EE 121 – Intro to Digital Communication Spring 2016

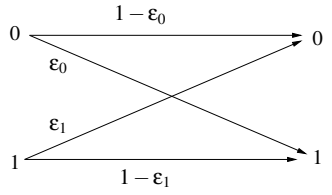
Homework 5

Assigned: Thur, 3/3/16 Due: Thur, 3/10/16

Instructor: Omur Ozel *GSI: Preetum Nakkiran*

1 Binary Asymmetric Channel

The binary asymmetric channel (BAC) is defined as in the following figure:



Formally, this is a binary input/ binary output channel defined by parameters ε_0 and ε_1 where

$$\Pr[Y = 1|X = 0] = \varepsilon_0, \qquad \Pr[Y = 0|X = 1] = \varepsilon_1$$

Note that if $\varepsilon_0 = \varepsilon_1$, then this channel is the binary symmetric channel we considered in the class. In this question, we consider optimal decoder over the BAC.

(a) Assume the encoder sends codewords $c_1 = 010$, $c_2 = 101$ and $c_3 = 110$ with equal probability and $\varepsilon_0 = \varepsilon_1 = \varepsilon$. Determine the optimal decoder, i.e., determine the optimal mapping from $\{0,1\}^3$ to $\{c_1, c_2, c_3\}$ that minimizes the probability of decoding error. Find expressions for the decoding error and the bit error rate for the optimal decoder. Does the optimal decoder change if the value of ε changes?

(b) Now, assume that the encoder sends codewords $c_1 = 010$ and $c_2 = 101$ with equal probability over the BAC and $\varepsilon_0 \neq \varepsilon_1$. Determine the optimal decoder in this case. Does the optimal decoder change if the values of ε_0 and ε_1 change?

2 Error Performance of A Linear Code

Consider a linear code $\mathcal{C}$ over the usual binary field $\mathbb{F}_2$ with parameters n, k and d_{min}. Let A_w denote the weight distribution of $\mathcal{C}$. Assume all messages are equally likely and the decoder implements MAP rule.

(a) Assume that the codewords are sent over the binary symmetric channel with crossover probability ε independent over time. Determine the probability of undetected error in terms of A_w conditioned on a specific codeword $\mathbf{c}$. Show that this probability does not depend on $\mathbf{c}$.

(b) In the setting of part-a, consider the probability of block error conditioned on a specific codeword $\mathbf{c}$. Does this probability depend on $\mathbf{c}$? Does the bit error rate depend on $\mathbf{c}$?

(c) Now, assume that the repetition code of length n is used over a binary asymmetric channel with probabilities $\varepsilon_0 = 0$ and $\varepsilon_1 \neq 0$ independent over time (this specific channel is called the Z-channel). Determine the probability of undetected error and the probability of block error conditioned on a specific codeword $\mathbf{c}$. Does these probabilities depend on $\mathbf{c}$?

3 Linear Source/Channel Coding Duality

In this problem, we will begin to explore the duality between source and channel coding. That is, we will see connections between the problem of coding for noisy channels (adding redundancy) and the problem of compressing a source (removing redundancy).

In fact, we have already seen an instance of this in the $(7, 4)$ Hamming code, with generator and parity-check matrices:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \qquad H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Recall that the $(7, 4)$ Hamming code has minimum-distance 3, and thus can correct all patterns of 1-bit errors. In fact, all noise patterns $\mathbf{e} \in \mathbb{F}_2^7$ of hamming-weight $w_H(\mathbf{e}) \leq 1$ result in different syndromes $\mathbf{e}H$. Thus, the parity check matrix defines a linear function $H : \mathbb{F}_2^7 \to \mathbb{F}_2^3$ that is injective when restricted to inputs with hamming-weight ≤ 1. That is, *H perfectly compresses* 7-bit strings of small hamming-weight into 3-bit strings. Such a function could be useful, for example, if we are trying to compress a source that outputs a stream of iid Bernoulli(p) bits, for some small p (thus we would expect the bit-strings produced have small hamming weight). In other words, we can think of H as *compressing the noise* introduced by the channel, into a 3-bit syndrome.

(a) Show that the phenomenon we observed for Hamming codes (above) holds in general. That is: Let $\mathcal{C}$ be an (n, k) t-error-correcting binary linear code, with parity-check matrix H. Show that all vectors $\mathbf{e}$ of hamming-weight $w_H(\mathbf{e}) \leq t$ map to distinct syndromes $\mathbf{e}H$. (That is, the map $H : \mathbb{F}_2^n \to \mathbb{F}_2^{n-k}$ is injective when restricted to vectors of weight $\leq t$). Thus, H can be used to perfectly compress a source that produces blocks of low hamming-weight.

(b) **(optional)** Now let us show a more robust version of this. Let $\mathcal{C}$ be an arbitrary binary linear code that is good for the binary symmetric channel BSC_p. That is, if a uniformly random codeword $c \in \mathcal{C}$ is sent across BSC_p (corrupted by noise n), it can be decoded such that $\Pr_{\mathbf{e}, c \in \mathcal{C}}[\{\text{decoding error}\}] \leq \varepsilon$.
Suppose we are trying to compress a source that produces independent and identically distributed Bernoulli(p) bits, for some small p (same as the channel noise). Let H be the pairity-check matrix of $\mathcal{C}$, and define the compression function as $Comp : \mathbf{e} \mapsto \mathbf{e}H$. Show that this compression can be decompressed, such that $\Pr_{\mathbf{e}}[\{\text{decompression error}\}] \leq \varepsilon$ (where the probability is over the randomness of the source $\mathbf{e}$, and of the decoder).
(Hint: This follows from the optimality of syndrome decoding. Recall that the MAP decoder finds the most-likely noise, given the received vector. And this depends only on the syndrome of the received vector – or equivalently on the syndrome of the noise. Thus, the optimal decoder estimates $\mathbf{e}$, given the syndrome $\mathbf{e}H$. This is exactly the role of the decompresser – given the compressed $\mathbf{e}H$, estimate $\mathbf{e}$...)
Remark: Applying this result, consider a random linear code for BSC_p. We know this achieves rate $R \approx 1 - H(p)$, meaning $k \approx (1 - H(p))n$. So the parity-check matrix H is $n \times nH(p)$, and it compresses n iid Bernoulli(p) bits to $\approx nH(p)$ bits. Note that $nH(p)$ is the *entropy* of a block of n iid Bernoulli(p) bits, and we have shown that we can compress this block into $nH(p)$ bits by applying a random linear projection.