CS 860 Topics in Coding Theory Sept. 23, 2025

Lecture 1

Lecturer: Elena Grigorescu *Scribe: Tiger Wu*

1 Random Error

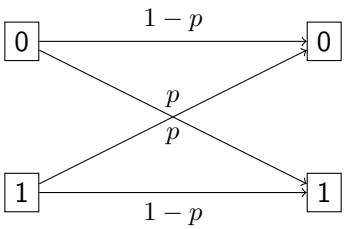
We consider two types of random errors:

- *Binary Symmetric Channel*(BSC)
- *Binary Eraser Channel*(BEC)

Binary Symmetric Channel (BSC)

Binary symmetric channel has an input space equal to the output space, $\{0, 1\}$. A BSC with crossover probability $0 \leq p \leq \frac{1}{2}$, denoted as BSC_p, flips an input with probability p.

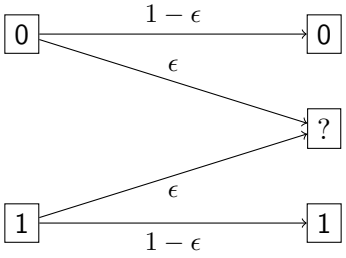
Figure 1: BSC_p



Binary Erasure Channel (BEC)

Binary erasure channel has input space $\{0, 1\}$ and ouput space $\{0, 1, ?\}$. Binary erasure channel with parameter $0 \leq \epsilon \leq \frac{1}{2}$, denotes as BEC_ϵ erase an input with probiablity ϵ.

Figure 2: BEC_ϵ



2 Basic Probability Facts

Let D be a probability distribution over U. A random variable $X \leftarrow D$ is drawn from U according to U. Let $f : U \to \mathbb{R}$ be a function.

Variance. The variance is defined as

$$\mathrm{Var}_{X \leftarrow D}[f(X)] = \mathbb{E}[f^2(x)] - \mathbb{E}[f(X)]^2.$$

Expecation. For random variables X_1 and X_1,

$$\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2)$$

and

$$\mathbb{E}(X_1 X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2) \text{ if and only } X_1 \text{ and } X_2 \text{ are independent.}$$

Expection. For events $\mathcal{E}_1$ and $\mathcal{E}_2$, the union bound states that

$$\Pr[\mathcal{E}_1 \cup \mathcal{E}_2] \leq \Pr[\mathcal{E}_1] + \Pr[\mathcal{E}_2].$$

Indicator Random Variable. For event $\mathcal{E}$, the indicator variable I_ϵ is a 0/1 random variable defind as

$$I_{\mathcal{E}}(X) = \begin{cases} 1 & \text{if } X \in \epsilon \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$\mathbb{E}[I_\epsilon(X)] = \Pr[\mathcal{E}].$$

Tail Bound/Markov Inequality. Let X be a non-negative random variable and $\alpha > 0$. Then

$$\Pr[X \geq \alpha] \leq \frac{\mathbb{E}[X]}{\alpha}.$$

E.g.

- If $\alpha = 2\mathbb{E}[X]$, $\Pr[X \geq \alpha] \leq \frac{1}{2}$.
- If $\alpha = \sqrt{\mathbb{E}[X]}$, $\Pr[X \geq \sqrt{\mathbb{E}[X]}] \leq \sqrt{\mathbb{E}[X]}$.

Chernoff Bounds for Bernoulli Random Variable. If $X_1, \ldots, X_n$ are i.i.d. $\{0,1\}$-valued random variables with $\Pr[X_i = 1] = p$. For $\epsilon > 0$, for large enought n,

$$\Pr\left[\sum X_i \geq (p+\epsilon)n\right] \leq 2^{-\frac{\epsilon^2}{2}n}$$

and

$$\Pr\left[\sum X_i \leq (p-\epsilon)n\right] \leq 2^{-\frac{\epsilon^2}{3}n}.$$

3 Shannon Capacity

The capacity of a BSC_p channel is

$$\lim_{n\to\infty} \frac{r}{n}$$

which depends only on the channel. We will show that the capcity of BSC_p is $1 - H(p)$ [1]. That is, we will show that there exists encoding and decoding maps $E : \{0,1\}^k \to \{\alpha k\}$ and $D : \{0,1\}^{\alpha k} \to \{0,1\}$ for $\alpha > 1$ such that with high probability over noise vector η decoding is successful.

Recall that $\eta \in \{0,1\}^n$ is a random variable with each word being 0 with probability p and 1 with probability $1 - p$.

Theorem 1 (Shannon's Capacity Theorem) *For all $p \in [0, \frac{1}{2})$, $0 \le \gamma \le \frac{1}{2} - p$, for large enough n, there exists $\beta = \beta(\gamma, p)$, let $\alpha = \frac{1}{1-H(p+\gamma)}$, and $E : \{0,1\}^k \to \{0,1\}^{\alpha k}$, $D : \{0,1\}^{\alpha k} \to \{0,1\}^k \cup \{\text{'fail'}\}$ such that*

$$\Pr_{\eta}[D(E(m) + \eta) = m] \ge 1 - 2^{-\Omega(\epsilon^2)n}.$$

Proof By the probabilistic method. We look at the random encoding functions. Let $\ell = k + 1$. Let $E : \{0,1\}^k \to \{0,1\}^\ell$ be randomly chosen from all such functions. This means that for a fixed m, $E(m)$ is uniformly random in $\{0,1\}^n$. Let $\epsilon = \epsilon(\gamma) > 0$ be small enough. The decoding function is deifned as:

$$D(y) = \begin{cases} m & \text{if } m \text{ is unique codeword such that } \Delta(y, E(m)) \le (p+\epsilon)n \\ \text{fail} & \text{otherwise.} \end{cases}$$

(D is not efficient).

Decoding is not successful when one of the two following events will happen:

1. Too many errors have been incurred. i.e., $\Delta(y, E(m)) > (p+\epsilon)n$, where $y = E(m)+\eta$.
2. More than 2 messages have encodings in $B(y, (p+\epsilon)n)$.

For fixed m, consider all E.
Case 1. Consider the first event.

$$\begin{aligned} \Pr_{\eta}[\Delta(y, E(m)) > (p+\epsilon)n] &= \Pr[\text{wt}(y - E(m)) > (p+\epsilon)n] \\ &= 2^{-\Omega(\epsilon^2)n} \qquad \text{(by Chernoff bound)} \end{aligned}$$

Case 2.. Consider the second event. Fix a y, (which fixes η and $m' \ne m$.

$$\begin{aligned} \Pr_{E}[E(m') \in B(y, (p+\epsilon)n)] &\le \frac{|B(y, p+\epsilon)|}{2^n} \\ &\le 2^{(H(p+\epsilon)+o(1))n} \end{aligned}$$

[1] binary entopy: $H(x) = -x \lg(x) - (1-x)\lg(1-x)$

By union bound over all m',

$$\begin{aligned}\Pr_E\left[\exists m', E(m') \in B(y,(p+\epsilon)n)\right] &\leq 2^{k+1} \cdot 2^{n(H(p+\epsilon)-1+o(1))} \\ &= 2 \cdot 2^{(1-H(p+\gamma))n+n(H(p+\epsilon)-1+o(1))} \\ &= 2 \cdot 2^{(-H(p+\gamma)+H(p+\epsilon)+o(1))n}\end{aligned}$$

By lineraity of expectation,

$$\begin{aligned}\mathbb{E}_E\left[\Pr_\eta[D(E(m)+\eta) \neq m]\right] &\leq 2^{-\Omega(\epsilon^2)n} + 2 \cdot 2^{(H(p+\epsilon)-H(p+\gamma)+o(1))n} \\ &< \frac{1}{2}2^{-\beta n}\end{aligned}$$

where β is a function of p and γ. This implies there exists an encoding function that is decoded correctly with high probability. Our goal is to say there exists such an E that is good for all messages m, and the union bound is not strong enough for this.

Since we have the above for each fixed m,

$$\begin{aligned}&\mathbb{E}_m\left[\mathbb{E}_E\left[\Pr_\eta[D(E(m)+\eta) \neq m]\right]\right] < \frac{1}{2}2^{-\beta n} \\ \implies &\mathbb{E}_E\left[\mathbb{E}_m\left[\Pr_\eta[D(E(m)+\eta) \neq m]\right]\right] < \frac{1}{2}2^{-\beta n}\end{aligned}$$

Therefore there exsits E^* such that

$$\mathbb{E}_m\left[\Pr_\eta\left[D(E^*(m)+\eta) \neq m\right]\right] < \frac{1}{2}2^{-\beta n}$$

Applying Markov's inequality,

$$\Pr_m[\Pr_\eta[D(E^*(m)+\eta) \neq m] > 2^{-\beta n}] < \frac{1}{2}.$$

Hence, for at least $1/2$ fraction of the messages, m, E^* fails on m with probability $< 2^{-\beta n}$. Removing all the other messages, which are fewer than $\frac{1}{2}e^\ell$, we are left with a code with $> 2^\ell \cdot \frac{1}{2} = 2^k$ many messages for which E^* decodes correctly with probability $1 - 2^{-\beta n}$. ■

Observation 2 *Neither E^* nor D^* are efficiently computable.*

3.1 Connection between Shannon and Hamming

Claim 3 *If $C \subseteq \{0,1\}^n$ and $\Delta(C) = 2p+\epsilon$, then we can communicate over BSC_p decoding correctly with $1 - 2^{-\Omega(\epsilon^2)}$.*

Proof By Chernoff bound, with high probability, the number of errors is

$$\leq \frac{\Delta(c)}{2} = \frac{p+\epsilon}{2}$$

So we have unique closest codeword. ■

4 Shannon's Converse Theorem

Theorem 4 (Shannon's Converse Theorem) *For all $p \in [0, \frac{1}{2})$, for all $\epsilon, \delta > 0$, for large enough n, for $k \geq (1 - H(p) + \epsilon)n$, for all $E : \{0,1\}^k \to \{0,1\}^n$, for all $D : \{0,1\}^n \to \{0,1\}k$*

$$\Pr_{\eta, m \leftarrow \{0,1\}^k}[D(E(m) + \eta) \neq m] \geq 1 - \delta.$$

In other words, if we are sending information at a rate $> 1 - H(p) + \epsilon$, decoding is on average erroneous for any encoding/decoding.

Intuition

For $1 - H(p)$ being the upperbound, suppose $D : \{0,1\}^n \to \{0,1\}^k \cup \{\text{'fail'}\}$ decodes with negligible error probability. Typical noise has weight $\in [(p - \epsilon)n, (p + \epsilon)n]$. So $E(m) + \eta$ takes approximately $2^{H(p)n}$ many possibilities. To decode correctly, D should map most of these $2^{H(p)}$ strings back to m. So $|C| \leq 2^n / 2^{H(p)n}$, therefore $R < 1 - H(p)$.