

On the noise sensitivity of monotone functions

Elchanan Mossel
Hebrew University of Jerusalem
and Microsoft Research
mossel@microsoft.com

Ryan O'Donnell
MIT Mathematics Department
odonnell@theory.lcs.mit.edu

May 28, 2003

Abstract

It is known that for all monotone functions $f : \{0, 1\}^n \rightarrow \{0, 1\}$, if $x \in \{0, 1\}^n$ is chosen uniformly at random and y is obtained from x by flipping each of the bits of x independently with probability $\epsilon = n^{-\alpha}$, then $\mathbf{P}[f(x) \neq f(y)] < cn^{-\alpha+1/2}$, for some $c > 0$.

Previously, the best construction of monotone functions satisfying $\mathbf{P}[f_n(x) \neq f_n(y)] \geq \delta$, where $0 < \delta < 1/2$, required $\epsilon \geq c(\delta)n^{-\alpha}$, where $\alpha = 1 - \ln 2 / \ln 3 = 0.36907\dots$, and $c(\delta) > 0$. We improve this result by achieving for every $0 < \delta < 1/2$, $\mathbf{P}[f_n(x) \neq f_n(y)] \geq \delta$, with:

- $\epsilon = c(\delta)n^{-\alpha}$ for any $\alpha < 1/2$, using the recursive majority function with arity $k = k(\alpha)$;
- $\epsilon = c(\delta)n^{-1/2} \log^t n$ for $t = \log_2 \sqrt{\pi/2} = .3257\dots$, using an explicit recursive majority function with increasing arities; and,
- $\epsilon = c(\delta)n^{-1/2}$, non-constructively, following a probabilistic CNF construction due to Talagrand.

We also study the problem of achieving the best dependence on δ in the case that the noise rate ϵ is at least a small constant; the results we obtain are tight to within logarithmic factors.

1 Introduction

1.1 Noise sensitivity and Fourier coefficients

The papers [KKL88, BL90] suggested the importance of the *Fourier expansion* and the *influence of variables* for the study of boolean functions. The ideas developed in these papers proved to be extremely fruitful in later work; see, e.g., [LMN93, FK96, F98, BKS99].

Let $\Omega_n = \{-1, +1\}^n$ be the Hamming cube endowed with the uniform probability measure $\mathbf{P}$. We consider boolean functions $f : \Omega_n \rightarrow \{-1, +1\}$. For $-1 \leq \eta \leq 1$ and $x \in \Omega_n$, define $N_\eta(x)$ to be a random element y of Ω_n which satisfies $\mathbf{E}[y_i x_i] = \eta$ (equivalently, $\mathbf{P}[x_i \neq y_i] = (1 - \eta)/2$), independently for all i . We will mostly be interested in the case $0 \leq \eta \leq 1$, i.e., in the case that y and x are correlated. It is natural to measure how stable f is under η -noise by the correlation between $f(x)$ and $f(N_\eta(x))$,

$$Z(f, \eta) := \mathbf{E}[f(N_\eta(x))f(x)] = 1 - 2\mathbf{P}[f(N_\eta(x)) \neq f(x)]. \quad (1)$$

For $0 < \eta < 1$ we expect that if f is stable under the noise operator N_η , then typically $f(x)$ and $f(N_\eta(x))$ should have the same value for and therefore $Z(f, \eta)$, the expression in (1), should be close to 1; if f is sensitive to noise, then $Z(f, \eta)$ should be close to 0.

The space Ω_n with the uniform probability measure naturally gives rise to an inner product space on all functions $f : \Omega_n \rightarrow \mathbb{R}$:

$$\langle f, g \rangle = \mathbf{E}[fg] = 2^{-n} \sum_{x \in \Omega_n} f(x)g(x).$$

For a set $S \subseteq [n]$, define $u_S(x) = \prod_{i \in S} x_i$. Since $u_S u_{S'} = u_{S \Delta S'}$, where Δ denotes symmetric difference, it follows that $(u_S)_{S \subseteq [n]}$ is an orthonormal basis. We call $\hat{f}(S) = \langle u_S, f \rangle$ the S *Fourier coefficient* of f , and $f = \sum_{S \subseteq [n]} \hat{f}(S) u_S$ the *Fourier expansion* of f .

The basis $(u_S)_{S \subseteq [n]}$ has very nice properties with respect to the noise operator; most notably, for all x and S , $\mathbf{E}[u_S(N_\eta(x))] = \eta^{|S|} u_S(x)$, which implies

$$Z(f, \eta) = \mathbf{E}[f(N_\eta(x))f(x)] = \sum_{S \subseteq [n]} \eta^{|S|} \hat{f}^2(S) \quad (2)$$

(see e.g. [BKS99, BJT99, O02]).

The stability of the function f under noise, $Z(f, \eta)$, is therefore closely related to how much of the ℓ_2 mass of the Fourier coefficients of f lies on coefficients $\hat{f}(S)$ for large sets S .

In addition to the sum in (2), it is common to study several other weighted sums of f 's squared Fourier coefficients. By Parseval's identity, $\sum_S \hat{f}^2(S) = 1$. The *average sensitivity* of f is defined by $I(f) = \sum_{k=1}^n I_k(f)$, where $I_k(f)$ is the probability that flipping the k th bit of a random input to f flips f 's value; it is shown in [KKL88] that $I(f) = \sum_S |S| \hat{f}^2(S)$. Note also that if f is monotone (we define monotonicity shortly), then $I_k(f) = |\hat{f}(\{k\})|$. Finally, we have the quantity $II(f) := \sum_{k=1}^n I_k^2(f)$, introduced in [BKS99].

1.2 Sensitivity of monotone functions

By any measure, the parity function $f = u_{[n]} = \oplus$ is the n -bit boolean function which is most sensitive to noise: $Z(f, \eta) = \eta^n$ is minimal, and $I(f) = n$ is maximal. In this paper we investigate the problem of extremal noise sensitivity when restricted to n -bit *monotone* boolean functions. Recall that a function f is *monotone* if for all $x, y \in \Omega_n$ we have $f(x) \leq f(y)$ whenever $x \leq y$ (in the sense $x_i \leq y_i$ for all i).

There are several reasons why it is interesting and natural to study the noise sensitivity of monotone functions. The first is that in some sense they are the “least noise sensitive” boolean functions. Benjamini, Kalai, and Schramm [BKS99] showed that applying combinatorial down-shifting (in the sense of Kleitman [K66]) to f only increases $Z(f, \eta)$, and repeated down-shifting eventually produces a monotone function.

A second reason for studying monotone functions is that they play an important role in the study of the influences of variables on boolean functions; see, e.g., [KKL88, T94, T96, FK96, T97, BKS99].

Finally, there are some well-known bounds on how noise-sensitive monotone functions can be. We now recall these bounds; for the remainder of the paper we will seek constructions which match or nearly match these bounds.

It is known that the majority function has maximal I among all monotone functions on n inputs (this is a simple isoperimetric inequality; for one proof, see Lemma 6.1 of [FK96]). Since this average sensitivity, I_n , is easily computed to be

$$I_n = (\sqrt{2/\pi} + o(1))\sqrt{n}, \quad (3)$$

we get that for all monotone f on n inputs,

$$I(f) \leq (\sqrt{2/\pi} + o(1))\sqrt{n}. \quad (4)$$

It remains to determine how small $Z_\eta(f)$ can be for monotone functions. A natural goal is to find a monotone function f on n bits such that $Z(f, 1 - \delta) \leq 1 - \Omega(1)$ for the smallest possible quantity δ . This problem was implicitly posed in [BKS99].

The following fact is considered folklore:

Proposition 1 *Let $f : \Omega_n \rightarrow \{-1, +1\}$ be monotone. Then*

- $Z(f, 1 - \delta) \geq (1 - \delta)^{I_n} = (1 - \delta)^{(\sqrt{2/\pi} + o(1))\sqrt{n}}$.
- *Therefore if $Z(f, 1 - \delta) \leq 1 - \epsilon$, then:*

$$\delta \geq \frac{\sqrt{\pi/2}}{\sqrt{n}} \epsilon + o(1/\sqrt{n}).$$

Proof: The second statement follows directly from the first one, which we now prove. For a monotone function f it holds that $\sum_S \hat{f}^2(S) = 1$, and $\sum_S |S| \hat{f}^2(S) \leq I_n$, where I_n is defined in (3). We want to bound $Z(f, \eta) = \sum_S \hat{f}^2(S) \eta^{|S|}$. Clearly a bound is the solution to the following linear programming problem:

$$\text{Minimize } \sum_{k=0}^n \eta^k x_k \quad \text{subject to: } \sum_{k=0}^n x_k = 1, \quad \sum_{k=0}^n k x_k \leq I_n, \quad 0 \leq x_k \leq 1 \text{ for all } k. \quad (5)$$

Since the optimum of a linear program occurs at a vertex, an optimal solution is of the form $x_m = 1$ and $x_i = 0$ for all $i \neq m$, for some m . The constraint $\sum_{k=0}^n k x_k \leq I_n$ implies that $m \leq I_n$. Therefore the value of $\sum_{k=0}^n \eta^k x_k$ is at least η^{I_n} , as claimed. $\square$

Prior to this work, the best sensitivity with respect to N_η was achieved via the recursive majority of 3 function (folklore, see [BL90, BKS99]; we formally define the recursive majority function in Subsection 2.2). This function satisfies $Z(f, 1 - \delta) \leq 1 - \Omega(1)$, for $\delta = n^{-\alpha}$, where $\alpha = 1 - \ln 2 / \ln 3 = 0.36907\dots$

1.3 Our results

Recursive majority functions seem to be sensitive to noise. Previous techniques for analyzing recursive majorities had suggested that recursive majorities of 5, 7, etc. might be less sensitive than recursive majority of 3. However, this is not the case.

Theorem 1 *For every constant $\alpha < 1/2$, and $0 < \epsilon < 1$, there exists an odd $k \geq 3$ such that for ℓ sufficiently large, and $n = k^\ell$, $f_n = \text{REC-MAJ-}k_\ell : \Omega_n \rightarrow \{-1, +1\}$ is a balanced function with*

$$Z(f_n, 1 - n^{-\alpha}) \leq \epsilon.$$

Note that this construction is explicit. Moreover, using k -majority gates, we obtain a read-once, log-depth circuit which implements the function. The proof technique we use is closely related to techniques in classical branching processes [AN72] (see also [M98]).

By relaxing the bounded degree property, and instead using majority gates of varying fan-in, we obtain an explicit read-once construction of log log-depth which is sensitive to a noise rate of about $n^{-1/2}$, up to a sub-logarithmic correction.

Theorem 2 *There exists an explicit infinite family of balanced monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ such that for every sufficiently small $\epsilon > 0$:*

$$Z(f_n, 1 - \epsilon/M) \leq 1 - \epsilon + O(\epsilon^2),$$

where $M = \sqrt{n}/\Theta(\log^t n)$, and $t = \log_2 \sqrt{\pi/2} = .3257\dots$

Finally, analyzing a probabilistic construction due to Talagrand [T96], we obtain a tight result up to constant factors.

Theorem 3 *There exists an infinite family of balanced monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ with the following property:*

$$Z(f_n, 1 - n^{-1/2}) \leq 1 - \Omega(1). \tag{6}$$

In the above theorems we aim for the smallest ϵ such that there is a monotone function f satisfying $Z(f, 1 - \epsilon) \leq 1 - \Omega(1)$. At the other end of the spectrum, one can ask for a function which is as sensitive as possible to large noise. In Section 5 we prove:

Theorem 4 *Let $\delta \leq 1 - \Omega(1)$. Then there is an infinite family of monotone functions $\{g_n\}$ satisfying:*

$$Z(g_n, \delta) \leq \frac{\log^{1+u'} n}{n},$$

where u' is any number exceeding $u = \log_{4/3} 3 = 3.818\dots$

Combining Theorem 4 with Theorem 3 and Theorem 2 (via Proposition 7 below), it is easy to deduce the following:

Corollary 5 *For every $\delta > 0$, there exists a constant $c(\delta) > 0$, such that*

- *There exists an explicit infinite family of balanced monotone $f_n : \Omega_n \rightarrow \{-1, +1\}$ such that*

$$Z(f_n, 1 - \epsilon) < \delta,$$

where $\epsilon = c(\delta) \log^t n / \sqrt{n}$, and $t = \log_2 \sqrt{\pi/2} = .3257\dots$

- *There exists an infinite family of monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ with the following property*

$$Z(f_n, 1 - c(\delta)n^{-1/2}) < \delta.$$

In Section 1.4 we briefly discuss how our results are related to problems in learning, hardness amplification, neural networks, and voting schemes. In Section 2 we analyze the sensitivity of the recursive majority function and prove Theorem 1. In Section 3 we prove Theorem 2. In Section 4 we analyze Talagrand’s construction and prove Theorem 3. In Sections 5 and 6, we discuss high sensitivity, and prove Theorem 4.

1.4 Related problems

Our constructions have interesting implications for computational learning theory, in particular the “PAC learning” model of Valiant [V84] and its restricted setting, “uniform PAC learning” (see, e.g., [LMN93, M94]). In particular, in Corollary 17 we give an explicit construction which shows the tightness of a learning algorithm by Bshouty and Tamon [BT96] for the class of monotone functions.

In [O02] the second author constructs noise-sensitive monotone functions in order to amplify the “hardness-on-average” of languages in NP. In fact, the proof of Proposition 16 is closely related to the construction in [O02].

Finally, Theorem 1 implies that “read-once” neural networks (see, e.g., [H99]) may have high sensitivity to noise compared to a single neuron (see [Pe98]). Similarly, it may also be interpreted as saying that in an i.i.d. voter model the simple majority scheme is more stable than a recursive majority scheme.

2 Sensitivity of majorities

2.1 Majority

We denote the majority function on k bits by MAJ_k , where k is assumed to be odd. The following lemma concerning the sensitivity of the majority function is used in the proof of Theorem 2.

Lemma 6 *Suppose $k \geq 3$ and $\delta \leq 1/k$. Say we pick a random input to MAJ_k — call it x — and then construct y by flipping each bit of x independently with probability δ . Then*

$$\mathbf{P}[\text{MAJ}_k(x) \neq \text{MAJ}_k(y)] \geq \sqrt{\frac{2}{\pi}} \sqrt{k} \delta \exp(-1/3k) \exp(-\delta k).$$

Proof: Clearly,

$$\mathbf{P}[\text{MAJ}_k(x) \neq \text{MAJ}_k(y)] \geq \mathbf{P}[\text{MAJ}_k(x) \neq \text{MAJ}_k(y) \mid \text{exactly one flip}] \times \mathbf{P}[\text{exactly one flip}], \quad (7)$$

and $\mathbf{P}[\text{exactly one flip}] = k\delta(1-\delta)^{k-1}$. By elementary calculus, $(1-\delta)^{k-1} \geq \exp(-\delta k)$ for $\delta \leq 1/k$. Therefore,

$$\mathbf{P}[\text{exactly one flip}] = k\delta(1-\delta)^{k-1} \geq k\delta \exp(-\delta k). \quad (8)$$

The probability that the majority flips given that there is exactly one flipped bit in x , is exactly the probability that the remaining input bits split evenly — i.e.,

$$\begin{aligned} \mathbf{P}[\text{MAJ}_k(x) \neq \text{MAJ}_k(y) \mid \text{exactly one flip}] \\ = \binom{k-1}{(k-1)/2} 2^{-(k-1)} \geq \sqrt{\frac{2}{\pi k}} (1 - 1/4k) \geq \sqrt{\frac{2}{\pi k}} \exp(-1/3k), \end{aligned} \quad (9)$$

where the first inequality follows by Stirling's formula and the second since $1 - 1/4k \leq \exp(-1/3k)$ for $k \geq 3$. Combining (7), (8) and (9) we obtain the required result. $\square$

2.2 Recursive majorities

We begin by defining a way to compose two boolean functions, and with the definition of the recursive majority function.

Definition 1 For $f : \Omega_n \rightarrow \{-1, +1\}$, $g : \Omega_m \rightarrow \{-1, +1\}$, we let $f \otimes g$ denote the function $f \otimes g : \Omega_{nm} \rightarrow \{-1, +1\}$ defined by

$$f \otimes g(x_1, \dots, x_{nm}) = f(g(x_1, \dots, x_m), \dots, g(x_{(n-1)m+1}, \dots, x_{nm})).$$

For ℓ a positive integer, we define $f^{\otimes \ell} = f$ if $\ell = 1$, and $f^{\otimes \ell} = f \otimes (f^{\otimes \ell-1})$ otherwise. For any odd k , we write $\text{REC-MAJ-}k_\ell$ for $\text{MAJ}_k^{\otimes \ell}$.

The following proposition is immediate, yet useful:

Proposition 7 If g is a balanced function and f is any function, then $Z(f \otimes g, \eta) = Z(f, Z(g, \eta))$.

We now present a simple asymptotic formula for calculating the noise sensitivity of $f^{\otimes \ell}$ as $\ell \rightarrow \infty$ for any balanced f . Following this, we will apply the formula to the majority function to recover Theorem 1

Proposition 8 Let $f : \Omega_k \rightarrow \{-1, +1\}$ be a balanced function, and let

$$a = I(f) = \sum_S |S| \hat{f}^2(S), \quad b = \sum_{|S|=1} \hat{f}^2(S).$$

(Note that if f is monotone then $b = II(f)$.) If $a > 1$ and $b < 1$, then $Z(f^{\otimes \ell}, 1 - \delta) \leq \epsilon$, for $\ell \geq \log_a(1/\delta)(1 + o_\delta(1)) + \log_{1/b}(1/\epsilon)(1 + o_\epsilon(1))$, where $o_\delta(1) \rightarrow 0$ as $\delta \rightarrow 0$ and $o_\epsilon(1) \rightarrow 0$ as $\epsilon \rightarrow 0$.

The proof of this is a straightforward analysis of the iteration of convex functions. We will use a simple lemma:

Lemma 9 *Let $p : [0, 1] \rightarrow [0, 1]$ be a positive increasing function.*

- *Suppose that $p(\eta) \leq b\eta + c\eta^2$ for all $\eta < \epsilon'$ where $0 \leq b < 1$ and $0 \leq c$. Let ϵ' be such that $b + c\epsilon' < 1$. Then for all $\eta < \epsilon'$, all $\epsilon > 0$ and $k \geq \ln(\epsilon/\eta)/\ln(b + c\epsilon')$, it holds that $p^{(k)}(\eta) \leq \epsilon$.*
- *Similarly, suppose that $p(1 - \delta) \leq 1 - a\delta + c\delta^2$, where $a > 1$ and $c \geq 0$. Let ϵ' be such that $a - c\epsilon' > 1$. Then for all $\delta < \epsilon'$, all $\epsilon > 0$ and $k \geq \ln(\epsilon/\delta)/\ln(a - c\epsilon')$, it holds that $p^{(k)}(1 - \delta) \leq 1 - \epsilon$.*

Proof: We prove the first claim; the second is proved in essentially the same way. If $\eta < \epsilon'$, then $p(\eta) \leq b\eta + c\eta^2 < (b + c\epsilon')\eta$. Therefore $p^{(k)}(\eta) \leq (b + c\epsilon')^k \eta$. So in order to get $p^{(k)}(\eta) \leq \epsilon$, it suffices to take $k \geq \ln(\epsilon/\eta)/\ln(b + c\epsilon')$. $\square$

Proof of Proposition 8: Let $f = \sum_S \hat{f}^2(S) u_S$ be the Fourier expansion of f . Letting $p(\eta) := Z(f, \eta) = \sum_S \hat{f}^2(S) \eta^{|S|}$, we see that $p(\eta)$ is a convex polynomial function of η which satisfies:

$$p(0) = 0, \quad p(1) = 1, \quad p'(0) = \sum_{|S|=1} \hat{f}^2(S) = \sum_{j=1}^k I_j^2(k) = b, \quad p'(1) = \sum_S |S| \hat{f}^2(S) = a. \quad (10)$$

Proposition 7 implies that

$$Z(f^{\otimes \ell}, \eta) = p^{(\ell)}(\eta) := \underbrace{p(p(\cdots p(\eta) \cdots))}_{\ell \text{ times}}. \quad (11)$$

The rest of the proof consists of studying the iterations (11) using properties (10) and Lemma 9.

For a function q which depends on ϵ and δ , write $q = O(1)$, if q is uniformly bounded, $q = o_\delta(1)$, if $q(\delta) \rightarrow 0$ as $\delta \rightarrow 0$ and similarly $q = o_\epsilon(1)$.

We are going to analyze the recursion in five steps.

- Note that $p(1 - \delta) \leq 1 - a\delta + c\delta^2$ for some constant c which depends on p only. Recall that $a > 1$ and let $r = (a - 1)/2c$. Let $\delta' = 1/\ln(1/\delta)$. We may assume without loss of generality that $\delta' < r$. By Lemma 9 it follows that $p^{(k)}(1 - \delta) \leq 1 - \delta'$ for k larger than

$$\frac{\ln(\delta'/\delta)}{\ln(a - c\delta')} < \frac{\ln(1/\delta)}{(1 + O(\delta')) \ln(a)} = (1 + o_\delta(1)) \log_a(1/\delta).$$

- Similarly $p^{(k)}(1 - \delta') \leq 1 - r$ for k larger than

$$\frac{\ln(r/\delta')}{\ln(a - cr)} < \frac{\ln(1/\delta')}{\ln((a + 1)/2)} = O(\ln \ln(1/\delta)) = o_\delta(\log_a(1/\delta)).$$

- By our assumption $p(\eta) \leq b\eta + c'\eta^2$ for some c' . Recall that $b < 1$ and let $r' = (1 - b)/2c'$. Since the function p is strictly convex with $p(0) = 0$ and $p(1) = 1$, there exists a constant k (which depends only on p) such that $p^{(k)}(1 - r) \leq r'$.

- Letting $\eta' = 1/\ln(1/\eta)$ and assuming that $\eta' < r'$, we obtain as before that $p^k(r') \leq \eta'$ for k larger than

$$\frac{\ln(\eta'/r)}{\ln(b + c'r')} < \frac{\ln(\eta')}{\ln((b+1)/2)} = O(\ln \ln(1/\eta)) = o_\eta(\log_{1/b}(1/\eta)).$$

- Finally we obtain $p^{(k)}(\eta') \leq \eta$ for k greater than

$$\frac{\ln(\eta/\eta')}{\ln(b + c\eta')} < \frac{\ln(\eta)}{(1 + O(\eta')) \ln(b)} = (1 + o_\eta(1)) \log_{1/b}(1/\eta).$$

Combining the above five steps, we obtain the required result. $\square$

Proof of Theorem 1: We now apply Proposition 8 to REC-MAJ- k_ℓ . It is easy to calculate (and well known) that for the majority function on $k = 2r + 1$ inputs, we have:

$$H(\text{MAJ}_k) = \frac{2r+1}{2^{4r}} \binom{2r}{r}^2, \quad I(\text{MAJ}_k) = \frac{2r+1}{2^{2r}} \binom{2r}{r}.$$

Note therefore that $I(\text{MAJ}_k) = \Theta(\sqrt{k})$. By substituting these values into Proposition 8, we get Theorem 1. $\square$

3 Sensitivity to small noise

In this section we prove Theorem 2. The construction in Theorem 2 again consists of recursive majorities, where now the number of inputs to the majority varies with the level. The estimates on the sensitivity of these majority functions are derived via Lemma 6.

Proof of Theorem 2: Since we are dealing with correlations close to 1, it will be more helpful to look at their difference from 1. In particular, we will prove the following equivalent formulation of the theorem: Let x be a randomly chosen input to f_n , and suppose we flip each bit of x independently with probability ϵ/M , forming y . Then the probability that $f_n(x) \neq f_n(y)$ is at least $\epsilon - O(\epsilon^2)$.

The function f_n will be given by recursive majorities of increasing arity: $f_n = \text{MAJ}_{k_1} \otimes \text{MAJ}_{k_2} \otimes \dots \otimes \text{MAJ}_{k_\ell}$. We will select $k_i = 3^{2^{i-1}+1}$, so “from the top down” the majorities have arity 9, 27, 243, etc. Note that $k_{i+1} = k_i^2/3$. With these choices, the number of inputs is $n = 3^{2^\ell + \ell - 1}$. Hence $\ell \leq \log_2 \log_3 n$.

Let $\delta_0 = \epsilon/M$, and recursively define δ_{i+1} to be the probability that the output of a $\text{MAJ}_{k_{\ell-i}}$ flips, given that each of its inputs is flipped independently with probability δ_i . Since all MAJ functions are balanced, Proposition 7 tells us that the probability the output of f_n is flipped is δ_ℓ . We will show that $\delta_\ell \geq \epsilon - O(\epsilon^2)$.

By Lemma 6,

$$\delta_{i+1} \geq g(k_{\ell-i}) \exp(-\delta_i k_{\ell-i}) \delta_i,$$

where:

$$g(t) := \frac{1}{\sqrt{\pi/2}} \sqrt{t} \exp(-1/3t).$$

Recursively define $\eta_0 = \eta'_0 = \delta_0$, and:

$$\eta_{i+1} = g(k_{\ell-i}) \exp(-\eta_i k_{\ell-i}) \eta_i, \quad \eta'_{i+1} = g(k_{\ell-i}) \eta'_i.$$

Since the probability that the output of MAJ flips is an increasing function of δ , we can conclude that $\delta_i \geq \eta_i$ for every i . But clearly $\eta'_i \geq \eta_i$ for every i . Hence, for every i , $\eta_{i+1} \geq g(k_{\ell-i}) \exp(-\eta'_i k_{\ell-i}) \eta_i$. It follows immediately that:

$$\begin{aligned} \eta_\ell &\geq \left(\prod_{i=0}^{\ell-1} g(k_{\ell-i}) \exp(-\eta'_i k_{\ell-i}) \right) \eta_0 \\ &= \left(\frac{1}{\sqrt{\pi/2}} \right)^\ell \prod_{j=1}^{\ell} \sqrt{k_j} \exp\left(-\frac{1}{3} \sum_{j=1}^{\ell} k_j^{-1}\right) \cdot \exp\left[\sum_{i=0}^{\ell-1} -\eta'_i k_{\ell-i}\right] \cdot \delta_0 \end{aligned}$$

Defining

$$M := \prod_{m=1}^{\ell} g(k_m) = \left(\frac{1}{\sqrt{\pi/2}} \right)^{\log_2 \log_3 n} \prod_{j=1}^{\ell} \sqrt{k_j} \exp\left(-\frac{1}{3} \sum_{j=1}^{\ell} k_j^{-1}\right) = \left(\frac{1}{\sqrt{\pi/2}} \right)^{\log_2 \log_3 n} \sqrt{n} \exp(-O(1)),$$

and recalling $\delta_0 = \epsilon/M$, we obtain:

$$\eta_\ell \geq M \cdot \exp\left[\sum_{i=0}^{\ell-1} -\eta'_i k_{\ell-i}\right] \cdot (\epsilon/M) = \epsilon \cdot \exp\left[\sum_{i=0}^{\ell-1} -\eta'_i k_{\ell-i}\right].$$

Since $\delta_\ell \geq \eta_\ell$, it remains to show:

$$\exp\left[\sum_{i=0}^{\ell-1} -\eta'_i k_{\ell-i}\right] \geq 1 - O(\epsilon).$$

By the recursive definition of η'_i , we immediately have $\eta'_i = (\prod_{j=0}^{i-1} g(k_{\ell-j})) \eta'_0$. Hence $\eta'_i = M(\prod_{m=1}^{\ell-i} g(k_m)^{-1}) \eta'_0 = \epsilon(\prod_{m=1}^{\ell-i} g(k_m)^{-1})$. Therefore:

$$\exp\left[\sum_{i=0}^{\ell-1} -\eta'_i k_{\ell-i}\right] = \exp\left[-\epsilon \sum_{m=1}^{\ell} \frac{k_m}{g(k_1)g(k_2) \cdots g(k_m)}\right].$$

Hence if we can show $\sum_{m=1}^{\ell} k_m / g(k_1)g(k_2) \cdots g(k_m) = O(1)$ then we're done. The first term in this sum is $k_1/g(k_1) = O(1)$. The ratio of the m th term to the $(m-1)$ th term is $k_m/k_{m-1}g(k_m)$. But $k_{m-1} = \sqrt{3}\sqrt{k_m}$ by definition, so this ratio is $\sqrt{k_m}/\sqrt{3}g(k_m) = \sqrt{\pi/2}/\sqrt{3} \exp(-1/3k_m) < 1$. Hence the terms in the sum decrease geometrically, so the sum is indeed $O(1)$. $\square$

4 Talagrand's function

In [T96], Talagrand gives a randomized construction of a monotone $f_n : \Omega_n \rightarrow \{-1, +1\}$ with the following property: at least an $\Omega(1)$ fraction of points x in Ω_n satisfy both $f_n(x) = -1$, and

$\#\{x' : \Delta(x, x') = 1 \text{ and } f(x) = +1\} \geq \Omega(n^{1/2})$, where Δ denotes Hamming distance. It is natural to conjecture that this function is sensitive to noise as small as $n^{-1/2}$, and indeed we prove this below.

The function of [T96] may not be balanced. However, the following Lemma implies that by slightly modifying the function, one can easily obtain a balanced function.

Lemma 10 *Let $f : \{-1, +1\}^n \rightarrow \{-1, +1\}$ be a monotone function and $\eta > 0$, such that $\mathbf{P}[f \neq f(N_\eta(x))] \geq \delta$. Then there exists a balanced monotone function $g : \{-1, +1\}^{n+1} \rightarrow \{-1, +1\}$ such that $\mathbf{P}[g_n(x) \neq g(N_\eta(x))] \geq \delta/4$.*

Proof: In order to define g from f , assume without loss of generality that $\mathbf{P}[f(x) = 1] \geq \mathbf{P}[f(x) = -1]$. We claim that there exists a balanced monotone function $g : \{-1, +1\}^n \rightarrow \{-1, +1\}$, such that for all z , $g(1z) = f(z)$. Note that the existence of g implies the required result, as

$$\mathbf{P}[g(yz) \neq g(N_\eta(yz))] \geq \mathbf{P}[y = 1, N_\eta(y) = 1] \times \mathbf{P}[f(z) \neq f(N_\eta(z))] \geq \delta/4.$$

To prove the existence of g , let $\tilde{h}(yz) = f(z)$ for all y and z ; and $h(1z) = f(z)$, $h(-1z) = -1$ for all z . Then both h and $\tilde{h}$ are monotone functions such that the value at $y = 1z$, is the same as the value of f at z . Moreover, for all y and z , $h(yz) \leq \tilde{h}(yz)$, and $\mathbf{P}[\tilde{h}(yz) = -1] = \mathbf{P}[f(z) = -1] \leq 1/2$, and $\mathbf{P}[h(yz) = -1] \geq 1/2$. It follows that there exists a monotone balanced function g “between” $\tilde{h}$ and h , as needed. $\square$

Talagrand’s function $f = f_n$ is a random CNF formula on its n inputs. Specifically, f is the $2^{\sqrt{n}}$ -wise AND of $\sqrt{n}$ -wise ORs, where each OR’s inputs are selected independently and uniformly at random (with replacement) from $[n]$. To prove Theorem 3, it suffices to prove that if we pick f , x , and $x' := N_{1-2\epsilon}(x)$ at random (where $\epsilon = n^{-1/2}$), then:

$$\mathbf{E}_f[\mathbf{P}[f(x) \neq f(N_{1-2\epsilon}(x))]] \geq \Omega(1).$$

Proof of Theorem 3:

For notational simplicity, let’s revert to using 0 and 1 for false and true, instead of +1 and -1. Consider two fixed inputs to f , x and x' . Let n_{00} denote the number of indices i such that $x_i = 0, x'_i = 0$, let n_{01} denote the number of indices i such that $x_i = 0, x'_i = 1$, and analogously define n_{10} and n_{11} . Also define $n_{0\star}$ to be $n_{00} + n_{01}$, the number of 0’s in x , and similarly $n_{\star 0}, n_{1\star}, n_{\star 1}$.

Consider any particular OR, call it $\vee$. Let p_{00} denote the probability — over the choice of f — that $\vee(x) = 0$ and $\vee(x') = 0$. Again, define $p_{01}, p_{10}, p_{11}, p_{0\star}$, etc. analogously.

We immediately get:

$$\begin{aligned} p_{00} &= \left(\frac{n_{00}}{n}\right)^{\sqrt{n}} \\ p_{10} &= \left(\frac{n_{\star 0}}{n}\right)^{\sqrt{n}} - \left(\frac{n_{00}}{n}\right)^{\sqrt{n}} \\ p_{01} &= \left(\frac{n_{0\star}}{n}\right)^{\sqrt{n}} - \left(\frac{n_{00}}{n}\right)^{\sqrt{n}}, \end{aligned} \tag{12}$$

and by subtracting these quantities from 1,

$$p_{11} = 1 - \left(\frac{n_{\star 0}}{n}\right)^{\sqrt{n}} - \left(\frac{n_{0\star}}{n}\right)^{\sqrt{n}} + \left(\frac{n_{00}}{n}\right)^{\sqrt{n}}.$$

All of the ORs are independent, so we may make a similar calculation for the main AND in f . Let q_{00} denote the probability — over the choice of f still — that $f(x) = 0, f(x') = 0$, and again define q_{01} , etc.

Calculating as before:

$$\begin{aligned} q_{11} &= p_{11}^{2\sqrt{n}} \\ q_{10} &= p_{1\star}^{2\sqrt{n}} - p_{11}^{2\sqrt{n}} \\ q_{01} &= p_{\star 1}^{2\sqrt{n}} - p_{11}^{2\sqrt{n}}. \end{aligned}$$

Now the probability that $f(x) \neq f(x')$ is simply $q_{01} + q_{10}$. Hence:

$$\begin{aligned} \mathbf{E}_f[\mathbf{P}[f(x) \neq f(N_{1-2\epsilon}(x))]] &= \mathbf{E}_{x, x':=N_{1-2\epsilon}(x)}[\mathbf{P}_f[f(x) \neq f(x')]] \\ &= \mathbf{E}_{x, x'}[q_{01} + q_{10}] \\ &= \mathbf{E}_{x, x'}[q_{01}] + \mathbf{E}_{x, x'}[q_{10}]. \end{aligned} \tag{13}$$

Since x and x' have the same distribution, $\mathbf{E}_{x, x'}[q_{01}] = \mathbf{E}_{x, x'}[q_{10}]$ by symmetry. Hence (13) = $2\mathbf{E}_{x, x'}[q_{10}]$. Thus it suffices to show:

$$\mathbf{E}_{x, x'}[q_{10}] = \mathbf{E}_{x, x'}[p_{1\star}^{2\sqrt{n}} - p_{11}^{2\sqrt{n}}] \geq \Omega(1).$$

We now focus on the quantity $(*) := p_{1\star}^{2\sqrt{n}} - p_{11}^{2\sqrt{n}}$. Let $g(t) = t^{2\sqrt{n}}$. By the mean value theorem, $g(b) - g(a) = (b - a)g'(c)$ for some $c \in [a, b]$. Thus:

$$(*) = (p_{1\star} - p_{11})2^{\sqrt{n}}c^{2^{\sqrt{n}}-1} = 2^{\sqrt{n}}p_{10}c^{2^{\sqrt{n}}-1},$$

for some $c \in [p_{11}, p_{1\star}]$. Since $c^{2^{\sqrt{n}}-1}$ is no smaller than $p_{11}^{2^{\sqrt{n}}}$, we conclude:

$$(*) \geq 2^{\sqrt{n}}p_{10}p_{11}^{2^{\sqrt{n}}}. \tag{14}$$

We proceed by conditioning on $(n_{00}, n_{01}, n_{10}, n_{11})$. Since $n_{0\star} \sim \text{Binomial}(n, 1/2)$, the probability that $n_{0\star}$ is outside the range $[\frac{n}{2} - \sqrt{n}, \frac{n}{2} + \sqrt{n}]$ is at most .05, for sufficiently large n (by a standard tail bound; $\sqrt{n}$ is two standard deviations). Assuming that $n_{0\star}$ is some fixed quantity in this range, $n_{00} \sim \text{Binomial}(n_{0\star}, 1 - \epsilon)$. By a similar tail bound, the probability that n_{00} is larger than $(1 - \epsilon + 2\sqrt{\epsilon/n_{0\star}})n_{0\star}$ is again at most .05. So assuming n is sufficiently large, we have that except with probability .1:

$$n_{0\star} \in \left[\frac{n}{2} - \sqrt{n}, \frac{n}{2} + \sqrt{n}\right] \tag{15}$$

$$n_{00}/n_{0\star} < 1 - \epsilon + 3\sqrt{\epsilon/n}, \tag{16}$$

where (16) uses the bound $n_{0\star} > n/2.1$.

Finally, just as $n_{0\star} \in [\frac{n}{2} - \sqrt{n}, \frac{n}{2} + \sqrt{n}]$ except with probability .05, so too may we conclude:

$$n_{\star 0} \in \left[\frac{n}{2} - \sqrt{n}, \frac{n}{2} + \sqrt{n} \right], \quad (17)$$

except with probability .05.

In conclusion, (15), (16), (17) hold, except with probability at most .15. Since $(*) \geq 0$ always,

$$\begin{aligned} \mathbf{E}_{x,x'}[(*)] &\geq \mathbf{E}_{x,x'}[(*) \mid (15), (16), (17)] \times \mathbf{P}[(15), (16), (17)] \\ &\geq .85 \mathbf{E}_{x,x'}[(*) \mid (15), (16), (17)]. \end{aligned}$$

Since we are only trying to prove $\mathbf{E}_{x,x'}[(*)] \geq \Omega(1)$, we will henceforth assume (15), (16), (17) hold, and it suffices to prove $\mathbf{E}_{x,x'}[(*)] \geq \Omega(1)$ conditioned on this assumption. I.e., all future expectations are conditioned on (15), (16), (17).

Continuing from (14),

$$\begin{aligned} (*) &\geq 2^{\sqrt{n}} p_{10} \left(1 - \left(\frac{n_{\star 0}}{n} \right)^{\sqrt{n}} - \left(\frac{n_{0\star}}{n} \right)^{\sqrt{n}} + \left(\frac{n_{00}}{n} \right)^{\sqrt{n}} \right)^{2^{\sqrt{n}}} \\ &\geq 2^{\sqrt{n}} p_{10} \left(1 - \left(\frac{n_{\star 0}}{n} \right)^{\sqrt{n}} - \left(\frac{n_{0\star}}{n} \right)^{\sqrt{n}} \right)^{2^{\sqrt{n}}} \\ &\geq 2^{\sqrt{n}} p_{10} \left(1 - (1/2 + 1/\sqrt{n})^{\sqrt{n}} - (1/2 + 1/\sqrt{n})^{\sqrt{n}} \right)^{2^{\sqrt{n}}} \quad (\text{by (15) and (17)}) \\ &\geq 2^{\sqrt{n}} p_{10} (1 - 2e/2^{\sqrt{n}})^{2^{\sqrt{n}}} \quad (\text{asymptotically}) \\ &\geq e^{-2e} 2^{\sqrt{n}} p_{10} \\ &\geq .004 \cdot 2^{\sqrt{n}} p_{10} \quad (\text{for all } n) \end{aligned}$$

Hence showing $\mathbf{E}_{x,x'}[(*)] \geq \Omega(1)$ amounts to showing $\mathbf{E}_{x,x'}[2^{\sqrt{n}} p_{10}] \geq \Omega(1)$. By (12):

$$\begin{aligned} 2^{\sqrt{n}} p_{10} &= \left(2 \frac{n_{\star 0}}{n} \right)^{\sqrt{n}} \left(1 - \left(\frac{n_{00}}{n_{\star 0}} \right)^{\sqrt{n}} \right) \\ &\geq (1 - 2/\sqrt{n})^{\sqrt{n}} \left(1 - \left(\frac{n_{00}}{n_{\star 0}} \right)^{\sqrt{n}} \right) \quad (\text{by (17)}) \\ &\geq e^{-2} \left(1 - \left(\frac{n_{00}}{n_{\star 0}} \right)^{\sqrt{n}} \right) \\ &\geq e^{-2} \left(1 - (1 - \epsilon + 2\sqrt{\epsilon/n})^{\sqrt{n}} \right) \quad (\text{by (16)}) \end{aligned}$$

When $\epsilon = 1/\sqrt{n}$, the quantity $(1 - \epsilon + 2\sqrt{\epsilon/n})^{\sqrt{n}}$ is asymptotically e^{-1} for large n . Hence:

$$\mathbf{E}_{x,x'}[2^{\sqrt{n}} p_{10}] \geq e^{-2}(1 - e^{-1}) \geq \Omega(1),$$

as needed. $\square$

5 Tribes and high sensitivity

We have mostly settled the question of how small ϵ can be, such that there is a monotone function f satisfying $Z(f, 1 - \epsilon) \leq 1 - \Omega(1)$. At the other end of the spectrum, one can ask for a function which is as sensitive as possible to large noise. An essentially optimal function for this problem (if the initial correlation is small enough) is the so-called tribes function of Ben-Or and Linial [BL90].

Let AND_k denote the AND function on k bits, and let OR_k denote the or function on k bits. For each $b \in \mathbb{N}$, define $n = n_b$ to be the smallest integral multiple of b such that $(1 - 2^{-b})^{n/b} \leq 1/2$, so n is very roughly $(\ln 2)b2^b$, and $b = \log_2 n - \log_2 \ln n + o(1)$. Now define the tribes function T_n to be $\text{OR}_{n/b} \otimes \text{AND}_b$. This function is monotone, and by construction it's near-balanced; it's easy to see that $\mathbf{P}[T_n = +1] = (1 - 2^{-b})^{n/b} = 1/2 - O(\log n/n)$.

There is a near-tight estimation of $Z(T_n, \eta)$ in [O02] based on Fourier coefficients. A simpler and exact analysis is given here.

Proposition 11

$$Z(T_n, \eta) = 1 - 4 \left[\left(1 - 2^{-b}\right)^{n/b} - \left(1 - \left(2 - \left(\frac{1}{2} + \frac{1}{2}\eta\right)^b\right) 2^{-b}\right)^{n/b} \right]$$

Proof: As in the proof of Theorem 3, we will temporarily switch to writing 0 and 1 for false and true, for notational ease.

Suppose x is a uniformly random input to T_n , and y is chosen so that $y_i = x_i$ independently with probability p . Call x 's length- b blocks $x^1, x^2, \dots, x^{n/b}$, and similarly for y . We have

$$\mathbf{P}[T_n(x) = 0] = (1 - 2^{-b})^{n/b}.$$

We now calculate

$$\mathbf{P}[T_n(y) = 1 \text{ and } T_n(x) = 0]. \quad (18)$$

Since x given y is distributed as y given x , by symmetry we have,

$$\mathbf{P}[\text{AND}_b(y^1) = 1 \text{ and } \text{AND}_b(x^1) = 0] = \mathbf{P}[\text{AND}_b(x^1) = 1 \text{ and } \text{AND}_b(y^1) = 0] = 2^{-b}(1 - p^b).$$

Also,

$$\mathbf{P}[\text{AND}_b(y^1) = 1 \text{ and } \text{AND}_b(x^1) = 1] = 2^{-b}p^b;$$

hence,

$$\begin{aligned} \mathbf{P}[\text{AND}_b(y^1) = 0 \text{ and } \text{AND}_b(x^1) = 0] &= 1 - 2^{-b}p^b - 2 \cdot 2^{-b}(1 - p^b) \\ &= 1 - (2 - p^b)2^{-b}. \end{aligned}$$

In order to have $T_n(y) = 1$ and $T_n(x) = 0$, we must have $\text{AND}_b(x^i) = 0$ for each i , but not have $\text{AND}_b(y^i) = 0$ for every i . The probability of this is thus:

$$(1 - 2^{-b})^{n/b} - (1 - (2 - p^b)2^{-b})^{n/b}.$$

Hence by symmetry again, the probability that $T_n(x) \neq T_n(y)$ is

$$2[(1 - 2^{-b})^{n/b} - (1 - (2 - p^b)2^{-b})^{n/b}].$$

Therefore by definition,

$$Z(T_n, \eta) = 1 - 4 \left[\left(1 - 2^{-b}\right)^{n/b} - \left(1 - \left(2 - \left(\frac{1}{2} + \frac{1}{2}\eta\right)^b\right) 2^{-b}\right)^{n/b} \right].$$

□

This is of course an unappealing expression. We estimate it thus:

Corollary 12

$$Z(T_n, \eta) \leq (1 + o(1)) \frac{\log_2^2 n}{n} \eta (1 + \eta)^{\log_2 n} + O(\log^2 n / n^2)$$

Proof: From the definition, one can see that $Z(f, 0) = 1 - 4\mathbf{P}[f = 0]\mathbf{P}[f = 1]$. Since we know $\mathbf{P}[T_n = 0] = 1/2 - O(\log n/n)$, we get $Z(T_n, 0) = O(\log^2 n/n^2)$. By the mean value theorem, $Z(T_n, \eta) - Z(T_n, 0) = \eta Z'(T_n, c)$, for some $0 \leq c \leq \eta$, where the derivative is with respect η . Differentiating Proposition 11, we get

$$Z'(T_n, c) = \left(\frac{1}{4} + \frac{1}{4}c\right)^b n \frac{4}{1+c} \left(1 - 2^{1-b} + \left(\frac{1}{4} + \frac{1}{4}c\right)^b\right)^{n/b-1}.$$

Even if c were as large as 1, the factor $\frac{4}{1+c}(1 - 2^{1-b} + (\frac{1}{4} + \frac{1}{4}c)^b)^{n/b-1}$ would be at most $1 + o(1)$. Using the upper bound $c \leq \eta$ for the $(\frac{1}{4} + \frac{1}{4}c)^b n$ factor, we conclude:

$$\begin{aligned} Z(T_n, \eta) &\leq \eta 4^{-b} n (1 + \eta)^b (1 + o(1)) + O(\log^2 n / n^2) \\ &= (1 + o(1)) \frac{\log_2^2 n}{n} \eta (1 + \eta)^{\log_2 n} + O(\log^2 n / n^2), \end{aligned}$$

as claimed. □

Making another simple estimate, we get:

Corollary 13 *If $\eta \leq O(1/\log n)$, then $Z(T_n, \eta) \leq O(\eta \log^2 n/n)$.*

We now prove Theorem 4, giving a monotone function for which $Z(f, \delta)$ is small when $\delta \leq 1 - \Omega(1)$.

Proof of Theorem 4: The idea is to first use REC-MAJ-3 to reduce δ to $\eta := 1/\log n$; then, apply a tribes function.

Let T_n be any tribes function. We will construct $g_{n'}$ on $n' := n \log^{u'} n$ inputs. Let ℓ be the REC-MAJ-3 depth necessary from Theorem 1 to reduce δ correlation down to $1/\log n$ correlation. Hence $\ell = (1 + o(1)) \log_{4/3}(\log n)$ (since $1 - \delta \geq \Omega(1)$). Put $h = \text{REC-MAJ-3}_\ell$, so h is a function on $3^\ell = \log^{u'} n$ inputs. Let $g_{n'} = T_n \otimes h$.

By construction, $Z(h, \delta) \leq 1/\log n$. By Corollary 13, $Z(T_n, 1/\log n) \leq O(\log n/n)$. Since h is balanced, by Proposition 7 we get $Z(g_{n'}, \delta) \leq O(\log n/n)$. The result follows, since as a function of n' , $O(\log n/n)$ is $\log^{1+u'} n'/n'$ (taking u' slightly larger to kill any constant factors). □

Remark 1 *It seems possible to improve Theorem 4 to $\log^{2+o(1)} n/n$ by using recursive tribes functions, possibly in the manner of Theorem 1. However the fact that the tribes functions are not perfectly balanced is a stumbling block.*

Theorem 4 and Corollary 13 are close to being tight, as we shall now show. In Theorem 3.1 of [KKL88] it is shown that for all $f : \Omega \rightarrow \{-1, +1\}$, if $\min\{\mathbf{P}[f = 1], \mathbf{P}[f = -1]\} = p$, then $II(f) \geq \Omega(p^2 \log^2 n/n)$, where $II(f)$ is as defined at the end of Section 1.1. Note that when f is monotone, $II(f) = \sum_{|S|=1} \hat{f}^2(S)$. Hence we conclude that for monotone f ,

$$Z(f, \eta) = \sum_S \eta^{|S|} \hat{f}^2(S) \geq \sum_{|S|=1} \eta \hat{f}^2(S) = \eta II(f) \geq \eta \Omega(p^2 \log^2 n/n).$$

But note also that for any f , $Z(f, \eta) \geq Z(f, 0) = \mathbf{E}[f]^2 = (1 - 2p)^2$. Hence, if $p < 1/4$, say, then $Z(f, \eta) \geq 1/4$, and this is much greater than $\eta \Omega(\log^2 n/n)$. Hence we conclude:

Proposition 14 *If $f : \Omega_n \rightarrow \{-1, +1\}$ is any monotone function, then $Z(f, \eta) \geq \Omega(\eta \log^2 n/n)$.*

This proposition shows that Theorem 4 is tight up to a factor $\log^{2.818} n$ factor, and Corollary 13 is tight up to a constant factor. In particular, when the initial correlation η is $O(1/\log n)$, the tribes function by itself is maximally sensitive among monotone functions (to within a constant factor).

6 Fourier concentration around $\sqrt{n}$

It seems natural to combine via Proposition 7 the functions from Theorems 1, 2, and 3 which are somewhat sensitive to very small amounts of noise with the function from Theorem 4 which is very sensitive to moderate noise.

Using the Talagrand construction, we prove:

Proposition 15 *There exists an infinite family of monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ such that for all large n , $Z(f_n, 1 - 1/Q) \leq \epsilon$, where:*

$$Q = \frac{\sqrt{n\epsilon}}{\log^{(1+u')/2}(1/\epsilon)},$$

and $(1 + u')/2 = 2.409 \dots$

Proof: Let $f_n = T_{n_2} \otimes \text{Tal}_{n_1}$ where $n_1 = Q^2$, $n_2 = n/Q^2 = \log^{1+u'}(1/\epsilon)/\epsilon$, Tal_{n_1} denotes the function on n_1 inputs from Theorem 3, and T_{n_2} denotes the tribes function on n_2 inputs. Applying Proposition 7 with Theorems 3 and 4, we get $Z(f_n, 1 - 1/Q) \leq \log^{1+u'}(n_2)/n_2 \leq \epsilon$ (changing u' slightly as necessary). $\square$

For certain applications (such as hardness amplification within NP as discussed in Section 1.4), one might prefer an explicit function family. Using the previous proof with the function from Theorem 2 in place of Talagrand's function, we get:

Proposition 16 *There exists an explicit infinite family of monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ computable in $\mathbf{P}$ such that for all large n : $Z(f_n, 1 - 1/Q') \leq \epsilon$, where:*

$$Q' = \frac{\sqrt{n\epsilon}}{\log^t(n\epsilon) \log^{(1+u')/2}(1/\epsilon)},$$

$t = .3257\dots$, and $(1 + u')/2 = 2.409\dots$

Using the relationship $Z(f, \eta) = \sum_S \eta^{|S|} \hat{f}^2(S)$, it's easy to derive the following from Proposition 15:

Corollary 17 *There exists an explicit infinite family of monotone functions $f_n : \Omega_n \rightarrow \{-1, +1\}$ satisfying for all large n :*

$$\sum_{|S| \leq Q} \hat{f}_n^2(S) \leq \epsilon,$$

where $Q = \tilde{\Omega}(\sqrt{n\epsilon})$ is the quantity from Proposition 16.

Remark 2 *Gil Kalai (private communication) noted to us that Corollary 17 is tight up to polylogarithmic factors. Indeed, let f be a monotone function and note that if $\sum_{|S| \leq Q} \hat{f}^2(S) \leq \epsilon$, then*

$$\sum_{|S|=1} |\hat{f}(S)| = \sum_S |S| \hat{f}^2(S) \geq Q(1 - \epsilon).$$

On the other hand, by Cauchy-Schwartz,

$$\sum_{|S|=1} |\hat{f}(S)| \leq \sqrt{n} \sqrt{\sum_{|S|=1} \hat{f}^2(S)} \leq \sqrt{n\epsilon}.$$

We thus obtain $\sqrt{n\epsilon} \geq Q(1 - \epsilon)$, or $Q \leq \frac{\sqrt{n\epsilon}}{1 - \epsilon}$, thus showing Corollary 17 is tight up to polylogarithmic factors.

Acknowledgments: We would like to thank Gil Kalai for providing encouragement to write this paper and for Remark 2, Yuval Peres for interesting discussions, and the anonymous referees for helpful comments.

References

- [AN72] K. Athreya, P. Ney. Branching processes. Springer-Verlag, New York-Heidelberg, 1972.
- [BDG88] J. Balczár, J. Díaz, J. Gabarró. Structural Complexity I, II. Springer-Verlag, Heidelberg, 1988.

- [BJT99] N. Bshouty, J. Jackson, T. Tamon. Uniform-distribution attribute noise learnability. *Workshop on Computational Learning Theory*, 1999.
- [BKS99] I. Benjamini, G. Kalai, O. Schramm. Noise sensitivity of boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, 1999.
- [BL90] M. Ben-Or, N. Linial. Collective coin flipping. In *Randomness and Computation*, S. Micali ed. Academic Press, New York, 1990.
- [BT96] N. Bshouty, C. Tamon. On the Fourier spectrum of monotone functions. *Journal of the ACM* 43(4), 1996.
- [DK00] D.-Z. Du, K.-I Ko. Theory of Computational Complexity. Wiley Interscience, New York, 2000.
- [F98] E. Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica* 18(1), 1998, 27–36.
- [FK96] E. Friedgut, G. Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.* 124, 1996, 2993–3002.
- [H99] S. Haykin. Neural Networks, 2nd Edition. Prentice Hall, 1999.
- [J97] J. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. *Journal of Computer and System Sciences*, 55(3), 1997.
- [KKL88] J. Kahn, G. Kalai, N. Linial. The influence of variables on boolean functions. *Foundations of Computer Science*, 1988.
- [K66] D. Kleitman. Families of non-disjoint subsets. *Journal of Combinatorial Theory*, 1966.
- [KOS02] A. Klivans, R. O’Donnell, R. Servedio. Learning intersections and thresholds of halfspaces. To appear.
- [LMN93] N. Linial, Y. Mansour, N. Nisan. Constant depth circuits, Fourier transform, and learnability. *J. Assoc. Comput. Mach.* 40, 1993, 607–620.
- [M94] Y. Mansour. Learning boolean functions via the Fourier transform. *Theoretical Advances in Neural Computing and Learning*, Kluwer Acad. Publ., Dordrecht (1994), 391–424.
- [M98] E. Mossel. Recursive reconstruction on periodic trees. *Random Structures Algorithms*, 13, 1998, no. 1, 81–97.
- [O02] R. O’Donnell. Hardness amplification within NP. Symposium on the Theory Of Computation, 2002.
- [Pa93] C. Papadimitriou. Computational Complexity. Addison Wesley, Reading, MA, 1993.
- [Pe98] Y. Peres. Personal communication, 1998.
- [T94] M. Talagrand. On Russo’s approximate 0-1 law. *Annals of Probability*, 1994.

- [T96] M. Talagrand. How much are increasing sets positively correlated? *Combinatorica* 16, 1996, no. 2, 243–258.
- [T97] M. Talagrand. On boundaries and influences. *Combinatorica*, 1997.
- [V84] L. Valiant. A theory of the learnable. *Communications of the ACM*, 40, 1994, no. 2, 445–474.