Measuring the Complexity of Boolean Functions and the Sensitivity Conjecture

Iden Kalemaj*

December 21, 2020

1 Introduction

The Sensitivity Conjecture of Nisan and Szegedy [NS94] states that two measures of the complexity of Boolean functions, namely *sensitivity* and *block sensitivity*, are polynomially related. This conjecture remained open for about thirty years until Hao Huang [Hua19] proved it in the positive with a remarkably short proof. It was known for a long time that block sensitivity was polynomially related to many other important measures of the complexity of Boolean functions [Nis91, NS94, BBC+01, Mid04, Tal13]. Until the result of Hao Huang, the sensitivity resisted such classification.

In this article we provide a proof of the Sensitivity Conjecture of [Hua19] which is self-contained with few exceptions. For an entirely self-contained survey of the results that comprise the proof of the Sensitivity Conjecture we refer the reader to [KSP20]. Our survey is much shorter, yet highlights the important contributions that went into this line of work. To familiarize the reader with measures of the complexity of Boolean functions, we provide some definitions and examples, while also stating some results on how these measures are related. Detailed surveys of such results can be found in [BdW02, HKP11]. Finally, towards the goal of making this exposition self-contained, we provide a summary of key facts from linear algebra and Fourier analysis of Boolean functions.

We consider Boolean functions $f\colon \{0,1\}^n \to \{0,1\}$ on the n-dimensional hypercube. Let Q_n denote the domain $\{0,1\}^n$. Sensitivity and block-sensitivity are two measures of the complexity of a function f. For $x \in Q_n$ and $S \subseteq [n]$, let x^S denote the point in Q_n obtained by flipping all of the bits x_i such that $i \in S$. If $S = \{i\}$ we will use the shorthand x^i instead of $x^{\{i\}}$.

Definition 1.1 (Sensitivity)**.** *The sensitivity of $f\colon \{0,1\}^n \to \{0,1\}$ at x, denoted $s_f(x)$, is the number of points y that differ from x in exactly one bit and satisfy $f(x) \neq f(y)$. The sensitivity $s(f)$ of f is the maximum over all $x \in Q_n$ of $s_f(x)$.*

Sensitivity measures how many bits of x have to be flipped in order to change the value of f. As an example, consider the OR-function $f(x) = x_1 \vee x_2 \vee \dots x_n$. Let $|x|$ denote the number of coordinates of x with value 1. If $f(x) = 0$, then $s_f(x) = n$, because changing any of the bits of x from 0 to 1 will change the function value. If $|x| \geq 2$, then $s_f(x) = 0$, because changing any single bit of x cannot change the function value. Therefore $s_f(x) = n$.

Definition 1.2 (Block Sensitivity)**.** *The block sensitivity of $f\colon \{0,1\}^n \to \{0,1\}$ at x, denoted $bs_f(x)$, is the maximum number of disjoint subsets of the coordinates $B_1, B_2, \dots, B_k \subseteq [n]$ such that $f(x) \neq f(x^{B_j})$ for all $j \in [k]$. The block sensitivity $bs(f)$ is the maximum value of $bs_f(x)$ over all $x \in Q_n$.*

Clearly, $s(f) \leq bs(f)$ since for each $x \in Q_n$ we can let each of the sets $B_1, B_2, \dots, B_k$ contain exactly one coordinate from the coordinates i such that $f(x) \neq f(x^i)$. To illustrate block sensitivity, we describe a

*Department of Computer Science, Boston University. Email: ikalemaj@bu.edu.

function of Rubinstein [Rub95] that gives a quadratic separation between $s(f)$ and $bs(f)$. This separation was later improved by [Vir11, AS11].

Example 1.3 ([Rub95]). *Let $n = k^2$, where k is even. Divide the n coordinates into $\sqrt{n} = k$ blocks $B_i, i \in [k]$ of k variables each, where $B_i = [(i-1)k+1, ik]$. Define $f(x) = 1$ if and only if there exists at least one block B_i and two consecutive coordinates $2j, 2j+1 \in B_i$ such that $x_{2j} = x_{2j+1} = 1$ and the remaining $k-2$ coordinates of x in B_i have value 0.*

Consider $x = 0^n$. Then $f(x) = 0$. We can define the disjoint sensitive sets of x to be the pairs of consecutive even-odd coordinates in each block B_i. Flipping any such pair of coordinates changes the value of f from 0 to 1. There are $k \cdot \frac{k}{2} = \frac{n}{2}$ such pairs. Thus $bs_f(0^n) = n/2$ and one can see that $bs(f) = n/2$. We argue that $s(f) = \sqrt{n}$. Consider x such that for each block B_i, $i \in [k]$, there is exactly one coordinate $2j \in B_i$ such that $x_{2j} = 1$ and all other coordinates of x in B_i are 0. Then $f(x) = 0$. For each block B_i, changing the coordinate $2j+1 \in B_i$ from 0 to 1 will change the value $f(x)$ from 0 to 1. Thus $s_f(x) = k = \sqrt{n}$, and one can see that $s(f) = \sqrt{n}$.

Hao Huang showed that sensitivity and block-sensitivity are polynomially-related.

Theorem 1.4 (Sensitivity Conjecture [Hua19]). *For every Boolean function $f\colon \{0,1\}^n \to \{0,1\}$, the sensitivity $s(f)$ and block-sensitivity $bs(f)$ satisfy:*

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

While the definition of block sensitivity might seem unusual, it was known for a long time that it is polynomially related to many other measures of the complexity of Boolean functions, such as decision tree depth, certificate complexity, and degree of the multilinear polynomial representing the Boolean function [Nis91, NS94, BBC+01, Mid04, Tal13]. These relations were known for deterministic, randomized, and quantum decision tree complexity. Sensitivity was the only measure which resisted classification until the result of [Hua19].

Huang [Hua19] proved the sensitivity conjecture by proving an equivalent result on induced subgraphs of the hypercube Q_n. For a graph G, we denote by $V(G)$ its vertex-set and $E(G)$ the set of edges (u, v) such that $u, v \in V(G)$ are adjacent. Viewed as a graph, Q_n contains 2^n vertices: the 0/1 vectors or strings of length n. For this graph, $(x, y) \in E(Q_n)$ if there exists some $i \in [n]$ such that $x_i \neq y_i$, and $x_i = y_j$ for all $j \in [n] \setminus \{i\}$.

Theorem 1.5 (Induced subgraphs [Hua19]). *Let H be an induced subgraph of the n-dimensional hypercube Q_n with $2^{n-1}+1$ vertices, and let $\Delta(H)$ be the maximum degree of H. Then*

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

Moreover, this inequality is tight when n is a perfect square.

Note that the subgraph of Q_n induced by the vertices with an even number of 1-coordinates has 2^{n-1} vertices and no edges, and thus has maximum degree 0. Adding just one extra vertex to this subgraph would make the maximum degree jump to $\sqrt{n}$. Theorem 1.5 is tight as shown by a construction of Chung, Füredi, Graham, and Seymour [CFGS88].

The equivalence between the problem of induced subgraphs of Q_n and the problem of complexity measures of Boolean functions was established by Gotsman and Linial [GL92]. For an induced subgraph H of Q_n, let $Q_n - H$ be the subgraph induced by the vertices $V(Q_n) \setminus V(H)$. Define $\Gamma(H)$ to be $\max(\Delta(H), \Delta(Q_n - H))$. For a Boolean function f, its degree $\deg(f)$ is the degree of the unique multilinear polynomial representing f (see Section 2 for how to obtain this polynomial via Fourier decomposition).

Theorem 1.6 (Equivalence theorem [GL92]). *The following are equivalent for any function $h\colon \mathbb{N} \to \mathbb{R}$.*

- **(a)** *For any induced subgraph H of Q_n with $|V(H)| \neq 2^{n-1}$, we have $\Gamma(H) \geq h(n)$.*
- **(b)** *For any Boolean function $f\colon Q_n \to \{0,1\}$, we have $s(f) \geq h(\deg(f))$.*

Huang applied Theorem 1.6 with the function $h(n) = \sqrt{n}$. Suppose H is an induced subgraph of Q_n with $V(H) \neq 2^{n-1}$. Then either H or $Q_n - H$ will have at least 2^{n+1} vertices. By Theorem 1.5, either $\Delta(H) \geq \sqrt{n}$ or $\Delta(Q_n - H) \geq \sqrt{n}$. We obtain $\Gamma(H) \geq \sqrt{n}$. As a corollary, we have:

Theorem 1.7 ([Hua19]). *For every Boolean function f,*

$$s(f) \geq \sqrt{\deg(f)}.$$

The final ingredient in the proof of Theorem 1.4 is a result of Nisan and Szegedy [NS94], strengthened by Tal [Tal13], who show that for every Boolean function $bs(f) \leq \deg(f)^2$. Combining this result with Theorem 1.7 concludes the proof of the Sensitivity Conjecture.

We remark that Theorem 1.7 is tight as demonstrated by the following folklore example.

Example 1.8 (AND-of-ORs). *Let f be a function on k^2 variables. Partition the coordinates in k blocks of k coordinates, so that block i contains coordinates $x_{i1}, x_{i2}, \ldots, x_{ik}$. Define*

$$f(x_{11}, x_{12}, \ldots, x_{kk}) = \bigwedge_{i=1}^{k} \bigvee_{j=1}^{k} x_{ij}.$$

Then f has degree k^2 and sensitivity k. If x is a lower endpoint of a sensitive edge, then there is a block i of coordinates $x_{i1}, x_{i2}, \ldots x_{ik}$ such that the coordinates of x are all 0 for that block, and x has at least one coordinate with value 1 in the other $k-1$ blocks. Then x is incident with exactly k sensitive edges, one for every coordinate in the block with all 0 coordinates.

However, the exponent of 4 in Theorem 1.4 may not be tight. Nisan and Szegedy suggested a quadratic upper bound, and the best separation between $s(f)$ and $bs(f)$ is quadratic, as we saw in Ex. 1.3.

Organization In Section 2 we recall a few concepts from linear algebra and Fourier analysis used in the proofs of the main theorems. We show a Fourier-analytic proof of Theorem 1.6 in Section 3. We prove Huangs's theorem on induced subgraphs (Theorem 1.5) in Section 4. In Section 5 we highlight some results on the connections between various measures of the complexity of Boolean functions.

2 Preliminaries

This section provides definitions and states results from Fourier analysis and linear algebra used in the proofs of the main theorems. The reader familiar with these concepts may skip the section.

2.1 Eigenvalues and Eigenvectors

Let's first recall the notion of vector spaces and dimension. A *vector space* V in $\mathbb{R}^n$ is a set of vectors $\mathbf{v} \in \mathbb{R}^n$ such that the all-zero vector is in V; for any two vectors $\mathbf{u}, \mathbf{v} \in V$, it also holds that $\mathbf{u} + \mathbf{v} \in V$; for any vector $\mathbf{v}$ and scalar $c \in \mathbb{R}$, it holds that $c\mathbf{v} \in V$. We denote all vectors by boldface lower-case letters. The all-zero vector is denoted by $\mathbf{0}$.

A set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d \in \mathbb{R}^n$ are called *linearly independent* if there do not exist $c_1, c_2, \ldots, c_d$ such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_d\mathbf{v}_d = \mathbf{0}$ and not all $c_1, c_2, \ldots, c_d$ are equal to zero. A set of linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d \in V$ is a *basis* for V if every vector in V can be expressed as a linear combination of the d vectors. The *dimension* of V is the size d of a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d$ for V.

Given a matrix $A \in \mathbb{R}^{n \times n}$, an *eigenvector* of A is a vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar $\lambda \in \mathbb{R}$. The scalar λ is *eigenvalue* corresponding to $\mathbf{v}$.

For an eigenvalue λ of A, let $V_\lambda = \{\mathbf{v} : A\mathbf{v} = \lambda\mathbf{v}\}$ be the set of eigenvectors corresponding to λ. Then V_λ is a vector space in $\mathbb{R}^n$, since:

- $A\mathbf{0} = \mathbf{0} = \lambda\mathbf{0}$, and so $\mathbf{0} \in V_\lambda$.

- If $\mathbf{u}, \mathbf{v} \in V_\lambda$ then $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \lambda\mathbf{u} + \lambda\mathbf{v} = \lambda(\mathbf{u}+\mathbf{v})$. Hence $\mathbf{u}+\mathbf{v} \in V_\lambda$.
- If $\mathbf{v} \in V_\lambda$ then $A(c\mathbf{v}) = cA\mathbf{v} = c\lambda\mathbf{v} = \lambda(c\mathbf{v})$, and so $c\mathbf{v} \in V_\lambda$.

The *multiplicity* of an eigenvalue λ is the dimension of the vector space V_λ.

The *trace* $\mathrm{tr}(A)$ of A is the sum of the diagonal entries of A, that is $\mathrm{tr}(A) = \sum_{i\in[n]} A_{ii}$. We make use of the following folklore fact about the trace.

Fact 2.1. *The trace of a matrix A equals the sum of its eigenvalues counted with multiplicity.*

2.2 Fourier Analysis of Boolean Functions on the Hypercube

We state some results needed for our discussion of the equivalence result of Gotsman and Linial (Theorem 1.6). In the discussion of Fourier analysis, it is convenient to represent Boolean functions as $f\colon \{-1,1\}^n \to \{-1,1\}$.

Fact 2.2. *Every function $f\colon \{-1,1\}^n \to \mathbb{R}$ can be expressed uniquely as a multilinear polynomial*

$$f(x) = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S(x),$$

where $\chi_S(x) = \prod_{i\in S} x_i$.

The coefficients $\widehat{f}(S)$ are the Fourier coefficients of f. The degree of f, $\deg(f)$, equals the size of the largest set $S \subseteq [n]$ such that $\widehat{f}(S) \neq 0$.

Fact 2.3. *The Fourier coefficients $\widehat{f}(S)$ of f are given by*

$$\widehat{f}(S) = \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)\chi_S(x)].$$

Fact 2.4. *For a function $f\colon \{-1,1\}^n \to \mathbb{R}$ it holds $\mathbb{E}_{x\sim\{-1,1\}^n}[f] = \widehat{f}(\emptyset)$.*

3 The Equivalence Theorem of Gotsman and Linial

In this section we prove the equivalence result of Gotsman and Linial for two problems on the hypercube as stated in Theorem 1.6 and restated below. Our discussion follows that of [KSP20], which differs slightly from the original proof of Gotsman and Linial.

Theorem 3.1 (Restated Theorem 1.6)**.** *The following are equivalent for any function $h\colon \mathbb{N} \to \mathbb{R}$.*

(**a**) *For any induced subgraph H of Q_n with $|V(H)| \neq 2^{n-1}$, we have $\Gamma(H) \geq h(n)$.*

(**b**) *For any Boolean function $f\colon Q_n \to \{0,1\}$, we have $s(f) \geq h(\deg(f))$.*

Proof of Theorem 1.6. We first show that (**b**) is equivalent to the the following:

(**b′**) For any Boolean function $f\colon Q_n \to \{0,1\}$ with $\deg(f) = n$ we have $s(f) \geq h(n)$.

Clearly (**b**) implies (**b′**). We show that (**b′**) $\Rightarrow$ (**b**).

Suppose f has $\deg(f) = d$, where $d \leq n$. Fix a monomial of degree d in the multilinear polynomial representation of f. Without loss of generality, assume this monomial is $x_1 \dots x_d$. Define $g(x_1 \dots x_d) = f(x_1 \dots x_d 00 \dots 0)$. Then $\deg(g) = d$. Consider an edge between $x = x_1x_2 \dots x_d$ and x^i for $i \in [d]$ in the hypercube Q_d. It corresponds to an edge between $x0\dots0$ and $x^i0\dots0$ in the hypercube Q_n. Note that $g(x) \neq g(x^i)$ if and only if $f(x0\dots0) \neq f(x^i0\dots0)$, by definition of g. It follows that $s(f) \geq s(g) \geq h(d)$.

(**a**) $\Rightarrow$ (**b′**). We prove the contrapositive. Suppose f is a Boolean function with $\deg(f) = n$ such that $s(f) < h(n)$. As an implication, we construct an induced subgraph H of Q_n with $|V(H)| \neq 2^{n-1}$ such that $\Gamma(H) < h(n)$.

The proof uses Fourier analysis, so for the remainder of the proof we assume that f is a Boolean function $f\colon \{-1,1\}^n \to \{-1,1\}$. Let $p(x)\colon \{-1,1\}^n \to \{-1,1\}$ be the parity function $p(x) = x_1x_2\ldots x_n$, so that $p(x) = 1$ if x has even parity and $p(x) = -1$ otherwise. The vertex set of H will be:

$$V(H) = \{x \in \{-1,1\}^n \mid f(x)p(x) = 1\}$$

H contains all the vertices such that the parity of that vertex equals the value of f. We show that $\widehat{f}(S) = \widehat{fp}([n] \setminus S)$ for all $S \subseteq [n]$. To see this, assume that $S = \{1, 2, \ldots, i\}$.

$$\begin{aligned}\widehat{fp}([n] \setminus S) = \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)p(x)\chi_{[n]\setminus S}(x)] &= \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)(x_1x_2\ldots x_n)(x_{i+1}x_{i+2}\ldots x_n)] \\ &= \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)x_1x_2\ldots x_i] \\ &= \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)\chi_S(x)] = \widehat{f}(S).\end{aligned}$$

The proof follows similarly for all S. As a consequence, $\widehat{f}([n]) = \widehat{fp}(\emptyset)$. Since $\deg(f) = n$, then $\widehat{f}(n) \neq 0$, by the Fourier decomposition of f. It follows that $\mathbb{E}_x[f(x)p(x)] = \widehat{fp}(\emptyset) \neq 0$ (see Fact 2.4). As a result, $|V(H)| \neq 2^{n-1}$. Note that all the edges in Q_n are between vertices with even parity and vertices with odd parity. The vertex x^i is adjacent to x in H if and only if $fp(x) = fp(x^i) = 1$, which happens if and only if $f(x) \neq f(x^i)$. Therefore $\deg_H(x) = s_f(x)$. By a similar argument, $\deg_{Q_n-H}(x) = s_f(x)$. It follows that $\Gamma(H) = s(f) < h(n)$. This concludes the proof.

$(\mathbf{b'}) \Rightarrow (\mathbf{a})$. To show this, note that all statements in the proof above also hold in the reverse direction. ■

4 Proof of the Sensitivity Conjecture

In this section we prove Theorem 1.5 which is the main result in the proof of the Sensitivity Conjecture. One key insight in the proof of Theorem 1.5 is that the maximum degree of a graph can be lower bounded by the largest eigenvalue of the adjacency matrix of that graph. Recall that the adjacency matrix A_H of an n-vertex graph H is an $n \times n$ symmetric matrix whose rows and columns are indexed by $V(H)$ and it has entries $A_{ij} = 1$ if $(i,j) \in E(H)$ and 0 otherwise. However, this insight alone is not enough, and Huang proves that the maximum degree of H is lower bounded by the largest eigenvalue of every matrix whose entries are obtained from A_H by flipping an arbitrary subset of the 1-entries to (-1). This is shown in Lemma 4.3.

Huang starts with the adjacency matrix of the hypercube graph Q_n and obtains a special matrix A_n by flipping a specific subset of the 1-entries to (-1). Crucially, the eigenvalues of A_n are precisely $\sqrt{n}$ and $-\sqrt{n}$, each of multiplicity 2^{n-1} (Lemma 4.2). For the induced subgraph H of Q_n, consider the submatrix A_H of A_n, which contains the rows and columns of A_n labeled by the vetices of H. Using the Cauchy Interlace Theorem (Lemma 4.1) one can bound the eigenvalues of A_H by the eigenvalues of A_n to finally obtain a bound on $\Delta(H)$.

In this section we also present a proof of Theorem 1.5 which does not use the Cauchy Interlace Theorem, and instead uses basic linear algebra to bound the largest eigenvalue of A_H. For that reason, we do not present a proof of the Interlace Theorem, and encourage the reader with rusty linear algebra to review Section 2.

Lemma 4.1 (Cauchy's Interlace Theorem)**.** *Let A be a symmetric $n \times n$ matrix and let B be a $m \times m$ submatrix of A, for some $m < n$, obtained by deleting the same set of rows and columns from A. Let $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ denote the eigenvalues of A, and let $\lambda_1(B) \geq \lambda_2(A) \geq \cdots \geq \lambda_m(B)$ be the eigenvalues of B. For all $i \in [m]$ we have*

$$\lambda_i(A) \geq \lambda_i(B) \geq \lambda_{i+n-m}(A).$$

Lemma 4.2. *Define a sequence of symmetric square matrices iteratively as follows,*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}.$$

Then A_n is a $2^n \times 2^n$ matrix with eigenvalues $\sqrt{n}$ and $-\sqrt{n}$, each of multiplicity 2^{n-1}.

Proof. We show that $A_n^2 = nI$ via an induction argument on n. This clearly holds for $n = 1$. By the induction hypothesis,

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

Thus, the claim is true for all $n \geq 1$.

If λ is an eigenvalue of some matrix B with respective eigenvector $\mathbf{v}$, then λ^2 will be an eigenvalue of B^2, since:

$$B^2\mathbf{v} = B(B\mathbf{v}) = \lambda B\mathbf{v} = \lambda^2\mathbf{v}.$$

The only eigenvalue of nI is n, since $nI\mathbf{v} = n\mathbf{v}$. Since A_n^2 has eigenvalue n of multiplicity n, then the eigenvalues of A_n will be either $\sqrt{n}$ or $-\sqrt{n}$. But note that the trace of A_n is 0 (this can be easily seen by the construction, or more formally shown by induction). The trace of a matrix equals the sum of its eigenvalues (counted with multiplicity). A_n is invertible (with inverse $\frac{1}{n}A_n$), so 0 is not an eigenvalue of A_n. It follows that A_n must have eigenvalues $\sqrt{n}$ and $-\sqrt{n}$ each of multiplicity 2^{n-1}. ■

Lemma 4.3. *Let H be a graph of m vertices. Let A be a symmetric matrix with values in $\{-1, 0, 1\}$ whose rows and columns are indexed by $V(H)$. Suppose $A_{u,v} = 0$ if u and v are not adjacent in H, and $A_{u,v} \in \{-1, 0, 1\}$ otherwise. Let λ_1 denote the largest eigenvalue of A. Then:*

$$\Delta(H) \geq |\lambda_1|. \tag{1}$$

Proof. Suppose $\mathbf{v} \in \mathbb{R}^m$ is the corresponding eigenvector for λ_1 so that $A\mathbf{v} = \lambda_1\mathbf{v}$. Let v_i be the coordinate of $\mathbf{v}$ with the largest absolute value, where $i \in [m]$. Then:

$$|\lambda_1 v_i| = |(A\mathbf{v})_i| = \left|\sum_{j=1}^{m} A_{ij}v_j\right| \leq \sum_{j=1}^{m} |A_{ij}||v_j|.$$

Note that $|A_{ij}| = 0$ whenever i and j are not adjacent in H, and in general $|A_{ij}| \leq 1$. Therefore:

$$\sum_{j=1}^{m} |A_{ij}||v_j| = \sum_{j \text{ s.t. } (i,j)\in E(H)} |A_{ij}||v_j| \leq \sum_{(i,j)\in E(H)}^{m} |v_j| \leq \Delta(H)|v_i|.$$

It follows that $\Delta(H) \geq |\lambda_1|$. ■

A first proof. With these lemmas in hand, we present the first proof of Theorem 1.5, restated below, which uses the Cauchy Interlace Theorem.

Theorem 4.4 (Restated Theorem 1.5)**.** *Let H be an induced subgraph of the n-dimensional hypercube Q_n with $2^{n-1} + 1$ vertices, and let $\Delta(H)$ be the maximum degree of H. Then*

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

Moreover, this inequality is tight when n is a perfect square.

Proof of Theorem 1.5. Let A_n be defined in Lemma 4.2. Consider another sequence of matrices defined as:

$$A_1' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_n' = \begin{bmatrix} A_{n-1}' & I \\ I & A_{n-1}' \end{bmatrix}. \tag{2}$$

Observe that A_n' is the adjacency matrix of the hypercube graph Q_n. One can see this by induction, considering two subcubes of Q_n, the one for which $x_n = 0$ and the one for which $x_n = 1$. Then A_{n-1}' is the adjacency matrix for each subcube, and the identity matrix I corresponds to the matching between the two subcubes, i.e. the edges from x to x^n. Note also that A'_n can be obtained from A_n by flipping all the (-1) values of A_n to 1. Therefore the matrix A_n and the graph Q_n satisfy the conditions of Lemma 4.3.

For an induced subgraph H of Q_n with $(2^{n-1}+1)$-vertices, let A_H be the submatrix of A_n obtained by only considering the columns and rows of A_n indexed by $V(H)$. Then A_H and H also satisfy the conditions of Lemma 4.3. As a result,

$$\Delta(H) \geq |\lambda_1(A_H)|. \tag{3}$$

From Lemma 4.2, the eigenvalues of A_n are $\sqrt{n}$ and $-\sqrt{n}$, each of multiplicity 2^{n-1}. The matrix A_H is a $(2^{n-1}+1) \times (2^{n-1}+1)$ submatrix of A_n. By the Cauchy Interlace Theorem:

$$\lambda_1(A_H) \geq \lambda_{1+2^n-2^{n-1}-1}(A_n) = \lambda_{2^{n-1}}(A_n) = \sqrt{n}. \tag{4}$$

It follows that $\Delta(H) \geq \sqrt{n}$ which completes the proof. ■

Alternate proof. It is possible to prove Theorem 1.5 using more basic linear algebra and without the Cauchy Interlace Theorem. The following proof formalizes the comment of Shalev Ben-David in blog post [Aro19].

Alternate Proof of Theorem 1.5. Let A_n be the matrix from Lemma 4.2 and let A_H be the submatrix obtained from A by considering only the rows and columns indexed by the vertices of a $(2^{n-1}+1)$-vertex induced subgraph H of Q_n.

We let $N = 2^n$ for ease of presentation. The matrix A_n is an $N \times N$ matrix that has an eigenvalue $\sqrt{n}$ with multiplicity $N/2$. Let U be the subspace of $\mathbb{R}^N$ consisting of all eigenvectors of A_n with eigenvalue $\sqrt{n}$, i.e.:

$$U = \{\mathbf{u} \in \mathbb{R}^N \mid A_n\mathbf{u} = \sqrt{n}\mathbf{u}\}$$

Then U has dimension $N/2$ by the multiplicity of the eigenvalue $\sqrt{n}$. Let $S = [N] \setminus V(H)$ denote the indices of the rows and the columns of A_n which do not belong to A_H (i.e. the rows and columns we remove from A_n to obtain A_H). For a vector $\mathbf{v} \in \mathbb{R}^N$, let $\mathbf{v}_{|S}$ denote the subvector obtained by removing entries of $\mathbf{v}$ whose indices are in S. Furthermore, consider the subspace V of vectors whose S-indices are 0:

$$V = \{\mathbf{v} \in \mathbb{R}^N \mid \mathbf{v}_i = 0 \text{ for all } i \in S\}.$$

Observe that for a vector $\mathbf{v} \in V$ we have $A_n\mathbf{v} = A_H\mathbf{v}_{|S}$. Therefore, if $\mathbf{v} \in V$ is an eigenvector of A_n then $\mathbf{v}_{|S}$ is an eigenvector of A_H with the same eigenvalue.

The subspace V has dimension $N/2+1$, since $|S| = N/2-1$. The dimension of U is $N/2$ and the dimension of $\mathbb{R}^N$ is N. It follows that $U \cap V$ has dimension at least 1, for $U \cap V$ is a subspace itself, and if it had dimension 0 we would obtain an $(N+1)$-dimensional subspace of $\mathbb{R}^N$ from the union of U and V. It follows that there is at least one non-zero vector $\mathbf{v} \in U \cap V$. By our argument, $\mathbf{v}_{|S}$ is an eigenvector of A_H with eigenvalue $\sqrt{n}$. It follows that $\lambda_1(H) \geq \sqrt{n}$.

■

5 Measuring the Complexity of Boolean Functions

Recall the definitions of degree, sensitivity, and block sensitivity. As we have mentioned, they are polynomially related to two measures of function complexity: decision tree depth and certificate complexity. In particular, decision tree depth is a very natural measure of computational complexity, since decision trees give an algorithm for computing the function. Some relationships between these measures are straightforward and we comment on them. We then summarize, without proofs, the known polynomial relations between these measures.

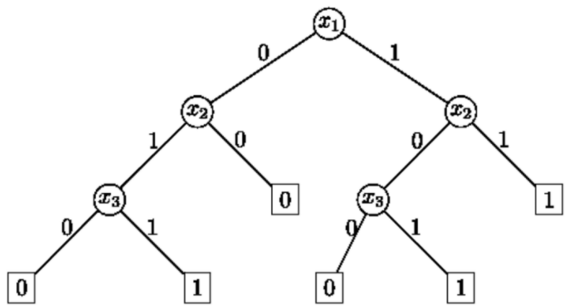

Figure 1: A decision tree for computing the majority function on 3 variables [KSP20]
.

Decision tree depth. Consider an algorithm for computing the value of $f\colon Q_n \to \{0,1\}$ at input x. The algorithm makes queries to the input x by asking for the value of x at index $i \in [n]$. We can represent such an algorithm as a rooted binary decision tree, whose nodes are labeled by variables x_i and edges are labeled by the values 0 or 1 of the corresponding variable. The leaf nodes are the values of the function f at the input x whose bits are indicated by the edge labels for the path ending at that leaf. See Fig. 1 for an example of a decision tree for the majority function on 3 variables.

Given a decision tree A_f for computing f, let $D(A_f)$ denote the depth of A_f, which is the longest path from the root to a leaf in the decision tree. The depth gives the worst-case query complexity of the algorithm A_f. There can be different decision trees for computing the function f. One way to quantify the complexity of computing f is to consider the depth of the optimal decision tree for computing f. The decision tree complexity of f is defined as:

$$D(f) = \min_{A_f} D(A_f).$$

We focus on deterministic decision trees, although many results in this article have randomized and quantum analogues.

To see that $s(f) \le bs(f) \le D(f)$, let $B_1, B_2, \ldots, B_{bs_f(x)}$ be disjoint sensitive blocks for $x \in \{0,1\}^n$. On input x, the decision tree must query at least one bit in each B_i, for otherwise we could flip the values of the bits of x in the unqueried block B_i and change the value of f so that the decision tree would not notice such change.

It is not too hard to show that we can obtain a multilinear polynomial for f from the decision tree A_f that computes f with degree at most the depth of A_f. Therefore $\deg(f) \le D(f)$.

Certificate complexity. The certificate complexity $C(f)$ of f measures how many of the n variables have to be assigned a value in order to fully determine the value of f. For a set $S \subseteq [n]$, let $x_{|S}$ denote the string in $\{0,1\}^{|S|}$, obtained by removing from x all bits x_i such that $i \notin S$. Then $C_f(x)$ is the minimum size of a set S such that $x_{|S} = y_{|S}$ implies $f(x) = f(y)$. This means that once the bits of x in S are revealed we can determine the value $f(x)$. The certificate complexity is defined as $C(f) = \max_x C_f(x)$.

Consider the OR-function $f(x) = x_1 \vee x_2 \cdots \vee x_n$. If $f(x) = 1$, then $C_f(x) = 1$. because revealing one of the bits i such that $x_i = 1$ we can fix $f(x) = 1$. If $f(x) = 0$ then $C_f(x) = n$, because we need to reveal all of the bits of x to fix $f(x) = 0$. Therefore $C(f) = n$.

Note that $C(f) \le D(f)$. Let S be a minimal certificate for x. On input x, if the decision tree does not query $|S|$ bits of x, then it cannot fix the value of $f(x)$.

For all $x \in \{0,1\}^d$, we have $s_f(x) \le bs_f(x) \le C_f(x)$, since a certificate for x must contain at least one bit from each sensitive block for x. Therefore $s(f) \le bs(f) \le C(f)$. On the other hand, Nisan [Nis91] showed that $C(f) \le s(f)bs(f)$. To see this, consider disjoint sensitive blocks $B_1, B_2, \ldots, B_b$ for x such that $b = bs_f(x)$. It is not too hard to show that $|B_i| \le s(f)$ for all $i \in [b]$ and $S = \cup B_i$ is a certificate for x.

Polynomial relations We summarize relations known between these complexity measures. Each line gives upper bounds for one measure.

1. $bs(f) \leq s(f)^4$ [Hua19], $bs(f) \leq C(f)$, $bs(f) \leq D(f)$, $bs(f) \leq \deg(f)^2$ [Tal13].
2. $s(f) \leq bs(f) \leq D(f)$, $s(f) \leq C(f)$, $s(f) \leq \deg(f)^2$ [Tal13].
3. $D(f) \leq s(f)bs(f)^2 \leq bs(f)^3$ [Nis91], $D(f) \leq \deg(f)^3$ [Mid04], $D(f) \leq C(f)^2$ [Mid04].
4. $C(f) \leq s(f)b(f) \leq bs(f)^2$ [Nis91], $C(f) \leq D(f) \leq \deg(f)^3$ [Mid04].
5. $\deg(f) \leq s(f)^2$ [Hua19], $\deg(f) \leq D(f) \leq C(f)^2$ [Mid04].

References

[Aro19] Scott Aronson. Sensitivity conjecture resolved, Jul 2019.

[AS11] Andris Ambainis and Xiaoming Sun. New separation between *s(f)* and *bs(f)*. *Electron. Colloquium Comput. Complex.*, 18:116, 2011.

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.

[BdW02] Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.

[CFGS88] Fan R. K. Chung, Zoltán Füredi, Ronald L. Graham, and Paul D. Seymour. On induced subgraphs of the cube. *J. Comb. Theory, Ser. A*, 49(1):180–187, 1988.

[GL92] Craig Gotsman and Nathan Linial. The equivalence of two problems on the cube. *J. Comb. Theory, Ser. A*, 61(1):142–146, 1992.

[HKP11] Pooya Hatami, Raghav Kulkarni, and Denis Pankratov. Variations on the sensitivity conjecture. *Theory Comput.*, 4:1–27, 2011.

[Hua19] Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals of Mathematics*, 190:949–955, 2019.

[KSP20] Rohan Karthikeyan, Siddharth Sinha, and Vallabh Patil. On the resolution of the sensitivity conjecture. *Bulletin of the American Mathematical Society*, 57:615–638, 2020.

[Mid04] Gatis Midrijanis. Exact quantum query complexity for total boolean functions. *Technical report, Cornell University ArXiv.org*, pages 615–638, 2004.

[Nis91] Noam Nisan. CREW prams and decision trees. *SIAM J. Comput.*, 20(6):999–1007, 1991.

[NS94] Noam Nisan and Mario Szegedy. On the degree of boolean functions as real polynomials. *Comput. Complex.*, 4:301–313, 1994.

[Rub95] David Rubinstein. Sensitivity vs. block sensitivity of boolean functions. *Comb.*, 15(2):297–299, 1995.

[Tal13] Avishay Tal. Properties and applications of boolean function composition. In *Innovations in Theoretical Computer Science, ITCS '13, Berkeley, CA, USA, January 9-12, 2013*, pages 441–454, 2013.

[Vir11] Madars Virza. Sensitivity versus block sensitivity of boolean functions. *Inf. Process. Lett.*, 111(9):433–435, 2011.