85b] F. Meyer auf der Heide, Lower Bounds for Solving Linear Diophantic Equations on Random Access Machines, J. ACM **32** (1985), pp. 929–937.

[M85c] F. Meyer auf der Heide, Simulating Probabilistic by Deterministic Algebraic Computation Trees, Theoretical Computer Science **41** (1985), pp. 325–330.

[MT82] U. Manber and M. Tompa, Probabilistic, Nondeterministic and Alternating Decision Trees, Proc. 14th ACM STOC (1982), pp. 234–244.

[S83] M. Snir, Lower Bounds for Probabilistic Linear Decision Trees, Research Report 83–6, Dept. of Computer Science, Hebrew University of Jerusalem, 1983.

[SY82] J. M. Steele and A. C. Yao, Lower Bounds for Algebraic Decision Trees, J. of Algorithms **3** (1982), pp. 1–8.

[T51] A. Tarski, A Decision Method for Elementary Algebra and Geometry, University of California Press, 1951.

[Y92] A. Yao, Algebraic Decision Trees and Euler Characteristics, Proc. 33rd IEEE FOCS (1992), pp. 268–277.

[Y94] A. Yao, Decision Tree Complexity and Betti Numbers, Proc. 26th ACM STOC (1994), pp. 615–624.