Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 84 (2015)

An Improved Upper Bound for the Most Informative Boolean Function Conjecture

Or Ordentlich, Ofer Shayevitz and Omri Weinstein *

Abstract

Suppose X is a uniformly distributed n-dimensional binary vector and Y is obtained by passing X through a binary symmetric channel with crossover probability α. A recent conjecture by Courtade and Kumar postulates that $I(f(X);Y) \leq 1 - h(\alpha)$ for any Boolean function f. So far, the best known upper bound was $I(f(X);Y) \leq (1-2\alpha)^2$. In this paper, we derive a new upper bound that holds for all balanced functions, and improves upon the best known bound for all $\frac{1}{3} < \alpha < \frac{1}{2}$.

1 Introduction

Let X be an n-dimensional binary vector uniformly distributed over $\{0,1\}^n$, and Y be the output of passing each component of X through a binary symmetric channel with crossover probability $\alpha \leq 1/2$. The following was recently conjectured by Courtade and Kumar [1].

Conjecture 1 *For any Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ it holds that*

$$I\left(f(X);Y\right) \leq 1 - h(\alpha), \tag{1}$$

*Or Ordentlich and Ofer Shayevitz are with the Department of Electrical Engineering - Systems at the Tel Aviv University, {ordent,ofersha}@eng.tau.ac.il. Omri Weinstein is with the Department of Computer Science, Princeton University, oweinste@cs.princeton.edu. The work of O. Ordentlich was supported by the Admas Fellowship Program of the Israel Academy of Science and Humanities. The work of O. Shayevitz was supported by an ERC grant no. 639573, a CIG grant no. 631983, and an ISF grant no. 1367/14. The work of O. Weinstein was supported by a Simons Fellowship award in TCS and a Siebel scholarship.

ISSN 1433-8092

where $I(f(X);Y)$ is the mutual information between $f(X)$ and Y, and $h(p) \triangleq -p\log p - (1-p)\log(1-p)$ is the binary entropy function.[1]

For a dictatorship function, $f(X) = X_i$ the conjectured upper bound (1) is attained with equality. Therefore, the conjecture can be interpreted as postulating that dictatorship is the most "informative" Boolean function, i.e., it achieves the maximal $I(f(X);Y)$.

So far, the best known bound that holds universally for all Boolean functions is

$$I\left(f(X);Y\right) \leq (1-2\alpha)^2. \tag{2}$$

This bound can be established through various techniques, including an application of Mrs. Gerber's Lemma [2–4], the strong data-processing inequality [5, 6] and standard Fourier analysis as described below.

In this paper, we derive an upper bound on $I(f(X);Y)$ that holds for all balanced functions, and improves upon (2) for all $\frac{1}{3} < \alpha < \frac{1}{2}$. Specifically, we obtain the following result.

Theorem 1 *For any balanced Boolean function $f(x) : \{0,1\}^n \mapsto \{0,1\}$, and any $\frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \leq \alpha \leq \frac{1}{2}$, we have that*

$$I(f(X);Y) \leq \frac{\log(e)}{2}(1-2\alpha)^2 + 9\left(1 - \frac{\log(e)}{2}\right)(1-2\alpha)^4. \tag{3}$$

For the proof of Theorem 1, we first lower bound the conditional entropy $H(f(X)|Y)$ in terms of the second and fourth moments of the random variable $(1-2P_Y^f)$, where $P_y^f \triangleq \Pr\left(f(X)=0|Y=y\right)$. Specifically, we show that

$$H(f(X)|Y) \geq 1 - \frac{\log(e)}{2}\mathbb{E}(1-2P_Y^f)^2 + \left(1 - \frac{\log(e)}{2}\right)\mathbb{E}(1-2P_Y^f)^4. \tag{4}$$

To upper bound the second and fourth moments in (4), we use basic Fourier analysis of Boolean functions along with a simple application of the Hypercontractivity Theorem [7–9], in order to derive universal upper bounds on

[1] All logarithms are taken to base 2.

$\mathbb{E}(1-2P_Y^f)^{2k}$ that hold for all balanced Boolean functions.[2] In particular, these bounds show that $\mathbb{E}(1-2P_Y^f)^2 \leq (1-2\alpha)^2$ and $\mathbb{E}(1-2P_Y^f)^4 \leq 9(1-2\alpha)^4$. Plugging these bounds in (4) yields the Theorem.

An appealing feature of the new upper bound, is that the ratio between the RHS of (3) and $1-h(\alpha)$ approaches 1 in the limit of $\alpha \to \frac{1}{2}$. For the bound (2), on the other hand, the same ratio does not approach 1. Rather

$$\lim_{\alpha \to \frac{1}{2}} \frac{(1-2\alpha)^2}{1-h(\alpha)} = \frac{2}{\log(e)} \approx 1.3863.$$

2 Preliminaries

To prove our results, it will be more convenient to map the additive group $\{0,1\}^n$ to the (isomorphic) multiplicative group $\{-1,1\}^n$. Specifically, let X be an n-dimensional binary vector uniformly distributed over $\{-1,1\}^n$ and Y be the output of passing each component of X through a binary symmetric channel with crossover probability $\alpha \leq 1/2$. Thus, for all $i \in \{1,\cdots,n\}$ we have $Y_i = X_i \cdot Z_i$, where $\{Z_i\}_{i=1}^n$ is an i.i.d. sequence of binary random variables statistically independent of $\{X_i\}_{i=1}^n$, with $\Pr(Z_i = -1) = \alpha$ and $\Pr(Z_i = 1) = 1-\alpha$. Note that Y is also uniformly distributed over $\{-1,1\}^n$.

Let $f : \{-1,1\}^n \mapsto \{-1,1\}$ be a *balanced* Boolean function, i.e.,

$$\Pr\left(f(X) = -1\right) = \frac{1}{2}.$$

Note that this condition is equivalent to $\mathbb{E}_X(f(X)) = 0$. For each $y \in \{-1,1\}^n$ define the posterior distribution of $f(X)$ given the observation y

$$P_y^f \triangleq \Pr(f(X) = -1 | Y = y),$$

and note that

$$\mathbb{E}(f(X)|Y=y) = 1 - 2P_y^f. \tag{5}$$

[2] We remark that using similar techniques it is possible to obtain an upper bound on $I(f(X);Y)$ of a similar form that holds for any q-biased function (i.e., any function for which $\mathbb{E}(f(X)) = 1-2q$). However, the obtained bound is not maximized at $q = \frac{1}{2}$, and therefore cannot be used to establish (3) for all Boolean functions.

In what follows, we will extensively use Fourier analysis of real functions on the hypercube $\{-1,1\}^n$. Let X be a random vector uniformly distributed on $\{-1,1\}^n$ and define $[n] \triangleq \{1,2,\ldots,n\}$. The Fourier-Walsh transform of a function $f:\{-1,1\}^n \mapsto \mathbb{R}$ is given by

$$f(x) = \sum_{S\subseteq[n]} \hat{f}(S) \prod_{i\in S} x_i, \tag{6}$$

where

$$\hat{f}(S) = \mathbb{E}_X \left(f(X) \prod_{i\in S} X_i \right) \tag{7}$$

is the correlation of f with the "parity" function on the subset S. It is easy to verify that the basis $\left\{\varphi_S(x) = \prod_{i\in S} x_i\right\}_{S\subseteq[n]}$ is orthonormal with respect to the inner product $< f,g >= \mathbb{E}\left(f(X)g(X)\right)$, which implies that for any two functions $f,g:\{-1,1\}^n \mapsto \mathbb{R}$ it holds that

$$\mathbb{E}\left(f(X)g(X)\right) = \sum_{S\subseteq[n]} \hat{f}(S)\hat{g}(S). \tag{8}$$

In particular, Parseval's identity gives $\sum_S \hat{f}^2(S) = \mathbb{E}\left(f^2(X)\right)$. Thus, if f is a Boolean function, i.e., $f:\{-1,1\}^n \mapsto \{-1,1\}$, we have $\sum_S \hat{f}^2(S) = 1$.

The following proposition gives three simple properties of the Fourier coefficients that will be useful in the sequel.

Proposition 1 (Basic properties of the Fourier transform) *Let $f:\{-1,1\}^n \mapsto \{-1,1\}$ be a Boolean function. We have that*

(i) f is balanced if and only if $\hat{f}(\emptyset) = 0$;

(ii) If $|\hat{f}(S)| > 0$ then $|\hat{f}(S)| \geq 2^{-n}$;

(iii) Any balanced function f which is not a dictatorship function must satisfy $\sum_{S:\,|S|\geq 2} \hat{f}^2(S) > 0$;

Proof.

(i) By definition $\hat{f}(\emptyset) = \mathbb{E}f(X) = 0$ for f balanced.

(ii) We note that the sum $\sum_{x\in\{-1,1\}^n} f(x)\prod_{i\in S} x_i$ is an integer, and the claim follows immediately.

(iii) Let f be a balanced Boolean function ($\hat{f}(\emptyset)=0$) with $\sum_{S:\,|S|\geq 2}\hat{f}^2(S)=0$. Then $f(X)=\sum_{i\in[n]}\hat{f}(\{i\})X_i$. Since there always exists some $x\in\{-1,1\}^n$ for which $f(x)=\sum_i|\hat{f}(\{i\})|$, we must have that $\sum_i|\hat{f}(\{i\})|=1$. On the other hand, by Parseval's identity we have $\sum_{i=1}\hat{f}^2(\{i\})=1$. Clearly, the last two equations can simultaneously hold if and only if there is a unique i for which $|\hat{f}(\{i\})|>0$. Hence, f must be a dictatorship.

■

For any function $f:\{-1,1\}^n\mapsto\mathbb{R}$ with Fourier coefficients $\{\hat{f}(S)\}$ and $\rho\in\mathbb{R}^+$, the *noise operator* $T_\rho f:\{-1,1\}^n\mapsto\mathbb{R}$ is defined as [7]

$$(T_\rho f)(x)\triangleq\sum_{S\subseteq[n]}\hat{f}(S)\rho^{|S|}\prod_{i\in S}x_i. \tag{9}$$

Recall that in our setting X and Y are the input and output of a binary symmetric channel with crossover probability α. Thus we can use the Fourier representation of f to write $\mathbb{E}\left(f(X)|Y=y\right)$ as follows [7]

$$\begin{aligned}
\mathbb{E}\left(f(X)|Y=y\right) &= \mathbb{E}\left(\sum_{S\subseteq[n]}\hat{f}(S)\prod_{i\in S}X_i\;\Big|Y=y\right)\\
&= \sum_{S\subseteq[n]}\hat{f}(S)\mathbb{E}\left(\prod_{i\in S}y_iZ_i\right)\\
&= \sum_{S\subseteq[n]}\hat{f}(S)\mathbb{E}\left(\prod_{i\in S}Z_i\right)\prod_{i\in S}y_i\\
&= \sum_{S\subseteq[n]}\hat{f}(S)\prod_{i\in S}\mathbb{E}\left(Z_i\right)\prod_{i\in S}y_i\\
&= \sum_{S\subseteq[n]}\hat{f}(S)(1-2\alpha)^{|S|}\prod_{i\in S}y_i, \qquad (10)\\
&= (T_{1-2\alpha}f)(y), \qquad (11)
\end{aligned}$$

where we have used the fact that $\{Z_i\}$ is an i.i.d. sequence. Recalling equation (5), this also yields

$$1 - 2P_y^f = (T_{1-2\alpha}f)(y) = \sum_{S\subseteq[n]} \hat{f}(S)(1-2\alpha)^{|S|} \prod_{i\in S} y_i. \tag{12}$$

The following well-known theorem will play an important role in the derivation of Theorem 1.

Theorem 2 (Hypercontractivity Theorem, [7–9]) *Let $1 \leq p < q < \infty$. Then for all $\rho \leq \sqrt{\frac{p-1}{q-1}}$, and all $g : \{-1,1\}^n \mapsto \mathbb{R}$, it holds that*

$$[\mathbb{E}(|(T_\rho g)(X)|^q)]^{\frac{1}{q}} \leq [\mathbb{E}(|g(X)|^p)]^{\frac{1}{p}}. \tag{13}$$

3 Proof of Theorem 1

In this section we prove Theorem 1, which gives a universal upper bound on the mutual information $I(f(X);Y)$, that holds for any *balanced* Boolean function. The mutual information can be expressed as

$$I(f(X);Y) = H(f(X)) - H(f(X)|Y) = 1 - \mathbb{E}_Y h(P_Y^f). \tag{14}$$

We note that $h(\cdot)$ admits the following Taylor series

$$h\left(\frac{1-p}{2}\right) = 1 - \sum_{k=1}^{\infty} \frac{\log(e)}{2k(2k-1)} p^{2k}, \tag{15}$$

and can be lower bounded by replacing p^{2k} with p^{2t} for all $k > t$, i.e.,

$$\begin{aligned} h\left(\frac{1-p}{2}\right) &\geq 1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)} p^{2k} - p^{2t} \sum_{k=t}^{\infty} \frac{\log(e)}{2k(2k-1)} \\ &= 1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)} p^{2k} - \left(1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)}\right) p^{2t}, \end{aligned} \tag{16}$$

where we have used the fact that $h(0) = 0$. Using (16), for any $t \in \mathbb{N}$, we can further lower bound $\mathbb{E}_Y h(P_Y^f)$ as

$$\begin{aligned}\mathbb{E}_Y h(P_Y^f) &= \mathbb{E}_Y h\left(\frac{1-(1-2P_Y^f)}{2}\right) \\ &\geq 1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)} \mathbb{E}_Y\left((1-2P_Y^f)^{2k}\right) \\ &\quad - \left(1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)}\right) \mathbb{E}_Y\left((1-2P_Y^f)^{2t}\right). \end{aligned} \tag{17}$$

Thus, any upper bound on the first t even moments $\mathbb{E}[(1-2P_Y^f)^{2k}]$, $k = 1, \ldots, t$, would directly translate to an upper bound on $I(f(X); Y)$.

We obtain an upper bound on all even moments of the random variable $1-2P_Y^f$, using a simple trick combined with the Hypercontractivity Theorem. Formally, we prove the following lemma.

Lemma 1 *Let $k \geq 1$ be an integer satisfying $(1-2\alpha)\sqrt{2k-1} \leq 1$. For any balanced Boolean function $f : \{-1,1\}^n \mapsto \{-1,1\}$ we have that*

$$\mathbb{E}_Y\left((1-2P_Y^f)^{2k}\right) \leq (2k-1)^k (1-2\alpha)^{2k}.$$

For the proof, we will need the following proposition.

Proposition 2 *For any balanced Boolean function $f : \{-1,1\}^n \mapsto \{-1,1\}$ and any $0 \leq \rho \leq 1$ we have that*

$$\mathbb{E}_Y\left((T_\rho f)^2(Y)\right) \leq \rho^2.$$

Proof. By the definition of the operator T_ρ:

$$\begin{aligned}\mathbb{E}_Y\left((T_\rho f)^2(Y)\right) &= \mathbb{E}_Y\left(\left(\sum_{S\subseteq[n]} \hat{f}(S)\rho^{|S|} \prod_{i\in S} Y_i\right)^2\right) \\ &= \mathbb{E}_Y\left(\sum_{S_1\subseteq[n]} \hat{f}(S_1)\rho^{|S_1|} \prod_{i\in S_1} Y_i \sum_{S_2\subseteq[n]} \hat{f}(S_2)\rho^{|S_2|} \prod_{j\in S_2} Y_j\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{S_1 \subseteq [n]} \hat{f}(S_1)\rho^{|S_1|} \sum_{S_2 \subseteq [n]} \hat{f}(S_2)\rho^{|S_2|} \mathbb{E}_Y \left(\prod_{i \in S_1} Y_i \prod_{j \in S_2} Y_j \right) \\
&= \sum_{S_1 \subseteq [n]} \hat{f}(S_1)\rho^{|S_1|} \sum_{S_2 \subseteq [n]} \hat{f}(S_2)\rho^{|S_2|} \mathbf{1}_{S_1 = S_2} \\
&= \sum_{S \subseteq [n]} \hat{f}^2(S)\rho^{2|S|}, \qquad (18)
\end{aligned}
$$

where $\mathbf{1}_{S_1=S_2}$ is the indicator function on the event $S_1 = S_2$. Recalling that $\sum_{S \subseteq [n]} \hat{f}^2(S) = 1$ and our assumptions that $\rho \leq 1$ and that $\hat{f}(\emptyset) = \mathbb{E}(f(X)) = 0$, the "weight" assignment $\hat{f}^2(S)$ that maximizes $\sum_{S \subseteq [n]} \hat{f}^2(S)\rho^{2|S|}$ puts all the weight on characters S whose cardinality is $|S| = 1$. Hence,

$$
\sum_{S \subseteq [n]} \hat{f}^2(S)\rho^{2|S|} \leq \rho^2,
$$

as desired. ■

Proof of Lemma 1. Let $f : \{-1,1\}^n \mapsto \{-1,1\}$ be a Boolean function and let $g \triangleq T_{1-2\alpha} f$. Let $k \geq 1$ be an integer, and let $\rho = \sqrt{1/(2k-1)} = \sqrt{(2-1)/(2k-1)}$. By (12), we have

$$
\begin{aligned}
\mathbb{E}_Y((1 - 2P_Y^f)^{2k}) &= \mathbb{E}_Y \left(g^{2k}(Y) \right) \\
&= \mathbb{E}_Y \left(\left(T_\rho (T_{1/\rho} g) \right)^{2k} (Y) \right) \\
&\leq \left[\mathbb{E}_Y \left((T_{1/\rho} g)^2 (Y) \right) \right]^k \qquad (19) \\
&= \left[\mathbb{E}_Y \left((T_{(1-2\alpha)\sqrt{2k-1}} f)^2 (Y) \right) \right]^k, \qquad (20)
\end{aligned}
$$

where (19) follows from the Hypercontractivity Theorem taken with $p = 2, q = 2k$ (as ρ satisfies the premise of the theorem), and (20) from the definition of the function g. Invoking Proposition 2 with $\rho = (1-2\alpha)\sqrt{2k-1} \leq 1$, we obtain that for any balanced Boolean function and any $k \geq 1$

$$
\mathbb{E}_Y \left((1 - 2P_Y^f)^{2k} \right) \leq (2k-1)^k (1-2\alpha)^{2k} \qquad (21)
$$

as desired. ■

The following is an immediate consequence of Lemma 1 and (17).

Proposition 3 *For any balanced Boolean function $f : \{-1,1\}^n \mapsto \{-1,1\}$, any integer $t \geq 1$ and any $\frac{1}{2}\left(1 - \frac{1}{\sqrt{2t-1}}\right) \leq \alpha \leq \frac{1}{2}$, we have that*

$$I(f(X);Y) \leq \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)}(2k-1)^k(1-2\alpha)^{2k} + \left(1 - \sum_{k=1}^{t-1} \frac{\log(e)}{2k(2k-1)}\right)(2t-1)^t(1-2\alpha)^{2t}. \quad (22)$$

Theorem 1 now follows by evaluating (22) with $t = 2$. Note that for balanced functions the upper bound $(1-2\alpha)^2$, which was the best known bound hitherto, is obtained as a special case of Proposition 3 by setting $t = 1$. It is easy to verify that for $\frac{1}{3} < \alpha < \frac{1}{2}$ the upper bound in Theorem 1 is tighter. See Figure 1 for a comparison between the bounds.

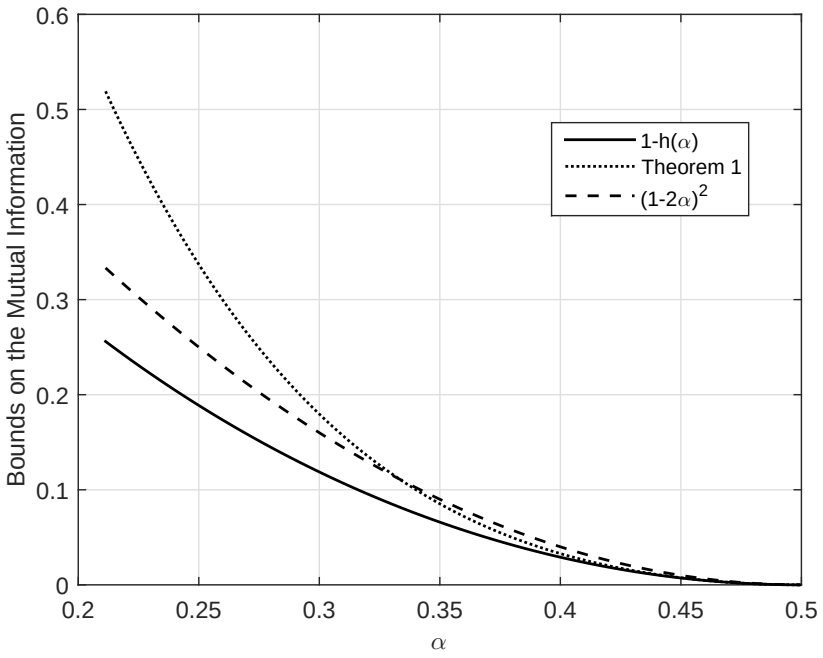

Figure 1: A comparison between the bound from Theorem 1, the conjectured bound and best known previous upper bound on $I(f(X);Y)$. Note that the bound from Theorem 1 is only valid for balanced functions.

Remark 1 *In [1, Appendix B, Remark 6], it is claimed that Conjecture 1 can be shown to hold in the limit of* $\alpha \to \frac{1}{2}$. *Theorem 1 demonstrates this fact for balanced functions, as for* $\alpha \to \frac{1}{2}$ *the ratio between the RHS of* (3) *and the conjectured bound tends to* 1. *As we discuss below, a slightly stronger statement can be shown to hold.*

The next simple corollary of Theorem 1 establishes the optimality of the dictatorship function among all balanced functions in the very noisy regime. The proof is essentially a consequence of the discreteness of the space of Boolean functions.

Corollary 1 (Dictatorship is optimal for very noisy channels) *Let* $f : \{0,1\}^n \mapsto \{0,1\}$ *be any balanced Boolean function. There exists* $\overline{\alpha}_n > 0$ *such that* $I(f(X);Y) \leq 1 - h(\alpha)$ *for all* $\alpha \in \left[\frac{1}{2} - \overline{\alpha}_n, \frac{1}{2}\right]$. *In particular, dictatorship is the most informative balanced function in the noise interval* $\alpha \in \left[\frac{1}{2} - \overline{\alpha}_n, \frac{1}{2}\right]$.

Proof. By equation (17) applied with $t = 2$, we have that for any balanced Boolean function (and any α),

$$I\left(f(X);Y\right) \leq \frac{\log(e)}{2}\mathbb{E}_Y\left((1-2P_Y^f)^2\right) + \left(1 - \frac{\log(e)}{2}\right)\mathbb{E}_Y\left((1-2P_Y^f)^4\right).$$

Lemma 1 (applied with $k = 2$) implies that for $\frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \leq \alpha \leq \frac{1}{2}$

$$\mathbb{E}_Y\left((1-2P_Y^f)^4\right) \leq (2\cdot 2-1)^2(1-2\alpha)^4 = 9(1-2\alpha)^4. \tag{23}$$

Furthermore, by equations (12) and (18), we have that for α in this range

$$\begin{aligned}\mathbb{E}_Y\left((1-2P_Y^f)^2\right) &= \sum_{S\subseteq[n]} \hat{f}^2(S)(1-2\alpha)^{2|S|} \\ &\leq \left(\sum_{|S|=1}\hat{f}^2(S)\right)(1-2\alpha)^2 + \left(\sum_{|S|\geq 2}\hat{f}^2(S)\right)(1-2\alpha)^4, \end{aligned} \tag{24}$$

where we have used the fact that $\hat{f}(\emptyset) = 0$ for balanced functions (Proposition 1). Combining the above three inequalities and using the fact that

$\sum_{S\subseteq[n]} \hat{f}^2(S) = 1$, yields

$$I\left(f(X);Y\right) \leq \frac{\log(e)}{2}\left(1-\sum_{|S|\geq 2}\hat{f}^2(S)\right)(1-2\alpha)^2+$$
$$+\left(9\left(1-\frac{\log(e)}{2}\right)+\frac{\log(e)}{2}\sum_{|S|\geq 2}\hat{f}^2(S)\right)(1-2\alpha)^4. \tag{25}$$

Now, suppose that f is a balanced Boolean function which is *not* a dictatorship function. This implies (in fact, equivalent to) $\sum_{|S|\geq 2}\hat{f}^2(S) > 0$. By Proposition 1, it therefore must be the case that

$$\sum_{|S|\geq 2}\hat{f}^2(S) \geq 2^{-2n} = 4^{-n}.$$

Therefore, for such f, the RHS of (25) can be upper bounded by

$$\frac{\log(e)}{2}(1-4^{-n})(1-2\alpha)^2+\left(9\left(1-\frac{\log(e)}{2}\right)+\frac{\log(e)}{2}4^{-n}\right)(1-2\alpha)^4. \tag{26}$$

It can be directly verified that for any $\alpha \in \left[\frac{1}{2}-\overline{\alpha}_n, \frac{1}{2}\right]$, where $\overline{\alpha}_n \triangleq \frac{1}{4}\cdot 2^{-n}$, the expression in (26) is smaller than $\frac{\log(e)}{2}(1-2\alpha)^2 < 1-h(\alpha)$, and thus by (25), for such α we have

$$I\left(f(X);Y\right) \leq 1-h(\alpha), \tag{27}$$

which completes the proof. ■

Remark 2 *In an unpublished work, Sushant Sachdeva, Alex Samorodnitsky and Ido Shahaf have shown, using different techniques, that dictatorship is optimal among all (not just balanced) Boolean functions from* $\{0,1\}^n$ *to* $\{0,1\}$ *for all* $\alpha \in \left[\frac{1}{2}-2^{-O(n)}, \frac{1}{2}\right]$.

4 Discussion

A natural question is: What are the limits of the approach pursued in this paper? To this end we note the following two limitations:

Firstly, the dictatorship function, which is conjectured to be optimal, satisfies $\mathbb{E}_Y\left((1-2P_Y^f)^{2k}\right) = (1-2\alpha)^{2k}$ for every $k \in \mathbb{N}$. The ratio between the bound in Lemma 1 on the k-th moment of any balanced Boolean function and $(1-2\alpha)^{2k}$ grows rapidly with k. For this reason, we only get mileage from applying the lemma with $k = 1, 2$ and not for higher moments.

The second limitation is that Lemma 1 upper bounds *each moment separately*, while we are seeking an upper bound on the entire distribution (weighted sum) of the moments: Quantifying the tradeoff between higher and lower moments seems to be one of the "brick walls" in proving the conjecture. For example, the dictatorship function has the largest second moment among all balanced functions, but it is not hard to see that the majority function, for example, has a much larger (relatively speaking) kth moment for very large values of k. To see this, note that for $k \gg 2^n$,

$$\mathbb{E}_Y\left((1-2P_Y^f)^{2k}\right) \gtrsim 2^{-n} \cdot \max_y |1-2P_y^f|^{2k}.$$

For the majority function, the maximum is attained at $y = (1, 1, , 1 \ldots, 1)$ for which $P_y^f \approx 2^{-nD\left(\frac{1}{2}||\alpha\right)}$ and consequently $\max_y |1-2P_y^{Maj}| \approx 1-2^{-nD\left(\frac{1}{2}||\alpha\right)}$. For dictatorship, on the other hand, $|1-2P_Y^f| = 1-2\alpha$ for every y, and therefore

$$\max_y |1-2P_y^{Maj}|^{2k} \gg \max_y |1-2P_y^{Dict}|^{2k}.$$

Therefore, one cannot hope to prove that there is a single function that simultaneously maximizes all moments; rather, the conjecture postulates that there is some tradeoff between these values and the largest mutual information is attained by functions that maximize lower moments at the expense of higher ones.

5 Acknowledgement

In a previous version of this paper (http://arxiv.org/abs/1505.05794v1), we have presented Corollary 1 as a new result, and in addition we have proved that dictatorship is the optimal function for all $\alpha \in \left[0, \frac{2^{-2n}}{16n^2}\right]$. We are grateful to Thomas Courtade for bringing to our attention that (slightly weaker versions of) these results were already known, as partially discussed in Remark 1.

References

[1] T. Courtade and G. Kumar, "Which Boolean functions maximize mutual information on noisy inputs?" *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4515–4525, Aug 2014.

[2] A. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications–I," *IEEE Transactions on Information Theory*, vol. 19, no. 6, pp. 769–772, Nov 1973.

[3] E. Erkip, "The efficiency of information in investment," Ph.D. dissertation, Stanford University, CA, 1996.

[4] V. Chandar and A. Tchamkerten, "Most informative quantization functions," in *Proc. ITA Workshop, San Diego, CA, USA*, Feb. 2014, available online http://perso.telecom-paristech.fr/~tchamker/CTAT.pdf.

[5] R. Ahlswede and P. Gacs, "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, vol. 4, no. 6, pp. 925–939, 1976.

[6] V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On hypercontractivity and the mutual information between Boolean functions," in *51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 13–19.

[7] R. O'Donnell, *Analysis of Boolean functions*. Cambridge University Press, 2014.

[8] A. Bonami, "tude des coefficients de fourier des fonctions de $l^p(g)$," *Annales de l'institut Fourier*, vol. 20, no. 2, pp. 335–402, 1970.

[9] W. Beckner, "Inequalities in fourier analysis on $\mathbb{R}^n$," *Proceedings of the National Academy of Sciences*, vol. 72, no. 2, pp. 638–641, 1975.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de