1

(scribe)

Tue 4/16

Last time: $R(\text{DISJ}) = \Omega(\sqrt{n})$ using product distributions.

Hellinger distance:

$$\Delta^2_{Hel}(p,q) \triangleq 1 - \sum_x \sqrt{p(x)\, q(x)}$$

$$\Delta^2_{Hel}(p,q) \leq \Delta_{TV}(p,q) \leq \sqrt{2}\, \Delta_{Hel}(p,q)$$

Today: $R(\text{DISJ}) = \Omega(n)$. [KS'87, Razborov '90, BJKS '04]

$$\text{DISJ}(x,y) = \bigwedge_i (\overline{x_i} \vee \overline{y_i}) = \bigwedge_i \text{NAND}(x_i, y_i).$$

High-level idea: protocol computing DISJ must carry enough information for estimating the individual NANDs.

Let $\sigma \in \{A,B\}^n$. Consider the input distribution $(X_1,Y_1), \dots, (X_n,Y_n)$, where (X_i,Y_i) ~~are~~ is independently sampled from η_A if $\sigma_i = A$ and from η_B if $\sigma_i = B$.

2

$$\begin{cases} \eta_A(1,0) = \eta_A(0,0) = \frac{1}{2}, & \eta_A(0,1) = \eta_A(1,1) = 0 \\ \eta_B(0,1) = \eta_B(0,0) = \frac{1}{2}, & \eta_B(1,0) = \eta_B(1,1) = 0. \end{cases}$$

"Alice active in η_A and Bob active in η_B".

NB. $\mathrm{DISJ}(X,Y) = 1$ w.p. 1.

[similar to the "fooling set argument"].

* Suppose the protocol Π communicates $\leq \delta n$ bits ($|\Pi| \leq \delta n$) and errs with prob $\leq \frac{1}{2} - \varepsilon$.

* $I(X,Y;\Pi) \leq H(\Pi) \leq |\Pi| \leq \delta n$.

Since (X_k, Y_k) is independent for different k,

$$I(X,Y;\Pi) \geq \sum_{k=1}^{n} I(X_k, Y_k; \Pi)$$

$$\Rightarrow \underset{\substack{k \\ \text{uniform in } [n]}}{\mathbb{E}} I(X_k, Y_k; \Pi) \leq \delta$$

$$\Rightarrow \underset{\substack{\sigma \\ \text{uniform random}}}{\mathbb{E}} \underset{k}{\mathbb{E}} \quad '' \quad \leq \delta \Rightarrow \underset{k}{\mathbb{E}} \underset{\sigma}{\mathbb{E}} I(X_k, Y_k; \Pi) \leq \delta$$

3

$\Rightarrow \exists$ fixed k s.t. $\mathbb{E}_{\sigma} I(X_k, Y_k ; \Pi) \leq \delta$.

Def: $\sigma_{-k} := (\sigma_1, \ldots, \sigma_{k-1}, \sigma_{k+1}, \ldots, \sigma_n)$.

$\Rightarrow \mathbb{E}_{\sigma_{-k}} \mathbb{E}_{\sigma_k} I(X_k, Y_k ; \Pi) \leq \delta$

$\Rightarrow \exists$ fixed σ_{-k} s.t. $\mathbb{E}_{\sigma_k} I(X_k, Y_k ; \Pi) \leq \delta$.

$\Rightarrow I(X_k, Y_k ; \Pi \mid \sigma_k = A) + I(X_k, Y_k ; \Pi \mid \sigma_k = B) \leq 2\delta$.

Now, from Π we construct a protocol for computing NAND of two bits $x, y \in \{0,1\}$:

* Alice and Bob set $X_k = x$, $Y_k = y$, and sample (X_i, Y_i), $i \neq k$, according to σ_{-k} and η_A, η_B. Then they run Π.

Note: $\mathrm{DISJ}(X, Y) = \mathrm{NAND}(x, y)$.

$\Rightarrow$ Alice and Bob compute $\mathrm{NAND}(x, y)$ with error $\leq \frac{1}{2} - \varepsilon$.

(4)

* Call the NAND protocol $\pi(x,y)$.

$$\Rightarrow \; I\big(X,Y\,;\,\pi(X,Y)\,\big|\,(X,X)\sim\eta_A\big) \;+$$
$$I\big(X,Y\,;\,\pi(X,Y)\,\big|\,(X,Y)\sim\eta_B\big) \;\le 2\delta.$$

def of η_A, η_B
$$\Rightarrow \; I\big(Z\,;\,\pi(Z,0)\big) + I\big(Z\,;\,\pi(0,Z)\big) \le 2\delta,$$

$Z \sim$ uniform random on $\{0,1\}$.

* Recall again from P.S. 1, Problem 6:

$$I\big(Z;\,\pi(Z,0)\big) \ge \cancel{\mathbb{E}}\; \frac{1}{2}\Big(\Delta^2_{TV}\big(\pi(Z,0),\pi(0,0)\big) + \Delta^2_{TV}\big(\pi(Z,0),\pi(1,0)\big)\Big)$$

(similarly for $\pi(0,Z)$)

$$\cancel{\ge \frac{1}{4}\Big(\Delta_{TV}\big(\pi(Z,0),\pi(0,0)\big) + \Delta_{TV}\big(\pi(Z,0),\pi(1,0)\big)\Big)^2}$$

$$\cancel{\ge \frac{1}{4}\Delta^2_{TV}\big(\pi(0,0),\pi(1,0)\big)}$$

$$\Rightarrow I\big(Z;\pi(Z,0)\big) + I\big(Z;\pi(0,Z)\big) \ge$$

$$\frac{1}{2}\Big(\Delta^2_{TV}\big(\pi(Z,0),\pi(0,0)\big) + \Delta^2_{TV}\big(\pi(Z,0),\pi(1,0)\big)$$
$$+\,\Delta^2_{TV}\big(\pi(0,Z),\pi(0,0)\big) + \Delta^2_{TV}\big(\pi(0,Z),\pi(0,1)\big)\Big)$$

Cauchy-Schwarz
$$\ge \frac{1}{8}\Big(\Delta_{TV}\big(\pi(Z,0),\pi(0,0)\big) + \Delta_{TV}\big(\pi(Z,0),\pi(1,0)\big) + \Delta_{TV}\big(\pi(0,Z),\pi(0,0)\big) + \Delta_{TV}\big(\pi(0,Z),\pi(0,1)\big)\Big)^2$$

5

Triangle Inequality

$$\geq \frac{1}{8}\left(\Delta_{TV}^2\left(\pi(1,0), \pi(0,1)\right)\right)$$

In fact, we could have worked with Hellinger distance, and it's true that for $Z \sim \{0,1\}$,

$$I(Z; f(Z)) \geq \Delta^2_{Hel}\left(\cancel{f_0(Z), f_2(Z)}\ f(0), f(1)\right)$$

where $f(\cdot)$ is a randomized function. ~~and~~ (Exercise)

~~$f_0(Z)$ is the distribution~~

$$\Rightarrow 2\delta \geq I\left(Z; \pi(Z,0)\right) + I\left(Z; \pi(0,Z)\right)$$

$$\geq \frac{1}{2}\Delta^2_{Hel}\left(\pi(1,0), \pi(0,1)\right)$$

* Recall: (from last time) $\Delta_{Hel}\left(\pi(1,0), \pi(0,1)\right) =$

$$\Delta_{Hel}\left(\pi(0,0), \pi(1,1)\right)$$

$$\Rightarrow 2\delta \geq \frac{1}{2}\Delta^2_{Hel}\left(\pi(0,0), \pi(1,1)\right)$$

(Hel → TV)

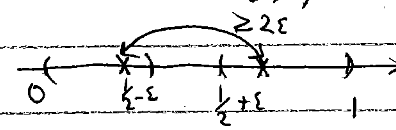
$$\geq \frac{1}{2}\Delta^2_{TV}\left(\pi(0,0), \pi(1,1)\right)$$

Note: $\Delta_{TV}(p,q) = \max_{S \subseteq supp(p)} \left| p(S) - q(S) \right|$.

$\Rightarrow \forall S, \ \Delta_{TV}(p,q) \geq \left| p(S) - q(S) \right|$

$$\Rightarrow \overset{\Delta_{TV}(\pi(0,0),\pi(1,1))}{\cancel{\delta}} \geq \Big| Pr\big(\text{Output of transcript } \pi(0,0) = \overset{NAND(0,0)}{0}\big) - Pr\big(\text{'' '' '' } \pi(1,1) = \underset{NAND(1,1)}{0}\big) \Big|$$

$$\geq 2\varepsilon.$$



$$\Rightarrow 2\delta \geq \frac{1}{4} (2\varepsilon)^2 = \varepsilon^2.$$

$$\Rightarrow R_{\frac{1}{2}-\varepsilon}(DISJ) \geq \frac{\varepsilon^2}{2} \cdot n$$

□

In fact, $R_{\frac{1}{2}-\varepsilon}(DISJ) = \Omega(\varepsilon \cdot n)$ (optimal)

(Braverman, Moitra '12)

<u>Simple application:</u> Moments in the streaming model.

<u>Setting:</u> A sequence $a_1, a_2, \dots, a_m$, $a_i \in [n]$, arrives as a stream.

$\forall i \in [n]$, $f_i \triangleq \left| \{ j \in [m],\ a_j = i \} \right|$.

Goal: Compute $\max_i f_i$.

Challenge: Use as little memory as possible.

Theorem: Any streaming algorithm requires $\Omega(n)$ space.

<u>Proof:</u> Reduction from disjointness.

Given (x, y) to DISJ, streaming algorithm A.

Alice: $x \mapsto$ sequence $a_x = \{ i \mid x_i = 1 \}$,

Alice Runs A on a_x and sends the state of A (C bits, $C \leq$ memory usage of A) to Bob.

Bob: $y \mapsto$ sequence $b_y = \{ i \mid y_i = 1 \}$.

Bob continues execution of A with sequence b_y.

Observe: $\begin{cases} \max_i f_i = \begin{cases} 2 & \text{if } DISJ(x,y) = 0 \\ 1 & \text{else.} \end{cases} \\ \text{Communication Cost} = C+1. \end{cases}$

$\Rightarrow C = \Omega(n)$. (by disjointness lower bound)

$\square$

Information Cost:

Recall that for the lower bounds we looked at:

Def: $IC_{ext}(\Pi, \mu) = I(X, Y; \Pi)$
$(X,Y) \sim \mu$.

That is, what an "external observer" learns from (X,Y) by observing the transcript Π.

★ We can also define a similar notion on what Alice and Bob learn about each other's input from Π:

Def: $IC(\Pi, \mu) := I(\Pi; Y \mid X) + I(\Pi; X \mid Y)$
$((X,Y) \sim \mu)$