CSS.413.1 Pseudorandomness 2 Dec, 2023

Problem Set 3

- Due Date: **10 December 2023 (hard-deadline!)**
- As the deadline is in just 8 days, this set only has two problems. The points for each problem is indicated on the side. The total for this set is **35** points.
- Feel free to use a subdivision of a problem for later subdivisions, even if you haven't solved it.
- Turn in your problem sets electronically (PDF; either LaTeXed or scanned etc.) on Piazza.
- Collaboration is encouraged, but all writeups must be done individually and must include names of all collaborators.
- Referring to sources other than the text book and class notes is strongly discouraged. But if you do use an external source (eg., other text books, lecture notes, or any material available online), ACKNOWLEDGE all your sources (including collaborators) in your writeup. This will not affect your grades. However, not acknowledging will be treated as a serious case of academic dishonesty.
- Be clear in your writing.

1. **[Concatenated codes (Problem 5.2 in Vadhan's notes)]** (3 + 7 + 5)

   We discussed the notion of *concatenated codes* in class, which were defined as follows: Suppose we have two codes $\mathrm{Enc}_1 : [M] \to \Sigma_1^{n_1}$, and another code $\mathrm{Enc}_2 : \Sigma_1 \to \Sigma_2^{n_2}$, then the concatenated code $\mathrm{Enc}_1 \circ \mathrm{Enc}_2 : [M] \to \Sigma_2^{n_1 n_2}$, obtained by applying Enc_1 and then encoding each of the n_1 symbols using Enc_2.

   (a) Show that if Enc_1 has minimum distance δ_1 and Enc_2 has minimum distance δ_2, then $\mathrm{Enc}_1 \circ \mathrm{Enc}_2$ has minimum distance at least $\delta_1 \delta_2$.

   (b) If Enc_1 is $(1-\varepsilon, \ell_1)$-list-decodable, and Enc_2 is (δ, ℓ_2)-list-decodable, then $\mathrm{Enc}_1 \circ \mathrm{Enc}_2$ is $((1-\varepsilon_1\ell_2)\delta, \ell_1\ell_2)$.

   (c) Construct an explicit code $\mathrm{Enc} : \{0,1\}^m \to \{0,1\}^n$ with $n = O({}^m/_{\varepsilon^2})$ that is $({}^1/_2 - \varepsilon, \mathrm{poly}({}^1/_\varepsilon))$ in $\mathrm{poly}(n)$ time.

2. **[Twenty questions (built off Problem 5.7 in Vadhan's notes)]** (3 + 5 + 5 + 2 + 5)

   Consider a game you are playing with a friend, and the friend has two unknown m-bit strings $s_1, s_2 \in \{0,1\}^m$. You are allowed to provide any function $f : \{0,1\}^m \to \{0,1\}$ and your friend will either return $f(s_1)$ or $f(s_2)$ (but not tell you which s_i was used).

   (a) Show that, no matter how many questions you ask, your friend can always play the game in a way that it is impossible for you to output a string s that is guaranteed to be one of your friend's strings. That is, your friend can play in a way that, no matter what string s you claim, they can always exhibit two other strings s_1, s_2 that are consistent with the answers provided.

   [Hint: Your friend can play with three secrets.]

Base version: (2023-12-02 06:04:34 +0530) , 74e3cf9

(b) Let Enc : $\{0,1\}^m \to \{0,1\}^n$ be a code with the property that for any four distinct codewords a, b, x, y, there is always an index $i \in [n]$ such that $a_i = b_i \neq x_i = y_i$. Then, in the above 20 questions game, show that you can ask your friend n questions and output a list of two strings $\{s'_1, s'_2\}$ such that one of your friend's strings is guaranteed to be in this list.

(c) Suppose Enc : $\mathbb{F}_2^m \to \mathbb{F}_2^n$ was infact a linear code and the above property *does not* hold. That is, there *do* exist four distinct codewords a, b, x, y such that there is no $i \in [n]$ with $a_i = b_i \neq x_i = y_i$. Without loss of generality (due to linearity), let us assume that a is the zero codeword. Then, show that

$$\Delta(0,x) + \Delta(0,y) + \Delta(b,x) + \Delta(b,y) = \mathrm{wt}(b) + \mathrm{wt}(x+y) + \mathrm{wt}(x+y+b)$$

where $\Delta(\cdot,\cdot)$ refers to the Hamming distance, and $\mathrm{wt}(\cdot)$ refers to the weight or the number of nonzero coordinates, and all the additions are over $\mathbb{F}_2$.

[Hint: Fix a coordinate i and do a case analysis for all configurations (other than $0 = b_i \neq x_i = y_i$) to see what that coordinate contributes to the LHS and the RHS.]

(d) Suppose Enc : $\mathbb{F}_2^m \to \mathbb{F}_2^n$ is a linear code that is "ε-biased", i.e. it has the property that for every nonzero codeword x, we have

$$n \cdot (1/2 - \varepsilon) \leq \mathrm{wt}(x) \leq n \cdot (1/2 + \varepsilon)\,.$$

If $\varepsilon < 1/14$, show that any linear ε-biased code must have the property described in item 2b.

(e) Based on what we have done in class, construct an "ε-biased code" with $n = O(m^2/\varepsilon^2)$.