CS395T Pseudorandomness and Combinatorial Constructions
Lecture 24, April 12, 2001
Lecturer: Prof. David Zuckerman, Scribe: K. J. Balakrishnan

1 Improvements to Trevisan's Extractor

Today we look at the improvements to Trevisan's Extractor by Ta-Shma, Umans and Zuckerman. We start by defining loss-less condensers and their properties.

Definition 1 *Loss-less Condensers*

$$C : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^{n'}$$

is an $(n,k) \to n'$ condenser if $\forall$ sources X with $H_\infty(X) \le k$, $U_t \circ C(X, U_t)$ is ϵ-close to $D_1 \circ D_2$, where $D_1 = U_t$ and $\forall y \in \{0,1\}^t$, $H_\infty(D_2|D_1 = y) \ge H_\infty(X)$.

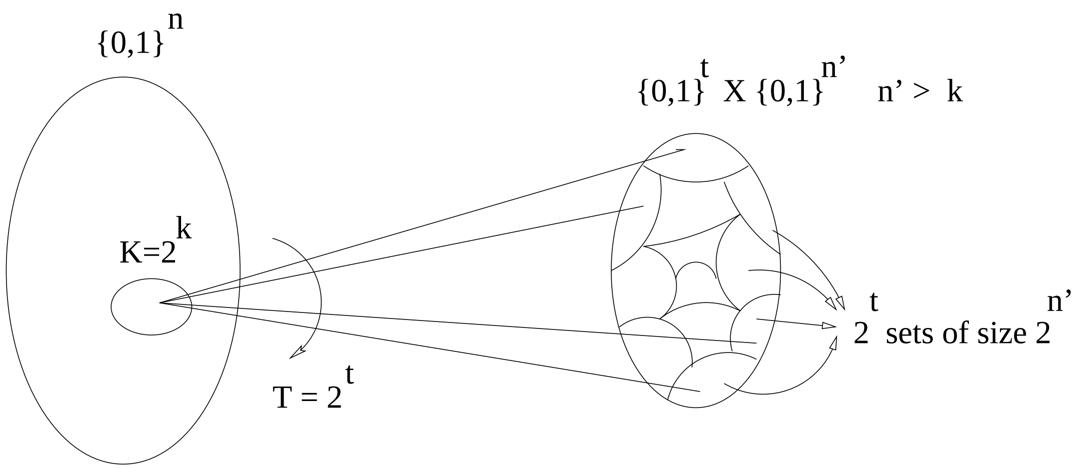

Figure 1: Lossless condenser

Recall that a flat source is a source which has equal probabilties on its support.

Proposition 1 *If X is a flat source, f a deterministic function such that $H_\infty(f(X)) \ge H_\infty(X)$, then f is one-one and $H_\infty(f(X)) = H_\infty(X)$.*

Using the above proposition, the K vertices representing X has $T = 2^t$ disjoint images of size K, and hence this set has $T \cdot K$ neighbors. We really only achieve an ϵ approximation, so K vertices will expand to $(1-\epsilon)T \cdot K$.

Lemma 1 *The following are equivalent:*

1. *C is an $(n,k) \to n'$ condenser*
2. *G_C is $(K = 2^k, (1-\epsilon)2^t)$ expanding, i.e. any set S, $|S| \le K$, has $|\Gamma(S)| \ge (1-\epsilon)2^t|S|$.*
3. *C has a $(k, 1-\epsilon)$ inverter.*

Definition 2 *C has a (k,δ) inverter if $\forall$ flat sources X with $H_\infty(X) \le k$,*

$$\exists I_X : \{0,1\}^t \times \{0,1\}^{n'} \to \{0,1\}^n, Pr[I_X(y; C(x,y)) = x] \ge \delta.$$

1.1 Condenser Construction

Trevisan's extractor T can be viewed as outputting an advice string in addition to the extractor output. Trevisan showed that if $T(x,y)$ is not close to uniform, then the advice string is a succinct description of the weak source input x, which is a contradiction if there are many possible weak source inputs.

To get a condenser, we can set the parameters so that $T(x,y)$ is not close to uniform. This implies that the advice string contains all the information about x, so the condenser can simply output the advice string. We ensure that $T(x,y)$ is far from uniform by outputting many more bits than the randomness in x.

To describe the condenser, recall that if $T(x,y)$ is not close to uniform, then $\exists$ next bit predictor, i.e., if Z is the output, it can predict Z_i from $Z_1, Z_2 \ldots Z_{i-1}$. Since $Z_j = \bar{X}(S_j)$, if we know $\bar{X}(S_j)$ for $j < i$, we can predict the entire function $\bar{X}$.

Condenser definition:

Input: $X \in \{0,1\}^n$ (from some weak random source), $\bar{X} = C(X)$ from $[\bar{n}, n, .3]_2$ code C

Auxillary Randomness: $\beta \in \{0,1\}^{d-l}$, $l = |S_i|$, $i \in [m]$, $m = k/\epsilon$.

(β will set all the bits outside of S_i), $j \in [1, i)$

Output: $\bar{X}(\beta^{\neq S_i} \circ \gamma^{S_i \cap S_j}|_{S_j})$, $\gamma \in \{0,1\}^{|S_i \cap S_j|}$

Length of output $= \sum_{j=1}^{i-1} 2^{|S_i \cap S_j|} = n'$

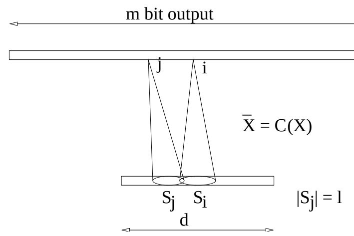

Figure 2: Condenser Based on Trevisan's Extractor

The proof of correctness for the condenser resembles the proof of correctness of Trevisan's extractor. Note that we do not use list decoding; we use unique decoding because we want to describe x uniquely (with high probability).

We also want to a next-bit predictor that works with high probability, not just significantly more than half. For this, we introduce entropy.

2 Entropy $H(X)$

Some basic definitions and properties of entropy.

Definition 3 *If X has probabilities p_i, then* $H(X) = E_X[\log_2 1/p] \; = -\sum_i p_i \log(1/p_i)$

Example 1: $H(U_l) = l$.

Example 2: If X has probabilities such that $p_1 = p_2 = \ldots p_{2^{l-1}} = 2^{-(l-1)}$ and $p_0 = 1/2$, $H(X) = \frac{1}{2}.1 + \frac{1}{2}(l-1) = l/2$.

Min Entropy $H_\infty(X) = 1$.

Property : $H(Y/X) = E_X[H(Y/X = x)]$

Proposition 2 $H(XY) = H(X) + H(Y/X)$.

Proof: (sketch) use $p(XY) = p(X) \cdot p(Y/X)$. $\square$

Lemma 2 *If $Y = Y_1, Y_2 \dots Y_m \in \{0,1\}^m$ has $H(Y) \leq \epsilon m$, then Y has a $(1-\epsilon)$ next bit predictor i.e. $Pr_{I \in [m] Y}[T(Y_1, Y_2 \dots Y_{I-1}) = Y_I] \geq 1 - \epsilon$.*

Proof: $I \ \epsilon_u \ [m]$. Let $Z = I \circ Y_1 \circ \dots \circ Y_{I-1}$. Therefore

$$H(Y_I/Z) = \frac{1}{m} \sum_{i=1}^{m} H(Y_I/Y_1 \dots Y_{I_1}) = \frac{1}{m} H(Y) \leq \epsilon$$

Let $p_z = Pr[Y_I = 1/Z = z]$. Predict $Y_I = 1$ if $p_z \geq 1/2$ and $Y_I = 0$ otherwise.
Error $= E[\min(p_Z, 1 - p_Z)]$. Now

$$\min(p, 1-p) \leq \min(p, 1-p) \log_2 (\frac{1}{\min(p, 1-p)} \leq p \log(1/p) + (1-p) \log(1/1-p) = H(p).$$

Therefore Error $\leq E[H(p_z)] = H(Y_I/Z = z) \leq \epsilon$. $\square$