Stability of Linear Threshold Functions

Sivakanth Gopi

B. Tech. Project

Advisors:
Prof. Prahladh Harsha (TIFR) Prof. Srikanth Srinivasan (IITB)

Department of Computer Science and Engineering
Indian Institute of Technology, Bombay

2013

Abstract

Boolean Half-spaces or Linear Threshold Functions (LTFs) are an important class of Boolean functions which come up in Learning Theory (Perceptrons), Social Choice Theory etc.. We will study lower bounds on the stability of LTFs.

Firstly, we look at the problem of finding the best lower bound for the total level-0 and level-1 weight of LTFs, denoted by $W^1[LTF]$, which is conjectured to be $\frac{2}{\pi}$. We will present the well-known bound of $W^1[LTF] \geq \frac{1}{2}$ [GL94], and then look at some recent progress [DDS12] which shows that $W^1[LTF] \geq \frac{1}{2} + c$ for some $c > 0$.

We then look at a stronger conjecture from [BKS98] which claims that Maj_n is the least stable among all LTFs on n(odd) variables. We give a counter example to this conjecture. We conjecture that for any LTF f, $\mathrm{Stab}_\rho[f] \geq \frac{2}{\pi} \arcsin \rho \; \forall \rho \in [0, 1]$ and then give evidence for this new conjecture.

Acknowledgements

I would like to thank my advisors Prof. Prahladh Harsha and Prof. Srikanth Srinivasan. Both of them were wonderful advisors; kind, patient and helpful. I have learnt a lot of things from them. Thanks to Srikanth for all those amazing discussions during our numerous journeys together. I also thank Swagato Sanyal for many stimulating discussions. I would like to thank Prof. Sivaramakrishnan Sivasubramanian for kindly accepting to be my examiner.

Thanks to the Department of Mathematics, IIT Bombay and School of Technology and Computer Science, TIFR for facilitating our discussions.

Contents

Chapter 1

Fourier Analysis of Boolean Functions

1.1 Boolean Functions and Boolean Half-spaces

Definition 1. *$f : \{-1,1\}^n \to \{-1,1\}$ are called Boolean functions. $f : \{-1,1\}^n \to \mathbb{R}$ are called real-valued functions on the Boolean hypercube.*

Definition 2. *$f : \{-1,1\}^n \to \{-1,1\}$ given by $f(x) = \operatorname{sgn}(l(x))$ where $l(x) = a_0 + a_1x_1 + \cdots + a_nx_n$ are called Boolean Half-spaces or Linear Threshold Functions (LTFs). We assume $l(x) \neq 0 \ \forall x$. If $a_0 = 0$, f is called a balanced LTF*

For example $\operatorname{Maj}_n(x) = \operatorname{sgn}(x_1 + \cdots + x_n)$ when n is odd is a balanced LTF called the majority on n-bits*.

Geometrically, LTFs are partitions of the hypercube by a hyperplane†. All the points in one half-space are labelled 1 and those in the other half-space are labelled -1 as shown in Figure 1.1. Balanced half-spaces have hyperplanes passing through the origin.

1.2 Fourier Analysis

1.2.1 Inner Product Space

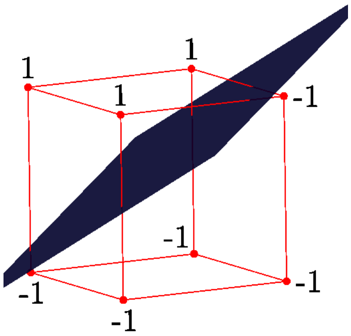
The set of all real-valued functions on the hypercube $\{-1,1\}^n$ forms a vector space $\mathcal{V}_n$ with field $\mathbb{R}$. We will define an inner product in this space as follows:

Definition 3. *Let $f, g \in \mathcal{V}_n$, $\langle f, g\rangle := \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n}[f(x)g(x)]$ where the expectation is over the uniform distribution on the hypercube*

* whenever Maj_n is mentioned, it is assumed that n is odd

†which we assume doesn't pass through any point on the hypercube

Figure 1.1: Partition of a hypercube by a hyperplane



Definition 4. *$\chi_S(x) := \prod_{i\in S} x_i$ where $S \subset [n]$, these are called Parity Functions on the hypercube.*

Proposition 1. *$\{\chi_S : S \subset [n]\}$ is an orthonormal basis for $(\mathcal{V}_n, \langle\cdot,\cdot\rangle)$*

Proof.

$$\langle \chi_S, \chi_T\rangle = \mathop{\mathbb{E}}_x[\prod_{i\in S} x_i \prod_{j\in T} x_j] = \mathop{\mathbb{E}}_x[\prod_{k\in S\Delta T} x_k \prod_{l\in S\cap T} x_l^2] = \mathop{\mathbb{E}}_x[\prod_{k\in S\Delta T} x_k] = \mathbb{1}_{S=T}$$

Thus $\{\chi_S : S \subset [n]\}$ is an orthonormal set and thus independent. Also,

$$|\,\{\chi_S : S \subset [n]\}\,| = 2^n = \dim(\mathcal{V}_n)$$

Thus $\{\chi_S : S \subset [n]\}$ is an orthonormal basis for $\mathcal{V}_n$ □

Since the parity functions form a basis, every function can be written uniquely as their linear combination.

Corollary 1. *Every $f : \{-1,1\}^n \to \mathbb{R}$ can be uniquely expressed as a multilinear polynomial $f(x) = \sum_{S\subset[n]} \hat{f}(S)\chi_S(x)$ where $\hat{f}(S) \in \mathbb{R}$*

$\left\{\hat{f}(S) : S \subset [n]\right\}$ are called the Fourier coefficients of f.

The Fourier coefficient $\hat{f}(S) = \langle f, \chi_S\rangle = \mathop{\mathbb{E}}_x[f(x)\chi_S(x)]$ is the correlation between f and χ_S.

For example $\mathrm{Maj}_3(x)$ can be written as a multilinear polynomial uniquely as,

$$\mathrm{Maj}_3(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_2x_3$$

Thus $\widehat{\mathrm{Maj}_3}(\emptyset) = 0, \widehat{\mathrm{Maj}_3}(\{1\}) = \widehat{\mathrm{Maj}_3}(\{2\}) = \widehat{\mathrm{Maj}_3}(\{3\}) = \frac{1}{2}, \widehat{\mathrm{Maj}_3}(\{1,2\}) = \widehat{\mathrm{Maj}_3}(\{2,3\}) = \widehat{\mathrm{Maj}_3}(\{3,1\}) = 0, \widehat{\mathrm{Maj}_3}(\{1,2,3\}) = -\frac{1}{2}$, The following is a simple proposition about Fourier coefficients of even and odd functions:

Proposition 2. *If f is an even function i.e. $f(-x) = f(x)$, then $\hat{f}(S) = 0$ whenever $|S|$ is odd. Similarly if f is an odd function i.e. $f(-x) = -f(x)$, then $\hat{f}(S) = 0$ whenever $|S|$ is even.*

Proof. $\chi_S(x)$ is odd(even) iff $|S|$ is odd(even). Thus $f(x)\chi_S(x)$ is odd in both the given cases. Thus

$$\hat{f}(S) = \mathop{\mathbb{E}}_x[f(x)\chi_S(x)] = \mathop{\mathbb{E}}_x[f(-x)\chi_S(-x)] = -\mathop{\mathbb{E}}_x[f(x)\chi_S(x)] = -\hat{f}(S)$$

Hence $\hat{f}(S) = 0$

□

For example, balanced LTFs are odd functions. So their Fourier coefficients at even levels are 0.

We will also introduce the following norms on $\mathcal{V}_n$.

Definition 5. $\|f\|_p := \mathop{\mathbb{E}}_x[|f(x)|^p]^{\frac{1}{p}}$ *for $p > 0$, in particular* $\|f\|_2 = \langle f, f\rangle^{\frac{1}{2}}$

1.2.2 Parseval and Level-k Weights

Proposition 3 (Plancherel Identity)**.** *Let $f, g \in \mathcal{V}_n$ then $\langle f, g\rangle = \sum_S \hat{f}(S)\hat{g}(S)$*

Proof. $\langle f, g\rangle = \sum_S \sum_T \hat{f}(S)\hat{g}(T)\langle \chi_S, \chi_T\rangle = \sum_S \hat{f}(S)\hat{g}(S)$ □

Corollary 2 (Parseval Identity)**.** *If $f : \{-1,1\}^n \to \{-1,1\}$ then $\sum_S \hat{f}(S)^2 = 1$*

Proof. $\sum_S \hat{f}(S)^2 = \langle f, f\rangle = \mathop{\mathbb{E}}_x[f^2] = 1$ □

Parseval identity implies that for Boolean functions, $\hat{f}(S)^2$ gives a probability distribution on subsets of $[n]$. This distribution is called the spectral distribution $\mathcal{S}_f$.

Definition 6. *Let $f \in \mathcal{V}_n$. $\hat{f}(S)^2$ is called the weight of f at S. The level-k weight of f, $W^k[f] := \sum_{|S|=k} \hat{f}(S)^2$. Similarly $W^{\leq k}[f] := \sum_{|S|\leq k} \hat{f}(S)^2$ and so on.*

Parseval can be rephrased as, for Boolean functions f, $\sum_{i=0}^n W^i[f] = 1$ i.e. their total weight 1 is distributed among the $n+1$ layers

Boolean functions having most of their weight concentrated on the lower layers are usually simple and easy to learn. Theorem 1 makes this precise.

Theorem 1. *(Low-Degree Algorithm) If $f : \{-1,1\}^n \to \{-1,1\}$ and $W^{>k} < \epsilon/2$ then f can be learned with ϵ-error in time* $\text{Poly}(n^k, 1/\epsilon)$ *using random examples $(x, f(x))$ where x is sampled from uniform distribution on $\{-1,1\}^n$ i.e. we can find $g : \{-1,1\}^n \to \{-1,1\}$ which is given as the sign of a degree k multilinear polynomial*[‡] *such that $\underset{x}{Pr}[f(x) \neq g(x)] < \epsilon$*

1.3 Stability

Definition 7. *Let $\rho \in [-1,1]$, $x, y \in \{-1,1\}^n$. We say (x,y) is a ρ-correlated pair of random strings if the pairs of random bits $\{(x_i, y_i) : i \in [n]\}$ are mutually independent and $\mathbb{E}[x_i] = \mathbb{E}[y_i] = 0, \mathbb{E}[x_i y_i] = \rho$. We also write this as $y \sim N_\rho(x)$*

When $\rho \in [0,1]$, this is equivalent to choosing $x \sim \{-1,1\}^n$ and then

$$y_i = \begin{cases} x_i & \text{w.p } \rho \\ 1 & \text{w.p } \frac{1-\rho}{2} \\ -1 & \text{w.p } \frac{1-\rho}{2} \end{cases}$$

So you can look at $y \sim N_\rho(x)$ as y being a noisy version of x where the noise is parametrized by ρ

Definition 8. $\text{Stab}_\rho[f] := \underset{(x,y):\rho-correlated}{\mathbb{E}}[f(x)f(y)]$

You can think of stability as the correlation between f and a noisy version of f. When f is a Boolean function,

$$\text{Stab}_\rho[f] = 1 - 2 \underset{(x,y):\rho-correlated}{Pr}[f(x) \neq f(y)]$$

We now give a Fourier based formula for stability.

Proposition 4. $\text{Stab}_\rho[f] = \sum_S \rho^{|S|} \hat{f}(S)^2 = \sum_{i=0}^n \rho^i W^i[f]$

Proof.

$$\begin{aligned} \underset{(x,y):\rho-correlated}{\mathbb{E}}[\chi_S(x)\chi_T(y)] &= \underset{(x,y):\rho-correlated}{\mathbb{E}}[\prod_{i \in S \cap T} x_i y_i \prod_{j \in S \setminus T} x_j \prod_{k \in T \setminus S} y_k] \\ &= \rho^{|S|} \mathbb{1}_{S=T} \end{aligned}$$

$$\begin{aligned} \text{Stab}_\rho[f] = \underset{(x,y):\rho-correlated}{\mathbb{E}}[f(x)f(y)] &= \underset{(x,y):\rho-correlated}{\mathbb{E}}[\sum_S \sum_T \hat{f}(S)\hat{f}(T)\chi_S(x)\chi_T(y)] \\ &= \sum_S \sum_T \hat{f}(S)\hat{f}(T) \underset{(x,y):\rho-correlated}{\mathbb{E}}[\chi_S(x)\chi_T(y)] \\ &= \sum_S \rho^{|S|} \hat{f}(S)^2 \end{aligned}$$

[‡] g can be calculated at any point in $\mathcal{O}(n^k)$ time

□

1.3.1 Stability of Majority

Theorem 2 (Multidimensional Central Limit Theorem [Tao]). *Let $X_1, \cdots, X_n$ be i.i.d random vectors in $\mathbb{R}^m$ with mean 0 (i.e. $\mathbb{E}[X_k] = 0$) and covariance matrix Σ (i.e. $\Sigma_{ij} = \mathbb{E}[X_{1i}X_{1j}]$) then $\frac{1}{\sqrt{n}} \sum_k X_k \underset{distribution}{\longrightarrow} \mathcal{N}(0, \Sigma)$*[§]

We will now calculate the stability of Maj_n for large n.

Proposition 5. $\mathrm{Stab}_\rho[\mathrm{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{2}{\pi} \arcsin \rho$

Proof.

$$\begin{aligned}
\mathrm{Stab}_\rho[\mathrm{Maj}_n] &= 1 - 2 \underset{(x,y):\rho-\text{correlated}}{Pr}[\mathrm{Maj}_n(x) \neq \mathrm{Maj}_n(y)] \\
&\underset{n\to\infty}{\longrightarrow} 1 - 2 \underset{(Z_1,Z_2):\rho-\text{correlated gaussians}}{Pr}[\mathrm{sgn}(Z_1) \neq \mathrm{sgn}(Z_2)] \ (2-\text{Dimensional CLT}) \\
&= \frac{2}{\pi} \arcsin \rho = \frac{2}{\pi}(\rho + \frac{1}{6}\rho^3 + \frac{3}{40}\rho^5 + \cdots)
\end{aligned}$$

□

From this we get $W^1[\mathrm{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{2}{\pi}, W^3[\mathrm{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{1}{3\pi}$ and so on.

1.4 Derivatives & Laplacian

We will define linear operators on the space $\mathcal{V}_n$ called the derivative and Laplacian operators.

Definition 9. *$D_i f(x) := \frac{1}{2}[f(x^{i\to 1}) - f(x^{i\to -1})]$ where $x^{i\to 1}$ is vector x with i^{th} bit fixed to 1 and similarly $x^{i\to -1}$. D_i is linear and is called the i^{th}-directional derivative.*

D_i acts like a formal derivative on polynomials.

Proposition 6. $D_i f(x) = \sum_{S\ni i} \hat{f}(S)\chi_{S\setminus\{i\}}(x)$

Proof.

$$D_i \chi_S = \begin{cases} 0 & \text{if } i \notin S \\ \chi_{S\setminus\{i\}} & \text{if } i \in S \end{cases}$$

The rest follows from linearity of D_i □

Definition 10. *$L_i f(x) := \frac{1}{2}[f(x) - f(x^{\oplus i})] = x_i D_i f(x)$ where $x^{\oplus i}$ is the vector x with its i^{th} bit flipped. L_i is linear and is called i^{th}-directional Laplacian. $L = \sum_{i=1}^n L_i$ is called the Laplacian.*

[§] Normal distribution with mean 0 and covariance matrix Σ

Proposition 7.

- $L_i f(x) = \sum_{S \ni i} \hat{f}(S)\chi_S(x)$
- $Lf(x) = \sum_S |S|\hat{f}(S)\chi_S(x)$
- $\langle f, Lf \rangle = \sum_S |S|\hat{f}(S)^2 = \sum_{i=1}^n iW^i[f]$

Proof.

$$L_i\chi_S = \begin{cases} 0 & \text{if } i \notin S \\ \chi_S & \text{if } i \in S \end{cases}$$

The rest follows from linearity of L_i and L □

The Laplacian as defined above is related to the graph Laplacian. Recall the graph Laplacian,

$$\langle f, Lf \rangle = \sum_{(u,v)\in E} (f_u - f_v)^2$$

which is the total variance across the edges.

In our definition,

$$\langle f, Lf \rangle = \frac{n}{4} \mathop{\mathbb{E}}_{(u,v)\in E}[(f(u) - f(v))^2]$$

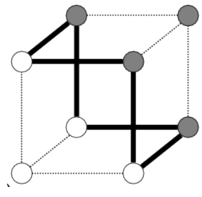
where E is the edge set of hypercube. Thus our Laplacian only differs from the standard Laplacian on the hypercube graph by a multiplicative constant.

1.5 Influence

Definition 11. *Let* $f : \{-1,1\}^n \to \{-1,1\}$. $\text{Inf}_i(f) := \Pr_x[f(x) \neq f(x^{\oplus i})]$ *where* $x^{\oplus i}$ *is* x *with* i^{th} *bit flipped.* $\text{Inf}(f) = \sum_{i=1}^n \text{Inf}_i(f)$

- $\text{Inf}_i(f)$ is the fraction of dimension-i edges which are boundary edges
- $\frac{1}{n}\text{Inf}(f)$ is the fraction of edges which are boundary edges

Figure 1.2: white points correspond to 1 and grey points correspond to -1. Blacks edges are the boundary (bichromatic) edges.



Definition 12. *A Boolean function* $f : \{-1,1\}^n \to \{-1,1\}$ *is said to be increasing in variable* i *if* $f(x^{i\to 1}) \geq f(x^{i\to -1})$ *i.e.* $D_i f \geq 0$*. A decreasing function is similarly defined.* f *is said to be monotone in variable* i *if it is either increasing in* i *or decreasing in* i*.*

Proposition 8.

1. $\mathrm{Inf}_i(f) = \|D_i f\|_2^2 = \sum_{S\ni i} \hat{f}(S)^2$
2. $\mathrm{Inf}(f) = \sum_{k=0}^{n} kW^k[f] = \frac{\mathrm{dStab}_\rho[f]}{\mathrm{d}\rho}|_{\rho=1}$
3. *If* f *is a Boolean function increasing in variable* i *then* $\mathrm{Inf}_i(f) = \hat{f}(i)$*. If* f *is decreasing in variable* i *then* $\mathrm{Inf}_i(f) = -\hat{f}(i)$

Proof.

1. $\mathrm{Inf}_i(f) = \mathop{\mathbb{E}}_x[\mathbb{1}_{f(x)\neq f(x^{\oplus i})}] = \mathop{\mathbb{E}}_x[(D_i f)^2] = \sum_{S\ni i} \hat{f}(S)^2$
2. $\mathrm{Inf}(f) = \sum_i \mathrm{Inf}_i(f) = \sum_i \sum_{S\ni i} \hat{f}(S)^2 = \sum_S |S|\hat{f}(S)^2 = \sum_k kW^k[f]$. The other part follows from $\mathrm{Stab}_\rho[f] = \sum_k W^k[f]\rho^k$
3. If f is increasing in variable i then $D_i f \in \{0,1\}$ and so $\mathrm{Inf}_i(f) = \mathop{\mathbb{E}}_x[D_i f] = \hat{f}(i)$. If f is decreasing in variable i then $D_i f \in \{-1,0\}$ and so $\mathrm{Inf}_i(f) = \mathop{\mathbb{E}}_x[-D_i f] = -\hat{f}(i)$

□

Proposition 9. $\hat{\mathrm{Maj}}_n(i) = \mathrm{Inf}_i(\mathrm{Maj}_n) = \frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}}$

Proof. Majority is increasing in each variable and thus $\hat{\mathrm{Maj}}_n(i) = \mathrm{Inf}_i(\mathrm{Maj}_n)$.

$$\mathrm{Inf}_i(\mathrm{Maj}_n) = \Pr_x[\mathrm{Maj}_n(x) \neq \mathrm{Maj}_n(x^{\oplus i})] = \Pr_x[\sum_{j\neq i} x_j = 0] = \frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}}$$

□

Chapter 2

Level-1 Weight of LTFs

2.1 Chow Parameters

Theorem 3 ([Cho61]). *Let f, g be LTFs. If $\hat{f}(S) = \hat{g}(S)\ \forall\ |S| \leq 1$ then $f = g$.*

Proof.

$$\text{Let } f(x) = \text{sgn}(l(x)), l(x) = l_0 + \sum_i l_i x_i, l(x) \neq 0\ \forall\ x$$

$$\hat{f}(\emptyset) l_0 + \sum_i \hat{f}(i) l_i = \mathbb{E}[f(x)l(x)] = \mathbb{E}[|l(x)|] \geq E[g(x)l(x)] = \hat{g}(\emptyset) l_0 + \sum_i \hat{g}(i) l_i$$

By hypothesis, the first and the last terms are equal. So

$$|l(x)| = g(x)l(x) \Rightarrow g(x) = \text{sgn}(l(x)) = f(x) \text{ since } l(x) \neq 0$$

□

The parameters $\hat{f}(\emptyset), \hat{f}(1), \cdots, \hat{f}(n)$* uniquely determine f among LTFs and are called Chow parameters

For balanced LTFs, which are odd functions, $\hat{f}(\emptyset) = \mathbb{E}_x[f(x)] = 0$. So $\hat{f}(1), \cdots, \hat{f}(n)$ uniquely determine balanced LTFs.

2.2 Level-1 Weight Conjecture

Let f be an LTF, then what is the best lower bound on $W^{\leq 1}[f]$?

There are two extreme cases:

- For dictators $\chi_i(x) = x_i$, $W^{\leq 1}[f] = 1$
- For majority $\text{Maj}_n(x) = \text{sgn}(x_1 + \cdots + x_n)$, $W^{\leq 1}[\text{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{2}{\pi}$

*we write $\hat{f}(i)$ instead of $\hat{f}(\{i\})$ for brevity

So intuitively, one of them should be the worst case which is Maj_n here. This is precisely the conjecture.

Conjecture 1. *$W^{\leq 1}[LTF] := \inf_f W^{\leq 1}[f] = \frac{2}{\pi}$ where the infimum is over all LTFs*

For balanced LTFs, $W^0[f] = 0$. It is enough to check the validity of conjecture 1 for balanced LTFs which follows from the following proposition.

Proposition 10. *Let f be an LTF, then there exists a balanced LTF g such that $W^1[g] = W^{\leq 1}[f]$*

Proof. Let $f : \{-1,1\}^n \to \{-1,1\}$ given by $f(x) = \mathrm{sgn}(l.x + \theta)$

Let $g : \{-1,1\}^{n+1} \to \{-1,1\}$ be given by $g(x, x_{n+1}) = \mathrm{sgn}(l.x + \theta x_{n+1})$,
then for $1 \leq i \leq n$,

$$\begin{aligned}\hat{g}(i) &= \mathop{\mathbb{E}}_{x,x_{n+1}}[\mathrm{sgn}(l.x + \theta x_{n+1})x_i] \\ &= \frac{1}{2}(\mathop{\mathbb{E}}_x[\mathrm{sgn}(l.x+\theta)x_i] + \mathop{\mathbb{E}}_x[\mathrm{sgn}(l.x-\theta)x_i]) \\ &= \frac{1}{2}(\mathop{\mathbb{E}}_x[\mathrm{sgn}(l.x+\theta)x_i] + \mathop{\mathbb{E}}_x[\mathrm{sgn}(-l.x-\theta)(-x_i)]) = \hat{f}(i)\end{aligned}$$

Similarly

$$\begin{aligned}\hat{g}(n+1) &= \mathop{\mathbb{E}}_{x,x_{n+1}}[\mathrm{sgn}(l.x + \theta x_{n+1})x_{n+1}] \\ &= \frac{1}{2}(\mathop{\mathbb{E}}_x[\mathrm{sgn}(l.x+\theta)] + \mathop{\mathbb{E}}_x[-\mathrm{sgn}(l.x-\theta)]) \\ &= \frac{1}{2}(\mathop{\mathbb{E}}_x[\mathrm{sgn}(l.x+\theta)] + \mathop{\mathbb{E}}_x[-\mathrm{sgn}(-l.x-\theta)]) = \hat{f}(\emptyset)\end{aligned}$$

So $W^1[g] = W^{\leq 1}[f]$ □

2.3 Gotsman-Linial

The following is a well-known bound on $W^{\leq 1}[f]$ when f is an LTF:

Theorem 4 ([GL94]). *Let $f = \mathrm{sgn}(l(x))$ where $l(x) = a_0 + a_1x_1 + \cdots + a_nx_n$, $\sum_i a_i^2 = 1$ then $W^{\leq 1}[f] \geq \|l\|_1^2$*

Proof.

$$\|l\|_1 = \mathop{\mathbb{E}}_x[|l(x)|] = \mathop{\mathbb{E}}_x[f(x)l(x)] = \langle f, l\rangle \leq \hat{f}(\emptyset)a_0 + \sum_{i=1}^{n}\hat{f}(i)a_i \leq \sqrt{W^{\leq 1}[f]}$$

where we have used Plancherel identity(3) and the last inequality is by Cauchy-Schwarz inequality. □

Combining this with the well-known Khintchine-Kahane inequality (Corollary 4) gives us a lower bound on $W^{\leq 1}[LTF]$.

Corollary 3 ([GL94]). *$W^{\leq 1}[LTF] \geq \frac{1}{2}$*

Proof. Let $f = \text{sgn}(l(x))$ be any LTF. By Khintchine-Kahane inequality (Corollary 4) $\|l\|_1 \geq \frac{1}{\sqrt{2}}$ and thus $W^{\leq 1}[f] \geq \|l\|_1^2 \geq \frac{1}{2}$ ☐

We will prove Khintchine-Kahane inequality in section 2.5

2.4 Regular-LTF

We have seen that in conjectures 1 and 2, Maj_n as $n \to \infty$ is the extreme case. Intuitively when a function is very close to majority, the inequalities in the conjectures should be approximately tight. We will prove this now by a notion of closeness called regularity.

Definition 13. *If $f = \text{sgn}(a_1x_1 + \cdots + a_nx_n)$ where $\sum_i a_i^2 = 1$ and $\max_i |a_i| \leq \epsilon$ then f is called an ϵ-regular LTF*

The smaller the ϵ, the closer it is to Maj_n which is $\frac{1}{\sqrt{n}}$-regular.

We will need a central limit theorem with explicit error bounds.

Theorem 5 (Berry-Esséen [O'Db]). *Let $X_1, \cdots, X_n$ be independent random variables satisfying $\mathbb{E}[X_i] = 0 \ \forall i \in [n]$, $\sum_i \mathbb{E}[X_i^2] = 1$ and $\sum_i \mathbb{E}[|X_i|^3] = \rho$, then $S = \sum_i X_i$ is ρ-close to $\mathcal{N}(0,1)$ in distribution. Also their first absolute moments are $\mathcal{O}(\rho)$-close i.e. for $Z \sim \mathcal{N}(0,1)$*

- $\forall\ t\ |Pr[S \leq t] - Pr[Z \leq t]| \leq \rho$
- $|\mathbb{E}[|S|] - \mathbb{E}[|Z|]| \leq \mathcal{O}(\rho)$

Theorem 6. *If $f = \text{sgn}(\sum_i a_ix_i)$ is an ϵ-regular LTF then $W^1[f] \geq \frac{2}{\pi} - \mathcal{O}(\epsilon)$*

Proof. Let $l(x) = \sum_i a_ix_i$, $X_i = a_ix_i$, then $\mathbb{E}[X_i] = 0$, $\sum_i \mathbb{E}[X_i^2] = \sum_i a_i^2 = 1$ and $\sum_i \mathbb{E}[|X_i|^3] = \sum_i |a_i|^3 \leq \max_i |a_i| \leq \epsilon$

$l(x) = \sum_i X_i$, so by Berry-Esséen Theorem, $\mathbb{E}[|l|] = \mathbb{E}[|\mathcal{N}(0,1)|] \pm \mathcal{O}(\epsilon) = \sqrt{\frac{2}{\pi}} \pm \mathcal{O}(\epsilon)$
By theorem 4, $W^1[f] \geq \|l\|_1^2 = \frac{2}{\pi} \pm \mathcal{O}(\epsilon)$ ☐

2.5 Khintchine-Kahane Inequality

We will first prove a weaker form of Khintchine-Kahane inequality which has a more natural proof.

Theorem 7 (Weak KK)**.** *Let* $l : \{-1,1\}^n \to \mathbb{R}$ *be given by* $l(x) = \sum_1^n l_i x_i + l_0$ *and* $\mathbb{E}[l^2] = \sum_{i=0}^n l_i^2 = 1$ *then* $\|l\|_1 = \mathbb{E}[|l|] \geq \frac{1}{\sqrt{3}}$

Proof. Let $L = |l|$

$$\mathbb{E}[L^4] = \sum_i l_i^4 + 6\sum_{i<j} l_i^2 l_j^2 \leq 3(\sum_i l_i^2)^2 = 3$$

Using Hölder's inequality on L^α and L^β with a, b norms such that $\alpha + \beta = 2, \frac{1}{a} + \frac{1}{b} = 1, a\alpha = 4, b\beta = 1$ i.e. $a = 3, b = 3/2$ gives

$$\mathbb{E}[L^2] \leq \mathbb{E}[L^4]^{1/3}\mathbb{E}[L]^{2/3} \Rightarrow \mathbb{E}[L] \geq \frac{1}{\sqrt{3}}$$

□

We will now prove the actual Khintchine-Kahane inequality with a more general hypothesis.

Theorem 8 (Khintchine-Kahane inequality)**.** *Let* $g : \{-1,1\}^n \to \mathbb{R}$ *be given by* $g(x) = \|\sum_{i=1}^n w_i x_i\|_*$ *where* $w_i \in V$*,* V *is a normed space with norm* $\|\cdot\|_*$ *then*

$$\|g\|_1 \geq \frac{1}{\sqrt{2}}\|g\|_2$$

Proof.

$$\forall\ x\ Lg(x) = \frac{1}{2}\sum_i [g(x) - g(x^{\oplus i})] = \frac{1}{2}[ng(x) - \sum_i g(x^{\oplus i})] \leq g(x)$$

where the last inequality follows from triangle inequality for $\|\cdot\|_*$ as follows:

$$\sum_i g(x^{\oplus i}) = \sum_i \|w \cdot x^{\oplus i}\|_* \leq \|w \cdot \sum_i x^{\oplus i}\|_* = (n-2)\|w \cdot x\|_* = (n-2)g(x)$$

$$\text{Now } \|g\|_2^2 = \langle g, g\rangle \geq \langle g, Lg\rangle = \sum_{i-even} iW^i[g] \geq 2\sum_{i=0}^n W^i[g] - 2W^0[g] = 2\|g\|_2^2 - 2\|g\|_1^2$$

where we have used the fact that g is even and thus by proposition 2, $W^i[g] = 0$ when i is odd. □

Corollary 4. *Let* $l(x) = w_0 + \sum_{i=1}^n w_i x_i$ *where* $\sum_{i=0}^n w_i^2 = 1$ *then* $\mathbb{E}[|l|] = \|l\|_1 \geq \frac{1}{\sqrt{2}}$

Proof. Let $l' : \{-1,1\}^{n+1} \to \mathbb{R}$ be given by $l'(x) = \sum_{i=0}^n w_i x_i$.

$$\mathbb{E}[|l'|] = \frac{1}{2}\mathbb{E}[|w_0 + w \cdot x| + |-w_0 + w \cdot x|] = \mathbb{E}[|w_0 + w \cdot x|] = \mathbb{E}[|l|]$$

By Parseval identity(2) $\|l'\|_2^2 = \sum_{i=0}^n w_i^2 = 1$ and thus by theorem 8 $\|l\|_1 = \|l'\|_1 \geq \frac{1}{\sqrt{2}}$ □

Geometric Interpretation

How do we geometrically interpret the above corollary? Let us assume that $w_0 = 0$, thus $l(x) = 0$ is a hyperplane that passes through the origin. Also $|l(x)| = |w \cdot x|$ is the distance of point x from the plane $l(x) = 0$. Thus the above corollary is proving that the average distance of a point on the hypercube from any plane that passes through the origin is atleast $\frac{1}{\sqrt{2}}$

Tight instances for Khintchine-Kahane inequality

When is Khintchine-Kahane inequality tight?

Definition 14. *We say* $a = (a_1, \cdots, a_n)$ *is canonical if* $a_i \geq 0,\ a_1 \geq \cdots \geq a_n$ *and* $\sum_i a_i^2 = 1$

Proposition 11. *Let* $g(x) = |w \cdot x|,\ w$ *is canonical. Then*

$$\|g\|_1 = \frac{1}{\sqrt{2}} \Leftrightarrow w = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \cdots, 0)$$

Proof Sketch. The two inequalities that we used in the proof of Khintchine-Kahane inequality (theorem 8) are:

- $\langle g, Lg \rangle \leq \langle g, g \rangle$
- $W^{\geq 4}[g] \geq 0$

For the first inequality to be tight, $\forall\ x\ Lg(x) = g(x)$ or $g(x) = 0$. This implies that the hyperplane $g(x) = 0$ cannot strictly intersect any edges of the hypercube. From this we can conclude that w should be of the form $(w_1, \cdots, w_1, 0, \cdots, 0),\ w_1 > 0$.

Now for the second inequality to be tight, the number of w_1 should be exactly two i.e. $w = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \cdots, 0)$ □

2.6 Gotsman-Linial is not tight

There was a recent result which made some progress on the conjecture, it was shown that Gotsman-Linial is not tight.

Theorem 9 ([DDS12]). $W^{\leq 1}[LTF] \geq \frac{1}{2} + c$ *for some absolute constant* $c > 0$

The two inequalities used in the proof of Gotsman-Linial (corollary 3) are:

- Cauchy-Schwarz inequality
- Khintchine-Kahane inequality

It can be shown that both cannot be simultaneously tight, atleast one of them has to have a constant slack. This is proved using a robust version of Khintchine-Kahane inequality.

Robust Khintchine-Kahane inequality

We have seen that the only tight instances for Khintchine-Kahane inequality in canonical form are $l(x) = \frac{1}{\sqrt{2}}(x_1 + x_2)$ (proposition 11). Intuitively if we are moving far away from this tight instance then Khintchine-Kahane inequality must have increasing slack.

Theorem 10 ([DDS12]). *If $l(x) = a \cdot x$, a is canonical, then $\|l\|_1 \geq \frac{1}{\sqrt{2}} + c\|a - a^*\|_2$ where $a^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \cdots, 0)$ and $c > 0$ is some absolute constant*

The proof of theorem 9 is then divided into two cases:

Case I: $\|a - a^*\|_2 \geq \tau$

Khintchine-Kahane inequality is not tight because of theorem 10

Case II: $\|a - a^*\|_2 \leq \tau$

Cauchy-Schwarz not tight

Chapter 3

Stability of LTFs

3.1 Majority is least stable Conjecture

There is an other related conjecture which is stronger than conjecture 1. Let us calculate stability for the two extreme cases of dictatorship and majority. Intuitively, one of them is an extreme case.

- $\mathrm{Stab}_\rho[\chi_i] = \rho$
- $\mathrm{Stab}_\rho[\mathrm{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{2}{\pi}\arcsin\rho = \frac{2}{\pi}(\rho + \frac{1}{6}\rho^3 + \frac{3}{40}\rho^5 + \cdots)$

Conjecture 2 ([BKS98]). *$\forall$ LTF f on n(odd) inputs,* $\mathrm{Stab}_\rho[f] \geq \mathrm{Stab}[\mathrm{Maj}_n]\ \forall\ \rho \in [0,1]$

Note that this conjecture implies conjecture 1 since for a balanced LTF f, $\mathrm{Stab}_\rho[f] \underset{\rho\to 0}{\longrightarrow} W^1[f]\rho$ and $\mathrm{Stab}_\rho[\mathrm{Maj}_n] \underset{\rho\to 0}{\longrightarrow} \frac{2}{\pi}\rho$

3.2 Counter Example

We now give a counter example to conjecture 2.

Let $f : \{-1,1\}^5 \to \{-1,1\}$ be given by $f(x) = \mathrm{sgn}(x_1 + x_2 + x_3 + 2x_4 + 2x_5)$. Clearly

$$f(x) = \mathrm{Maj}_5(x) - 2\mathbb{1}_{\{x_1=x_2=x_3=1, x_4=x_5=-1\}} + 2\mathbb{1}_{\{x_1=x_2=x_3=-1, x_4=x_5=1\}}$$

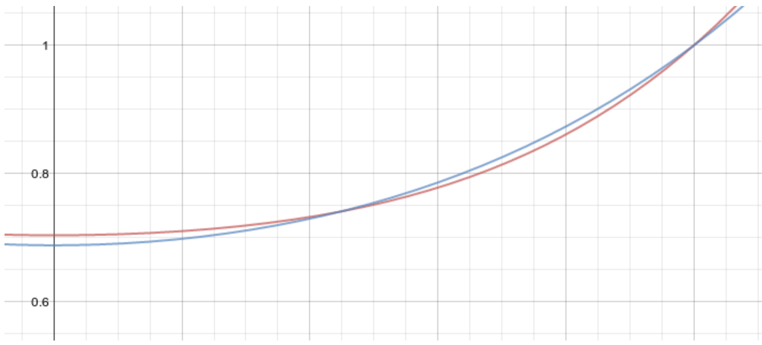
We can calculate the stabilities to be,

$$\mathrm{Stab}_\rho[f] = 0.6875\rho + 0.25\rho^3 + 0.0625\rho^5$$

$$\mathrm{Stab}_\rho[\mathrm{Maj}_5] = 0.703125\rho + 0.28125\rho^3 + 0.015625\rho^5$$

Figure 3.1 shows the plot of $\frac{\mathrm{Stab}_\rho[f]}{\rho}$(blue) and $\frac{\mathrm{Stab}_\rho[\mathrm{Maj}_5]}{\rho}$(red). Observe that the level-1 weight of f is less than that of Maj_5

Figure 3.1: Graphs of $\frac{\mathrm{Stab}_\rho[f]}{\rho}$ and $\frac{\mathrm{Stab}_\rho[\mathrm{Maj}_5]}{\rho}$



We now extend this example to general odd $n > 5$.

Let $f : \{-1,1\}^{2t+1} \to \{-1,1\}$ be a LTF given by $f(x) = \mathrm{sgn}(a\sum_{i=1}^{t+1} x_i + b\sum_{i=t+2}^{2t+1} x_i)$ where $a, b \in \mathbb{N}$ are chosen such that $(t+1)a - tb < 0, ta - a + b - (t-1)b > 0$ and $a+b$ is odd. When $a+b$ is odd, $a\sum_{i=1}^{t+1} x_i + b\sum_{i=t+2}^{2t+1} x_i \neq 0\ \forall x$. So we need $\frac{t+1}{t} < \frac{b}{a} < \frac{t-1}{t-2}$. But since $\frac{t+1}{t} < \frac{t-1}{t-2}\ \forall t > 2$, we can always find such a pair (a, b). So,

$$\begin{aligned} f(x) &= \mathrm{Maj}_{2t+1}(x) - 2\mathbb{1}_{\{x_1=\cdots=x_{t+1}=1, x_{t+2}=\cdots=x_{2t+1}=-1\}} + 2\mathbb{1}_{\{x_1=\cdots=x_{t+1}=-1, x_{t+2}=\cdots=x_{2t+1}=1\}} \\ &= \mathrm{Maj}_{2t+1}(x) - 2\prod_{i=1}^{t+1}\left(\frac{1+x_i}{2}\right)\prod_{i=t+2}^{2t+1}\left(\frac{1-x_i}{2}\right) + 2\prod_{i=1}^{t+1}\left(\frac{1-x_i}{2}\right)\prod_{i=t+2}^{2t+1}\left(\frac{1+x_i}{2}\right) \end{aligned}$$

Theorem 11. *For f defined as above $W^1[f] < W^1[\mathrm{Maj}_{2t+1}]$ and thus $\mathrm{Stab}_\rho[f] < \mathrm{Stab}_\rho[Maj_{2t+1}]$ where $\rho \in [0, \epsilon]$ for some $\epsilon > 0$*

Proof. Let $n = 2t+1$.

$$\hat{f}(i) = \begin{cases} \hat{\mathrm{Maj}}_n(i) - \frac{1}{2^{n-2}} & 1 \le i \le t+1 \\ \hat{\mathrm{Maj}}_n(i) + \frac{1}{2^{n-2}} & t+2 \le i \le 2t+1 \end{cases}$$

By proposition 9, $\hat{\mathrm{Maj}}_n(i) = \frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}}\ \forall i$

$$\begin{aligned} W^1[f] &= (t+1)[\hat{\mathrm{Maj}}_n(1) - \frac{1}{2^{n-2}}]^2 + t[\hat{\mathrm{Maj}}_n(1) + \frac{1}{2^{n-2}}]^2 \\ &= W^1[\mathrm{Maj}_n] - \frac{\binom{n-1}{\frac{n-1}{2}} - n}{2^{2(n-2)}} < W^1[\mathrm{Maj}_n] \end{aligned}$$

□

3.3 New Conjecture

We now introduce a new conjecture.

Conjecture 3. *$\forall$ LTF f, $\mathrm{Stab}_\rho[f] \geq \frac{2}{\pi}\arcsin\rho \quad \forall\rho \in [0,1]$*

Here again, we need only look at balanced LTF.

Proposition 12. *Let $f = \mathrm{sgn}(a_0 + \sum_{i=1}^n a_i x_i)$ be an LTF and let $f' = \mathrm{sgn}(\sum_{i=0}^n a_i x_i)$ then $\mathrm{Stab}_\rho[f] \geq \mathrm{Stab}_\rho[f']$*

Proof.

$$\begin{aligned}
\mathrm{Stab}_\rho[f'] &= \mathop{\mathbb{E}}_{(x,y)\sim\rho}[sgn(a_0x_0 + \sum_{i=1}^n a_ix_i)sgn(a_0y_0 + \sum_{i=1}^n a_iy_i)] \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\rho}[sgn(a_0 + \sum_{i=1}^n a_ix_i)sgn(a_0 + \sum_{i=1}^n a_iy_i)]\left(\frac{1+\rho}{2}\right) \\
&- \mathop{\mathbb{E}}_{(x,y)\sim\rho}[sgn(a_0 + \sum_{i=1}^n a_ix_i)sgn(a_0 - \sum_{i=1}^n a_iy_i)]\left(\frac{1-\rho}{2}\right) \\
&= \mathrm{Stab}_\rho[f]\left(\frac{1+\rho}{2}\right) - \mathrm{Stab}_{-\rho}[f]\left(\frac{1-\rho}{2}\right) \\
&= \sum_{k\ odd} W^k[f]\rho^k + \rho\sum_{k\ even} W^k[f]\rho^k \leq \sum_k W^k[f]\rho^k = \mathrm{Stab}_\rho[f]
\end{aligned}$$

□

3.4 Evidence for the conjecture

3.4.1 Majority

By proposition 5, we know that Maj_n tends to $\frac{2}{\pi}\arcsin\rho$ for large n. Now we will show that it converges from above.

Theorem 12. *For $k \leq n$ and n odd, $W^k[\mathrm{Maj}_n] > W^k[\mathrm{Maj}_{n+2}]$*

We can evaluate $W^k[\mathrm{Maj}_n]$. Refer section 5.3 of [O'Da] for a proof.

Theorem 13. *Let n be odd, then $\mathrm{Stab}_\rho[\mathrm{Maj}_n] \geq \mathrm{Stab}_\rho[\mathrm{Maj}_{n+2}]$ and thus $\mathrm{Stab}_\rho[\mathrm{Maj}_n] \geq \frac{2}{\pi}\arcsin\rho \quad \forall\rho \in [0,1]$*

Proof. * Let $f(\rho) = \text{Stab}_\rho[\text{Maj}_n] - \text{Stab}_\rho[\text{Maj}_{n+2}]$. Clearly $f(0) = f(1) = 0$.

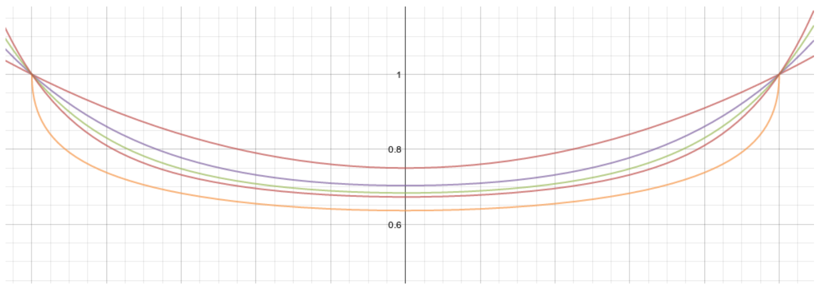
$$\begin{aligned}\rho f'(\rho) &= \sum_{k=0}^{n} k(W^k[\text{Maj}_n] - W^k[\text{Maj}_{n+2}])\rho^k - (n+2)W^{n+2}[\text{Maj}_{n+2}]\rho^{n+2} \\ &\le \sum_{k=0}^{n} (n+2)(W^k[\text{Maj}_n] - W^k[\text{Maj}_{n+2}])\rho^k - (n+2)W^{n+2}[\text{Maj}_{n+2}]\rho^{n+2} \\ &= (n+2)f(\rho)\end{aligned}$$

Since f is a smooth function on $[0,1]$, it attains its minimum say $f(\rho^*)$. If $f(\rho^*) < 0$ then $\rho^* \neq 0, 1$ and thus $f'(\rho^*) = 0$. But $0 = \rho^* f'(\rho^*) \le (n+2)f(\rho^*) < 0$ which is a contradiction. Thus $f(\rho) \ge 0 \ \forall \rho \in [0,1]$.

By proposition 5 we know that $\text{Stab}_\rho[\text{Maj}_n] \underset{n\to\infty}{\longrightarrow} \frac{2}{\pi}\arcsin\rho$, so we get $\text{Stab}_\rho[\text{Maj}_n] \ge \frac{2}{\pi}\arcsin\rho$ □

Figure 3.2 shows Maj_n for $n = 3, 4, 5, 6$ and $\frac{2}{\pi}\frac{\arcsin\rho}{\rho}$(orange)

Figure 3.2: $\frac{\text{Stab}_\rho[Maj_n]}{\rho}$ for n=3,5,7,9 and $\frac{2}{\pi}\frac{\arcsin\rho}{\rho}$



3.4.2 Regular-LTF

We have seen regular-LTFs in section 2.4. Now we will prove a stronger version of theorem 6.

Theorem 14. *If* $f = \text{sgn}(\sum_i a_i x_i)$ *is an* ϵ*-regular LTF then* $|\text{Stab}_\rho[f] - \frac{2}{\pi}\arcsin\rho| \le \mathcal{O}(\frac{\epsilon}{\sqrt{1-\rho^2}})$

*Proof due to Swagato Sanyal (http://www.tcs.tifr.res.in/people/swagato-sanyal)

Proof. By Berry-Esséen theorem, $\sum_i a_i x_i \approx_\epsilon \mathcal{N}(0,1)$

$$\begin{aligned}\mathrm{Stab}_\rho[f] &= 1 - 2 \underset{(x,y)\ \rho\text{−correlated}}{Pr}[f(x) \neq f(y)] \\ \underset{2d\ Berry-Esseen}{\rightarrow} & 1 - 2 \underset{(Z_1,Z_2)\ \rho\text{−corr}\ \mathcal{N}(0,1)}{Pr}[\mathrm{sgn}(Z_1) \neq \mathrm{sgn}(Z_2)] \pm \mathcal{O}(\frac{\epsilon}{\sqrt{1-\rho^2}}) \\ &= \frac{2}{\pi} \arcsin \rho \pm \mathcal{O}(\frac{\epsilon}{\sqrt{1-\rho^2}})\end{aligned}$$

□

3.4.3 Behaviour near $\rho = 1$

Proposition 13. *∀ LTF f on n(odd) variables,* $\mathrm{Inf}(f) \leq \mathrm{Inf}(\mathrm{Maj}_n)$

Proof. Let $f(x) = \mathrm{sgn}(w_0 + \sum_{i=1}^n w_i x_i)$. We can assume $w_i \geq 0$ for $i \in [n]$, this is because influences do not change by changing signs of w_i. Now f is increasing in each variable, so by proposition 8

$$\mathrm{Inf}(f) = \sum_i \hat{f}(i) = \mathbb{E}[f(x)(\sum_i x_i)] \leq \mathbb{E}[|\sum_i x_i|] = \mathbb{E}[\mathrm{sgn}(\sum_i x_i)(\sum_i x_i)] = \mathrm{Inf}(\mathrm{Maj}_n)$$

□

Corollary 5. *∀ LTF* $f : \{-1,1\}^n \to \{-1,1\}$, $\mathrm{Stab}_\rho[f] \geq \mathrm{Stab}_\rho[\mathrm{Maj}_n] \geq \frac{2}{\pi}\arcsin\rho$ *for* ρ *in some neighbourhood of 1.*

Proof. $\mathrm{Stab}_1[f] = 1 = \mathrm{Stab}_1[\mathrm{Maj}_n]$. By proposition 8,

$$\frac{\mathrm{dStab}_\rho[f]}{\mathrm{d}\rho}|_{\rho=1} = \mathrm{Inf}[f] \leq \mathrm{Inf}[\mathrm{Maj}_n] = \frac{\mathrm{dStab}_\rho[\mathrm{Maj}_n]}{\mathrm{d}\rho}|_{\rho=1}$$

So by a Taylor expansion around $\rho = 1$ we get the desired result. □

Note that the above theorem doesn't give a uniform neighbourhood of 1 in which the conjecture is true, the neighbourhood depends on the function.

A more precise estimate can be given about the stability near $\rho = 1$. The conjecture implies that

$$\mathrm{Stab}_{1-\epsilon}[f] \geq \frac{2}{\pi}\arcsin(1-\epsilon) = 1 - \frac{2}{\pi}\sqrt{2\epsilon} - \mathcal{O}(\epsilon^{3/2})$$

A similar albeit weaker bound is indeed known.

Theorem 15 ([Per04])**.** *∀ LTF f,* $\mathrm{Stab}_{1-\epsilon}[f] \geq 1 - \sqrt{\frac{2}{\pi}}\sqrt{2\epsilon} - \mathcal{O}(\epsilon^{3/2})$

Refer section 5.5 of [O'Da] for a proof.

3.4.4 Behaviour near $\rho = 0$

The conjecture is true near $\rho = 0$ iff the level-1 conjecture(1) is true. This is because

$$\mathrm{Stab}_\rho[f] = W^1[f]\rho + \mathcal{O}(\rho^2) \geq \frac{2}{\pi}\arcsin\rho \geq \frac{2}{\pi}\rho$$

Bibliography

[BKS98] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of boolean functions and applications to percolation. 1998.

[Cho61] C. K. Chow. On the characterization of threshold functions. In *SWCT (FOCS)*, pages 34–38, 1961.

[DDS12] Anindya De, Ilias Diakonikolas, and Rocco A. Servedio. A robust Khintchine-Kahane. *CoRR*, abs/1207.2229, 2012.

[GL94] Craig Gotsman and Nathan Linial. Spectral Properties of Threshold Functions. *Combinatorica*, 14(1):35–50, 1994.

[O'Da] Ryan O'Donnell. Analysis of Boolean Functions. http://www.analysisofbooleanfunctions.org.

[O'Db] Ryan O'Donnell. Berry Esséen Theorem. http://www.cs.cmu.edu/~odonnell/boolean-analysis/lecture21.pdf.

[Per04] Yuval Peres. Noise stability of weighted majority. http://arxiv.org/abs/math/0412377, 2004.

[Tao] Terence Tao. Central Limit Theorem. http://terrytao.wordpress.com/2010/01/05/254a-notes-2-the-central-limit-theorem.