Notes on lower bounding formulas

November 3, 2015

1 Generalize the PRG from shrinkage

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

Lemma 1.1. *Let $\mathcal{F} : \{0,1\}^n \to \{0,1\}$ be a class of de Morgan formulas with an associated size function $s : \mathcal{F} \to \mathbb{N}$. Then, there exists a function $h : \{0,1\}^{s^{2/3}} \to \{0,1\}^n$ such that $s^{1/3}-$ wise independent input $G_{s^{1/3}}$ could fool $f \circ h(\cdot)$ within ϵ, that is*

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

Lemma 1.2. *For sufficiently large t, there exists shrinkage functions $h_1, \ldots, h_t$ such that the distribution $\oplus_{i=1}^t f \circ h_i(U_{s^{2/3}})$ is $\epsilon-$close to the distribution $f(U_n)$*

2 Main Goal

Observation 2.1. *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 2.2 (KRW conjecture)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$ 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. $g \circ f = g(f(\overline{x_1}), \ldots, f(\overline{x_m}))$, in which $\overline{x_i}$ is the vector of n distinct variables.

Our first theorem shows that,

Theorem 2.3. *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 2.4.
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 4. I am not sure whether it is workable?*
4. *What can be revealed for the formula from the random formula restriction?*

2.1 Related work

In the paper [**?**], 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 $F(f_1, \ldots, f_n) = t^{d+1}/polylog(t)$ 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.

3 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....

4 Assuming F is AC0 and has average-case hardness

By [**?**], we know

Theorem 4.1. *There is an explicit Boolean function $h : \{0,1\}^n \to \{0,1\}$ 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 [**?**],

Theorem 4.2. *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 μ 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 4.3. *$r(m,d,\epsilon)$-independence ϵ-fools depth-d AC0 circuits of size m, where*

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

.

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

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

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

Concluding with the above two theorems, we get

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

where ϵ, δ 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 4.4. *If there are two functions $f : \{0,1\}^n \to \{0,1\}$ and $g : \{0,1\}^n \to \{0,1\}$ such that f approximates g within $1/2 + \epsilon$, and for any s-sized formula Φ,*

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

then we know

$$L(g) \geq s$$

4.1 Approximate Boolean function by using PRG

Definition 4.5. *For two functions $f, g : \{0,1\}^l \to \{0,1\}$ and a number $0 \leq \rho \leq 1$ we say that g approximates f within a factor ρ if f and g agree on at least a fraction ρ of their domain, i.e.*

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

Theorem 4.6 (Main Theorem To prove)**.** *If there is a PRG $G : \{0,1\}^k \to \{0,1\}^n$ that δ-fools a class of Boolean functions C with n variables, then for each function $f \in C$, there exists a function $h_f : \{0,1\}^n \to \{0,1\}^k$ 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 transfrom of the function f. □