CS 229r: IT & CS

2/21/19

~~LECTURE 8~~

TODAY

- POLAR CODES
  - OVERVIEW
  - KEY NOTIONS
  - ENCODING + DECODING
  - MAIN LEMMAS $\Rightarrow$ THEOREM

Theorem: [Polar Coding Main, GX13, HAU13]

$\forall p \in [0,1] \; \forall \epsilon > 0 \quad \exists n_0 = \text{poly}(1/\epsilon)$ s.t. $\forall n \geq n_0$

$\exists$ code of rate $(1 - h(p) - \epsilon)$ of length n that is decodable in time $\text{poly}(n)$ and corrects the $BSC(p)$ w.p. $1 - \frac{1}{n^{10}}$

Theorem': [Polar Compression]

$\forall p \in [0,1], \forall \epsilon > 0 \quad \exists n_0 = \text{poly}(1/\epsilon)$ s.t. $\forall n \geq n_0$

$\exists$ linear compressor $H \in \mathbb{F}_2^{m \times n}$, $m \leq (h(p) + \epsilon) n$,

& efficient decompressor $D : \mathbb{F}_2^m \to \mathbb{F}_2^n$

s.t. $\Pr_{Z \sim \text{Bern}(p)^n}[D(HZ) \neq Z] \leq 1 - \frac{1}{n^{10}}$

Other Equivalent (*) versions

Theorem[2]: $\forall p \; \forall \epsilon > 0 \; \exists n_0 \leq \text{poly}(1/\epsilon)$ & ~~codes of~~ linear compressor + eff. decoder... * for length n_0 *

[No promises for larger lengths!]

Theorem[3]: $\forall p \; \exists \delta > 0$ s.t. $\forall n$ can compress n bits to $H(p) \cdot n + O(n^{1-\delta})$ bits

---

[for reference "optimal" non-linear compressor compresses to

$H(p) \cdot n + O(n^{1/2})$ bits

Lempel-Ziv ... maybe $H(p) \cdot n + O\left(\frac{n}{\log\log\log n}\right)$ ...

Will try to go for Theorem[3] version.

Main Construction ... Polar Codes ... due to Arikan 2008

Analysis Guruswami-Xia 2013

Hassani-Alishahi-Urbanke -2013.

This analysis: Blasiok - Guruswami - Nakkiran - Rudra - S.

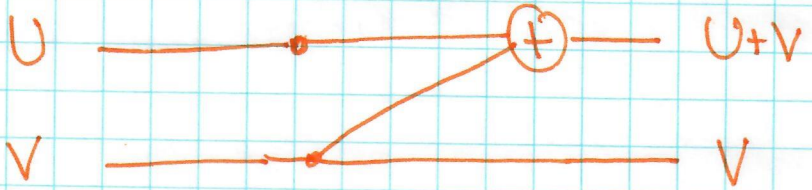
Key Idea: Can't wait to collect $1/\epsilon^2$ bits before "squeezing" them.

Will start the squeezing much earlier?

How?

Idea with <u>two</u> independent identical bits

!!



$$(U,V) \longrightarrow (U+V, V)$$

① Invertible transform; no compression $[H(U,V) = H(U+V,V)]$

② But ~~"segregates"~~ "differentiates"

① $U+V$ more "entropic" than U (or V).

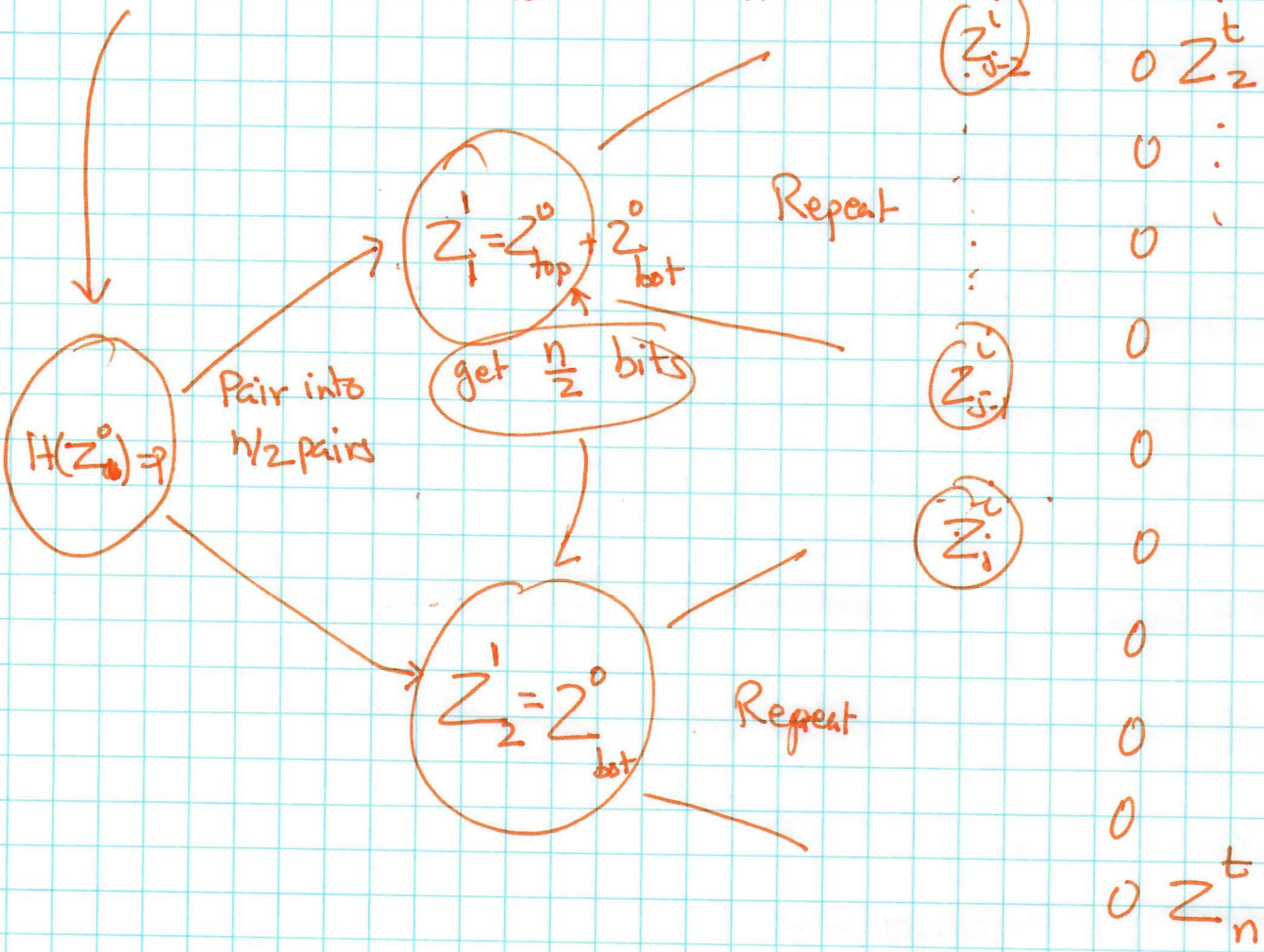
② $V \mid U+V$ less entropic !!

$$H(U) + H(V) = H(U+V) + H(V \mid U+V)$$

Idea: - start with many iid bits
- Keep applying such transforms till we get a bunch of bits that have cond. entry $\nearrow 1$ $\searrow 0$

"Repeating"

Start with $n \to \infty$ ~~bit~~ i.i.d bits



$t = \log n$

At Stage i $\quad H(Z_j^i \mid Z_{<j}^i) = ?$

Very Informal Claim: ~~As $t \to \infty$ for random j $H(Z_j^i$~~

$$\lim_{t\to\infty} H(Z_j^t \mid Z_{<j}^t) \to \{0,1\}.$$

for random $j \in \{1 \cdots 2^t\}$.

Various "approximations"

Formally: $\Pr_j\left[H(X_j^t \mid X_{<j}^t) \notin (\delta, 1-\gamma)\right] \le \epsilon$

$\tau(t), \gamma(t), \epsilon(t) \to 0$ as $t \to \infty$.

POLARIZATION

Suppose $\tau, \gamma, \epsilon = 0$. Why is this good?

$H(Z_1^0 \dots Z_n^0) = n \cdot H(p)$

$H(Z_1^t \dots Z_n^t) = n \cdot H(p)$

But if $\forall j \quad H(Z_j^t \mid Z_{<j}^t) \in \{0,1\}$

then $\#\{j \mid H(Z_j^t \mid Z_{<j}^t) = 1\} = n \cdot H(p)$.

Let $S \triangleq \{j \mid H(Z_j^t \mid Z_{<j}^t) = 1\}$

Then to compress $\bar{Z} = (Z_1^0 \dots Z_n^0)$

compute $\bar{W} = (Z_1^t \dots Z_n^t)$

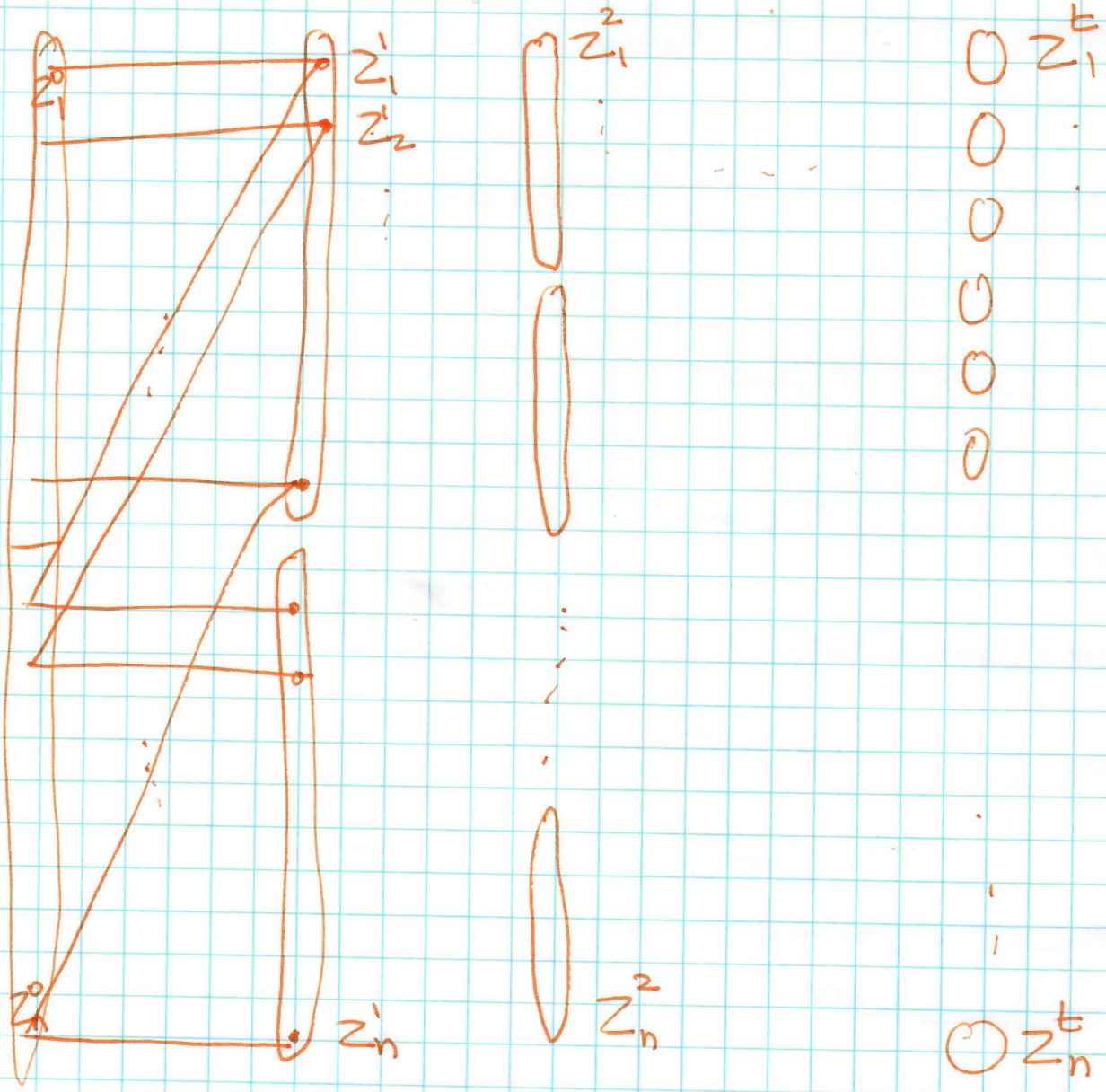
output $\bar{W}_S = \{Z_i^t \mid i \in S\}$.

$H(\bar{W}_{\bar{S}} \mid \bar{W}_S) = 0 \Rightarrow \bar{W}_S \Rightarrow \bar{W}_{\bar{S}} \Rightarrow \bar{W}$

$\Downarrow$

$\bar{Z}$.

Expanded Picture



Issues to deal with

(1) $\epsilon(t), \tau(t), \gamma(t) \neq 0$ [only going to zero]

need to define S correctly; What is ~~$|S|$~~ $|S|$?

$= n \cdot H(p) + n^{1-\delta}$?

(2) Need to compute S efficiently.

(3) Encode? Easy ✓

(4) Decode?

Fixing (1)

Suppose $\Pr\left[H(Z_j^t \mid Z_{<j}^t) \in (\tau, 1-\gamma) \right] \leq \epsilon$

Then $S \triangleq \{ j \mid H(Z_j^t \mid Z_{<j}^t) > \tau \}$

Claim: $|S| \leq \frac{H(p) \cdot n}{1-\gamma} + \epsilon \cdot n$

$$\leq H(p)n + \frac{\gamma \cdot n}{1-\gamma} + \epsilon \cdot n$$

Need $(\gamma+\epsilon)n \leq n^{1-\delta}$

$\Rightarrow \gamma, \epsilon \leq n^{-\delta} = (2^t)^{-\delta} = 2^{-\delta t} \approx (1-\delta')^t$

$\boxed{\gamma(t), \epsilon(t) \text{ need to shrink exponentially with } t.}$

How about $\tau(t)$?

"Decoding Plan"

- Given W_S will reconstruct W
- Write $W \in \{0,1,?\}^n$ $(W_i = ? \text{ if } i \notin S)$
- Will reconstruct W_j ~~one~~ for $j = 1 \ldots n$

  Use $W_{<j}$ to guess $\hat{W}_j$. Can do if $H(W_j \mid W_{<j})$ $\leq$ small.

Process wrong if $\exists j$ s.t. W_j not guessed correctly

$$Pr\left[\hat{W}_j \neq W_j\right] \leq H(W_j \mid W_{<j}) \leq \tau$$

$$Pr\left[\exists j \text{ s.t. } \hat{W}_j \neq W_j\right] \leq n \cdot \tau$$

Need $n \cdot \tau \leq \frac{1}{n^{10}} \Rightarrow \tau \leq \frac{1}{n^{11}} = \left(\frac{1}{2^{11}}\right)^t$

exponentially with really tiny base!!

Strong Polarization Theorem

$\forall \alpha > 0 \quad \exists \beta < 1$ s.t.

$$Pr\left[H(Z_j^t \mid Z_{<j}^t) \notin (\alpha^t, 1-\alpha^t)\right] \leq \beta^t$$

Fixes ①

② : Will not fix $\Rightarrow$ Preprocessing time $= \exp(n)$

③ Encoding given $S \Rightarrow$ easy.

④ Decoding =? [So far: not efficient procedure].

In next lectures....