# ERROR - CORRECTING CODES

MESSAGE SOURCE $\xrightarrow{x = x_1 \dots x_k \text{ message}}$ ENCODER $\xrightarrow{c = c_1 \dots c_m \text{ codeword}}$ CHANNEL $\xrightarrow{y = c + e \text{ received vector}}$ DECODER $\xrightarrow{\hat{x} \text{ estimate of message}}$ RECEIVER

$e = e_1 \dots e_m$ error (noise)

The fundamental problem in coding theory is to determine what msg. was sent on the basis of what is received.

## Linear codes

$\mathbb{F}_q = GF(q)$ [finite field with $q$ elements]

- Let $\mathbb{F}_q^m$ denote the vector space of all m-tuples over $\mathbb{F}_q$

(m, M) code C - over $\mathbb{F}_q$ is a subset of $\mathbb{F}_q^m$ of size M.

$(a_1, a_2 \dots a_m)$ - vector in $\mathbb{F}_q^m$

elements in C are called „codewords".

ex: $\mathbb{F}_2$ has binary codes

- A generator matrix for an [m,k] code C is any k×m matrix G whose rows form a basis for C.

obs: We can have many generator matrices for any code.

- For any set of k independent columns of a generator matrix G, the corresponding set of coordinates forms an information set for C.
- The remaining set $r = m - k$ is called the redundancy of C.

If the first k coordinates form an information set, the code has an unique generator matrix of the form $[I_k | A]$ where $I_k$ is k×k identity matrix

- We can define an $(m-k) \times m$ matrix H (a parity matrix) for the [m,k] code C, defined by

$$C = \{ x \in \mathbb{F}_q^m \mid Hx^T = 0 \}$$

$H = [-A^T | I_{m-k}]$ ([I|A] is the generator matrix for C.

# THE HAMMING DISTANCE → the higher the minimum distance the more errors we can correct.

An important invariant of a code is the minimum distance between codewords.

The Hamming distance $d(x,y)$ between two vectors $x,y \in \mathbb{F}_2^m$ is the nr. of coordinates in which $x$ and $y$ differ.

ex:

$$\left.\begin{matrix}0000\\0001\end{matrix}\right\} 1 \quad (1 \text{ bit differs}) \qquad \left.\begin{matrix}0111\\1000\end{matrix}\right\} 4$$

$$\left.\begin{matrix}0001\\0010\end{matrix}\right\} 2$$

<u>Obs:</u>

The distance $d(x,y)$ satisfies the four properties that make it a metric on $\underset{\text{vect. space}}{\mathbb{F}_2^m}$

(i) $d(x,y) \geq 0 \quad \forall x,y \in \mathbb{F}_2^m$

(ii) $d(x,y) = 0 \iff x = y$

(iii) $d(x,y) = d(y,x)$

(iv) $d(x,z) \leq d(x,y) + d(y,z)$ [triangle inequality]

# Examples

Finding the corrupted bit:

①

| | $2^0$ | $2^1$ | | $2^2$ | | | | $2^3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ? | ? | | ? | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| message | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |

$p_1 =$ ? 0 1 0 0 1 = 0 (two 1's) ✗ (1)

$p_2 =$ ? 0 10 11 = 1 ✓

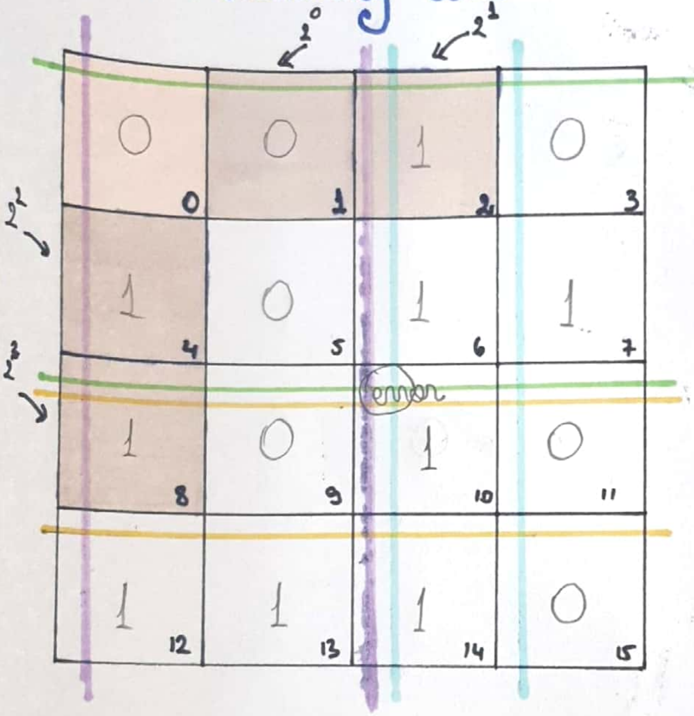
$p_4 =$ 0110 0 = 0 ✓

$p_8 =$ ? 0110 = 0 ✗ (8)

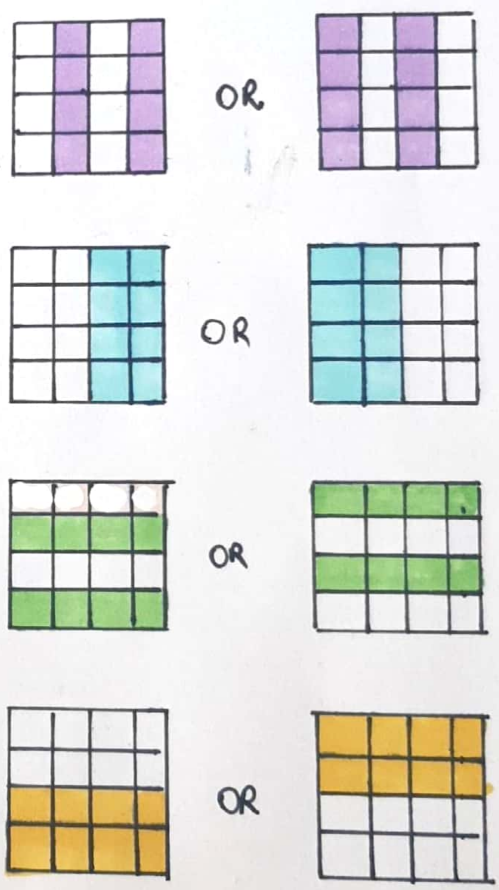
1 + 8 = 9
↳ the corrupted bit

②

(15, 11) Hamming code




Find the corrupted bit.



$p_1$: OR (on pos $2^0$)

$p_2$ OR (on pos $2^1$)

$p_3$ OR (on pos $2^2$)

$p_4$ OR (on pos $2^3$)

$P_1 =$ „2" = 0 ✓

$P_2 =$ „4" ≠ 0 ✗

$P_3 =$ „3" = 1 ✓

$P_4 =$ „4" ≠ 0 ✗