Linear Transformations of Monotone Functions on the Discrete Cube

Nathan Keller and Haran Pilpel
Einstein Institute of Mathematics, Hebrew University
Jerusalem 91904, Israel
{nkeller,haranp}@math.huji.ac.il

July 17, 2008

Abstract

For a function $f : \{0,1\}^n \to \mathbb{R}$ and an invertible linear transformation $L \in GL_n(2)$, we consider the function $Lf : \{0,1\}^n \to \mathbb{R}$ defined by $Lf(x) = f(Lx)$. We raise two conjectures: First, we conjecture that if f is Boolean and monotone then $I(Lf) \geq I(f)$, where $I(f)$ is the total influence of f. Second, we conjecture that if both f and $L(f)$ are monotone, then $f = L(f)$ (up to a permutation of the coordinates). We prove the second conjecture in the case where L is upper triangular.

1 Introduction

Definition 1 *A function $f : \{0,1\}^n \to \mathbb{R}$ is monotone if for all $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$,*

$$(\forall i : x_i \leq y_i) \Rightarrow (f(x) \leq f(y)).$$

Monotone functions on the discrete cube, and especially monotone Boolean functions, were intensively investigated over the last decades (see, for example, [9]). Despite the extensive research, the structure of such functions is far from being understood. For example, there is no simple criterion to determine whether a Boolean function is monotone or not.

In this paper we raise two conjectures regarding the application of linear transformations to monotone functions on the discrete cube. The first conjecture deals with the *total influence* of the function.

Definition 2 *For a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and for $1 \leq i \leq n$, the influence of the i-th coordinate on f is*

$$I_i(f) = \mu(\{(x_1, \ldots, x_n) | f(x_1, \ldots, x_n) \neq f(x_1, \ldots, x_{i-1}, 1 - x_i, x_{i+1}, \ldots, x_n)\}),$$

where μ is the uniform measure on the discrete cube. The total influence of f is

$$I(f) = \sum_{i=1}^{n} I_i(f).$$

Influences of Boolean functions have been extensively studied and have applications in numerous fields, including Combinatorics, Theoretical Computer Science, Statistical Physics, Social Choice Theory, etc. (see, for example, the survey article [7]). Functions with a low total influence are of special importance, since they essentially depend on a small number of coordinates [4]. Our first conjecture asserts that the total influence of a monotone Boolean function cannot be reduced by applying a linear transformation.

Conjecture 1 *If f is Boolean and monotone, then $I(f) \leq I(Lf)$, for all $L \in GL_n(2)$.*

We have verified the conjecture for all the Boolean functions on $n \leq 5$ variables.

The second conjecture deals with general monotone functions on the discrete cube.

Conjecture 2 *If both f and Lf are monotone, then $f = Lf$, up to a permutation of the coordinates.*

We have verified Conjecture 2 for all monotone Boolean functions on $n \leq 5$ variables. Further, we have proved the following particular case:

Theorem 1 *If $L \in GL_n(2)$ is an upper triangular matrix, and both f and Lf are monotone, then $f = Lf$.*

We present the proof of Theorem 1 in Section 2. In Section 3 we discuss the relation of Conjecture 1 to the Entropy-Influence conjecture [5] and consider related questions dealing with other properties of the Fourier-Walsh expansions of the functions f and Lf.

2 Proof of Theorem 1

The proof is by inverse induction on the number of coordinates L preserves (i.e., the number of i's such that $(Lx)_i = x_i$ for all $x \in \{0,1\}^n$). If L preserves all the coordinates, then L is the identity matrix. Assume now the claim for any L that preserves at least $n-k+1$ coordinates, and let L_0 be a transformation that preserves $n-k$ coordinates. Without loss of generality we can assume that L_0 preserves the last $n-k$ coordinates. We want to show that for all $(x_1, \ldots, x_n)$, we have $L_0 f(x) = L_1 f(x)$, where L_1 is identical to L_0 except for the k-th row, and in the k-th row, L_1 is equal to the identity matrix. Then, by the induction hypothesis, we have $L_0 f = L_1 f = f$.

We divide the discrete cube into subcubes according to the values of the last $n-k$ coordinates. That is, for every $(v_{k+1},\ldots,v_n)\in\{0,1\}^{n-k}$ we define

$$W=W(v_{k+1},\ldots,v_n)=\{x\in\{0,1\}^n|x_j=v_j,\forall j\geq k+1\}.$$

By the assumption, each $W=W(v_{k+1},\ldots,v_n)$ is invariant under L_0 and L_1. Hence, it is sufficient to prove the assertion for each $W=W(v_{k+1},\ldots,v_n)$ separately. From now on we fix W, and leave $(v_{k+1},\ldots,v_n)$ implicit.
The proof is based on two observations:

1. If $h:\{0,1\}^k\to\mathbb{R}$ is monotone, then for all $(x_1,\ldots,x_{k-1})\in\{0,1\}^{k-1}$ we have $h(x_1,\ldots,x_{k-1},0)\leq h(x_1,\ldots,x_{k-1},1)$. Thus,

$$\sum_{(x_1,\ldots,x_{k-1})\in\{0,1\}^{k-1}} h(x_1,\ldots,x_{k-1},0)\leq \sum_{(x_1,\ldots,x_{k-1})\in\{0,1\}^{k-1}} h(x_1,\ldots,x_{k-1},1).$$

   Moreover, equality holds if and only if $h(x_1,\ldots,x_{k-1},0)=h(x_1,\ldots,x_{k-1},1)$ for all $(x_1,\ldots,x_{k-1})\in\{0,1\}^{k-1}$, or equivalently, if the k-th coordinate does not influence the output of h.

2. Since L_0 and L_1 are identical except for the k-th row, the values L_0x and L_1x can differ only in the k-th coordinate. Since L_0 is upper triangular and since the values $(x_{k+1},\ldots,x_n)$ are fixed for all $x\in W$, we have either $(L_0x)_k=x_k$ for all $x\in W$ or $(L_0x)_k=1-x_k$ for all $x\in W$. If $(L_0x)_k=x_k$ for all $x\in W$ then $L_0f=L_1f$ on W, as asserted. Hence, we may assume that $(L_0x)_k=1-x_k$ for all $x\in W$.

Define
$$f'(x_1,\ldots,x_k)=f(x_1,\ldots,x_k,v_{k+1},\ldots,v_n),$$
and
$$L_0'(x_1,\ldots,x_k)=(y_1,\ldots,y_k),$$
where $(y_1,\ldots,y_n)=L_0(x_1,\ldots,x_k,v_{k+1},\ldots,v_n)$. Note that f' and $L_0'(f')$ are the restrictions of f and L_0f to W. In particular, since L_0 preserves the last $n-k$ coordinates, the monotonicity of f and L_0f implies that both f' and $L_0'(f')$ are monotone.
Define the sets

$$S_0=\{x\in\{0,1\}^k|x_k=0\},\qquad S_1=\{x\in\{0,1\}^k|x_k=1\}.$$

Since f' is monotone, by Observation 1 we have

$$\sum_{x\in S_0} f'(x)\leq\sum_{x\in S_1} f'(x). \tag{1}$$

Similarly, since $L_0'(f')$ is monotone,

$$\sum_{x\in S_0} f'(L_0'x)\leq\sum_{x\in S_1} f'(L_0'x). \tag{2}$$

By Observation 2, $L'_0(S_0) \subseteq S_1$ and $L'_0(S_1) \subseteq S_0$. Since L'_0 is the restriction to W of L_0 which is invertible, it follows that its restrictions $L'_0 : S_0 \to S_1$ and $L'_0 : S_1 \to S_0$ are injective and surjective. Hence, Equation 2 is equivalent to

$$\sum_{x \in S_1} f'(x) \leq \sum_{x \in S_0} f'(x). \tag{3}$$

Combining Equations 1 and 3 we get

$$\sum_{x \in S_0} f'(x) = \sum_{x \in S_1} f'(x). \tag{4}$$

By Observation 1, Equation 4 implies that the k-th coordinate does not influence the output of f'. Therefore, $L_1 f(x) = L_0 f(x)$ for all $x \in W$ (since $L_0 x$ and $L_1 x$ differ only in the k-th coordinate that does not affect the output of f for $x \in W$). This completes the proof of the theorem.

Remark 1 *We note that the equality $f = Lf$ does not imply that L is the identity matrix. For example, if $f = maj_5$, the majority function on 5 variables, and L is the linear transformation which preserves the first four variables and replaces the fifth with the XOR of all five, then it is easy to check that $Lf = f$.*

3 Related Questions

We conclude the paper with several questions related to Conjecture 1.

3.1 Other Analytic Properties of Functions on the Discrete Cube

In Conjecture 1, the total influence can be replaced by other analytic properties of the function, i.e., other properties of its Fourier-Walsh expansion.

Definition 3 *Let $f : \{0,1\}^n \to \mathbb{R}$. The Fourier-Walsh expansion of f is*

$$f(S) = \sum_{T \in \{0,1\}^n} \hat{f}(T) u_T(S),$$

where elements of $\{0,1\}^n$ are identified with subsets of $\{1, 2, \ldots, n\}$, the characters are $u_T(S) = (-1)^{|S \cap T|}$, and

$$\hat{f}(T) = \langle f, u_T \rangle = \frac{1}{2^n} \sum_{T' \in \{0,1\}^n} f(T') u_T(T').$$

As was first observed in [6], the total influence is given by the formula

$$I(f) = 4 \sum_{S \in \{0,1\}^n} |S| \hat{f}(S)^2.$$

It can be shown, using the linearity of the Fourier-Walsh expansion, that the Fourier-Walsh coefficients of Lf satisfy

$$\widehat{Lf}(S) = \hat{f}((L^T)^{-1}(S)).$$

That is, the Fourier-Walsh coefficients of Lf are a permutation of the Fourier-Walsh coefficients of f, defined by the matrix $(L^T)^{-1}$. Hence, Conjecture 1 asserts, qualitatively, that if f is monotone then the *level* of the Fourier-Walsh coefficients of f is lower than the level of the coefficients of Lf (where the level of the coefficient $\hat{f}(S)$ is $|S|$).

In the same spirit, one can raise the following question:

Question 1 *If f is Boolean and monotone, $\alpha > 0$, and $0 < \epsilon < 1$, is it true that:*

1.

$$\sum_{S \in \{0,1\}^n} |S|^\alpha \hat{f}(S)^2 \leq \sum_{S \in \{0,1\}^n} |S|^\alpha \widehat{Lf}(S)^2$$

2.

$$\sum_{S \in \{0,1\}^n} \epsilon^{|S|} \hat{f}(S)^2 \geq \sum_{S \in \{0,1\}^n} \epsilon^{|S|} \widehat{Lf}(S)^2$$

The expression $N_\epsilon(f) = \sum_{S \in \{0,1\}^n} \epsilon^{|S|} \hat{f}(S)^2$ represents the *noise sensitivity* of f (see [2]). The higher is $N_\epsilon(f)$, the less is f sensitive to a small change of the coordinates. Hence, Part 2 of the question asks, qualitatively, whether it is true that f is less sensitive to noise than Lf, for all L.

Remark 2 *We note that the similar question: Is it true that if f is Boolean and monotone then for all k,*

$$\sum_{|S| \leq k} \hat{f}(S)^2 \geq \sum_{|S| \leq k} \widehat{Lf}(S)^2$$

has a negative answer. The counterexample is the majority function on $2m+1$ coordinates, with $k = 2m$.

3.2 Relation to the Entropy-Influence Conjecture

Definition 4 *The spectral entropy of a Boolean function f is defined by*

$$E(f) = \sum_{S \in \{0,1\}^n} \hat{f}(S)^2 \log_2 \frac{1}{\hat{f}(S)^2}.$$

The Entropy-Influence conjecture [5] asserts the following:

Conjecture 3 *There exists an absolute constant C such that for every Boolean function f, $E(f) \leq CI(f)$.*

There exist classes of monotone functions for which the inequality in the conjecture is tight, i.e., there exists a universal constant C' such that $I(f) \leq C'E(f)$ for functions in the class. An example of such a class is the *tribes* functions, introduced in [1]. For such functions, applying the Entropy-Influence conjecture to the function Lf and using the fact that $E(f) = E(Lf)$ for all $L \in GL_n(2)$, we get that there exists a universal constant C'' such that $I(f) \leq C''I(Lf)$ for all $L \in GL_n(2)$.

Hence, a weaker form of Conjecture 1 for several classes of functions follows from the Entropy-Influence conjecture.

3.3 Non-Monotone Functions

If the function f is not monotone, then $I(Lf)$ can be much smaller than $I(f)$. For example, let $f(x_1, \ldots, x_n) = x_1 \oplus \ldots \oplus x_n$ be the parity function, and let L be the linear transformation that represents the change of coordinates: $y_1 = x_1 \oplus \ldots \oplus x_n$, $y_2 = x_2, \ldots, y_n = x_n$. Then $Lf(x) = x_1$ is the dictatorship function. While $I(f) = n$ is the maximal possible total influence amongst Boolean functions on n variables, $I(Lf) = 1$ is, by the Edge Isoperimetric Inequality (see [3], Theorem 16.2), the minimal possible total influence amongst balanced Boolean functions.

For the parity function, the only non-constant Fourier-Walsh coefficient of f is on the n-th level, and the only non-constant coefficient of Lf is on the first level. This "level reduction" can be generalized:

Proposition 1 *Let $f : \{0, 1\} \to \mathbb{R}$ be a function such that all the Fourier-Walsh coefficients of f are concentrated on the k lowest and the k highest levels. Let L be the matrix representing the change of coordinates*

$$y_1 = x_1 \oplus x_2, y_2 = x_2 \oplus x_3, \ldots, y_{n-1} = x_{n-1} \oplus x_n, y_n = x_n.$$

Then the Fourier-Walsh coefficients of Lf are concentrated on the $2k+1$ lowest levels.

It was shown in [8] that there exists a universal constant C such that if f is a Boolean function all whose Fourier-Walsh coefficients are concentrated on the k lowest levels, then f depends on at most C^k coordinates. Hence, Proposition 1 implies that if all the Fourier-Walsh coefficients of f are on the k highest and the k lowest levels, then Lf depends on at most C^{2k+1} coordinates.

References

[1] M. Ben-Or and N. Linial, Collective Coin Flipping, in: **Randomness and Computation** (S. Micali Ed.), Academic Press, New York, 1989, pp. 91–115.

[2] I. Benjamini, G. Kalai, and O. Schramm, Noise Sensitivity of Boolean Functions and Applications to Percolation, *Inst. Hautes Etudes Sci. Publ. Math.* **90** (1999), pp. 5–43.

[3] B. Bollobas, **Combinatorics**, Cambridge University Press, 1986.

[4] E. Friedgut, Boolean Functions with Low Average Sensitivity Depend on Few Coordinates, *Combinatorica* **18** (1998), no. 1, pp. 27–35.

[5] E. Friedgut and G. Kalai, Every Monotone Graph Property Has a Sharp Threshold, *Proc. Amer. Math. Soc.* **124** (1996), no. 10, pp. 2993–3002.

[6] J. Kahn, G. Kalai, and N. Linial, The Influence of Variables on Boolean Functions, Proc. 29-th Annual Symposium on Foundations of Computer Science (1988), pp. 68–80.

[7] G. Kalai and S. Safra, Threshold Phenomena and Influence: Perspectives from Mathematics, Computer Science, and Economics, in: **Computational Complexity and Statistical Physics** (A.G. Percus, G. Istrate and C. Moore, eds.), St. Fe Inst. Stud. Sci. Complex., Oxford Univ. Press, New York, 2006, pp. 25–60.

[8] G. Kindler and M. Safra, Noise Insensitive Boolean Functions are Juntas, manuscript.

[9] A.D. Korshunov, Monotone Boolean Functions, *Russian Math. Surveys* **58** (2003), no. 5, pp. 929–1001.