09/1/2020 Lecture 2 (Examples of Codes, Framework of Block Codes).

message.

$\underline{u}$ - vector of length K bits

$\underline{c}$ - codeword of length n bits.

$\underline{y}$ - noisy version of $\underline{c}$.

$\hat{\underline{c}}$ - Estimate of the codeword at the output of the decoder.

Encoder itself is a one-one map.

$\hat{\underline{c}}$ will uniquely map to a $\hat{\underline{u}}$.

3-fold repetition code.

	$\underline{u}$	$\underline{c}$
Encoder	0	000
	1	111

Decoder: Majority Rule (assuming that the channel is BSC with $p<0.5$).

Probability of error calculation for this code.

r-fold repetition code

(say r odd). $\underline{u}$ $\underline{c}$

Encoder

$0 \longrightarrow \underbrace{00\ldots0}_{r}$ r zeros

$1 \longrightarrow \underbrace{11\ldots1}_{r}$ r ones.

Decoder: Majority rule. (Assuming BSC with $p<0.5$).

What is the expression for prob. of error?

Error event: $\geq \frac{r+1}{2}$ bits are flipped.

$$P_e(\underline{c}) = \sum_{j=\frac{r+1}{2}}^{r} \binom{r}{j} p^j (1-p)^{r-j}.$$

$$P_e = \sum_{j=\frac{r+1}{2}}^{r} \binom{r}{j} p^j (1-p)^{r-j}$$ (Averaged over the two codewords)

$P_e \downarrow$ as r increases.

Error detection capability and error correction capability of a code.

The max no. of errors which can be introduced anywhere in the codeword & you can still detect it → Meaning.

Whether the $\underline{y}$ is a codeword or not.

00...0 ⟶ Any r-1 errors.
11...1

If you introduce r errors;

$\underline{y}$ → 11111-1 was transmitted w/o any error.
→ 00...0 with r errors.

Max no. of errors which r-fold repetition code can detect = r-1.

Important:- To Understand. Error detection & error correction capability of a code, probabilities related to the channel do not come into picture.

r=3. 000 ⟶ max 2 errors: 011 101 110. 000: 010 001 100

111 → 111, 110, 011, 101, 010, 001, 100.

110.

Error Correction Capability of r-fold repetition Code.

Look at $\underline{y}$ and also estimate $\underline{\hat{c}}$ should match $\underline{c}$.

Max no. of errors which you can introduce s.t. when you look at $\underline{y}$ & use decoder of your choice, the estimate $\underline{\hat{c}}$ has to match $\underline{c}$.

3-fold repetition Code.

max 2 errors

$000 \rightarrow 000, 010, 100, 001, 011, 101, 110.$

$111 \rightarrow 111, 110, 101, 011, 010, 100, 001.$

If $\underline{y} = 100$ → 000 is transmitted & 1 error was introduced.

→ 111 was transmitted & 2 errors are introduced

max 1 error

$000 \rightarrow 000, 010, 100, 001$

$111 \rightarrow 111, 110, 101, 011$

Error Correction capability of a r-fold repetition code = $\frac{r-1}{2}$.

Simple parity check code

$\underline{u}$ $\longrightarrow$ $\underline{c}$.

binary vector of length 6 — binary vector of length 7.

$u_1 \ldots u_6$ $\quad$ $c_1 c_2 \ldots c_7$.

$c_i = u_i \quad 1 \le i \le 6$.

$c_7 = \sum_{i=1}^{7} u_i$ (XOR of all u_i's). mod 2 addition.

Error detection capability of simple parity check code is 1.

If there are 2 errors, then I get another codewords.

$\underline{u}$		$\underline{c}$.
101001	$\rightarrow$	1010011
		$\downarrow$ 2 errors
011001	$\leftarrow$	0110011

You can detect any odd no. of errors.

What is the error correction capability $\rightarrow$ Zero.

Hamming Code: (Code of 3 circles)

↳ (Informal).

u — binary vector of length 4. → c — binary vector of length 7.

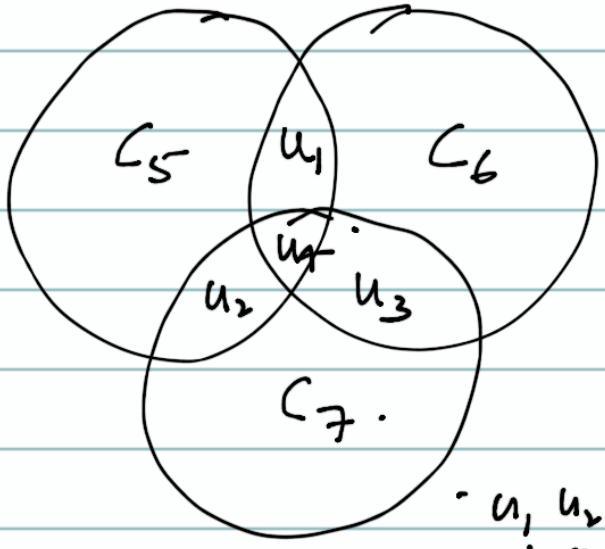
$c_i = u_i \quad 1 \le i \le 4.$

$c_5 = u_1 \oplus u_2 \oplus u_4.$

$c_6 = u_1 \oplus u_4 \oplus u_3.$

$c_7 = u_2 \oplus u_4 \oplus u_3.$

Encoder is completely specified.



Error detection capability of this code: 2.

Suppose if you flip u_1, u_2, and u_3.

$u_1\ u_2\ u_3\ u_4$ | $c_5\ c_6\ c_7$.

1 0 1 0 | 1 0 1.

1010 → 1010101.

↓ 3 flips.

0100 ← 0100101.

$c_5 = u_1 \oplus u_2 \oplus u_4.$

What is the error correction capability of Hamming Code?

Ans: 1

step 1:- Look at all circles which do not satisfy parity ($\underline{y} = y_1 y_2 y_3 y_4 y_5 y_6 y_7$)

step 2:- Take the bit at the intersection of all circles which do not satisfy the parity. Flip that bit.

This decoder corrects one error in any bit position.

Formal definitions →
- Block Code with parameters
- Channel Model.
- Prob. of error.
- Error detection Capability
- Error Correction Capability.

Block Codes.

Input bit stream is first split into blocks of k bits each.

Code Alphabet:- Set from which the vector $\underline{c}$ is drawing symbols from.
It is denoted by $\mathbb{F}$ or $\mathbb{F}_q$.
We have seen only examples of binary code alphabet. $\underline{c}$ is taking symbols from a binary alphabet (field) $\mathbb{F}_2$.

$\underline{c} \in \mathbb{F}^n$. $\rightarrow$ $\underline{c}$ is a vector of length n over code alphabet $\mathbb{F}$.

Block Code: An (n, M) block code $\mathbb{C}$ is a non-empty subset of $\mathbb{F}^n$ with $|\mathbb{C}| = M$.
$\rightarrow$ Elements of $\mathbb{C}$ are called codewords
$\rightarrow$ n is length of the code / blocklength.
$\rightarrow$ M is the size of the code, the no. of codewords in the code.

→ Rate of a Code

$$R = \frac{1}{n} \log_q M \quad \text{where } q = |F|.$$

How many information symbols are you transmitting per code symbol.

If $C = F^n$; then $R = \frac{1}{n} \log_q q^n = 1$.
(Uncoded transmission).

Examples:

r-fold repetition (over F_2).

$n = r$, $|C| = M = 2$.

Rate of the Code $= \frac{1}{r} \log_2 |M|$

$= \frac{1}{r} \log_2 2 = \frac{1}{r}$.

Simple parity check code.

$n = k+1$; $|C| = M = 2^k$.

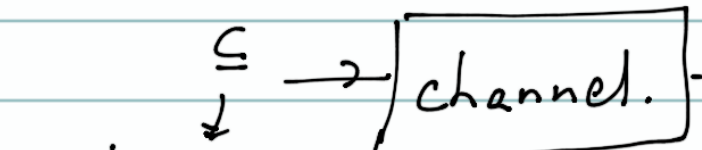
Rate of the Code $= \frac{1}{k+1} \log_2 (M)$

$= \frac{1}{k+1} \log_2 2^k = \frac{k}{k+1}$

Hamming Code:

$n = 7$, $|C| = M = 16$, Rate of the Code $= 4/7$.

Channel Model in general.



vector of length n over alphabet $\mathbb{F}$.

Channel is given by a triple $(\mathbb{F}, \mathcal{Y}, Pr)$.

→ $\mathbb{F}$ is the input alphabet

→ $\mathcal{Y}$ is output alphabet.

→ $Pr(\cdot|\cdot)$ are the transition probabilities:

$$Pr\left(\underline{y} \middle| \underline{c}\right).$$

BSC

Input alphabet = $\mathbb{F}_2$.

output alphabet = $\mathbb{F}_2$

$$\boxed{Pr\left(\underline{y} \middle| \underline{c}\right) = \prod_{i=1}^{n} Pr\left(y_i \middle| c_i\right)}$$

Memoryless property

$Pr(0|0) = 1-p \quad Pr(1|0) = p$

$Pr(0|1) = p. \quad Pr(1|1) = 1-p$

BEC

Input alphabet: $\mathbb{F}_2$.

output alphabet: $\{\mathbb{F}_2, ?\}$.

$$Pr\left(\underline{y} \middle| \underline{c}\right) = \prod_{i=1}^{n} Pr\left(y_i \middle| c_i\right)$$

$Pr(0|0) = 1-\epsilon, \quad Pr(?|0) = \epsilon$

$Pr(1|0) = 0$

$Pr(0|1) = 0, \quad Pr(?|1) = \epsilon$

$Pr(1|1) = 1-\epsilon.$

Decoder: $\mathcal{D}: \underline{\mathcal{Y}}^n \rightarrow \underline{\mathbb{C}}$.

Probability of error.

Let $\underline{c} \in \mathbb{C}$;

$$P_e(\underline{c}) = \sum_{\underline{y}: \mathcal{D}(\underline{y}) \neq \underline{c}} P_r(\underline{y} \mid \underline{c}).$$

$$P_e = \sum_{c \in \mathbb{C}} P(\underline{c}) P_r(\underline{y} \mid \underline{c}).$$

(Maximum A posteriori) → MAP
(Maximum Likelihood) → ML

} Decoders which are optimal in the sense of minimizing the prob. of error.