CSC 2414H - Assignment 3

Due Apr 1 , 2008

General rules : In solving this you may consult books and you may also consult with each other, but you must each write your own solution. In each problem list the people you consulted. This list will not affect your grade.

1. Let $G = (L, R, E)$ be a d-regular, bipartite graph on the sets L and R, where $|L| = |R| = n$. Let $N = |E|(= dn)$. For $A \subset L, B \subset R$ we let $e(A, B)$ denote the number of edges between A and B. It can be shown that certain expander graphs satisfy that for all $A \subset L, B \subset R$

$$\left| e(A,B) - \frac{d|A||B|}{n} \right| \leq \Lambda d \sqrt{|A| \cdot |B|}. \tag{1}$$

Here $0 < \Lambda \leq 1$ is some positive constant. Notice that inequality (1) says that the two sets have roughly as many edges connecting them as would be expected by random graph with the same density.

Now, let C_0 be a $[d, d \cdot r_0, d \cdot \delta_0]_2$ linear code to which we have a decoding algorithm that corrects up to $d \cdot \delta_0/2$ errors. We now present a new code C, based on G and on C_0.

Assume that the vertices in L and R are ordered in some arbitrary way. An assignment of bits to the edges of G, $\{c_e\}_{e \in E}$, where $c_e \in \{0, 1\}$ for every $e \in E$, is a received corrupted codeword of a code that will be shortly defined. For a vertex v of G, let E_v be the set of edges adjacent to v. Now, define $c|_v = (c_e)_{e \in E_v}$. We can view $c|_v$ as a vector in $\{0, 1\}^d$ which is the restriction of c to E_v. The orderings of the vertices of L and R induce the order of the coordinates of $c|_v$. Now, define

$$C = \{c \in \{0,1\}^N \ : \ c|_v \in C_0 \ \forall v \in V(G)\}.$$

(a) Show that C is a linear code and bound its rate from below.

We shall assume from here on that $\Lambda \leq \delta_0/3$. Our goal is to show an efficient decoding algorithm for C that corrects upto $\frac{\delta_0^2}{8}(1 - 2\Lambda/\delta_0)N$ errors. Consider the following algorithm.

Decoding Algorithm:

i. Given a word $w \in \{0, 1\}^N$ correct w to $w^{(1)}$ so that $w^{(1)}|_v \in C_0$ for all $v \in L$. This is done by running the decoding algorithm for C_0 on the words $w|_v$ for $v \in L$.

ii. Repeat the same process on the R vertices starting with the word $w^{(1)}$. The word produced $w^{(2)}$ satisfies $w^{(2)}|_v \in C_0$ for all $v \in R$.

iii. Go back to step i. until you get $w^{(j)} \in C$.

We will show that the algorithm terminates with the correct codeword after $O(\log n)$ rounds. Here is a suggested plan to do it:

Set V_i to L when i is odd and to R when i is even. Let X_i be the set of erroneous edges after round i (so X_0 is the set of the originally corrupted edges). Let $S_i = \{v \in V_i | E_v \cap X_i \neq \emptyset\}$.

(b) Let $i > 0$. Prove that for all $v \in S_i$, $|E_v \cap X_{i-1}| \geq \delta_0 d/2$ and that for all $v \in S_i$, $|E_v \cap X_i| \geq \delta_0 d$.

(c) Deduce that $|S_1| \leq \frac{n}{2}(\delta_0/2 - \Lambda)$.

(d) Using inequality (1) for appropriate sets and the bound on S_1 to show that $|S_2| \leq \frac{2}{3}|S_1|$; continue inductively (using the obtained upper bound on S_i) to show that $|S_{i+1}| \leq \frac{2}{3}|S_i|$.

(e) Conclude that the decoding algorithm terminates with the correct code word after $O(\log n)$ rounds.

2. (a) Recall that in the discussion about Delsarte LP, we have defined α_i as the value of a symmetrized version of $\mathbf{1}_C * \mathbf{1}_C$, where C is a code. Specifically, for $x \in \{0,1\}^n$ we define

$$g(x) = \frac{1}{n!} \sum_{\sigma \in S^n} (\mathbf{1}_C * \mathbf{1}_C)(\sigma(x)),$$

where $\sigma(x)$ is defined as the permuted version of x by σ. In other words $\sigma(x)_i = x_{\sigma(i)}$. Now, let $\alpha_i = g(x)$ for $|x| = i$. What quantity of C is captured by α_i?

The Krawtchouk polynomial is defined as $K_t(x) = \sum_{i=0}^{t} (-1)^i \binom{x}{i} \binom{n-x}{t-i}$. Notice that we can define $\binom{y}{k}$ where y is *real* and k is a nonnegative integer as $\binom{y}{k} = \frac{y \cdot (y-1) \cdot \ldots \cdot (y-k+1)}{k!}$ (and notice that $\binom{y}{0}$ is always 1). The Delsarte Linear Program we defined in class is

$$\begin{aligned} & \max \sum_{i=0}^{n} \alpha_i \\ \text{s.t.} \quad & \\ & \alpha_i \geq 0 \text{ for all } i \\ & \alpha_1 = \alpha_2 = \ldots = \alpha_{d-1} = 0 \\ & \alpha_0 = 1 \\ & \sum_{j=0}^{n} \alpha_j K_t(j) \geq 0 \text{ for all } 0 \leq t \leq n \end{aligned}$$

We have shown that if C is a code of block length n and distance d then $|C|$ is at most the solution to the LP. Therefore, to bound the size of a code it is possible

to analyze the LP and provide an optimal solution. We show an alternative formulation that allows to get the bound by presenting one particular solution to a (different) LP.

(b) Let $\beta(x) = 1 + \sum_{t=1}^{n} \beta_t K_t(x)$ be any polynomial with $\beta_t \geq 0$ $(1 \leq t \leq n)$ such that $\beta(j) \leq 0$ for $j = d, d+1 \ldots, n$. **Show** that $|C| \leq \beta(0)$ whenever C is a code of block length n and distance d.

(c) Using the above, Show that a code with block length n with distance at least $n/2$ cannot contain more than $2n$ words. (We saw this already as part of Plotkin bound). Hint: use the polynomial $\beta(x) = 1 + K_1(x) + \frac{2}{n}K_2(x)$.

3. In class we saw that if C' is a code with dual distance d (namely $\hat{\mathbf{1}_{C'}}(S) = 0$ for $0 < |S| < d$) , and if $B \subset \{0,1\}^n$ with $\lambda_B \geq n - 2d + 1$ then

$$\left| \bigcup_{z \in C'} (z + B) \right| \geq 2^n / n.$$

(a) Show that if C' is a linear code with distance d then its dual code as dual distance d.

(b) From the above deduce that if C is a linear code with distance d, and if B is a set with $\lambda_B \geq n - 2d + 1$ then

$$|C| \leq n|B|.$$