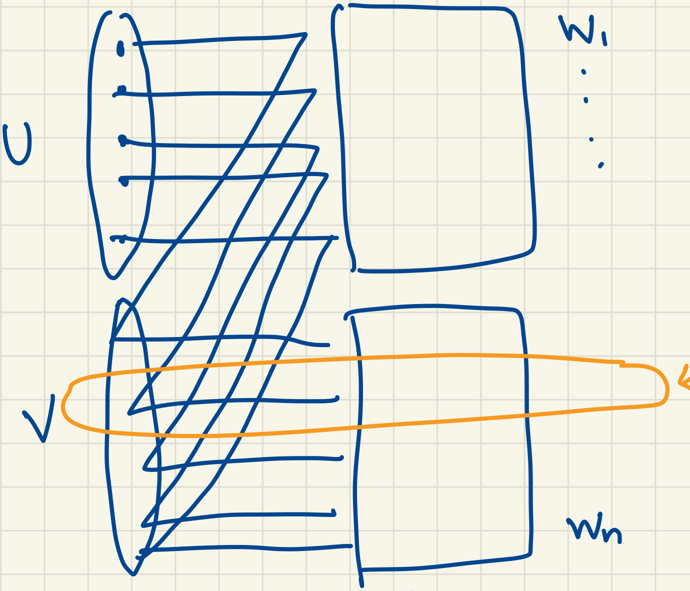
LECTURE 18

Today

- Analysis of Polarization
  - Martingales + Arikan Martingale
  - Strong vs. Local Polarization
  - Local Polarization of Arikan Martingale

Review: $P_n(U,V) \triangleq \left(P_{\frac{n}{2}}(U+V),\ P_{\frac{n}{2}}(V)\right)$



$H(W_i \mid W_{<i})$ polarizes?

Want to show: $\Pr_{i \in [n]}\left[H(W_i \mid W_{<i}) \in \left(\frac{1}{n^{100}},\ 1-\frac{1}{n^{100}}\right)\right] \leq \frac{1}{n^{.001}}$

7 lectures to go

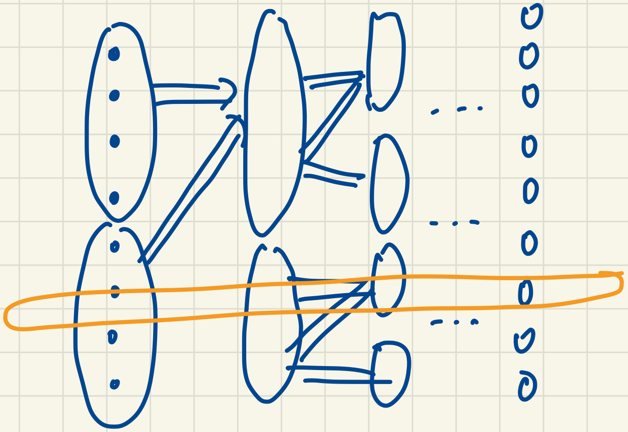
1. Coding for interactive comm
   (2 lectures)

2. Coding for edit distance
   (2 lectures)

3. Locality in coding theory
   (3 lectures)

4. Complexity $\Longleftrightarrow$ Coding
   (3 - 7 lectures)

Key Notion



$t = \log n$

$Z^{(0)}$ $Z^{(1)}$ $Z^{(2)}$ $Z^{(t)}$

x_0 x_1 ... x_t

Pick $i \sim \text{Unif}[n]$; $\boxed{X_j \triangleq H(Z^{(j)}_i \mid Z^{(j)}_{<i})}$

---x---

X_j = "Arikan Martingale"

---x---

Defn: $X_0, X_1, \ldots X_j, \ldots$ is a martingale if

$\forall\ j,\ \forall\ a_0 \ldots a_{j-1}$

$$\mathbb{E}\left[X_j \mid X_0 = a_0, \ldots X_{j-1} = a_{j-1}\right] = a_{j-1}$$

"conditioned on past future expectation is "no change" "

---x---

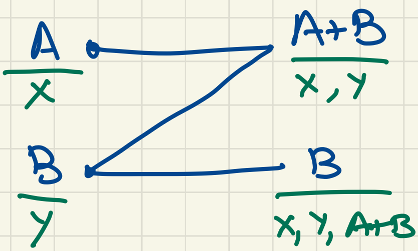
Q: - is our X_t = martingale ?

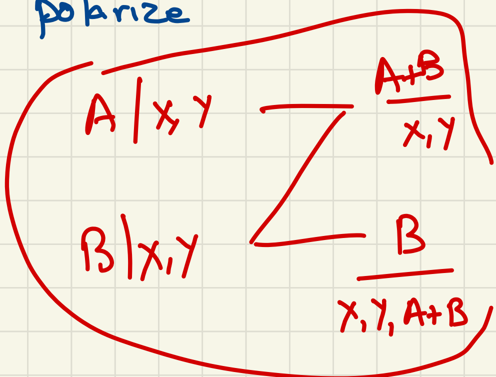
- why do martingales help.

$$X_j \triangleq H(Z_i^{(j)} \mid Z_{<i}^{(j)})$$

Why is $X_0, X_1, \dots X_j \dots$ a martingale

Suppose at time j we polarize





Claim: Given entropy history (A, X) indist. from (B, Y)

$$\Rightarrow \quad H(A|X) = H(B|Y)$$

Furthermore

$$H(A+B \mid X, Y) + H(B \mid X, Y, A+B)$$
$$= H(A|X) + H(B|Y)$$

$\boxtimes$

Martingales & Polarization

$t = \log n \qquad n = 2^t$

- What we want

STRONG POLARIZATION

$$\Pr\left[X_t \in \left(2^{-100t},\ 1-2^{-100t}\right)\right] \leq 2^{-.001t}$$

$$\cancel{\left[H(W_i \mid W_{<i}) \in \left(\tfrac{1}{n^{100}},\ 1-\tfrac{1}{n^{100}}\right)\right] \leq \tfrac{1}{n^{.001}}}$$

- What we know:

if $X_j = \alpha \Rightarrow \exists$ dist $(A,S) =_d (B,T)$

s.t. $H(A|S) = H(B|T) = \alpha$

A = bit

B = collection of bits

(A,S), (B,T) iid

$$X_{j+1} = H(A+B \mid S,T) \quad \text{w.p. } \tfrac{1}{2}$$
$$= H(B \mid A+B, S, T) \quad \text{w.p. } \tfrac{1}{2}$$

(very local behaviour; not too simple!)

- Q. What local behavior implies STRONG POLARIZATION?

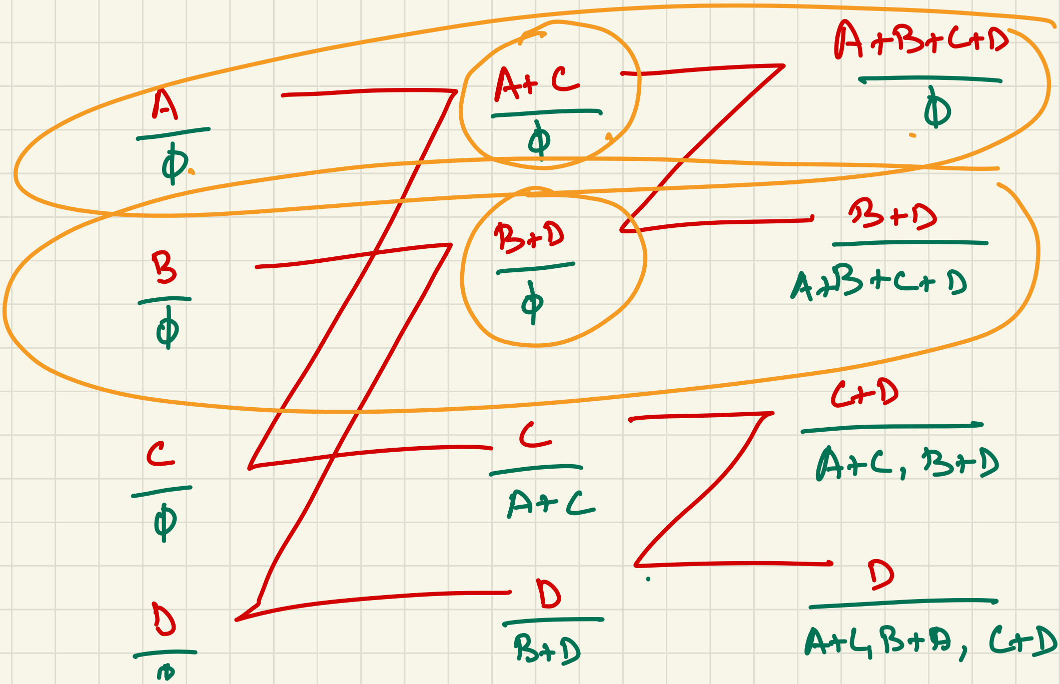
- Examples

$$X_{t+1} = X_t + 2^{-(t+2)} \quad \text{w.p. } \tfrac{1}{2}$$
$$= X_t - 2^{-(t+2)} \quad \text{w.p. } \tfrac{1}{2}$$

$$X_{t+1} = X_t^2 \quad \text{w.p. } \tfrac{1}{2}$$
$$= 2X_t - X_t^2 \quad \text{w.p. } \tfrac{1}{2}$$

$$X_{t+1} = X_t \pm \tfrac{1}{2}\cdot \min\{X_t, 1-X_t\}$$
each w.p. $\tfrac{1}{2}$

Which of these polarize? strongly?

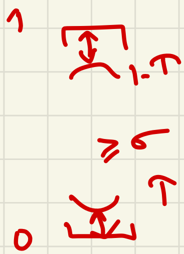


$\frac{A}{\phi}$
$\frac{B}{\phi}$
$\frac{C}{\phi}$
$\frac{D}{\phi}$
$\frac{A+C}{\phi}$
$\frac{B+D}{\phi}$
$\frac{C}{A+C}$
$\frac{D}{B+D}$
$\frac{A+B+C+D}{\phi}$
$\frac{B+D}{A+B+C+D}$
$\frac{C+D}{A+C,\ B+D}$
$\frac{D}{A+C,\ B+D,\ C+D}$

Barriers to (Strong) Polarization

1. Not enough variance as time $t \to \infty$
2. (Weak attraction to outer boundary $\{0,1\}$)

Defn: LOCAL POLARIZATION



1. Variance in middle:

$\forall \tau \; \exists \sigma > 0$ s.t. $\forall t$

$$\operatorname{Var}\left[X_t \;\middle|\; X_{t-1} = a_{t-1} \in (\tau, 1-\tau) \right] \geq \sigma^2$$

2. Suction at ends

low-end case; high-end similar

$\exists \theta > 0 \; \forall c \; \exists \tau > 0$ s.t.

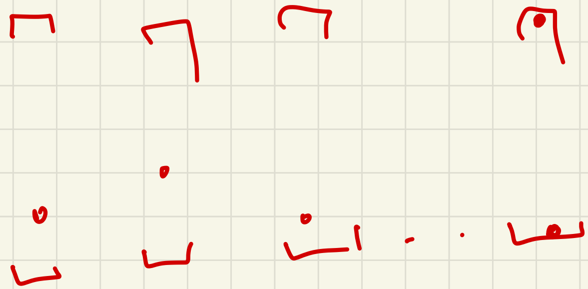
$$\Pr\left[X_t < \frac{X_{t-1}}{c} \;\middle|\; X_{t-1} < \tau \right] \geq \theta .$$

Theorem: LOCAL $\Rightarrow$ STRONG POLARIZATION

Theorem: ARIKAN MARTINGALE shows LOCAL P.

(2 Theorems $\Rightarrow$ Polar codes ...)

$x_t \in [0, 1]$



- Sketch of Theorem "LOCAL ⇒ STRONG"

$$\phi_t \triangleq \min \{ \sqrt{X_t}, \sqrt{1-X_t} \}$$

- Claim: $\exists \beta < 1$ s.t. $\forall t, X_{t-1}$

$$E[\phi_t \mid X_{t-1}] \leq \beta \cdot \phi_{t-1}$$

Proof of Claim omitted. Calculation

- Claim $\Rightarrow$ $E[\phi_t] \leq \beta^t \cdot X_0$

$$\Rightarrow \Pr[\phi_t > \beta^{t/2} X_0] < \beta^{t/2}$$

$$\Rightarrow \Pr[H(W_i | W_{<i}) \in (2^{-.001t}, 1-2^{-.001t})] \leq 2^{-.001t}$$

Not good enough. GOOD!

Idea: Two step analysis

- first $t/2$ steps: $\Pr[\_ \in [\beta^{t/2}, 1-\beta^{t/2}]] < \beta^{t/2}$

- Second $t/2$ steps:

① $\Pr[\text{not stay below } T_0 \mid \text{start} < \beta^{t/2}] < \frac{\beta^{t/2}}{T_0}$

② if always below T_0 then double w.p. $< 1/2$; drop by c w.p. $\frac{1}{2}$.

$\Rightarrow \log X_t$ has negative drift.

Part 2 analysis

for us

$T = T_0$ absolute const.

$X_0 = \beta^{t/2}$

① Doob's Inequality for martingales

if $X_t \geq 0$ $\quad X_0, X_1, \ldots X_t \ldots$

then $\forall\, T, X_0, t$

$$\Pr\left[\exists\, j \in \{1..t\} \text{ s.t. } X_j \geq T\right] \leq \frac{X_0}{T}.$$

② $\log X_t$ goes down by $\sim \Theta \log c$ in each step.

$$\Rightarrow \quad E[\log X_t] \leq -t\,\Theta \log c.$$

$$\Rightarrow \quad \Pr\left[\log X_t \geq -\frac{t\,\Theta}{2} \log c\right] \leq \exp(-t)$$

$$\Rightarrow \Pr\left[X_t \geq \exp\left(-\frac{t\,\Theta}{2} \log c\right)\right] \leq \ldots$$

Arikan Martingale Polarizes Locally

$H(A|S)$

- Mostly skipped
- Tedious calculations .... But .... Idea below.

  ignore conditioning

$$X_t = h(p)$$

$$\Rightarrow \quad X_{t+1} = h(2p-p^2) \quad \text{w.p. } \tfrac{1}{2}$$

$$= 2h(p) - h(2p-p^2)$$

Variance: (1) $2p-p^2$ bounded away from p

(2) $h(\ )$ continuous

Suction at high end: $p = \frac{1}{2} - \epsilon$

- $X_t = h(p) = 1 - \Theta(\epsilon^2)$
- $2p - p^2 = \frac{1}{2} - \Theta(\epsilon^2)$
- $X_{t+1} = 1 - \Theta(\epsilon^4)$ w.p. $\frac{1}{2}$

$\epsilon^2 \to \epsilon^4$ smaller than any c-factor reduction.

Suction at low end:

$$h(p) \approx p \log \frac{1}{p}$$

$$2p - p^2 \approx 2p$$

$$2h(p) - h(2p) \approx 2p \log \frac{1}{p} - 2p \log \frac{1}{2p}$$

$$= 2p$$

$$= \frac{h(p)}{\log \frac{1}{p}} \approx \frac{h(p)}{\log h(p)}$$

$$\Rightarrow \quad X_{t+1} \le \frac{X_t}{\log X_t} \quad \Rightarrow \text{ subconstant drop.}$$

$$\text{w.p. } \tfrac{1}{2}$$

Summary

Arikan Martingale shows local polarization

$\Rightarrow$ strong polarization

$\Rightarrow$ $\Pr\left[H(W_i \mid W_{<i}) \in (n^{-100}, 1 - n^{-100})\right] \le n^{-.001}$

$\Rightarrow$ Polar codes achieve poly gap to capacity.

$$X_i = X_{i-1} + N\left(0, \cancel{1}\ \frac{X_{i-1}}{\sqrt{i}}\right)$$

This is a martingale

<u>Thm</u>: $\left\{ \frac{X_n}{\sqrt{n}} \right\} \longrightarrow N(0,1)$