Channel Coding:

*) Chp. 13.

*) Idea: add redundancy into the transmitted data stream to protect the information against channel impairments (e.g. additive noise).

$0 \to 0$ $\quad$ $0 \longrightarrow 000$

$1 \to 1$ $\quad$ $1 \longrightarrow 111$

$P_e = p$ $\quad$ $P_e = \binom{3}{2} p^2(1-p) + \binom{3}{3} p^3$

$= 3p^2 - 2p^3$

Why does it help?

An example: Assume that four messages are being transmitted.

Case 1: uncoded transmission with BPSK.

$0\ 0 \longrightarrow -1\ -1$

$0\ 1 \longrightarrow -1\ +1$

$1\ 0 \longrightarrow +1\ -1$

$1\ 1 \longrightarrow +1\ +1$

Case 2: coded transmission

$0\ 0 \longrightarrow 1\ 1\ 1 \longrightarrow +1\ +1\ +1$

$0\ 1 \longrightarrow 1\ 0\ 0 \longrightarrow +1\ -1\ -1$

$1\ 0 \longrightarrow 0\ 0\ 1 \longrightarrow -1\ -1\ +1$

$1\ 1 \longrightarrow 0\ 1\ 0 \longrightarrow -1\ +1\ -1$

let's compare the performance of the two schemes over an AWGN channel.

Case 1. Average energy = 2.

min. squared Euclidean distance = 4.

$$\frac{d_{ij}^2}{E_{av}} = 2.$$

Case 2. Ave. energy = 3.

min. squared Euclidean distance = 8.

$$\frac{d_{ij}^2}{E_{av}} = \frac{8}{3}$$

At high SNRs the coded scheme is better by $10 \log_{10}\left(\frac{8/3}{2}\right) = \underline{1.25 \text{ dB}}.$

—— // ——

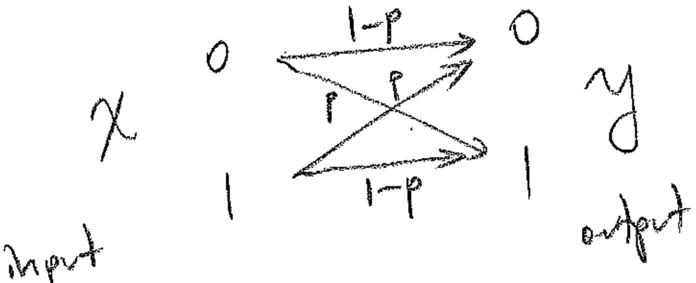
*) More sophisticated coding schemes would provide more gains

*) Cost : Instead of two bits, we transmit three bits (in the above example) ⇒ bandwidth expansion!

Ultimate limits:

*) For noisy channels, there exists a quantity called channel capacity (denote it by C) for which reliable communication (with arbitrarily low prob. of error) at rates R<C is possible. And, for R>C, error rate is bounded away from zero.

Ex) Binary symmetric channel. (BSC)

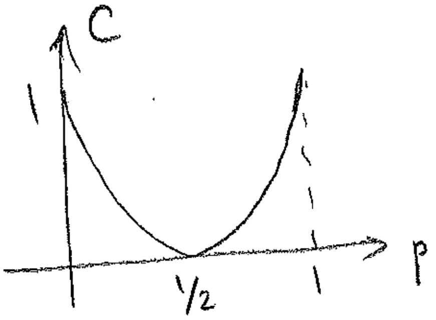


p: crossover probability

$1 - p\log_2\frac{1}{p} - (1-p)\log_2\left(\frac{1}{1-p}\right)$

$$C = 1 - H_b(p) \ \frac{\text{bits}}{\text{channel use}}$$

$H_b(\cdot)$: binary entropy function (in bits)



e.g. for $p = 0.1 \Rightarrow C = 0.53$ bits/use.

EX) Discrete time AWGN channel.

$$Y = X + N$$

X : input, power constraint P.
$(E[X^2] \leq P)$

$N \sim \mathcal{N}(0, P_N)$.

$$C = \frac{1}{2} \log\left(1 + \frac{P}{P_N}\right). \quad \text{bits/use.}$$

EX) Bandlimited Gaussian Waveform Channel

$$y(t) = x(t) + \eta(t)$$

Gaussian noise process with PSD: $\frac{N_0}{2}$.

input with
power constraint P
& bandwidth constraint W

$$C = W \log\left(1 + \frac{P}{N_0 W}\right) \quad \text{bits/sec.}$$

e.g. $W = 3\,\text{kHz}$, $SNR = \frac{P}{N_0 W} = 39\,\text{dB}$
$\Rightarrow \boxed{C = 38.8\ \text{kbps}}$

*) Channel coding is the way to approach the information theoretic channel capacity limits.

//

There are two main classes of codes:

- linear block codes.
- convolutional codes.

Linear Block Codes:

*) (n,k) block code is a collection of $M=2^k$ binary sequences of length n.

Codewords: $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_M$. (each an n-tuple)

Code rate: k/n.

Bandwidth expansion factor: $\frac{n}{k}$.

*) If the set of codewords $\underline{c}_1, \ldots, \underline{c}_M$ form a subspace of all n-tuples, then the code is called a linear block code.

An equivalent defn: If the modulo-2 sum of any two codewords is also a codeword, then the code is a linear block code.

*) <u>Note</u>: All zero sequence is always a codeword for a linear block code. (why?)

<u>ex)</u> Even parity code.

Add one parity bit to a sequence of message bits (of length k) to make the total number of 1's even.

e.g. $\underbrace{01101}_{\text{message bits}}\ \underbrace{1}$ $\qquad k=5, \quad n=6$

rate $= 5/6$.

With this code we can <u>detect</u> single bit errors in the transmission.

<u>ex)</u>

message bits

$$\begin{array}{ccccc|c}
1 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 \\
\hline
1 & 1 & 0 & 0 & 0 &
\end{array}$$

parity bits.

- even parity on each row & each column.
- rate $\frac{25}{35} = \frac{5}{7}$.
- Can correct single bit errors.

ex A (5,2) code with codewords $\{00000, 10100, 01111, 11011\}$ is a linear block code (verify!)

For encoding, the mapping of

$00 \rightarrow 00000$
$01 \rightarrow 01111$
$10 \rightarrow 10100$
$11 \rightarrow 11011$.

Generator and Parity Check Matrices:

*) Generator matrix $\underline{G}$ is a $k \times n$ matrix (of 0's & 1's) where the rows form a basis for the k-dimensional code subspace.

*) Define
$$\underline{e}_1 = (1,0,0,\cdots,0)$$
$$\underline{e}_2 = (0,1,0,\cdots,0)$$
$$\vdots$$
$$\underline{e}_k = (0,0,0,\cdots,1)$$
($1 \times k$ row vectors)

Then, if we pick $\underline{g}_i$ as the codeword corresponding to $\underline{e}_i$ $(i=1,2,\cdots,k)$, we can form a generator matrix for the code as
$$\underline{G} = \begin{bmatrix} \underline{g}_1 \\ \vdots \\ \underline{g}_k \end{bmatrix}.$$

*) Consider a message vector $\underline{x}$ as,

$$\underline{x} = (x_1, x_2, \cdots, x_k).$$

We can write: $\underline{x} = \sum_{i=1}^{k} x_i \underline{e}_i$, and as the codeword corresponding to $\underline{x}$ we can use:

$$\underline{c} = \sum_{i=1}^{k} x_i \underline{g}_i .$$

i.e., we have $\underline{c} = \underline{x}\,\underline{G}$.

*) Note: all operations are in the binary field.

ex) For the prev. ex:

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

e.g. for $\underline{x} = [1\ 1]$, $\underline{c} = [1\ 1]\,\underline{G} = [1\,1\,0\,1\,1]$.

Systematic Codes:

n-k parity check bits are added to the information bits to form the n-bit codewords.

In this case: $\underline{G} = [\,\underline{I}_k \,\vdots\, \underline{P}\,]$

$\underline{I}_k$: $k \times k$ identity matrix

$\underline{P}$: $k \times n-k$.

&

$$\underline{c} = \underline{x}\,\underline{G} = [\,\underbrace{\underline{x}}_{\text{message bits}} \,\vdots\, \underbrace{\underline{x}\,\underline{P}}_{\text{parity check bits}}\,]$$

i.e., $c_i = x_i \qquad i = 1, 2, \dots, k$

$$c_i = \sum_{j=1}^{k} P_{ji}\, x_j \quad \text{if} \quad i = k+1, k+2, \dots, n.$$

ex) $(4,2)$ code with $c_1 = x_1$, $c_2 = x_2$, $c_3 = x_1 + x_2$, $c_4 = x_1$ is a systematic code, with

$$\underline{G} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

parity check matrix.

*) parity check matrix $\underline{H}$ is an $n-k \times n$ matrix of 0's & 1's whose rows form a basis for the null space of $\underline{G}$.

*) For any codeword $\underline{c}$: $\underline{c} \cdot \underline{H}^t = \underline{0}$

*) $\underline{G} \cdot \underline{H}^t = \underline{0}$.

*) For systematic codes: $\underline{G} = [\underline{I}_k \vdots \underline{P}]$

& $\underline{H} = [\underline{P}^t \vdots \underline{I}_{n-k}]$.

[to see this: $\underline{G} \cdot \underline{H}^t = [\underline{I}_k \vdots \underline{P}] [\underline{P}^t \vdots \underline{I}_{n-k}]^t$

$$= [\underline{I}_k \vdots \underline{P}] \begin{bmatrix} \underline{P} \\ \underline{I}_{n-k} \end{bmatrix}$$

$$= \underline{P} + \underline{P} = \underline{0}]$$

ex) Prev. example: $\underline{G} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$

$$\Rightarrow \underline{H} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Minimum Distance of a Code:

*) Def: Hamming distance between any two codewords $\underline{c}_i, \underline{c}_j$ is the number of components at which the two codewords differ. Denoted by $d(\underline{c}_i, \underline{c}_j)$.

e.g. $\left.\begin{array}{l} \underline{c}_i = [1\ 0\ 1\ 1\ 0\ 1\ 1] \\ \underline{c}_j = [0\ 1\ 0\ 0\ 1\ 0\ 1] \end{array}\right\} \Rightarrow d(\underline{c}_i, \underline{c}_j) = 6.$

*) Def. Hamming weight of a codeword $\underline{c}_i$ is the number of 1's in the codeword. Denoted by $w(\underline{c}_i)$.

*) Def. Minimum distance of a code:

$$d_{min} = \min_{\substack{\underline{c}_i, \underline{c}_j \\ i \neq j}} d(\underline{c}_i, \underline{c}_j)$$

(min. Hamming distance between any two codewords)

*) Def. Minimum weight of a code:

$$w_{min} = \min_{\underline{c}_i \neq \underline{0}} w(\underline{c}_i).$$

*) In any linear block code $d_{min} = w_{min}$.
(needs proof, see the textbook, p. 696)

*) d_{min} is the smallest number of columns of $\underline{H}$ that add to zero.

[to see this: recall that $\underline{c} \cdot \underline{H}^t = \underline{0}$, i.e., if $\underline{c} = [c_1\ c_2 \cdots c_n]$ is a codeword & $\underline{H} = [\underline{h}_1\ \underline{h}_2 \cdots \underline{h}_n]$, we have

$$[c_1\ c_2 \cdots c_n]\,[\underline{h}_1 \cdots \underline{h}_n]^t = 0$$

$$\Rightarrow \sum_{i=1}^{n} c_i \underline{h}_i^t = \underline{0} \Rightarrow \sum_{i=1}^{n} c_i \underline{h}_i = \underline{0}.]$$

Hamming Codes:

*) A class of linear block codes with

$n = 2^m - 1$

$k = 2^m - m - 1$

where $m \geqslant 3$ is an integer.

i.e., (7,4), (15,11), (31,26), (63,57) ... Hamming codes exist.

$$\text{Code rate} = \frac{k}{n} = \frac{2^m - m - 1}{2^m - 1}.$$

*) Can correct exactly <u>one</u> bit error.

*) Parity check matrix consists of all $2^m - 1$ nonzero m-tuples.

ex) $m=3$, $n=7$, $k=4$, (7,4) Hamming code.

$$\underline{H} = \left[\begin{array}{cccc|ccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \quad \& \quad \underline{G} = \left[\begin{array}{c|ccc} & 0 & 1 & 1 \\ I_4 & 1 & 0 & 1 \\ & 1 & 1 & 0 \\ & 1 & 1 & 1 \end{array}\right]$$

(this is an example of a systematic Hamming code).

Decoding of Linear Block Codes:

Consider transmission of a codeword over an AWGN channel (assume that the coded bits are modulated using BPSK). I.e., $\underline{c}_i = [c_{i1}, c_{i2} \cdots c_{in}]$, c_{ij}'s are BPSK modulated. (i.e., $0 \to -1$, $1 \to +1$). Received
(or, $c_{ij} \to 2c_{ij}-1$)

signal: $y_i = (2c_{ij}-1)\cdot\sqrt{E_s} + \eta_i$ $\quad i=1,2,\dots,n.$

E_s: energy per coded bit.

$E_b = \frac{E_s}{R_c} = \frac{n}{k} E_s$ energy per ~~coded~~ info. bit.

$\eta_i \sim \mathcal{N}(0, \frac{N_0}{2})$. i.i.d. noise samples.

Soft decision decoding:

Given $\underline{y} = [y_1\ y_2 \cdots y_n]$, what is the most likely codeword?

(ML decoding)

Solution: Minimize the squared Euclidean distance between the received sequence & the BPSK modulated versions of codewords. I.e.,

$$\hat{\underline{c}}_{opt} = \underset{\underline{c}_i}{\arg\min} \sum_{j=1}^{n} \left(y_j - \sqrt{E_s}(2c_{ij}-1)\right)^2 .$$

*) Soft decision decoding is usually very difficult for linear block codes, and a more frequently used decoding scheme is "hard decision decoding."

Hard Decision Decoding:

*) **Idea:** Map the components of the received signal $\underline{y} = [y_1\ y_2 \cdots y_n]$ to 0's & 1's (i.e., find $\hat{\underline{c}} = [\hat{c}_1, \ldots, \hat{c}_n]$ s.t. $\hat{c}_j = 0$ if $y_j < 0$ & $\hat{c}_j = 1$ if $y_j > 0$),

& then find the codeword which is closest to $\hat{\underline{c}}$ in the Hamming distance sense. That is,

$$\hat{\underline{c}}_{opt} = \underset{\underline{c}_i}{\arg\min}\ d(\underline{c}_i, \hat{\underline{c}}).$$

*) For a general block code (not necessarily linear), this is a very hard problem. We need to search over all codewords (there are 2^k of them).

*) The decoding problem is simplified for the case of linear block codes. We can use the standard array or syndrome table decoding.

*) Assume that the codeword $\underline{c}$ is transmitted, and the binary n-tuple $\hat{\underline{c}}$ is received.

Error vector: $\underline{e} = \underline{c} + \hat{\underline{c}}$, $\boxed{\hat{\underline{c}} = \underline{c} + \underline{e}}$

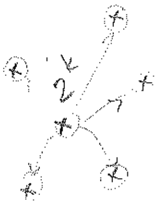
(binary componentwise addition).

* Syndrome of $\hat{\underline{c}}$: $\boxed{\underset{1\times(n-k)}{\underline{s}} = \underset{1\times n}{\hat{\underline{c}}} \cdot \underset{n\times(n-k)}{\underline{H}^t}}$ (1 × n-k vector of 0's & 1's.)

•) $\underline{s} = (\underline{c} + \underline{e})\,\underline{H}^t = \underline{e}\,\underline{H}^t$ (syndrome depends only on the error pattern)

•) $\underline{s} = \underline{0}$ if and only if $\hat{\underline{c}}$ is a codeword.

•) If $\underline{s} \neq \underline{0}$, presence of errors in the transmission has been detected.



•) Clearly, there are $2^k - 1$ itself undetectable error patterns

•) $d(\hat{\underline{c}}, \underline{c}) = d(\underline{c}+\underline{e}, \underline{c}) = d(\underline{e}, \underline{0}) = w(\underline{e})$

•) There are 2^k error patterns that result in the same syndrome. $\boxed{\underset{1\times(n-k)}{\underline{s}} = \underset{?}{\underline{e}}\, \underset{n\times(n-k)}{H^T}} \longrightarrow$ n unknown, n-k equations

Standard array:

*) 2^k codewords are placed on a row (with $\underline{c}_1=\underline{0}$ as the leftmost element).

*) A new row is obtained by choosing an unused n-tuple $\underline{e}$ with the lowest number of 1's by placing $\underline{e}+\underline{c}_i$ under $\underline{c}_i$.

*) Continue until all n-tuples are covered.

That is, (let $M=2^k$)

$\underline{c}_1=\underline{0}$	$\underline{c}_2$	$\underline{c}_3$	......	$\underline{c}_M$
$\underline{e}_1$	$\underline{e}_1+\underline{c}_2$	$\underline{e}_1+\underline{c}_3$	.....	$\underline{e}_1+\underline{c}_M$
$\underline{e}_2$	$\underline{e}_2+\underline{c}_2$	$\underline{e}_2+\underline{c}_3$	...	$\underline{e}_2+\underline{c}_M$
$\vdots$				
$\underline{e}_{2^{n-k}-1}$	$\underline{e}_{2^{n-k}-1}+\underline{c}_2$	$\underline{e}_{2^{n-k}-1}+\underline{c}_3$	...	$\underline{e}_{2^{n-k}-1}+\underline{c}_M$

ex) (3,1) repetition code

0 → 000

1 → 111.

Standard array:

000	111
001	110
010	101
100	011

*) 2^{n-k} rows of the standard array are called the cosets of the code. The first n-tuple in each row is the coset leader.

Facts:

*) No two n-tuples in the same row are identical.

*) Every n-tuple appears in one and only one row.

*) If $\underline{e}$ & $\underline{e}'$ have the same syndrome then they differ by a nonzero codeword. Thus, two n-tuples have the same syndrome if and only if they are in the same coset (row).

Decoding based on the standard array:

Find $\underline{\hat{c}}$ in the standard array. The coset leader is the most likely error pattern & the column header is the maximum likelihood codeword.

ML decoding:

- Calculate syndrome : $\underline{s} = \underline{\hat{c}} H^T$
- Find $\underline{e}$ s.t. $\underline{e} H^T = \underline{s}$ ⇝ multiple solutions
- Choose $\underline{e}$ with minimum $w(\underline{e})$.
- Add $\underline{e}$ to $\underline{\hat{c}}$.

Syndrome Table Decoding:

*) Compute the syndrome of the received vector $\underline{\hat{c}}$, i.e., $\underline{s} = \underline{\hat{c}} \cdot \underline{H}^t$, look up the error pattern with the same syndrome & find $\underline{\hat{c}} + \underline{\hat{e}}$ to find the most likely codeword.

$$\boxed{\underline{s} = \underset{1\times 3}{\underline{\hat{c}}}\, \underset{3\times 2}{H^T}} \rightarrow 1\times 2$$

ex) (3,1) repetition code.

$$\underline{G} = [1\ 1\ 1]_{1\times 3} \qquad \underline{H} = \left[\begin{array}{c:cc} 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}\right]_{2\times 3}$$

error pattern	syndrome
0 0 0	0 0
0 0 1	0 1
0 1 0	1 0
1 0 0	1 1

if $\underline{\hat{c}} = [1\ 1\ 0]$

⇒ $\underline{s} = [0\ 1]$

⇒ $\underline{e} = [0\ 0\ 1]$

⇒ [1 1 1] is the most likely codeword.

$$\underset{\substack{\downarrow \\ 1\times 3}}{\underline{s} = \underline{\hat{c}}}\; \underset{\substack{\downarrow \\ 3\times 2}}{H^T} \longrightarrow 1\times 2 \rightsquigarrow 2^2 = 4 \text{ syndrome:}$$

$(k=1 \quad n=3)$

Codewords:

$\underline{c} = \underset{\substack{\downarrow \\ 0 \\ 1}}{\underline{x}}\, G$ ←

Syndrome Table:

$\underline{s}$	$\overset{0}{\underline{c_1}}+\underline{e}$	$\underline{c_2}+\underline{e}$
00	000	111
01	$\boxed{001}$	110
10	010	101
11	100	011

↑ Syndrome (coset) leader

$$\underline{e}H^T = \underline{s} \longrightarrow [e_1\ e_2\ e_3]\begin{bmatrix}1 & 1\\ 1 & 0\\ 0 & 1\end{bmatrix} = [0\quad 0] \rightsquigarrow \begin{aligned} e_1 \oplus e_2 &= 0 \\ e_1 \oplus e_3 &= 0 \end{aligned}$$

$\underline{e} = [1\ 1\ 1]$

$\underline{e} = [0\ 0\ 0]$ ← min. weight ✓

$$\rightarrow [e_1\ e_2\ e_3]\begin{bmatrix}1 & 1\\ 1 & 0\\ 0 & 1\end{bmatrix} = [0\quad 1] \rightsquigarrow \begin{aligned} e_1 \oplus e_2 &= 0 \\ e_1 \oplus e_3 &= 1 \end{aligned}$$

$\underline{e} = [1\ 1\ 0]$

$\underline{e} = [0\ 0\ 1]$ ← min. weight,

exercise ⤳ Similarly for $\underline{s} = [1\ 0]$ & $\underline{s} = [1\ 1]$

e.g. $\underline{\hat{c}} = [1\ 1\ 0]$

$$\underline{s} = \underline{\hat{c}}H^T = [1\ 1\ 0]\begin{bmatrix}1 & 1\\ 1 & 0\\ 0 & 1\end{bmatrix} = [0\ 1] \xrightarrow[\text{table}]{\text{Syndrome}} \underline{e} = [0\ 0\ 1]$$

So, $\underline{\hat{c}} \oplus \underline{e} = \boxed{[1\ 1\ 1]}$

Syndrome Table:

$\underline{s}$	$\longleftarrow$		2^k		$\longrightarrow$	
$\underline{0}$	$\underline{0}$	$\underline{c}_2$	. . .	$\underline{c}_{2^k}$		
$\underline{s}_2$	$\underline{e}_2$	$\underline{e}_2+\underline{c}_2$	. . .	$\underline{e}_2+\underline{c}_{2^k}$		
⋮	⋮	⋮		⋮		2^{n-k}
$\underline{s}$	$\underline{e}$	$\underline{e}+\underline{c}_2$	. . .	$\underline{e}+\underline{c}_{2^k}$		
⋮	⋮	⋮		⋮		
$\underline{s}_{2^{n-k}}$	$\underline{e}_{2^{n-k}}$	$\underline{e}_{2^{n-k}}+\underline{c}_2$	. . .	$\underline{e}_{2^{n-k}}+\underline{c}_{2^k}$		

↑ Syndrome (coset) leaders

Ex. If $G=\begin{bmatrix}1&1&0&1&0\\1&0&1&0&1\end{bmatrix}$ and $H=\begin{bmatrix}1&0&0&1&1\\0&1&0&1&0\\0&0&1&0&1\end{bmatrix}$,

obtain syndrome table and decode $\hat{\underline{c}}=[1\ 1\ 1\ 0\ 1]$.

$2^2=4$ codewords $2^{5-2}=8$ syndromes

$\underline{c}=\underline{x}\,G \rightsquigarrow \underline{c}_1=[0\ 0\ 0\ 0\ 0]$ $\underline{c}_2=[1\ 0\ 1\ 0\ 1]$ $\underline{c}_3=[1\ 1\ 0\ 1\ 0]$ $\underline{c}_4=[0\ 1\ 1\ 1\ 1]$.

Syndrome Table:

$\underline{s}$	$\underline{e}+\overset{0}{\underline{c}_1}$	$\underline{e}+\underline{c}_2$	$\underline{e}+\underline{c}_3$	$\underline{e}+\underline{c}_4$
000	00000	10101	11010	01111
001	00100	10001	11110	01011
$\boxed{010}$	01000	$\boxed{11101}$	10010	00111
011	01100	11001	10110	00011
100	10000	00101	01010	11111
101	00001	10100	11011	01110
110	00010	10111	11000	01101
111	01001 or, 00110	11100	10011	00110

e.g. find $\underline{e}$ for $\underline{s}=[1\ 1\ 1]$:

$$\overset{\underline{e}}{[e_1\ e_2\ e_3\ e_4\ e_5]}\overset{H^T}{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\1&1&0\\1&0&1\end{bmatrix}}=\overset{\underline{s}}{[1\ 1\ 1]}$$

$$\left.\begin{aligned}e_1\oplus e_4\oplus e_5&=1\\ e_2\oplus e_4&=1\\ e_3\oplus e_5&=1\end{aligned}\right\}\begin{aligned}&\boxed{\begin{aligned}\underline{e}&=[0\ 1\ 0\ 0\ 1]\\ \underline{e}&=[0\ 0\ 1\ 1\ 0]\end{aligned}}\ \text{min. weight}\\ &\underline{e}=[1\ 1\ 1\ 0\ 0]\\ &\quad\vdots\end{aligned}$$

$\hat{\underline{c}}=[1\ 1\ 1\ 0\ 1] \Rightarrow \underline{s}=\hat{\underline{c}}H^T=[0\ 1\ 0] \rightsquigarrow \underline{e}=[0\ 1\ 0\ 0\ 0]$

$$\hat{\underline{c}}\oplus\underline{e}=[1\ 0\ 1\ 0\ 1]=\underline{c}_2$$

From H, $d_{min}=3$.

Or, $d_{min}=\min\limits_{\substack{\underline{c}_i\\ \underline{c}_i\neq\underline{0}}} w(\underline{c}_i)=3$

} can detect all errors up to 2 bits, can correct all errors of 1 bit.

Error Correction & Detection Capabilities:

*) A block code with min. distance d_{min} is guaranteed to detect all error patterns of $d_{min}-1$ or fewer 1's.

*) An (n,k) linear block code can detect exactly 2^n-2^k error patterns. There are 2^k-1 undetectable error patterns.

$$\left(2^n-\overset{\underline{e}=000..}{\overset{\nearrow}{1}}\right)-\left(2^k-1\right)$$
$$=2^n-2^k$$

$\underline{c}_i \to \underline{c}_j$

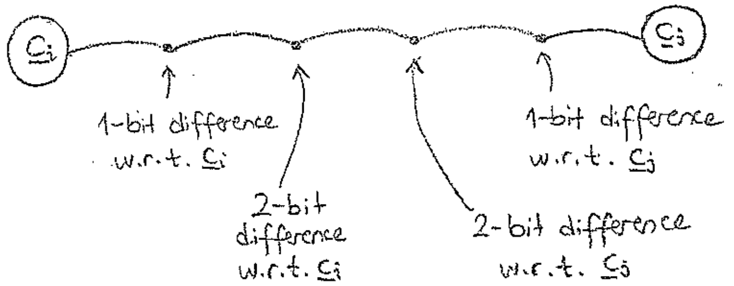
*) If a linear block code is used for random error correction, then all error patterns with t or fewer "1"s with $2t+1 \le d_{min} \le 2t+2$ can be corrected. I.e., $\boxed{t=\left\lfloor \frac{d_{min}-1}{2} \right\rfloor}$.

*) There exists an error pattern with $\boxed{t+1}$ 1's that cannot be corrected.

*) A total of $2^{n-k}-1$ non-zero error patterns can be corrected (coset leaders in syndrome table decoding).

Appendix

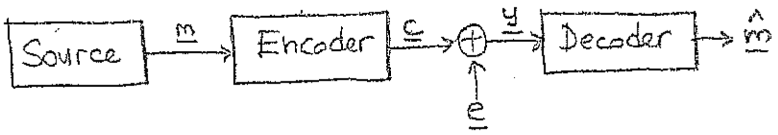
e.g. $d_{min} = d(\underline{c}_i, \underline{c}_j) = 5$



2-bit errors can be corrected by selecting the closest codeword.

- **Error Correction Rule:** Given a vector $\underline{y}$, decode it into the codeword $\underline{c}$ s.t. $d(\underline{y}, \underline{c})$ is minimum over all codewords.

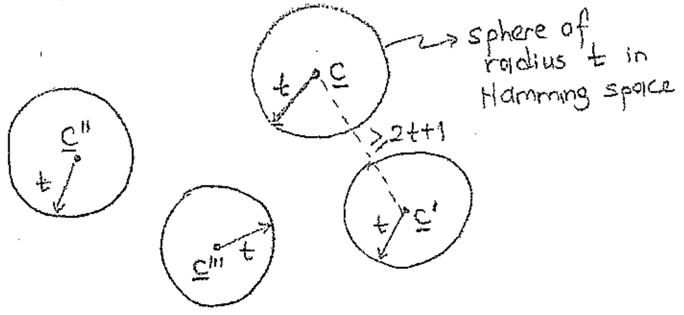
  ⇒ Minimum distance decoding.



Replace $\underline{y}$ with $\hat{\underline{c}}$ on this page.

$\underline{y} \longleftrightarrow \hat{\underline{c}}$

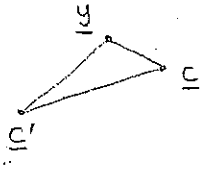
Hamming space



Spheres do not intersect. So errors with $\leq t$ can be corrected.

- If $d(\underline{c}, \underline{y}) \leq t$, then $d(\underline{y}, \underline{c}') > t \quad \forall \underline{c}' \neq \underline{c}$. ← for a code with $d_{min} = 2t+1$.

Proof: Triangle inequality: $d(\underline{c}, \underline{y}) + d(\underline{y}, \underline{c}') \geq d(\underline{c}, \underline{c}') \geq 2t+1$



$$d(\underline{y}, \underline{c}') \geq 2t+1 - \underbrace{d(\underline{c}, \underline{y})}_{\leq t} \geq t+1 \quad ✓$$

- For a linear code, if $\underline{c}$ and $\underline{c}'$ are codewords, then $\underline{c} - \underline{c}'$ is also a codeword.

  Proof: $\underline{c} H^T = \underline{0} \quad \underline{c}' H^T = \underline{0} \Rightarrow (\underline{c} - \underline{c}') H^T = \underline{0}$ ✓