A Spectral Method for the Sensitivity Conjecture

Yuchong Pan

April 16, 2021

Abstract

In this short essay, we introduce several complexity measures for Boolean functions and present a spectral method, due to Huang [9], for proving the long-standing sensitivity conjecture $bs(f) \leq \text{poly}(s(f))$ posed by Nisan and Szegedy [12] in 1992. We also show that the example of Rubinstein [13] gives a quadratic separation between sensitivity and block sensitivity. Finally, several related questions that remain unsolved are collected.

1 Introduction

There are several measures of complexity for Boolean functions. Sensitivity is one of the simplest and the most fundamental complexity measure for Boolean functions. It measures how many neighbours of an input have different function values. Rigorously, we define the sensitivity of a Boolean function as follows. For $x \in \{0,1\}^n$ and $S \subseteq [n]$, we denote by x^S the binary vector obtained from x by flipping indices from S. Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the *local sensitivity* $s(f,x)$ of f on the input x is defined as the number of indices i such that $f(x) \neq f(x^{\{i\}})$, and the *sensitivity* $s(f)$ of f is defined as the maximum of $s(f,x)$ over all inputs x.

The sensitivity of a Boolean function can be regarded as a discrete analogue of the smoothness of a continuous function, which measures how gradually the function changes locally [5]. Let $\delta(\cdot,\cdot)$ be the *normalized Hamming metric*, i.e., $\delta(x,y) = \frac{1}{n}\|x-y\|_1$ for all $x, y \in \{0,1\}$. Let $x \in \{0,1\}^n$, $\delta_0 \in \{0, \frac{1}{n}, \ldots, 1\}$ and $d_0 = \delta_0 n$. For $y \in \{0,1\}^n$ with $x \neq y$, we denote by y' the binary vector obtained from y by flipping the last index at which x and y differ. For $d \subseteq \{0,\ldots,n\}$, we denote by $S(x,d)$ the set of binary vectors that differ from x by d indices. By the triangle inequality,

$$\begin{aligned}
\mathop{\mathbb{E}}_{y:\delta(x,y)=\delta_0} |f(x)-f(y)| &= \frac{1}{\binom{n}{d_0}} \sum_{y:\delta(x,y)=\delta_0} |f(x)-f(y)| = \frac{1}{\binom{n}{d_0}} \sum_{y\in S(x,d_0)} |f(x)-f\left(y'\right)+f\left(y'\right)-f(y)| \\
&\leq \frac{1}{\binom{n}{d_0}} \sum_{y\in S(x,d_0)} \left(|f(x)-f\left(y'\right)|+|f\left(y'\right)-f(y)|\right) \\
&\leq \frac{1}{\binom{n}{d_0}} \sum_{y'\in S(x,d_0-1),x_n=y'_n} |f(x)-f\left(y'\right)| \sum_{y\in S(y',1)} |f\left(y'\right)-f(y)| \\
&\leq \frac{1}{\binom{n}{d_0}} \sum_{y'\in S(x,d_0-1),x_n=y'_n} |f(x)-f\left(y'\right)|\, s(f) \leq \frac{s(f)}{\binom{n}{d_0}} \sum_{y'\in S(x,d_0-1),x_n=y'_n} 1 \\
&= \frac{s(f)}{\binom{n}{d_0}}\binom{n-1}{d_0-1} = \frac{d_0}{n}s(f) = \delta_0 s(f).
\end{aligned}$$

Recall that, given two metric spaces $(X,d_X),(Y,d_Y)$, a function $f : X \to Y$ is called *Lipschitz continuous* if there exists $L > 0$ such that, for all $x_1, x_2 \in X$,

$$d_Y\left(f\left(x_1\right), f\left(x_2\right)\right) \leq L d_X\left(x_1, x_2\right),$$

where any such L is called a *Lipschitz constant* for the function f. Therefore, $s(f)$ can be regarded as an analogue of of the Lipschitz constant of a Boolean function f.

Block sensitivity is a related yet less intuitive measure of complexity for Boolean functions introduced by Nisan [11]. Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the *local block sensitivity* $bs(f,x)$ of f on the input x is defined as the maximum number of disjoint subsets $B_1, \ldots, B_k \subseteq [n]$ such that $f(x) \neq f(x^{B_i})$ for each $i \in [k]$, and the *block sensitivity* $bs(f)$ of f is defined as the maximum of $bs(f,x)$ over all inputs x. Since $\{1\}, \ldots, \{n\}$ are disjoint subsets of $[n]$, then $bs(f) \geq s(f)$ for any Boolean function f.

Two complexity measures α, β for Boolean functions are called *polynomially related* if there exist polynomials p_1, p_2 such that for any Boolean function f,

$$\alpha(f) \leq p_1(\beta(f)), \qquad \beta(f) \leq p_2(\alpha(f)).$$

It is known that block sensitivity is polynomially related to many other important measures of complexity for Boolean functions, such as *deterministic decision tree complexity*, *certificate complexity*, *degree*, *approximate degree*, *randomized query complexity*, *quantum query complexity*, etc. We refer the readers to the surveys of Buhrman and de Wolf [3] and Hatami et al. [7] for comprehensive expositions. It had been a long-standing open question, posed by Nisan and Szegedy [12] in 1992, asking whether sensitivity and block sensitivity are polynomially related until Huang [9] solved it in 2019. The positive answer to this question was previously known as the *sensitivity conjecture*.

Theorem 1.1 (Huang [9])**.** *For every Boolean function f,*

$$bs(f) \leq s(f)^4.$$

This theorem implies that sensitivity is also polynomially related to the other aforementioned complexity measures for Boolean functions. Therefore, low-sensitivity Boolean functions are easy to compute in simple computational models such as the deterministic decision tree model, and have low degrees as real multilinear polynomials.

2 Cauchy's Interlacing Theorem

In this section, we state and prove Cauchy's interlacing theorem, which plays an essential role in Huang [9]. Various proofs have been found for Cauchy's interlacing theorem. The one we provide below is based upon the Rayleigh quotient theorem. Alternatively, Cauchy's interlacing theorem can be proved from the Courant-Fischer minimax theorem and from Sylvester's law of inertia, respectively (indeed, these three theorems are equivalent). It is noteworthy that Fisk [4] gives a two-sentence proof of Cauchy's interlacing theorem from polynomial interlacing, which simplifies an elementary proof of Hwang [10] based upon the intermediate value theorem.

We first state and prove the Rayleigh quotient theorem. The proof we present below can be found in [8].

Theorem 2.1 (Rayleigh)**.** *Let $A \in \mathcal{M}_n(\mathbb{C})$ be Hermitian. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A with $\lambda_1 \geq \ldots \geq \lambda_n$. Let $i_1, \ldots, i_k \in [n]$ with $i_1 < \ldots < i_k$. Let $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_k}$ be orthonormal and such that $A\mathbf{v}_{i_j} = \lambda_{i_j}\mathbf{v}_{i_j}$ for each $j \in [k]$. Let $S = \operatorname{span}\{\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_p}\}$. Then*

$$\lambda_{i_1} = \max_{\substack{\mathbf{v} \in S \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}}, \qquad \lambda_{i_k} = \min_{\substack{\mathbf{v} \in S \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}}.$$

Proof. For $\mathbf{v} \in S$ with $\mathbf{v} \neq \mathbf{0}$, we denote by $\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|_2}$ the unit vector corresponding to $\mathbf{v}$. Note that

$$\frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}} = \frac{(\|\mathbf{v}\|_2 \hat{\mathbf{v}})^H A (\|\mathbf{v}\|_2 \hat{\mathbf{v}})}{\|\mathbf{v}\|_2^2} = \frac{\|\mathbf{v}\|_2^2 \cdot \hat{\mathbf{v}}^H A \hat{\mathbf{v}}}{\|\mathbf{v}\|_2^2} = \hat{\mathbf{v}}^H A \hat{\mathbf{v}}.$$

Let $\mathbf{v} \in S$ with $\|\mathbf{v}\|_2 = 1$. Since $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_k}$ are orthonormal, then there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\mathbf{v} = \sum_{j=1}^k \alpha_j \mathbf{v}_{i_j}$. By the orthonormality of $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_k}$,

$$1 = \|\mathbf{v}\|_2^2 = \mathbf{v}^H \mathbf{v} = \left(\sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} \right)^H \sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} = \sum_{j,j' \in [k]} \overline{\alpha_j} \alpha_{j'} \mathbf{v}_{i_j}^H \mathbf{v}_{i_{j'}} = \sum_{j=1}^k \overline{\alpha_j} \alpha_j \mathbf{v}_{i_j}^H \mathbf{v}_{i_j} = \sum_{j=1}^k |\alpha_j|^2 .$$

Therefore,

$$\begin{aligned}
\mathbf{v}^H A \mathbf{v} &= \left(\sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} \right)^H A \left(\sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} \right) = \left(\sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} \right)^H \sum_{j=1}^k \alpha_j A \mathbf{v}_{i_j} \\
&= \left(\sum_{j=1}^k \alpha_j \mathbf{v}_{i_j} \right)^H \sum_{j=1}^k \alpha_j \lambda_{i_j} \mathbf{v}_{i_j} = \sum_{j,j' \in [k]} \overline{\alpha_j} \alpha_{j'} \lambda_{i_{j'}} \mathbf{v}_{i_j}^H \mathbf{v}_{i_{j'}} \\
&= \sum_{j=1}^k \overline{\alpha_j} \alpha_j \lambda_{i_j} \mathbf{v}_{i_j}^H \mathbf{v}_{i_j} = \sum_{j=1}^k |\alpha_j|^2 \lambda_{i_j}.
\end{aligned}$$

In other words, $\mathbf{v}^H A \mathbf{v}$ is a convex combination of $\lambda_{i_1}, \ldots, \lambda_{i_k}$. Hence,

$$\lambda_{i_k} = \min_{j=1,\ldots,k} \lambda_{i_j} \leq \mathbf{v}^H A \mathbf{v} \leq \max_{j=1,\ldots,k} \lambda_{i_j} = \lambda_{i_1}.$$

By the orthonormality of $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_k}$, we have ${\mathbf{v}_{i_1}}^H A \mathbf{v}_{i_1} = \lambda_{i_1}$ and $\mathbf{v}_{i_k}^H A \mathbf{v}_{i_k} = \lambda_{i_k}$. Hence,

$$\lambda_{i_1} = \max_{\substack{\mathbf{v} \in S \\ \|\mathbf{v}\|_2 = 1}} \mathbf{v}^H A \mathbf{v} = \max_{\substack{\mathbf{v} \in S \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}}, \qquad \lambda_{i_k} = \min_{\substack{\mathbf{v} \in S \\ \|\mathbf{v}\|_2 = 1}} \mathbf{v}^H A \mathbf{v} = \min_{\substack{\mathbf{v} \in S \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}}.$$

This completes the proof. □

The Rayleigh quotient theorem implies Cauchy's interlacing theorem. For $A \in \mathcal{M}_n(\mathbb{C})$, a *principal submatrix* of A is obtained by deleting the same set of rows and columns from A.

Theorem 2.2 (Cauchy)**.** *Let $A \in \mathcal{M}_n(\mathbb{C})$ be Hermitian. Let B be an $m \times m$ principal submatrix of A for some $m \in [n]$. If the eigenvalues of A are $\lambda_1 \geq \ldots \geq \lambda_n$, and the eigenvalues of B are $\mu_1 \geq \ldots \geq \mu_m$, then for $i \in [m]$,*

$$\lambda_i \geq \mu_i \geq \lambda_{i+n-m}.$$

Proof. By relabelling the rows and the columns of A, we assume without loss of generality

$$A = \begin{bmatrix} B & C \\ C^H & D \end{bmatrix}.$$

Since A is Hermitian and since B is a principal submatrix of A, then B is Hermitian. Note that Hermitian matrices have a orthonormal eigenbasis. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be orthonormal eigenvectors

of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. Let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be orthonormal eigenvectors of B corresponding to the eigenvalues $\mu_1, \ldots, \mu_m$, respectively. Let $i \in [m]$. Let

$$V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+n-m}\}, \quad W = \operatorname{span}\{\mathbf{w}_i, \ldots, \mathbf{w}_m\}, \quad \widetilde{W} = \left\{ \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^n, \mathbf{w} \in W \right\}.$$

By the orthonormality of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and of $\mathbf{w}_1, \ldots, \mathbf{w}_m$, we have $\dim V = i + n - m$ and $\dim W = \dim \widetilde{W} = m - i + 1$. Since $\dim V + \dim \widetilde{W} > n$, then there exists $\tilde{\mathbf{w}} = [\mathbf{w}\ \mathbf{0}]^T \in V \cap \widetilde{W}$ for some $\mathbf{w} \in W$ with $\mathbf{w} \neq \mathbf{0}$. Note $\tilde{\mathbf{w}}^H \tilde{\mathbf{w}} = \|\tilde{\mathbf{w}}\|_2^2 = \|\mathbf{w}\|_2^2 = \mathbf{w}^H \mathbf{w}$. Therefore,

$$\tilde{\mathbf{w}}^H A \tilde{\mathbf{w}} = \begin{bmatrix} \mathbf{w}^H & \mathbf{0}^H \end{bmatrix} \begin{bmatrix} B & C \\ C^H & D \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{w}^H & \mathbf{0}^H \end{bmatrix} \begin{bmatrix} B\mathbf{w} \\ C^H \mathbf{w} \end{bmatrix} = \mathbf{w}^H B \mathbf{w}.$$

By the Rayleigh quotient theorem,

$$\lambda_{i+n-m} = \min_{\substack{\mathbf{v} \in V \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}} \leq \frac{\tilde{\mathbf{w}}^H A \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}} = \frac{\mathbf{w}^H B \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \leq \max_{\substack{\mathbf{v} \in W \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H B \mathbf{v}}{\mathbf{v}^H \mathbf{v}} = \mu_i.$$

On the other hand, let

$$V = \operatorname{span}\{\mathbf{v}_i, \ldots, \mathbf{v}_n\}, \quad W = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_i\}, \quad \widetilde{W} = \left\{ \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^n, \mathbf{w} \in W \right\}.$$

By the orthonormality of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and of $\mathbf{w}_1, \ldots, \mathbf{w}_m$, we have $\dim V = n - i + 1$ and $\dim W = \dim \widetilde{W} = i$. Since $\dim V + \dim \widetilde{W} > n$, then there exists $\tilde{\mathbf{w}} = [\mathbf{w}\ \mathbf{0}]^T \in V \cap \widetilde{W}$ for some $\mathbf{w} \in W$ with $\mathbf{w} \neq \mathbf{0}$. Note $\tilde{\mathbf{w}}^H \tilde{\mathbf{w}} = \mathbf{w}^H \mathbf{w}$ and $\tilde{\mathbf{w}}^H A \tilde{\mathbf{w}} = \mathbf{w}^H B \mathbf{w}$ as before. By the Rayleigh quotient theorem,

$$\lambda_i = \max_{\substack{\mathbf{v} \in V \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H A \mathbf{v}}{\mathbf{v}^H \mathbf{v}} \geq \frac{\tilde{\mathbf{w}}^H A \tilde{\mathbf{w}}}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}} = \frac{\mathbf{w}^H B \mathbf{w}}{\mathbf{w}^H \mathbf{w}} \geq \min_{\substack{\mathbf{v} \in W \\ \mathbf{v} \neq \mathbf{0}}} \frac{\mathbf{v}^H B \mathbf{v}}{\mathbf{v}^H \mathbf{v}} = \mu_i.$$

This completes the proof. □

3 The Gotsman-Linial Equivalence Theorem

Another pioneering work on which the proof of Huang [9] is based is the equivalence theorem of Gotsman and Linial [6], by which it suffices to prove a problem on the n-dimensional Boolean hypercube in order to prove Theorem 1.1.

Before we delve into the statement and the proof of the equivalence theorem of Gotsman and Linial [6], we introduce an additional complexity measure for Boolean functions—the *degree* of a Boolean function as a multilinear polynomial. For $S \subseteq [n]$, the n-variable function $X_S = \prod_{i \in S} x_i$ on $x_1, \ldots, x_n$ is called a *monomial* whose *degree* is defined as $|S|$. A *multilinear polynomial* is a function $p : \mathbb{R}^n \to \mathbb{R}$ such that $p(x) = \sum_{S \subseteq [n]} \alpha_S X_S$ for some coefficients $\alpha_S \in \mathbb{R}, S \subseteq [n]$, whose *degree* is defined as the largest degree of its mononials with non-zero coefficients, i.e., $\max\{|S| : S \subseteq [n], \alpha_S \neq 0\}$. For $S \subseteq [n]$, we denote by $\chi^S \in \{0,1\}^n$ be the characteristic vector of S, i.e., $\chi_i^S = 1$ if and only if $i \in S$ for each $i \in [n]$. The next two lemmas imply that each Boolean functions has a unique multilinear polynomial representation. For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the *degree* $\deg(f)$ of f is defined as the degree of the unique multilinear polynomial that represents f.

Lemma 3.1. *Let $p, q : \mathbb{R}^n \to \mathbb{R}$ be two multilinear polynomials. Suppose that $p(x) = q(x)$ for all $x \in \{0,1\}^n$. Then $p = q$.*

Proof. Suppose for the sake of contradiction that $p - q \not\equiv 0$. Then $p - q = \sum_{S \subseteq [n]} \alpha_S X_S$ for some coefficients $\alpha_S \in \mathbb{R}, S \subseteq [n]$ such that at least one of the coefficients is non-zero. Let $S^* \subseteq [n]$ be a minimal set of vertices such that $\alpha_{S^*} \neq 0$. By the minimality of S^*, $(p-q)(\chi^{S^*}) \neq 0$. This implies that $p(\chi^{S^*}) \neq q(\chi^{S^*})$, a contradiction. □

Lemma 3.2. *Let $f : \{0,1\}^n \to \{0,1\}$. Then there exists a multilinear polynomial $p : \mathbb{R}^n \to \mathbb{R}$ such that $p(x) = f(x)$ for all $x \in \{0,1\}^n$.*

Proof. We proceed by induction on $\mathbb{Z}_{\geq 0}$. The base case is vacuously true. Let $n \in \mathbb{N}$. Suppose that for all $(n-1)$-variable Boolean functions $f : \{0,1\}^{n-1} \to \{0,1\}$, there exists a multilinear polynomial $p : \mathbb{R}^{n-1} \to \mathbb{R}$ such that $p(x) = f(x)$ for all $x \in \{0,1\}^{n-1}$. Let $g : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Let $f_1, f_2 : \{0,1\}^{n-1} \to \{0,1\}$ be defined by

$$f_1(x_1, \ldots, x_{n-1}) = g(x_1, \ldots, x_{n-1}, 1), \qquad f_2(x_1, \ldots, x_{n-1}) = g(x_1, \ldots, x_{n-1}, 0).$$

Then there exist multilinear polynomials $p_1, p_2 : \mathbb{R}^{n-1} \to \mathbb{R}$ such that $p_1(x) = f_1(x)$ and $p_2(x) = f_2(x)$ for all $x \in \{0,1\}^{n-1}$. Let $p : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$p(x_1, \ldots, x_n) = x_n p_1(x_1, \ldots, x_{n-1}) + (1 - x_n) p_2(x_1, \ldots, x_{n-1}).$$

Therefore, p is a multilinear polynomial. It is straightforward to verify that $p(x) = g(x)$ for all $x \in \{0,1\}^n$. This completes the induction step. □

Tal [14] proves the following relation between the block sensitivity of a Boolean function and its degree as a multilinear polynomial, which improves the relation $bs(f) \leq 2\deg(f)^2$ of Nisan and Szegedy [12] by a factor of 2.

Theorem 3.3 (Tal [14])**.** *For every Boolean function f,*

$$bs(f) \leq \deg(f)^2.$$

Now, we state and prove the equivalence theorem of Gotsman and Linial [6]. We denote by $\mathbb{B}^n$ the n-dimensional Boolean hypercube. Given an undirected graph G, we denote by $d_G(v)$ the degree of each $v \in V(G)$, and denote by $\Delta(G)$ the maximum degree of G. For an undirected graph G and an induced subgraph H of G, we denote by $G \setminus H$ the subgraph of G induced by the vertex set $V(G) \setminus V(H)$. Moreover, given an induced subgraph H of $\mathbb{B}^n$, we define

$$\Gamma(H) = \max\{\Delta(H), \Delta(\mathbb{B}^n \setminus H)\}.$$

Theorem 3.4 (Gotsman and Linial [6])**.** *The following are equivalent for any monotone function $h : \mathbb{N} \to \mathbb{R}$.*

1. *For any induced subgraph H of $\mathbb{B}^n$ with $|V(H)| \neq 2^{n-1}$, we have $\Gamma(H) \geq h(n)$.*
2. *For any Boolean function f, we have $s(f) \geq h(\deg(f))$.*

Proof. To simplify the argument, we note that Boolean functions can be equivalently written in the form of $\{-1,1\}^n \to \{-1,1\}$, and that each Boolean function has a unique multilinear polynomial representation that coincides on $\{-1,1\}^n$ by slight modifications to Lemma 3.1 and Lemma 3.2.

Let $h : \mathbb{N} \to \mathbb{R}$ be monotone. We denote by $\mathbb{E}(g)$ the average value of a Boolean function g. Moreover, We first prove that (1) is equivalent to (1'), and that (2) is equivalent to (2'), where (1') and (2') are stated below:

1'. For any Boolean function g, $\mathbb{E}(g) \neq 0$ implies $s(g,x) \leq n - h(n)$ for some $x \in \{-1,1\}^n$.

2'. For any Boolean function f, $s(f) < h(n)$ implies $\deg(f) < n$.

We show that (1) and (1') are equivalent. We define a bijection between induced subgraphs H of $\mathbb{B}^n$ and Boolean functions by associating with each induced subgraph H of $\mathbb{B}^n$ the Boolean function $g_H : \{-1,1\}^n \to \{-1,1\}$ such that $g_H(x) = 1$ if and only if $x \in V(H)$. Hence, $\mathbb{E}(g_H) \neq 0$ if and only if $|V(H)| \neq 2^{n-1}$. Moreover, this implies that $d_H(x) = n - s(g,x)$ for each $x \in V(H)$, and $d_{\mathbb{B}^n \setminus H}(x) = n - s(g,x)$ for each $x \in V(\mathbb{B}^n \setminus H)$. Therefore, $\Gamma(H) \geq h(n)$ if and only if $\min s(g,x) \leq n - h(x)$, i.e., there exists $x \in \{-1,1\}^n$ such that $s(g,x) \leq n - h(n)$. This proves the equivalence between (1) and (1').

We show that (2) and (2') are equivalent. Since h is monotone, then $\deg(f) = n$ and (2) imply $h(n) = h(\deg(f)) \leq s(f)$, so (2) implies (2'). It remains to show that (2') implies (2). Let f be a Boolean function with $\deg(f) = d$. Then there exists a non-vanishing monomial of degree d in the unique multilinear polynomial representation of f. By relabelling indices, we assume without loss of generality that this monomial is $x_1 \dots x_d$. Let $g : \{-1,1\}^d \to \{-1,1\}$ be defined by $g(x_1,\dots,x_d) = f(x_1,\dots,x_d,1,\dots,1)$. Therefore, (2') implies $s(f) \geq s(g) \geq h(d) = h(\deg(f))$. This proves that (2') implies (2).

It remains to show that (1') and (2') are equivalent. Let $f : \{-1,1\}^n \to \{-1,1\}$. Let $p : \mathbb{R}^n \to \mathbb{R}$ be the unique multilinear representation of f given by $p = \sum_{S \subseteq [n]} \alpha_S X_S$ for some coefficients $\alpha_S \in \mathbb{R}, S \subseteq [n]$. We denote by $\hat{p}(S)$ be the coefficient of X_S in p. We show $\mathbb{E}(f) = \hat{p}(\varnothing)$. Note

$$\begin{aligned}
\mathbb{E}(f) &= \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x) = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \sum_{S \subseteq [n]} \alpha_S \prod_{i \in S} x_i \\
&= \frac{1}{2^n} \left(\sum_{x \in \{-1,1\}^n} \alpha_\varnothing + \sum_{\substack{\varnothing \neq S \subseteq [n] \\ S = \{i_1,\dots,i_k\}}} \alpha_S \sum_{x_{i_1} \in \{-1,1\}} \cdots \sum_{x_{i_k} \in \{-1,1\}} x_{i_k} \right) \\
&= \frac{1}{2^n} \left(2^n \alpha_\varnothing + \sum_{\substack{\varnothing \neq S \subseteq [n] \\ S = \{i_1,\dots,i_k\}}} \alpha_S \cdot 0 \cdots 0 \right) = \alpha_\varnothing = \hat{p}(\varnothing).
\end{aligned}$$

Let $g : \{-1,1\}^n \to \{-1,1\}$ be defined by $g(x) = f(x)\pi(x)$, where $\pi : \{-1,1\}^n \to \{-1,1\}$ is such that $\pi(x) = -1$ if and only if x has an odd number of 1-components (called the *parity function*). It is straightforward that $s(g,x) = n - s(f,x)$ for each $x \in \{-1,1\}^n$ and the unique multilinear polynomial representation of π is $\prod_{i=1}^n x_i$. Let $q : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$q(x_1,\dots,x_n) = p(x_1,\dots,x_n) \prod_{i=1}^n x_i = \sum_{S \subseteq [n]} \alpha_S \prod_{i \in S} x_i \prod_{i=1}^n x_i = \sum_{S \subseteq [n]} \alpha_S \prod_{i \in S} x_i^2 \prod_{i \in [n] \setminus S} x_i.$$

Since $x_i^2 = 1$ if $x_i \in \{-1,1\}$ and since each Boolean function has a unique multilinear polynomial representation, then

$$q(x_1,\dots,x_n) = \sum_{S \subseteq [n]} \alpha_S \prod_{i \in [n] \setminus S} x_i.$$

Hence, q is the unique multilinear representation of g, and $\hat{p}(S) = \hat{q}([n] \setminus S)$ for each $S \subseteq [n]$.

We first show that (1') implies (2'). Let $f : \{-1,1\}^n \to \{-1,1\}$ be such that $\deg(f) = n$. Let g, p, q be defined as above. Then $\hat{p}([n]) \neq 0$. This implies $\mathbb{E}(g) = \hat{q}(\varnothing) = \hat{p}([n]) \neq 0$. Therefore, (1')

implies $n - s(f,x) = s(g,x) \leq n - h(n)$ and hence $s(f) \geq s(f,x) \geq h(n)$ for some $x \in \{-1,1\}^n$. Now, we show that (2') implies (1'). Let $f : \{-1,1\}^n \to \{-1,1\}$ be such that $s(f,x) > n - h(n)$ for all $x \in \{-1,1\}^n$. Let g, p, q be defined as above. Then $s(g,x) = n - s(f,x) < h(n)$ for all $x \in \{-1,1\}^n$. This implies $s(g) < h(n)$. By (2'), $\deg(g) < n$. Therefore, $\mathbb{E}(f) = \hat{p}(\varnothing) = \hat{q}([n]) = 0$. Hence, (1') and (2') are equivalent, completing the proof. □

4 The Spectral Method

Having stated the two classic results upon which the proof of Huang [9] is based, we are now ready to finally introduce the spectral method for Theorem 1.1. By the equivalence theorem of Gotsman and Linial [6], Huang [9] proves the following graph-theoretic theorem which implies Theorem 1.1.

Theorem 4.1 (Huang [9])**.** *For $n \in \mathbb{N}$, if H is an induced subgraph of $\mathbb{B}^n$ with $|V(H)| = 2^{n-1} + 1$,*

$$\Delta(H) \geq \sqrt{n}.$$

To see why Theorem 4.1 implies Theorem 1.1, we note that for any induced subgraph H of $\mathbb{B}^n$ with $|V(H)| \neq 2^{n-1}$, one of H and $\mathbb{B}^n \setminus H$ contains at least $2^{n-1} + 1$ vertices. Since $h(n) = \sqrt{n}$ is monotone, then the equivalence theorem of Gotsman and Linial [6] implies $s(f) \geq \sqrt{\deg(f)}$ for any Boolean function f. Recall that Theorem 3.3 states $bs(f) \leq \deg(f)^2$. Combining these two inequalities gives $bs(f) \leq s(f)^4$ for any Boolean function f, completing the proof of Theorem 1.1.

The proof of Theorem 4.1 in [9] rests upon the notion of *signed adjacency matrices* of an undirected graph. Given an undirected graph G with $|V(G)| = m$ (without loss of generality, we assume $V(G) = [m]$), a matrix $A \in \mathcal{M}_m(\{-1,0,1\})$ with entries a_{ij} for $i,j \in [m]$ is called a *signed adjacency matrix* of G when $a_{ij} = 0$ if and only if i,j are not adjacent in G. Moreover, given a symmetric matrix $A \in \mathcal{M}_m(\mathbb{R})$, we order the real eigenvalues of A such that $\lambda_1(A) \geq \ldots \geq \lambda_m(A)$, counting multiplicity. An immediate lemma lower-bounds the maximum degree of an undirected graph G by the largest eigenvalue of a signed adjacency matrix of G.

Lemma 4.2. *Let G be an m-vertex undirected graph. If A be a signed adjacency matrix of G, then*

$$\Delta(G) \geq \lambda_1(A).$$

Proof. Let $\lambda_1 = \lambda_1(A)$. Let $\mathbf{v} = [v_1 \ \ldots \ v_m]^T$ be an eigenvector of A corresponding to $\lambda_1(A)$. Then $A\mathbf{v} = \lambda_1 \mathbf{v}$. Let $j \in [m]$ be such that v_j is the component of $\mathbf{v}$ with the largest absolute value. Let the entries of A be a_{ij} for $i,j \in [m]$. By the triangle inequality,

$$|\lambda_1| \cdot |v_j| \leq \left|(\lambda_1 \mathbf{v})_j\right| = |(A\mathbf{v})_j| = \left|\sum_{i=1}^{m} a_{ji} v_i\right| \leq \sum_{i=1}^{m} |a_{ji} v_i| \leq \sum_{i=1}^{m} |a_{ji}| \cdot |v_j| \leq |v_j| \sum_{i=1}^{m} |a_{ji}|$$

$$= |v_j| \left(\sum_{i \in \delta_G(j)} 1 + \sum_{i \in V(G) \setminus \delta_G(j)} 0 \right) = |v_j| \cdot |\delta_G(j)| = |v_j| \cdot d_G(j) \leq |v_j| \, \Delta(G).$$

This implies that $\lambda_1 \leq |\lambda_1| \leq \Delta(G)$. □

Let A be a signed adjacency matrix of $\mathbb{B}^n$. If H is an induced subgraph of $\mathbb{B}^n$ with $|V(H)| = 2^{n-1} + 1$, then the principal submatrix A_H of A corresponding to H (i.e., obtained by keeping only the columns and the rows corresponding to the vertices in H) is a signed adjacency matrix of H. Hence, Cauchy's interlacing theorem together with Lemma 4.2 implies

$$\delta(H) \geq \lambda_1 (A_H) \geq \lambda_{2^n - (2^{n-1}+1)+1}(A) = \lambda_{2^{n-1}}(A).$$

Therefore, it suffices to find a signed adjacency matrix A of $\mathbb{B}^n$ with $\lambda_{2^{n-1}}(A) \geq \sqrt{n}$. Huang [9] gives such a construction.

Lemma 4.3. *Let*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \qquad A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}, n \geq 2.$$

Then the eigenvalues of $A_n \in \mathcal{M}_{2^n}(\mathbb{R})$ are $\sqrt{n}$ of multiplicity 2^{n-1} and $-\sqrt{n}$ of multiplicity 2^{n-1}.

Proof. We prove by induction on $n \in \mathbb{N}$ that $A_n^2 = nI$. For $n = 1$,

$$A_1^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

This proves the base case. Suppose that $A_{n-1}^2 = (n-1)I$ for some $n \in \mathbb{N}, n \geq 2$. Then

$$\begin{aligned} A_n^2 &= \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix} \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix} = \begin{bmatrix} A_{n-1}^2 + I^2 & A_{n-1}I - IA_{n-1} \\ IA_{n-1} - A_{n-1}I & I^2 + A_{n-1}^2 \end{bmatrix} \\ &= \begin{bmatrix} (n-1)I + I & 0 \\ 0 & I + (n-1)I \end{bmatrix} = \begin{bmatrix} nI & 0 \\ 0 & nI \end{bmatrix} = nI. \end{aligned}$$

This completes the induction step. Therefore, $A_n^2 = nI$ for all $n \in \mathbb{N}$. Let $n \in \mathbb{N}$. Since the eigenvalues of $I \in \mathcal{M}_n(\mathbb{R})$ are 1 of multiplicity n, then the eigenvalues of $nI \in \mathcal{M}_n(\mathbb{R})$ are n of multiplicity n. Therefore, the eigenvalues of $A_n \in \mathcal{M}_{2^n}(\mathbb{R})$ are either $\sqrt{n}$ or $-\sqrt{n}$. We have $\operatorname{tr} A_n = 0$ by construction. Note

$$\operatorname{tr} A_n = \sum \{\lambda : \lambda \text{ is an eigenvalue of } A_n, \text{ counting multiplicity}\}.$$

Hence, exactly half of the eigenvalues of A are $\sqrt{n}$, and the rest are $-\sqrt{n}$, completing the proof. □

This completes the proof of Theorem 4.1. Indeed, this theorem is stronger than the original sensitivity conjecture. Aaronson et al. [1] point out an additional implication of this stronger theorem that any randomized algorithm to guess the parity of an n-bit string, which succeeds with probability at least $\frac{2}{3}$ on the majority of strings, must make at least $\Omega(\sqrt{n})$ queries to the string, and any such quantum algorithm must make at least $\Omega(n^{1/4})$ queries.

5 A Quadratic Lower Bound

Huang [9] provides an upper bound of block sensitivity in terms of sensitivity. One can ask the following natural question: How large can the block sensitivity of a Boolean function be in terms of its sensitivity, asympototically? That is, what is the largest asympototic separation between block sensitivity and sensitivity? Rubinstein [13] provides an infinite family of Boolean functions f such that $bs(f) = \frac{1}{2}s(f)^2$, showing a quadratic separation between block sensitivity and sensitivity. This lower bound has been improved to $bs(f) = \frac{1}{2}s(f)^2 + \frac{1}{2}s(f)$ by Virza [15] and subsequently to $bs(f) = \frac{2}{3}s(f)^2 - \frac{1}{3}s(f)$ by Ambainis and Sun [2], but it remains quadratic. It is an open question to close the current gap between the quadratic lower bound and the quartic upper bound.

In this section, we present the Boolean function of Rubinstein [13] to prove a quadratic lower bound of block sensitivity in terms of sensitivity. We note that with more technical effort one can get a more precise bound $bs(f) = \frac{1}{2}s(f)^2$.

Theorem 5.1. *There exists an infinite family of Boolean functions f such that*

$$bs(f) = \Omega\left(s(f)^2\right).$$

Proof. For even $n \in \mathbb{N}$, let $f : \{0,1\}^{n^2} \to \{0,1\}$ be defined by

$$f\left(x_{11}, \ldots, x_{nn}\right) = \bigvee_{i=1}^{n} g\left(x_{i1}, \ldots, x_{in}\right),$$

where $g\left(x_1, \ldots, x_n\right) = 1$ if and only if there exists $j \in [n-1]$ such that $x_j = x_{j+1} = 1$ and $x_k = 0$ for all $k \in [n] \setminus \{j, j-1\}$. Intuitively, the input to f can be regarded as an $n \times n$ array, of which each row outputs 1 if and only if there exist two consecutive 1's only, and f outputs 1 if and only if at least one row outputs 1.

First, we show that $bs(f) = \Omega(n^2)$. We denote by $\vec{0} \in \{0,1\}^{n^2}$ the zero vector. Then $f(\vec{0}) = 0$. Note that the blocks $\{(i, 2j-1), (i, 2j)\}$ for $i \in [n], j \in [\frac{n}{2}]$ are disjoint and $f(\vec{0}^{\{(i,2j-1),(i,2j)\}} = 1$ for all $i \in [n], j \in [\frac{n}{2}]$. Hence, $bs(f) \geq bs(f, 0) = \frac{1}{2}n^2 = \Omega(n^2)$.

Second, we show that $s(f) = O(n)$. Let $x \in \{0,1\}^{n^2}$. If $f(x) = 0$, each row outputs 0. Note that there exist at most two indices of each row such that flipping each of these indices changes the output of the row, which case is attained when the row contains exactly one "internal" 1. On the other hand, if $f(x) = 1$, we have the following two subcases. If two rows output 1, then flipping a single index cannot change the value of f, so $s(f, x) = 0$. If only one row otuputs 1, then there are at most n indices such that flipping each index *may* change the value of f. This shows $s(f) \leq 2n = O(n)$. This completes the proof. □

6 Conclusion

In this short essay, we present in detail a spectral method for proving Theorem 1.1 which asserts a quartic upper bound of the block sensitivity of a Boolean function in terms of its sensitivity, confirming the long-standing sensitivity conjecture posed by Nisan and Szegedy [12]. In addition, we prove that Rubinstein [13]'s function gives a quadratic lower bound. Many related and interesting questions remain open in the field of Boolean functions, Boolean hypercubes and graph theory [9].

- For Boolean functions, as mentioned above, we have a quadratic lower bound and a quartic upper bound for block sensitivity in terms of sensitivity. It is an open question to close this gap. In particular, it is conjectured that $bs(f) \leq O(s(f)^2)$.

  Huang gives the following linear algebraic conjecture which implies $bs(f) \leq O(s(f)^2)$: A principal submatrix of A_n that includes rows and columns corresponding to pairwise disjoint sets $S_1, \ldots, S_t \subseteq [n]$, but not $\varnothing$, has an eigenvalue of absolute value at most $\sqrt{n-t}$.[1]

- For Boolean hypercubes, let $g(n, k)$ be defined as the minimum $t \in [2^n]$ such that every induced subgraph H of $\mathbb{B}^n$ with $|V(H)| = t$ has $\Delta(H) \geq k$. Huang [9] proves $g(n, \sqrt{n}) = 2^{n-1} + 1$. Determining $g(n, k)$ for all values of k would be interesting.

- More generally, let G be a "nice" undirected graph with high symmetry (e.g., a Boolean hypercube). We denote by $\alpha(G)$ the independence number of G, i.e., the largest size of an independent vertex set of G. Moreover, we denote by $f(G)$ the minimum of $\Delta(H)$ over all induced subgraphs H with $|V(H)| = \alpha(G) + 1$. It would be interesting to determine $f(G)$ for certain classes of undirected graphs G.

[1] See Huang's talk at Simons Institute, `https://www.youtube.com/watch?v=EJoe4qH6kLs`.

- The eigenvalues of the adjacency matrix of $\mathbb{B}^n$ are $-n, -n+2, \ldots, n-2, n$, each of which has multiplicity $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n-1}, \binom{n}{n}$, respectively. The "signing" of the adjacency matrix of a Boolean hypercube "compresses" the positive and negative eigenvalues in some sense. Formally defining such operations would be interesting and useful for other open problems in theoretical computer science and combinatorics.

Acknowledgment

The author would like to thank Professor Joel Friedman for providing this opportunity to explore this intriguing area of Boolean functions as a course project of CPSC 531F "Topics in Computer Science Theory: Applications of Linear Algebra" at the University of British Columbia.

References

[1] S. Aaronson, A. Ambainis, K. Balodis, and M. Bavarian. Weak parity. In *International Colloquium on Automata, Languages, and Programming*, pages 26–38. Springer, 2014.

[2] A. Ambainis and X. Sun. New separation between $s(f)$ and $bs(f)$. *arXiv preprint arXiv:1108.3494*, 2011.

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

[4] S. Fisk. A very short proof of cauchy's interlace theorem for eigenvalues of hermitian matrices. *The American Mathematical Monthly*, 112(2):118, 2005.

[5] 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, 2016.

[6] C. Gotsman and N. Linial. The equivalence of two problems on the cube. *Journal of Combinatorial Theory, Series A*, 61(1):142–146, 1992.

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

[8] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[9] H. Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals of Mathematics*, 190(3):949–955, 2019.

[10] S.-G. Hwang. Cauchy's interlace theorem for eigenvalues of hermitian matrices. *The American Mathematical Monthly*, 111(2):157–159, 2004.

[11] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

[12] N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of Computing*, pages 462–467, 1992.

[13] D. Rubinstein. Sensitivity vs. block sensitivity of boolean functions. *Combinatorica*, 15(2): 297–299, 1995.

[14] A. Tal. Properties and applications of boolean function composition. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 441–454, 2013.

[15] M. Virza. Sensitivity versus block sensitivity of boolean functions. *Information Processing Letters*, 111(9):433–435, 2011.