Discrete Mathematics 312 (2012) 3364–3372

Contents lists available at SciVerse ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

# A note on the Entropy/Influence conjecture

Nathan Keller [a], Elchanan Mossel [b,a,*], Tomer Schlank [c]

[a] Weizmann Institute of Science, Israel
[b] U.C. Berkeley, United States
[c] Hebrew University, Israel

ARTICLE INFO

*Article history:*
Received 6 June 2011
Received in revised form 24 May 2012
Accepted 25 July 2012
Available online 20 August 2012

*Keywords:*
Entropy
Influence
Discrete Fourier analysis
Probabilistic combinatorics

ABSTRACT

The Entropy/Influence conjecture, raised by Friedgut and Kalai (1996) [9], seeks to relate two different measures of concentration of the Fourier coefficients of a Boolean function. Roughly saying, it claims that if the Fourier spectrum is "smeared out", then the Fourier coefficients are concentrated on "high" levels. In this note we generalize the conjecture to biased product measures on the discrete cube.



## 1. Introduction

**Definition 1.1.** Consider the discrete cube $\{0,1\}^n$ endowed with the product measure $\mu_p = (p\delta_{\{1\}} + (1-p)\delta_{\{0\}})^{\otimes n}$, denoted in the sequel by $\{0,1\}^n_p$, and let $f : \{0,1\}^n_p \to \mathbb{R}$. The Fourier–Walsh expansion of $f$ with respect to the measure $\mu_p$ is the unique expansion

$$f = \sum_{S \subset \{1,2,\ldots,n\}} \alpha_S u_S,$$

where for any $T \subset \{1, 2, \ldots, n\}$,[1]

$$u_S(T) = \left(-\sqrt{\frac{1-p}{p}}\right)^{|S\cap T|} \left(\sqrt{\frac{p}{1-p}}\right)^{|S\setminus T|}.$$

In particular, for the uniform measure (i.e., $p = 1/2$), $u_S(T) = (-1)^{|S\cap T|}$. The coefficients $\alpha_S$ are denoted by $\hat{f}(S)$,[2] and the level of the coefficient $\hat{f}(S)$ is $|S|$.

Properties of the Fourier–Walsh expansion are one of the main objects of study in discrete harmonic analysis. The Entropy/Influence (EI) conjecture, raised by Friedgut and Kalai [9] in 1996, seeks to relate two measures of concentration of

* Corresponding author at: U.C. Berkeley, United States.
*E-mail addresses:* nathan.keller@weizmann.ac.il (N. Keller), mossel@stat.berkeley.edu (E. Mossel), tomer.schlank@gmail.com (T. Schlank).

1 Throughout the paper, we identify elements of $\{0,1\}^n$ with subsets of $\{1, 2, \ldots, n\}$ in the natural way.

2 Note that since the functions $\{u_S\}_{S\subset\{1,\ldots,n\}}$ form an orthonormal basis, the representation is indeed unique, and the coefficients are given by the formula $\hat{f}(S) = \mathbb{E}_{\mu_p}[f \cdot u_S]$.

0012-365X/$ – see front matter © 2012 Published by Elsevier B.V.
doi:10.1016/j.disc.2012.07.031

the Fourier coefficients (i.e. coefficients of the Fourier–Walsh expansion) of Boolean functions. The first of them is *spectral entropy*.

**Definition 1.2.** Let $f : \{0,1\}_p^n \to \{-1,1\}$ be a Boolean function. The spectral entropy of $f$ with respect to the measure $\mu_p$ is

$$\mathrm{Ent}_p(f) = \sum_{S\subset\{1,\ldots,n\}} \hat{f}(S)^2 \log\left(\frac{1}{\hat{f}(S)^2}\right),$$

where the Fourier–Walsh coefficients are computed w.r.t. to $\mu_p$.

Note that by Parseval's identity, for any Boolean function we have $\sum_S \hat{f}(S)^2 = 1$, and thus, the squares of the Fourier coefficients can be viewed as a probability distribution on the set $\{0,1\}^n$. In this notation, the spectral entropy is simply the entropy of this distribution, and intuitively, it measures how much the Fourier coefficients are "smeared out".

The second notion is *total influence*.

**Definition 1.3.** Let $f : \{0,1\}_p^n \to \{0,1\}$. For $1 \le i \le n$, the influence of the $i$-th coordinate on $f$ with respect to $\mu_p$ is

$$\mathrm{I}_i^p(f) = \Pr_{x\sim\mu_p}[f(x) \neq f(x \oplus e_i)],$$

where $x \oplus e_i$ denotes the point obtained from $x$ by replacing $x_i$ with $1 - x_i$ and leaving the other coordinates unchanged. The total influence of the function $f$ is

$$\mathrm{I}_p(f) = \sum_{i=1}^{n} \mathrm{I}_i^p(f).$$

Influences of variables on Boolean functions were studied extensively in the last decades, and have applications in a wide variety of fields, including Theoretical Computer Science, Combinatorics, Mathematical Physics, Social Choice Theory, etc. (See, e.g., the survey [12].) As observed in [10], the total influence can be expressed in terms of the Fourier coefficients:

**Observation 1.4.** *Let $f : \{0,1\}_p^n \to \{-1,1\}$. Then*

$$\mathrm{I}_p(f) = \frac{1}{4p(1-p)} \sum_S |S|\hat{f}(S)^2. \tag{1}$$

*In particular, for the uniform measure $\mu_{1/2}$, $\mathrm{I}_{1/2}(f) = \sum_S |S|\hat{f}(S)^2$.*

Thus, in terms of the distribution induced by the Fourier coefficients, the total influence is (up to normalization) the *expectation of the level* of the coefficients, and it measures the question whether the coefficients are concentrated on "high" levels.

The Entropy/Influence conjecture asserts the following:

**Conjecture 1.5** (*Friedgut and Kalai*)**.** *Consider the discrete cube $\{0,1\}^n$ endowed with the uniform measure $\mu_{1/2}$. There exists a universal constant $c$, such that for any $n$ and for any Boolean function $f : \{0,1\}_{1/2}^n \to \{-1,1\}$,*

$$\mathrm{Ent}_{1/2}(f) \le c \cdot \mathrm{I}_{1/2}(f).$$

The conjecture, if confirmed, has numerous significant implications. For example, it would imply that for any property of graphs on $n$ vertices, the sum of influences is at least $c(\log n)^2$ (which is tight for the property of containing a clique of size $\approx \log n$). The best currently known lower bound, by Bourgain and Kalai [5], is $\Omega((\log n)^{2-\epsilon})$, for any $\epsilon > 0$.

Another consequence of the conjecture would be an affirmative answer to a variant of a conjecture of Mansour [14] stating that if a Boolean function can be represented by a DNF formula of polynomial size in $n$ (the number of coordinates), then most of its Fourier weight is concentrated on a polynomial number of coefficients (see [15] for a detailed explanation of this application). This conjecture, raised in 1995, is still wide open.

The main object of this note is to generalize the Entropy/Influence conjecture to the product measure $\mu_p$ on the discrete cube. We state a generalization of the conjecture to the biased case:

**Conjecture 1.6.** *There exists a universal constant $c$, such that for any $0 < p < 1$, for any $n$ and for any Boolean function $f : \{0,1\}_p^n \to \{-1,1\}$,*

$$\mathrm{Ent}_p(f) \le cp\log(1/p) \cdot \mathrm{I}_p(f).$$

We prove that Conjecture 1.6 follows from the original EI conjecture, and that it is tight for the graph property of containing a clique of fixed size (at the critical probability). This answers a question raised by Kalai [11].

In addition, we remark on three questions related to the EI conjecture:

1. We cite an unpublished result of Bourgain and Kalai which proves a variant of the conjecture for functions whose Fourier weight beyond some level decays exponentially, and suggest a possible way to attack the entire conjecture via a strengthening of Bourgain and Kalai's result.
2. We give an easy proof of a weaker upper bound on the entropy of Boolean functions.
3. We show a connection between the EI conjecture and Friedgut's characterization of functions with a low total influence [8], and formulate a conjecture which is in a sense "intermediate" between Friedgut's theorem and the EI conjecture.

## 2. Entropy/Influence conjecture for the product measure $\mu_p$ on the discrete cube

In this section we consider the space $\{0, 1\}_p^n$, for $0 < p < 1$. First we formulate a variant of the EI conjecture for the biased measure and prove that it follows from the original conjecture. Then we show that it is tight for the graph property of containing a copy of a complete graph $K_r$ as an induced subgraph, for random graphs distributed according to the model $G(n, p)$, at the critical probability $p_c$.

**Proposition 2.1.** *Assume that the EI conjecture holds. Then there exists a universal constant c such that for any $0 < p < 1$, for any n and for any $f : \{0, 1\}_p^n \to \{-1, 1\}$, we have*

$$\mathrm{Ent}_p(f) \leq cp\log(1/p) \cdot \mathrm{I}_p(f).$$

Our proof is based on a standard reduction from the biased measure $\mu_p$ to the uniform measure $\mu_{1/2}$ first considered in [4]. Let $p \leq 1/2$, and assume that $p = t/2^m$.[3] For any function $f : \{0, 1\}^n \to \mathbb{R}$ we define a function $\mathrm{Red}(f) = g : \{0, 1\}^{mn} \to \mathbb{R}$ as follows: each $y \in \{0, 1\}^{mn}$ is considered as a concatenation of $n$ vectors $y^i \in \{0, 1\}^m$, and each such vector is translated to a natural number $0 \leq \mathrm{Bin}(y^i) < 2^m$ through its binary expansion (i.e., $\mathrm{Bin}(y^i) = \sum_{j=0}^{m-1} 2^j \cdot y^i_{m-j}$). Then, for any $y \in \{0, 1\}^{mn}$,

$$g(y) = g(y^1, y^2, \ldots, y^n) := f\left(h(y^1), h(y^2), \ldots, h(y^n)\right),$$

where $h : \{0, 1\}^m \to \{0, 1\}$ is given by

$$h(y^i) = \begin{cases} 1, & \mathrm{Bin}(y^i) \geq 2^m - t \\ 0, & \mathrm{Bin}(y^i) < 2^m - t. \end{cases}$$

We use two simple properties of the reduction. The first, proved by Friedgut and Kalai [9], relates the total influence of $g$ (w.r.t. $\mu_{1/2}$) to that of $f$ (w.r.t. to $\mu_p$).

**Lemma 2.2** (*Friedgut and Kalai*)**.** *Let $f : \{0, 1\}_p^n \to \{-1, 1\}$, and let $g = \mathrm{Red}(f)$. Then*

$$\mathrm{I}_{1/2}(g) \leq 6p\lfloor \log(1/p) \rfloor \mathrm{I}_p(f). \tag{2}$$

The second property relates the Fourier coefficients of $f$ (w.r.t. $\mu_p$) to corresponding coefficients of $g$ (w.r.t. $\mu_{1/2}$).

**Lemma 2.3.** *Let $f : \{0, 1\}_p^n \to \mathbb{R}$, and let $g = \mathrm{Red}(f)$. For any $S \subset \{1, 2, \ldots, mn\}$, denote $S_i = S \cap \{(i-1)m + 1, (i-1)m + 2, \ldots, im\}$, and for $S' \subset \{1, 2, \ldots, n\}$, let*

$$V(S') = \{S \subset \{1, 2, \ldots, mn\} : \{i : |S_i| > 0\} = S'\}.$$

*Then:*

$$\sum_{S \in V(S')} \hat{g}(S)^2 = \hat{f}(S')^2. \tag{3}$$

**Proof.** For each $S' \subset \{1, 2, \ldots, n\}$, let $f_{S'} : \{0, 1\}_p^n \to \mathbb{R}$ be defined by $f_{S'} = \hat{f}(S')u_{S'}$. We claim that

$$\mathrm{Red}(f_{S'}) = \sum_{S \in V(S')} \hat{g}(S)u_S. \tag{4}$$

[3] It is clear that there is no loss of generality in assuming that $p$ is diadic, as the results for general $p$ follow immediately by approximation.

This claim implies the assertion, as by the Parseval identity, Eq. (4) implies:

$$\sum_{S\in V(S')} \hat{g}(S)^2 = \|\mathrm{Red}(f_{S'})\|_2^2 = \|f_{S'}\|_2^2 = \hat{f}(S')^2.$$

(The first and third equalities use the Parseval identity, and the middle equality holds since by the structure of the reduction, it preserves all $L_p$ norms.)
In order to prove Eq. (4), we use Proposition 2.2 in [13] that describes the exact relation between the Fourier coefficients of $\mathrm{Red}(f)$ and the corresponding coefficients of $f$. By the proposition, for all $S \in V(S')$,

$$\widehat{\mathrm{Red}(f)}(S) = c(S,p)\cdot \hat{f}(S'),$$

where $c(S,p)$ depends on $S$ and $p$ but not on $f$. Hence, for all $S \in V(S')$, we have $\widehat{\mathrm{Red}(f_{S'})}(S) = \widehat{\mathrm{Red}(f)}(S)$ (since both are determined by $S$, $p$, and $\hat{f}(S')$). Similarly, for all $S \notin V(S')$, $\widehat{\mathrm{Red}(f_{S'})}(S) = 0$, since $\widehat{f_{S'}}(S'') = 0$ for all $S'' \neq S'$. Therefore, the Fourier expansion of $\mathrm{Red}(f_{S'})$ is:

$$\mathrm{Red}(f_{S'}) = \sum_{S\in V(S')} \widehat{\mathrm{Red}(f)}(S)u_S,$$

as asserted. □

Now we are ready to prove Proposition 2.1.

**Proof.** Let $f:\{0,1\}_p^n \to \{-1,1\}$, and let $g = \mathrm{Red}(f)$. By Eq. (3),

$$\begin{aligned}\mathrm{Ent}_{1/2}(g) &= \sum_{S\subset\{1,2,\ldots,mn\}} \hat{g}(S)^2 \log\frac{1}{\hat{g}(S)^2} = \sum_{S'\subset\{1,2,\ldots,n\}}\sum_{S\in V(S')} \hat{g}(S)^2 \log\frac{1}{\hat{g}(S)^2}\\ &\geq \sum_{S'\subset\{1,2,\ldots,n\}}\sum_{S\in V(S')} \hat{g}(S)^2 \log\frac{1}{\hat{f}(S')^2} = \sum_{S'\subset\{1,2,\ldots,n\}} \hat{f}(S')^2 \log\frac{1}{\hat{f}(S')^2} = \mathrm{Ent}_p(f). \qquad (5)\end{aligned}$$

(The inequality follows since by (3), we have $\hat{g}(S)^2 \leq \hat{f}(S')^2$ for all $S \in V(S')$.) Combining Eq. (5) with Eq. (2) and applying the EI conjecture to $g$, we get:

$$\mathrm{Ent}_p(f) \leq \mathrm{Ent}_{1/2}(g) \leq c\cdot \mathrm{I}_{1/2}(g) \leq c\cdot 6p\lfloor\log(1/p)\rfloor \mathrm{I}_p(f),$$

and therefore,

$$\mathrm{Ent}_p(f) \leq c'p\log(1/p)\mathrm{I}_p(f),$$

as asserted. □

Consider the random graph model $G(n,p)$. Recall that in this model, the probability space is $\{0,1\}_p^N$, where $N = \binom{n}{2}$, the coordinates correspond to the edges of a graph on $n$ vertices, and each edge exists in the graph with probability $p$, independently of the other edges. It is well-known that for the graph property of containing the complete graph $K_r$ as an induced subgraph, there exists a threshold at $p_t = \Theta(n^{-2/(r-1)})$. This means that if $p \ll n^{-2/(r-1)}$ then $\Pr[K_r \subset G | G \in G(n,p)]$ is close to zero, and if $p \gg n^{-2/(r-1)}$ then $\Pr[K_r \subset G | G \in G(n,p)]$ is close to one. We choose a value $p_0$ in the critical range, consider the characteristic function $f$ of this graph property in $G(n,p_0)$, and show that the assertion of Proposition 2.1 is tight for $f$. In order to simplify the computation, we choose $p_0$ such that the expected number of copies of $K_r$ in $G(n,p_0)$ is "nice". However, the same argument holds for any value of $p$ in the critical range.

**Proposition 2.4.** *Let $n, r$ be integers such that $r < \log n$. Consider the random graph $G(n,p_0)$ where $p_0$ is chosen such that $\binom{n}{r}\cdot p_0^{\binom{r}{2}} = 1/2$. Let $f$ be defined by:*

$$f(G) = 1 \Leftrightarrow G \text{ contains a copy of } K_r \text{ as an induced subgraph},$$

*and $f(G) = 0$ otherwise. Then*

$$\mathrm{Ent}_{p_0}(f) \geq c\cdot p_0\log(1/p_0)\cdot \mathrm{I}_{p_0}(f),$$

*where $c$ is a universal constant.*

**Proof.** The result is a combination of an upper bound on $\mathrm{I}_{p_0}(f)$ with a lower bound on $\mathrm{Ent}_{p_0}(f)$.

In order to bound $\mathrm{I}_{p_0}(f)$ from above, note that a necessary (but not sufficient) condition for an edge $e = (v,w)$ to be pivotal for $f$ at a graph $G$[4] is that there exists a set $S$ of $r$ vertices including $v$ and $w$ such that all $\binom{r}{2}$ edges inside $S$ except

4 An edge $e$ is pivotal for the property $f$ at a graph $G$ if $f(G) = 1$ and $f(G\setminus\{e\}) = 0$.

for $e$ appear in $G$. Hence, a simple union bound yields that for any edge $e$,

$$\mathrm{I}_e^{p_0}(f) \le \binom{n-2}{r-2} \cdot p_0^{\binom{r}{2}-1} = \frac{r(r-1)}{n(n-1)p_0} \cdot \binom{n}{r} p_0^{\binom{r}{2}} = \frac{r(r-1)}{2n(n-1)p_0},$$

and thus,

$$\mathrm{I}_{p_0}(f) = \sum_e \mathrm{I}_e^{p_0}(f) \le \frac{1}{p_0} \cdot \frac{r(r-1)}{4}. \tag{6}$$

In order to bound $\mathrm{Ent}_{p_0}(f)$ from below, we show that at least a constant portion of the Fourier weight of $f$ is concentrated on coefficients that correspond to copies of $K_r$ in $\{0, 1\}^N$. Concretely, we show that if $S$ corresponds to a copy of $K_r$, then:

$$\hat{f}(S)^2 \ge c' \cdot \binom{n}{r}^{-1}. \tag{7}$$

As the number of such coefficients is $\binom{n}{r}$, it will follow that:

$$\mathrm{Ent}_{p_0}(f) \ge \sum_{\{S:\, S \text{ is a copy of } K_r\}} \hat{f}(S)^2 \log\left(\frac{1}{\hat{f}(S)^2}\right) \ge c' \cdot \log\binom{n}{r} \ge c'' \cdot r\log(n), \tag{8}$$

where the rightmost inequality holds since $r < \log n$. Finally, a combination of Eq. (8) with Eq. (6) will imply:

$$\mathrm{Ent}_{p_0}(f) \ge c'' \cdot r\log(n) \ge c'' \cdot \frac{r(r-1)}{2} \cdot \log(1/p_0) \ge c'' \cdot p_0 \log(1/p_0) \cdot \mathrm{I}_{p_0}(f),$$

as asserted.

To prove Eq. (7), consider a specific copy $H$ of $K_r$ and denote its set of edges by $S = E(H)$. By the definition of the Fourier coefficients, we have:

$$\begin{aligned}
\hat{f}(S) &= \sum_{T\in\{0,1\}^N} \mu_{p_0}(T) \left(-\sqrt{\frac{1-p_0}{p_0}}\right)^{|S\cap T|} \left(\sqrt{\frac{p_0}{1-p_0}}\right)^{\binom{r}{2}-|S\cap T|} f(T) \\
&= \sum_{T\in\{0,1\}^N} \mu_{p_0}(T\setminus S) \cdot (p_0(1-p_0))^{\binom{r}{2}/2} (-1)^{|S\cap T|} f(T) \\
&= (p_0(1-p_0))^{\binom{r}{2}/2} \sum_{T\in\{0,1\}^N} \mu_{p_0}(T\setminus S)(-1)^{|S\cap T|} f(T),
\end{aligned} \tag{9}$$

where $\mu_{p_0}(T \setminus S)$ denotes the induced measure of the graph $T \setminus S$. Note that the total contribution to $\hat{f}(S)$ of $\{T \in \{0, 1\}^N : S \subset T\}$ is

$$(-1)^{|S|} \cdot (p_0(1-p_0))^{\binom{r}{2}/2}, \tag{10}$$

since $f(T) = 1$ for all $T \supset S$. On the other hand, if $f(T) = 1$ and $T \supsetneq S$, then $T$ contains a copy of $K_r$, in which $k \le r-1$ vertices are included in $V(H)$, and the remaining $r-k$ vertices are not included in $V(H)$. Hence, the total contribution to $\hat{f}(S)$ of $\{T \in \{0, 1\}^N : S \subsetneq T\}$ is bounded from above (in absolute value) by:

$$(p_0(1-p_0))^{\binom{r}{2}/2} \cdot \sum_{k=0}^{r-1} \binom{n}{r-k} \binom{r}{k} p_0^{\binom{r}{2}-\binom{k}{2}} = (p_0(1-p_0))^{\binom{r}{2}/2} \cdot (1/2 + o_n(1)), \tag{11}$$

since for our choice of $p_0$, the term corresponding to $k = 0$ equals $\binom{n}{r} p_0^{\binom{r}{2}} = 1/2$, and the other terms are negligible. Combining estimates (10) and (11), we get:

$$\hat{f}(S)^2 \ge (1 - 1/2 - o_n(1))^2 (p_0(1-p_0))^{\binom{r}{2}} \ge c p_0^{\binom{r}{2}} = (c/2) \cdot \binom{n}{r}^{-1}. \tag{12}$$

This completes the proof. □

We conclude this section by noting that if $p$ is inverse polynomially small as a function of $n$, then one can easily prove a statement which is only slightly weaker than the EI conjecture. In [15] it was shown that with respect to the uniform measure, we have

$$\mathrm{Ent}_{1/2}(f) \le (\log n + 1)\mathrm{I}_{1/2}(f) + 1,$$

for any Boolean $f$. The statement generalizes easily to a general biased measure $\mu_p$, and yields the following:

**Claim 2.5.** *There exists a universal constant $c$ such that for any $0 < p < 1$, for any $n$ and for any $f : \{0,1\}^n_p \to \{-1,1\}$, we have*

$$\mathrm{Ent}_p(f) \le cp(1-p)\log(n) \cdot \mathrm{I}_p(f).$$

**Proof.** Assume w.l.o.g. that $p \le 1/2$. As shown in [15], we have:

$$\mathrm{Ent}_p(f) \le (\log n + 1)\sum_S |S|\hat{f}(S)^2 + \epsilon\log(1/\epsilon) + 2\epsilon,$$

where $1 - \epsilon = \hat{f}(\emptyset)^2$. (Note that this part of the proof of Theorem 3.2 in [15] holds without any change for the biased measure.) In order to bound the term $\epsilon\log(1/\epsilon) + 2\epsilon$, it was shown in Proposition 3.6 of [15] that by the edge isoperimetric inequality on the cube, it is bounded from above by $2\mathrm{I}_{1/2}(f)$. By Eq. (2), this implies that for the measure $\mu_p$, we have

$$\epsilon\log(1/\epsilon) + 2\epsilon \le 12p\lfloor\log(1/p)\rfloor\mathrm{I}_p(f)$$

(since the reduction from the biased measure to the uniform measure preserves the expectation). Thus, by Eq. (1),

$$\mathrm{Ent}_p(f) \le (\log n + 1) \cdot 4p(1-p)\mathrm{I}_p(f) + 12p\lfloor\log(1/p)\rfloor\mathrm{I}_p(f) \le cp(1-p)\log(n)\mathrm{I}_p(f),$$

as asserted. □

For $p$ that is inverse polynomially small in $n$, the statement of Claim 2.5 differs from the assertion of the EI conjecture only by a constant factor.

## 3. Three remarks

In this section we remark on three questions related to the Entropy/Influence conjecture.

### 3.1. Functions with a low Fourier weight on the high levels

One of the important classes of Boolean functions for which one can hope to find a simpler proof of the EI conjecture are functions whose Fourier coefficients decay rapidly beyond some level. In the basic case where the coefficients beyond some level $k$ vanish, a variant of the conjecture was shown by Diakonikolas et al. [7]:

**Proposition 3.1** ([7]). *Let $f : \{-1,1\}^n_{1/2} \to \{-1,1\}$, such that all the Fourier weight of $f$ is concentrated on the first $k$ levels. Then all the Fourier coefficients of $f$ are of the form*

$$\hat{f}(S) = a(S) \cdot 2^{-k+1},$$

*where $a(S) \in \mathbb{Z}$. In particular, $\mathrm{Ent}_{1/2}(f) \le 2(k-1)$.*

We present a different proof of a slightly more general statement. Note that we replace the domain by $\{-1,1\}^n$ here, so that the characters are given by the simpler formula

$$u_{\{i_1,\ldots,i_r\}}(x) = x_{i_1} \cdot x_{i_2} \cdot \cdots \cdot x_{i_r}.$$

**Proposition 3.2.** *Let $f : \{-1,1\}^n_{1/2} \to \mathbb{Z}$, such that all the Fourier weight of $f$ is concentrated on the first $k$ levels. Then all the Fourier coefficients of $f$ are of the form*

$$\hat{f}(S) = a(S) \cdot 2^{-k},$$

*where $a(S) \in \mathbb{Z}$. In particular, $\mathrm{Ent}_{1/2}(f) \le 2k$.*

**Remark 3.3.** Note that Proposition 3.1 follows from Proposition 3.2 by moving from a function $f : \{-1,1\}^n \to \{-1,1\}$ to the function $(1+f)/2$ whose range is contained in $\mathbb{Z}$ and whose Fourier coefficients are halved.

**Proof.** The proof is by induction on $k$. The case $k = 0$ is trivial. Assume that the assertion holds for all $k \leq d-1$, and let $f$ be a function of Fourier degree $d$ (i.e., all its Fourier coefficients are concentrated on the $d$ lowest levels). For $1 \leq i \leq n$, let $f^i$ be the discrete derivative of $f$ with respect to the $i$th coordinate, i.e.,

$$f^i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = \frac{f(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, -1, x_{i+1}, \ldots, x_n)}{2}.$$

It is easy to see that if $f = \sum_S \hat{f}(S) u_S$, then the Fourier expansion of $f^i$ is given by:

$$f^i = \sum_{S \subset (\{1,2,\ldots,n\} \setminus \{i\})} \hat{f}(S \cup \{i\}) u_S. \tag{13}$$

Hence, $f^i$ is of Fourier degree at most $d-1$. Note that by the definition of $f^i$, we have $2f^i(x) \in \mathbb{Z}$ for all $x \in \{-1,1\}^{n-1}$, and thus by the induction hypothesis, the Fourier coefficients of $f^i$ satisfy $2\hat{f^i}(S) = a(S) \cdot 2^{-d+1}$, where $a(S) \in \mathbb{Z}$. This holds for any $1 \leq i \leq n$, and therefore, by Eq. (13), all the Fourier coefficients of $f$ (except, possibly, for $\hat{f}(\emptyset)$), are of the form $\hat{f}(S) = a(S) \cdot 2^{-d}$, where $a(S) \in \mathbb{Z}$. Finally, $\hat{f}(\emptyset)$ must also be of this form, since otherwise $f(x)$ cannot be an integer. This completes the proof. □

A more interesting case in which the Fourier coefficients beyond some level $k$ decay exponentially was covered in an unpublished work of Bourgain and Kalai [6].

**Theorem 3.4.** *There exists a function $C : (0, 1/2) \times (0, \infty) \to (0, \infty)$ such that the following holds: For any $n \in \mathbb{N}$, let $f : \{-1,1\}^n \to \{-1,1\}$ and assume that there exist $c_0 > 0$, $0 < a < 1/2$, and $0 < k < n$, such that for all $0 \leq t \leq n$,*

$$\sum_{\{S:|S|>t\}} \hat{f}(S)^2 \leq e^{c_0 k} \cdot e^{-at},$$

*then for any $\alpha > 1$, there exists a set $B_\alpha = B(f, \alpha, c_0, a, k)$, such that:*

1. $\log |B_\alpha| \leq C(c_0, a) \cdot \alpha k$.
2. $\sum_{S \notin B_\alpha} \hat{f}(S)^2 \leq n^{-\alpha}$.

The theorem asserts that if the Fourier weight of $f$ beyond the $k$th level decays exponentially, then most of the Fourier weight of $f$ is concentrated on $\exp(Ck)$ coefficients, and thus, $\mathrm{Ent}_{1/2}(f) \leq C'k$ (for an appropriate choice of $C'$). The proof uses the $d$th discrete derivative of $f$ (like our proof of Proposition 3.2 above), and the Bonami–Beckner hypercontractive inequality [2,1]. We note that the exact dependence of $C$ on $a$ (i.e., the rate of the exponential decay) in the assertion of the theorem, which is important if $a$ is allowed to be a function of $n$, is of order $C = \Theta(a^{-1}\log(a^{-1}))$.

We would like to suggest a possible way to attack the entire EI conjecture via a combination of a *tensorisation technique* and an extension of Theorem 3.4 (and Proposition 3.1).

In [11], Kalai observed that the EI conjecture tensorises, in the following sense. For $f : \{-1,1\}^l_{1/2} \to \{-1,1\}$ and $g : \{-1,1\}^m_{1/2} \to \{-1,1\}$, define $f \otimes g : \{-1,1\}^{l+m}_{1/2} \to \{-1,1\}$ by:

$$f \otimes g(x_1, \ldots, x_{l+m}) = f(x_1, \ldots, x_l) \cdot g(x_{l+1}, \ldots, x_{l+m}).$$

Furthermore, let

$$f^{\otimes N} = f \otimes f \otimes \cdots \otimes f,$$

where the tensorisation is performed $N$ times. It is easy to see that $\mathrm{I}_{1/2}(f^{\otimes N}) = N \cdot \mathrm{I}_{1/2}(f)$ and $\mathrm{Ent}_{1/2}(f^{\otimes N}) = N \cdot \mathrm{Ent}_{1/2}(f)$. Hence, proving the EI conjecture for any "tensor power" of $f$ is equivalent to proving the conjecture for $f$ itself.

This observation was used in [15] to deduce that it is sufficient to prove a seemingly weaker version of the conjecture: $\mathrm{Ent}_{1/2}(f) \leq c\mathrm{I}_{1/2}(f) + o(n)$, where $n$ is the number of variables.

We observe that by the Law of Large Numbers, as $N \to \infty$, the level of the Fourier coefficients of $f^{\otimes N}$ is concentrated around its expectation, which is $N \cdot \mathrm{I}_{1/2}(f)$. In particular, it is easy to check that the total weight of coefficients above level $2N \cdot \mathrm{I}_{1/2}(f)$ goes to 0 as $N \to \infty$.

If it were the case that for $N$ large enough the weight above level $N \cdot \mathrm{I}_{1/2}(f)$ is zero, then Proposition 3.1 would imply the EI conjecture in the form $\mathrm{Ent}_{1/2}(f^{\otimes N}) \leq 2N \cdot \mathrm{I}_{1/2}(f)$. Bourgain–Kalai's Theorem 3.4 allows to replace the unreasonable condition that there is no weight above level $2N \cdot \mathrm{I}_{1/2}(f)$ by a weaker condition on the decay of the Fourier spectrum of $f^{\otimes N}$.

Unfortunately, even Bourgain–Kalai's Theorem 3.4 (applied for $k = \Theta(N \cdot \mathrm{I}_{1/2}(f))$ together with a large deviation estimate) is not sufficient to obtain the Entropy/Influence conjecture. However, we believe that either a strengthening of Bourgain–Kalai's Theorem or a stronger utilization of the tensor structure may lead to a proof of the conjecture along the lines suggested here.

*3.2. A weaker upper bound on the entropy of Boolean functions that can be easily proved*

As mentioned in Section 2, it was shown in [15] that with respect to the uniform measure, one can easily prove the following weaker upper bound on the entropy of any Boolean function:

$$\mathrm{Ent}_{1/2}(f) \leq (\log n + 1)\mathrm{I}_{1/2}(f) + 1. \tag{14}$$

We provide an independent proof of an upper bound which is slightly stronger than (14).

**Claim 3.5.** *For any $n$ and for any $f : \{0,1\}^n_{1/2} \to \mathbb{R}$, we have*

$$\mathrm{Ent}_{1/2}(f) \leq \sum_{i=1}^{n} h(\mathrm{I}_i^{1/2}(f)) \leq 2\mathrm{I}_{1/2}(f)(1 + \log n - \log \mathrm{I}_{1/2}(f)),$$

*where*

$$h(x) = -x\log x - (1-x)\log(1-x).$$

**Proof.** As the proof deals only with the uniform measure on the discrete cube, we write $\mathrm{Ent}(f)$ and $\mathrm{I}(f)$ instead of $\mathrm{Ent}_{1/2}(f)$ and $\mathrm{I}_{1/2}(f)$ during the proof.

Let $S \subset \{1, 2, \ldots, n\}$ be chosen according to the Fourier distribution (i.e., $\Pr[S = S_0] = \hat{f}(S_0)^2$), and let $X_i = 1_{\{i \in S\}}$. Then by the basic rules of entropy,

$$\mathrm{Ent}(f) = H(S) = H(X_1, \ldots, X_n) \leq \sum_{i=1}^{n} H(X_i) = \sum_{i=1}^{n} h(\mathrm{I}_i(f)),$$

thus obtaining the first inequality. Note that if $\mathrm{I}_i(f) \geq 0.5$, then $h(\mathrm{I}_i(f)) \leq 2\mathrm{I}_i(f)$, and otherwise, $h(\mathrm{I}_i(f)) \leq -2\mathrm{I}_i(f)\log \mathrm{I}_i(f)$. Therefore,

$$\begin{aligned}\frac{1}{2}\mathrm{Ent}(f) &\leq \mathrm{I}(f) + \sum_{i=1}^{n} \mathrm{I}_i(f)(-\log \mathrm{I}_i(f)) \\ &= \mathrm{I}(f)\left(1 + \sum_{i=1}^{n} \frac{\mathrm{I}_i(f)}{\mathrm{I}(f)} \cdot \left(-\log\frac{\mathrm{I}_i(f)}{\mathrm{I}(f)} - \log \mathrm{I}(f)\right)\right).\end{aligned}$$

We note that the expression $\sum_{i=1}^{n} \mathrm{I}_i(f)/\mathrm{I}(f)(-\log \mathrm{I}_i(f)/\mathrm{I}(f))$ is the entropy of the random variable $Y$ defined by $\Pr[Y = i] = \mathrm{I}_i(f)/\mathrm{I}(f)$ which is supported on $\{1, 2, \ldots, n\}$, and is therefore bounded by $\log n$. We thus conclude that

$$\frac{1}{2}\mathrm{Ent}(f) \leq \mathrm{I}(f)(1 + \log n - \log \mathrm{I}(f))$$

as asserted. □

**Remark 3.6.** It is easy to see that the bound of Claim 3.5 is stronger than (14) in some cases, in particular when there is variability in the influences of different coordinates.

We note that the proof does not use the fact that $f$ is Boolean and indeed it could not provide a proof of the EI conjecture, as can be seen, e.g., for the majority function, where $\mathrm{I}_{1/2}(f)$ is of order $\sqrt{n}$ while $\sum_{i=1}^{n} h(\mathrm{I}_i^{1/2}(f))$ is of order $\sqrt{n}\log n$.

*3.3. Relation to Friedgut's characterization of functions with a low influence sum*

In [8], Friedgut showed that any Boolean function $f : \{0,1\}^n_p \to \{0,1\}$ essentially depends on at most $C(p)^{\mathrm{I}(f)}$ coordinates, where $C(p)$ depends only on $p$. The main step of the proof is to show that most of the Fourier weight of the function is concentrated on sets that contain one of these coordinates. A stronger claim one may hope to prove is that most of the Fourier weight is concentrated on at most $C(p)^{\mathrm{I}(f)}$ coefficients. Formally, we raise the following conjecture that resembles the assertion of Bourgain–Kalai's theorem:

**Conjecture 3.7.** *For any $0 < p < 1$ and any $0 < \epsilon < 1$, there exists a constant $C(p,\epsilon) > 0$ such that for any $n$ and for any $f : \{-1,1\}^n_p \to \{-1,1\}$, there exists a set $B_\epsilon(f) \subset \{0,1\}^n$ such that:*

1. $\log|B_\epsilon(f)| \leq C(p,\epsilon) \cdot \mathrm{I}(f)$*, and*
2. $\sum_{S \notin B_\epsilon} \hat{f}(S)^2 < \epsilon$.

This conjecture looks quite strong and even implies a variant of Mansour's conjecture [14] (since as shown in [3], if a Boolean function $f$ can be represented by an $m$-term DNF, then $\mathrm{I}_{1/2}(f) = O(\log m)$), but it still does not imply the EI conjecture, since the remaining Fourier coefficients (whose total Fourier weight is at most $\epsilon$) can still contribute $n \cdot \epsilon$ to $\mathrm{Ent}_{1/2}(f)$.

We note that if Conjecture 3.7 holds, it clearly can be combined with Friedgut's theorem to obtain an additional condition on the set $B_\epsilon(f)$:

$$\log \left| \bigcup B_\epsilon(f) \right| \leq C(p, \epsilon) \cdot \mathrm{I}(f), \tag{15}$$

where $\bigcup B_\epsilon(f) \subset \{1, 2, \ldots, n\}$ is the set of all coordinates included in at least one of the sets $B_\epsilon(f)$. That is, not only the Fourier weight is concentrated on a "small" set of coefficients, these coefficients also depend on a "small" set of coordinates.

## Acknowledgments

We are grateful to Gil Kalai for introducing us to his unpublished work with Jean Bourgain [6], and to Ryan O'Donnell for sending us his recent work [15].

The first author was partially supported by the Koshland Center for Basic Research. The second author was supported by a DMS 0548249 (CAREER) award, by a DOD ONR grant N000141110140, by an ISF grant 1300/08 and by a Minerva Grant. The third author was partially supported by the Hoffman program for leadership.

## References

[1] W. Beckner, Inequalities in Fourier analysis, Ann. of Math. 102 (1975) 159–182.
[2] A. Bonami, Etude des coefficients Fourier des fonctiones de $L^p(G)$, Ann. Inst. Fourier 20 (1970) 335–402.
[3] R. Boppana, The average sensitivity of bounded-depth circuits, Inform. Process. Lett. 63 (5) (1997) 257–261.
[4] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, N. Linial, The influence of variables in product spaces, Israel J. Math. 77 (1992) 55–64.
[5] J. Bourgain, G. Kalai, Influences of variables and threshold intervals under group symmetries, Geom. Funct. Anal. 7 (1997) 438–461.
[6] J. Bourgain, G. Kalai, Unpublished Manuscript, 2000.
[7] I. Diakonikolas, H. Lee, K. Matulef, K. Onak, R. Rubinfeld, R. Servedio, A. Wan, Testing for concise representations, in: Proc. 48-th Ann. Symp. on Foundations of Comp. Sci., FOCS, 2007, pp. 549–558.
[8] E. Friedgut, Boolean functions with low average sensitivity depend on few coordinates, Combinatorica 18 (1) (1998) 27–35.
[9] E. Friedgut, G. Kalai, Every monotone graph property has a sharp threshold, Proc. Amer. Math. Soc. 124 (1996) 2993–3002.
[10] J. Kahn, G. Kalai, N. Linial, The influence of variables on Boolean functions, in: Proc. 29-th Ann. Symp. on Foundations of Comp. Sci., FOCS, Computer Society Press, 1988, pp. 68–80.
[11] G. Kalai, The entropy/influence conjecture, blog entry, 2007. Available on-line at: http://terrytao.wordpress.com/2007/08/16/gil-kalai-the-entropyinfluence-conjecture/.
[12] G. Kalai, M. Safra, Threshold phenomena and influence, in: A.G. Percus, G. Istrate, C. Moore (Eds.), Computational Complexity and Statistical Physics, Oxford University Press, New York, 2006, pp. 25–60.
[13] N. Keller, A simple reduction from a biased measure on the discrete cube to the uniform measure, 2010. Available online at: http://arxiv.org/abs/1001.1167.
[14] Y. Mansour, An $O(n^{\log\log n})$ learning algorithm for DNF under the uniform distribution, J. Comput. System Sci. 50 (3) (1995) 543–550.
[15] R. O'Donnell, J. Wright, Y. Zhou, The Fourier entropy-influence conjecture for certain classes of Boolean functions, in: Proceedings of ICALP'11, in: Lecture Notes in Computer Science, vol. 6755, Spinger-Verlag, 2011, pp. 330–341.