

# Notes on lower bounding formulas

November 5, 2015

## 1 Instance compression for formula lower bounds

**Theorem 1.1** (Main Theorem To prove). *For a boolean formula  $f$  with  $n$  variables, there exists a function  $h_f : \{0, 1\}^n \rightarrow \{0, 1\}^{n^\alpha}$  where  $\alpha \leq 1$  and a set  $A \subset \{0, 1\}^{n^\alpha}$  such that  $L(h_f) = o(L(f)^{1-\alpha})$  and with probability  $1/2 + \Omega(\delta)$ ,*

$$h_f(x) \in A \leftrightarrow f(x) = 1.$$

Denote  $I_A$  is the indicator function for the set  $A$ , that is  $I_A(x) = 1 \leftrightarrow x \in A$ . Then the function  $I_A(h_f(\cdot))$   $(1/2 + \Omega(\delta))$ -agrees with  $f(\cdot)$ . If  $f$  is average hard for  $n^3$ -sized formula and  $k = o(n)$ , then

$$L(I_A) \geq L(f)/L(h_f) = n^{3\alpha}/o(1) \gg n^{3\alpha}$$

. Note that  $I_A$  depends on  $n^\alpha$  variables, which means the size of  $I_A$  is larger than the cube of its size.

### 1.0.1 Negative results for Instance compression

**Theorem 1.2** ([2]). *If there is a set  $A$  and function  $h$  such that given  $m$  Boolean formula  $\phi_1, \dots, \phi_m$  where each  $\phi_i$  has length at most  $n$ ,  $h$  has the following properties:*

1.  $h$  is computable in time polynomial in  $m$  and  $n$ ,
2.  $h(\phi_1, \dots, \phi_m) \in A$  if and only if at least one of the  $\phi_i$  is satisfiable, and
3.  $|h(\phi_1, \dots, \phi_m)|$  is bounded by a polynomial in  $n$ ,

*then  $NP$  is contained in  $coNP/poly$ .*

To apply the above theorem to our setting, we need transfer the satisfiable problem to computable problem...

**Theorem 1.3** (Probabilistic version [2]). *If  $OR-SAT$  is probabilistically compressible with randomness complexity  $O(\log(n))$ , then  $PH$  collapses.*

## 2 Generalize the PRG from shrinkage

First of all, from the paper [4], we clean out two main lemmas by viewing the restriction as a special function.

**Lemma 2.1.** *Let  $\mathcal{F} : \{0, 1\}^n \rightarrow \{0, 1\}$  be a class of de Morgan formulas with an associated size function  $s : \mathcal{F} \rightarrow \mathbb{N}$ . Then, there exists a function  $h : \{0, 1\}^{s^{2/3}} \rightarrow \{0, 1\}^n$  such that  $s^{1/3}$ -wise independent input  $G_{s^{1/3}}$  could  $\epsilon$ -fool  $f \circ h(\cdot)$ , that is*

$$|\mathbb{E}[f \circ h(G_{s^{1/3}})] - \mathbb{E}[f \circ h(U_{s^{2/3}})]| \leq \epsilon.$$

**Lemma 2.2.** *For sufficiently large  $t$ , there exists shrinkage functions  $h_1, \dots, h_t$  such that the distribution  $\oplus_{i=1}^t f \circ h_i(U_{s^{2/3}})$  is  $\epsilon$ -close to the distribution  $f(U_n)$ .*

In the Lemma ??, view  $h(G_{s^{1/3}})$  and  $h(U_{s^{2/3}})$  as two different  $n$ -bits codes  $C_1, C_2$ . Then Lemma ?? shows that no  $n^3$ -sized de Morgan formula  $f$  can distinguish  $C_1, C_2$  within  $\epsilon$ .

**Question 2.3.** *For any constant  $\epsilon < 0$ , are there two codes  $C_1, C_2 \subset \{0, 1\}^n$  such that  $|C_1| = o(|C_2|)$  and no  $n^4$ -sized formula  $f$  could distinguish them within error  $\epsilon$ , namely*

$$|E_{x \in C_1}[f(x)] - E_{x \in C_2}[f(x)]| \leq \epsilon$$

## 3 Main Goal

**Observation 3.1.** *For any formula  $f$  on  $k$  bits and PRG  $G$  whose input is  $T$  bits and output is  $k$  bits, if  $f(G(x|_\rho)) = f(U_k)$ , then we have  $L(f(G)) \geq (T/k)^2 L(f)$ .*

However, if  $L(f) = k^d$ , then we can only get  $L(f(G)) \geq T^2 k^{d-2} = O(T^d)$ . That is, this method cannot improve our conclusion. But it sheds light on how to composite functions to increase the formula complexity. Furthermore, the Andreev function is the special case when  $k = \log n$  and  $T = n$  where  $L(f) = \log n$ .

Actually, there has been a similar conjecture:

**Conjecture 3.2** (KRW conjecture). *Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  and  $g : \{0, 1\}^m \rightarrow \{0, 1\}$  be non-constant functions. Then*

$$D(g \circ f) \simeq D(g) + D(f),$$

where  $D(f)$  denotes the depth of  $f$  and  $g \circ f$  represents the composition of functions, i.e.  $g \circ f = g(f(\bar{x}_1), \dots, f(\bar{x}_m))$ , in which  $\bar{x}_i$  is the vector of  $n$  distinct variables.

Our first theorem shows that,

**Theorem 3.3.** *For any formula  $F$  with  $k$  variables and formula  $G$  with  $T$  variables, if  $F \circ G(x|_\rho) = F(U_k)$  and  $L(G) \leq T^2$ , then the KRW conjecture is true.*

**Question 3.4.** 1. How about  $L(G) \geq T^2$ ? Can we develop new technique to prove it, instead of using the Hastad's shrinkage argument?

2. Maybe the PRG can have more property, rather than only fixing the bit-fixing source.
3. Furthermore, to prove a better lower bound, we don't need such strong conjecture. If  $F$  has more properties, e.g.  $F \in AC^0$  and  $F$  is average hardness, we may avoid the difficulty of proving this conjecture. This part is covered in Section 5. I am not sure whether it is workable?
4. What can be revealed for the formula from the random formula restriction?

### 3.1 Related work

In the paper [3], Gavinsky, Meir, Weinstein and Avi Wigderson tried to use information-theoretic approach to prove it.

They have observed the Andreev functions is just  $L(g \circ \oplus_m) = \Omega(L(g) \cdot L(\oplus_m))$ . Thus, the  $F(f_1, \dots, f_n) = t^{d+1}/polylog(t)$  where  $t = n \log n$

And they made a new conjecture about  $\oplus_m \circ f$ , However, this is a significant difference: when the parity function is at the bottom, one can easily apply random restrictions to the function, but it is not clear how to do it when the parity function is at the top.

In Andreev function, we only need  $\log n$  variables left but now we need  $n$ .

## 4 Another Way— bound the variance of shrinkage

First of all, we know the shrinkage size  $L(\phi|_\rho)$  is a random variable with expected size  $p^2 L$  and has a high upper concentration result. We proceed to ask what about  $Var[L(\phi|_\rho)]$

We can use some non-malleable code to construct a function with high variance after shrinkage. But it is a bit involved to bound the variance of shrinkage....

## 5 Assuming $F$ is AC0 and has average-case hardness

By [5], we know

**Theorem 5.1.** *There is an explicit Boolean function  $h : \{0, 1\}^n \rightarrow \{0, 1\}$  and a constant  $c \geq 8$  and a constant  $c \geq 8$  such that for any parameter  $r$  such that  $c \log(n) \leq r \leq n^{1/3}$ , any formula of size  $\frac{n^3 - o(1)}{r^2}$  computes  $h$  correctly on a fraction of at most  $1/2 + 2^{-\Omega(r)}$  of the inputs.*

Furthermore, by the [1],

**Theorem 5.2.** *Let  $s \geq \log m$  be any parameter. Let  $F$  be a Boolean function computed by a circuit of depth  $d$  and size  $m$ . Let  $\mu$  be an  $r$ -independent distribution where*

$$r \geq r(s, d) = 3 \cdot 60d + 3 \cdot (\log m)$$

*then*

$$|Pr_{x \sim U_{poly \log n}}[f(x = 1)] - Pr_{x \sim U_n}[f(x) = 1]| = |E_\mu[F] - E[F]| < \epsilon(s, d),$$

*where  $\epsilon(s, d) = 0.82s \cdot (10m)$ .*

**Theorem 5.3.**  $r(m, d, \epsilon)$ -independence  $\epsilon$ -fools depth- $d$  AC0 circuits of size  $m$ , where

$$r(m, d, \epsilon) = \log\left(\frac{m}{\epsilon}\right)^{O(d^2)}$$

Now, we consider the following composition  $f \circ G$  on polylog many bits, where  $G$  is pseudorandom generator to generate polylog-wise independent distribution from  $\text{poly log } n$  inputs. That is,

$$\left| \Pr_{x \sim U_{\text{poly log } n}}[f(x) = 1] - \Pr_{y \in \{0,1\}^{\text{poly log } n}}[f(G(y)) = 1] \right| \leq \delta$$

Note that instead of fooling the formula on uniform random distribution, what we want to fool here is the formula over  $\text{poly log } n$ -wise independent distribution. This should give an exponentially error bound, rather than the polynomially small compared with the uniform distribution case.

Concluding with the above two theorems, we get

$$\left| \Pr_{x \sim U_n}[f(x) = 1] - \Pr_{y \in \{0,1\}^{\text{poly log } n}}[f(G(y)) = 1] \right| \leq \delta + \epsilon.$$

where  $\epsilon, \delta$  could be both exponentially small. Then we hope to prove the following theorem, and we could get a much better lower bound than  $\Omega(n^3)$ .

Then, by the paper [?], we want to show the indistinguishability can imply a good approximation for each  $f$ . If so, we can have the following property.

**Theorem 5.4.** *If there are two functions  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  and  $g : \{0, 1\}^n \rightarrow \{0, 1\}$  such that  $f$  approximates  $g$  within  $1/2 + \epsilon$ , and for any  $s$ -sized formula  $\Phi$ ,*

$$\Pr_x[f(x) = \Phi(x)] \leq 1/2 + \epsilon;$$

*then we know*

$$L(g) \geq s$$

## 5.1 Approximate Boolean function by using PRG

**Definition 5.5.** *For two functions  $f, g : \{0, 1\}^l \rightarrow \{0, 1\}$  and a number  $0 \leq \rho \leq 1$  we say that  $g$  approximates  $f$  within a factor  $\rho$  if  $f$  and  $g$  agree on at least a fraction  $\rho$  of their domain, i.e.*

$$\Pr[f(x) = g(x)] \geq \rho$$

**Theorem 5.6** (Main Theorem To prove). *If there is a PRG  $G : \{0, 1\}^k \rightarrow \{0, 1\}^n$  that  $\delta$ -fools a class of Boolean functions  $C$  with  $n$  variables, then for each function  $f \in C$ , there exists a function  $h_f : \{0, 1\}^n \rightarrow \{0, 1\}^k$  such that  $L(h_f) = o(L(f))$  and  $f(G(h_f(\cdot)))$  approximates  $f$  within  $1/2 + \Omega(\delta)$ .*

*Proof.* That is, we need to bound

$$\Pr[h_f(x) \in (f(G))^{-1}(1) | f(x) = 1] + \Pr[h_f(x) \in (f(G))^{-1}(0) | f(x) = 0].$$

And what we have is a precise estimate of  $E[f(x)]$  by using PRG on  $E[f(G(x))]$ . Could Fourier transform help us here? Since we know the first coefficient in the Fourier transform of the function  $f$ . □

## References

- [1] Mark Braverman. Polylogarithmic independence fools ac 0 circuits. *Journal of the ACM (JACM)*, 57(5):28, 2010.
- [2] Lance Fortnow and Rahul Santhanam. Infeasibility of instance compression and succinct pcps for np. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 133–142. ACM, 2008.
- [3] Dmitry Gavinsky, Or Meir, Omri Weinstein, and Avi Wigderson. Toward better formula lower bounds: An information complexity approach to the krw composition conjecture. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 213–222. ACM, 2014.
- [4] Russell Impagliazzo, Raghu Meka, and David Zuckerman. Pseudorandomness from shrinkage. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 111–119. IEEE, 2012.
- [5] Ilan Komargodski, Ran Raz, and Avishay Tal. Improved average-case lower bounds for demorgan formula size. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 588–597. IEEE, 2013.