15-859 : Homework 3 Solutions

1 Question 1

See Section 3.1 of `http://www.cs.cmu.edu/afs/cs/user/dwoodruf/www/wNow3.pdf`

2 Question 2

2.1 Part 1

S be a Cauchy matrix used in estimating ℓ_1-norm of a stream. As Alice and Bob both have shared randomness, they can both generate this matrix S. Alice computes Sx, Sa and sends both Sx, Sa to Bob. This requires a communication of $O(\log(n)/\varepsilon^2)$ bits. Bob then computes $Sx + Sy$ and $Sa + Sb$ and computes $Z = \text{median}(|Sx + Sy|) + \text{median}(|Sa + Sb|)$. We have from class that with probability $9/10$, the following hold simultaneously,

$$(1-\varepsilon)\|x+y\|_1 \leq \text{median}(|S(x+y)|) \leq (1+\varepsilon)\|x+y\|_1$$
$$(1-\varepsilon)\|a+b\|_1 \leq \text{median}(|S(a+b)|) \leq (1+\varepsilon)\|a+b\|_1.$$

Therefore Z computed by Bob satisfies, with probability $\geq 9/10$, that

$$(1-\varepsilon)(\|x+y\|_1 + \|a+b\|_1) \leq Z \leq (1+\varepsilon)(\|x+y\|_1 + \|a+b\|_1). \tag{1}$$

An issue with Cauchy random variables is that they take real values and thus cannot be communicated efficiently. Note that ε is assumed to be $\Omega(1/\text{poly}(n))$. Now rounding the entries of the vectors Sa, Sx to nearest multiple of $1/\text{poly}(n)$ will modify the value of Z computed by at most $1/\text{poly}(n)$ and therefore for all $(x, y), (a, b)$ such that $\|x + y\|_1 + \|a + b\|_1 \geq 1$, rounding of the vectors Sx, Sa still means that the value Z computed satisfies (1). In the case when $x + y = 0$ and $a + b = 0$, rounding of the vectors might lead to Bob computing a different value but as rounding changes the norm by at most $1/\text{poly}(n)$, Z computed is such that $Z \leq 1/\text{poly}(n)$. Then Bob can infer that $\|x + y\|_1 + \|a + b\|_1 = 0$. Note that $\|x + y\|_1 + \|a + b\|_1$ can only take integer values and thus we covered all the cases.

2.2 Part 2

We prove a lowerbound by reducing from the index problem. Alice gets a set $S \subseteq [n]$ and Bob gets an index $i \in [n]$ and wants to determine if $i \in S$ or not. Alices sets the vectors $x = a$ to be the vectors that corresponds to the set S. Bob sets y to be a vector such that $y_i = +1$ and y_j and b to be a vectors such that $b_i = -1$ and $b_j = 0$ for all $j \neq i$. Then it is easy to see that

$$\|x+y\|_1 - \|a+b\|_1 = \begin{cases} 0 & \text{if } i \notin S \\ 2 & \text{if } i \notin S. \end{cases}$$

Thus, by checking the value of Z computed with $\varepsilon = 1/2$, Bob can check if $i \in S$ and will be correct with probability $\geq 9/10$. But as index problem has a lower-bound of $\Omega(n)$ one-way communication complexity, we obtain that this problem also has a $\Omega(n)$ lowerbound.

3 Part 3

Let $\tilde{Z}$ be such that $(1-\varepsilon/2)\|a+b\|_2^2 \leq \tilde{Z} \leq (1+\varepsilon/2)\|a+b\|_2^2$. Suppose $\|a+b\|_2^2 \geq 2$. We have $\tilde{Z} \geq 2-\varepsilon$. In this case $\tilde{Z}-1$ satisfies

$$(1-\varepsilon/2)\|a+b\|_2^2 - 1 \leq Z - 1 \leq (1+\varepsilon/2)\|a+b\|_2^2 - 1.$$

We have

$$(1+\varepsilon)(\|a+b\|_2^2 - 1) - ((1+\varepsilon/2)\|a+b\|_2^2 - 1) = (\varepsilon/2)\|a+b\|_2^2 - \varepsilon \geq 0.$$

Similarly,

$$(1-\varepsilon/2)\|a+b\|_2^2 - 1 - (1-\varepsilon)(\|a+b\|_2^2 - 1) = \varepsilon/2\|a+b\|_2^2 - \varepsilon \geq 0.$$

Therefore we have

$$(1-\varepsilon)(\|a+b\|_2^2 - 1) \leq \tilde{Z} - 1 \leq (1+\varepsilon)(\|a+b\|_2^2 - 1)$$

if $\|a+b\|_2^2 \geq 2$. If $\|a+b\|_2^2 < 2$, there are only two possibilities as a, b have integer coordinates.

- Case 1: $\|a+b\|_2^2 = 1$. We have $(1-\varepsilon/2) \leq \tilde{Z} \leq (1+\varepsilon/2) < 2-\varepsilon$. So if $\tilde{Z}$ computed by Bob is $< 2-\varepsilon$ and nonzero, then we can conclude that $\|a+b\|_2^2 = 1$ and $|\|a+b\|_2^2 - 1| = 0$.
- Case 2: $\|a+b\|_2^2 = 0$. Then $\tilde{Z}$ computed is also equal to 0 and we can conclude that $|\|a+b\|_2^2 - 1| = |-1| = 1$.

Therefore based on value of $\tilde{Z}$, Bob can compute $1 \pm \varepsilon$ approximation to $|\|a+b\|_2^2 - 1|$ and such a $\tilde{Z}$ can be computed using $O(\log(n)/\varepsilon^2)$ bits of communication as seen in class. Bob can also compute $1 \pm \varepsilon$ approximation to $\|x+y\|_1$ by sketching with Cauchy matrix and a communication of $O(\log(n)/\varepsilon^2)$ bits. Now multiplying both values gives a $1 \pm O(\varepsilon)$ approximation and appropriately scaling ε lets Bob compute a Z such that

$$(1-\varepsilon)(\|x+y\|_1 \cdot |\|a+b\|_2^2 - 1|) \leq Z \leq (1+\varepsilon)(\|x+y\|_1 \cdot |\|a+b\|_2^2 - 1|)$$

using a protocol that has communication complexity $O(\log(n)/\varepsilon^2)$. Rounding issues can be similarly solved as in Part 1.

4 Part 4

We again show a lowerbound by reducing from the index problem. Alice obtains a set $S \subseteq [n]$ and Bob obtains an index $i \in [n]$ and Bob wants to determine if $i \in S$ or not. As we've seen, this problem has $\Omega(n)$ lower bound on the one-way communication complexity.

Alice sets the vector a to correspond to the set S i.e., $a_j = 1$ if $j \in S$ and 0 otherwise and sets $x = 0$. Bob sets the vector $b = 0$ and y to be the vectors such that $y_i = 1$ and $y_j = 0$ for all $j \neq i$. We now have $\|(D_a - D_b) \cdot (x - y)\|_1 = 1$ if $i \in S$ and 0 otherwise which can be distinguished using a value of $\varepsilon = 1/2$. Lowerbound on Index problem implies a $\Omega(n)$ lower bound on this problem.