1

# Parallel Repetition, continued:

* Game $\mathcal{G}$, $\mathcal{G}^n$: n-repeated game.

  $\omega(\mathcal{G}) = 1-\delta$, $q = \#$ possible answers.

  $\Rightarrow \omega(\mathcal{G}^n) \leq e^{-\Omega_{\delta,q}(n)}$.

* $\Pr(W_1 \wedge \cdots \wedge W_n) = ?$ $W_i =$ event that you win on the $i^{th}$ game.

* Need to show: $\exists j$, $\Pr(W_j \mid W_1 \wedge \cdots \wedge W_m)$ is small

  (unless $\Pr(W_1 \wedge \cdots \wedge W_m)$ itself is small) $\leq 1 - \delta/2$.
  (or $m \geq \mu n$)

Idea: For some $j \geq m$, show the following can be achieved:

1. Prover 1 (Alice) on input $x$, produces $\overline{x}^n$ s.t. $\overline{x}^n_j = x$

2. Bob, on input $y$, produces $\overline{y}^n$ s.t. $\overline{y}^n_j = y$.

3 $\left\| (\bar{x}^n, \bar{y}^n) - (\tilde{x}^n, \tilde{y}^n) \right\| < \varepsilon,$

(distribution)

TV distance

where

$$(\tilde{x}^n, \tilde{y}^n) \triangleq (x^n, y^n \mid w_1 \wedge w_2 \wedge \cdots \wedge w_m)$$

where $(x, y) \leftarrow (X, Y)$

This implies that, given such an embedding, get a strategy for $\mathcal{G}$ with value $\geq 1 - \delta/2 - \varepsilon$.

(Get a contradiction by choosing $\varepsilon < \delta/2$).

★ *Necessary Condition:*

$$\| (X, Y) - (\tilde{X}_j, \tilde{Y}_j) \|_{TV} < \varepsilon$$

(otherwise there obviously is no hope!)

$(X^n, Y^n \mid w_1 \wedge \cdots \wedge w_n)$

3

**Lemma:** If $\mathcal{U} = \mathcal{U}_1 \ldots \mathcal{U}_n$ is a product distribution, and $\mathcal{E}$ is an event,

$$\frac{1}{n} \sum_{j=1}^{n} \| \mathcal{U}_j - \mathcal{U}_{j|\mathcal{E}} \|_{TV} \leq \sqrt{\frac{1}{n} \log \frac{1}{\Pr(\mathcal{E})}}$$

**Proof:** 1. $\mathcal{D}(p \| q) \geq \| p - q \|_{TV}^2$ (Pinsker)

2. If $\bar{q} = q_1 \ldots q_n$ is a product distribution and $\bar{p} = p_1 \ldots p_n$ is <u>any</u> distribution,

then $\mathcal{D}(\bar{p} \| \bar{q}) \geq \sum_{i=1}^{n} \mathcal{D}(p_i \| q_i)$.

$p_i$: marginal of $\bar{p}$ on coord. $i$.

$$\mathcal{D}(p(x,y) \| q(x) q(y)) \geq \mathcal{D}(p(x) \| q(x)) + \mathcal{D}(p(y) \| q(y))$$

Pf: $$\sum_{x,y} p(x,y) \log \frac{p(x,y)}{q(x)q(y)} = \sum_{x,y} p(x) \left[ \underbrace{\log \frac{p(x,y)}{p(x)p(y)}}_{(\geq 0)} + \log \frac{p(x)}{q(x)} + \log \frac{p(y)}{q(y)} \right]$$

$$\geq \mathcal{D}(p(x) \| q(x)) + \mathcal{D}(p(y) \| q(y)).$$

Now, $$LHS^2 \underset{\text{Cauchy Schwarz}}{\leq} \frac{1}{n} \sum_i \| \mathcal{U}_i - \mathcal{U}_{i|\mathcal{E}} \|^2 \overset{\text{Pinsker}}{\leq} \frac{1}{n} \sum \mathcal{D}(\mathcal{U}_{i|\mathcal{E}} \| \mathcal{U}_i)$$

④

$$\leq \frac{1}{n} \mathcal{D}(\mathcal{U}_{|\mathcal{E}} \| \mathcal{U})$$

$$= \frac{1}{n} \sum_u \Pr(\mathcal{U}=u|\mathcal{E}) \log \frac{\Pr(\mathcal{U}=u|\mathcal{E})}{\Pr(\mathcal{U}=u)}$$

$$= \frac{1}{n} \sum_u \Pr(\mathcal{U}=u|\mathcal{E}) \log \frac{\Pr(\mathcal{U}=u \wedge \mathcal{E})}{\Pr(\mathcal{E})\Pr(\mathcal{U}=u)}$$

$$= \frac{1}{n} \log \frac{1}{\Pr(\mathcal{E})} + \underbrace{\frac{1}{n} \sum_u \Pr(\mathcal{U}=u|\mathcal{E}) \log \frac{\Pr(\mathcal{U}=u \wedge \mathcal{E})}{\Pr(\mathcal{U}=u)}}_{\leq 0}.$$

□

* Variant: Suppose $T$ is some other r.v. s.t.

$\forall t \in \mathrm{supp}(T)$,

$\mathcal{U}_{1|T=t}, \ldots, \mathcal{U}_{n|T=t}$ are independent.

$$\Rightarrow \sum_t \Pr(T=t|\mathcal{E}) \left[\frac{1}{n} \sum_{j=1}^{n} \| \mathcal{U}_{j|T=t} - \mathcal{U}_{j|\mathcal{E} \wedge T=t} \| \right] \leq \sqrt{\frac{1}{n} \log \frac{1}{\Pr(\mathcal{E})}}$$

Pf: Apply the lemma for each fixing of $t$, and use concavity. □

5

## Back to parallel repetition:

**Lemma (A):** $\forall m \le n$ & $k > m$.

Let $\Pr\left[W_k \mid W_1 \wedge \dots \wedge W_m\right] = \omega(\mathcal{G}) + \varepsilon_k$.
$(= 1 - \delta + \varepsilon_k)$

Then,

$$\frac{1}{n-m} \sum_{k=m+1}^{n} \varepsilon_k \lesssim \sqrt{\frac{1}{n-m}\left[\log \frac{1}{p_m} + m \log q\right]},$$

where $p_m = \Pr\left(W_1 \wedge \dots \wedge W_m\right)$

(follows by just calculations and using the following lemma B)

Easier to condition on both

$W = W_1 \wedge \dots \wedge W_m$ and answers to the

first $m$ questions: $V \in [q]^m \times [q]^m$

6

**Lemma B:** $\forall m \leq n$, $\forall k > m$,

(Notation: $E(V) = W_1 \wedge \dots \wedge W_m$ & answers equal $V$)

if $\Pr(W_k \mid E(V)) = 1 - \delta + \varepsilon_k$,

Then, $$\frac{1}{n-m} \sum_{k=m+1}^{n} \varepsilon_k \lesssim \sqrt{\frac{1}{n-m} \log \frac{1}{\Pr(E(V))}}$$

$\square$

**Correlated Sampling:**

How will $T$ look like?

$$T: (T_1, \dots, T_n), \quad T_i = (X_i, Y_i) \sim (X, Y)$$

for $i = 1, \dots, m$

$\forall j \geq m$, $T_j = (b_j, Z_j)$,

$$\begin{cases} b_j = 0 \Rightarrow Z_j : X_j \sim X \\ b_j = 1 \Rightarrow Z_j : Y_j \sim Y. \end{cases}$$

Alice and Bob need to sample $T = t \mid E(V)$

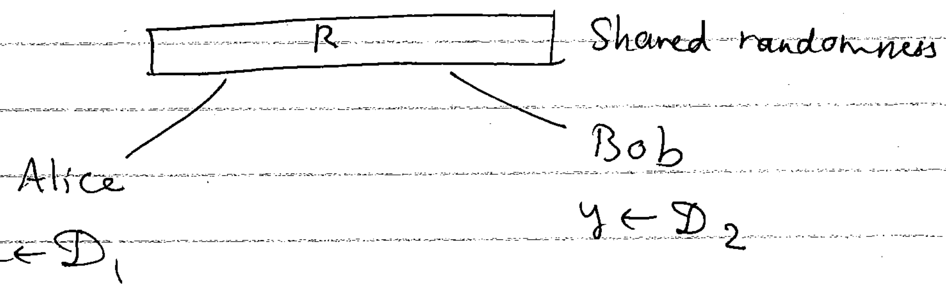
without communication and using shared randomness.

7

**Def.** $T_{-k} = (T_1, \ldots, T_{k-1}, T_{k+1}, \ldots, T_n)$.

$$\begin{cases} \text{Alice is in the world } T_{-k} \mid E(v) \wedge X_k = x \\ \text{Bob } '' \; '' \; '' \; '' \; T_{-k} \mid E(v) \wedge Y_k = y. \end{cases}$$

Correlated Sampling:




Theorem: $\exists$ strategy for Alice & Bob to sample

$$x = \text{Alice}(\mathcal{D}_1, R), \quad y = \text{Bob}(\mathcal{D}, R)$$

s.t. $x \sim \mathcal{D}_1$, $y \sim \mathcal{D}_2$,

and $\Pr(x \neq y) \leq 2 \, \| \mathcal{D}_1 - \mathcal{D}_2 \|_{TV}$.

Pf: $\mathcal{U}$ = universe on which $\mathcal{D}_1, \mathcal{D}_2$ are supported

They Interpret $R$ as a sequence:

$(u_1, p_1), (u_2, p_2), \ldots$

8

Where each $u_i$ is uniform from $\mathcal{U}$,

each $p_i$ is uniform in $[0,1]$.

Strategy:

<u>Alice:</u> Find the first $i$ for which

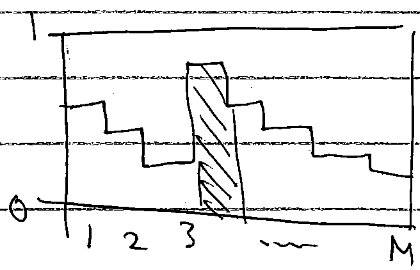
$$p_i \leq \mathcal{D}_1(u_i)$$

return $u_i$ for the sample.

<u>Bob:</u> Similarly find the first $j$ for which

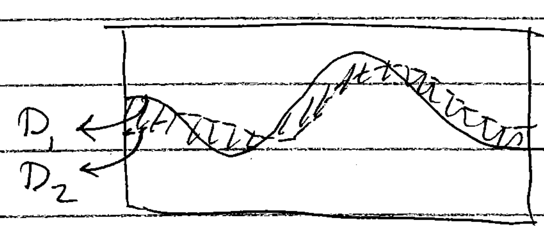
$$p_j \leq \mathcal{D}_2(u_j) \text{ , return } u_j.$$

Histogram of $\mathcal{D}_1$:




Clearly, Pr. of each element is exactly what we want. (we are throwing darts at the rectangle, picking the first that lies below the graph).

We can now draw the graphs for $\mathcal{D}_1$ and $\mathcal{D}_2$:




The two samples disagree when the two darts fall in between the two graphs. (Prob. $\leq 2 \cdot$ statistical distance!) □