Linear Block Codes (1)
Introduction

Pierangelo Migliorati

DII - University of Brescia

Pioneers of Coding Theory

Bell Telephone Laboratories

Richard Hamming

Claude Shannon

Shannon basic ideas …

- If R < C it is possible to think to P(E) close to zero; if R > C the performances are unacceptable (!!! not completely intuitive !!!)

- To reach good performances, we need:
  - N, the dimension of the signal vector space, very very big (in principle approaching to infinity)
  - SOFT decision (for the decision at the receiver, we have to use real (soft) values and not quantized (hard) values)

- Shannon chose the signals randomly and independently each other, and evaluated the average P(E) over all codes …
  - He was not anyway able to find a real implementable system with the promised dramatic performance !!!

- From 1948 to 1993, to now, the researchers tried to find the promised optimal system
  - 1993: introduction of Turbo codes …
  - After few years, the LDPC codes (already introduced by Gallager in 1962) have been reconsidered, discovering that their performances are very good (better than turbo codes !!!)

Traditional modulation

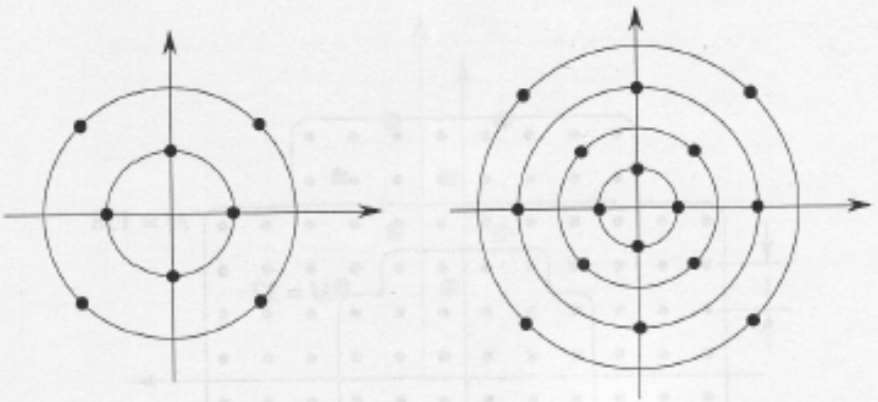
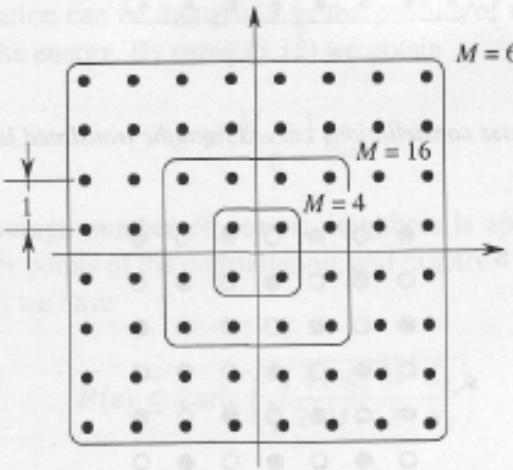


Figure 5.10: *Two QAM constellations with* $M = 8$ *and* $M = 16$.



Classic codes

- Using the classic modulation systems, the performance stay far from the limits
  - to increase d, we need more Power or more BT, or both …
  - increasing N the system get to much complicated …

- To improve P(E) we need a concatenation (in the time domain) of many symbols and we have to introduce some “rules” (i.e., a code) able to increase d_min
  - Clearly the decoding system should be physically realizable (i.e., practically feasible), with a computational complexity possibly linear with N …

- Classical solution (also named Forward Error Correction, FEC)
  - Linear Block Codes (also named Parity Check Codes, Cyclic Redundancy Codes (CRC), …)
  - Convolutional Codes (also named Trellis Codes, Tree Codes, …)

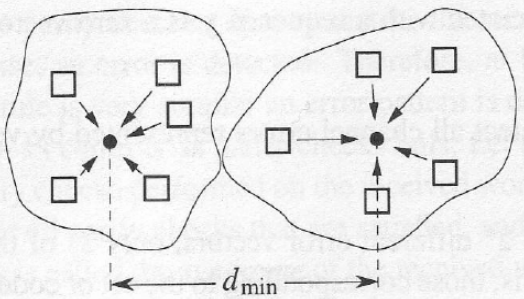
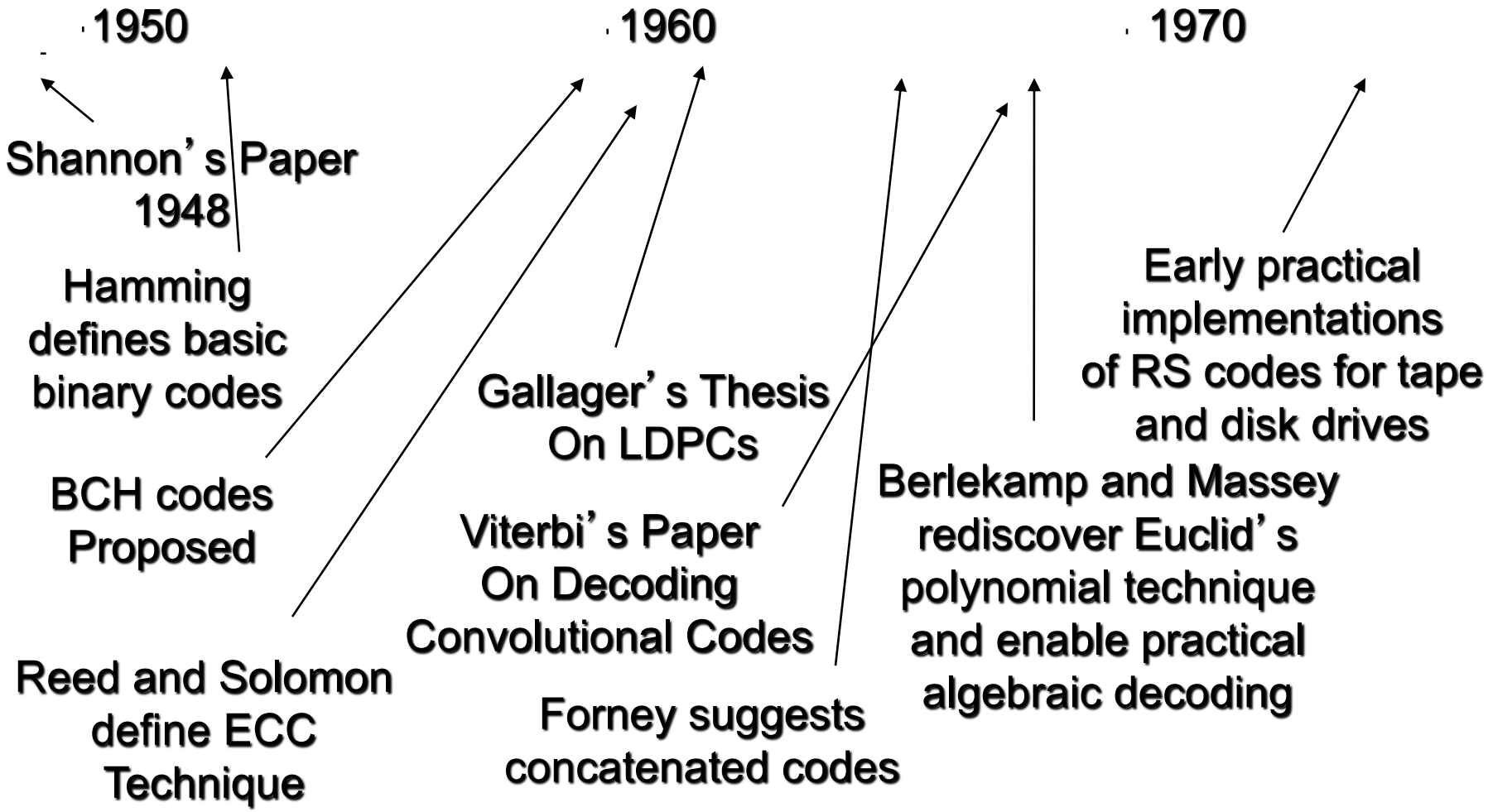
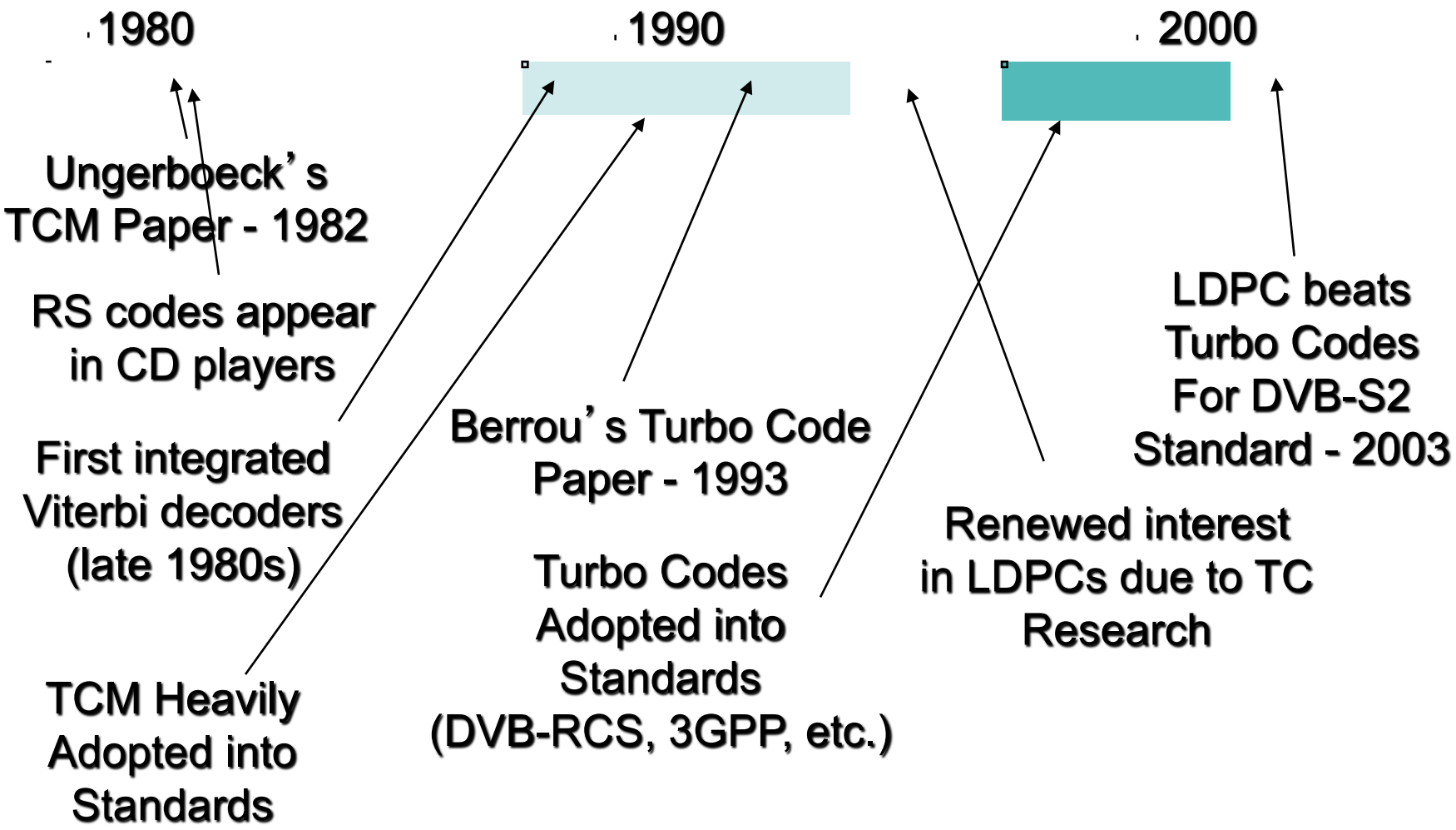


Figure 10.8: *Qualitative representation of the decision regions assigned to code words.*

Forward Error Correction (FEC) Historical Pedigree



FEC Historical Pedigree II



An Old Proverb

- Almost all codes are good codes except for the codes which are decodable.
  - good code : codes yields near Shannon limit performance.
  - decodable : The code can be decoded with linear time.

Good Codes

- Random
- If we set 1000 bits per word
- 10^{301} , astronomical number
- No way with conventional coding schemes

Basic idea of cannel coding

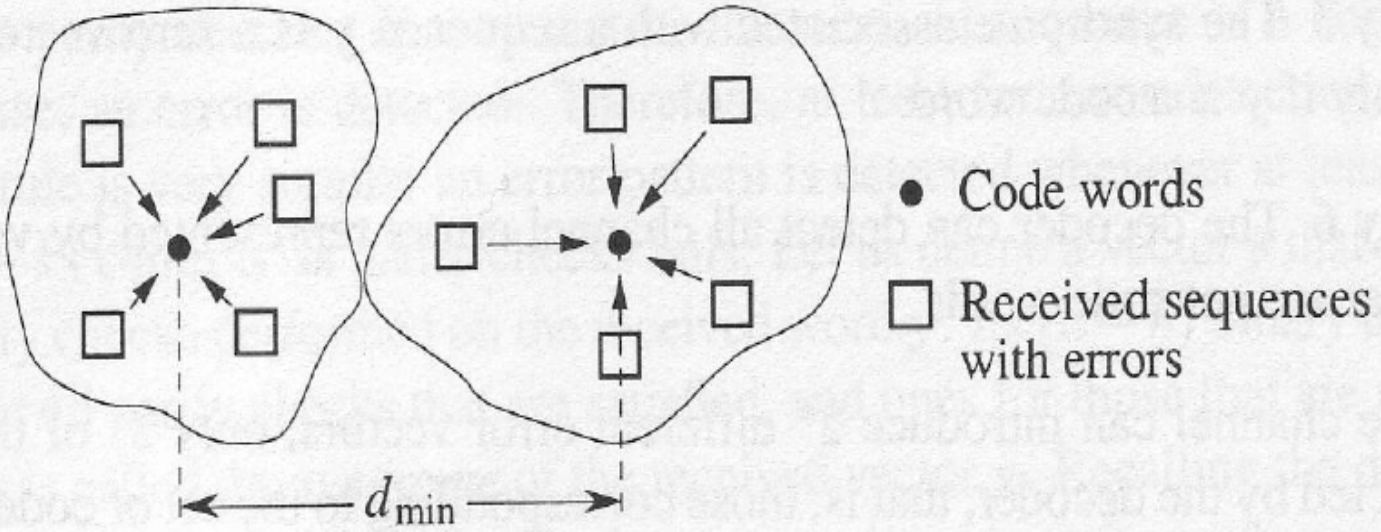
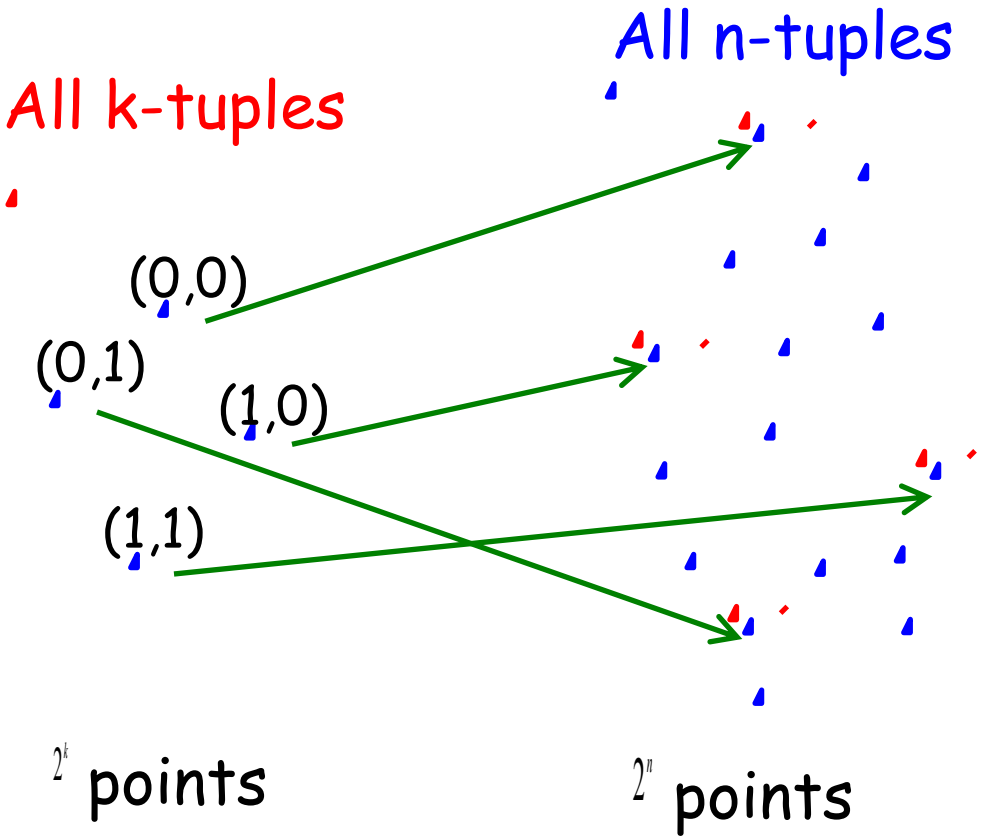


Figure 10.8: *Qualitative representation of the decision regions assigned to code words.*

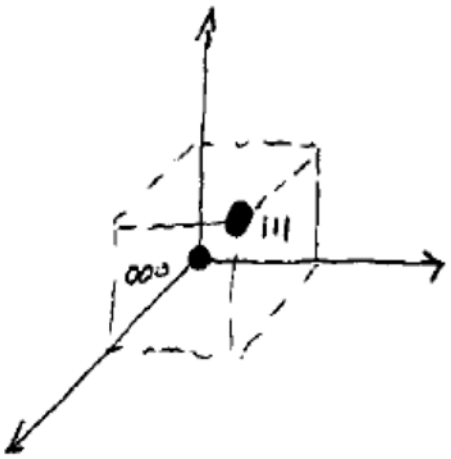
Block Codes

$$(x_1, x_2, \ldots, x_k) \mapsto (y_1, y_2, \ldots, y_n)$$

Valid codewords

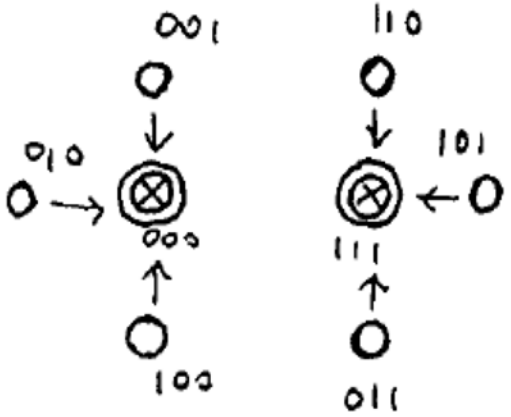


Example (repetition code) ...

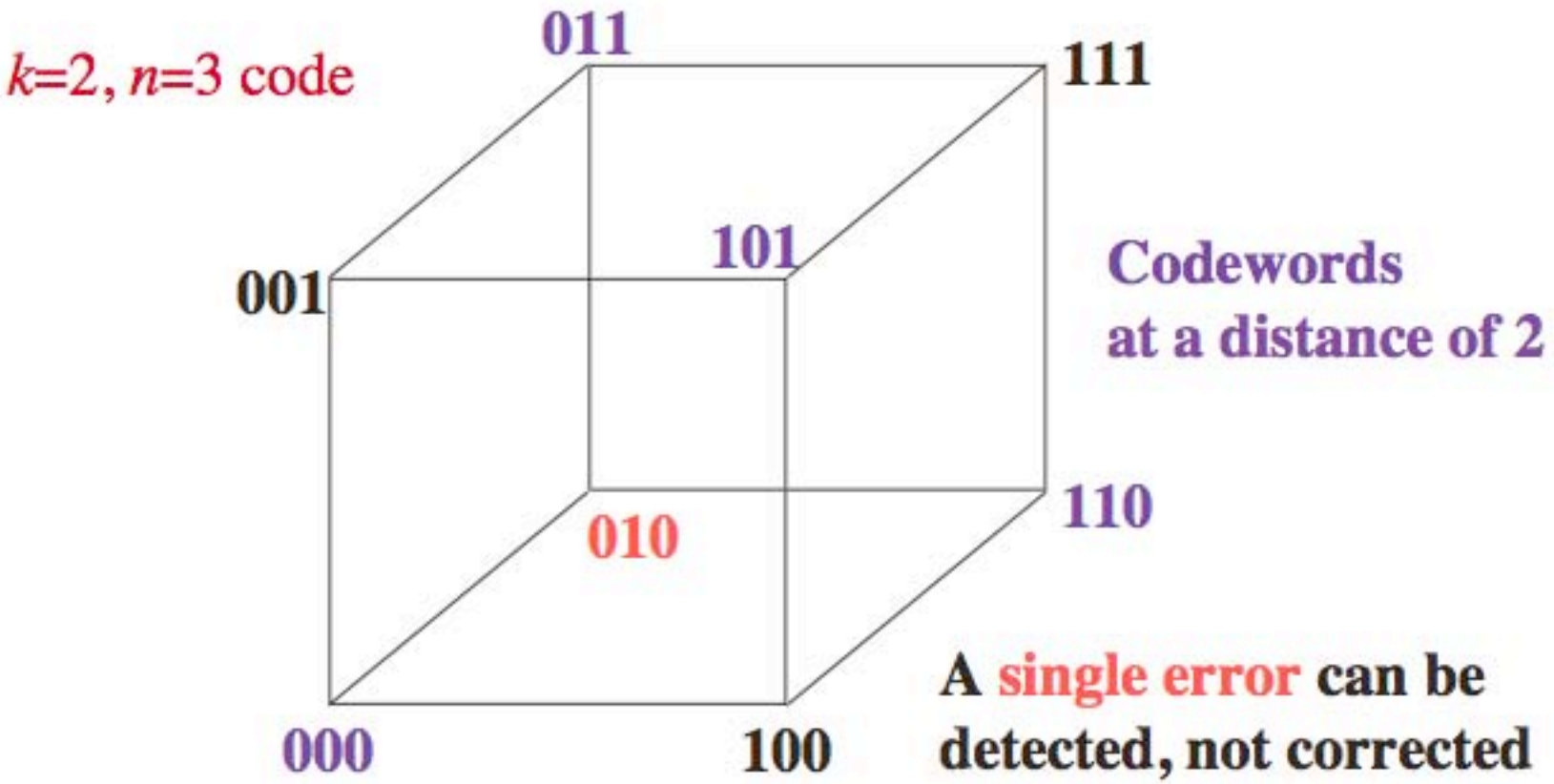


$N = 3 \qquad K = 1$

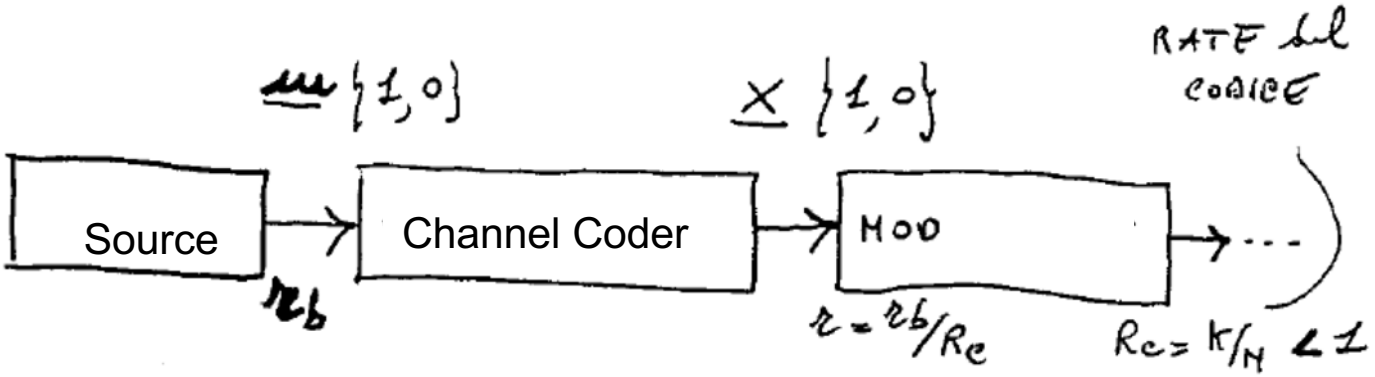
u	x
$0 \rightarrow$	$0\,0\,0$
$1 \rightarrow$	$1\,1\,1$



Example (simple parity check) ...



Basic coding scheme



CLASSI DI CODICI

- CODICI A BLOCCO
- CODICI CONVOLUZIONALI (AD ALBERO)

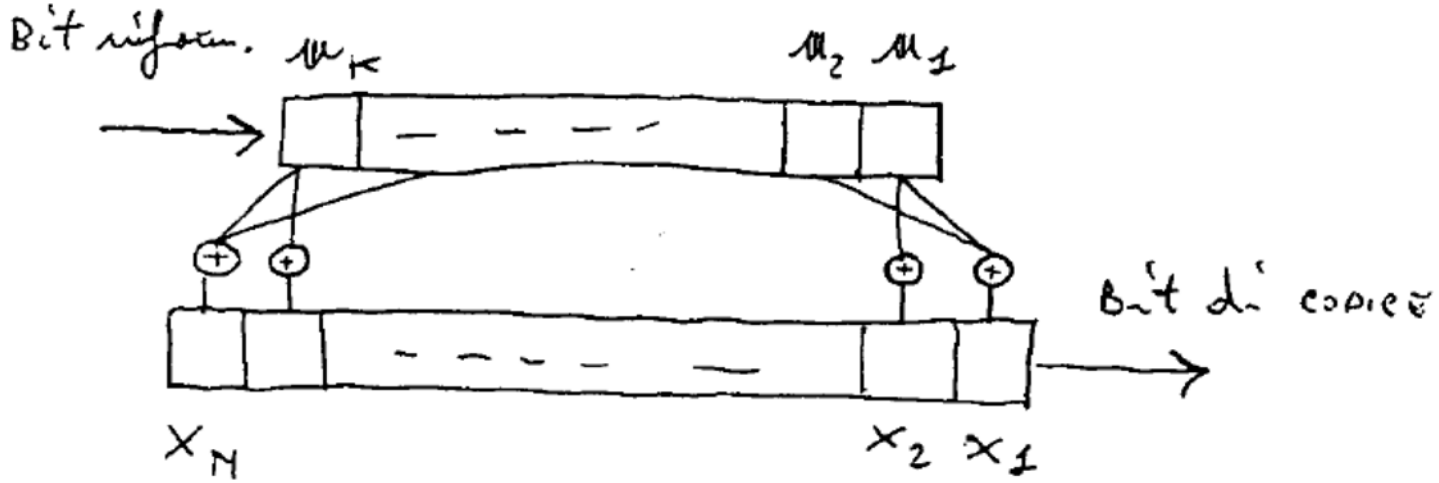
- CODICI "RIVELATORI" DI ERRORE
- CODICI "CORRETTORI" DI ERRORI

$$R_c = \frac{K}{N}$$

Code classes:
Block / Convolutional;
Error Detecting / Error Correcting.

Rc=K/N < 1, code rate

Almost all block codes used today belong to a subset called linear block codes. A linear block code is a code in which the exclusive OR (addition modulo-2) of two valid codewords creates another valid codeword.



Le parole di codice si possono costruire combinando linearmente i bit di informazione (in aritmetica MODULO DUE) (1+1=0)

The code words (x, c, …) are obtained by a linear combination (in algebra modulo 2, ex-or, 1+1=0) of the information bits (m, u, i, …)

The code words (x, c, …) are obtained by a linear combination (in algebra modulo 2, ex-or, 1+1=0) of the information bits (m, u, i, …)

$$x_i = \sum_{1}^{K} g_{ji} \, u_j \qquad g_{ji} = \{1, 0\} \qquad i = 1, \ldots, N$$

Using the matrix notation …
G_(K,N)=Generator matrix

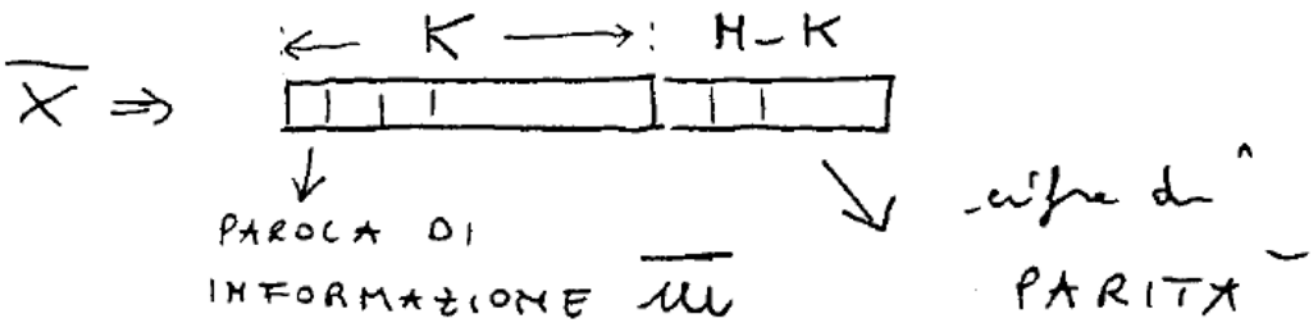
$$\overline{x} = \overline{u}\, G$$

G = matrice "generatrice" del codice (K, N)

$$G = \begin{vmatrix} g_{11} & g_{12} & \cdots & g_{1N} \\ g_{21} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ g_{K1} & \cdots & \cdots & g_{KN} \end{vmatrix}$$

Systematic Codes (also named: Parity Check Codes)

CODICE SISTEMATICO (N, K)



In questo caso la matrice G assume la forma

$$G = | I_K \quad P |$$

In this case G assume this form …
I_k = K order unity matrix,
P = Parity Check Matrix

I_K = matrice UNITA' di ordine K

P = matrice $(K \cdot (N-K))$ di PARITA'

$$G = | I_K \quad P |$$

I_K = matrice UNITA' di ordine K

P = matrice $(K \cdot (N-K))$ di PARITA'

$$\overline{X} = | \overline{m}; \overline{X}_P | \qquad \overline{X}_P = \overline{m} P \qquad \text{"bit di parità"}$$

Esempio

$$-G = \left| \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right| = | I_K \; P |$$

I_K P

$$\overline{X} = \overline{m}\, G$$

Se codice è sistematico se

$$\overline{X} = \overline{m}\, G = |\overline{m}, \overline{m}\, P| \Rightarrow G = |I_K \; P|$$

Ponendo $H^T = \begin{vmatrix} P \\ I_{N-K} \end{vmatrix}$ → Parity Check Matrix

si ha la relazione fondamentale

$$\boxed{\overline{X}\, H^T = 0}$$

→ (N-K) verifiche di parità

si ha infatti:

$$\overline{X}\, H^T = |\overline{m}, \overline{m}\, P| \begin{vmatrix} P \\ I_{N-K} \end{vmatrix} = \overline{m}P + \overline{m}P = 0$$

$$\overline{x}\, H^T = \left| \overline{u}, \overline{u} P \right| \begin{vmatrix} P \\ I_{N-K} \end{vmatrix} = \overline{u} P + \overline{u} P = 0$$

Esempio:

$$G = \left| \begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right| \Rightarrow H^T = \left| \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right|$$

$$\begin{cases} x_2 + x_3 + x_4 + x_5 = 0 \\ x_1 + x_3 + x_4 + x_6 = 0 \\ x_1 + x_2 + x_4 + x_7 = 0 \end{cases}$$

Linear Block Codes

- the number of codeworde is 2^k since there are 2^k distinct messages.
- The set of vectors $\{g_i\}$ are linearly independent since we must have a set of unique codewords.
- linearly independent vectors mean that no vector g_i can be expressed as a linear combination of the other vectors.
- These vectors are called baises vectors of the vector space C.
- The dimension of this vector space is the number of the basis vector which are *k*.
- $G_i \in C \rightarrow$ the rows of G are all legal codewords.

The Hamming code (7, 4) is defined by the relations

$$\begin{aligned} x_i &= u_i\,, \qquad i = 1, 2, 3, 4 \\ x_5 &= u_1 + u_2 + u_3 \\ x_6 &= u_2 + u_3 + u_4 \\ x_7 &= u_1 + u_2 + u_4\,. \end{aligned} \tag{10.3}$$

The corresponding systematic encoder is shown in Fig. 10.7. It is defined by the correspondence

Data words	Code words
0000	0000 000
0001	0001 011
0010	0010 110
0011	0011 101
0100	0100 111
0101	0101 100
0110	0110 001
0111	0111 010
1000	1000 101
1001	1001 110
1010	1010 011
1011	1011 000
1100	1100 010
1101	1101 001
1110	1110 100
1111	1111 111

Basic Definitions (cont'd)

- Def: The weight of a codeword $\mathbf{c}_i$, denoted by $w(\mathbf{c}_i)$, is the number of of nonzero elements in the codeword.
- Def: The minimum weight of a code, $w_{\min}$, is the smallest weight of the nonzero codewords in the code.
- Theorem: In any linear code, $d_{\min} = w_{\min}$

Any linear block code can be put in systematic form

Linear Block Codes

$$\underline{c} = \underline{i}G \quad , \quad \underline{c}H^T = \underline{0}$$

$\underline{i}$ Vector 1xK : information bits

$\underline{c}$ Vector 1xN : codeword bits

G Matrix KxN : code generator matrix

H Matrix (N-K)xN : parity check matrix (pay attention to transposition)

All codes are can be put in systematic form

$$G = \begin{bmatrix} I_k & P \end{bmatrix} \Rightarrow H^T = \begin{bmatrix} P \\ I_{N-K} \end{bmatrix}$$

$$\Rightarrow \quad \underline{c} = \underbrace{c_1 c_2 \cdots c_K}_{\underline{i} = i_1\ i_2\ i_3 \cdots i_K} c_{K+1} \cdots c_N$$

Some definitions

$d_H(\underline{a},\underline{b})$ Hamming distance: number of different bits in $\underline{a}$ and $\underline{b}$

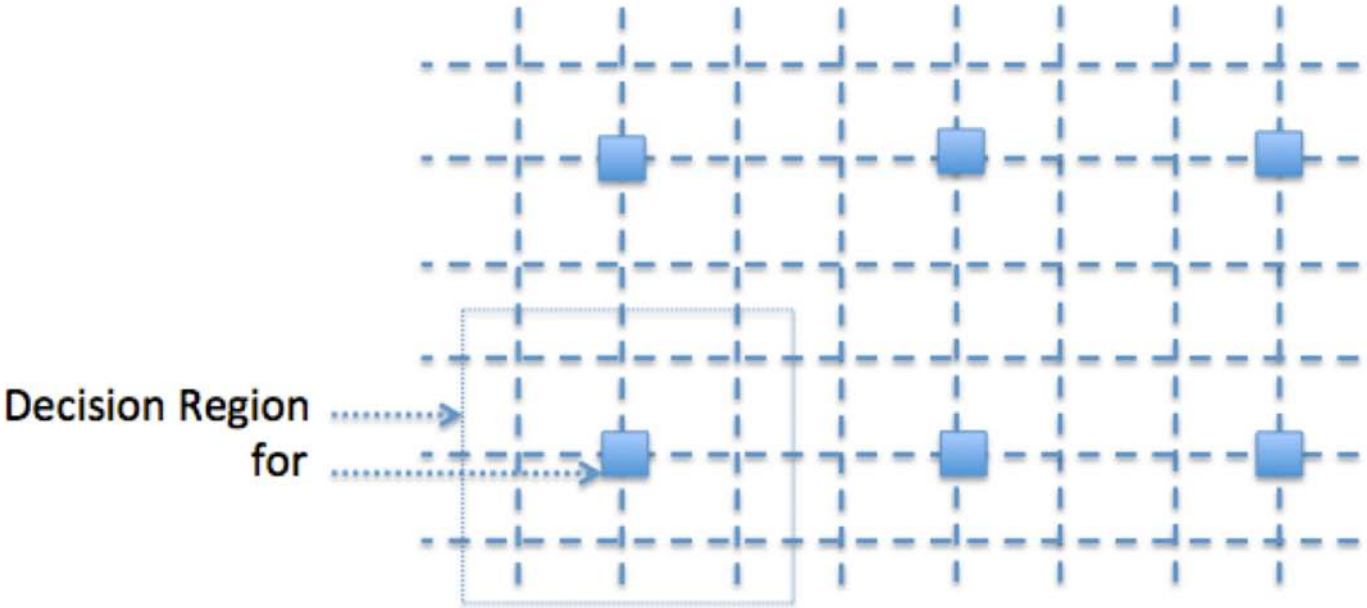
$w_H(\underline{a})$ Hamming weight: number of bits set to `1' in $\underline{a}$

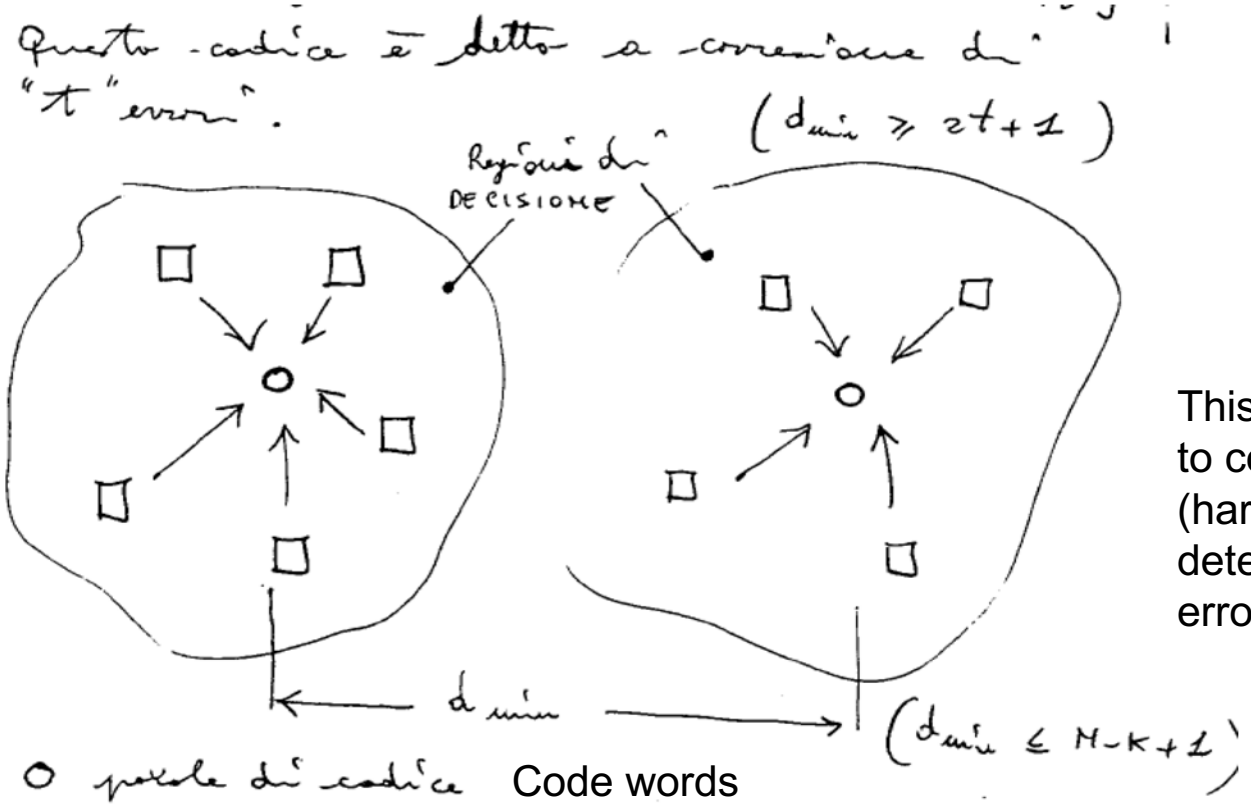
Maximum Likelihood Decoding Strategy (Hard Decision)

$$\Rightarrow \quad \hat{\underline{e}} = \arg\min_{\underline{e}:\, \underline{v}-\underline{e}\in C} w_H(\underline{e}) \qquad \ldots \qquad \text{ma} \qquad w_H(\underline{e}) = d_H(\underline{c},\underline{v})$$

$$\Rightarrow \quad \boxed{\hat{\underline{c}} = \arg\min_{\underline{c}:\, \underline{c}\in C} d_H(\underline{c},\underline{v}) \qquad \text{Minimum distance decoding}}$$

As for the ML decoder, but working with Hamming distances



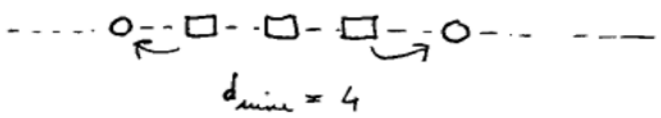


This code is able to correct “t” errors (hard dec.), or to detect d_min-1 errors …

Code words

Received word with possible errors …

$t = \frac{d_{min}-1}{2}$ (d_{min} dispari) ; $\frac{d_{min}}{2}-1$ (d_{min} pari)



Minimum Distance

It is important to find the minimum distance between codewords

$$d^* = \min_{i,j} d_H(\underline{c}_i, \underline{c}_j)$$

Since the code is linear, we can "shift" all codewords by a fixed codeword

$$\begin{aligned} d^* &= \min_{i,j} d_H(\underline{c}_i - \underline{c}_j, \underline{c}_j - \underline{c}_j) \\ &= \min_{h} d_H(\underline{c}_h, \underline{0}) \\ &= \min_{h} w_H(\underline{c}_h) \end{aligned}$$

$\Rightarrow$ The minimum distance is the same as the minimum codeword weight

Let us consider the information word $\underline{i}$=1,0,0,0,0…. for a systematic code
The associated codeword $\underline{c}=\underline{i}G$ contains at the most
- one '1' within the information bits
- *N-K* '1' within the parity bits

$$\Rightarrow \quad d^* \leq N - K + 1$$

Minimum distance

Let us consider the parity check over a vector $\underline{a}$

$$\underline{a}H^T = \underline{0} \quad \Leftrightarrow \quad \underline{a} \in C \quad \Rightarrow \quad w_H(\underline{a}) \geq d^*$$

But the evaluation of $\underline{a}H^T$ is a linear combination of $w_H(\underline{a})$ rows of H^T

$\Rightarrow$ The minimum distance is the minimum number of dependent rows in H^T

Correcting power

A code can correct h errors if and only if $d^* \geq 2h+1$

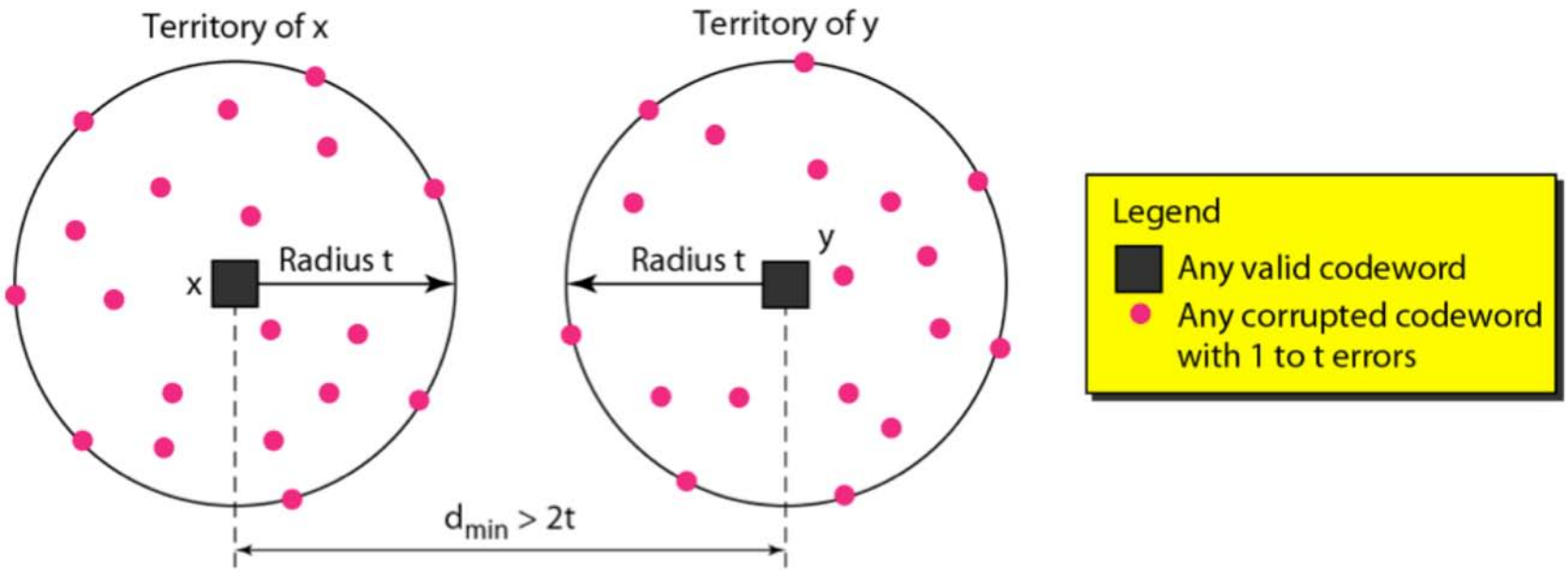
$\Rightarrow$ Correcting power of the code $t = \left\lfloor \frac{d^* - 1}{2} \right\rfloor$

$\Rightarrow$ Every combination of no more than t rows of H^T can be obtained in a unique way

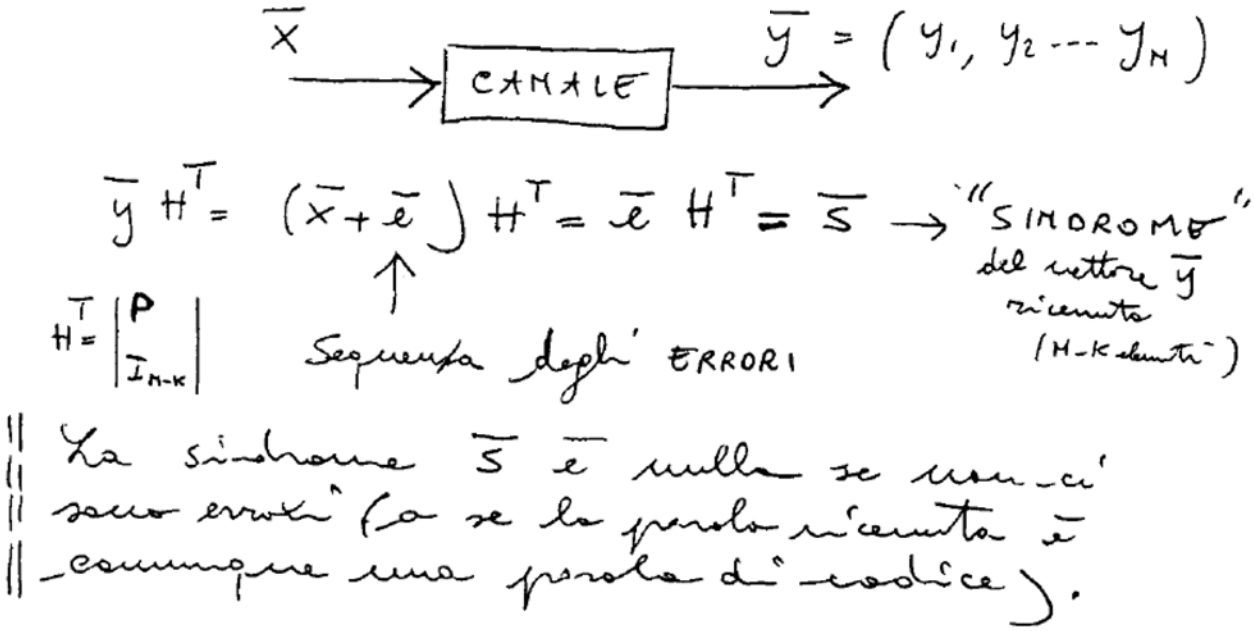
Detecting Power

If a code is not used to correct, it can detect up to d^*-1 errors

Figure 10.9 *Geometric concept for finding d_{min} in error correction*



Syndrome ...

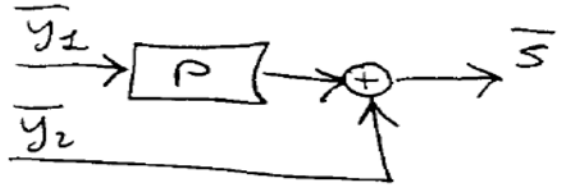


The syndrome (s) of a received word (y) is equal to zero if there are no errors (e), or if the received word is itself a code word …

PAROLA TRASMESSA $\overline{x}$ $| x_1 x_2 \cdots x_K \,\vdots\, x_{K+1} \cdots x_N |$ ($\overline{u}$ | $\overline{up}$)

PAROLA RICEVUTA $\overline{y}$ $| y_1 y_2 \cdots y_K \,\vdots\, y_{K+1} \cdots y_N |$ ($\overline{y}_1$ | $\overline{y}_2$)

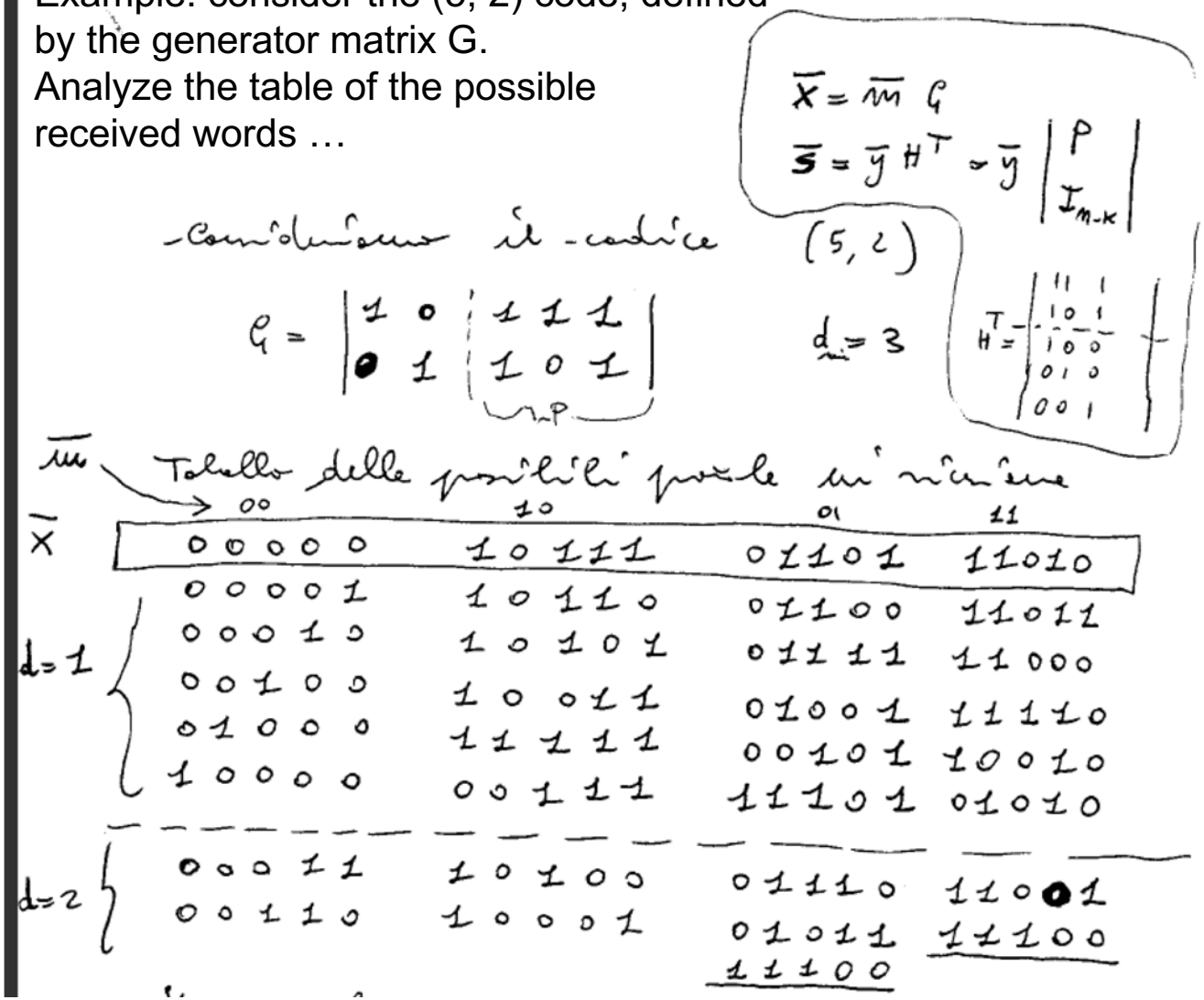
SINDROME $\overline{s}$ $\overline{s} = \overline{y}_1 P + \overline{y}_2$

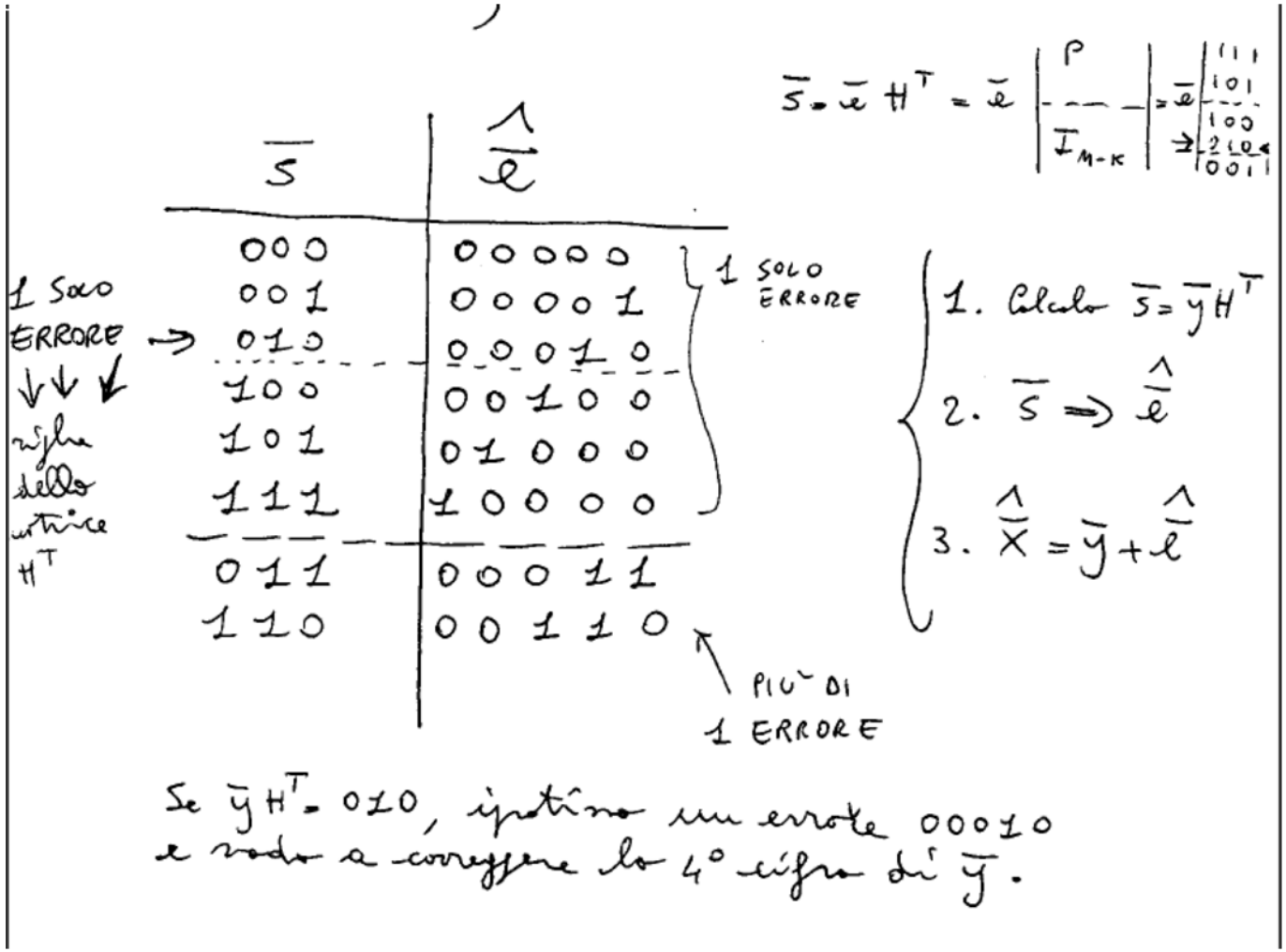


→ Il sistema può "rilevare gli errori che non sono" parole del codice.

The system can “detect” the errors which are not code words !!!
Otherwise s=0 … and nothing can be assumed …

Example: consider the (5, 2) code, defined by the generator matrix G.
Analyze the table of the possible received words …





Assuming a maximum of 1 error, s is one row of H^T.
In this case, if y H^T = 010, we assume the error pattern 00010, and therefore we correct the 4th bit of y.

Summary …

Linear Block Codes

$$\underline{c} = \underline{i}G \quad , \quad \underline{c}H^T = \underline{0}$$

$\underline{i}$	Vector 1xK : information bits
$\underline{c}$	Vector 1xN : codeword bits
G	Matrix KxN : code generator matrix
H	Matrix (N-K)xN : parity check matrix (pay attention to transposition)

All codes are can be put in systematic form

$$G = \begin{bmatrix} I_k & P \end{bmatrix} \Rightarrow H^T = \begin{bmatrix} P \\ I_{N-K} \end{bmatrix}$$

$$\Rightarrow \quad \underline{c} = \underbrace{c_1 c_2 \cdots c_K}_{\underline{i} = i_1\ i_2\ i_3 \cdots i_K} c_{K+1} \cdots c_N$$

Some definitions

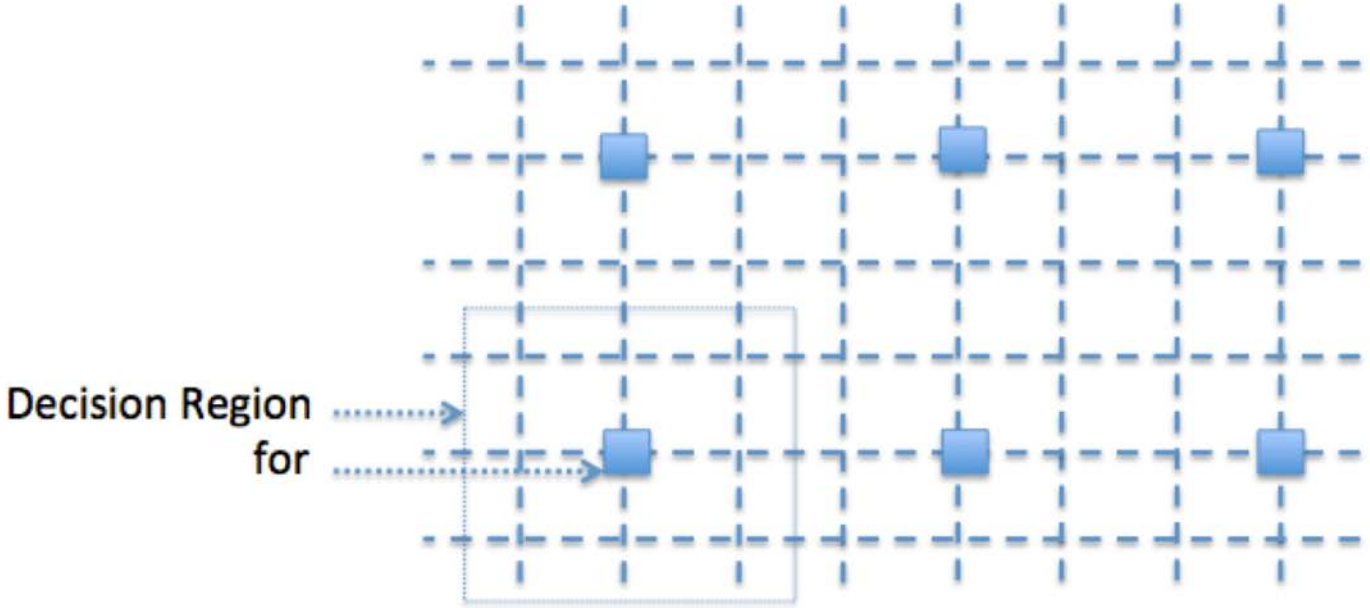
$d_H(\underline{a},\underline{b})$	Hamming distance: number of different bits in $\underline{a}$ and $\underline{b}$
$w_H(\underline{a})$	Hamming weight: number of bits set to `1' in $\underline{a}$

ML Decoding Strategy

$$\Rightarrow \quad \hat{\underline{e}} = \arg\min_{\underline{e}:\, \underline{v}-\underline{e}\in C} w_H(\underline{e}) \qquad \ldots \qquad \text{ma} \qquad w_H(\underline{e}) = d_H(\underline{c},\underline{v})$$

$$\Rightarrow \quad \boxed{\hat{\underline{c}} = \arg\min_{\underline{c}:\, \underline{c}\in C} d_H(\underline{c},\underline{v}) \qquad \text{Minimum distance decoding}}$$

As for the ML decoder, but working with Hamming distances



Minimum Distance

It is important to find the minimum distance between codewords

$$d^* = \min_{i,j} d_H(\underline{c}_i, \underline{c}_j)$$

Since the code is linear, we can "shift" all codewords by a fixed codeword

$$\begin{aligned} d^* &= \min_{i,j} d_H(\underline{c}_i - \underline{c}_j, \underline{c}_j - \underline{c}_j) \\ &= \min_{h} d_H(\underline{c}_h, \underline{0}) \\ &= \min_{h} w_H(\underline{c}_h) \end{aligned}$$

$\Rightarrow$ The minimum distance is the same as the minimum codeword weight

Let us consider the information word $\underline{i}$=1,0,0,0,0…. for a systematic code
The associated codeword $\underline{c}=\underline{i}G$ contains at the most
- one '1' within the information bits
- *N-K* '1' within the parity bits

$$\Rightarrow \quad d^* \leq N - K + 1$$

Minimum distance

Let us consider the parity check over a vector $\underline{a}$

$$\underline{a}H^T = \underline{0} \quad \Leftrightarrow \quad \underline{a} \in C \quad \Rightarrow \quad w_H(\underline{a}) \geq d^*$$

But the evaluation of $\underline{a}H^T$ is a linear combination of $w_H(\underline{a})$ rows of H^T

$\Rightarrow$ The minimum distance is the minimum number of dependent rows in H^T

Correcting power

A code can correct h errors if and only if $d^* \geq 2h+1$

$\Rightarrow$ Correcting power of the code $t = \left\lfloor \frac{d^* - 1}{2} \right\rfloor$

$\Rightarrow$ Every combination of no more than t rows of H^T can be obtained in a unique way

Detecting Power

If a code is not used to correct, it can detect up to d^*-1 errors

Syndrome

Consider the parity check on a received word $\underline{v}$

$$\underline{v} = \underline{c} + \underline{e} \quad \Rightarrow \quad \underline{v}H^T = \left(\underline{c} + \underline{e}\right)H^T$$
$$= \underline{c}H^T + \underline{e}H^T$$
$$= \underline{e}H^T$$

We define the **syndrome** of a received word as

$$\underline{s} = \underline{v}H^T$$

which corresponds thus to the syndrome of the error word

Decoding

The decoding process is an estimate of $\underline{c}$ and thus an estimate of $\underline{e}$

We want to estimate $\underline{e}$ as the minimum weight error word compatible with $\underline{s}$

$$\underline{s} = \underline{e}H^T \quad \rightarrow \quad \hat{\underline{e}} = \underset{\underline{e}:\underline{e}H^T = \underline{s}}{\arg\min}\left\{w_H(\underline{e})\right\}$$

The decoding is thus obtained as $\hat{\underline{c}} = \underline{v} - \hat{\underline{e}} = \underline{v} + \hat{\underline{e}}$

Esempio

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix} \Longrightarrow \quad \underline{c} = \underline{i}G$$

Information words → Codewords

$$\{\underline{i}\} = \begin{Bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{Bmatrix} \Longrightarrow \{\underline{c}\} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{Bmatrix}$$

$$H^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\{\underline{c}\}H^T = \underline{\underline{0}} \Rightarrow \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{Bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{\underline{0}}$$

Minimum distance of the code

$$d^* = \min_i w_H(\underline{c}_i) = 3 \quad \Rightarrow \quad t = 1$$

$\Rightarrow$ Correct 1 error, detects 2 errors

Minimum distance decoding

	$\underline{c}_1$					$\underline{c}_2$					$\underline{c}_3$					$\underline{c}_4$				
	0	0	0	0	0	1	0	1	1	1	0	1	1	0	1	1	1	0	1	0
$d = 1$ Error corrected	0	0	0	0	1	1	0	1	1	0	0	1	1	0	0	1	1	0	1	1
	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1	0	0	0
	0	0	1	0	0	1	0	0	1	1	0	1	0	0	1	1	1	1	1	0
	0	1	0	0	0	1	1	1	1	1	0	0	1	0	1	1	0	0	1	0
	1	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0	1	0	1	0
$d = 2$ Error detected	0	0	0	1	1	1	0	1	0	0	0	1	1	1	0	1	1	0	0	1
	0	0	1	1	0	1	0	0	0	1	0	1	0	1	1	1	1	1	0	0

Decoding

$$\underline{s} = \underline{v}H^T \quad \rightarrow \quad \hat{\underline{e}} = \arg\min_{\underline{e}:\,\underline{e}H^T = \underline{s}} \{w_H(\underline{e})\} \quad \rightarrow \quad \hat{\underline{c}} = \underline{v} + \hat{\underline{e}}$$

Error patterns deduced from the syndrome

	$\underline{s}$	$\hat{\underline{e}}$
One error ⇓ $\underline{s}$ Is a row of H uniquely determined	0 0 0	0 0 0 0 0
	0 0 1	0 0 0 0 1
	0 1 0	0 0 0 1 0
	1 0 0	0 0 1 0 0
	1 0 1	0 1 0 0 0
	1 1 1	1 0 0 0 0
Two errors ⇓	0 1 1	0 0 0 1 1 o 1 0 1 0 0 ?
	1 1 0	0 0 1 1 0 o 1 0 0 0 1 ?

$\underline{s}$ is the sum of è two rows of H
Not uniquely determined!