E0 224 Computational Complexity Theory **Indian Institute of Science, Bangalore**
Fall 2014 **Department of Computer Science and Automation**

Lecture 28: Sep 19, 2014

Lecturer: Chandan Saha <chandan@csa.iisc.ernet.in> *Scribe: Pawan Kumar, Jaiprakash*

28.1 Proof of PCP theorem(weaker version)

In this lecture we will continue the proof of PCP theorem stated in the lecture 27. Notations are borrowed from the previous lecture. Recall from the previous lecture, certificate π will consist of g_1 and g_2, where $g_1 = f_u \in \{0,1\}^{2^n}$ and $g_2 = f_{u\otimes u} \in \{0,1\}^{2^{n^2}}$ are Walsh-Hadamard encoding of strings u and $u \otimes u$ respectively. The verification process consists of the following steps.

- Step 1: Verify that both g_1 and g_2 are W-H(Walsh-Hadamard) codes of some strings u and w i.e. check if
  - $g_1 = f_u$, where $u \in \{0,1\}^n$
  - $g_2 = f_w$, where $w \in \{0,1\}^{n^2}$
- Step 2: Check if $w = u \otimes u$ where u and w are described in step 1.
- Step 3: Check if u is indeed a satisfying assignment for the input system of quadratic equations over $\mathbb{F}_2$ (see previous lecture)

Theorem 28.1. *NP $\subseteq$ PCP(poly(n),1)*

Proof. Proof is continued from the last lecture. □

Claim 28.2. *Walsh-Hadamard codewords of length 2^q are precisely linear functions on q-length strings.*

Proof. See lecture 27 for the prove. □

Step 3 has already been proved in previous lecture, here we will show step 1 and step 2 by using the above claim i.e. it suffices to check whether the given 2^q-length string is a truth table for a linear function on q-length strings. For checking whether a function say f is linear or not, i.e. $(f(x+y) = f(x) + f(y), \forall x, y \in \{0,1\}^q$, where addition of vectors x and y is coordinate wise over $\mathbb{F}_2$) we have to read all 2^q values of f. We will define a test that on one hand accepts every linear function, and on the other hand rejects with high probability every function that is far from linear. For this we need to formally define the closeness between two functions.

Definition 28.3. ***Closeness between two boolean functions:*** *Let g and h be two boolean function (i.e. $g, h : \{0,1\}^q \to \{0,1\}$). We say g is $(1-\epsilon)-$close for $\epsilon \in [0,1]$ to h if*

$$\Pr_{x \in_R \{0,1\}^q} [g(x) = h(x)] \geq 1 - \epsilon$$

Definition 28.4. *We say that a function $g : \{0,1\}^q \to \{0,1\}$ is $(1-\epsilon)-$close to a linear function if there exists a linear function $h : \{0,1\}^q \to \{0,1\}$ s.t. g is $(1-\epsilon)-$close to h.*

28.2 Linearity testing : BLR(Blum, Luby, Rubinfeld '90)

Theorem 28.5. *Let* $g : \{0,1\}^q \to \{0,1\}$. *If*

$$\Pr_{x,y \in_R \{0,1\}^q} [g(x+y) = g(x) + g(y)] \geq 1 - \epsilon \qquad \text{for } \epsilon \in [0,1]$$

then g *is* $(1-\epsilon)-$*close to a linear function.*

Note: If we repeat this test $\mathcal{O}(\frac{1}{\epsilon})$ times the error probability will reduce to $(1-\epsilon)^{\frac{100}{\epsilon}}$, (say).

Remark: The BLR test implies that verifier can check if g_1 and g_2 are $(1-\epsilon)-$close to linear function using $\mathcal{O}(1)$ queries to the proof π and time $poly(q)$. After step 1 verification, verifier knows that w.h.p (say .99) g_1 is $\frac{9}{10}$ close to f_u for some $u \in \{0,1\}^n$, and g_2 is $\frac{9}{10}$ close to f_w for some $w \in \{0,1\}^{n^2}$.

28.3 Local decoding of W-H code

Given $g \in \{0,1\}^{2^q}$ and suppose g is $(1-\epsilon)-$close to f_z for some $z \in \{0,1\}^q$. Since g is $(1-\epsilon)-$ close to f_z our task is to find $f_z(r)$ from g given $r \in \{0,1\}^q$ by using $\mathcal{O}(1)$ queries to g.

- Pick $r_1 \in_R \{0,1\}^q$
- Output $g(r_1) + g(r + r_1)$

Claim 28.6. $f_z(r) = g(r_1) + g(r + r_1)$ *with probability atleast* $1 - 2\epsilon$

Proof. Since f_z is $(1-\epsilon)-$close to g, we have

$$\Pr_{r_1 \in \{0,1\}^q} \{g(r_1) \neq f_z(r_1)\} < \epsilon$$

Note that both r_1 and $r + r_1$ are uniformly distributed over $\{0,1\}^q$.Hence

$$\Pr_{r_1 \in \{0,1\}^q} \{g(r + r_1) \neq f_z(r + r_1)\} < \epsilon$$

Therefore with probability atleast $1 - 2\epsilon$ (by union bound)

$$\begin{aligned} g(r_1) + g(r + r_1) &= f_z(r_1) + f_z(r + r_1) \\ &= f_z(r_1) + f_z(r) + f_z(r_1) && f_z \text{ is linear function} \\ &= f_z(r) && f_z(r_1) + f_z(r_1) = 0 \text{ (over } \mathbb{F}_2 \text{)} \end{aligned}$$

□

Remark: By above claim w.l.g we can assume that we can read arbitrary bits f_u and f_w correctly w.h.p. from g_1 and g_2 where g_1 is $(1-\epsilon)-$close to f_u and g_2 is $(1-\epsilon)-$close to f_w.

Now we indeed to show that $w = u \otimes u$. Let us assume

$$W = (w_{ij})_{i,j \in [n]} \in \mathbb{F}_2^{n \times n}$$

Similarly we can think of $u \otimes u = U \in \mathbb{F}_2^{n \times n}$ we need to check whether $W \stackrel{?}{=} U$. Here is the procedure

Table 28.1: Truth tables before and after notational switch

Boolean world	real world
0 + 0 = 0	1 * 1 = 1
0 + 1 = 1	1 * -1 = -1
1 + 0 = 1	-1 * 1 = -1
1 + 1 = 0	-1 * -1 = 1

- Pick two vectors $r^{'}$ and $r^{''}$ randomly form $\{0,1\}^n$
- Check if $f_u(r^{'})f_u(r^{''}) = f_w(r^{'} \otimes r^{''})$

$$\text{L.H.S.} = f_u(r^{'})f_u(r^{''})$$
$$f_u(r^{'})f_u(r^{''}) = (\sum_{i=1}^{n} u_i r_i^{'})(\sum_{j=1}^{n} u_j r_j^{''})$$
$$= \sum_{i,j\in[n]} u_i u_j r_i^{'} r_j^{''}$$
$$= r^{'T} U r^{''}$$

$$\text{R.H.S.} = f_w(r^{'} \otimes r^{''})$$
$$f_w(r^{'} \otimes r^{''}) = \sum_{i,j\in[n]} w_{ij} r_i^{'} r_j^{''}$$
$$= r^{'T} W r^{''}$$

Hence the above check implies that

$$r^{'T} U r^{''} = r^{'T} W r^{''}$$
$$\Rightarrow r^{'T}(U+W)r^{''} = 0$$
$$r^{'T} V r^{''} = 0 \qquad \text{say } V = U + W$$

If $U \neq W$ then $V \neq O$ in which case $r^{'T} V r^{''} \neq 0$ w.p. $\frac{1}{4}$ as $r^{'}$ and $r^{''}$ are chosen independently.
Remark : after step 2, verifier is convinced that $w = u \otimes u$ w.h.p. using $\mathcal{O}(1)$ reads of π.

28.4 A detour into Fourier analysis

Notational Switch: In the boolean world we were operating over $\mathbb{F}_2$ with variables taking boolean values 0,1. In the real world 0 is mapped to 1, 1 is mapped to -1 and + over $\mathbb{F}_2$ is mapped to multiplication i.e. if v_1, v_2 &v_3 are boolean vectors s.t. $v_1 + v_2 = v_3$, where $+$ is coordinate wise addition over $\mathbb{F}_2$, then after the notational switch $(0 \to 1,\ 1 \to -1)$, $v_1\ o\ v_2 = v_3$ where o is coordinate wise multiplication over reals.

Definition 28.7. ***Linear Functions on q-length strings:*** *A function $f : \{-1,1\}^q \to \{-1,1\}$ is said to be linear function iff $\forall x, y \in \{-1,1\}^q$*

$$f(x\ o\ y) = f(x)f(y)$$

Definition 28.8. ***Equivalent definition of a linear function on q -length strings:*** *A function $f : \{-1,1\}^q \to \{-1,1\}$ if $\exists S \subseteq [q]$ s.t. for every $x \in \{-1,1\}^q$*

$$f = \prod_{i \in S} x_i \quad =: \mathcal{X}_S \qquad \textit{(where } x_i \textit{ is the } i^{th} \textit{ coordinate of } x\textit{)}$$

Definition 28.9. *Nice Inner Product: Let f and g be two real valued function i.e. $f : \{-1,1\}^q \to \mathbb{R}$, $g : \{-1,1\}^q \to \mathbb{R}$. Define an inner product of f and g, denoted by $< f, g >$, as*

$$< f, g > = \frac{1}{2^q} \sum_{x \in \{-1,1\}^q} f(x)g(x) \quad = \mathbb{E}_x \left[f(x)g(x) \right]$$

Remark: Treating f and g as 2^q-dimensional vector in $\mathbb{R}^{2^q}$, the operation $< f, g >$ defines an inner product.
Properties of Inner Product : Let v_1 and $v_2 \in \mathbb{R}^{2^q}$, $\beta \in \mathbb{R}$ then

- $< \beta.v_1, v_2 >= \beta. < v_1, v_2 >=< v_1, \beta.v_2 >$
- $< v_1 + v_2, v_3 >=< v_1, v_3 > + < v_2, v_3 >$
- $< v_1, v_2 + v_3 >=< v_1 + v_2 > + < v_1 + v_3 >$
- $< v_1, v_2 >=< v_2, v_1 >$
- $< v_1, v_1 >= 0$ iff $v_1 = 0$

Lemma 28.10. *If $S, T \subseteq [q]$ and $S \neq T$ then $< \mathcal{X}_S, \mathcal{X}_T >= 0$.*

Proof. Given $\mathcal{X}_S = \prod_{i \in S} x_i, \mathcal{X}_T = \prod_{j \in T} x_j$

$$\begin{aligned}
< \mathcal{X}_S, \mathcal{X}_T > &= \mathbb{E}_x \left[\mathcal{X}_S(x)\mathcal{X}_T(x) \right] \\
&= \mathbb{E}_x \left[\prod_{i \in S \triangle T} x_i \prod_{j \in S \cap T} x_j^2 \right] && \triangle \text{ is symmetric difference} \\
&= \mathbb{E}_x \left[\prod_{i \in S \triangle T} x_i \right] && \text{as } x_j^2 = 1 \\
&= \prod_{i \in S \triangle T} \mathbb{E} \left[x_i \right] && \text{as } x_i's \text{ are independent} \\
&= 0 && \mathbb{E} \left[x_i \right] = 0
\end{aligned}$$

□

Lemma 28.11. *If $S \subseteq [q]$ then $< \mathcal{X}_S, \mathcal{X}_S >= 1$.*

Proof. Given $\mathcal{X}_S = \prod_{i \in S} x_i, \mathcal{X}_T = \prod_{j \in T} x_j$

$$\begin{aligned}
< \mathcal{X}_S, \mathcal{X}_S > &= \mathbb{E}_x \left[\mathcal{X}_S(x)\mathcal{X}_S(x) \right] \\
&= \mathbb{E}_x \left[\prod_{i \in S} x_i^2 \right] \\
&= \mathbb{E}_x \left[1 \right] \\
&= 1
\end{aligned}$$

□

Claim 28.12. *$\mathcal{X}_\phi, ..., \mathcal{X}_{[q]}$ is an orthonormal basis of the space $\mathbb{R}^{2^q}$.*

Proof. Follows from the above lemmas. □

The above claim implies that any vector $f \in \mathbb{R}^{2^q}$ can be uniquely expressed as

$$f = \sum_{S \subseteq [q]} \alpha_S \mathcal{X}_S \qquad \text{where } \alpha_S \in \mathbb{R}$$

Note: This expression is the fourier transform of f and α_S are the fourier coefficients.

The analysis of BLR test will be done in the next lecture i.e. we will prove the following theorem.

Theorem 28.13. *If $Pr\{$BLR accepts $f\} \geq 1 - \epsilon$ then f is $(1-\epsilon)-$close to a linear function.*

References

[M1] S. ARORA and B. BARAK "Computational Complexity: A Mordern Approach," *Cambridge University Press*, 2009