CS6845 - Pseudorandomness
Assignment 3

Due date: March 13, in class

Instructions

- For each day's delay in submission, you will lose 25% marks.
- Discussing the questions on the assignment with your classmates is allowed, but **no** collaboration is allowed while writing up the solution. If you discuss an assignment problem with someone, please acknowledge that while writing the solution. This will **not** affect your grade.
- Checking online sources for answers is **strongly discouraged**. If you need any clarification, you are welcome to discuss with me or Dinesh. Any academic dishonesty will result in zero marks for the assignment.
- Include all the steps/arguments in your answer/proof.

---

1. (2 points) In the last problem set we saw that when $\mathbb{F}$ is a finite field we can construct a graph G on $|\mathbb{F}|^2$ vertices with degree $|\mathbb{F}|$ such that $\lambda(G) \leq 1/\sqrt{|\mathbb{F}|}$. Show that for a sufficiently large (yet still constant-sized) field $\mathbb{F}$, we can construct a $(D^8, D, 1/8)$-spectral expander from G using squaring, tensoring and the zig-zag product, where D is a constant.

2. For two probability distribution X and Y defined over a universe $\mathcal{U}$, the statistical difference $\Delta(X,Y) = \max_{T \subseteq \mathcal{U}} |\Pr[X \in T] - \Pr[Y \in T]|$.
   (a) (2 points) Show that $\Delta(X,Y) = \frac{1}{2}\sum_{u \in \mathcal{U}} |\Pr[X = u] - \Pr[Y = u]|$.
   (b) (2 points) Show that for every function f, $\Delta(f(X), f(Y)) \leq \Delta(X,Y)$.
   (c) (2 points) Show that $\Delta(X,Y) = \max_f |\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]|$.

3. A random variable $X = (X_1, X_2, \ldots, X_t)$ is a $(k_1, k_2, \ldots, k_t)$ *block source* if for every $x_1, x_2, \ldots, x_{i-1}$, the random variable $X_i |_{X_1 = x_1, \ldots, X_{i-1} = x_{i-1}}$ is a k_i-source.
   (a) (2 points) Show that if $X = (X_1, X_2, \ldots, X_t)$ is a $(k_1, k_2, \ldots, k_t)$ block source, then X is also a $(k_1 + k_2 + \ldots + k_t)$-source.
   (b) Suppose that X is an $(n - \Delta)$-source taking values in $\{0,1\}^n$, and we let X_1 consist of the first n_1 bits of X and X_2 the remaining $n_2 = n - n_1$ bits.
      i. (2 points) Show that X_1 is an $(n_1 - \Delta)$-source, and X_2 is an $(n_2 - \Delta)$-source.
      ii. (2 points) Show that for any $\epsilon > 0$, with probability at least $1 - \epsilon$ over the choice of x_1, the conditional distribution $X_2 |_{X_1 = x_1}$ is an $(n_2 - \Delta - \log(1/\epsilon))$-source.

4. (4 points) Let $A(w; r)$ be a randomized algorithm for computing a function f using m random bits such that $A(w; \mathcal{U}_m)$ has error probability at most 1/3 (the algorithm A has two-sided error). Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a $(k, 1/7)$-extractor. Define $A'(w; x) = \mathrm{maj}_{y \in \{0,1\}^d} A(w; \mathrm{Ext}(x,y))$ (breaking ties arbitrarily). Show that for every $(k+t)$-source X, $A'(w; X)$ has error probability at most 2^{-t}.

5. (4 points) A family $\mathcal{H}$ of functions mapping $[N]$ to $[M]$ is said to have collision probability at most δ if for every $x_1 \neq x_2 \in [N]$, we have $\Pr_{h \in_r \mathcal{H}}[h(x_1) = h(x_2)] \leq \delta$. The family $\mathcal{H}$ is ϵ-almost universal if it has collision probability at most $(1+\epsilon)/M$. Show that if $\mathcal{H} = \{h : [N] \to [M]\}$ is ϵ^2-almost universal, then $\mathrm{Ext}(x, h) = (h, h(x))$ is a (k, ϵ)-extractor for $k = m + 2\log(1/\epsilon) + O(1)$, where $m = \log M$.

6. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code and let $\mathcal{C}^\perp$ be the null space of $\mathcal{C}$.
   (a) (1 point) Show that if $\mathcal{C} = \mathcal{C}^\perp$, then $\mathcal{C}$ has dimension $n/2$.
   (b) (1 point) Show that the code $\mathcal{C} = \{(\mathbf{x}, \mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^k\}$ has the property that $\mathcal{C} = \mathcal{C}^\perp$.
   (c) (2 points) Show that for $\mathcal{C} = (\mathcal{C}^\perp)^\perp$.