Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 174 (2012)

On the Noise Stability of Small De Morgan Formulas

Anat Ganor* Ilan Komargodski* Troy Lee† Ran Raz‡

Abstract

We show a connection between the De Morgan formula size of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and the noise stability of the function. Specifically, we prove that for $0 < p \leq 1/2$ it holds that

$$\mathbf{NS}_p(f) \geq 1 - 2p\sqrt{L(f) \cdot \|f\|^2 \cdot (1 - \|f\|^2)}$$

where $\mathbf{NS}_p(f)$ is the noise stability of f with noise parameter p, $\|f\|$ is the $\mathcal{L}_2$ norm of f, and $L(f)$ is the De Morgan formula size of f. This result stems from a generalization of Khrapchenko's bound [Khr71], that might be of independent interest.

Our main result implies the following lower bound:

$$\sum_{S\subseteq[n]} \delta^{|S|}\widehat{f}(S)^2 \geq \|f\|^2 - \frac{1-\delta}{2}\sqrt{L(f)\|f\|^2(1 - \|f\|^2)}$$

for $0 \leq \delta \leq 1$, where $\widehat{f}(S)$ is the Fourier coefficient of f at S. In particular, this bound implies a concentration result on the spectrum of Boolean functions that can be computed by small De Morgan formulas. Specifically, for any $\varepsilon > 0$, we show that $\sum_{S\subseteq[n], |S|<k} \widehat{f}(S)^2 \geq \|f\|^2 (1-\varepsilon)$ where k is roughly $\frac{1}{2\varepsilon}\sqrt{L(f)\frac{1-\|f\|^2}{\|f\|^2}}$. We observe that this concentration result also stems from a relation between the average sensitivity of f and the original Khrapcheko bound.

In addition, we show that the De Morgan formula size in the results mentioned above can be replaced by the square of the non-negative quantum adversary bound, thus giving a (possibly) tighter bound.

*Weizmann Institute of Science, Rehovot 76100, Israel. Email: {anat.ganor, ilan.komargodski}@weizmann.ac.il. Research supported by an Israel Science Foundation grant and by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation.

†Centre for Quantum Technologies, National University of Singapore, Singapore 117543. Email: troyjlee@gmail.com.

‡Weizmann Institute of Science, Rehovot 76100, Israel and Institute for Advanced Study, Princeton, NJ. Email: ran.raz@weizmann.ac.il. Research supported by an Israel Science Foundation grant, by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation, and by NSF grants number CCF-0832797, DMS-0835373.

ISSN 1433-8092

1 Introduction

Noise stability is a central tool for measuring the complexity of Boolean functions. Noise stability appears in many applications in various fields of theoretical computer science such as social choise [Kal02] and hardness of approximation [Kho02]. Moreover, noise stability has many interesting connections to isoperimetric inequalities [Tal95, BKS98] and to Boolean function analysis [MOO05].

Intuitively, the noise stability with parameter $p \in [0,1]$ of a function $f : \{0,1\}^n \to \{0,1\}$ answers the following question. Given $f(x)$ for $x \in \{0,1\}^n$, what can we say about $f(x \oplus e)$ where $e \in \{0,1\}^n$ is a vector such that e_i (the i^{th} coordinate of e) is 1 with probability p and is 0 otherwise? In other words, noise stability (with some noise parameter p) of a function tells us how the output of a function reacts to noise applied to its input. If the noise stability is close to 1, then the function is stable and its output is not affected much by noise applied to its input. On the other hand, if a function has noise stability that is close to 0, then the output of the function is very sensitive to noise applied to its input.

Definition 1.1 (Noise Stability). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For $p \in [0,1]$ and $x \in \{0,1\}^n$, define $N_p(x)$ to be the distribution of a random element $y \in \{0,1\}^n$ which satisfies $\Pr[x_i \neq y_i] = p$, independently for all $i \in [n]$. The p-noise stability of f is*

$$\mathbf{NS}_p(f) = \Pr_{\substack{x \in \{0,1\}^n, \\ y \sim N_p(x)}} [f(x) = f(y)] .$$

In this work, we show a connection between the *De Morgan formula size* of a Boolean function and the *noise stability* of the function. A *De Morgan formula* is a Boolean formula over the basis $B_2 = \{\vee, \wedge, \neg\}$ with fan in at most 2. A De Morgan formula is represented by a tree such that every leaf is labeled by an input variable and every internal node is labeled by an operation from B_2. A formula is said to compute a function $f : \{0,1\}^n \to \{0,1\}$ if on input $x \in \{0,1\}^n$ it outputs $f(x)$. The computation is done in the natural way from the leaves to the root. The *size* of a formula is defined as the number of leaves it contains. For a Boolean function f let $L(f)$ be the size of the smallest De Morgan formula that computes f. Let $\|f\|$ be the $\mathcal{L}_2$ norm of f.

In [Khr71] Khrapchenko gives a technique for proving lower bounds on the De Morgan formula size of Boolean functions. In this paper, we prove a generalization of Khrapchenko's technique and use it in order to prove the following theorem.

Theorem 1.2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For every $p \in (0, 1/2]$*

$$\mathbf{NS}_p(f) \geq 1 - 2p\sqrt{L(f) \cdot \|f\|^2 \cdot (1 - \|f\|^2)}.$$

Theorem 1.2 implies a lower bound (that depends on $L(f)$) on the sum $\mathcal{W}_\delta(f) = \sum_{S \subseteq [n]} \delta^{|S|} \widehat{f}(S)^2$ for every $0 \leq \delta \leq 1$, where $\widehat{f}(S)$ is the Fourier coefficient of f at S. Bounds on $\mathcal{W}_\delta(f)$ may be very useful and appear in many works (e.g., [KKL88, Tal95]). Formally, we prove the following corollary.

Corollary 1.3. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For every $0 \le \delta \le 1$*

$$\sum_{S\subseteq[n]} \delta^{|S|}\widehat{f}(S)^2 \ge \|f\|^2 - \frac{1-\delta}{2}\sqrt{L(f)\|f\|^2(1-\|f\|^2)}.$$

Corollary 1.3 implies a concentration result on the spectrum of functions that can be computed by small De Morgan formulas.

Theorem 1.4. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function computable by a De Morgan formula of size s. Then, for any $\varepsilon > 0$, letting $k = \frac{1}{2\varepsilon}\sqrt{s\frac{1-\|f\|^2}{\|f\|^2}} - \frac{1-\|f\|^2-\varepsilon}{\varepsilon}$ it holds that*

$$\sum_{\substack{S\subseteq[n],\\ |S|<k}} \left(\widehat{f}(S)\right)^2 \ge \|f\|^2 (1-\varepsilon).$$

We observe that Theorem 1.4 also stems from a relation between the average sensitivity of f and the original Khrapcheko bound, as we show in Appendix A.

In addition, we show that the De Morgan formula size in the results mentioned above (Theorem 1.2, Corollary 1.3 and Theorem 1.4) can be replaced by the square of the *non-negative quantum adversary bound* (see Section 5). By replacing the De Morgan formula size with the square of the non-negative quantum adversary bound, we get (possibly) stronger bounds since for every Boolean function f it holds that $L(f) \ge \mathbf{Adv}(f)^2$ [LLS06], where $\mathbf{Adv}(f)$ is the non-negative quantum adversary bound of f.

1.1 Previous Work

As we have already mentioned, our main result follows from a generalization of Khrapchenko's bound on the size of De Morgan formulas [Khr71]. Various generalizations of Khrapchenko's bound were used in the past in numerous works. Zwick [Zwi91] extended the definition of formula size to handle weighted input variables and generalized Khrapchenko's bound to cover the new definition. Koutsoupias [Kou93] was able to extend Khrapchenko's bound with a spectral version to give better lower bounds for specific functions. Håstad [Hås98] showed that the shrinkage exponent of Boolean De Morgan formulas (for the exact definition see [Hås98]) is 2. One of the components in his proof is a lower bound on the De Morgan formula size that depends on the probability that some restrictions occur (for the exact formulation see [Hås98]). Håstad proves that indeed this lower bound is a generalization of Khrapchenko's bound. Laplante, Lee and Szegedy [LLS06] introduce a new complexity measure for Boolean functions that is a lower bound on the De Morgan formula size. They show that several De Morgan formula size lower bounds (including [Khr71, Kou93, Hås98]) are, in fact, a special case of their method.

In comparison with our Theorem 1.4, previous works give an upper bound on the degree of an approximating polynomial using tools from quantum complexity. Specifically, for every Boolean function f, Beals et al. [BBC+01] show that if f has a q-query bounded-error quantum algorithm (in the black box model), then there exists a polynomial of degree at most $2q$ that approximates f. Moreover, in a line of works in quantum query complexity [FGG08, ACR$^+$07, RS08, Rei09, Rei11] it is shown that if a Boolean function $f : \{0,1\}^n \to \{0,1\}$ can be computed by a De Morgan formula of size s, then there is a quantum black box algorithm that computes f in $O(\sqrt{s})$ queries, suffering from a point-wise error of $1/3$. By repeating independent applications of the algorithm, one can increase the number of queries to $O(t \cdot \sqrt{s})$ and reduce the point-wise error to 2^{-t}. Combining both of these results proves that every function $f : \{0,1\}^n \to \{0,1\}$ that can be computed by a De Morgan formula of size s can be approximated by a polynomial of degree $O(t \cdot \sqrt{s})$ up to point-wise error of 2^{-t}.

2 Preliminaries

We start with some general notation. We denote by $[n]$ the set of numbers $\{1, 2, \ldots, n\}$. For $i \in [n]$ and for $x \in \{0,1\}^n$, denote by x_i the i-th bit of x. We denote by $\Delta(x, y)$ the Hamming distance between two strings $x, y \in \{0,1\}^n$ (i.e., the number of coordinates in which x and y differ). In addition, for simplicity, we define $\frac{0}{0} = 0$.

2.1 De Morgan Formulas

Throughout the paper we will only consider De Morgan formulas and not always explicitly mention it.

Definition 2.1. *A De Morgan formula is a Boolean formula with AND, OR and NOT gates with fan in at most 2.*

Definition 2.2. *The size of a formula F is the number of leaves in it and is denoted by $L(F)$. For a function $f : \{0,1\}^n \to \{0,1\}$, we will denote by $L(f)$ the size of the smallest formula computing the function f.*

2.2 Fourier Analysis

For each $S \subseteq [n]$, define $\chi_S : \{0,1\}^n \to \{-1,1\}$ as $\chi_S(x) = \prod_{i \in S} (-1)^{x_i}$. It is well known that the set $\{\chi_S\}_{S \subseteq [n]}$ is an orthonormal basis (called the Fourier basis) for the space of all functions $f : \{0,1\}^n \to \mathbb{R}$. It follows that every function $f : \{0,1\}^n \to \mathbb{R}$ can be represented as

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \chi_S(x)$$

where $\widehat{f} : \{0,1\}^n \to \mathbb{R}$, and $\widehat{f}(S)$ is called the Fourier coefficient of f at $S \subseteq [n]$.

Definition 2.3. *We define the inner product $\langle \cdot, \cdot \rangle$ on pairs of functions $f, g : \{0,1\}^n \to \mathbb{R}$ by*

$$\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) g(x) = \mathop{\mathbb{E}}_{x \in \{0,1\}^n} [f(x) g(x)]$$

Proposition 2.4. *For $f : \{0,1\}^n \to \mathbb{R}$ and $S \subseteq [n]$, the Fourier coefficient of f at S is*

$$\widehat{f}(S) = \langle f, \chi_S \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)\chi_S(x)$$

Proposition 2.5. *Consider functions $f, g : \{0,1\}^n \to \mathbb{R}$. Since $\{\chi_S\}_{S \subseteq [n]}$ forms an orthonormal basis for the space of functions from $\{0,1\}^n$ to $\mathbb{R}$, we get Plancherel's theorem*

$$\langle f, g \rangle = \sum_{S,T \subseteq [n]} \widehat{f}(S)\widehat{g}(T) \langle \chi_S, \chi_T \rangle = \sum_{S \subseteq [n]} \widehat{f}(S)\widehat{g}(S)$$

In particular, the orthonormality of the basis gives the known Parseval theorem

$$\sum_{S \subseteq [n]} \left(\widehat{f}(S)\right)^2 = \langle f, f \rangle = \|f\|^2$$

where $\|f\|$ denotes the $\mathcal{L}_2$-norm of f.

2.3 Matrix Analysis

Definition 2.6 (Matrix Norm)**.** *The norm of an m by n matrix A over the reals is defined as (see e.g., [HJ90])*

$$\|A\| = \max_{u \in \mathbb{R}^m, v \in \mathbb{R}^n} \frac{u^T A v}{\ell_2(u) \cdot \ell_2(v)}$$

and for $x \in \mathbb{R}^n$, $\ell_2(x) = \sqrt{\sum_{i=1}^n x_i^2}$ (where x_i is the i^{th} coordinate of x).

Definition 2.7 (Kronecker Product)**.** *Let $A \in \mathbb{R}^{k \times l}$ and $B \in \mathbb{R}^{m \times n}$ be matrices. The* Kronecker product *(also known as the* tensor product*) of A and B is a matrix in $\mathbb{R}^{km \times ln}$ and is defined as*

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1l}B \\ a_{21}B & a_{22}B & \dots & a_{1l}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1}B & a_{k2}B & \dots & a_{kl}B \end{bmatrix}$$

We also denote by $A^{\otimes t} \in \mathbb{R}^{k^t \times l^t}$ the Kronecker product of A with itself t times.

3 Generalization of Khrapchenko's Bound

In this section we generalize Khrapchenko's bound on the size of De Morgan formulas. We begin by recalling the original Khrapchenko bound.

Theorem 3.1 ([Khr71]). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and let $A \subseteq f^{-1}(1), B \subseteq f^{-1}(0)$. Denote $H(A,B) = \{(a,b) | a \in A, b \in B, \Delta(a,b) = 1\}$. It holds that*

$$L(f) \geq \mathcal{K}(A,B) = \frac{|H(A,B)|^2}{|A| \cdot |B|}$$

In this section we prove a lower bound for formula size that can be interpreted as a generalized version of Khrapchenko's theorem. Let $A, B \subseteq \{0,1\}^n$ and $p \in [0,1]$, we define

$$H_p(A,B) = \sum_{a \in A, b \in B} p^{\Delta(a,b)} (1-p)^{n-\Delta(a,b)}$$

Theorem 3.2 (Generalized Khrapchenko bound). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and let $A \subseteq f^{-1}(1), B \subseteq f^{-1}(0)$. It holds that for any $0 < p \leq 1$,*

$$L(f) \geq \mathcal{K}_p(A,B) = \frac{(H_p(A,B))^2}{|A| \cdot |B| \cdot p^2}$$

Proof. The proof follows the lines of the proof of Khrapchenko's bound from [Weg87] (see Section 8.8 there).

For $f : \{0,1\}^n \to \{0,1\}$ denote $K_p(f) = \max_{A \subseteq f^{-1}(1), B \subseteq f^{-1}(0)} \{\mathcal{K}_p(A,B)\}$. It is enough to prove that $K_p(f)$ is a formal complexity measure (see Lemma 8.1 in [Weg87]). In order to do so, we prove 3 properties of $K_p(f)$, following the original proof, as follows:

- $\forall i \in [n] : K_p(x_i) \leq 1$. Each vector in A (or B, symmetrically) contributes at most $\sum_{i=0}^{n-1} \binom{n-1}{i} p^{i+1} (1-p)^{n-i-1} = p$ to $H_p(A,B)$. It follows that $K_p(x_i) \leq 1$.
- $K_p(\neg f) = K_p(f)$. The definition of $K_p(f)$ is symmetric with respect to A and B.
- $K_p(f \vee g) \leq K_p(f) + K_p(g)$. We choose $A \subseteq (f \vee g)^{-1}(1)$ and $B \subseteq (f \vee g)^{-1}(0)$ such that $\mathcal{K}_p(A,B) = K_p(f \vee g)$. Since $B \subseteq (f \vee g)^{-1}(0)$, then $B \subseteq f^{-1}(0)$ and $B \subseteq g^{-1}(0)$. Partition A into disjoint $A_f \subseteq f^{-1}(1)$ and $A_g \subseteq g^{-1}(1)$. Then $H_p(A,B) = H_p(A_f,B) + H_p(A_g,B)$. Then,

$$K_p(f \vee g) = \frac{(H_p(A_f,B) + H_p(A_g,B))^2}{(|A_f| + |A_g|)|B|p^2}$$

$$K_p(f) + K_p(g) \geq \frac{(H_p(A_f,B))^2}{|A_f||B|p^2} + \frac{(H_p(A_g,B))^2}{|A_g||B|p^2}$$

The claim now follows (as done in [Zwi91]) since for every $a_1, a_2 \in \mathbb{R}$ and every $b_1, b_2 > 0$ it holds that

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \geq \frac{(a_1 + a_2)^2}{b_1 + b_2}$$

□

Remark 3.3. *We consider our bound as a generalization of Khrapchanko's bound since for every Boolean function* $f : \{0,1\}^n \to \{0,1\}$ *and* $A \subseteq f^{-1}(1), B \subseteq f^{-1}(0)$ *it holds that*

$$\lim_{p \to 0} \mathcal{K}_p(A, B) = \mathcal{K}(A, B).$$

We end this section with a lemma that will be useful for the rest of the paper.

Lemma 3.4. *Let* $f : \{0,1\}^n \to \{0,1\}$ *be a Boolean function such that* $L(f) = s$. *For* $A \subseteq f^{-1}(1)$, $B \subseteq f^{-1}(0)$ *and* $C = \{0,1\}^n \setminus (A \cup B)$ *it holds that*

$$H_p(A, A) = |A| - H_p(A, B) - H_p(A, C)$$

and thus

$$H_p(A, A) \quad \geq \quad |A| - \sqrt{s \cdot |A| \cdot |B|} \cdot p - H_p(A, C)$$

Notice that when $C = \emptyset$ *(that is* $A = f^{-1}(1)$ *and* $B = f^{-1}(0)$*), it holds that* $H_p(A, C) = 0$.

Proof of Lemma 3.4. First, it is clear, by the definition, that $H_p(A, A) = H_p(A, A \cup B \cup C) - H_p(A, B) - H_p(A, C)$. Second, we notice that $H_p(A, A \cup B \cup C) = |A| \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = |A|$, which proves the equality of the lemma. For the second part, using Theorem 3.2 we get that $s = L(f) \geq \frac{(H_p(A,B))^2}{|A| \cdot |B| \cdot p^2}$. So $\sqrt{s \cdot |A| \cdot |B|} \cdot p \geq H_p(A, B)$ which proves the inequality of the lemma. □

4 Generalized Khrapchenko and Noise Stability

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, let $A \subseteq f^{-1}(1)$, $B \subseteq f^{-1}(0)$ and $C = \{0,1\}^n \setminus (A \cup B)$. In this subsection we bound $\mathbf{NS}_p(f)$ in terms of $H_p(A, B)$ and $H_p(A, C)$.

Lemma 4.1. *Let* $f : \{0,1\}^n \to \{0,1\}$ *be a Boolean function and* $p \in [0, 1]$. *Let* $A \subseteq f^{-1}(1)$, $B \subseteq f^{-1}(0)$ *and* $C = \{0,1\}^n \setminus (A \cup B)$. *It holds that*

$$\mathbf{NS}_p(f) \geq \left(1 - \frac{|C|}{2^n}\right) - \frac{2H_p(A, B) + H_p(A \cup B, C)}{2^n}$$

Specifically, if $C = \emptyset$, *it holds that*

$$\mathbf{NS}_p(f) = 1 - \frac{2H_p(A, B)}{2^n}$$

Proof. By the definition of noise stability

$$\begin{aligned} \mathbf{NS}_p(f) &= \Pr_{\substack{x \in \{0,1\}^n, \\ y \sim N_p(x)}} [f(x) = f(y)] \\ &\geq \Pr_{\substack{x \in \{0,1\}^n, \\ y \sim N_p(x)}} [x \in A \wedge y \in A] + \Pr_{\substack{x \in \{0,1\}^n, \\ y \sim N_p(x)}} [x \in B \wedge y \in B] \end{aligned} \tag{4.1}$$

Using simple manipulations we get that

$$\begin{aligned}
\Pr_{\substack{x\in\{0,1\}^n,\\ y\sim N_p(x)}}[x\in A\wedge y\in A] &= \sum_{x'\in\{0,1\}^n}\Pr_{\substack{x\in\{0,1\}^n,\\ y\sim N_p(x)}}[x\in A\wedge y\in A|x=x']\Pr_{x\in\{0,1\}^n}[x=x'] \\
&= \frac{1}{2^n}\sum_{x'\in\{0,1\}^n}\Pr_{y\sim N_p(x)}[x\in A\wedge y\in A|x=x'] \\
&= \frac{1}{2^n}\sum_{x'\in A}\Pr_{y\sim N_p(x')}[y\in A] \\
&= \frac{1}{2^n}\sum_{x'\in A}\sum_{y'\in\{0,1\}^n}\Pr_{y\sim N_p(x')}[y\in A|y=y']\Pr_{y\sim N_p(x')}[y=y'] \\
&= \frac{1}{2^n}\sum_{x'\in A}\sum_{y'\in A}\Pr_{y\sim N_p(x')}[y=y'] \\
&= \frac{1}{2^n}\sum_{x'\in A}\sum_{y'\in A}p^{\Delta(x',y')}(1-p)^{n-\Delta(x',y')} \\
&= \frac{1}{2^n}H_p(A,A)
\end{aligned}$$

An analogous calculation shows that

$$\Pr_{\substack{x\in\{0,1\}^n,\\ y\sim N_p(x)}}[x\in B\wedge y\in B] = \frac{1}{2^n}H_p(B,B)$$

Plugging these back into equation (4.1), we get that

$$\begin{aligned}
\mathbf{NS}_p(f) &\geq \frac{1}{2^n}\left(H_p(A,A)+H_p(B,B)\right) \\
&= \frac{1}{2^n}\left(|A|+|B|-2H_p(A,B)-H_p(A\cup B,C)\right) \\
&= \left(1-\frac{|C|}{2^n}\right)-\frac{2H_p(A,B)+H_p(A\cup B,C)}{2^n}
\end{aligned}$$

where the first equality follows from Lemma 3.4.

Notice that if $C=\emptyset$, then the inequality in equation (4.1) becomes an equality (from which the equality in the lemma follows). □

4.1 Proof of Theorem 1.2

Let $f:\{0,1\}^n\to\{0,1\}$ be a Boolean function computable by a De Morgan formula of size s. Denote $A=f^{-1}(1)$, $B=f^{-1}(0)$. Using Lemma 4.1 (where $C=\emptyset$) and Theorem 3.2, we get that

$$\begin{aligned}\mathbf{NS}_p(f) &= 1 - \frac{2H_p(A,B)}{2^n} \\ &\geq 1 - \frac{2 \cdot p\sqrt{s \cdot |A| \cdot |B|}}{2^n}\end{aligned}$$

which proves Theorem 1.2.

4.2 Proof of Corollary 1.3

A well known relation between the noise stability of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ and its Fourier expansion (see e.g., [BKS98, BJT99, O'D02]) is the following.[1]

Proposition 4.2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For every $p \in (0, 1/2]$ it holds that*

$$\mathbf{NS}_p(f) = 1 - 2\widehat{f}(\emptyset) + 2\sum_{S \subseteq [n]} (1-2p)^{|S|} \left(\widehat{f}(S)\right)^2.$$

Combining Theorem 1.2 together with Proposition 4.2 and letting $\delta = 1 - 2p$ we get Corollary 1.3.

5 Generalized Khrapchenko and the Adversary Bound

In this section we show that the generalized Khrapchenko bound (Theorem 3.2) is at most the square of the non-negative quantum adversary bound. The non-negative quantum adversary bound is defined as follows.

Definition 5.1 (Non-Negative Quantum Adversary Bound)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. Let Γ be a matrix with rows indexed by elements of $f^{-1}(0)$ and columns indexed by elements from $f^{-1}(1)$. For every $i \in [n]$ let $\Gamma_i(x,y) = \Gamma(x,y)$ if $x_i \neq y_i$ and 0 otherwise. Then*

$$\mathbf{Adv}(f) = \max_{\Gamma \neq 0, \Gamma \geq 0} \frac{\|\Gamma\|}{\max_{i \in [n]} \|\Gamma_i\|}$$

where the definition of the norm of an m by n matrix appears in Definition 2.6.

Lemma 5.2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function, $A \subseteq f^{-1}(1)$, $B \subseteq f^{-1}(0)$ and $p \in (0,1]$. Then*

$$\mathbf{Adv}(f) \geq \frac{H_p(A,B)}{\sqrt{|A||B|} \cdot p}$$

where $H_p(A,B) = \sum_{a \in A, b \in B} p^{\Delta(a,b)} (1-p)^{n-\Delta(a,b)}$, as defined in Section 3.

[1] We state this connection for Boolean function from $\{0,1\}^n$ to $\{0,1\}$ and not from $\{-1,1\}^n$ to $\{-1,1\}$, as it is usually done. The proof of Proposition 4.2 can be derived following the same lines as the proof of O'Donnell [O'D02] for the $\{-1,1\}$ case.

Using this lemma, the De Morgan formula size in our results (Theorem 1.2, Corollary 1.3 and Theorem 1.4) can be replaced by the square of the non-negative quantum adversary bound (following the same proofs). As we have already mentioned, by replacing the De Morgan formula size with the square of the non-negative quantum adversary bound, we get stronger bounds since for every Boolean function f it holds that $L(f) \geq \mathbf{Adv}(f)^2$ [LLS06].

Note that a simple corollary of Lemma 5.2 is that $\mathbf{Adv}(f)^2$ is lower bounded by the (standard) Khrapchenko bound (taking $p \to 0$. See Remark 3.3). This lower bound is known due to [LLS06].

Proof of Lemma 5.2. Define the matrix Γ as follows:

$$\Gamma(x,y) = \begin{cases} p^{\Delta(x,y)}(1-p)^{n-\Delta(x,y)} & x \in A \text{ and } y \in B \\ 0 & \text{otherwise} \end{cases}.$$

Denote by $\mathbf{1}_A$ and $\mathbf{1}_B$ the characteristic vectors of the sets A and B, respectively. It holds that

$$\|\Gamma\| \geq \frac{\mathbf{1}_A^T \Gamma \mathbf{1}_B}{\sqrt{|A||B|}} = \frac{H_p(A,B)}{\sqrt{|A||B|}}$$

where $\mathbf{1}_A^T$ denotes the vector $\mathbf{1}_A$ transposed.

Now we upper bound $\|\Gamma_i\|$ (see Definition 5.1). Let

$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & p \\ p & 0 \end{bmatrix}$$

Clearly the spectral norm of P is 1 and of Q is p. Moreover, notice that $P^{\otimes n}$ is a 2^n by 2^n matrix where, if we label rows and columns by binary strings of length n in lexicographic order, then the (x,y)-entry is $p^{\Delta(x,y)}(1-p)^{n-\Delta(x,y)}$ (recall Definition 2.7). Thus, Γ is a submatrix of $P^{\otimes n}$. Similarly, Γ_i is a submatrix of $R_i = P^{\otimes i-1} \otimes Q \otimes P^{\otimes n-i}$. Thus the spectral norm of Γ_i is at most the spectral norm of R_i which is p. This proves the lemma. □

Acknowledgements

We thank Shengyu Zhang for helpful remarks and specifically for his contribution to Appendix A.

References

[ACR+07] Andris Ambainis, Andrew M. Childs, Ben Reichardt, Robert Spalek, and Shengyu Zhang. Any AND-OR formula of size n can be evaluated in time $n^{1/2+o(1)}$ on a quantum computer. In *FOCS*, pages 363–372. IEEE Computer Society, 2007.

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.

[BJT99] Nader H. Bshouty, Jeffrey C. Jackson, and Christino Tamon. Uniform-distribution attribute noise learnability. In *Workshop on Computational Learning Theory*, pages 75–80, 1999.

[BKS98] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of Boolean functions and applications to percolation, 1998.

[FGG08] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum algorithm for the hamiltonian NAND tree. *Theory of Computing*, 4(1):169–190, 2008.

[GKR12] Anat Ganor, Ilan Komargodski, and Ran Raz. The spectrum of small deMorgan formulas. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:174, 2012.

[Hås98] Johan Håstad. The shrinkage exponent of de Morgan formulae is 2. *SIAM J. Comput.*, 27(1):48–64, 1998.

[HJ90] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

[Kal02] G. Kalai. A fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Advances in Applied Mathematics*, 29(3):412–426, 2002.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *IEEE Conference on Computational Complexity*, page 25, 2002.

[Khr71] V.M. Khrapchenko. A method of determining lower bounds for the complexity of π schemes. *Matematicheskie Zametki*, 10:83–92, 1971. In Russian.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on Boolean functions (extended abstract). In *FOCS*, pages 68–80. IEEE Computer Society, 1988.

[Kou93] E. Koutsoupias. Improvements on Khrapchenko's theorem. *Theoretical Computer Science*, 116(2):399–403, August 1993.

[LLS06] Sophie Laplante, Troy Lee, and Mario Szegedy. The quantum adversary method and classical formula size lower bounds. *Computational Complexity*, 15(2):163–196, 2006.

[MOO05] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *FOCS*, pages 21–30. IEEE Computer Society, 2005.

[O'D02] Ryan O'Donnell. Hardness amplification within np. In John H. Reif, editor, *STOC*, pages 751–760. ACM, 2002.

[Rei09] Ben Reichardt. Span programs and quantum query complexity: The general adversary bound is nearly tight for every Boolean function. In *FOCS*, pages 544–551. IEEE Computer Society, 2009.

[Rei11] Ben Reichardt. Reflections for quantum query algorithms. In Dana Randall, editor, *SODA*, pages 560–569. SIAM, 2011.

[RS08] Ben Reichardt and Robert Spalek. Span-program-based quantum algorithm for evaluating formulas. In Cynthia Dwork, editor, *STOC*, pages 103–112. ACM, 2008.

[Tal95] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathematiques de l'IHES*, 81(1):73–205, 1995.

[Weg87] Ingo Wegener. *The complexity of Boolean functions*. Wiley-Teubner, 1987.

[Zwi91] Uri Zwick. An extension of Khrapchenko's theorem. *Inf. Process. Lett.*, 37(4):215–217, 1991.

A Concentration using Khrapchenko's Bound

We show that the Fourier spectrum of Boolean functions that can be computed by small De Morgan formulas is concentrated. Recall that for a function $f : \{0,1\}^n \to \{0,1\}$ the $\mathcal{L}_2$ norm of f is denoted $\|f\|$, and for a subset $S \subseteq [n]$, the Fourier coefficient of f at S is denoted $\widehat{f}(S)$.

Theorem 1.4 (Restated). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function computable by a De Morgan formula of size s. Then, for any $\varepsilon > 0$, letting $k = \frac{1}{2\varepsilon}\sqrt{s\frac{1-\|f\|^2}{\|f\|^2}} - \frac{1-\|f\|^2-\varepsilon}{\varepsilon}$ it holds that*

$$\sum_{\substack{S\subseteq[n],\\ |S|<k}} \left(\widehat{f}(S)\right)^2 \geq \|f\|^2 (1-\varepsilon).$$

The idea of the proof of this theorem was communicated to us by Shengyu Zhang. This proof improves the result that appeared in a previous version of the paper [GKR12].

We begin with the definition of average sensitivity.

Definition A.1 (Average Sensitivity)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. The average sensitivity (also known as total influence) of f is*

$$\mathbf{AS}(f) = \sum_{i=1}^{n} \Pr_{x\in\{0,1\}^n}\left[f(x) \neq f(x^{(i)})\right]$$

where $x^{(i)}$ is x with the i^{th} bit flipped.

Proof of Theorem 1.4. Using the fact that $\mathbf{AS}(f) = 4\sum_{S\subseteq[n]} |S|\widehat{f}(S)^2$ [KKL88] and splitting the sum of the right-hand side to "light" Fourier coefficients and "heavy" ones, we get that

$$\begin{aligned}
\frac{\mathbf{AS}(f)}{4} &= \sum_{\substack{S\subseteq[n],\\ 0<|S|<k}} |S|\widehat{f}(S)^2 + \sum_{\substack{S\subseteq[n],\\ |S|\geq k}} |S|\widehat{f}(S)^2 \\
&\geq \sum_{\substack{S\subseteq[n],\\ 0<|S|<k}} \widehat{f}(S)^2 + \sum_{\substack{S\subseteq[n],\\ |S|\geq k}} k\widehat{f}(S)^2 \\
&= \sum_{\substack{S\subseteq[n],\\ |S|<k}} \widehat{f}(S)^2 - \|f\|^4 + k\left(\|f\|^2 - \sum_{\substack{S\subseteq[n],\\ |S|<k}} \widehat{f}(S)^2\right)
\end{aligned}$$

where we used the fact that $\widehat{f}(\emptyset) = \|f\|^2$ and Parseval's theorem (see Proposition 2.5). Therefore,

$$\begin{aligned}
\sum_{\substack{S\subseteq[n],\\ |S|<k}} \widehat{f}(S)^2 &\geq \frac{1}{k-1}\left(k\|f\|^2 - \|f\|^4 - \frac{\mathbf{AS}(f)}{4}\right) \\
&= \frac{\|f\|^2}{k-1}\left(k - \|f\|^2 - \frac{\mathbf{AS}(f)}{4\|f\|^2}\right) \\
&= \|f\|^2\left(1 - \frac{1}{k-1}\left(\frac{\mathbf{AS}(f)}{4\|f\|^2} - (1-\|f\|^2)\right)\right).
\end{aligned}$$

Let $A = f^{-1}(1)$ and $B = f^{-1}(0)$. By the definition of average sensitivity together with Khrapchenko's bound (Theorem 3.1) we get that

$$\mathbf{AS}(f) = \sum_{i=1}^{n} \Pr_{x\in\{0,1\}^n}\left[f(x) \neq f(x^{(i)})\right] = \frac{2|H(A,B)|}{2^n} \leq \frac{2\sqrt{s\cdot|A||B|}}{2^n}$$

and therefore,

$$\sum_{\substack{S\subseteq[n],\\ |S|<k}} \widehat{f}(S)^2 \geq \|f\|^2\left(1 - \frac{1}{k-1}\left(\frac{1}{2}\sqrt{s\frac{1-\|f\|^2}{\|f\|^2}} - (1-\|f\|^2)\right)\right).$$

□

ECCC ISSN 1433-8092

http://eccc.hpi-web.de