Probability and Computation: Problem Sheet 6

Question 1. *Let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ be any vector. Now consider a random vector $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, where each y_i, $1 \leq i \leq n$ is in $\{-1, +1\}$ uniformly at random.*

1. *Compute $\mathbf{E}\left[\, x^T y \,\right]$ and $\mathbf{E}\left[\, (x^T y)^2 \,\right]$.*
2. *We are now trying to generalise these results. Which properties of the distribution the y_i's are sampled from are needed to recover the results from Part 1?*
3. *What is the challenge in proving concentration for $(x^T y)^2$, for example, why cannot we apply a Chernoff Bound?*

Question 2. *(easy) What is the time complexity for applying the Random Projection Method to a set of P input vectors?*

Question 3. *Demonstrate that it is essential for the Random Projection Method to choose the linear function f randomly. Specifically, prove that for any linear function $f : \mathbb{R}^N \to \mathbb{R}^M$ there are two vectors $x_1, x_2 \in \mathbb{R}^N$, $x_1 \neq x_2$, such that $f(x_1) = f(x_2)$.*
Hint: *Your proof should at some point exploit that $N > M$.*

Question 4. *Let $A, B \in \{0,1\}^{n \times n}$ be two binary matrices, and let $C := A \cdot B$. Design a randomised algorithm of running time $O(n^2 \cdot \log n)$ that computes a matrix $\widetilde{C}$ such that with probability at least $1 - n^{-1}$ the following holds for all $1 \leq i, j \leq n$:*

$$\left\lfloor \frac{C_{i,j}}{4 \log n} \right\rfloor \leq \widetilde{C}_{i,j} \leq C_{i,j} + \max\{C_{i,j}, 4 \log n\}.$$

Question 5. *Consider the deterministic weighted majority algorithm. [LECTURE 15]*

1. *Prove or disprove that the algorithm always makes a number of mistakes M^T within T rounds that satisfies: $M^T \geq \min_{i \in [n]} m_i^T$.*
2. *Show that for any deterministic algorithm, there exists an input such that $M^T \geq 2 \cdot \min_{i \in [n]} m_i^T$.*