HW 2

Error-Correcting Codes (Spring 2016)
Rutgers University
Swastik Kopparty

Due: March 24, 2016

You may skip any one problem of your choosing.

1. The binary erasure channel with parameter p, denoted $BEC(p)$, is the channel over a $\{0,1\}$ alphabet which independently *erases* each symbol sent on it with probability p (i.e. with probability p any symbol 0 or 1 which is sent on the channel is received as a ?, and with probability $1-p$ the symbol goes through unharmed).

   Show that for all $\epsilon > 0$, there are codes of length n and rate $1-p-\epsilon$ which can be used to transmit messages on $BEC(p)$ such that the probability of incorrect decoding is $\exp(-n)$.

   Show that one cannot transmit at a rate $1-p+\epsilon$.

2. Let C be a Reed-Solomon code over $\mathbb{F}_q$ with length N and distance D.

   (a) Let $c \in C$. Suppose x is a received word obtained from c after r errors and s erasures occur.
   Give a polynomial time algorithm, which on input x can recover c, provided:

   $$r + \frac{s}{2} < \frac{D}{2}.$$

   (b) Let $c \in C$. Let $x \in \mathbb{F}_q^N$ and $u \in [0,1]^N$: we will view u_i as the amount of "uncertainty" in the symbol x_i ($u_i = 1$ is like an erasure). For each $i \in [N]$, define err_i by:

   $$err_i = \begin{cases} 1 - u_i/2 & x_i \neq c_i \\ u_i/2 & x_i = c_i \end{cases}$$

   Give a polynomial time algorithm, which on input x and u can recover c, provided:

   $$\sum_{i \in [N]} err_i < \frac{D}{2}.$$

   A hint for this available at the end of the problem set.

   (c) Let $C_{in} \subseteq \{0,1\}^n$ be a binary code with q codewords. Let d be the minimum distance of C_{in}. Let V be the concatenated code obtained by concatenating C with C_{in}. Recall that V has minimum distance $\geq D \cdot d$.
   Here is an algorithm for decoding V from $\frac{D \cdot d}{2}$ errors.

      i. Let $y_1, y_2, \ldots, y_N \in \{0,1\}^n$ be the blocks of the received vector y.

ii. Decode each y_i from up to $d/2$ errors to obtain a codeword $c_i \in C_{in}$. Let $a_i = \Delta(y_i, c_i)$.

iii. Let $x_i \in \mathbb{F}_q$ be the $\mathbb{F}_q$-symbol corresponding to c_i. Let $u_i = \frac{a_i}{d/2}$.

iv. Then (x, u) satisfy the hypothesis for the previous part of this problem. Decode this to obtain the codeword c.

Show that this algorithm works.

3. For each $R \in (0, 1)$, show that there exist linear codes $C \subseteq \mathbb{F}_2^n$ such that both C and $C^\perp$ meet the Gilbert-Varshamov bound.

4. A code $C \subseteq \{0,1\}^n$ is called a covering code with covering radius r if for every $x \in \{0,1\}^n$, there exists some $c \in C$ with $\Delta(x, c) \leq r$.

   Let $\rho \in (0, 1/2)$ be a constant. Show that every covering code $C \subseteq \{0,1\}^n$ with covering radius ρn has rate $R \geq 1 - H(\rho) - o(1)$. Show that there exist covering codes C with covering radius ρn with rate $R \leq 1 - H(\rho) + o(1)$.

   Thus the the main combinatorial questions for covering codes are much easier than for error-correcting codes! In fact, one can even construct such covering codes efficiently! We may see this in a future problem set.

5. Let c be a constant. Show that a random code (and even a random linear code) in $\{0,1\}^n$ with n^c codewords has minimum distance at most $\frac{n}{2} - \Theta(\sqrt{cn \log n})$ with high probability.

   Recall that dual-BCH codes with n^c codewords have distance around $\frac{n}{2} - \Theta(c\sqrt{n})$.

6. **(Not to be turned in)** Compute the asymptotics as $\epsilon \to 0$ of the upper and lower bounds on the rate R of codes with minimum distance $\delta = \frac{1}{2} - \epsilon$.

   Note the big gap.

Hint for weighted Reed-Solomon decoding: reduce to errors-and-erasures decoding.