A Lower Bound for Randomized Algebraic Decision Trees

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

November 25, 1996

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, $\Omega(n^2)$ randomized lower bound for the *knapsack problem* (which was previously only known for deterministic algebraic decision trees).

[*] 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 Non-distinctness* ($\Omega(n \log n)$ lower bound).

The paper is organized as follows. Section 2 introduces a necessary algebraic terminology. Section 3 discusses local cases of randomized computation trees and section 4 formulates our main lower bound theorem, and gives its applications to concrete problems.

2 Preliminaries

We use standard algebraic and topological notations:

Given a polynomial $f(x_1, \ldots, x_n) \in \mathbb{R}[X_1, \ldots, X_n]$ and a point $v \in \mathbb{R}^n$. Also, let the vectors $a_1 = ({a_1}^{(1)}, \ldots, {a_1}^{(n)}), \ldots, a_n = ({a_n}^{(1)}, \ \ldots \ , {a_n}^{(n)}) \in \mathbb{R}^n$ be given, we define an $n \times m$ matrix $A = ({a_i}^{(j)})$. Introduce new variables $Y_1, \ldots, Y_n$ and consider a polynomial $f^{(v;a_1,\ldots,a_n)}(Y_1, \ldots, Y_n) = f(v + A(Y_1, \ldots, Y_n))$, which obviously is the expansion of f with, the origin v and the coordinates, being the vectors $a_1, \ldots, a_n$.

Denote for brevity $g = f^{(v;a_1,\ldots,a_n)}$ 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 $\widetilde{\mathbb{R}(\boldsymbol{\varepsilon})}$ for $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_\ell)$, $0 \leq \ell < n$. Then the sign $\mathrm{sgn}(g(\varepsilon_1, \ldots, \varepsilon_n)) = \mathrm{sgn}(lm(g)(\varepsilon_1, \ldots, \varepsilon_n))$ and on the other hand 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 $\varepsilon_\ell^{(0)}$, $0 \leq l \leq n-1$. But then one should specify, what does it mean "considerably smaller", and it is more convenient to use infinitesimals.

3 Randomized Computation Trees: Local Case

Recall (see, e. g., [B83]) that a (deterministic) computation tree (CT) contains the nodes of two types: the computation nodes and the branching query ones. At the computation node one can compute a new polynomial by either addition or multiplication of two previously (on the path of the tree) computed polynomials. One can also replace one or both of the previously computed polynomials by a real constant or by one of the input variables $X_1, \ldots, X_n$. At the branching node, to which a previously computed polynomial f is assigned a CT branches to one of the three sons of the node according to the sign of $f(x_1, \ldots, x_n)$. Herewith $(x_1, \ldots, x_n) \in \mathbb{R}^n$ is an input vector. Every leaf is labeled by "yes" or "no".

A randomized computation tree (RCT) is a collection $T = \{T_\alpha\}$ of CT T_α each chosen with its own probability p_α, such that $\sum_\alpha p_\alpha = 1$. The depth is defined as the maximum of depths. We say that T tests a (semialgebraic) set $S \subset \mathbb{R}^n$ if T gives the correct answer for any point $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with the probability $> \frac{2}{3}$ (cf. [M85a]).

For a family of polynomials $f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n]$, vectors

$a_1, \ldots, a_n \in \mathbb{R}^n$ and a point $v \in \mathbb{R}^n$ denote by $\mathrm{Var}(f_1, \ldots, f_s) = \mathrm{Var}^{(v;a_1,\ldots,a_n)}(f_1, \ldots, f_s)$ (we omit $a_1, \ldots, a_n$ and v from the notation since in this section they would be fixed) the number of the variables among $Y_1, \ldots, Y_n$ occuring in the leading terms $lm(f_1^{(v;a_1,\ldots,a_n)}), \ldots, lm(f_s^{(v;a_1,\ldots,a_n)})$. For a CT T_α denote by $\mathrm{Var}(T_\alpha)$ the maximum of $\mathrm{Var}(f_1, \ldots, f_s)$, where $f_1, \ldots, f_s$ are the testing polynomials along one path in T_α, over all the paths of T_α. For RCT T by $E(Var(T))$ we denote the expectation of $\mathrm{Var}(T_\alpha)$, also denote

$\mathrm{Var}(T) = max_\alpha\{\mathrm{Var}(T_\alpha)\}$

For the next lemma fix $v; a_1, \ldots, a_n$.

Lemma 1 The depth of T is greater or equal to $\frac{1}{2} Var\,(T)$.

Proof. There exists CT T_α and a certain path of it with the testing polynomials $f_1, \ldots, f_s$ along it, such that $\mathrm{Var}(f_1, \ldots, f_s) = \mathrm{Var}(T)$. Observe that for each $1 \leq \ell \leq s$ the polynomials $f_1, \ldots, f_\ell$ depend on at most 2ℓ variables among $Y_1, \ldots, Y_n$. It is easy to prove be induction in ℓ noticing that each polynomial among $f_1, \ldots, f_s$ could introduce into the game at most 2 new variables due to the definition of CT. Hence $s \geq \frac{1}{2}\mathrm{Var}(f_1, \ldots, f_s)$. □

Remark For d-degree decision trees T' (see the next section one easily obtains the lower bound $\frac{1}{d}\mathrm{Var}(T')$.

Denote $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) : x_i \geq 0, 1 \leq i \leq n\}$ and $\mathbb{R}^n_0 = (\mathbb{R} \setminus \{0\})^n$.

Theorem 1 Any RCT which recognizes

a) $\mathbb{R}^n_+$,

b) $\mathbb{R}^n_0$

has the depth greater or equal to $n/4$.

Proof. Observe that by the Tarski's transfer principle ([T51]) the same RCT recognizes the set

$F^n_+ = \{(x_1, \ldots, x_n) \in F^n : x_i \geq 0, 1 \leq i \leq n\}$ (respectively the set $F^n_0 = (F \backslash \{0\})^n$) if to consider RCT over the real closure $F = \widetilde{\mathbb{R}(\varepsilon)}$. Below we take $v = 0$, and a_i is i-th ort,$1 \leq i \leq n$ (see the section 2).

Let RCT $T^{(+)}$ recognize $\mathbb{R}^n_+$. Consider the points $E = (\varepsilon_1, \ldots, \varepsilon_n)$, $E_i^{(+)} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, -\varepsilon_i, \varepsilon_{i+1}, \ldots, \varepsilon_n)$, $1 \leq i \leq n$. There exists CT $T_\alpha^{(+)}$ for which the output for E is correct (i. e. "yes") and for at least of $n/2$ among the points $E_i^{(+)}$, $1 \leq i \leq n$ the outputs are correct (i. e. "no") as well. Take one of such $1 \leq i_0 \leq n$ and consider a path in $T_\alpha^{(+)}$ which provides the output for the point E. Denote by $f_1, \ldots, f_s$ the testing polynomials along this path. We claim that X_{i_0} occurs in one of the leading terms $lm(f_1), \ldots, lm(f_s)$. Indeed, otherwise $\mathrm{sgn}(f_\ell(E_{i_0}^{(+)})) = \mathrm{sgn}(lm(f_\ell(E_{i_0}^{(+)}))) =$

$\mathrm{sgn}(lm(f_\ell(E))) = \mathrm{sgn}(f_\ell(E))$, $1 \le \ell \le s$, 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. Thus $\mathrm{Var}(f_1, \dots, f_s) \ge n/2$ and lemma 1 implies that the depth of $T^{(+)}$ is greater or equal to $n/2$ that proves the theorem in case a).

In the case b) consider the points $E_i^{(0)} = (\varepsilon_1, \dots, \varepsilon_{i-1}, 0, \varepsilon_{i+1}, \dots, \varepsilon_n)$, $1 \le i \le n$ and argue as in the case a), replacing $E_i^{(+)}$ by $E_i^{(0)}$, $1 \le i \le n$. □

Remark For d-degree randomized decision tree one can get the lower bound $\frac{1}{2d}n$ following the proof of the theorem 1 and replacing the reference to lemma 1 by the remark after lemma 1.

4 Randomized Decision Trees: Global Case and Applications

Recall (see e. g. [B83]) that to every node of a d-decision tree (d - DT) a testing polynomial of degree at most d is assigned. Similar as in the section 1 one defines a randomized d-decision tree (d - RDT), see e. g. [M85a].

Let $H_1, \dots, H_m \subset \mathbb{R}^n$ be hyperplanes. We consider recognizing by d - RDT of one of two following sets; either the complement to the arrangement $S = \mathbb{R}^n \setminus \bigcup_{1 \le i \le m} H_i$ (cf. [M85a], [GK94]) or the polyhedron $P = \bigcap_{1 \le i \le m} \{H_i \ge 0\}$, (cf. [GKV95]), although the main result could be extended to more general sets constructed by means of the hyperplanes.

Observe that in [BKL93] a RCT is exhibited which recognizes in $0(n)$ time the set $\Big\{(x_1, \dots, x_n, y_1, \dots, y_n) \epsilon \mathbb{R}^{2n}$, where $(y_1, \dots, y_n)$ is a permutation of $(x_1, \dots, x_n)\}$. This shows that the results of this section ascertained for RDT can not be directly extended to RCT.

Under a k-face L_k of S (cf [M85a]) we mean k-dimensional plane of the form $\cap H_i$ for a certain subset $I \subset \{1, \dots, m\}$. Throughout this section we fix the following unique representation of L_k as an intersection. Take the maximal possible i_{n-k} such that $L_k \subset H_{i_{n-k}}$. Then take the next maximal possible i_{n-k-1} such that $L_k \subset H_{i_{n-k-1}}$ and $\dim\left(H_{i_{n-k-1}} \bigcap H_{i_{n-k}}\right) < \dim H_{i_{n-k}}$, obviously $i_{n-k-1} < i_{n-k}$. If $i_{l+1} < i_{l+2} < \dots < i_{n-k}$ are already yielded, where $l \ge 1$, take the maximal possible i_l such that $L_k \subset H_{i_l}$ and $\dim\left(H_{i_l} \bigcap H_{i_{l+1}} \bigcap \dots \bigcap H_{i_{n-k}}\right) < \dim\left(H_{i_{l+1}} \bigcap \dots \bigcap H_{i_{n-k}}\right)$, obviously $i_l < i_{l+1}$. Similarly we define a k-face of the polyhedron P.

Now we need an extension of Var function used in the section 1 from the points (so, 0-faces), to k-faces. The next consideration concerns both cases of the sets S and P.

Let a point v_{L_k} belong to a k-face L_k defined by an intersection $H_{i_1} \bigcap \dots \bigcap H_{i_{n-k}}, i_1 < \dots < i_{n-k}$ as yielded above, and v_{L_k} does not belong to the faces of smaller dimensions. We choose a coordinate system $(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k})$ in a neighbourhood of v_{L_k} taking the last $n-k$ orts orthogonal to the hyperplanes $H_{i_1}, \dots H_{i_{n-k}}$ respectively, and the first k orts in L_k in an arbitrary way. Expanding each polynomial $f \epsilon \mathbb{R}[X_1, \dots X_n]$ in the variables $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$ we define its leading term (cf. section 1) $lm^{(v,Z_1,\dots,Z_k,Y_1,\dots,Y_{n-k})}(f)$ considering f as a polynomial from $\mathbb{R}(Z_1, \dots, Z_k)[Y_1, \dots, Y_{n-k}]$ i. e. first take the terms with the least degree in Y_{n-k}, after that among them with the least degree in Y_{n-k-1} and so on till the variable Y_1.

Notice that $lm(g_1 g_2) = lm(g_1) lm(g_2)$ for any $g_1, g_2 \epsilon \mathbb{R}[X_1, \dots, X_n]$ (we omit here the indices $v, Z_1, \dots, Y_{n-k}$). For a family of polynomials $f_1, \dots, f_s \epsilon \mathbb{R}[X_1, \dots, X_n]$ we define $var_k^{(v,Z_1,\dots,Z_k,Y_1,\dots,Y_{n-k})}(f_1, \dots, f_s)$ as the number of all variables among $Y_1, \dots, Y_{n-k}$ which occur in at least one of the terms $lm^{(v,Z_1,\dots,Z_k,Y_1,\dots,Y_{n-k})}(f_\ell), 1 \leq \ell \leq s$ (cf. section 1). We usually omit $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$ in the notations if it does not lead to the ambiguity.

Now we fix $0 \leq k < n$ and on any k-face $L = L_k$ we fix a point v_L, at each point v_L fix a coordinate system as above. Suppose, a d-decision tree T' is given which recognizes either the complement S to an arrangement or P (cf.above). For each point v_L we define $var^{(v_L)}(T')$ as

$max_{f_1,\dots,f_s}\{var^{(v_L)}(f_1, \dots, f_s)\}$ where $(f_1, \dots, f_s)$ are the testing polynomials along a path of T', and the maximum is taken over all paths.

Lemma 2 Assume that for some $c > 0$ there are at least M k-faces L such that $var^{(v_L)}(T') \geq c(n-k)$. Then the depth of T' is greater than $\Omega((n-k)\log m)$, provided that $M > m^{(n-k)(1-c+c_0)} d^{(c+c_0)(n-k)}$ for a certain $c_0 > 0$.

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

By a flag of L we mean the sequence of imbedded 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 polynimials $f_1, \ldots, f_s$ 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 $f_1, \ldots, f_s$ 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 $f_1, \ldots, f_s$, we do not label this vertex of $\mathcal{T}$. To complete the construction of the first level of $\mathcal{T}$, we represent the polynomial $f_j = \tilde{f}_j Y_{n-k}^{m_j} \mathcal{L}_{H_{r_1}}^{m_{j,1}} \ldots \mathcal{L}_{H_{r_p}}^{m_{j,p}}$, $1 \leq j \leq s$ as a product for maximal possible $m_j, m_{j,1}, \ldots, m_{j,p}$ where $i_{n-k} < r_1 < \ldots < r_p$ and $\mathcal{L}_{H_{r_1}}, \ldots, \mathcal{L}_{H_{r_p}}$ are all linear polynomials determining hyperplanes $H_{r_1}, \ldots, H_{r_p}$ which divide f_j with the indices $r_1, \ldots, r_p$ greater than i_{n-k}. We assign to the constructed vertex the polynomials $f_j^{(1)}(Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-1}) = \widetilde{f_j}(Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-1}, 0)$, $1 \leq j \leq s$. One could view the polynomial $f_j^{(1)}$ as being defined on the hyperplane $H_{i_{n-k}}$.

Observe that the linear polynomials $\mathcal{L}_{H_{r_1}} \ldots \mathcal{L}_{H_{r_p}}$ 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 $Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k}$ 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

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

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 $H = H_{i_{n-k}} \bigcap \cdots \bigcap H_{i_{n-k-\ell+1}}$.

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 $f_j^{(\ell)}\left(Z_1, \ldots, Z_k, Y_1, \ldots, Y_{n-k-\ell}\right), 1 \leq j \leq s$ 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 $H_{i_{n-k}} \bigcap \cdots \bigcap H_{i_{n-k-\ell+1}} \bigcap H_{i_{n-k-\ell}} = H \bigcap H_{i_{n-k-\ell}}$.

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

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

which was to be shown.

Substituting now

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

we obtain a bound

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

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) and proved that to this path at most

$$m^{(n-k)(1-c)} \left(\frac{sd}{c(n-k)} \right)^{c(n-k)} 2^{c_1(n-k)}$$

k-faces L for an appropriate $c_1 > 0$ could correspond. 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

$$M \leq 3^t m^{(n-k)(1-c)} \left(\frac{td}{c(n-k)} \right)^{c(n-k)} 2^{c_1(n-k)}$$

which implies the lower bound $\Omega((n-k)\log m)$, taking into account the assumptions of the lemma. □

Now we apply lemma 2 to obtain lower bound on the depth of d - RDT T, which recognizes either the complement S to an arrangement or a polyhedron P. Still we fix $0 \leq k = c_1 n < n$, $c_1 > 0$. Take a k-face L. Then at any of its points v we can choose the coordinates $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$. For any polynomial $f \in \mathbb{R}[X_1, \dots, X_n]$ considering it as an element of the ring $\mathbb{R}(Z_1, \dots, Z_k)[Y_1, \dots, Y_{n-k}]$ (cf. above) we take its leading term in the variables $Y_1, \dots, Y_{n-k}$ let

$$lm^{(v)}(f) = \widetilde{f} Y_1^{q_1} \cdot \ldots \cdot Y_{n-k}^{q_{n-k}},$$

where $\widetilde{f} \in \mathbb{R}[Z_1, \dots, Z_k]$. If we take any other point $v_1 \in L$ then the leading term $lm^{(v_1)}(f) = \widetilde{f}^{(v_1)} Y_1^{q_1} \cdot \ldots \cdot Y_{n-k}^{q_{n-k}}$ with respect to the coordinates in a neighbourhood of v_1. In particular, if the value of $\widetilde{f}$ at v_1 is not zero then $\widetilde{f}^{(v_1)}$ contains a constant nonzero term, then the leading term of f with respect to all variables $\widetilde{Z}_1, \dots, \widetilde{Z}_k, Y_1, \dots, Y_{n-k}$ in a neighbourhood of v_1 is $Y_1^{q_1} \cdot \ldots \cdot Y_{n-k}^{q_{n-k}}$. Thus, the variety of points in L in which the leading term of f vanishes, has the dimension less than k.

For each k-face L choose a point $v_L \epsilon L$ such that for any testing polynomial from any of the d - DT T_α from the collection determining d - RDT T its leading term does not vanish at v_L. Choose now the coordinates in the neighbourhood of v_L as described above: $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$, in particular, the coordinates of v_L are all zeroes.

Similar to the section 1 take a point $E_L = (0, \dots, 0, \varepsilon_1, \dots, \varepsilon_{n-k})$ and in the case of the polyhedron P take the points $E_{L,i}^{(+)} = (0, \dots, 0, \varepsilon_1, \dots, \varepsilon_{i-1}, -\varepsilon_i, \varepsilon_{i+1}, \dots, \varepsilon_{n-k}), 1 \leq i \leq n-k$. The point $E_L \in$

$P, E_{L,i}^{(+)} \notin P, 1 \leq i \leq n-k$. Using the theorem on the diminishing of the error of d - RDT at the expence of the increasing the depth [M85a], we can assume *w.l.o.g.* that for arbitrary fixed in advance $c < 1$ the probability that the output of T for the point E_L is correct (so, "yes") and the outputs for at least $c(n-k)$ points among $E_{L,i}^{(+)}, 1 \leq i \leq n-k$ are correct (so, "no"), is greater than $\frac{2}{3}$.

In the case of the complement S to arrangement we consider the points $E_{L,i}^{(0)} = (0, \dots, 0, \varepsilon_1, \dots, \varepsilon_{i-1}, 0, \varepsilon_{i+1}, \dots, \varepsilon_{n-k}), 1 \leq i \leq n-k$ (cf. section 1). The same arguing as in the previous paragraph holds for $E_{L,i}^{(0)}$ instead of $E_{L,i}^{(+)}, 1 \leq i \leq n-k$.

Let the set P (resp. S) contain M_1 k-faces. Because the described property for the points $E_L, E_{L,i}^{(+)}$ (resp. $E_i, E_{L,i}^{(0)}$) is valid for all k-faces L, we deduce that there exists d - DT T_α for which this property holds for at least $\frac{1}{2}M_1$ of k-faces. Take any k-face L among these $\frac{1}{2}M_1$ and a path (which is unique) of T_α along which T_α runs for the (input) point E_L. Let $f_1, \dots, f_s$ be the testing polynomials along this path. We claim that for each $1 \leq i \leq n-k$ such that the output of T_α for $E_{L,i}^{(+)}$ (resp. $E_{L,i}(0)$) is correct, the variable Y_i occurs in one of the leading terms $lm^{(v_L)}(f_j)$, $1 \leq j \leq s$ (in the notations of lemma 2 this means that $\mathrm{Var}^{(v_L)}(T_\alpha) \geq c(n-k)$). Indeed, otherwise, arguing as in the proof of the theorem 1, we get that $\mathrm{sgn}(f_j(E_{L,i}^{(*)})) = \mathrm{sgn}(lm^{(v_L)}(f_j(E_{L,i}^{(*)}))) = \mathrm{sgn}(lm^{(v_L)}(f_j(E_L))) = \mathrm{sgn}(f_j(E_L))$ where $*$ stands for either $+$ or 0, respectively, taking into account the choice of the points v_L.

Finally, we apply lemma 2 to the set of $M = \frac{1}{2}M_1$ k-faces and obtain the following theorem.

Theorem 2 If d-RDT T recognizes either the complement to an arrangement or a polyhedron generated by $m \geq n$ hyperplanes, which has M k-faces, for some $k = nc_1$, $c_1 < 1$, then the depth of T is greater than $\Omega(n \log m)$, provided that $M > \Omega(m^{(n-k)(1-c+c_0)} d^{(n-k)(c+c_0)})$ for certain $c_0 > 0, 0 < c < 1$.

Corollary 1 When $d = const$ the statement of the theorem holds, provided that $M > \Omega(m^{nc_0})$ for a certain $c_0 > 0$.

To prove the corollary notice that c could be taken as close to 1 as desired.

As a consequence we get $\Omega(n^2)$ lower bound for the depth of d - RDT, recognizing the knapsack (cf. [M85a]), when $d = const$. Similarly, for the distinctness problem $\{(x_1, \dots, x_n) : x_i \neq x_j, i \neq j\} \in \mathbb{R}^n$ we get $\Omega(n \log n)$ lower bound (cf. [M85a]), when $d = const$.

Corollary 2 For degree d fixed:

(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 Non-Distinctness Problem* is $\Omega(n\ log\ n)$.

5 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 half-spaces. 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 $O(n)$ 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 Theorem 2 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.

[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.

[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 $\underline{3}$ (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.