Poly-logarithmic independence fools AC^0 circuits

Mark Braverman

Let μ be a distribution on the $\{0,1\}^n$. We denote by $\mathbf{E}_\mu[F]$ the expected value of F on inputs drawn according to μ . For an event X, we denote by $\mathbf{P}_\mu[X]$ its probability under μ . If the subscript is missing, uniform distribution is considered.

The distribution μ on $\{0,1\}^n$ is r-independent if every restriction of μ to any r coordinates is uniform on $\{0,1\}^r$.

A distribution μ is said to ε-fool a function F if

$$|\mathbf{E}_\mu[F] - \mathbf{E}[F]| \le \varepsilon.$$

An AC^0 circuit is a circuit with AND, OR and NOT gates, where the fan-in of the gates is unbounded. The depth of a circuit is the maximum number of AND/OR gates between input and output.

Main Problem: How large does $r = r(m,d,\varepsilon)$ have to be in order for every r-independent distribution μ on $\{0,1\}^n$ to ε-fool every function F that is computed by a depth-d AC^0 circuit of size $\le m$?

Theorem *L.M.J. Bazzi, Polylogarithmic independence can fool DNF formulas]***:**

$$r(m,2,\varepsilon) = \mathcal{O}\left(\log^2 \frac{m}{\varepsilon}\right).$$

Main Theorem: Let $s \ge \log m$ be any parameter, F be a boolean function computed by a circuit of depth d and size m, let μ be an $r(s,d)$-independent distribution. Then $|\mathbf{E}_\mu[F] - \mathbf{E}[F]| \le \varepsilon(s,d)$ where

$$r(s,d) = 3 \cdot 60^{d+3} \cdot (\log m)^{(d+1)(d+3)} \cdot s^{d(d+3)} \quad \text{and} \quad \varepsilon(s,d) = 0.82^s \cdot (15m).$$

Corollary of Main Theorem: By taking $s = 5\log(15m/\varepsilon)$ we get

$$r(m,d,\varepsilon) = 3 \cdot 60^{d+3} \cdot (\log m)^{(d+1)(d+3)} \cdot \left(5\log\frac{15m}{\varepsilon}\right)^{d(d+3)} = \left(\log \frac{m}{\varepsilon}\right)^{\mathcal{O}(d^2)}.$$

Proposition 1: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a degree-r polynomial and let μ be an r-independent distribution. Then f is completely fooled by μ : $\mathbf{E}_\mu[f] = \mathbf{E}[f]$.

Lemma 2 *[LMN93] (f with small $||f||_2^2$)***:** If $F : \{0,1\}^n - > \{0,1\}$ if a boolean function computable by a depth-d circuit of size m, then for every t there is a degree t polynomial $\tilde{f}$ with

$$||F - \tilde{f}||_2^2 = 2^{-n} \sum_{x \in \{0,1\}^n} |F(x) - \tilde{f}(x)|^2 \le 2m \cdot 2^{-t^{1/d}/20}.$$

Lemma 3 *(f with small $\mathbf{P}_\nu[f \ne F]$)***:** Let ν be any probability distribution of $\{0,1\}^n$. For a circuit of depth d and size m computing a function F, for any s, there is a degree $r = (s \cdot \log m)^d$ polynomial f and a depth $< d+3$ boolean function $\mathcal{E}_\nu$ of size $\mathcal{O}(m^2 r)$ such that

- $\mathbf{P}_\nu[\mathcal{E}_\nu(x) = 1] < 0.82^s m$,
- whenever $\mathcal{E}_\nu(x) = 0$ then $f(x) = F(x)$,
- for $s \ge \log m$, $||f||_\infty < (2m)^{\deg(f)-2} = (2m)^{(s\log m)^d - 2}$.

Lemma 4 *($F' \approx F$ and f' with small both $\mathbf{P}_\nu[F' \ne f']$ and $||F' - f'||_2^2$)***:** Let F be computed by a circuit of depth d and size m. Let s_1, s_2 be two parameters with $s_1 \ge \log m$ and let μ be any probability distribution on $\{0,1\}^n$. Set $\nu = 1/2(\mu + \mathcal{U}_{\{0,1\}^n})$. Let $\mathcal{E}_\nu$ be the function from Lemma 3 with $s = s_1$. Set $F' = F \vee \mathcal{E}_\nu$. Then there is a polynomial f' of degree $r \le (s_1 \cdot \log m)^d + s_2$, such that

- $\mathbf{P}[F \ne F'] < 2 \cdot 0.82^{s_1} m$,
- $\mathbf{P}_\mu[F \ne F'] < 2 \cdot 0.82^{s_1} m$,
- $||F' - f'||_2^2 < 0.28^{s_1} \cdot (4m) + 2^{2.9(s_1 \cdot \log m)^d \cdot \log m - s_2^{1/(d+3)}/20}$,
- $f'(x) = 0$ whenever $F'(x) = 0$.

Lemma 5 *($F' \approx F$ and f'_l with small $\mathbf{E}[F' - f'_l]$)*: For every boolean circuit F of depth d and size m and any $s \geq \log m$ and for any probability distribution μ on $\{0,1\}^n$ there is a boolean function F' and a polynomial f'_l of degree less than $r = 3 \cdot 60^{d+3} \cdot (\log m)^{(d+1)(d+3)} \cdot s^{d(d+3)}$ such that

- $\mathbf{P}[F \neq F'] < \varepsilon(s,d)/3$,
- $\mathbf{P}_\mu[F \neq F'] < \varepsilon(s,d)/3$,
- $f'_l \leq F'$ on $\{0,1\}^n$,
- $\mathbf{E}[F' - f'_l] < \varepsilon(s,d)/3$,

for $\varepsilon(s,d) = 0.82^s \cdot (15m)$.

Lemma 6 *(one-sided ε-fooling)*: Let $s \geq \log m$ be any parameter, F be a boolean function computed by a circuit of depth d and size m, let μ be an r-independent distribution where $r \geq 3 \cdot 60^{d+3} \cdot (\log m)^{(d+1)(d+3)} \cdot s^{d(d+3)}$. Then

$$\mathbf{E}_\mu[F] > \mathbf{E}[F] - \varepsilon(s,d)$$

where $\varepsilon(s,d) = 0.82^s \cdot (15m)$.

Constant Depth Circuits, Fourier Transform, and Learnability, *[LMN93]*

Boolean functions on n variables will be considered as real valued functions $f : \{0,1\}^n \to \{-1,1\}$. The set of all real functions on a cube is a 2^n-dimensional real vector space with scalar product defined as $\langle g, f \rangle = 2^{-n} \sum_{x \in \{0,1\}^m} f(x)g(x) = \mathbf{E}[gf]$.

For S a subset of $\{1,\ldots,n\}$ we define $\chi_S(x_1,\ldots,x_n) = [\sum_{i \in S} x_i \text{ is odd}]$. Then χ_S forms an orthogonal basis of real-valued functions on a cube, so every such $f = \sum_S \tilde{f}(S)\chi_S$, where $\tilde{f}(S) = \langle f, \chi_s \rangle$. Orthonormality of the basis implies $||f||^2 = \sum_S \tilde{f}(S)^2$. Finally, the degree of a Boolean function, deg(f), is the size of the largest set S such that $\tilde{f}(S) \neq 0$. This equals the degree of f as a multi-linear polynomial.

A random restriction ρ with parameter p is the mapping of variables to 0, 1 and *, where probability of * is p and the probability of 0 and 1 is $(1-p)/2$. The function obtained from $f(x_1,...,x_n)$ by applying a random restriction ρ is f_ρ, its variables are those x_i which $\rho(x_i) = *$, all other variables set according to ρ .

Lemma 1 *(Hastad)*: Let f is a CNF formula where each clause has size at most t. Then with probability at lest $1 - (5pt)^s$ can f_ρ be expressed as a DNF formula each clause of which has size at most s and all the clauses accept disjoint sets of inputs.

Lemma 2 *(iterated Hastad)*: Let f be a Boolean function computed by a circuit of size m and depth d. Then $\mathbf{P}[\deg(f_\rho) > s] \leq m2^{-s}$ where ρ is a random restriction with $p = 10^{-d}s^{-(d-1)}$.

Lemma 3: Let f be a Boolean function and let S be arbitrary subset. For any $B \subset S$ we have $\sum_{C \subset S^c} \tilde{f}(B \cup C)^2 = 2^{-|S^c|} \sum_{R \in \{0,1\}^{S^c}} \tilde{f}_{S^c \leftarrow R}(B)^2$.

Lemma 4: Let f be a Boolean function, S arbitrary subset and k an integer. Then $\sum_{A, |A \cap S| > k} \tilde{f}(A)^2 = \mathbf{E}_R[\sum_{|B|>k} \tilde{f}_{S^c \leftarrow R}(B)^2] \leq \mathbf{P}_R[\deg(f_{S^c \leftarrow R}) > k]$, with R a random 0-1 assignment to the variables in S^c.

Lemma 5: Let f be a Boolean function, $t \in \mathbb{N}$, $0 < p < 1$. Then $\sum_{|A|>t} \tilde{f}(A)^2 \leq 2\mathbf{E}_S[\sum_{|A \cap S| > pt/2} \tilde{f}(A)^2]$, where S is chosen such that each variable appears in it independently with probability p, $pt > 8$.

Lemma 6: Let f be a Boolean function computed by a circuit of depth d ans size m and let t be any integer. Then $\sum_{|A|>t} \tilde{f}(A)^2 \leq 2m \cdot 2^{-t^{1/d}/20}$.