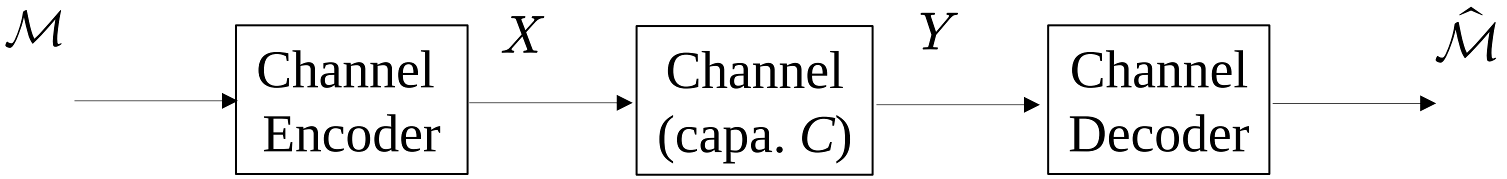
I. Formulation – Assumption – Proof – Practice

II. The Art of Channel Coding

Channel Coding Theorem

I. Formulation

How to achieve channel capacity (C)?



Channel Coding Theorem

I. Formulation

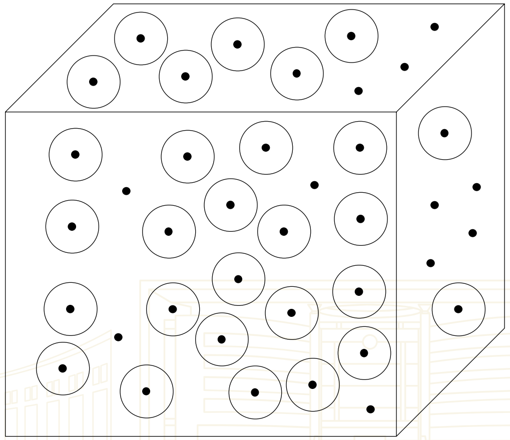
$$\mathcal{M}(m^k) \longrightarrow \boxed{\text{Channel Encoder}} \xrightarrow{X(x^n)} \boxed{\text{Channel (capa. } C)} \xrightarrow{Y(y^n)} \boxed{\text{Channel Decoder}} \longrightarrow \hat{\mathcal{M}}(\hat{m}^k)$$

$$(y^n = x^n + e^n)$$

Channel Coding: Adding redundancy into message can correct errors introduced by the channel

$$\dim(\mathcal{M}) < \dim(X)$$

$$\text{rate} = \frac{\dim(\mathcal{M})}{\dim(X)} = \frac{k}{n} = r$$



$$\{n-\text{dim vector space}\} = \{\text{legal subspace}\} \cup \{\text{illegal subspace}\}$$

$$\{y^n\} = \{x^n\} \cup \{x^n + e^n\}$$

Channel Coding Theorem

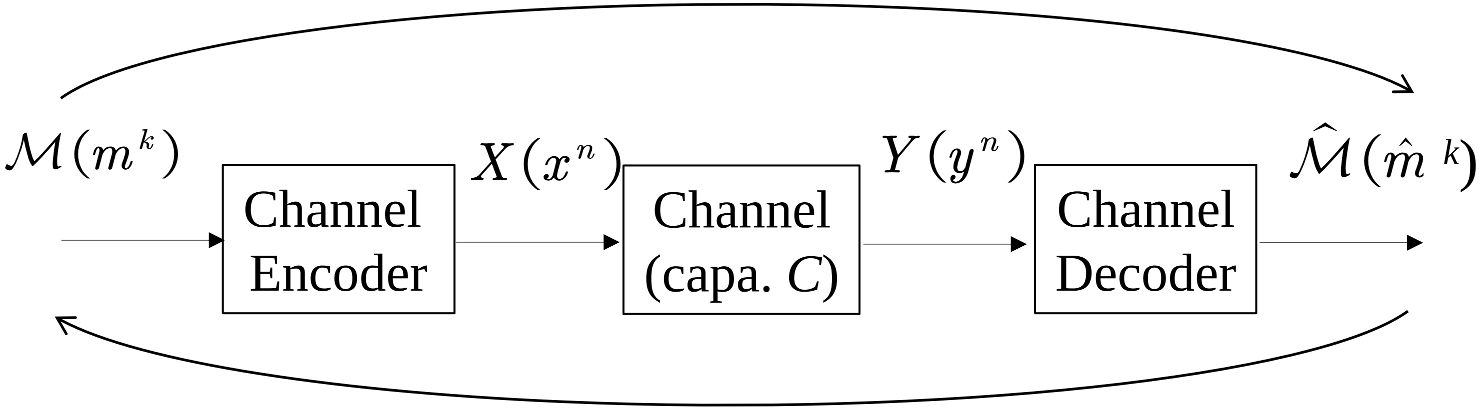
I. Formulation

Shannon's Channel Coding Theorem

$$P_e = P(\hat{m}^k \neq m^k | y^n)$$

Achievability: If $r < C$, $P_e \to 0$.

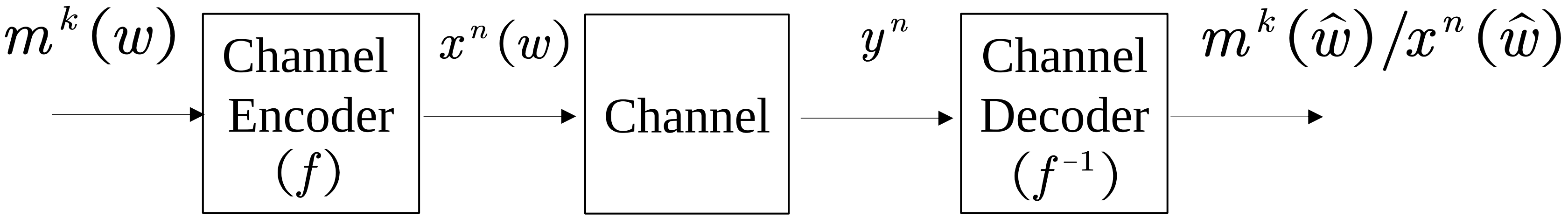
Proof Pivot:
Jointly Typical Sequences (JTS)



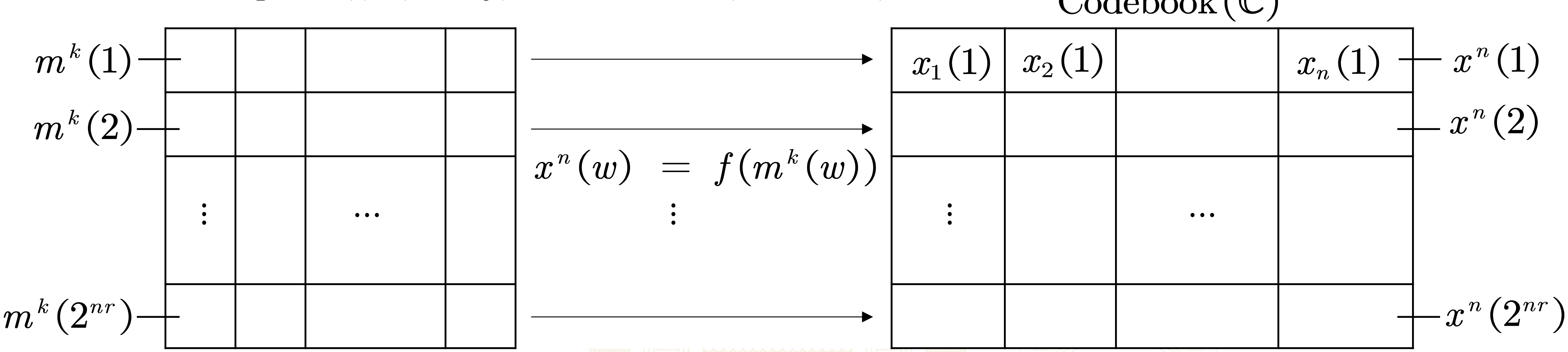
Converse: If $P_e \to 0$, $r < C$.

Proof Pivot:
Fano's Inequality

I. Formulation - Assumption



Assumption (i): (binary) random code (codebook)

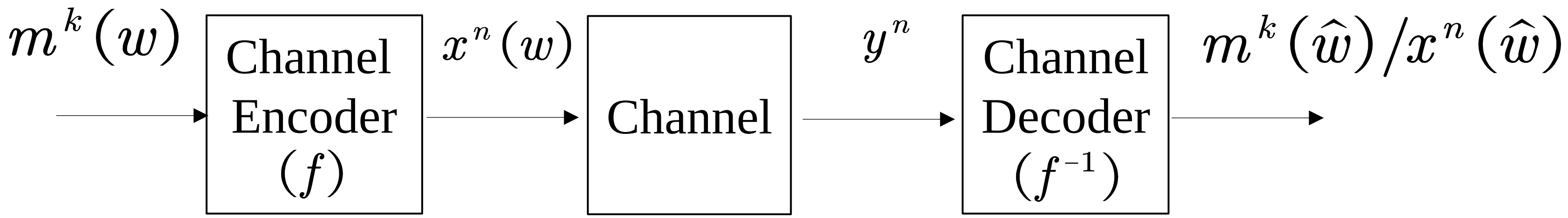


Encoding function f generates codebook $\mathbb{C}$, s.t.

$$P(\mathbb{C}) = \prod_{w=1}^{2^{nr}} \prod_{i=1}^{n} P(x_i(w))$$

Channel Coding Theorem

I. Formulation - Assumptions



Assumption (ii): Both sides know the channel

Assumption (iii): $m^k(w)$ $(x^n(w))$ are uniformly chosen for transmission

$$P(x^n(w)) = P(m^k(w)) = \frac{1}{2^{nr}}$$

Assumption (iv): The channel is discrete memoryless

$$P(y^n|x^n(w)) = \prod_{i=1}^{n} P(y_i|x_i(w))$$

Channel Coding Theorem

I. Assumptions - Proof

Achievability: If $r < C$, $P_e \to 0$.

$$y^n \longrightarrow \boxed{\text{Channel Decoder } (f^{-1})} \longrightarrow m^k(\hat{w})/x^n(\hat{w})$$

$$\boxed{y^n \ \& \ x^n(\hat{w}) \text{ are a pair of JTS}}$$

JTS: Given $\epsilon \to 0$, y^n and $x^n(\hat{w})$ are JTS if

$$\left| -\frac{1}{n}\log_2 P(x^n(\hat{w})) - H(X) \right| < \epsilon$$

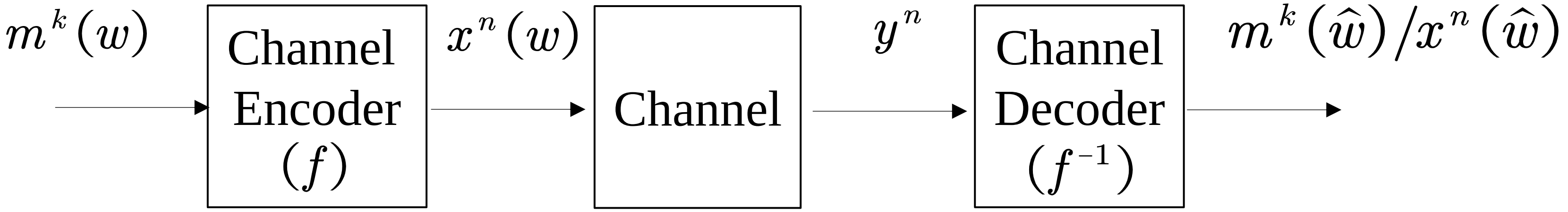
$$\left| -\frac{1}{n}\log_2 P(y^n) - H(Y) \right| < \epsilon$$

$$\left| -\frac{1}{n}\log_2 P(x^n(\hat{w}), y^n) - H(X,Y) \right| < \epsilon$$

Channel Coding Theorem

I. Assumptions - Proof

Achievability: If $r < C$, $P_e \to 0$.



JTS property ①: If $x^n(w)$ and y^n are drawn i.i.d, s.t.

$$P(x^n(w),\ y^n) = \prod_{i=1}^{n} P(x_i(w),\ y_i),$$

ensured by Assumptions (i) (iv)

When $n \to \infty$, $P(x^n(w)$ and y^n are JTS$) = 1 - \epsilon$,

JTS property ②: If $x^n(w)$ and y^n are independent, i.e.,

$$P(x^n(w),\ y^n) = P(x^n(w)) \cdot P(y^n),$$

$$P(x^n(w) \text{ and } y^n \text{ are JTS}) \leqslant 2^{-n(I(X;Y)-3\epsilon)}$$

I. Assumptions - Proof

Achievability: If $r < C$, $P_e \to 0$.

$$
\begin{aligned}
P_e &= \sum_{\mathbb{C}} P(\mathbb{C}) P_e(\mathbb{C}) \\
&= \frac{1}{2^{nr}} \sum_{\mathbb{C}} P(\mathbb{C}) \cdot \sum_{w=1}^{2^{nr}} P_{e,w}(\mathbb{C}) \\
&= \sum_{\mathbb{C}} P(\mathbb{C}) \cdot P_{e,1}(\mathbb{C}) \\
&= P_{e,1}
\end{aligned}
$$

code construction symmetry

$P_e(\mathbb{C})$ - Error probability of code $\mathbb{C}$

$P_{e,w}(\mathbb{C})$ – Error probability of codeword $x^n(w)$

I. Assumptions - Proof

Achievability: If $r < C$, $P_e \to 0$.

E_w: $x^n(w)$ and y^n are JTS.

$$P_{e,1} = P\left(E_1^C \cup E_2 \cup \cdots \cup E_{2^{nr}}\right)$$

$$\leqslant \epsilon + \sum_{w=2}^{2^{nr}} 2^{-n(I(X;Y)-3\epsilon)}$$ Properties ①② of JTS

$$= \epsilon + (2^{nr}-1) \cdot 2^{-n(I(X;Y)-3\epsilon)}$$

$$< \epsilon + 2^{-n(I(X;Y)-r-3\epsilon)}$$

IF input distribution is ideal, s.t. $I(X;Y) = C$ Assumption (i)

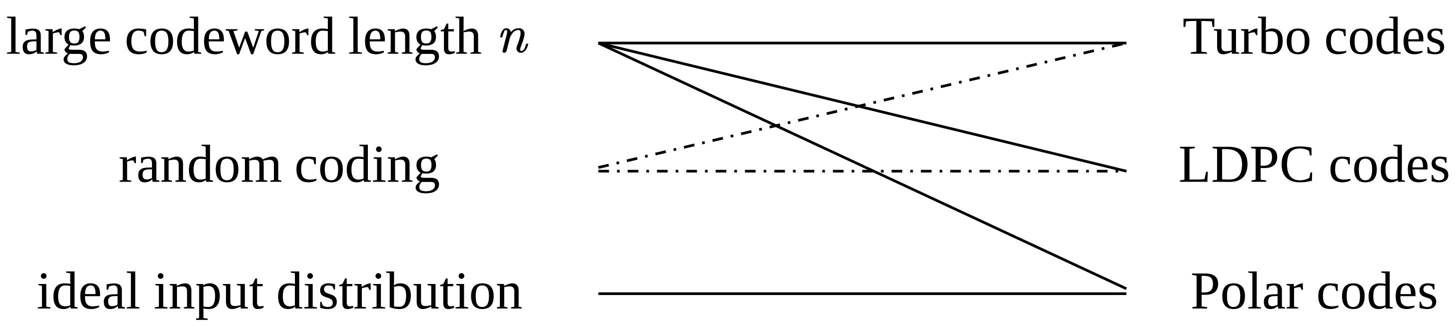
IF $r < C$ & $n \to \infty$

$$P_e = P_{e,1} = 2\epsilon$$

Channel Coding Theorem

I. Proof - Application

Disparity between Assumptions and Practice.



I. Assumptions - Proof

Converse: If $P_e \to 0,\ r < C.$

$$I(m^k(w);\ y^n) = H(m^k(w)) - H(m^k(w)|y^n)$$

$$H(m^k(w)) = \log_2 2^{nr} = nr \qquad \boxed{\text{Assumption (iii)}}$$

$$I(m^k(w);\ y^n) \leqslant I(x^n(w);\ y^n) \qquad \boxed{\text{Data Processing Inequality}}$$

$$\leqslant n \cdot C \qquad \boxed{\text{Assumptions (i) (iv)}}$$

$$H(m^k(w)|y^n) = H(c^n(w)|y^n)$$

$$\leqslant H(P_e) + P_e \log_2(2^{nr} - 1) \qquad \boxed{\text{Fano's Inequality}}$$

$$\lesssim 1 + P_e \cdot nr$$

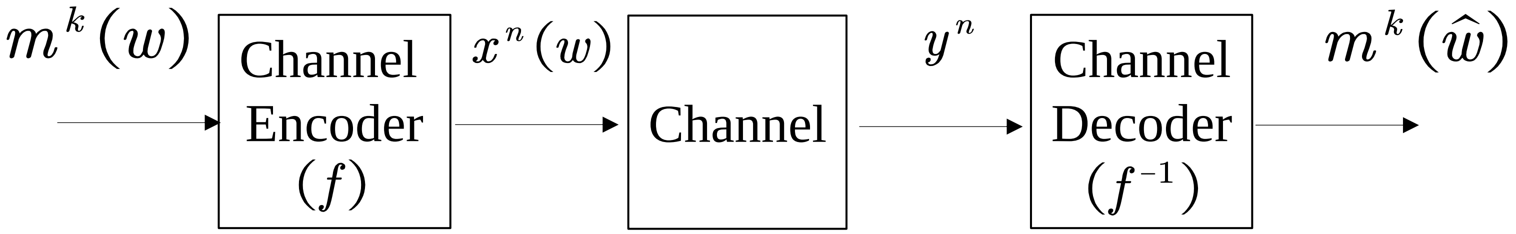
$$nr \leqslant 1 + P_e \cdot nr + nC \text{ ,or } r \leq C + \frac{1}{n} + P_e \cdot r$$

IF $P_e \to 0$ & $n \to \infty$

$$r \leq C$$

Channel Coding Theorem

II. The Art of Channel Coding



Random coding is Good for proof, but Bad for Practice

$$f\colon\ m^k(w) \mapsto x^n(w)$$

$$f^{-1}\colon\ y^n \mapsto m^k(\hat{w})/x^n(\hat{w})$$

The art of channel coding arises as we refrain f from randomness

II. The Art of Channel Coding

In light of linear block codes,

$$f : \mathbf{G} = \begin{bmatrix} x^n(1) \\ x^n(2) \\ \vdots \\ x^n(k) \end{bmatrix}$$

a basis of k linearly independent codewords

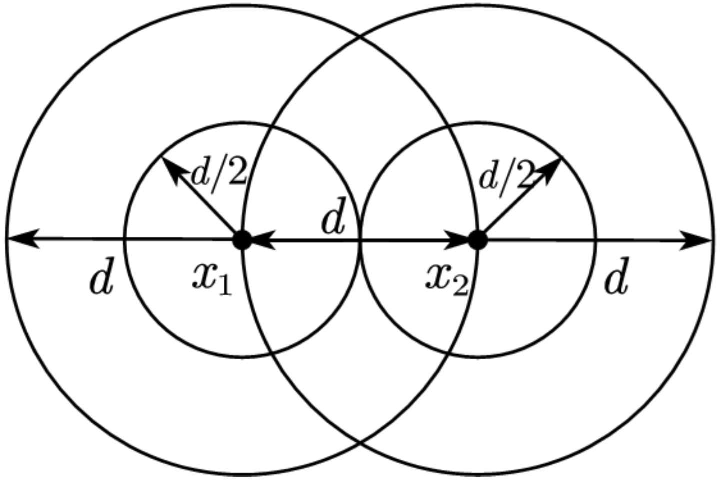
$$x^n(w) = m^k(w) \cdot \mathbf{G}.$$

$$\mathbb{C} = \{m^k(w) \cdot \mathbf{G}, \ \forall w\}$$

If $\mathbf{G}$ is a semi-definite and tridiagonal matrix, the code becomes a convolutional code

II. The Art of Channel Coding

For an (n, k) block code $(\mathbb{C})$, $d_{\text{Ham}}(\mathbb{C}) = \min\{d_{\text{Ham}}(x^n(w), x^n(w'))\}$



The number of correctable errors: $\longrightarrow \tau = \left\lfloor \frac{d_{\text{Ham}}(\mathbb{C})}{2} \right\rfloor$

Singleton bound: $d_{\text{Ham}}(\mathbb{C}) \leq n - k + 1$

II. The Art of Channel Coding

$(n,\ 1,\ n)$ Repetition code $\qquad \mathbf{G} = (1,\ 1,\ \cdots,\ 1)$

Majority voting realizes maximum likelihood (ML) decoding

II. The Art of Channel Coding

Symmetric Code for Poetry

$$(\text{中秋圆月}) \cdot \mathbf{G} = (\text{中秋圆月 月圆秋中})$$

$$(\text{围席似月}) \cdot \mathbf{G} = (\text{围席似月 月似席围})$$

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

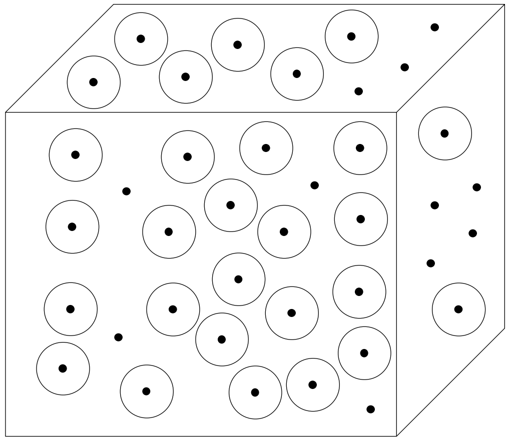
II. The Art of Channel Coding

$$(n,\ k)\ \text{block code} \perp (n,\ n-k)\ \text{block code}$$

(dual codes)

$$f_{\perp} : \mathbf{G}_{\perp} = \begin{bmatrix} x_{\perp}^{n}(1) \\ x_{\perp}^{n}(2) \\ \vdots \\ x_{\perp}^{n}(n-k) \end{bmatrix}$$

(a basis of $(n - k)$ linearly independent dual codewords)



$$x_{\perp}^{n}(w) = m^{n-k}(w) \cdot \mathbf{G}_{\perp},\ x^{n}(w) \cdot (x_{\perp}^{n}(w))^{\mathrm{T}} = 0$$

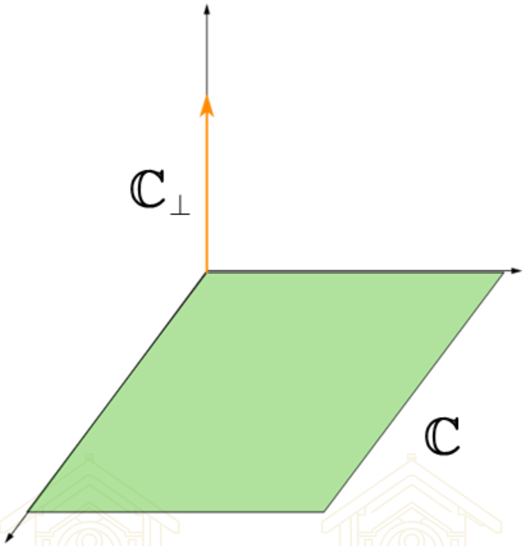
$$\mathbb{C}_{\perp} = \{m^{n-k}(w) \cdot \mathbf{G}_{\perp}, \forall\, w\}$$

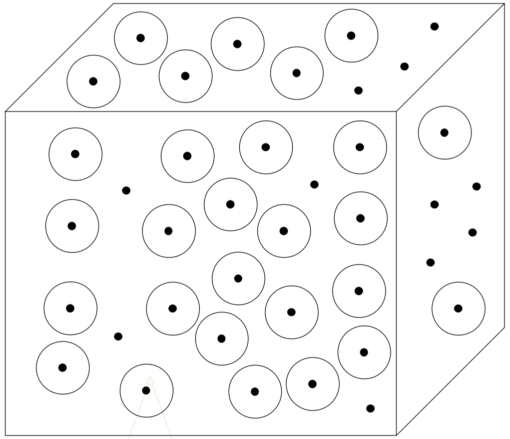
II. The Art of Channel Coding

$(n,\ k)$ block code $\perp$ $(n,\ n-k)$ block code

(dual codes) $\mathbb{C}$ and $\mathbb{C}_{\perp}$ are orthogonal

f^{-1} : $\mathbf{H} = \mathbf{G}_{\perp}$





The art in f^{-1} is tantamount to that in f

II. The Art of Channel Coding

Art in f	Art in f^{-1}
Polar	RS
LDPC	BCH
Turbo	RM
Concatenated	Hamming
AG	
Convolutional	

Channel Coding Theorem

References:

1. Website: www.chencode.cn
2. Email: chenli55@mail.sysu.edu.cn
3. Videos: https://www.xuetangx.com/course/sysu0807bt

中山大学：1924 - 2024 -

谢谢！