Boolean Function Approximation by a Flat Polynomial

Thesis submitted by

Ankit Gupta
M.Tech(CS)

under the guidance of

Prof. Sourav Chakraborty
Indian Statistical Institute, Kolkata

in partial fulfilment of the requirements
for the award of the degree of

Master of Technology

ACMU
Indian Statistical Institute, Kolkata

09 July 2021

CERTIFICATE

This is to certify that the dissertation entitled **"Boolean Function Approximation by a Flat Polynomial"** submitted by **Ankit Gupta** to Indian Statistical Institute, Kolkata, in partial fulfillment for the award of the degree of **Master of Technology in Computer Science** is a bonafide record of work carried out by him under my supervision and guidance. The dissertation has fulfilled all the requirements as per the regulations of this institute and, in my opinion, has reached the standard needed for submission.

9/7/21

Prof. Sourav Chakraborty
Associate Professor,
Advanced Computing and Microelectronics Unit (ACMU) , Kolkata,
Indian Statistical Institute,
Kolkata-700108, India.

Acknowledgements

I would like to express my special thanks of gratitude to my guide Dr. Sourav Chakraborty who gave me an opportunity to do this wonderful project, It also gave me glimpse of how to do Research and I came to know about so many new things in this area.

Ankit Gupta

Ankit Gupta
M.Tech(Computer Science),
Indian Statistical Institute,
Kolkata-700108, India.

Abstract

Boolean functions $f : \{-1,1\}^n \to \{-1,1\}$ arise in many areas of theoretical computer science and mathematics, for example: complexity theory, quantum computing and graph theory etc and Fourier analysis is a powerful technique used to analyze problems in these areas. One of the most important and longstanding open problems in this field is the Fourier Entropy-Influence (FEI) conjecture [EG96], first formulated by Ehud Friedgut and Gil Kalai; The FEI conjecture connects two fundamental properties of boolean function f, i.e. influence and entropy. FEI conjecture says, for all boolean functions $f : \{-1,1\}^n \to \{-1,1\}$, $H[\hat{f}^2] \le CI[f]$ where $H[\hat{f}^2]$ is the spectral entropy of f and if we fix $\epsilon = \frac{1}{3}$ and consider polynomials p such that $|p(x) - f(x)| \le \frac{1}{3}$ where f is boolean function then these polynomials have many applications in theoretical computer science.

In particular, this work attempts to address the following problem:

Suppose, the FEI conjecture is true, what can be said about the approximating polynomials. We have a flat polynomial of degree d and sparsity $2^{\omega(d)}$. The proposed conjecture [SSM+20] says that No flat polynomial of degree d and sparsity $2^{\omega(d)}$ can $\frac{1}{3}-$ approximate a boolean function.[The degree of a function is the maximum d such that $\hat{f}(S) \neq 0$ for some set S of size d]. It is useful to understand better the structure of polynomials that $\epsilon-$approximate Boolean functions on the Boolean cube. Such polynomials have proved to be powerful and found diverse applications in theoretical computer science. Here, we restrict ourselves to a class of polynomials called flat polynomials over $\{-1,1\}$, i.e., polynomials whose non-zero coefficients have the same magnitude. This conjecture is true by assuming FEI conjecture and it is also true for a class of polynomials without assuming FEI conjecture.

Contents

Chapter 1

Introduction

1.1 Introduction

We define a Boolean function as $f : \{0,1\}^n \to \{0,1\}$ or $f : \{-1,1\}^n \to \{-1,1\}$ which means f maps each length n binary vector or string into a single binary value or bit. The domain of a Boolean function can be defined as the hamming cube. It is also known as hypercube/n-cube/Boolean cube/discrete cube. Generally, we are interested in hamming distance between $x, y \in \{-1,1\}^n$ which is defined as $\Delta(x,y) = \#\{x_i \neq y_i\}$, where x denotes a bit string and x_i denotes its i^{th} co-ordinate.

The Fourier expansion of a Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ is represented as a real multilinear polynomial which generally means no variable x_i appears squared or cubed etc.

For example, $\max_2(x_1, x_2) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2$ where function $\max_2$ is defined on 2 bits with $\max_2(+1,+1) = +1, \max_2(-1,+1) = +1, \max_2(+1,-1) = +1, \max_2(-1,-1) = -1$

Every Boolean function has Fourier expansion and this Fourier expansion can be written as multilinear polynomial **uniquely**. It is same as for $\{0,1\}^n \to \{0,1\}$ where 0 is encoded as 1 and 1 is encoded as -1.

There are some definitions based on which we can get the Fourier expansion of various Boolean functions easily.

Definition 1.1.1 (Fourier Spectrum). *Every $f : \{-1,+1\}^n \to \mathbb{R}$ can be represented as a multilinear polynomial uniquely as:*

$$f(x) = \Sigma_{S \subseteq \{1,2,\ldots,n\}} \hat{f}(S)\chi_S(x) = \Sigma_{S \subseteq \{1,2,\ldots,n\}} \hat{f}(S)\Pi_{i \in S} x_i$$

where Fourier coefficient $\hat{f}(S) = \frac{1}{2^n}\Sigma_{x \in \{-1,1\}^n} f(x)\Pi_{i \in S} x_i$.

Definition 1.1.2. ***By using Lagrange Interpolation*** *Here, we interpolate the 2^n values that f assigns to the strings $\{-1,1\}^n$. For each $a = (a_1, a_2, \ldots, a_n)$, there is an* ***Indicator Polynomial****:*

$$\mathbb{I}_{\{a\}}(x) = \left(\frac{1 + a_1x_1}{2}\right)\left(\frac{1 + a_2x_2}{2}\right) \ldots \left(\frac{1 + a_nx_n}{2}\right)$$

And hence,

$$f(x) = \Sigma_{a \in \{-1,1\}^n} f(a)\mathbb{I}_{\{a\}}(x)$$

This is a familiar method for finding a polynomial that interpolates the 2^n values that f assigns to the points $\{-1,1\}^n \subset \mathbb{R}^n$. Here, For each point $a = (a_1, a_2, ..., a_n) \in \{-1,1\}^n$, the indicator polynomial $\mathbb{1}_{\{a\}}(x)$ takes value 1 when $x = a$ and value 0 when $x \in \{-1,1\}^n \backslash \{a\}$.

For example, we can write $\max_2 x$ as:
$(+1)\left(\frac{1+x_1}{2}\right)\left(\frac{1+x_2}{2}\right) + (+1)\left(\frac{1-x_1}{2}\right)\left(\frac{1+x_2}{2}\right) + (+1)\left(\frac{1+x_1}{2}\right)\left(\frac{1-x_2}{2}\right) + (-1)\left(\frac{1-x_1}{2}\right)\left(\frac{1-x_2}{2}\right)$
$= \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_1x_2.$

This interpolation procedure works for $\{-1,1\}^n \to \mathbb{R}$ also.

There are some Boolean functions which are very useful in Fourier analysis and definition of these Boolean functions are as follows:

Definition 1.1.3. ***Majority Function:***

Majority function $'f'$ is defined on the n Boolean variables as

$$f(x_1, x_2, ..., x_n) = \begin{cases} 1 & \Sigma_{i=1}^n x_i \geq 0 \\ -1 & o/w \end{cases}$$

Definition 1.1.4. ***Parity Function :***

For $x \in \{-1,+1\}^n \to \{-1,+1\}$ It is defined as :

$$\chi_S(x) = \Pi_{i \in S} x_i$$

So, χ_S is a Boolean function and it computes the logical parity or ex-or(XOR) of bits $(x_i)_{i \in S}$

Any f can be represented as a linear combination of parity function over the reals as

$$f = \Sigma_{S \subseteq \{1,2,....,n\}} \hat{f}(S) \chi_S$$

Definition 1.1.5. ***Inner Product Function :***

Considering the 2^n dimensional functions of all functions

$$f : \{0,1\}^n \to \mathbb{R},$$

and we define an inner product of 2 functions on this space as

$$< f, g >= \frac{1}{2^n} \Sigma_{x \in \{0,1\}^n} f(x) g(x) = \mathbb{E}[f.g].$$

Now, for each $S \subseteq [n]$, define a function,

$$\chi_S : \{0,1\}^n \to \{1,-1\}$$

as

$$\chi_S(x) = (-1)^{S.x} = (-1)^{\Sigma_{i \in S} x_i}$$

Definition 1.1.6. ***Maximum Function :***

It takes n bit string as an input and gives maximum value on these n bits and this Maximum function represents logical AND function on n bits i.e. AND_n.

Definition 1.1.7. ***Linear Threshold Function(LTF)***

It is a Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$ *which is defined as*

$$f(x) = sgn(a_0 + a_1x_1 + ... + a_n\, x_n)$$

where constants $a_0, a_1, ..., a_n \in \mathbb{R}$ *and* $sgn(0) = 1$

These LTFs play an important role in learning theory and in circuit complexity.

Influence

Influence of $i \in [n]$ *is* defined as $\mathbb{P}_{x \sim \{-1,1\}^n}[f(x) \neq f(x^{\oplus i})]$. Here, x is chosen uniformly at random and $f(x^{\oplus i})$ means i^{th} voter has reversed their vote(Flipping the i^{th} bit changes the function value).

So, we can define the Total Influence as $\mathbb{I}[f] = \Sigma_{i=1}^{n} Inf_i[f]$.

Flat Polynomial

The Class of polynomials over $\{-1, 1\}$ whose non-zero coefficients have the same magnitude is called the Flat polynomial.

Littlewood polynomial is a polynomial, all of whose coefficients are $+1$ or -1. They are named after *J. E. Littlewood* who studied them in the 1950s.

A flat multilinear polynomial may or may not be Boolean function and similarly, a Boolean function may or may not be flat. A flat polynomial which satisfies the Perseval's identity i.e. Sum of the square of the Fourier coefficients is 1, may or may not be Boolean function.

Example: $f(x) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_1x_2$ [Put $x_1 = 1, x_2 = -1$]

1.1.1 Conjectures

In 1996, Friedgut and Kalai made the Fourier Entropy–Influence Conjecture [EG96] which says $\exists C \forall f, \mathbb{H}[\hat{f}^2] \leq C\mathbb{I}[f]$ i.e.

$$\Sigma_{S\subseteq[n]}\hat{f}(S)^2\log_2\left(\frac{1}{\hat{f}(S)^2}\right) \leq C\Sigma_{S\subseteq[n]}\hat{f}(S)^2|S|$$

Left side of the inequality represents the spectral entropy or Fourier entropy of f and measures how "spread out" $f's$ Fourier spectrum is. The right side of the inequality represents the total influence or average sensitivity of f. Both quantities have range between 0 and n. This conjecture also implies the famous KKL Theorem. The result showing that the FEI Conjecture holds for random DNFs is the only published progress on the FEI Conjecture since it was posed. Since then, there have been many significant steps taken in the direction of resolving the FEI conjecture.

The Fourier Entropy–Influence Conjecture would imply $\mathbb{H}[\hat{f}^2] \leq C.\mathcal{O}(\log m)$ from which one may take $K = \mathcal{O}(\frac{C}{\epsilon})$ in Mansour's Conjecture [Man95]. In 1994, Y. Mansour conjectured that for every DNF formula on n variables with t terms, there exists a polynomial p with $t^{\mathcal{O}\log(1/\epsilon)}$ non-zero coefficients such that $\mathbb{E}_{x\in\{0,1\}^n}[(p(x)-f(x))^2] \leq \epsilon$. Mansour's Conjecture is important because if it is true then the query algorithm of Gopalan, Kalai, and Klivans would agnostically DNF formulas under the uniform distribution to any constant accuracy in polynomial time. Establishing such a result is a major open problem in computational learning theory.

The FEI Conjecture is superficially similar to the well-known Logarithmic Sobolev Inequality [Gro75] for the Boolean cube which says, for $\{-1,1\}^n \to \mathbb{R}$, $Ent[f^2] \leq 2\mathbb{I}[f]$ where $Ent[f] = \mathbb{E}[f\log f] - \mathbb{E}[f]\log(\mathbb{E}[f])$. Note that FEI Conjecture requires $f : \{-1,1\}^n \to \{-1,1\}$ to be Boolean-valued, and it definitely fails for real-valued f.

The FEI Conjecture holds for "the usual examples" that arise in analysis of Boolean functions i.e. Parities (for which the conjecture is trivial), ANDs and ORs, Majority, Tribes, and Inner-Product-mod-2 function.There have been many works on proving the FEI conjecture for specific classes of Boolean functions. Assuming the FEI conjecture, a flat polynomial of degree d and sparsity $2^{\omega(d)}$ cannot $\frac{1}{3}-$approximate a Boolean function. It is not clear to us how to obtain the same conclusion unconditionally i.e. without assuming that the FEI conjecture is true and that's why the conjecture: No flat polynomial of degree d and sparsity $2^{\omega(d)}$ can $\frac{1}{3}-$approximate a Boolean function is posed.

1.2 Contribution of this thesis

The main contribution of this thesis is to find the Fourier Spectrum of various boolean functions.

In Chapter 1, We have given the definitions of various boolean functions such as AND, Majority, Parity, Inner Product, Linear Threshold Function(LTF). Also, in this chapter, we have defined the Influence and flat polynomial and have given various conjectures.

In Chapter 2, We have represented various notations and preliminaries where we have shown the linear algebra persepective for these boolean functions and we have also covered the Parseval's theorem, mean and variance of these boolean functions.

In Chapter 3, We have computed the Fourier Coefficients of various boolean functions by using direct computation and Lagrange Interpolation method. We have represented these boolean functions as a multilinear polynomial and verified by various examples for different values of n. This is useful in the Fourier analysis, for example, based on the first Fourier coefficient, we can tell whether the boolean function is balanced or not.

In Chapter 4, We have given a proof for the statement "A flat multilinear polynomial of degree $'d'$ and sparsity(number of monomials with non-zero Fourier-Coefficients) more than 2^d , can't be a boolean function." We also have written about the KKL Theorem and FEI Conjecture and its implication on the approximating polynomials.

In Chapter 5, We have written a conclusive note about the approximating polynomials and work along the direction of resolving FEI conjecture for some interesting classes of boolean functions.

Chapter 2

Notations and Preliminaries

2.1 Notations

We denote the set $\{1, 2, ..., n\}$ by $[n]$ and we can write the monomial corresponding to the subsets $S \subseteq [n]$ as $x^S = \Pi_{i \in S} x_i$ with $x^\phi = 1$ and we use the notation $\hat{f}(S)$ for the coefficients on monomial x^S in the multilinear representation of f. We also use the notation $\chi_S(x)$ for the character function which is defined as $\chi_S(x) = \Pi_{i \in S} x_i$ and therefore, we can also write f as $f(x) = \Sigma_{S \subseteq [n]} \hat{f}(S) \chi_S(x)$

We represent an inner product on the pair of functions $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$ as $< f, g >= 2^{-n} \Sigma_{x \in \{-1,1\}^n} f(x) g(x) = \mathbb{E}_{x \sim \{-1,1\}^n}[f(x)g(x)]$ where $x \sim \{-1, 1\}^n$ denotes that x is uniformly chosen random string from $\{-1, 1\}^n$. A boolean function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ has $< f, f >= 1$ i.e. a unit vector. There is one more notation which we use,

$$||f||_p = \mathbb{E}[|f(x)|^p]^{1/p}$$

and so, $||f||_2 = \sqrt{< f, f >}$.

2.2 Preliminaries

There is a Fourier Expansion theorem which says every function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ can be uniquely expressed as a multilinear polynomial as $f(x) = \Sigma_{S \subseteq [n]} \hat{f}(S) x^S$ and it is called the Fourier expansion of f and the real number $\hat{f}(S)$ is called the Fourier coefficients(Fourier Spectrum) of f on S. We can think of monomial x^S as a function on $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$.

The character function $\chi_S : \{-1, 1\}^n \rightarrow \{-1, 1\}$ is a boolean function which computes the logical parity or exclusive-or(XOR) of bits $(x_i)_{i \in S}$ and so, if we write $f = \Sigma_{S \subseteq [n]} \hat{f}(S) \chi_S$ then it means f can be written as a linear combination of parity functions over the real numbers.

The set of all functions $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ forms a vector space V and it is 2^n dimensional space and we can think of the functions in this vector space as vectors in $\mathbb{R}^{2^n}$ where we are stacking the 2^n values of $f(x)$ into a column vector in some fixed order.

Consider a boolean function Max_2. So,

x_1	x_2	$f(x_1, x_2)$
1	1	1
-1	1	1
1	-1	1
-1	-1	-1

$$f(x_1, x_2) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \in \mathbb{R}^4 \text{ can be written as : } \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$$

Here, on RHS, first vector corresponds to χ_ϕ, second vector corresponds to $\chi_{\{1\}}$, third vector corresponds to $\chi_{\{2\}}$ and fourth vector corresponds to $\chi_{\{1,2\}}$. The parity functions $(\chi_S)_{S\subseteq[n]}$ spans the vector space V, it means every function in this vector space V is a linear combination of parity functions (χ_S).

The above defined parity functions make the spanning set for the vector space V. Since, the number of parity functions is 2^n which is the dimension of vector space V and in fact, they are a linearly independent basis for vector space V and it justifies the uniqueness of the Fourier expansion.

The 2^n parity functions χ_S form an orthonormal basis for the vector space V of functions $\{-1,1\}^n \to \mathbb{R}$ i.e. $< \chi_S, \chi_T >= 1$ if $S = T$ and $< \chi_S, \chi_T >= 0$ if $S \neq T$ because $< \chi_S, \chi_T >= \mathbb{E}[\chi_S(x)\chi_T(x)]$

The "coordinates of f" in the χ_S direction can be represented as $< f, \chi_S >$ and the Fourier coefficient of f on S is given by $\hat{f}(S) =< f, \chi_S >= \mathbb{E}_{x\sim\{-1,1\}^n}[f(x)\chi_S(x)]$ because $< f, \chi_S >=< \Sigma_{T\subseteq[n]}\hat{f}(T)\chi_T, \chi_S >= \Sigma_{T\subseteq[n]}\hat{f}(T) < \chi_T, \chi_S >= \hat{f}(S)$. From this formula, we can easily calculate the Fourier coefficients.

Parseval's Theorem is used in Fourier analysis of Boolean function which says, For any $f : \{-1,1\}^n \to \mathbb{R}$, $< f, f >= \mathbb{E}_{x\sim\{-1,1\}^n}[f(x)^2] = \Sigma_{S\subseteq[n]}\hat{f}(S)^2$. So, if $f : \{-1,1\}^n \to \{-1,1\}$ is Boolean-valued then $\Sigma_{S\subseteq[n]}\hat{f}(S)^2 = 1$.The set $\{\hat{f}^2(S)\}_{S\subseteq[n]}$ is the Fourier distribution denoted by $\hat{f}$.

More Specifically, If we are given two functions $f, g : \{-1,1\}^n \to \mathbb{R}$, we can compute $< f, g >$ by taking the dot product of their co-ordinates in the orthonormal basis of the parities and we get the **Plancheral's Theorem** as a result which says, for any $f, g : \{-1,1\}^n \to \mathbb{R}$, $< f, g >= \mathbb{E}_{x\sim\{-1,1\}^n}[f(x)g(x)] = \Sigma_{S\subseteq[n]}\hat{f}(S)\hat{g}(S)$.

Proof is simple as we can write:

$$< f, g >=< \Sigma_{S\subseteq[n]}\hat{f}(S)\chi_S, \Sigma_{T\subseteq[n]}\hat{g}(T)\chi_T >= \Sigma_{S,T\subseteq[n]}\hat{f}(S)\hat{g}(T) < \chi_S, \chi_T >= \Sigma_{S\subseteq[n]}\hat{f}(S)\hat{g}(S)$$

We can also write $< f, g >$ as $Pr(f(x) = g(x)) - Pr(f(x) \neq g(x)) = 1 - 2\ dist(f, g)$ where $dist(f, g) = Pr_x(f(x) \neq g(x))$ is called the relative hamming distance. It is the fraction of inputs on which they disagree.

The mean of $f : \{-1, 1\}^n \to \mathbb{R}$ is $\mathbb{E}[f]$ and when it is zero then we say, f is balanced or unbiased. If f is defined as $f : \{-1, 1\}^n \to \{-1, 1\}$ is boolean-valued then mean $\mathbb{E}[f] = Pr(f = 1) - Pr(f = -1)$. So, we can say, f is unbiased iff it takes value 1 on exactly half of the points of the hamming cube.

If $f : \{-1, 1\}^n \to \mathbb{R}$ then $\mathbb{E}[f] = \hat{f}(\phi)$ because $\mathbb{E}[f] = \mathbb{E}[f \times 1] =< f, 1 >$ and $< f, \chi_S >= \hat{f}(S)$, If $S = \phi$ then $\chi_S = 1$ and so, $< f, 1 >= \hat{f}(\phi)$

The variance of $f : \{-1, 1\}^n \to \mathbb{R}$[real-valued boolean function] is defined as:

$$Var[f] =< f - \mathbb{E}[f], f - \mathbb{E}[f] >= \mathbb{E}[f^2] - \mathbb{E}[f]^2 = \mathbb{E}[f^2] - (\hat{f}(\phi))^2 =$$

$$< f, f > -(\hat{f}(\phi))^2 = \Sigma_{S\subseteq[n]}\hat{f}(S)^2 - (\hat{f}(\phi))^2 = \Sigma_{S\neq\phi}\hat{f}(S)^2.$$

A boolean-valued function f has variance 1 if it's unbiased and it has variance zero when it is constant.

Chapter 3

Fourier Analysis of Some Boolean Functions

3.1 Fourier Spectrum of $\mathbf{AND_n}$ Function

Theorem 3.1.1. *The Fourier coefficients of the function* $\mathbf{AND_n}$ *are as follows:*

- $\widehat{\mathbf{AND_n}}(\emptyset) = 1 - \frac{1}{2^{n-1}}$
- $\widehat{\mathbf{AND_n}}(S) = \frac{(-1)^{k+1}}{2^{n-1}}$, *for* $|S| = k$ *and* $S \neq \phi$

3.1.1 Proof of Theorem 3.1.1

We will prove Theorem 3.1.1 by computing the Fourier Coefficents of the function $\mathsf{AND_n}$.

Method 1: Direct computation

Here, character functions for different S are:

$\chi(\phi) = 1$
$\chi(\{1\}) = x_1$
$\chi(\{2\}) = x_2$
....
....
$\chi(n) = x_n$
$\chi(1, 2) = x_1x_2$
....
....
$\chi(\{1, 2, 3, ..., n\} = x_1x_2.....x_n)$

So, $\chi(S) = \Pi_{i\in S}x_i$

Now, according to the definition 1, Fourier coefficients will be,

$\hat{f}(\phi) = \frac{1}{2^n}\left((2^n - 1)(+1) - 1\right) = \frac{2^n-2}{2} = 1 - \frac{1}{2^{n-1}}$

$\hat{f}(\{1\}) = \frac{1}{2^n}\left(\frac{2^n}{2}(+1) + (\frac{2^n}{2} - 1)(-1) + 1\right) = \frac{1}{2^{n-1}}$

$\hat{f}(\{2\}) = \frac{1}{2^n}\left(\frac{2^n}{2^2}(+1) + \frac{2^n}{2^2}(-1) + \frac{2^n}{2^2}(+1) + (\frac{2^n}{2^2} - 1)(-1) + 1\right) = \frac{1}{2^{n-1}}$

$\hat{f}(\{1,2\}) = \frac{1}{2^n}\left(\frac{2^n}{2^2}(+1) + \frac{2^n}{2^2}(-1) + \frac{2^n}{2^2}(-1) + (\frac{2^n}{2^2} - 1)(+1) - 1\right) = \frac{-1}{2^{n-1}}$

$\hat{f}(\{1,3\}) = \frac{1}{2^n}\left(\frac{2^n}{2^3}(+1) + \frac{2^n}{2^3}(-1) + \frac{2^n}{2^3}(+1) + \frac{2^n}{2^3}(-1) + \frac{2^n}{2^3}(-1) + (\frac{2^n}{2^3} - 1)(+1) - 1\right) = \frac{-1}{2^{n-1}}$

Now, observe the pattern and why terms are canceling out, similarly we can write,

$\hat{f}(\{2,3\}) = \frac{-1}{2^{n-1}}$

Now, For $|S| = 3$,

$\hat{f}(\{1,2,3\}) = \frac{1}{2^n}\left(\frac{2^n}{2^3}(+1) + \frac{2^n}{2^3}(-1) + \frac{2^n}{2^3}(-1) + \frac{2^n}{2^3}(+1) + \frac{2^n}{2^3}(-1) + \frac{2^n}{2^3}(+1) + \frac{2^n}{2^3}(+1) + (\frac{2^n}{2^3} - 1)(-1)\right)$
$= \frac{1}{2^{n-1}}$

Now, here if we observe how the terms are canceling out in the expressions of Fourier coefficients, we can write the generalized Fourier coefficient for the AND_n functions as

$$\hat{f}(S) = \frac{(-1)^{k+1}}{2^{n-1}}$$

For $|S| = k$ and $S \neq \phi$

and $\hat{f}(\phi) = 1 - \frac{1}{2^{n-1}}$

Reason is that when $|S| = even$ then $\hat{f}(S) = \frac{1}{2^n}((-1)(+1) - 1) = \frac{-2}{2^n} = \frac{-1}{2^{n-1}}$ and in this expression all others terms will be cancelled out and only terms will be survived when all n boolean variables are even, so product will give "+1" and output will be "-1" and one "-1" will also be there due to $\frac{2^n}{2^k} - 1$ whose output is "$+1$", so $-1-1=-2$, and so, Fourier coefficient will be $\frac{-1}{2^{n-1}}$.

When $|S| = odd$ then "-1" odd number of times gives "-1" and output is "-1", so, product will be "$+1$ and so, Fourier coefficient for odd $|S|$ will be $\frac{2}{2^n} = \frac{1}{2^{n-1}}$.

Method 2: Interpolation

$\frac{(1+x_1)(1+x_2)...(1+x_n)}{2^n}(+1) + \frac{(1+x_1)(1+x_2)...(1-x_n)}{2^n}(+1) + \frac{(1+x_1)(1+x_2)...(1-x_{n-1})(1+x_n)}{2^n}(+1) + ...$
$+ \frac{(1-x_1)(1-x_2)(1-x_3)...(1+x_n)}{2^n}(+1) + ... + \frac{(1-x_1)(1-x_2)(1-x_3)...(1-x_n)}{2^n}(-1)$

Here, coefficient of "x_1" $= \frac{2^n}{2}\frac{1}{2^n} + (\frac{2^n}{2} - 1)(\frac{-1}{2^n}) + \frac{1}{2^n} = \frac{1}{2^{n-1}}$

coefficient of "x_1x_2" $= \frac{2^n}{4}\frac{1}{2^n} + \frac{2^n}{4}(\frac{-1}{2^n}) + \frac{2^n}{4}(\frac{-1}{2^n})(\frac{2^n}{4} - 1)(\frac{1}{2^n}) - \frac{1}{2^n} = \frac{-1}{2^{n-1}}$

Similarly, we can find the other Fourier coefficients which will be same as above.

Examples

Here, I have given 2 examples and we can verify the above formulae for Fourier coefficients for these 2 examples means whether it is working or not.

Example 3.1.1. *For $n = 3$,*

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	1	1	1
1	1	-1	1
1	-1	1	1
1	-1	-1	1
-1	1	1	1
-1	1	-1	1
-1	-1	1	1
-1	-1	-1	-1

Here, Fourier coefficients (according to the above definition 1) are:

$\hat{f}(\phi) = \frac{3}{4}$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = \frac{1}{4}$
$\hat{f}(\{1, 2\}) = \hat{f}(\{1, 3\}) = \hat{f}(\{2, 3\}) = \frac{-1}{4}$
$\hat{f}(\{1, 2, 3\}) = \frac{1}{4}$

Hence,

$$f(x) = \frac{3}{4} + \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_3 - \frac{1}{4}x_1x_2 - \frac{1}{4}x_1x_3 - \frac{1}{4}x_2x_3 + \frac{1}{4}x_1x_2x_3$$

Now, we can verify this multilinear polynomial by putting values of boolean variables x_1, x_2, x_3 from above truth table and check whether it is giving correct output value or not.

Example 3.1.2. *For $n = 4$,*

x_1	x_2	x_3	x_4	$f(x_1, x_2, x_3, x_4)$
1	1	1	1	1
1	1	1	-1	1
1	1	-1	1	1
1	1	-1	-1	1
1	-1	1	1	1
1	-1	1	-1	1
1	-1	-1	1	1
1	-1	-1	-1	1
-1	1	1	1	1
-1	1	1	-1	1
-1	1	-1	1	1
-1	1	-1	-1	1
-1	-1	1	1	1
-1	-1	1	-1	1
-1	-1	-1	1	1
-1	-1	-1	-1	-1

Here, Fourier coefficients (according to the above definition 1) are:

$\hat{f}(\phi) = \frac{7}{8}$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = \hat{f}(\{4\}) = \frac{1}{8}$
$\hat{f}(\{1,2\}) = \hat{f}(\{1,3\}) = \hat{f}(\{2,3\}) = \hat{f}(\{2,4\}) = \hat{f}(\{1,4\}) = \hat{f}(\{3,4\}) = \frac{-1}{8}$
$\hat{f}(\{1,2,3\}) = \hat{f}(\{1,2,4\}) = \hat{f}(\{2,3,4\}) = \hat{f}(\{1,3,4\}) = \frac{1}{8}$
$\hat{f}(\{1,2,3,4\}) = \frac{-1}{8}$

Hence,

$f(x) = \frac{7}{8} + \frac{1}{8}x_1 + \frac{1}{8}x_2 + \frac{1}{8}x_3 + \frac{1}{8}x_4 - \frac{1}{8}x_1x_2 - \frac{1}{8}x_1x_3 - \frac{1}{8}x_1x_4 - \frac{1}{8}x_2x_3 - \frac{1}{8}x_2x_4 - \frac{1}{8}x_3x_4 + \frac{1}{8}x_1x_2x_3 + \frac{1}{8}x_1x_2x_4 + \frac{1}{8}x_1x_3x_4 + \frac{1}{4}x_2x_3x_4 - \frac{1}{8}x_1x_2x_3x_4$

Now, we can verify this multilinear polynomial by putting values of boolean variables x_1, x_2, x_3, x_4 *from above truth table and check whether it is giving correct output value or not.*

3.2 Fourier Spectrum of Majority Function $\mathbf{MAJ_n}$

Theorem 3.2.1. *The Fourier coefficients of the Majority Function* $\mathbf{MAJ_n}$ *are as follows:*

- $\widehat{\mathbf{MAJ_n}}(S) = \begin{cases} 0 \textit{ if } |S| = even \\ (-1)^{\frac{k-1}{2}} \frac{\binom{\frac{n-1}{2}}{\frac{k-1}{2}}}{\binom{n-1}{k-1}} \frac{2}{2^n} \binom{n-1}{\frac{n-1}{2}} \textit{ if } |S| = odd \textit{ and } |S| = k \end{cases}$

3.2.1 Proof of Theorem 3.2.1

1) MAJ_n is a **symmetric** function because order does not matters if we permute the sequence of n boolean variables and so, output remains the same.
Hence, Fourier Coefficients $\widehat{MAJ_n(S)}$ only depends on $|S|$.

2) MAJ_n is also an **odd** function because if we change the values of each boolean variable from -1 to 1 or from 1 to -1 then outcome gets flipped.
Hence, $\widehat{MAJ_n(S)} = 0$ when $|S|$ is even because $x_1 x_2 ... x_n(even\ number\ of\ times) \Rightarrow \pm 1$ and flipping(negating) each boolean variable, we get the same outcome because after negation, product of "-ve" even number of times make "+ve".

So, we only need to determine $\widehat{MAJ_n(S)}$ when $|S| = odd$.

Claim 3.2.1. *For* $|S| = k(odd)$

$$\widehat{MAJ_n(S)} = (-1)^{\frac{k-1}{2}} \frac{\binom{\frac{n-1}{2}}{\frac{k-1}{2}}}{\binom{n-1}{k-1}} \frac{2}{2^n} \binom{n-1}{\frac{n-1}{2}}$$

Proof. When $n = odd$ then Majority function MAJ_n has value "-1" when $+1$ and -1 are equally divided from length $(n-1)$ by the definition of majority function. That's why we are choosing $\frac{n-1}{2}$ boolean variables out of $n-1$ boolean variables. So, we have included the term $\binom{n-1}{\frac{n-1}{2}}$ in the above formula and we do the same for output value "$+1$" because according to the definition we have to consider all input data points for each Fourier coefficients and according to the definition 1, we also have to divide whole thing by 2^n, so, we have included the term $\frac{2}{2^n}$.

Now, for $|S| = k$ we have to consider only "$k-1$" number of boolean variables out of $n-1$ boolean variables for majority (same logic as choosing $n-1$ from n), So, number of ways will be $\binom{n-1}{k-1}$. Now, we equally divide $+1$ and -1 of $\frac{n-1}{2}$ length sub-sequence of boolean variable and so, we included the term $\frac{\binom{\frac{n-1}{2}}{\frac{k-1}{2}}}{\binom{n-1}{k-1}}$ and at the end, for sign, we included the term $(-1)^{\frac{k-1}{2}}$. ∎

Examples

Example 3.2.1. *For $n = 3$,*

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	1	1	1
1	1	-1	1
1	-1	1	1
1	-1	-1	-1
-1	1	1	1
-1	1	-1	-1
-1	-1	1	-1
-1	-1	-1	-1

Method 1 (By Direct Computation) : *Here, Fourier coefficients (according to the above formula) are:*

$\hat{f}(\phi) = 0$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = (-1)^0 \frac{\binom{1}{0}}{\binom{2}{0}} \frac{2}{2^3} \binom{2}{1} = \frac{1}{2}$
$\hat{f}(\{1,2\}) = \hat{f}(\{1,3\}) = \hat{f}(\{2,3\}) = 0$
$\hat{f}(\{1,2,3\}) = (-1)^1 \frac{\binom{1}{1}}{\binom{2}{2}} \frac{2}{2^3} \binom{2}{1} = \frac{-1}{2}$

Hence,

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

which is same as above.

Method 2 (By Interpolation)

$f(x) = (\frac{1+x_1}{2})(\frac{1+x_2}{2})(\frac{1+x_3}{2})(+1) + (\frac{1+x_1}{2})(\frac{1+x_2}{2})(\frac{1-x_3}{2})(+1) +$
$(\frac{1+x_1}{2})(\frac{1-x_2}{2})(\frac{1+x_3}{2})(+1) + (\frac{1+x_1}{2})(\frac{1-x_2}{2})(\frac{1-x_3}{2})(-1) + (\frac{1-x_1}{2})(\frac{1+x_2}{2})(\frac{1+x_3}{2})(+1) +$
$(\frac{1-x_1}{2})(\frac{1+x_2}{2})(\frac{1-x_3}{2})(-1) + (\frac{1-x_1}{2})(\frac{1-x_2}{2})(\frac{1+x_3}{2})(-1) + (\frac{1-x_1}{2})(\frac{1-x_2}{2})(\frac{1-x_3}{2})(-1)$

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

3.3 Fourier Spectrum of Parity Function

Theorem 3.3.1. *The Fourier coefficients of the Parity Function are as follows:*

- $\hat{f}(S) = \begin{cases} 1, \textit{ if } S = \{1, 2, ..., n\} = [n] \\ 0, \ o/w \end{cases}$

3.3.1 Proof of Theorem 3.3.1

$\chi_S(x)$ **is the parity function which is** -1 **iff an odd number of the input bits are** -1.

Here, if $S = \{1, 2, ..., n\} = [n]$ then

$$\hat{f}(S) = 1$$

Because according to the definition , odd number of "-1" gives output as "-1". So, product $= (\underbrace{-1 \times -1 \times ..... \times -1}_{odd\ number\ of\ times}) \times (-1) = 1$

So, $\hat{f}(S) = \frac{1}{2^n}\Sigma_{S=[n]} f(x)\chi_S(x) = \frac{1}{2^n}[\underbrace{1 + 1 + 1 + ... + 1}_{2^n\ number\ of\ times}] = 1$

Since, $\Sigma(\hat{f}(S))^2 = 1$ and for $S = [n]$, $(\hat{f}(S))^2 = 1$, so, other Fourier coefficients will be zero (or) we can say, $\forall\ T \neq S$, and so, $\hat{f}(T) =< \chi_S, \chi_T >= 0$(By using orthogonality).

Hence,

$$f(x) = \Pi_{i=1}^n x_i$$

Now, if parity function is defined as :

$$\chi_{[n]} : \mathbb{F}_2^n \to \mathbb{F}_2$$

where, False = 0 and True = 1

Then,

$$\chi_{[n]}(x) = x_1 + x_2 + ... + x_n$$

which is 1 if odd number of $1s$ and 0 otherwise.

Here, for $a \in \mathbb{F}_2^n$, Indicator function will be :

$$\mathbb{I}_{\{a\}}(x) = \Pi_{i:a_i=1} x_i \Pi_{i:a_i=0}(1 - x_i)$$

So, Fourier expansion of $f(x)$ in this case will be :

$$f(x) = \Sigma_{a \in \mathbb{F}_2^n} f(a) \mathbb{I}_{\{a\}}(x)$$

$$f(x) = \Sigma_{a \in \mathbb{F}_2^n} f(a) \Pi_{i:a_i=1} x_i \Pi_{i:a_i=0}(1 - x_i)$$

Examples

Example 3.3.1. *Suppose,* $f : \{-1, 1\}^n \to \{-1, 1\}$

Then, For $n = 3$*,*

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
1	1	1	1
1	1	-1	-1
1	-1	1	-1
1	-1	-1	1
-1	1	1	-1
-1	1	-1	1
-1	-1	1	1
-1	-1	-1	-1

Now, ***By Interpolation****, we get,*

$(\frac{1+x_1}{2})(\frac{1+x_2}{2})(\frac{1+x_3}{2})(+1)+(\frac{1+x_1}{2})(\frac{1+x_2}{2})(\frac{1-x_3}{2})(-1)+(\frac{1+x_1}{2})(\frac{1-x_2}{2})(\frac{1+x_3}{2})(-1)+(\frac{1+x_1}{2})(\frac{1-x_2}{2})(\frac{1-x_3}{2})(1)+$
$(\frac{1-x_1}{2})(\frac{1+x_2}{2})(\frac{1+x_3}{2})(-1)+(\frac{1-x_1}{2})(\frac{1+x_2}{2})(\frac{1-x_3}{2})(+1)+(\frac{1-x_1}{2})(\frac{1-x_2}{2})(\frac{1+x_3}{2})(1)+(\frac{1-x_1}{2})(\frac{1-x_2}{2})(\frac{1-x_3}{2})(-1)$

$$f(x) = x_1 x_2 x_3$$

and Fourier coefficients will be :

$\hat{f}(\phi) = \frac{1}{2^3}[0] = 0$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = 0$
$\hat{f}(\{1, 2\}) = \hat{f}(\{2, 3\}) = \hat{f}(\{1, 3\}) = 0$
$\hat{f}(\{1, 2, 3\}) = 1$

Hence,

$$f(x) = x_1x_2x_3$$

Now, if

$$\chi_{[3]} : \mathbb{F}_2^3 \to \mathbb{F}_2$$

Then

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

Here, ***By Interpolation,*** *we get,*

$$f(x) = (1-x_1)(1-x_2)x_3 + (1-x_1)x_2(1-x_3) + x_1(1-x_2)(-x_3) + x_1x_2x_3$$

$$f(x) = x_1 + x_2 + x_3$$

3.4 Fourier Spectrum of Inner Product Function

Theorem 3.4.1. *The Fourier coefficients of the function Inner Product are as follows:*

- $\hat{f}(\phi) = \frac{1}{2^{2n}}\left(\frac{2^{2n-2^n}}{2}\right)$ *for* $f : \mathbb{F}_2^{2n} \to \mathbb{F}_2$

3.4.1 Proof of Theorem 3.4.1

Here, for any $f : \{0,1\}^n \to \mathbb{R}$,

Fourier coefficients are defined as:

$$\hat{f}(S) =< f, \chi_S >$$

$$\hat{f}(S) = \frac{1}{2^n}\Sigma_{x\in\{0,1\}^n} f(x)\chi_S(x)$$

$$\hat{f}(S) = \frac{1}{2^n}\Sigma_{x\in\{0,1\}^n} f(x)(-1)^{\Sigma_{i\in S} x_i}$$

Hence, Fourier expansion will be :

$$f(x) = \Sigma_S \hat{f}(S)\chi_S(x)$$

$$f(x) = \Sigma_S \hat{f}(S)(-1)^{\Sigma_{i\in S} x_i}$$

Now, here, For $f : \mathbb{F}_2^{2n} \to \mathbb{F}_2$, we define the **inner product mod-2** boolean function as :

$$f(x_1, x_2, ..., x_n, x_{n+1}, x_{n+2}, ..., x_{2n}) = x_1x_{n+1} + x_2x_{n+2} + ... + x_nx_{2n}$$

So, for this function Fourier coefficients will be :

$$\hat{f}(\phi) = \frac{1}{2^{2n}}\left(\frac{2^{2n-2^n}}{2}\right)$$

To get this formula, we have to count number of $1's$ and if there is 0 then there will 3 ways to get it for x_ix_j.

So, if $n = 1$, we can get it in only one way i.e. $x_1 = x_2$ to get 1. If $n = 2$, then number of ways to get $1s$ in f for $x_1x_3 + x_2x_4$ is 6 because $0 + 1 = 1$ will give 3 ways and $1 + 0$ will give 3 ways, so total 6 ways. If $n = 3$ then for $x_1x_4 + x_2x_5 + x_3x_6$, we get 1 if we can have 2 $0s$ *and* 1 $1s$ or 3 $1s$. So, total ways $= 3 \times 3 \times 3 = 27 + 1 = 28$.

So, observe the pattern of $1, 6, 28, ...$, we get the formula or we can also solve the $\binom{n}{1}3^{n-1} + \binom{n}{3}3^{n-3} + ... + \binom{n}{n}3^{n-n}$, we get the formula.

Now, to get the other Fourier coefficients, we have to use $\hat{f}(S) = \frac{1}{2^n}\Sigma_{x\in\{0,1\}^n} f(x)(-1)^{\Sigma_{i\in S} x_i}$. and then compute the Fourier expansion as :

$f(x) = \Sigma_S \hat{f}(S)(-1)^{\Sigma_{i\in S} x_i}$.

Here, if we consider $f : \{-1, 1\}^{2n} \to \{-1, 1\}$ and if we take the above definition of inner product then there are 2 possibilities :

1. $f(x_1, x_2, y_1, y_2) = f(x_2, x_1, y_2, y_1)$

2. $f(x_1, x_2, -y_1, -y_2) = -f(x_1, x_2, y_1, y_2)$

Both are correct But

- By 1, $f(1,1,1,-1) = f(1,1,-1,1)$

- By 2, $f(1,1,1,-1) = -f(1,1,-1,1)$

So $f(1,1,1,-1) = -f(1,1,1,-1) \Rightarrow f(1,1,1,-1) = 0$. But the only values allowed are 1 and -1.

Hence, we can't do this without either giving up one possibility out of two.

Examples

Example 3.4.1. *For $n = 1$,*

For $\mathbb{F}_2^{2n} \to \mathbb{F}_2$,

x_1	x_2	$f(x_1, x_2)$
0	0	0
0	1	0
1	0	0
1	1	1

*Here, **By Interpolation,***

$f(x) = x_1 x_2$.

Now, computing the Fourier coefficients :

$\hat{f}(\phi) = \frac{1}{2^2}(0+0+0+1) = \frac{1}{4}$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \frac{1}{4}(1 \times (-1)^1) = \frac{-1}{4}$
$\hat{f}(\{1,2\}) = \frac{1}{4}(1 \times (-1)^{1+1}) = \frac{1}{4}$.

Hence,

$$f(x) = \frac{1}{4} - \frac{1}{4}(-1)^{x_1} - \frac{1}{4}(-1)^{x_2} + \frac{1}{4}(-1)^{x_1+x_2}$$

Example 3.4.2. *For $n = 3$,*

For $\mathbb{F}_2^{2n} \to \mathbb{F}_2$,

x_1	x_2	x_3	x_4	$f(x_1, x_2, x_3, x_4)$
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	1
1	0	0	0	0
1	0	0	1	0
1	0	1	0	1
1	0	1	1	1
1	1	0	0	0
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Here, Fourier coefficients are :

$\hat{f}(\phi) = \frac{1}{16}(sum\ of\ values\ of\ f) = \frac{6}{16} = \frac{3}{8}$
$\hat{f}(\{1\}) = \hat{f}(\{2\}) = \hat{f}(\{3\}) = \hat{f}(\{4\}) = \frac{-2}{16} = \frac{-1}{8}$
$\hat{f}(\{1,2\}) = \frac{-2}{16} = \frac{-1}{8}$
$\hat{f}(\{1,3\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{1,4\}) = \frac{-2}{16} = \frac{-1}{8}$
$\hat{f}(\{2,3\}) = \frac{-2}{16} = \frac{-1}{8}$
$\hat{f}(\{2,4\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{3,4\}) = \frac{-2}{16} = \frac{-1}{8}$
$\hat{f}(\{1,2,3\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{1,2,4\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{1,3,4\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{2,3,4\}) = \frac{2}{16} = \frac{1}{8}$
$\hat{f}(\{1,2,3,4\}) = \frac{-2}{16} = \frac{-1}{8}$

Hence, Fourier expansion will be :

$f(x) = \frac{3}{8} - \frac{1}{8}(-1)^{x_1} - \frac{1}{8}(-1)^{x_2} - \frac{1}{8}(-1)^{x_3} - \frac{1}{8}(-1)^{x_4} - \frac{1}{8}(-1)^{x_1+x_2} + \frac{1}{8}(-1)^{x_1+x_3} - \frac{1}{8}(-1)^{x_1+x_4} - \frac{1}{8}(-1)^{x_2+x_3} + \frac{1}{8}(-1)^{x_2+x_4} - \frac{1}{8}(-1)^{x_3+x_4} + \frac{1}{8}(-1)^{x_1+x_2+x_3} + \frac{1}{8}(-1)^{x_1+x_2+x_4} + \frac{1}{8}(-1)^{x_1+x_3+x_4} +$

$\frac{1}{8}(-1)^{x_2+x_3+x_4} - \frac{1}{8}(-1)^{x_1+x_2+x_3+x_4}$.

Chapter 4

Boolean Function Approximation

4.1 Implication of FEI conjecture on approximating polynomials

KKL Theorem [JGN88] : $Inf_i(f) \geq \Omega(\frac{\ln n}{n}) Var(f)$[Special case: When f is balanced i.e. $\hat{f}(\phi) = 0$ and so, $Var(f) = 1$ because $Var(f) = 1 - \hat{f}(\phi)^2$]

Fourier-Entropy-Influence Conjecture ['96]:

$\forall f : \{-1,1\}^n \to \{-1,1\}$, $\mathbb{H}[\hat{f}^2] \leq C\mathbb{I}[f]$, Where, $\mathbb{H}[\hat{f}^2]$ is spectral(Shannon) entropy of f which is equal to $\hat{f}(S)^2 \log(\frac{1}{\hat{f}(S)^2})$ and C is a universal constant.
Informally, the FEI conjecture states that Boolean function whose Fourier distribution is well "spread out"(i.e. functions with larger Fourier entropy) must have significant Fourier weights on the high-degree monomials(i.e their total influence is large)

Conjecture: No flat polynomial of degree d and sparsity $2^{\omega(d)}$ can $\frac{1}{3}-$ approximate a boolean function[The degree of a function is the maximum d such that $\hat{f}(S) \neq 0$ for some set S of size d].

If f, g are boolean-valued functions then we say, functions f, g are $\epsilon-$ close if $dist(f, g) \leq \epsilon$ otherwise they are $\epsilon-$ far where $dist(f, g) = Pr_x(f(x) \neq g(x))$.

This conjecture is true by assuming FEI conjecture and it is also true for a class of polynomials without assuming FEI conjecture.

$\epsilon-$ approximation polynomial means, Say, we have a family of functions $\mathcal{F}_n = \{f : \{-1,1\}^n \to \{-1,1\}\}$ and consider subset of it as $\mathcal{B}_n \subseteq \mathcal{F}_n$ and we are interested in approximating $f \in \mathcal{B}_n$ by a $\hat{f} \in \mathcal{F}_n$ with smallest possible degree.
Let, $\epsilon \in (0, \frac{1}{2})$. The $\epsilon-$ approximate polynomial degree of $f \in \mathcal{B}_n$, denoted by $\hat{deg}_\epsilon(f)$, is the smallest integer k such that there exists $\hat{f} \in \mathcal{F}_n$ with $deg(\hat{f}) = k$ and $|\hat{f}(x) - f(x)| \leq \epsilon$, $\forall x \in \{-1,1\}^n$

If we restrict ourselves to the class of **block-multilinear** polynomial [An n-variate polynomial is said to be block-multilinear if the input variables can be partitioned into disjoint blocks such that every monomial in the polynomial has atleast one variable from each block]

Bohnenblust & Hille showed [SSM+20] that for every degree$-d$ block-multilinear $p : (\mathbb{R}^n)^d \to \mathbb{R}$,

$$\left(\Sigma_{i_1,i_2,...,i_d=1}^{n} |\hat{p}_{i_1,i_2,...i_d}|^{\frac{2d}{d+1}}\right)^{\frac{d+1}{d}} \leq C_d \max_{x^1,...x^d \in [-1,1]^n} |p(x^1,...,x^d)|$$

where C_d is constant and it depends on d.

The best upper bound on C_d is polynomial in d. Using the best upper bound on C_d in BH-inequality implies that a flat multilinear polynomial of degree-d and sparsity $2^{\omega(d\log d)}$ can't $\frac{1}{3}-$ approximate a boolean function.
FEI conjecture implies the following theorem :
If p is a flat block-multilinear polynomial of degree d and sparsity $2^{\omega(d)}$ then p can't approximate a boolean function. This theorem is also implied when C_d is assumed to be universal constant.

Proposition 4.1.1. *A flat multilinear polynomial of degree $'d'$ and sparsity(number of monomials with non-zero Fourier-Coefficients) more than 2^d , can't be a boolean function.*

Proof. To prove it by contradiction, try and assume that the statement is false, proceed from there and at some point you will arrive to a contradiction.

We have a multilinear polynomial $f(x)$ and since, it is a flat polynomial means all non-zero Fourier coefficients have same magnitude, say $'p'$. Now,
Suppose, $f(x)$ is a boolean function.

We can write $f(x)$ as $f(x) = p$(sum of monomials). Now, Since, $f(x)$ is $\{-1,1\}^n \to \{-1,1\}$, so, sum of monomials must be an integer. It might be positive, negative or zero but it can't be in fraction because putting values of -1 and $+1$ in monomials will generate an integer because multiplication of integers is integer and when we sum of all monomials then again, we get an integer. Now, since, f is a boolean function, it means $f(x) = p$(sum of monomials) must be either $+1$ or -1 and say, sum of monomials is $\pm s$ then $p = \frac{1}{s}$ where s is an integer.
So, p can be $1, \frac{1}{2}, \frac{1}{3}, ...$ but it can't be like $2, 3, ...$ (Taking absolute value only).

Now, since, f is a degree d multilinear polynomial, it means with d boolean variables, maximum number of possible monomials are $2^{d+1} - 1$ because if we add 1 more monomial then number of boolean variables will be more than d.
Degree d polynomial must have at least one monomial with d boolean variables. With d boolean variables, maximum number of monomials, we can make as 2^d i.e. $\binom{d}{0} + \binom{d}{1} + ... + \binom{d}{d}$, where each term shows: total number of zero degree monomial, total number of one degree

monomials,..etc. A flat polynomial with degree d and sparsity $\leq 2^d$ may or may not be a boolean function but here our assumption was $f(x)$ is a degree d flat multilinear polynomial with values -1 and $+1$ with more than 2^d number of monomials and we have to contradict it to prove the statement.

With degree d and 2^d number of monomials, minimum number of variables are d and maximum number of variables are $d2^d$. For Sparsity more than 2^d, Then, Either 2^d number of monomials have more than d variables or when 2^d number of monomials have maximum d boolean variables and say, $sum = s$ then adding one or more monomial with new variable will give sum as $s+t$ and $s-t$ with $t > 0$, $s \neq 0$ for at least 2 instances out of $2^{>d}$ instances because if we include new variables then with the positive sign of new variables, it will give new value and with the negative sign of new variables, it will give another new value in magnitude but our assumption was that f is boolean and for that each value of f must have same magnitude, say s and so, $p = \frac{1}{s}$ to get value as $+1$ or -1 for all values of f but here, we are getting 2 different magnitude values for at least 2 instances of f. So, Contradiction. ■

Let's understand it with some examples:

Example 4.1.1. *Suppose, degree-2 flat multilinear polynomial $f(x) = \frac{1}{3}(1+x_1+x_2-x_1x_2+x_3)$. Now, if we take $2^d = 4$ monomials as $1, x_1, x_2, x_1x_2$ and if we assign $x_1 = -1, x_2 = -1$ then $1+x_1+x_2-x_1x_2 = -2$ and so, for sparsity more than $2^d = 4$ i.e considering monomials $1, x_1, x_2, x_1x_2, x_3$, and for input variables $\{x_1, x_2, x_3\}$, out of $2^3 = 8$ instances, there must be 2 possible values of newly introduced variable x_3 i.e. -1 and $+1$ and so, we get the value of $f(x) = -2+1 = -1$ when $x_3 = 1$ and $f(x) = -2-1 = -3$ when $x_3 = -1$ and these 2 instances must be there but for $f(x)$ to be a boolean function, sum of monomials must be same which is not the case here and hence, it can't be an boolean function.*

If we consider $2^d = 4$ monomials as $x_1, x_2, -x_1x_2, x_3$, Now, for more than $2^d = 4$ monomials i.e. $1, x_1, x_2, x_1x_2, x_3$, same logic as above applies. Think it as we initially have $1, x_1, x_2, x_1x_2$ monomials and introdued a new variables which gives 2 different magnitude values of f.

Example 4.1.2. *Consider a degree $-$ 2 flat multilinear polynomial as:*
$f(x) = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3 + \frac{1}{3}x_1x_2 + \frac{1}{3}x_2x_3 = \frac{1}{3}(x_1 + x_2 + x_3 + x_1x_2 + x_2x_3)$
Now, again take the monomials which includes 2 boolean variables i.e. x_1, x_2, x_1x_2 . It has maximum sum $= 3$ when $x_1 = 1, x_2 = 1$ and minimum sum $= -1$. Now, For sparsity more than $2^d = 4$, if we take $x_1 = 1, x_2 = 1$ then with newly introduced variable x_3, $f(x) = 3+2 = 5$ when $x_3 = 1$ and $f(x) = 3-2 = 1$ when $x_3 = -1$. And also, for $x_1 = -1$ and $x_2 = 1$ and with $x_3 = \pm 1$, $f(x) = -1+2 = 1$ and $f(x) = -1-2 = -3$.

Since, $f(x)$ is having 2 different magnitude values due to $s+t$ and $s-t$ for $t > 0, s \neq 0$ as we have discussed above and so, $f(x)$ can't be a boolean function.
Note: *With 5 monomials and possible summation values are $\pm 1, \pm 3, \pm 5$ because we have*

5 places which we have to fill with $+1$ or -1.

Example 4.1.3. *Consider, a degree $-$ 2 flat monomial as:*
$f_(x) = x_1x_2 + x_2x_3 + x_2x_4 + x_3x_4 + x_3x_5 + x_4x_5$
Now, take the $2^d = 4$ monomials as: $x_1x_2, x_2x_3, x_2x_4, x_3x_4$ and if we introduce a new variable x_5 then considering the sparsity more than $2^d = 4$ which has 6 monomials i.e. $x_1x_2, x_2x_3, x_2x_4, x_3x_4, x_3x_5, x_4x_5$. Now, for $x_1 = x_2 = x_3 = x_4 = -1, f(x) = 4 + 2 = 6$ when $x_5 = -1$ and $f(x) = 4 - 2 = 2$ when $x_5 = 1$. Also, instead of 6 monomials if there were 5 monomials then also f has valuese $4 + 1 = 5$ and $4 - 1 = 3$.

So, based on the pattern, for more than 2^d monomials, f has 2 different magnitude values in the form of $s + t$ and $s - t$ where $t > 0, s \neq 0$ for at least 2 different instances of f.

Some More Analysis:

(1) Consider one more *degree* $-$ 2 multilinear flat polynomial with degree d as :
$f(x) = \frac{1}{2} + \frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_1x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_1x_3 = \frac{1}{2}(1 + x_1 + x_2 + x_1x_2 + x_3 - x_1x_3)$
Now, as above, if we consider $2^d = 4$ monomials with 2 distinct variables x_1, x_2 as $1 = x_1^0, x_2^0, x_1, x_2, x_1x_2$ and by introducing one new variable x_3 and taking more than $2^d = 4$ monomials as $1, x_1, x_2, x_1x_2, x_3, x_1x_3$.

Observe that $s = 0$ or $s = 4$. $s = 4$ is possible when $x_1 = 1, x_2 = 1$ and then with $x_3 = \pm 1, f(x) = 4 + 0$ and $f(x) = 4 - 0$, so here $t = 0$ and since, $s = 0$ and $t = 0$ are possible we are unable to detect whether it is a boolean function or not. But if we assume, a *degree* $- d$ multilinear flat(all Fourier coefficients are same and say, p) polynomial $f(x)$ is a boolean function then According to Parseval's identity for sparsity $= 2^d$,
$2^dp^2 = 1 \Rightarrow p = \pm\frac{1}{2^{d/2}}$

Hence, if magnitude of the Fourier coefficients are $\frac{1}{2^{d/2}}$ then *degree* $- d$ flat multilinear polynomial with sparsity more than 2^d , can't be a boolean function.
Therefore, for the above example, $p = \frac{1}{2} = \frac{1}{2^{2/2}}$ so, it can't be a boolean function.

(2) Ignoring p and analyzing the value of sum of monomials, $'s'$.
With d boolean variables, sparsity more than 2^d means number of monomials in $f(x)$ is from 1 to $2^{d+1} - 1$ because for sparsity 2^{d+1}, a monomial of *degree* $- d + 1$ must be there and so, we can add monomial terms from $2^d + 1$ to $2^{d+1} - 1$ in the previous f to get f of degree d. So, for d boolean variables, $f(x)$ contains number of monomials from $2^d + 1$ to $2^{d+1} - 1$ for sparsity $> 2^d$ with *degree* $- d$. With $\leq 2^d$ number of monomials, $f(x)$ has minimum vallue as $1 - 2^d$ and max value as 2^d.

Now, number of terms between $2^d + 1$ to $2^{d+1} - 1$ are $2^{d+1} - 1 - 2^d - 1 + 1 = 2^d - 1$
Hence, With d boolean variables, for *degree* $- d$ flat multilinear polynomial: Maximum value $= 2^d + 2^d - 1 = 2^{d+1} - 1$ and Minimum value $= -2^d + 1 - (2^d - 1) = -2^{d+1} + 2$

Chapter 5

Discussion and Conclusion

We have shown the Fourier analysis of various Boolean functions and proved that a flat multilinear polynomial of degree d and sparsity (number of monomials with non-zero Fourier-Coefficients) more than 2^d, can't be a Boolean function. Along the same direction, one may try to resolve the conjecture "a flat polynomial of degree d and sparsity $2^{\omega(d)}$ cannot $\frac{1}{3}-$approximate a Boolean function."

Since the general Fourier Entropy–Influence Conjecture seems difficult to resolve, so, it might be possible that one may try to prove it for additional interesting classes of Boolean functions for example linear threshold functions for a possibly tractable case as Andrew Wan suggested in his paper "Talk at the Center for Computational Intractability[2010]".

Bibliography

[EG96] Friedgut E. and Kalai G. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society*, 124(10):2993–3002, 1996.

[Gro75] Leonard Gross. Logarithmic sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975.

[JGN88] Kahn J., Kalai G., and Linial N. The influence of variables on boolean functions. *In Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 68–80, 1988.

[Man95] Yishay Mansour. An $n^{O(\log\log n)}$ learning algorithm for dnf under the uniform distribution. *Journal of Computer and System Sciences (JCSS)*, 50(3):543–550, 1995.

[SSM+20] Arunachalam S., Chakraborty S., Koucký M., Saurabh N., and de Wolf R. Improved bounds on fourier entropy and min-entropy. *International Symposium on Theoretical Aspects of Computer Science*, 2020.

References

1. Srinivasan Arunachalam, Sourav Chakraborty, Michal Koucký, Nitin Saurabh, Ronald de Wolf, Improved bounds on Fourier entropy and Min-entropy, *International Symposium on Theoretical Aspects of Computer Science* (March 2020).

2. O'Donnell, R., Analysis of Boolean Functions, *Cambridge University Press* (2014).