CS 6817: Special Topics in Complexity Theory **Spring 2025**

Lecture 13: March 6, 2025

Lecturer: Eshan Chattopadhyay *Scribe: Santiago Lai*

1 Recap from Lecture 12

Proposition 1.1. *Let f be computed by a size s DNF. Then f is ϵ-close to a function g computed by a DNF with width $w = \log(s/\epsilon)$.*

Proposition 1.2. *Let f be computed by a width w DNF. Then its influence satisfies $I(f) \leq 2w$.*

Corollary 1.3. *Let f be computed by a width w DNF. Then for $\epsilon > 0$, the Fourier coefficients of f is ϵ-concentrated on degree up to $I(f)/\epsilon = 2w/\epsilon$.*

Combining Proposition 1.1 and Corollary 1.3 yields the following theorem:

Theorem 1.4. *For any size s DNF, its Fourier coefficients are ϵ-concentrated up to degree $k = O(1/\epsilon \cdot \log(s/\epsilon))$.*

Corollary 1.5. *There is a low-degree algorithm that runs in time poly$(n^{O(1/\epsilon \cdot \log(s/\epsilon))}, 1/\epsilon)$, and PAC learns DNFs of size s with error 2ϵ.*

The goal of this lecture is to improve the bound in Theorem 1.4 from $O(1/\epsilon \cdot \log(s/\epsilon))$ to $O(\log(1/\epsilon) \cdot \log(s/\epsilon))$. A direct consequence is that when ϵ is a small constant and $s = \text{poly}(n)$, then $n^{O(1/\epsilon \cdot \log(s/\epsilon))} \approx n^{\log(n)}$, and $n^{O(\log(1/\epsilon) \cdot \log(s/\epsilon))}$, which is has a much better dependence on the error parameter. In fact the machinery we will develop can be used to derive a PAC learning algorithm (in the query model) that runs in time $n^{O(\log \log n)}$.

2 Improving the bound via Restrictions

The technique we are going to use is combining random restrictions with Fourier Analysis.

Definition 2.1. *A restriction ρ is a pair (J, z) where $J \subseteq [n]$ is the set of unrestricted variables, and $z \in \{-1, 1\}^n$ represents the restricted values of variables outside J.*

Definition 2.2. *Consider $f : \{-1, 1\}^n \to \mathbb{R}$. The function under restriction of ρ is $f_\rho : \{-1, 1\}^n \to \mathbb{R}$, defined as $f_\rho(x) = f(x_J, z_{\bar{J}})$. Here $y = (x_J, z_{\bar{J}})$ is defined as*

$$y_i = \begin{cases} x_i & \text{if } i \in J, \\ z_i & \text{if } i \notin J. \end{cases}$$

What we are more interested in is random restriction, formally defined as follows.

Definition 2.3. *For $n \geq 0, \delta \in [0, 1]$, consider sampling $\rho = (J, z)$ in the following way. For each $i \in [n]$, i is added to J independently w.p. δ, and $z \sim \{-1, 1\}^n$ is sampled uniformly at random. Such distribution is called a δ-random restriction, denoted by $\rho_{n,\delta}$.*

From the definitions, we can easily compute the Fourier coefficients of f_ρ:

Claim 2.4. $\hat{f}_\rho(S) = (\sum_{T\subseteq \bar{J}} \hat{f}(S\cup T)\chi_T(z_{\bar{J}}))\cdot 1_{S\subseteq J}$.

Proof. Using the definition of Fourier expansion, we have

$$f_\rho(x) = f(x_J, z_{\bar{J}}) = \sum_{S\subseteq J, T\subseteq \bar{J}} \hat{f}(S\cup T)\chi_S(x_J)\chi_T(z_{\bar{J}})$$
$$= \sum_{S\subseteq J}(\sum_{T\subseteq \bar{J}} \hat{f}(S\cup T)\chi_T(z_{\bar{J}}))\chi_S(x_J).$$

The result follows. □

After that, we can compute the expectations of $\hat{f}_\rho(S)$ and $\hat{f}_\rho(S)^2$ over the distribution $\rho_{n,\delta}$.

Claim 2.5. $\mathbb{E}_{\rho\sim\rho_{n,\delta}}[\hat{f}_\rho(S)] = \delta^{|S|}\hat{f}(S)$.

Proof. From Claim 2.4, $\mathbb{E}_\rho[\hat{f}_\rho(S)] = \mathbb{E}_{J,z}[\sum_{T\subseteq\bar{J}}\hat{f}(S\cup T)\chi_T(z_{\bar{J}})\cdot 1_{S\subseteq J}]$. Note that if $S\not\subseteq J$, then

$$\sum_{T\subseteq \bar{J}} \hat{f}(S\cup T)\chi_T(z_{\bar{J}})1_{S\subseteq J} = 0.$$

If $S\subseteq J$, then

$$\mathbb{E}_z[\sum_{T\subseteq\bar{J}}\hat{f}(S\cup T)\chi_T(z_{\bar{J}})1_{S\subseteq J}] = \mathbb{E}_z[\hat{f}(S)\chi_\emptyset(z_{\bar{J}})] + \mathbb{E}_z[\sum_{T\neq\emptyset, T\subseteq\bar{J}}\hat{f}(S\cup T)\chi_T(z_{\bar{J}})]$$
$$= \hat{f}(S) + 0 = \hat{f}(S).$$

Finally, we have

$$\mathbb{E}_{\rho\sim\rho_{n,\delta}}[\hat{f}_\rho(S)] = \mathbb{E}_J[\mathbb{E}_z[\hat{f}_\rho(S)]] = \mathbb{E}_J[\hat{f}_\rho(S)\cdot 1_{S\subseteq J}] = |\delta|^{|S|}\hat{f}_\rho(S).$$

□

Claim 2.6. $\mathbb{E}_{\rho\sim\rho_{n,\delta}}[\hat{f}_\rho(S)^2] = \sum_{W\subseteq[n]}\hat{f}(W)^2\Pr_J[W\cap J = S]$.

Proof. Again by Claim 2.4, $\mathbb{E}_\rho[\hat{f}_\rho(S)^2] = \mathbb{E}_{J,z}[\sum_{T_1,T_2\subseteq\bar{J}}\hat{f}(S\cup T_1)\hat{f}(S\cup T_2)\chi_{T_1}(z_{\bar{J}})\chi_{T_2}(z_{\bar{J}})\cdot 1_{S\subseteq J}]$. If $S\not\subseteq J$ then

$$\sum_{T_1,T_2\subseteq\bar{J}}\hat{f}(S\cup T_1)\hat{f}(S\cup T_2)\chi_{T_1}(z_{\bar{J}})\chi_{T_2}(z_{\bar{J}})\cdot 1_{S\subseteq J} = 0.$$

If $S\subseteq J$ then

$$\mathbb{E}_z[\sum_{T_1,T_2\subseteq\bar{J}}\hat{f}(S\cup T_1)\hat{f}(S\cup T_2)\chi_{T_1}(z_{\bar{J}})\chi_{T_2}(z_{\bar{J}})\cdot 1_{S\subseteq J}] = \mathbb{E}_z[\sum_{T_1,T_2\subseteq\bar{J}}\hat{f}(S\cup T_1)\hat{f}(S\cup T_2)\chi_{T_1\triangle T_2}(z_{\bar{J}})]$$
$$= \sum_{T\subseteq\bar{J}}\hat{f}(S\cup T)^2 = \sum_{W\subseteq[n]}\hat{f}(W)^2\cdot 1_{W\cap J=S}.$$

Combining the above estimates, we have

$$\mathbb{E}_\rho[\hat{f}_\rho(S)^2] = \mathbb{E}_{J,z}[\hat{f}_\rho(S)^2] = \sum_{W\subseteq[n]}\hat{f}(W)^2\Pr_J[W\cap J = S].$$

□

We conclude by presenting Hastad's Switching Lemma. In the next lecture we will combine it with Claim 2.5 and 2.6 to achieve the desired bound.

Lemma 2.7 (Hastad's Switching Lemma)**.** *Suppose that f is computable by width w DNF. Then, for any $d\geq 0$, $\Pr_{\rho\sim\rho_{n,\delta}}[DT(f_\rho)\geq d]\leq(7\delta w)^d$.*