

The Entropy Influence Conjecture Revisited

Bireswar Das ^{*} Manjish Pal [†] Vijay Visavaliya [‡]

Abstract

In this paper, we prove that most of the boolean functions, $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ satisfy the Fourier Entropy Influence (FEI) Conjecture due to Friedgut and Kalai (Proc. AMS'96)[2]. The conjecture says that the Entropy of a boolean function is at most a constant times the Influence of the function. The conjecture has been proven for families of functions of smaller sizes. O'donnell, Wright and Zhou (ICALP'11)[8] verified the conjecture for the family of symmetric functions, whose size is 2^{n+1} . They are in fact able to prove the conjecture for the family of d -part symmetric functions for constant d , the size of whose is $2^{O(n^d)}$. Also it is known that the conjecture is true for a large fraction of polynomial sized DNFs (COLT'10)[6]. Using elementary methods we prove that a random function with high probability satisfies the conjecture with the constant as $(2 + \delta)$, for any constant $\delta > 0$.

1 Introduction

The Entropy Influence Conjecture due to Friedgut and Kalai [2] says that for every boolean function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ the following holds,

$$\sum_{S \subseteq [n]} \widehat{f}(S)^2 \log_2 \frac{1}{\widehat{f}(S)^2} \leq C \cdot \sum_{S \subseteq [n]} \widehat{f}(S)^2 |S|$$

for some universal constant $C > 0$ where $\widehat{f}(S)$ is the coefficient of $\chi_S(x)$ in the Fourier expansion of f . The conjecture is of profound importance because of its potential impacts in areas like Learning Theory, Threshold Phenomena in monotone graph properties, metric embeddings etc. For a detailed description of the impact and background of the FEI conjecture the reader is recommended to read the Introduction section of [8] and the blog post by Gil Kalai [4].

The conjecture can be verified for simple functions like AND, OR, MAJORITY, Tribes etc. Although posed about 15 years back we are not aware of a significantly large (of

^{*}Indian Institute of Technology Gandhinagar, India. bireswar@iitgn.ac.in

[†]Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel. This work was done when the author was at Indian Institute of Technology Gandhinagar, India. manjishster@gmail.com

[‡]Vishwakarma Government Engineering College Gandhinagar, India. visavaliavijay@gmail.com

doubly exponential size) family of functions which satisfies this conjecture. Klivans et al. [6] proved recently that a large fraction of polynomial sized DNF formulae satisfy the Mansour's conjecture [7] which in turn implies that FEI conjecture is also true for those functions, a class which has a size of $2^{\text{poly}(n)}$. In a recent resurrection of the FEI conjecture, O'Donnell et al. [8] proved the conjecture for the family of symmetric functions and d -part symmetric functions for constant d . The sizes of these families are 2^{n+1} and $2^{O(n^d)}$ respectively, again only exponential in size. They also verified it for read-once decision trees which are of exponential size as well. Thus, one is not aware of an explicit or non-explicit family of doubly exponential size that satisfies the conjecture. Very recently, in a note, Keller et al [5] managed to prove a variant of the conjecture for functions which have low Fourier weight on characters of large size.

2 Preliminaries

In this section we introduce the basic preliminaries of discrete Fourier Analysis which will be of interest for us.

Definition 1. Let $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ be a boolean function. The Fourier expansion of f is written as

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

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

Definition 2. Let $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ be a boolean function. The entropy of f is defined as

$$\mathbf{H}(f) = \sum_{S \subseteq [n]} \hat{f}(S)^2 \log_2 \frac{1}{\hat{f}(S)^2}$$

Notice by Parseval's identity, $\sum_{S \subseteq [n]} \hat{f}(S)^2 = 1$ which implies that the Fourier coefficients can be thought of as a probability distribution and hence $\mathbf{H}(f)$ gives us the entropy of that distribution. The following fact (see page 40 of [3]) gives an upper bound on $\mathbf{H}(f)$ which will be of use for us later.

Fact 1. For an arbitrary boolean function f , $\mathbf{H}(f) \leq n$.

Definition 3. For a boolean function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, $\mathbf{Inf}_i(f)$, the Influence of coordinate i is defined as

$$\mathbf{Inf}_i(f) = \sum_{S: i \in S} \hat{f}(S)^2$$

and the Influence of f , $\mathbf{Inf}(f)$, is defined as

$$\mathbf{Inf}(f) = \sum_{i=1}^n \mathbf{Inf}_i = \sum_{S \subseteq [n]} \hat{f}(S)^2 |S|$$

An equivalent combinatorial interpretation of the influence of the i th coordinate is given by $\mathbf{Inf}_i = \Pr_x[f(x) \neq f(x^{(i)})]$ where $x^{(i)}$ is x with the i th coordinate flipped.

We will also use the following well-known facts regarding the Fourier coefficients.

Fact 2. For a boolean function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, the following holds for any subset $S \subseteq [n]$

1. $\widehat{f}(S) = \frac{1}{2^n} \sum_x f(x) \chi_S(x)$.
2. $\sum_x \chi_S(x) = 0$ if $S \neq \phi$, and 2^n if $S = \phi$.

3 The Result

The main contribution of our paper is to prove that there is a large (non-explicit) family of functions that satisfies the entropy influence conjecture. The size of this family is significantly larger than the size of any of the known families for which the FEI conjecture is known to be true. More precisely, we show that there is a family of functions whose size is

$$\left[1 - 4 \left(1 + \frac{2}{\delta} \right)^2 \frac{1}{2^{n+1}n} \right] \cdot 2^{2^n}$$

satisfies the conjecture with $C = 2 + \delta$ for any constant $\delta > 0$. It is worth mentioning here is that there are functions which need the constant C to be at least 4.6 [8] so taking δ very small will not work for some functions.

Theorem 1. A random function satisfies the FEI conjecture with high probability with $C = 2 + \delta$, for any constant $\delta > 0$.

Consider random function which puts values 1 or -1 independently on every point of $\{1, -1\}^n$ with an equal probability of $1/2$. Clearly, every function is obtained with a probability of $\frac{1}{2^{2^n}}$. Let H_r and I_r be random variables denoting the entropy and influence of a randomly chosen function as above. We will prove Theorem 1 using a simple application of Chebyshev Inequality. We will use the following two lemmas in our proof. A proof of Lemma 1 and Lemma 2 via the notion of average sensitivity (which is same as the Influence) is available in a technical report due to Bernasconi, Codenotti and Simon [1]. For the sake of completeness we include slightly different proofs for these two lemmas.

Lemma 1. $E[I_r] = \frac{n}{2}$.

Proof.

$$\begin{aligned}
\mathbf{E}[I_r] &= \mathbf{E} \left[\sum_{S \subseteq [n]} \widehat{f}(S)^2 |S| \right] \\
&= \sum_{S \subseteq [n]} \mathbf{E}[\widehat{f}(S)^2] |S| \\
&= \sum_{S \subseteq [n]} \frac{1}{2^{2n}} \mathbf{E} \left[\sum_x f(x)^2 \chi_S(x)^2 + \sum_{x \neq y} f(x)f(y) \chi_S(x) \chi_S(y) \right] |S| \\
&= \frac{1}{2^{2n}} \sum_{S \subseteq [n]} \left(\sum_x 1 + \sum_{x \neq y} \mathbf{E}[f(x)f(y)] \chi_S(x) \chi_S(y) \right) |S| \\
&= \frac{1}{2^{2n}} \sum_{S \subseteq [n]} 2^n |S| + 0 \text{ (this follows because for } x \neq y, \mathbf{E}[f(x)f(y)] = \mathbf{E}[f(x)]\mathbf{E}[f(y)] = 0 \text{)} \\
&= \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} k = \frac{n2^{n-1}}{2^n} = \frac{n}{2}
\end{aligned}$$

□

Lemma 2. $\text{Var}[I_r] = \frac{n}{2^{n+1}}$.

Proof. We have already calculated $\mathbf{E}[I_r]$. To calculate,

$$\text{Var}[I_r] = \mathbf{E}[I_r^2] - (\mathbf{E}[I_r])^2$$

we need to calculate $\mathbf{E}[I_r^2]$.

$$\begin{aligned}
\mathbf{E}[I_r^2] &= \mathbf{E} \left[\left(\sum_{S \subseteq [n]} \widehat{f}(S)^2 |S| \right)^2 \right] \\
&= \mathbf{E} \left[\sum_{S_1, S_2 \subseteq [n]} \widehat{f}(S_1)^2 \widehat{f}(S_2)^2 |S_1| |S_2| \right] \\
&= \sum_{S_1, S_2 \subseteq [n]} \mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] |S_1| |S_2|
\end{aligned}$$

$$\begin{aligned}
\mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] &= \frac{1}{2^{4n}} \mathbf{E} \left[\left(\sum_x f(x) \chi_{S_1}(x) \right)^2 \left(\sum_x f(x) \chi_{S_2}(x) \right)^2 \right] \\
&= \frac{1}{2^{4n}} \mathbf{E} \left[\sum_{x_1, y_1, x_2, y_2} f(x_1) f(y_1) f(x_2) f(y_2) \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2) \right] \\
&= \frac{1}{2^{4n}} \sum_{x_1, y_1, x_2, y_2} \mathbf{E} [f(x_1) f(y_1) f(x_2) f(y_2)] \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2)
\end{aligned}$$

To calculate the above sum, consider the following sets,

$$\begin{aligned}
A_1 &= \{(x_1, y_1, x_2, y_2) | x_1 = y_1, x_2 = y_2\} \\
A_2 &= \{(x_1, y_1, x_2, y_2) | x_1 = x_2, y_1 = y_2\} \\
A_3 &= \{(x_1, y_1, x_2, y_2) | x_1 = y_2, x_2 = y_1\}
\end{aligned}$$

Notice the following properties of A_1, A_2, A_3 , $|A_1| = |A_2| = |A_3| = 2^{2n}$, $A_1 \cap A_2 = A_2 \cap A_3 = A_3 \cap A_1 = A_1 \cap A_2 \cap A_3$ and $|A_1 \cap A_2 \cap A_3| = 2^n$. It is easy to verify that if $(x_1, y_1, x_2, y_2) \notin A_1 \cup A_2 \cup A_3$, then $\mathbf{E} [f(x_1) f(y_1) f(x_2) f(y_2)] = 0$. Otherwise it is $\chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2)$. Using the above properties and inclusion exclusion principle we have $\mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right]$ equal to

$$\begin{aligned}
&\frac{1}{2^{4n}} \left[\sum_{A_2} \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2) + \sum_{A_3} \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2) \right] + \\
&\frac{1}{2^{4n}} \left[\sum_{A_1} \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2) \right] - 2 \cdot \frac{1}{2^{4n}} \left[\sum_{A_1 \cap A_2 \cap A_3} \chi_{S_1}(x_1) \chi_{S_1}(y_1) \chi_{S_2}(x_2) \chi_{S_2}(y_2) \right] \\
&= \frac{1}{2^{4n}} \left[\sum_{x, y} \chi_{S_1}(x) \chi_{S_1}(y) \chi_{S_2}(x) \chi_{S_2}(y) + \sum_{x, y} \chi_{S_1}(x) \chi_{S_1}(y) \chi_{S_2}(y) \chi_{S_2}(x) \right] + \\
&\frac{1}{2^{4n}} \left[\sum_{x, y} \chi_{S_1}(x) \chi_{S_1}(x) \chi_{S_2}(y) \chi_{S_2}(y) \right] - 2 \cdot \frac{1}{2^{4n}} \cdot 2^n \\
&= \frac{1}{2^{4n}} \left[2 \sum_{x, y} \chi_{S_1 \Delta S_2}(x) \chi_{S_1 \Delta S_2}(y) + 2^{2n} - 2 \cdot 2^n \right].
\end{aligned}$$

Thus, using the fact that $\sum_x \chi_{S_1 \Delta S_2}(x) = 0$ if $S_1 \neq S_2$ and 2^n otherwise, we have

$$\mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] = \begin{cases} \frac{1}{2^{4n}} \cdot (2^{2n} - 2 \cdot 2^n) & \text{if } S_1 \neq S_2. \\ \frac{1}{2^{4n}} \cdot (3 \cdot 2^{2n} - 2 \cdot 2^n) & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned}
\mathbf{E}[I_r^2] &= \sum_{S_1, S_2 \subseteq [n]} \mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] |S_1||S_2| \\
&= \sum_{S_1=S_2} \mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] |S_1||S_2| + \sum_{S_1 \neq S_2} \mathbf{E} \left[\widehat{f}(S_1)^2 \widehat{f}(S_2)^2 \right] |S_1||S_2| \\
&= \sum_{S_1=S_2} \frac{3 \cdot 2^{2n} - 2 \cdot 2^n}{2^{4n}} |S_1||S_2| + \sum_{S_1 \neq S_2} \frac{2^{2n} - 2 \cdot 2^n}{2^{4n}} |S_1||S_2| \\
&= \sum_{S_1=S_2} \frac{2 \cdot 2^{2n}}{2^{4n}} |S_1||S_2| + \sum_{S_1, S_2} \frac{2^{2n} - 2 \cdot 2^n}{2^{4n}} |S_1||S_2| \\
&= \frac{2 \cdot 2^{2n}}{2^{4n}} \sum_{k=0}^n \binom{n}{k} k^2 + \frac{2^{2n} - 2 \cdot 2^n}{2^{4n}} \left(\sum_{k=0}^n \binom{n}{k} k \right)^2 \\
&= \frac{2(n(n-1)2^{n-2} + n2^{n-1})}{2^{2n}} + \left(\frac{1}{2^{2n}} - \frac{2}{2^{3n}} \right) (n2^{n-1})^2 \\
&= \frac{n}{2^{n+1}} + \frac{n^2}{4}.
\end{aligned}$$

Hence,

$$\mathbf{Var}[I_r] = \frac{n}{2^{n+1}}.$$

□

We are now ready to prove Theorem 1.

Proof. (Theorem 1) We will prove this using simple applications of Chebyshev inequality. As mentioned earlier we pick a random function which puts values -1 or 1 independently on every point of $\{-1, 1\}^n$ with an equal probability of $1/2$. Recall that, H_r and I_r are random variables denoting the entropy and influence of a randomly chosen function. Let us define the event \mathcal{E}_{FEI} indicating that the randomly chosen boolean function satisfies the FEI conjecture with the constant $C = 2 + 2\epsilon$ for $\epsilon > 0$. More precisely, \mathcal{E}_{FEI} is the event that $H_r \leq (2 + 2\epsilon)I_r$. Our aim is to prove that almost surely \mathcal{E}_{FEI} takes place. From Fact 1, we have $H_r \leq n$. Therefore,

$$\Pr[H_r > n] = 0.$$

Consider the following events:

$$\mathcal{E}_1 := H_r > n$$

$$\mathcal{E}_2 := (2 + 2\epsilon)I_r \leq n$$

Now if $H_r > (2 + 2\epsilon)I_r$, then either \mathcal{E}_1 or \mathcal{E}_2 happens. Therefore by union bound,

$$\Pr[H_r > (2 + 2\epsilon)I_r] \leq \Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2]$$

Since from Lemma 1, $\mathbf{E}[I_r] = n/2$, $\mathbf{E}[(2+2\epsilon)I_r] = (1+\epsilon)n$. Now we upper bound $\Pr[\mathcal{E}_2]$.
 $\Pr[\mathcal{E}_2] = \Pr[(2+2\epsilon)I_r \leq n] = \Pr[(2+2\epsilon)I_r - (1+\epsilon)n \leq -\epsilon n] \leq \Pr[|(2+2\epsilon)I_r - (1+\epsilon)n| \geq \epsilon n]$

Using Chebyshev Inequality,

$$\Pr[|(2+2\epsilon)I_r - (1+\epsilon)n| \geq \epsilon n] \leq \frac{\mathbf{Var}[(2+2\epsilon)I_r]}{\epsilon^2 n^2} = 4 \left(1 + \frac{1}{\epsilon}\right)^2 \frac{\mathbf{Var}[I_r]}{n^2}.$$

From Lemma 2, this in turn implies

$$\Pr[H_r > (2+2\epsilon)I_r] \leq 4 \left(1 + \frac{1}{\epsilon}\right)^2 \frac{1}{2^{n+1}n}$$

Hence, it follows that $\Pr[\mathcal{E}_{FEI}] \geq 1 - 2 \left(1 + \frac{1}{\epsilon}\right)^2 \frac{1}{2^{n+1}n}$ with constant $C = 2 + 2\epsilon$ for arbitrary constant $\epsilon > 0$. \square

4 Conclusion

In this paper we gave a simple proof of the fact a random function will almost surely satisfy the FEI conjecture for $C = 2 + \delta$ for $\delta > 0$. Although our proof is non-constructive, this is the only doubly exponential sized family for which it is known that the FEI conjecture is true.

It would be interesting to get a large ($\omega(2^{\text{poly}(n)})$) explicit family of functions that satisfy FEI conjecture. One possible candidate is the class of functions $f(x_1, x_2, \dots, x_p)$ which are invariant under the action of the cyclic permutation group $C_p \leq \text{Sym}(p)$ where p is prime. It can be verified that the size of this class is $2^{\frac{2^p-2}{p}+2}$. Because of the highly structured nature of the functions which are invariant under C_p it might be plausible to verify the conjecture for these functions.

5 Acknowledgement

We thank Kunal Dutta and Justin Salez for pointing out that our result can be extended to a high probability statement.

References

- [1] A. Bernasconi, B. Codenotti and J. Simon. *On the Fourier Analysis of Boolean Functions*, IMC B4-97-03, 1996.
- [2] E. Friedgut and G. Kalai. *Every monotone graph property has a sharp threshold*, Proceedings of AMS, 124(10), 1996, pp. 2993 - 3002.

- [3] T. M. Cover and J. A. Thomas. *Elements of Information Theory*, John Wiley & Sons, Inc., N. Y., 1991.
- [4] G. Kalai, *The entropy/influence conjecture*, <http://terrytao.wordpress.com/2007/08/16/gil-kalai-the-entropyinfluence-conjecture/>
- [5] N. Keller, E. Mossel, and T. Schlam, *A Note on the Entropy/Influence Conjecture*, arXiv:1105.2651v1, 2011
- [6] A. Klivans, H. Lee, and A. Wan. *Mansour's Conjecture is true for random DNF formulas*, COLT, 2010, pp. 368-380.
- [7] Y. Mansour, *Learning Boolean functions via the Fourier transform*, Kluwer Academic Publishers, 1994, pp. 391-424.
- [8] R. O'donnell, J. Wright and Y. Zhou, *The Fourier Entropy-Influence Conjecture for certain classes of Boolean functions*, ICALP, 2011, pp. 330-341.