

# Chapter 8. Digital Communications

Summary: Brief introduction of the components of digital communication systems, the fundamental rules in digital communication designs.

Assignment Project Exam Help

8.1 Source Coding

8.2 Channel coding

8.3 Binary Shift Keying Modulation

8.4 M-ary Digital Modulation Scheme

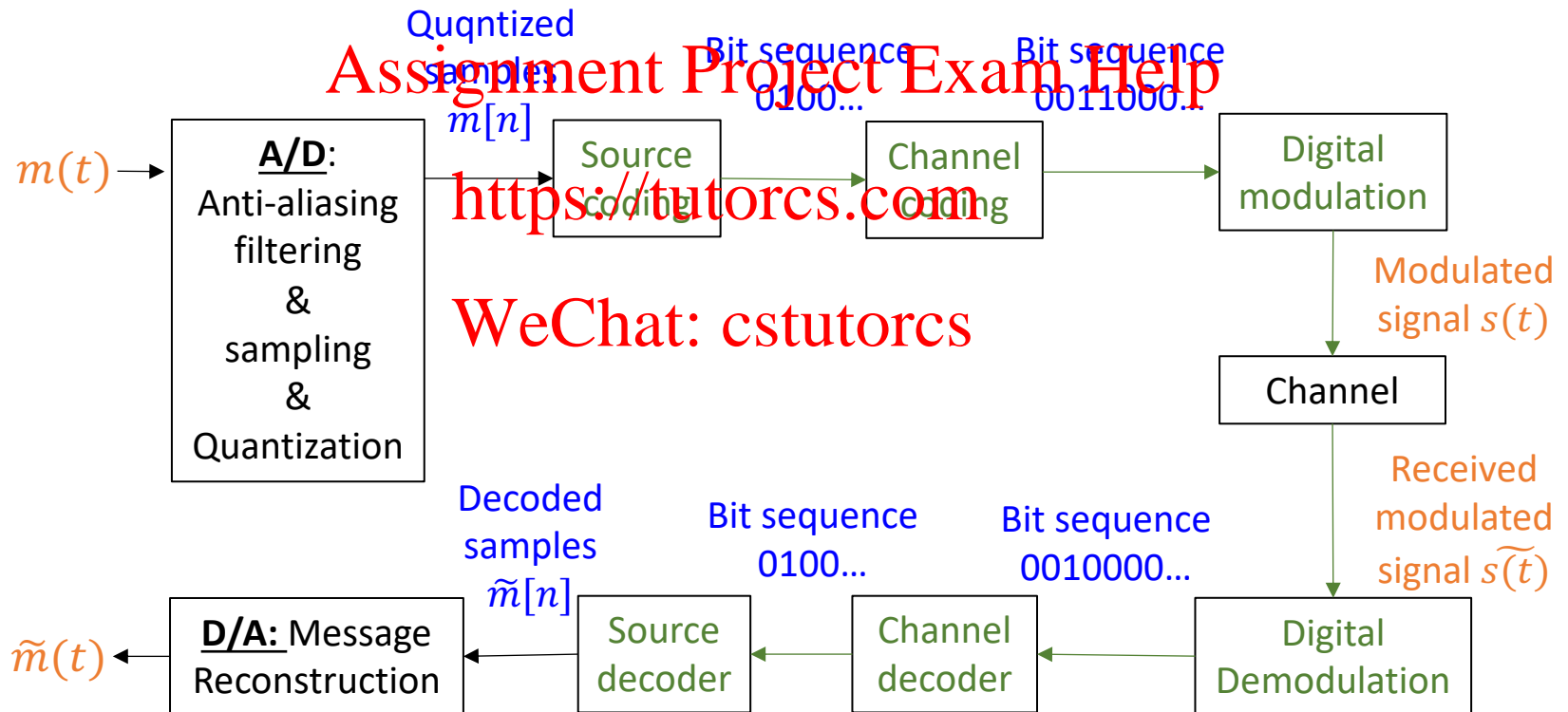
8.5 Constellation Design

8.6 Detection Design

<https://tutorcs.com>

WeChat: tutorcs

**Digital communication:** Signal to be communicated is digital and discrete-time. Communication of symbols with finite possible values.



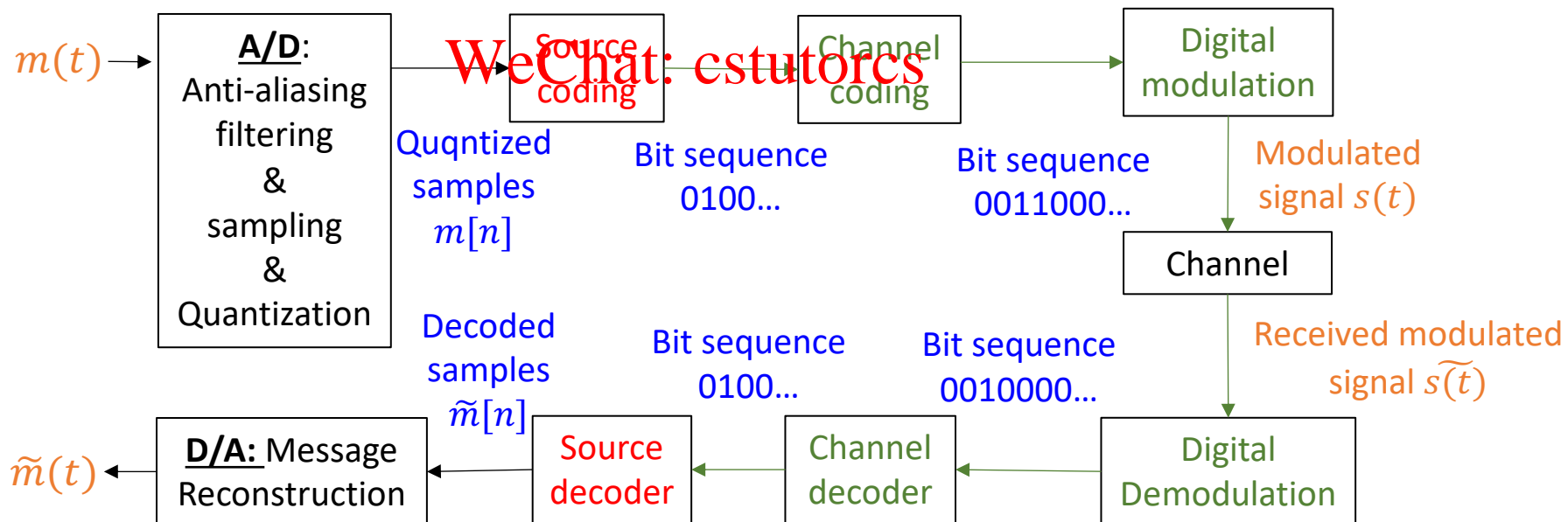
## 8.1 Source Coding

**Source encoding/data compression:** Efficiently convert message into a binary sequence.

- Remove redundancy in message & achieve low bitrate

**Source decoding:** Interpret the information sequence into message (Reconstruct the original message).

<https://tutorcs.com>



## Coding categories:

- Fixed-length coding
- Variable-length coding

### Fixed-length binary code:

**Example.** Extended ASCII (American Standard Code for Information Interchange) code.

Alphabet	Codeword	Length
A	01000001	8
O	01001111	8
...	...	8
1	00110001	8
...	...	8
\	01011100	8
...	...	8

## Variable-length binary code:

Statistics of message source can be used to design good source encoding algorithms.

- principles: use long codewords for less frequent symbols; short codewords for more frequent elements

Assignment Project Exam Help

<https://tutorcs.com>

Average length of codeword:  $\bar{l} = \sum_i p(i)l_i$

$\mathcal{A}$	Prob.	Code 1	Code 2
$a$	$1/2$	00	0
$b$	$1/4$	01	10
$c$	$1/8$	10	110
$d$	$1/8$	11	111
Ave. len. $\bar{l}$		2	1.75

WeChat: cstutorcs

e.g. for codes on the left,

$$\bar{l}_{code1} = \frac{1}{2} * 2 + \frac{1}{4} * 2 + \frac{1}{8} * 2 + \frac{1}{8} * 2 = 2$$

$$\bar{l}_{code2} = \frac{1}{2} * 1 + \frac{1}{4} * 2 + \frac{1}{8} * 3 + \frac{1}{8} * 3 = 1.75$$

## Variable-length binary code:

Huffman coding:

- Example of using long codewords for less frequent symbols; short codewords for more frequent elements.
- Greedy algorithm: Build a Huffman tree where two nodes with the smallest probabilities merge to an intermediate.

$\mathcal{A}$	Prob.
$a$	$1/2$
$b$	$1/4$
$c$	$1/8$
$d$	$1/8$

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

**Example:** Please design a Huffman code for the following source  
(each letter is shown with its probability in the bracket)  
x(3/13), m(3/13), g(3/13), s(1/13), h(1/13), p(1/13), r(1/13)

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

**Entropy:** For a source  $X$  with certain alphabet and distribution. Its entropy  $H(X)$  is defined as

$$H(X) = E[-\log_2 p(X)].$$

**Intuitive example:** Consider weather forecast with 4 possible outcomes: cloudy, snowy, rainy, and sunny.

- For place A, all possible weathers occurs with equal probability;

→ need 2 bits to delivery a message

- For place B, snowy and raining never happen but cloudy and sunny occurs evenly.

→ Need 1 bit only to delivery a message because less information needs to communicate.



**Entropy:** For a source  $X$  with certain alphabet and distribution. Its entropy  $H(X)$  is defined as

$$H(X) = E[-\log_2 p(X)].$$

**The meaning of entropy (Shannon's theorem)**

- A source can be compressed without distortion to  $H(X)$  bits per symbol.
- $H(X)$  bits per symbol is the lowest that a source can be compressed without distortion.
- The entropy  $H(X)$  provides a quantitative measure on the amount of information.

In general, for a **random source** **X** which generates alphabet

$$\mathcal{A} = \{x_1, x_2, \dots, x_N\}$$

with the following probability mass function (PMF):

$$p(x_1) = p_1, \quad p(x_2) = p_2, \quad \dots, \quad p(x_N) = p_N$$

**the entropy** of **X** is:

$$\begin{aligned} H(X) &= \mathbb{E}[-\log_2 p(X)] = - \sum_{x \in \mathcal{A}} p(x) \log_2 p(x) = \sum_{x \in \mathcal{A}} p(x) \log_2 [1/p(x)] \\ &= -p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_N \log_2 p_N. \end{aligned}$$

- Entropy is non-negative.
- With base-2 log, the unit is bit.
- With the log-function, the joint entropy of 2 independent sources equals the sum of their entropies individually.

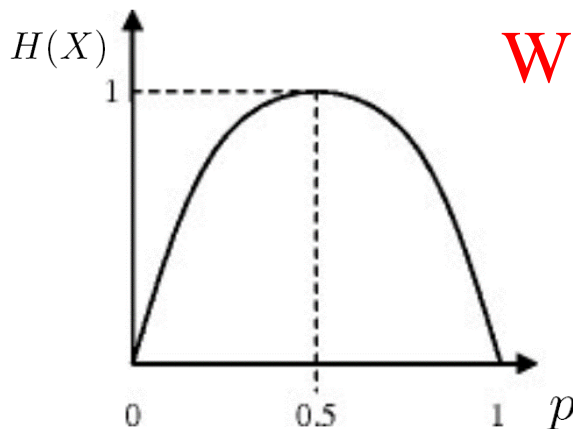
**Example:** a binary random source  $X$  generates binary bits randomly following some probabilities:

$$p(0) = P[X = 0] = p, \quad p(1) = P[X = 1] = 1 - p.$$

The entropy of the *binary* source is

$$H(X) = -p \log_2 p - (1 - p) \log_2 (1 - p).$$

<https://tutorcs.com>



WeChat: estutorcs

- When  $p \approx 0$  or  $p \approx 1$ , little information is in  $X$ . The entropy achieves its minimum 0.
- When  $p \approx 1/2$ , the maximum amount of information is in  $X$ . The entropy is about its maximum 1.

**Example:** A source whose alphabet  $\mathcal{A}$  and its probability distribution are as follows. Calculate the entropy.

$$\mathcal{A} = \{a, b, c, d\}.$$

$$p(a) = 1/2, p(b) = 1/4, p(c) = 1/8, p(d) = 1/8.$$

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

**Example:** Consider the following source  $X$  and codes.

$\mathcal{A}$	Prob.	Code 1	Code 2	Code 3	Code 4
$a$	$1/2$	00	0	0	0
$b$	$1/4$	01	1	10	01
$c$	$1/8$	10	0	110	011
$d$	$1/8$	11	1	111	111
Ave. len. $\bar{l}$		2	1	1.75	1.75

$$H(X) = 1.75 \text{ bits/symbol.}$$

**Code 1:** Uniquely decodable. Fixed-length codeword. Not the shortest.

**Code 2:** Not uniquely decodable. Shortest.

**Code 3:** Uniquely decodable.

Prefix-free code: no codeword is a prefix of another.

Instantaneously decodable.

Optimal in the sense of achieving the entropy bound:  $\bar{l} = H(X)$ .

**Code 4:** Uniquely decodable. Not prefix-free code.

Not instantaneously decodable: need to see following bits to decode.

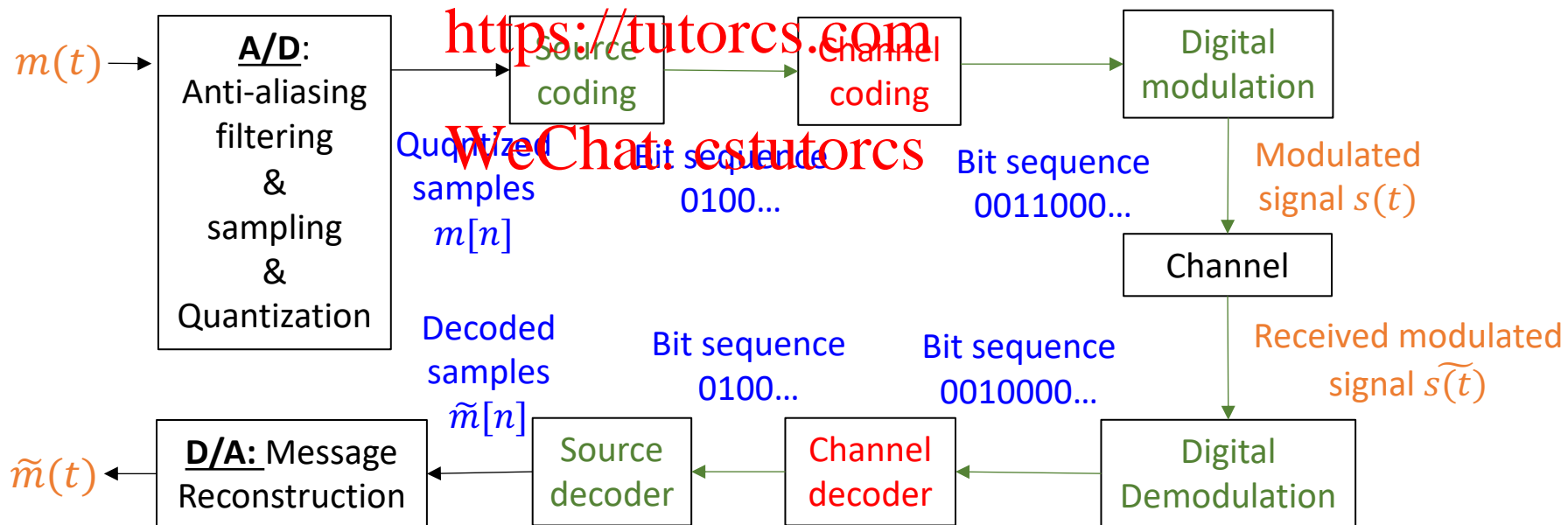
Optimal:  $\bar{l} = H(X)$ .

Data compression idea: Use long codewords for less frequent elements; short codewords for more frequent elements.

## 8.2 Channel Coding

**Channel encoding:** Introduce controlled redundancy (extra information) to combat noise/distortions in channels.

**Channel decoding:** Remove the controlled redundancy and Reconstruct the compact bit sequence.



## Repetition codes : most basic error-correcting codes

**Example:** 3-bit repetition codes

- Encoding: 0 → 000, 1 → 111

- Decoding: Majority vote

Original bit sequence	0	0	1	0	1
After repetition codes	000	000	111	000	111
Received bit sequence	000	001	111	000	101
after decoding	0	0	1	0	1

**Information rate** = # message bit / # bit after channel coding

- Introducing more bits in channel coding may help to handle error, with the penalty of lower information rate and longer bit sequence to be sent.



## (7,4) Hamming code (1950)

- Take 4 information bits:  $d_1, d_2, d_3, d_4$
- Add 3 parity check bits:  $p_1, p_2, p_3$
- Form a 7-bit codeword:  $d_1 d_2 d_3 d_4 p_1 p_2 p_3$
- Can correct 1 bit error.

XOR operation:  $\oplus$

$$0 \oplus 0 = 1 \oplus 1 = 0$$

$$1 \oplus 0 = 0 \oplus 1 = 1$$

Assignment Project Exam Help

Encoding procedure:

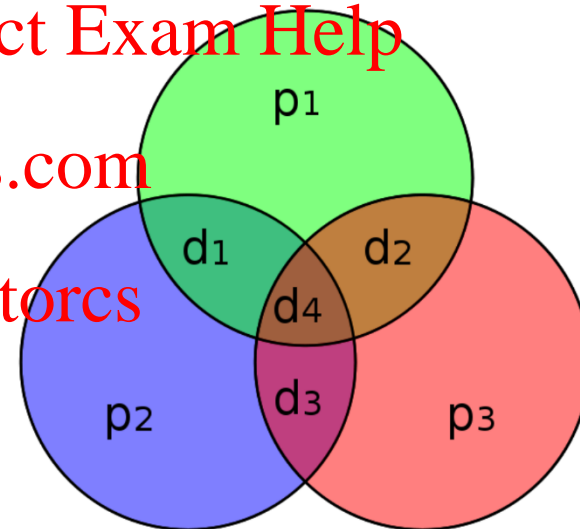
$$p_1 = d_1 \oplus d_2 \oplus d_4$$

$$p_2 = d_1 \oplus d_3 \oplus d_4$$

$$p_3 = d_2 \oplus d_3 \oplus d_4$$

<https://tutorcs.com>

WeChat: cstutorcs

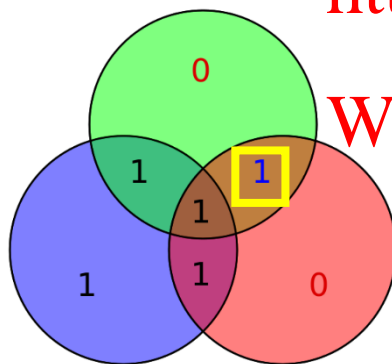
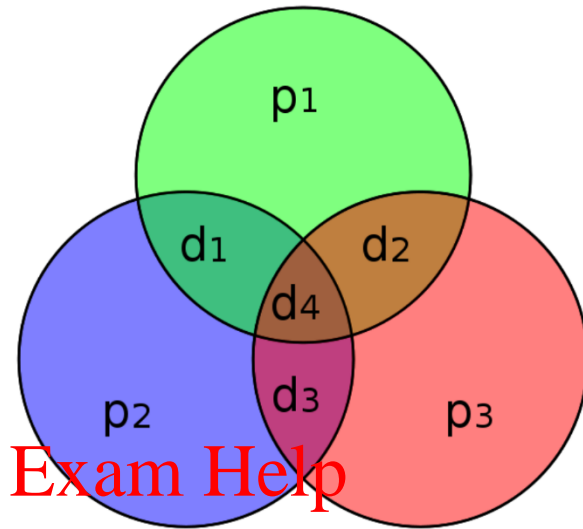


Every circle has an even number of 1's.

[https://en.wikipedia.org/wiki/Hamming\(7,4\)](https://en.wikipedia.org/wiki/Hamming(7,4))

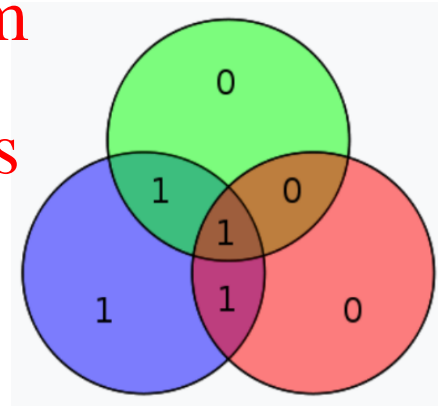
## Decoding procedure:

Put bits in their corresponding positions in the figure. See which circles violates even-parity check. Flip the corresponding bit (1 bit only) to ensure the parity checks.



<https://tutorcs.com>

WeChat: cstutorcs



Data sequence: 1011 → Encoded sequence: 1011010  
→ received sequence (1 bit error) 1111010  
→ decoded sequence: 1011010

Comments:

- Code rate

= number of info. bits / total number of bits of a codeword

=4/7.

Assignment Project Exam Help

- Can correct 1 bit error but no more.

<https://tutorcs.com>

- Repetition code would need 12 bits to protect 4 bits against 1 error.

WeChat: cstutorcs

Applications of Hamming codes:

- DRAM memory chips
- Satellite communication

**Example:** (a) Using (7,4) Hamming code to transmit 0110.  
What is the coded segment?

(b) Assume 1001111 is received. Please find the decoding result. Is there any error in the received signal?

(c) If the probability of each bit error is 1% and bit errors are independent. Please calculate the probability that the decoding results is correct in (7,4) Hamming code.

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

## How much redundancy can be introduced?

There is an upper limit of amount of redundancy constrained by **channel capacity**.

**Channel capacity  $C$ :** the maximum bit rate of a digital communication systems.

- Characteristic of a channel
- It is a fundamental limit regardless of the coding.

**WeChat: cstutorcs**

**For example:** After the source coding, the bitrate of a message is 0.6Mbps. If we want to transmit the message through a channel with capacity  $C = 1\text{Mbps}$ . Then the maximum redundancy we can introduced by the channel coding is  $1 - 0.6 = 0.4\text{Mbps}$ .

## Communication system design in coding ? (Cont.)

3 factors to be considered:

- The channel bandwidth  $B$ 
  - Measured by Hz
- The number of levels  $M$  in digital signals
- The quality of the channel
  - Characterized by channel capacity  $C$
  - Closely related to channel noise level SNR

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Two theorems to relate above three factors:

- Nyquist's channel theorem
- Shannon's channel coding theorem

## Communication system design in coding ? (Cont.)

### Nyquist's Channel Theorem:

For a noiseless channel, there is an upper limit of bit rate (called the **Nyquist bit rate**) for error-free communication.

$$\text{Nyquist bit rate} = 2 * B * \log_2 M$$

<https://tutores.com>

Nyquist bit-rate is the maximum bit rate to represent a message represented by M levels for reliable communication in noiseless channel.

Ex. Consider a noiseless channel with a bandwidth of 3000Hz. We want to transmit a message with 4 signal levels. What is the Nyquist bit rate?  $\rightarrow 2 * 3000 * \log_2 4 = 12000\text{bps}$

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs



## Communication system design in coding ? (Cont.)

### Shannon's Channel Coding Theory:

For any given noisy channel, it is possible to communicate digital information nearly error-free when the communicate bit rate is kept below the channel capacity.

<https://tutorcs.com>

The channel capacity is usually determined by the noise level and available bandwidth.

WeChat: cstutorcs

$$C = B * \log_2(1 + SNR)$$

State-of-the-art: Since around 2000, with advancements in low-density parity-check (LDPC) codes and turbo codes, we are approaching the fundamental limit.

## More about SNR

$$SNR = \frac{\text{Power of signal (in Watt)}}{\text{power of noise (in Watt)}} = \frac{S}{N} \quad SNR_{dB} = 10 \log_{10}(SNR)$$

Q1: Consider a communication system for which the SNR at the receiver is  $SNR^0$ .

- (a) If  $SNR^0$  is doubled, how much the  $SNR_{dB}^0$  increases?
- (b)  $SNR^0$  should be multiplied by which factor to reduce  $SNR_{dB}^0$  by -3 dB?

Comment: Each 3 dB(-3 dB) increment (decrement) in  $SNR_{dB}$ , increases (decreases) the  $SNR$  by a factor of 2 (1/2).

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

**Example:** A telephone line normally has a bandwidth of 3000Hz (300Hz-3300Hz) assigned for data communication. The SNR is usually at the level of 35dB. If we want to transmit an signal through this channel,

- (a) Calculate the theoretical channel capacity;
- (b) Calculate the maximum signal level.

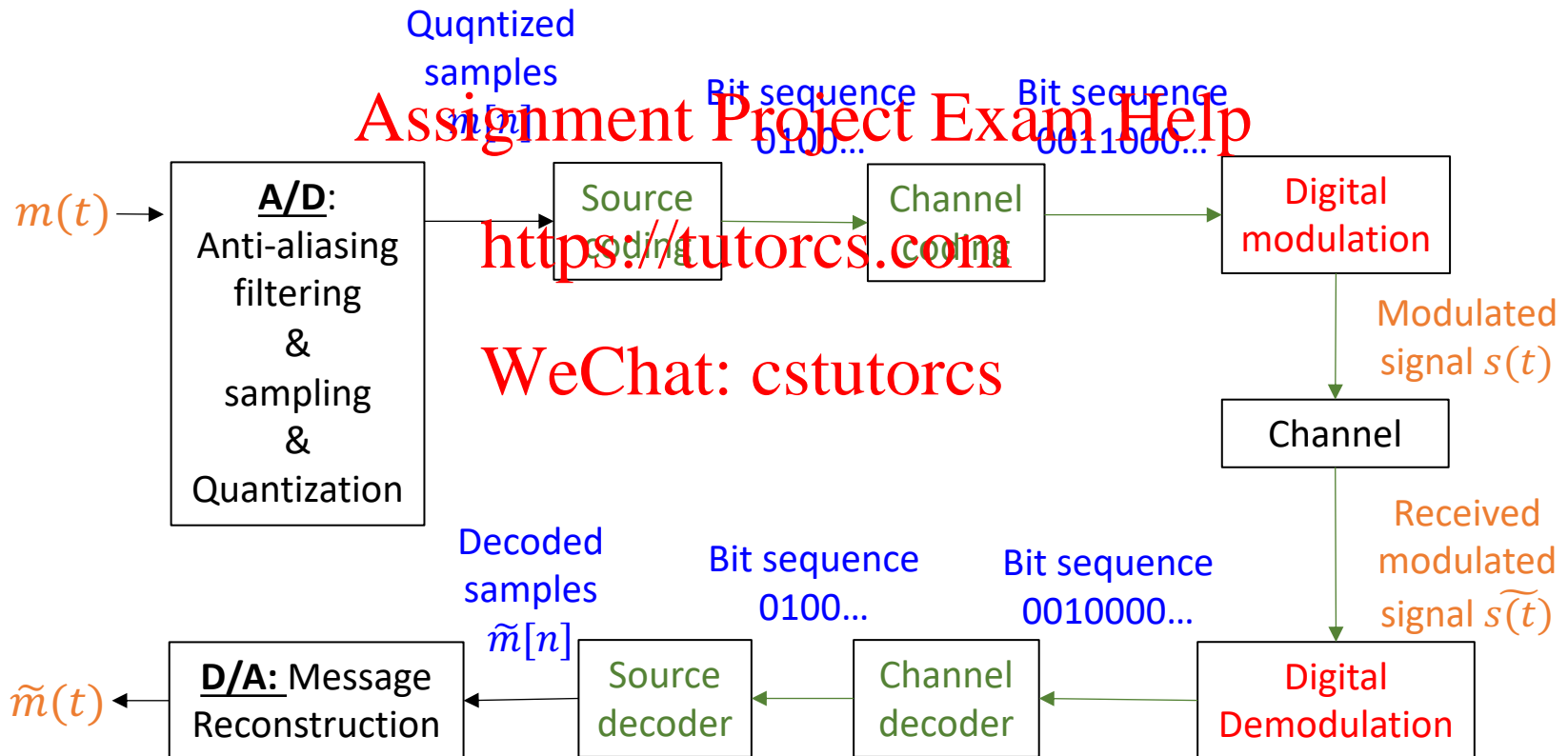
Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

## 8.3 Binary Shift Keying Modulation

Shift keying modulation: to convert digital information to analog signals for transmission



## Binary Shift Keying modulation

- Input: binary sequence  $b_i = 0/1$
- Carrier:  $c(t) = A_c \cos(2\pi f_c t + \varphi_c)$ 
  - Without any loss of generality,  $\varphi_c = 0$
- Output: analog modulated wave whose amplitude, phase, or frequency varying with  $b_i$

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

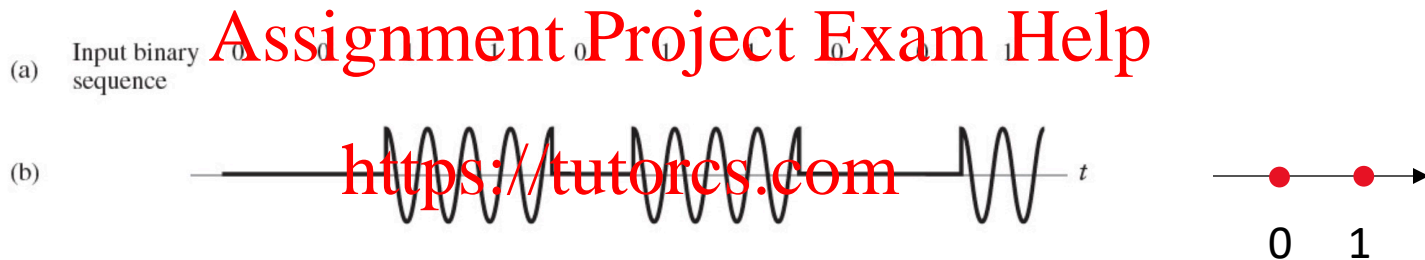


**FIGURE 7.1** The three basic forms of signaling binary information. (a) Binary data stream. (b) Amplitude-shift keying. (c) Phase-shift keying. (d) Frequency-shift keying with continuous phase.

## Binary amplitude Shift Keying (BASK)

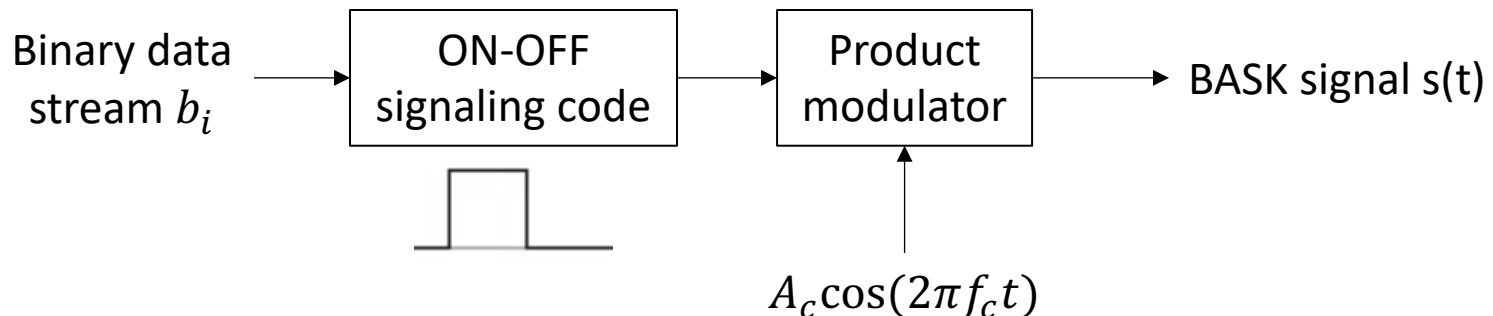
- Amplitude of the carrier varies in accordance with  $b_i = 0/1$

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t) & \text{for } b_i = 1 \\ 0 & \text{for } b_i = 0 \end{cases}$$



WeChat: cstutorcs

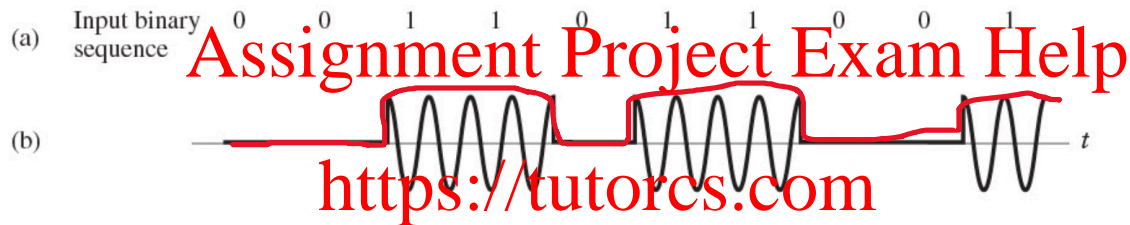
### BASK modulation



# Binary amplitude Shift Keying (BASK)

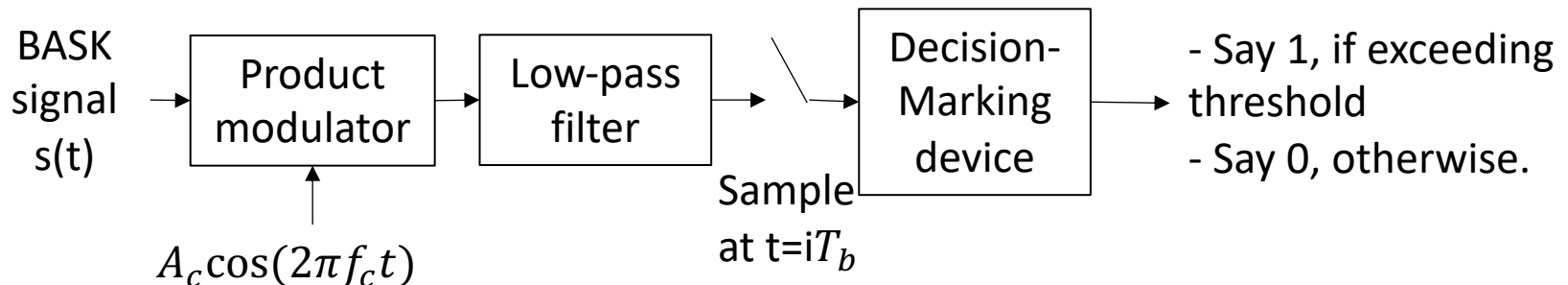
## BASK demodulation

- Envelop detector



tracing the nonconstant-envelope of BASK signals

- Coherence detector – synchronize phase and frequency

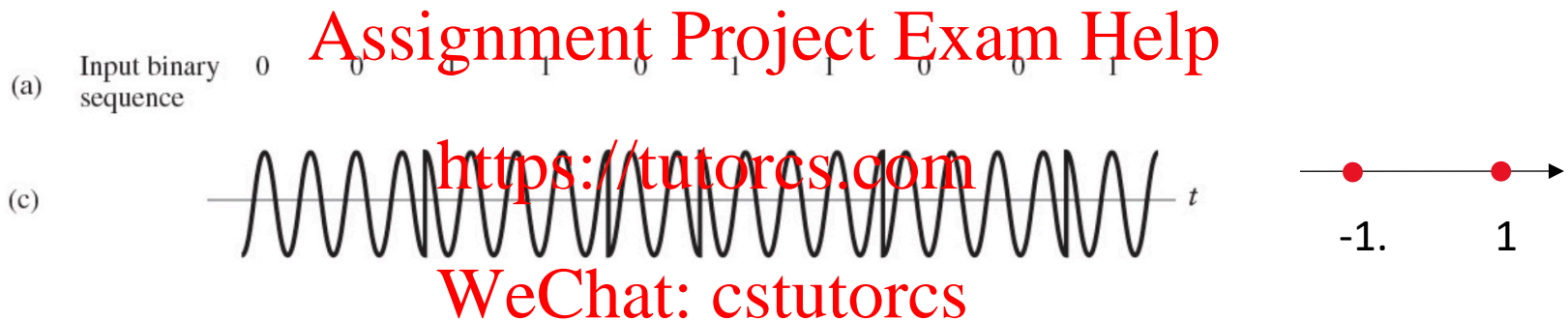




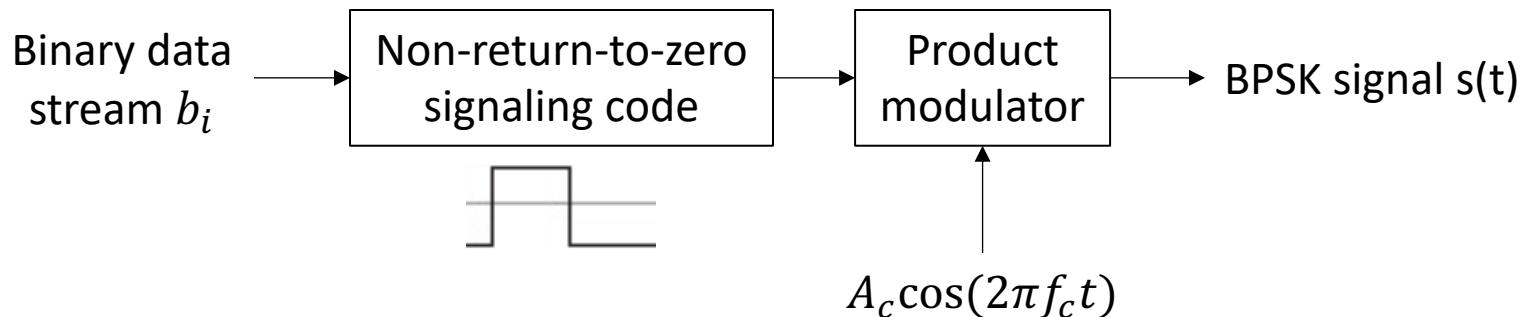
## Binary Phase Shift Keying (BPSK)

- Phase of the carrier varies in accordance with  $b_i = 0/1$

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t) & \text{for } b_i = 1 \\ A_c \cos(2\pi f_c t - \pi) & \text{for } b_i = 0 \end{cases} = \begin{cases} A_c \cos(2\pi f_c t) & \text{for } b_i = 1 \\ -A_c \cos(2\pi f_c t) & \text{for } b_i = 0 \end{cases}$$



### BPSK modulation



## Binary Frequency Shift Keying (BFSK)

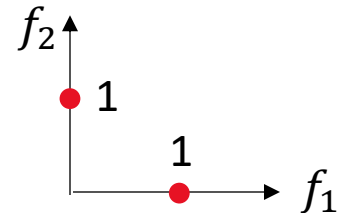
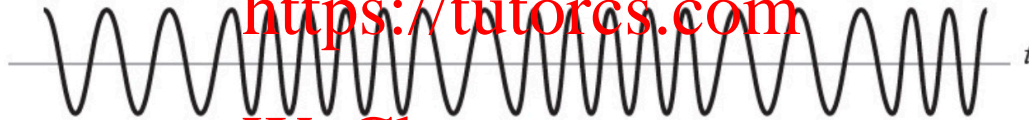
- Frequency of the carrier varies in accordance with  $b_i = 0/1$

$$s(t) = \begin{cases} A_c \cos(2\pi f_1 t) & \text{for } b_i = 1 \\ A_c \cos(2\pi f_2 t) & \text{for } b_i = 0 \end{cases}$$

(a) Input binary sequence

0 0 1 0 1 1 0 0 1

(d)



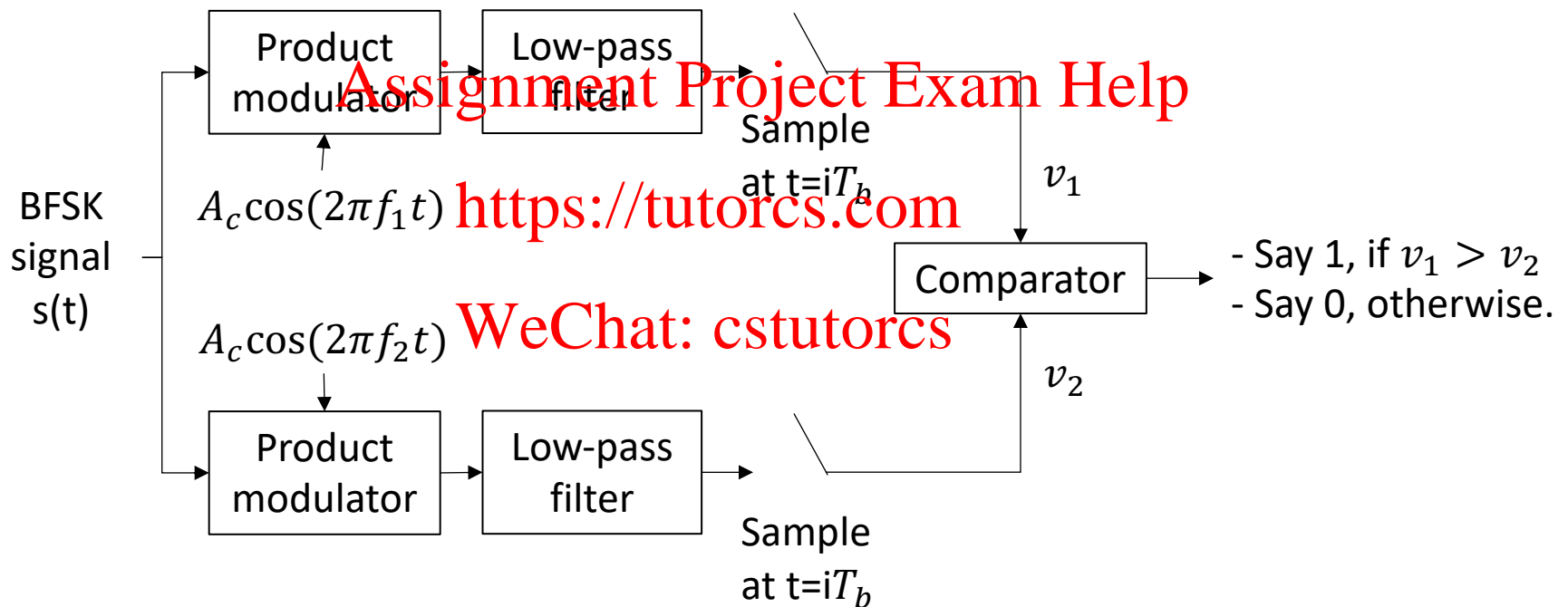
## Minimum-shift keying (MSK)

Let  $1/T_b$  and  $f_c$  be the bit rate and carrier frequency. In MSK,

$$f_1 = f_c + \frac{1}{4T_b}, f_2 = f_c - \frac{1}{4T_b}. \text{ That is, } f_1 - f_2 = \frac{1}{2T_b}.$$

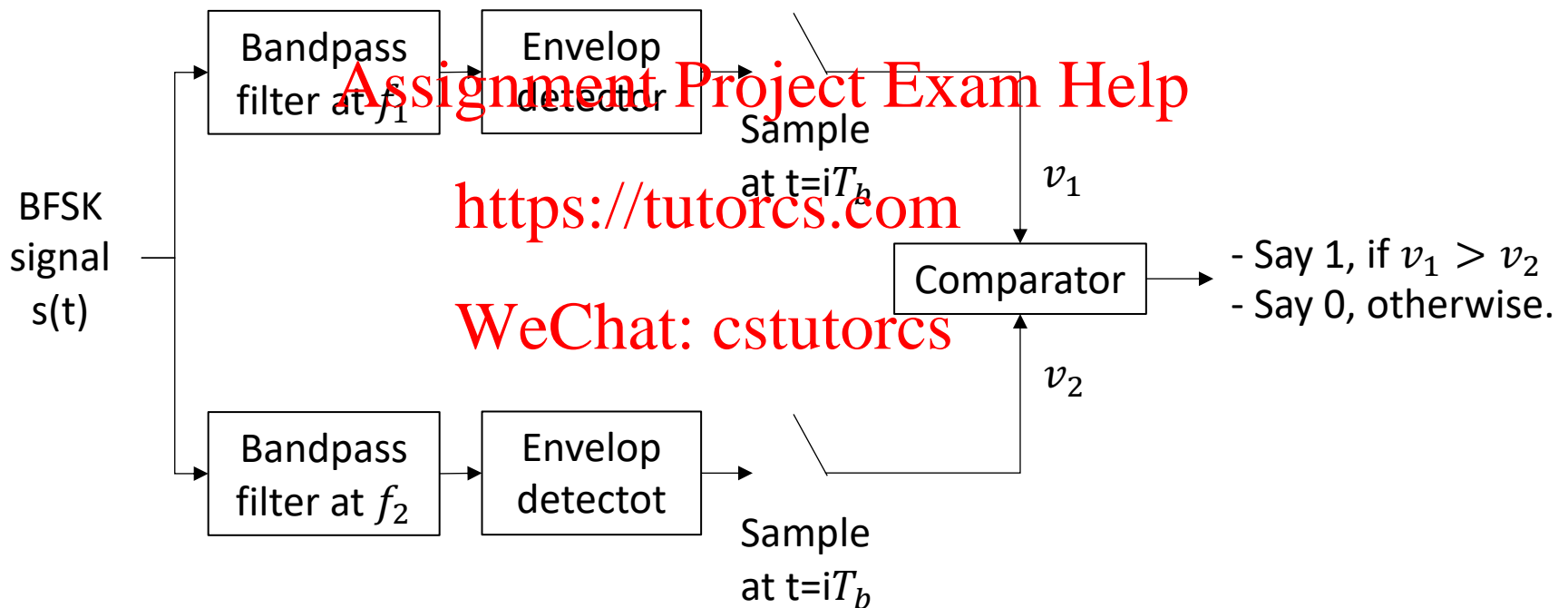
# Binary Frequency Shift Keying (BFSK)

## Coherent detection



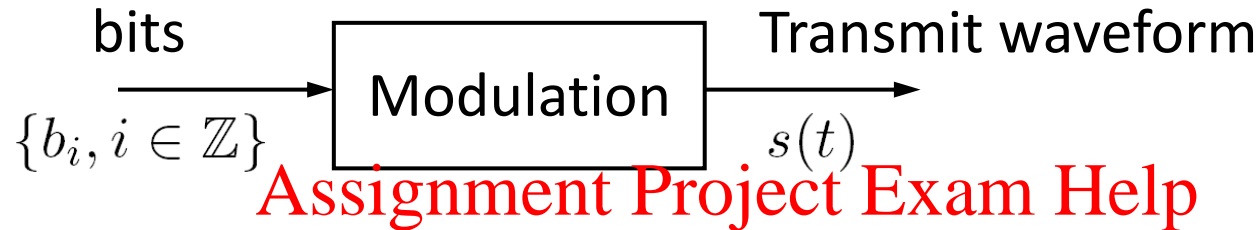
# Binary Frequency Shift Keying (BFSK)

## Noncoherent detection



## 8.4 M-ary Shift Keying Modulation

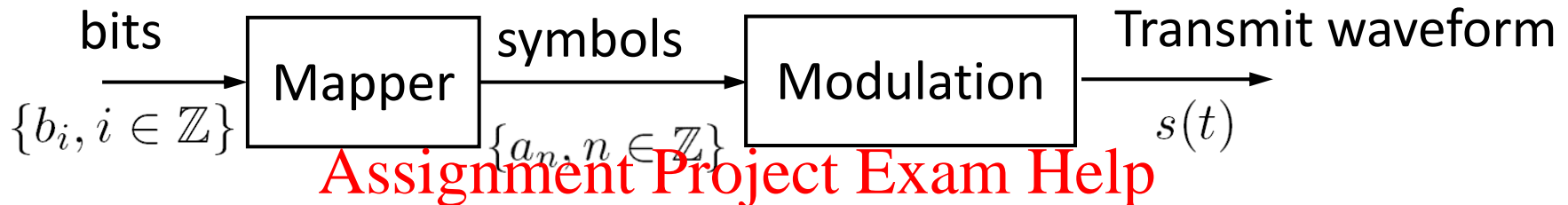
Binary shift keying is simply, but not bandwidth efficient.



Question: can we extend binary shift keying in some manner to improve the bandwidth efficiency?  
<https://tutorcs.com>  
WeChat: cstutorcs

Instead, we send one of  $M$  possible signals during each signaling (symbol) interval of duration.

## M-ary shift keying modulation

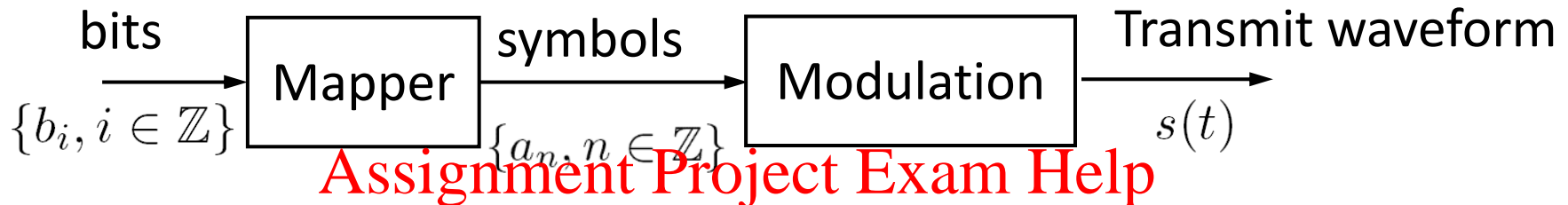


**Mapper:** Takes a group of bits and converts them to a symbol from an alphabet whose elements are real (or complex) numbers.

Depending on the mapper, we may have

- M-ary amplitude shift keying
- M-ary phase shift keying
- M-ary frequency shift keying

## M-ary Amplitude shift keying modulation



**Mapper:** Takes a group of bits and converts them to a symbol with different amplitude  $A_i$ .

WeChat: cstutorcs

**Example:** Every group of 2 bits can be converted to a symbol

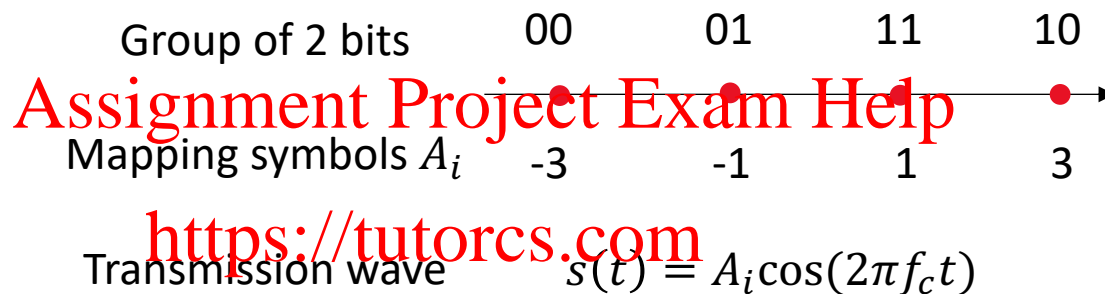
Group of 2 bits	00	01	11	10	
Mapping symbols $A_i$	-3	-1	1	3	
Transmission wave	$s(t) = A_i \cos(2\pi f_c t)$ for different symbol interval				

4-ASK

## M-ary Amplitude shift keying modulation

**Example:** Every group of 2 bits can be converted to a symbol

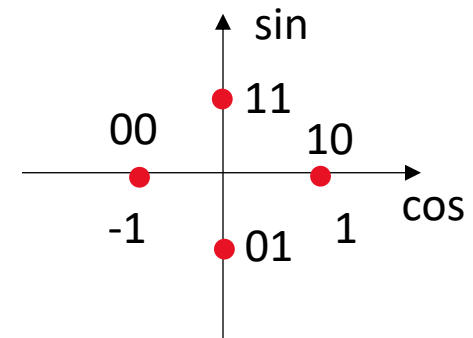
### 4-ASK



WeChat: cstutorcs

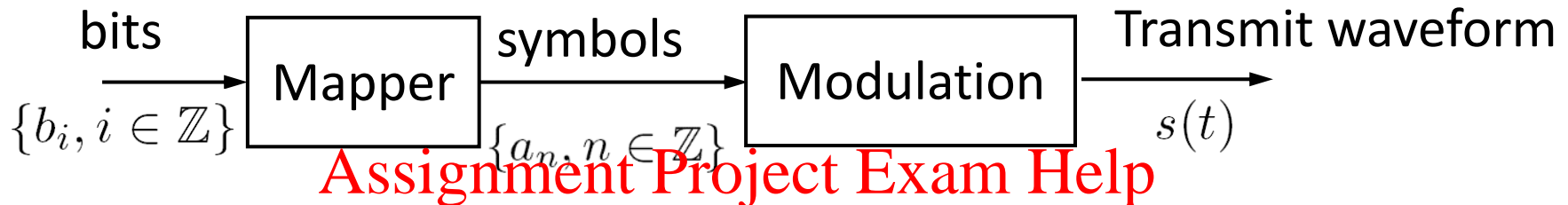
### Quadrature Amplitude modulation(QAM)

- 2 orthogonal carriers:  $A \cos(2\pi f_c t)$  and  $A \sin(2\pi f_c t)$
- Transmission wave is a superposition of the two carriers
- With the same bandwidth efficiency, QAM is more power-efficient than 4-ASK.





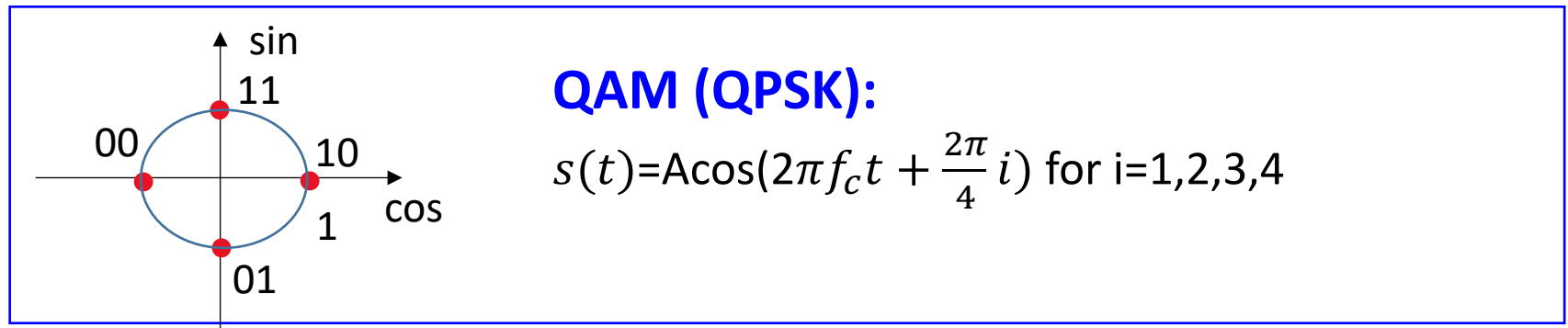
## M-ary Phase shift keying modulation



**Mapper:** Takes a group of bits and converts them to a symbol with different phase on the unit circle in the complex plane.

WeChat: cstutorcs

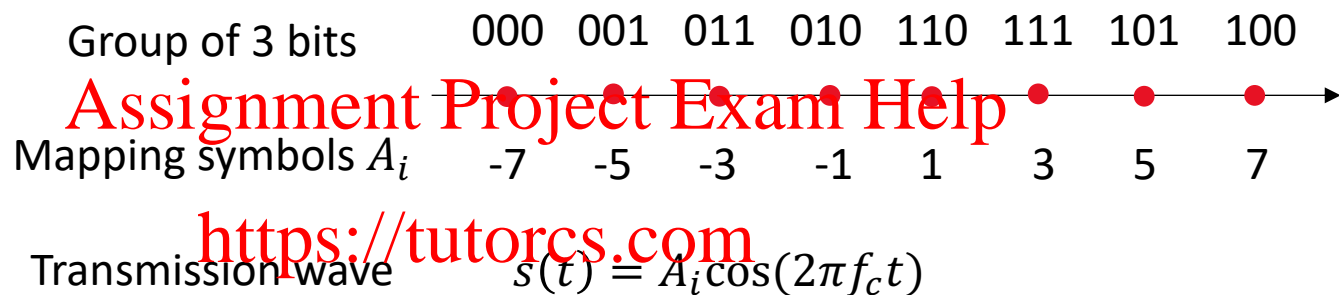
**Example:** Every group of 2 bits can be converted to a symbol



## M-ary Phase shift keying modulation

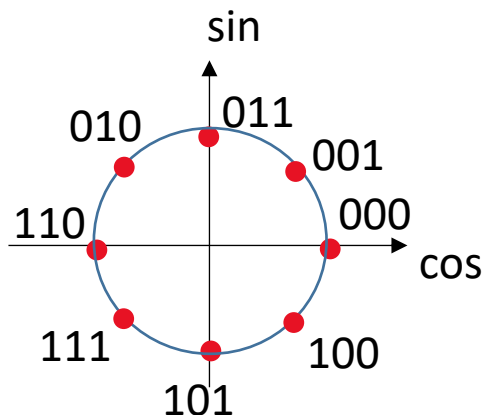
**Example:** Design 8-ASK and 8-PSK

### 8-ASK



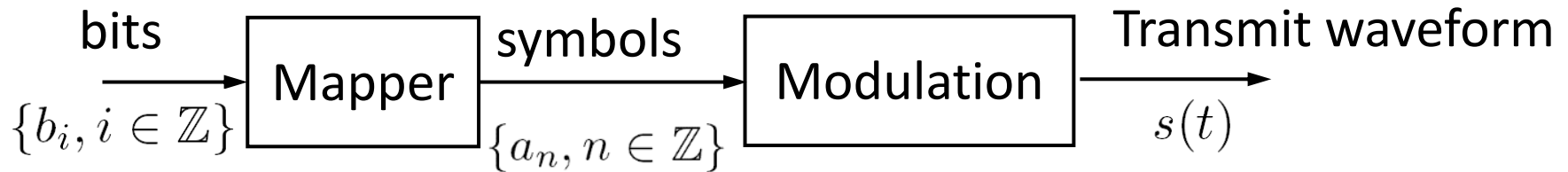
WeChat: cstutorcs

### 8-PSK



$$s(t) = A \cos(2\pi f_c t + \frac{2\pi}{4} i) \text{ for } i=1, \dots, 8$$

## M-ary Frequency shift keying modulation



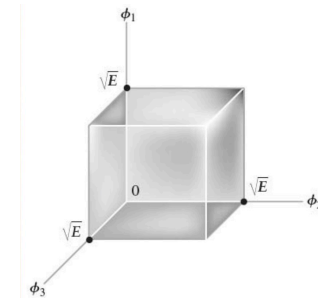
**Mapper:** Takes a group of bits and converts them to symbols modulated by different carriers

<https://tutorcs.com>

**Note,** in M-ary FSK, any two carriers should be orthogonal:

$$\int_0^T s_i(t)s_j(t)dt = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

where  $s_i(t) = \cos(2\pi f_{c_i}t)$ .



M-ary frequency shift keying is less bandwidth efficiency, in general.

## 8.5 Constellation Design

Design of the alphabet or symbol set  $\mathcal{A}$ . In other words, design of elements in  $\mathcal{A}$  for the best efficiency and/or reliability in communications.

Assignment Project Exam Help

**Design factors:**

- average energy per symbol;
- minimum distance between elements (determines the reliability).

<https://tutorcs.com>

WeChat: cstutorcs

## To minimize average energy per symbol

Given a constellation

$$\mathcal{A} = \{x_1, x_2, \dots, x_N\},$$

its average transmit energy is

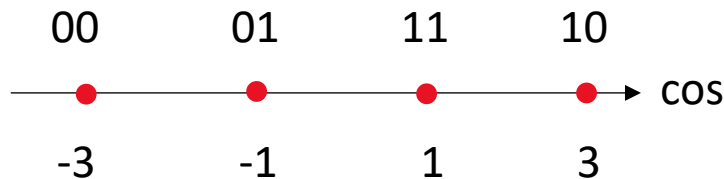
$$E_{\mathcal{A}} = \frac{1}{N} \sum_{n=1}^N |x_n|^2 = \frac{|x_1|^2 + \dots + |x_N|^2}{N}$$

Assignment Project Exam Help

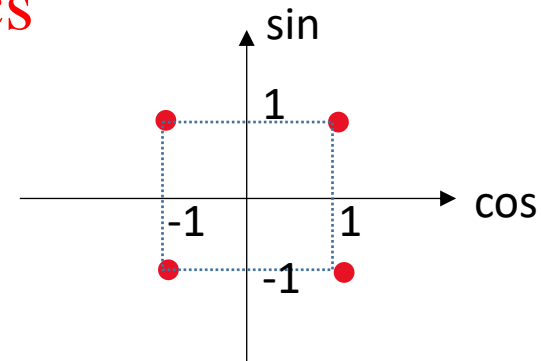
<https://tutorcs.com>

**Example.** Compare power efficiency of the two constellations.

WeChat: cstutorcs



$$E_{A1} = (9 + 1 + 1 + 9)/4 = 5$$



$$E_{A2} = (2 + 2 + 2 + 2)/4 = 2$$



## To maximize distance between symbols.

Given a constellation

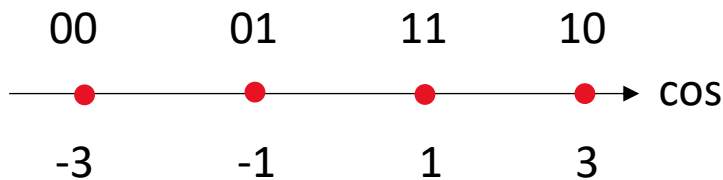
$$\mathcal{A} = \{x_1, x_2, \dots, x_N\},$$

the distance between two elements  $x_i, x_j$  is  $|x_i - x_j|$

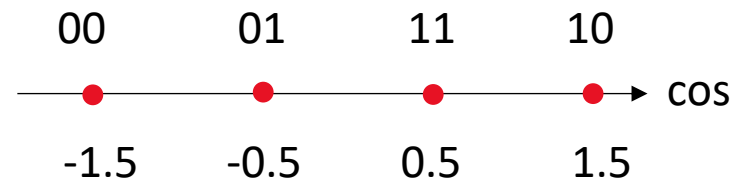
The minimum distance is

$$d_{\min, \mathcal{A}} = \min_{i \neq j} |x_i - x_j|$$

**Example.** Calculate the minimum distance of two 4-PAMs.



$$d_{\min, A1} = 2$$



$$d_{\min, A2} = 1$$

## To maximize distance between symbols.

Example. Comparison of two 4-PAMs.



<https://tutorcs.com>

Modulated waveform  $s(t) = A_i \cos(2\pi f_c t)$

- The left modulation is less energy efficient but can survive with maximal noise level of 2 because  $d_{min,A1} = 2$ .
- The right modulation method is more energy efficient but cannot survive when noise level is larger than 1 (i.e.  $d_{min,A2} = 1$ )

Large minimum distance suggests reliable the system.

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Constellation design:

- Given average transmit energy, try to maximize the minimum distance.
- Given minimum distance, try to minimize the average transmit energy.

Assignment Project Exam Help

**Example:** Please design 8-symbol constellations with minimum distance of 1.

(a) 8-ASK;

(b) 8-PSK;

(c) Is there any constellation that is more power efficient?

WeChat: cstutorcs

<https://tutorcs.com>



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

## 8.6 Detector Design

So far, we discuss demodulation methods without considering noise at the receiver.

In reality, noise is inevitable. With noise,  $y(mT) = a_m + n_m$ .

where the noise  $n_m$  follows some distribution, usually Gaussian.

**Q:** How to recover  $a_m$  from  $y(mT)$ ?

This is the **detector design problem**.



$y$  : Received value, corrupted by noise

$n$  : noise. Usually assume to be Gaussian following  $\mathcal{N}(0, \sigma^2)$ .

$a$  : transmitted symbol. An element in  $\mathcal{A}$ .

$\hat{a}$  : Detection result. An element in  $\mathcal{A}$ .

**Detector design: Design the decision rule.**

**Objectives:**

- Probability of error  $P_e = P[\hat{a} \neq a]$ .
- Complexity.

## Detection with minimum distance rule.

$$\hat{a} = \arg \min_{a \in \mathcal{A}} |y - a| = \arg \min_{a \in \mathcal{A}} |y - a|^2.$$

Decision regions.

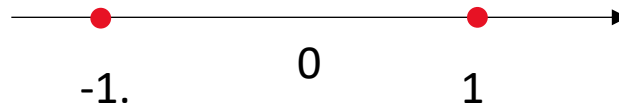
Probability of error calculation.

Assignment Project Exam Help

**Example:** For BPSK, please draw the decision region.

<https://tutorcs.com>

WeChat: cstutorcs





**Example:** Given a 4-ASK constellation as shown in the figure.

The output of the receiver is  $y = a + n$ , where  $a$  is the transmitted symbol and  $n \in N(0,1)$  is the Gaussian noise.

- (1) If  $y = 0.2$ , estimate the transmit symbol using the minimum distance rule.
- (2) Find the general decision rule and region for all  $y \in R$ .
- (3) What is the probability of error with this decision rule assuming  $a = 1$ .

WeChat: cstutorcs



Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs

Assignment Project Exam Help

<https://tutorcs.com>

WeChat: cstutorcs