On the Influences of Variables on Boolean Functions in Product Spaces

NATHAN KELLER
Einstein Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
(e-mail: nkeller@math.huji.ac.il)

Received 23 April 2009; revised 24 May 2010; first published online 9 July 2010

In this paper we consider the influences of variables on Boolean functions in general product spaces. Unlike the case of functions on the discrete cube, where there is a clear definition of influence, in the general case several definitions have been presented in different papers. We propose a family of definitions for the influence that contains all the known definitions, as well as other natural definitions, as special cases. We show that the proofs of the BKKKL theorem and of other results can be adapted to our new definition. The adaptation leads to generalizations of these theorems, which are tight in terms of the definition of influence used in the assertion.

1. Introduction

Influences of variables on Boolean functions have been extensively studied in the last decades. This study has led to important applications in theoretical computer science, combinatorics, mathematical physics, social choice theory, and other areas. The basic results on influences were obtained for Boolean functions on the discrete cube, with respect to the uniform probability measure. However, some of the applications (including applications to theoretical computer science, to statistical physics, and to random graph theory) require generalization of the results to more general probability measures (e.g., [7, 11]), and even to more general product spaces (e.g., [13, 15]). Unlike the discrete case, where there exists a single natural definition of influence, for general product spaces several definitions have been presented in different papers. All the definitions are based on dividing the space into subspaces called fibres.

Notation. In the following definitions, $X$ denotes a product space $X = X_1 \times X_2 \times \cdots \times X_n$, endowed with a product measure $\mu = \mu_1 \otimes \mu_2 \otimes \cdots \otimes \mu_n$. Throughout the paper, $\log n$ denotes $\log_2 n$.

Definition 1. For $x = (x_1, \ldots, x_n) \in X$ and for $1 \leq k \leq n$, the fibre of $x$ in the $k$th direction is $s_k(x) = \{y \in X : y_i = x_i, \forall i \neq k\}$. 
The influence of the $k$th variable on a function $f : X \to \{0, 1\}$ is the expectation, over all fibres in the $k$th direction, of the influence of the variable on each fibre. The original definition of influences in product spaces, introduced by Bourgain, Kahn, Kalai, Katznelson and Linial [6], is as follows.

**Definition 2.** For $f : X \to \{0, 1\}$, and for $1 \leq k \leq n$, the influence of the $k$th variable on $f$ is
\[ I_k(f) = \mu[\{x \in X : f \text{ is non-constant on } s_k(x)\}] . \]

From now on we use a slightly modified version of this definition proposed by Grimmett [12], in which a function is considered constant on a fibre if it is constant, except on a set of measure zero. Another widely used definition (see, e.g., [15, 19]), is as follows.

**Definition 3.** For a function $f : X \to \{0, 1\}$, an element $x = (x_1, \ldots, x_n) \in X$, and for $1 \leq k \leq n$, denote the restriction of $f$ to the fibre $s_k(x)$ by $f^x_k : X_k \to \{0, 1\}$. That is, for all $t \in X_k$,
\[ f^x_k(t) = f(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n) . \]

The influence of the $k$th variable on $f$ is
\[ \tilde{I}_k(f) = \mathbb{E}_{x \in X} [\text{Var}(f^x_k)] . \]

The difference between the definitions is demonstrated by the following example.

**Example.** Consider $X = [0, 1]^n$, endowed with the Lebesgue measure $\lambda$. Let $f : [0, 1]^n \to \{0, 1\}$ be defined by
\[ f(x) = 1 \iff x_i > 1/n, \forall 1 \leq i \leq n . \]

It is easy to see that for all $1 \leq k \leq n$ we have $I_k(f) = (1 - 1/n)^{n-1} \approx 1/e$, while $\tilde{I}_k(f) = I_k(f) \cdot \frac{n-1}{n} \approx 1/en$.

In general, for every function $f$ we have $\tilde{I}_k(f) \leq I_k(f)/4$, but in many cases $\tilde{I}_k(f)$ is much smaller than $I_k(f)$.

The most well-known theorem concerning influences in general product spaces is the BKKKL theorem [6].

**Theorem 1.1 (Bourgain, Kahn, Kalai, Katznelson and Linial).** For every $f : X \to \{0, 1\}$ such that $\mathbb{E}f = p$, there exists a variable $k$ such that $I_k(f) \geq cp(1 - p) \log n/n$, where $c$ is a universal constant.

The BKKKL theorem has been used to obtain several important results. For example, it was a central tool in showing that in the random graph model $G(n, p)$, any monotone graph property has a sharp threshold [11].

In this paper we present a new definition of influences in general product spaces, which contains Definitions 2 and 3 as special cases.
Definition 4. Let $h : [0, 1] \to [0, 1]$. For every $f : X \to \{0, 1\}$ and for each $1 \leq k \leq n$, the $h$-influence of the $k$th coordinate on $f$ is

$$I^h_k(f) = \mathbb{E}_{x \in X}[h(\mathbb{E}f^x_k)].$$

Note that since the range of $f$ is $\{0, 1\}$, we have $\mathbb{E}f_k = \mathbb{P}[f_k = 1]$. Furthermore, if $f$ is monotone then $\mathbb{P}[f_k = 1]$ determines $f_k$ completely, and thus $h(\mathbb{E}f_k)$ reflects the entire structure of $f_k$.

Definition 2 is obtained from the new definition by substituting $h(t) = 1$ if $t \neq 0, 1$, and $h(t) = 0$ otherwise. Definition 3 is obtained by substituting $h(t) = t(1 - t)$.\footnote{We note that influences toward zero and one (see, e.g., [10]), can also be expressed as $h$-influences. Influence toward one is obtained by $h(t) = 1 - t$ for $t \neq 0$, and $h(0) = 0$. Influence toward zero is obtained by $h(t) = t$ for $t \neq 1$, and $h(1) = 0$. We also note that after this paper was written, our results were applied in [17] to prove a variant of the BKKKL theorem for a newly proposed definition of influences in product spaces of continuous distributions, called ‘geometric influences’.

We show that the BKKKL theorem can be generalized by replacing the influence $I_k(f)$ with a smaller $h$-influence, and obtain a full characterization of the $h$-influences for which the BKKKL theorem holds. As in [6], we prove our results in the case $X = [0, 1]^n$, endowed with the Lebesgue measure $\lambda$. This case is considered fairly general (see [6, 15]), since many cases of interest can be easily reduced to it.\footnote{We note that in [12], the BKKKL theorem is generalized to any product measure space satisfying some separability conditions. It seems possible that our results can also be generalized to such spaces, but the proof given in [12] does not imply such generalization directly.}

Theorem 1.2. Denote the entropy function $H(t) = -t \log t - (1 - t) \log(1 - t)$ by $\text{Ent}(t)$. Let $h : [0, 1] \to [0, 1]$ such that $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. Then, for every $f : [0, 1]^n \to \{0, 1\}$ with $\mathbb{E}f = p$, there exists $1 \leq k \leq n$ such that the $h$-influence of the $k$th variable on $f$ satisfies

$$I^h_k(f) \geq cp(1 - p) \log n/n,$$

where $c$ is a universal constant.

The advantage of Theorem 1.2 over the BKKKL theorem is demonstrated by the function $f(x) = 1 \iff (x_i > 1/n, \forall 1 \leq i \leq n)$, presented above. For this function we have $\mathbb{E}(f) \approx 1/e$, and for all $k$, $I_k(f) \approx 1/e$, and hence the BKKKL theorem is far from being tight in this case. On the other hand, for $h(t) = \text{Ent}(t)$ we get $I^h_k(f) \approx \text{Ent}(1/n)/e \approx c' \log n/n$, and thus Theorem 1.2 is tight in this case.

Furthermore, by examining variants of the tribes function presented by Ben-Or and Linial [3], we show that Theorem 1.2 is tight in the following sense.

Proposition 1.3. Let $h : [0, 1] \to [0, 1]$ and let $\epsilon > 0$. If there exists $0 < q < 1$ such that $h(q) \leq \epsilon \text{Ent}(q)$, then for all $n \geq n_0(q)$ there exists a function $f : [0, 1]^n \to \{0, 1\}$ such that $\mathbb{E}f = \Theta(1)$, and for all $1 \leq k \leq n$, the $h$-influence of the $k$th variable on $f$ satisfies $I^h_k(f) \leq c \epsilon \log n/n$, where $c$ is a universal constant.
The proof of Theorem 1.2 for monotone functions is a simple adaptation of the proof of the BKKKL theorem given by Friedgut [10]. The reduction from general functions to monotone functions is a bit more complicated in our case, and involves a new monotonization argument that exploits the concavity of the entropy function. We note that statements similar to Theorem 1.2 have been proved by Talagrand [22], and independently, by Friedgut and Kalai [11, Theorem 3.1] (both in the special case of functions on the discrete cube endowed with a product measure \( \mu_q \)), and by Bollobás and Riordan [4, Theorem 5] (for a product measure on \( \{-1, 0, 1\}^n \)).

We show that several known results on influences in product spaces can be easily generalized to the setting of \( h \)-influences. In particular, we generalize a lower bound on the vector of influences obtained by Talagrand [22] for functions on the discrete cube endowed with the product measure \( \mu_q \). The lower bound on \( h \)-influences we obtain is presented in Proposition 3.4 below. We show that Proposition 3.4 follows easily from Talagrand’s result for the uniform measure \( \mu_{1/2} \). Since Talagrand’s result for a general measure \( \mu_q \) follows immediately from Proposition 3.4, our technique can replace the major part of Talagrand’s proof (containing a proof of a biased version of the Bonamie–Beckner hypercontractive inequality [5, 2]).

This paper is organized as follows. In Section 2 we present the monotonization and discretization techniques for \( h \)-influences. The proofs of the main results are presented in Section 3. In Section 4 we discuss the tightness of our assertions.

2. Monotonization and discretization for \( h \)-influences

In this section we generalize to \( h \)-influences two central steps in the proof of the BKKKL theorem. The first step is monotonization, a shifting technique allowing us to replace a Boolean function on the continuous cube by a monotone Boolean function with the same expectation and non-higher influences. Whereas for Definition 2 of the influences, the shifting argument is standard and easy (see, e.g., [8]), for \( h \)-influences the proof is a bit more complicated.\(^3\) We present this proof, involving discretization and a new argument exploiting the concavity of the function \( h \), in Section 2.1. The second step is discretization, allowing us to approximate a monotone function \( f : [0, 1]^n \to \{0, 1\} \) by a function \( g : \{0, 1\}^n \to \{0, 1\} \) having the following property: the \( ln \) coordinates can be divided into \( n \) sets of size \( l \) each, such that the sum of influences of the coordinates in each set on \( g \) is not much bigger than the influence of the corresponding coordinate on \( f \). We show that the same can be performed with \( h \)-influences, given that, for all \( x \), we have \( h(x) \geq \text{Ent}(x) \). The proof of this claim, presented in Section 2.2, is an easy adaptation of the proof of Claim 2.8 in [10].

2.1. Monotonization for \( h \)-influences

Definition 5. A function \( f : [0, 1]^n \to \mathbb{R} \) is monotone if, for all \( x, y \in [0, 1]^n \),

\[
\forall i(x_i \geq y_i) \implies f(x) \geq f(y).
\]

\(^3\) We note that the proof works only under the assumption that \( h \) is concave and continuous. However, most of the natural definitions of \( h \)-influences satisfy this assumption.
On the Influences of Variables on Boolean Functions in Product Spaces

Figure 1. The matrices representing $f$ and $Mf$.

**Theorem 2.1.** Let $h : [0, 1] \rightarrow [0, 1]$ be a concave continuous function. For every Borel measurable function $f : [0, 1]^n \rightarrow \{0, 1\}$, there exists a monotone function $g : [0, 1]^n \rightarrow \{0, 1\}$ such that:

1. $\mathbb{E}f = \mathbb{E}g$.
2. For all $1 \leq k \leq n$, we have $I^h_k(f) \geq I^h_k(g)$.

**Proof.** The proof of Theorem 2.1 is composed of three steps.

**2.1.1. Reduction to functions on the unit square.** The monotonization procedure, which allows us to obtain $g$ given $f$, is the same standard shifting procedure as presented in [6]. The monotonization is performed coordinate-wise. For the $k$th coordinate, we consider the fibres in the $k$th direction, and for each fibre $s_k(x)$, we replace the function $f^x_k$ by the function

$$Mf^x_k(t) = \begin{cases} 1 & t > 1 - \mathbb{E}(f^x_k), \\ 0 & \text{otherwise}. \end{cases}$$

Clearly, the resulting function $Mf$ satisfies $\mathbb{E}(Mf) = \mathbb{E}(f)$. Furthermore, it is clear that performing this monotonization procedure in all the coordinates sequentially leads to a monotone function. Hence, the only claim we have to prove is that $I^h_j(f) \geq I^h_j(Mf)$ for all $1 \leq j \leq n$, where $Mf$ denotes the monotonization of $f$ in the $k$th direction.

For $j = k$ we have $I^h_j(f) = I^h_j(Mf)$, since the expectation of $f^x_k$ on each fibre is preserved. Thus, we assume that $j \neq k$. It is sufficient to show that for every fixed value of the $n - 2$ remaining coordinates, the contribution of the family of fibres corresponding to that fixed value to $I^h_j(f)$ is not less than the respective contribution to $I^h_j(Mf)$. Hence, without loss of generality we can fix the remaining coordinates and thus assume that $n = 2$, i.e., $f : [0, 1]^2 \rightarrow \{0, 1\}$.

**2.1.2. Proof of a discrete version of Theorem 2.1.** Let $r$ be an integer. We assume that the value of $f$ is constant on squares of the form $\left(\frac{l}{r}, \frac{l+1}{r}\right) \times \left(\frac{m}{r}, \frac{m+1}{r}\right)$, for all $0 \leq l, m < r$. Hence, $f$ can be represented by an $r \times r$ matrix $A$, where $A[i, j]$ denotes the value of $f(x, y)$ for all $(x, y) \in \left(\frac{l-1}{r}, \frac{l}{r}\right) \times \left(\frac{m-1}{r}, \frac{m}{r}\right)$. The monotonization (assuming, without loss of generality, that it is performed according to the rows of the matrix) consists of moving all the 1s in each row to the rightmost positions in the same row, as described in Figure 1.
The influence in the ‘columns’ direction is

\[ I_{\text{columns}}^h(f) = \frac{1}{r} \sum_{j=1}^{r} h\left( \frac{\sum_{i=1}^{r} A[i, j]}{r} \right) . \]

We show that if \( h \) is concave, the influence is not increased by our monotonization. We use two simple observations:

1. if two entire columns are interchanged, the influence does not change,
2. if a ‘1’ is moved from a column with a smaller (or equal) number of 1s to a column with a bigger (or equal) number, then the influence is not increased.

The first observation is clear. The second observation follows from the concavity of \( h \). Indeed, denote the numbers of 1s in the columns before the operation by \( a \) and \( b \), where \( b \geq a \). Since the contribution of all the untouched columns to the influence remains unchanged, we want to prove that

\[ h\left( \frac{b}{r} \right) + h\left( \frac{a}{r} \right) \geq h\left( \frac{b + 1}{r} \right) + h\left( \frac{a - 1}{r} \right) , \]

or equivalently

\[ h\left( \frac{a}{r} \right) - h\left( \frac{a - 1}{r} \right) \geq h\left( \frac{b + 1}{r} \right) - h\left( \frac{b}{r} \right) . \]

The latter means that the average gradient of \( h \) on the segment \( \left[ \frac{a-1}{r}, \frac{a}{r} \right] \) is not smaller than the average gradient on the segment \( \left[ \frac{b}{r}, \frac{b+1}{r} \right] \), and this indeed holds for any concave function.

The transformation from \( f \) to \( Mf \) can be treated as a concatenation of steps of the two classes discussed in the observations above. Indeed, the transformation can be represented by the following algorithm.

For \( i = 0, 1, 2, \ldots, r - 1 \):

1. Consider columns 1, 2, \ldots, \( r - i \) of the matrix.
2. Reorder the columns, such that the rightmost column will have the maximal number of 1s.
3. Consider the zero values in the rightmost column (i.e., column \( r - i \)). For each such value, look at the corresponding row of the matrix (restricted to the first \( r - i \) columns), and if there exist ‘1’ values in the row, interchange one of them with the zero value in the rightmost column.

For example, in Figure 1, the step of the algorithm corresponding to \( i = 0 \) consists of interchanging the ‘1’ in the cell \( A[3, 5] \) with the zero in the cell \( A[3, 6] \), and interchanging the ‘1’ in the cell \( A[5, 3] \) with the zero in the cell \( A[5, 6] \). (There is no reordering of the columns since the rightmost column already has the maximal number of 1s.) In the remaining steps of the algorithm, the \( r \)th column is left untouched.

In the output of the algorithm, all the 1s in each row are concentrated in the rightmost cells. Since, during the algorithm, only cells in the same row are interchanged, the output of the algorithm is the matrix representing \( Mf \). The steps of the algorithm are indeed of the two classes discussed in the observations above (reordering the columns and moving a ‘1’ from a column with a smaller (or equal) number of 1s to a column with a bigger (or equal) number of 1s).
Therefore, the \( h \)-influence is not increased by the transformation from \( f \) to \( Mf \). This proves the assertion of Theorem 2.1 for a discrete version of the function \( f \).

2.1.3. Reduction from the general case to the discrete case. Let \( f : [0,1]^2 \rightarrow \{0,1\} \) be a Borel measurable function, and let \( \epsilon > 0 \). We want to show that \( I_2^h(f) \geq I_2^h(Mf) - \epsilon \), where \( Mf \) denotes the monotonization of \( f \) in the first direction (i.e., with respect to the \( x \)-coordinate). We would like to approximate \( f \) by a ‘discrete’ function \( f_s \) and use the result on discrete functions obtained in the previous step of the proof.

The function \( h \) is continuous on \([0,1]\), and thus there exists \( \delta \) such that if \(|x-y| \leq \delta\), then \(|h(x) - h(y)| \leq \epsilon/4\). Since, by the construction of the product measure on the unit square, the set of functions which are constant on products of intervals of the form \((\frac{1}{2^r}, \frac{1}{2^{r+1}}) \times (\frac{m}{2^r}, \frac{m+1}{2^r})\) for some \( s \) is dense in \(L^1([0,1]^2)\) (see, e.g., [1], Theorem 5.7), there exists \( s \) and a function \( f_s \) which is constant on squares of the form \((\frac{1}{2^r}, \frac{1}{2^{r+1}}) \times (\frac{m}{2^r}, \frac{m+1}{2^r})\), such that

\[
\int_0^1 \int_0^1 |f(x,y) - f_s(x,y)| \, dx \, dy \leq \delta \epsilon/4.
\]  
(2.1)

Let

\[
S_1 = \left\{ x \in [0,1] : \int_0^1 |f(x,y) - f_s(x,y)| \, dy \leq \delta \right\},
\]

and \( S_2 = [0,1] \setminus S_1 \). By the Fubini theorem and (2.1), we have \( \lambda(S_2) \leq \epsilon/4 \), where \( \lambda \) is the Lebesgue measure on \( \mathbb{R} \). On the other hand, if \( x \in S_1 \), then

\[
\left| \int_0^1 f(x,y) \, dy - \int_0^1 f_s(x,y) \, dy \right| \leq \int_0^1 |f(x,y) - f_s(x,y)| \, dy \leq \delta,
\]

and hence,

\[
|h\left(\int_0^1 f(x,y) \, dy\right) - h\left(\int_0^1 f_s(x,y) \, dy\right)| \leq \epsilon/4.
\]

Therefore, using the triangle inequality, we get

\[
|I_2^h(f) - I_2^h(f_s)| \leq \int_0^1 \left|h\left(\int_0^1 f(x,y) \, dy\right) - h\left(\int_0^1 f_s(x,y) \, dy\right)\right| \, dx \leq (\epsilon/4) \cdot \lambda(S_1) + 1 \cdot \lambda(S_2) \leq \epsilon/2.
\]  
(2.2)

In order to bound the term \( |I_2^h(Mf) - I_2^h(Mf_s)| \), we note that since for each \( y \) the two sets \( A_y = \{ x : Mf(x,y) = 1 \} \) and \( B_y = \{ x : Mf_s(x,y) = 1 \} \) are nested (i.e., one of them contains the other), we have (for each \( y \))

\[
\int_0^1 |Mf(x,y) - Mf_s(x,y)| \, dx \leq \int_0^1 |f(x,y) - f_s(x,y)| \, dx,
\]

and thus, by the Fubini theorem and (2.1),

\[
\int_0^1 \int_0^1 |Mf(x,y) - Mf_s(x,y)| \, dx \, dy \leq \int_0^1 \int_0^1 |f(x,y) - f_s(x,y)| \, dx \, dy \leq \delta \epsilon/4.
\]

Hence, by the argument applied above to \( |I_2^h(f) - I_2^h(f_s)| \),

\[
|I_2^h(Mf) - I_2^h(Mf_s)| \leq \epsilon/2.
\]  
(2.3)
Finally, combining (2.2) and (2.3) with the proof of the theorem in the discrete case, we get
\[ I_2^h(f) - I_2^h(Mf) = (I_2^h(f) - I_2^h(f_i)) + (I_2^h(f_i) - I_2^h(Mf)) + (I_2^h(Mf) - I_2^h(Mf)) \]
\[ \geq (-\epsilon/2) + 0 + (-\epsilon/2) = -\epsilon. \]  
(2.4)

This completes the proof of Theorem 2.1.

2.2. Discretization for h-influences

Proposition 2.2. Let \( h : [0, 1] \rightarrow [0, 1] \) be a continuous function such that \( h(t) \geq \text{Ent}(t) \) for all \( 0 \leq t \leq 1 \), and let \( f : [0, 1]^n \rightarrow \{0, 1\} \) be a monotone function. For any \( \epsilon > 0 \), there exists \( l = l(f, h, \epsilon) \) and a monotone function \( g : [0, 1]^n \rightarrow \{0, 1\} \), such that:

- \( |E(g) - E(f)| < \epsilon \),
- the coordinates of \( \{0, 1\}^n \) are divided into \( n \) sets \( S_1, S_2, \ldots, S_n \) of size \( l \) each, and for all \( 1 \leq i \leq n \),
  \[ \sum_{j \in S_i} I_j(g) \leq 4I_i^h(f) + \epsilon, \]

where the influence in the left-hand side is with respect to the uniform measure on \( \{0, 1\}^n \).

Proof. As shown in Section 2.1.3,\(^4\) for any measurable function \( f : [0, 1]^n \rightarrow \{0, 1\} \) and for any \( \alpha > 0 \), there exists \( l = l(f, \alpha) \) and a function \( \tilde{f} : [0, 1]^n \rightarrow \{0, 1\} \), which is constant on the \( 2^{ln} \) subcubes obtained by subdividing each base interval of \( [0, 1]^n \) into \( 2^l \) equal parts, such that
\[ \mathbb{E} |\tilde{f} - f| < \alpha. \]

Moreover, it is easy to see that if \( f \) is monotone, then \( \tilde{f} \) can be chosen to be monotone. Indeed, for each fixed \( l \), the function \( \tilde{f} \) defined by \( \tilde{f} |_{\text{subcube}} = 1 \) if and only if \( E_{\text{subcube}}[f] \geq 1/2 \) (which is the closest to \( f \) in \( L^1 \)-norm amongst all functions which are constant on the \( 2^{ln} \) subcubes) is clearly monotone. Since \( h \) is continuous on \([0, 1]\), for any \( \epsilon > 0 \) there exists \( \delta > 0 \) such that if \( |x - y| < \delta \) then \( |h(x) - h(y)| < \epsilon \). Choose \( l = l(f, h, \epsilon) \) and \( \tilde{f} \) as described above, such that
\[ \mathbb{E} |\tilde{f} - f| < \delta \epsilon/4. \]

An argument similar to that used in the proof of (2.2) in Section 2.1.3 shows that, in this case, for all \( 1 \leq i \leq n \),
\[ |I_i^h(\tilde{f}) - I_i^h(f)| < \epsilon. \]

The function \( \tilde{f} \) corresponds to a function \( g : [0, 1]^n \rightarrow \{0, 1\} \), by replacing the interval \([m2^{-l}, (m + 1)2^{-l}] \) with the binary expansion of \( m \). Each of the \( n \) coordinates of \([0, 1]^n \) corresponds to a set of \( l \) coordinates of \([0, 1]^n \), and we denote the set corresponding to coordinate \( i \) by \( S_i \). Clearly, \( E(g) = E(\tilde{f}) \). Thus, by the construction of \( \tilde{f} \), the assertion of Proposition 2.2 would follow once we show that for all \( 1 \leq i \leq n \),
\[ \sum_{j \in S_i} I_j(g) \leq 4I_i^h(\tilde{f}). \]  
(2.5)

\(^4\) See (2.1). Actually, the proof in Section 2.1.3 considers the special case \( n = 2 \) of the statement here, and the general case is similar.
We note that it is sufficient to prove the same inequality for the contribution of any single fibre $s_i(x)$ to the influences in the two sides of (2.5). On the right-hand side, this contribution is given by $h(\mathbb{E}[\tilde{f}_{i}])$, where $\tilde{f}_i$ is the restriction of $\tilde{f}$ to the fibre $s_i(x)$. In order to represent the contribution to the left-hand side, we fix $(n-1)l$ coordinates in $\{0,1\}^n$ (i.e., all coordinates except for those in $S_i$) according to the corresponding coordinates of $x$, and consider the restriction of $g$ to the $l$ coordinates in $S_i$. We denote this restriction by $g_i^x: \{0,1,2,\ldots,2^l-1\} \rightarrow \{0,1\}$, and alert the reader to the slight abuse of notation. In terms of this definition, the contribution to the left-hand side is given by $\sum_{j=1}^{l} I_j(g_i^x)$.

Denote $t_0 = \mathbb{E}(f_i)$, and assume that $t_0 \geq 1/2$ (the case $t_0 < 1/2$ is similar). Since $\tilde{f}_i$ is monotone, we have $\tilde{f}_i(t) = 0$ if $t < 1 - t_0$ and $\tilde{f}_i(t) = 1$ if $t > 1 - t_0$. By the construction of $g$, the function $g_i^x$ satisfies $g_i^x(m) = 0$ if $m < (1 - t_0)2^l$ and $g_i^x(m) = 1$ if $m \geq (1 - t_0)2^l$. It is well known (as part of the proof of the Edge Isoperimetric Inequality on the cube: see [14]) that functions of this form have the minimal sum of influences amongst Boolean functions with the same expectation, and that their sum of influences is of order $\Theta((1 - t_0) \log(1/(1 - t_0)))$ (see [11, Proof of Theorem 3.1]). For the sake of completeness, we present the short proof here.

Let $k$ satisfy $2^k \leq (1 - t_0)2^l < 2^{k+1}$. Choosing $m$ uniformly at random from $0, 1, 2, \ldots, 2^l-1$, we have

$$\mathbb{P}[g_i^x(m) \neq g_i^x(m \oplus e_j)] \leq \begin{cases} 2^{j+1-l} & j \leq k, \\ 2(1 - t_0) & j > k. \end{cases}$$

Indeed, for $j \leq k$, the relation $[g_i^x(m) \neq g_i^x(m \oplus e_j)]$ can be satisfied only for $(1 - t_0)2^l - 2^j \leq m < (1 - t_0)2^l + 2^j$, and hence the probability is bounded from above by $2^{j+1-l}$. For $j > k$, the restriction is even stricter: $m$ must satisfy either $0 \leq m < (1 - t_0)2^l$ or $2^l \leq m < (1 - t_0)2^l + 2^j$, and thus the probability is bounded by (and actually is equal to) $2(1 - t_0)$. Therefore,

$$\sum_{j=1}^{l} I_j(g_i^x) \leq \sum_{j \leq k} 2^{j+1-l} + \sum_{j > k} 2(1 - t_0) \leq 2^{k+2-l} + 2(1 - t_0)(l - 1 - k). \quad (2.6)$$

Here we consider two cases, according to the value of $t_0$. For $t_0 \geq 3/4$, we have

$$2^{k+2-l} + 2(1 - t_0)(l - 1 - k) \leq 4(1 - t_0) + 2(1 - t_0) \log \frac{1}{1 - t_0} \leq 4(1 - t_0) \log \frac{1}{1 - t_0} \leq 4\text{Ent}(t_0),$$

and thus,

$$\sum_{j=1}^{l} I_j(g_i^x) \leq 4\text{Ent}(t_0). \quad (2.7)$$

For $1/2 \leq t_0 < 3/4$, we have $k = l - 2$ and Ent$(t_0) > 0.8$, and thus (2.6) implies (2.7) trivially. Finally, since by assumption $h(t) \geq \text{Ent}(t)$ for all $t$, (2.5) follows from (2.7). This completes the proof. \hfill \Box

Remark. We note that $T_l(f) = \sum_{j \in S_l} I_j(g)$ (where $g$ is as defined above) was treated in [6] as an alternative definition of influences in the continuous case. It follows from the proof of Proposition 2.2 that this definition is equivalent, up to a multiplicative constant, to $h$-influence
with \( h(t) = \text{Ent}(t) \). As follows from Theorem 1.2, this is the ‘optimal’ definition of influence for which the BKKKL theorem holds.

3. Generalized BKKKL theorem and other results

In this section we show that the results of Section 2 allow us to adapt to \( h \)-influences the proofs of the BKKKL theorem and of several other known results concerning influences in product spaces. The results we consider include the following:

- a lower bound on the vector of influences obtained by Talagrand [22] (Section 3.2),
- a characterization of functions with a low sum of influences, obtained by Dinur [10], Friedgut [9, 10], and Hatami [15] (Section 3.3),
- a relation between the measure of the boundary of a subset of the discrete cube and its influences, obtained by Margulis [18] and Talagrand [21] (Section 3.4).

We note that the strongest form of the assertions in this section is obtained for the function \( h(t) = \text{Ent}(t) \). However, for the sake of generality, we state the results for general \( h \)-influences satisfying the condition ‘\( h(t) \geq \text{Ent}(t) \) for all \( 0 \leq t \leq 1 \)’.

3.1. BKKKL theorem for \( h \)-influences: Proof of Theorem 1.2

In the proof of the theorem, we follow the simple proof of the BKKKL theorem presented by Friedgut [10]. The proof uses the following generalization of the KKL theorem presented in [11].

**Theorem 3.1 (Friedgut and Kalai).** There exists a constant \( c > 0 \) such that the following holds. Consider the discrete cube \( \{0, 1\}^n \) endowed with the uniform measure. Let \( f : \{0, 1\}^n \to \{0, 1\} \) with \( \mathbb{E} f = p \). If, for all \( k \), \( I_k(f) \leq \delta \), then

\[
\sum_k I_k(f) \geq cp(1 - p) \log(1/\delta),
\]

and in particular, there exists \( k \) such that \( I_k(f) \geq cp(1 - p) \log(1/\delta)/n \).

The proof of Theorem 1.2 is a straightforward combination of Theorem 2.1, Proposition 2.2, and Theorem 3.1, as follows.

**Proof of Theorem 1.2.** Clearly, it is sufficient to prove the assertion for \( h(t) = \text{Ent}(t) \). Let \( f : \{0, 1\}^n \to \{0, 1\} \), such that \( \mathbb{E} f = p \). We want to find a lower bound on the \( h \)-influences of \( f \). By Theorem 2.1, we can replace \( f \) by a monotone function with the same expectation and lower \( h \)-influences, and thus we can assume without loss of generality that \( f \) is monotone. By Proposition 2.2, applied with \( \epsilon = c_0 p(1 - p) \log n/n \) (for some \( c_0 \leq 1 \) to be determined below), there exists \( g : \{0, 1\}^m \to \{0, 1\} \), such that \( |\mathbb{E}(g) - \mathbb{E}(f)| < \epsilon \), and for all \( i \), \( \sum_{j \in S_i} I_j(g) \leq 4(I_i^h(f) + \epsilon) \).

Assume that all the \( h \)-influences of \( f \) satisfy \( I_i^h(f) \leq \log n/n \). In this case, for all \( j \) we have \( I_j(g) \leq 8 \log n/n \), and thus, by Theorem 3.1, there exists \( 1 \leq j_0 \leq n \) such that

\[
I_{j_0}(g) \geq c_2 p(1 - p) \log n/n,
\]
for some constant $c_2$. Let $i_0$ be such that $j_0 \in S_{i_0}$. Since $I^{h}_{j_0}(f) \geq \frac{1}{4} \left( \sum_{j \in S_{i_0}} I_j(g) - 4\epsilon \right)$, choosing $c_0 = \min(1, c_2/8)$ yields the assertion (with $c = c_0$).

A slight modification of the proof yields the following generalization of Theorem 3.1 to $h$-influences, which will be used later.

**Proposition 3.2.** Let $h : [0, 1] \rightarrow [0, 1]$ be such that $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. There exists a constant $c > 0$ such that the following holds. Let $f : [0, 1]^n \rightarrow \{0, 1\}$ with $\mathbb{E}f = p$. If, for all $k$, $I^h_k(f) \leq \delta$, then

$$\sum_k I^h_k(f) \geq cp(1 - p) \log(1/\delta).$$

3.2. Generalization of Talagrand’s lower bound on the vector of influences

In [22], Talagrand proved the following strengthening of the KKL theorem for functions on the discrete cube endowed with a general product measure.

**Theorem 3.3 (Talagrand).** Consider the discrete cube $\{0, 1\}^n$ endowed with the product measure $\mu_q$ defined by $\mu_q(x) = q^{\sum x_i}(1 - q)^{n - \sum x_i}$. There exists a universal constant $K > 0$ (which does not depend on $q$) such that, for any function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ with $\mathbb{E}(f) = p$,

$$p(1 - p) \leq Kq(1 - q) \log \frac{2}{q(1 - q)} \sum_{i \leq n} \frac{I_i(f)}{\log \frac{1}{q(1 - q)I_i(f)}}. \tag{3.1}$$

The proof of Theorem 3.3 uses the biased Fourier–Walsh expansion for functions on the discrete cube, and a biased version of the Bonamie–Beckner hypercontractive inequality [5, 2], which is proved in [22].

Theorem 3.3 generalizes easily to $h$-influences.

**Proposition 3.4.** Let $h : [0, 1] \rightarrow [0, 1]$ be such that $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. There exists a constant $K > 0$ such that for any function $f : [0, 1]^n \rightarrow \{0, 1\}$ with $\mathbb{E}(f) = p$,

$$p(1 - p) \leq K \sum_{i \leq n} \frac{I^h_i(f)}{\log \frac{I^h_i(f)}{2}} \tag{3.2}$$

**Proof.** Clearly, it is sufficient to prove the assertion for $h(t) = \text{Ent}(t)$. Let $f : [0, 1]^n \rightarrow \{0, 1\}$. Without loss of generality, we may assume that $I^h_i(f) > 0$ for all $i$, since otherwise we can treat $f$ as a function defined on a lower-dimensional cube. As in the proof of Theorem 1.2, for any

\[ \text{(5.1)} \]

The biased hypercontractive inequality proved in [22, Lemma 2.1] yields a hypercontractivity constant of $q(1 - q)$. As was shown later in [20], the optimal constant is bigger (of order $\text{Ent}(q)$). However, this sub-optimality affects the assertion of Theorem 3.3 only in the constant factor $K$, which is not specified in [22].

Another generalization of Theorem 3.3 to the continuous case was obtained recently by Hatami [15], for a combination of Definitions 2 and 3 of the influences. Hatami’s result does not follow from Proposition 3.4, but also does not imply it.
\(\epsilon > 0\), monotonization and discretization allow us to replace \(f\) by a function \(g : \{0, 1\}^n \to \{0, 1\}\), such that \(|\mathbb{E}(g) - \mathbb{E}(f)| < \epsilon\), and for all \(i\), \(\sum_{j \in S_i} I_j(g) \leq 4(I_i^h(f) + \epsilon)\). Consider such a function \(g\) corresponding to

\[
\epsilon = \min \left( \frac{p}{2}, \frac{1-p}{2}, \min_{1 \leq i \leq n} [I_i^h(f)] \right).
\]

Applying Theorem 3.3 in the case \(q = 1/2\) to \(g\), we get

\[
\mathbb{E}(g)(1 - \mathbb{E}(g)) \leq K \sum_{1 \leq i \leq n} \sum_{j \in S_i} \frac{I_j(g)}{\log \frac{4}{I_j(g)}} \leq 4K \sum_{1 \leq i \leq n} \sum_{j \in S_i} \frac{I_j(g)}{\log \frac{16}{I_j(g)}}.
\]

Note that for all \(1 \leq i \leq n\),

\[
\sum_{j \in S_i} \frac{I_j(g)}{\log \frac{16}{I_j(g)}} \leq \sum_{j \in S_i} \frac{I_j(g)}{\log \frac{16}{\sum_{j \in S_i} I_j(g)}} = \sum_{j \in S_i} \frac{I_j(g)}{\log \frac{16}{\sum_{j \in S_i} I_j(g)}}.
\]

Since the function \(\varphi(x) = x / \log(16/x)\) is monotone increasing in \(x\) in \((0, 16)\), we get

\[
\sum_{1 \leq i \leq n} \sum_{j \in S_i} \frac{I_j(g)}{\log \frac{16}{I_j(g)}} \leq \sum_{1 \leq i \leq n} \sum_{j \in S_i} \frac{8I_i^h(f)}{\log \frac{2}{I_i^h(f)}},
\]

where the last inequality holds since \(\epsilon \leq \min_{1 \leq i \leq n}[I_i^h(f)]\). Finally, since \(\epsilon \leq \min(p/2, (1-p)/2)\), we get \(\mathbb{E}(g)(1 - \mathbb{E}(g)) \geq p(1-p)/2\), and thus substitution into (3.3) yields the assertion.

We note that Theorem 3.3 (for a general \(q\)) follows from Proposition 3.4, using a standard transformation from the biased measure on the discrete cube to the Lebesgue measure on the continuous cube. Indeed, consider the discrete cube \(\{0, 1\}^n\) endowed with the measure \(\mu_q\). Define \(G : [0, 1] \to \{0, 1\}\) by \(G(x) = 0\) if \(x \leq 1 - q\), and \(G(x) = 1\) if \(x > 1 - q\). For a function \(f : \{0, 1\}^n \to \{0, 1\}\), define \(\tilde{f}(x_1, x_2, \ldots, x_n) = f(G(x_1), G(x_2), \ldots, G(x_n))\).

It is easy to see that \(\mathbb{E}(\tilde{f}) = \mathbb{E}(f)\), where the expectation in the left-hand side is with respect to the Lebesgue measure on the continuous cube, and the expectation in the right-hand side is with respect to the measure \(\mu_q\) on the discrete cube. Applying Proposition 3.4 to the function \(\tilde{f}\), we get

\[
p(1-p) \leq K \sum_{1 \leq i \leq n} \frac{I_i^h(\tilde{f})}{\log \frac{2}{I_i^h(\tilde{f})}},
\]

where \(p = \mathbb{E}(\tilde{f}) = \mathbb{E}(f)\). Due to the construction of \(\tilde{f}\), the contribution of each non-constant fibre to an influence of \(\tilde{f}\) is \(h(q)\), and thus taking \(h(t) = \text{Ent}(t)\), we have (for all \(1 \leq i \leq n\))

\[
I_i^h(\tilde{f}) = \text{Ent}(q)I_i(f),
\]
where the influence in the right-hand side is with respect to the measure $\mu_q$ on the discrete cube. Substituting into (3.4), we get

$$p(1 - p) \leq K \text{Ent}(q) \sum_{1 \leq i \leq n} \frac{I_i(f)}{\log \frac{2}{\text{Ent}(q)I_i(f)}}.$$  (3.5)

Finally, there exist constants $K_1$ and $K_2$ such that, for all $q$ and all $I_i(f)$,

$$\text{Ent}(q) \leq K_1 q(1 - q) \log \frac{2}{q(1 - q)},$$

and

$$\log \frac{2}{\text{Ent}(q)I_i(f)} \geq K_2 \log \frac{1}{q(1 - q)I_i(f)}.$$

Therefore, the assertion of Theorem 3.3 follows from (3.5).

This shows that the biased hypercontractive inequality used in the proof in [22] can be replaced by the original hypercontractive inequality (i.e., the inequality for the uniform measure), combined with our argument presented above.

### 3.3. Functions with a low sum of influences

One of the most useful results concerning influences of variables on functions on the discrete cube is the following theorem, due to Friedgut [9], asserting that if the sum of influences is small then the function essentially depends on a few coordinates.

**Theorem 3.5 (Friedgut).** Consider the discrete cube $\{0, 1\}^n$ endowed with the uniform measure. Let $f : \{0, 1\}^n \to \{0, 1\}$ be such that $\sum I_i(f) = B$, and let $\epsilon > 0$. Denote $M = B/\epsilon$. There exists a Boolean function $g$ depending only on

$$\exp\left(2 + \sqrt{2 \log(4M) \frac{\log(4M)}{M}}\right)$$

variables, such that $P[f \neq g] \leq \epsilon$.

Dinur and Friedgut [10] observed that, using discretization, Theorem 3.5 can be generalized to monotone functions on the continuous cube.

**Theorem 3.6 (Dinur and Friedgut).** There exists a constant $c > 0$ such that the following holds. Let $\epsilon > 0$, and let $f : [0, 1]^n \to \{0, 1\}$ be a monotone function. If $\sum I_i(f) \leq B$, then there exists a set $J \subset \{1, 2, \ldots, n\}$ with $|J| \leq \exp(cB/\epsilon)$ and a function $g : [0, 1]^n \to \{0, 1\}$ depending only on the coordinates in $J$, such that $\|f - g\|_2^2 \leq \epsilon$.

Dinur and Friedgut [10] conjectured that the assertion of Theorem 3.6 holds even without the monotonicity assumption. This conjecture was disproved by Hatami [15]. On the other hand, Hatami proved (for general Boolean functions on the continuous cube) that if the sum of influences of $f$ is small, then $f$ can be approximated by a function having a decision tree of bounded depth (see [15] for the definitions).
Theorem 3.7 (Hatami). Let $f : [0, 1]^n \rightarrow \{0, 1\}$ satisfy $\sum_i I_i(f) \leq B$. Then, for every $\epsilon > 0$, there exists a function $g : [0, 1]^n \rightarrow \{0, 1\}$ such that $\|f - g\|_2^2 \leq \epsilon$, and $g$ has a decision tree of depth at most $\exp(cB/\epsilon^2)$, where $c$ is a universal constant.

Using the techniques presented above, we can generalize Theorems 3.6 and 3.7 to $h$-influences.

Proposition 3.8. Let $h : [0, 1] \rightarrow [0, 1]$ satisfy $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. There exists a constant $c > 0$ such that the following holds. Let $\epsilon > 0$, and let $f : [0, 1]^n \rightarrow \{0, 1\}$ be a monotone function. If $\sum_i I^h_i(f) \leq B$, then there exists a set $J \subset \{1, 2, \ldots, n\}$ with $|J| \leq \exp(cB/\epsilon)$ and a function $g : [0, 1]^n \rightarrow \{0, 1\}$ depending only on the coordinates in $J$, such that $\|f - g\|_2^2 \leq \epsilon$.

Proposition 3.8 follows immediately from Theorem 3.5, using Proposition 2.2. As in Section 3.2, we can apply Proposition 3.8 with $h(t) = \text{Ent}(t)$ to functions on the discrete cube endowed with the measure $\mu_q$, to get the following result.

Proposition 3.9. There is an absolute constant $c > 0$ such that the following holds. Consider the discrete cube $\{0, 1\}^n$ endowed with the measure $\mu_q$. If a monotone function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ satisfies $\sum_i I_i(f) \leq B$, then there exists $g : \{0, 1\}^n \rightarrow \{0, 1\}^n$, depending on at most $\exp(c\text{Ent}(q)B/\epsilon)$ coordinates, such that $\|f - g\|_2^2 \leq \epsilon$.

For monotone functions, Proposition 3.9 gives a more precise result than [9, Theorem 4.1], in which the exact dependence on $q$ was not specified.

The generalization of Theorem 3.7 is as follows.

Proposition 3.10. Let $h : [0, 1] \rightarrow [0, 1]$ be such that $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. Let $f : [0, 1]^n \rightarrow \{0, 1\}$ satisfy $\sum_i I^h_i(f) \leq B$. Then, for every $\epsilon > 0$, there exists a function $g : [0, 1]^n \rightarrow \{0, 1\}$ such that $\|f - g\|_2^2 \leq \epsilon$, and $g$ has a decision tree of depth at most $\exp(cB/\epsilon^2)$, where $c$ is a universal constant.

The proof of Proposition 3.10 is a minor modification of the proof of Theorem 3.7 (i.e., of Theorem 3.2 in [15]). The only two changes are replacing ordinary influences with $h$-influences throughout the proof, and replacing "Theorem B" used in the proof by the following proposition, which follows immediately from Proposition 3.2.

Proposition 3.11. Let $h : [0, 1] \rightarrow [0, 1]$ be such that $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. For any $f : [0, 1]^n \rightarrow \{0, 1\}$ with $\mathbb{E}(f) = p$, there exists $1 \leq i \leq n$, such that

$$I^h_i(f) \geq e^{-\frac{c}{p(1-p)} \sum_{j=1}^n I^h_j(f)},$$

where $c$ is a universal constant.

In Section 4 we show that for some cases of interest, the advantage of Proposition 3.8 over Theorem 3.6 is significant.
3.4. Relation between the sum of influences and the size of the boundary

**Notation.** In the following subsection we consider a geometric realization of Boolean functions on product spaces as characteristic functions of sets in the respective spaces. For a product space $X$ and a set $A \subset X$, we denote by $I_i(A)$ the $i$th influence of the characteristic function $1_A$, and similarly for $h$-influences.

**Definition 6.** Consider a product space $X = X_1 \times X_2 \times \cdots \times X_n$ endowed with a product measure $\mu = \mu_1 \otimes \cdots \otimes \mu_n$, and let $A \subset X$. The boundary of $A$ is

$$\partial A = \{ x \in A : \exists (1 \leq i \leq n), \text{ such that the function } 1_A \text{ is non-constant on the fibre } s_i(x) \}.$$  

Note that for the discrete cube, $\partial A$ is simply the vertex boundary of $A$. In [18], Margulis proved that for subsets of the discrete cube endowed with the uniform measure, the size of the boundary and the sum of influences cannot be small simultaneously. For monotone subsets of the discrete cube, Talagrand [21] gave the following precise form to this statement.

**Theorem 3.12 (Talagrand).** Consider the discrete cube $\{0,1\}^n$ endowed with the uniform measure $\mu$. For any monotone subset $A \subset \{0,1\}^n$ with $\mu(A) \leq 1/2$,

$$\mu(\partial A) \sum_i I_i(A) \geq c \mu(A)^2 \log \frac{e}{\mu(A)},$$

where $c$ is a universal constant.

Theorem 3.12 can be generalized to subsets of the continuous cube, as follows.

**Proposition 3.13.** Let $h : [0,1] \to [0,1]$ be a function satisfying $h(t) \geq \text{Ent}(t)$ for all $0 \leq t \leq 1$. Consider the continuous cube $[0,1]^n$ endowed with the Lebesgue measure $\lambda$. For any monotone subset $A \subset [0,1]^n$ with $\lambda(A) \leq 1/2$,

$$\lambda(\partial A) \sum_i I^h_i(A) \geq c \lambda(A)^2 \log \frac{e}{\lambda(A)}, \quad (3.6)$$

where $c$ is a universal constant.

**Proof.** Clearly, it is sufficient to prove the assertion for $h(t) = \text{Ent}(t)$. Let $A \subset [0,1]^n$ be a monotone set. As in the proof of Proposition 3.4, we can assume without loss of generality that $I^h_i(A) > 0$ for all $1 \leq i \leq n$. By Proposition 2.2, for any $\epsilon > 0$ there exists $l = l(A, \epsilon)$ and a function $g : \{0,1\}^l \to \{0,1\}$, such that

$$|E(g) - \lambda(A)| < \epsilon, \quad (3.7)$$

and for all $1 \leq i \leq n$,

$$\sum_{j \in S_i} I_j(g) \leq 4(I^\text{Ent}_i(A) + \epsilon). \quad (3.8)$$
Denote the subset of the discrete cube corresponding to $g$ by $B$. We would like to show that $g$ can be chosen such that

\[ \mu(\partial B) \leq \lambda(\partial A) + \epsilon, \quad (3.9) \]

where $\mu$ denotes the uniform measure on the discrete cube.

Recall that in the proof of Proposition 2.2, $g$ is constructed as the ‘discrete version’ of a function $\tilde{f} : [0, 1]^n \to \{0, 1\}$, which is constant on small subcubes. Denote the subset of $[0, 1]^n$ corresponding to $\tilde{f}$ by $B_0$. From the definition of the boundary, it is clear that for any two sets $A_1, A_2 \subset [0, 1]^n$,

\[ |\lambda(\partial A_1) - \lambda(\partial A_2)| \leq \sum_{i=1}^{n} (\lambda(A_{1,i} \setminus A_{2,i}) + \lambda(A_{2,i} \setminus A_{1,i})), \]

where

\[ A_{1,i} = \{ x \in A_1 : \text{the function } 1_{A_1} \text{ is non-constant on the fibre } s_i(x) \}, \]

and similarly for $A_{2,i}$. Thus, due to the monotonicity of $A$, $l$ and $\tilde{f}$ can be chosen such that in addition to the conditions (3.7) and (3.8), the following condition will be satisfied:

\[ |\lambda(\partial B_0) - \lambda(\partial A)| < \epsilon. \]

(See the proof of Lemma 4.53 in [12] for a detailed presentation of a similar argument.) Now consider the transition from $\tilde{f}$ to $g$. It is easy to see that

\[ \mu(\partial B) \leq \lambda(\partial B_0). \]

Indeed, if $x \in \partial B$ then there exists $j$ such that either ($x \in B$ and $x \oplus e_j \notin B$) or ($x \notin B$ and $x \oplus e_j \in B$). Let $i$ be such that $j \in S_i$. Recall that the pre-image of $x$ (in the discretization procedure) is a subcube of the continuous cube, denoted by $A_x$, such that $\lambda(A_x) = 2^{-nl}$. For all $y \in A_x$, the function $\tilde{f}$ is non-constant on the fibre $s_i(y)$, and hence $A_x \subset \partial B_0$. Thus, $\tilde{f}$ can be chosen such that (3.9) will also be satisfied.

Applying Theorem 3.12 to the set $B$, we get

\[ \mu(\partial B) \sum_{j} I_j(g) \geq c \mu(B)^2 \log \frac{e}{\mu(B)}. \quad (3.10) \]

Choosing

\[ \epsilon = \min \left( \frac{\lambda(A)}{2}, \frac{\lambda(\partial A)}{\lambda(\partial A)}, \min_{1 \leq i \leq n} [I_{\text{Ent}}^i(A)] \right), \]

and combining conditions (3.7), (3.8) and (3.9), the assertion of Proposition 3.13 (with the constant $c/64$) follows by substitution to (3.10).

We note that a stronger version of Theorem 3.12 was proved by Talagrand in [23]. It seems challenging to find a generalization of this version to the continuous cube.
4. Tightness of results

The tightness of most of our results can be shown using variants of the tribes function presented in [3]. We note that our function is the standard transformation to the continuous cube of the biased tribes function that was used to show the tightness of the results in [22] and [11]. In the construction below we assume that \( q \leq 1/2 \). The case \( q > 1/2 \) is treated at the end of this section.

4.1. Tightness of Theorem 1.2

Consider the discrete cube \( \{0,1\}^n \) endowed with the product measure \( \mu_q \). Partition the set \( \{1,2,\ldots,n\} \) into sets \( \{T_1, T_2, \ldots, T_{n/r}\} \) of size

\[
r = \left\lfloor \frac{\log n - \log \log n + \log \log(1/q)}{\log(1/q)} \right\rfloor
\]

each,\(^7\) and define \( f : \{0,1\}^n \rightarrow \{0,1\} \) by setting \( f(x) = 1 \) if and only if there exists \( i \) such that \( x_j = 1 \) for all \( j \in T_i \). This function can be transformed to a function on the continuous cube as follows. Define \( G : [0,1] \rightarrow \{0,1\} \) by \( G(x) = 0 \) if \( x \leq 1 - q \), and \( G(x) = 1 \) if \( x > 1 - q \), and let \( \tilde{f} : [0,1]^n \rightarrow \{0,1\} \) be defined by

\[
\tilde{f}(x_1, x_2, \ldots, x_n) = f(G(x_1), G(x_2), \ldots, G(x_n)).
\]

It is easy to see that

\[
\mathbb{E}\tilde{f} = 1 - (1 - q^n)^{n/r} \approx 1 - 1/e
\]

where the expectation is with respect to the Lebesgue measure on the continuous cube. The function \( \tilde{f} \) is non-constant on a fibre \( s_k(x) \), where \( k \in T_i \) if, for all \( j \in T_i \), we have \( x_j > 1 - q \), and for each \( i' \neq i \), there exists \( j \in T_{i'} \) such that \( x_j \leq 1 - q \). In each of the non-constant fibres we have \( \mathbb{E}(\tilde{f}_k) = q \). Thus, for all \( h : [0,1] \rightarrow [0,1] \) and for all \( 1 \leq k \leq n \),

\[
I^h_k(\tilde{f}) = q^{r-1}(1 - q^r)^{(n/r)-1} h(q) \leq \frac{2\log n}{(en)\log(1/q)} h(q) \leq \frac{2\log n}{(en)\text{Ent}(q)} h(q),
\]

where the last inequality holds since we have \( q \log(1/q) \geq (1 - q) \log(1/(1 - q)) \) for all \( q \leq 1/2 \). This example proves Proposition 1.3 (for \( q \) fixed and \( n \) large enough, as asserted in the proposition), thus showing that the assertion of Theorem 1.2 is tight.

4.2. Tightness of Proposition 3.4

The same example shows the tightness of Proposition 3.4. Indeed, if there exists \( q \) such that \( h(q) \leq \epsilon \text{Ent}(q) \), then for the corresponding function \( \tilde{f} \) we have

\[
I^h_k(\tilde{f}) \leq \frac{2\epsilon \log n}{en},
\]

\(^7\) If \( n \) is not divisible by \( r \), the division into sets is applied to the first \( n - (n \mod r) \) coordinates, and the other coordinates simply do not influence the output of the function. It is clear that for \( n \) big enough (as a function of \( q \)), as in the assertion of Proposition 1.3, this modification affects only the constant in the assertion.
for all $1 \leq k \leq n$. Since the function $g(x) = x/\log(2/x)$ is monotone increasing in $x$, we have
\[ \sum_{k=1}^{n} \frac{I_k^h(\tilde{f})}{\log \frac{2}{I_k^h(\tilde{f})}} = \sum_{k=1}^{n} g(I_k^h(\tilde{f})) \leq n g \left( \frac{2e \log n}{en} \right) \approx e \frac{\log n}{\log(n/e)} \leq ce. \]

For $\epsilon$ sufficiently small, this contradicts (3.2) since for $\tilde{f}$, the left-hand side of the inequality is $\Theta(1)$.

4.3. Tightness of Proposition 3.8

In order to show the advantage of Proposition 3.8 over Theorem 3.6, we consider a slight modification of the tribes function examined above. For $0 < q \leq 1/2$ and for $m < n$, we construct the tribes function $\tilde{f} : [0, 1]^m \rightarrow \{0, 1\}$, as above. Then we define $\tilde{f} : [0, 1]^n \rightarrow \{0, 1\}$ by $\tilde{f}(x_1, x_2, \ldots, x_n) = \tilde{f}(x_1, x_2, \ldots, x_m)$. Clearly, $\tilde{f}$ depends on $m$ variables. By the calculation presented above, we have
\[ \sum_k I_k^h(\tilde{f}) = \sum_k I_k^h(\tilde{f}) \leq \frac{2 \log m}{e \text{Ent}(q)} h(q), \]

and thus, for $h(t) = \text{Ent}(t)$, we get $\sum_k I_k^h(\tilde{f}) \leq 2 \log m/e$. Therefore, Proposition 3.8 implies that $\tilde{f}$ can be approximated by a function depending on at most $\exp(c \log m/e)$ variables, which is tight, up to the $(c/e)$ factor in the exponent. For comparison, note that $\sum_k I_k(\tilde{f}) \approx m/e$, and hence Theorem 3.6 implies approximation by a function depending on $\exp(cm/e)$ variables, which is far from being tight.

4.4. Tightness of Proposition 3.13

Finally, in order to show the tightness of Proposition 3.13, we consider a balanced threshold function on the biased discrete cube. Consider the discrete cube $\{0, 1\}^n$ endowed with the measure $\mu_q$, and let
\[ A = \left\{ x \in \{0, 1\}^n : \sum_i x_i > \lfloor nq \rfloor \right\}. \]

Using the function $G$ defined above, the set $A$ can be transformed to $\tilde{A} \subset [0, 1]^n$. It is well known that $\lambda(\tilde{A}) = \mu_q(A) = \Theta(1)$. We have
\[ \partial A = \left\{ x \in \{0, 1\}^n : \sum_i x_i = \lfloor nq \rfloor + 1 \right\}, \]

and hence it can be shown using Stirling’s formula that
\[ \lambda(\partial \tilde{A}) = \mu_q(\partial A) \approx \frac{1}{\sqrt{2\pi nq(1-q)}}. \]

The function $1_{\tilde{A}}$ can be non-constant on the fibre $s_1(x_1, x_2, \ldots, x_n)$ only if
\[ \lfloor nq \rfloor \leq \sum_{j=1}^{n} G(x_j) \leq \lfloor nq \rfloor + 1, \]
and the expectation of the function on each non-constant fibre is $q$. Thus, it follows from Stirling’s formula that

$$
\sum_{i=1}^{n} I^h_i(\tilde{A}) \leq \frac{2n}{\sqrt{2\pi n q(1-q)}} h(q).
$$

Taking $h(q) = \text{Ent}(q)$, we get

$$
\lambda(\partial \tilde{A}) \sum_{i} I^h_i(\tilde{A}) \leq \frac{1}{\sqrt{2\pi n q(1-q)}} \frac{2n}{\sqrt{2\pi n q(1-q)}} \text{Ent}(q) = \frac{2\text{Ent}(q)}{2\pi q(1-q)} \leq \frac{4\log(1/q)}{\pi}.
$$

Since for $\tilde{A}$, the right-hand side of (3.6) is $\Theta(1)$, this implies that Proposition 3.13 is tight, up to a multiplicative factor of $(\log(1/q))$.

### 4.5. The case $q > 1/2$

In order to show the tightness for $q > 1/2$, we use the dual function.

**Definition 7.** For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the dual function $\bar{f} : \{0,1\}^n \to \{0,1\}$ is defined by

$$
\bar{f}(x_1, x_2, \ldots, x_n) = 1 - f(1-x_1, 1-x_2, \ldots, 1-x_n).
$$

It is easy to see that for all $0 < q < 1$,

$$
\mathbb{E}_q(f) = 1 - \mathbb{E}_{1-q}(\bar{f}),
$$

where $\mathbb{E}_q(\cdot)$ denotes expectation with respect to the measure $\mu_q$. Similarly, it is easy to see that for all $1 \leq i \leq n$,

$$
I^q_i(f) = I^{1-q}_i(\bar{f}),
$$

where $I^q_i(f)$ denotes influence with respect to the measure $\mu_q$. Therefore, the dual functions of the functions described above show the tightness of our results for $q > 1/2$.

### Acknowledgements

It is a pleasure to thank Benjy Weiss for helping to prove Theorem 2.1, and Gil Kalai for numerous fruitful discussions. We are grateful to Geoffrey Grimmett, to Micha Perles, to Orr Dunkelman, and to the anonymous referee, for valuable suggestions.

### References

[1] Bauer, H. (2001) *Measure and Integration Theory*, Walter de Gruyter.

[2] Beckner, W. (1975) Inequalities in Fourier analysis. *Ann. of Math.* **102** 159–182.

[3] Ben-Or, M. and Linial, N. (1990) Collective coin flipping. In *Randomness and Computation* (S. Micali, ed.), Academic Press, pp. 91–115.

[4] Bollobás, B. and Riordan, O. (2006) The critical probability for random Voronoi percolation on the plane is 1/2. *Probab. Theory Rel. Fields* **136** 417–468.
[5] Bonamie, A. (1970) Etude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier* **20** 335–402.

[6] Bourgain, J., Kahn, J., Kalai, G., Katznelson, Y. and Linial, N. (1992) The influence of variables in product spaces. *Israel J. Math.* **77** 55–64.

[7] Dinur, I., Friedgut, E., Kindler, G. and O’Donnell, R. (2007) On the Fourier tails of bounded functions over the discrete cube. *Israel J. Math.* **160** 389–412.

[8] Frankl, P. (1987) The shifting technique in extremal set theory. In *Surveys in Combinatorics* (C. W. Whitehead, ed.), Cambridge University Press, pp. 81–110.

[9] Friedgut, E. (1998) Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica* **18** 27–35.

[10] Friedgut, E. (2004) Influences in product spaces: KKL and BKKKL revisited. *Combin. Probab. Comput.* **13** 17–29.

[11] Friedgut, E. and Kalai, G. (1996) Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.* **124** 2993–3002.

[12] Grimmett, G. R. (2010) *Probability on Graphs*, Cambridge University Press, to appear. Draft version available online at: http://www.statslab.cam.ac.uk/~grg/preprints.html.

[13] Grimmett, G. R. and Graham, B. (2006) Influence and sharp-threshold theorems for monotonic measures. *Ann. Probab.* **34** 1726–1745.

[14] Hart, S. (1976) A note on the edges of the $n$-cube. *Discrete Math.* **14** 157–163.

[15] Hatami, H. (2009) Decision trees and influence of variables over product probability spaces. *Combin. Probab. Comput.* **18** 357–369.

[16] Kahn, J., Kalai, G. and Linial, N. (1988) The influence of variables on Boolean functions. In *Proc. 29th Annual Symposium on Foundations of Computer Science* (FOCS), Computer Society Press, pp. 68–80.

[17] Keller, N., Mossel, E. and Sen, A. (2009) Geometric influences. Preprint, available online at: http://arxiv.org/abs/0911.1601.

[18] Margulis, G. A. (1977) Probabilistic characteristics of graphs with large connectivity. *Problems Info. Trans.* **10** 174–179.

[19] Mossel, E., O’Donnell, R. and Oleszkiewicz, K. (2010) Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.* **171**(1) 295–341.

[20] Oleszkiewicz, K. (2003) On the non-symmetric Khinchine–Kahane inequality. In *Stochastic Inequalities and Applications* (E. Giné, C. Houdré and D. Nualart, eds), Birkhäuser, pp. 157–168.

[21] Talagrand, M. (1993) Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis’ graph connectivity theorem. *Geom. Funct. Anal.* **3** 295–314.

[22] Talagrand, M. (1994) On Russo’s approximate zero-one law. *Ann. Probab.* **22** 1576–1587.

[23] Talagrand, M. (1997) On boundaries and influences. *Combinatorica* **17** 275–285.