錯誤更正碼
(ECC:Error Correcting Code)

課程目標: 本課程主要介紹錯誤更正碼之編碼理論, 其可用來作數據資料"傳輸"及"儲存"之錯誤更正控制.

評量方式： 作業及小考(40%), 期中考(30%), 期末考或期末報告(30%)

Outline

- Coding for Reliable Digital Transmission and Storage
- Linear Block Codes
- Introduction to Algebra
- Cyclic Codes
- Binary BCH Codes
- Nonbinary BCH Codes, Reed-Solomon Codes, and Decoding Algorithms
- Majority-Logic Decodable and Finite Geometry Codes
- Trellises for Linear Block Codes

- Textbook : Shu Lin and Daniel J. Costello, "Error Control Coding – Fundamentals and Applications," Second Edition, 2004. (歐亞書局代理)
- References:

*1. "Theory and Practice of Error Control Codes", Richard E. Blahut, 1983.

2. "Error-Control Coding for Data Networks", I.S. Reed & X.Chen, 新智, 1999.

*3. "Error Correction Coding-Mathematical Methods and Algorithms", Todd K. Moon, 全華,2005

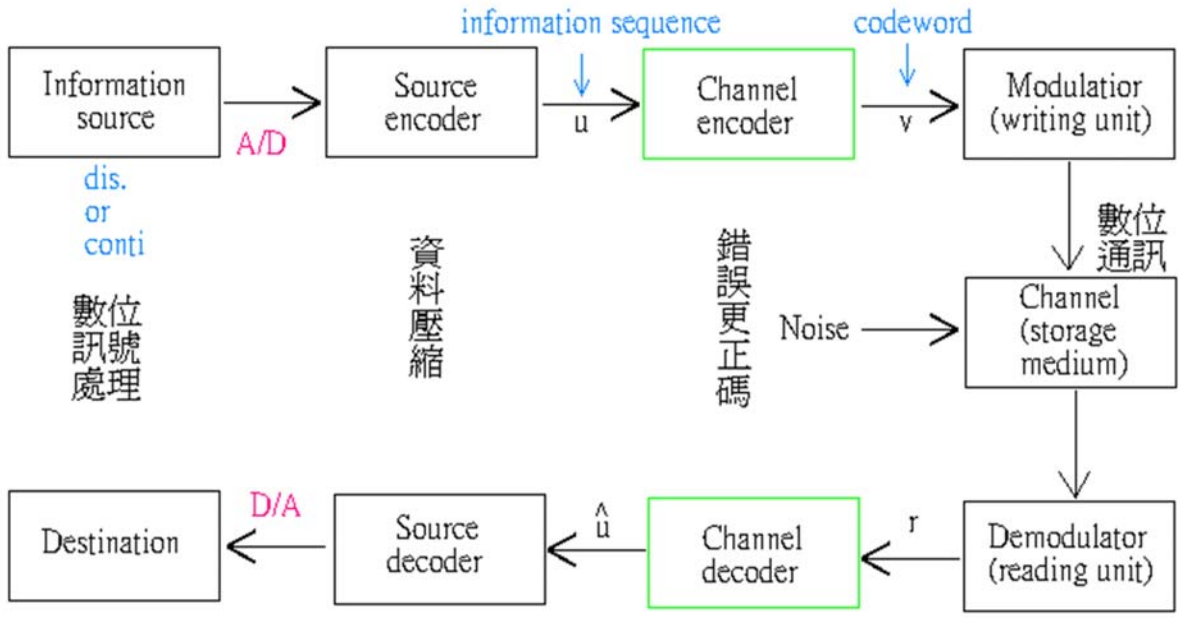
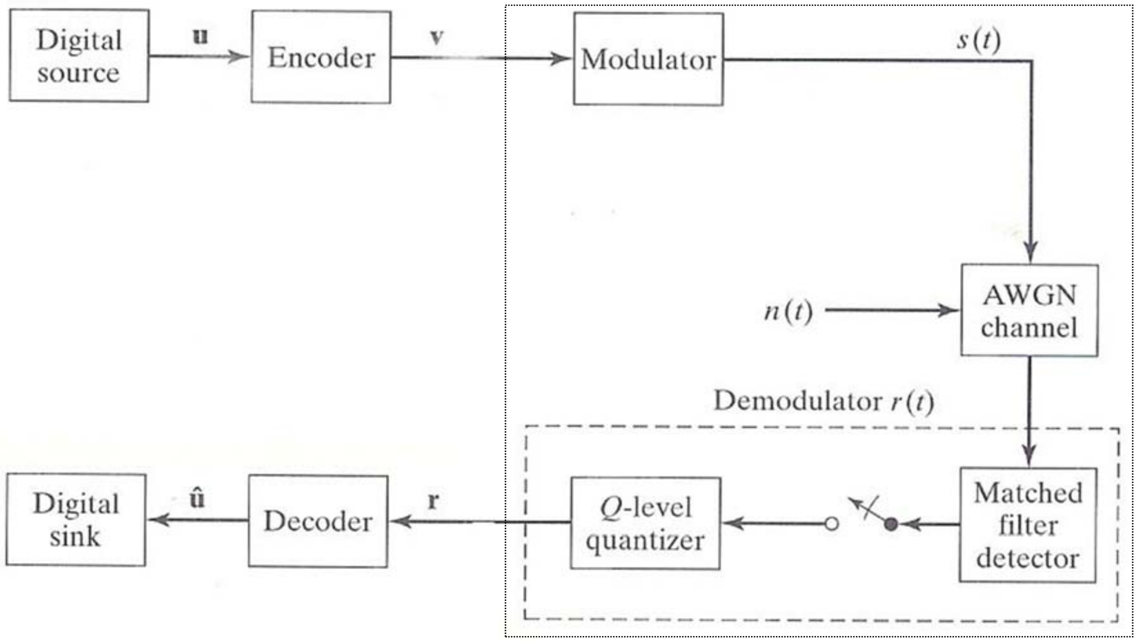


Figure 1.1 Block diagram of a typical data transmission or storage system

A coded system on an additive white Gaussian noise channel



Discrete Memory Channel (DMC)

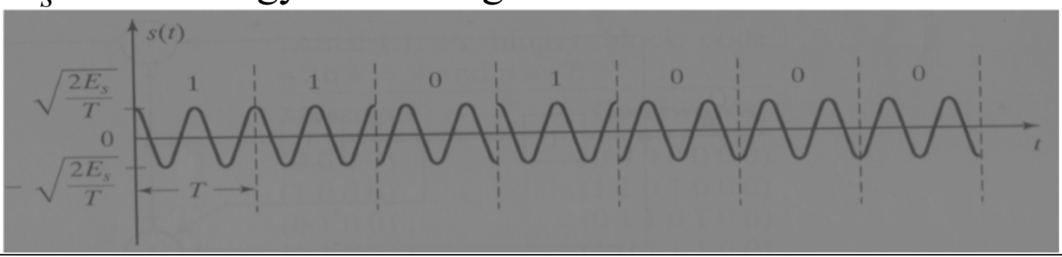
Binary-Phase-Shift-keying (BPSK)

$$S_1(t) = \sqrt{\frac{2E_s}{T}} \cos 2\pi f_0 t, \quad 0 \le t \le T$$

$$S_2(t) = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_0 t + \pi\right)$$

$$= -\sqrt{\frac{2E_s}{T}} \cos 2\pi f_0 t, \quad 0 \le t \le T,$$

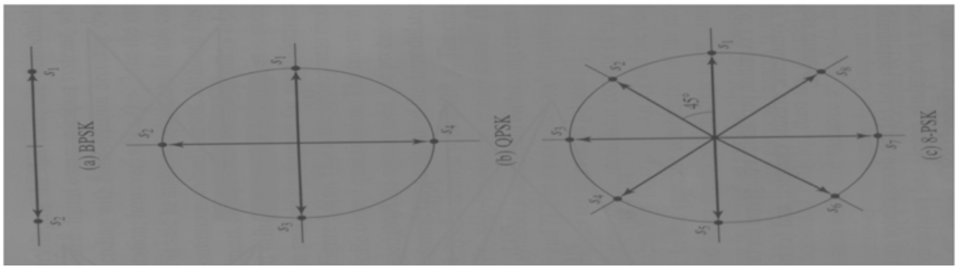
where the carrier frequency f_0 is a multiple of $1/T$, and E_s is the energy of each signal.



M-ary Phase-Shift-keying (MPSK)

$$S_i(t) = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_0 t + \phi_i\right), \quad 0 \le t \le T,$$

$$\text{where } \phi_i = 2\pi(i-1)/M \text{ for } 1 \le i \le M.$$



AWGN (Additive White Gaussian Noise) Channel

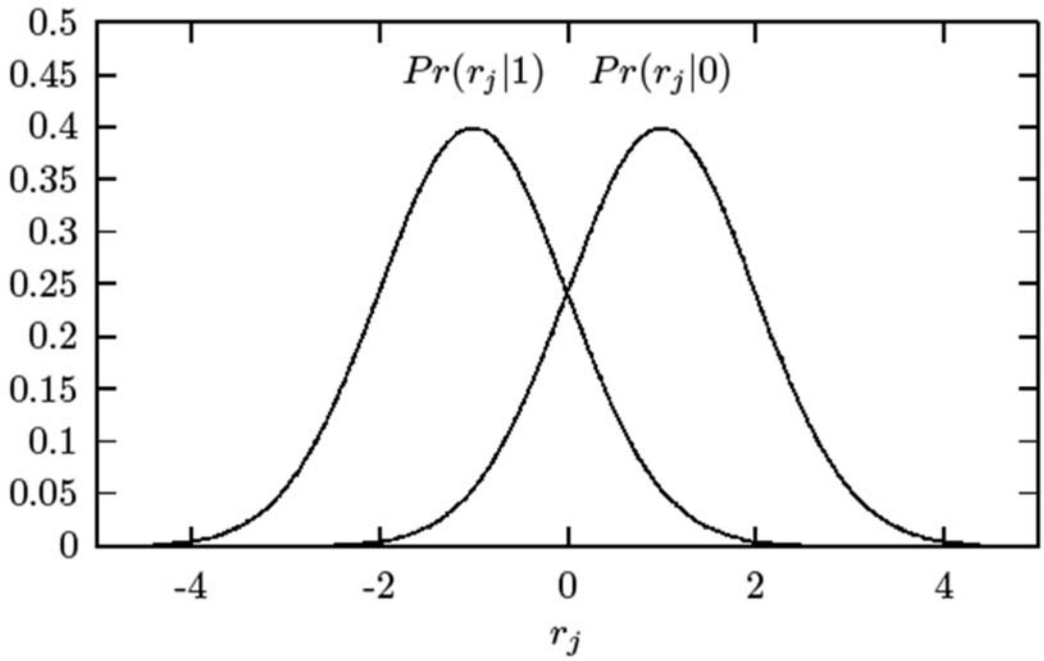
- A 0 is transmitted as $+\sqrt{E}$ and a 1 is transmitted as $-\sqrt{E}$, where E is the signal energy per channel bit.
- $r_i = (-1)^{s_i}\sqrt{E} + n_i$, where s_i is the transmitted bit, r_i is the received bit, and n_i is a noise sample of a Gaussian process with single-sided noise power per hertz N_0.
- The variance of n_i is $N_0/2$ and the signal-to-noise ratio (SNR) for the channel is E/N_0.

$$\Pr(r_i \mid s_i) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r_i - (-1)^{s_i}\sqrt{E})^2}{N_0}}, \text{where } s_i = 0 \text{ or } 1, r_i \in \text{real number.}$$

$$[n(x;\mu;\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ with } \sigma^2 = N_0/2]$$

Probability distribution function for r_j

The signal energy per channel bit E has been normalized to 1.



Binary Symmetric Channel

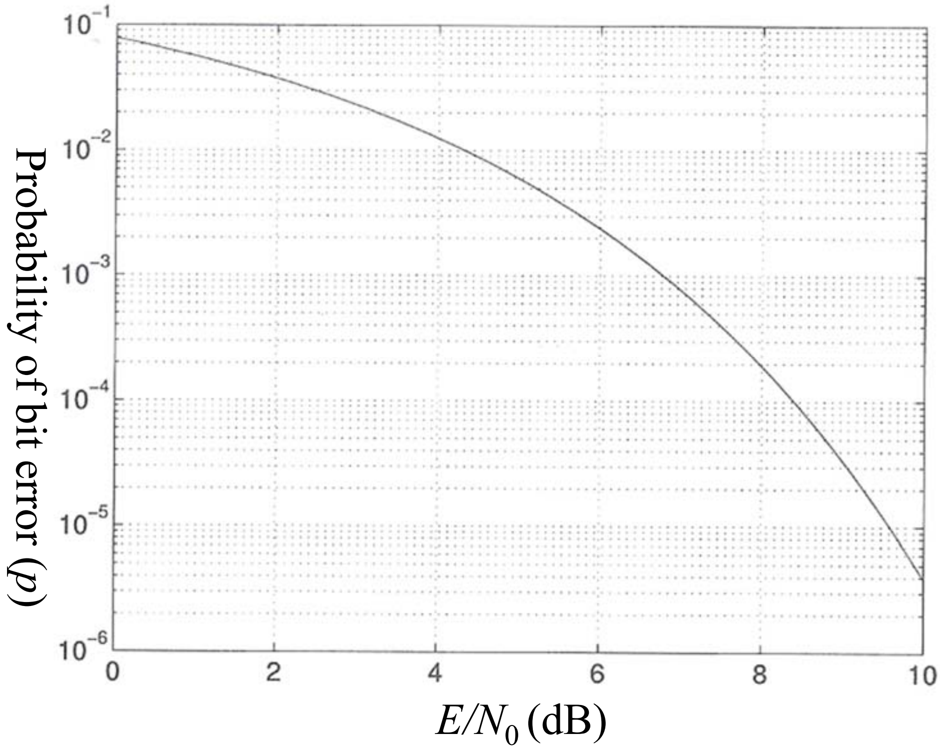
- BSC is characterized by a probability *p* of bit error such that the probability *p* of a transmitted bit 0 being received as a 1 is the same as that of a transmitted 1 being received as a 0.
- When BPSK modulation is used on an AWGN channel with optimum coherent detection and binary output quantization

$$p = \int_0^{\infty} p_r(r_i \mid 1)dr_i = \int_{-\infty}^{0} p_r(r_i \mid 0)dr_i$$

$$= \int_0^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r_i+\sqrt{E})^2}{N_0}} dr_i = \int_{\sqrt{2E/N_0}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = Q(\sqrt{2E/N_0}),$$

$$\text{where } Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Probability of error for BPSK signaling




```
/******AWGN (transfer information to receive sequence)******/
void awgn_channel(double Y, double R, char information[], double receive[], int size){// Y: SNR(dB) R: code rate
    double u1,u2;
    double r_max;
    int i;

    i = 0;
    #ifdef PC
      r_max = RAND_MAX;
    #endif
    while (i < size){
        #ifdef PC
          u1 = rand()/r_max;
          u2 = rand()/r_max;
        #else
          u1 = drand48();
          u2 = drand48();
       #endif
        receive[i] = sqrt(-2 * log(u1)) * cos(2 * Pi * u2);
        i++;
        if (i<size) {
            receive[i] = sqrt(-2 * log(u1)) * sin(2 * Pi * u2);
            i++;
        }
    }

    for (i = 0; i < size; i++) {
        receive[i] = receive[i] / sqrt(2.0 * R * pow(10.0, Y / 10.0));
        if(information[i] == '1')
            receive[i] = receive[i] + One;
        else
            receive[i] = receive[i] + Zero;
    }
}
```

- p15 = (pow(2,15)-1.0)/2.0;
- do {u1=(float)rand()/p15-1.0; u2=(float)rand()/p15-1.0; z=u1*u1+u2*u2;} while(z > 1.0 || z == 0.0);
- z = sqrt(-2.0*log(z)/z); u1 = u1*z; u2 = u2*z;
- ➔ u1, u2 will be two iid standard Gaussian rv's

$X, Y : Gaussian$

$R = \sqrt{X^2 + Y^2} : Rayleigh \quad W = \tan^{-1}(\frac{Y}{X}) : uniform$ on $[0, 2\pi)$

$\Rightarrow X = R\cos W, \quad Y = R\sin W$

(1)

$U1, U2 : uniform \quad R = \sqrt{-2\log U1} : Rayleigh$

$X = R\cos(2\pi U2) : Gaussian$

$X = R\sin(2\pi U2) : Gaussian$

(2)

$U1, U2 : uniform \quad Z = U1^2 + U2^2$ with Z<1 :*uniform*

$R = \sqrt{-2\log Z} : Rayleigh$

$X = R \times \frac{U1}{\sqrt{Z}} : Gaussian \quad Y = R \times \frac{U2}{\sqrt{Z}} : Gaussian$

Hard-decision decoding

- The output of the matched filter for each signaling interval is quantized in two levels, denoted as 0 and 1.
- e.g. algebraic decodings :using algebraic structures of the codes
- A hard decision of a received signal results in a loss of information, which degrades performance

Soft-decision decoding

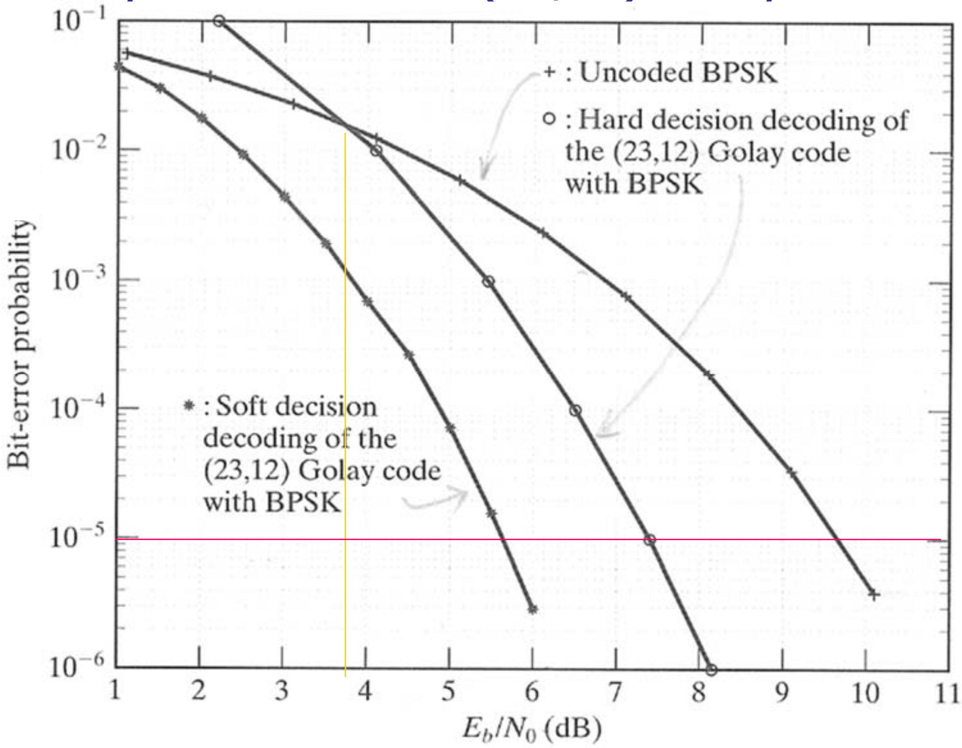
- If the outputs of the matched filter are unquantized or quantized in more than two levels, we say that the demodulator makes *soft decisions*. A sequence of soft-decision outputs of the matched filter is referred to as a *soft-decision received sequence*. Decoding by processing this soft-decision received sequence is called *soft decision decoding*.
- The decoder uses the additional information contained in the unquantized (or multilevel quantized) received samples to recover the transmitted codewords, soft-decision decoding provides better error performance than hard-decision decoding.

- In general, soft-decision maximum likelihood decoding (MLD) of a code has about 3 dB of coding gain over algebraic decoding of the code; however, soft-decision decoding is much harder to implement than algebraic decoding and requires more computational complexity.
- These decoding algorithms can be classified into two major categories:
  - *Reliability-based* (or probabilistic)
  - *Code structure-based*

- Code rate R=k/n
- Coding gain : the reduction in the E_b/N_0 required to achieve a specific bit error probability (rate) (BER) for a coded communication system compared to an uncoded system.
- Coding threshold : There always exists an E_b/N_0 below which the code loses its effectiveness and actually makes the situation worse.
- E_b : the signal energy per information bit

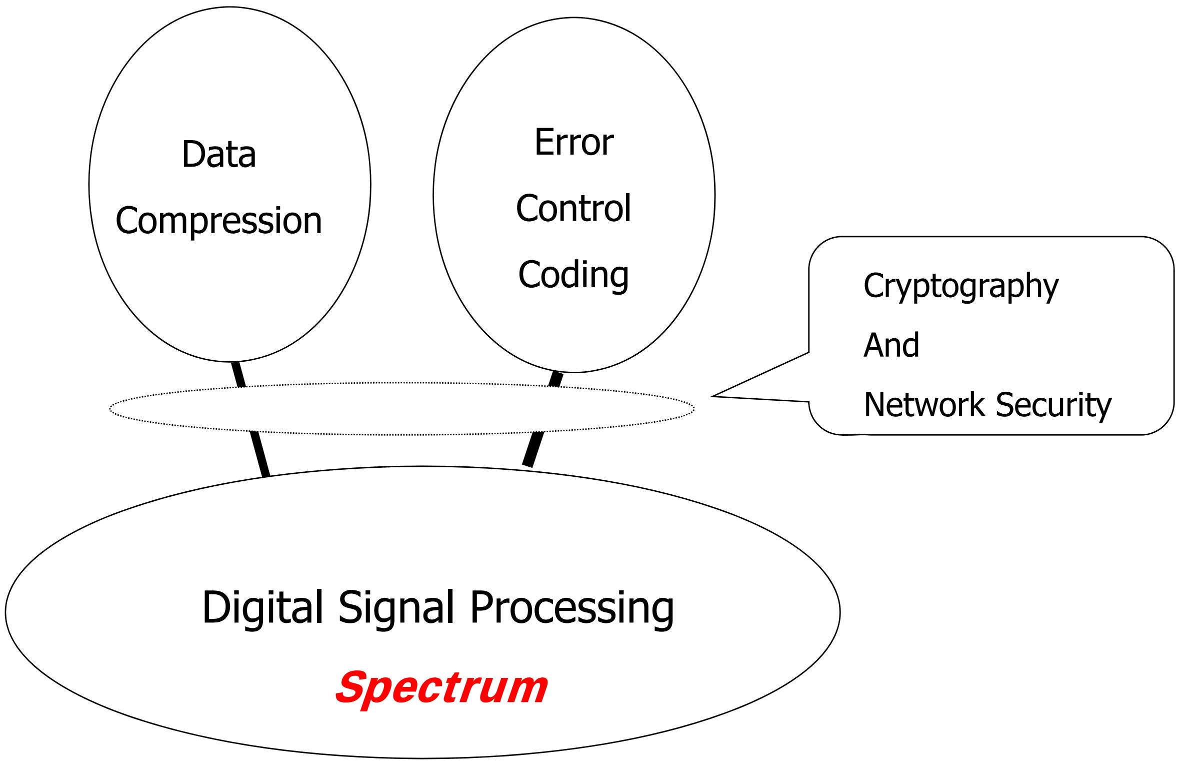
$$E_b = E / \mathrm{R}$$

Bit-error performance of a coded communication system with the (23,12) Golay code



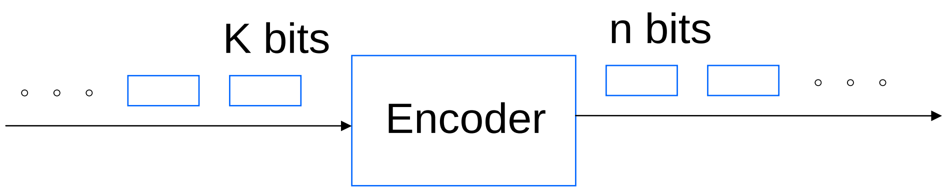
- Transmission channels : telephone line, high-frequency radio link, microwave link, satellite link, etc..
- Storage media : semiconductor memory, disk file, optical memory.

- Source Coding
  - \# of bits per unit time is minimized
  - Non-ambiguous decoding
- Channel Coding
  - Encoder : to combat the noisy environment in which code words must be transmitted or stored.
  - Decoder : that minimize the probability of decoding error

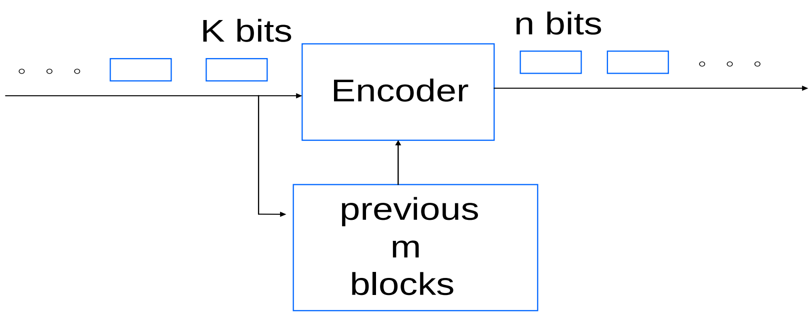


Types of Codes

- Linear Block Code (n,k) 2^k codewords
  - divide the information sequence into message blocks of k information bits each.
    - i.e. $u=(u_1,u_2,\dots,u_k)$ message
    - $v=(v_1,v_2,\dots,v_n)$ codeword
  - memoryless
  - combinational logic circuit

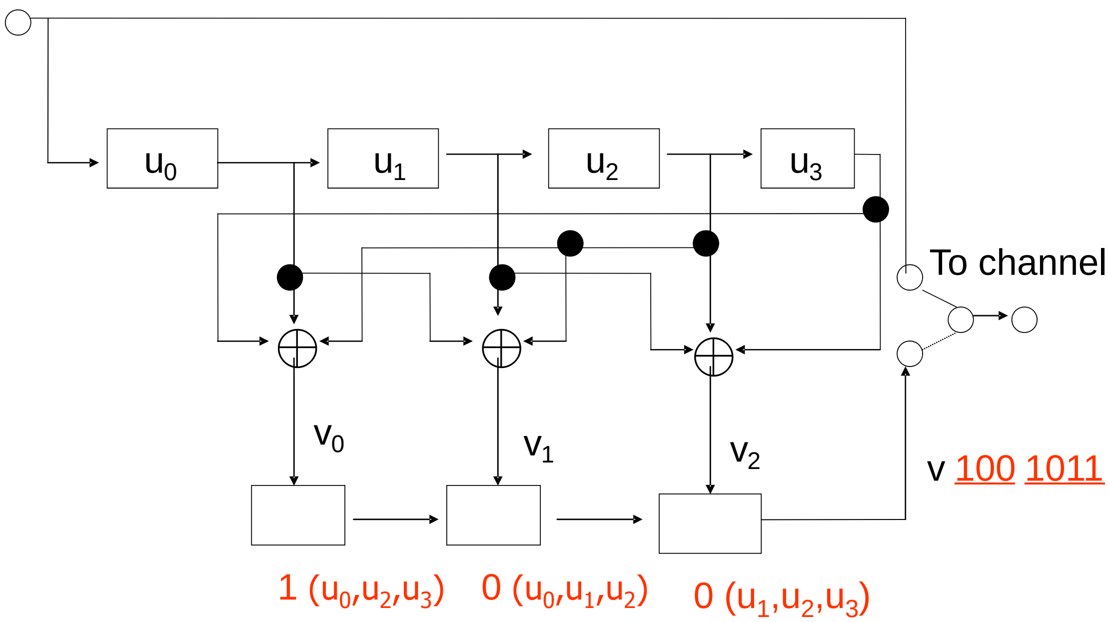


- Convolutional Code (n,k,m)
  - u,v : two sequences of blocks
  - memory order m
  - sequential logic circuit
- R = k/n code rate



* (7,4) block code

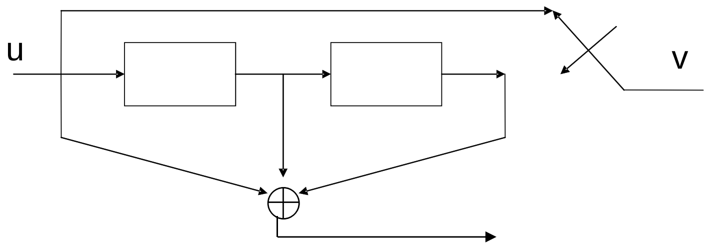
Input u (1011)



Q: 其它組合?

* (2,1,2) Convolutional code

- let u=(1 1 0 1 0 0 0 0 0)
  v=(11,10,00,10,01,01,00,00,00)



…0001011
…0001011
…0001011
…0001011
…0001011
…0001011
…0001011

- u=(…,m_{-1},m_0,m_1,…)
- v=(…,$C_{-1}^{(1)}$, $C_{-1}^{(2)}$, $C_0^{(1)}$, $C_0^{(2)}$, $C_1^{(1)}$, $C_1^{(2)}$,…)

Linear Time Invariant System

-2 -1 0 1 2

$$\delta(n)=\begin{cases}1 & n=0\\ 0 & n\neq 0\end{cases} \qquad \delta(n-k)=\begin{cases}1 & n=k\\ 0 & n\neq k\end{cases} \qquad L:\delta(n)\rightarrow h(n)$$

where $\delta(n)$: impulse sequence and $h(n)$: impulse response sequence

$$\Rightarrow f(n)=\sum_{k=-\infty}^{\infty} f(k)\delta(n-k)$$

0 K

$--- + f(-1)\delta(n+1) + f(0)\delta(n) + f(1)\delta(n-1) + ---$

$$\Rightarrow g(n)=L\{f(n)\}=L\{\sum_{k=-\infty}^{\infty} f(k)\delta(n-k)\}$$

$$\overset{\text{linear}}{=}\sum_{k=-\infty}^{\infty} f(k)L\{\delta(n-k)\}\overset{\text{time invariant}}{=}\sum_{k=-\infty}^{\infty} f(k)h(n-k)\equiv f(n)*h(n)$$

$$C_i^{(1)} = \sum_k g_k^{(1)} \cdot m_{i-k} = m_i$$

$$C^{(1)} = m \otimes g^{(1)}$$

$$C_i^{(2)} = \sum_k g_k^{(2)} \cdot m_{i-k} = m_i \oplus m_{i-1} \oplus m_{i-2}$$

$$C^{(2)} = m \otimes g^{(2)}$$

linear convolution

where $\begin{cases} g_0^{(1)} = 1 \text{ , others } = 0 \\ g_0^{(2)} = g_1^{(2)} = g_2^{(2)} = 1 \text{ , others } = 0 \end{cases}$

Impulse response

Type of Errors

* random-error-correcting codes
  - random-error channels (memoryless)
    . deep-space channel
    . satelite channel

* burst-error-correcting codes
  - burst-error channels (with memory)
    . radio channel

* burst-and-random-error-correcting codes
  - compound channels

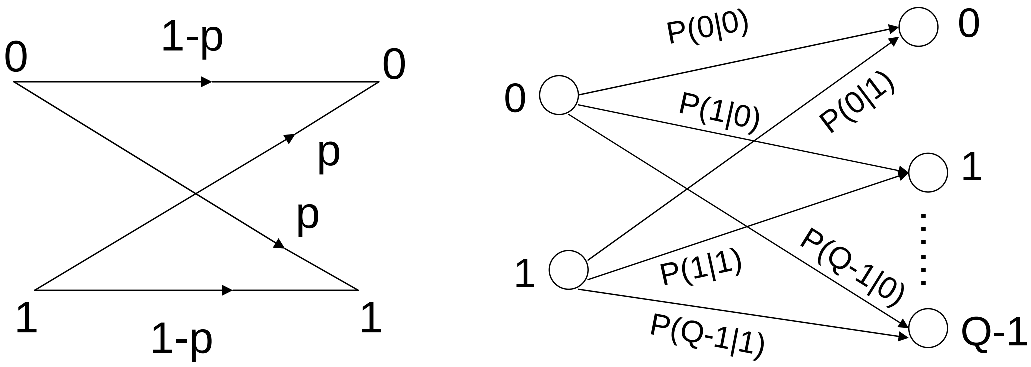
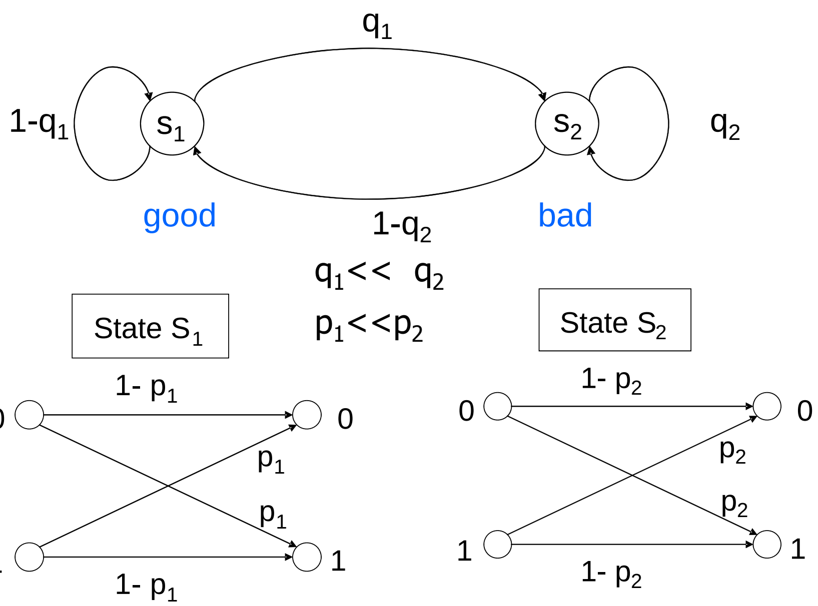


Figure 1.2 Transition probability diagrams :
(a) binary symmetric channel ;
(b) binary-input , Q-ary-output discrete memoryless channel



$q_1 << q_2$

$p_1 << p_2$

e.g.
$p_1 \approx 0$
$p_2 \approx 0.5$

Figure 1.7 Simplified model of a channel with memory

e.g. radio channel : error bursts are caused by signal fading

Error Control Strategies

- Forward Error Correction (FEC)
  - one-way system
    - e.g. magnetic tape storage system
    - deep-space communication system
- Automatic Repeat Request (ARQ)
  - error detection and retransmission
  - two-way system
    - e.g. telephone channel
  - types
    - stop-and wait ARQ
    - continuous ARQ
      - e.g. go-back-N ARQ ; select-repeat ARQ

- Advantage of ARQ over FEC
  - simpler decoding equipment
- For high-speed data network, which is the better ?