ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 26 Advanced Digital Communications
Homework 11 Dec. 5, 2016

PROBLEM 1. In this example we will explore some of the basic properties of binary linear block codes. A binary linear block code is a subspace of $\{0,1\}^n$ for some n and therefore has a dimension k, $0 \le k \le n$. We can therefore represent such a code $\mathcal{C}$ as

$$\mathcal{C} := \{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x} = \boldsymbol{u}G; \boldsymbol{u} \in \{0,1\}^k\},$$

where $G \in \{0,1\}^{k\times n}$ is called the *generator matrix*. Define the set of words $\mathcal{C}^\perp$ as

$$\mathcal{C}^\perp := \{\boldsymbol{y} \in \{0,1\}^n : G\boldsymbol{y}^T = \mathbf{0}^T\}.$$

(a) Show that $\mathcal{C}^\perp$ is a linear subspace of $\{0,1\}^n$ and has dimension $n-k$.

(b) From the (a) conclude that $\mathcal{C}^\perp$ has a representation of the form

$$\mathcal{C}^\perp := \{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x} = \boldsymbol{u}H; \boldsymbol{u} \in \{0,1\}^{n-k}\}.$$

(c) Show that $\boldsymbol{x} \in \mathcal{C}$ if and only if $H\boldsymbol{x}^T = \mathbf{0}^T$. H is called the *parity check matrix.*

PROBLEM 2. The weight of a binary sequence of length n is the number of 1's in the sequence. The Hamming distance between two binary sequences of length n is the weight of their modulo 2 sum. Let $\boldsymbol{x}_1$ be an arbitrary codeword in a linear binary code of block length n and let $\boldsymbol{x}_0$ be the all-zero codeword. Show that for each $d \le n$, the number of codewords at distance d from $\boldsymbol{x}_1$ is the same as the number of codewords at distance d from $\boldsymbol{x}_0$.

PROBLEM 3.

(a) Show that in a binary linear code, either all codewords contain an even number of 1's or half the codewords contain an odd number of 1's and half an even number.

(b) Let $x_{m,i}$ be the i^{th} digit in the m^{th} codeword of a binary linear code. Show that for any given i, either half or all of the $x_{m,i}$ are zero. If all of the $x_{m,i}$ are zero for a given n, explain how the code could be improved.

(c) Show that the average number of ones per codeword, averaged over all codewords in a linear binary code of block-length n, can be at most $n/2$.

(d) A linear code is called *proper*, if its generator matrix has no all zero column. Prove that if a codeword chosen uniformly at random from a binary linear code then each digit of the codeword is uniformly distributed on $\{0,1\}$.

PROBLEM 4. As we discussed in class, one way to design good codes is to look at their distance profile. In particular, the minimum distance of a code $\mathcal{C}$ defined as

$$d_{\min}(\mathcal{C}) = \min_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{C}\\ \boldsymbol{x}\neq\boldsymbol{y}}} \mathrm{d_H}(\boldsymbol{x},\boldsymbol{y})$$

turns out to be an important characterizing factor of its performance. In the above $\mathrm{d_H}(\boldsymbol{x},\boldsymbol{y})$ is the Hamming distance between two codewords, i.e., the number of positions they differ, or in the case of binary codes, the number of ones in their modulo-2 sum. In this problem we look at some basic properties of minimum distance.

(a) Prove that if $\mathcal{C}$ is a linear code,

$$d_{\min} = \min_{\substack{\boldsymbol{x} \in \mathcal{C} \\ \boldsymbol{x} \neq \mathbf{0}}} \mathrm{w_H}(\boldsymbol{x})$$

where $\mathrm{w_H}(\boldsymbol{x})$ denotes the Hamming weight of $\boldsymbol{x}$, the number of positions $\boldsymbol{x}$ is non-zero

(b) Prove that any binary code of block-length n with M codewords (not even necessarily linear) with minimum distance $d_{\min}$ must satisfy

$$\sum_{i=1}^{\lfloor (d_{\min}-1)/2 \rfloor} \binom{n}{i} \leq \frac{2^n}{M}.$$

Make sure to carefully formulate your argument.

The inequality you proved in (b) (known as *Hamming bound*) says if we wish to increase the minimum distance of a code we need to decrease the number of codewords at a fixed block-length n. Now we would like to see how fast the minimum distance can grow with n assuming the code rate $R := \log(M)/n$ is fixed (i.e., M, when the number of codewords exponentially with n.)

(c) Let $h_2(p) := -p \log_2(p) - (1-p)\log_2(1-p)$, $0 \leq p \leq 1$, be the *binary entropy function.* Starting from the bound in (a), show that if $R := \lim_{n\to\infty} \frac{\log_2(M)}{n}$ and $\delta := \lim_{n\to\infty} \frac{d_{\min}}{n}$, then

$$h_2(\delta/2) \leq 1 - R.$$

That is, $d_{\min}$ can grow linearly with n but there is a trade-off between code rate R and δ the slope of this growth.

Hint. Using Stirling's approximation it can be shown that $\lim_{n\to\infty} \frac{1}{n} \log_2 \binom{n}{np} = h_2(p)$.

Problem 5. Show that the message passing decoder for the BEC is suboptimal by finding a simple graph and a particular codeword such that the ML decoder will succeed but such that the iterative algorithm will fail. What is the smallest example you can find?