Error Correcting Codes: Combinatorics, Algorithms and Applications Spring 2009

HOMEWORK 4
Due Friday, April 2, 2009 in class

You can collaborate in groups of up to 3. However, the write-ups must be done individually, that is, your group might have arrived at the solution of a problem together but everyone in the group has to write up the solution in their own words. Further, you **must** state at the beginning of your homework solution the names of your collaborators. Just to be sure that there is no confusion, the group that you pick has to be for all problems [i.e. you cannot pick different groups for different problems :-)]

If you are not typesetting your homework, please make sure that your handwriting is legible. Illegible handwriting will most probably lose you points.

Unless stated otherwise, for all homeworks, you are **only** allowed to use notes from the course: this includes any notes that you might have taken in class or any scribed notes from Fall 07 or Spring 09 version or the current version of the course. Doing otherwise will be considered cheating. Note that if your collaborator cheats and you use his/her solution, then you have cheated too (ignorance is not a valid excuse).

Please use the comment section of the post on HW 4 on the blog if you have any questions and/or you need any clarification.

In total you can use at most **two** pages for this homework.

I encourage you to start thinking on the problems **early**.

---

1. (**Rate of linear list-decodable codes**) $(4 + 6 = 10$ points$)$

   (a) Let $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{F}_q^a$ be *linearly-independent* vectors. Then prove that for a random $a \times b$ matrix $M \in \mathbb{F}_q^{a \times b}$, the vectors $\mathbf{v}_i \cdot M$ $(1 \leq i \leq \ell \leq a \leq b)$ are random *independent* vectors in $\mathbb{F}_q^b$.
   (*Hint:* First prove the above for the special case when $\mathbf{v}_i = \mathbf{e}_i$, where $\mathbf{e}_i \in \mathbb{F}_q^a$ is the vector which is 0 in all locations except i where it is 1. Then "reduce" the general linearly independent case to this special case.)

   (b) For $0 < p < 1$ and a positive integer L, call a code $C \subset \Sigma^n$ to be (p, L)-list decodable if every Hamming ball of radius pn (in the space Σ^n) has at most L codewords of C. Prove that for every finite field $\mathbb{F}_q$, $0 < p < 1 - 1/q$, integer $L \geq 1$, and large enough n, there is a (p, L)-list decodable linear code $C \subseteq \mathbb{F}_q^n$ that has rate at least $1 - H_q(p) - \frac{1}{\log_q(L+1)} - o(1)$.
   (*Hint*: Apply the usual random coding method of picking a generator matrix at random. In estimating the probability that L nonzero messages all get mapped into a ball of radius pn, these L events are not all independent (and this is the difference compared to picking a general random code). But at least how many of these events are independent of one

another? Part (a) will be useful in answering this question.)
(*Note:* As usual for part (b), you can assume part (a) as given.)

2. (**Intractability of Maximum Likelihood Decoding**) I have mentioned a few times in class that MLD is a notoriously hard to implement any faster than exponential time. In this problem we will show that doing MLD for linear codes in general is NP-hard.
(*This problem is for your cognitive pleasure only; no need to turn this problem in*)

   Given an undirected graph $G = (V, E)$, consider the binary code $C_G \subseteq \{0,1\}^{|E|}$, where every codeword in C_G corresponds to a cut in G. More precisely, every position in any vector in $\{0,1\}^{|E|}$ is associated with an edge in E. Let $\mathbf{c} \in C_G$ be a codeword. Let $E_{\mathbf{c}} = \{i \in E | c_i = 1\}$. Then $E_{\mathbf{c}}$ must correspond to exactly the edges in some cut of G.

   (a) Prove that C_G is a linear code.

   (b) Prove that if one can do MLD on G in polynomial time then one can solve the Max-Cut problem[1] on G in polynomial time. Conclude that solving the MLD problem on linear codes in general is NP-hard.
   (*Hint*: Try to think of a vector $\mathbf{y} \in \{0,1\}^{|E|}$ such that solving MLD with received word $\mathbf{y}$ for C_G is equivalent to solving the Max-Cut problem on G.)

[1]Given a graph $G = (V, E)$, a cut is a partition of the vertices into sets $S \subseteq V$ and $\overline{S} = V \setminus S$. The size of the cut is the number of edges that have exactly one end-point in S and the other in $\overline{S}$. The Max-Cut of G is a cut with the maximum possible size. Max-Cut is a well known NP-hard problem.