Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 161 (2016)

Robust sensitivity

Shachar Lovett*
University of California, San Diego
slovett@cs.ucsd.edu

Avishay Tal†
Institute for Advanced Study, Princeton
avishay.tal@gmail.com

Jiapeng Zhang‡
University of California, San Diego
jpeng.zhang@gmail.com

January 8, 2017

Abstract

The sensitivity conjecture is one of the central open problems in Boolean complexity. A recent work of Gopalan et al. [CCC 2016] conjectured a robust analog of the sensitivity conjecture, which relates the decay of the Fourier mass of a Boolean function to moments of its sensitivity. We prove this robust analog in this work.

*Research supported by NSF CAREER award 1350481, NSF CCF award 1614023 and a Sloan fellowship.
†Research supported by the Simons Collaboration on Algorithms and Geometry, and by NSF CCF award 1412958.
‡Research supported by NSF CAREER award 1350481.

ISSN 1433-8092

1 Introduction

The sensitivity conjecture is a central open problem in Boolean complexity theory. Let $f : \{0,1\}^n \to \{-1,1\}$ be a Boolean function. The sensitivity of f at $x \in \{0,1\}^n$, denoted $s(f,x)$, is the number of neighbours of x in the Boolean hypercube where f takes the opposite value. That is, it is the number of indices $i \in [n]$ such that $f(x \oplus e_i) \neq f(x)$. The sensitivity of f is defined as the maximum sensitivity of a vertex, $s_{\max}(f) = \max_{x \in \{0,1\}^n} s(f,x)$.

The sensitivity conjecture speculates that functions of low sensitivity must be “simple”. This can be phrased in several equivalent formulations. For our purposes, we will express this by the Fourier degree of f (see also [8] for other notions in which low sensitivity functions are simple). We say that f has Fourier degree d if $\widehat{f}(S) = 0$ for all sets $|S| > d$. Equivalently, f can be computed by a real polynomial of degree d.

Conjecture 1.1 (Sensitivity conjecture)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a Boolean function. If the sensitivity of f is $s = s_{max}(f)$ then the Fourier degree of f is at most poly(s).*

Despite much research [1–6, 8–17], the sensitivity conjecture remains wide open, where the best upper bounds on the degree are exponential in the sensitivity, and the best separations are quadratic. The survey [10] provides a good account of the conjecture, many of its equivalent formulations and consequences, and the progress so far.

A recent work of Gopalan, Servedio, Tal and Wigderson [9] suggested relaxing Conjecture 1.1, and instead of showing that all of the Fourier mass of f appears in low levels, show that for most of the Fourier mass. Their main result gives such tight bounds, under the assumption of low maximal sensitivity.

Theorem 1.2 (Theorem 1.2 in [8])**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ be a Boolean function. If the sensitivity of f is $s = s_{max}(f)$ then for every $d \geq 1$,*

$$\sum_{S \subseteq [n], |S| \geq d} \widehat{f}(S)^2 \leq 2^{-\Omega(d/s)}.$$

Gopalan et al. [9] conjectured a stronger variant of Theorem 1.2 may be true. They consider two distributions over integers $0, \ldots, n$:

1. The Fourier distribution of f, where one chooses a Fourier coefficient S with probability $\widehat{f}(S)^2$ and computes its degree $|S|$.
2. The sensitivity distribution of f, where one chooses a random point $x \in \{0,1\}^n$ and computes its sensitivity $s(f,x)$.

They conjectured that moments of the sensitivity distribution bound the respective moments of the Fourier distribution. In other words, they conjectured that if **most** inputs to a Boolean function f have low sensitivity, then **most** of the Fourier mass of f is concentrated on low levels.

Conjecture 1.3 (Conjecture 1.3 in [8])**.** *For every $d \geq 1$ there exists a constant a_d such that the following holds. For any $n \geq 1$ and any Boolean function $f : \{0,1\}^n \to \{-1,1\}$ it holds that*

$$\sum_{S \subseteq [n]} \widehat{f}(S)^2 |S|^d \leq a_d \cdot \mathbb{E}_{x \in \{0,1\}^n}[s(f,x)^d].$$

It is easy to verify that Conjecture 1.3 with a good enough constant a_d (concretely, $a_d = d^d 2^{O(d)}$) implies Theorem 1.2, even if we replace the assumption that the maximum sensitivity of f is at most s, with the weaker assumption that the d-th moment of the sensitivity is at most s^d. In this work we prove this conjecture. The following is our main theorem, which is a slight re-formulation of Conjecture 1.3.

Theorem 1.4 (Main theorem). *Let $f : \{0,1\}^n \to \{-1,1\}$. For any $d \geq 1$ it holds that*

$$\sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d} \leq a_d \cdot \mathbb{E}_{x \in \{0,1\}^n} \left[s(f,x)^d \right]$$

where $a_d \leq 2^d \cdot d^{12d}$.

We conjecture that the bound on a_d can be improved to $a_d \leq 2^{O(d)}$. If so, this will imply the strongest quantitative form of Conjecture 1.3, and in particular imply Theorem 1.2 under the weaker assumptions on the moments of the sensitivity, as discussed above.

Organization. We present some basic preliminary definitions in Section 2. Our starting point, described in Section 3, is a tight relation between the d-th moments of the Fourier distribution, and the number of d-dimensional sub-cubes of $\{0,1\}^n$ for which the restriction of f has maximal degree d. This relation was also utilized (in a somewhat different form) in [9]. Given this relation, we proceed in Section 4 to bound the number of such sub-cubes, where we build upon and extend the arguments of [9]. We first present a simplified bound of $a_d \leq 2^{O(d^2)}$ in this section (see Theorem 4.8), and then proceed in Section 5 to give the improved bound of $a_d \leq d^{O(d)}$, which yields Theorem 1.4. We discuss open problems in Section 6.

2 Preliminaries

Boolean hypercube. We denote $[n] = \{1, \ldots, n\}$. For $I \subseteq [n]$ we denote by $e_I \in \{0,1\}^n$ the indicator vector for I. For $i \in [n]$ we shorthand $e_i = e_{\{i\}}$. We denote by $\mathcal{H}_n$ the n-dimensional hypercube, whose vertices are $V(\mathcal{H}_n) = \{0,1\}^n$ and edges are $E(\mathcal{H}_n) = \{(x, x \oplus e_i) : x \in \{0,1\}^n, i \in [n]\}$. Given two vectors $x, y \in \{0,1\}^n$, we shorthand $x + y$ for $x \oplus y$ whenever the context is clear. In particular, edges of the hypercube are written as $(x, x + e_i)$. We say that an edge $(x, x + e_i)$ has direction i.

Fourier analysis. The Fourier decomposition of $f : \{0,1\}^n \to \mathbb{R}$ is

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S)(-1)^{\langle x, e_S \rangle}.$$

3 Fourier moments and max degree cubes

Definition 3.1 (Sub-cubes). *For $v \in \{0,1\}^n$ and $I \subset [n]$ let*

$$C(v, I) := \{x \in \{0,1\}^n : x_i = v_i \ \forall i \notin I\}$$

denote a sub-cube. The dimension of the sub-cube is $|I|$. Note that $C(v, I) = C(v', I)$ for all $v' \in C(v, I)$. We denote by $\mathcal{C}(n, d)$ the set of all d-dimensional cubes in $\{0,1\}^n$.

Given $C = C(v, I) \in \mathcal{C}(n, d)$, the restriction of $f : \{0,1\}^n \to \{-1,1\}$ to C is $f|_C : \{0,1\}^I \to \{-1,1\}$ given by $f|_C(x) = f(y)$ where $y_i = x_i$ for $i \in I$ and $y_i = v_i$ for $i \notin I$. We say that $f|_C$ has *max degree* if its degree as a multilinear real polynomial over $\{x_i : i \in I\}$ is maximal, namely d. This is equivalent to $\widehat{f|_C}(I) \neq 0$.

The following lemma connects the Fourier moments of f and the number of maximal degree cubes in f. It appears in a slightly different formulation as Theorem 3.2 in [9]. For completeness, we give a proof of Lemma 3.2 in the appendix.

Lemma 3.2. *Let $f : \{0,1\}^n \to \{-1,1\}$. Fix $d \geq 1$. Define*

$$A := 2^n \sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d}$$

and

$$B := |\{C \in \mathcal{C}(n,d) : f|_C \text{ has max degree}\}|.$$

Then

$$2^{-d} B \leq A \leq 2^d B.$$

4 Bounding the number of max degree cubes

Let $f : \{0,1\}^n \to \{-1,1\}$. Given Lemma 3.2, we focus on bounding the number of d-dimensional cubes C such that $f|_C$ has max degree.

4.1 Sensitivity graph and related notions

The following definitions are from [9].

Definition 4.1 (Sensitivity graph)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$. The sensitivity graph G_f of f is the sub-graph of $\mathcal{H}_n$ whose edges are*

$$E(G_f) := \{(x, x+e_i) : x \in \{0,1\}^n, i \in [n], f(x) \neq f(x+e_i)\}.$$

Edges of G_f are called "sensitive edges" of $\mathcal{H}_n$ with respect to f.

Definition 4.2 (Sensitive nodes / edges / directions of walks in the hypercube)**.** *Let P be a walk (i.e. a path) in $\mathcal{H}_n$, whose vertices are $v_0, v_1, \ldots, v_m \in \{0,1\}^n$. Let $i_1, \ldots, i_m \in [n]$ be the directions of the edges of P, namely $v_i = v_{i-1} + e_i$. An edge (v_j, v_{j+1}) is said to be a* sensitive edge *of the walk if there is no $j' < j$ for which $i_{j'} = i_j$. Namely, the edge (v_j, v_{j+1}) is the first edge in the walk in direction i_j. In such a case, we also say that v_j is a* sensitive node*. We further define:*

- *Sensitive nodes of P: $V(P) = (v_{j_1}, \ldots, v_{j_d})$.*
- *Sensitive directions of P: $I(P) = (i_{j_1}, \ldots, i_{j_d})$.*
- *Dimension of P: $\dim(P) = |V(P)| = |I(P)|$.*

Definition 4.3 (Walk sensitive for a function)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$. A walk P in $\mathcal{H}_n$ is sensitive for f if the sensitive edges of P are also sensitive edges for f.*

Definition 4.4 (Proper walk)**.** *Let $f : \{0,1\}^n \to \{-1,1\}$ and $1 \leq d \leq n$. A proper walk P with respect to f, of dimension d, is given by:*

- *Its sensitive nodes* $V(P) = (v_1, \ldots, v_d)$*, where* $v_1, \ldots, v_d \in \{0,1\}^n$*.*
- *Its sensitive directions* $I(P) = (i_1, \ldots, i_d)$*, where* $i_1, \ldots, i_d \in [n]$ *are distinct.*

Such that they satisfy:

- $f(v_j) \neq f(v_j + e_{i_j})$ *for* $j = 1, \ldots, d$*.*
- $v_j \in C(v_1, \{i_1, \ldots, i_{j-1}\})$ *for* $j = 2, \ldots, d$*.*

A proper walk can be extended to a walk in $\mathcal{H}_n$ *with sensitive nodes* $V(P)$ *and sensitive directions* $I(P)$*, by connecting each* v_j *to* v_{j+1} *using some shortest walk. By definition, this part of the walk will only use edges with directions in* $\{i_1, \ldots, i_j\}$*. The resulting walk is sensitive for* f*.*

Given a proper walk P with $V(P) = (v_1, \ldots, v_d)$ and $I(P) = (i_1, \ldots, i_d)$, we say that it *realizes* the sub-cube $C(P) := C(v_1, I(P))$. Equivalently, $C(P)$ is the minimal sub-cube which contains all the edges $(v_j, v_j + e_{i_j})$.

4.2 Proper walks in maximal degree cubes

Let $f : \{0,1\}^n \to \{-1,1\}$. Gopalan et al. [9] proved that if $f|_C$ has maximal degree, then C is realized by some proper walk (in fact, they prove that there exists such a proper walk with a succinct description, which allows for better quantitative bounds; for now, we ignore this aspect, and re-inspect it in Section 5). We will ask for a proper walk where the first node has maximal sensitivity.

Definition 4.5 (First-maximal proper walk)**.** *Let* P *be a proper walk with respect to* f*, with sensitive nodes* $V(P) = (v_1, \ldots, v_d)$*. We say that* P *is* first-maximal *if* $s(f, v_1) \geq s(f, v_i)$ *for all* $i = 2, \ldots, d$*.*

Lemma 4.6. *Let* $f : \{0,1\}^n \to \{-1,1\}$*,* $C \in \mathcal{C}(n,d)$ *such that* $f|_C$ *has maximal degree* d*. Then* C *is realized by a first-maximal proper walk with respect to* f*.*

Proof. Let $g = f|_C$. For a sensitive edge (x, x') for g, define its weight as $w(x,x') = \max(s(f,x), s(f,x'))$. We will prove that there exists a d-dimensional proper walk P with respect to g, with sensitive nodes $V(P) = (v_1, \ldots, v_d)$ and sensitive directions $I(P) = (i_1, \ldots, i_d)$, such that

$$w(v_1, v_1 + e_{i_1}) \geq w(v_2, v_2 + e_{i_2}) \geq \ldots \geq w(v_d, v_d + e_{i_d}).$$

We first observe that this suffices for the lemma. We may assume that $s(f, v_1) \geq s(f, v_1 + e_{i_1})$, as otherwise we can set the starting point to be $v_1 + e_{i_1}$ without changing any of the properties of the proper walk. Then by design for every $j = 2, \ldots, d$ we have

$$s(f, v_1) = w(v_1, v_1 + e_{i_1}) \geq w(v_j, v_j + e_{i_j}) \geq s(f, v_j).$$

Next, we prove the existence of such a walk by induction on d. For $d = 1$ this is obvious, so assume $d \geq 2$. Let (y, y') be a sensitive edge in G_g with minimal weight $w(y, y')$. Assume that $y' = y + e_\ell$. If g has maximal degree d, then at least one of the restrictions $g|_{x_\ell = 0}$ or $g|_{x_\ell = 1}$ must have maximal degree $d-1$ in their respective sub-cube. Assume without loss of generality that this holds for $g|_{x_\ell=0}$ and that $y_\ell = 0$. By induction there is a proper walk with the required conditions, realizing the sub-cube $\{x : x_\ell = 0\}$ of dimension $d-1$, given by sensitive nodes $v_1, \ldots, v_{d-1}$ and sensitive directions $i_1, \ldots, i_{d-1}$. To complete the walk we set $v_d = y$ and $i_d = \ell$. $\square$

4.3 Putting it together

Let $f : \{0,1\}^n \to \{-1,1\}$. By Lemma 4.6, in any d-dimensional sub-cube C where $f|_C$ has maximal degree, we can find a first-maximal proper walk realizing it. Thus, instead of counting maximal degree sub-cubes, we will count first-maximal proper walks.

Claim 4.7. *The number of d-dimensional first-maximal proper walks in G_f, which start at a given node x, is at most*

$$2^{\binom{d}{2}} s(f,x)^d.$$

Proof. We wish to count d-dimensional proper walks P with respect to f. Let $V(P) = (v_1, \ldots, v_d)$ and $I(P) = (i_1, \ldots, i_d)$. We assume $v_1 = x$, hence there are $s(f,x)$ possible values for i_1. Given that we already defined $v_1, \ldots, v_{j-1}$ and $i_1, \ldots, i_{j-1}$, we have by assumption that $v_j \in C(v_1, \{i_1, \ldots, i_{j-1}\})$, and hence it has at most 2^{j-1} different possibilities. Given a choice of v_j, the number of choices for i_j is at most $s(f, v_j) \le s(f,x)$. Thus we can bound the number of such walks by

$$2^{1+2+\ldots+d-1} \cdot s(f,x)^d = 2^{\binom{d}{2}} s(f,x)^d.$$

□

We now obtain a proof of Theorem 1.4 with a weaker quantitative bound on a_d.

Theorem 4.8. *Let $f : \{0,1\}^n \to \{-1,1\}$. For any $d \ge 1$ it holds that*

$$\sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d} \le a_d \cdot \mathbb{E}_{x \in \{0,1\}^n} \left[s(f,x)^d \right]$$

where $a_d \le 2^{\binom{d}{2}+d}$.

Proof. Let $A = 2^n \sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d}$, $B = |\{C \in \mathcal{C}(n,d) : f|_C \text{ has max degree}\}|$ and $D = \sum_{x \in \{0,1\}^n} s(f,x)^d$. By Lemma 3.2 we have $A \le 2^d B$. By Lemma 4.6 we can bound B by the number of d-dimensional first-maximal proper walks with respect to f, and by Claim 4.7 this number is bounded by $2^{\binom{d}{2}} D$. Thus

$$2^n \sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d} \le 2^d B \le 2^{\binom{d}{2}+d} \sum_{x \in \{0,1\}^n} s(f,x)^d.$$

The theorem follows by dividing both sides by 2^n. □

5 Improving the parameters

The goal in this section is to improve the parameters in Theorem 4.8. A keen reader (or one familiar with previous work [9]) can see that the main reason for the loss of parameters in Theorem 4.8 is the number of potential first-maximal proper walks in a max degree function, which we naively bounded by $2^{\binom{d}{2}}$. In order to obtain a better bound, we need to define more carefully what do we mean by a "description" of a proper walk. This notion was studied implicitly in [9] (see Lemma 5.5), and we define it here explicitly.

Definition 5.1 (Signature of a walk)**.** *Let P be a d-dimensional walk in $\{0,1\}^n$. Let $V(P) = (v_1, \ldots, v_d)$ and $I(P) = (i_1, \ldots, i_d)$. By construction, we have $v_{j+1} \in C(v_1, \{i_1, \ldots, i_j\})$ for all $j = 1, \ldots, d-1$. This means that there exists $r_{i,j} \in \{0,1\}$ such that*

$$v_{j+1} = v_1 + r_{j,1} \cdot e_{i_1} + \ldots + r_{j,j} \cdot e_{i_j}.$$

The signature *of P is*

$$R(P) = (r_{i,j} : 1 \le i \le j \le d-1) \in \{0,1\}^{\binom{d}{2}}.$$

We next define when a family of walks has a succinct description.

Definition 5.2 (Signature of a family of paths)**.** *Let $\mathcal{P}$ be a family of walks in $\{0,1\}^d$. The signatures of $\mathcal{P}$ are*

$$R(\mathcal{P}) = \{R(P) : P \in \mathcal{P}\} \subset \{0,1\}^{\binom{d}{2}}.$$

If $|R(\mathcal{P})| \le 2^b$ then we say that $\mathcal{P}$ can be described using b bits.

We also need to extend the notion of first-maximal proper walks, in a way that breaks the relation between the sub-cube and the global sensitivity of the function on $\mathcal{H}_n$.

Definition 5.3 (first-maximal with respect to a weight function)**.** *Let P be a d-dimensional walk whose sensitive nodes are $V(P) = (v_1, \ldots, v_d)$. Let $w : \{0,1\}^d \to \mathbb{R}$ be some weight function on the nodes of the hypercube. We say that P is* first-maximal with respect to w *if $w(v_1) \ge w(v_i)$ for all $i = 2, \ldots, d$.*

In the applications we will use $g = f|_C$ with weight function $w(x) = s(f,x)$. However, making the general definition allows to focus on the restricted function $f|_C$ and forget about the function f. The following definition isolates our notion of "efficient description" of a first-maximal proper walks.

Definition 5.4 (Description of first-maximal proper walks)**.** *Fix $d \ge 1$. We say that first-maximal proper walks in d dimensions can be described using b bits if the following holds. For any function $g : \{0,1\}^d \to \{-1,1\}$ of maximal degree d, and any weight function $w : \{0,1\}^d \to \mathbb{R}$, there exist a d-dimensional walk $P_{g,w}$ which is proper with respect to g, and first-maximal with respect to w, such that the family*

$$\mathcal{P}_{proper,\ first-maximal} := \{P_{g,w}\}$$

can be described using b bits.

One can verify that Lemma 4.6 can be extended to an arbitrary weight function. Thus, it establishes that first-maximal proper walks in d dimensions can be described using $\binom{d}{2}$ bits. This motivates the question of looking for the minimal such description length. This is further motivated by the following lemma.

Lemma 5.5. *Assume that first-maximal proper walks in dimension d can be described using b bits. Then Theorem 1.4 holds with the bound $a_d = 2^{d+b}$.*

Proof. Let $\mathcal{R} \subset \{0,1\}^{\binom{d}{2}}$ be a set of size $|\mathcal{R}| \le 2^b$, such that for any function $g : \{0,1\}^d \to \{-1,1\}$, and any weight function $w : \{0,1\}^d \to \mathbb{R}$, there exists a d-dimensional walk $P_{g,w}$ which is proper with respect to g, and first-maximal with respect to w, such that $R(P_{g,w}) \in \mathcal{R}$.

The only change needed in the proof of Theorem 1.4 is in Claim 4.7, where instead of allowing for an arbitrary first-maximal proper walk, we only allow for walks P for which $R(P) \in \mathcal{R}$. Thus the number of first-maximal proper walks starting at node x can be bounded by $s(f,x)^d|\mathcal{R}|$ and the rest of the proof remains as is. □

Gopalan et al. [9] proved (Lemma 5.5) that if we remove the requirement that the walk is first-maximal, then proper walks can be described using $4d$ bits. However, their proof does not give the first-maximal condition, which is why their proof only works assuming a bound on the maximal sensitivity of f. We conjecture that such a bound can be obtained also with the first-maximal condition.

Conjecture 5.6. *For any $d \geq 1$, first-maximal proper walks in dimension d can be described using $O(d)$ bits.*

Conjecture 5.6 would give optimal bounds in Theorem 1.4. Below, we give a nearly tight bound.

Theorem 5.7. *For any $d \geq 1$, first-maximal proper walks in dimension d can be described using $12d \log d$ bits.*

Theorem 1.4 follows immediately from Theorem 5.7 and Lemma 5.5. Below, we give the details necessary to prove Theorem 5.7. We start with some more definitions from [9].

5.1 Sensitive trees

Let $g : \{0,1\}^d \to \{-1,1\}$. Its corresponding sensitivity graph is G_g. We will generally assume that g has max degree, although the following statements also follow from a weaker assumption that g has maximal decision tree depth d.

Definition 5.8 (Sensitive tree). *Let $g : \{0,1\}^d \to \{-1,1\}$. A sensitive tree for g is a sub-tree T of G_g such that all edges of T have distinct directions. We denote by $V(T)$ the nodes of T, by $I(T)$ the directions of the edges of T, and by $C(T)$ the minimal sub-cube that contains T.*

The following claim is Lemma 5.3 in [9], which shows how to get a proper walk from a sensitive tree. It also shows that such walks can be succinctly described.

Claim 5.9 (Proper walk from a sensitive tree). *Let $g : \{0,1\}^d \to \{-1,1\}$, and let T be a sensitive tree for g. Then for every $v \in V(T)$ there exists a proper walk $P = P_{tree}(v;T)$ with respect to g, such that v is the first node in P, $V(P) \subseteq V(T)$ and $I(P) = I(T)$. Furthermore, let*

$$\mathcal{P}_{tree} := \{P_{tree}(v;T) : g : \{0,1\}^d \to \{-1,1\},\ T \text{ sensitive tree for } g,\ v \in V(T)\}.$$

Then $\mathcal{P}_{tree}$ can be described using $2d$ bits.

Proof. Given a sensitive tree T with respect to g, consider the walk obtained by performing a depth first search on T starting at v. This gives the required proper walk. To analyze the signatures of $\mathcal{P}_{tree}$, note that if T is a tree with k edges, then a depth first search in T is a path of length $2k$ which can be described as a sequence of length $2k$ with two types of operations: "follow next sensitive edge" or "backtrack". Moreover, there are exactly k of each type. This determines the signature of the walk. Thus the total number of different signatures in $\mathcal{P}_{tree}$ is at most

$$|R(\mathcal{P}_{tree})| \leq \sum_{k=1}^{d} \binom{2k}{k} \leq 2^{2d}. \qquad \square$$

Definition 5.10 (Shifting a sensitive tree). *Let $g : \{0,1\}^d \to \{-1,1\}$ and let T be a sensitive tree for g. We say that T can be shifted in direction $J \subseteq [d]$, where $J \cap I(T) = \emptyset$, if $f(x) = f(x + e_J)$ for all nodes x of T. In such a case, we denote by $T + e_J$ the tree obtained by shifting all nodes and edges of T by e_J. Observe that $T + e_J$ is also a sensitive tree for g.*

Definition 5.11 (Sensitive tree invariant to shifts). *Let $g : \{0,1\}^d \to \{-1,1\}$ and let T be a sensitive tree for g. Let $I \subset [d]$ disjoint from $I(T)$. We say that T is invariant to shifts supported on directions I, if for any $J \subseteq I$ we can shift T in direction J. Equivalently, if $f(x) = f(x + e_J)$ for all $x \in V(T)$ and all $J \subseteq I$. In the case that $I = [d] \setminus I(T)$ we say that T is* maximally invariant to shifts.

The following claim is essentially Lemma 4.6 in [9].

Claim 5.12. *Let $g : \{0,1\}^d \to \{-1,1\}$. Let T be a sensitive tree with respect to g. Let $I \subset [d]$ disjoint from $I(T)$. Then there exists $I' \subseteq I$, and a sensitive tree T' with respect to g, such that the following holds:*

- *$I(T') = I(T) \cup I'$.*
- *T' is invariant to shifts supported on directions $I \setminus I'$.*
- *There exists $J \subseteq I'$ such that $T + e_J$ is a sub-tree of T'.*

Proof. We build T' greedily. Set initially $T' = T$ and $I' = \emptyset$. If T' is invariant to shifts supported on $I \setminus I'$, we are done. Otherwise, let J be minimal such that $g(v + e_J) \neq g(v)$ for some $v \in V(T')$. Choose some arbitrary $j \in J$. By assumption T' can be shifted in direction $J \setminus \{j\}$, so set $T' = T' + e_{J\setminus\{j\}}$. Now, there exists some $v \in V(T')$ for which $g(v) \neq g(v + e_j)$. Thus, we can add a new sensitive edge $(v, v + e_j)$ to T', and add j to I'. Repeat this process until it terminates. □

Let $v \in \{0,1\}^d$. We say that a sensitive tree T agrees with v on coordinates $I \subset [n]$, where $I \cap I(T) = \emptyset$, if $v_i = x_i$ for all $x \in C(T)$ and all $i \in I$. Note that if a sensitive tree T is invariant to shifts supported on directions I, then for any v there exists some shift $T' = T + e_J$ for $J \subseteq I$ such that T' agrees with v on I.

5.2 Sensitive tree chains

Definition 5.13 (Sensitive tree chain)**.** *Let $g : \{0,1\}^d \to \{-1,1\}$. A sequence of sensitive trees $T_1, \dots, T_m$ with respect to g is called a* sensitive tree chain *if for each $i = 2, \dots, m$, $V(T_i) \cap C(T_{i-1})$ is nonempty. We define $V(T_1, \dots, T_m) := V(T_1) \cup \dots \cup V(T_m)$ and $I(T_1, \dots, T_m) := I(T_1) \cup \dots \cup I(T_m)$.*

Note that if $T_1, \dots, T_m$ is a sensitive tree chain with respect to g, then so is any sub-sequence. Namely, for any $i \leq j$ we have that $T_i, \dots, T_j$ is also a sensitive tree chain with respect to g.

Claim 5.14 (Proper walk from a sensitive tree chain)**.** *Let $g : \{0,1\}^d \to \{-1,1\}$. Let $T_1, \dots, T_m$ be a sensitive tree chain for g. For every $v \in V(T_1)$ there exists a proper walk $P = P_{chain}(v; T_1, \dots, T_m)$ with respect to g, such that v is the first node in P, $V(P) \subseteq V(T_1, \dots, T_m)$ and $I(P) = I(T_1, \dots, T_m)$.*

Proof. Let $v_1 = v$, and for $i > 1$ fix some $v_i \in V(T_i) \cap C(T_{i-1})$. Consider the following path: start with a tree path $P_{tree}(v_1; T_1)$, which traverses T_1 and starts and ends with v_1. Then choose a shortest path from v_1 to v_2, which by assumption only uses directions in $I(T_1)$. Proceed with a tree path $P_{tree}(v_2; T_2)$, which traverses T_2 and starts and ends with v_2. Then choose a shortest path from v_2 to v_3, which by assumption only uses directions in $I(T_2)$. Iterate this procedure until we cover all trees. □

Definition 5.15 (Disjoint sensitive tree chain)**.** *Let $g : \{0,1\}^d \to \{-1,1\}$. Let $T_1, \dots, T_m$ be a sensitive tree chain with respect to g. It is said to be* disjoint *if $I(T_1), \dots, I(T_m)$ are pairwise disjoint.*

Gopalan et al. [9] proved that for any function of maximal degree, there exists a disjoint sensitive tree chain which cover all directions.

Lemma 5.16 (Lemma 5.2 in [9])**.** *Let $g : \{0,1\}^d \to \{-1,1\}$ of maximal degree. There exists a disjoint sensitive tree chain $T_1, \dots, T_m$ with respect to g, such that $I(T_1, \dots, T_m) = [d]$.*

Gopalan et al. [9] also showed that for these disjoint sensitive tree chains, their corresponding proper walks can be descried using $4d$ bits.

Lemma 5.17 (Lemma 5.5 in [9]). *Define*

$$\mathcal{P}_{disjoint} := \{P(v; T_1, \ldots, T_m) \mid g : \{0,1\}^d \to \{-1,1\},\ v \in V(T_1),\\ T_1, \ldots, T_m \text{ disjoint sensitive tree chain for } g\}.$$

Then $\mathcal{P}_{disjoint}$ can be described using $4d$ bits.

Proof sketch. We show that $O(d)$ bits are enough, where with some optimizations this can be made $4d$. Let $d_i = \dim(T_i)$ where by the disjointness assumption $\sum d_i \leq d$. Fix $v_1 \in V(T_1)$ and $v_i \in V(T_i) \cap C(T_{i-1})$. Each path $P_{tree}(v_i; T_i)$ can be encoded using $2d_i$ bits, as we saw in Claim 5.9. The shift from $v_i \in V(T_i)$ to $v_{i+1} \in V(T_{i+1})$ can be encoded using additional d_i bits. In addition, we need symbols to denote when a description of a tree starts and ends, and when the description of a shift starts and ends. Each of these is repeated at most d times. □

The main problem with tree chains $T_1, \ldots, T_m$ is that they allow to "move" only in one direction, that is following the sequence $T_1, T_2, \ldots, T_m$, but not in the reverse direction. In the next section, we introduce reversible tree chains, which allow to move in both directions. These will turn out to be crucial for the purpose of designing first-maximal proper walks.

5.3 Reversible tree chains

Given trees $T_1, \ldots, T_m$, we define by $C(T_1, \ldots, T_m)$ the smallest sub-cube that contains all their edges. The following definition is a weak form of a sensitive tree chain, that will be important for us.

Definition 5.18 (Weak sensitive tree chain). *Let $g : \{0,1\}^d \to \{-1,1\}$. A sequence of sensitive trees $T_1, \ldots, T_m$ with respect to g is called a* weak sensitive tree chain *if for each $i = 2, \ldots, m$, $V(T_i) \cap C(T_1, \ldots, T_{i-1})$ is nonempty (as opposed to $V(T_i) \cap C(T_{i-1}) \neq \emptyset$ in Definition 5.13).*

Claim 5.19 (Proper walk from a weak sensitive tree chain). *Let $g : \{0,1\}^d \to \{-1,1\}$. Let $T_1, \ldots, T_m$ be a weak sensitive tree chain for g. For every $v \in V(T_1)$ there exists a proper walk $P = P_{weak-chain}(v; T_1, \ldots, T_m)$ with respect to g, such that v is the first node in P, $V(P) \subseteq V(T_1, \ldots, T_m)$ and $I(P) = I(T_1, \ldots, T_m)$.*

Proof. The proof is identical to that of Claim 5.14, except that after the traversal on T_i we may change coordinates in $I(T_1) \cup \ldots \cup I(T_i)$ to get to T_{i+1} (as opposed to just changing the coordinates in $I(T_i)$, as done in Claim 5.14). □

Definition 5.20 (Reversible sensitive tree chain). *Let $g : \{0,1\}^d \to \{-1,1\}$. A reversible sensitive tree chain for g is comprised of:*

- *A disjoint sensitive tree chain $T_1, \ldots, T_m$ where $I(T_1, \ldots, T_m) = [d]$.*
- *A weak sensitive tree chain $T'_m, \ldots, T'_1$ (in this order!) where $I(T'_m, \ldots, T'_1) = [d]$.*

Such that

- *Each T_i is a sub-tree of T'_i*
- *The sets $I(T'_i) \setminus I(T_i)$ for $i = 1, \ldots, m$ are pairwise disjoint.*

Reversible sensitive tree chains allow us to construct first-maximal walks, as they support proper walks which start at any node of $T'_1, \ldots, T'_m$.

Claim 5.21. *Let $g : \{0,1\}^d \to \{-1,1\}$. Assume that there exists a reversible sensitive tree chain $(T_1, \ldots, T_m; T'_m, \ldots, T'_1)$ for g. Then, for any weight function $w : \{0,1\}^d \to \mathbb{R}$, there exists a walk $P = P_{g,w}$ which is proper with respect to g, and first-maximal with respect to w. In addition:*

- $V(P) \subset V(T'_1, \ldots, T'_m)$.
- $I(P) = [d]$.
- *The length of P is at most $12d$.*

Proof. Let $V = V(T'_1, \ldots, T'_m)$. Let $v \in V$ for which $w(v)$ is maximal. We will construct a path P as above starting at v. Assume that $v \in V(T'_j)$. The path P is composed of:

- The shortest path in the tree T'_j from v to some $v' \in V(T_j)$.
- The path $P_{chain}(v'; T_j, \ldots, T_m)$, ending at some $v'' \in V(T_m)$.
- The path $P_{weak-chain}(v''; T'_m, \ldots, T'_1)$.

The first two claims clearly hold. We next bound the length of P.

In order to bound the length of walk, the first part has length at most $\dim(T'_i) \le d$. The second path has length bounded by $\sum_{i=j}^{m} 3\dim(T_i) \le 3d$, which follows as we assume that $I(T_1), \ldots, I(T_m)$ are disjoint, and that $C(T_i) \cap V(T_{i+1}) \neq \emptyset$ for $i = 1, \ldots, m-1$. The length of the third part can be bounded as follows.

The path in the third part $P_{weak-chain}(v''; T'_m, \ldots, T'_1)$ selects vertices $v'_{m-1}, \ldots, v'_1$ in $V(T'_{m-1}), \ldots, V(T'_1)$, respectively, such that for all $i = m-1, \ldots, 1$ we have $v'_i \in C(T'_{i+1}, \ldots, T'_m)$. The path starts at $v'_m := v'' \in V(T'_m)$, and explores T'_m using $P_{tree}(v'_m; T'_m)$ that starts and ends at v'_m. We then take the shortest walk in $\mathcal{H}_d$ from v'_m to v'_{m-1}(we explain why the walk is proper below). From v'_{m-1} explore T'_{m-1} using $P_{tree}(v'_{m-1}; T'_{m-1})$, and then take the shortest walk in $\mathcal{H}_d$ from v'_{m-1} to v'_{m-2}. We continue this way until we reach v'_1, where we explore T'_1 using $P_{tree}(v'_1; T'_1)$.

First, we argue that the walk is proper. Recall that when moving from v'_{i+1} to v'_i, for $i = m-1, \ldots, 1$, we take the shortest path in $\mathcal{H}_d$ between the two vertices. Since $v'_i \in C(v'_{i+1}, I(T'_{i+1}, \ldots, T'_m))$, we only change coordinates in $I(T'_{i+1}, \ldots, T'_m)$ which means that the walk is indeed proper.

Next, we wish to bound the length of the shortest path from v'_{i+1} to v'_i, i.e., the distance between v'_i and v'_{i+1} in $\mathcal{H}_d$. Denote by $d_H(u,v)$ the distance between two nodes u and v in $\mathcal{H}_d$ (i.e., their Hamming distance). To bound $d_H(v'_i, v'_{i+1})$, we use the fact that $T_1, \ldots, T_m$ is a disjoint sensitive tree chain (that is, we are using the forward chain to bound the length of the backward walk!). Since $T_1, \ldots, T_m$ is a disjoint sensitive tree chain, there exist $v_i \in V(T_i)$ and $v_{i+1} \in V(T_{i+1})$ with distance at most $\dim(T_i)$ between them (simply take $v_{i+1} \in V(T_{i+1}) \cap C(T_i)$ and any $v_i \in V(T_i)$). By the triangle inequality,

$$d_H(v'_i, v'_{i+1}) \le d_H(v'_i, v_i) + d_H(v_i, v_{i+1}) + d_H(v_{i+1}, v'_{i+1}) \le \dim(T'_i) + \dim(T_i) + \dim(T'_{i+1}),$$

where we used the fact that $v'_i, v_i \in V(T'_i)$ to bound the first summand and that $v_{i+1}, v'_{i+1} \in V(T'_{i+1})$ to bound the third. Thus, the total length of the third part is at most

$$\sum_{j=1}^{m} \dim(T_j) + \sum_{j=1}^{m} 4\dim(T'_j) \le 9d,$$

where $\sum_{j=1}^m \dim(T_j) = d$ and where $\sum_{j=1}^m \dim(T'_j) \le 2d$ by our assumption that $I(T'_1) \setminus I(T_1), \dots, I(T'_m) \setminus I(T_m)$ are disjoint. Thus we can bound the length of the total walk by $12d$. $\square$

The following lemma shows how, starting from a sensitive tree chain, we can construct a reversible sensitive tree chain.

Lemma 5.22. *Let $g : \{0,1\}^d \to \{-1,1\}$. Assume that there exists a disjoint sensitive tree chain $T_1, \dots, T_m$ for g such that $I(T_1, \dots, T_m) = [d]$. Then there exists a reversible sensitive tree chain for g.*

Before proving Lemma 5.22 we need an extension of Claim 5.12 to a weak sensitive tree chain.

Claim 5.23. *Let $g : \{0,1\}^d \to \{-1,1\}$. Let $T_1, \dots, T_m$ be a weak sensitive tree chain with respect to g. Let $I \subset [d]$ disjoint from $I(T_1, \dots, T_m)$. Then there exists $I' \subseteq I$, and a weak sensitive tree chain $T'_1, \dots, T'_m$ with respect to g, such that the following hold:*

- *$I(T'_i) = I(T_i) \cup I'_i$, where $I'_1, \dots, I'_m$ is a partition of I'.*
- *For all $i = 1, \dots, m$, T'_i is invariant to shifts supported on directions $I \setminus I'$.*
- *There exists $J \subseteq I'$ such that for all i, $T_i + e_J$ is a sub-tree of T'_i.*

Proof. The proof is nearly identical to that of Claim 5.12. Let initially $T'_i = T_i, I' = \emptyset$. If all of $T'_1, \dots, T'_m$ are invariant to shifts supported on directions $I \setminus I'$, we are done. Otherwise, pick minimal $J \subset I \setminus I'$ for which some T'_i cannot be shifted in direction J, and pick $j \in J$. Replace each T'_i with $T'_i + e_{J \setminus \{j\}}$, and observe that $T'_1, \dots, T'_m$ is still a weak sensitive tree chain with respect to g. Choose $v \in V(T'_i)$ such that $g(v) \ne g(v + e_j)$, add the edge $(v, v + e_j)$ to T'_i, and add j to I'. Repeat this process until it terminates. $\square$

Proof of Lemma 5.22. Let $T_1, \dots, T_m$ be the initial disjoint sensitive tree chain. Throughout the proof, we will modify $T_1, \dots, T_m$ by the following operations: for some $i \in [m]$ we will choose $J \subseteq I(T_i)$, and replace $T_{i+1}, \dots, T_m$ with $T_{i+1} + e_J, \dots, T_m + e_J$, while assuring that the latter are also sensitive trees for g. Observe that such operations maintain the property that $T_1, \dots, T_m$ is a disjoint sensitive tree chain, and that they do not change $I(T_j)$ for any j.

We construct $T'_m, \dots, T'_1$ in this order. In the i-th iteration (where $i = m, \dots, 1$), we will construct T'_i, and along the way also change $T'_{i+1}, \dots, T'_m$ and $T_{i+1}, \dots, T_m$. We will obtain the following invariant at the end of the i-th iteration (and the beginning of the $i-1$ iteration):

- T_j is a sub-tree of T'_j for all $j = i, \dots, m$.
- $T_1, \dots, T_m$ is a disjoint sensitive tree chain.
- $T'_m, \dots, T'_i$ is a weak sensitive tree chain.
- $C(T'_m, \dots, T'_i) = C(T_i, \dots, T_m)$.

The first iteration, for $i = m$, is very simple: take $T'_m = T_m$. At the beginning of the i-th iteration, for $i < m$, we have already constructed $T'_m, \dots, T'_{i+1}$ that satisfy the requirements above. Apply Claim 5.23 to the weak sensitive tree chain $T'_m, \dots, T'_{i+1}$ with $I = I(T_i)$ (which by induction is disjoint from $I(T_m, \dots, T_{i-1}) = I(T'_m, \dots, T'_{i-1})$). This results in a weak sensitive tree chain $T''_m, \dots, T''_{i+1}$ and a set $I' = I(T_i) \cap I(T''_{i+1}, \dots, T''_m)$ such that

- There exists $J' \subseteq I'$ such $T'_j + e_{J'}$ is a sub-tree of T''_j for all $j = i+1, \dots, m$.

- The directions $I(T''_{i+1}) \cap I(T_i), \ldots, I(T''_m) \cap I(T_i)$ are disjoint and partition I'.
- For all $j = i+1, \ldots, m$, T''_j is invariant to shifts supported on directions $I \setminus I'$.

Next, choose some $v_i \in V(T_i)$. Let $I'' = I \setminus I'$. Let $J'' \subseteq I''$ be a shift so that $T''_{i+1} + e_{J''}$ will agree with v_i on the coordinates I''. Define $T'''_j = T''_j + e_{J''}$ for $j = i+1, \ldots, m$. Note that for $J = J' \cup J''$ we have that $T_j + e_J$ is a sub-tree of T'''_j. Perform the following operations:

- Set $T'_j = T'''_j$ for $j = i+1, \ldots, m$.
- Set $T_j = T_j + e_J$ for $j = i+1, \ldots, m$.
- Set $T'_i = T_i$.

We claim that this satisfies the required conditions for the end of the i-iteration.

First, we have that T_j is a sub-tree of T'_j for $j = i, \ldots, m$. Second, $T_1, \ldots, T_m$ is still a disjoint sensitive tree chain, as we shifted $T_{i+1}, \ldots, T_m$ by some $J \subseteq I(T_i)$. Next, we need to show that $T'_m, \ldots, T'_i$ is a weak sensitive tree chain.

Recall that by definition that means that $C(T'_m, \ldots, T'_{j+1}) \cap V(T'_j) \neq \emptyset$ for all $j = m-1, \ldots, i$. First, we claim that this holds for $j = m-1, \ldots, i+1$. This is true since it held at the beginning of the i-th iteration, and the only change is that we shifted all trees $T'_{i+1}, \ldots, T'_m$ by the same shift e_J, and potentially replaced them by larger sensitive trees containing them. So, it also holds at the end of the i-th iteration. Next, we show that for $j = i$.

Recall that we chose the shift J so that for some $v_i \in V(T_i)$, $C(T'_{i+1})$ agrees with v_i on $I'' = I(T_i) \setminus I(T'_m, \ldots, T'_{i+1})$. By the assumption that $C(T'_m, \ldots, T'_{i+1}) = C(T_{i+1}, \ldots, T_m)$ which held at the beginning of the i-th iteration, and since we only shifted and extended $T'_{i+1}, \ldots, T'_m$ by some directions in $I(T_i)$, we have that $T'_{i+1}, \ldots, T'_m \subset C(T_i, \ldots, T_m)$. As each T_j is a sub-tree of T'_j, and as $T'_i = T_i$, this implies that $C(T'_m, \ldots, T'_i) = C(T_i, \ldots, T_m)$. But then $C(T'_{i+1})$ also agrees with v_i on $I(T_1, \ldots, T_{i-1})$. This then implies that $v_i \in C(T'_m, \ldots, T'_{i+1})$. Thus $V(T_i) \cap C(T'_m, \ldots, T'_{i+1})$ is nonempty, as claimed. □

5.4 Completing the proof

We conclude the proof of Theorem 5.7. Let $g : \{0,1\}^d \to \{-1,1\}$ of maximal degree. By Lemma 5.16 there exists a disjoint sensitive tree chain $T_1, \ldots, T_m$ for g, such that $I(T_1, \ldots, T_m) = [d]$. We may thus apply Lemma 5.22, which shows the existence of a reversible sensitive tree chain for g. Claim 5.21 then shows that there exists a proper walk for g of length at most $12d$. To conclude, observe that for any length $\ell \geq 1$, if we define a family of paths in $\mathcal{H}_d$ of length ℓ,

$$\mathcal{P}_{length\ \ell} := \{P \text{ path of length } \ell \text{ in } \mathcal{H}_d\}$$

then $\mathcal{P}$ can be described using $\ell \cdot \log d$ many bits, simply by giving the edges in the path. This shows that first-maximal proper walks in dimension d can be described using $12d \log d$ bits. [1]

6 Open problems

Our main result is Theorem 1.4, which proves Conjecture 1.3. As we discussed in the introduction, we suspect that our quantitative bounds are sub-optimal. The conjectured bound below will allow us to match the results of [9], which assumed a bound on the maximal sensitivity.

[1] The constant 12 is not optimal. We chose to compromise optimizing the constant, in order to make the presentation simpler.

Conjecture 6.1. *Theorem 1.4 holds with a bound of* $a_d \leq 2^{O(d)}$.

The main technical component of our paper is the structure of the sensitivity graph for functions of maximal degree. Our techniques, however, apply equally well under the weaker assumption that the decision tree complexity of the function is maximal. In fact, any complexity measure where if f has maximal complexity then, for any bit x_i, one of the restrictions $f|_{x_i=0}$ or $f|_{x_i=1}$ has maximal complexity would do. We note that the same phenomena holds for the results and techniques of [9]. Thus, this invites the question of developing techniques which directly rely on the degree; in particular, as there are examples (see e.g. [7]) where the degree and decision tree depth are quadratically separated.

References

[1] A. Ambainis, K. Prūsis, and J. Vihrovs. Sensitivity versus certificate complexity of boolean functions. In *International Computer Science Symposium in Russia*, pages 16–28. Springer, 2016.

[2] N. Balaji, S. Datta, R. Kulkarni, and S. Podder. Graph properties in node-query setting: effect of breaking symmetry. *arXiv preprint arXiv:1510.08267*, 2015.

[3] S. Ben-David. Low-sensitivity functions from unambiguous certificates. *arXiv preprint arXiv:1605.07084*, 2016.

[4] H. Buhrman and R. De Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002.

[5] S. Chakraborty. *Sensitivity, block sensitivity and certificate complexity of Boolean functions.* PhD thesis, Masters thesis, University of Chicago, USA, 2005. 23, 2005.

[6] A. Gál and A. Rosén. A theorem on sensitivity and applications in private computation. *SIAM Journal on Computing*, 31(5):1424–1437, 2002.

[7] M. Göös, T. Pitassi, and T. Watson. Deterministic communication vs. partition number. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1077–1088. IEEE, 2015.

[8] P. Gopalan, N. Nisan, R. A. Servedio, K. Talwar, and A. Wigderson. Smooth boolean functions are easy: efficient algorithms for low-sensitivity functions. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 59–70. ACM, 2016.

[9] P. Gopalan, R. Servedio, A. Tal, and A. Wigderson. Degree and sensitivity: tails of two distributions. *arXiv preprint arXiv:1604.07432*, 2016.

[10] P. Hatami, R. Kulkarni, and D. Pankratov. Variations on the sensitivity conjecture. *arXiv preprint arXiv:1011.0354*, 2010.

[11] K. He, Q. Li, and X. Sun. A tighter relation between sensitivity and certificate complexity. *arXiv preprint arXiv:1609.04342*, 2016.

[12] C. Karthik and S. Tavenas. On the sensitivity conjecture for disjunctive normal forms. *arXiv preprint arXiv:1607.05189*, 2016.

[13] C. Kenyon and S. Kutin. Sensitivity, block sensitivity, and -block sensitivity of boolean functions. *Information and Computation*, 189(1):43–53, 2004.

[14] C. Lin and S. Zhang. Sensitivity conjecture and log-rank conjecture for functions with small alternating numbers. *arXiv preprint arXiv:1602.06627*, 2016.

[15] N. Nisan. Crew prams and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

[16] N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. *Computational complexity*, 4(4):301–313, 1994.

[17] H.-U. Simon. A tight $\Omega(\log\log n)$-bound on the time for parallel ram's to compute nondegenerated boolean functions. In *International Conference on Fundamentals of Computation Theory*, pages 439–444. Springer, 1983.

A Proof of Lemma 3.2

We restate and prove Lemma 3.2.

Lemma A.1. *Let $f : \{0,1\}^n \to \{-1,1\}$. Fix $d \geq 1$. Define*

$$A := 2^n \sum_{S \subseteq [n]} \widehat{f}(S)^2 \binom{|S|}{d}$$

and

$$B := |\{C \in \mathcal{C}(n,d) : f|_C \textit{ has max degree}\}|.$$

Then

$$2^{-d} B \leq A \leq 2^d B.$$

Proof. For a function $g : \{0,1\}^n \to \mathbb{R}$ define its directional derivative in direction $i \in [n]$ as $\Delta_i g : \{0,1\}^n \to \mathbb{R}$ given by $\Delta_i g(x) = g(x+e_i) - g(x)$. For a set of directions $I = \{i_1, \dots, i_d\}$ the iterated derivative is defined as

$$\Delta_I f(x) = (\Delta_{i_1} \dots \Delta_{i_d} f)(x) = \sum_{J \subseteq I} (-1)^{|I|-|J|} f(x+e_J).$$

In particular, the iterative derivative does not depend on the order of $i_1, \dots, i_d$, making $\Delta_I f$ well defined. Define

$$T := \{(x, I) : x \in \{0,1\}^n, I \subset [n], |I| = d, \Delta_I f(x) \neq 0\}.$$

We will see that $|T|$ is directly related to B, while A is related to the expression

$$\sum_{(x,I) \in T} (\Delta_I f(x))^2.$$

We first show that $B = 2^{-d}|T|$. To see that, fix a d-dimensional cube $C = C(v, I)$ and consider $f|_C$. Note that $\Delta_I f(v)$ is the sum with alternating signs of the points of C. In particular, if we let $f|_C(x) = \sum_{J \subseteq I} \widehat{f|_C}(J)(-1)^{\langle x, e_J \rangle}$ be the Fourier decomposition of $f|_C$, then

$$\Delta_I f(v) = \pm 2^d \cdot \widehat{f|_C}(I).$$

(the sign can be computed explicitly as $(-1)^{\langle v, e_I \rangle}$, but we don't need it). In particular, $f|_C$ has max degree iff $\Delta_I f(v) \neq 0$; namely exactly when $(v, I) \in T$. As this holds for any $v' \in C$ we have that

$$2^d B = |T|.$$

Next we relate T to A. To that end, we explore the effect of derivatives on the Fourier decomposition. It is easy to see that the Fourier decomposition of $\Delta_i f$ is

$$\Delta_i f(x) = 2 \sum_{S \subseteq [n]: i \in S} \widehat{f}(S)(-1)^{\langle x, e_S \rangle}.$$

Applying this iteratively for $I \subset [n]$ of size $|I| = d$ gives

$$\Delta_I f(x) = 2^d \sum_{S \subseteq [n]: I \subseteq S} \widehat{f}(S)(-1)^{\langle x, e_S \rangle}.$$

Thus we have

$$\sum_{x \in \{0,1\}^n} (\Delta_I f(x))^2 = 2^n \cdot 2^{2d} \sum_{S \subseteq [n]: I \subseteq S} \widehat{f}(S)^2.$$

Summing over all sets I with $|I| = d$, and restricting to $(x, I) \in T$ (otherwise by definition $\Delta_I f(x) = 0$ contributes nothing to the sum) gives

$$\sum_{(x,I) \in T} (\Delta_I f(x))^2 = 2^n \cdot 2^{2d} \sum_S \widehat{f}(S)^2 \binom{|S|}{d} = 2^{2d} A.$$

To conclude, note that whenever $(x, I) \in T$ then $1 \leq (\Delta_I f(x))^2 \leq 2^{2d}$, where the lower bound follows from $\Delta_I f(x)$ being a nonzero integer, and the upper bound from the fact that $\Delta_I f(x)$ is the sum with alternating signs of 2^d evaluations of a Boolean function f. Thus

$$2^d B = |T| \leq \sum_{(x,I) \in T} (\Delta_I f(x))^2 \leq 2^{2d} |T| = 2^{3d} B$$

and hence

$$2^{-d} B \leq A \leq 2^d B. \qquad \square$$

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il