CS 671: Graph Streaming Algorithms and Lower Bounds **Rutgers: Fall 2020**

Problem set 7

Due: 11:59PM, October 27, 2020

Problem 1. In this question, we prove a lower bound for the Index problem over a different distribution than what we used before.

1. Alice has a string $x \in \{0,1\}^n$ such that each $x_i = 1$ independently and with probability $p < 1/2$.
2. Bob has an index $i \in [n]$ chosen uniformly at random.

In this distribution, even without any communication, Bob can output the answer correctly with error probability p (by always outputting 0). We are now going to prove that if however Bob wants to output the correct answer with probability of error $o(p)$, then $\Omega(np \cdot \log(1/p))$ communication is necessary.

1. Let A be a binary random variable, B be an arbitrary random variable and suppose that there is a function $g : supp(B) \rightarrow supp(A)$ such that $\Pr(A \neq g(B)) = \delta$. Then, prove that

$$\mathbb{H}(A \mid B) \leq H_2(\delta) := \delta \cdot \log \frac{1}{\delta} + (1-\delta) \cdot \log \frac{1}{1-\delta};$$

   (this inequality is known as Fano's inequality).

   Hint: Define a random variable $\Theta \in \{0,1\}$ such that $\Theta = 1$ iff $g(B) \neq A$. Prove that

$$\mathbb{H}(A \mid B) \leq \mathbb{H}(A, \Theta \mid B) \leq \mathbb{H}(\Theta) = H_2(\delta).$$

   Remember to justify every inequality/equality you use by citing the lecture notes or proving it directly.

2. Let π be a protocol for Index over this distribution with probability of error $\delta = o(p)$. Use part 1 to prove that $\mathbb{I}(X_I; M \mid I) \geq H_2(p) - H_2(\delta)$ where X, M, I are the random variables for the string x, message m, and index i, respectively.
3. Use part 2 to prove that π needs to communicate a message of length $n \cdot (H_2(p) - H_2(\delta))$ and further show that this is $\Omega(np \cdot \log(1/p))$ as desired.
4. Let us also prove the tightness of this bound: show that there is a protocol for solving Index over this distribution with probability of error $o(1)$ and communication $O(np \cdot \log(1/p) + \log n)$ bits.