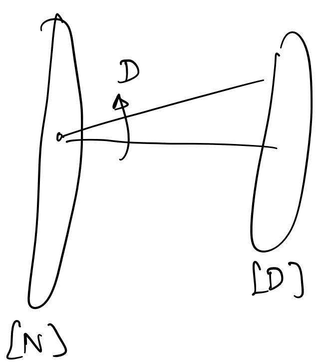
Pseudorandomness: Lecture 23.

Instructor: Ramprasad.
Date: 2021-11-18
Lecture: #23

Agenda:
- Ramsey graphs
- An exposition of constr. of 2-source extractors.

Recap:
- Extractors: $\mathrm{Ext}: [N] \times [D] \to [M]$ is a (k,ε)-extractor if for every k-source X, $\mathrm{Ext}(X, \mathcal{U}_d) \approx_\varepsilon \mathcal{U}_m$.

Thm: [GUV] For all $k \leq n \in \mathbb{N}$, $\alpha, \varepsilon > 0$, there is an explicit (k,ε)-strong-extractor $\mathrm{Ext}: [N] \times [D] \to [M]$ with $m = (1-\alpha)k + O(\log n/\varepsilon)$ and $d = O(\log n/\varepsilon)$.

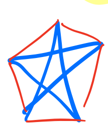


$$|\mathrm{List}_\Gamma(T, \mu+\varepsilon)| \leq K.$$
for all $T \subseteq [M]$

- Strong-extractors: $(\mathrm{Ext}(X,Y), Y)_{Y \sim \mathcal{U}_d} \approx_\varepsilon \mathcal{U}_{m+d}$.

Ramsey Graphs:

Obs: Any graph on 6 vertices either contains a clique or ind. set of size 3.



Qn: For some $k>0$, is there an n large enough such that any graph on n vertices must either have a k-clique or a k-ind set?

$R(k)$ = smallest n s.t all n-vertex graphs have either a k-clique or a k-ind set.

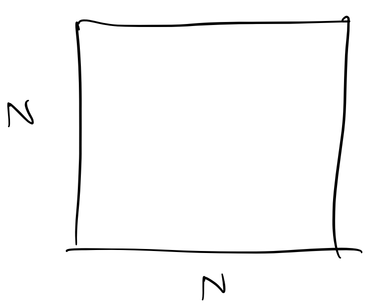
How large is $R(k)$?

$$2^{\frac{k}{2}} (\quad) \leq R(k) \leq 4^k / (\quad).$$

Defn: A k-Ramsey graph is one that has no k-clique or k-independent set.

A bipartite k-Ramsey graph is one with no $K_{k \times k}$ subgraph or no $k \times k$ indep. set.

As a matrix:



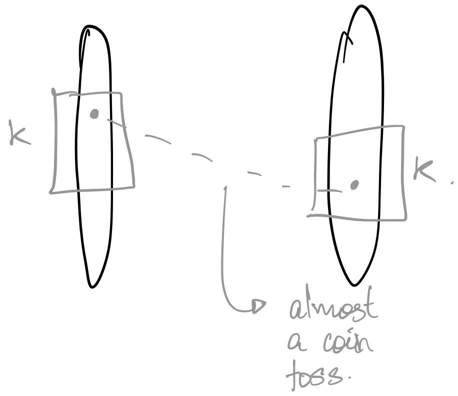
Made of 0s & 1s s.t there is no $k \times k$ submatrix that is constant.

How large can we make k?

$\exists$ such matrices for $k = 2 \log N$.

Can you find an explicit such matrix?

Stronger ask: Every $k \times k$ submatrix is nearly balanced.



(k,k,ε)-extractor with a 1-bit output.

Two-source Extractors:

$\mathrm{Ext}: [N_1] \times [N_2] \to [M]$

is a (k_1, k_2, ε) - 2-source ext

if for every k_1-src X_1 & k_2-src X_2

s.t. X_1, X_2 indep, $\mathrm{Ext}(X_1, X_2) \approx_\varepsilon \mathcal{U}$.

$\therefore$ If we can build 2-source extractors, then we can find Ramsey graphs.

History:

Chor-Goldreich 88: $\left(> \frac{n}{2}, > \frac{n}{2}\right)$-extractor. $\quad k = \sqrt{N}$.

Bourgain 05 : $(> 0.49n, > 0.49n)$-extractor. $\quad k = N^{0.49\ldots}$

Raz 05 : $(> 0.5n,\ c\log n)$-ext.

Chattopadhyay-Zuckerman 15 : $(\mathrm{poly}\log(n), \mathrm{poly}\log(n))$-ext.

⋮

Li 16 : $(\log n \log\log n,\ \log n \log\log n)$ $\quad k: (\log N)^{\log\log\log N}$.

$\downarrow$ extracts $0.9k$ bits.

A puzzle: You have D players but some b of them are malicious. The honest players choose a uniform bit. and the malicious players choose their bits after the honest players.

You compute $f(z_1, .., z_D)$ and would like this to be as unbiased as possible.

f = PARITY. How many bad players can you tolerate? Not even 1!

f - MAJ $O(\sqrt{D})$.

Defn: f is (b, ε)-resilient if $f(\bar{z})$ is ε-unbiased even in the presence of b bad players

[Ajtai-Linial] There are $f: \{0,1\}^D \to \{0,1\}$ that are b-resilient for $b \approx O\left(\frac{n}{\log n}\right)$.

(Tribes fn (sort of)).

[Meka] An explicit monotone fn f that is $O\left(\frac{n}{\log n}\right)$ resilient.

Key ideas:

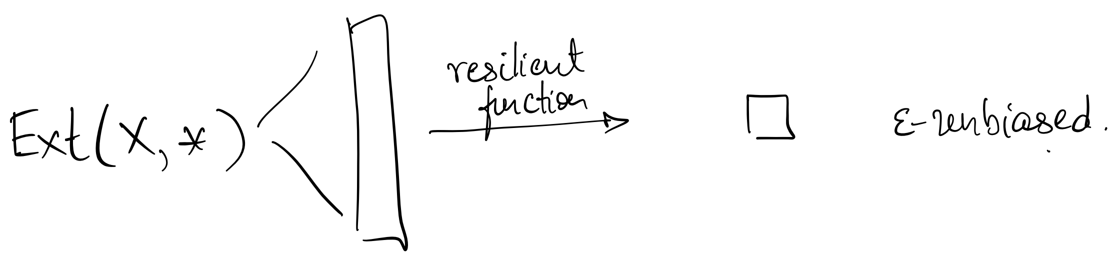
Revisiting strong extractors:

$$(Y, \mathrm{Ext}(X, Y))_{Y \sim \mathcal{U}_d} \approx_\varepsilon (\mathcal{U}_d, \mathcal{U}_m)$$

$\boxed{\quad X \quad} \longrightarrow$ $\mathrm{Ext}(X, y_1)$
$\vdots$
$\mathrm{Ext}(X, y_D)$

Claim: For any k-src X, there are $(1-\varepsilon)$ frac of the y's s.t $\mathrm{Ext}(X, y) \approx_\varepsilon \mathcal{U}_d$

Idea 1: Run over all seeds, and apply a resilient fn on them.



Will this work? Each of the good rows are close to uniform but they are all correlated...

Eg: $f = \mathrm{Maj}$:
$z_1, \bar{z}_1, z_2, \bar{z}_2, \ldots, z_r, \bar{z}_r$, (bad z)
Totally dead.

Idea 2: What if we somehow knew that the z's were t-wise independent?

If # bad players $\le \sqrt{t}$, then Majority still works! [Viola].

Can we arrange for this situation to happen?

Defn (Non-malleable extractors). $Ext : [N] \times [D] \to [M]$ is a (t, k, ε)-n.m. ext if for every k-source X, there is a set $G \subseteq [D]$ of density $\geq (1-\varepsilon)$ s.t for every $y \in G$ and $y_1, .., y_t \in [D] \setminus \{y\}$, we have that $Ext(X, y)$, even conditioned on $Ext(X, y_1), .., Ext(X, y_t)$, looks ε-close to uniform.

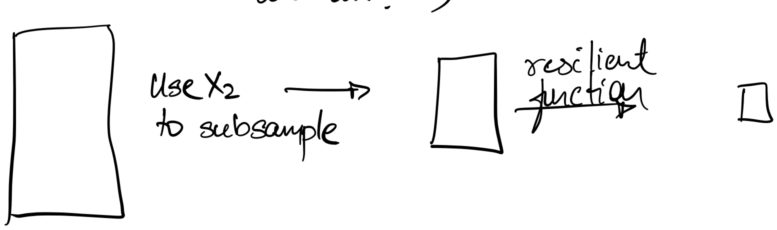
[Dodis-Wichs]

Turns out, explicit such extractors exist.

Modified approach: Use a n.m extractor on X, run over all seeds y, use a suitable resilient function on top of that.

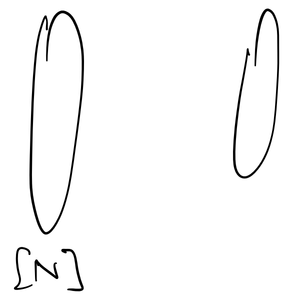
... doesn't work out. Cannot work! (We haven't used source 2 at all!)

[CZ]: $nmExt(X_1, *) =$



An alternate approach: [BCDLT]

Dispersers:



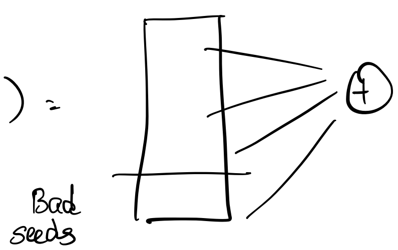
is a (K_1, K_2)-disperser if any set $S \subseteq [N]$ of size K_1 has $> K_2$ vertices on the right.

$nmE : [N_1] \times [D] \to \{0,1\}^m$, t-n.m extractor. with error ε.

$\Gamma : [N_2] \times [t+1] \to [D]$, a $(\varepsilon K_2, \varepsilon D)$-disperser

$$\mathrm{Ext}(x_1, x_2) = \bigoplus_{i=1}^{t+1} nm\mathrm{Ext}(x_1, \Gamma(x_2, i))$$

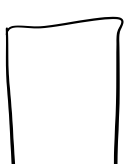
$\mathrm{Ext}(x_1, *) =$



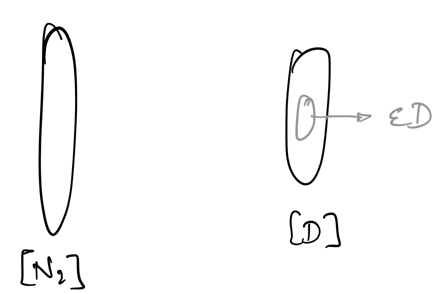
$\Gamma(x_2, *) =$ some $t+1$ rows

No resilient fns!

Why should this work?



D rows, and at most εD bad rows. (B)



$\therefore \left|\{x_2 : \Gamma(x_2, *) \subseteq B\}\right| < \varepsilon K_2$.

Since $x_2 \sim X_2$, a k_2-source, w.p $1-\varepsilon$ we will "hit" a good row. : say row y.

$\therefore$ $nmExt(X, y)$ is ε-close to uniform even cond. on t other rows!

$\therefore$ The parity fn ensures that final output is close to uniform.