A Lower Bound for Randomized Algebraic Decision Trees *

Dima Grigoriev † Marek Karpinski ‡

Friedhelm Meyer auf der Heide § Roman Smolensky ¶

Abstract

We prove the first *nontrivial* (and *superlinear*) lower bounds on the depth of *randomized* algebraic decision trees (with two-sided error) for problems being finite unions of hyperplanes and intersections of halfspaces, solving a long standing open problem. As an application, among other things, we derive, for the first time, an $\Omega(n^2)$ *randomized* lower bound for the *Knapsack Problem*, and an $\Omega(n \log n)$ *randomized* lower bound for the *Element Distinctness Problem* which were previously known only 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 lower bound technique for deterministic algebraic decision trees without making use of Milnor's bound on Betti number of algebraic varieties.

*A preliminary version of this paper appeared in [GKMS96]

†Dept. of Computer Science and Mathematics, Penn State University, University Park. Research partially supported by NSF Grant CCR-9424358. Email: dima@cse.psu.edu

‡Dept. of Computer Science, University of Bonn, 53117 Bonn. Research partially supported by DFG Grant KA 673/4-1, by the ESPRIT BR Grants 7097 and EC-US 030. Email: marek@cs.uni-bonn.de

§Heinz Nixdorf Institute and Computer Science Department, University of Paderborn, 33098 Paderborn. Supported partially by DFG Grant ME 872/4-1 and ESPRIT BR Grant 7141 (ALCOMII). Email: fmadh@uni-paderborn.de

¶Dept. of Computer Science, University of Bonn, 53117 Bonn. Roman Smolensky died on 19 October, 1995 in New York. This paper is also dedicated to the memory of Roman by the other co-authors.

1 Introduction.

Starting with [MT85], [S85], [M85a] and [M85b] there has been a continued effort in the last decade to understand the 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. [Y81], [SY82], [B83], [BLY92], [Y92], [GKV97], [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 finite 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 problems being finite unions of hyperplanes as well as intersections of halfspaces. This provides also a new method for proving lower bounds 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 randomized algebraic decision and computation trees. Section 3 overviews the known results in the area. Section 4 summarizes our results and applies them for a number of concrete problems. Section 5 gives an outline of the lower bound proof, and Sections 6 and 7 give the proof of the Main Theorem. Section 8 contains the complexity lower bound for deterministic decision trees under less restrictive conditions than in the Main Theorem for their randomized counterparts.

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 depending 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*.

We deal in this paper with *randomized algebraic decision trees* of degree d, *d-RDTs* for short. There are several variants of this model known in the literature (see, e.g. [M85a], [M85b], [M85c], [MT85]). One of the most natural variants allows coin flipping nodes and charges for the random bits used. Other variants ignore the costs for random choices, or pull flipping nodes out of the decision trees, and view a *d-RDT* as a finite collection $\{T_\alpha\}$ of *d-DTs* T_α with the assigned (rational) probabilities p_α, $\sum_\alpha p_\alpha = 1$, of choosing T_α out of the collection $\{T_\alpha\}$. In the latter case we also do not charge for the randomization. This variant is easily seen to be equivalent (up to constant factors in the depth) to the "equal likelihood" variant with all trees T_α having equal probabilities p_α ($p_\alpha = \frac{1}{|\{T_\alpha\}|}$). In this paper we use the last model mentioned above and define a *d-RDT* as a finite collection $\{T_\alpha\}$ of *d-DT* 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 (two-sided) *error probability.*

We note that the class of sets $S \subseteq \mathbb{R}^n$ *recognized* by *d-RDT*s is closed under complement.

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]).

We note that *RDTs* are robust under changes of the error probability. This fact was first observed in [BG81], and then explicitely formulated and proven for *RDTs* in [M85c](Claim 1.3):

Fact 1 (Amplification): *Let $\gamma', \gamma \in (0, \frac{1}{2})$ be constants. Every d-RDT* *of depth T recognizing $L \subseteq \mathbb{R}^n$ with error probability γ' can be simulated by a d-RDT of depth $O(T)$ recognizing L with error probability γ.*

A standard notion of runtime of randomized algorithms is *worst case expected time*, i.e., the maximum over the expected runtimes of the algorithm started with input x, maximum taken over all inputs of length n. In [MT85] it is shown that any *d-RDT* with worst case expected time T can be simulated by one with depth $O(T)$. Using Fact 1, the error probability can remain unchanged.

3 Known Results.

The most important results in connection to this research are the variants of the component counting lower bounds for deterministic computations: Let $L \subseteq \mathbb{R}^n$ have q connected components. Then each LDT for L has depth $\Omega(\log(q))$ [DL78], each *d-DT* for L has depth $\Omega(\frac{\log(q)}{\log(d)} - n)$ (can be deduced from [B83], see also [SY82]), 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 problems.

Consider, e.g., an Integer Programming Problem $L_{n,s} = \{x \in \mathbb{R}^n, \exists a \in \{0, \ldots, s\}^n : xa = s\}$ cf. [M85a, M85b]. For each $s \geq 1$, the family $\{L_{n,s}, n \in \mathbb{N}\}$, restricted to integer inputs, is *NP-complete*, for $s = 1$ this is the famous *Knapsack Problem*.

As shown in [DL78] for $s = 1$ and in [M85b] for arbitrary s, $\mathbb{R}^n - L_{n,s}$ has $(s+1)^{\Omega(n^2)}$ many connected components, yielding lower bounds $\Omega(n^2 \log(s+1))$ in the above models. In [M84] and [M93] it is shown that all these NP-complete problems have *polynomial* depth LDTs, for their n-dimensional restrictions.

A further important example is the *Element Distinctness Problem* [B83], with the connected components bound $n!$, and therefore a deterministic lower bound $\Omega(n \log n)$. Other related problems are: *Set Disjointness*, *Resultant*, and the *ε-Approximation Knapsack Problem* (cf. [B83]).

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 [GK94].)

In [BKL93] it is shown that there are benefits if the randomization is used in *CTs*: Consider the language $\{(x, y) \in \mathbb{R}^{2n} : y$ is a 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 closed halfspaces. S is often called a *linear arrangement*, S^+ is a *polyhedron.* A *k-face* L of S is a k-dimensional plane 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^+. 0-*faces* are also called *vertices.*

We prove the following lower bound.

Main Theorem: *For any constants $0 \leq \gamma < \frac{1}{2}, c_1 > 0$ and $\tau > \delta \geq 0$, there exists a constant $c_0 > 0$ satisfying the following. Let $H_1, \ldots, H_m$ be hyperplanes in $\mathbb{R}^n$, $S = \bigcup_{i=1}^{m} H_i, S^+ = \bigcap_{i=1}^{m} H_i^+$ for $m \geq n$. If S or S^+ has at least $m^{\tau(n-k)}$ k-faces for some $k \in \{0, \ldots, n-1\}$, then each d-RDT for S or S^+ with error probability γ has depth greater than $c_0(n-k)\log(m)$, provided that $d \leq c_1 m^{\delta}$.*

Obviously, in a *generic* arrangement (when all the intersections of subfamilies of the hyperplanes are pairwise distinct) the number of k-faces equals $\binom{m}{n-k}$ which is bounded from below by $m^{\Omega(n-k)}$, provided that $m \geq \Omega(n^{1+\sigma_0})$ for some $\sigma_o > 0$, thus the bound on the number of k-faces from the theorem is attainable. In case of a polyhedron, the upper bound $O(m^{\min\{n-k,[\frac{n}{2}]\}})$ [E87] on the number of k-faces is also attainable (for example for the dual to the cyclic polyhedra, see [M71], also [GKV97]).

Our Main Theorem yields directly the following concrete applications:

Corollary 1: *For $\delta > 0$ sufficiently small, the following randomized lower bounds hold.*

(1.1) $\Omega(n^2\, log(s+1))$ *is a lower bound for the depth of any d-RDT (with $d = O((s+1)^{\delta n})\delta < \frac{1}{16}$) recognizing the Integer Programming Problem $L_{n,s}$.*

(1.2) $\Omega(n^2)$ *is a lower bound for the depth of any d-RDT (with $d = O(2^{\delta n})$, $\delta < \frac{1}{16}$.) recognizing the Knapsack Problem or the ε-Approximation Knapsack Problem.*

Corollary 2: $\Omega(n\, log\, n)$ *is a lower bound for the depth of any d-RDT (with $d = O(n^\delta), \delta < \frac{1}{2}$) for the following problems:*

(2.1) *Element Distinctness,*

(2.2) *Set Disjointness,*

(2.3) *Resultant* (Decision Version).

Proof of Corollary 1: Observe that (1.1) entails (1.2).
We prove now (1.1). Without loss of generality we assume that $n = 8r$ for certain integer r. Our purpose is to prove that $L_{n,s} = \{x \in \mathbb{R}^n, \exists a \in \{0, \ldots, s\}^n : xa = s\}$ contains at least $(s+1)^{\frac{n^2}{16}}$ vertices. Consider the following system of linear equations $E_j, 1 \leq j \leq \frac{n}{2}$ in $\frac{n}{2}$ variables $x_1, \ldots, x_{\frac{n}{2}}$:
$E_1 : sx_1 = s; E_2 : sx_2 = s; E_3 : sx_1 + sx_2 + sx_3 = s; E_4 : sx_1 + sx_2 + sx_4 = s.$
For any $0 \leq l \leq r-2$:
$E_{4l+5} : sx_1 + sx_{4l+3} + x_{4l+4} + x_{4l+5} = s;$
$E_{4l+6} : sx_1 + sx_{4l+3} + x_{4l+4} + x_{4l+6} = s;$
$E_{4l+7} : sx_1 + x_{4l+5} + x_{4l+7} = s;$
$E_{4l+8} : sx_1 + x_{4l+5} + x_{4l+8} = s$
Obviously, the system $K(X_1, \ldots, X_{\frac{n}{2}}) = (s, \ldots, s)^T$ has a unique solution $x_{4l+1} = x_{4l+2} = (s+1)^l; x_{4l+3} = x_{4l+4} = -(s+1)^l$ for $l = 0, \ldots, r-1$.

Let C be any $\frac{n}{2} \times \frac{n}{2}$ matrix with the property that all its entries being integers in the range from 0 to s such that its columns with the numbers $4l+1$, $4l+2$, $4l+3$ for all $0 \le l \le \frac{n}{8} - 1$, consist of zeroes. Consider the following $n \times n$ matrix

$$A(C) = \begin{pmatrix} I & C \\ 0 & K \end{pmatrix}$$

where I is the unit $\frac{n}{2} \times \frac{n}{2}$ matrix.

Evidently, the linear $n \times n$ system $A(C) \cdot X = (s, \dots, s)^\top$ has a unique solution $X(C) = (x_1, \dots, x_n)$. Therefore, $X(C)$ is a vertex of $L_{n,s}$.

Moreover, we claim that for different choices of above matrices C, C' the respective vectors $X(C)$ and $X(C')$ are different. Indeed, $(x_{\frac{n}{2}+1}, \dots, x_n) = (1, 1, -1, -1, (s+1), (s+1), -(s+1), -(s+1), \dots, (s+1)^{r-1}, (s+1)^{r-1}, -(s+1)^{r-1}, -(s+1)^{r-1})$. For $i < \frac{n}{2}$ we have $x_i = s - \sum_{0 \le j \le r-1} a_j (s+1)^j$, $a_j \in \{0, \dots, s\}$, herewith i-th row of C equals to $(a_0, 0, 0, 0, a_1, 0, 0, 0, a_2, 0, 0, 0, \dots, a_{r-1}, 0, 0, 0)$. Since $0 \le a_j \le s$ we get that the i-th components of the vectors $X(C)$, $X(C')$ differ, provided that C and C' differ in the i-th row.

Thus, we can choose $(s+1)^{\frac{n^2}{16}}$ distinct matrices of the form C and thereby so many distinct vertices of $L_{n,s}$ of the form $X(C)$. In other words, in the application of the main theorem we have $m = (s+1)^n$, $k = 0$ with at least $(s+1)^{\frac{n^2}{16}} = m^{\frac{1}{16}n}$ of 0-faces in the arrangement $L_{n,s}$. □

Proof of Corollary 2: The proof of (2.2) is similar to the proof of (2.1), and obviously (2.2) implies (2.3) (cf. [B83]).

We give now a proof of (2.1). Observe that the faces of the arrangement $\bigcup_{i<j}\{x_i = x_j\}$ (which represents the *Distinctness Problem*) are in one-to-one correspondence with all partitions of the set $\{1, \dots, n\}$ (except the partition into singletons $\{1\}, \dots, \{n\}$). Indeed, the partition $A_1 \cup \dots \cup A_k = \{1, \dots, n\}$ corresponds to the face being the intersection of all the hyperplanes $\{x_i = x_j\}$ for $i, j \in A_l$, $1 \le l \le k$, and the dimension of this face equals to k. Taking $k = \frac{n}{2}$ (let n be even w. l. o. g.) we obtain at least $k!$ partitions $A_1, \dots, A_k$ of the following form: for each permutation $\pi \in S_k$, put $A_i = \{i, k + \pi(i)\}$, $1 \le i \le k$. Thus, in the application of the theorem we have $k = \frac{n}{2}$ (provided that n is even), $m = \binom{n}{2}$, and the number of k-faces equal to $k! \ge m^{\delta(n-k)}$, for arbitrary $\delta < \frac{1}{2}$. □

5 Outline of the Lower Bound Proof.

Assume that the arrangement $s = \bigcup_{1\le i\le m} H_i$ has a vertex v, then after a certain permutation one can suppose that the hyperplanes $H_i = \{x \in \mathbb{R}^n : a_i x = b_i\}, 1 \le i \le n$ are affinely independent and $\{v\} = H_1 \cap \cdots \cap H_n$.

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 $a_1, \ldots, a_n$; $f^{(v;H_1,\ldots,H_n)}\ (Y_1, \ldots, Y_n) := f(v + A^{-1}(Y_1, \ldots, Y_n))$. 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. [L84]) 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 [L84]). 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 the less number of 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 a generic point $v \in L$ 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 (in Lemma 1 we utilize just introduced notations). The following chapters contain their proofs.

Lemma 1: *Let T be a d-RDT (or an RCT) recognizing*
a) an arrangement $S = \bigcup_{1 \le i \le m} H_i$ for some hyperplanes $H_1, \ldots, H_m$ such that $L = \bigcap_{1 \le i \le n-k} H_i$ is a k-face of S , or
b) a polyhedron $S^+ = \bigcap_{1 \le i \le m} H_i^+$ such that for each $1 \le l \le n-k$ $\bigcap_{l \le i \le n-k} H_i$ is $(k+l-1)$-face of S^+,
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 $\mathbb{R}^n_0 = (\mathbb{R} \setminus \{0\})^n$. 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 3: Any RCT which recognizes $\mathbb{R}^n_+$ or $\mathbb{R}^n_0$ must have the depth greater than or equal to $\frac{1}{2}(1-2\gamma)^2 n$.

Corollary 4: Any d-RDT which recognizes $\mathbb{R}^n_+$ or $\mathbb{R}^n_0$ must have the depth greater or equal to $\frac{1}{d}(1-2\gamma)^2 n$.

Proofs of Corollaries 3 and 4: Because of Lemma 1 there is a CT (resp. d-DT) T_α from the definition of an RCT (resp. d-RDT) satisfying the condition $Var^{(v;X_1,\ldots,X_n)}(T_\alpha) \ge (1-2\gamma)^2 n$ (here $X_1, \ldots, X_n$ denote the coordinate hyperplanes). Take a path in T_α with the testing polynominals $f_1, \ldots, f_s$ along it for which $Var^{(v;X_1,\ldots,X_n)}(f_1, \ldots, f_s) \ge (1-2\gamma)^2 n$.

To complete the proof of the Corollary 3 we easily show by induction on l that the polynominals $f_1, \ldots, f_l$ $(1 \le l \le s)$ depend on at most $2l$ variables among $X_1, \ldots, X_n$, since any computation step along the path in T_α could introduce into the game at most two new variables (one could evidently view the sequence of computation steps along the path as a straight-line program in which each step is an arithmetic operation with at most two arguments). Therefore, $s \ge \frac{1}{2} Var^{(v;X_1,\ldots,X_n)}(f_1, \ldots, f_s)$. To complete the proof of the corollary 4 we notice that each term $lm(f_i^{(v;X_1,\ldots,X_n)})$, $1 \le i \le s$, depends on at most d variables, hence $s \ge \frac{1}{d} Var^{(v;X_1,\ldots,X_n)}(f_1, \ldots, f_s)$.

$\square$

Let $S = \cup_{i=1}^m H_i$ or $S^+ = \cap_{i=1}^m H_i^+$ for hyperplanes $H_1, \dots, H_m \subset \mathbb{R}^n$.

For a k-face L of S let $1 \le i_1 < \dots < i_{n-k} \le m$ be the (inverse lexicographically smallest) sequence of $(n-k)$ indices such that $L = H_{i_1} \cap \dots \cap H_{i_{n-k}}$, i.e. i_{n-k} is the maximal possible index such that $H_{i_{n-k}} \supset L$, and $i_{n-k-1} < i_{n-k}$ is the maximal possible index such that $H_{i_{n-k-1}} \supset L$ and $\dim H_{i_{n-k}} \cap H_{i_{n-k-1}} = n-2$, and so on. We fix this canonical representation of $L = H_{i_1} \bigcap \dots \bigcap H_{i_{n-k}}$ by means of the indices $i_{n-k} > \dots > i_1$.

Now we give the canonical representation of a k-face L of the polyhedron $S^+ = \bigcap_{1 \le i \le m} H_i^+$. W.l.o.g. one could assume that $\dim S^+ = n$ (indeed, otherwise one can replace *d-RDT* under consideration by *d-RDT* obtained by restricting this *d-RDT* on the plane, being the linear hull of S^+). For any k_1-dimensional face L_1 of S^+ there exist hyperplanes $H_{j_1}, \dots, H_{j_{n-k_1}}$ such that $L_1 = H_{j_1} \bigcap \dots \bigcap H_{j_{n-k_1}}$. Under a hyperface of a l-dimensional polyhedron we mean a $(l-1)$-plane which is its face of the dimension $(l-1)$. W.l.o.g. one could assume that all hyperplanes $H_1, \dots, H_m$ are hyperfaces of S^+.

By $\mathcal{H}_0$ denote the family of all hyperplanes H_i such that $H_i \supset L$ and H_i is a hyperface of S^+. Since S^+ is a convex polyhedron, any its face is an intersection of some its hyperfaces, in particular, any its face L_1 which contains $L \subset L_1$, could be represented as $L_1 = \cap_{H_i \in \mathcal{H}_0'} H_i$ for a suitable subfamily $\mathcal{H}_0' \subset \mathcal{H}_0$.

Assume that by recursion on l it is already produced a sequence of indices $i_{n-k} > \dots > i_{n-k-l+1}$, $0 \le l \le n-k-1$, such that $H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l_1+1}}$ is $(n-l_1)$-face of S^+ for every $0 \le l_1 \le l$. Denote the polyhedron $S_l^+ = (H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap S^+$. In addition, a family $\mathcal{H}_l \subset \{H_1, \dots, H_{i_{n-k-l+1}-1}\}$ is produced such that for any $H_i \in \mathcal{H}_l$ $(H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap H_i$ contains L and is a hyperface of S_l^+, and vice versa any hyperface of S_l^+ has the form $(H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap H_i$ for a certain $H_i \in \mathcal{H}_l$. Hence any face $L_1 \supset L$ of S_l^+ has the form $L_1 = (H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap (\bigcap_{H_i \in \mathcal{H}_l'} H_i)$ for a suitable subfamily $\mathcal{H}_l' \subset \mathcal{H}_l$.

To carry out the recursive step, take as i_{n-k-l} the maximal index such that $H_{i_{n-k-l}} \in \mathcal{H}_l$ (obviously, $i_{n-k-l} < i_{n-k-l+1}$). Then $L_0 = H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}} \bigcap H_{i_{n-k-l}}$ is a hyperface of S_l^+. Denote the polyhedron $S_{l+1}^+ = H_{i_{n-k-l}} \bigcap S_l^+$. Take as $\mathcal{H}_{l+1}$ the family of all $H_i \in \mathcal{H}_l$ such that

$L_0 \bigcap H_i$ is a hyperface of S^+_{l+1} (evidently, $L_0 \bigcap H_i \supset L$ since $H_{i_{n-k-l}}$, $H_i \in \mathcal{H}_l$). Due to the choice of i_{n-k-l} we have $\mathcal{H}_{l+1} \subset \{H_1, \dots, H_{i_{n-k-l}-1}\}$.

It remains to prove that for any hyperface L_2 of S^+_{l+1} such that $L_2 \supset L$, there exists $H_i \in \mathcal{H}_{l+1}$ such that $L_2 = L_0 \bigcap H_i$. According to the property of $\mathcal{H}_l$ there exist $H_{j_1}, H_{j_2} \in \mathcal{H}_l$ such that $L_2 = (H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap H_{j_1} \bigcap H_{j_2}$ and $L_2^{(1)} = (H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap H_{j_1}$, $L_2^{(2)} = (H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) \bigcap H_{j_2}$ are the hyperfaces of S^+_l. For any face of codimension 2 of a convex polyhedron there exist exactly two its hyperfaces containing this face, and furthermore this face coincides with the intersection of these two hyperfaces. Since the face L_2 of S^+_l lies in three its hyperfaces $L_2^{(1)}, L_2^{(2)}, L_0$, we have either $L_2^{(1)} = L_0$ or $L_2^{(2)} = L_0$. Let for definiteness $L_2^{(2)} = L_0$; then $H_{j_1} \in \mathcal{H}_{l+1}$ by the definition of $\mathcal{H}_{l+1}$, and $L_2 = L_0 \bigcap H_{j_1}$, that completes the recursive step.

Thus, at the end of the recursion we obtain a flag of L (which we treat as the claimed canonical representation of the k-face L) $H_{i_{n-k}} \supset H_{i_{n-k}} \bigcap H_{i_{n-k-1}} \supset \dots \supset H_{i_{n-k}} \bigcap H_{i_{n-k-1}} \bigcap \dots \bigcap H_{i_1} = L$ such that for each $0 \le l \le n-k$ $H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}$ is $(n-l)$-face of S^+ (the recursion on l implies that $\dim(H_{i_{n-k}} \bigcap \dots \bigcap H_{i_{n-k-l+1}}) = n-l$, cf. lemma 1). Just to satisfy these properties of the flag of L, the above inductive construction was required.

Let v_L be a generic point of L which belongs to no lower dimensional face of S or S^+. We choose a coordinate system $(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k})$ in a

an arbitrary way. Expanding each polynomial $f \in \mathbb{R}[X_1, \dots, X_n]$ in the variables $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k}$ we define its leading monomial (cf. beginning of this section) $lm^{(v_L)}(f)$ considering f as a polynomial from $\mathbb{R}[Z_1, \dots, Z_k][Y_1, \dots, Y_{n-k}] \subset \mathbb{R}(Z_1, \dots, Z_k)[Y_1, \dots, Y_{n-k}]$, i.e. first taking the monomials with the least degree in the variable Y_{n-k}, after that among them with the least degree in Y_{n-k-1} and so on till the variable Y_1. Let T' be an $d - DT$. We abbreviate $\mathrm{Var}_k^{(v_L, H_{i_1}, \dots, H_{i_{n-k}})}(T')$ by $\mathrm{Var}^{(v_L)}(T')$.

Lemma 2: *Assume that, for some $c > 0$, there are at least M k-faces L of S or S^+ with $Var_k^{(v_L)}(T') \ge c(n-k)$. Then the depth t of T' fulfils $M \le 3^t \cdot m^{(1-c)(n-k)} \cdot (\frac{td}{c(n-k)})^{c(n-k)} \cdot 6^{n-k}$.*

Using these lemmas it is easy to conclude the Main Theorem:

First consider a d-RDT for S or S^+ with error probability $\gamma \in (0, \frac{1}{2})$ small enough such that $\beta := \tau - (1 - (1-\delta)c) > 0$, where $c := (1-2\gamma)^2$ and τ, δ chosen as in the Main Theorem.

Then Lemma 1 yields for each k-face L of S or S^+: $Var_k^{(v_L)}(T_\alpha) \geq c \cdot (n-k)$ for at least a fraction of $\frac{1-2\gamma}{2-2\gamma}$ of the T_α's. Thus there is a T_α with $Var_k^{(v_L)}(T_\alpha) \geq (1-2\gamma)^2(n-k)$ for $\frac{1-2\gamma}{2-2\gamma} \cdot N$ many k-faces L of S or S^+.

Therefore Lemma 2 implies that

$\frac{1-2\gamma}{2-2\gamma} m^{\beta(n-k)} \leq 3^t (\frac{t}{c(n-k)})^{c(n-k)} \cdot c_1^{c(n-k)} 6^{n-k}$ if $d \leq c_1 m^\delta$.

This proves the lower bound claimed in the Main Theorem for the choice of γ described above. Using Fact 1 yields the Main Theorem for arbitrary $\gamma \in (0, \frac{1}{2})$. $\square$

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 $H_n \supset H_n \cap H_{n-1} \supset \cdots \supset H_n \cap \cdots \cap H_1 = \{(0, \ldots, 0)\}$ is the canonical representation (flag) of 0-face $\{(0, \ldots, 0)\}$ (in both cases a), b)), and herewith the H_i's, $1 \leq i \leq n$, are defined by $\{x \in \mathbb{R}^n : x_i = 0\}$, in other words the linear transformation $(y_1, \ldots, y_n) \mapsto v + A(y_1, \ldots, y_n)$ is the identity.

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

Consider the points $E = (\varepsilon_1, \ldots, \varepsilon_n)$ and $E_i^{(+)} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, -\varepsilon_i, \varepsilon_{i+1}, \ldots, \varepsilon_n), i = 1, \ldots, n$. We show that $E \in S^+$. Take any hyperplane $H_l = \{\beta_1 X_1 + \cdots + \beta_n X_n + \beta_0 = 0\}$, $n+1 \leq l \leq m$ given by a linear function L_{H_l} with $\beta_j \in \mathbb{R}$, $0 \leq j \leq n$. We need to show that $L_{H_l}(E) \geq 0$. Denote by $0 \leq j_0 \leq n$ the index such that $\beta_0 = \cdots = \beta_{j_0-1} = 0$, $\beta_{j_0} \neq 0$. It suffices to show that $\beta_{j_0} > 0$, this would entail that $sgn(L_{H_l}(E)) = sgn(\beta j_0) > 0$. Pick out an arbitrary point $v_{n-j_0} = (x_1^{(n-j_0)}, \ldots, x_{j_0}^{(n-j_0)}, 0, \ldots, 0) \in ((H_n \cap \cdots \cap H_{j_0+1}) \cap S^+) \setminus H_{j_0}$. Therefore, $x_{j_0}^{(n-j_0)} \neq 0$ and $0 \leq L_{H_{j_0}}(v_{n-j_0}) = x_{j_0}^{(n-j_0)}$ since $v_{n-j_0} \in S^+$.

Hence $0 \leq sgn(L_{H_l}(v_{n-j_0})) = sgn(\beta_{j_0} \cdot x_{j_0}^{(n-j_0)})$, therefore $sgn(\beta_{j_0}) > 0$, this implies that $E \in S^+$. Evidently, $E_i^{(+)} \notin S^+$.

Easy counting yields that the probability of T_α's that classify E and at least $(1-2\gamma)^2 n$ many $E_i^{(+)}$'s correctly is greater than $\frac{1-2\gamma}{2-2\gamma}$. Indeed, the probability of the set of all α for which T_α classifies wrong at least $(1-(1-2\gamma)^2)n$ among $E_i^{(+)}, 1 \leq i \leq n$ is less or equal to $\frac{\gamma n}{(1-(1-2\gamma)^2)n} = \frac{1}{4(1-\gamma)}$. Hence the probability of α such that T_α classifies E and at least $(1-2\gamma)^2 n$ among $E_i^{(+)}$, $1 \leq i \leq n$ correctly is greater or equal to $\left(1-\gamma-\frac{1}{4(1-\gamma)}\right) > \frac{1-2\gamma}{2-2\gamma}$.

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 b) for $\bigcap_{i=1}^m H_i^+$.

a) In case of T recognizing $S = \bigcup_{i=1}^m H_i$ consider the points $E_i^{(0)} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, 0, \varepsilon_{i+1}, \ldots, \varepsilon_n) \in S$, $1 \leq i \leq n$ and argue as above, replacing $E_i^{(+)}$ by $E_i^{(0)}$, $1 \leq i \leq n$, and observing that the point E does not lie in the arrangement S. $\square$

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 $\mathrm{Var}^{(v_L)}(T') = \mathrm{Var}^{(v_L)}(f_1, \ldots, f_s)$.

Consider the flag of L

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

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 neighborhood of v_L). For the hyperplane $H_{i_{n-k}}$ we construct a vertex of the tree $\mathcal{T}$ being a son of the root of $\mathcal{T}$, and mark it with this hyperplane $H_{i_{n-k}}$. We label this vertex if and only if Y_{n-k} divides one of $f_1, \dots, f_s$. 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 $f_j^{(1)}(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-1}) = \widetilde{f_j}(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-1}, 0)$, $1 \leq j \leq s$ where the coordinates $Y_1, \dots, Y_{n-k}$ correspond to hyperplanes $H_{i_1}, \dots, H_{i_{n-k}}$, respectively, and $Z_1, \dots, Z_k$ are arbitrary coodinates in L with the origin at v_L. 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 $\ell < n-k$ 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 the 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 $f_j^{(\ell)} = \tilde{f}_j^{(\ell)} Y_{n-k-\ell}^{q_j} \mathcal{L}_{H \bigcap H_{t_1}}^{q_{j,1}} \cdots \mathcal{L}_{H \bigcap H_{t_\pi}}^{q_{j,\pi}}, 1 \leq j \leq s$ for the maximal possible $q_j, q_{j,1}, \dots, q_{j,\pi}$ where $i_{n-k-\ell} < t_1 < \dots < t_\pi$ and $\mathcal{L}_{H \bigcap H_{t_1}}, \dots, \mathcal{L}_{H \bigcap H_{t_\pi}}$ are all the linear polynomials in the plane H determining hyperplanes $H \bigcap H_{t_1}, \dots, H \bigcap H_{t_\pi}$ (in H) which divide $f_j^{(\ell)}$ with the indices $t_1, \dots, t_\pi$ greater than $i_{n-k-\ell}$. We assign to the constructed vertex the polynomials $f_j^{(\ell+1)} = \tilde{f}_j^{(\ell)}(Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-\ell-1}, 0), 1 \leq j \leq s$. 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 $\mathcal{L}_{H \bigcap H_{t_1}}, \dots, \mathcal{L}_{H \bigcap H_{t_\pi}}$ 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 $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-l}$ 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 $Z_1, \dots, Z_k, Y_1, \dots, Y_{n-k-\ell}$). Thus, $lm^{(v_L)} f_j^{(\ell)}$ coincides with $lm^{(v_L)} \left(\tilde{f}_j^{(l)} Y_{n-k-l}^{q_j} \right)$ 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 $f_1, \dots, f_s$, 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

$$\deg\left(\prod_{1 \leq j \leq s} f_j^{(\ell+1)}\right) \leq \deg\left(\prod_{1 \leq j \leq s} f_j^{(\ell)}\right) - (p - q_0 + 1) \qquad (1)$$

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.l.o.g. that $Q \leq R$ (if $Q > R$ then set $M(R,Q,D)=0$). When $R=Q$ we have $M(R,R,D) \leq m \cdot M(R-1,R-1,D)$ and thereby $M(R,R,D) \leq m^R$ by induction on R. When $R>Q$ 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 $M(R,Q,D) \leq m \cdot M(R-1,Q-1,D) + \sum_{1\leq p\leq D} M(R-1,Q,D-p)$ 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 (1) 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

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

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

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

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).

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 (by *Stirling*'s formula) that

$$M \leq 3^t m^{(n-k)(1-c)}(\frac{td}{c(n-k)})^{c(n-k)} \cdot 6^{n-k},$$

which proves Lemma 2.

□

8 Lower bound on the complexity of deterministic decision trees.

For $d-DT$ T' which recognizes either an arrangement S or a polyhedron S^+, we can give the similar complexity lower bound $\Omega(\log N)$ as in the main theorem, where N is the number of k-faces of S or S^+, without the restriction $N \geq m^{\Omega(n-k)}$, imposed in the theorem. This implies the theorem from [GKV97] in case of a polyhedron S^+. For arrangements S, it gives new proofs of the results from [SY82], [B83], without making use of Milnor's bound on the Betti numbers of algebraic varieties.

Theorem 1: *Any $d-DT$ T' recognizing S or S^+ with N k-faces, has the depth t greater than $\frac{1}{2}(\log_3 N - (n-k)\log_3 d)$.*

The proof follows the proof of the Main Theorem with considerable simplifications. Namely, in Lemma 1 one states that

$$Var^{(v;H_1,\ldots,H_{n-k})}(T') = n-k$$

.

In Lemma 2 we have the bound $N \leq 3^t(\frac{td}{n-k})^{n-k}3^{n-k}$ due to the estimation $M(n-k, 0, td) \leq (\frac{td}{n-k})^{n-k}3^{n-k}$.

$\square$

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 Lemma 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 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.

[BG81] C.H. Bennett, J.Gill, Relative to a Random Oracle A, $P^A \neq NP^A \neq co-NP^A$ with Probability 1, SIAM J.Comput. **10** (1981), pp. 96-113.

[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.

[E87] H. Edelsbrunner, Algorithms in Computational Geometry, Springer, 1987.

[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.

[GKMS96] D. Grigoriev, M. Karpinski, F. Meyer auf der Heide, R. Smolensky, A Lower Bound for Randomized Algebraic Decision Trees, Proc. ACM STOC (1996), pp. 612-619.

[GKV97] D. Grigoriev, M. Karpinski, N. Vorobjov, Lower Bound on Testing Membership to a Polyhedron by Algebraic Decision Trees, Discrete Comput. Geom. 17 (1997), pp. 191-215.

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

[L84] S. Lang, Algebra, Addison–Wesley, New York, 1984.

[M71] P. McMullen and G. Shephard, Convex Polytopes and the Upper Bound Conjecture, Cambridge University Press, Cambridge (1971).

[M93] S. Meiser, Point Location in Arrangements of Hyperplanes, Information and Computation **106** (1993), pp. 286 – 303.

[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.

[MT85] U. Manber and M. Tompa, Probabilistic, Nondeterministic and Alternating Decision Trees, J. ACM, Vol. 32 (1985),pp. 720-732.

[S85] M. Snir, Lower Bounds for Probabilistic Linear Decision Trees, Theor. Comput. Sci., Vol. 38 (1985), pp. 69-82.

[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.

[Y81] A. Yao, A Lower Bound to Finding Convex Hulls, J. ACM **28** (1981), pp. 780–787.

[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.