ON THE FOURIER ANALYSIS OF BOOLEAN FUNCTIONS AND THE LINIAL-MANSOUR-NISAN THEOREM

EDWARD KIM

Abstract. We briefly summarize the basic tools of Fourier Analysis applied to the finite abelian group $\mathbb{Z}_2^n$. By viewing boolean functions $f : \{0,1\}^n \to \{0,1\}$ as square-integrable maps in the Hilbert space $L^2(\mathbb{Z}_2^n)$ under the uniform probability measure, we derive the Linai-Mansour-Nisan theorem concerning the tail of the Fourier spectrum of boolean functions in AC^0. Finally, we highlight some applications of the LMN theorem in the context of sensitivity and influence of functions in AC^0.

1. Fourier Analysis of Boolean Functions

1.1. **Characters.**

We introduce some stepping stones defining the tools used in the Fourier Analysis of boolean functions. Many of these tools can be defined using techniques from Representation Theory of Finite Groups, but for the sake of briefity, we will not delve into those concepts here in detail. Instead, we will provide remarks and snippets on relevant topics along with references to some literature at the end of the section. Our presentation follows that of [2].

Let us first define some general concepts for any finite abelian group G. A group homomorphism from $\chi : G \to \mathbb{C}^*$ is called a *character* of G where $\mathbb{C}^* = \mathbb{C}/\{0\}$ is the multiplicative group of the complex numbers. We call the homomorphism $\chi_0 : G \to \mathbb{C}^*,\ \chi_0 = 1$ the *trivial character* of G. By definition, for $a, b \in G$,

$$\chi(a+b) = \chi(a)\chi(b) \tag{1.1}$$

for all characters χ.

Consider the set whose elements are the *characters* of G and whose binary operation is *pointwise multiplication* of complex-valued functions on G. This set is also an *abelian group* by the following theorem:

Theorem 1.1. *If G is a finite abelian group, then the characters of G form an abelian group under pointwise multiplication.*

Proof. We first show that if χ, ϕ are characters of G, then $\chi\phi$ is also a character of G. But this immediately follows from 1.1:

$$\chi\phi(a+b) = \chi(a)\chi(b)\phi(a)\phi(b) = [\chi(a)\phi(a)][\chi(b)\phi(b)] = \chi\phi(a)\chi\phi(b)$$

Hence, $\chi\phi$ is also a character of G. Furthermore, $\chi^{-1} = \frac{1}{\chi} = \overline{\chi}$ is also a character of G since

$$\chi^{-1}(a+b) = \frac{1}{\chi(a+b)} = \frac{1}{\chi(a)\chi(b)} = \frac{1}{\chi(a)}\frac{1}{\chi(b)} = \chi^{-1}(a)\chi^{-1}(b)$$

showing that the set contains inverses. □

Remark 1.2. We abuse notation a bit and just note that χ^{-1} is not the inverse of the χ as a *map*. Rather, χ^{-1} will refer to the *group inverse* of character χ from hereon.

Denote this set as $\hat{G}$. The group $\hat{G}$ is known as the *character group* of G. It also will turn out that there is a bijective map between G and $\hat{G}$ such that the map is also a *group isomorphism.* We will prove this later on.

Let us now consider the simplest finite abelian group, $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ for some $n \in \mathbb{N}$. It follows from basic properties of cyclic groups that the only characters of $\mathbb{Z}_n$ are:

$$\chi_j(x) = e^{2\pi ijx/n} \quad j \in [n],\ x \in \mathbb{Z}_n \tag{1.2}$$

We can decompose any finite abelian group G into a direct product of finite cyclic groups $G \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \cdots \times \mathbb{Z}_{n_k}$. If we decompose $x = \sum_{i \in [k]} x_i$ for $x_i \in \mathbb{Z}_{n_i},\ x \in G$, then the characters of G are defined as

$$\chi_a(x) = \prod_{i \in [k]} e^{2\pi i a_k x_k / n_k} \tag{1.3}$$

where x_i refers to the i^{th} index of x.

We see that every element $a \in G$ has an unique associated character χ_a. Furthermore

$$\chi_{a+b}(x) = \chi_a(x)\chi_b(x) \tag{1.4}$$

To define the Fourier transform, we must consider the Hilbert space $L_2(G)$ where G endowed with the uniform probability measure i.e the discrete measure mapping each subset $H \subseteq G$ to $\frac{|H|}{|G|}$. These will be the complex-valued maps $f : G \to \mathbb{C}$ such that

$$\frac{1}{|G|} \sum_{x \in G} |f(x)|^2 < \infty \tag{1.5}$$

which yields the standard inner product for maps $f, g \in L_2(G)$:

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{x \in G} f(x)\overline{g(x)} = \mathbb{E}_x f(x)\overline{g(x)} \tag{1.6}$$

This induces the norm $||f||_2^2 = \langle f, f \rangle = \mathbb{E}_x |f(x)|^2$ on $L_2(G)$. It is simple to check that the characters of G indeed lie in $L_2(G)$. However, we can prove something stronger: the characters of G, χ_a form an *orthonormal basis* of $L_2(G)$. First, we prove a lemma:

Lemma 1.3. *Let χ be a non-trivial character of G. Then $\sum_{x \in G} \chi(x) = 0$*

Proof. Given any non-trivial character χ, the following holds for all $y \in G$:

$$\chi(y) \sum_{x \in G} \chi(x) = \sum_{x \in G} \chi(y + x) = \sum_{x \in G} \chi(x)$$

If $\sum_{x \in G} \chi(x) \neq 0$, then this would force $\chi(y) = 1$ for all $y \in G$. However, this would contradict the non-triviality of the χ. □

Theorem 1.4. *The characters $\chi \in \hat{G}$ form an orthonormal basis for $L_2(G)$*

Proof. We first note that $\langle \chi, \chi \rangle = \mathbb{E}_x |\chi(x)|^2 = 1$ since the image of character χ lies on the complex unit circle by the discussions above. Now given two non-trivial characters χ, ϕ

$$\langle \chi, \phi \rangle = \mathbb{E}_x \chi(x)\overline{\phi(x)} = \frac{1}{|G|} \sum_{x \in G} \chi\phi^{-1}(x) = 0$$

The last equality follows from Lemma 1.3 since $\chi\phi^{-1}$ must also be a non-trivial character if χ, ϕ are non-trivial characters. The case where one of the characters is trivial also follows from the lemma. Finally, observe that $L_2(G)$ is $|G|$-dimensional as a vector space by considering the dual of G. We know that there are at least $|G|$-characters, showing that the characters span $L_2(G)$. However, from the lemma, the characters must also be pairwise orthgonal to each other. This yields that the characters of G form an *orthonormal basis* of $L_2(G)$ and that there must be exactly $|G|$ characters in $\hat{G}$. Finally, by using 1.4, it follows that the map $a \mapsto \chi_a$ must be a group isomorphism, so $G \cong \hat{G}$. □

To finish, we restrict our attention to the finite abelian group $\mathbb{Z}_2^n$ as a natural group to define boolean functions $f : \{0,1\}^n \to \{0,1\}$. By the observations above, we conclude that our characters for $\mathbb{Z}_2^n$ are defined as:

$$\chi_a(x) = (-1)^{\sum_{i\in[n]} x_i a_i} = (-1)^{\sum_{i\in[n],\ a_i=1} x_i} \tag{1.7}$$

where x_i is the i^{th} bit of $x \in \{0,1\}^n$.

Notation 1.5. Instead of considering $x \in \{0,1\}^n$ as a bit-string, we would like to use the following notation: Let $S \subseteq [n]$ be such that $S_x = \{i \mid x_i = 1\}$. As there is a bijective correspondence between all such subsets S_x and $x \in \{0,1\}^n$, we will sometimes identify bit strings with their subset counterparts for notational convenience. This changes the notation for characters of $\mathbb{Z}_2^n$ as below

$$\chi_A(x) = (-1)^{\sum_{i\in A} x_i}, \quad A \subseteq [n]$$

1.2. **The Fourier Transform.** We have shown above that the members of the character group $\hat{G}$ form an orthonormal basis of $L_2(G)$. This bestows us the power of decomposing any $f \in L_2(G)$ into its orthogonal counterparts. The restriction to the case where $G = \mathbb{Z}_2^n$ is what motivates us as any n-ary boolean function $f : \{0,1\}^n \to \{0,1\}$ must lie in $L_2(\mathbb{Z}_2^n)$. As before, we begin by considering the general case for any finite abelian group G

Let $\hat{f} : \hat{G} \to \mathbb{C}$ be the complex-valued function defined as:

$$\hat{f}(\chi_a) = \langle f, \chi_a\rangle,\ a \in G \tag{1.8}$$

The function $\hat{f}$ is deemed as the *Fourier Transform of* f. By iterating this for all characters, we get a direct sum decomposition for any $f \in L_2(G)$ which we call the *Fourier Inversion Formula*

$$f = \sum_{a\in G} \hat{f}(\chi_a)\chi_a \tag{1.9}$$

It can be shown that the map $f \mapsto \hat{f}$ is linear by the left-linearity of the inner product. This linear map is the *Fourier Transform on G* and (1.4) shows that the decomposition is unique. Hence the Fourier Transform can be viewed as a linear isomorphism $\hat{}: L_2(G) \to L_2(\hat{G})$. The complex value $\hat{f}(\chi_a)$ is called the *Fourier coefficient* associated to χ_a.

Notation 1.6. We will sometimes denote $\hat{f}(a)$ for $\hat{f}(\chi_a)$.

From the above definition, we directly calculate that

$$\hat{f}(0) = \langle f, \chi_0\rangle = \mathbb{E}_x f(x) \tag{1.10}$$

Additionally, a simple application of the Fourier Transform yields an useful identity which will be instrumental in our formulation of the Fourier tail of a boolean function.

Theorem 1.7. *(Parseval's Identity) Let $f \in L_2(G)$. Then*

$$||f||_2^2 = \sum_{a\in G} |\hat{f}(a)|^2$$

Proof.

$$||f||_2^2 = \langle f, f\rangle = \langle \sum_{a\in G} \hat{f}(a)\chi_a, \sum_{b\in G} \hat{f}(b)\chi_b\rangle = \sum_{a,b\in G} \hat{f}(a)\overline{\hat{f}(b)}\langle \chi_a, \chi_b\rangle = \sum_{a\in G} |\hat{f}(a)|^2$$

The last equality is just the orthonormality of the characters (1.4). □

By combining Parseval's Identity with the observation that $||\hat{f}||_2^2 = \langle \hat{f}, \hat{f}\rangle = \frac{1}{|G|}\sum_{a\in G} |\hat{f}(a)|^2$, we arrive at the *Plancheral Formula*

Theorem 1.8. *(Plancheral Formula)*

$$||\hat{f}||_2 = \sqrt{|G|}\cdot ||f||_2$$

Example 1.9. *(Fourier coefficients for Parity)* Let $G = \mathbb{Z}_2^n$ and $f = \oplus_n$, the parity function on n bits. We wish to calculate the Fourier coefficients of f. We begin by calculating $\hat{f}(0^n)$ and $\hat{f}(1^n)$

$$\hat{f}(0^n) = \frac{1}{2^n}\sum_{x\in G} f(x) = \frac{2^{n-1}}{2^n} = \frac{1}{2} \tag{1.11}$$

$$\hat{f}(1^n) = \frac{1}{2^n}\sum_{x\in G} (-1)^{|x|} f(x) = \frac{-2^{n-1}}{2^n} = -\frac{1}{2} \tag{1.12}$$

Now for $a \neq 0^n, 1^n$, there must exist two indices i, j such that $a_i = 0$ and $a_j = 1$. Define e_i, e_j to be the n-bit strings such that all indices are set to zero except for the i^{th}, j^{th} places respectively. Note that:

$$\hat{f}(a) = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{\chi_a(x)} = \frac{1}{|G|}\sum_{x\in G} f(x + e_0 + e_1)\chi_a(x + e_0 + e_1)$$

since translating by $e_0 + e_1$ rearranges the lefthand sum. This yields the equality:

$$\hat{f}(a) = \frac{1}{2}\mathbb{E}_x[f(x)\chi_a(x) + f(x + e_0 + e_1)\chi_a(x + e_0 + e_1)]$$

As flipping bits at indices i, j does not change the parity for any $x \in \{0,1\}^n$, $f(x) = f(x+e_0+e_1)$. If $x_j = 1$, then adding e_j will flip the bit to $x_j = 0$. Recalling the definition of the character $\chi_a(x) = (-1)^{\sum_{i\in[n],a_i=1}}$ shows that a factor of -1 disappears from the product since $a_j = 1$. Similarly, if initially $x_j = 0$, then the product gains a factor of -1. This proves that $\chi_a(x + e_0 + e_1) = -\chi_a(x)$ for all $x \in \{0,1\}^n$ which in turn proves that $\hat{f}(a) = 0$ for all such $a \neq 0^n, 1^n$. By Parseval's identity:

$$|| \oplus_n ||_2 = \frac{1}{\sqrt{2}}$$

Example 1.10. *(Fourier coefficients for Maj$_3$)* We proceed through a similar analysis for the parity case. Let $f = \mathrm{Maj}_3$.

$$\hat{f}(0^3) = \mathbb{E}_x[f(x)] = \frac{1}{2} \tag{1.13}$$

$$\hat{f}(\{001, 010, 100\}) = -\frac{1}{4} \tag{1.14}$$

$$\hat{f}(\{011, 110, 101\}) = 0 \tag{1.15}$$

$$\hat{f}(1^3) = \frac{1}{8} \sum_{x \in \{0,1\}^3} (-1)^{|x|} f(x) = \frac{1}{4} \tag{1.16}$$

There is an exact formula for calculating the Fourier coefficients for Maj$_n$. It can be found in [6] page 108.

We end with a definition. Recall that we will identify n-bit strings with subsets $S \subseteq [n]$.

Definition 1.11. Let $f : \{0,1\}^n \to \{0,1\}$ be an n-ary boolean function. The *Fourier degree* of f, denoted by $\mathrm{def}_{\mathcal{F}}(f)$, is the largest $|S|$ such that $\hat{f}(S) \neq 0$.

It's worth stressing that the Fourier Transform here is over the group $\mathbb{Z}_2^n$.

Remark 1.12. *(Further Topics)* There are still many essential concepts in the Fourier Analysis of boolean functions which we have not covered in this chapter. For instance, one can endow a multiplication operator to $L_2(G)$ called the *convolution operator*

$$(f * g)(x) = \frac{1}{|G|} \sum_{\substack{h,j \in G \\ hj = x}} f(h)g(j) = \frac{1}{|G|} \sum_{y \in G} f(xy^{-1})g(y), \quad f, g \in L_2(G)$$

The convolution $f * g$ can be interpreted as the average of $f(h) \cdot g(j)$ over all elements $h, j \in G,\ h + j = x$. Many natural, interesting functions can be defined through the convolution of two boolean functions. For instance, let $\mathbf{1}_{B_r}$ be the indictator function for a closed ball centered at 0 of radius r in respect to the Hamming metric on $\mathbb{Z}_2^n$. Then for any n-ary boolean function $f : \{0,1\}^n \to \{0,1\}$, $f * \mathbf{1}_{B_r}(x)$ will be the expected value of f over the closed ball of radius r centered at x. Additionally, the convolution operator satisfies many fundamental algebraic properties such as associativity and commutativity, turning $L_2(G)$ into a commutative $\mathbb{C}$-algebra. For a more detailed introduction to these topics, see [2],[6]. For a more general approach to harmonic analysis over groups, see [8], [1].

Remark 1.13. *(Generalizations of Fourier Analysis)* The above properties on characters can be shown through techniques in Representation Theory. For instance, we know representations of finite groups are group homomorphisms into the group of linear automorphisms of a suitable finite-dimensional complex vector space $\rho : G \to GL(\mathbb{C}^n)$. The *characters* of ρ are defined as $\chi(x) = Tr(\rho(x))$ where Tr is the trace operator. It turns out that the representations of finite abelian groups are one-dimensional i.e isomorphic to $\mathbb{C}$. Thus, $\rho(x)$ will just be multiplying $\mathbb{C}$ by some scalar $\lambda_x \in \mathbb{C}^*$. In this case, we can show that indeed for any $a, b \in G$ and χ any character of G:

$$\chi(a + b) = Tr(\rho(a + b)) = Tr(\rho(a) \circ \rho(b)) = \lambda_a \lambda_b = Tr(\rho(a))Tr(\rho(b)) = \chi(a)\chi(b)$$

Note that this relation does not hold for arbitrary finite groups since the trace is generally not multiplicative. The genealizations of Fourier Analysis generally lie in *Harmonic Analysis.* For more information on the Representation Theory of Finite Groups, see [9], [13].

2. Linial-Mansour-Nisan Theorem

We now present the LMN (Linial-Mansour-Nisan) Theorem concerning the Fourier tail of boolean functions in AC^0.

Definition 2.1. Let k be a positive integer and $f : \mathbb{Z}_2^n \to \mathbb{C}$ a complex-valued function on $\mathbb{Z}_2^n$. We define:

$$f^{\leq k} := \sum_{|S| \leq k} \hat{f}(S)\chi_S$$

The notation symbols $f^{=k}, f^{\geq k}$ are defined in the same manner.

Let us state the relevant theorem first before proceeding further

Theorem 2.2. *(Linial, Mansour, Nisan [4]) Let f be a boolean fucntion computed by a circuit of depth d and size M and let t be any non-negative integer. Then*

$$\sum_{|S|>t} |\hat{f}(S)|^2 \leq 2M2^{-t^{1/d}/20} \tag{2.1}$$

The theorem reveals that the t-tails of the Fourier spectrum, i.e strings indexed by sets $|S| > t$, become exponentially small in t for boolean functions in AC^0. To prove the theorem, we need version of the Håstad's switching lemma shown in [3].

Theorem 2.3. *(Håstad, [3]) Let f be given by a CNF-formula where each clause has size at most t, and choose a random restriction ρ with parameter p such that $Pr[\rho(x_i)] = p$ for all input variables x_i. With probability of at least $1 - (5pt)^s$, f_ρ can be expressed as a DNF formula where each clause has size of at most s, and the clause all accept disjoint sets of inputs i.e no string $x \in \{0,1\}^n$ satisfies more than one clause.*

From Håstad's switching lemma, we derive two integral corollaries:

Corollary 2.3.1. *Let f be a boolean function computed by a CNF of bottom fan-in of at most t, and ρ is a p-random restriction, then*

$$Pr[deg_{\mathcal{F}}(f_\rho) > s] < (5pt)^s \tag{2.2}$$

Proof. By Håstad's switching lemma, we can convert the CNF computing f, into a DNF of bottom fan-in of at most s with probability of at least $1 - (5pt)^s$, such that the clauses accept disjoint sets of inputs. Suppose that we had a subset $S \subseteq [n]$ such that $|S| > s$. Pick one such $\wedge$-clause C and let $x_C \subseteq [n]$ denote the bit-indices of the input variables. By our assumptions, $|x_C| \leq s$. Now w.l.o.g set S to be some subset of $[n]$ such that $x_C \subset S$ and $|S| = |x_C| + 1$. Since each $\wedge$-clause in the DNF accepts disjoint sets of inputs, they cannot contribute to the coefficient. Calculating the Fourier coefficients shows us that $\hat{f}(S) = 2, -2 \neq 0$, so $\deg_{\mathcal{F}}(f_\rho) > s$. □

Corollary 2.3.2. *Let f be a boolean function computed by a circuit of size M and depth d. Then*

$$Pr[deg(f_\rho) > s] \leq M2^{-s}$$

where ρ is a random restriction where $p = \frac{1}{10^d s^{d-1}}$

Proof. First, we will repeatedly iterate the switching lemma to reduce the depth of f through random restrictions such that the bottom fan-in at each step is at most s. This is accomplished by sampling the first round of restrictions ρ_0 to parameter $p_0 = \frac{1}{10}$ and by sampling the subsequent $d-1$ rounds ρ_i according to parameter $p_i = \frac{1}{10s}$. We begin by claiming that, after the first restriction, the bottom gates' fan-ins are at most s with probability of at least $1 - 2^{-s}$. This can be decomposed into two cases:

(1) Suppose a gate initially has fan-in of at least $2s$ and (w.l.o.g) suppose that it is a $\wedge$-gate. The probability of ρ_0 setting at least of the input variables to 0 is at most $(\frac{1}{10} + \frac{9}{20})^{2s} = (0.55)^{2s} < 2^{-s}$.
(2) Suppose a gate initially has fan-in of at most $2s$. The probability of ρ setting at least s of the input variables to $*$ is $\binom{2s}{s}(1/10)^s < 2^{-s}$

Since we assume that the AC^0 circuit computing f is in alternating normal form, once we invoke the switching lemma for the bottom two layer gates, we can merge the two $\vee$-gates to reduce the depth by one. The switching tells us that with probability of at least $1 - (5\frac{1}{10s}s)^s < 2^{-s}$, we can reduce the depth in each step such that the gates of distance two from the input variables have bottom fan-in of at most s. By the end of our depth-reduction routine, we are left with a CNF with bottom fan-in of at most s. Invoking Corollary (2.3.1) tells us that our final restricted function $f_{\rho_0\rho_1\ldots\rho_{d-1}}$ has Fourier degree of at most s with probability $< 2^{-s}$. The final bound arises from the observation that we applied the beginning analysis for each of the bottom gates exactly once, the switching lemma exactly once for each gate above the bottom gates except for the output gate, and we invoked the corollary for the output gate. Each step contributed probability $< 2^{-s}$. □

Corollary 2.3.2 tells us that the probability of the Fourier degree exceeding s decays exponentially with s given that we set our probability parameter to a specific quantity $p = \frac{1}{10^d s^{d-1}}$.

Proof. (Proof of Theorem 2.2) We begin by first setting our probability parameter $p \leq \frac{1}{10^d k^{d-1}}$. The values p, k will be fixed later to invoke Corollary 2.3.2. We sample a random restriction ρ by sampling some $V \subseteq$ to be the indices which are not set to $*$. Each index has a $1-p$ probability of being set to either $0, 1$. For each index in V, we uniformly sample some bit string in $\{0,1\}^{|V|}$ to fix the indices contained in V. If we recall that our characters are defined as $\chi_A(x) = (-1)^{\sum_{i\in A} x_i}$ and let x_V be the restriction of string x on the input indices to those found in V

$$\chi_S(x) = (-1)^{\sum_{i\in S} x_i} = (-1)^{\sum_{i\in S\cap V} x_i + \sum_{i\in S/V} x_i} \tag{2.3}$$

$$= (-1)^{\sum_{i\in S\cap V} x_i}(-1)^{\sum_{i\in S/V} x_i} = \chi_{S\cap V}(x_V)\chi_{S/V}(x_{\overline{V}}) \tag{2.4}$$

Let $f_{x_V} = f(x_V, *)$ as a function $f_{x_V} : \{0,1\}^{|x_{\overline{V}}|} \to \{0,1\}$. Rearrange the Fourier decomposition of any function f as follows:

$$f(x) = \sum_{S\subseteq[n]} \hat{f}(S)\chi_S(x) = \sum_{V\sqcup\overline{V}\subseteq[n]} \hat{f}(S)\chi_{S\cap V}(x_V)\chi_{S/V}(x_{\overline{V}}) \tag{2.5}$$

$$= \sum_{H\subseteq\overline{V}} \left(\sum_{J\subseteq V} \hat{f}(H\cup J)\chi_J(x_V) \right) \chi_H(x_{\overline{V}}) \tag{2.6}$$

The last equality just rearranges the sum via iterating through all subsets $S \subseteq [n]$ by considering susbets of each partition $V, \overline{V}$ separately. Since the Fourier decomposition is unique for f_{x_V},

$$\widehat{f_{x_V}}(H) = \sum_{J\subseteq V} \hat{f}(H\cup J)\chi_J(x_V), \quad H \subseteq \overline{V} \tag{2.7}$$

By Parseval's identity on the function $x_T \mapsto \widehat{f_{x_T}}(H)$,

$$\mathbb{E}_{x_V}|\widehat{f_{x_V}}(H)|^2 = \langle \hat{f}_-(H), \hat{f}_-(H)\rangle = \sum_{J\subseteq V} |\hat{f}(H\cup J)|^2 \tag{2.8}$$

This along with another application of Parseval's identity yields the string of equalities:

$$\mathbb{E}_{x_V}||f_{x_V}^{>k}||_2^2 = \mathbb{E}_{x_V}\sum_{\substack{H\subseteq\overline{V}\\ |H|>k}} |\widehat{f_{x_V}}(H)|^2 = \sum_{\substack{H\subseteq\overline{V}\\ |H|>k}}\sum_{J\subseteq V} |\hat{f}(H\cup J)|^2 \tag{2.9}$$

$$= \sum_{\substack{S\subseteq[n]\\ |S\cap\overline{V}|>k}} |\hat{f}(S)|^2 \tag{2.10}$$

Sampling over all such p-restrictions further shows that:

$$\mathbb{E}_V\mathbb{E}_{x_V}||f_{x_V}^{>k}||_2^2 \leq Pr[\deg_{\mathcal{F}}(f_\rho) > k] \leq M2^{-k} \tag{2.11}$$

The second-to-last inequality follows since $||f_\rho^{>k}||_2^2 \leq 1$. So we can just replace $||f_\rho^{>k}||_2^2$ with an indicator random variable flipping to one if $\deg_{\mathcal{F}}(f_\rho) > k$ and zero otherwise. The last inequality is from Corollary (2.3.2).

Lemma 2.4. *For any boolean function f and $0 < p < 1$,*

$$\sum_{|S|>t} |\hat{f}(S)|^2 \leq 2\cdot\mathbb{E}_T \sum_{\substack{S\subseteq[n]\\ |S\cap\overline{V}|>k}} |\hat{f}(S)|^2$$

for $k = pt/2$

Proof. We know that:

$$\mathbb{E}_T \sum_{\substack{S\subseteq[n]\\ |S\cap\overline{V}|>k}} |\hat{f}(S)|^2 = \sum_{S\subseteq[n]} Pr[|S\cap\overline{T}| > k]|\hat{f}(S)|^2 \geq \sum_{|S|>t} Pr[|S\cap\overline{T}| > k]|\hat{f}(S)|^2 \tag{2.12}$$

By invoking a Chernoff bound on the probability that $|A\cap\overline{T}|$ will be $\leq pt/2$

$$Pr[|A\cap\overline{T}| < (1 - \frac{2|S|-t}{2|S|})p|S|] \leq e^{-\frac{pt}{8}}$$

Thus, $Pr[|A\cap\overline{T}| > \frac{pt}{2}] \geq 1 - e^{-\frac{pt}{8}} \geq \frac{1}{2}$. This shows that:

$$\sum_{|S|>t} \frac{1}{2}|\hat{f}(S)|^2 \leq \mathbb{E}_T \sum_{\substack{S\subseteq[n]\\ |S\cap\overline{V}|>k}} |\hat{f}(S)|^2$$

by (2.12). □

To finish the proof, set $p = \frac{1}{10t^{(d-1)/d}}$ and $k = t^{1/d}/20$. Direct calculation shows that $p \leq \frac{1}{10^d k^{d-1}}$. This allows to us to invoke Corollary (2.3.2) through (2.11) to yield the desired outcome:

$$\sum_{|S|>t} |\hat{f}(S)|^2 \leq 2M2^{-t^{1/d}/20} \tag{2.13}$$

□

3. Basic Applications

3.1. **Approximation by Low-degree Polynomials.** Theorem (2.2) shows that we can approximate functions in $f \in \mathsf{AC}^0$ by taking our polynomials to be $f^{\leq k}$ for some sufficiently large k. Here we make the distinction that our Fourier expansion of f will be

$$f = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} x_i$$

as a polynomial with complex coefficients.

Lemma 3.1. *Let $f \in \mathsf{AC}^0$ be a boolean function of polynomial size and depth d. Then there exists a complex polynomial of degree $\mathcal{O}((\log n/\epsilon)^d)$ such that $||f - p||_2 < \epsilon$*

Proof. This follows from setting $t = \mathcal{O}((\log n/\epsilon)^d)$ and invoking the LMN theorem to conclude that $||f - f^{\leq t}|| < \epsilon$. □

This gives another scheme of approximating functions in AC^0 through low-degree polynomials, complementing the results of [7], [10]. Here we make the extra distinction that our polynomials are also of low Fourier degree.

3.2. **Sensitivity and Influence.** LMN also shows that functions in AC^0 have low average sensitivity i.e its output is not very sensitive to changes to the input.

Definition 3.2. Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function. We define the *sensitivity* of an input $x \in \{0,1\}^n$ in respect to f, $s_f(x)$ to be the number of indices i such that $f(x) \neq f(x + e_i)$ where e_i is the bit string with zeros everywhere except for the i^{th} index.

Definition 3.3. Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function. The *influence* of f, I_f is defined as the average sensitivity over all input bit strings

$$I_f = \mathbb{E}_x[s_f(x)] \tag{3.1}$$

The influence of f can be expressed in terms of its Fourier coefficients:

$$I_f = 4 \sum_{S \subseteq [n]} |S||\hat{f}(s)|^2 \tag{3.2}$$

By combining this equivalence with Theorem (2.2), we deduce the following upper bound on the influence of a function in AC^0

Lemma 3.4. *Let $f \in \mathsf{AC}^0$ be of depth d. Then*

$$I_f = \mathcal{O}((\log n)^d) \tag{3.3}$$

Lemma (3.4) shows that functions in AC^0 are not suitable for constructing universal hash functions ([5]) and pseudorandom function generators ([4]). We give a rough definition and sketch based on the treatment found in ([4].

Definition 3.5. A function $f : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}$ is called a *pseudorandom function generator* if there exists no polynomial-time oracle Turing machine which can distinguish between the outputs from a true random oracle versus $f(s,)$ for some random seed $s \in \{0,1\}^m$.

Lemma 3.6. *No pseudorandom function generators exist in* AC^0

By taking advantage of the low average sensitivity, we can simply perturb the input slightly and check if the output of f changes. If it doesn't, there is a good chance that it lies in AC^0.

Remark 3.7. (*Some concluding remarks*) The bounds shown by Linai-Mansour-Nisan have been improved by Tal ([12]). Specifically, he showed that for any n-ary boolean function f with size M and depth d

$$\sum_{|S|>t} |\hat{f}(S)|^2 = 2^{\Omega(k/\log^{d-1} M)} \tag{3.4}$$

Furthermore, he showed a converse to the proof direction employed by Linial et. al [4]. Recall that in our proof we started with a switching lemma which bound the probability of the randomly-restricted function's Fourier degree exceeding a certain value. We then used this switching lemma to bound the tail Fourier spectrum. Tal showed a form of the opposite direction in [11].

Theorem 3.8. (Tal [11]) *Let $f\{1,-1\}^n \to \{-1,1\}$ be a Boolean function, let $t, C > 0$ such that for all k, $\sum_{|S|>k} |\hat{f}(S)|^2 \leq Ce^{k/t}$ and let ρ be a p-random restriction, then for all d, $Pr[deg(f_\rho)] \leq C \cdot (4pt)^d$.*

His work uses results in Quantum Query Complexity to ultimately culminate into another proof of the shrinkage exponent of Demorgan Formulae under random restrictions:

Theorem 3.9. (Tal [11]) *Let f be a Boolean function. For $p > 0$,*

$$\mathbb{E}_{\rho\sim\mathcal{R}_p}[\mathcal{L}(f_\rho)] = \mathcal{O}\left(p^2\mathcal{L}(f) + p\sqrt{\mathcal{L}(f)}\right) \tag{3.5}$$

References

1. T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli, *Discrete harmonic analysis: Representations, number theory, expanders, and the fourier transform*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2018.
2. Hamed Hatami, *Harmonic analysis of boolean functions.*
3. Johan Torkel Håstad, *Computational limitations for small-depth circuits*, MIT Press, Cambridge, MA, USA, 1987.
4. Nathan Linial, Yishay Mansour, and Noam Nisan, *Constant depth circuits, fourier transform, and learnability*, J. ACM **40** (1993), no. 3, 607–620.
5. Yishay Mansour, Noam Nisan, and Prasoon Tiwari, *The computational complexity of universal hashing*, Theoretical Computer Science **107** (1993), no. 1, 121–133.
6. Ryan O'Donnell, *Analysis of boolean functions*, Cambridge University Press, 2014.
7. Alexander A Razborov, *Lower bounds for the size of circuits of bounded depth with basis f^; g*, Math. notes of the Academy of Sciences of the USSR **41** (1987), no. 4, 333–338.
8. W. Rudin, *Fourier analysis on groups*, Dover Books on Mathematics, Dover Publications, 2017.
9. Jean-Pierre Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics, vol. 42, Springer, 1977.
10. Roman Smolensky, *Algebraic methods in the theory of lower bounds for boolean circuit complexity*, Proceedings of the nineteenth annual ACM symposium on Theory of computing, 1987, pp. 77–82.
11. Avishay Tal, *Shrinkage of de morgan formulae by spectral techniques*, 2014 IEEE 55th Annual Symposium on Foundations of Computer Science, IEEE, 2014, pp. 551–560.
12. Avishay Tal, *Tight Bounds on the Fourier Spectrum of AC0*, 32nd Computational Complexity Conference (CCC 2017) (Dagstuhl, Germany) (Ryan O'Donnell, ed.), Leibniz International Proceedings in Informatics (LIPIcs), vol. 79, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017, pp. 15:1–15:31.
13. P. Webb, *A course in finite group representation theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2016.

Email address: `ehkim@cs.unc.edu`