CS 880: Advanced Complexity Theory	2/22/06
Lecture 12: Pseudorandom Generators	
Instructor: Dieter van Melkebeek	Scribe: Vinay Choudhary

The main aim of this lecture is to give another proof of the theorem we saw in the last lecture. Earlier we proved the theorem using the "eigen-value method". This time we will use extractors. First we will see that every expander is essentially an extractor and then we see how to use this idea to get a different proof of the theorem. The exact bounds obtained will be slightly different. Most of the notation we use is from the previous lecture and we do not redefine it here.

1 Extractors

An extractor is a function that takes an input which is sampled from a distribution that can be far from uniform and produces an output which is distributed uniformly. The output has number of bits roughly equal to the amount of randomness in the source. It can be shown that this cannot be achieved by a deterministic method. Extractors typically take a small number of additional random bits.

Definition 1.

$$f : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$$

is a (k, ϵ) extractor if, for any distribution X on $\{0,1\}^n$ with $H_\infty(X) \geq k$, we have,

$$\| f(X, U_t) - U_m \|_1 \leq \epsilon.$$

f can be viewed as extracting the randomness from X which can be thought of as obtained from some physical process that produces some, not completely independent, random bits. k is typically a polynomial or a linear fraction of n while t is typically logarithmic in n.

We will first prove that the expanders are actually extractors with certain parameters.

1.1 Expanders and Extractors

Lemma 1. *For any k,*

$$g_r : \{0,1\}^n \times \{0,1\}^{\log d} \to \{0,1\}^n$$

is a (k, ϵ) extractor for $\epsilon = \lambda_2 2^{\frac{n-k}{2}}$.

Proof. Recall our analysis for random walks on expanders. We showed that when we take one step in a random walk, the 1-norm of the distribution goes down by a factor of λ_2.

$$\| g_r(X, U_{\log d}) - U_n \|_1 \leq \lambda_2 \| X - U \|_1$$

Since U and $X - U$ are orthogonal,

$$\| X - U \|_2^2 + \| U \|_2^2 = \| X \|_2^2$$

Now, $\| X \|_2^2 \leq 2^k 2^{-2k}$ because the min-entropy of X is at least k and $\| X \|_2^2$ is maximized when 2^k components of X are equal to 2^{-k} and the rest are zero. Using this and the equations above, we get

$$\begin{aligned} \| g_r(X, U_{\log d}) - U_n \|_1 &\leq 2^{\frac{n}{2}} \lambda_2 (2^{-k} - 2^{-n})^{\frac{1}{2}} \\ &\leq \lambda_2 2^{\frac{n-k}{2}} \end{aligned}$$

□

We note that these extractors are useful when k is close to n.

2 The Theorem

Theorem 1. *For any distribution X on $\{0,1\}^S$,*

$$\| Z(X, U_r, U_r) - Z(X, g(U_r, U_{\log d})) \| \leq (\lambda_2 2^S)^{\Omega(1)}$$

where $g = (g_l, g_r) : (\rho, \sigma) \to (\rho, \sigma^{th}$ neighbor of $\rho)$

Proof. First let's see the intuition behind the proof. Recall the process from the last lecture where we combined two "blocks" into one. The first block takes input x, random bits ρ_1 and produces output y. The second block takes y, randomness ρ_2 and gives output z. If we do not use any PRG, ρ_1 and ρ_2 are chosen to be completely independent. When we use a PRG to combine these two blocks together, ρ_2 is obtained from ρ_1 (and some small amount of additional randomness). In the last lecture, we obtained ρ_2 by taking one step of a random walk from ρ_1 in an expander. The additional randomness needed was just the edge to be followed.

In this proof, we want to replace ρ_2 by something that is obtained by applying the extractor to ρ_1. This will work provided that the min-entropy (of what?) is small. This can be see as follows

$$Pr[\rho_1 | Y(x, \rho_1) = y] = \begin{cases} \frac{Pr[\rho_1]}{Pr[Y(x,\rho_1)=y]}, & \text{if } Y(x, \rho_1) = y \\ 0, & \text{otherwise} \end{cases}$$

So we are in a good position if the y is "popular" i.e. the probability that the first step produces y is large. If y is "unpopular", then the most intuitive thing to do would be to just ignore it. This can be done if the number of "unpopular" y's is not too large. This is where the space bound will be useful. Now we proceed to the formal proof.

Fix x. Let $H = \{y | Pr[Y(x, \rho_1) - y] \geq \beta\}$ where β is to be chosen later. Fix any $y \in H$. Then using the equation above,

$$\begin{aligned} & H_\infty(\rho_1 | Y(x, \rho_1) = y) \geq n - \log \frac{1}{\beta} \\ \Rightarrow \quad & \| g(\{\rho_1 | Y(x, \rho_1) = y\}, U_{\log d}) - U_n \|_1 \leq \frac{\lambda_2}{\sqrt{\beta}} \qquad \text{(Using Lemma 1)} \end{aligned}$$

Let D_y be the uniform distribution on $\{\rho_1 | Y(x, \rho_1) = y\}$, then

$$\| Z(D_y, U_n) - Z(g(D_y, U_{\log d})) \|_1 \leq \frac{\lambda_2}{\sqrt{\beta}}$$

The first scenario corresponds to picking ρ_1 uniformly but conditioned on the event that the output is the fixed y and then picking ρ_2 uniformly. The second scenario corresponds to picking ρ_1 in the same way but picking ρ_2 using the extractor applied to ρ_1.

Note that it is important to apply Z in the inequality above. In particular, the following need not be true.

$$\| (D_y, U_n) - (g(D_y, U_{\log d})) \|_1 \leq \frac{\lambda_2}{\sqrt{\beta}}$$

This is because for Z, only the output of the first block, i.e. y matters. If we do not apply Z, we are looking at the distribution of ρ_1's which may be widely different.

Since the above holds for all "popular" y's, we get,

$$\| Z(D_H, U_n) - Z(g(D_H, U_{\log d})) \|_1 \leq \frac{\lambda_2}{\sqrt{\beta}}$$

This is because this difference is a convex combination of the differences in the earlier inequality and hence the upper bound holds.

Now for the y's that are "unpopular", we simply have the following

$$N(\overline{H}) \leq 2^S \beta$$

Using these two inequalities,

$$\begin{aligned} \| Z(x, U_n, U_n) - Z(x, g(U_n, U_{\log d})) \|_1 &\leq \mu(H)\frac{\lambda_2}{\sqrt{\beta}} + \mu(\overline{H}) \cdot 2 \\ &\leq \frac{\lambda_2}{\sqrt{\beta}} + 2^{S+1}\beta \end{aligned}$$

The right hand side is minimized when the two terms in the sum are equal. The resulting bound is

$$\| Z(x, U_n, U_n) - Z(x, g(U_n, U_{\log d})) \|_1 \leq 2\lambda_2^{\frac{2}{3}} \cdot 2^{\frac{S+1}{3}}$$

This completes the proof. □

Let's look again the construction of the pseudo-random generator (see Figure 1). Note that the construction does not depend on the space bounds of each of the disjoint processes, but rather on the amount of space they use to communicate with each other. Therefore, we can also use this construction for amplification. If each of these processes were individual runs of a randomized machine with one or two-sided error, we would have a new means of performing amplification or error reduction. In this instance, the amount of communication between processes would be limited to a single bit (indicating success or failure) for machines with one-sided error, or a vote count, for machines with two-sided error.

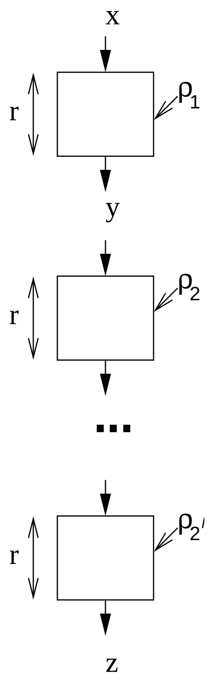

Figure 1: The Leaf Level of the PRG Construction from Lecture 11