ITC CS: CS 229r

2/28/2019

Lecture 10

Today:

- Polar Codes Wrapup
- Start Communication Complexity
  - Model
  - Example Problems
  - Non-Trivial Protocols (with Randomness)
  - Lower Bounds
    - Rank
    - Discrepancy
    - Information Theory.

———— x ————

POLAR CODES

~~Seen~~:

- **Definition**: $X_0, X_1, \ldots X_t \ldots$ $[0,1]$-Martingale **Polarizes Strongly** if $\forall c \; \exists \beta < 1 \; \forall t$

$$\Pr_{X_0 \ldots X_t}\left[X_t \in (c^{-t}, 1-c^{-t})\right] \leq \beta^t$$

- **Defn**: $X_0 \ldots X_t \ldots$ $[0,1]$-Martingale **Polarizes Locally** if it has variance in middle ① + suction at ends ②.

① $\forall \tau > 0 \; \exists \sigma > 0$ s.t. $\forall t$ if $X_{t-1} \in (\tau, 1-\tau)$ then $\mathrm{Var}(X_t \mid X_{t-1}) \geq \sigma^2$

② ~~$\nexists$~~ $\exists \theta > 0$, $\forall c \; \exists \tau > 0$ s.t. if $X_{t-1} < \tau$ then $\Pr\left[X_t \leq \frac{X_{t-1}}{c}\right] \geq \theta$

[similarly for $\tilde{X}_t = 1 - X_t$]

- Defn: Arikan Martingale ~~$\nexists$~~ $X_t \triangleq$ ~~$H(Z_{j_1}^{j_t} \mid Z_{i<j_t}^{j_t})$~~ $H(Z_{j_t}^{t} \mid Z_{i<j_t}^{t})$

- Seen: if A.M. polarizes strongly then Encoding + Decoding work in time $O(n \log n)$ error $\frac{1}{n^{10}}$, $n \leq \mathrm{poly}(1/\epsilon)$

- Sketched: Thm: Local Polarization $\Rightarrow$ Strong Polarization.

Today

Sketch of Arikan Martingale is locally Polarizating.

1. Need to reason about e.g.

$H(X+Y|A,B)$ where $(X,A),(Y,B)$ are i.i.d.

vs.

$H(X|A)$, $H(Y|B)$

2. But for intuition, ignore conditioning

- Suppose $X,Y \sim Bern(p)$ ind. [assume $p<\frac{1}{2}$ wlog]

- if $h(p) \in (\tau, 1-\tau) \Rightarrow p \in (\tau', \frac{1}{2}-\tau')$

$X+Y \sim Bern(2p(1-p))$

$2p(1-p) - p \geq \tau''$ for some $\tau''>0$

$\Rightarrow \underbrace{H(X+Y) - H(X) \geq \tau'''}$ for some $\tau'''>0$

Variance in middle

- $h(p) < \tau \Rightarrow H(X) = h(p) \approx p\log\frac{1}{p} \Rightarrow$ ~~p~~

$2p(1-p) \approx 2p$

$H(Y|X+Y) = H(X)+H(Y)-H(X+Y)$

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

$\lessapprox \frac{2 \cdot H(X)}{\log\frac{1}{H(X)}}$

[Suction at low end]

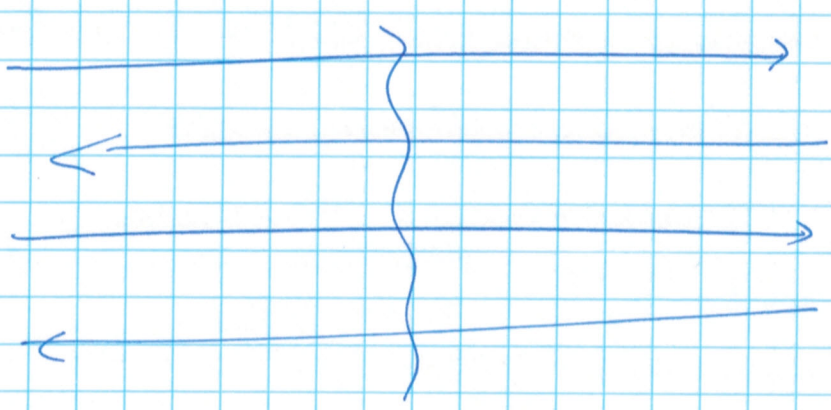
Suction at high end; Removing conditioning ~~finite~~. [Exercise/BET]

Communication Complexity

Model [Yao '80]

$f: \{0,1\}^n \times \{0,1\}^n \to S$

Alice $\leftarrow X \in \{0,1\}^n$ $\qquad$ $Y \in \{0,1\}^n \rightarrow$ Bob



↑ Exchange bits

$b_1 = f_1(X)$

$b_2 = f_2(Y, b_1)$

$b_3 = f_3(X, b_1, b_2)$

$\vdots$

etc.

} Π

- End goal: Determine $f(x,y)$
- Complexity measure: # bits exchanged (no restriction on complexity of $f_1, f_2, \ldots$)
- Randomness: - Can allow private randomness (Alice / Bob on their own)
  - Public Randomness (Alice + Bob share with no cost to sharing).
  - get right answer $\forall x,y$ w.p. $\geq 2/3$

Examples

Trivial: $f(X,Y) = \bigoplus_{i=1}^{n} (X_i \oplus Y_i)$

Use Associativity + Commutivity to get

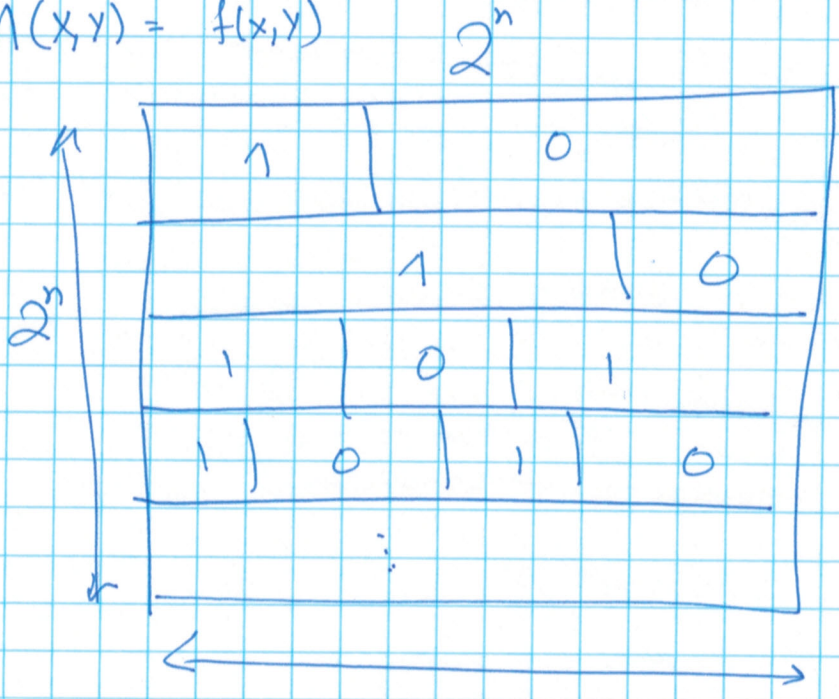
$$f(X,Y) = \underbrace{\left(\bigoplus_{i=1}^{n} X_i\right)}_{b_1} \oplus \underbrace{\left(\bigoplus_{i=1}^{n} Y_i\right)}_{b_2}$$

Solved with 2 bits.

Anything more interesting? ... not really.

Underlying Matrix

$M(X,Y) = f(X,Y)$



k bits of communication suffice

$\Rightarrow$ 2^k bits from Alice $\to$ Bob + 1 bit from Bob $\to$ Alice suffice

$\Rightarrow$ M divisible into 2^k horizontal blocks s.t. rows within block identical

O(1) deterministic communication uninteresting

Thm [Yao]: Det Comm $(M) \geq \log$ rank M

Proof: Above reasoning showed rank$(M) \leq 2^k$. (over any field)

Conjecture: $\forall f$ Det. Comm$(M) \leq (\log \text{rank } M)^{O(1)}$

Best known: [Lovett]: Det Comm$(M) \leq (\text{rank } M)^{1/2}$

Functions with high Det. Comm. Comp.: M = Identity!

Rank$(M_n) = 2^n \Rightarrow$ Det. $\mathcal{C} \geq n+1$

M_n corresponds to $EQ(x,y) = 1$ if $x=y$

$= 0$ o.w.

EQ is really hard? No... actually one of easiest.

With randomness.

Randomized protocol

Let $E: \{0,1\}^n \to \{0,1\}^N$ be a "error-correcting code"

(& satisfying $\forall x \neq y, |\{i : E(x)_i \neq E(y)_i\}| \geq \frac{N}{10}$)

(Such exist if $N \geq 2n$).

Alice picks $i \in [N]$ at random

$i, E(x)_i \rightarrow$

$\leftarrow$ 1 if $E(x)_i = E(y)_i$; 0 o.w.

] $\log n + O(1)$ bits; $\Pr[\text{err}] \leq .9$.

Even better

if Alice & Bob share i then comm. comp $= O(1)$.
[Public randomness]

Theorems: ① if Priv.-cc$(f) = k$ then Det.-cc$(f) \leq 2^k$

② $\not\exists$ Priv-cc$(f) \leq$ Public-cc$(f) + \Omega(\log n)$.

EQ: show both tight

Other nice examples

① Hamming: if $\Delta(x,y) \triangleq \#\{i \mid x_i \neq y_i\} \leq k \Rightarrow 1$
else $\Rightarrow 0$

Public-cc$(\text{Hamming}_k) = O(k \log k)$ [Huang et al.]

② Small Set Disj: $\not\exists$ $wt(x) = wt(y) = k$ $wt(x) \triangleq |\{i \mid x_i = 1\}$

1 if $wt(x), wt(y) \leq k$ & $\exists i$ s.t. $x_i = y_i$

0 o.w. [Hastad Wigderson]: R-cc$(f) = \Theta(k)$

③ Integer Inner Product: $X, Y \in [\pm M]^t$. if $\langle x,y \rangle = 0$?

[Chattopadhyay, Lovett, ?]: Public-cc $= O(t \log M)$.

Main Challenge: Prove high (linear) lower bounds on prob. comm. compl. (with public randomness) of natural functions.

Which ones? How?

Functions

(1) $\mathbb{F}_2$ Inner product : $IP(X,Y) = \sum X_i Y_i \pmod 2$

(2) Disjointness : $\exists i$ s.t. $X_i = Y_i = 1 = \bigvee_i (X_i \wedge X_i)$

(3) Gap Hamming : (Promise Problem) :

$$GH(X,Y) = 1 \quad \text{if } \Delta(X,Y) \geq \frac{N}{2} + \sqrt{N}$$

$$= 0 \quad \text{if } \Delta(X,Y) \leq \frac{N}{2} - \sqrt{N}$$

$$= * \text{ (don't care) otherwise.}$$

How to Prove Lower Bounds?

Usually "Distributional Lower bounds"

(1) Fix distribution ~~$P_{X,Y}$ on (X,Y)~~ μ on (X,Y)

(2) Show ϵ-error deterministic protocols for $(X,Y) \sim \mu$ have large cc, say k

$$\text{error}_{\mu}(\Pi, f) = \Pr_{(X,Y)\sim\mu}[\Pi(X,Y) \neq f(X,Y)]$$

↑ function computed by Π on input (X,Y).

(3) $\Rightarrow$ Public-cc$(f) \geq k$.

Example Discrepancy Method

w.l.o.g. $f: \{0,1\}^n \times \{0,1\}^n \to \{-1,+1\}$

$M_f \in \{-1,1\}^{2^n \times 2^n}$

$$\mathrm{Disc}_\mu(M_f) \triangleq \max_{\substack{S \subseteq \{0,1\}^n \\ T \subseteq \{0,1\}^n}} \left| \sum_{\substack{x \in S \\ y \in T}} \mu(x,y) \cdot M(x,y) \right|$$

Intuition / Motivation: ① if f has k bit protocol with zero-error then $\exists\ S,T$ $|S| \cdot |T| \geq \frac{4^n}{2^k}$

with $M(x,y) = +1 \quad \forall\ (x,y) \in S \times T$

$$\Rightarrow \sum_{\substack{x \in S \\ y \in T}} \mu(x,y)\, M(x,y) = 2^{-k}.$$

② if f is k-bits long & is correct w.p. $\frac{1+\epsilon}{2}$

Then $\mathrm{Disc}_\mu(M_f) \geq \epsilon \cdot 2^{-k}$

$\Rightarrow$ Can Prove CC-lower bound using Disc upper bound

Exercise: Prove $\mathrm{Disc}_{\mathrm{unif}}(\mathrm{IP}_2) \leq \lambda_2(\mu \cdot M)$

$\leq 2^{-n/2}$.