CS 153 Communication Complexity **Timur Kuzhagaliyev**

Problem Set #2 Solutions

Out: May 22 *Due:* **June 1**

1. (a) We can show that the inequality in the question holds by considering the structure of M_1 and M_2. The breakdown of matrix of M into M_1 and M_2looks like

$$\begin{pmatrix} \times & X_1 \times T & \times \\ S \times Y_1 & S \times T & S \times Y_2 \\ \times & X_2 \times T & \times \end{pmatrix} \tag{1}$$

where the highlighted row is matrix M_1, the highlighted column is the matrix M_2 and X_1, S, X_2 with Y_1, T, Y_2 are partitions of X and Y respectively. Note that due to the structure of given matrices, the standard rank of M_1 is bound above by $|S|$ and the standard rank of M_2 is bound above by $|T|$, and the same bounds apply to triangle-rank.

Note also that $S \times T$ is a zero matrix - this means that regardless of how rows or columns from $S \times T$ are permuted within M_1 or M_2, they will not change the triangle-rank. As such, there is no reason to permute rows indexed by S in M_2 and no reason to permute columns indexed by T in M_1. This creates a convenient "mask" that lets us permute rows and columns of M and build desired square submatrices inside M_1 and M_2 without undoing any progress.

Now we can put the two points from above together to define a procedure that proves the desired inequality. We start off with M. Next we permute rows and columns of M to build the largest square submatrix with ones in diagonals on zeros below the diagonal in both M_1 and M_2 (showing triangle-rank of these matrices). As discussed previously, we can do that without any clashes.

Denote such square submatrices S_1 in M_1 and S_2 in M_2. Finally, we permute columns of resultant M to put S_1 into the top left corner of block $S \times Y_2$, and then permute rows of resultant M to put S_2 into the bottom right corner of block $X_1 \times T$. As a result of this operation, we have transformed M in such a way that it now contains a square matrix that is a tensor product of S_1 and S_2 (with the exception of the top right corner which contains arbitrary values and doesn't affect triangle-rank). Hence we have shown that if matrices M_1 and M_2 have triangle-rank of m and n respectively, we can permute rows and columns of M to build a square submatrix of side length $m + n$ with ones on the diagonal and zeroes below the diagonal, showing that the triangle-rank of matrix M must be at least $m + n$.

(b) We can prove the desired bound using induction on triangle-rank. We can start by fixing a 0-cover on M_f.

2. In class, we have covered a theorem which stated that for any boolean function f with rank$(M_f) = r$ and a monochromatic rectangle R in M_f with size $|R| \geq 2^{-c(r)}|X||Y|$, the deterministic complexity is

$$D(f) \leq O(\log^2 r) + O\left(\sum_{i=1}^{\log r} c(\frac{r}{2i})\right)$$

.

Clearly, showing that $c(r) = \log(r)$ implies that the log-rank conjecture holds. To show that the conjecture holds with our notion of Boolean-rank, we can show that any $f : X \times Y \to \{0, 1\}$ with Boolean-rank$(M_f) = r$ admits a monochromatic rectangle R of size $|R| \geq \frac{1}{r}|X||Y|$.

TODO: *Can show this by proving that there must a big monochromatic rectangle of size $\geq 2^{-c(r)}|X||Y|$, then showing that we can recursively say yes or no to recurse into one of the smaller rectangles, at every step either halving the rank of making the matrix much smaller.*

3. (a) First, we can observe that the definition of discrepancy implies that we can invert Boolean outputs of the function without changing the discrepancy, since we're using the absolute difference between probabilities.

We can start by looking at the full $X_1 \times \ldots \times X_k$ cylinder first. With this choice of S, the calculation for

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

reduces into just

$$\left|\Pr_{\mu}[f(x_1, \ldots, x_k) = 0] - \Pr_{\mu}[f(x_1, \ldots, x_k) = 1]\right|.$$

It's clear that the maximum possible value is 1, when the function either outputs always 1 or always 0. When we don't have this trivial case, the two terms begin to cancel each other out. That is, if we WLOG assume $\Pr_{\mu}[f(\ldots) = 1]$ is the largest of the two, we know that

$$\left|\Pr_{\mu}[f(\ldots) = 1]\right| > \left|\Pr_{\mu}[f(\ldots) = 0] - \Pr_{\mu}[f(\ldots) = 1]\right|.$$

Due to this fact, we can aim to maximise the value by somehow getting rid of the smaller component. We can rewrite the definition of discrepancy as follows:

$$\text{disc}_{\mu}(f) = \max_{S} \left|\Pr_{\mu}\left(f(x^k) = 1\right) \Pr_{\mu}\left(x^k \in S \mid f(x^k) = 1\right) - \Pr_{\mu}\left(f(x^k) = 0\right) \Pr_{\mu}\left(x^k \in S \mid f(x^k) = 0\right)\right|$$

From now on, we will WLOG focus on 1-monochromatic cylinder intersections and assume given f has nondeterministic complexity t. Identical argument applies to the co-nondeterminstic version. If we WLOG consider only 1-monochromatic cylinder intersections, it's clear the second term becomes zero and we're left off with just $\Pr_{\mu}\left(f(x^k) = 1\right) \Pr_{\mu}\left(x^k \in S \mid f(x^k) = 1\right)$.

We're given that f has nondeterministic complexity t, which implies that there exists a non-disjoint cover of f with 1-monochromatic cylinder intersections, with the size of this cover being 2^t. Since these cylinder intersections cover *all* 1s and there are 2^t of them, we know that the largest cylinder intersection in the cover must cover at least 2^{-t} fraction of inputs x^k s.t. $f(x^k) = 1$. As a result, we can conclude that there exists a 1-monochromatic cylinder intersection S such that $\Pr_{\mu}\left(x^k \in S \mid f(x^k) = 1\right) \geq 2^{-t}$.

The function f and given distribution μ are fixed, so we can conclude that

$$\Omega(\Pr_{\mu}\left(f(x^k) = 1\right) \Pr_{\mu}\left(x^k \in S \mid f(x^k) = 1\right))$$

simplifies to just $\Omega(\Pr_{\mu}\left(x^k \in S \mid f(x^k) = 1\right))$, and, finally, using the result from above, to $\Omega(2^{-t})$ when we use the largest 1-monochromatic cylinder intersection. This shows that the discrepancy is bounded below by $\Omega(2^{-t})$ given that nondeterministic or co-nondeterministic communication complexity of f is t.

TODO: *Check that this works for all distributions, not just uniform?*

(b) We can view the input for the problem DISJ$_k$ as a $k \times n$ Boolean matrix, where each player i holds the string corresponding to ith row. In this formulation, the co-DISJ$_k$ problem reduces to finding a column of all 1's (which implies a global intersection).

An all powerful prover can point out the index of the column that contains all 1's, which would require $\log n$ bits to represent. Then, the first player can check everyone else's values in that column and output 1 if they are indeed all 1's. The second player can check first player's value and output 1 in the same case, requring $O(\log n)$ bits in total to solve co-DISJ$_k$.