| **CS 153 Current topics in theoretical computer science** | **Spring 2018** |
|---|---|
| Problem Set 2 | |
| *Out: May 22* | *Due:* **June 1** |

You are encouraged to work in groups of two or three; however you must turn in your own write-up and note with whom you worked. Please don't consult online solutions or original research papers or surveys containing such solutions while doing this problem set. Please attempt all problems.

1. Define the *triangle-rank* of a Boolean matrix to be the sidelength of the largest square submatrix (possibly after permuting the rows and columns) with ones on the diagonal and zeros below the diagonal. Notice that triangle-rank is at most the standard rank.

   (a) Consider a matrix $M$ with rows and columns indexed by sets $X$ and $Y$. For subsets $S \subset X$ and $T \subset Y$, denote by $M_1$ the submatrix of $M$ restricted to rows indexed by $S$ and denote by $M_2$ the submatrix of $M$ restricted to columns indexed by $T$. Show that

   $$\text{triangle-rank}(M_1) + \text{triangle-rank}(M_2) \leq \text{triangle-rank}(M).$$

   (b) Let $M_f$ be the matrix associated with function $f : X \times Y \to \{0, 1\}$. Prove that

   $$D(M_f) \leq (\log_2(\text{triangle-rank}(M_f) + 1) + 1)(N^0(M_f) + 1).$$

   Hint: Induction on triangle-rank. Fix a 0-cover. Alice should focus on rectangles $S \times T$ in this cover for which restricting $M$ to the rows indexed by $S$ reduces the triangle rank by at least a factor 2; Bob should focus on rectangles $S \times T$ in this cover for which restricting $M$ to the columns indexed by $T$ reduces the triangle rank by at least a factor of 2.

2. Say that the Boolean-rank of an $N \times N$ Boolean matrix $M$ is the smallest $r$ for which

   $$M = UV,$$

   where $U$ is an $N \times r$ matrix, and $V$ is an $r \times N$ matrix, and $U, V$ are both Boolean matrices. Prove that the log-rank conjecture holds for this notion of rank: show that for $f : X \times Y \to \{0, 1\}$, it holds that

   $$D(M_f) \leq O(\log^2 \text{Boolean-rank}(M_f)).$$

3. This problem concerns the $k$-party number-on-forehead model. Recall that the discrepancy of $f : X_1 \times \cdots \times X_k \to \{0, 1\}$ with respect to a distribution $\mu$ on the inputs, is $\text{disc}_\mu(f)$

$$= \max_S \left| \Pr_\mu[f(x_1, \ldots, x_k) = 0 \wedge (x_1, \ldots, x_k) \in S] - \Pr_\mu[f(x_1, \ldots, x_k) = 1 \wedge (x_1, \ldots, x_k) \in S] \right|$$

where the maximum is taken over all cylinder intersections $S$. Recall also that small discrepancy for some distribution $\mu$, implies large randomized communication complexity in the $k$-party model:

$$R_{1/2-\epsilon}(f) \geq \log_2 \left( \frac{2\epsilon}{\mathrm{disc}_\mu(f)} \right).$$

For example we showed in class that

$$\mathrm{disc}_{\mathrm{uniform}}(\mathrm{GIP}_k) \leq \exp\left(-n/2^k\right)$$

which led to an asymotically tight lower bound on the randomized communication complexity. We now turn to the $k$-party version of disjointness, denoted $\mathrm{DISJ}_k$, in which player $i$ holds an $n$-bit Boolean string $x_i$, viewed as a set, and the output of the protocol is 1 iff there is no global intersection, i.e., if $\cap_i x_i = \emptyset$.

(a) Prove that for any function $f : X_1 \times \cdots \times X_k \to \{0, 1\}$ with nondeterministic or co-nondeterministic communication complexity $t$, we have

$$\mathrm{disc}_\mu(f) \geq \Omega(2^{-t})$$

for all distributions $\mu$. Recall that nondeterministic communication complexity is the base-2 log of the cardinality of the smallest (non-disjoint) cover of $f^{-1}(1)$ with 1-monochromatic cylinder intersections, and the co-nondeterministic communication complexity is the base-2 log of the cardinality of the smallest (non-disjoint) cover of $f^{-1}(0)$ with 0-monochromatic cylinder intersections. Hint: consider the "full" cylinder $X_1 \times \cdots \times X_k$ first.

(b) Prove that $\mathrm{DISJ}_k$ has co-nondeterministic communication complexity at most $\log_2 n$.

We conclude that the discrepancy method cannot prove strong lower bounds on the the randomized communication complexity of disjointness in the $k$-party number-on-forehead model. In contrast, the best known lower bound, by Sherstov, is $\omega\left(\frac{\sqrt{n}}{2^k k}\right)$.