Edge-isoperimetric inequalities and influences

Dvir Falik* Alex Samorodnitsky†

Abstract

We give a combinatorial proof of the result of Kahn, Kalai, and Linial [19], which states that every balanced boolean function on the n-dimensional boolean cube has a variable with influence of at least $\Omega\left(\frac{\log n}{n}\right)$.

The methods of the proof are then used to recover additional isoperimetric results for the cube, with improved constants.

We also state some conjectures about optimal constants and discuss their possible implications.

1 Introduction

This paper deals with isoperimetric problems on graphs. Given a graph $G = (V, E)$, the vertex boundary of a subset $S \subseteq V$ contains the vertices of S which have neigbours outside S

$$\mathcal{B}_v(S) = \{x \in S : \exists y \in S^c \text{ such that } (x, y) \in E\}$$

The edge boundary of S is the set of edges crossing from S to its complement.

$$\mathcal{B}_e(S) = \{(x, y) \in E : x \in S \text{ and } y \in S^c\}$$

The question about the smallest possible boundary a set of given cardinality can have is an important combinatorial question with obvious connections to the classical isoperimetry. Good estimates of the minimal boundary size are also very useful in applications. We briefly mention two. Lower bounds on the vertex boundary show how fast a neighbourhood of a set has to grow when the allowed distance from the set increases, and this leads to concentration of measure results for Lipschitz functions on the graph [23]. Lower bounds on the edge boundary suggest that a simple random walk on the graph doesl not remain in any subset for too long, and this leads to upper bounds on its mixing time [15].

Early isoperimetric results on graphs include isoperimetric theorems for the boolean cube $\{0, 1\}^n$. This is a graph with 2^n vertices indexed by boolean strings of length n. Two vertices are connected by an edge if they differ only in one coordinate. The metric defined by this graph is called the Hamming distance. Two vertices x and y are at distance d if they differ in

*School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel.
†School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel.

d coordinates. A Hamming ball is a ball in this metric. A subcube is a subset of the vertices obtained by fixing the value in some of the coordinates. The number of fixed coordinates is called the co-dimension of the subcube. It turns out [13], [12, 14] that the vertex boundary of a Hamming ball is smallest among all sets of equal size, and the same is true for the edge boundary of a subcube. (It is also possible to interpolate these results for all the intermediate subset sizes.)

While many other exact vertex and edge isoperimetric results are known (see [2] for a survey), in most cases exact results seem to be hard to obtain. In many of these cases they could be replaced by sufficiently strong approximate isoperimetric results [16, 24, 29].

Given a solution of an isoperimetric problem, one can ask about its stability. Namely, should a set whose boundary is not much larger than minimal be close to the optimal set? Such results turn out to be especially useful and interesting [8, 7, 17].

In this paper we focus on edge-isoperimetric questions in the boolean cube. A major result in this area was obtained by Kahn, Kalai, and Linial [19] who showed that any balanced boolean function has a variable with large *influence*. We proceed to describe this result, starting with some background.

For a subset $A \subseteq \{0,1\}^n$ and an index $1 \leq i \leq n$ let $I_i(A)$ be the fraction of edges in direction i between A and its complement A^c. This means that $2^{n-1} \cdot I_i(A)$ counts the edges with one vertex in A and another in A^c, the vertices disagreeing in i-th coordinate. $\sum_{i=1}^n I_i(A)$ is the total (normalized) cardinality of the edge boundary of A.

The familiar edge-isoperimetric inequality in the cube states that for any subset $A \subseteq \{0,1\}^n$ of cardinality at most 2^{n-1} holds [1]

$$\sum_{i=1}^n I_i(A) \geq \frac{2}{\log 2} \cdot \frac{|A|}{2^n} \log \frac{2^n}{|A|}.$$

This is tight if A is a subcube of (arbitrary) co-dimension $1 \leq t \leq n$.

Let f be the characteristic function of A, with expectation $\mu = \mathbb{E}_{x \in \{0,1\}^n} f(x) = \frac{|A|}{2^n}$. The edge-isoperimetric inequality asserts that for $\mu \leq 1/2$

$$\sum_{i=1}^n I_i(f) \geq \frac{2}{\log 2} \cdot \mu \log \frac{1}{\mu}. \tag{1}$$

Here $I_i(f)$ stands for the influence of the i-th variable on the support of f.[2]

The inequality (1) has several easy proofs [12, 14]. The one most relevant to this discussion is by induction on dimension. To illustrate its outlay and its simplicity, here it is (a sketch): the base $n = 1$ is easy . Assume for dimension $n - 1$ and consider the case of dimension n. Write $A = A_0 \cup A_1$, where A_i contains all the elements of A with i in the n'th coordinate. Think

[1]We use natural logarithms throughout the paper.

[2]We interchange freely between a set and its characteristic function. Whenever this does not cause confusion we do not mention either, and simply write I_i.

about A_i as subsets of $(n-1)$-dimensional cube, and observe $I_n(A) \geq \left(1/2^{n-1}\right) \cdot \Big||A_0| - |A_1|\Big|$. Taking $a_i = |A_i|$ it remains to check that for any nonnegative a_0, a_1 holds $2 \log 2a_0 \log \frac{2^{n-1}}{a_0} + 2 \log 2a_1 \log \frac{2^{n-1}}{a_1} + |a_0 - a_1| \geq 2 \log 2(a_0 + a_1) \log \frac{2^n}{a_0+a_1}$, which is easily verified, using the properties of the logarithm.

Things become more complicated when we ask for more detailed information. Interpreting the set A as the set of positive outcomes of a game with n players, the number $I_i(A)$ acquires a game-theoretic interpretation as the *influence* of i-th player on the outcome of the game, namely the probability that the outcome of the game remains uncertain if the decisions of other players are chosen at random. Motivated by questions from computational game theory Ben-Or and Linial [3] conjectured that for any balanced game (namely $|A| = 2^{n-1}$) there is a player with influence of at least $\Omega\left(\frac{\log n}{n}\right)$.

This conjecture was proved by Kahn, Kalai, and Linial in [19].

Theorem 1.1: *Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function with expectation $\mathbb{E}f = \mu$. Then*

$$\sum_{i=1}^{n} I_i^2(f) \geq \Omega\left(\frac{\mu^2(1-\mu)^2 \log^2 n}{n}\right) \tag{2}$$

In particular, there is i with $I_i \geq \Omega\left(\frac{\mu(1-\mu)\log n}{n}\right)$.

[19] is one of the first papers to use Fourier analysis on $\mathbb{Z}_2^n$ in a combinatorial setting. Rather surprisingly, a crucial tool in the proof is an inequality [1, 5, 11] which is easiest to describe in Fourier-analytic terms. Let $\{w_S\}$ be the Walsh-Fourier basis of the vector space of real-valued functions on the cube. For a function $f : \{0,1\}^n \to \mathbb{R}$, $f = \sum_{S \in \{0,1\}^n} \hat{f}(S) w_s$, and a nonnegative real ϵ, let $T_\epsilon(f) = \sum_{S \in \{0,1\}^n} \epsilon^{|S|} \hat{f}(S) w_s$. Then

$$\|T_\epsilon f\|_2 \leq \|f\|_{1+\epsilon^2}$$

Following its application in [19], this inequality, known (for historical reasons) as the Bonami-Beckner inequality, became a very important tool in combinatorics and theory of computer science. Still it is very different from the familiar combinatorial tools, and its appearance in the proof is somewhat mysterious. Thus it seemed of interest to look for a combinatorial proof of theorem 1.1, possibly along the lines of the forementioned proof of (1). Let us mention two papers dealing with this problem along very different routes. The first of these papers [10] gives a combinatorial (entropic) proof of the Bonami-Beckner inequality for $\epsilon = \sqrt{3}/3$. This special case is already strong enough to be instrumental in the proof of theorem 1.1. The second paper [28] presents an inductive proof that the maximal influence of a balanced function is at least $\Omega\left(\frac{\log^\alpha(n)}{n}\right)$ for some $0 < \alpha < 1$.

In this paper we give a fully combinatorial proof of theorem 1.1. After completing our work, we learned that a very similar proof was recently obtained by [27].

We start with a functional form of inequality (1). For a nonnegative function $f : \{0,1\}^n \to \mathbb{R}$

$$\mathbb{E}_x \sum_{y \sim x} (f(x) - f(y))^2 \geq 2 \cdot \mathbb{E}f^2 \log \frac{\mathbb{E}f^2}{\mathbb{E}^2 f} \tag{3}$$

Functional forms of isoperimetric inequalities are widely used in local theory of Banach spaces [20]. They turn out to be useful in our setting too. We show theorem 1.1 to be a simple consequence of inequality (3).

This inequality can be proved by induction on dimension (see Appendix A), similarly to (1), though the proof is somewhat more complicated. The isoperimetric constant $C_3 = 2$ is tight, if we want it to be independent of the dimension. In section 4 we give examples of functions satisfying (3) with equality if the constant 2 is replaced by $2 + o_n(1)$. These are symmetric functions (a function f on the cube is symmetric if $f(x)$ depends only on the distance of x from zero) closely related to a classical family of orthogonal polynomials of discrete variable - the Krawchouk polynomials.

We also suggest a reason behind the relevance of the Bonami-Beckner inequality. It is well-known that this inequality is equivalent to the logarithmic Sobolev inequality

$$\mathbb{E}_x \sum_{y \sim x} (f(x) - f(y))^2 \geq 2 \cdot Ent\left(f^2\right) = 2 \cdot \left(\mathbb{E}f^2 \log f^2 - \mathbb{E}f^2 \log \mathbb{E}f^2\right) \tag{4}$$

in the cube. It turns out that this inequality implies inequality (3). In this sense Bonami-Beckner's inequality can be thought of as a refined form of the edge-isoperimetric inequality in the discrete cube.

The actual result we prove seems to be somewhat stronger than theorem 1.1. We show that for a boolean function with expectation μ holds

Theorem 1.2:

$$\sum_{i=1}^{n} I_i^2(f) \geq 4\mu(1-\mu) \exp\left\{ -\frac{1/2}{\mu(1-\mu)} \sum_{i=1}^{n} I_i(f) \right\} \tag{5}$$

This inequality implies (2) with a constant $c_2 = 4$, improving the estimate $c_2 = \frac{4}{5 \log 2} \approx 1.2$ of [19].

Theorem 1.2 is a special case of our main result, theorem 2.2, stated and proved in section 2. This theorem presents a more general inequality valid for real-valued functions on the discrete cube endowed with an arbitrary measure. The theorem and the approach used in its proof seem to provide convenient tools for dealing with a certain type of isoperimetric statements in the cube. We illustrate this by giving simple proofs of two results from [7] and [9], with better isoperimetric constants.

It should be mentioned that inequality (5), in its turn, is implied, *up to a constant in the exponent*, by an inequality of Talagrand [30]. A special case of this inequality asserts that for a boolean function f with expectation μ

$$\sum_{i=1}^{n} \frac{I_i}{\log\left(e/I_i\right)} \geq \Omega\left(\mu(1-\mu)\right) \tag{6}$$

It is not hard to see that this gives (5) if $1/2$ in the exponent is replaced by a sufficiently large constant.

Next, we focus our attention on the best possible constants for the above-mentioned inequalities. Specifically, we are interested in the exact constant C_5 that should appear in the exponent in the right hand side of (5). We point out an interesting phenomenon in that obtaining the (conjectured) optimal constant for this inequality wouldt lead to a stability result for the basic inequality (1).

To be more specific, by theorem 1.2 $C_5 \le \frac{1}{2}$. On the other hand, taking f to be a characteristic function of a subcube of large co-dimension, shows $C_5 \ge \frac{\log 2}{2}$. We believe the lower bound to be the right one.

Conjecture 1.3:

$$C_5 = \frac{\log 2}{2}$$

∎

In particular, we conjecture small subcubes to be (nearly)-isoperimetric sets for this inequality.

If conjecture 1.3 holds, it would, in particular, improve the estimate on the optimal constant C_2 in inequality (2). It is easy to see that $C_2 \ge \frac{1}{C_5^2}$.[3] Therefore the conjecture would imply $C_2 \ge \frac{4}{\log^2 2} \approx 8.3$. We remark that the best known candidate to be an isoperimetric function for inequality (2), the "tribes" function of Ben-Or and Linial [3] shows $C_2 \le 16$.

Now, consider functions which are nearly isoperimetric in the sense of the basic inequality (1). Kahn and Kalai conjecture ([17]) that such functions behave similarly to subcubes, in the following precise sense.

Conjecture 1.4: [17] Let $K > 0$ be a real number. There are positive real numbers K', δ depending on K such that the following assertion holds: If a monotone boolean function $f: \{0,1\}^n \to \{0,1\}$ with expectation $\mu \le 1/2$ satisfies

$$\sum_{i=1}^{n} I_i \le K \cdot \frac{2}{\log 2} \cdot \mu \log \frac{1}{\mu} \tag{7}$$

then there is a set of at most $K' \cdot \log \frac{1}{\mu}$ coordinates such that the expectation of f restricted to the subcube obtained by setting all the coordinates in this set to 1 is at least $(1+\delta) \cdot \mu$. ∎

It seems that a weaker version of this conjecture, which claims the same conclusion from a stronger assumption that the multiplicative factor K in (7) is close to 1, i.e., $K = 1 + \epsilon$ for a small $\epsilon > 0$, is also interesting [18].

Here we prove an even weaker result in this direction, conditioned on conjecture 1.3. Let $o_{\mu,\epsilon}(1)$ denote a quantity which goes to zero when *both* μ and ϵ do.

[3] More precisely, $C_2 \ge \frac{1}{C_5^2} - o_n(1)$. Here and in the rest of this paper we ignore negligible factors when comparing constants.

Proposition 1.5: *Assume conjecture 1.3. Let $f: \{0,1\}^n \rightarrow \{0,1\}$ be a monotone boolean function with expectation $\mu \le 1/2$, and assume*

$$\sum_{i=1}^{n} I_i \le (1+\epsilon) \cdot \frac{2}{\log 2} \cdot \mu \log \frac{1}{\mu}$$

Then there is a set of $O\left(\left(\frac{1}{\mu}\right)^{(1+o_{\mu,\epsilon}(1))\cdot\epsilon}\right)$ coordinates such that the expectation of f restricted to the subcube obtained by setting all the coordinates in this set to 1 is at least 2μ.

We conclude this section by saying a few words about a possible approach to the proof of conjecture 1.3. We will say more about this in section 3. The main step in the proof of theorem 1.2 is a variant of the logarithmic Sobolev inequality for the discrete cube. This inequality applies to general real-valued functions on the cube, and is tight with constant $c = 2$. To prove the conjecture we need to take into account the specific structure of boolean functions. The familiar approach using tensorization does not seem to be convenient for this. We give a proof of the inequality for general functions which works by induction on the dimension, similar to the proof of (1), and seems to be more conducive for this purpose.

The paper is organized as follows: in the next section we prove the main theorem 2.2. In section 3 several corollaries are derived from theorem 2.2, and the main technical conjecture is stated. Section 4 constructs nonnegative real-valued functions which are almost isoperimetric for inequality (8) and hence for several other inequalities in this paper, including (3). Inductive proofs of (3) and a logarithmic Sobolev inequality for the cube are given in the Appendices.

2 The main theorem

We start with some definitions and notation.

Let $\mathcal{F}_j$, for $0 \le j \le n$, be the algebra of subsets of $\{0,1\}^n$ generated by the first j bits. More precisely, $\mathcal{F}_j$ is generated by the atoms $\left\{A_{\epsilon_1 \ldots \epsilon_j} : \epsilon_i \in \{0,1\}\right\}$ where $A_{\epsilon_1 \ldots \epsilon_j} = \{x : x_1 = \epsilon_1 \ldots x_j = \epsilon_j\}$. Then $\{\mathcal{F}_j\}$ is an increasing sequence of algebras. In particular $\mathcal{F}_0 = \{\emptyset, \{0,1\}^n\}$ and $\mathcal{F}_n = 2^{\{0,1\}^n}$.

For a function $f: \{0,1\}^n \rightarrow \mathbb{R}$, let $f_i = \mathbb{E}(f|\mathcal{F}_i)$, the conditional expectation of f given the algebra $\mathcal{F}_i$. This means that $f_i(x)$ is the average of f over the points y that coincide with x in the first i coordinates. In particular, $f_0 = \mathbb{E}f$, $f_n = f$. The sequence $f_0, ..., f_n$ is a martingale with respect to $\{\mathcal{F}_j\}$. [4] Let d_i, $i = 1...n$ be the sequence of martingale differences. $d_i = f_i - f_{i-1}$, $1 \le i \le n$.

Let $\mathcal{E}(f,f) = \mathbb{E}_x \sum_{y \sim x} (f(x) - f(y))^2$. Let us mention that $\mathcal{E}(f,g) = E_x \sum_{y \sim x} (f(x) - f(y))(g(x) - g(y))$ is sometimes called the canonical Dirichlet form on $\{0,1\}^n$ [4].

The following lemma is simple and well-known [25]. For completeness, we will give a proof at the end of this section.

[4]Essentially the only martingale property we use is the fact that conditional expectation is an orthogonal projection on a subspace.

Lemma 2.1:

$$\mathcal{E}(f,f) = \sum_{i=1}^{n} \mathcal{E}\left(d_i, d_i\right)$$

Let μ be a measure on $\{0,1\}^n$. Let C be the best constant in the logarithmic Sobolev inequality for $\{0,1\}^n$ with μ. This is to say that C is maximal such that for any function $f:\ \{0,1\}^n \to \mathbb{R}$ holds $\mathcal{E}(f,f) \geq C \cdot Ent\left(f^2\right)$.

For a function $f:\ \{0,1\}^n \to \mathbb{R}$, let $\sigma^2(f) = \mathbb{E}f^2 - \mathbb{E}^2 f$.

Our main result is:

Theorem 2.2:

$$\sum_{i=1}^{n} \mathbb{E}^2|d_i| \geq \sigma^2(f) \exp\left\{-\frac{\mathcal{E}(f,f)}{C\sigma^2(f)}\right\} \tag{8}$$

Proof: First a simple lemma.

Lemma 2.3: *For a nonnegative function f holds $Ent(f^2) \geq \mathbb{E}f^2 \log \frac{\mathbb{E}f^2}{\mathbb{E}^2 f}$.*

Proof: (Of the lemma) Since both sides of the inequality are 2-homogeneous, this amounts to showing $\mathbb{E}f^2 \log f^2 + \log \mathbb{E}^2(f) \geq 0$, given $\mathbb{E}f^2 = 1$. This is the same as $\log \mathbb{E}f - \mathbb{E}f^2 \log \frac{1}{f} \geq 0$. And this is true since logarithm is concave. ∎

Now we can conclude the proof of theorem 2.2. Using (4) and lemma 2.3,

$$\mathcal{E}(f,f) = \sum_{i=1}^{n} \mathcal{E}\left(d_i, d_i\right) \geq C \cdot \sum_{i=1}^{n} Ent\left(d_i^2\right) \geq C \cdot \sum_{i=1}^{n} \mathbb{E}d_i^2 \log \frac{\mathbb{E}d_i^2}{\mathbb{E}^2|d_i|}.$$

Observe $\sum_{i=1}^{n} \mathbb{E}d_i^2 = \mathbb{E}f^2 - \mathbb{E}^2 f =: \sigma^2(f)$. By the convexity of the minus logarithm, the last sum is

$$C\sigma^2(f) \sum_{i=1}^{n} \frac{\mathbb{E}d_i^2}{\sigma^2(f)} \log \frac{\mathbb{E}d_i^2}{\mathbb{E}^2|d_i|} \geq C\sigma^2(f) \log\left(\frac{\sigma^2(f)}{\sum_{i=1}^{n} \mathbb{E}^2|d_i|}\right).$$

∎

We remark that instead of the logarithmic Sobolev inequality (4) it is possible to use isoperimetric inequality (3) directly. Hence the logarithmic Sobolev constant C in the statement can be replaced by potentially bigger isoperimetric constant C'.

Proof of lemma 2.1

For a function g and an index $1 \leq i \leq n$ let g^i be a function defined by $g^i(x) = g(x \oplus e_i)$. Here e_i is the vector with 1 in i'th coordinate, and zero in the other coordinates. Note that g^i is $\mathcal{F}_k$-measurable iff g is.

$$\mathcal{E}\left(d_j, d_j\right) = \sum_{i=1}^{n} \|d_j - d_j^i\|_2^2 = \sum_{i=1}^{n} \left\langle d_j - d_j^i, d_j - d_j^i \right\rangle = \sum_{i=1}^{n} \left\langle d_j - d_j^i, (f_j - f_{j-1}) - \left(f_j^i - f_{j-1}^i\right)\right\rangle =$$

$$\sum_{i=1}^{n}\left\langle d_j-d_j^i,f_j-f_j^i\right\rangle=\sum_{i=1}^{n}\left\langle\left(f_j-f_j^i\right)-\left(f_{j-1}-f_{j-1}^i\right),f_j-f_j^i\right\rangle=$$

$$\sum_{i=1}^{n}\left\langle f_j-f_j^i,f_j-f_j^i\right\rangle-\sum_{i=1}^{n}\left\langle f_j-f_j^i,f_{j-1}-f_{j-1}^i\right\rangle=$$

$$\sum_{i=1}^{n}\left\langle f_j-f_j^i,f_j-f_j^i\right\rangle-\sum_{i=1}^{n}\left\langle f_{j-1}-f_{j-1}^i,f_{j-1}-f_{j-1}^i\right\rangle=\mathcal{E}\left(f_j,f_j\right)-\mathcal{E}\left(f_{j-1},f_{j-1}\right).$$

The proof is concluded by observing $\mathcal{E}\left(f_0,f_0\right)=0$. ∎

3 Some corollaries for product measures

In this section we derive theorem 1.1, proposition 1.5, and the theorems of Friedgut and Friedgut-Kalai from theorem 2.2. We also state our main technical conjecture (9).

We will assume the measure μ to be a product probability measure, that is $\mu=\otimes_{k=1}^{n}\mu_k$, with $\mu_k(1)=p_k$, and $\mu_k(0)=1-p_k$. In this case $\mathbb{E}|d_i|$ has a simple upper bound.

Lemma 3.1: *For a product measure μ,*

$$\mathbb{E}|d_i|\leq 2p_i(1-p_i)\cdot\mathbb{E}_x\Big|f(x)-f(x\oplus e_i)\Big|$$

Proof:

$$\mathbb{E}|d_i|=\mathbb{E}|f_i-f_{i-1}|=\mathbb{E}\Big|\mathbb{E}\left(f|\mathcal{F}_i\right)-\mathbb{E}\left(f|\mathcal{F}_{i-1}\right)\Big|.$$

Let $\mathcal{G}_i$ be the algebra of subsets of $\{0,1\}^n$ generated by all the bits but j. That is for x with $x_i=0$ holds $\mathbb{E}\left(f|\mathcal{G}_i\right)(x)=\mathbb{E}\left(f|\mathcal{G}_i\right)(x\oplus e_i)=(1-p_i)f(x)+p_if(x\oplus e_i)$.

Then for a product measure μ holds $\mathbb{E}\left(\mathbb{E}\left(f|\mathcal{F}_i\right)\Big|\mathcal{G}_i\right)=\mathbb{E}\left(f|\mathcal{F}_{i-1}\right)$. Therefore

$$\mathbb{E}\Big|\mathbb{E}\left(f|\mathcal{F}_i\right)-\mathbb{E}\left(f|\mathcal{F}_{i-1}\right)\Big|=\mathbb{E}\Big|\mathbb{E}\left(f-\mathbb{E}\left(f|\mathcal{G}_i\right)\Big|\mathcal{F}_i\right)\Big|\leq\mathbb{E}\Big|f-\mathbb{E}\left(f|\mathcal{G}_i\right)\Big|=2p_i(1-p_i)\cdot\mathbb{E}_x\Big|f(x)-f\left(x\oplus e_i\right)\Big|$$

We have used the well-known fact that conditional expectation decreases the ℓ_1-norm. ∎

3.1 Uniform measure

The best constant C for a cube endowed with the uniform measure is $C=2$. Hence the theorem gives, for a real-valued function f on the discrete cube, $\sum_{i=1}^{n}\mathbb{E}^2|d_i|\geq\sigma^2(f)\exp\left\{-\frac{\mathcal{E}(f,f)}{2\sigma^2(f)}\right\}$ This may be somewhat simplified for monotone functions, for which $\mathbb{E}|d_i|=\hat{f}(\{i\})$

$$\sum_{i=1}^{n}\hat{f}^2(\{i\})\geq\sigma^2\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}I_i(f)\right\}.$$

Section 4 presents a construction of monotone functions for which this inequality is essentially tight.

Our main concern are boolean functions. We present several easy implications of theorem 2.2 and a related conjecture.

Proof of theorem 1.2.

For a boolean function f we have $\mathbb{E}_x|f(x) - f(x \oplus e_i)| = I_i$ and using the theorem together with lemma 3.1 yields $\sum_{i=1}^n I_i^2(f) \geq 4\sigma^2(f)\exp\left\{-\frac{1}{2\sigma^2(f)}\sum_{i=1}^n I_i(f)\right\}$ proving (5) and theorem 1.2. ∎

Example 3.2: Let f be the characteristic function of a subcube of dimension $n-t$. Then f has t non-zero influences of size 2^{-t+1}. Assume t is large enough so that $\sigma^2 = \frac{2^t-1}{4^t}$ may be replaced with 2^{-t}. Then (5) gives

$$\frac{4t}{4^t} \geq \frac{1}{2^t}\exp\left\{-2^{t-1}\cdot\frac{2t}{2^t}\right\} = \frac{1}{2^t}\exp\{-t\} \qquad or \qquad 4t2^{-t} \geq e^{-t}.$$

∎

We conjecture (conjecture 1.3) that for boolean functions a stronger inequality $\sum_{i=1}^n I_i^2(f) \geq 4\sigma^2(f)\exp\left\{-\frac{\log 2}{2\sigma^2(f)}\sum_{i=1}^n I_i(f)\right\}$ holds. Such an inequality would be tight by the example above.

This inequality would follow from the following version of the logarithmic Sobolev inequality for boolean functions.

Conjecture 3.3: For a boolean function f on the discrete cube

$$\mathcal{E}(f,f) \geq \frac{2}{\log 2}\cdot\sum_{i=1}^n Ent\left(d_i^2\right) \tag{9}$$

∎

Discussion. The inequality holds with a constant $c = 2$ for real-valued functions on $\{0,1\}^n$. This is proved by applying the logarithmic Sobolev inequality to functions d_i, $1 \leq i \leq n$. This is also tight, as shown by functions constructed in section 4. To improve the constant for boolean functions a different approach seems to be required, one that "remembers" that $\{d_i\}$ are difference functions of a martingale defined by a boolean function. In particular the familiar tensorization approach might not be sufficient here since it does not keep track of the combinatorial structure of the functions involved.

We give an inductive proof of the inequality $\mathcal{E}(f,f) \geq 2\cdot\sum_{i=1}^n Ent\left(d_i^2\right)$ in Appendix B. This proof seems to be better suited for handling functions f with a specific structure, such as boolean functions.

Proof of theorem 1.1

We show

$$\sum_{i=1}^{n} I_i^2(f) \geq \left(\frac{4}{\log 2} - o_n(1)\right) \cdot \frac{\mu^2(1-\mu)^2 \log^2 n}{n}$$

For a boolean function f with expectation μ holds $\sigma^2(f) = \mu(1-\mu)$. Let $\epsilon(n) \gg \frac{\log\log n}{\log n}$. There are two cases to consider. First, $\sum_{i=1}^{n} I_i < (2-\epsilon(n)) \cdot \mu(1-\mu)\frac{\log n}{n}$. In this case (5) implies $\sum_{i=1}^{n} I_i^2(f) \gg \mu(1-\mu)\frac{\log^2 n}{n}$. The second case is $\sum_{i=1}^{n} I_i \geq (2-\epsilon(n)) \cdot \mu(1-\mu)\frac{\log n}{n}$. The Cauchy-Schwarz inequality now implies $\sum_{i=1}^{n} I_i^2 \geq (2-\epsilon(n))^2 \cdot \mu^2(1-\mu)^2\frac{\log^2 n}{n}$. ∎

Proof of a theorem of Friedgut

Theorem 3.4: *[7] For a boolean function f and an arbitrary $\epsilon > 0$ there is a function g depending only on $(\sum_{i=1}^{n} I_i) \cdot \exp\left\{\frac{\sum_{i=1}^{n} I_i}{(2-o_\epsilon(1))\epsilon}\right\}$ coordinates*[5] *(a junta) such that $\|f-g\|_2^2 \leq \epsilon$.*

Proof: Let $K = \sum_{i=1}^{n} I_i$ and take $\alpha = \exp\left\{-\frac{K}{(2-o_\epsilon(1))\epsilon}\right\}$. The error term $o_\epsilon(1)$ will be chosen later.

Without loss of generality assume the influences I_i to decrease with i, and let r be the maximal index with $I_r \geq \alpha$. Clearly $r \leq \frac{K}{\alpha}$ and $\sum_{i=r+1}^{n} I_i^2 \leq K\alpha$. Take $g = \mathbb{E}\left(f|\mathcal{F}_r\right)$. The function g depends only on r variables. We will show $\|f-g\|_2^2 \leq \epsilon$.

Let $h = f - g$. Take $h_i = \mathbb{E}\left(h|\mathcal{F}_i\right)$, $i = 1...n$, to be the martingale defined by h, and let $d_i(h)$ be its difference functions. Then $d_i(h) = d_i(f)$ for $i > r$ and $d_i(h) = 0$ otherwise. Note that $\mathbb{E}h = 0$ and therefore $\sigma^2(h) = \|h\|_2^2$. By theorem 2.2

$$K = \sum_{i=1}^{n} I_i = \mathcal{E}(f,f) \geq \mathcal{E}(h,h) \geq 2\|h\|_2^2 \log\left(\frac{\|h\|_2^2}{\sum_{i=r+1}^{n} I_i^2}\right) \geq 2\|h\|_2^2 \log\left(\frac{\|h\|_2^2}{K\alpha}\right)$$

Recalling the definition of α, it is now easy to choose the error term $o_\epsilon(1)$ appropriately, so that the last inequality implies $\|h\|_2^2 \leq \epsilon$. ∎

Proof of proposition 1.5

We proceed similarly to the preceding proof. Let $K = \sum_{i=1}^{n} I_i$ and let $\alpha = \mu^{1+(1+o_{\mu,\epsilon}(1))\cdot\epsilon}$, where $o_{\mu,\epsilon}(1)$ is an error term which goes to zero when both μ and ϵ do. We will set it later.

Assume the influences to decrease, and define the index r and and functions g and h as above. Now we need a simple lemma

Lemma 3.5: *If $\|h\|_2^2 \leq \mu - 2\mu^2$ then*

$$Pr\Big\{f(x) = 1 \;\Big|\; x_1 = ... = x_r = 1\Big\} \geq 2\mu$$

[5]The original proof in [7] has a somewhat larger estimate for the required number of variables in the junta. This estimate has a constant 2 instead of 1/2 in the exponent.

Proof: For $y = (y_1 ... y_r) \in \{0,1\}^r$ let K_y be the subcube $\{x \in \{0,1\}^n : \ x_1 = y_1, \ ... \ x_r = y_r\}$. Let f_y be the function f restricted to K_y. Let $\mu_y = \mathbb{E} f_y$. Alternatively μ_y is the value of $g = \mathbb{E}(f|\mathcal{F}_r)$ on K_y. Let $\mathbf{1} \in \{0,1\}^r$ be the vector of all ones. $\mu_{\mathbf{1}}$ is the quantity we want to lower bound. Since f is a monotone function, so is g. In particular $\mu_{\mathbf{1}}$ is the largest among all μ_y.

We have $\mathbb{E}_y \mu_y = \mathbb{E} g = \mathbb{E} f = \mu$. On the other hand,

$$\mu - 2\mu^2 \geq \|h\|_2^2 = \|f - g\|_2^2 = \mathbb{E}_y \mu_y (1 - \mu_y)$$

Therefore $2\mu^2 \leq \mathbb{E}_y \mu_y^2 \leq \mu_{\mathbf{1}} \cdot \mathbb{E}_y \mu_y = \mu_{\mathbf{1}} \cdot \mu$. ∎

Now, by conjecture 1.3

$$(1+\epsilon)\frac{2}{\log 2}\mu \log \frac{1}{\mu} \geq \sum_{i=1}^{n} I_i = \mathcal{E}(f,f) \geq \mathcal{E}(h,h) \geq$$

$$\frac{2}{\log 2}\|h\|_2^2 \log\left(\frac{\|h\|_2^2}{\sum_{i=r+1}^{n} I_i^2}\right) \geq \frac{2}{\log 2}\|h\|_2^2 \log\left(\frac{\|h\|_2^2}{K\alpha}\right)$$

Recalling the definition of α, it is now easy to choose the error term $o_{\mu,\epsilon}(1)$ appropriately, so that the last inequality implies $\|h\|_2^2 \leq \mu - 2\mu^2$. ∎

3.2 The measure μ_p

Let μ_p be a product distribution, $\mu_p = \otimes_{k=1}^{n} \mu$, with $\mu(1) = p$, $\mu(0) = 1 - p$. Assume $p \leq \frac{1}{2}$. The best constant C in the logarithmic Sobolev inequality in this case is known [6] to be $C(p) = \frac{1-2p}{p(1-p)} \cdot \frac{1}{\log(1-p) - \log p}$.

Proof of a theorem of Friedgut and Kalai

Theorem 3.6: *[9] Let f be a boolean function with expectation μ on $\{0,1\}^n$ endowed with the measure μ_p. Assume that $1 \gg p \geq n^{-o_n(1)}$. Then there is a variable with influence at least $\Omega\left(\frac{\mu(1-\mu)}{p \log \frac{1}{p}} \cdot \frac{\log n}{n}\right)$ on f.*

Proof: For a boolean function f theorem 2.2 gives

$$\sum_{i=1}^{n} I_i^2(f) \geq \frac{\sigma^2(f)}{4p^2(1-p)^2} \exp\left\{-\frac{1}{C(p)\sigma^2(f)} \sum_{i=1}^{n} I_i(f)\right\}$$

The expression on the right hand side is somewhat complicated. It simplifies for $p \ll 1$, for which $C(p) \approx \frac{1}{p \log \frac{1}{p}}$, and we get (ignoring negligible errors)

$$\sum_{i=1}^{n} I_i^2(f) \geq \frac{\sigma^2(f)}{4p^2} \exp\left\{-\frac{p \log \frac{1}{p}}{\sigma^2(f)} \sum_{i=1}^{n} I_i(f)\right\}$$

Proceeding similarly to the proof of theorem 1.1, we get that

$$\sum_{i=1}^{n} I_i^2(f) \geq \frac{\mu^2(1-\mu)^2}{p^2 \log^2 \frac{1}{p}} \cdot \frac{\log^2 n}{n} \tag{10}$$

In particular, there is a variable i with influence at least

$$I_i \geq \frac{\mu(1-\mu)}{p \log \frac{1}{p}} \cdot \frac{\log n}{n} \tag{11}$$

∎

We remark that this proof provides (11) with an explicit constant 1, and does not rely on the assumption $p \geq n^{-o(1)}$.

4 Construction of 'isoperimetric' functions

In this section we construct nonnegative functions f_s on the cube endowed with the uniform measure, for which inequality (8) is almost tight. This directly implies that these functions are 'isoperimetric' for inequalities (3), (5), and inequality (9) with constant $c = 2$.

The functions f_s were constructed in [26] (for a different purpose). Here we repeat parts of this construction for completeness.

Let s be an integer, $\sqrt{n} \ll s \ll n$. We first construct an auxiliary function k_s. This function will be *symmetric*, namely its value at a point will depend only on the distance of the point from zero. Such a function, of course, is fully defined by its values $k_s(0), ..., k_s(n)$ at distances $0...n$. Set $k_s(-1) = 0$ and $k_s(0) = 1$, and define $k_s(r)$ for $1 \leq r \leq n$ so that the relation $(n-2s)k_s(r) = (r-1)k_s) + (n-r)k_s(r+1)$ is satisfied for $r = 0...n-1$. The univariate function $k_s(r)$ we have defined on the integer points $r = 0...n$ coincides with a normalized Krawchouk polynomial K_s (see [21] for detailed information on Krawchouk polynomials).

Krawchouk polynomials $\{K_s\}_{s=0}^{n}$ are a family of polynomials orthogonal with respect to a measure supported on $0...n$. Hence their roots are simple and are located in the interval $(0, n)$ [31]. Let x_s be the first root of K_s. We now define $f = f_s$ to be a symmetric function on $\{0,1\}^n$ defined by $f(x) = k_s(x)$ for points whose distance from zero is at most x_s, and $f_s(x) = 0$ otherwise.

We require an asymptotic estimate $x_s = \frac{n}{2} - \sqrt{sn} + o(\sqrt{sn})$ [21]. This means, in particular, that the support of f is small (of cardinality e^{-s}), and therefore $\frac{\mathbb{E}^2 f}{\mathbb{E} f^2} \leq \frac{|supp(f)|}{2^n}$ is small, so that $\sigma^2(f)$ can be, for all practical reasons, replaced with $\mathbb{E} f^2$.

For $x \in \{0,1\}^n$, let $N(x) = \sum_{y:\ y \sim x} f(y)$. Then it is not hard to see ([26], lemma 3.4) that $N(x) \geq (n-2s)f(x)$ and therefore $\mathcal{E}(f,f) = \mathbb{E}_x \sum_{y \sim x} (f(x) - f(y))^2 = 2n\mathbb{E} f^2 - 2\langle f, N \rangle \leq 4s\mathbb{E} f^2$.

Hence the right hand side in (8) can be estimated from below by $\mathbb{E} f^2 \cdot e^{-2sn}$.

Now to the left hand side. The function f is symmetric and (easy to see, cf. also [21]) monotone. Therefore $\mathbb{E}|d_i| = I_i = \hat{f}(\{i\})$ have the same value for all indices $1 \le i \le n$. Let I denote this common value. We want to upper bound I. Let $f(x)$ denote the value of f in points at distance x from zero. Let $m = \lfloor x_s \rfloor$. Then

$$nI = \sum_{i=1}^{n} \hat{f}(\{i\}) = \frac{1}{2^{2n}} \sum_{x=0}^{n} \binom{n}{x} (n-2x) f(x) = \frac{1}{2^{2n}} \sum_{x=0}^{m} \binom{n}{x} (n-2x) f(x) \le \frac{n}{2^{2n}} \sum_{x=0}^{m} \binom{n}{x} f(x)$$

Since $\mathbb{E}f^2 = \frac{1}{2^n} \sum_{x=0}^{n} \binom{n}{x} f^2(x)$, we have $f(x) \le \sqrt{\frac{2^n \mathbb{E} f^2}{\binom{n}{x}}}$, and so

$$\frac{1}{2^n} \sum_{x=0}^{m} \binom{n}{x} f(x) \le \sqrt{2^n \mathbb{E} f^2} \sum_{x=0}^{m} \sqrt{\binom{n}{x}} \le n \sqrt{\binom{n}{m} 2^n \mathbb{E} f^2}$$

Therefore $I \le n \sqrt{\frac{\binom{n}{m} \mathbb{E} f^2}{2^n}}$, and the left hand side of (8) is $nI^2 \le n^2 \frac{\binom{n}{m}}{2^n} \mathbb{E} f^2$. Now observe [22] that $\binom{n}{m} \le 2^{nH(m/n)} \le 2^n e^{-(1-o_n(1))2sn}$.

Hence the left hand side in (8) can be estimated from below by $\mathbb{E} f^2 \cdot n^2 e^{-(1-o_n(1))2sn}$. The estimates for both sides are sufficiently close to show the constant $1/2$ in the exponent on the right hand side of (8) to be best possible.

5 Appendix A - an inductive proof of an isoperimetric inequality

In this section we give an inductive proof of the inequality (3)

$$\mathcal{E}(f,f) \ge 2 \cdot \mathbb{E} f^2 \log \frac{\mathbb{E} f^2}{\mathbb{E}^2 f}$$

for a real nonnegative function $f : \{0,1\}^n \to \mathbb{R}$. By homogeneity, we may and will assume $\mathbb{E} f = 1$.

The proof is by induction on the dimension n.

For $n = 1$, let $f(0) = a$, and $f(1) = 2 - a$. Then $\mathbb{E} f^2 = \frac{1}{2}\left(a^2 + (2-a)^2\right) = 1 + (1-a)^2$; and $\mathbb{E}_x \sum_{y \sim x} (f(x) - f(y))^2 = (2-2a)^2 = 4(1-a)^2$. Let $x = 1 - a$. It remains to verify that $2x^2 \ge (1+x)\log(1+x)$ for $-1 \le x \le 1$, which is easily seen to be true. In fact a stronger inequality $2x^2 \ge (1+x)\log_2(1+x)$ is also valid in this interval, since the right hand side is a convex function which is 0 at zero and 2 at one.

Assume the inequality to hold for $n-1$. Let f_0 and f_1 be the restrictions of f to $(n-1)$-dimensional half-cubes determined by value of the n-th coordinate. Let μ_i be the expectations of f_i, and v_i the second moments of f_i for $i = 0, 1$.

Then

$$\mathcal{E}(f,f) = \frac{1}{2} \cdot \left(\mathcal{E}(f_0, f_0) + \mathcal{E}(f_1, f_1)\right) + \|f_0 - f_1\|_2^2. \tag{12}$$

Note that the expectations and the distance in this formula are computed on $(n-1)$-dimensional cubes.

By the induction hypothesis we can lower bound the first summand by

$$v_0 \log \frac{v_0}{\mu_0} + v_1 \log \frac{v_1}{\mu_1}.$$

For the second summand we need a simple lemma.

Lemma 5.1: *Let f_0 and f_1 be two functions with expectations μ_0, μ_1 and variances $\sigma_0^2 = v_0 - \mu_0^2$, $\sigma_1^2 = v_1 - \mu_1^2$. Then*

$$\|f_0 - f_1\|^2 \geq (\sigma_0 - \sigma_1)^2 + (\mu_0 - \mu_1)^2.$$

Proof: Let $g_i = f_i - \mu_i$, $i = 0, 1$. Then $\mathbb{E} g_i = 0$ and therefore

$$\|f_0 - f_1\|^2 = \langle g_0 - g_1, g_0 - g_1 \rangle + (\mu_0 - \mu_1)^2 = \|g_0 - g_1\|^2 + (\mu_0 - \mu_1)^2 \geq (\sigma_0 - \sigma_1)^2 + (\mu_0 - \mu_1)^2.$$

▮

Going back, and substituting in (12),

$$\mathcal{E}(f, f) \geq v_0 \log \left(\frac{v_0}{\mu_0^2} \right) + v_1 \log \left(\frac{v_1}{\mu_1^2} \right) + \left[v_0 + v_1 - 2\sqrt{v_0 - \mu_0^2}\sqrt{v_1 - \mu_1^2} - 2\mu_0\mu_1 \right].$$

So it suffices to show that under the assumptions

1. $\mu_0, \mu_1, v_0, v_1 \geq 0$,
2. $\frac{\mu_0 + \mu_1}{2} = 1$, and
3. $\frac{v_0 + v_1}{2} = v := \mathbb{E} f^2$

holds

$$v_0 \log \left(\frac{v_0}{\mu_0^2} \right) + v_1 \log \left(\frac{v_1}{\mu_1^2} \right) + \left[v_0 + v_1 - 2\sqrt{v_0 - \mu_0^2}\sqrt{v_1 - \mu_1^2} - 2\mu_0\mu_1 \right] \geq 2v \log v. \quad (13)$$

The next few steps swap variables to simplify this expression.

Take $t = \frac{\mu_0 - \mu_1}{2}$. Then $\mu_0 = 1 + t$ and $\mu_1 = 1 - t$. Similarly take $v_0 = v(1 + y)$ and $v_1 = v(1 - y)$. Note that $-1 \leq t, y \leq 1$.

Substituting in (2.2), and dividing out by $2v$ it needs to be seen that

$$\frac{1+y}{2} \log \left(\frac{1+y}{(1+t)^2} \cdot v \right) + \frac{1-y}{2} \log \left(\frac{1-y}{(1-t)^2} \cdot v \right) + 1 \geq$$

$$\log v + \frac{1}{v} \cdot \left[\sqrt{v(1+y) - (1+t)^2}\sqrt{v(1-y) - (1-t)^2} + (1 - t^2) \right],$$

or

$$\frac{1+y}{2}\log\left(\frac{1+y}{(1+t)^2}\right)+\frac{1-y}{2}\log\left(\frac{1-y}{(1-t)^2}\right)+1\geq$$
$$\frac{1}{v}\cdot\left[\sqrt{v(1+y)-(1+t)^2}\sqrt{v(1-y)-(1-t)^2}+(1-t^2)\right].$$

We first take on the right hand side and show it to be at most $\sqrt{1-y^2}$. Indeed, it suffices to show

$$\sqrt{v^2\left(1-y^2\right)-v(1+y)(1-t)^2-v(1-y)(1+t)^2+\left(1-t^2\right)^2}\leq v\sqrt{1-y^2}+\left(1-t^2\right).$$

Squaring both expressions, and rearranging, we get to

$$(1+y)(1-t)^2+(1-y)(1+t)^2\geq 2\sqrt{1-y^2}\left(1-t^2\right).$$

This inequality is a special case of the Arithmetic-Geometric inequality.

Now to the left hand side. Let $H(x)=-x\log x-(1-x)\log(1-x)$ be the (natural) entropy function. For $0\leq p,q\leq 1$ let $D(p||q)=p\log\frac{p}{q}+(1-p)\log\frac{1-p}{1-q}$ denote the divergence between two-point distributions $(p,1-p)$ and $(q,1-q)$. It is well-known (and is a simple consequence of the concavity of logarithm) that divergence is nonnegative. Now,

$$1+\left[\frac{1+y}{2}\log\left(\frac{(1+y)^2}{(1+t)^2}\right)+\frac{1-y}{2}\log\left(\frac{(1-y)^2}{(1-t)^2}\right)\right]-\left[\frac{1+y}{2}\log(1+y)+\frac{1-y}{2}\log(1-y)\right]=$$
$$1+2\left[\frac{1+y}{2}\log\left(\frac{(1+y)}{(1+t)}\right)+\frac{1-y}{2}\log\left(\frac{(1-y)}{(1-t)}\right)\right]+H\left(\frac{1-y}{2}\right)-\log 2=$$
$$2D\left(\frac{1-y}{2}\bigg|\bigg|\frac{1-t}{2}\right)+H\left(\frac{1-y}{2}\right)+(1-\log 2)\geq H\left(\frac{1-y}{2}\right)+(1-\log 2).$$

Therefore we need to show

$$1-\sqrt{1-y^2}\geq\log 2-H\left(\frac{1-y}{2}\right)$$

for all $-1\leq y\leq 1$.

We will need two well-known facts: the function $\phi(t)=\frac{1}{2}-\sqrt{t(1-t)}$ is an involution on $[0,\frac{1}{2}]$; and the function $R(x)=H(\phi(x))$ is convex on $[0,\frac{1}{2}]$.

Since $R(0)=\log 2$ and $R'(0)=-2$, by convexity $R(x)\geq\log 2-2x$ for $x\in[0,\frac{1}{2}]$. Rearranging and taking $x=\phi(z)$,

$$2\phi(z)\geq\log 2-H(z)$$

Substituting $y=1-2z$

$$1-\sqrt{1-y^2}\geq\log 2-H\left(\frac{1-y}{2}\right)$$

and we are done.

6 Appendix B - an inductive proof of a logarithmic Sobolev inequality

In this section we give an inductive proof of the inequality

$$\mathcal{E}(f,f) \geq 2 \cdot \sum_{i=1}^{n} Ent(d_i^2)$$

for a real function $f\ :\{0,1\} \rightarrow \mathbb{R}$, with d_i the difference functions of f (cf. section 2).

Observe that the right hand side might depend on the ordering of coordinates.

The proof is by induction on n. For $n=1$, $d_1 = f - \mathbb{E}f$, and therefore d_1^2 is a constant function with zero entropy. The claim follows. Assume the claim for $n-1$, and consider it for n.

Let the influence $I_i(f)$ of the i-th bit on a real-valued function f to be given by $\mathbb{E}_x(f(x) - f(x \oplus e_i))^2$. Thus $\mathcal{E}(f,f) = \sum_{i=1}^n I_i$.

Let f_0, f_1 be the restrictions of f to subcubes defined by the value of the n-th coordinate. We write $f \leftrightarrow (f_0, f_1)$. These are functions on $n-1$ variables. Let their influences, their conditional expectations, and their difference functions with respect to the natural ordering $1...n-1$ of the coordinates be denoted by $I_{i,0}$, $I_{i,1}$, $f_{i,0}$, $f_{i,1}$, $d_{i,0}$, $d_{i,1}$ correspondingly. Then, by the induction hypothesis

$$\sum_{i=1}^{n} I_i = \frac{1}{2}\left(\sum_{i=1}^{n-1} I_{i,0} + \sum_{i=1}^{n-1} I_{i,1}\right) + I_n \geq \sum_{i=1}^{n-1} Ent\left(d_{i,0}^2\right) + \sum_{i=1}^{n-1} Ent\left(d_{i,1}^2\right) + I_n$$

Consider now a slightly different ordering of the coordinates for f, which is $n, 1, 2..., n-1$. Let f_i and d_i be the conditional expectations and the difference functions in this new ordering. We will show

$$\sum_{i=1}^{n-1} Ent\left(d_{i,0}^2\right) + \sum_{i=1}^{n-1} Ent\left(d_{i,1}^2\right) + I_n \geq 2 \cdot \sum_{i=1}^{n} Ent\left(d_i^2\right)$$

This will prove the inequality for the ordering $n, 1, 2..., n-1$ of the coordinates.[6]

Observe $f_i \leftrightarrow (f_{i-1,0},\ f_{i-1,1})$ for $i = 2...n$, and similarly for difference functions. Note also that d_1^2 is constant, and therefore has zero entropy.

Lemma 6.1: *Let $k \geq 0$, $k \leftrightarrow (g,h)$. Then*

$$Ent(k) = \frac{1}{2}(Ent(g) + Ent(h)) + \frac{1}{2}\left(\mathbb{E}g \log \mathbb{E}g + \mathbb{E}h \log \mathbb{E}h - (\mathbb{E}g + \mathbb{E}h) \log \frac{\mathbb{E}g + \mathbb{E}h}{2}\right)$$

[6] Alternatively, we could have insisted on the 'natural' order of the coordinates for f, and changed the order of coordinates for f_0, f_1.

Proof: ▮
Therefore we need to show

$$\frac{1}{2}\cdot I_n \geq \sum_{i=2}^{n}\left(Ent\left(d_i^2\right)-\frac{1}{2}\left(Ent\left(d_{i-1,0}^2\right)+Ent\left(d_{i-1,1}^2\right)\right)\right)=$$

$$\frac{1}{2}\sum_{i=1}^{n-1}\left(\mathbb{E}d_{i,0}^2\log\mathbb{E}d_{i,0}^2+\mathbb{E}d_{i,1}^2\log\mathbb{E}d_{i,1}^2-\left(\mathbb{E}d_{i,0}^2+\mathbb{E}d_{i,1}^2\right)\log\frac{\mathbb{E}d_{i,0}^2+\mathbb{E}d_{i,1}^2}{2}\right)$$

Wor this purpose we need information on the joint behaviour of the sequences $\mathbb{E}d_{i,0}^2$ and $\mathbb{E}d_{i,1}^2$.

Applying the Cauchy-Schwarz inequality twice,

$$\left(\mathbb{E}d_{i,0}^2-\mathbb{E}d_{i,1}^2\right)^2=\mathbb{E}^2\left(d_{i,0}-d_{i,1}\right)\left(d_{i,0}+d_{i,1}\right)\leq\mathbb{E}^2|d_{i,0}-d_{i,1}||d_{i,0}+d_{i,1}|\leq$$

$$\mathbb{E}\left(d_{i,0}-d_{i,1}\right)^2\mathbb{E}\left(d_{i,0}+d_{i,1}\right)^2\leq\left(\mathbb{E}d_{i,0}^2+\mathbb{E}d_{i,1}^2+2\sqrt{\mathbb{E}d_{i,0}^2\mathbb{E}d_{i,1}^2}\right)\cdot\mathbb{E}\left(d_{i,0}-d_{i,1}\right)^2$$

Observe that $d_{i,0}-d_{i,1}$ is a difference sequence for f_0-f_1 and therefore $\sum_{i=1}^{n-1}\mathbb{E}\left(d_{i,0}-d_{i,1}\right)^2\leq(f_0-f_1)^2=I_n$.

Take $a_i:=\mathbb{E}d_{i,0}^2$, $b_i:=\mathbb{E}d_{i,1}^2$, and consider an optimization problem

$$\text{Maximize}\quad\sum_{i=1}^{n-1}\left(a_i\log a_i+b_i\log b_i-(a_i+b_i)\log\frac{a_i+b_i}{2}\right)$$

$$\text{Given}\quad\sum_{i=1}^{n-1}\frac{(a_i-b_i)^2}{a_i+b_i+2\sqrt{a_ib_i}}\leq I_n,\quad a_i,b_i\geq 0$$

Let $m_i=min\{a_i,b_i\}$ and $c_i=|a_i-b_i|$. Then an equivalent formulation is

$$\text{Maximize}\quad\sum_{i=1}^{n-1}\left((m_i+c_i)\log(m_i+c_i)+m_i\log m_i-(2m_i+c_i)\log\frac{2m_i+c_i}{2}\right)$$

$$\text{Given}\quad\sum_{i=1}^{n-1}\frac{c_i^2}{2m_i+c_i+2\sqrt{m_i(m_i+c_i)}}\leq I_n,\quad m_i,c_i\geq 0$$

Assume $c_i>0$ for all i since removing coordinates with $c_i=0$ does not effect neither the target function nor the constraint. Therefore we are allowed to consider $r_i=\frac{m_i}{c_i}$, leading to the following formulation

$$\text{Maximize}\quad\sum_{i=1}^{n-1}c_i\cdot\left((1+r_i)\log(1+r_i)+r_i\log r_i-(1+2r_i)\log\frac{1+2r_i}{2}\right)$$

$$\text{Given}\quad\sum_{i=1}^{n-1}\frac{c_i}{1+2r_i+2\sqrt{r_i(1+r_i)}}\leq I_n,\quad r_i,c_i\geq 0$$

The following technical claim completes the analysis.

Lemma 6.2: *For any $r \geq 0$ holds*

$$(1+r)\log(1+r) + r\log r - (1+2r)\log\frac{1+2r}{2} \leq \frac{1}{1+2r+2\sqrt{r(1+r)}}$$

In addition

$$\left(1+2r+2\sqrt{r(1+r)}\right)\cdot\left((1+r)\log(1+r) + r\log r - (1+2r)\log\frac{1+2r}{2}\right) \to_{r\to\infty} 1$$

Therefore the supremum of the above maximization problem is bounded by I_n (and it can easily be seen that it actually equals I_n), completing the proof.

Proof: (Of the lemma)

Let $g(r) = 1 + 2r + 2\sqrt{r(1+r)} = \left(\sqrt{r} + \sqrt{1+r}\right)^2$, and $h(r) = (1+r)\log(1+r) + r\log r - (1+2r)\log\frac{1+2r}{2}$. We want to show $(gh)(r) \leq 1$, for all $r \geq 0$. At zero, $(gh)(0) = \log 2 < 1$, at infinity, $g(r) \sim 4r$ and $h(r) \sim \frac{1}{4r}$, and thus $(gh)(r) \to_{r\to\infty} 1$, proving the second part of the lemma. Thus it is sufficient to show gh is increasing, or $\frac{g'h}{g} \geq -h'$.

Computing, $\frac{g'h}{g} = \frac{h}{\sqrt{r(1+r)}}$, and $-h' = \log\frac{(1+2r)^2}{4r(1+r)}$. It remains to show $h(r) \geq \sqrt{r(1+r)} \cdot \log\frac{(1+2r)^2}{4r(1+r)}$ Rewriting $h(r)$ as $\log\frac{2+2r}{1+2r} - r\log\frac{(1+2r)^2}{4r(1+r)}$, this is the same as

$$\log\frac{2+2r}{1+2r} \geq \left(\sqrt{r(1+r)} + r\right)\cdot\log\frac{(1+2r)^2}{4r(1+r)} = \frac{r}{\sqrt{r(1+r)} - r}\cdot\log\frac{(1+2r)^2}{4r(1+r)}$$

Or

$$\left(\sqrt{\frac{1+r}{r}} - 1\right)\cdot\log\frac{2+2r}{1+2r} \geq \log\frac{(1+2r)^2}{4r(1+r)}.$$

Let $t = \sqrt{\frac{1+r}{r}}$. Then $t \in (1,\infty)$. Rewriting in terms of t, we want to have

$$(t-1)\log\frac{2t^2}{t^2+1} \geq 2\log\frac{t^2+1}{2t}$$

This holds at one. Comparing the derivatives, it suffices to show

$$\log\frac{2t^2}{t^2+1} \geq \frac{2t-2}{t^2+1}.$$

Once again, this holds at one. Comparing the derivatives for the final time, one has to show

$$\frac{1}{t} \geq \frac{-t^2+2t+1}{t^2+1},$$

or $t^3 + 1 \geq t^2 + t$, which is immediate for $t \geq 1$. ∎

7 Acknowledgements

We are grateful to Ehud Friedgut for his suggestions which led to a significant simplification of the proof of theorem 2.2. We also thank Gil Kalai and Nati Linial for many valuable conversations and remarks.

References

[1] W. Beckner, *Inequalities in Fourier Analysis*, Annals of Math., 102(1975), pp. 159-182.

[2] S. Bezrukov, *Isoperimetric problems in discrete spaces*, in **Extremal Problems for Finite Sets**, Bolyai Soc. Math. Stud. 3, (P. Frankl, Z. Fu"redi, G. Katona, D. Miklos eds.) 1994,pp. 59-91.

[3] M. Ben-Or, N. Linial, *Collective Coin Flipping*, in **Randomness and Computation** (S. Micali, ed.), Academic Press.

[4] S. G. Bobkov, P. Tetali, *Modified logarithmic Sobolev inequalities in discrete settings*, Proceedings of STOC 2003, pp. 287-296.

[5] A. Bonami, *Etude des coefficients Fourier des fonctiones de $L^p(G)$*, Ann. Inst. Fourier (Grenoble) 20:2 (1970), pp. 335-402.

[6] P. Diaconis, L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Prob., 6 (1996), pp. 695-750.

[7] E. Friedgut, *Boolean functions with low average sensitivity depend on few coordinates*, Combinatorica 18(1), 1998, pp. 27-36.

[8] E. Friedgut, *Sharp Thresholds of Graph Proprties, and the k-sat Problem.*, J. Amer. Math. Soc. 12 (1999), no. 4, pp. 1017–1054.

[9] E. Friedgut, G. Kalai, *Every monotone graph property has a sharp threshold*, Proc. AMS 124 (1996), pp. 2993-3002.

[10] E. Friedgut, V. Rodl, *Proof of a Hypercontractive Estimate via Entropy*, Israel J. Math. 125 (2001), pp. 369–380 .

[11] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. of Math., 97 (1975), pp. 1061-1083.

[12] L. H. Harper, *Optimal assignment of numbers to vertices*, J. Soc. Ind. Appl. Math., 12, 1964, pp. 131-135.

[13] L. H. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Comb. Theory 1, 1966, pp. 385-393.

[14] S. Hart, *A note on the edges of the n-cube*, Discr. Math., 14, 1976, pp. 157-163.

[15] M. Jerrum, **Counting, sampling and integrating: algorithms and complexity**, Lectures in Mathematics, ETH Zurich. Birkhauser Verlag, Basel, 2003.

[16] M. Jerrum, A. Sinclair *The Markov chain Monte Carlo method: an approach to approximate counting and integration* in **Approximation Algorithms for NP-hard Problems**, (Dorit Hochbaum, ed.), PWS, 1996.

[17] J. Kahn, G. Kalai, *A discrete isoperimetric conjecture with probabilistic applications*, preprint.

[18] G. Kalai, personal communication.

[19] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on boolean functions*, FOCS 1988, pp. 68-80.

[20] R. Latala, *On some inequalities for Gaussiam measures*, in Proceedings of the International Congress of Mathematicians, vol II, pp. 813-822.

[21] V. I. Levenshtein, *Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces*, IEEE Trans. Inform. Theory, vol. IT-41, 1995, 1303-1321.

[22] J.H. van Lint, **Introduction to Coding Theory**, Springer-Verlag, Berlin, 1999.

[23] V. D. Milman, G. Schechtman, **Asymptotic theory of finite dimensional normed spaces**, Springer-Verlag, New York, 1986.

[24] E.Mossel, R. ODonnell, K. Oleszkiewicz *Noise stability of functions with low influences: invariance and optimality*, FOCS 2005, pp. 21-30.

[25] E. Mossel, personal communication.

[26] M. Navon, A. Samorodnitsky, *On Delsarte's Linear Programming Bounds for Binary Codes*, Proceedings of FOCS 2005.

[27] R. Rossignol, *Threshold for monotone symmetric properties through a logarithmic Sobolev inequality*, arXiv:math.PR/0511607

[28] M. Talagrand, *On boundaries and influences*, Combinatorica 17, 1997, pp. 275-285.

[29] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publ. Math. I.H.E.S. 81, 1995, 73-203.

[30] M. Talagrand, *On Russo's approximate* $0-1$ *law*, The Annals of Prob., 22, 3 (1994), pp. 1576-1587.

[31] G. Szegö, **Orthogonal Polynomials**, American Mathematical Society, 1939.