

Shannon's Theorem, Linear Codes

1 Shannon's Theorem

To prove Shannon's Theorem we will make use of the following two lemmas:

Lemma 1.1 $\binom{n}{np} = 2^{H(p)+o(1)}$

Lemma 1.2 $\Pr_{\eta} \left(\left| \frac{wt(\eta)}{n} - p \right| > \lambda \right) \leq e^{-\frac{\lambda^2 n}{2}}$, where η is a vector of length n , each element being either a 0 or a 1.

We wish to prove that

Theorem 1

$$\Pr\{D(E(x) \oplus \eta) = x\} < c^{-n}$$

for some constant c and every coding, decoding functions E, D , given $|x| = k$, every error vector η , and we know that $k/n > 1 - H(p)$.

Proof: First, we examine the probability that the error vector η is some fixed vector b . We know that there are $\binom{n}{pn} = 2^{nH(p)}$ total error vectors. It follows then that

$$\Pr\{\eta = b\} \leq \frac{1}{\binom{n}{pn}} = 2^{-nH(p)+o(1)}$$

for big enough n .

Next, we examine the space of all codewords $\{0, 1\}^n$. We know that the length of a word of data before coding is k bits, so the total number of data words is 2^k . Also, through the Encoding function $E : 0, 1^k \rightarrow 0, 1^n$, we map every data word to a subset of the total codeword space S_x as seen in Figure 1

In this context, when $E(x_i) \oplus \eta$ is a vector $y \in S_{x_i}$, we will be able to decode y to x , otherwise we will have an error. Therefore,

$$\Pr\{D(E(x) \oplus \eta) = x\} = \Pr_{x_i, y \in S_{x_i}} \{E(x_i) \oplus \eta = y\}$$

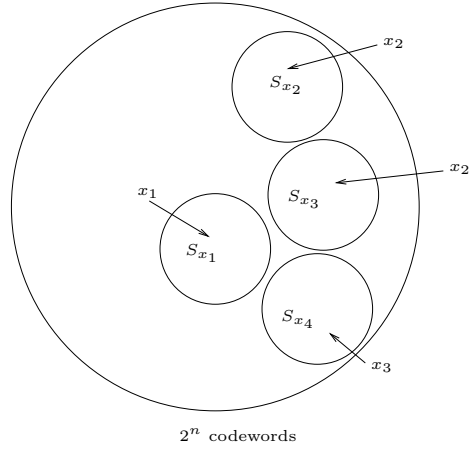


Figure 1: Every possible data word x_i is mapped to a subset of $\{0, 1\}^n$ S_{x_i}

Note as well that for a given η , there is only one y that will be the result of $E(x_i) \oplus \eta$, so this is the same as writing $\eta = E(x_i) \oplus y$.

Now we are ready to get the expression for

$$\begin{aligned}
\Pr\{D(E(x) \oplus \eta) = x\} &= \frac{1}{2^k} \sum_{x_i} \sum_{y \in S_{x_i}} \Pr\{(E(x) \oplus \eta) = y\} \\
&= \frac{1}{2^k} \sum_{x_i} \sum_{y \in S_{x_i}} \Pr\{\eta = E(x_i) \oplus y\} \\
&\leq \frac{1}{2^k} \sum_{x_i} \sum_{y \in S_{x_i}} 2^{-n(H(p)+o(1))} \\
&= \frac{2^n}{2^k} 2^{-n(H(p)+o(1))} \\
&= 2^{n-k} 2^{-n(H(p)+o(1))}
\end{aligned}$$

But remember that $k/n > 1 - H(p)$, and $n > k$ implying that $n - k = n(H(p) - \epsilon)$ for some $\epsilon > 0$:

$$\begin{aligned}
\Pr\{D(E(x) \oplus \eta) = x\} &\leq 2^{n(H(p)-\epsilon)} 2^{-n(H(p)+o(1))} \\
&= 2^{-\epsilon n}
\end{aligned}$$

■

2 Linear Codes

2.1 Hamming Code

Hamming codes are defined as a family of (n, k, d) block error-correcting codes, where:

- $n = 2^l - 1$, n being the block length
- $k = 2^l - l - 1$, k being the number of data bits
- $n - k = m$, m being the number of check bits
- $d_{min} = 3$, where d_{min} is the minimum distance

For the Hamming codes we know that the channel can mess up to t bits. Let d be the minimum distance between 2 codewords. In the case we want to perform:

- error-correction: $d \geq 2t + 1$
- error-detection: $d \geq t + 1$

The parity check matrix is the following:

$$H = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 \\ & & & \dots & & & \\ 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{pmatrix}$$

where H has m columns and $2^m - 1$ rows and the i^{th} row is the m -bit binary representation of i .

If H is the parity check matrix, then for every codeword y , $y \times H = [0]$. Remember that by codeword we mean the original string we want to send.

2.2 Hadamard Code

Hadamard matrices are used for spread spectrum chipping sequences.

$$G = \begin{pmatrix} 0 & 0 & 0 & \dots & 1 & 1 & 1 \\ & & & \cdot & & & \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 0 & 0 & 0 & \dots & 1 & 1 & 1 \\ 0 & 0 & 1 & \dots & 0 & 1 & 1 \\ 0 & 1 & 0 & \dots & 1 & 0 & 1 \end{pmatrix}$$

In this case G acts as a generator matrix for Hadamard code, where G has m rows and 2^m columns.