香港中文大學
The Chinese University of Hong Kong

7.2 The Channel Coding Theorem

- **Direct Part** Information can be communicated through a DMC with an arbitrarily small probability of error at any rate less than the channel capacity.

- **Direct Part** Information can be communicated through a DMC with an arbitrarily small probability of error at any rate less than the channel capacity.
- **Converse** If information is communicated through a DMC at a rate higher than the capacity, then the probability of error is bounded away from zero.

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

$$f : \{1, 2, \cdots, M\} \to \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

$$f : \{1, 2, \cdots, M\} \to \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

$$f : \{1, 2, \cdots, M\} \to \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

- **Block length** n

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

$$f : \{1, 2, \cdots, M\} \to \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \to \{1, 2, \cdots, M\}.$$

- **Block length** n
- **Message Set** $\mathcal{W} = \{1, 2, \cdots, M\}$

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

$$f : \{1, 2, \cdots, M\} \rightarrow \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \rightarrow \{1, 2, \cdots, M\}.$$

- **Block length** n
- **Message Set** $\mathcal{W} = \{1, 2, \cdots, M\}$
- **Codewords** $f(1), f(2), \cdots, f(M)$

Definition of a Channel Code

Definition 7.9 An (n, M) code for a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is defined by an encoding function

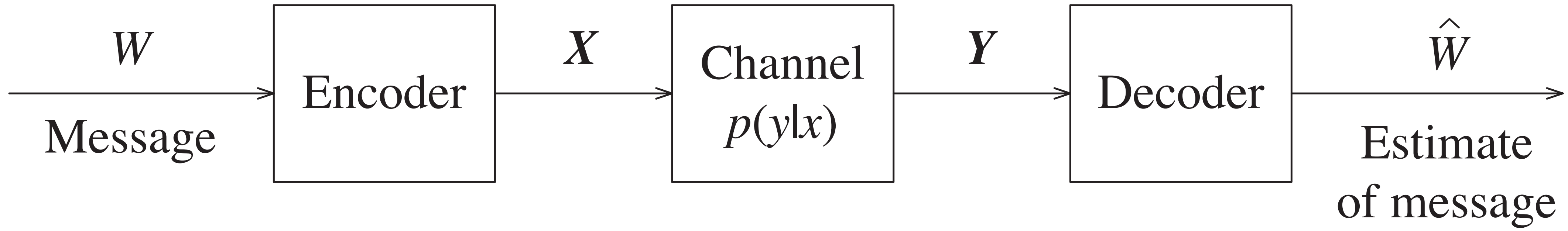
$$f : \{1, 2, \cdots, M\} \rightarrow \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \rightarrow \{1, 2, \cdots, M\}.$$

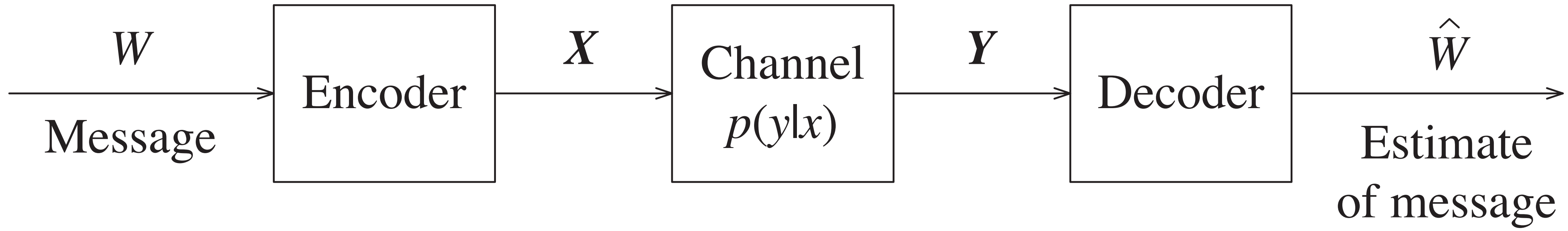
- **Block length** n
- **Message Set** $\mathcal{W} = \{1, 2, \cdots, M\}$
- **Codewords** $f(1), f(2), \cdots, f(M)$
- **Codebook** the set of all codewords

Assumptions and Notations



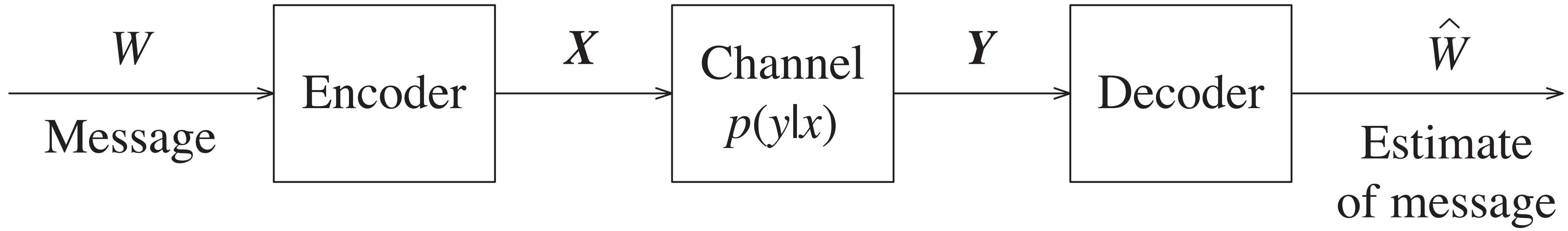
Assumptions and Notations

- W is randomly chosen from the message set $\mathcal{W}$, so $H(W) = \log M$.



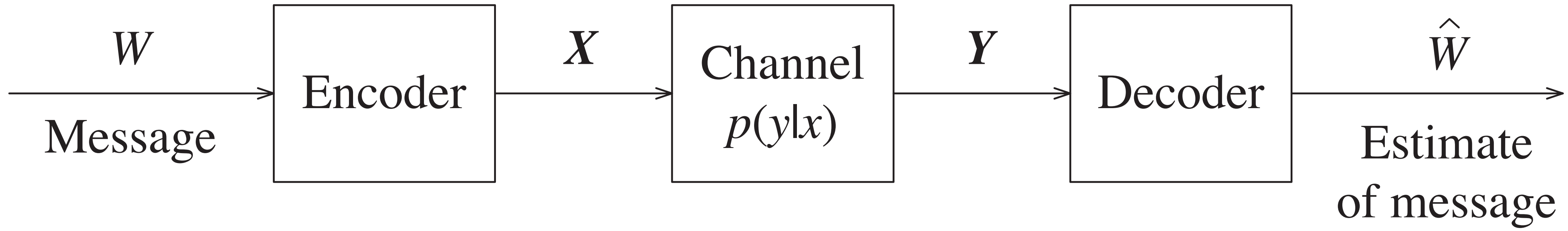
Assumptions and Notations

- W is randomly chosen from the message set $\mathcal{W}$, so $H(W) = \log M$.
- $\mathbf{X} = (X_1, X_2, \cdots, X_n)$; $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$



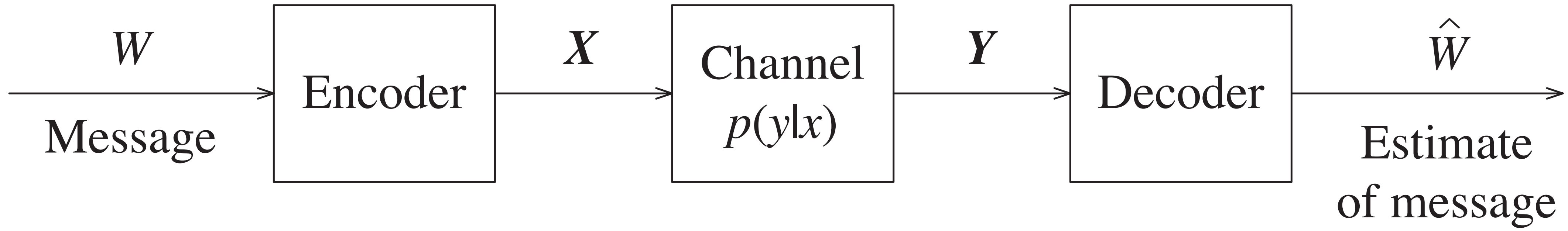
Assumptions and Notations

- W is randomly chosen from the message set $\mathcal{W}$, so $H(W) = \log M$.
- $\mathbf{X} = (X_1, X_2, \cdots, X_n)$; $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$
- Thus $\mathbf{X} = f(W)$.



Assumptions and Notations

- W is randomly chosen from the message set $\mathcal{W}$, so $H(W) = \log M$.
- $\mathbf{X} = (X_1, X_2, \cdots, X_n)$; $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$
- Thus $\mathbf{X} = f(W)$.
- Let $\hat{W} = g(\mathbf{Y})$ be the estimate on the message W by the decoder.



Error Probabilities

Error Probabilities

Definition 7.10 For all $1 \leq w \leq M$, let

$$\lambda_w = \Pr\{\hat{W} \neq w | W = w\} = \sum_{\mathbf{y} \in \mathcal{Y}^n : g(\mathbf{y}) \neq w} \Pr\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = f(w)\}$$

be the conditional probability of error given that the message is w.

Error Probabilities

Definition 7.10 For all $1 \leq w \leq M$, let

$$\lambda_w = \Pr\{\hat{W} \neq w | W = w\} = \sum_{\mathbf{y} \in \mathcal{Y}^n : g(\mathbf{y}) \neq w} \Pr\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = f(w)\}$$

be the conditional probability of error given that the message is w.

Definition 7.11 The maximal probability of error of an (n, M) code is defined as

$$\lambda_{max} = \max_w \lambda_w.$$

Error Probabilities

Definition 7.10 For all $1 \leq w \leq M$, let

$$\lambda_w = \Pr\{\hat{W} \neq w | W = w\} = \sum_{\mathbf{y} \in \mathcal{Y}^n : g(\mathbf{y}) \neq w} \Pr\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = f(w)\}$$

be the conditional probability of error given that the message is w.

Definition 7.11 The maximal probability of error of an (n, M) code is defined as

$$\lambda_{max} = \max_w \lambda_w.$$

Definition 7.12 The average probability of error of an (n, M) code is defined as

$$P_e = \Pr\{\hat{W} \neq W\}.$$

P_e vs λ_{max}

P_e vs λ_{max}

- Consider

P_e vs λ_{max}

- Consider

$$P_e = \Pr\{\hat{W} \neq W\}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \Pr\{\underline{W = w}\}\Pr\{\hat{W} \neq W | W = w\} \end{aligned}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \Pr\{\underline{W = w}\}\Pr\{\hat{W} \neq W | \underline{W = w}\} \end{aligned}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \underline{\Pr\{W = w\}} \Pr\{\hat{W} \neq W | W = w\} \end{aligned}$$

P_e vs λ_{max}

- Consider

$$
\begin{aligned}
P_e &= \Pr\{\hat{W} \neq W\} \\
&= \sum_w \underline{\Pr\{W = w\}} \Pr\{\hat{W} \neq W | W = w\} \\
&= \sum_w \underline{\frac{1}{M}} \Pr\{\hat{W} \neq w | W = w\}
\end{aligned}
$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \Pr\{W = w\} \underline{\Pr\{\hat{W} \neq W | W = w\}} \\ &= \sum_w \frac{1}{M} \Pr\{\hat{W} \neq w | W = w\} \end{aligned}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \Pr\{W = w\}\underline{\Pr\{\hat{W} \neq W | W = w\}} \\ &= \sum_w \frac{1}{M}\underline{\Pr\{\hat{W} \neq w | W = w\}} \end{aligned}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned}
P_e &= \Pr\{\hat{W} \neq W\} \\
&= \sum_w \Pr\{W = w\} \Pr\{\hat{W} \neq W | W = w\} \\
&= \sum_w \frac{1}{M} \underline{\Pr\{\hat{W} \neq w | W = w\}} \\
&= \frac{1}{M} \sum_w \underline{\lambda_w}.
\end{aligned}$$

P_e vs λ_{max}

- Consider

$$\begin{aligned} P_e &= \Pr\{\hat{W} \neq W\} \\ &= \sum_w \Pr\{W = w\}\Pr\{\hat{W} \neq W | W = w\} \\ &= \sum_w \frac{1}{M}\Pr\{\hat{W} \neq w | W = w\} \\ &= \frac{1}{M}\sum_w \lambda_w. \end{aligned}$$

- Therefore,

$$P_e \leq \max_w \lambda_w = \lambda_{max}.$$

Rate of a Channel Code

Rate of a Channel Code

Definition 7.13 The rate of an (n, M) channel code is $n^{-1} \log M$ in bits per use.

Rate of a Channel Code

Definition 7.13 The rate of an (n, M) channel code is $n^{-1} \log M$ in bits per use.

Definition 7.14 A rate R is (asymptotically) achievable for a discrete memoryless channel if for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon$$

and

$$\lambda_{max} < \epsilon.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.