Communication Complexity 23 Nov, 2011 (@ TIFR)

28. A discrepancy result on Multiparty Set Disjointness

Lecturer: Jaikumar Radhakrishnan *Scribe: Sagnik Mukhopadhyay*

Last time we derived a lower bound for the multiparty set disjointness problem in the number in forehead model. The input to the problem was k subsets of $[n]$, given to k players, represented as an $n \times k$ matrix, each column of which is the characteristic vector of each subset. We assumed that each player can see other players' input but not her own (Number in the forehead model), i.e., player i can see the whole input matrix except column i. Now, finding out whether these sets are disjoint amounts to the same problem as to find out whether there is an all 1's row in the input matrix. So the protocol gives answer $+1$ if there is an all 1's row in the input matrix, -1 otherwise. Keeping this model in mind, we proved the following result by Shrestov.

$$R_{1/3}(\mathsf{DISJ}_{\mathsf{n,k}}) = \Omega\left(\ \tfrac{n}{4^k} \ \right)^{\frac{1}{4}}.$$

In this lecture we will analyze the k-party discrepancy of disjointness model, which is nothing but AND of m independent copies of disjointness problem, as we have seen in the last lecture. We call each copy as X_i for $i \in [m]$. Each X_i is an $n_i \times (k+1)$ matrix, whose first column, x^i, is chosen from an uniform distribution over n_i-length boolean vector (i.e., from distribution $\mathcal{U}_{n_i}$) and the rest $n_i \times k$ submatrix W^i chosen from an uniform distribution on those matrices in $\{0,1\}^{n_i \times k}$ that have exactly one row composed of all 1's (i.e., from distribution $\mu_{n_i,k}$). For positive integers $n_1, ..., n_m$, define

$$\Gamma_k(n_1, ..., n_m) = \max_{\mathcal{X} \in \mathcal{C}} \left| E_{X_1,\cdots,X_n} \left[\mathcal{X}(X_1, \cdots, X_n) \prod_{i=1}^{m} \mathsf{DISJ}_{\mathsf{n_i,k+1}}(X_i) \right] \right|$$

where $X_i \sim \mathcal{U}_{n_i} \times \mu_{n_i,k}$ independently for each i and the maximum is taken over all $(k+1)$-dimensional cylinder intersections. We will proved the following theorem by Sherstov.

Theorem 28.1. *For all positive integers $n_1, ..., n_m$ and k*

$$\Gamma_k(n_1, ..., n_m) \leq \frac{(2^k - 1)^m}{\sqrt{\prod_{i=1}^{m} n_i}}$$

.

If we replace n_i by r for all i, then we get the expression for the bound on $\alpha(S)$ from the previous lecture which was left to prove.

28.1 Proof of Theorem 28.1

This is a proof by induction on k. We first prove the base case, which is the following lemma.

Lemma 28.2. *For all positive integers $n_1, ..., n_m$*

$$\Gamma_1(n_1, ..., n_m) \leq \frac{1}{\sqrt{\prod_{i=1}^{m} n_i}}$$

.

28.1.1 Proof of Lemma 28.2

For $k = 1$, each X_i is an $n_i \times 2$ matrix, where the first column is generated by $\mathcal{U}_{n_i}$ and the second column is generated by $\mu_{n_i,2}$, i.e., the second column has exactly one 1. Now let us consider the matrix D_i, the disjointness matrix for X_i, the rows of which are indexed by the choices of x_i and columns are indexed by choices of W_i of X_i. Clearly, for $k = 2$, D_i is a $2^{n_i} \times n_i$ matrix. The $\langle m, n \rangle$th entry of D_i is $+1$ if X_i, composed of x_m and W_n has a all 1's row, -1 otherwise.

For $\prod_{i=1}^{m} \mathsf{DISJ}_{n_i,k+1}(X_i)$, the disjointness matrix D can be thought of a $2^{\sum n_i} \times \prod n_i$ matrix, which is nothing but tensor product of D_i's for all $i \in [m]$.

$$D = \bigotimes_{i=1}^{m} D_i.$$

We are trying to compute the dot product of D with a cylinder (which is a rectangle for $k = 2$). More formally, given a rectangle $\mathcal{X} = S \times T$, we want to compute $|1_S^t D 1_T|$ normalized by number of possibilities of X_i's for all $i \in [m]$. Hence we get,

$$\begin{aligned}
\Gamma_1(n_1, ..., n_m) &\leq \frac{\sqrt{2^{n_1+...+n_m}}.\sqrt{n_1...n_m}}{2^{n_1}.n_1...2^{n_m}.n_m}.\sigma_{max}(D) \\
&\leq \frac{\sqrt{2^{n_1+...+n_m}}.\sqrt{n_1...n_m}}{2^{n_1}.n_1...2^{n_m}.n_m}.||D|| \\
&\leq \frac{\sqrt{2^{n_1+...+n_m}}.\sqrt{n_1...n_m}}{2^{n_1}.n_1...2^{n_m}.n_m}.\sqrt{2^{n_1}}...\sqrt{2^{n_m}} \\
&= \frac{1}{\sqrt{\prod_{i=1}^{m} n_i}}
\end{aligned}$$

We used the fact that the largest singular value of matrix is bounded by the Frobenius norm of the matrix, and $\langle D, D \rangle$ is a diagonal matrix, ith entry of the diagonal being 2^{n_i} for all $i \in [m]$.

28.1.2 Induction Step

Fix $k \geq 2$ and consider the cylinder intersection $\mathcal{X}(X_1, ..., X_m)$ which can be written as $\mathcal{X}((x^1, Y^1, u^1), ..., (x^m, Y^m, u^m))$, where u^i is the last column of W^i which is a $n_i \times (k+1)$ matrix. Among the cylinder which take part in this intersection, some cylinders involve the last column ,i.e., $u^1, ..., u^m$, and others don't. Consider the cylinder which involves the last column. For a fixing of $u^1, ..., u^m$, that cylinder depends on only $(x^1, Y^1), ..., (x^m, Y^m)$. So can represent this cylinder intersection in the following way.

$$\mathcal{X}((x^1,Y^1,u^1),...,(x^m,Y^m,u^m)) = \mathcal{X}_{u^1,...,u^m}((x^1,Y^1),...,(x^m,Y^m)).\xi((x^1,Y^1),...,(x^m,Y^m)).$$

where ξ is some function into $\{0,1\}$. So we have,

$$\begin{aligned}
\Gamma_k(n_1,...,n_m) &= \left| E_{(x^1,Y^1),...,(x^m,Y^m)} E_{u^1,...,u^m} \left[\mathcal{X}.\prod_{i=1}^{m} D(x^i,Y^i,u^i) \right] \right| \\
&\leq E_{(x^1,Y^1),...,(x^m,Y^m)} \left| E_{u^1,...,u^m} \left[\mathcal{X}.\prod_{i=1}^{m} D(x^i,Y^i,u^i) \right] \right| \quad \text{[Jensen's inequality]} \\
&= E_{(x^1,Y^1),...,(x^m,Y^m)} \left| E_{u^1,...,u^m} \left[\mathcal{X}_{u^1,...,u^m}.\xi.\prod_{i=1}^{m} D(x^i,Y^i,u^i) \right] \right| \\
&\leq E_{(x^1,Y^1),...,(x^m,Y^m)} \left| E_{u^1,...,u^m} \left[\mathcal{X}_{u^1,...,u^m}.\prod_{i=1}^{m} D(x^i,Y^i,u^i) \right] \right| \quad \text{[as } \xi \text{ is into } \{0,1\}\text{]} \\
&\leq E_{(x^1,Y^1),...,(x^m,Y^m)} \left[E_{u^1,...,u^m} \left[\mathcal{X}_{u^1,...,u^m}.\prod_{i=1}^{m} D(x^i,Y^i,u^i) \right]^2 \right]^{1/2} \quad \text{[Cauchy Schwarz]} \\
&= E_{(x^1,Y^1),...,(x^m,Y^m)} \left[E_{u,v} \left[\mathcal{X}_u.\mathcal{X}_v.\prod_{i=1}^{m} D(x^i,Y^i,u^i) D(x^i,Y^i,v^i) \right] \right]^{1/2} \quad \text{[Expanding square term]}
\end{aligned}$$

We have used the shorthand u and v to represent the vectors $u^1,...,u^m$ and $v^1,...,v^m$ respectively. Note that the above technique is borrowed from Babai, Nisan and Szegedy.

For analyzing the previous inequality we need the following lemma. We define $\lambda_{n,k}$ to be the probability distribution on $\{0,1\}^{n\times(k-1)} \times \{0,1\}^n \times \{0,1\}^n$ in which one first chooses Y according to the marginal distribution $\mu_{n,k}(Y) = \sum_u \mu_{n,k}(Y,u)$ and then, given Y, chooses u and v independently according to the conditional distribution $\mu_{n,k}(u|Y)$ defined as follows.

$$\mu_{n,k}(u|Y) = \frac{\mu_{n,k}(Y,u)}{\mu_{n,k}(Y)}.$$

More formally,

$$\lambda_{n,k}(Y,u,v) = \mu_{n,k}(Y)\mu_{n,k}(u|Y)\mu_{n,k}(v|Y).$$

Lemma 28.3. *For each (Y,u,v) in the support of $\lambda_{n,k}$ and each $x \in \{0,1\}^n$,*

$$D(x,Y,u)D(x,Y,v) = \begin{cases} D(x,Y,u \wedge \overline{v})D(x,Y,\overline{u} \wedge v) & \textit{if } D(Y,u,v) = -1 \\ 1 & \textit{otherwise} \end{cases}$$

Proof. As the definition of $\lambda_{n,k}$ suggests, the matrix (Y,u) has only one row, call it i, composed of all 1's. Similarly, we assume that the jth row of the matrix (Y,v) composed of all 1's. If $i=j$, no matter what the value of x_i is, the left hand side will always evaluate to 1.

Now there can be four patterns arrising in different coordinates of u and v, namely $u \wedge v$, $u \wedge \overline{v}$, $\overline{u} \wedge v$ and $\overline{u} \wedge \overline{v}$, based on whether u and v have a 0 or 1 in a specific coordinate. If the all 1's row in Y appear in $u \wedge v$, the left hand side reduces to 1 no matter what the value of the first column is in that coordinate and that is the case we have analyzed. We can rule out the irrelevant part $\overline{u} \wedge \overline{v}$ as the all 1's row of (x, Y, u) occurs either in $u \wedge v$ or in $u \wedge \overline{v}$ and similarly, all 1's row of (x, Y, v) occurs wither in $u \wedge v$ or in $\overline{u} \wedge v$. If the left hand side evaluates to -1, then there can be two possibilities, - either $D(x, Y, u) = 1$ and $D(x, Y, v) = -1$, which happens if the all 1's row is in $u \wedge \overline{v}$, or $D(x, Y, u) = -1$ and $D(x, Y, v) = 1$, which happens if the all 1's row is in $\overline{u} \wedge v$. Hence proved. □

Using this lemma and conditioning on u^i, v^i, $Y^i|_{u^i \wedge v^i}$ and $Y^i|_{\overline{u^i} \wedge \overline{v^i}}$, we get the following expression.

$$E_{(x^1,Y^1),\ldots,(x^m,Y^m)}\left[E_{u,v}\left[\mathcal{X}_u.\mathcal{X}_v.\prod_{i=1}^{m} D(x^i,Y^i,u^i)D(x^i,Y^i,v^i)\right]\right]$$

$$= \sum_{z\in\{-1,+1\}^m} E_{(x^1,Y^1),\ldots,(x^m,Y^m)}\left[E_{u,v}\left[\mathcal{X}_u.\mathcal{X}_v.\prod_{i:z_i=-1} D((x^i,Y^i)|_{u^i\wedge\overline{v^i}})D((x^i,Y^i)|_{\overline{u^i}\wedge v^i})\right]\right] \times \prod_{i=1}^{m}\mathbb{I}[D(Y^i|_{u^i\wedge v^i}) = z_i]$$

For bounding the tems of the previous expression, we need the following lemma.

Lemma 28.4. *Let $(Y, u, v) \sim \lambda_{n,k}$. Conditioned on fixed values of u, v, $Y|_{u\wedge v}$ and $Y|_{\overline{u}\wedge\overline{v}}$ with $D(Y|_{u\wedge v}) = -1$, the remaining parts $Y|_{u\wedge\overline{v}}$ and $Y|_{\overline{u}\wedge v}$ are independent and distributed according to $\mu_{|u\wedge\overline{v}|,k-1}$ and $\mu_{|\overline{u}\wedge v|,k-1}$ respectively.*

Consider $\mathcal{X}' = \mathcal{X}_u.\mathcal{X}_v$. It is a cylinder intersection for fixed u and v and hence is a k dimensional cylinder intersection. So we can apply the induction hypothesis to it, thereby bounding the right hand side of the previous expression in absolute value by

$$\sum_{z\in\{-1,+1\}^m}\prod_{i:z_i=-1}\frac{(2^{k-1}-1)^2.\mathbb{I}[D(Y^i|_{u^i\wedge v^i}) = -1]}{\sqrt{|u^i\wedge\overline{v^i}||\overline{u^i}\wedge v^i|}}.\prod_{i:z_i=1}\mathbb{I}[D(Y^i|u^i\wedge v^i) = 1].$$

Passing to expectation,

$$\Gamma_k(n_1,\ldots,n_m)^2 \le \sum_{z\in\{-1,+1\}^m}\prod_{i:z_i=-1}E_{\lambda_{n_i,k}}\left[\frac{(2^{k-1}-1)^2.\mathbb{I}[D(Y^i|_{u^i\wedge v^i}) = -1]}{\sqrt{|u^i\wedge\overline{v^i}||\overline{u^i}\wedge v^i|}}\right] \times \prod_{i:z_i=1}\Pr_{\lambda_{n_i,k}}[D(Y^i|u^i\wedge v^i) = 1]$$

$$= \prod_{i=1}^{m}\left((2^{k-1}-1)^2.E_{\lambda_{n_i,k}}\left[\frac{\mathbb{I}[D(Y^i|_{u^i\wedge v^i}) = -1]}{\sqrt{|u^i\wedge\overline{v^i}||\overline{u^i}\wedge v^i|}}\right] + \Pr_{\lambda_{n_i,k}}[D(Y^i|u^i\wedge v^i) = 1]\right)$$

The probability and the expectation terms of the final expression are evaluated in the following lemmas.

Lemma 28.5. *For* $\lambda_{n,k}$*,*

$$E_{\lambda_{n,k}}\left[\frac{\mathbb{I}[D(Y|_{u\wedge v}) = -1]}{\sqrt{|u \wedge \overline{v}||\overline{u} \wedge v|}}\right] \leq \frac{4}{n}.\frac{2^k - 1}{2^k - 2}$$

Lemma 28.6. *For* $\lambda_{n,k}$*,*

$$\Pr_{\lambda_{n,k}}[D(Y|u \wedge v) = 1] \leq \frac{2^k - 1}{n}$$

Using the previous lemma, it is easy to see that $\Gamma_k(n_1, ..., n_m)^2$ is bounded from above by $\frac{(2^k-1)^m}{\sqrt{\prod_{i=1}^m n_i}}$, which completes the proof.