Goldreich- Levin theorem

Let $f:\{0,1\}^n \to \{0,1\}$ be available by query access. There is an algorithm running in time poly$(n, \frac{1}{\epsilon}, \frac{1}{\tau})$ ~~sat~~ that returns a list L of sets $S \subseteq [n]$ with the following property

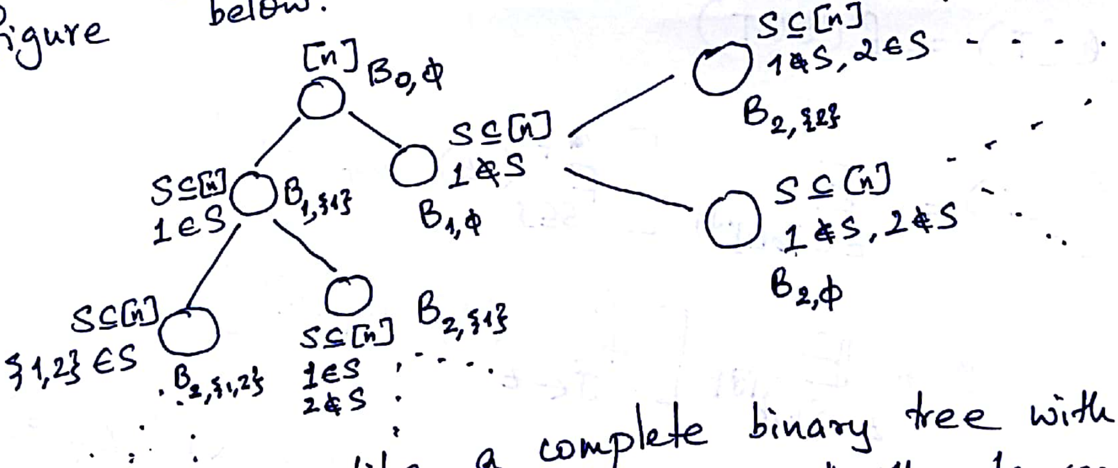
(i) If $|\hat{f}(S)| > \tau$, then $S \in L$

(ii) If $S \in L$, then $|\hat{f}(S)| \geq \tau/2$.

(This is the Kushilevitz-Mansour version of the G-L theorem). It was proved in the context of designing learning algorithms for ~~sp~~ functions with sparse Fourier spectrum.

Proof Idea: We are going to maintain a collection of "buckets", where each bucket contains a few subsets of $[n]$, with the property that for a bucket B, $\sum_{S \in B} \hat{f}(S)^2 > \frac{\tau^2}{2}$.

The way these buckets are maintained as shown in the figure below.



This looks like a complete binary tree with 2^n leaves. We are actually interested in all the leaves which have large Fourier weight. There cannot be many of these (Why?)

Here ~~B_k~~ $B_{1,\{1\}}$ denotes all subsets that contain the element 1. If $W[B_{1,\{1\}}] > \frac{\tau^2}{2}$, we keep going down that subtree, else we remove the bucket $B_{1,\{1\}}$

At the k^{th} level of this binary ~~search~~ tree we have $W[B_{k,S}]$, where $S \subseteq \{1,2,\dots,k\}$

$$W[B_{k,S}] = \sum_{T \subseteq \{k+1,\dots,n\}} \hat{f}(S \cup T)^2 \quad \text{— (1)}$$

So, if we can estimate $\sum \hat{f}(S \cup T)^2$ then we can proceed with the algorithm.

Restrictions

Let $f: \{\pm 1\}^n \to \{\pm 1\}$ & let $J \subseteq [n]$ & $\bar{J} = [n] \setminus J$

"The restriction of f to J using $z \in \bar{J}$" is the function $f_{J \leftarrow z} : \{\pm 1\}^{|J|} \to \{\pm 1\}$

- Let $f: \{\pm 1\}^n \to \mathbb{R}$ & $(J, \bar{J})$ is a partition of $[n]$. The function $F^f_{S \subseteq J} : \{\pm 1\}^{|\bar{J}|} \to \mathbb{R}$ is defined as follows

$$F^f_{S \subseteq J}(z) = \hat{f}_{J \leftarrow z}(S) \quad , \quad \text{Recall that } f_{J \leftarrow z} : \{\pm 1\}^{|J|} \to \{\pm 1\}$$

Lemma $\hat{F}^f_{S \subseteq J}(T) = \hat{f}(S \cup T)$

Proof:
$$\hat{F}^f_{S \subseteq J}(T) = \mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[F^f_{S \subseteq J}(z) \, \chi_T(z) \right]$$
$$= \mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\hat{f}_{\bar{J} \leftarrow z}(S) \, \chi_T(z) \right]$$
$$= \mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\mathop{\mathbb{E}}_{x \in \{\pm 1\}^{|J|}} \left[f_{\bar{J} \leftarrow z}(x) \, \chi_S(x) \right] \chi_T(z) \right]$$
$$= \mathop{\mathbb{E}}_{\substack{z \in \{\pm 1\}^{|\bar{J}|} \\ x \in \{\pm 1\}^{|J|}}} \left[f_{\bar{J} \leftarrow z}(x) \, \chi_S(x) \chi_T(z) \right] \qquad \begin{array}{l} S \subseteq J \\ T \subseteq \bar{J} \\ S \cap T = \phi \end{array}$$
$$= \mathop{\mathbb{E}}_{Y \in \{\pm 1\}^n} \left[f(\underset{J}{\underset{\downarrow}{x}}, \underset{\bar{J}}{\underset{\downarrow}{z}}) \, \chi_{S \cup T}(Y) \right] = \hat{f}(S \cup T)$$

$x \cdot z = Y$

Finally, we are interested in the following:

Lemma: Let $f: \{\pm 1\}^n \to \mathbb{R}$, and let $(J, \bar{J})$ be a partition of $[n]$, Suppose that $S \subseteq J$, then

$$\mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\hat{f}_{\bar{J} \leftarrow z}(S)^2 \right] = \sum_{T \subseteq \bar{J}} \hat{f}(S \cup T)^2$$

Proof: We know that

$$F^f_{S \subseteq J}(z) = \hat{f}_{\bar{J} \leftarrow z}(S)$$

$$\mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\hat{f}_{\bar{J} \leftarrow z}(S)^2 \right] = \mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\left(F^f_{S \subseteq J}(z) \right)^2 \right]$$

$$= \| F^f_{S \subseteq J} \|_2^2 = \sum_{T \subseteq \bar{J}} \hat{F}^f_{S \subseteq J}(T)^2$$

$$= \sum_{T \subseteq J} \hat{f}(S \cup T)^2$$

- Estimating the weight of the buckets.

To estimate the weight of a bucket $B_{k,S}$ (where $S \subseteq \{1, 2, .., k\}$) is to estimate the sum

$$\sum_{T \subseteq \{k+1, .., n\}} \hat{f}(S \cup T)^2.$$

In the formulation above, we have a partition $(J, \bar{J})$ where $J = \{1, 2, .., k\}$,

So, $$\sum_{T \subseteq \{k+1, .., n\}} \hat{f}(S \cup T)^2 = \mathop{\mathbb{E}}_{z \in \{\pm 1\}^{|\bar{J}|}} \left[\hat{f}_{\bar{J} \leftarrow z}(S)^2 \right]$$

where $\bar{J} = \{k+1, .., n\}$

(4)

$$\mathbb{E}_{z\in\{\pm1\}^{|\bar{J}|}}\left[\hat{f}_{\bar{J}\leftarrow z}(S)^2\right] = \mathbb{E}_{z\in\{\pm1\}^{|\bar{J}|}}\left[\left(\mathbb{E}_{x\in\{\pm1\}^{|J|}}\left[f_{\bar{J}\leftarrow z}(x)\,\chi_S(x)\right]\right)^2\right]$$

$$= \mathbb{E}_{z\in\{\pm1\}^{|\bar{J}|}}\left[\mathbb{E}_{x\in\{\pm1\}^{|J|}}\left[f_{\bar{J}\leftarrow z}(x)\chi_S(x)\right]\cdot \mathbb{E}_{y\in\{\pm1\}^{|J|}}\left[f_{\bar{J}\leftarrow z}(y)\chi_S(y)\right]\right]$$

$$= \mathbb{E}_{\substack{z\in\{\pm1\}^{|\bar{J}|}\\ x\in\{\pm1\}^{|J|}\\ y\in\{\pm1\}^{|J|}}}\left[f(x,z)\chi_S(x)\, f(y,z)\,\chi_S(y)\right]$$

The random variable $f(x,z)\cdot f(y,z)\cdot \chi_S(x)\cdot \chi_S(y)$ lies in the set $\{\pm1\}$. So, we can use Hoeffding's to estimate it a very small error.