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

Robust sensitivity

Shachar Lovett* Jiapeng Zhang†
University of California, San Diego
{slovett,jiz173}@cs.ucsd.edu

October 26, 2016

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.

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 [7] 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 [1–16], this conjecture remains wide open, where the best bounds on the degree are exponential in the sensitivity. The survey [9] provides a good account of the conjecture, many of it's equivalent formulations and consequences, and the progress so far.

*Research supported by NSF CAREER award 1350481, NSF CCF award 1614023 and a Sloan fellowship.
†Research supported by NSF CAREER award 1350481.

ISSN 1433-8092

A recent work of Gopalan, Servedio, Tal and Wigderson [8] 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 [7]). *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. [8] 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.

Conjecture 1.3 (Conjecture 1.3 in [7]). *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 that Theorem 1.2 still holds, even if we replace the assumption the 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. 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 $a_d \leq 2^{O(d^{3/2})}$.

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 [8]. 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 [8]. We first present a simplified bound of $a_d \le 2^{O(d^2)}$ in this section (see Theorem 4.6), and then proceed in Section 5 to give the improved bound of $a_d \le 2^{O(d^{3/2})}$, which yields Theorem 1.4.

2 Preliminaries

Boolean hypercube. We denote $[n] = \{1, \dots, n\}$. For $I \subseteq [n]$ we denote by e_I 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 $\{0,1\}^n$ and edges are $(x, x \oplus e_i)$ for $x \in \{0,1\}^n, i \in [n]$. 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 polynomial over $\{x_i : i \in I\}$ is maximal, namely d.

The following lemma connects the Fourier moments of f and the number of maximal degree cubes in f. It is a variant of an argument appearing in [8].

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

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

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 $\Delta_i g(x) = g(x \oplus e_i) - g(x)$. For a set of directions $I = \{i_1, \ldots, i_d\}$ the iterated derivative is defined as

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

In particular, the iterative derivative does not depend on the order of $i_1, \ldots, i_d$, 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 than 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

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

□

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 [8].

Definition 4.1 (Sensitivity graph)**.** *The sensitivity graph G_f of f is the sub-graph of $\mathcal{H}_n$ whose edges are $(x, x + e_i)$ where $f(x) \neq f(x + e_i)$. Edges of G_f are called "sensitive edges" of $\mathcal{H}_n$ with respect to f.*

Definition 4.2 (Proper walk)**.** *A proper walk in $\mathcal{H}_n$ with respect to f, is a directed path P with vertices $v_0, v_1, \ldots, v_m$ in $\mathcal{H}_n$ with the following property. Let $i_1, \ldots, i_m$ be the directions of the edges of this path. For any $i \in \{i_1, \ldots, i_m\}$, if $i_j = i$ is the first edge that a direction i appears in the path, then we require that the corresponding edge $(v_j, v_j + e_{i_j})$ is sensitive for f. If $j_1 < \ldots < j_d$ are the indices of the sensitive directions, we define*

- *The dimension of P is $d = |\{i_1, \ldots, i_m\}|$.*
- *The sensitive nodes of P are $v_{j_1}, \ldots, v_{j_d}$.*
- *The sensitive directions of P are $I(P) = (i_{j_1}, \ldots, i_{j_d})$.*

Observe that if P is a proper walk with sensitive nodes $v_1, \ldots, v_d$ and sensitive directions $i_1, \ldots, i_d$, then by definition $v_{j+1} \in C(v_\ell, \{i_1, \ldots, i_j\})$ for all $1 \le \ell \le j \le d-1$.

We define the cube defined by a proper walk to be the minimal sub-cube containing it, denoted $C(P)$. Equivalently, for any node v in P we have $C(P) = C(v, I(P))$. We say that $C(P)$ is *realized* by the proper walk P. We say that two proper walks are *equivalent* if they have the same sensitive nodes and sensitive directions. To simplify notation, from now on when we discuss a proper walk, we actually mean an equivalence class of proper walks. Otherwise put, we will only care about their sensitive nodes and sensitive directions, and ignore how the path exactly connects them.

4.2 Proper walks in maximal degree cubes

Let $f : \{0,1\}^n \to \{-1,1\}$. Gopalan et al. [8] 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 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 4.3 (Descending proper walk)**.** *Let P be a proper walk with respect to f. Let the sensitive nodes of P be $v_1, \ldots, v_d$. We say that P is* descending *if $s(f, v_1) \ge s(f, v_i)$ for all $i = 2, \ldots, d$.*

Lemma 4.4. *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 descending proper walk.*

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 $v_1, \ldots, v_d$ and sensitive directions $i_1, \ldots, i_d$, such that

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

We first observe that this suffices for the lemma. We may assume that $s(f, v_1) \ge s(f, v_1 + e_{i_1})$, 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

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

Next, we prove the existence of such a walk by induction on d. For $d = 1$ this is obvious, so assume $d \ge 2$. 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 4.4, in any d-dimensional sub-cube C where $f|_C$ has maximal degree, we can find a descending proper walk realizing it. Thus, instead of counting maximal degree sub-cubes, we will count descending proper walks.

Claim 4.5. *The number of d-dimensional descending 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 in G_f with sensitive nodes $x = v_1, v_2, \ldots, v_d$ and sensitive directions $i_1, \ldots, i_d$. We have $v_1 = x$ and $s(f,x)$ options 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 $v_j \in C(v_{j-1}, \{i_1, \ldots, i_{j-1}\})$, and hence it has at most 2^{j-1} different possibilities. Given that we obtained v_j, the number of choices for i_j is at most $s(f, v_j) \leq s(f,x)$. Thus we can bound the number of such walks by

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

$\square$

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

Theorem 4.6. *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 $a_d \leq 2^{\binom{d}{2}+d}$.

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 $A \leq 2^d B$. By Lemma 4.4 we can bound B by the number of d-dimensional decreasing proper walks with respect to f, and by Claim 4.5 this number is bounded by $2^{\binom{d}{2}} D$. Thus

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

$\square$

5 Improving the parameters

The goal in this section is to improve the parameters in Theorem 4.6. A keen reader (or one familiar with previous work [8]) can see that the main reason for the loss of parameters in

Theorem 4.6 is the number of potential descending 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 [8] (see Lemma 5.5 in the arxiv version), and we define it here explicitly.

Definition 5.1 (Description of a proper walk). *Let $f : \{0,1\}^n \to \{-1,1\}$. A d-dimensional proper walk P with respect to f can be described by three components:*

- *A starting node v.*
- *Its sensitive directions $I(P) = (i_1, \ldots, i_d)$.*
- *A description of each sensitive node given the previously defined sensitive nodes and directions. This is given by $R = (r_{i,j} : 1 \le j \le i \le d-1)$ where $r_{i,j} \in \{0,1\}$. Note that $R \in \{0,1\}^{\binom{d}{2}}$.*

The sensitive nodes of P are then defined as follows:

- $v_1 = v$.
- $v_2 = v + r_{1,1} \cdot e_{i_1}$.
- $v_3 = v + r_{2,1} \cdot e_{i_1} + r_{2,2} \cdot e_{i_2}$.
- *In general,* $v_{j+1} = v + r_{j,1} \cdot e_{i_1} + \ldots + r_{j,j} \cdot e_{i_j}$.

We define this proper walk by $P(v, I, R)$.

We also need to extend the notion of descending 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 5.2 (descending with respect to a weight function). *Let $g : \{0,1\}^d \to \{-1,1\}$ be a boolean function, and let P be a d-dimensional proper walk with respect to g. Let the sensitive nodes of P be $v_1, \ldots, v_d$. Let furthermore $w : \{0,1\}^d \to \mathbb{R}$ be some weight function. We say that P is* descending 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 descending proper walks.

Definition 5.3 (Family supporting descending proper walks). *Let $\mathcal{R} \subset \{0,1\}^{\binom{d}{2}}$. We say that $\mathcal{R}$* supports descending proper walks *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*

- $v \in \{0,1\}^d$

- $I = (i_1, \ldots, i_d)$ *a permutation of* $[d]$
- $R \in \mathcal{R}$

such that the walk $P = P(v, I, R)$ *is a proper walk with respect to* g *and descending with respect to* w.

One can verify that Lemma 4.4 can be extended to an arbitrary weight function. Thus, it establishes that $\mathcal{R} = \{0,1\}^{\binom{d}{2}}$ supports descending proper walks. This motivates the question of looking for the minimal such $\mathcal{R}$. This is further motivated by the following lemma.

Lemma 5.4. *Let* $\mathcal{R} \subset \{0,1\}^{\binom{d}{2}}$ *be a family which supports descending proper walks. Then Theorem 1.4 holds with the bound* $a_d = 2^d|\mathcal{R}|$.

Proof. The only change needed in the proof of Theorem 1.4 is in Claim 4.5, where instead of allowing for an arbitrary descending proper walk, we only allow for walks of the form $P = P(x, I, R)$ with $R \in \mathcal{R}$. Thus the number of 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. $\square$

We note that Lemma 5.5 in [8] proves the existence of family supporting proper walks, without the descending condition, of size $|\mathcal{R}| = 2^{4d}$. However, their proof does not give the descending 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 descending condition.

Conjecture 5.5. *There exists a family* $\mathcal{R}$ *supporting descending proper walks of size* $|\mathcal{R}| = 2^{O(d)}$.

Conjecture 5.5 would give optimal bounds in Theorem 1.4. Below, we give an intermediate bound. This can be taken as evidence that the relatively simple argument for Theorem 4.6 is not tight.

Theorem 5.6. *There exists a family* $\mathcal{R}$ *supporting descending proper walks of size* $|\mathcal{R}| = 2^{O(d^{3/2})}$.

Theorem 1.4 follows immediately from Theorem 5.6 and Lemma 5.4. Below, we give the details necessary to prove Theorem 5.6. We start with some more definitions from [8].

5.1 Sensitive trees

Let $g : \{0,1\}^d \to \{-1,1\}$ be a function. 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 5.7 (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 $V(T) \subset \{0,1\}^d$ the nodes of T, and by $I(T) \subset [n]$ the directions of the edges of T. Note that given any node $v \in V(T)$, by performing a depth-first search on T, we define a proper walk whose sensitive nodes are a subset of $V(T)$ and its sensitive directions are $I(T)$. This is summarized by the following lemma.

Lemma 5.8 (Lemma 5.3 in [8]). *Let $d \geq 1$. There exists $\mathcal{R}_{tree} = \mathcal{R}_{tree}(d) \subset \{0,1\}^{\binom{d}{2}}$ of size $|\mathcal{R}_{tree}| = 2^{2d}$ such that the following holds. Let $g : \{0,1\}^d \to \{-1,1\}$, T be a sensitive tree for g such that $I(T) = [d]$. Then, for any node $v \in V(T)$, there exists a d-dimensional proper walk $P = P(v, I, R)$ where I is a permutation of $[d]$ and $R \in \mathcal{R}_{tree}$.*

Proof. The length of a depth-first search walk in P, starting at node v, is $2d$. The walk can be described as a sequence of operations of two types: "follow next sensitive edge" or "backtrack". This can be encoded by $\mathcal{R}_{\text{tree}}$ of size $|\mathcal{R}_{\text{tree}}| = 2^{2d}$. □

Definition 5.9 (Shifting a sensitive tree). *Let T be a sensitive tree for g. We say that T can be shifted in direction J, 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 5.10 (Maximal sensitive tree). *Let T be a sensitive tree for g. We say that T is maximal if T can be shifted in all directions J where $J \cap I(T) = \emptyset$.*

Lemma 5.11 (Lemma 4.6 in [8]). *Any sensitive tree is either maximal, or otherwise, some shift of it is contained in a larger sensitive tree.*

Corollary 5.12. *Let T_0 be a sensitive tree for g. Then there exists a maximal sensitive tree T for g such that $I(T_0) \subseteq I(T)$.*

Proof. If T_0 is not maximal, by Lemma 5.11 we can shift it so that it is contained in a larger sensitive tree T_1. Note that $I(T_0) \subset I(T_1)$. Doing so iteratively will eventually result in a maximal sensitive tree. □

Lemma 5.13 (Theorem 4.9 in [8]). *If g has decision tree complexity d then there exists a maximal sensitive tree with then $|T| \geq \sqrt{2d} - 1$. In particular, this holds when g has max degree.*

5.2 Sensitive covers

Lemma 5.13 shows that if $g : \{0,1\}^d \to \{-1,1\}$ has max degree, then it contains a relatively large sensitive tree. We will iterate this procedure to contain a sequence of maximal sensitive trees which cover all the variables.

Definition 5.14 (Sensitive cover). *Let $g : \{0,1\}^d \to \{-1,1\}$. A* sensitive cover *for g is a collection of maximal sensitive trees $T_1, \ldots, T_m$ such that together they cover all directions, namely $I(T_1) \cup \ldots I(T_m) = [d]$.*

Lemma 5.15. *Let $g : \{0,1\}^d \to \{-1,1\}$ be a function of maximal decision tree depth d. Then there exists a sensitive cover for g consisting of $m = O(\sqrt{d})$ trees.*

Proof. We first argue that for any subset of the variables $I \subset [n]$, there exists a maximal sensitive tree T for g such that

$$|I(T) \cap I| \geq \sqrt{2|I|} - 1.$$

To see that, let $C = C(v, I)$ be a sub-cube such that $h = g|_C$ has maximal decision tree depth $|I|$ (if it doesn't exist, then g cannot have maximal decision tree depth). By Lemma 5.13, there exists a sensitive tree T_h for h with $|I(T_h)| \geq \sqrt{2|I|} - 1$. Next, we embed T_h as a sub-tree in G_g, and note that it is still a sensitive tree (but possibly not maximal). By Corollary 5.12 there exists a maximal sensitive tree T for g such that $I(T_h) \subset I(T)$.

We now apply this iteratively to construct a sequence of maximal sensitive trees $T_1, T_2, \ldots$ for g, where when we construct T_i, we apply the argument above to the set of yet-to-be covered directions, namely $[d] \setminus (I(T_1) \cup \ldots \cup I(T_{i-1}))$. Let $n_i = |I(T_i) \setminus (I(T_1) \cup \ldots I(T_{i-1})|$ be the number of new variables covered by T_i. Then we have

$$n_i \geq \sqrt{2(d - (n_1 + \ldots + n_{i-1}))} - 1.$$

In particular, as long as $n_1 + \ldots + n_{i-1} \leq d/2$, we have $n_i \geq \sqrt{d} - 1$, which shows that after at most $\sqrt{d}$ trees we covered at least $d/2$ variables. The total number of trees we need until covering all variables is thus at most $\sqrt{d} + \sqrt{d/2} + \sqrt{d/4} + \ldots = O(\sqrt{d})$. $\square$

We finish with an observation that sensitive covers are invariant to shifts. Let $T_1, \ldots, T_m$ be a sensitive cover for g. Recall that a maximal tree T' is a shift of a maximal tree T if $T' = T + e_J$ where $J \cap I(T) = \emptyset$. Maximality implies that $I(T') = I(T)$. In particular, if T'_i is a shift if T_i for $i = 1, \ldots, m$ then $T'_1, \ldots, T'_m$ is also a sensitive cover for g.

5.3 From sensitive covers to descending proper walks

Let $g : \{0,1\}^d \to \{-1,1\}$ be a boolean function, and let Let $T_1, \ldots, T_m$ be a sensitive cover for g. The next lemma shows that by possibly shifting each maximal sensitive tree in the cover, we can obtain a descending proper walk for any given weight function.

For simplicity of exposition, we make the following definitions. Given a sensitive cover $T_1, \ldots, T_m$, we denote by $C(T_1, \ldots, T_m)$ the minimal sub-cube containing it. Note that $C(T_1, \ldots, T_m) = C(v, I(T_1) \cup \ldots \cup I(T_m))$ for any node v of any of $T_1, \ldots, T_m$. For a weight function $w : \{0,1\}^n \to \mathbb{R}$, we denote by $w(T) := \max_{v \in V(T)} w(v)$ the maximal weight of a node in T. For a sensitive cover we denote $w(T_1, \ldots, T_m) := \max_{i=1,\ldots,m} w(T_i)$ the maximal weight of a tree in the cover.

Lemma 5.16. *Let $g : \{0,1\}^d \to \{-1,1\}$ a function, $T_1, \ldots, T_m$ a sensitive cover for g. Let $w : \{0,1\}^n \to \mathbb{R}$ be a weight function. Then there exists a sensitive cover $T'_1, \ldots, T'_m$ for g such that*

(i) There exists a permutation π over $[m]$ such that T'_i is a shift of $T_{\pi(i)}$.

(ii) For each $i = 2, \ldots, m$, $C(T'_1, \ldots, T'_{i-1}) \cap V(T'_i) \neq \emptyset$.

(iii) $w(T'_1) = w(T'_1, \ldots, T'_m)$.

Proof. First, note that we are allowed to shift each tree T_i, as well as permute the order of the trees, and still satisfy (i). These will be the operations that we do below.

Next, observe that if T is a maximal sensitive tree and $v \in \{0,1\}^d$, then there is a shift T' of T such that $v \in C(T')$. Thus condition (ii) can be satisfied. In fact, we can even guarantee the stronger condition that $C(T'_{i-1}) \cap V(T'_i) \neq \emptyset$. In order to achieve that, set $T'_m = T_m$, and for $i = m-1, \ldots, 1$, shift T_i to T'_i so that $C(T'_i)$ contains some arbitrary node of T'_{i+1}. Thus, we can accomplish (i)+(ii). We next move to accomplish (iii) as well.

Define the *index* of $T'_1, \ldots, T'_m$ to be the minimal index of a tree with maximal weight,

$$\text{index}(T'_1, \ldots, T'_m) := \min\{i \in [m] : w(T'_i) = w(T'_1, \ldots, T'_m)\}.$$

Let $T'_1, \ldots, T'_m$ be a sensitive cover which satisfy both (i) and (ii); among all these, choose the one which minimizes its weight $w(T'_1, \ldots, T'_m)$; and among all of these, choose the one that minimizes its index $\text{index}(T'_1, \ldots, T'_m)$. We will prove that $\text{index}(T'_1, \ldots, T'_m) = 1$, thus proving the lemma.

Assume towards contradiction that $\text{index}(T'_1, \ldots, T'_m) = i$ for some $i > 1$. Let $I^* = (I(T'_1) \cup \ldots \cup I(T'_{i-1})) \setminus I(T'_i)$. Observe that for any $J \subseteq I^*$ we have $w(T'_i + e_J) \geq w(T'_i)$. Otherwise, we could replace T'_i with $T'_i + e_J$, which would preserver (i)+(ii) and would either decrease the weight of the tree cover, or otherwise decrease its index. Consider the following sensitive cover:

$$T''_1 = T'_i + e_J, T''_2 = T'_1, \ldots, T''_i = T'_{i-1}, T''_{i+1} = T'_{i+1}, \ldots, T''_m = T'_m.$$

We choose J so that, for some arbitrary chosen node $v \in T'_1$, it holds that $v \in C(T''_1)$. Observe that $T''_1, \ldots, T''_m$ indeed satisfy (i)+(ii).

Next, we claim that $J \subseteq I^*$. This is since $T'_1, T'_i \in C(T'_1, \ldots, T'_i)$, and hence all their nodes agree on coordinates outside $I(T'_1) \cup \ldots \cup I(T'_i)$. Moreover, J must be disjoint from $I(T'_i)$. Hence $J \subseteq I^*$. This completes the proof: we have that conditions (i)+(ii) still hold and that in addition

$$w(T''_1) \geq w(T'_i) = w(T'_1, \ldots, T'_m) \geq w(\{T''_j : j > 1\}).$$

Hence (iii) must hold. $\square$

We next show how to use this sensitive cover obtained in Lemma 5.16 to obtain a decreasing proper walk.

Lemma 5.17. *Let $d, m \geq 1$. There exists $\mathcal{R} = \mathcal{R}(d,m) \subset \{0,1\}^{\binom{d}{2}}$ of size $|\mathcal{R}| \leq 2^{4dm}$ such that the following holds. Let $g : \{0,1\}^d \to \{-1,1\}$ a function, $w : \{0,1\}^n \to \mathbb{R}$ be a weight function, $T_1, \ldots, T_m$ a sensitive cover for g that satisfies*

(i) For each $i = 2, \ldots, m$, $C(T_1, \ldots, T_{i-1}) \cap V(T_i) \neq \emptyset$.

(iii) For each $i = 2, \ldots, m$, $w(T_1) \geq w(T_i)$.

Then there exists a decreasing proper walk P for g. Moreover, $P = P(v, I, R)$ where I is a permutation of $[d]$ and $R \in \mathcal{R}$.

Proof. Let $C_i := C(T_i)$ be the sub-cube containing T_i, and let $d_i := |I(T_i)|$. Fix nodes $v_i \in T_i$ such that $w(v_1) = w(T_1)$ and $w(v_i) \in C(T_1, \ldots, T_{i-1})$ for $i > 1$. By Lemma 5.8 applied to $f|_{C_i}$, there exists a proper walk P_i which realizes C_i, such that $P_i = P_i(v_i, I_i, R_i)$ where I_i is a permutation of $I(T_i)$ and $R_i \in \mathcal{R}_{\text{tree}}(d_i)$. Let P be the walk obtained by concatenating $P_1, \ldots, P_m$. Observe that it is a proper walk by our assumption, and furthermore it is descending as it starts at v_1, which has maximal weight in T_1, which in tune has maximal weight among all $T_1, \ldots, T_m$. The walk P description is composed of:

- The descriptions of each P_i, each has size $2^{2d_i} \leq 2^{2d}$.
- The descriptions of each v_i, $i > 1$, with respect to $I_{<i} := I(T_1) \cup \ldots \cup I(T_{i-1})$. This has size $2^{|I_{<i}|} \leq 2^{2d}$.

These can be all be encoded by a single $\mathcal{R} \subset \{0,1\}^{\binom{d}{2}}$ of size $|\mathcal{R}| \leq 2^{4dm}$. □

Theorem 5.6 now follows. By Lemma 5.15 there always exists a sensitive cover of size $m = O(\sqrt{d})$. By Lemma 5.16, it can be shuffled to obtain the required conditions to apply Lemma 5.17.

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] 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.

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

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

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

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

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

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

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

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

[16] 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.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de