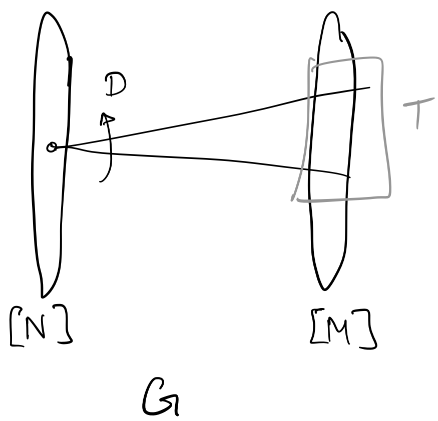
Pseudorandomness : Lecture 22

Instructor: Ramprasad.
Date : 2021-11-16
Lecture # : 22

Agenda: - Unified view of pseudorandom objs.
- Trevisan's Extractor.



$\Gamma : [N] \times [D] \to [M]$

Each of the PROs we studied ask for some property of these graphs.

Defn: (List decoding view). For any $T \subseteq [M]$, define

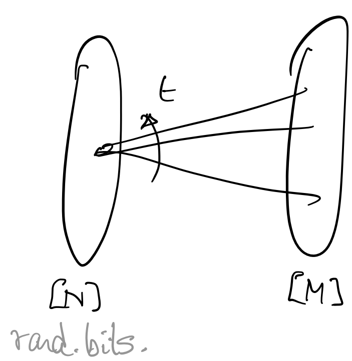
$$\mathrm{List}_\Gamma(T, \varepsilon) = \left\{ x \in [N] : \Pr_{y \in [D]} [\Gamma(x,y) \in T] > \varepsilon \right\}.$$

Generalising to $f : [M] \to [0,1]$, define

$$\mathrm{List}_\Gamma(f, \varepsilon) = \left\{ x \in [N] : \mathop{\mathbb{E}}_y [f(x,y)] > \varepsilon \right\}.$$

What is a Sampler in this language?

$\mathrm{Samp} : \underset{\text{seed}}{[N]} \to \underset{\text{sample.}}{[M]^t}$



$\Gamma(x,y) = \left(\mathrm{Samp}(x)\right)_y$

$\Gamma : [N] \times [t] \to [M].$

If Sampler is a (δ, ε)-averaging sampler, then.

$$\Pr_x\left[\frac{1}{t}\sum_y f(\Gamma(x,y)) > \mu(f)+\varepsilon\right] < \delta.$$

for every $f: [M] \to [0,1]$.

Or equivalently:

$$\left|\mathrm{List}_\Gamma(f, \mu(f)+\varepsilon)\right| \leq \delta \cdot N \quad \text{for any } f: [M] \to [0,1].$$

(if sampler only for sets (aka. boolean averaging sampler), replace above with $T \subseteq [M]$).

Typical setting: $N = \mathrm{poly}(M/\delta)$, $D = O\left(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}\right)$

Coding theory: $\mathrm{Enc}: \mathbb{F}_q^a \to \mathbb{F}_q^b$

Think of this as $\Gamma: \mathbb{F}_q^a \times [b] \to [b] \times \mathbb{F}_q$.

$$\Gamma(x,i) = (i, \mathrm{Enc}(x)_i)$$

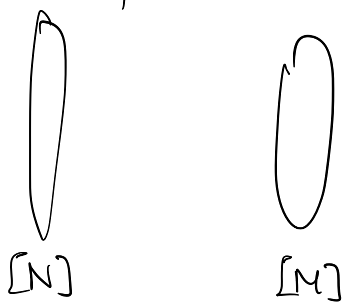
Given a received word, we don't want too many close codewords ...

If $T = \{(i, y_i) : i \in [b]\}$, want

$$\left|\mathrm{List}_\Gamma\left(T, \text{"agreement threshold"}\right)\right| \leq \text{"small"}.$$

RHS: $[N] = \mathbb{F}_q^a$. $D = b$ $M = \mathbb{F}_q \cdot b$.

Vertex expanders:



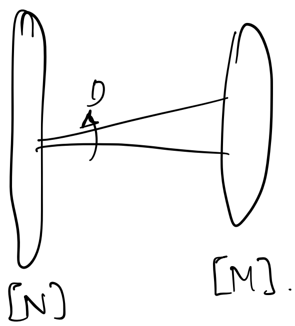
$(=k, A)$ - vertex expander

$\Rightarrow$ all sets of size k on left have $\geq Ak$ neighbours

$\Updownarrow$

$|T| < Ak \Rightarrow |List_\Gamma(T, 1)| < k$.

Extractors:



(k, ε)- extractor

$\Rightarrow$ any $S \subseteq [N]$ of size $\geq k$ must be close to uniform on the right.

Claim: (a) If $Ext: [N] \times [D] \to [M]$ is a (k, ε)-extractor, then for every $f: [M] \to [0,1]$, we have

$$\left| List_\Gamma(f, \mu(f) + \varepsilon) \right| < k.$$

(b) Suppose $|List_\Gamma(T, \mu(T) + \varepsilon)| < K$ for all $T \subseteq [M]$, then Ext is a $(k + \log\frac{1}{\varepsilon}, 2\varepsilon)$- extractor.

Pf: (a). Say $f: [M] \to [0,1]$ with $|List_\Gamma(f, \mu(f) + \varepsilon)| \geq k$.

Take X to be the uniform dist on $\uparrow$.

Minentropy $(X) \geq k$.

$$\mathop{\mathbb{E}}_{\substack{x \sim X \\ y \sim [D]}} \left[f(E(x, y)) \right] > \mu(f) + \varepsilon = \mathbb{E}[f] + \varepsilon \Rightarrow \Leftarrow$$

(b) Fix a $k+\log\frac{1}{\varepsilon}$ - source X. Fix $T \subseteq [M]$.
and let $L = \mathrm{List}_\Gamma(T, \mu(T)+\varepsilon)$

$$\Pr\left[\mathrm{Ext}(X, U_d) \in T\right] \le \Pr_{x\sim X}\left[x \in L\right] + \Pr\left[\mathrm{Ext}(\ ) \in T \mid x \notin L\right]$$

$$\le K \cdot \frac{1}{2^{k+\log 1/\varepsilon}} + \mu(T) + \varepsilon = \mu(T) + 2\varepsilon \qquad \square.$$

$$\mathrm{Ext} \overset{\approx}{\iff} \left|\mathrm{List}_\Gamma(T, \mu(T)+\varepsilon)\right| < K \qquad \forall\ T \subseteq [M].$$

Cor: Samplers $\iff$ extractors for suitable parameters.

(No wonder expanders & hash fns had extractor like properties).

PRGs, however, appear to be different... there is a "computational requirement".
(Or is it?)

Let's revisit the NW PRG. We start with $f: \{0,1\}^k \to \{0,1\}$.

$$G^f_{NW}(y_1, .., y_l) = \left(f(y|_{S_1}), .., f(y|_{S_m})\right)$$

where $S_1, .., S_m \subseteq [l]$ is a (k,a)-comb. design.

$|S_i| = k$ $\qquad$ $|S_i \cap S_j| < a$.

Rough proof: If we have a distinguisher C, then we can use C to build a circuit C' that computes f
If C is small, then so is C' and that contradicts hardness of f.

$\{0,1\}^k \to \{0,1\}$

Defn: (Blackbox PRG constructions) A generator $G_f : \{0,1\}^{\ell} \to \{0,1\}^m$, defined for every $f \in [H]$ is said to be a (t, r, ε)-BB-PRG if, there is an oracle procedure Recon running in time t s.t for every $f \in [H]$ and an ε-distinguisher C, there is an advice string $z \in [R]$ such that

$$C' = \mathrm{Recon}^C(y, z) = f_y \quad \text{for all } y.$$

(NW appears to fit into this right? More on this later.)

Trevisan: BB-PRGs immediately yield extractors!

Thm (Trevisan): Let $G^f : \{0,1\}^{\ell} \to \{0,1\}^m$, def for all $f \in [H]$, be an (∞, r, ε)-BB PRG. Then, the map

$$\Gamma : [H] \times [L] \to [M]$$

$$\Gamma : (f, y) = G^f(y)$$

is an $(r + \log 1/\varepsilon, 2\varepsilon)$-extractor. $\Leftarrow\Rightarrow \left|\mathrm{List}_\Gamma(T, \mu(t)+\varepsilon)\right| \leq R$

Pf: $T \subseteq [M]$ $\qquad \mathrm{List}_\Gamma(T, \mu(T)+\varepsilon) = \left\{ f : \Pr[\Gamma(f,y) \in T] > \mu + \varepsilon \right\}$

$\therefore$ If $f \in \mathrm{List}$, then T is an ε-distinguisher.

$\Rightarrow \exists\, z \in [R]$ s.t $\mathrm{Recon}^T(\cdot, z) = f$

$\Rightarrow \left|\mathrm{List}_\Gamma(T, \mu(T)+\varepsilon)\right| \leq R$

$\square$.

Is the NW PRG a blackbox construction?

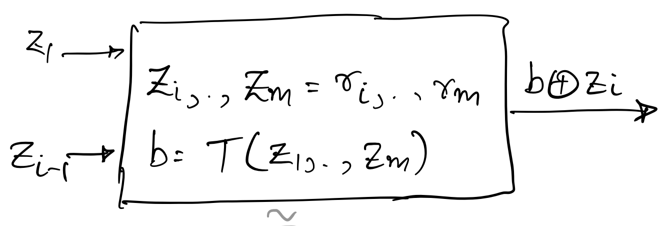
$f: \{0,1\}^k \to \{0,1\}$. $S_1, .., S_m$ is a $(k, \gamma \log m)$-NW design. $S_i \subseteq [d]$ with $d = \frac{2k^2}{\gamma \cdot \log m}$.

$G^f: [D] \to [M]$

$G^f(y) = (f(y|_{S_1}), .., f(y|_{S_m}))$

Suppose $T: \{0,1\}^m \to \{0,1\}$ is an ε-distinguisher.

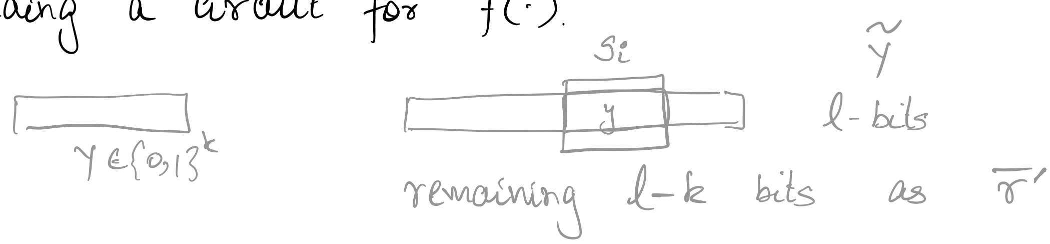
NBP:



$\tilde{T}$

$$\underbrace{\exists\, i \in [m]\ r_i, \ldots r_m}_{\text{advice}} : \Pr_{z_1 .. z_{i-1}} [\tilde{T}(z_1, .., z_{i-1}) = z_i] \geq \frac{1}{2} + \frac{\varepsilon}{m}$$

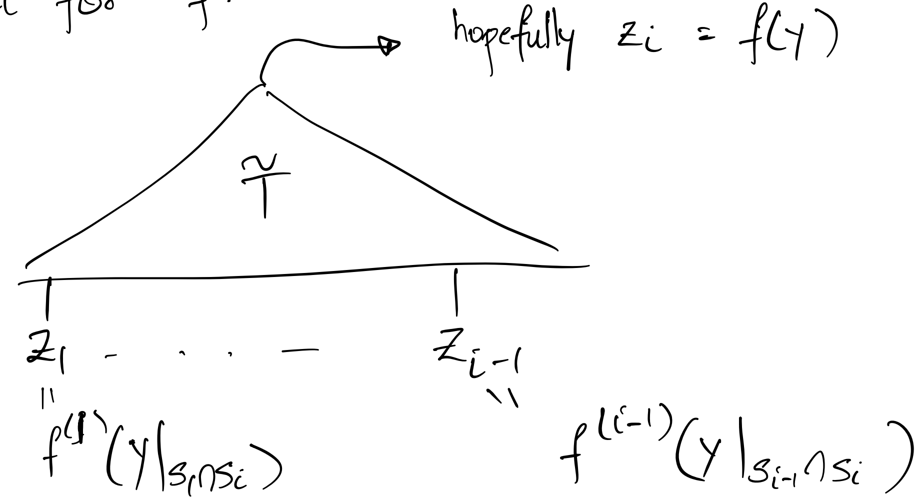
Building a circuit for $f(\cdot)$.



$$z_j = f(\tilde{y}|_{S_j}) \quad \text{if } j < i$$

$$= f^{(j)}(y|_{S_i \cap S_j}) \quad \text{for some boolean fn } f^{(j)}: \{0,1\}^a \to \{0,1\}, \ a = \gamma \log m.$$

Circuit for f:



$\Rightarrow \exists\, r'$ and fns $f^{(1)} \dots f^{(i-1)} : \{0,1\}^a \to \{0,1\}$ s.t

$$\underbrace{\tilde{T}\left(f^{(1)}(y|_{S_1 \cap S_i}), \dots, f^{(i-1)}(y|_{S_{i-1} \cap S_i}) \right)}_{C'} = f(y) \quad \underset{\text{over } y.}{\text{w.p}} \geq \frac{1}{2} + \frac{\varepsilon}{m}$$

But C' only computes f on $1/2 + \delta$ locations . . .

Idea: Use an ECC on f.

$$f : \underbrace{\{0,1\}^k}_{2^k} \to \{0,1\} \xrightarrow{\text{ECC}} g : \underbrace{\{0,1\}^{\tilde{k}}}_{2^{\tilde{k}}} \to \{0,1\}$$

$$\left| \left\{ f : \text{ECC}(f) \text{ agrees with some } g : \{0,1\}^{\tilde{k}} \to \{0,1\} \text{ on } \tfrac{1}{2} + \delta \text{ fraction.} \right\} \right| \leq \text{"small"}.$$

Fact: There exists such explicit codes with $\tilde{k} = O(k + \log 1/\delta)$ and list size "small" $\leq O(1/\delta^2)$

Modified construction:

$$\tilde{G}^f(y) = G^{\tilde{f}}(y) \qquad \text{where } \tilde{f} = \text{ECC}(f).$$

$y \in [D]$ with $d = O((k+\log(1/\delta))^2 / \gamma \log m)$

Advice for the Reconstruction:

NBP
- ▷ $i \in [m]$, $r_i, \ldots, r_m \in \{0,1\}$ | $\log m + m$
- ▷ $r' \in \{0,1\}^{l-k}$, $f^{(1)} \ldots f^{(i-1)} : \{0,1\}^a \to \{0,1\}$ | $d + m \cdot 2^a = d + m^{1+\gamma}$
- ▷ An index within the decoded list of size $1/\delta^2 = (m/\varepsilon)^2$ | $2 \log \frac{m}{\varepsilon}$

Total advice length : $m^{1+\gamma} + O(\log \frac{m}{\varepsilon} + d)$

Thm: For any $\gamma > 0$, $\varepsilon > 0$ and $l, m \in \mathbb{N}$, we have an (t, r, ε)-BB PRG construction $\tilde{G}^f : [D] \to [M]$ for every $f : \{0,1\}^k \to \{0,1\}$ with

[BB-PRG analogue of NW]

- ▷ $d = O((k + \log \frac{m}{\varepsilon})^2 / \log m)$
- ▷ $t = \text{poly}(m, 1/\varepsilon)$
- ▷ Advice length $r = m^{1+\gamma} + O(l + \log \frac{m}{\varepsilon})$

Cor: For any $\gamma, \varepsilon > 0$ and $k \leq n \in \mathbb{N}$, the map

[Trevisan]

$$\Gamma : [N] \times [D] \to [M]$$
$$\Gamma(f, y) = G^f(y)$$

is a (k, ε)-extractor with $d = O((\log \frac{n}{\varepsilon})^2 / \log k)$ and $m = k^{1-\gamma}$.

(Not the best choice of parameters, but a cool connection!)

Next: An exposition on 2-source extractors & connections to Ramsey graphs.