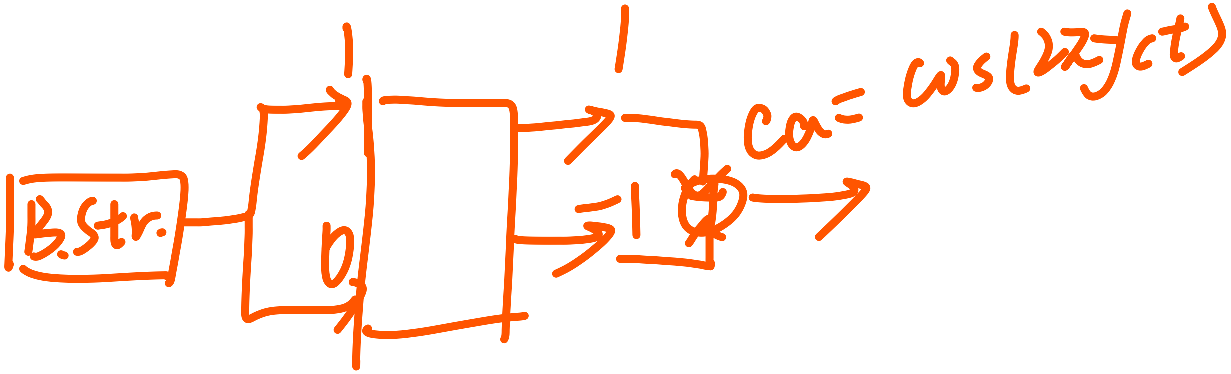
UESTC4004
Digital Communications

QPSK – Capacity, Entropy & Source Coding

Let's refresh!

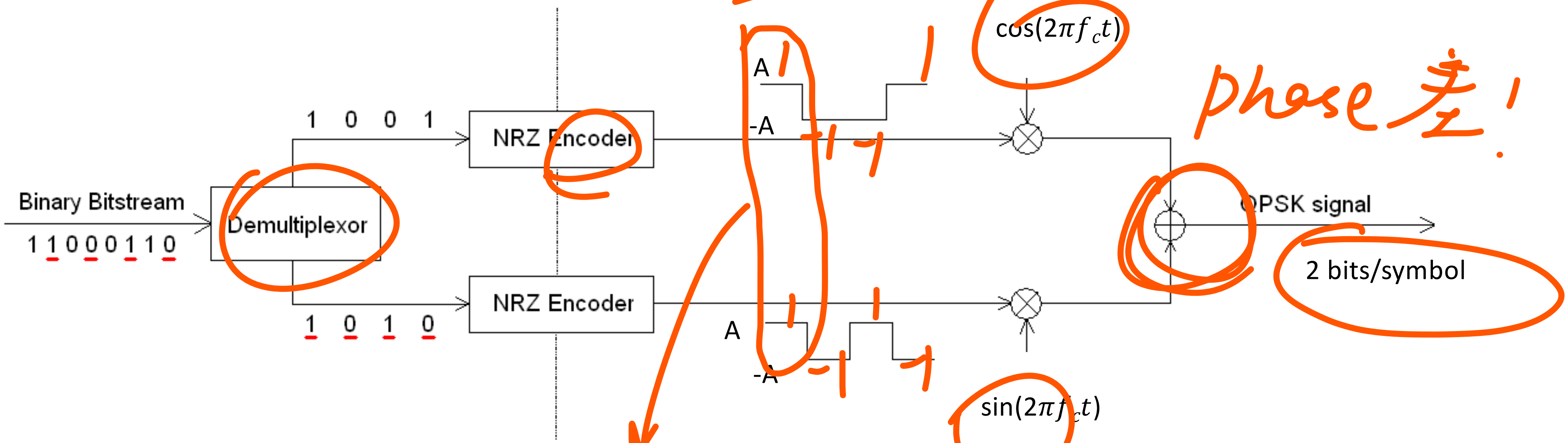
- Draw the diagram for the Binary PSK modulator



4-ary PSK or QPSK (Quadrature Phase Shift Keying)

Similar to transmitting 2 BPSK signals simultaneously

11000110



Phase 差!

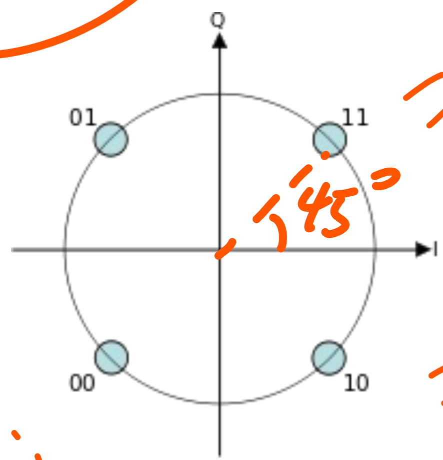
Source: Wikipedia https://en.wikipedia.org/wiki/Phase-shift_keying

价码元.

QPSK –Constellation Diagram

-1+j

= 1+j



45°

-1-j

= 1-j

Source: Wikipedia https://en.wikipedia.org/wiki/Phase-shift_keying

QPSK

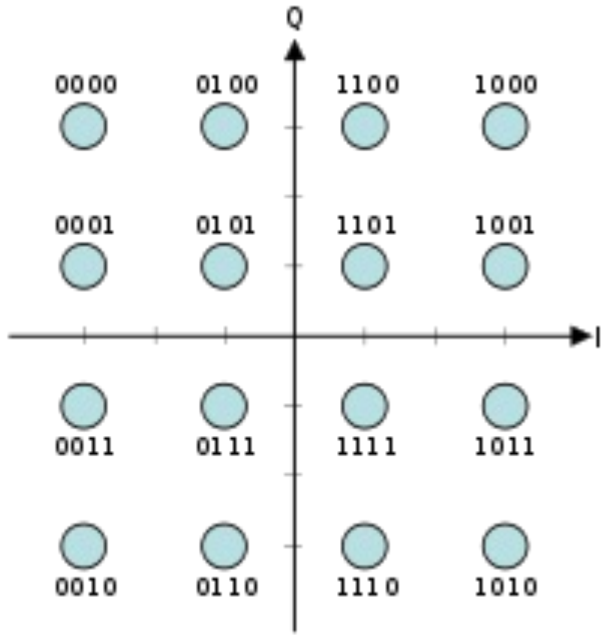
- **Advantages:** higher data rate than BPSK (2 bits per symbol interval), while bandwidth occupancy remains the same.
- 4-PSK can be easily extended to 8-PSK or more.
- However, higher rate PSK schemes are limited by the ability of equipment to distinguish small differences in phase.

QPSK Example

- A binary stream is modulated using binary Phase Shift Keying (PSK). The same bit stream is modulated using Quadrature Phase Shift Keying (QPSK).
  - Which one of these two modulated signals would require the least channel bandwidth, if not same? Justify your answer.
  - Which modulation would provide better data rate, if not same? Justify your answer.

Quadrature Amplitude Modulation (QAM)

- Information in both amplitude and phase



16 - QAM

University of Glasgow

Goals in designing a Digital Communication Systems

- Maximizing the transmission bit rate ↑
- Minimizing probability of bit error ↓
- Minimizing the required power ↑
- Minimizing required system bandwidth ↓
- Maximizing system utilization ↑
- Minimize system complexity ↓

- One of the limitations to achieve these goals is
  - The **Shannon-Hartley capacity** theorem (and the Shannon limit)

R 传输速率
BER 误码率 P_B
E_b/N_0 信息能量
B(W) 信道带宽
users 系统使用
C 系统复杂度

bits/s/Hz
效率

Shannon limit

信道容量

- Channel capacity: The maximum data rate at which error-free communication over the channel is performed.

最大上限速率

- The Shannon theorem puts a limit on the transmission data rate
- Channel capacity of AWGN channel (Shannon-Hartley capacity theorem):

SNR

$$\boxed{C = W \log_2\left(1+\frac{S}{N}\right) \quad \text{[bits/s]}}$$

W [Hz] Bandwidth

$S = E_b C$ [Watt]: Average received signal power

$N = N_0 W$ [Watt]: Average noise power

Shannon limit

$$C = W \log_2\left(1+\frac{S}{N}\right) \qquad \text{[bits/s]}$$

Since

$S = E_b C$ $\qquad$ $N = N_0 W$

$$\boxed{\frac{C}{W} = \log_2\left(1+\frac{E_b}{N_0}\frac{C}{W}\right)}$$

$$\boxed{\frac{S}{N} = \frac{E_b C}{N_0 W}}$$

- The ratio C/W represents the bandwidth efficiency (bits/second/Hz) of the communication system and the ratio E_b/N_0 represents the power efficiency of the communication system.

University of Glasgow

Example

For a signal to noise ratio of *12dB* and available channel bandwidth of *2 kHz*, we wish to transmit information at a rate of *10 kbps*. By calculating the Shannon capacity for the given channel, analyse if the channel can support error free communication

$$20\log_{10}\left(\frac{S}{N}\right) = 12dB \Rightarrow \frac{S}{N} = 10^{\frac{12}{20}}$$

$$W = 2kHz$$

$$\Rightarrow C = W\log_2\left(1+\frac{S}{N}\right)$$

$$= 2\times10^3 \log_2\left(1+10^{\frac{12}{20}}\right) = 4633 \text{ bit/s}$$

$$= 4.633 \text{ kbps}$$

消息中的信息量 I(x) 定义为熵

Entropy

•The average amount of information per message *I(X)* is also referred to as the source entropy, or the communication entropy of the source. It is usually denoted by the letter *H*. Thus

$$\boxed{H(X) = I(X) = \sum_{k=1}^{n} p_k I_k = -\sum_{k=1}^{n} p_k \log_2 p_k}$$

•Entropy is a measure of uncertainty, i.e., high entropy implies high uncertainty and accordingly, a high amount of information that can be transmitted

•For the binary source where the symbols have probabilities *α* and *1- α,* i.e., $X = \{0,1\}$ with $P\{0\} = a$ and $P\{1\} = 1 - a$ we have

$$\boxed{H(X) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)}$$

$$-(P_1 \log_2 P_1 + P_2 \log_2 P_2) = -(\alpha \log_2 \alpha + (1-\alpha)\log_2(1-\alpha))$$

Entropy

- The entropy is based on the probability, p, of an event.
- This can also be looked at as the randomness of successive events or how correlated individual events are.
- Note that maximum entropy is achieved when the probability is 50%
  - A sample provides no information about a succeeding sample.

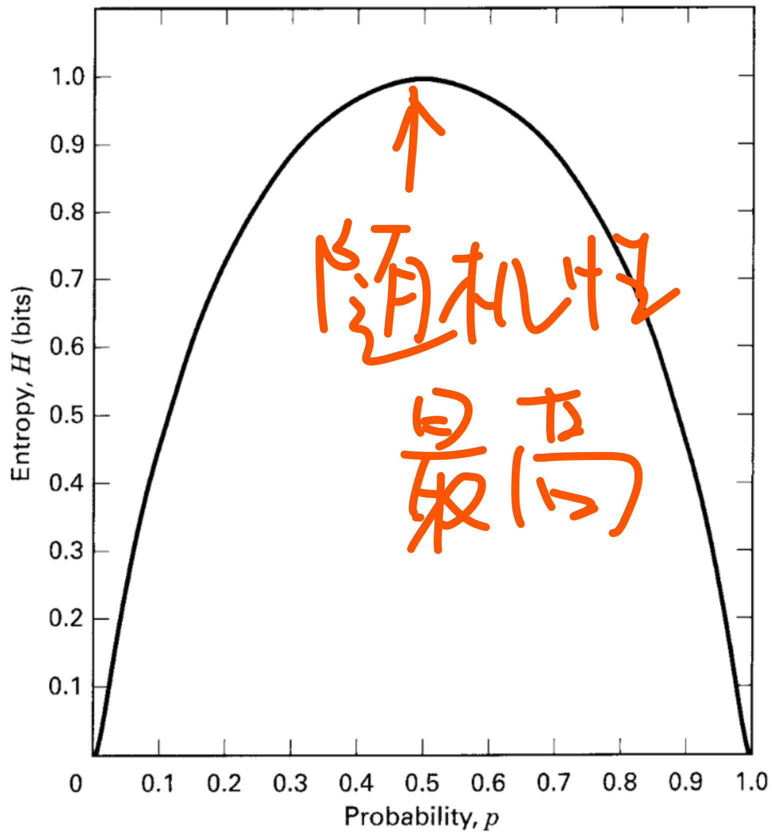


Figure 9.5 Entropy versus probability (two events).

Example

Compare the entropies of two sources, Source A and Source B, of 4 messages each. Source A transmits the 4 messages with equal probabilities. While Source B transmits the messages with probabilities {0.4, 0.3, 0.2, 0.1}.

$$H_A(X) = -\sum p_k \log_2 p_k = -(\cancel{0.25} \log_2 0.25 \times \cancel{4}) = -\log_2 2^{-2}$$

$$= 2$$

$$H_B(X) = -\sum p_k \log_2 p_k = -(0.4 \log_2 0.4 + 0.3 \log_2 0.3$$

$$+ 0.2 \log_2 0.2 + 0.1 \log_2 0.1)$$

$$= 1.85$$

University of Glasgow

Source Coding

$H(X) < \bar{L} \Rightarrow eff < 1$

•If the entropy of the source is less than the average word length of the original alphabet, this means that the original coding is redundant and that the original information may be compressed by an efficient source coding algorithm. The source coding efficiency is defined as the ratio of the source entropy to the average word length.

编码效率

$$\boxed{\text{Efficiency} = \frac{H(X)}{\bar{L}}}$$

平均字长

•The main idea behind the compression is to create such a code, for which the average length of the encoding vector (word) will not exceed the source entropy. In general, this means the codes that are used for compression are not uniform, i.e., the words have different number of bits.

redundant ⟷ efficient

Source Coding

•The most common source coding algorithms are the Shannon-Fano and Huffman used for discrete memoryless information sources.

•Shannon-Fano source coding is the first established and widely used coding method.

•Huffman source coding constructs binary codes with minimum redundancy (i.e., maximum efficiency) for a set of discrete messages.

•Both coding algorithms belong to the category of lossless coding as the source data can be perfectly reconstructed from the compressed data.

University of Glasgow

Shannon – Fano Source Coding

•We consider the original messages that are to be transmitted with their corresponding probabilities, i.e.,

$$[X]=[x_1, x_2, \ldots, x_n] \qquad [P]=[p_1, p_2, \ldots, p_n]$$

•More frequent messages are coded by a shorter sequence and less frequent messages are coded by a longer sequence. 核心特点：概率越大越短编码

•The messages of the information source must be arranged in order from most probable to least probable. 排序

•Then the initial set of messages must be divided into two subsets whose total probabilities are as close as possible to being equal; symbols in the first set receive "0" and symbols in the second set receive "1". 分组

•The same process is repeated on those subsets, to determine the next digits of a symbol.

•When a subset has been reduced to one symbol, this means the symbol's code is complete.

Shannon – Fano Source Coding - Example

1 − 0.5 = 0.5

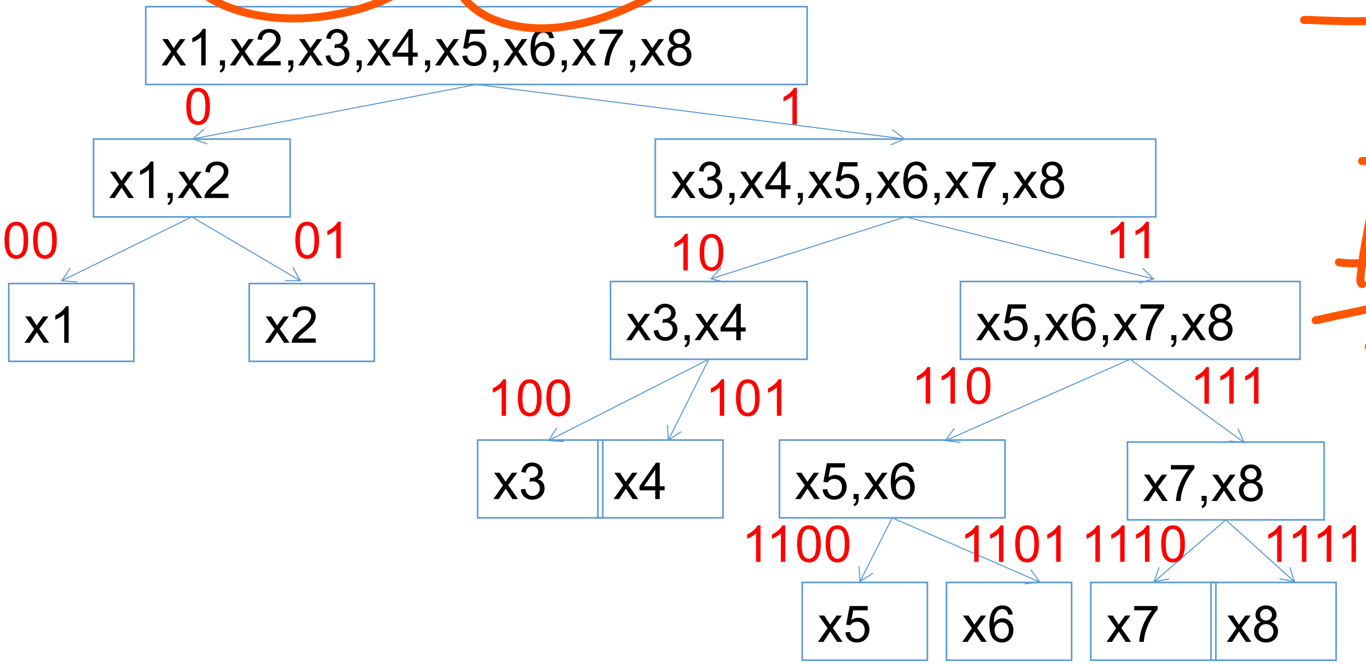
0.5

2

Message	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Probability	0.25	0.25	0.125	0.125	0.0625	0.0625	0.0625	0.0625

排序

再分组



University of Glasgow

Shannon – Fano Source Coding - Example

Message	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Probability	0.25	0.25	0.125	0.125	0.0625	0.0625	0.0625	0.0625
Encoding sequence	00	01	100	101	1100	1101	1110	1111

Source entropy:

$$H(X) = -\sum P\{x_i\}\log P\{x_i\} = -\left[2\cdot\left(\frac{1}{4}\log_2\frac{1}{4}\right)+2\cdot\left(\frac{1}{8}\log_2\frac{1}{8}\right)+4\cdot\left(\frac{1}{16}\log_2\frac{1}{16}\right)\right] = 2.75$$

Average length of the encoding vector:

定义

$$\boxed{\bar{L} = \sum P\{x_i\} n_i} = \left[2\cdot\left(\frac{1}{4}\cdot 2\right)+2\cdot\left(\frac{1}{8}\cdot 3\right)+4\cdot\left(\frac{1}{16}\cdot 4\right)\right] = 2.75$$

n_i 是编码长度 of 每个 mess.

•The efficiency for this example is 100%. We also see that the direct uniform plain binary encoding (3 bits/symbol) is redundant.

$$\bar{L} = \sum_{n=1} P_i n_i$$

University of Glasgow

Example

先排序再分组

- Consider the symbols “a”, “b”, “c”, “d” with the following probabilities of occurrence: p(a)=0.45, p(b)=0.3, p(c)=0.15, p(d)=0.1. Calculate the compression efficiency when the symbols are coded using Shannon-Fano source coding algorithm.

X	a	b	c	d
P	0.45	0.3	0.15	0.1

$$H(x) = -\sum p_k \log_2 p_k$$

$$=$$

$$\bar{L} = \sum p_k n_k =$$

0 a

1 b c d

10 b

11 c d

110 c

111 d

$$\frac{S}{N} = \frac{E_b C}{N_0 W}$$

每比特平均能量 → E_b

信道容量 → C

信道宽度 → W

noise能量 → N_0

Up next!

- Dr Kayode will take next **5 lectures**
- I’ll see you at the end of **October**
- Keep posting on **Moodle Forum** for any questions or discussions

带宽效率

$$\frac{C}{W} = \log_2\left(1+\frac{E_b C}{N_0 W}\right)$$

① 香农极限：信道容量 $C = W\log_2\left(1+\frac{S}{N}\right)$

② 信息熵：$H(x) = I(x) = -\sum P_k \log_2 P_k$

③ 信源编码：香农法诺编码：排序+分组 (方式?)

效率 $E = \frac{H(x)}{L}$ ⇒ 平均字长 $= \sum P_k n_k$ (编码长)