Problem Set 2

- Due Date: **28 Oct (Fri), 2011 (TIFR)**
- If you submit handwritten solutions, start each problem on a fresh page.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Referring sources other than the lectures is strongly discouraged. But if you do use an outside source (eg., other text books, lecture notes, any material available online), ACKNOWLEDGE it in your writeup.
- The points for each problem are indicated on the side.
- If you don't know the answer to a problem, then just don't answer it. Do not try to convince yourself or others into believing a false proof.
- Be clear in your writing.

---

1. [**Corruption bound and inner product**] (8+7)

   (a) Let $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a Boolean function and μ be a distribution on $\mathcal{X} \times \mathcal{Y}$ such that for every rectangle $R = S \times T \subseteq \mathcal{X} \times \mathcal{Y}$ with $\mu(R) > \rho$, we have that

   $$\mu(R \cap f^{-1}(1)) > \varepsilon \cdot \mu(R \cap f^{-1}(0)).$$

   In other words, every large rectangle is 1-corrupted. Then, prove that for every $\delta > 0$, we have

   $$2^{\mathrm{R}_\delta(f)} \geq \frac{1}{\rho} \cdot \left(\mu(f^{-1}(0)) - \frac{\delta}{\varepsilon} \right).$$

   (b) Using the above corruption bound or otherwise, prove that

   $$\mathrm{R}_{\frac{1}{2}-\varepsilon}(\mathsf{IP}) \geq n - O\left(\log \frac{1}{\varepsilon}\right) - O(1).$$

2. [**Private coins vs. Public coins for the zero-error case**] (18)

   Prove that for any $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, we have

   $$\mathrm{R}_0(f) = O\left(\mathrm{R}_0^{\mathrm{pub}}(f) + \log n\right).$$

   [Hint: Modify the public coins to private coins transformation (see Theorem 3.14 in Kushilevitz-Nisan book) for bounded error protocols.]

3. [**Augmented Index function**] (10)

   Consider the following 2-party randomized one-way communication problem aIND. Alice is given a n bit string $x \in \{0,1\}^n$ while Bob is given an index $i \in [n]$ and the first $(i-1)$ bits of x, i.e., $x_1, \ldots, x_{i-1}$. The goal of the protocol is for Bob to deterimine x_i. Prove that the lower bound proved in class for the index function continues to hold even for this augmented version of the problem. More precisely, show that

   $$\mathrm{R}^{A \to B}_{\frac{1}{2}-\varepsilon}(\mathsf{aIND}) \geq 2 \log e \cdot \varepsilon^2 n.$$

4. **[Multi-party disjointness]** (5+10+5)

Recall the proof of the multi-party disjointness UDISJ proved in lecture. We showed that the private coins randomized communication complexity of $\mathsf{UDISJ}_{n,t}$ in the number-in-hand broadcast model is $\Omega(n/t^2)$. An intermediate step in the proof of this result involved showing the following geometric inequality:

$$\sum h^2\left(\Pi_{\overline{0}}, \Pi_{e_i}\right) \geq t \cdot h^2\left(\Pi_{\overline{0}}, \Pi_{\overline{1}}\right),$$

where $h^2(\cdot,\cdot)$ denotes the squared-Hellinger distance and Π_z denotes the transcript distribution on input z. In this problem, we will improve the lower bound to $\Omega(n/t)$ by improving the above geometric inequality to

$$\sum h^2\left(\Pi_{\overline{0}}, \Pi_{e_i}\right) \geq O(1) \cdot h^2\left(\Pi_{\overline{0}}, \Pi_{\overline{1}}\right),$$

(a) Prove that for any $n+1$ vectors $v_0, v_1, \ldots, v_n$, we have

$$\sum_{i=1}^{n} \|v_0 - v_i\|_2^2 \geq \frac{1}{n} \cdot \sum_{1 \leq i < j \leq n} \|v_i - v_j\|_2^2.$$

[Hint: Consider the sum $\sum_{i,j \in \{0,1,\ldots,n\}} b_i \cdot b_j \cdot \|v_i - v_j\|_2^2$ where $\sum_{i=0}^{n} b_i = 0$.]

(b) Suppose $A_1, \ldots, A_n$ are pairwise disjoint collection of $n = 2^k$ subsets of $[t]$. Set $A = \bigcup_i A_i$. Show that

$$\sum_{i=1}^{n} h^2\left(\Pi_{\emptyset}, \Pi_{A_i}\right) \geq h^2\left(\Pi_{\emptyset}, \Pi_A\right) \cdot \prod_{l=1}^{k} \left(1 - \frac{1}{2^l}\right).$$

[Hint: Use induction on k and Part ?? for each step of the induction.]

(c) Conclude that the private coins randomized communication complexity of $\mathsf{UDISJ}_{n,t}$ in the number-in-hand broadcast model is $\Omega(n/t)$.

[Hint: You may use the fact that $\prod_{l=1}^{\infty}\left(1 - \frac{1}{2^l}\right)$ is a constant $(= 0.288788\ldots)$, also called the digital search tree constant (cf., http://oeis.org/A048651).]

5. **[streaming algorithms for Dyck variations?]** (12)

Consider the following variation of the context-free grammar that generates 2-Dyck language:

$$S \to SS, (S), (S], [S], \varepsilon.$$

Let L be the language generated by the above grammar. Note that L is the set of well-formed paratheses of 2-types with the modification being that (can be closed by either) or] while [can be closed only by]. Show that any r-pass (randomized) streaming algorithm for L requires space at least $\Omega(n/r)$.

6. **[Aborting index function problem]** (10)

Suppose Π is a randomized one-round protocol (Alice send a message, Bob guesses) for the index function problem with satisfying the following conditions. Let $p \in [0,1]$ and $q \in [\frac{1}{2}, 1]$. For every input,

(a) Bob may abort without giving a 0-1 answer, but he may do so with probability at most $1-p$;

(b) Conditioned on Bob declaring the answer, the probability that the answer is correct is at least q.

Show that Alice must send at least $p(1-H(q))n$ bits in any such protocol.

7. **[Chasing the pointer's head]** (15)

Consider the following variant of the pointer chasing problem P_k. There are $k+1$ layers of vertices: $L_0, L_1, \ldots, L_k$, (L_0 has only one vertex v_0). Pointers go from one layer to the next and define a unique path $v_0, v_1, \ldots, v_k$ from L_0 to L_k. Alice has the pointers leaving the even layers and Bob has the pointers leaving the odd layers. The goal is to determine the msb of v_k. Suppose Alice starts the communication. Show that there is an $O(kn)$-bit $\left\lfloor \frac{k}{2} \right\rfloor$-round deterministic protocol for this problem.