《信道编码》
《Channel Coding》

Li Chen
Professor, School of Electronics and Communication Engineering
Sun Yat-sen University
Office: 631C, SEIT Building
Email: chenli55@mail.sysu.edu.cn
website: www.chencode.cn

《 Channel Coding 》

Textbooks:

1.《Elements of Information Theory》, by T. Cover and J. Thomas, Wiley (and introduced by Tsinghua University Press), 2003.

2. 《Non-binary error control coding for wireless communication and data storage》, by R. Carrasco and M. Johnston, Wiley, 2008.

3. 《Error Control Coding》, by S. Lin and D. Costello, Prentice Hall, 2004.

4. 《信息论与编码理论》， 王育民、李晖著，高等教育出版社，2013.

Outlines

Chapter 1: Fundamentals of Information Theory (3 W)

Chapter 2: An Introduction of Channel Coding (2 W)

Chapter 3: Convolutional Codes and Trellis Coded Modulation (5 W)

Chapter 4: Turbo Codes (2 W)

Chapter 5: Low-Density Parity-Check Codes (3 W)

Chapter 6: Reed-Solomon Codes (3 W)

Evolution of Communications

Analogue comm.

Late 80s to early 90s

Digital comm.

Information theory and coding techniques

1G

2G

2.5G

EE+CS

3G

4G

Chapter 1 Fundamentals of Information Theory

- 1.1 An Introduction of Information
- 1.2 Entropy
- 1.3 Mutual Information
- 1.4 Channel Capacity

§ 1.1 An Introduction of Information

- What is information?
- How do we measure information?

Let us look at the following sentences:

1) I will be one year older next year.

    No information — Boring!

2) I was born in 1993.

    Some information — Being frank!

3) I was born in 1990s.

    More information — Interesting, so which year?

The number of *possibilities* should be linked to the information!

§ 1.1 An Introduction of Information

Let us do the following game:

Throw a die once

You have 6 possible outcomes.
{1, 2, 3, 4, 5, 6}

Throw three dies

You have 6^3 possible outcomes.
{(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 1, 4)
……
(2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 1, 4)
……
(6, 6, 3), (6, 6, 4), (6, 6, 5), (6, 6, 6)}

Information should be *'additive'*.

§ 1.1 An Introduction of Information

Let us look at the following problem.

If there are 30 students in our class, and we would like to use binary bits to distinguish each of them, how many bits do we need?

Solution: 30 possibilities.

requires

$\log_2 30 = 4.907$ bits.

we need at least 5 bits to represent each of us.

Q: There are 7 billion people on our planet, how many bits do we need?

We can use *'logarithm'* to scale down the a huge amount of possibilities.

Number (binary bit) permutations are used to represent all possibilities.

§ 1.1 An Introduction of Information

Finally, let us look into the following game.

Pick one ball from the hat randomly,

The probability of picking up a white ball, $\frac{1}{4}$ (25%).

Representing the probability needs

$$\log_2 \frac{1}{1/4} = 2 \text{ bits.}$$

The probability of picking up a black ball, $\frac{3}{4}$ (75%).

Representing the probability needs

$$\log_2 \frac{1}{3/4} = 0.415 \text{ bits.}$$

§ 1.1 An Introduction of Information

- How do we measure the overall event? (On average, how many bits do we need to represent an outcome?)

$$\frac{1}{4}\cdot\log_2\frac{1}{1/4}+\frac{3}{4}\log_2\frac{1}{3/4}=0.811 \text{ bits.}$$

- The measure of information should be

$$\sum_{i=1}^{N} P_i \log_2 P_i^{-1} = -\sum_{i=1}^{N} P_i \log_2 P_i$$

  - P_i: probability of the ith possible event.
  - N: Total number of possible events.

Measure of information should consider the *probabilities of various possible events.*

§ 1.2 Entropy

- Information: knowledge not precisely known by the recipient, as it is a measure of unexpectedness.
- Amount of information ∝ (probability of occurance)$^{-1}$
- Messages: $M_1 \quad M_2 \quad M_3 \ldots\ldots \quad M_q$

  Prob of occur: $P_1 \quad P_2 \quad P_3 \ldots\ldots \quad P_q \qquad (P_1 + P_2 + P_3 + \cdots + P_q = 1)$

Measure the amount of information carried by each message by

$$I(M_i) = log_x P_i^{-1}, \quad i = 1, 2, \ldots, q$$

$x = 2$, $I(M_i)$ in bits

$x = e$, $I(M_i)$ in nats

$x = 10$, $I(M_i)$ in Hartley.

- Observations: … …

§ 1.2 Entropy

Observations:

1) $I(M_i) \to 0, \quad if \quad P_i \to 1;$
2) $I(M_i) \geq 0, \quad when \quad 0 \leq P_i \leq 1;$
3) $I(M_i) > I(M_j), \quad if \quad P_j > P_i$
4) Given M_i and M_j are statistically independent,

   $I(M_i \& M_j) = I(M_i) + I(M_j).$

§ 1.2 Entropy

Example 1.1: A source outputs five possible messages. The probabilities of these messages are:

$$P_1 = \frac{1}{2} \quad P_2 = \frac{1}{4} \quad P_3 = \frac{1}{8} \quad P_4 = \frac{1}{16} \quad P_5 = \frac{1}{16}.$$

Determine the information contained in each of these messages.

Solution:

$$I(M_1) = \log_2 \frac{1}{1/2} = 1 \text{ bit}$$

$$I(M_2) = \log_2 \frac{1}{1/4} = 2 \text{ bit}$$

$$I(M_3) = \log_2 \frac{1}{1/8} = 3 \text{ bit}$$

$$I(M_4) = \log_2 \frac{1}{1/16} = 4 \text{ bit}$$

$$I(M_5) = \log_2 \frac{1}{1/16} = 4 \text{ bit}$$

Total amount of information = 14 bits. Is it right?

§ 1.2 Entropy

Given a source vector of length N, and it has U possible symbols $S_1, S_2, \dots S_U$, each of which has probability of $P_1, P_2, \dots P_U$ of occurrence.

To represent the source vector, we need

$$I = \sum_{i=1}^{U} N\, P_i log_2 P_i^{-1} \text{ bits.}$$

So on average, how many information bits do we need for a source symbol?

$$H = \frac{I}{N} = \sum_{i=1}^{U} P_i log_2 P_i^{-1} \text{ bits/symbol}$$

H is called the source entropy – average number of information per source symbol.

§ 1.2 Entropy

Example 1.2: A source vector contains symbols of four possible outcomes A, B, C, D. They occur with probabilities of $\frac{1}{4}$, $\frac{1}{3}$, $\frac{1}{3}$ and $\frac{1}{12}$, respectively. Determine the entropy of the source vector.

$$H = \frac{1}{4}\log_2\frac{1}{1/4} + \frac{2}{3}\log_2\frac{1}{1/3} + \frac{1}{12}\log_2\frac{1}{1/12}$$

$$= 1.856 \text{ bits/symbol}$$

§ 1.2 Entropy

Entropy of a binary source: The source vector has only two possible symbols, i.e., 0 and 1. Let $P(0)$ denote the probability of a source symbol being 0, and $P(1)$ denote the probability of a source symbol being 1, we have

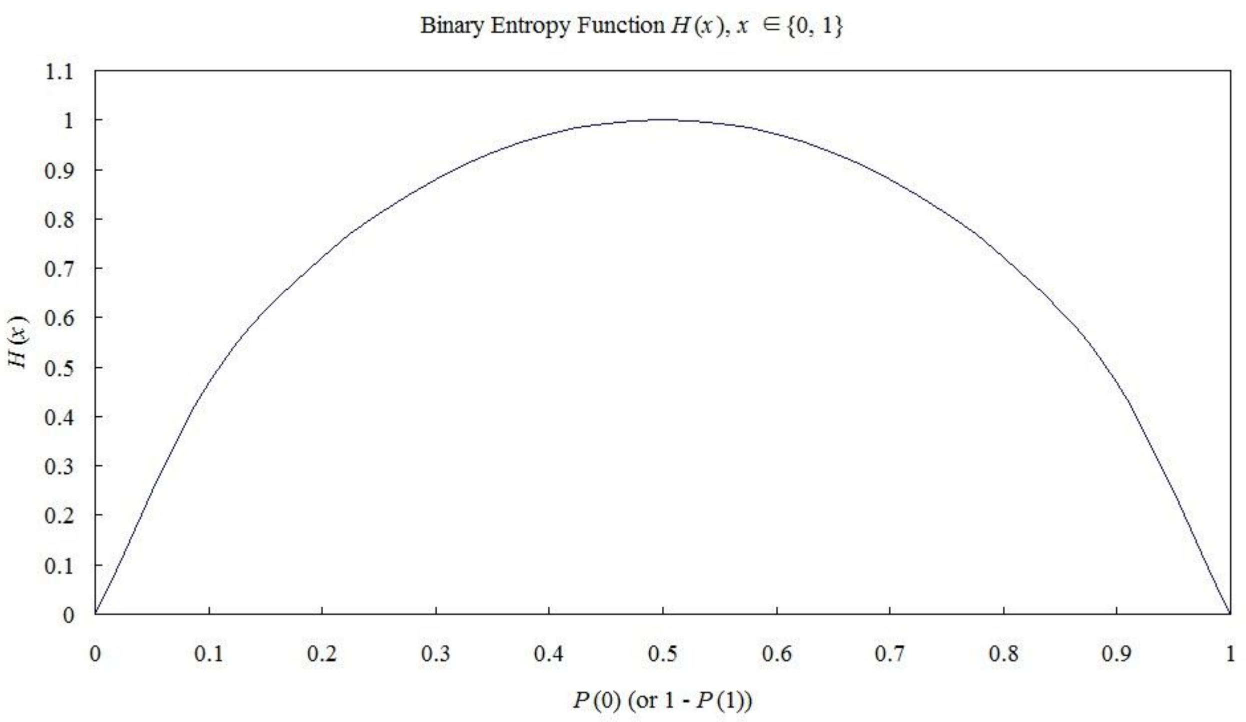
$$H = P(0) \cdot \log_2 P(0)^{-1} + P(1) \log_2 P(1)^{-1}$$

or

$$H = P(0) \cdot \log_2 P(0)^{-1} + (1 - P(0)) \cdot \log_2 (1 - P(0))^{-1}$$

Binary Entropy Function

§ 1.2 Entropy



Binary Entropy Function $H(x)$, $x \in \{0, 1\}$
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
$H(x)$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
$P(0)$ (or 1 - $P(1)$)

§ 1.2 Entropy

- Entropy for two random variables *X* and *Y*.
- Realizations of *X* and *Y* are *x* and *y*.
- Distributions of *X* and *Y* are *P*(*x*) and *P*(*y*).

Joint Entropy $H(X, Y)$: Given a joint distribution $P(x, y)$,

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 P(x,y)$$

Condition Entropy $H(Y|X)$:

$$H(Y|X) = \sum_{x\in X} P(x)H(Y|X=x)$$

$$= -\sum_{x\in X}\sum_{y\in Y} P(x)P(y|x)\log_2 P(y|x)$$

$$= -\sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 P(y|x)$$

§ 1.2 Entropy

The Chain Rule (Relationship between Joint Entropy and Conditional Entropy)

$$H(X,Y) = H(X) + H(Y|X)$$

$$= H(Y) + H(X|Y)$$

Proof:

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(x,y)$$

$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 (P(y|x)P(x))$$

$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(x) - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(y|x)$$

$$= -\sum_{x \in X} P(x) \log_2 P(x) - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(y|x)$$

$$= H(X) + H(Y|X)$$

§ 1.3 Mutual Information

- Two random variables X and Y.
- Realizations of X and Y are x and y.
- Distributions of X and Y are $P(x)$ and $P(y)$.
- Joint distribution of X and Y is $P(x, y)$.
- Conditional distribution of X is $P(x|y)$.

Mutual Information between X and Y:

$$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 \frac{P(x|y)}{P(x)}$$

§ 1.3 Mutual Information

Mutual Information's Relationship with Entropy:

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

Proof:

$$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 \frac{P(x,y)}{P(x)P(y)}$$

$$= \sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 P(x,y) - \sum_{x\in X} P(x)\log_2 P(x) - \sum_{y\in Y} P(y)\log_2 P(y)$$

$$= H(X) + H(Y) - H(X,Y)$$

Remark: The above proof also shows the symmetry of mutual information as

$$I(X,Y) = I(Y,X)$$

§ 1.3 Mutual Information

Mutual Information's Relationship with Entropy:

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

This relationship can be visualized in the Venn diagram

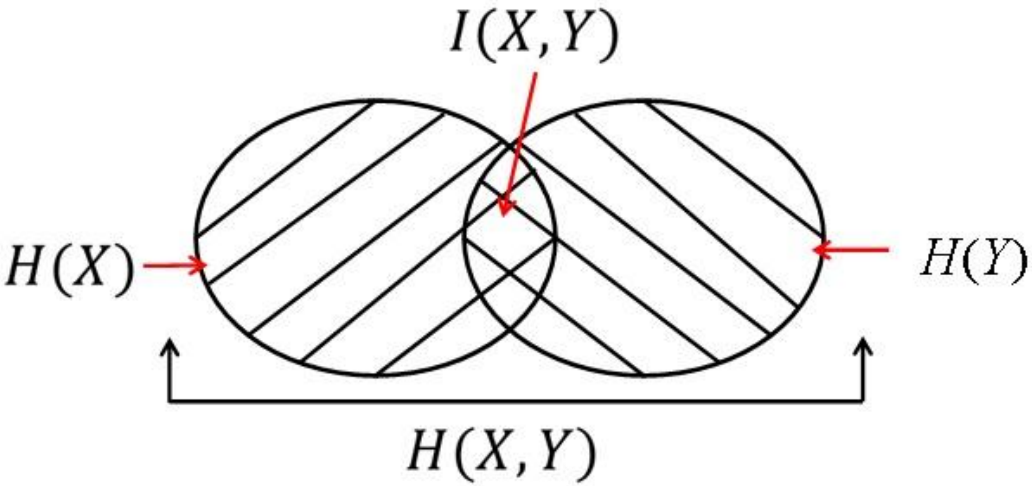


Fig. A Venn diagram

§ 1.3 Mutual Information

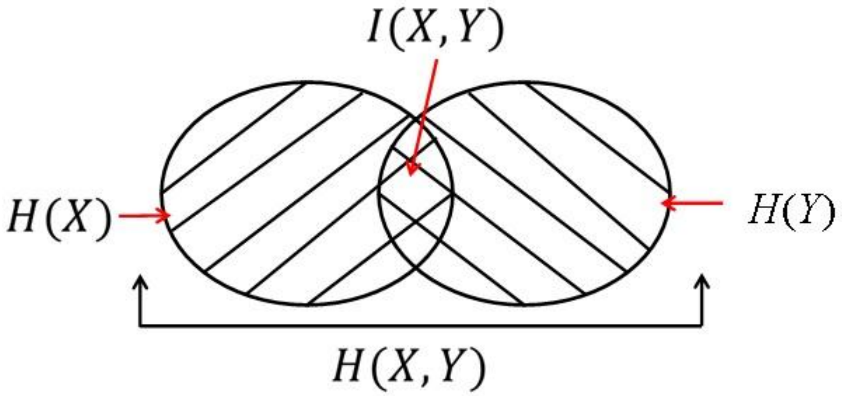


Fig. A Venn diagram

Corollary:

$$\begin{aligned} I(X,Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$

This can also be concluded using the Chain Rule.

Bounds on $I(X,Y)$

$$0 \le I(X,Y) \le \min\{H(X), H(Y)\}$$

§ 1.3 Mutual Information

Mutual Information of A Channel

- Consider X is the transmitted signal, Y is the received signal.
- Y is a variant of X where the discrepancy is introduced by channel.

$$H(X) - H(X|Y)$$

How much we don't know BEFORE the channel observations.

−

How much we still don't know AFTER the channel observations.

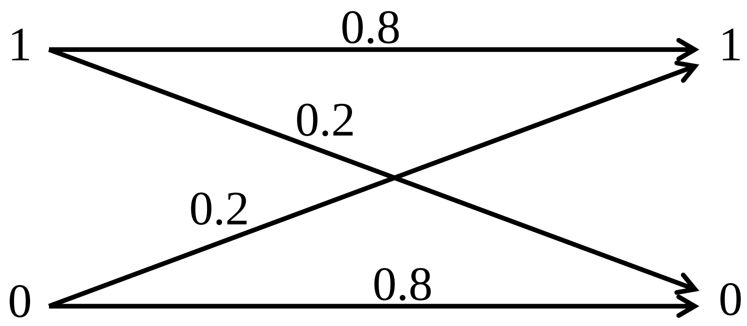
⇓

How much information is carried by the channel, and this is called the **Mutual Information** of the channel, denoted as $I(X,Y)$.

Remark: Mutual information $I(X, Y)$ describes the amount of information one variable X contains about the other Y, or vice versa as in $I(Y, X)$.

§ 1.3 Mutual Information

Example 1.3: Given the binary symmetric channel shown as



We know $P(x=0)=0.3$, $P(x=1)=0.7$, $P(y=1|x=1)=0.8$,
$P(y=1|x=0)=0.2$, $P(y=0|x=1)=0.2$ and $P(y=0|x=0)=0.8$.

Please determine the mutual information of such a channel.

Solution:

- Entropy of the binary source is

$$H(x) = -P(x=0)\log_2 P(x=0) - P(x=1)\log_2 P(x=1)$$
$$= 0.3\cdot\log_2\frac{1}{0.3} + 0.7\cdot\log_2\frac{1}{0.7}$$
$$= 0.881 \text{ bits}$$

§ 1.3 Mutual Information

- With $P(x)$ and $P(y|x)$, we know

$$P(y=1)=P(y=1|x=1)P(x=1)+P(y=1|x=0)P(x=0)$$
$$=0.62$$
$$P(y=0)=P(y=0|x=1)P(x=1)+P(y=0|x=0)P(X=0)$$
$$=0.38$$
$$P(x=0,y=0)=P(y=0|x=0)\cdot P(x=0)=0.24$$
$$P(x=0|y=0)=\frac{P(x=0,y=0)}{P(y=0)}=0.63$$
$$P(x=1,y=0)=P(y=0|x=1)\cdot P(x=1)=0.14$$
$$P(x=1|y=0)=\frac{P(x=1,y=0)}{P(y=0)}=0.37$$
$$P(x=0,y=1)=P(y=1|x=0)P(x=0)=0.06$$
$$P(x=0|y=1)=\frac{P(x=0,y=1)}{P(y=1)}=0.10$$
$$P(x=1,y=1)=P(y=1|x=1)P(x=1)=0.56$$
$$P(x=1|y=1)=\frac{P(x=1,y=1)}{P(y=1)}=0.90$$

§ 1.3 Mutual Information

- Hence, the conditional entropy is:

$$H(X|Y)=P(x=0,y=0)\log_2\frac{1}{P(x=0|y=0)}+P(x=1,y=0)\log_2\frac{1}{P(x=1|y=0)}$$

$$+P(x=0,y=1)\log_2\frac{1}{P(x=0|y=1)}+P(x=1,y=1)\log_2\frac{1}{P(x=1|y=1)}$$

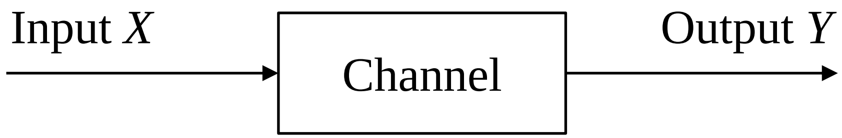
$$=0.24\log_2\frac{1}{0.63}+0.14\log_2\frac{1}{0.37}+0.06\log_2\frac{1}{0.10}+0.56\log_2\frac{1}{0.90}$$

$$=0.644\text{bits/sym}$$

- The mutual information is:

$$I(X,Y)=H(X)-H(X|Y)=0.237\text{bits}$$

§ 1.4 Channel Capacity



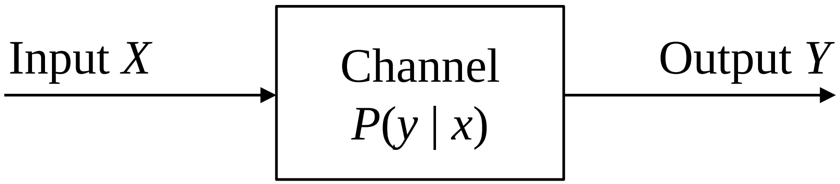
- In a communication system, with the observation of Y, we aim to recover X.
- Mutual Information $I(X,Y) = H(X) - H(X|Y)$

  $$= H(Y) - H(Y|X)$$

  It defines the amount of uncertainty about X that has been reduced thanks to the knowledge of Y, and vise versa. This uncertainty discrepancy is introduced by the channel.

- Channel capacity describes the channel's best capability in reducing the uncertainty.

§ 1.4 Channel Capacity



- Let the realization of input *X* and output *Y* be *x* and *y*, respectively.
- Channel transition probability $P(y \mid x)$: knowing *x* was transmitted, the probability of observing *y*. It defines the quality of channel.
- Channel Capacity

$$C = \max_{P(x)} \{ I(X, Y) \}$$

The maximum mutual information $I(X,Y)$ that can be realized over all distribution of the input $P(x)$.

§ 1.4 Channel Capacity

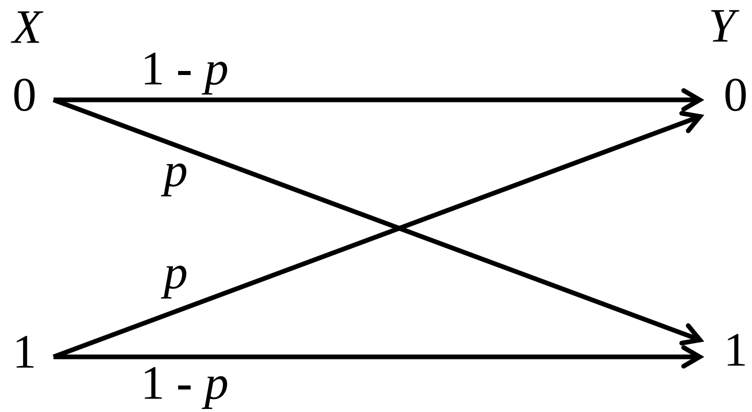
- Channel Capacity: $C = \max_{P(x)}\{I(X,Y)\}$
- In a wireless communication system, it is the maximum number of information bits that can be carried by a modulated symbol such that the information can be recovered with an arbitrarily low probability of error.
- To realize this reliable communications, channel coding is needed. Given k information symbols (or bits), redundancy is added to obtain n ($n > k$) codeword symbols (or bits). The code rate is $r = \frac{k}{n}$. Using binary modulation, e.g., BPSK, reliable communications is possible if $r < C$.

§ 1.4 Channel Capacity

- Why input distribution $P(x)$ matters?
- Consider the data transmission as human flows from Shenzhen to Hong Kong

§ 1.4 Channel Capacity

- **Binary Symmetric Channel (BSC)**



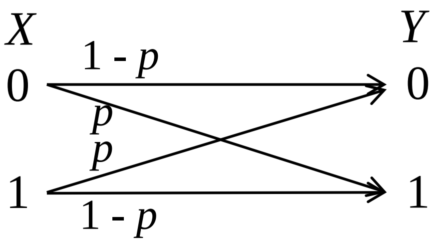
- Input: 0 1 0 0 0 1 1 0 1 0 …

  Output: 0 1 1 1 0 0 1 0 0 0 …
- Input and output are discrete
- $P(y=1|x=0) = P(y=0|x=1) = p$

  $P(y=0|x=0) = P(y=1|x=1) = 1-p$
- It is the simplest model of channel that introduces errors. Many wireless channels can be abstracted to BSC.

§ 1.4 Channel Capacity

- **Binary Symmetric Channel (BSC)**



- Analytic intuition

$$I(X,Y) = H(Y) - H(Y|X)$$

$I(X,Y)$ will be maximized if $H(Y)$ is maximized and $H(Y|X)$ is minimized.

(1) $H(Y) \leq 1$.

(2) $H(Y|X) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 P(y|x)$

$$= -\sum_{x\in X}\sum_{y\in Y} P(y|x)P(x)\log_2 P(y|x)$$

$$= -P(x=0)\sum_{y\in\{0,1\}} P(y|x=0)\log_2 P(y|x=0)$$

$$-P(x=1)\sum_{y\in\{0,1\}} P(y|x=1)\log_2 P(y|x=1)$$

$$= -P(x=0)((1-p)\log_2(1-p) + p\log_2 p) - P(x=1)(p\log_2 p + (1-p)\log_2(1-p))$$

$$= -(1-p)\log_2(1-p) - p\log_2 p$$

- When $P(x=0) = P(x=1) = \frac{1}{2}$, $H(Y) = 1$ and

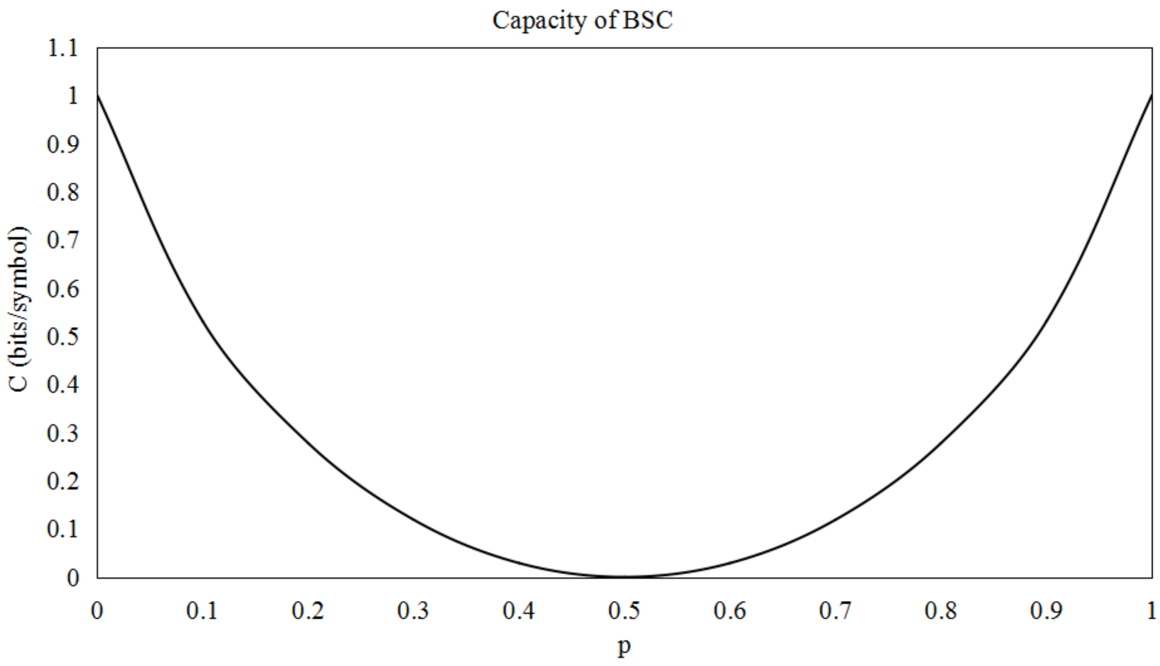
$$C = 1 - H(Y|X) \text{ bits/symbol.}$$

§ 1.4 Channel Capacity

- **Binary Symmetric Channel (BSC)**
- Intuition: If 0 and 1 experience the same degree of channel impairment, i.e., $P(y=1|x=0) = P(y=0|x=1)$, there is no need to prioritize either 0 or 1 for transmission and $P(x=0) = P(x=1) = \frac{1}{2}$.
- $C = 1 - H(Y|X)$, if $P(x=0) = P(x=1) = \frac{1}{2}$.
- $$\begin{aligned} H(Y|X) &= -P(y=0|x=0)\cdot\frac{1}{2}\cdot\log_2 P(y=0|x=0) \\ &= -P(y=1|x=0)\cdot\frac{1}{2}\cdot\log_2 P(y=1|x=0) \\ &= -P(y=0|x=1)\cdot\frac{1}{2}\cdot\log_2 P(y=0|x=1) \\ &= -P(y=1|x=1)\cdot\frac{1}{2}\cdot\log_2 P(y=1|x=1) \\ &= -p\log_2 p - (1-p)\log_2(1-p) \end{aligned}$$
- $C = 1 + p\log_2 p + (1-p)\log_2(1-p)$ bits/symbol

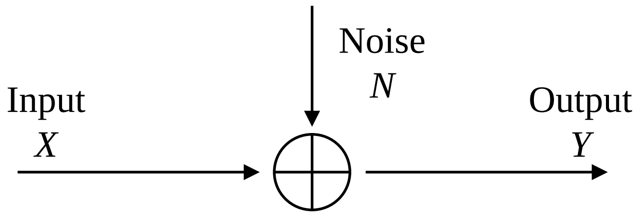
§ 1.4 Channel Capacity

- **Binary Symmetric Channel (BSC)**
- $C = 1 + p\log_2 p + (1-p)\log_2(1-p)$ bits/symbol



§ 1.4 Channel Capacity

- **<u>Additive White Gaussian Noise (AWGN) Channel</u>**



- Channel model $y_i = x_i + n_i$

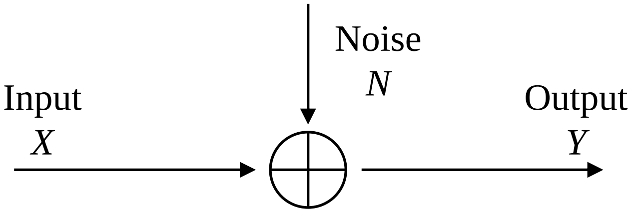
  x_i: discrete input signal, a modulated signal

  n_i: white Gaussian noise as $\mathcal{N}(0, \sigma_N^2)$, independent of x_i

  y_i: continuous output signal, a variation of x_i
- It is a more realistic wireless channel model where the transmitted signal is impaired by noise.
- It is adopted to represent the space communication channel where light-of-sight (LoS) transmission is always ensured.
- It is also often used as a common platform for channel code comparison.

§ 1.4 Channel Capacity

- **<u>Additive White Gaussian Noise (AWGN) Channel</u>**



- Channel model $y_i = x_i + n_i$
- Mutual Information: $I(X,Y) = H(Y) - H(Y|X)$

$$= H(Y) - H(X+N|X)$$

$$= H(Y) - H(N|X)$$

$$= H(Y) - H(N)$$

- Capacity: $C = \max_{P(x)}\{I(X,Y)\}$

$$= \max_{P(x)}\{H(Y) - H(N)\}$$

§ 1.4 Channel Capacity

- **<u>Additive White Gaussian Noise (AWGN) Channel</u>**
- For AWGN: $\mathcal{N}(0, \sigma_N^2)$. Its pdf is

$$P(n) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{n^2}{2\sigma_N^2}\right)$$

$$\begin{aligned} H(N) &= -\int_{-\infty}^{+\infty} P(n) \log_2 P(n) \mathrm{d}n \\ &= -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{n^2}{2\sigma_N^2}\right) \log_2\left(\frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{n^2}{2\sigma_N^2}\right)\right) \mathrm{d}n \\ &= \frac{1}{2}\log_2(2\pi e \sigma_N^2) \quad \text{bits/symbol} \end{aligned}$$

- If input X is normal distributed (continuous) as $\mathcal{N}(\mu_X, \sigma_X^2)$, $I(X, Y)$ will be maximized and

$$C = H(Y) - H(N)$$

§ 1.4 Channel Capacity

- **Additive White Gaussian Noise (AWGN) Channel**
- For Input: $\mathcal{N}(\mu_X, \sigma_X^2)$. Its pdf is

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)$$

$$\begin{aligned} H(X) &= -\int_{-\infty}^{+\infty} P(x)\log_2 P(x)\mathrm{d}x \\ &= -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right) \log_2\left(\frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)\right) \mathrm{d}x \\ &= \frac{1}{2}\log_2(2\pi e \sigma_X^2) \quad \text{bits/symbol} \end{aligned}$$

- Since $Y = X + N$ and X and N are independent

Output: $\mathcal{N}(\mu_X, \sigma_X^2 + \sigma_N^2) = \mathcal{N}(\mu_X, \sigma_Y^2)$

$$H(Y) = \frac{1}{2}\log_2(2\pi e(\sigma_X^2 + \sigma_N^2)) \quad \text{bits/symbol}$$

§ 1.4 Channel Capacity

- **Additive White Gaussian Noise (AWGN) Channel**
- Channel model: $y_i = x_i + n_i$
- Capacity: $C = H(Y) - H(N)$

$$= \frac{1}{2}\log_2\left(2\pi e(\sigma_X^2 + \sigma_N^2)\right) - \frac{1}{2}\log_2(2\pi e\sigma_N^2)$$

$$= \frac{1}{2}\log_2\left(1 + \frac{\sigma_X^2}{\sigma_N^2}\right) \text{ bits/symbol}$$

- σ_X^2 is the power of the transmitted signal, while σ_N^2 is the power of noise. Hence, $\frac{\sigma_X^2}{\sigma_N^2}$ is often defined as the signal-to-noise ratio (SNR).
- This only defines the inachievable transmission limit since in practice, X will not be normal distributed.

§ 1.4 Channel Capacity

- **<u>Band Limited AWGN Channel</u>**
- In a practical system, sampling is needed at the receiver to reconstruct the received signal as Fig. 1.

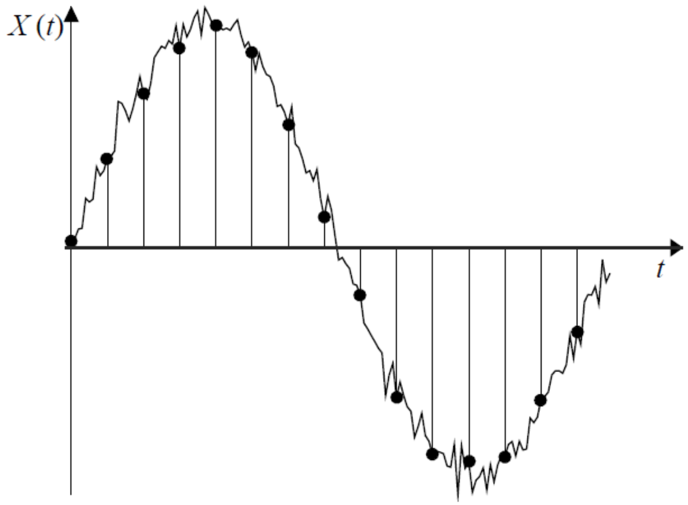


Fig. 1 Received Signal and Sampling

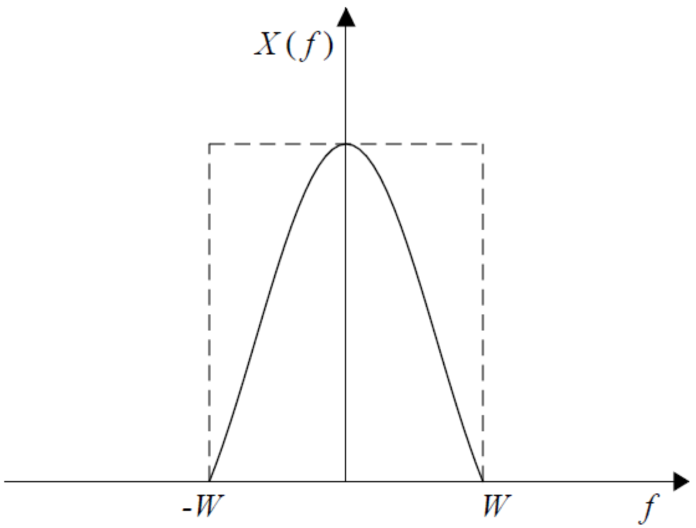


Fig. 2 Signal Sampling in frequency domain

- If the signal has a frequency of W, the sampling frequency should be at least $2W$ for perfect signal reconstruction. (Fig. 2)

§ 1.4 Channel Capacity

- **Band Limited AWGN Channel**
- With the sampling, we now have a series of time discrete Gaussian samples and the channel model becomes

$$y\left(t=\frac{s}{2W}\right)=x\left(t=\frac{s}{2W}\right)+n\left(t=\frac{s}{2W}\right), s=1,2,\cdots$$

- Signal $x\left(t=\frac{s}{2W}\right)$ has variance σ_X^2

  Noise $n\left(t=\frac{s}{2W}\right)$ has variance $\frac{N_0}{2}$, where N_0 is the noise power

- Capacity for each time discrete Gaussian channel

$$C_s=\frac{1}{2}\log_2\left(1+\frac{2\sigma_X^2}{N_0}\right) \text{ bits/symbol}$$

§ 1.4 Channel Capacity

- **Band Limited AWGN Channel**
- Capacity of this band limited AWGN channel can be determined by

$$C = \frac{\sum_{s=1}^{2WT} C_s}{T}$$, T–sampling duration

- Since the average signal power

$$E = \frac{2WT \cdot \sigma_X^2}{T} = 2W\sigma_X^2$$

$$C_s = \frac{1}{2}\log_2\left(1 + \frac{E}{WN_0}\right)$$ bits/symbol

- Capacity of band limited AWGN channel becomes

$$C = \frac{2WT \cdot \frac{1}{2}\log_2\left(1 + \frac{E}{WN_0}\right)}{T} = W\log_2\left(1 + \frac{E}{WN_0}\right)$$ bits/second

§ 1.4 Channel Capacity

- **Shannon Limit:** Error free transmission over the Gaussian channel is possible if the signal-to-noise ratio $\frac{E_b}{N_0}$ is at least -1.6 dB.

Proof:

- This possibility is sealed by the use of channel code (information length *k* bits, codeword length *n* bits).
- Let E_b and E_c denote the energy of each information bit and each coded bit, respectively. It is required

$$k \cdot E_b = n \cdot E_c$$

  so that adding redundancy does not increase the transmission energy.

- Consider each coded bit is carried by a modulated signal, e.g., using binary phase shift keying (BPSK),

$$E = E_c = \frac{E_b \cdot k}{n} = E_b \cdot r$$

§ 1.4 Channel Capacity

Continue the Proof

- Assume the signal frequency $W \to \infty$

$$C = \lim_{W \to \infty} W \log_2\left(1 + \frac{E}{N_0 W}\right)$$

$$= \frac{E}{N_0 \ln 2}$$

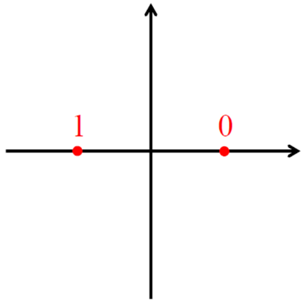
$$= \frac{E_b \cdot r}{N_0 \ln 2} \text{ bits/second}$$

- For error free transmission, it is required

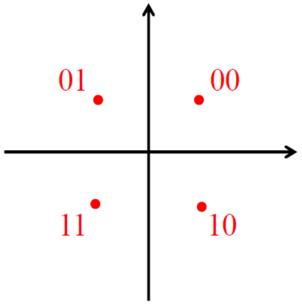
$$r < C \Longrightarrow \frac{E_b}{N_0} > \ln 2 = 0.69 = -1.6\text{dB}$$

§ 1.4 Channel Capacity

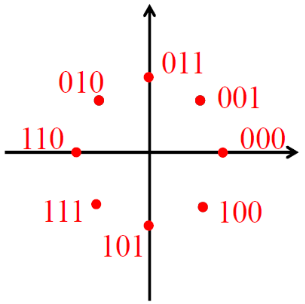
- **<u>AWGN Channel with Finite Modulation Alphabets</u>**
- In a wireless communication system, digital signal is modulated (mapped) to an analog signal for transmission.
- Commonly used modulation schemes include:



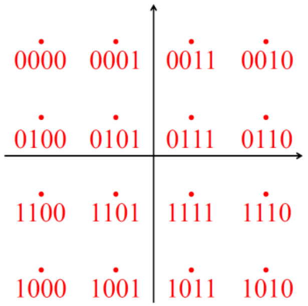
BPSK



QPSK



8PSK



16QAM

§ 1.4 Channel Capacity

- **AWGN Channel with Finite Modulation Alphabets**
- Input $X \in \{x_1, x_2, \ldots, x_M\}$, e.g., BPSK $M = 2$, QPSK $M = 4$, 8PSK $M = 8$, 16QAM $M = 16$, … .
- Channel Capacity

$$C = \max_{P(x)} \left\{ \sum_{i=1}^{M} \int_{y:-\infty}^{+\infty} P(x_i, y) \log_2 \frac{P(x_i|y)}{P(x_i)} \mathrm{d}y \right\}$$

Since

$$P(x_i, y) = P(y|x_i)P(x_i)$$

$$P(x_i|y) = \frac{P(y|x_i)P(x_i)}{P(y)}$$

$$P(y) = \sum_{i'=1}^{M} P(y|x_{i'})P(x_{i'})$$

§ 1.4 Channel Capacity

- **AWGN Channel with Finite Modulation Alphabets**

$$C = \max_{P(x_i)} \left\{ \sum_{i=1}^{M} P(x_i) \int_{y:-\infty}^{+\infty} P(y|x_i) \log_2 \frac{P(y|x_i)}{\sum_{i'=1}^{M} P(x_{i'}) P(y|x_{i'})} \mathrm{d}y \right\}$$

- Assume each modulated symbol is equally likely to be transmitted

$$P(x_i) = P(x_{i'}) = \frac{1}{M}.$$

- Capacity:

$$C = \frac{1}{M} \sum_{i=1}^{M} \int_{y:-\infty}^{+\infty} P(y|x_i) \log_2 \frac{P(y|x_i)}{\frac{1}{M} \sum_{i'=1}^{M} P(y|x_{i'})} \mathrm{d}y \ \text{bits/symbol}$$

§ 1.4 Channel Capacity

- **AWGN Channel with Finite Modulation Alphabets**
- Over the AWGN Channel $y = x_i + n$

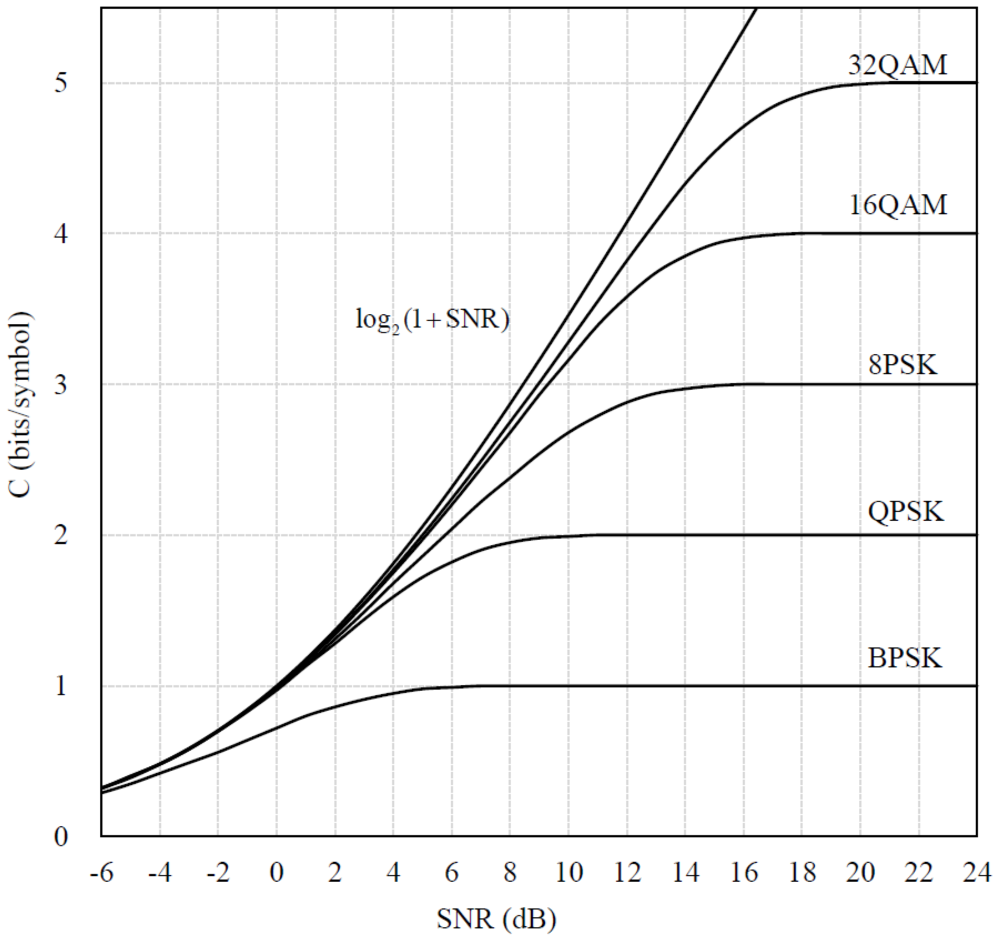
$$P(y|x_i) = \frac{1}{\sqrt{2\pi}\sigma_N}\exp\left(-\frac{|y-x_i|^2}{2\sigma_N^2}\right)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_N}\exp\left(-\frac{|n|^2}{2\sigma_N^2}\right)$$

- Capacity:

$$C = \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left[\log_2\frac{P(y|x_i)}{\frac{1}{M}\sum_{i'=1}^{M}P(y|x_{i'})}\right]$$
$$= \log_2 M - \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left[\log_2\sum_{i'=1}^{M}\exp\left(-\frac{|x_i+n-x_{i'}|^2-|n|^2}{2\sigma_N^2}\right)\right] \text{ bits/symbol}$$

§ 1.4 Channel Capacity

- **AWGN Channel with Finite Modulation Alphabets**



References:

[1] Elements of Information Theory, by T. Cover and J. Thomas.

[2] Scriptum for the lectures, Applied Information Theory, by M. Bossert.