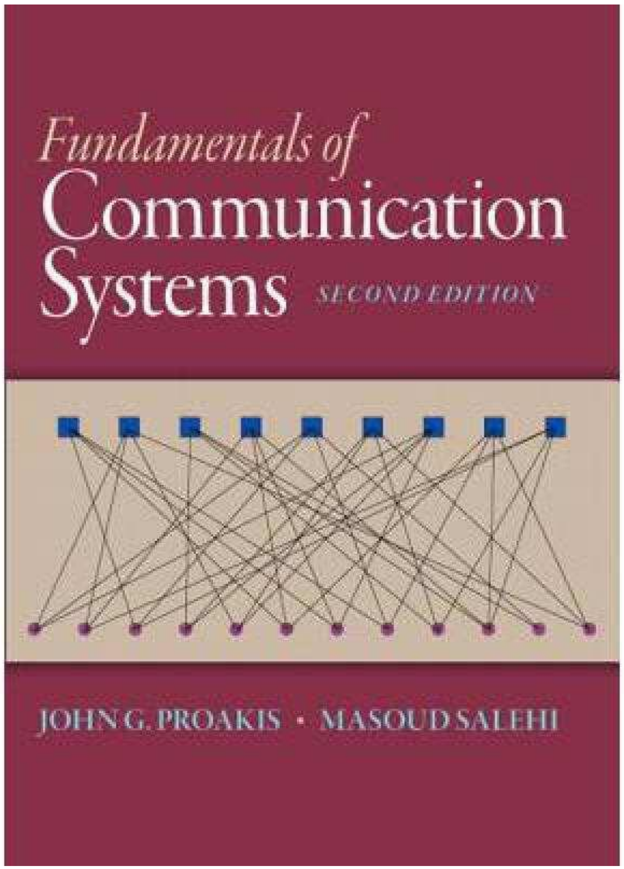
School of Commun. and Elec. Engineering

Communication Theory: Digital Communications

Prof. Deli Qiao
Department of Communications Engineering

Instructor Information

- Office: 245
- Email: dlqiao@ce.ecnu.edu.cn
- Web: http://faculty.ecnu.edu.cn/s/2915/main.jspy
- Phone: 54345492
- Office Hours: Tues. & Thu. 0:30pm~1:30pm
- TA: Na Ge & Zhilei Chen, Room 147, Information Buiding



J. Proakis

樊昌信改编

Reading Materials

- 通信原理（Principles of Communications）, Lecture Notes, Prof. Meixia Tao, Shanghai Jiao Tong University.[http://iwct.sjtu.edu.cn/Personal/mxtao/teaching.html]
- Introduction to analog and digital communications, Lecture Notes, Ohio State University. [http://www2.ece.ohio-state.edu/~schniter/ee501/index.html]
- Principles of Communications, Lecture Notes, City University of Hong Kong [http://www.ee.cityu.edu.hk/~lindai/teaching-EE3008.htm]
- Communications Theory and Systems, Lecture Notes, Washington University in St. Louis[https://span.engineering.wustl.edu/teaching/ese471/LectureNotes-s21-All.pdf]

Reading Materials

- 《Communication Systems Principles Using MATLAB》，J. W. Leis, Wiley, 2018
- 《Communication Systems》(5th ed), Simon Haykin, Wiley, 2009
- Introduction to Matlab, Lecture Notes, University College London, [http://web4.cs.ucl.ac.uk/teaching/3085/archive/2010/matlab_tutorial/matlab_booklet.pdf]

Course Objective

- The primary objective of this course is
  - to introduce the basic techniques used in modern communication systems, and
  - to provide fundamental tools and methodologies in analysis and design of these systems
- After this course, the students are expected to
  - understand the information flow in communication systems and the theories and techniques of modulation, coding and transmission, and
  - analyze the merits and demerits of current communication systems and to eventually perform research and development (R&D) related to new systems

[5G][VR/AR][6G]

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

Grades

- **Quiz (10%)**
  - About 10-15 times, each **one and only one** problem.
  - Attendance.
- **Homework (10%)**
  - About 4-7 times
- **Project Report (10%)**
  - About 3-5 times, Matlab simulation problems.
- **Mid-term Exam (20%)**
  - In-class.
- **Final Exam (50%)**

In-Class Rules

- **Shutdown smartphone and donot put on desk!**
- **No food/drink in class and put your drinking bottle aside the desk!**
- **No mutual conversation !**
- **Prepare, make notes and review frequently!**
- **Practice makes perfect!**

Maxim (格言)

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

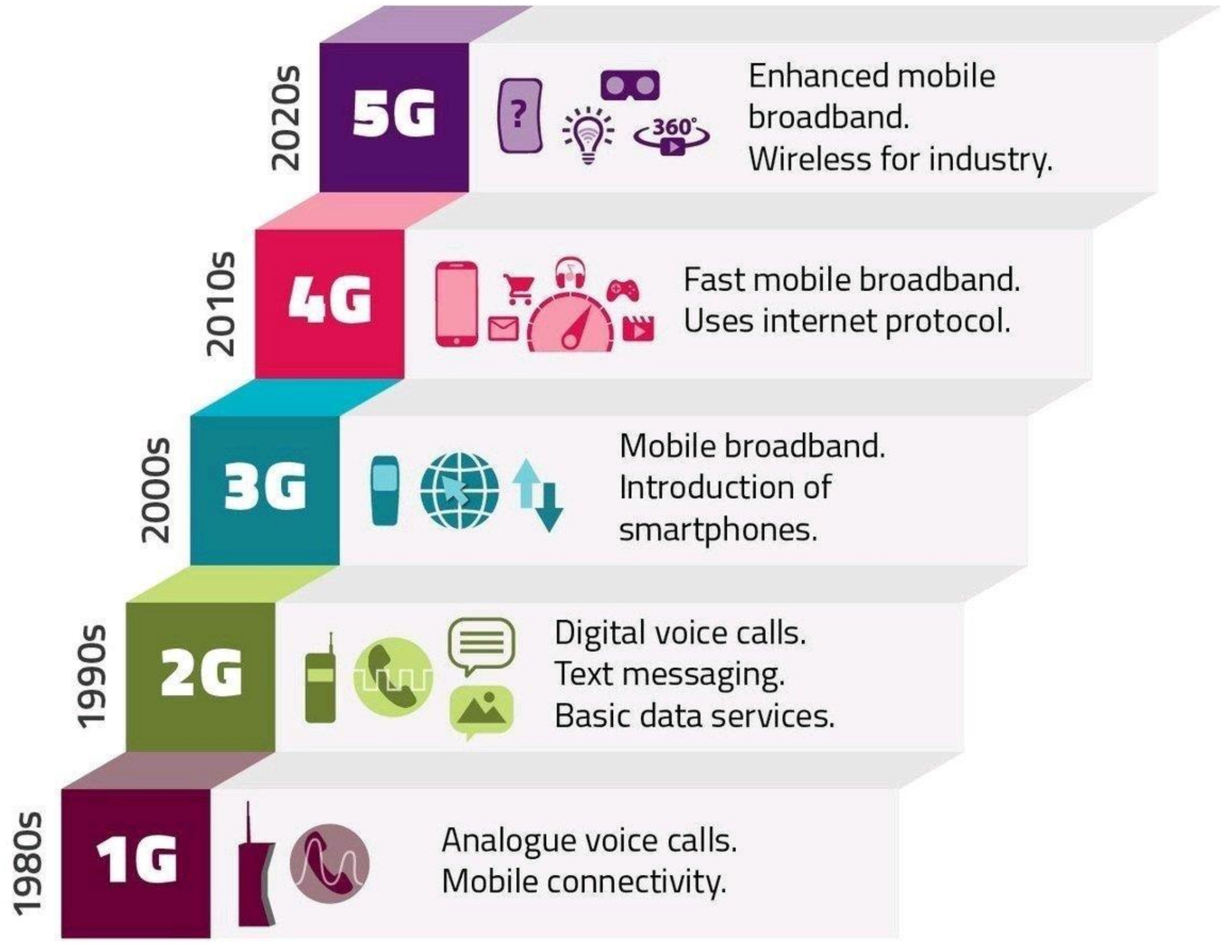
Introduction

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.

---Claude E. Shannon 1948

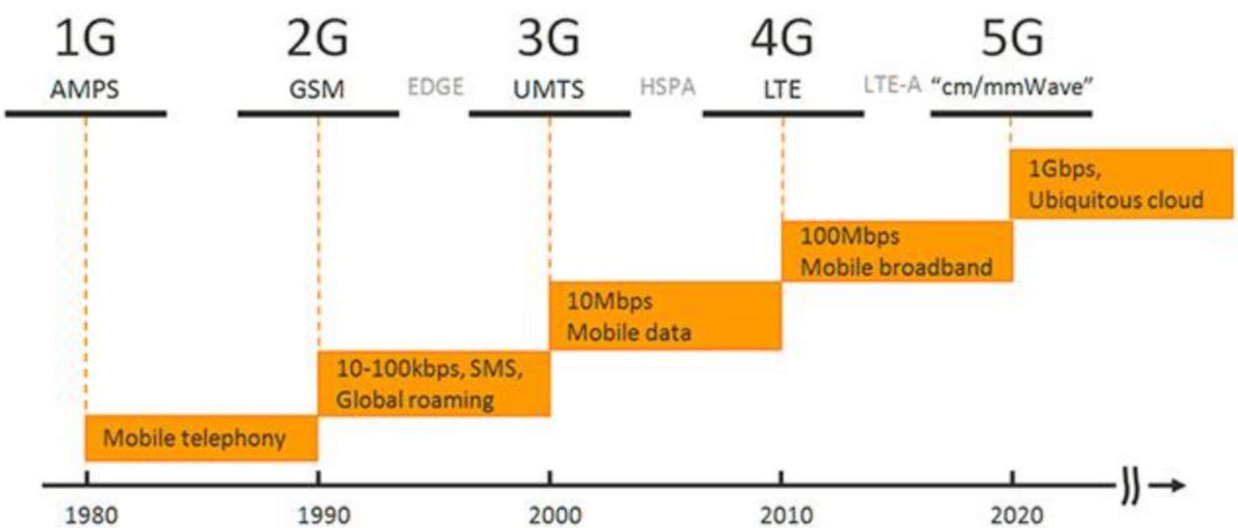
Introduction

- Review of Digital Communications



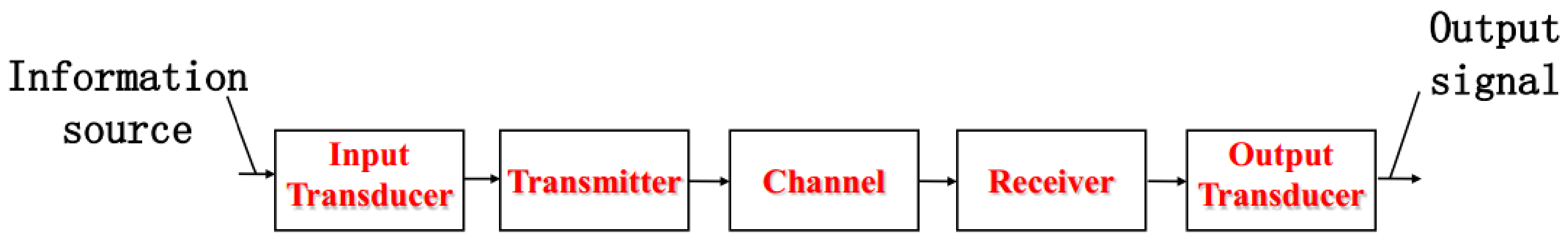
Introduction

- General questions
  - Is there a general methodology for designing communication systems?
  - Is there a limit to how fast one can communicate?

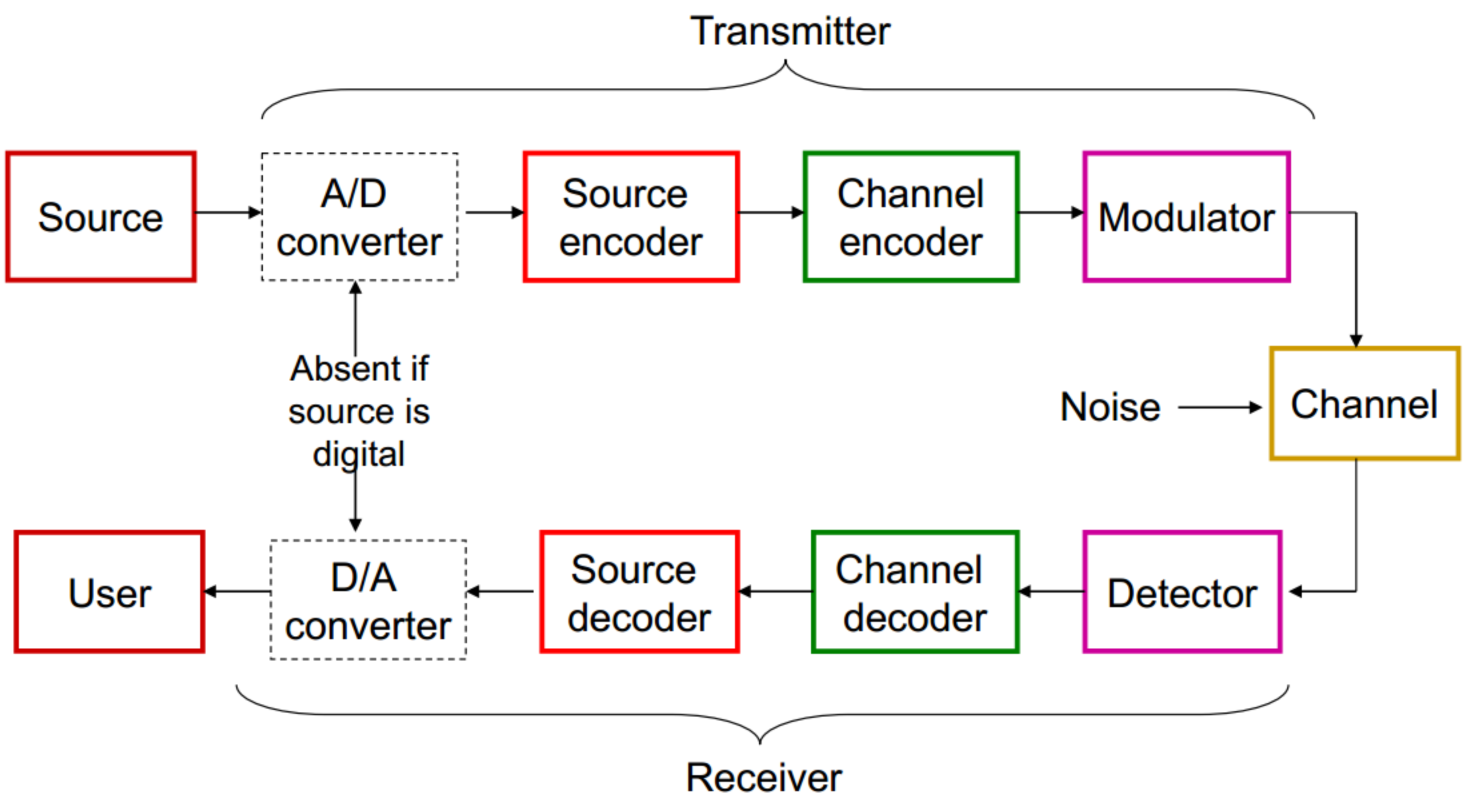


Introduction

- System diagram



Digital Communication Systems



Introduction

- Why digital systems?
  - Robustness to channel noise and external interference
  - Security of information during its transmission from source to destination
  - Integration of diverse source information into a common format
  - Low cost DSP chips by very cheap VLSI designs

Introduction

- Performance metrics of communication systems
  - Reliability: SNR for analog; Bit error rate (BER) for digital
  - Efficiency: Spectral efficiency vs. Energy efficiency

$$\frac{\text{data rate } R}{\text{bandwidth } W} \text{ bits/sec/Hz}$$

$$\updownarrow$$

$$\frac{\text{bit energy}}{\text{noise power spectral density}} = E_b/N_o$$

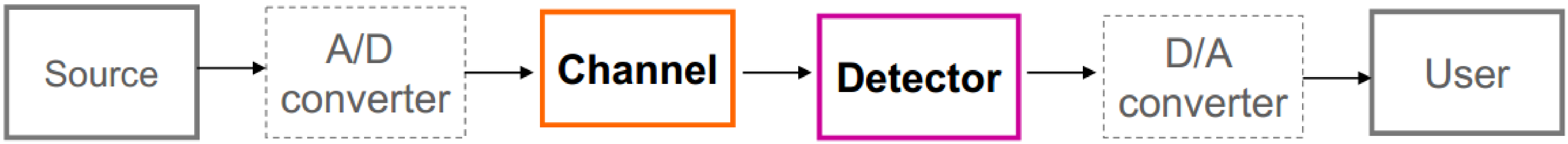
Introduction

- Review of Probability and Stochastic Processes
  - Bayes' Rule
  - CDF
  - PDF
  - Common Distributions
  - Q-function
  - Statistics of multiple r.v.s
  - Autocorrelation Function and Covariance
  - WSS
  - PSD
  - WSS through LTI systems

Outline

- Introduction
- **Digital transmission through baseband channels**
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

Digital transmission through bandlimited channels



- Digital waveforms over band-limited baseband channels
- Band-limited channel and Inter-symbol interference
- Signal design for band-limited channel
- System design
- Channel equalization

Chapter 10.1-10.4

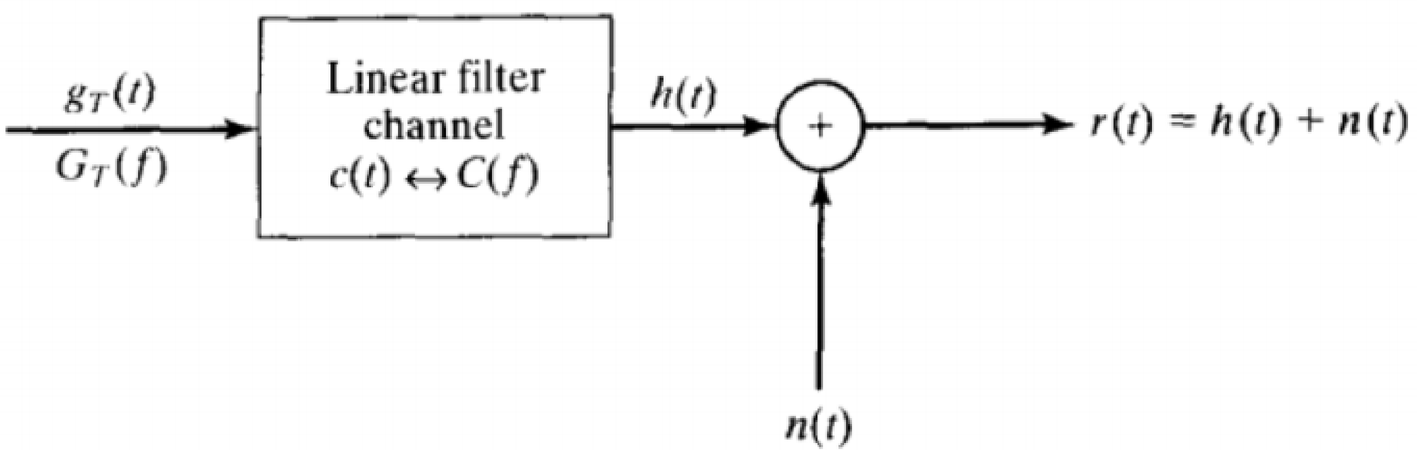
Digital transmission through bandlimited channels

- Band-limited channel

20MHz

5MHz, 10MHz
15MHz, 20MHz

- Modeled as a linear filter with frequency response limited to certain frequency range



Digital transmission through bandlimited channels

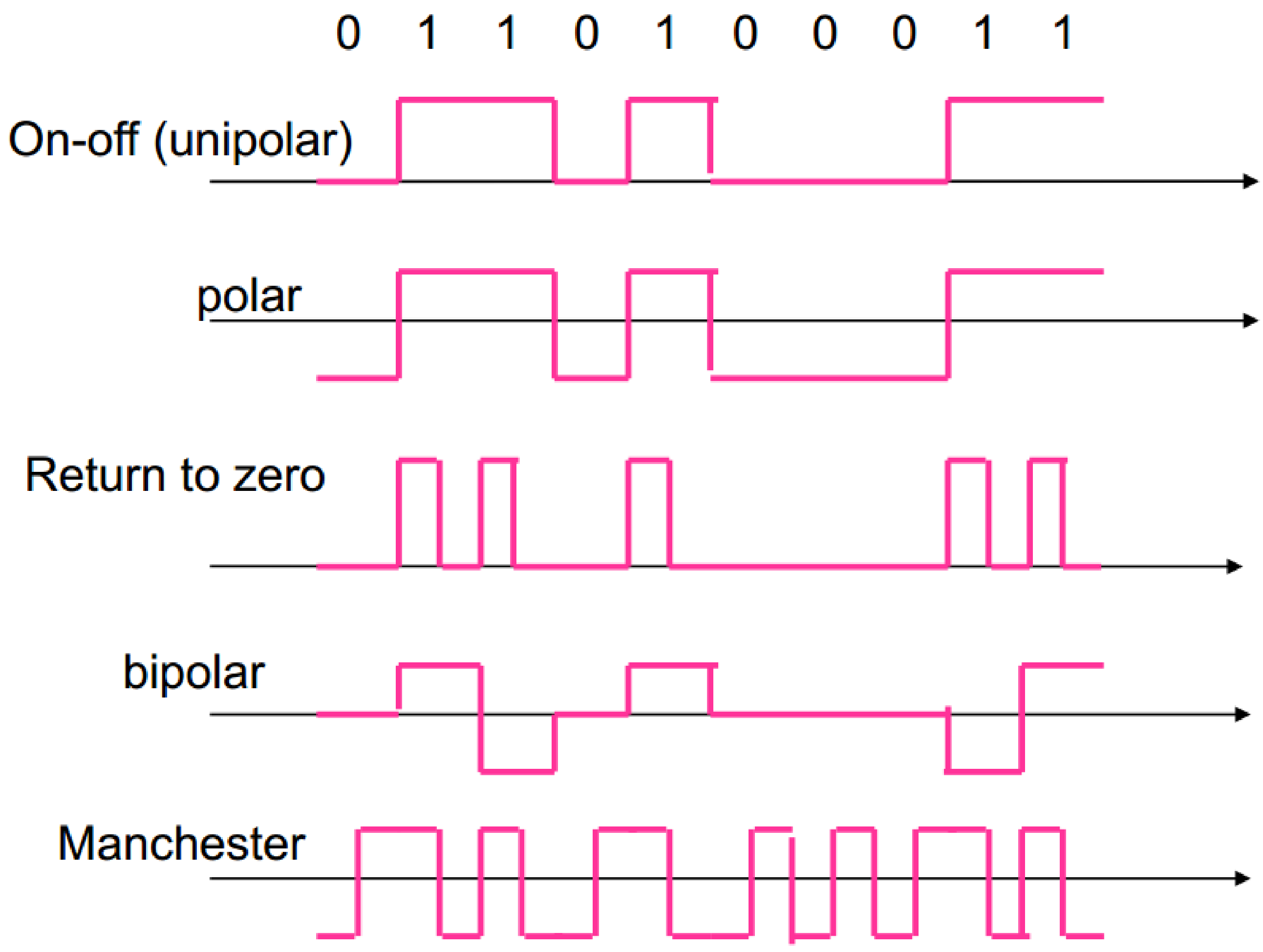
- Baseband signaling waveforms
  - To send the encoded digital data over a baseband channel, we require the use of format or waveform for representing the data
  - System requirement on digital waveforms
    1. Easy to synchronize
    2. High spectrum utilization efficiency
    3. Good noise immunity
    4. No DC component and little low frequency component
    5. Self-error-correction capability
    6. Et al.

Digital transmission through bandlimited channels

- Basic waveforms
  - On-off or unipolar signaling
  - Polar signaling
  - Return-to-zero signaling
  - Bipolar signaling (useful because no DC)
  - Split-phase or Manchester code (no DC)
  - Et al.

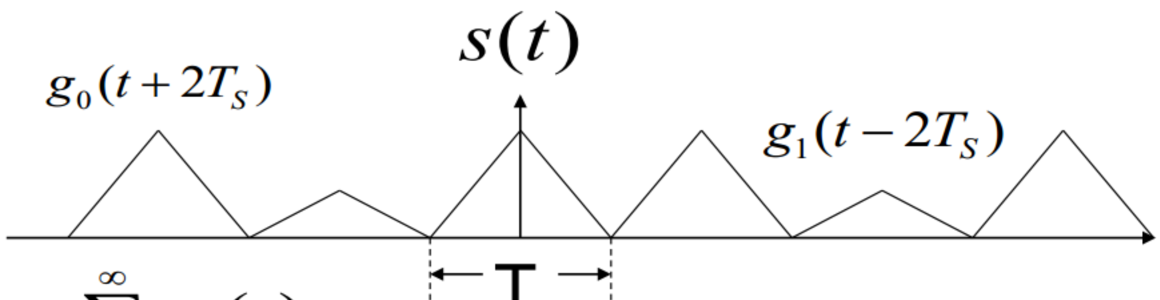
Digital transmission through bandlimited channels

- Basic waveforms



Digital transmission through bandlimited channels

- Spectra of baseband signals
  - Consider a random binary sequence $g_0(t)$ – 0, $g_1(t)$ – 1
  - The pulses $g_0(t)$ and $g_1(t)$ occur independently with probabilities given by p and 1-p, respectively. The duration of each pulse is given by Ts.



$$s(t) = \sum_{n=-\infty}^{\infty} s_n(t)$$

$$s_n(t) = \begin{cases} g_0(t-nT_s), & \text{with prob. } P \\ g_1(t-nT_s), & \text{with prob. } 1-P \end{cases}$$

Digital transmission through bandlimited channels

- Spectra of baseband signals
  - PSD of the baseband signal s(t) is[1]

$$S(f)=\frac{1}{T_s}p(1-p)\left|G_0(f)-G_1(f)\right|^2+\frac{1}{T_s^2}\sum_{m=-\infty}^{\infty}\left|pG_0(\frac{m}{T_s})+(1-p)G_1(\frac{m}{T_s})\right|^2\delta(f-\frac{m}{T_s})$$

    1st term is the continuous freq. component

    2nd term is the discrete freq. component
  - For polar signaling with $g_0(t)=-g_1(t)=g(t)$ and p=1/2

$$S(f)=\frac{1}{T}\left|G(f)\right|^2$$

  - For unipolar signaling with $g_0(t)=0$ $g_1(t)=g(t)$ and p=1/2, and g(t) is a rectangular pulse

$$G(f)=T\left[\frac{\sin\pi fT}{\pi fT}\right] \Rightarrow S_x(f)=\frac{T}{4}\left[\frac{\sin\pi fT}{\pi fT}\right]^2+\frac{1}{4}\delta(f)$$

1. R.C. Titsworth and L. R. Welch, "**Power spectra of signals modulated by random and pseduraandom sequences**," JPL, CA, Technical Report, Oct. 1961.

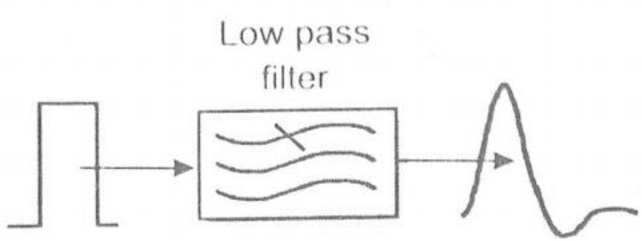
Digital transmission through bandlimited channels

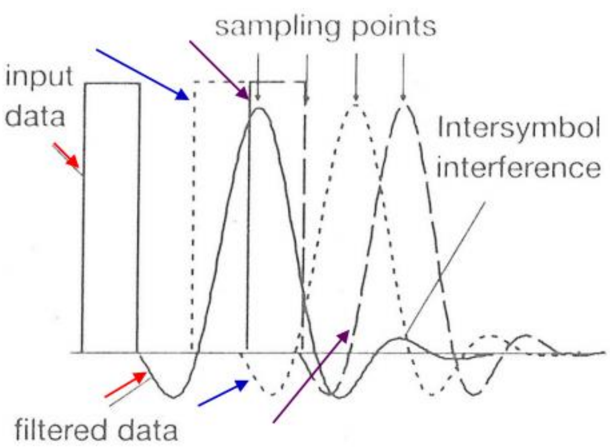
- Spectra of baseband signals
  - For return-to-zero unipolar signaling $\tau = T/2$

$$S_x(f) = \frac{T}{16}\left[\frac{\sin \pi fT/2}{\pi fT/2}\right]^2 + \frac{1}{16}\delta(f) + \frac{1}{4}\sum_{\text{odd } m}\frac{1}{[m\pi]^2}\delta(f - \frac{m}{T})$$

Digital transmission through bandlimited channels

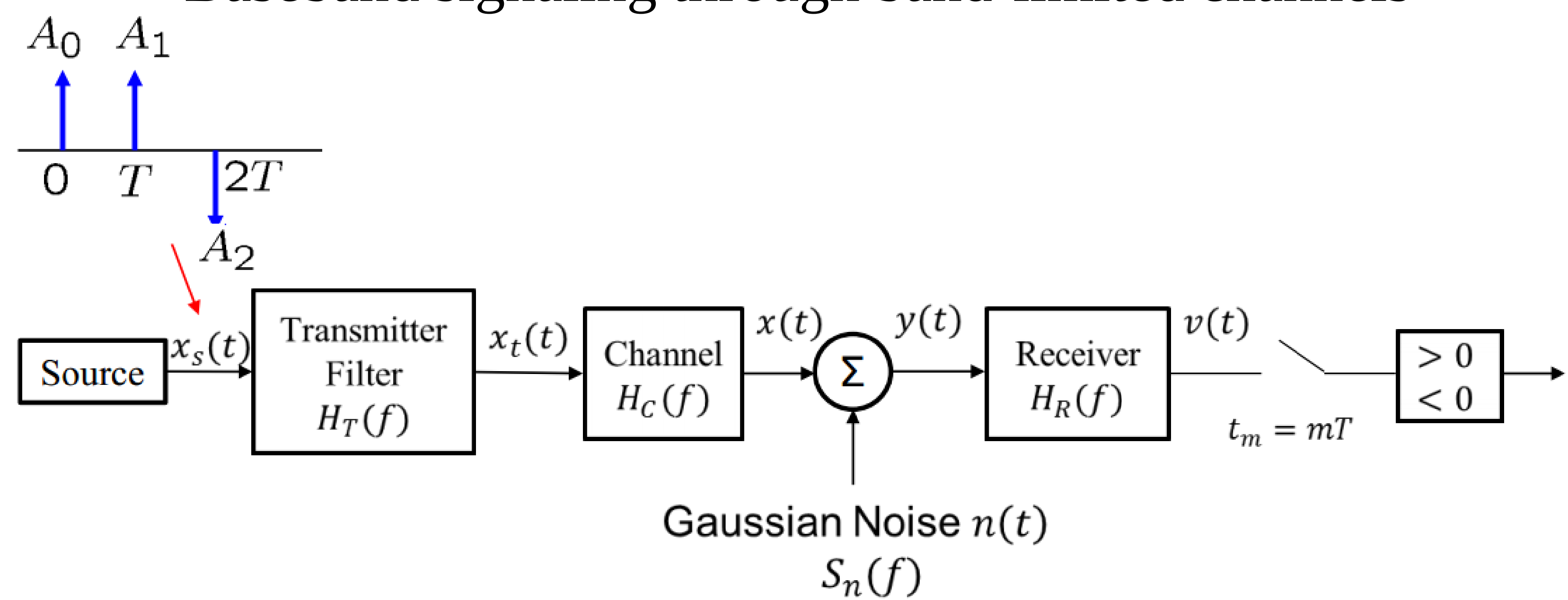
- Inter-symbol interference
  - ➢ The filtering effect of the band-limited channel will cause a spreading of individual data symbols passing through
  - ➢ For consecutive symbols, this spreading causes part of symbol energy to overlap with neighboring symbols, causing inter-symbol interference





Digital transmission through bandlimited channels

- Baseband signaling through band-limited channels



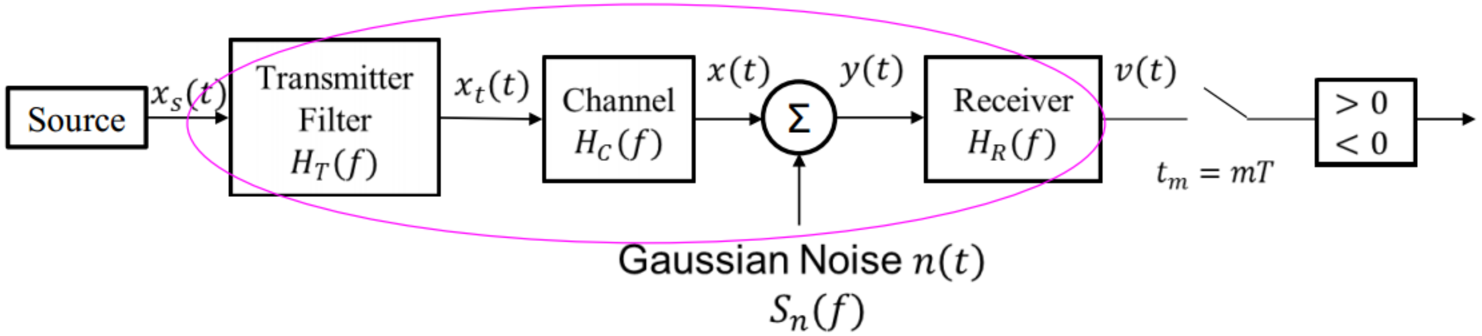
Input to tx filter
$$x_s(t) = \sum_{i=-\infty}^{\infty} A_i \delta(t - iT)$$

Output of tx filter
$$x_t(t) = \sum_{i=-\infty}^{\infty} A_i h_T(t - iT)$$

Output of rx filter
$$v(t) = x_s(t) * h_T(t) * h_c(t) * h_R(t) + n(t) * h_R(t)$$

Digital transmission through bandlimited channels

- Baseband signaling through band-limited channels



- Pulse shape at the receiver filter output

$$p(t) = h_T(t) * h_c(t) * h_R(t)$$

- Overall frequency response

$$P(f) = H_T(f)H_C(f)H_R(f)$$

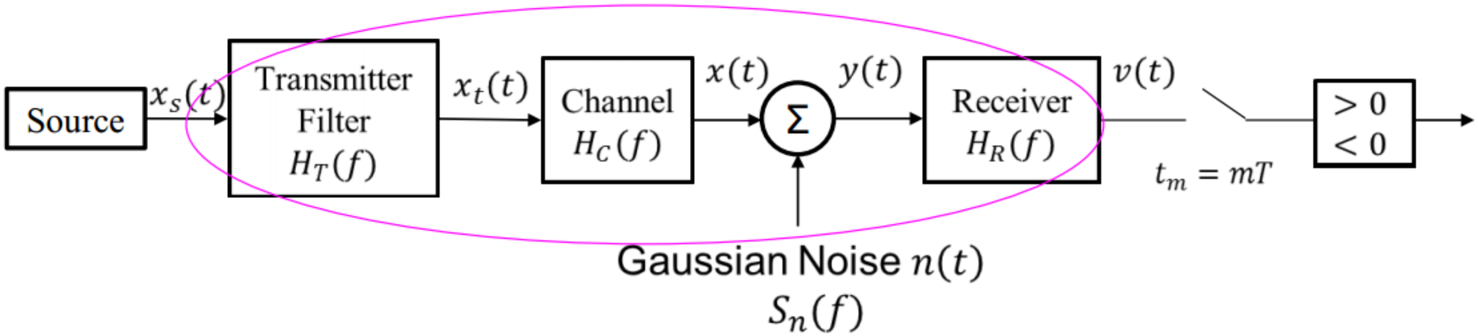
- Receiving filter output

$$v(t) = \sum_{k=-\infty}^{\infty} A_k p(t-kT) + n_o(t)$$

$$n_o(t) = n(t) * h_R(t)$$

Digital transmission through bandlimited channels

- Baseband signaling through band-limited channels

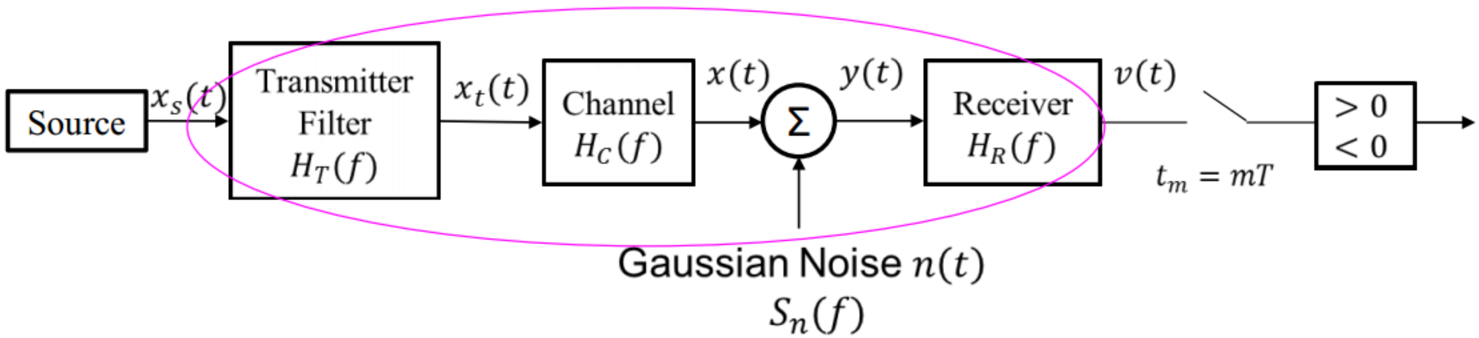


- ➢ Sample the receiver filter output v(t) at t_m=mT to detect A_m

$$v(t_m) = \sum_{k=-\infty}^{\infty} A_k p(mT - kT) + n_o(t_m)$$

$$= \underbrace{A_m p(0)}_{\text{Desired signal}} + \underbrace{\sum_{\substack{k=-\infty \\ k\neq m}}^{\infty} A_k p[(m-k)T]}_{\text{intersymbol interference (ISI)}} + \underbrace{n_o(t_m)}_{\text{Gaussian noise}}$$

- Baseband signaling through band-limited channels



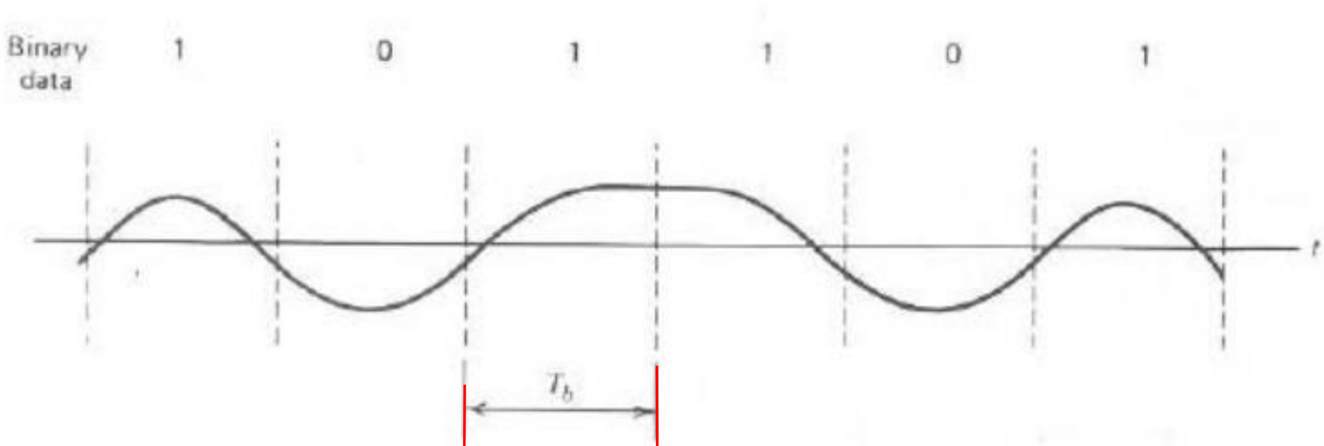
- Sample the receiver filter output v(t) at $t_m=mT$ to detect A_m

$$v(t_m) = \sum_{k=-\infty}^{\infty} A_k p(mT - kT) + n_o(t_m)$$

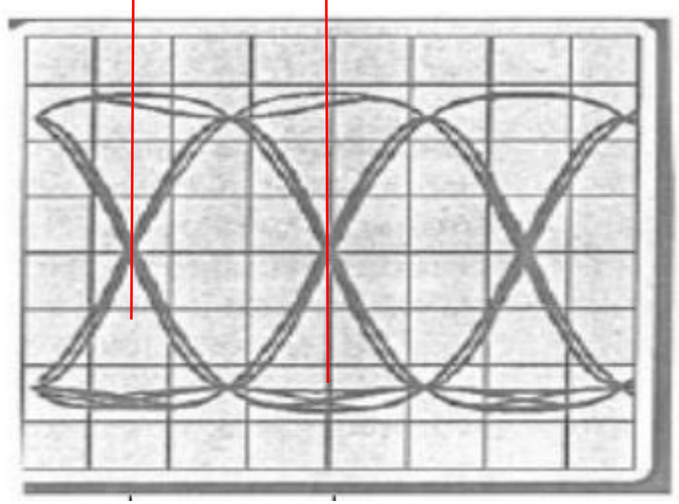
$$= \underbrace{A_m p(0)}_{\text{Desired signal}} + \underbrace{\sum_{\substack{k=-\infty \\ k\neq m}}^{\infty} A_k p[(m-k)T]}_{\text{intersymbol interference (ISI)}} + \underbrace{n_o(t_m)}_{\text{Gaussian noise}}$$

Digital transmission through bandlimited channels

- Eye diagram
  - Distorted binary wave



  - Eye diagram



Digital transmission through bandlimited channels

- ISI minimization
  - Choose transmitter and receiver filters which shape the received pulse function to eliminate or minimize interference between adjacent pulses, hence not to degrade the bit error rate performance of the link

Digital transmission through bandlimited channels

- Signal design for band-limited channel zero ISI
  - Nyquist condition for zero ISI for pulse shape p(t)

$$p(nT) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$$

Echos made to be zero at sampling points

or $$\sum_{k=-\infty}^{\infty} P\left(f + \frac{k}{T}\right) = \text{constant}$$

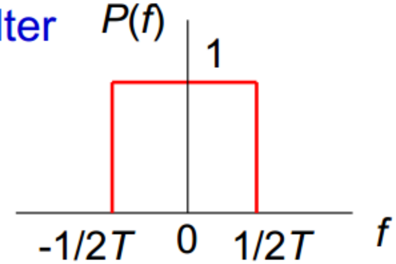
  - With the above condition, the receiver output simplifies to

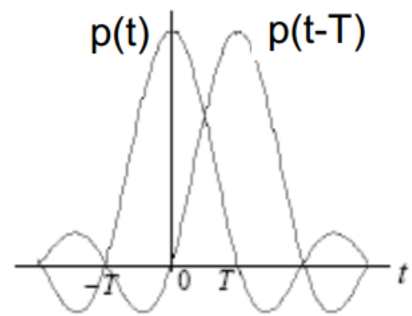
$$v(t_m) = A_m + n_o(t_m)$$

- Signal design for band-limited channel zero ISI
  - Nyquist's first method for eliminating ISI is to use

$$P(f) = \begin{cases} 1 & |f| < \frac{1}{2T} \\ 0 & \text{otherwise} \end{cases} \quad \Longleftrightarrow \quad p(t) = \frac{\sin(\pi t/T)}{\pi t/T} = \mathrm{sinc}\left(\frac{t}{T}\right)$$

"brick wall" filter





$$B_0 = \frac{1}{2T} = \frac{R_s}{2}$$ = Nyquist bandwdith,

> The minimum transmission bandwidth for zero ISI. A channel with bandwidth B_0 can support a maximal transmission rate of $2B_0$ symbols/sec

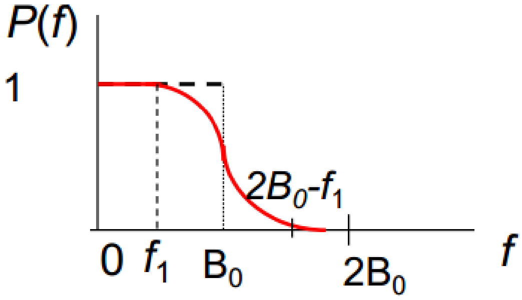
Digital transmission through bandlimited channels

- Signal design for band-limited channel zero ISI
  - Challenges of designing such p(t) or P(f)
    1. P(f) is physically unrealizable due to the abrupt transitions at B0
    2. p(t) decays slowly for large t, resulting in little margin of error in sampling times in the receiver
    3. This demands accurate sample point timing-a major challenge in current modem/data receiver design
    4. Inaccuracy in symbol timing is referred to as timing jitter.

Digital transmission through bandlimited channels

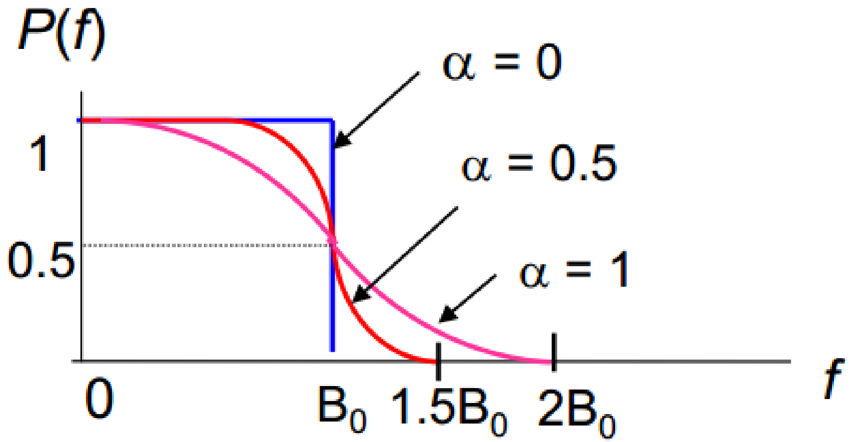
- Signal design for band-limited channel zero ISI
  - Raised cosine filter.
  - P(f) is made up of 3 parts: pass band, stop band, and transition band. The transition band is shaped like a cosine wave.

$$P(f)=\begin{cases} 1 & 0\le |f| < f_1 \\ \dfrac{1}{2}\left\{1+\cos\left[\dfrac{\pi(|f|-f_1)}{2B_0-2f_1}\right]\right\} & f_1\le |f| < 2B_0-f_1 \\ 0 & |f|\ge 2B_0-f_1 \end{cases}$$



Digital transmission through bandlimited channels

- Signal design for band-limited channel zero ISI
  - Raised cosine filter.



Roll-off factor

$$\alpha = 1 - \frac{f_1}{B_0}$$

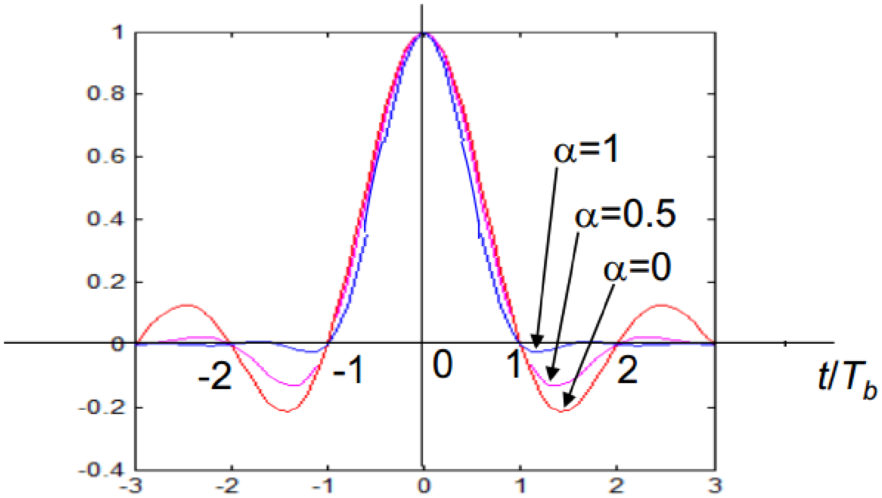
  - The sharpness of the filter is controlled by *a*.
  - Required bandwidth $B = B_0(1+a)$.

- Signal design for band-limited channel zero ISI
  - Raised cosine filter.
  - Taking the inverse Fourier transform

$$p(t) = \text{sinc}(2B_0 t)\frac{\cos(2\pi\alpha B_0 t)}{1-16\alpha^2 B_0^2 t^2}$$

Ensures zero crossing at desired sampling instants

Decreases as $1/t^2$, such that the data receiving is relatively insensitive to sampling time error



Digital transmission through bandlimited channels

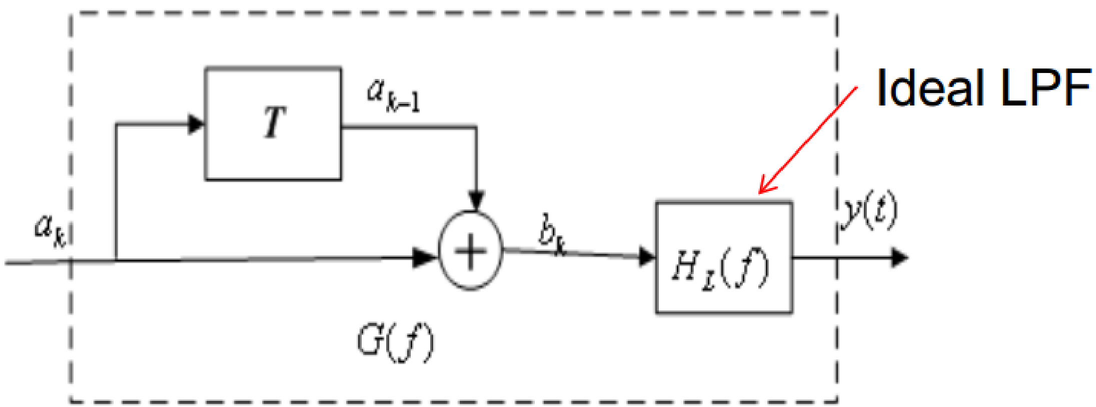
- Signal design for band-limited channel zero ISI
  - Raised cosine filter.
  - Small a: higher bandwidth efficiency
  - Large a: simpler filter with fewer stages hence easier to implement; less sensitive to symbol timing accuracy

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Relax the condition of zero ISI and allow a controlled amount of ISI
  - Then we can achieve the maximal symbol rate of 2W symbols/sec
  - The ISI we introduce is deterministic or controlled; hence it can be taken into account at the receiver

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Duobinary signal.
  - Let $\{a_k\}$ be the binary sequence to be transmitted. The pulse duration is T.
  - Two adjacent pulses are added together, i.e., $b_k=a_k+a_{k-1}$



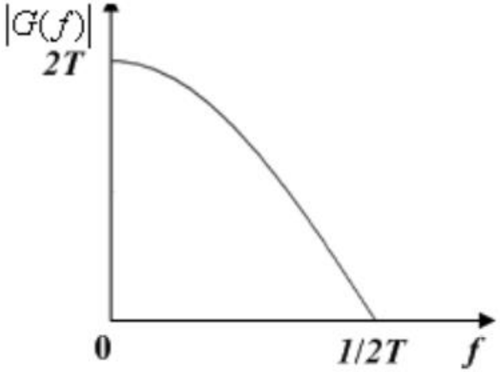
  - The resulting sequence $\{b_k\}$ is called duobinary signal

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Duobinary signal: frequency domain.

$$G(f) = \left(1 + e^{-j2\pi fT}\right) H_L(f) \qquad H_L(f) = \begin{cases} T & (|f| \le 1/2T) \\ 0 & (\text{otherwise}) \end{cases}$$

$$= \begin{cases} 2Te^{-j\pi fT} \cos \pi fT & (|f| \le 1/2T) \\ 0 & (\text{otherwise}) \end{cases}$$



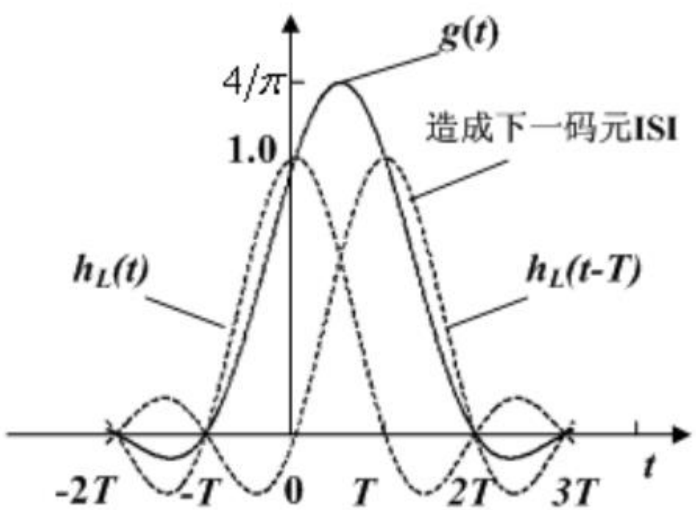
- Signal design with controlled ISI – partial response signals
  - Duobinary signal: time domain.

$$g(t)=[\delta(t)+\delta(t-T)]*h_L(t) \quad =\frac{\sin \pi t/T}{\pi t/T}+\frac{\sin \pi(t-T)/T}{\pi(t-T)/T}$$

$$=\operatorname{sinc}\left(\frac{t}{T}\right)+\operatorname{sinc}\left(\frac{t-T}{T}\right)=\frac{T^2}{\pi t}\cdot\frac{\sin \pi t/T}{(T-t)}$$

  - g(t) is called a duobinary signal pulse
  - $g(0)=g_0=1$
  - $g(T)=g_1=1$
  - $g(iT)=g_i=0, i \neq 1$

(b)



Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Duobinary signal: decoding.
  - Without noise, the received signal is the same as the transmitted signal

$$y_k = \sum_{i=0}^{\infty} a_i g_{k-i} = a_k + a_{k-1} = b_k$$

A 3-level sequence

  - When {ak} is a polar sequence with values +1 or -1

$$y_k = b_k = \begin{cases} 2 & (a_k = a_{k-1} = 1) \\ 0 & (a_k = 1, a_{k-1} = -1 \text{ or } a_k = -1, a_{k-1} = 1) \\ -2 & (a_k = a_{k-1} = -1) \end{cases}$$

  - When {ak} is a unipolar sequence with values 1 or 0

$$y_k = b_k = \begin{cases} 0 & (a_k = a_{k-1} = 0) \\ 1 & (a_k = 0, a_{k-1} = 1 \text{ or } a_k = 1, a_{k-1} = 0) \\ 2 & (a_k = a_{k-1} = 1) \end{cases}$$

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Duobinary signal: decoding.
  - To recover the transmitted sequence, we can use

$$\hat{a}_k = b_k - \hat{a}_{k-1} = y_k - \hat{a}_{k-1}$$

    although the detection of the current symbol relies on the detection of the previous symbol→ error propagation will occur
  - How to solve the ambiguity problem and error propagation?
  - Precoding: Apply differential encoding on {ak} so that

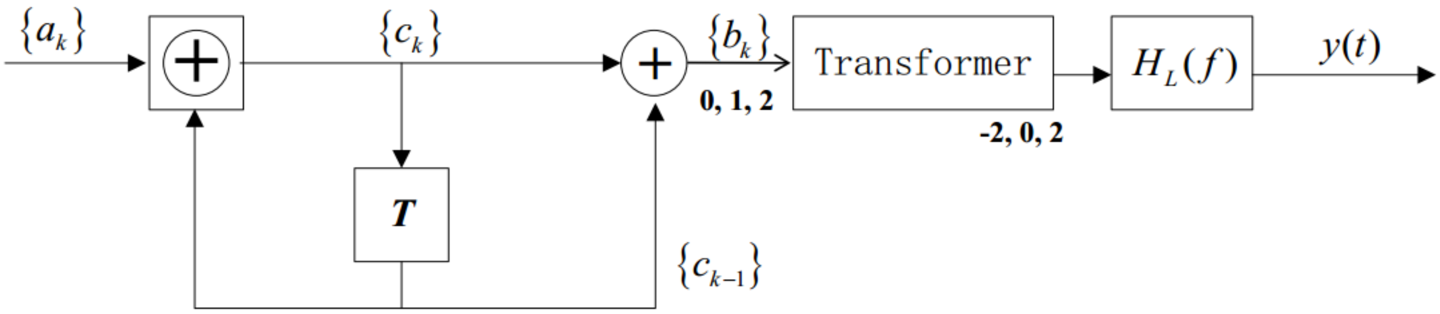
$$c_k = a_k \oplus c_{k-1}$$

Then the output of the duobinary signal system is

$$b_k = c_k + c_{k-1}$$

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Duobinary signal: decoding.
  - Block diagram of precoded duobinary signal



Digital transmission through bandlimited channels

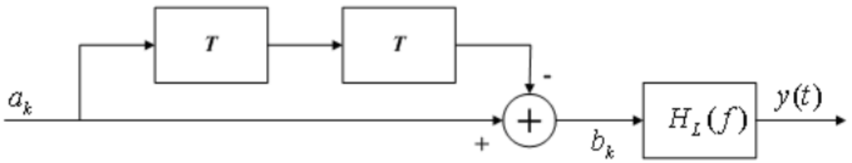
- Signal design with controlled ISI – partial response signals
  - Modified duobinary signal

$$b_k = a_k - a_{k-2}$$

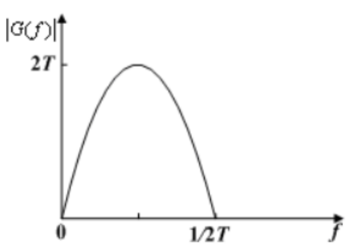
  - After LPF H(f), the overall response is

$$G(f) = (1 - e^{-j4\pi fT})H_L(f) \quad = \begin{cases} 2Tje^{-j2\pi fT}\sin 2\pi fT & (|f| \le 1/2T) \\ 0 & \text{otherwise} \end{cases}$$

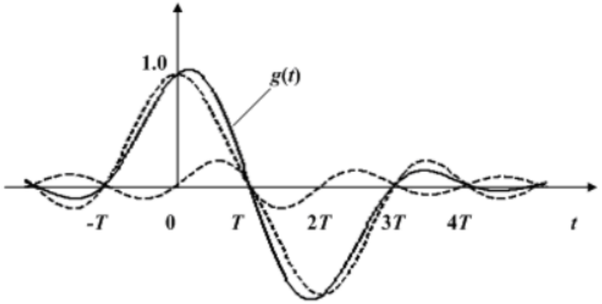
$$g(t) = \frac{\sin \pi t / T}{\pi t / T} - \frac{\sin \pi (t-2T)/T}{\pi (t-2T)/T} \quad = -\frac{2T^2 \sin \pi t / T}{\pi t (t-2T)}$$



(a)



(b)



Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Modified duobinary signal.
  - The magnitude spectrum is a half-sin wave and hence easy to implement
  - No DC component and small low freq. component
  - At sampling interval T, the sampled values are

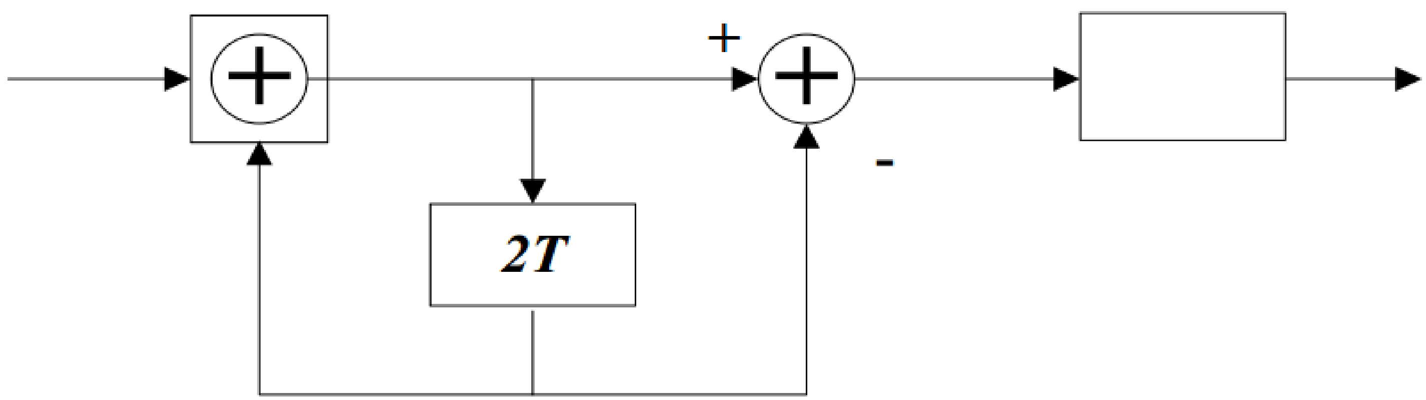
$$g(0) = g_0 = 1$$
$$g(T) = g_1 = 0$$
$$g(2T) = g_2 = -1$$
$$g(iT) = g_i = 0, i \neq 0, 1, 2$$

  - g(t) decays as $1/t^2$. But time offset may cause significant problem.

Digital transmission through bandlimited channels

- Signal design with controlled ISI – partial response signals
  - Modified duobinary signal: decoding.
  - To overcome error propagation, precoding is also needed $c_k = a_k \oplus c_{k-2}$
  - The coded signal is

$$b_k = c_k - c_{k-2}$$



Digital transmission through bandlimited channels

- Update
  - We have discussed
    1. Pulse shapes of baseband signal and their power spectrum
    2. ISI in band-limited channels
    3. Signal design for zero ISI and controlled ISI
  - We will discuss system design in the presence of channel distortion
    1. Optimal transmitting and receiving filters
    2. Channel equalizer

Digital transmission through bandlimited channels

- Optimal transmit/receive filter
  - Recall that when zero-ISI condition is satisfied by p(t) with raised cosine spectrum P(f), then the sampled output of the receiver filter is Vm=Am+Nm (assume p(0)=1)
  - Consider binary PAM transmission: Am = $\pm$d
  - Variance of $N_m = \sigma^2 = \int_{-\infty}^{\infty} S_n(f)|H_R(f)|^2 df$ with $P(f) = H_T(f)H_C(f)H_R(f)$ $\quad p(t) = h_T(t) * h_C(t) * h_R(t)$

$$\Longrightarrow \quad P_e = Q\left(\frac{d}{\sigma}\right)$$

Error Probability can be minimized through a proper choice of H_R(f) and H_T(f) so that d/σ is maximum (assuming H_C(f) fixed and P(f) given)

Digital transmission through bandlimited channels

- Optimal transmit/receive filter
  - Compensate the channel distortion equally between the transmitter and receiver filters

$$\begin{cases} |H_T(f)| = \dfrac{\sqrt{P(f)}}{|H_c(f)|^{1/2}} \\ |H_R(f)| = \dfrac{\sqrt{P(f)}}{|H_c(f)|^{1/2}} \end{cases} \quad \text{for } |f| \leq W$$

  - Then, the transmit signal energy is given by

$$E_{av} = \int_{-\infty}^{\infty} d^2 h_T^2(t)dt = \int_{-\infty}^{\infty} d^2 H_T^2(f)df = \int_{-W}^{W} \frac{d^2 P(f)}{|H_C(f)|}df$$

By Parseval's theorem

  - Hence $d^2 = E_{av} \cdot \left[\int_{-W}^{W} \frac{P(f)}{|H_C(f)|}df\right]^{-1}$

Digital transmission through bandlimited channels

- Optimal transmit/receive filter
  - Noise variance at the output of the receive filter is

$$\sigma^2 = \frac{N_0}{2}\int_{-\infty}^{\infty} |H_R(f)|^2 df = \frac{N_0}{2}\int_{-W}^{W} \frac{P(f)}{|H_C(f)|} df$$

$$\Rightarrow \quad P_{e,\min} = Q\left[\sqrt{\frac{2E_{av}}{N_0}\underbrace{\left\{\int_{-W}^{W}\frac{P(f)}{|H_c(f)|}df\right\}^{-1}}_{\text{Performance loss due to channel distortion}}}\right]$$

  - Special case: Hc(f)=1 for |f|≤W
    1. This is the idea case with “flat” fading
    2. No loss, same as the matched filter receiver of AWGN channel

Digital transmission through baseband channels

- Optimal transmit/receive filter
  - Exercise.
  - Determine the optimum transmitting and receiving filters for a binary communication system that transmits data at a rate R=1/T=4800 bps over a channel with a frequency response $|H_c(f)| = \frac{1}{\sqrt{1+(\frac{f}{W})^2}}$, $|f| \leq W$ where W=4800 Hz
  - The additive noise is zero mean white Gaussian with spectral density $N_0/2=10^{-15}$ Watt/Hz

Digital transmission through baseband channels

- Optimal transmit/receive filter
  - Exercise.
  - Since W=1/T=4800, we use a signal pulse with a raised cosine spectrum and a roll-off factor =1.
  - Thus,

$$P(f)=\frac{1}{2}\left[1+\cos(\pi T\,|\,f\,|)\right]=\cos^2\left(\frac{\pi\,|\,f\,|}{9600}\right)$$

  - Therefore

$$|H_T(f)|=|H_R(f)|=\cos\left(\frac{\pi|f|}{9600}\right)\left[1+\left(\frac{f}{4800}\right)^2\right]^{1/4}, \text{for } |f|\leq 4800$$

  - One can now use these filters to determine the amount of transmit energy required to achieve a specified error probability.

Digital transmission through bandlimited channels

- Performance with ISI
  - If zero-ISI condition is not met, then

$$V_m = A_m + \sum_{k \neq m} A_k p[(m-k)T] + N_m$$

  - Let

$$A_I = \sum_{k \neq m} I_k = \sum_{k \neq m} A_k p[(m-k)T]$$

  - Then

$$V_m = A_m + A_I + N_m$$

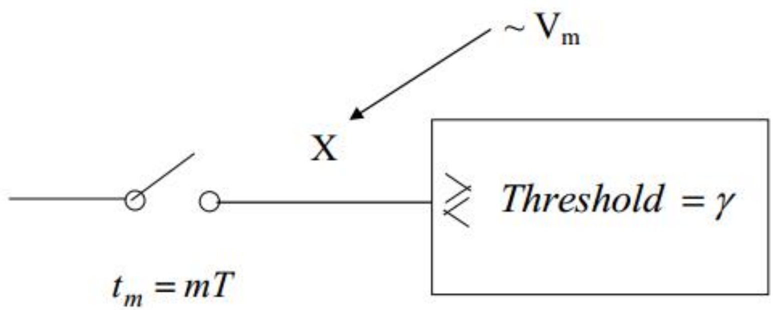
  - Often only 2M significant terms are considered. Hence

$$V_m = A_m + A_I' + N_m \quad \text{with} \quad A_I' = \sum_{k=-M}^{M} A_k p[(m-k)T]$$

  - Finding the probability of error?

Digital transmission through bandlimited channels

- Performance with ISI
  - ➢ Monte Carlo simulation.



Let

$$I(x) = \begin{cases} 1 & \textit{error occurs} \\ 0 & \textit{else} \end{cases}$$

$$\therefore \quad \left\| P_e = \frac{1}{L} \sum_{l=1}^{L} I\left(X^{(l)}\right) \right.$$

where $X^{(1)}, X^{(2)}, \ldots, X^{(L)}$ are i.i.d. (*independent and identically distributed*) random samples

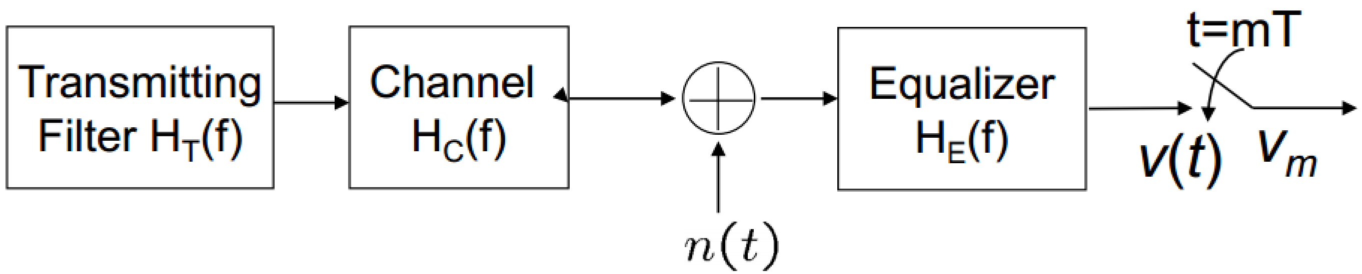
- Performance with ISI
  - Monte Carlo simulation.
  - If one want Pe to be within 10% accuracy, how many independent simulation runs do we need?
  - If Pe~10-9 (this is typically the case for optical communication systems), and assume each simulation run takes 1 ms, how long will the simulation take?

Digital transmission through bandlimited channels

- Update
  - Monte Carlo simulation.
  - We have shown that by properly designing the transmitting and receiving filters, one can guarantee zero ISI at sampling instants, thereby minimizing Pe.
  - Appropriate when the channel is precisely known and its characteristics do not change with time.
  - In practice, the channel is unknown or time-varying
  - We next consider channel equalizer.

Digital transmission through bandlimited channels

- Equalizer
  - A receiving filter with adjustable frequency response to minimize/eliminate inter-symbol interference



  - Overall frequency response

$$H_o(f) = H_T(f)H_C(f)H_E(f)$$

  - Nyquist criterion for zero-ISI

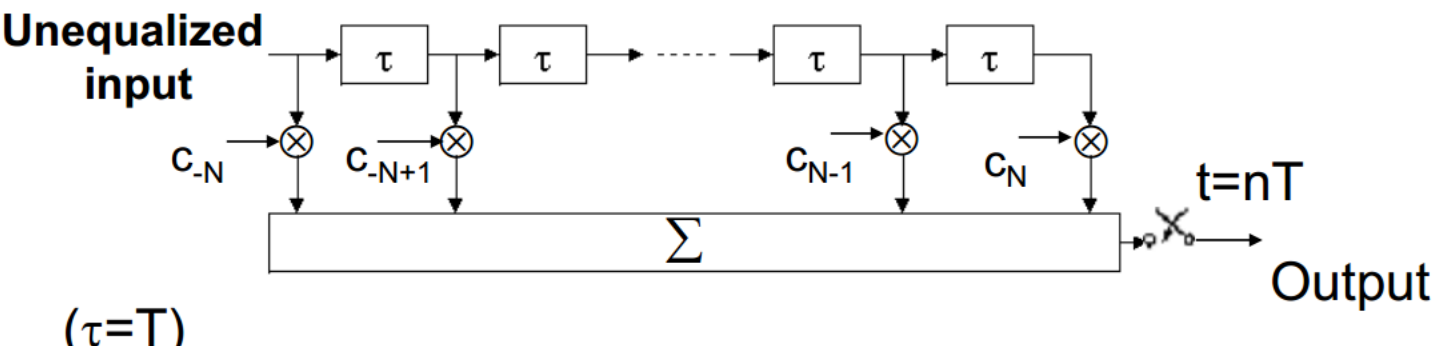
$$\sum_{k=-\infty}^{\infty} H_o\left(f + \frac{k}{T}\right) = \text{constant}$$

  - Thus, ideal zero-ISI equalizer is an inverse channel filter

$$H_E(f) \propto \frac{1}{H_T(f)H_C(f)} \qquad |f| \leq 1/2T$$

- Equalizer
  - ➢ Linear transversal filter
  - ➢ Finite impulse response (FIR) filter

$$h_E(t) = \sum_{n=-N}^{N} c_n \delta(t - nT)$$

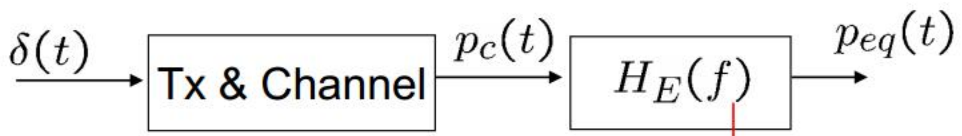


(2N+1)-tap FIR equalizer

- $\{c_n\}$ are the adjustable $2N+1$ equalizer coefficients
- N is sufficiently large to span the length of ISI

- Equalizer
  - Zero-forcing (ZF) equalizer

$P_c(t)$ the received pulse from a channel to be equalized



$$h_E(t) = \sum_{n=-N}^{N} c_n \delta(t - nT)$$

$$p_{eq}(t) = p_c(t) * h_E(t) = \sum_{n=-N}^{N} c_n p_c(t - nT)$$

$$t = mT$$

$$p_{eq}(mT) = \sum_{n=-N}^{N} c_n p_c[(m-n)t] = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \ldots, \pm N \end{cases}$$

To suppress 2N adjacent interference terms

Digital transmission through bandlimited channels

- Equalizer
  - Zero-forcing (ZF) equalizer
  - In matrix form

$$\mathbf{p}_{eq} = \mathbf{P}_c \cdot \mathbf{c}$$

$(2N+1)\times(2N+1)$ channel response matrix

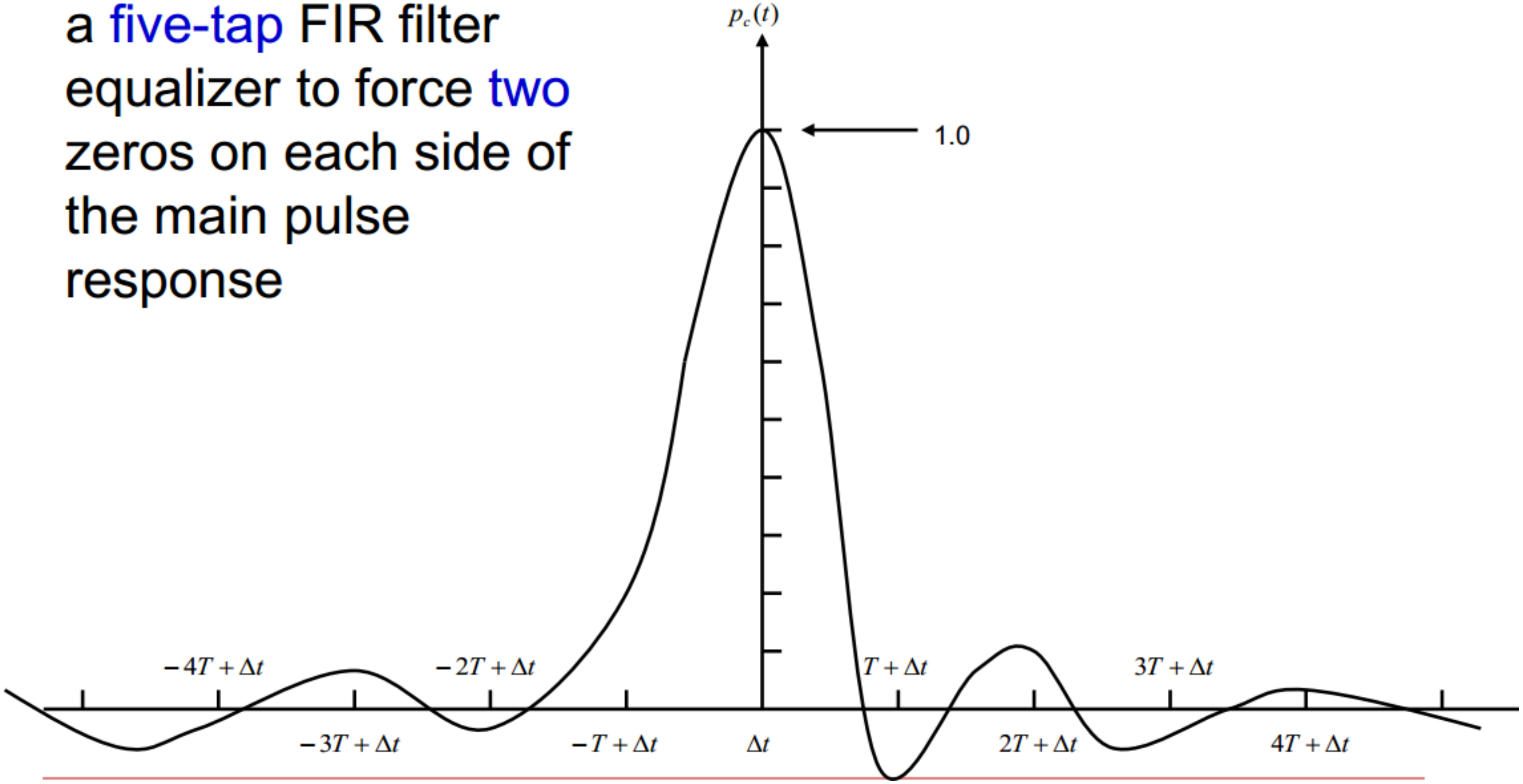
$$\mathbf{p}_{eq} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} c_{-N} \\ c_{-N+1} \\ \vdots \\ c_{-1} \\ c_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix} \quad \mathbf{P}_c = \begin{bmatrix} p_c(0) & p_c(-1) & \cdots & p_c(-2N) \\ p_c(1) & p_c(0) & \cdots & p_c(-2N+1) \\ \vdots & \vdots & \ddots & \vdots \\ p_c(2N) & p_c(2N-1) & \cdots & p_c(0) \end{bmatrix}$$

$\Rightarrow$ $c = \mathbf{P}_c^{-1}\mathbf{p}_{eq}$ or the middle-column of $\mathbf{P}_c^{-1}$

Digital transmission through bandlimited channels

- Equalizer
  - Example.

- Find the coefficients of a five-tap FIR filter equalizer to force two zeros on each side of the main pulse response



- Equalizer
  - Example.
  - By inspection

$$p_c(-4) = -0.02 \qquad p_c(0) = 1$$
$$p_c(-3) = 0.05 \qquad p_c(1) = -0.1$$
$$p_c(-2) = -0.1 \qquad p_c(2) = 0.1$$
$$p_c(-1) = 0.2 \qquad p_c(3) = -0.05$$
$$p_c(4) = 0.02$$

  - The channel response matrix

$$[P_c] = \begin{bmatrix} 1.0 & 0.2 & -0.1 & 0.05 & -0.02 \\ -0.1 & 1.0 & 0.2 & -0.1 & 0.05 \\ 0.1 & -0.1 & 1.0 & 0.2 & -0.1 \\ -0.05 & 0.1 & -0.1 & 1.0 & 0.2 \\ 0.02 & -0.05 & 0.1 & -0.1 & 1.0 \end{bmatrix}$$

Digital transmission through bandlimited channels

- Equalizer
  - Example.
  - The inverse of this matrix

$$P^{-1} = \begin{bmatrix} 0.9687 & -0.1796 & 0.1297 & 0.0085 & 0.0396 \\ 0.1179 & 0.9442 & -0.1562 & 0.1231 & -0.0851 \\ -0.0908 & 0.1343 & 0.9356 & -0.1664 & 0.1183 \\ 0.0279 & -0.0955 & 0.1338 & 0.9480 & -0.1709 \\ -0.0016 & 0.0278 & -0.0906 & 0.1174 & 0.9660 \end{bmatrix}$$

  - Therefore, $C_{-2}=0.1297$, $C_{-1}=-0.1562$, $C_0=0.9356$, $C_1=0.1338$, $C_2=-0.0906$,
  - Equalized pulse response $p_{eq}(m) = \sum_{n=-2}^{2} c_n p_c(m-n)$
  - It can be verified

$$p_{eq}(0) = 1.0 \qquad p_{eq}(m) = 0,\ m = \pm 1, \pm 2$$

Digital transmission through bandlimited channels

- Equalizer
  - Example.
  - Note that values of $p_{eq}(n)$ for $n<-2$ or $n>2$ are not zero. For example

$$p_{eq}(-3)=0.2*0.1297+(-0.1)*(-0.1562)+0.05*0.9356+(-0.02)*0.1338+(-0.05)*(-0.0906)=0.0902$$

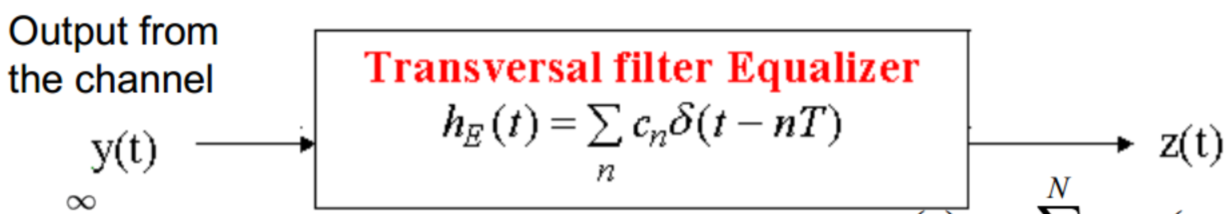
$$p_{eq}(-3)=0.005*0.1297+0.02*(-0.1562)+(-0.05)*0.9356+0.1*0.1338+(-0.1)*(-0.0906)=-0.0268$$

Digital transmission through bandlimited channels

- Equalizer
  - Minimum mean-square error equalizer.
  - Drawback of ZF equalizer: ignores the additive noise
  - Suppose we relax zero ISI condition, and minimize the combined power in the residual ISI and additive noise at the output of the equalizer.
  - Then, we have MMSE equalizer, which is a channel equalizer optimized based on the minimum mean square error (MMSE) criterion

Digital transmission through bandlimited channels

- Equalizer
  - ➢ Minimum mean-square error equalizer.



$$y(t) = \sum_{n=-\infty}^{\infty} A_n g_c(t - nT) + n(t)$$

$$z(t) = \sum_{n=-N}^{N} c_n y(t - nT)$$

  - ➢ The output is sampled at t=mT:

$$z(mT) = \sum_{n=-N}^{N} c_n y[(m-n)T]$$

  - ➢ Let Am=desired equalizer output

$$MSE = E\left[\left(z(mT) - A_m\right)^2\right] = \text{Minimum}$$

- Equalizer
  - Minimum mean-square error equalizer.

$$MSE = E\left[\left(\sum_{n=-\infty}^{\infty} c_n y[(m-n)T] - A_m\right)^2\right]$$
$$= \sum_{n=-N}^{N}\sum_{k=-N}^{N} c_n c_k R_Y(n-k) - 2\sum_{k=-N}^{N} c_k R_{AY}(k) + E(A_m^2)$$

where

$$\begin{cases} R_Y(n-k) = E[y(mT-nT)y(mT-kT)] \\ R_{AY}(k) = E[y(mT-kT)A_m] \end{cases}$$

E is taken over A_m and the additive noise

- MMSE solution is obtained by $\dfrac{\partial MSE}{\partial c_n} = 0$

$$\Rightarrow \sum_{N}^{N} c_n R_Y(n-k) = R_{AY}(k), \quad \text{for } k = 0, \pm 1, \ldots, \pm N.$$

Digital transmission through bandlimited channels

- Equalizer
  - MMSE equalizer vs. ZF equalizer.
  - Both can be obtained by solving similar equations.
  - ZF equalizer does not consider the effects of noise
  - MMSE equalizer is designed so that mean-square error (consisting of ISI terms and noise at the equalizer output) is minimized
  - Both equalizers are known as linear equalizers

Digital transmission through bandlimited channels

- Equalizer
  - ➢ Example of channels with ISI.

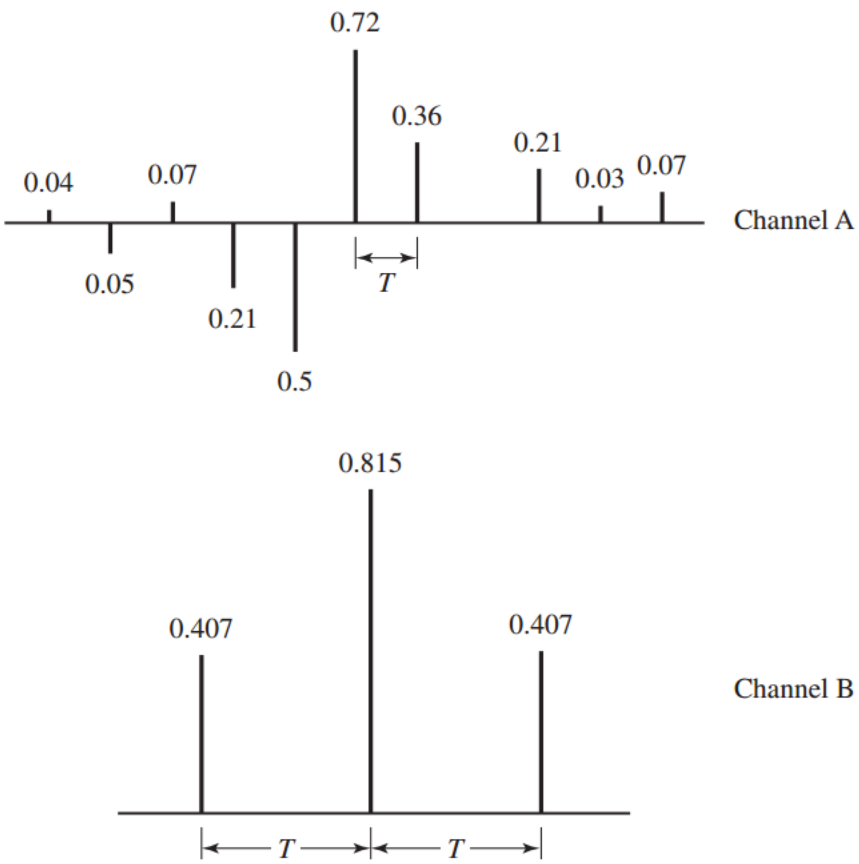


Figure 8.42 Two channels with ISI.

Digital transmission through bandlimited channels

- Equalizer
  - Frequency response.

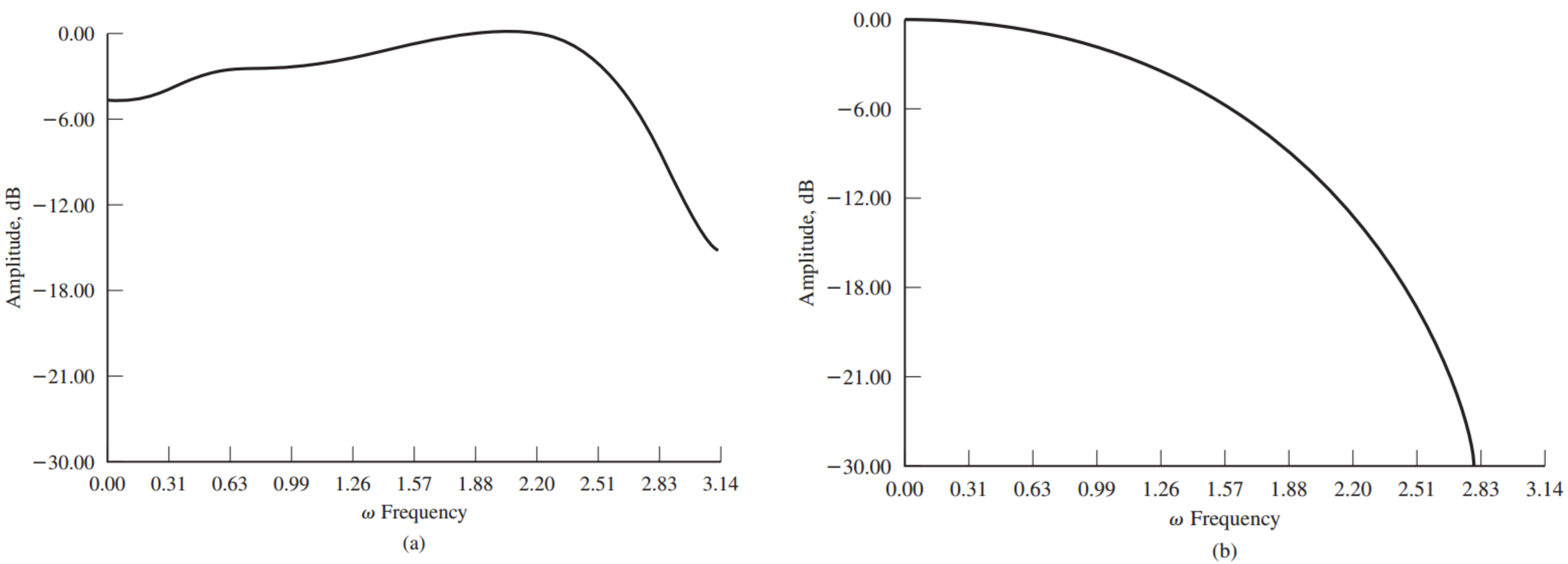


Figure 8.43 Amplitude spectra for (a) channel A shown in Figure 8.42(a) and (b) channel B shown in Figure 8.42(b).

Digital transmission through bandlimited channels

- Equalizer
  - Performance of MMSE equalizer.

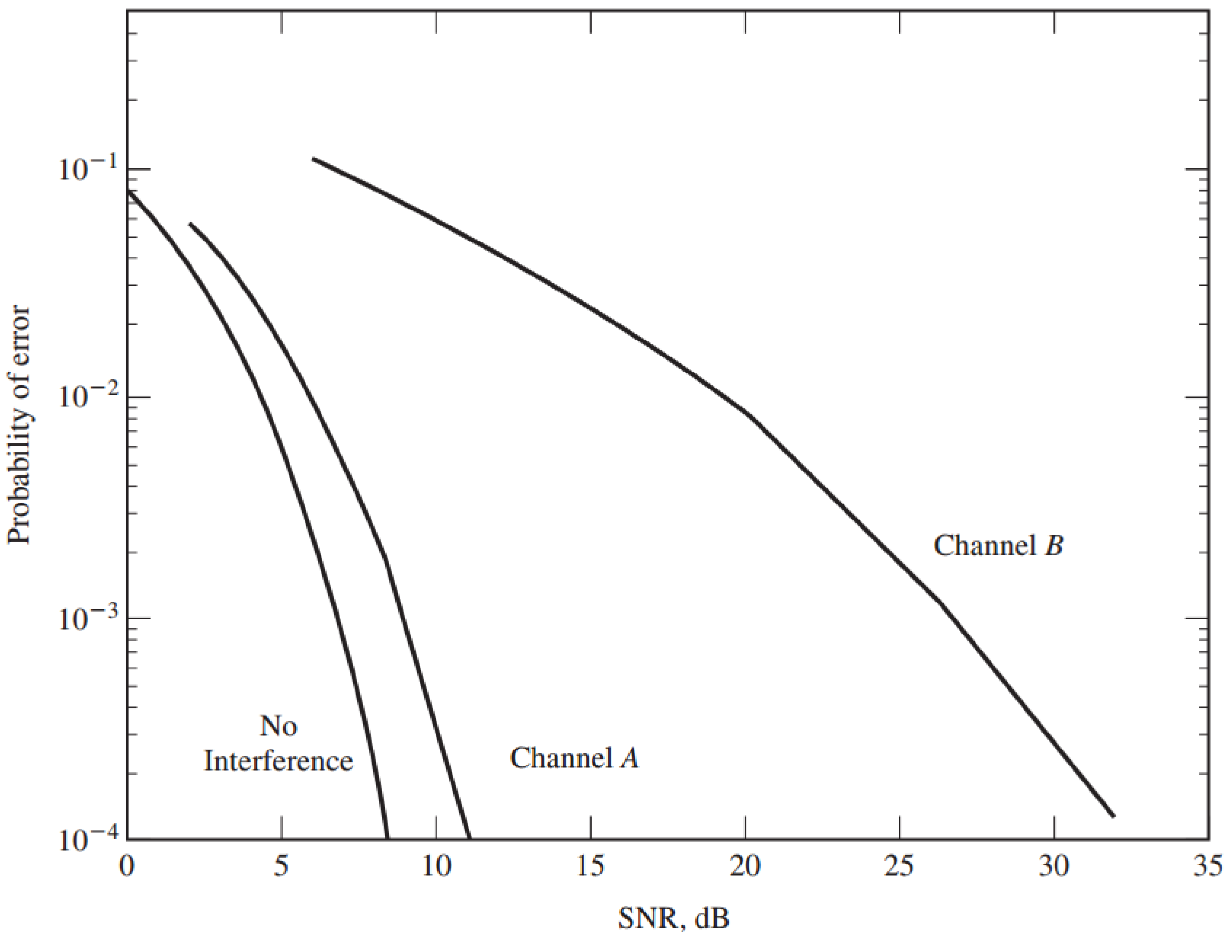
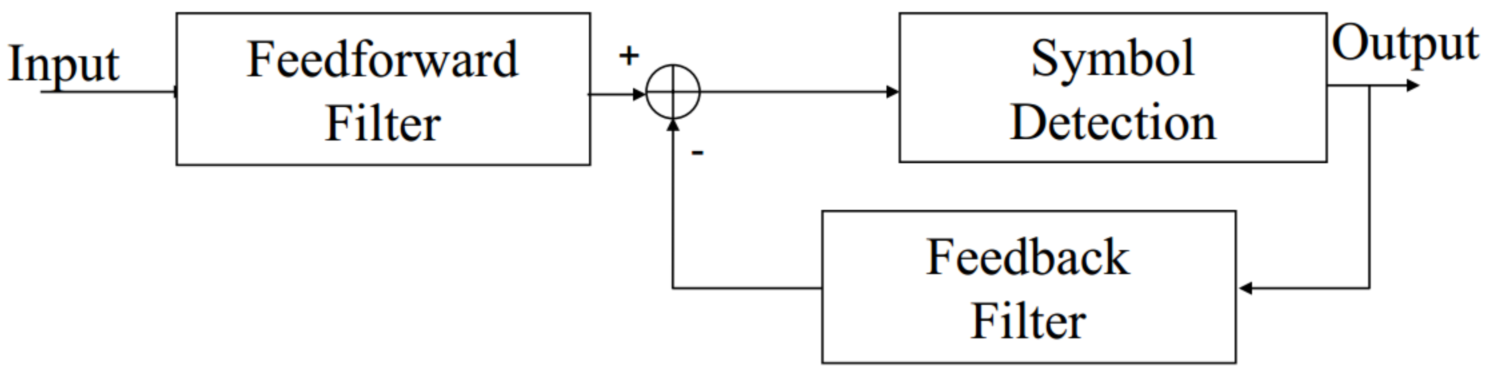


Figure 8.44 Error-rate performance of linear MSE equalizer.

Digital transmission through bandlimited channels

- Equalizer
  - Decision feedback equalizer (DFE)
  - DFE is a nonlinear equalizer which attempts to subtract from the current symbol to be detected the ISI created by previously detected symbol



- Equalizer
  - Performance of DFE equalizer.

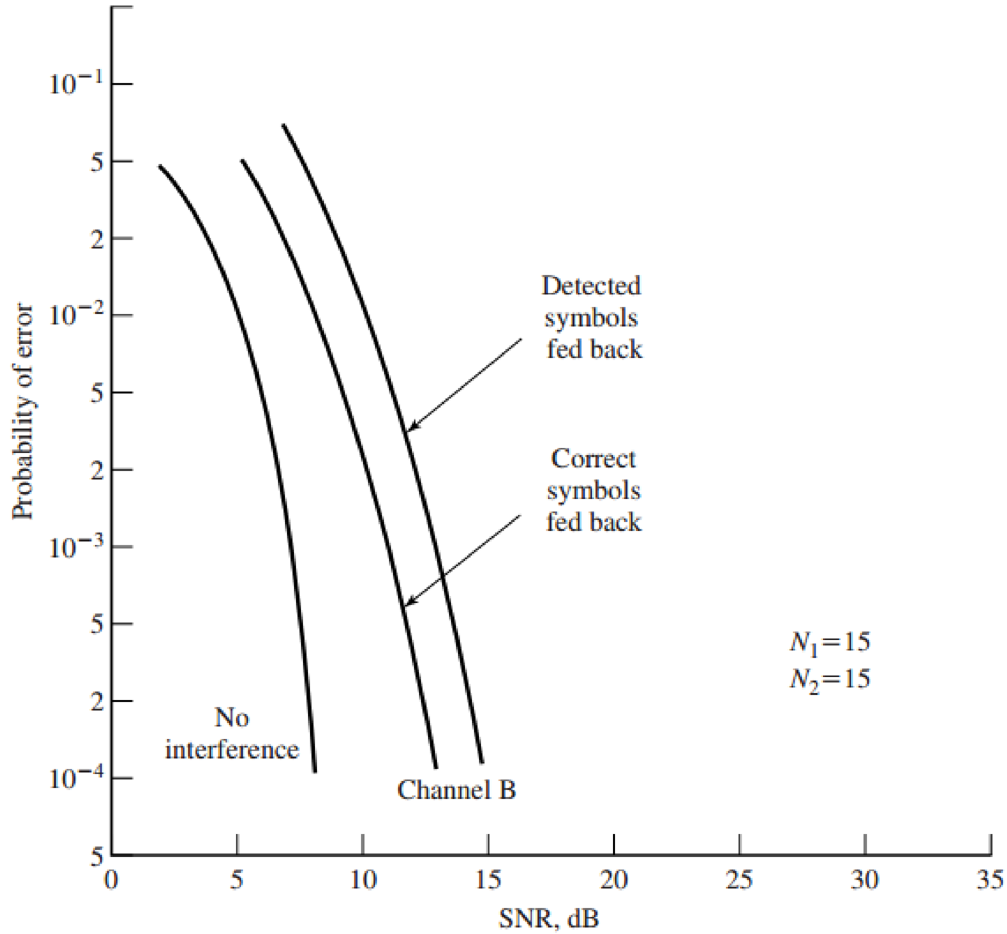
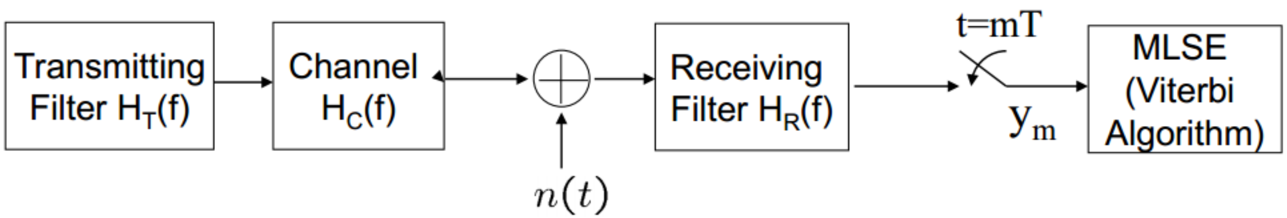


Figure 8.47 Performance of DFE with and without error propagation.

Digital transmission through bandlimited channels

- Equalizer
  - Maximum likelihood sequence estimation (MLSE).



  - Let the transmitting filter have a square root raised cosine frequency response

$$|H_T(f)| = \begin{cases} \sqrt{P(f)} & |f| \le W \\ 0 & |f| > W \end{cases}$$

  - The receiving filter is matched to the transmitter filter with

$$|H_R(f)| = \begin{cases} \sqrt{P(f)} & |f| \le W \\ 0 & |f| > W \end{cases}$$

  - The sampled output from receiving filter is

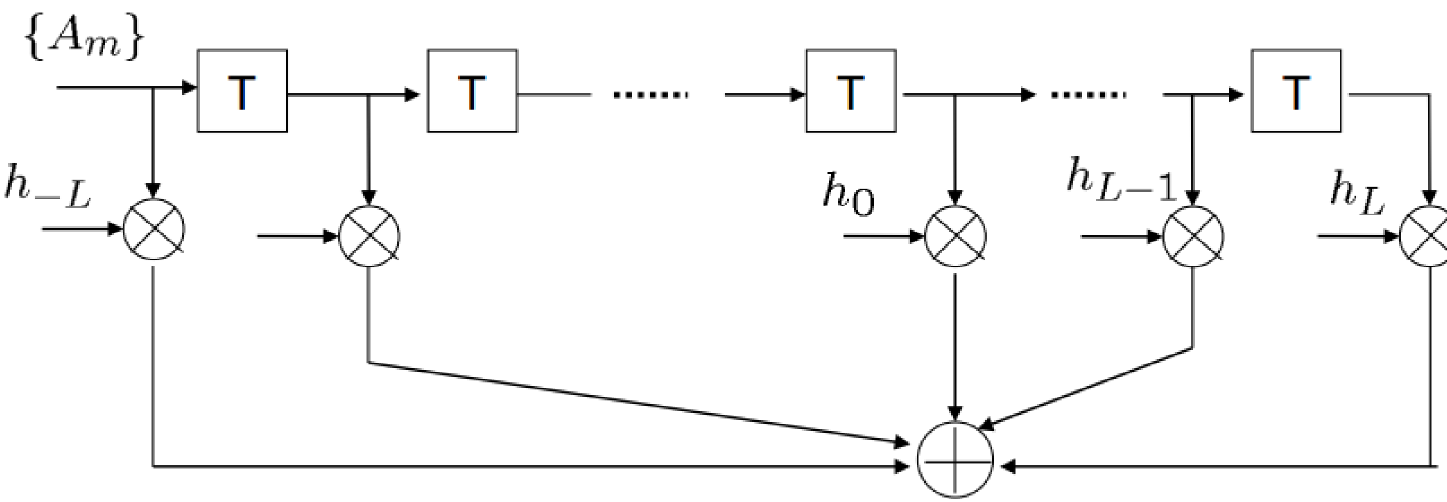
$$y_m = h_0 A_m + \sum_{\substack{n=-\infty \\ n \neq m}}^{\infty} h_{m-n} A_n + v_m$$

Digital transmission through bandlimited channels

- Equalizer
  - Maximum likelihood sequence estimation (MLSE).
  - Assume ISI affects finite number of symbols with

$$h_n = 0 \text{ for } |n| > L$$

  - Then, the channel is equivalent to a FIR discrete-time filter



Finite-state machine

$$\text{Output} = \sum_{k=-L}^{L} h_k A_{m-k}$$

Digital transmission through bandlimited channels

- Equalizer
  - ➢ Performance of MLSE

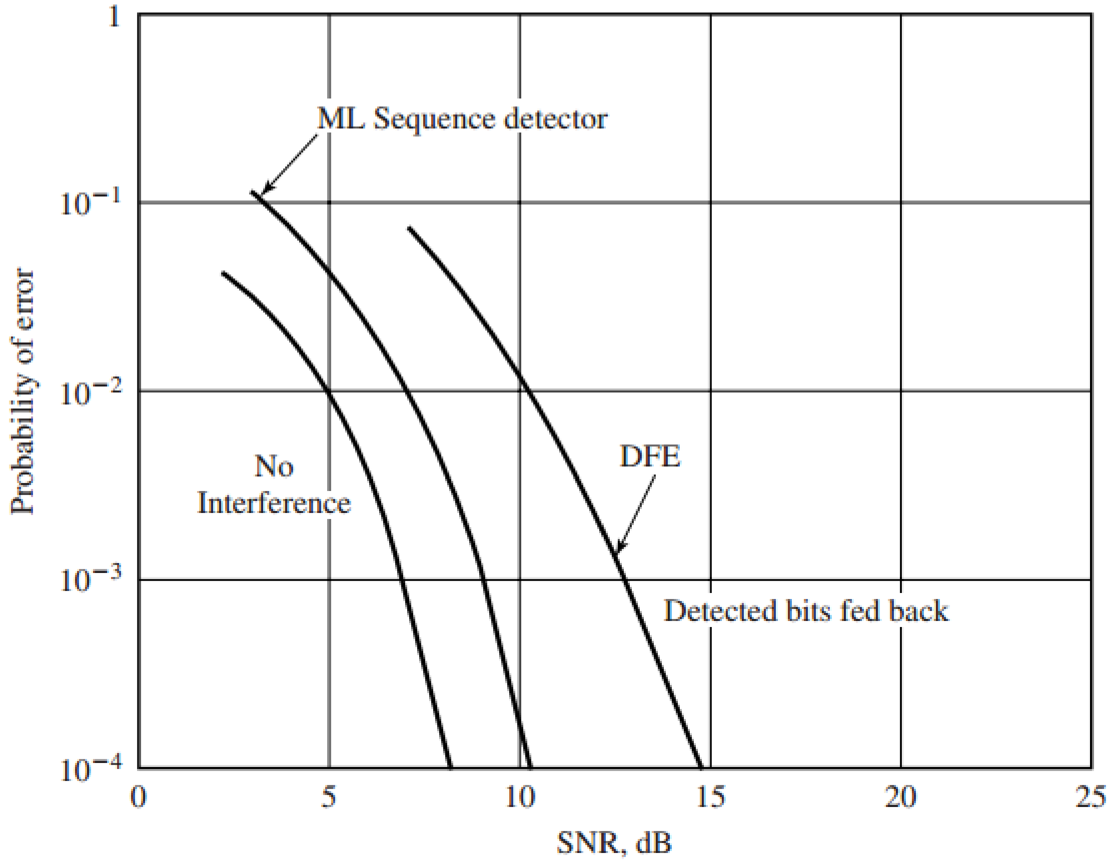
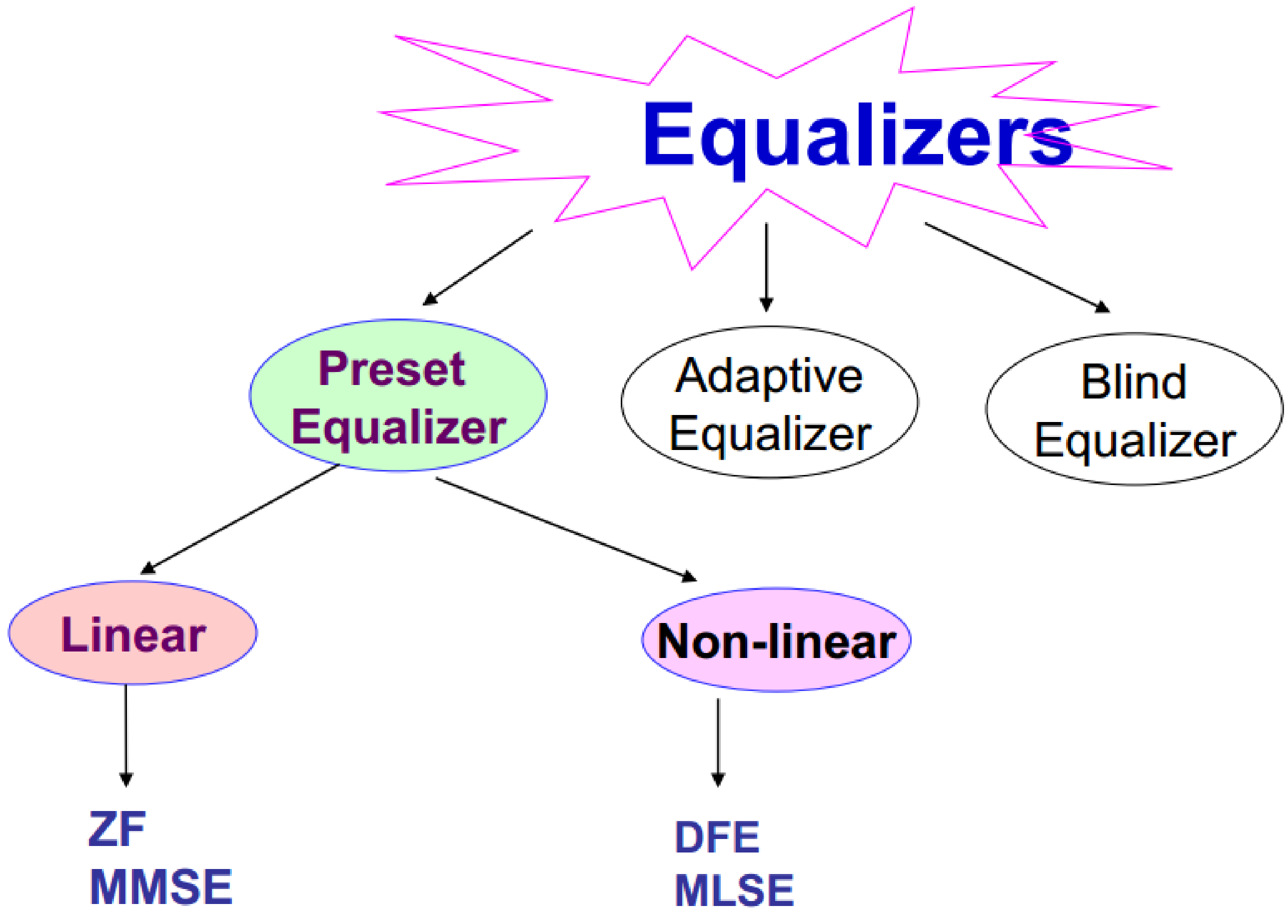


Figure 8.48 Performance of Viterbi detector and DFE for channel B.

Digital transmission through bandlimited channels

- Equalizer



Outline

- Introduction
- Digital transmission through baseband channels
- **Signal space representation**
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

Signal space representation

- Concepts
  - The key to analyzing and understanding the performance of digital transmission is the realization that **signals used in communications can be expressed and visualized graphically (constellation)**
  - Thus, we need to understand signal space concepts applied to digital communications

Chapter 8.1

Signal space representation

- Traditional bandpass signal representation
  - Baseband signals are the message signal generated at the source
  - Passband (Bandpass) signals refer to the signals after modulating with a carrier. The bandwidth of these signals are usually small compared with the carrier frequency f_c
  - Passband signals can be represented in three forms

1. Magnitude and phase representation
2. Quadrature representation
3. Complex envelope representation

Signal space representation

- Magnitude and phase representation

$$s(t) = a(t)\cos[2\pi f_c t + \theta(t)]$$

where $a(t)$ is the magnitude of $s(t)$

$\theta(t)$ is the phase of s(t)

Signal space representation

- Quadrature (I/Q) representation

$$s(t) = x(t)\cos(2\pi f_c t) - y(t)\sin(2\pi f_c t)$$

where x(t) and y(t) are real-valued baseband signals called the **in-phase** and **quadrature** components of s(t)

Signal space is a more convenient way than I/Q representation to study modulation scheme in digital communications

Signal space representation

- Vectors and space
  - Consider an n-dimensional space with unity basis vectors

$$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$$

  - Any vector **a** in the space can be written as

$$\mathbf{a} = \sum_{i=1}^{n} a_i \mathbf{e}_i \Rightarrow \mathbf{a} = (a_1, a_2, \ldots, a_n)$$

$n \triangleq$ Dimension = Minimum number of vectors that is necessary and sufficient for representation of any vector in space

Signal space representation

- Vectors and space
  - Inner product $$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i$$
  - **a** and **b** are orthogonal if $\mathbf{a} \cdot \mathbf{b} = 0$
  - Norm $$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} = \sqrt{\sum_{i=1}^{n} a_i^2}$$
  - A set of vectors are orthonormal if they are mutually orthogonal and all have unit norm

Signal space representation

- Vectors and space
  - The set of basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of a space are chosen such that

1. Should be complete or span the vector space, i.e., any vector a can be expressed as a linear combination of these vectors
2. Should be orthonormal vectors

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0, \ \forall i \neq j \qquad \|e_i\| = 1, \ \forall i$$

  - A set of basis vectors satisfying these properties is also said to be a complete orthonormal basis

Signal space representation

- Signal space
  - Basic idea: If a signal can be represented by n-tuple, then it can be treated in much the same way as a n-dim vector.
  - Let $\phi_1(t), \phi_2(t), ..., \phi_n(t)$ be n signals
  - Consider a signal x(t) and suppose that

$$x(t) = \sum_{i=1}^{n} x_i \phi_i(t)$$

  - If every signal can be written as above, then

$\{\phi_1(t), \ldots, \phi_n(t)\}$ ~ **basis functions (基函数)**

Signal space representation

- Orthonormal basis
  - Signal set $\{\phi_k(t)\}$ is an orthogonal set if

$$\int_{-\infty}^{\infty} \phi_j(t)\phi_k(t)dt = \begin{cases} 0 & j \neq k \\ c_j & j = k \end{cases}$$

  - If $c_j \equiv 1 \forall j \to \{\phi_k(t)\}$ is an orthonormal set.
  - In this case,

$$x_k = \int_{-\infty}^{\infty} x(t)\phi_k(t)dt$$

$$x(t) = \sum_{i=1}^{n} x_i \phi_i(t)$$

$$\mathbf{x} = (x_1, x_2, ..., x_n)$$

Signal space representation

- Orthonormal basis
  - Given the set of the orthonormal basis

$$\{\phi_1(t), \ldots, \phi_n(t)\}$$

  - Let x(t) and y(t) be represented as

$$x(t) = \sum_{i=1}^{n} x_i \phi_i(t) \;, \qquad y(t) = \sum_{i=1}^{n} y_i \phi_i(t)$$

with $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$

  - Then the inner product of **x** and **y** is

$$\boxed{\mathbf{x} \cdot \mathbf{y} = \int_{-\infty}^{\infty} x(t) y(t) dt}$$

Signal space representation

- Orthonormal basis
  - Proof

$$\int_{-\infty}^{\infty} x(t)y(t)dt = \int_{-\infty}^{\infty} \left[\sum_{i=1}^{n} x_i\phi_i(t)\right]\left[\sum_{j=1}^{n} y_j\phi_j(t)\right] dt$$

$$= \sum_{k=1}^{n} x_k y_k \triangleq \mathbf{x}\cdot\mathbf{y}$$

Since $$\int_{-\infty}^{\infty} \phi_i(t)\phi_j(t)dt = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

E_x = Energy of $x(t)$ = $\int_{-\infty}^{\infty} x^2(t)dt$

$$\Rightarrow \quad \boxed{E_x = \mathbf{x}\cdot\mathbf{x} = \|\mathbf{x}\|^2}$$

Signal space representation

- Basis functions for a signal set
  - Consider a set of M signals (M-ary symbol) $\{s_i(t), i = 1, 2, ..., M\}$ with finite energy, i.e.,

$$\int_{-\infty}^{\infty} s_i^2(t)dt < \infty$$

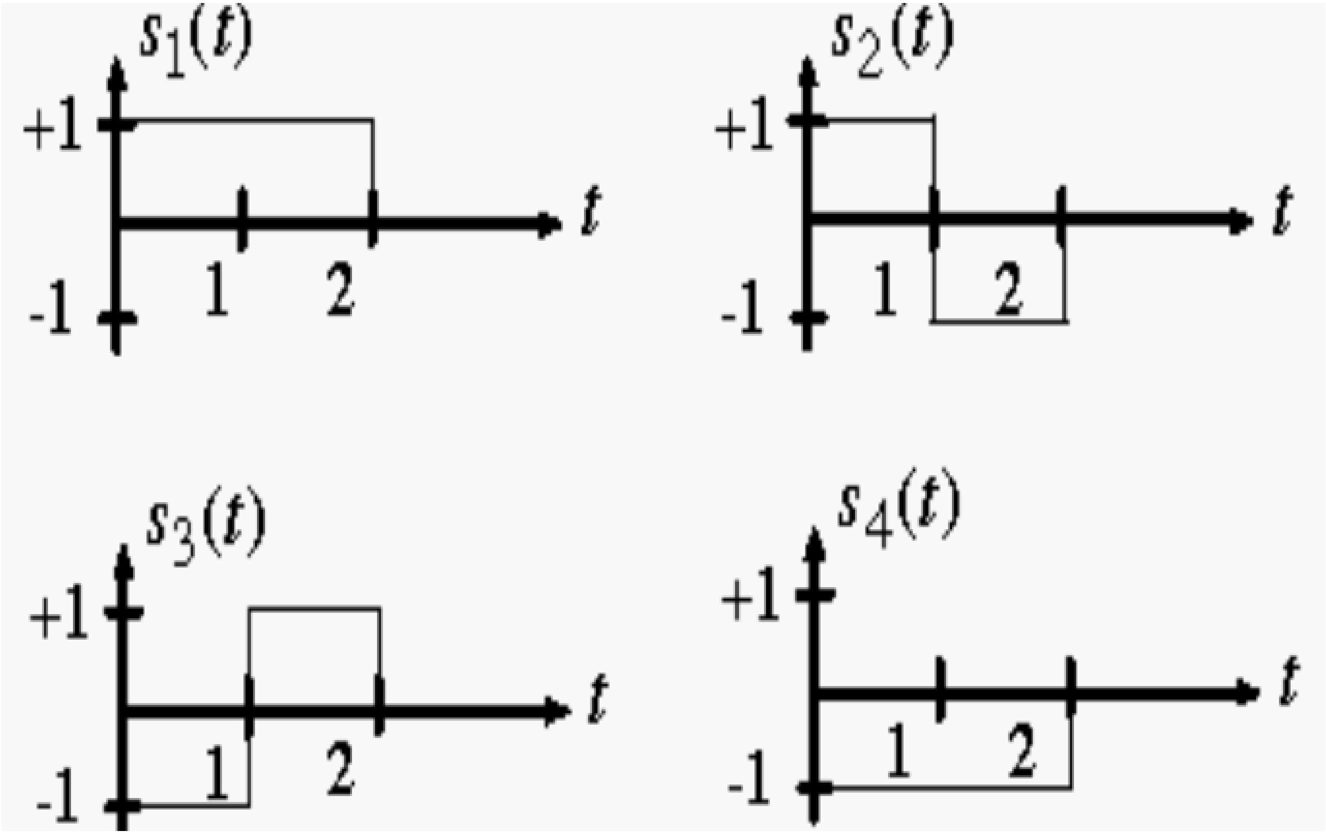
  - Then, we can express each of these waveforms as weighted linear combination of orthonormal signals:

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t) \qquad \text{for } i = 1, \ldots, M$$

where $N \leq M$ is the dimension of the signal space and $\{\phi_j(t)\}_1^N$ are called the orthonormal basis functions

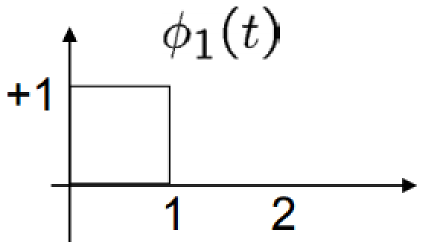
Signal space representation

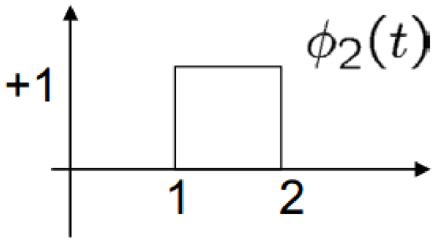
- Basis functions for a signal set
  - Consider the following signal set



Signal space representation

- Basis functions for a signal set
  - By inspection, the signals can be expressed in terms of the following two basis functions:





$$s_1(t) = 1 \cdot \phi_1(t) + 1 \cdot \phi_2(t) \qquad s_3(t) = -1 \cdot \phi_1(t) + 1 \cdot \phi_2(t)$$

$$s_2(t) = 1 \cdot \phi_1(t) - 1 \cdot \phi_2(t) \qquad s_4(t) = -1 \cdot \phi_1(t) - 1 \cdot \phi_2(t)$$

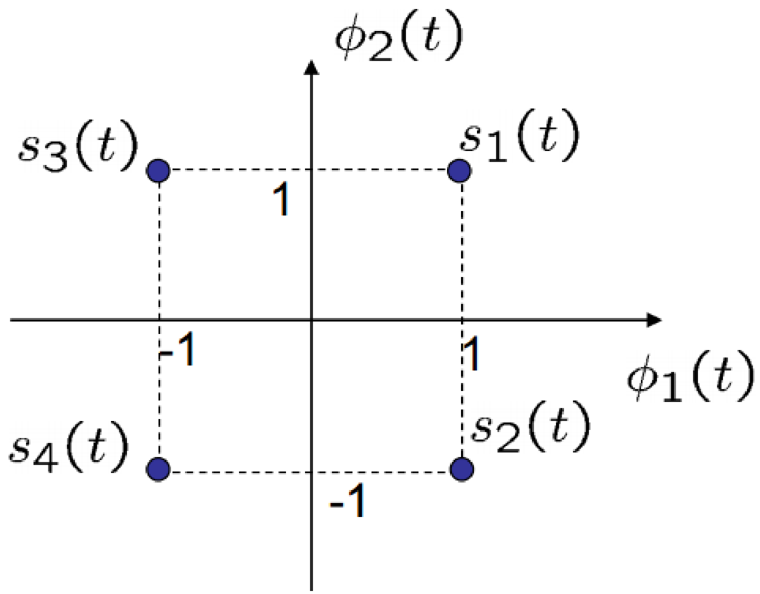
  - Note that the basis is orthonormal

$$\int_{-\infty}^{\infty} \phi_1(t)\phi_2(t)dt = 0$$

$$\int_{-\infty}^{\infty} |\phi_1(t)|^2 dt = \int_{-\infty}^{\infty} |\phi_2(t)|^2 dt = 1$$

Signal space representation

- Basis functions for a signal set
  - Constellation diagram is a representation of a **digital modulation** scheme in the **signal space**
  - The axes are labeled with $\phi_1(t)$ and $\phi_2(t)$
  - Possible signals are plotted as points, called **constellation points**



Signal space representation

- Basis functions for a signal set
  - Suppose our signal set can be represented in the following form

$$s(t) = \pm\sqrt{\frac{2}{T}}\cos(2\pi f_c t) \pm \sqrt{\frac{2}{T}}\sin(2\pi f_c t)$$

with $t \in [0, T)$ and $f_c T >> 1$

  - We can choose the basis functions as follows

$$\phi_1(t) = \sqrt{\frac{2}{T}}\cos(2\pi f_c t) \qquad \phi_2(t) = \sqrt{\frac{2}{T}}\sin(2\pi f_c t)$$

$t \in [0, T)$

Since

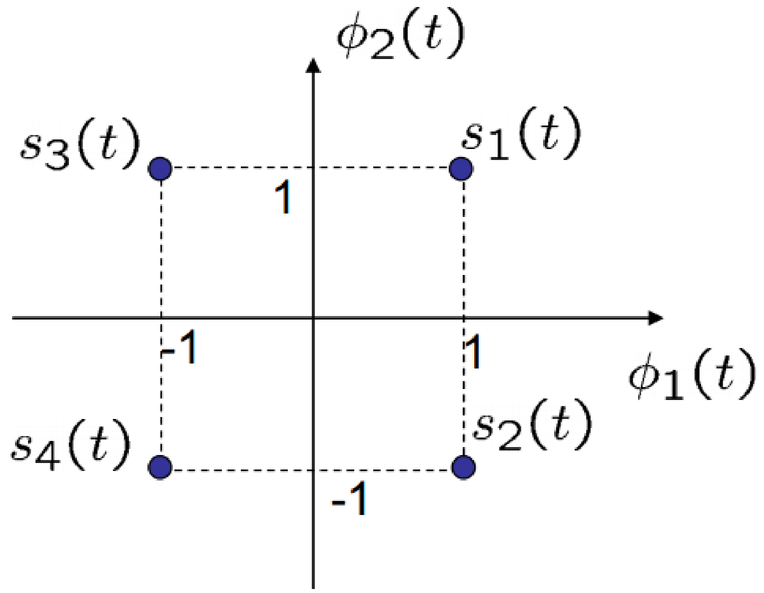
$$\begin{aligned}\int_0^T \phi_1(t)\phi_2(t)dt &= \int_0^T \sqrt{\tfrac{2}{T}}\cos(2\pi f_c t)\sqrt{\tfrac{2}{T}}\sin(2\pi f_c t)dt \\ &= \tfrac{2}{T}\int_0^T \tfrac{1}{2}[\sin(0) + \sin(4\pi f_c t)]dt \\ &= -\frac{1}{4\pi f_c T}[\cos(4\pi f_c t)]_0^T \approx 0, \text{for } f_c T \ll 1\end{aligned}$$

and

$$\int_0^T |\phi_1(t)|^2 dt = \int_0^T |\phi_2(t)|^2 dt = \tfrac{2}{T}\int_0^T \tfrac{1}{2}[1 + \cos(4\pi f_c t)]dt \approx 1$$

Signal space representation

- Basis functions for a signal set
  - The above example is exactly **QPSK** modulation. Its constellation diagram is also



Widely used in commercial communication systems, including 4G-LTE, Wifi, and incoming 5G.

Signal space representation

- Basis functions for a signal set
  - Two entirely different signal sets can have the same **geometric representation**.
  - The underlying geometry will determine the **performance** and the **receiver structure**
  - In general, is there any method to find a complete orthonormal basis for an arbitrary signal set?

Gram-Schmidt Orthogonalization (GSO) Procedure

Signal space representation

- GSO procedure
  - Suppose we are given a signal set $\{s_1(t), \ldots, s_M(t)\}$
  - Find the orthogonal basis functions for this signal set

    $\{\phi_1(t), \ldots, \phi_K(t)\}$ with $K \leq M$
  - **Step 1**: Compute the energy in signal 1

$$E_1 = \int_{-\infty}^{\infty} s_1^2(t)dt$$

  - The first basis function is just a normalized version of $s_1(t)$

$$\boxed{\phi_1(t) = \frac{1}{\sqrt{E_1}} s_1(t)}$$

$$\Rightarrow \quad s_1(t) = s_{11}\phi_1(t) = \sqrt{E_1}\phi_1(t)$$

$$s_{11} = \int_{-\infty}^{\infty} s_1(t)\phi_1(t)dt = \sqrt{E_1}$$

Signal space representation

- GSO procedure
  - **Step 2**: Compute the correlation between signal 2 and basic function 1

$$s_{21} = \int_{-\infty}^{\infty} s_2(t)\phi_1(t)dt$$

  - Subtract off the correlation portion

$$g_2(t) = s_2(t) - s_{21}\phi_1(t) \Longrightarrow g_2(t) \text{ is orthogonal to } \phi_1(t)$$

  - Compute the energy in the remaining portion

$$E_{g_2} = \int_{-\infty}^{\infty} [g_2(t)]^2 \, dt$$

  - Normalize the remaining portion

$$\boxed{\phi_2(t) = \frac{1}{\sqrt{E_{g_2}}} g_2(t)}$$

$$\Longrightarrow s_{22} = \int_{-\infty}^{\infty} s_2(t)\phi_2(t)dt = \sqrt{E_{g_2}}$$

Signal space representation

- GSO procedure
  - **Step 3**: For signal $s_k(t)$, compute
    $$s_{ki} = \int_{-\infty}^{\infty} s_k(t)\phi_i(t)dt$$
  - Define
    $$g_k(t) = s_k(t) - \sum_{i=1}^{k-1} s_{ki}\phi_i(t)$$
  - Compute the energy of $g_k(t)$
    $$E_{g_k} = \int_{-\infty}^{\infty} [g_k(t)]^2\, dt$$
  - k-th basis function
    $$\boxed{\phi_k(t) = \frac{1}{\sqrt{E_{g_k}}} g_k(t)}$$

$$\Rightarrow \quad s_{kk} = \int_{-\infty}^{\infty} s_k(t)\phi_k(t)dt = \sqrt{E_{g_k}}$$

Signal space representation

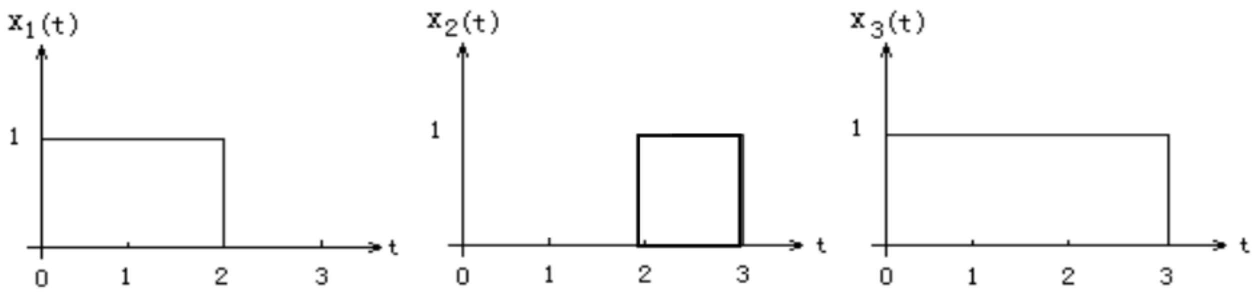
- GSO procedure
  - **Summary**

1. 1^{st} basis function is normalized version of the first signal
2. Successive basis functions are found by removing portions of signals that are correlated to previous basis functions and normalizing the result
3. The procedure is repeated until all basis functions are found (if $g_k(t)=0$, no new basis functions is added)

The order in which signals are considered is arbitrary

Signal space representation

- GSO procedure
  - Consider an example.
  - Use the Gram-Schmidt procedure to find a set of orthonormal basis functions corresponding to the signals shown below



  - Express x_1, x_2, x_3, in terms of the orthonormal basis functions found previously
  - Draw the constellation diagram for this signal set

Signal space representation

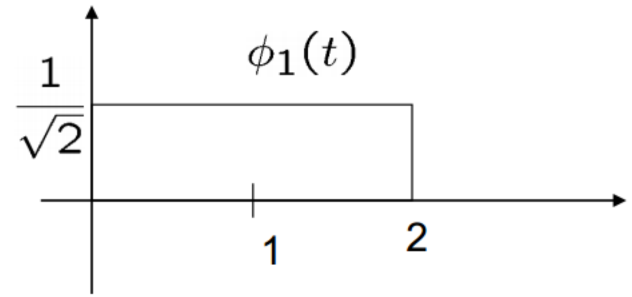
- GSO procedure
  - Example

Step 1:

$$E_1 = \int_{-\infty}^{\infty} x_1^2(t)dt = 2$$

$$\phi_1(t) = \frac{1}{\sqrt{2}} x_1(t)$$

$$x_{11} = \sqrt{2}$$



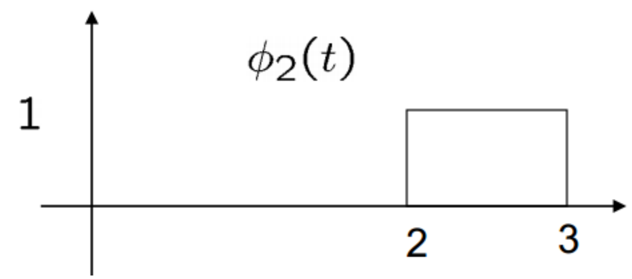
Step 2:

$$x_{21} = \int_{-\infty}^{\infty} x_2(t)\phi_1(t)dt = 0$$

$g_2(t) = x_2(t)$ and $E_{g_2} = 1$

$$\phi_2(t) = x_2(t)$$

$$x_{22} = 1$$



Signal space representation

- GSO procedure
  - Example

Step 3: $$x_{31} = \int_{-\infty}^{\infty} x_3(t)\phi_1(t)dt = \sqrt{2}$$

$$x_{32} = \int_{-\infty}^{\infty} x_3(t)\phi_2(t)dt = 1$$

$$g_3(t) = x_3(t) - x_{31}f_1(t) - x_{32}f_2(t) = 0$$

=> No more new basis functions

Procedure completes

$$\begin{cases} \phi_1(t) = \dfrac{1}{\sqrt{2}}x_1(t) \\ \\ \phi_2(t) = x_2(t) \end{cases}$$

Signal space representation

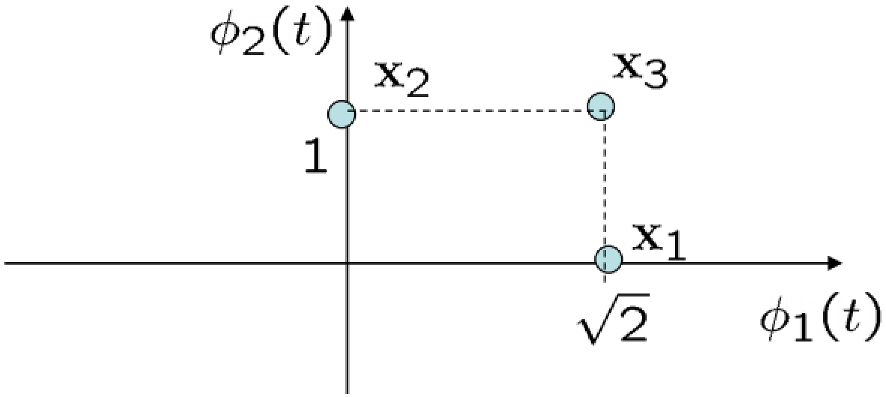
- GSO procedure
  - Example

Express x_1, x_2, x_3 in basis functions

$$x_1(t) = \sqrt{2}\phi_1(t), \quad x_2(t) = \phi_2(t)$$

$$x_3(t) = \sqrt{2}\phi_1(t) + \phi_2(t)$$

Constellation diagram



Signal space representation

- GSO procedure
  - Exercise

Given a set of signals (8PSK modulation)

$$s_i(t) = A\cos\left(2\pi f_c t + \frac{\pi}{4} i\right)$$

$$i = 0, 1, \ldots, 7 \quad \text{and } 0 \leq t < T$$

- Find the orthonormal basis functions using Gram Schmidt procedure
- What is the dimension of the resulting signal space ?
- Draw the constellation diagram of this signal set

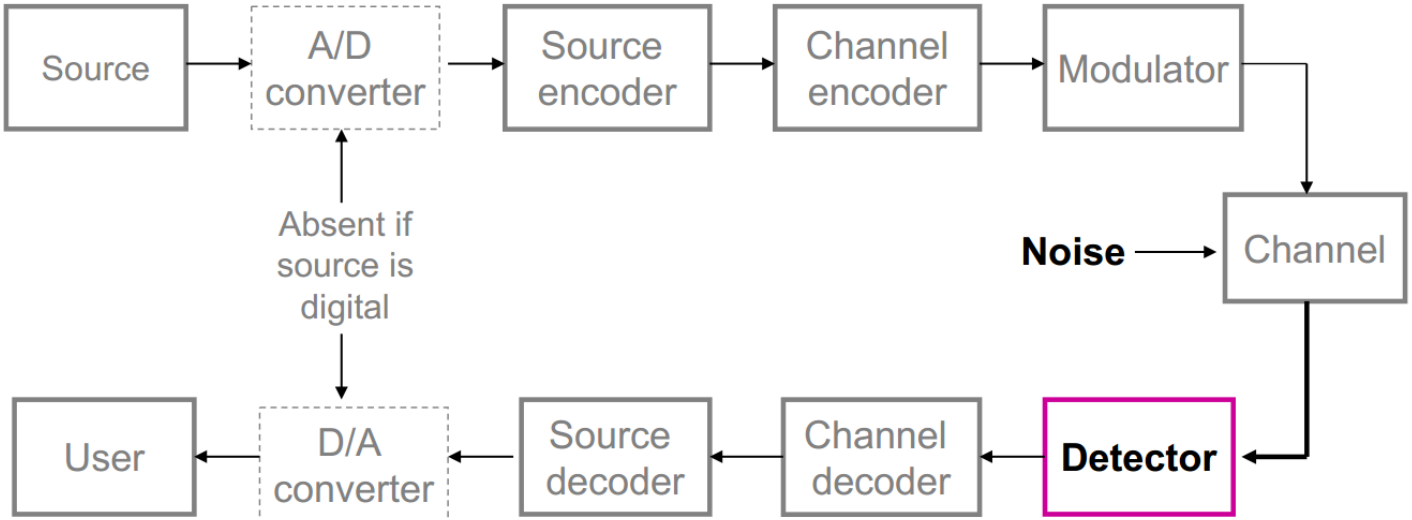
Signal space representation

- GSO procedure
  - A signal set may have many different sets of basis functions
  - A change of basis functions is essentially a rotation of the signal points around the origin
  - The order in which signals are used in GSO procedure affects the resulting basis functions
  - The choice of basis functions does not affect the performance of the modulation scheme

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- **Optimal receivers**
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

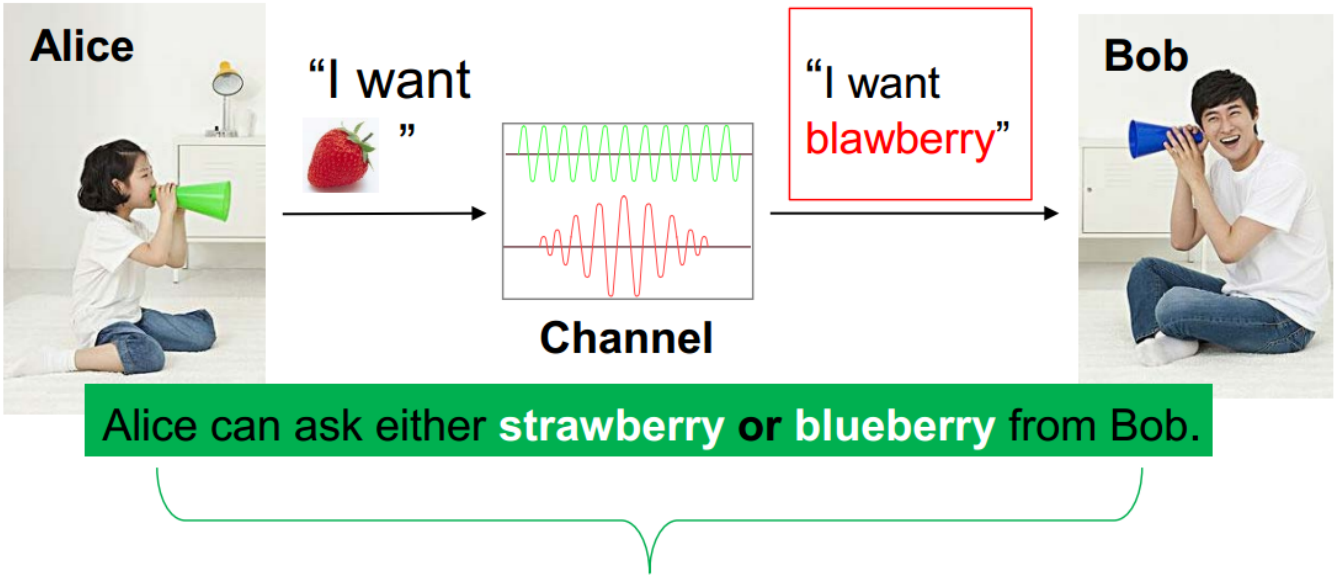
Optimal receivers



- Detection theory
- Optimal receiver structure
- Matched filter
- Decision regions
- Error probability analysis

Chapter 8.2-8.4, 8.5.3

Optimal receivers



- In digital communications, hypotheses are the possible messages and observations are the output of a channel
- Based on the observed values of the channel output, we are interested in the best decision making rule in the sense of minimizing the probability of error

Detection theory

- Given M possible hypotheses H_i (signal m_i) with probability

$$P_i = P(m_i) \quad , \quad i = 1, 2, \ldots, M$$

where P_i represents the priori knowledge concerning the probability of the signal m_i (priori probability)

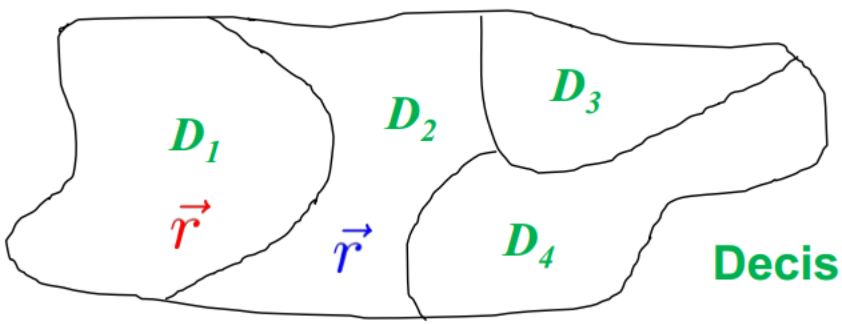
- The observation is some collection of N real values denoted by $\vec{r} = (r_1, r_2, \ldots, r_N)$ with conditional pdf

$f(\vec{r}|m_i)$ -- conditional pdf of observation $\vec{r}$ given the signal m_i

- Our goal is to find the best decision-making rule in the sense of minimizing the probability of error

Detection theory

- In general, $\vec{r}$ can be regarded as a point in some observation space
- Each hypothesis H_i is associated with a decision region D_i:
  If $\vec{r}$ falls into D_i, the decision is H_i
- Error occurs when a decision is in favor of another when the signal $\vec{r}$ falls outside the decision region D_i



Decision Space M=4

Detection theory

- Consider a decision rule based on the computation of the posterior probabilities defined as

$$P(m_i|\vec{r}) = P(\text{ signal } m_i \text{ was transmitted given } \vec{r} \text{ observed })$$
$$\text{for } i = 1, \ldots, M$$

- A posterior since the decision is made after (or given) the observation
- Different from the a priori where some information about the decision known before the observation
- By Bayes' Rule: $P(m_i|\vec{r}) = \dfrac{P_i f(\vec{r}|m_i)}{f(\vec{r})}$
- Minimizing the probability of detection error given $\vec{r}$ is equivalent to maximize the probability of correct detection
- Maximum a posterior (MAP) decision rule:

$$\boxed{\begin{array}{l}\text{Choose } \hat{m} = m_k \text{ if and only if} \\ P_k f(\vec{r}|m_k) \geq P_i f(\vec{r}|m_i); \text{ for all } i \neq k\end{array}}$$

Detection theory

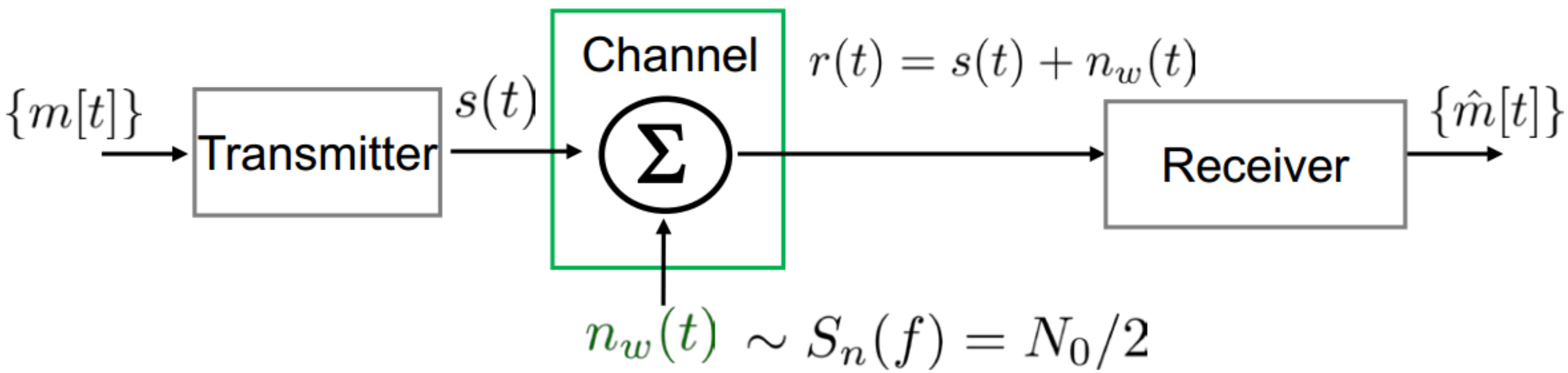
- If $p_1 = p_2 = \ldots = p_M$, the signals are equiprobable, finding the signal that maximizes $P(m_k|\vec{r})$ is equivalent to finding the signal that maximizes $f(\vec{r}|m_k)$
- The conditional pdf $f(\vec{r}|m_k)$ is usually called the likelihood function. The decision criterion based on the maixmum of $f(\vec{r}|m_k)$ is called the maximum likelihood (ML) dectection
- ML decision rule:

Choose $\hat{m} = m_k$ if and only if

$$f(\vec{r}|m_k) \geq f(\vec{r}|m_i); \text{ for all } i \neq k$$

Optimal receiver structure

- Signal model
  - Transmitter transmits a sequence of symbols or messages from a set of M symbols $m_1, m_2, \ldots, m_M$ with priori probabilities
    $$p_1 = P(m_1),\ p_2 = P(m_2),\ p_M = P(m_M)$$
  - The symbols are represented by finite energy waveforms $s_1(t), s_2(t), \ldots, s_M(t)$ defined in intervals [0,T]
  - The signal is assumed to be corrupted by additive



Optimal receiver structure

- Signal space representation
  - Signal space of $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ is assumed to be of dimension N ($N \le M$)
  - $\phi_k(t)$ for k=1,…,N will denote the orthonormal basis functions
  - Then each transmitted signal waveform can be represented as

$$s_m(t) = \sum_{k=1}^{N} s_{mk}\phi_k(t) \quad \text{where} \quad s_{mk} = \int_0^T s_m(t)\phi_k(t)dt$$

  - Note that the noise $n_w(t)$ can be written as

$$n_w(t) = \underbrace{n_0(t)}_{} + \underbrace{\sum_{k=1}^{N} n_k\phi_k(t)}_{\text{Projection of } n_w(t) \text{ on the N-dim space}} \quad \text{where} \quad n_k = \int_0^T n_w(t)\phi_k(t)dt$$

orthogonal to the space, falls outside the signal space spanned by $\{\phi_k(t), k = 1, \ldots N\}$

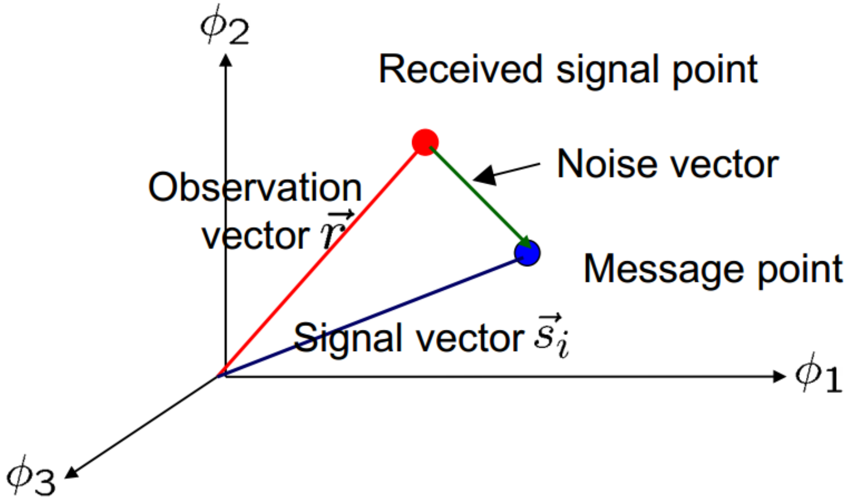
Optimal receiver structure

- Signal space representation
  - The received signal can thus be represented as

$$
\begin{aligned}
r(t) &= s(t) + n_w(t) \\
&= \sum_{k=1}^{N} s_{mk}\phi_k(t) + \sum_{k=1}^{N} n_k\phi_k(t) + n_0(t) \\
&= \underbrace{\sum_{k=1}^{N} r_k\phi_k(t)}_{\text{Projection of r(t) on N-dim signal space}} + n_0(t) \qquad \text{where} \quad r_k = s_{mk} + n_k
\end{aligned}
$$

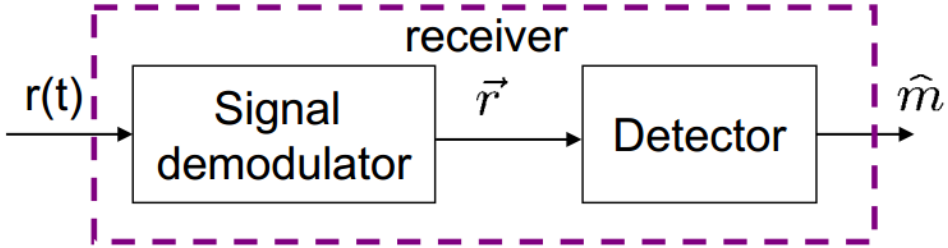
  - In vector form, we have

$$\vec{r} = \vec{s}_i + \vec{n}$$



Optimal receiver structure

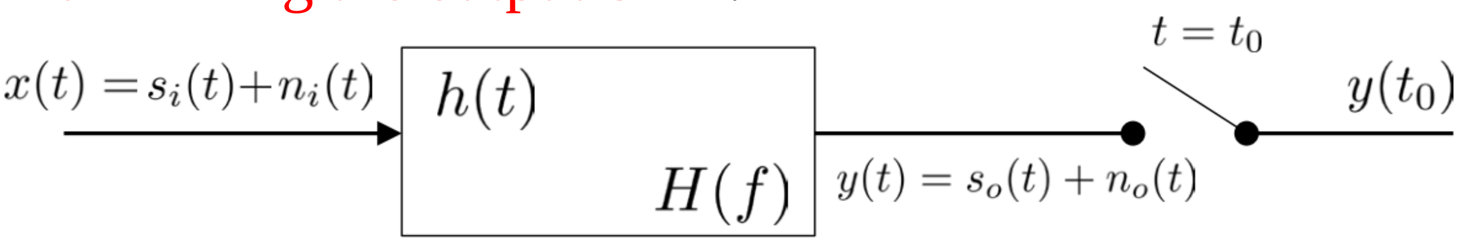
- Receiver structure
  - Signal demodulator: to convert the received wave form r(t) into an N-dim vector $\vec{r} = (r_1, r_2, \ldots, r_N)$
  - Detector: to decide which of the M possible signal waveforms was transmitted based on the observation vector $\vec{r}$



  - Two realizations of the signal demodulator: correlation type and matched-filter type

Matched filter

- Derivation
  - The matched-filter (MF) is the optimal linear filter for maximizing the output SNR.



  - Input signal component $s_i(t) \quad \leftrightarrow \quad A(f) = \int_{-\infty}^{\infty} s_i(t) e^{-j\omega t} dt$
  - Input noise component $n_i(t)$ with PSD $S_{n_i}(f) = N_0/2$
  - Output signal component

$$s_o(t) = \int_{-\infty}^{\infty} s_i(t-\tau) h(\tau) d\tau$$

$$= \int_{-\infty}^{\infty} A(f) H(f) e^{j\omega t} df$$

  - Sample at $t = t_0$

Matched filter

- Derivation
  - At the sampling instance $t = t_0$, $s_o(t_0) = \int_{-\infty}^{\infty} A(f)H(f)e^{j\omega t_0}df$
  - Average power of the output noise is

$$N = E\{n_o^2(t)\} = \frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2 df$$

  - Output SNR

$$\boxed{d = \frac{s_o^2(t_0)}{E\{n_o^2(t)\}} = \frac{\left[\int_{-\infty}^{\infty} A(f)H(f)e^{j\omega t_0}df\right]^2}{\frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2 df}}$$

Find $H(f)$ that can maximize d

Matched filter

- Derivation
  - Schwarz's inequality

$$\int_{-\infty}^{\infty}|F(x)|^2 dx\int_{-\infty}^{\infty}|Q(x)|^2 dx \ge \left|\int_{-\infty}^{\infty}F^*(x)Q(x)dx\right|^2$$

equality holds when $F(x)=CQ(x)$

  - Let $\begin{cases} F^*(x)=A(f)e^{j\omega t_0} \\ Q(f)=H(f) \end{cases}$, then

***E* : signal energy**

$$\boxed{d \le \frac{\int_{-\infty}^{\infty}|A(f)|^2 df\int_{-\infty}^{\infty}|H(f)|^2 df}{\frac{N_0}{2}\int_{-\infty}^{\infty}|H(f)|^2 df}=\frac{\int_{-\infty}^{\infty}|A(f)|^2 df}{\frac{N_0}{2}}=\frac{2E}{N_0}}$$

Matched filter

- Derivation
  - When the max output SNR 2E/N0 is achieved, we have

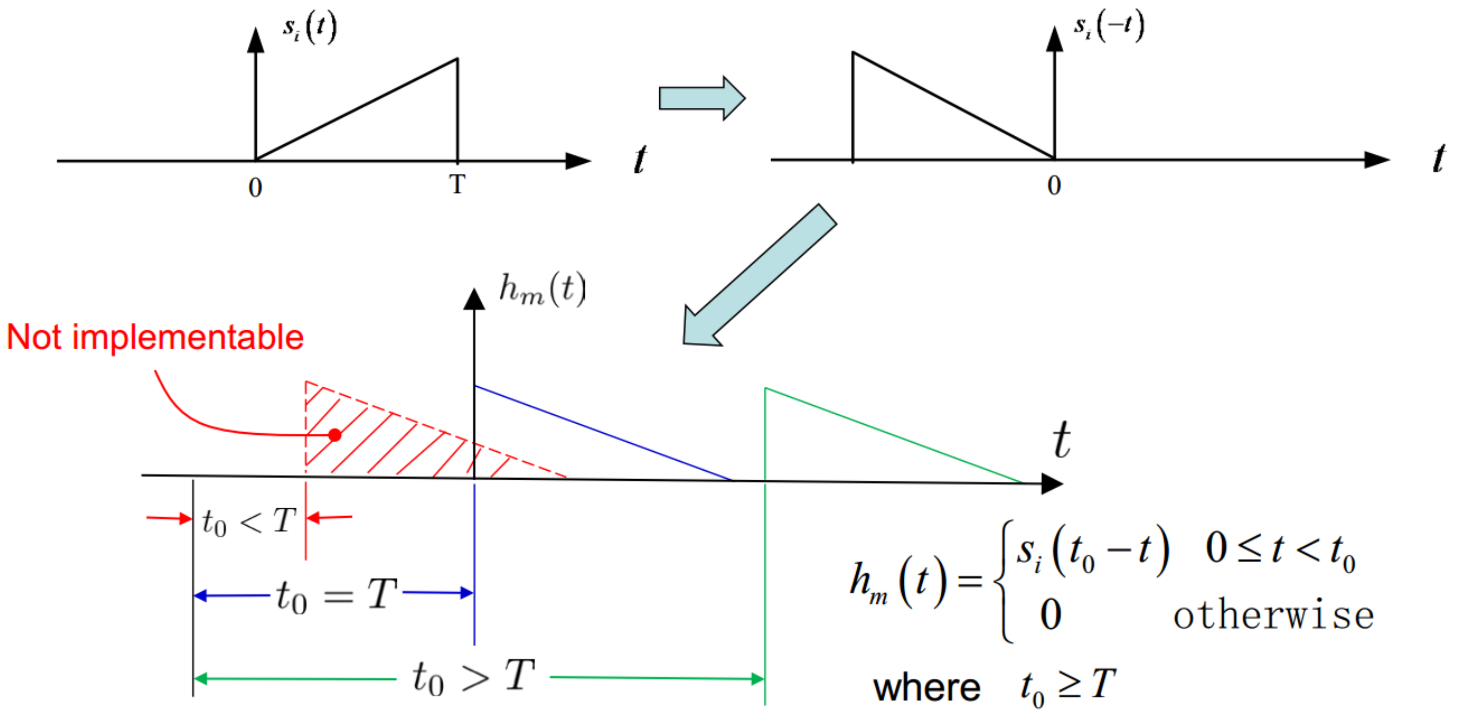
$$\boxed{\begin{aligned} H_m(f) &= A^*(f)\,e^{-j\omega t_0} \\ &\updownarrow \\ h_m(t) &= s_i^*(t_0 - t) \end{aligned}}$$

$$\begin{aligned} h_m(t) &= \int_{-\infty}^{\infty} H_m(f)\,e^{j\omega t}\,df \\ &= \int_{-\infty}^{\infty} A^*(f)\,e^{-j\omega(t_0 - t)}\,df \\ &= s_i^*(t_0 - t) \end{aligned}$$

  - Transfer function: complex conjugate of the input signal spectrum
  - Impulse response: time-reversal and delayed version of the input signal s(t)

Matched filter

- Properties
  - Choice of t_0 versus the causality



$$h_m(t) = \begin{cases} s_i(t_0 - t) & 0 \le t < t_0 \\ 0 & \text{otherwise} \end{cases}$$

where $t_0 \ge T$

Matched filter

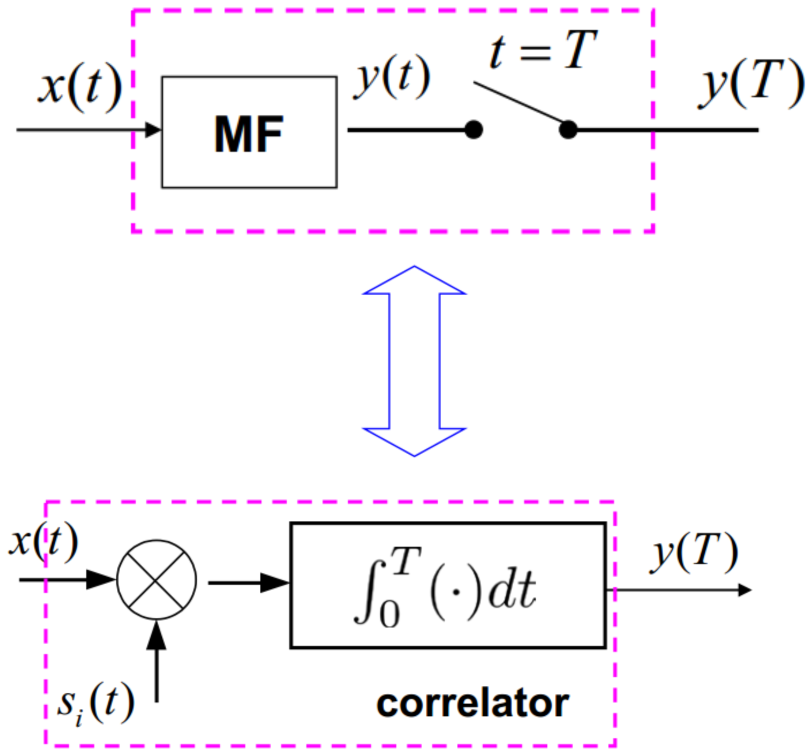
- Properties
  - Equivalent form in correlator
  - Let $s_i(t)$ be within [0,T]

$$y(t) = x(t) * h_m(t) = x(t) * s_i(T-t)$$

$$= \int_0^T x(\tau) s_i(T - t + \tau) d\tau$$

  - Observe at sampling time t=T

$$y(T) = \int_0^T x(\tau) s_i(\tau) d\tau = \int_0^T x(t) s_i(t) dt$$



Correlation integration
(相关积分）

Matched filter

- Properties
  - Correlation function

$$R_{12}(\tau)=\int_{-\infty}^{\infty} s_1(t)s_2(t+\tau)dt=\int_{-\infty}^{\infty} s_1(t-\tau)s_2(t)dt=R_{21}(-\tau)$$

  - Auto-correlation function

$$R(\tau)=\int_{-\infty}^{\infty} s(t)s(t+\tau)dt$$

1. $R(\tau)=R(-\tau)$
2. $R(0)\geq R(\tau)$
3. $R(0)=\int_{-\infty}^{\infty} s^2(t)dt=E$
4. $R(\tau)\leftrightarrow|A(f)|^2 \quad R(0)=\int_{-\infty}^{\infty} s^2(t)dt=\int_{-\infty}^{\infty}|A(f)|^2 df$

Matched filter

- Properties
  - MF output is the auto-correlation function of input signal
    $$s_o(t)=\int_{-\infty}^{\infty} s_i(t-u)h_m(u)du=\int_{-\infty}^{\infty} s_i(t-u)s_i(t_0-u)du$$
    $$=\int_{-\infty}^{\infty} s_i(\mu)s_i[\mu+t-t_0]d\mu=R_{s_0}(t-t_0)$$
  - The peak value of $s_0(t)$ happens
    $$s_o(t_0)=\int_{-\infty}^{\infty} s_i^2(\mu)du=E$$
  - $s_0(t)$ is symmetric at $t=t_0$
    $$A_o(f)=A(f)H_m(f)=|A(f)|^2 e^{-j\omega t_0}$$

Matched filter

- Properties
  - MF output noise
  - The statistical auto-correlation of $n_0(t)$ depends on the auto-correlation of $s_i(t)$

$$R_{n_o}(\tau) = E\{n_o(t)n_o(t+\tau)\} = \frac{N_0}{2}\int_{-\infty}^{\infty} h_m(u)h_m(u+\tau)du$$
$$= \frac{N_0}{2}\int_{-\infty}^{\infty} s_i(t)s_i(t-\tau)dt$$

  - Average power

$$E\{n_o^2(t)\} = R_{n_o}(0) = \frac{N_0}{2}\int_{-\infty}^{\infty} s_i^2(\mu)du \quad \textbf{Time domain}$$
$$= \frac{N_0}{2}\int_{-\infty}^{\infty}|A(f)|^2 df = \frac{N_0}{2}\int_{-\infty}^{\infty}|H_m(f)|^2 df \quad \textbf{Frequency domain}$$
$$= \frac{N_0}{2}E$$

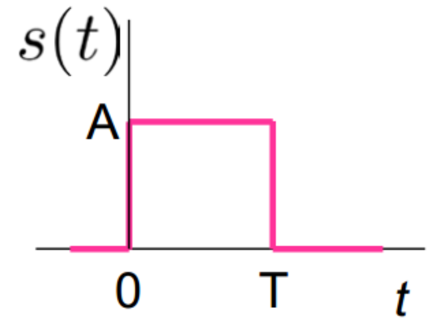
Matched filter

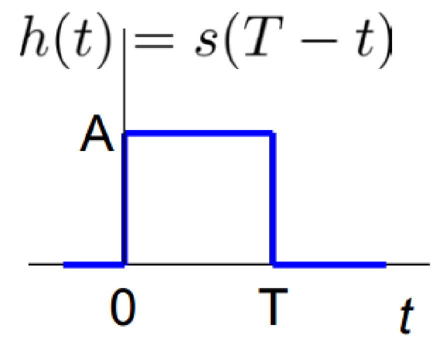
- Example
  - Consider a rectangular pulse s(t)

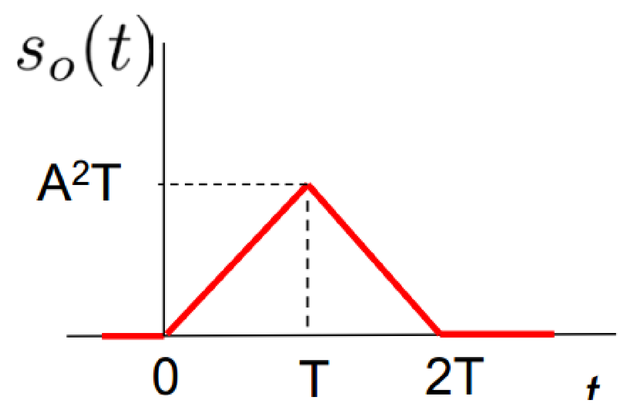
$$E_s = A^2 T$$

  - The impulse response of a filter matched to s(t) is also a rectangular pulse
  - The output of the matched filter $s_0(t)$ is $h(t) * s(t)$
  - The output SNR is

$$(SNR)_o = \frac{2}{N_0}\int_0^T s^2(t)dt = \frac{2A^2T}{N_0}$$



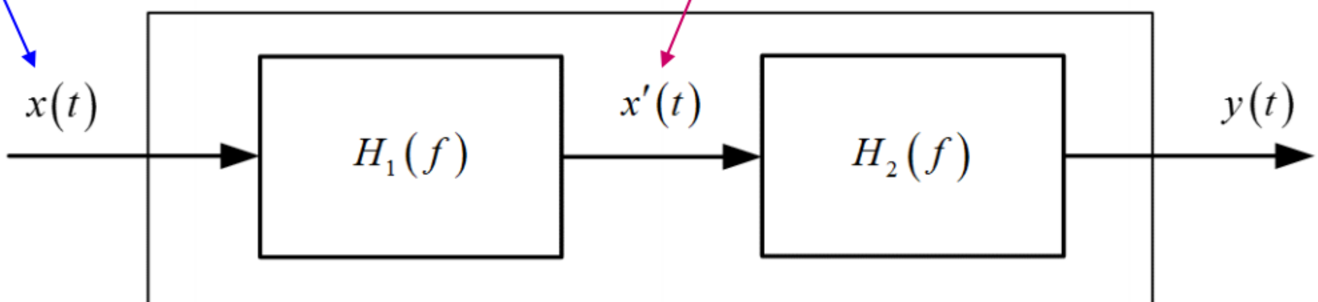




- Colored noise
  - In case of colored noise, we need to preprocess the combined signal and noise such that the non-white noise becomes white noise- Whitening Process

$x(t) = s_i(t) + n(t)$ where $n(t)$ is colored noise with PSD $S_n(f)$

$x'(t) = s'(t) + n'(t)$



Choose $H_1(f)$ so that $n'(t)$ is white, i.e.

$$S_n'(f) = \left|H_1(f)\right|^2 S_n(f) = C$$

Matched filter

- Colored noise
  - We choose

$$H_1(f):\ |H_1(f)|^2=\frac{C}{S_n(f)}$$

$H_2(f)$ **should match with** $S'(t)$ $\quad A'(f)=H_1(f)A(f)$

$$H_2(f)=A'^*(f)e^{-j2\pi f t_0}=H_1^*(f)A^*(f)e^{-j2\pi f t_0}$$

  - Therefore, the overall transfer function of the cascaded system:

$$H(f)=H_1(f)\cdot H_2(f)=H_1(f)H_1^*(f)A^*(f)e^{-j2\pi f t_0}$$

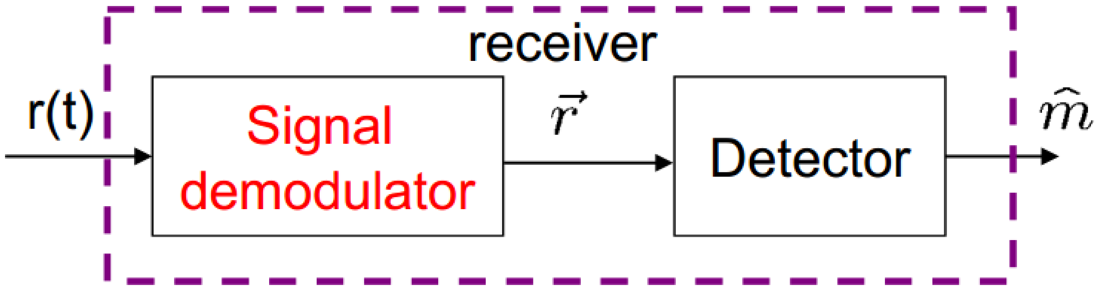
$$=|H_1(f)|^2 A^*(f)e^{-j2\pi f t_0}$$

$$\boxed{H(f)=\frac{A^*(f)}{S_n(f)}e^{-j2\pi f t_0}}$$

MF for colored noise

Updates on the receiver

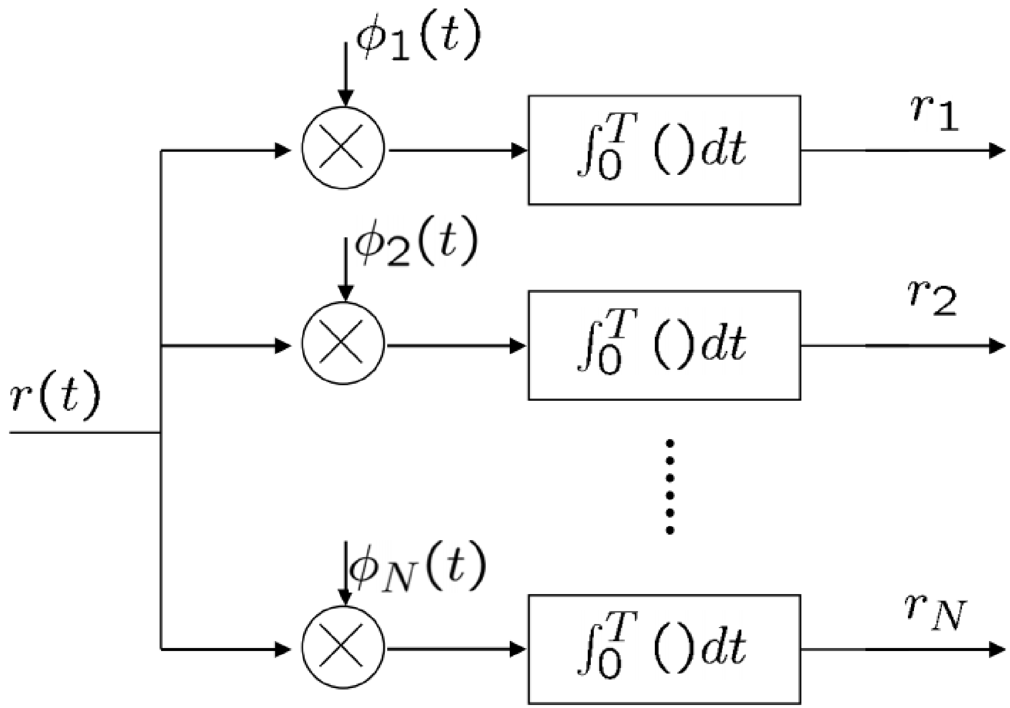
- We have talked about matched filter
- Consider the optimal receiver structure again



  - Two realizations of the signal demodulator: correlation type and matched-filter type

Updates on the receiver

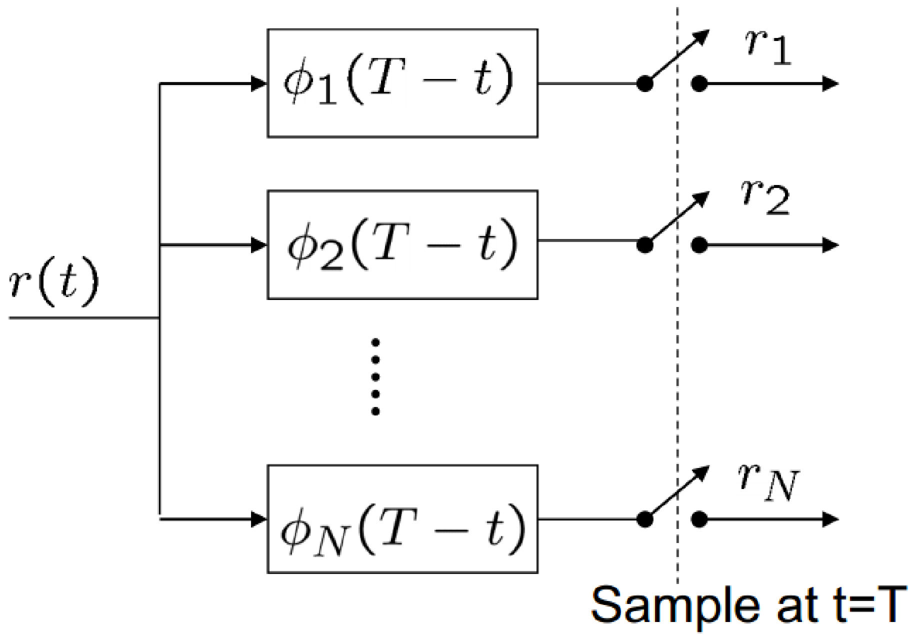
- Correlation type demodulator
  - The received signal r(t) is passed through a parallel bank of N cross correlators which basically compute the projection of r(t) onto the N basis functions $\{\phi_k(t), k = 1, \ldots N\}$



Updates on the receiver

- Matched filter type demodulator
  - Alternatively, we may apply the received signal r(t) to a bank of N matched filters and sample the output of filters at t=T. The impulse responses of the filters are

$$h_k(t) = \phi_k(T-t), \quad 0 \leq t \leq T$$



Updates on the receiver

- For a signal transmitted over an AWGN channel, either a correlation type demodulator or a matched filter type demodulator produces the vector $\vec{r} = (r_1, r_2, \ldots, r_N)$ which contains all the necessary information in r(t)
- The next step is to design a signal detector that makes a decision of the transmitted signal in each signal interval based on the observation of $\vec{r}$, such that the probability of error is minimized (or correct probability is maximized)
- Decision rules:

likelihood function $f(\vec{r}|m_k)$

MAP decision rule:

choose $\hat{m} = m_k$ if and only if

$$P_k f(\vec{r}|m_k) > P_i f(\vec{r}|m_i); \text{ for all } i \neq k$$

ML decision rule

choose $\hat{m} = m_k$ if and only if

$$f(\vec{r}|m_k) > f(\vec{r}|m_i); \text{ for all } i \neq k$$

Likelihood function

- Distribution of the noise vector
  - Since $n_w(t)$ is a Gaussian random process, the noise component of output $n_k = \int_0^T n_w(t)\phi_k(t)dt$ is Gaussian r.v.
  - Mean:

$$E[n_k] = \int_0^T E[n_w(t)]\phi_k(t)dt = 0 \quad , \quad k = 1, \ldots, N$$

  - Correlation between n_j and n_k

$$
\begin{aligned}
E[n_j n_k] &= E\left[\int_0^T n_w(t)\phi_j(t)dt \cdot \int_0^T n_w(\tau)\phi_k(\tau)d\tau\right] \\
&= E\left[\int_0^T \int_0^T n_w(t)n_w(\tau)\phi_j(t)\phi_k(\tau)dtd\tau\right] \\
&= \int_0^T \int_0^T E[n_w(t)n_w(\tau)]\phi_j(t)\phi_k(\tau)dtd\tau \\
&= \int_0^T \int_0^T \frac{N_0}{2}\delta(t-\tau)\phi_j(t)\phi_k(\tau)dtd\tau \\
&= \frac{N_0}{2}\int_0^T \phi_j(\tau)\phi_k(\tau)d\tau = \begin{cases} \frac{N_0}{2}, & j = k \\ 0, & j \neq k \end{cases}
\end{aligned}
$$

PSD of $n_w(t)$ is

$S_n(f) = N_0/2$

Likelihood function

- Distribution of the noise vector
  - Therefore, n_j and n_k ($j \neq k$) are uncorrelated Gaussian r.v.s, and hence independent with zero mean and variance $N_0/2$
  - The joint pdf of $\vec{n} = (n_1, \ldots, n_N)$

$$
\begin{aligned}
p(n_1, \ldots, n_N) &= \prod_{k=1}^{N} p(n_k) = \prod_{k=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left(-n_k^2/N_0\right) \\
&= (\pi N_0)^{-N/2} \exp\left(-\sum_{k=1}^{N} n_k^2/N_0\right)
\end{aligned}
$$

Likelihood function

- Conditional probability
  - If m_k is transmitted, $\vec{r} = \vec{s}_k + \vec{n}$ with $r_j = s_{kj} + n_j$

    and $E[r_j|m_k] = s_{kj} + E[n_j] = s_{kj}$

    Transmitted signal values in each dimension represent the mean values for each received signal

    and $Var[r_j|m_k] = Var[n_j] = N_0/2$
  - Therefore, the conditional pdf of

$$f(\vec{r}|m_k) = \prod_{j=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(r_j - s_{kj})^2}{N_0}\right)$$

$$= (\pi N_0)^{-N/2} \exp\left(-\frac{\sum_{j=1}^{N} (r_j - s_{kj})^2}{N_0}\right)$$

Likelihood function

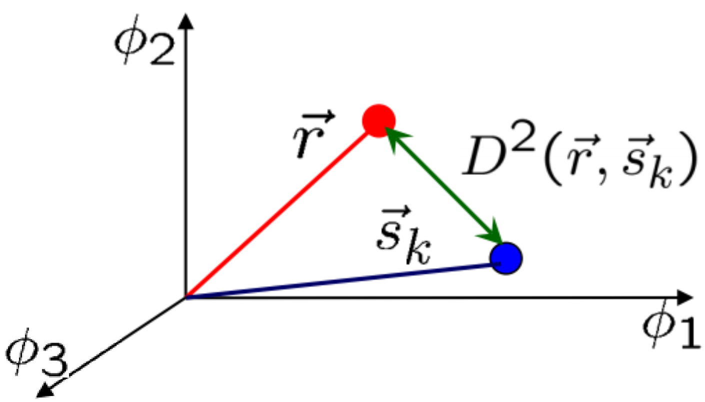
- Log-likelihood function
  - To simplify the computation, we take the natural logarithm of $f(\vec{r}|m_k)$, which is a monotonic function. Thus,

$$\ln f(\vec{r}|m_k) = -\frac{N}{2}\ln(\pi N_0) - \frac{1}{N_0}\sum_{j=1}^{N}(r_j - s_{kj})^2$$

  - Let

$$D^2(\vec{r},\vec{s}_k) = \sum_{j=1}^{N}(r_j - s_{k,j})^2 \quad = \|\vec{r} - \vec{s}_k\|^2$$

denote the Euclidean distance between $\vec{r}$ and $\vec{s}_k$ in the N-dim signal space. It is also called the distance metric.



Likelihood function

- Optimal detector
  - MAP rule:

$$\hat{m} = \arg\max_{\{m_1,\dots,m_M\}} f(\vec{r}|m_k)P(m_k)$$
$$= \arg\max_{\{m_1,\dots,m_M\}} \ln\left[f(\vec{r}|m_k)P(m_k)\right]$$
$$= \arg\max_{\{m_1,\dots,m_M\}} \left\{-\frac{1}{N_0}\|\vec{r}-\vec{s}_k\|^2 + \ln P_k\right\}$$
$$\boxed{= \arg\min_{\{m_1,\dots,m_M\}} \left\{\|\vec{r}-\vec{s}_k\|^2 - N_0 \ln P_k\right\}}$$

  - ML rule:

$$\boxed{\hat{m} = \arg\min_{\{m_1,\dots,m_M\}} \|\vec{r}-\vec{s}_k\|^2}$$

ML detector chooses $\hat{m} = m_k$ iff received vector $\vec{r}$ is closer to $\vec{s}_k$ in terms of Euclidean distance than to any other $\vec{s}_i$ for i ≠ k

Minimum distance detection

(will discuss more in decision region)

Likelihood function

- Optimal receiver structure
  - With the above expression, we can develop a receiver structure using the following derivation

$$-\sum_{j=1}^{N}(r_j - s_{kj})^2 + N_0 \ln P_k = -\sum_{j=1}^{N} r_j^2 - \sum_{j=1}^{N} s_{kj}^2 + 2\sum_{j=1}^{N} r_j s_{kj} + N_0 \ln P_k$$

$$= -\|\vec{r}\|^2 - \|\vec{s}_k\|^2 + 2\vec{r}\cdot\vec{s}_k + N_0 \ln P_k$$

with

$\|\vec{s}_k\|^2 = \int_0^T s_k^2(t)dt = E_k$ = signal energy

$\vec{r}\cdot\vec{s}_k = \int_0^T s_k(t)r(t)dt$ = correlation between the received signal vector and the transmitted signal vector

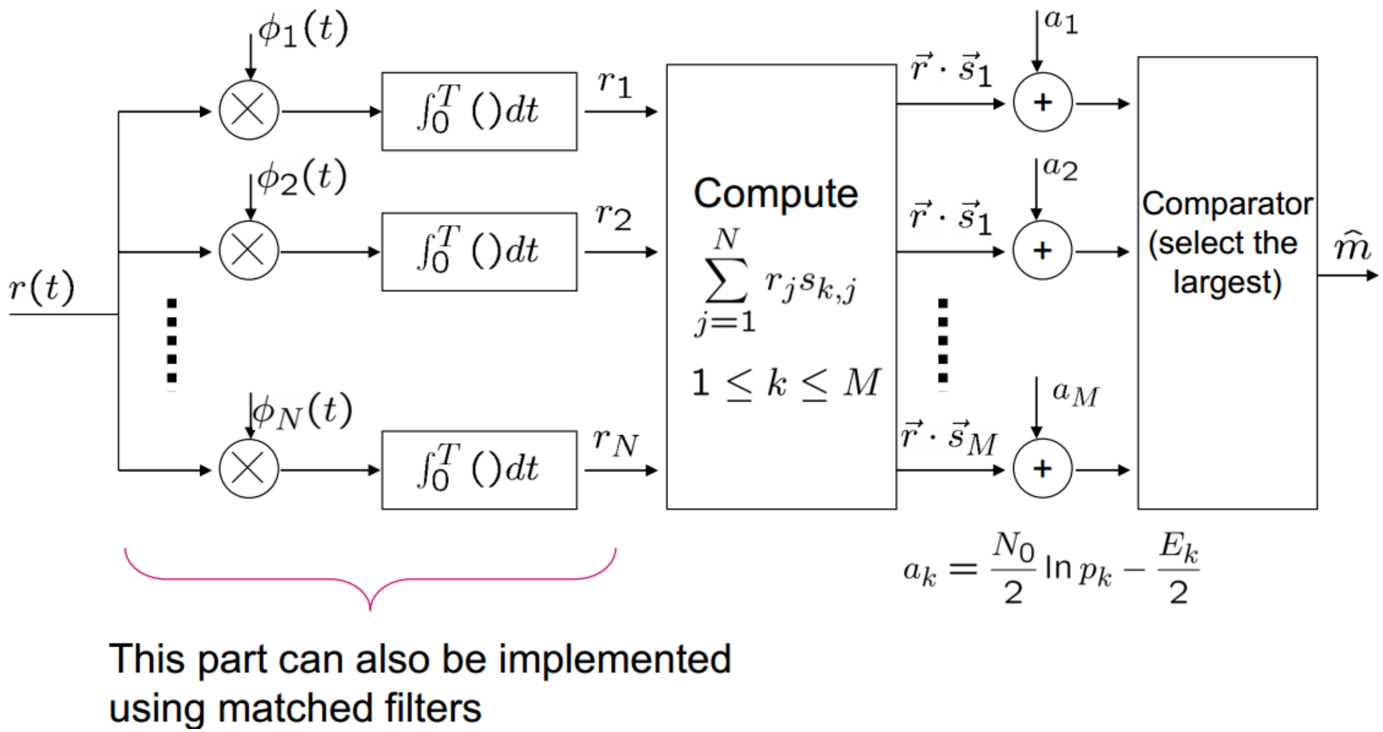
$\|\vec{r}\|^2$ = common to all M decisions and hence can be ignored

Likelihood function

- Optimal receiver structure
  - Hence, we have

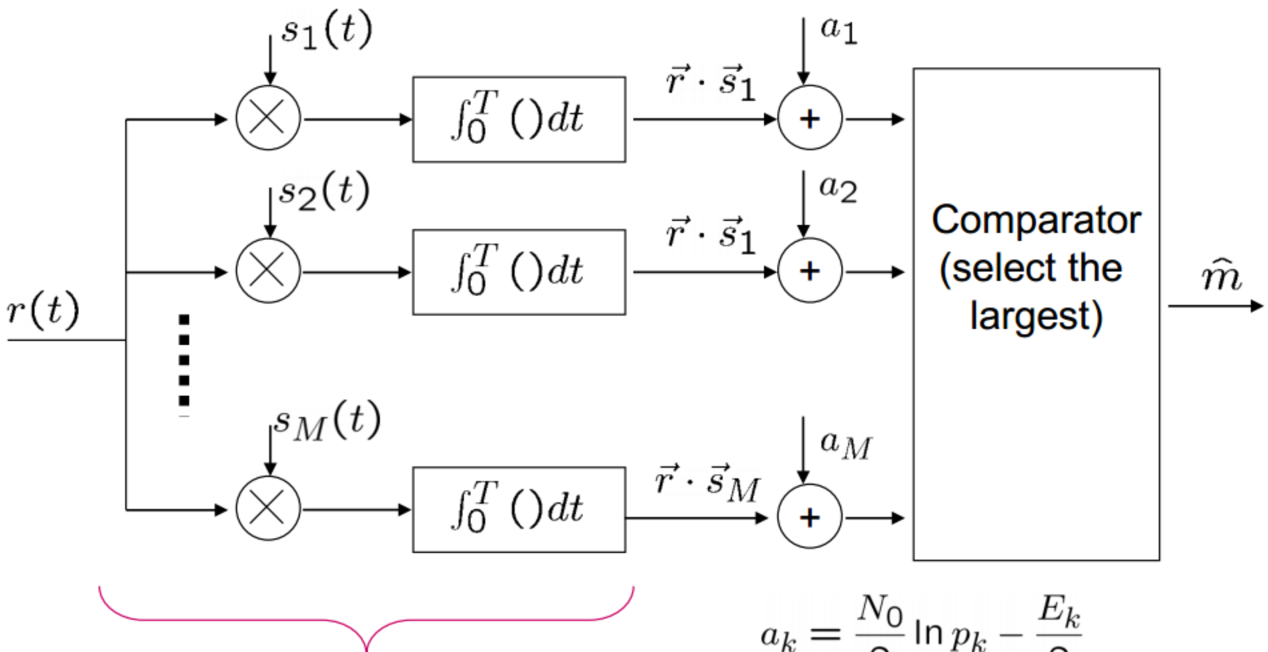
$$\hat{m} = \arg \max_{m_1,\dots,m_M} \left\{ \vec{r} \cdot \vec{s}_k - \frac{E_k}{2} + \frac{N_0}{2} \ln P_k \right\}$$

  - The diagram of MAP receiver can be (Method 1)



Likelihood function

- Optimal receiver structure
  - The diagram of MAP receiver can also be (Method 2)



$$a_k = \frac{N_0}{2}\ln p_k - \frac{E_k}{2}$$

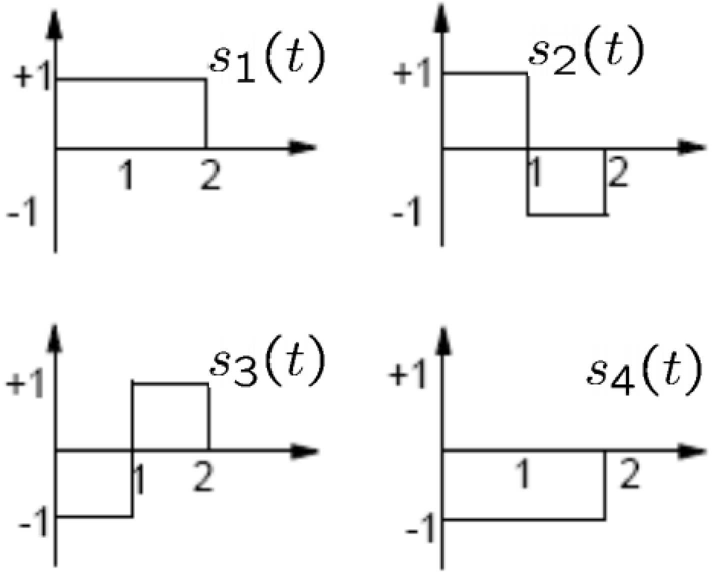
This part can also be implemented using matched filters

$$\hat{m} = \arg\max_{m_1,\ldots,m_M}\left\{\vec{r}\cdot\vec{s}_k - \frac{E_k}{2} + \frac{N_0}{2}\ln P_k\right\}$$

- The structures are general and can be simplified in certain cases.

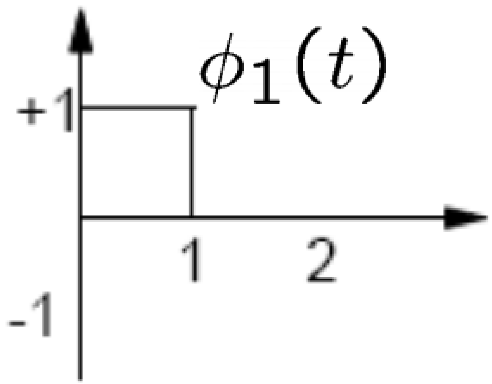
Likelihood function

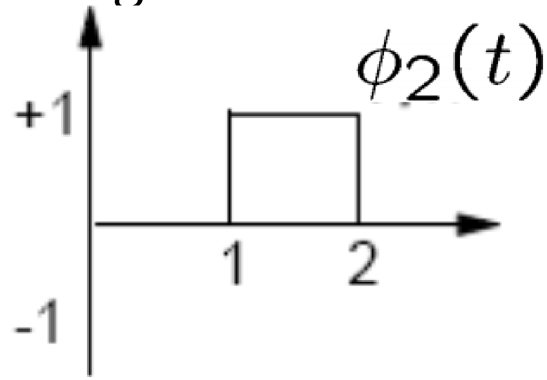
- Optimal receiver structure
  - Both receivers perform identically
  - Choice depends on circumstances
  - For instance, if N<M and $\{\phi_j(t)\}$ are easier to generate than $\{s_k(t)\}$, then the choice is obvious
  - Consider for example the following signal set



Likelihood function

- Optimal receiver structure
  - Suppose that we use the following basis functions





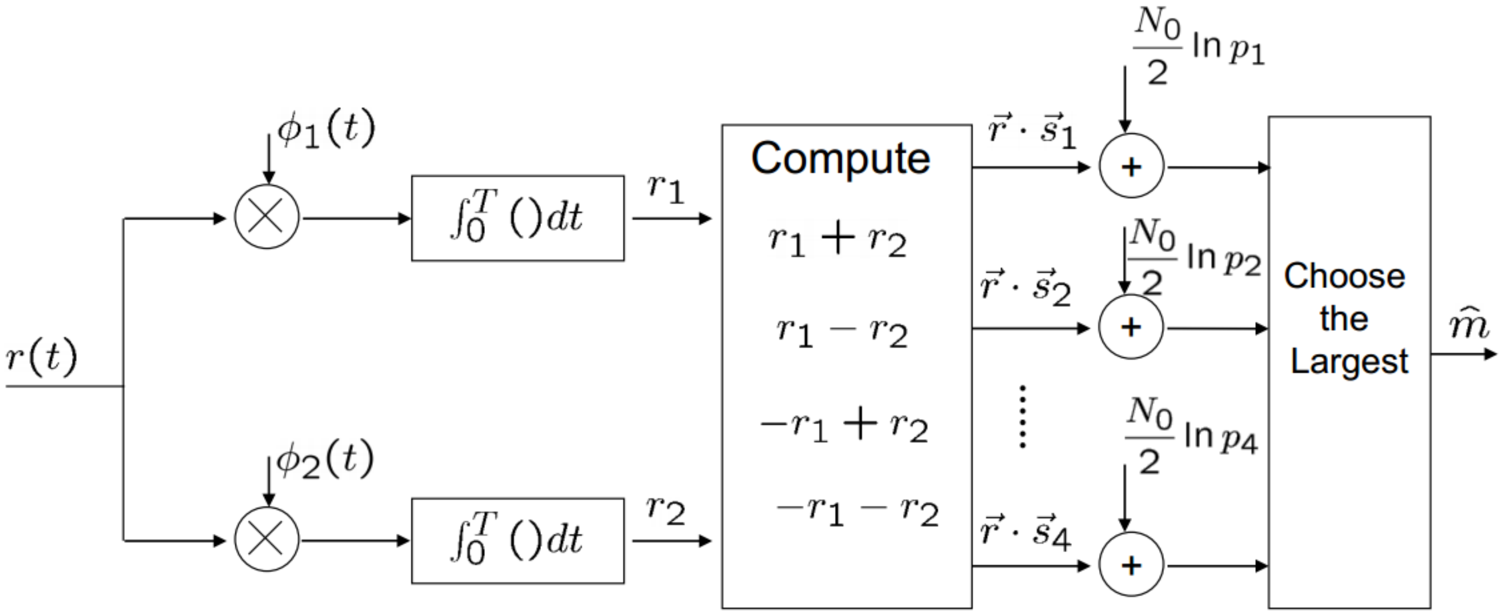
$$s_1(t) = 1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_2(t) = 1\,\phi_1(t) - 1\cdot\phi_2(t)$$

$$s_3(t) = -1\,\phi_1(t) + 1\cdot\phi_2(t) \qquad s_4(t) = -1\,\phi_1(t) - 1\,\phi_2(t)$$

  - Since the energy is the same for all four signals, we can drop the energy term, and hence $a_k = \frac{N_0}{2} \ln p_k$

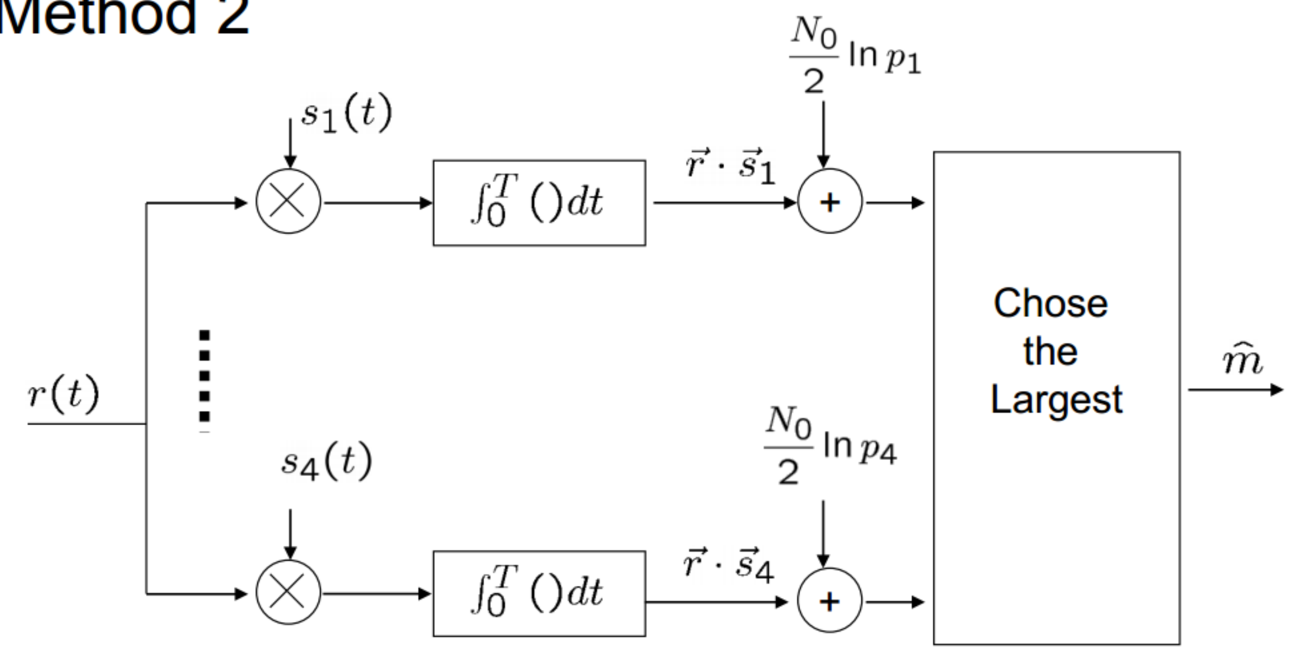
Likelihood function

- Optimal receiver structure
  - We have (Method 1)



Likelihood function

- Optimal receiver structure
  - Method 2



Likelihood function

- Optimal receiver structure
  - Exercise: In an additive white Gaussian noise channel with a noise power-spectral density of $N_0/2$, two equiprobable messages are transmitted by

$$s_1(t) = \begin{cases} \dfrac{At}{T} & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases}$$

$$s_2(t) = \begin{cases} A - \dfrac{At}{T} & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases}$$

  - Determine the optimal receiver structure.

Decision regions

- Graphical interpretation
  - Signal space can be divided into M disjoint decision regions R_1, R_2, …, R_M.

If $\vec{r} \in R_k$ $\Rightarrow$ decide m_k was transmitted

  - Select the decision regions so that P_e is minimized
  - Recall that the optimal receiver sets $\hat{m} = m_k$ iff $\|\vec{r} - \vec{s}_k\|^2 - N_0 \ln P_k$ is minimized
  - For simplicity, if one assumes $p_k = 1/M$ for all k, then the optimal receiver sets $\hat{m} = m_k$ iff $\|\vec{r} - \vec{s}_k\|^2$ is minimized

Decision regions

- Graphical interpretation
  - Geometrically, we take projection of r(t) in the signal space (i.e., $\vec{r}$). Then, decision is made in favor of signal that is the closest to $\vec{r}$ in the sense of minimum Euclidean distance.
  - Specifically, the observations with $\|\vec{r}-\vec{s}_k\|^2 < \|\vec{r}-\vec{s}_i\|^2$ for all $i \neq k$ should be assigned to decision region R_k
  - Consider for example the binary data transmission over AWGN channel with PSD $S_n(f)$=N0/2 using

$$s_1(t) = -s_2(t) = \sqrt{E}\phi(t)$$

  - Assume $P(m_1) \neq P(m_2)$. Determine the optimal receiver and the optimal decision regions.

Decision regions

- Graphical interpretation
  - For the above example, the optimal decision making

$$\|\vec{r}-\vec{s}_1\|^2 - N_0 \ln P(m_1) \underset{\text{Choose } m_2}{\overset{\text{Choose } m_1}{\lessgtr}} \|\vec{r}-\vec{s}_2\|^2 - N_0 \ln P(m_2)$$

  - Let $d_1 = \|\vec{r}-\vec{s}_1\|$ and $d_2 = \|\vec{r}-\vec{s}_2\|$
  - Equivalently,

$$d_1^2 - d_2^2 \underset{\text{Choose } m_2}{\overset{\text{Choose } m_1}{\lessgtr}} N_0 \ln \frac{P(m_1)}{P(m_2)} \quad \text{Constant c}$$

  - Therefore

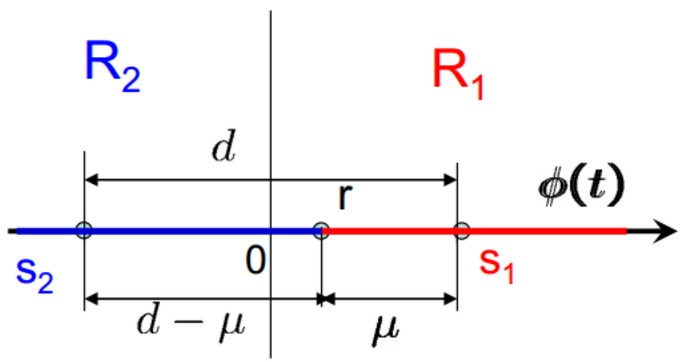
$$\text{R}_1\text{: } d_1^2 - d_2^2 < c \quad \textbf{and} \quad \text{R}_2\text{: } d_1^2 - d_2^2 > c$$

Decision regions

- Graphical interpretation
  - Now consider the example with $\vec{r}$ on the decision boundary

$$\begin{cases} d = d_1 + d_2 \\ d_1^2 = \mu^2 \\ d_2^2 = (d-\mu)^2 \end{cases} \Rightarrow \quad d_1^2 - d_2^2 = 2d\mu - d^2 \equiv c$$

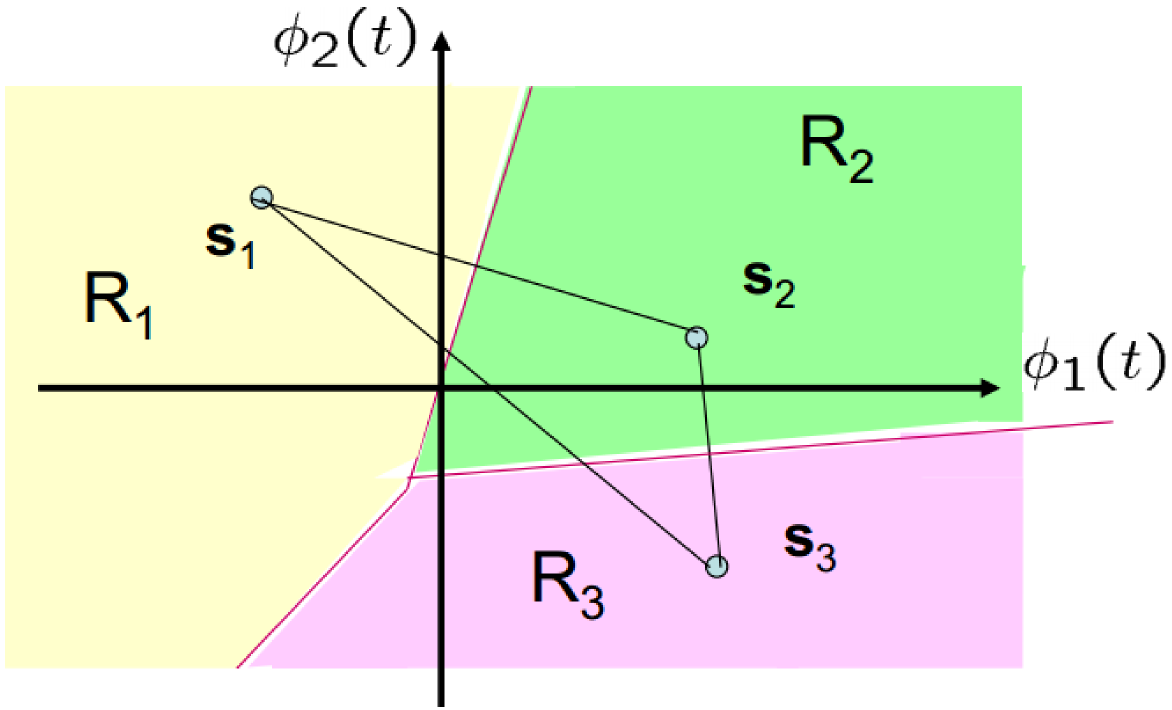
$$\mu = \frac{c + d^2}{2d} = \frac{d}{2} + \frac{N_0}{2d} \ln \frac{P(m_1)}{P(m_2)}$$



$$\mu \begin{cases} = d/2 & \text{if } P(m_1) = P(m_2) \\ > d/2 & \text{if } P(m_1) > P(m_2) \\ < d/2 & \text{if } P(m_1) < P(m_2) \end{cases}$$

Decision regions

- Graphical interpretation
  - In general, boundaries of decision regions are perpendicular bisectors of the lines joining the original transmitted signals
  - Example: three equiprobable 2-dim signals



Decision regions

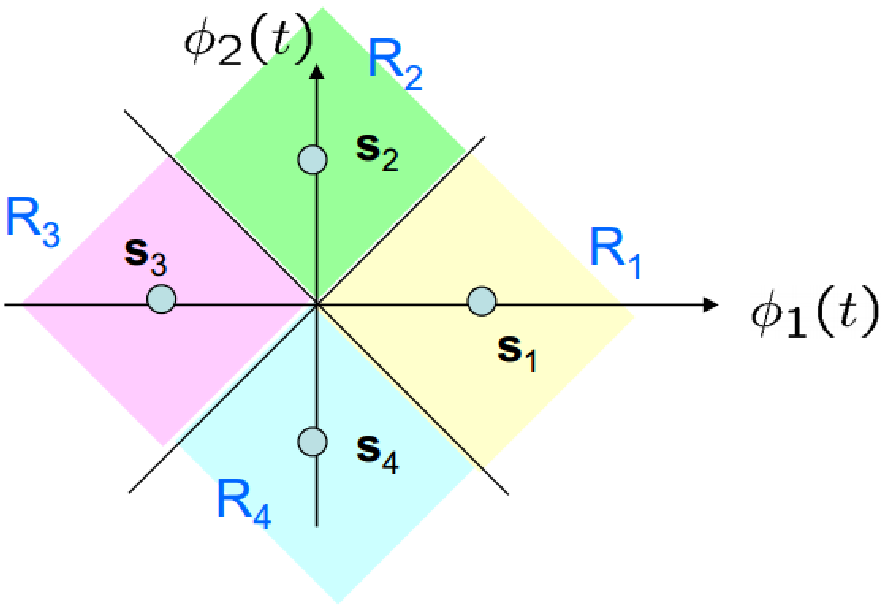
- Graphical interpretation
  - Now consider for example the decision regions for QPSK
  - Assume all signals are equally likely and all 4 signals could be written as the linear combination of two basis functions
  - Constellations of 4 signals

$\mathbf{s}_1$=(1,0)

$\mathbf{s}_2$=(0,1)

$\mathbf{s}_3$=(-1,0)

$\mathbf{s}_4$=(0,-1)



Decision regions

- Graphical interpretation
  - Exercise: Three equally probable messages m_1, m_2, and m_3 are to be transmitted over an AWGN channel with noise power-spectral density N0/2. The messages are

$$s_1(t)=\begin{cases}1 & 0\le t\le T\\ 0 & otherwise\end{cases}\qquad s_2(t)=-s_3(t)=\begin{cases}1 & 0\le t\le \frac{T}{2}\\ -1 & \frac{T}{2}\le t\le T\\ 0 & otherwise\end{cases}$$

  - What is the dimensionality of the signal space?
  - Find an appropriate basis for the signal space (Hint: you donot need to perform Gram-Schmidt procedure.)
  - Draw the signal constellation for this problem.
  - Sketch the optimal decisions R_1, R_2, and R_3.

Error probability analysis

- Probability of error
  - Suppose m_k is transmitted and $\vec{r}$ is received
  - Correct decision is made when $\vec{r} \in R_k$ with probability

$$P(C|m_k) = P(\vec{r} \in R_k | m_k \text{ is sent})$$

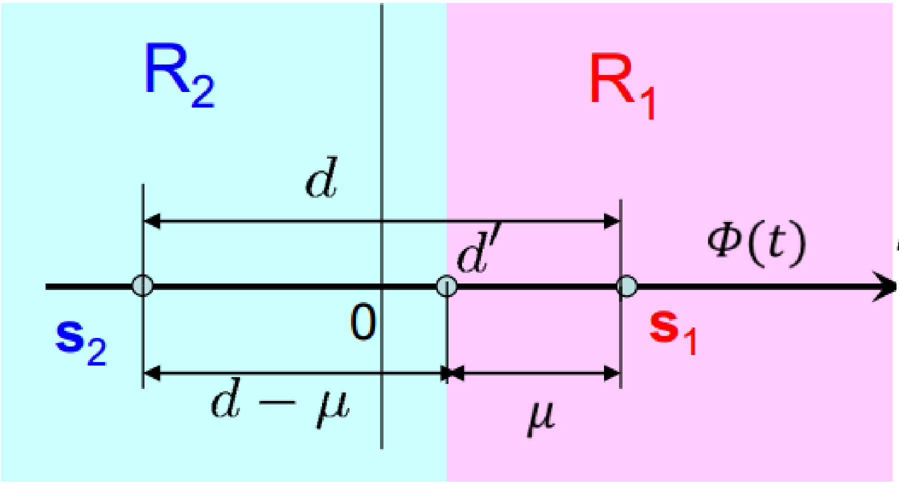
  - Averaging over all possible transmitted symbols, we obtain the average probability of making correct decision

$$P(C) = \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)$$

  - Average probability of error

$$\boxed{P_e = 1 - P(C) = 1 - \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)}$$

- Probability of error
  - Consider for example the binary data transmission



$$\mu = \frac{d}{2} + \frac{N_0}{2d}\ln\frac{P(m_1)}{P(m_2)}$$

  - Given m1 is transmitted, then

$$\begin{aligned} P(C|s_1) &= P(r \in R_1|s_1) \\ &= P(s_1 + n > d') \\ &= P(n > -\mu) \end{aligned}$$

  - Since n is Gaussian with variance $N_0/2$

$$P(C|s_1) = 1 - Q\left(\frac{\mu}{\sqrt{N_0/2}}\right)$$

Error probability analysis

- Probability of error
  - Similarly, we have

$$P(C|s_2) = P(s_2+n < d') = P(n < d-u) = 1-Q\left(\frac{d-\mu}{\sqrt{N_0/2}}\right)$$

  - Thus,

$$P(C) = P(m_1)\left\{1 - Q\left[\frac{\mu}{\sqrt{N_0/2}}\right]\right\} + P(m_2)\left\{1 - Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]\right\}$$

$$= 1 - P(m_1)Q\left[\frac{\mu}{\sqrt{N_0/2}}\right] - P(m_2)Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]$$

$$\Rightarrow \boxed{P_e = P(m_1)Q\left[\frac{\mu}{\sqrt{N_0/2}}\right] + P(m_2)Q\left[\frac{d-\mu}{\sqrt{N_0/2}}\right]}$$

where $d = 2\sqrt{E}$

$$\mu = \frac{N_0}{4\sqrt{E}}\log\left[\frac{P(m_1)}{P(m_2)}\right] + \sqrt{E}$$

Error probability analysis

- Probability of error
  - Note that when $P(m_1) = P(m_2)$

$$\mu = \sqrt{E} = \frac{d}{2}$$

$$\boxed{P_e = Q\left[\frac{d/2}{\sqrt{N_0/2}}\right] = Q\left[\sqrt{\frac{d^2}{2N_0}}\right] = Q\left[\sqrt{\frac{2E}{N_0}}\right]}$$

  - This shows us that:

1. When optimal receiver is used, P_e does not depend on the specific waveform used.
2. P_e depends only on their geometrical representation in signal space
3. In particular, P_e depends on signal waveforms only through their energies (distance)

Error probability analysis

- Graphical interpretation
  - Exercise: Three equally probable messages m_1, m_2, and m_3 are to be transmitted over an AWGN channel with noise power-spectral density $N_0/2$. The messages are

$$s_1(t)=\begin{cases}1 & 0\le t\le T\\ 0 & otherwise\end{cases}\qquad s_2(t)=-s_3(t)=\begin{cases}1 & 0\le t\le \frac{T}{2}\\ -1 & \frac{T}{2}\le t\le T\\ 0 & otherwise\end{cases}$$

  - Which of the three messages is more vulnerable to errors and why? In other words, which of the probability of error $p(Error\,|\,m_i \quad transmitted)$ is larger?

Error probability analysis

- General expression
  - Average probability of symbol error

$$P_e = 1 - P(C) = 1 - \sum_{k=1}^{M} P(\vec{r} \in R_k | m_k \text{ is sent}) P(m_k)$$

  - Since

$$P(\vec{r} \in R_k | m_k \text{ is sent}) = \underbrace{\int_{R_k}}_{\text{N-dim integration}} \overbrace{f(\vec{r}|m_k)}^{\text{Likelihood function}} d\vec{r}$$

  - Thus, we can rewrite Pe in terms of likelihood functions, assuming that symbols are equally likely to be sent

$$P_e = 1 - \frac{1}{M} \sum_{k=1}^{M} \int_{R_k} f(\vec{r}|m_k) d\vec{r}$$

  - Multi-dimension integrals are quite difficult to evaluate. To overcome the difficulty, we resort to the use of bounds. Then, we can obtain a simple and yet useful bound of P_e, called **union bound**.

Error probability analysis

- General expression
  - Let A_{kj} denote the event that $\vec{r}$ is closer to $\vec{s}_j$ than to $\vec{s}_k$ in the signal space when m_k ($\vec{s}_k$) is sent
  - Conditional probability of symbol error when m_k is sent

$$P(error|m_k) = P(\vec{r} \notin R_k|m_k) = P\left(\bigcup_{j\neq k} A_{kj}\right)$$

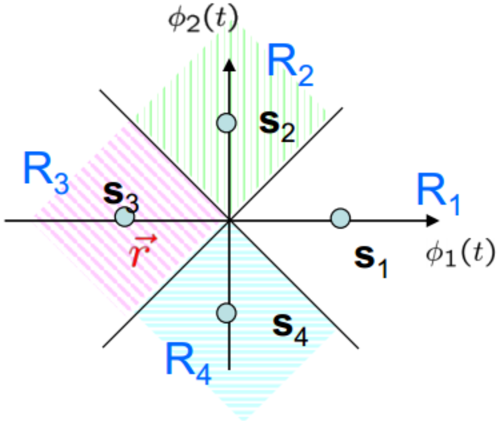
  - Note that

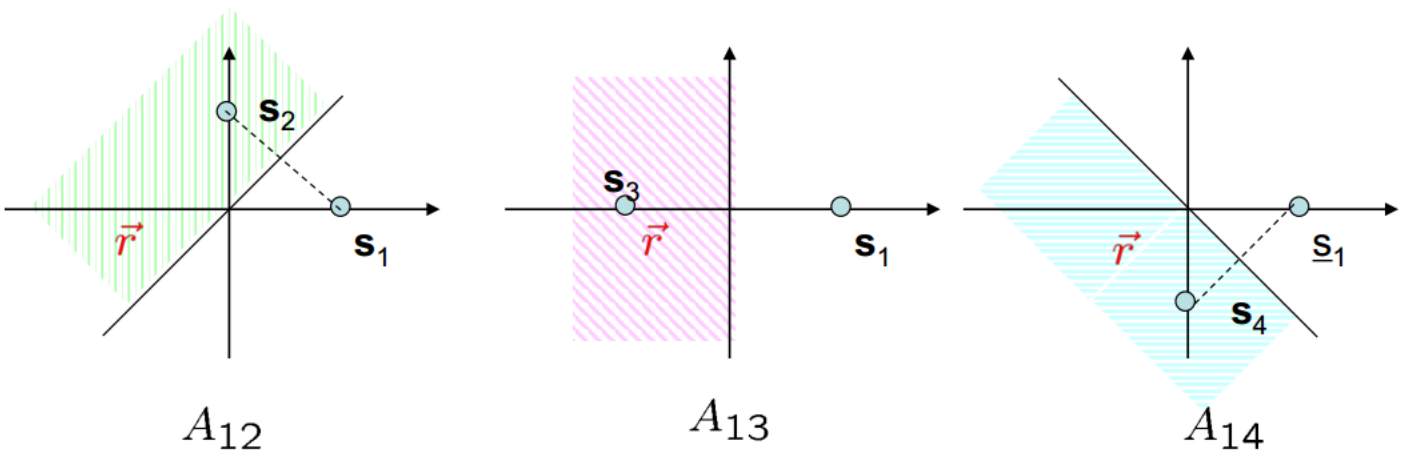
$$\boxed{P\left(\bigcup_{j\neq k} A_{kj}\right) \leq \sum_{\substack{j=1 \\ j \neq k}}^{M} P\left(A_{kj}\right)}$$

Error probability analysis

- General expression
  - ➢ Consider for example

$$A_{12} \cup A_{13} \cup A_{14}$$





Error probability analysis

- General expression
  - Define the pair-wise error probability as

$$P(\vec{s}_k \to \vec{s}_j) = P(A_{kj})$$

  - It is equivalent to the probability of deciding in favor of $\vec{s}_j$ when $\vec{s}_k$ was sent in a simplified binary system that involves the use of two equally likely messages $\vec{s}_k$ and $\vec{s}_j$
  - Then

$$P\left(\vec{s}_k \to \vec{s}_j\right) = P\left(n > d_{kj}/2\right) = Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right)$$

where $d_{kj} = \|\vec{s}_k - \vec{s}_j\|$ is the Euclidean distance between $\vec{s}_k$ and $\vec{s}_j$

  - Therefore the conditional error probability

$$P(error|m_k) \leq \sum_{\substack{j=1 \\ j \neq k}}^{M} P(\vec{s}_k \to \vec{s}_j) = \sum_{\substack{j=1 \\ j \neq k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right)$$

Error probability analysis

- General expression
  - Finally, with M equally likely messages, the average probability of symbol error is upperbounded by

$$\boxed{\begin{aligned} P_e &= \frac{1}{M}\sum_{k=1}^{M} P(error|m_k) \\ &\le \frac{1}{M}\sum_{k=1}^{M}\sum_{\substack{j=1\\ j\neq k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right) \end{aligned}}$$

The most general formulation of union bound

  - Let d_{min} denote the minimum distance, i.e., $d_{\text{min}} = \min\limits_{\substack{k,j\\ k\neq j}} d_{k,j}$

  - Since Q-function is a monotone decreasing function

$$\sum_{\substack{j=1\\ j\neq k}}^{M} Q\left(\sqrt{\frac{d_{kj}^2}{2N_0}}\right) \le (M-1)Q\left(\sqrt{\frac{d_{\text{min}}^2}{2N_0}}\right)$$

Error probability analysis

- General expression
  - Consequently, we may simplify the union bound as

$$\boxed{P_e \le (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right)}$$

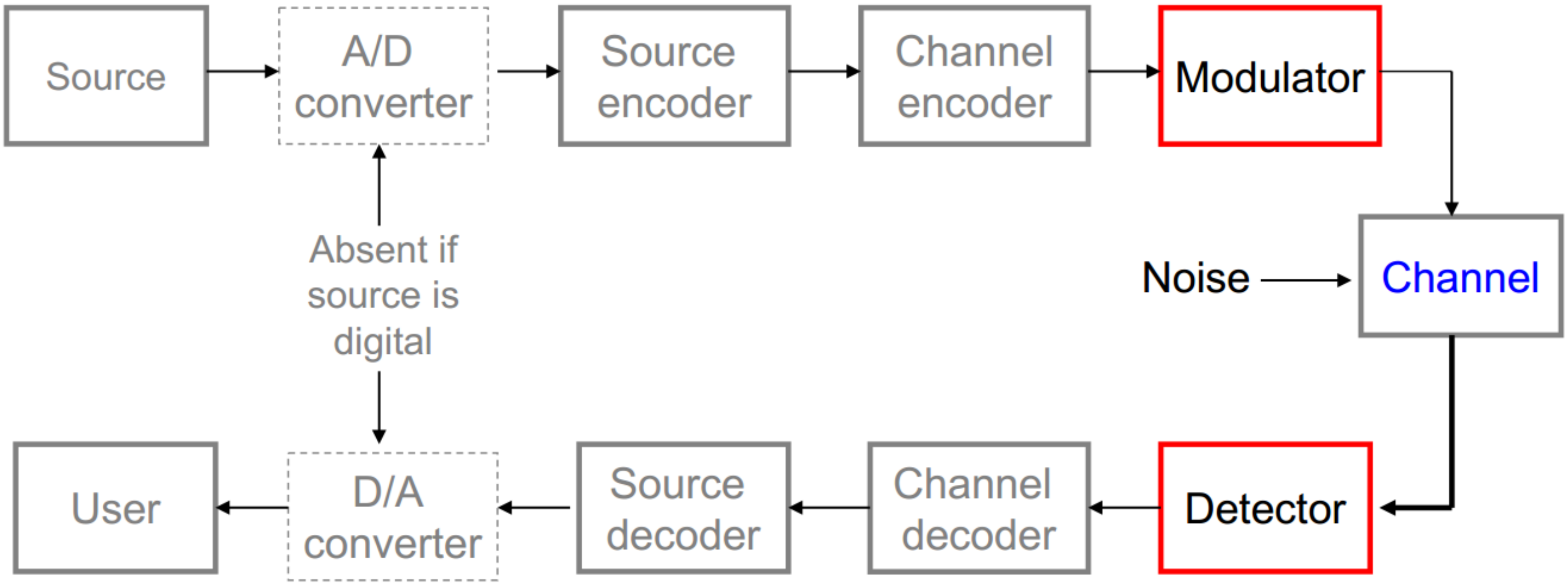
Simplified form of union bound

  - Think about: What is the design criterion of a good signal set?

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- **Digital modulation techniques [Matlab]**
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

Digital modulation techniques



- Binary digital modulation
- M-ary digital modulation
- Comparison study

Chapter 8.2， 8.3.3， 8.5-8.7， 9.1-9.5， 9.7

Digital modulation techniques

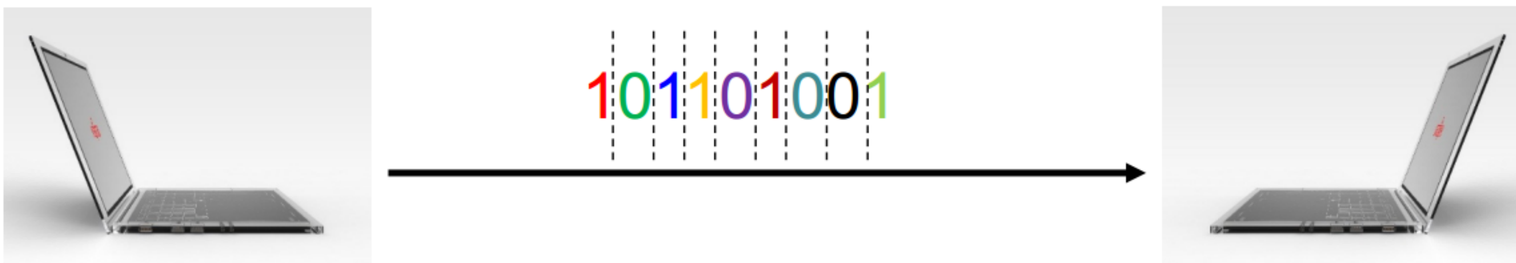
- In digital communications, the modulation process corresponds to **switching** or **keying** the amplitude, frequency, or phase of a **sinusoidal** carrier wave corresponding to incoming digital data
- Three basic digital modulation techniques

1. Amplitude-shift keying (ASK) - special case of AM
2. Frequency-shift keying (FSK) - special case of FM
3. Phase-shift keying (PSK) - special case of PM

- We use signal space approach in receiver design and performance analysis

Binary digital modulation

- In binary signaling, the modulator produces one of two distinct signals in response to one bit of source data at a time.

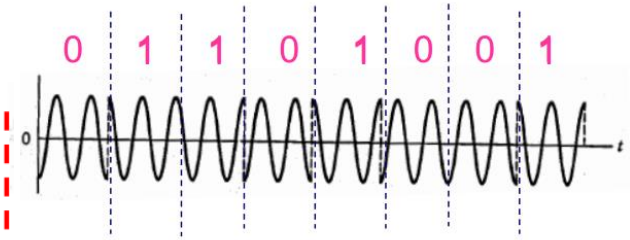


- Binary modulation type

Binary PSK | **Binary FSK** | **Binary ASK**

Binary digital modulation

- Binary Phase-Shift Keying (**BPSK**)
  - Modulation



"1" $\longrightarrow s_1(t) = \sqrt{\frac{2E_b}{T_b}}\cos\boxed{(2\pi f_c t)}$

"0" $\longrightarrow s_2(t) = \sqrt{\frac{2E_b}{T_b}}\cos\boxed{(2\pi f_c t + \pi)} = -\sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_c t)$

  - $0 \le t < T_b$, T_b bit duration
  - f_c : carrier frequency, chosen to be n_c/T_b for some fixed integer n_c or $f_c >> 1/T_b$
  - E_b : transmitted signal energy per bit, i.e.,

$$\int_0^{T_b} s_1^2(t)dt = \int_0^{T_b} s_2^2(t)dt = E_b$$

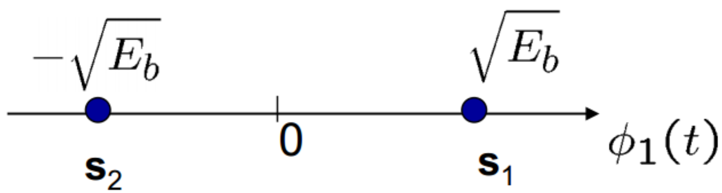
  - The pair of signals differ only in a 180-degree phase shift

Binary digital modulation

- Binary Phase-Shift Keying (**BPSK**)
  - Signal space representation:

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t) \quad \text{with} \quad 0 \le t < T_b$$

  - So $s_1(t) = \sqrt{E_b}\phi_1(t)$ and $s_2(t) = -\sqrt{E_b}\phi_1(t)$
  - A binary PSK system is characterized by a signal space that is one-dimensional (N=1), and has two message points (M=2)



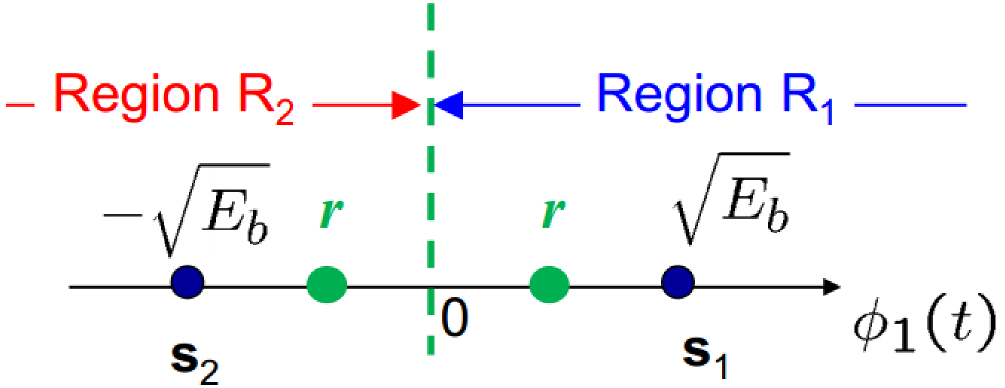
$$d_{12} = 2\sqrt{E_b}$$

  - Assume that the two signals are equally likely, i.e.,

$$P(s_1) = P(s_2) = 0.5$$

Binary digital modulation

- Binary Phase-Shift Keying (**BPSK**)
  - The optimal decision boundary is the midpoint of the line joining these two message points



  - Decision rule:

1. Guess signal $s_1(t)$ (or binary 1) was transmitted if the received signal point r falls in region R_1 $(r > 0)$
2. Guess signal $s_2(t)$ (or binary 0) was transmitted otherwise $(r \leq 0)$

Binary digital modulation

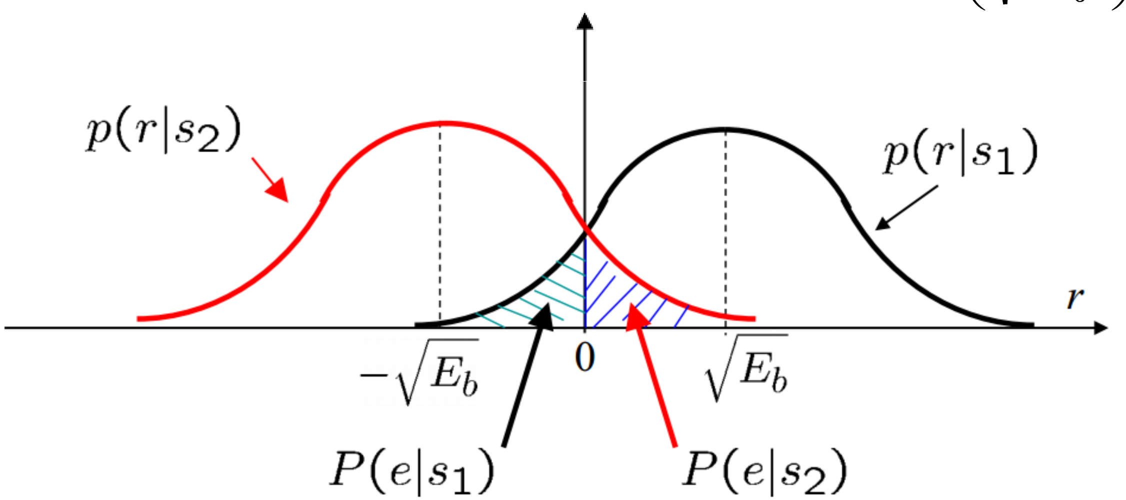
- Binary Phase-Shift Keying (**BPSK**)
  - Probability of error analysis.
  - The conditional probability of the receiver deciding in favor of $s_2(t)$ given that $s_1(t)$ was transmitted is

$$P(e|s_1) = P(r < 0|s_1)$$

$$= \int_{-\infty}^{0} \frac{1}{\sqrt{\pi N_0}} \exp\left\{-\frac{(r-\sqrt{E_b})^2}{N_0}\right\} dr = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

  - Due to symmetry

$$P(e|s_2) = P(r > 0|s_2) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$



Binary digital modulation

- Binary Phase-Shift Keying (**BPSK**)
  - Probability of error analysis.
  - Since the signals $s_1(t)$ and $s_2(t)$ are equally likely to be transmitted, the average probability of error is

$$P_e = 0.5P(e|\mathbf{s}_1) + 0.5P(e|\mathbf{s}_2) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

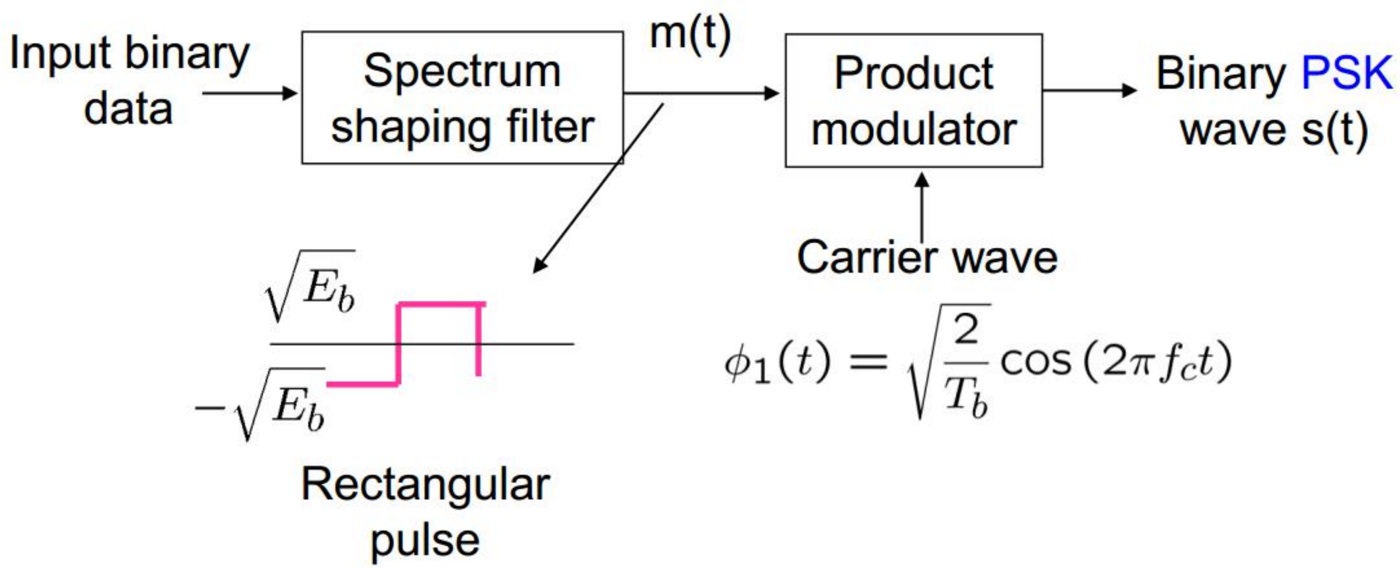
$$\Downarrow$$

$$\boxed{P_e \text{ depends on ratio } \frac{E_b}{N_0}}$$

  - This ratio is normally called bit energy to noise density ratio (SNR/bit)

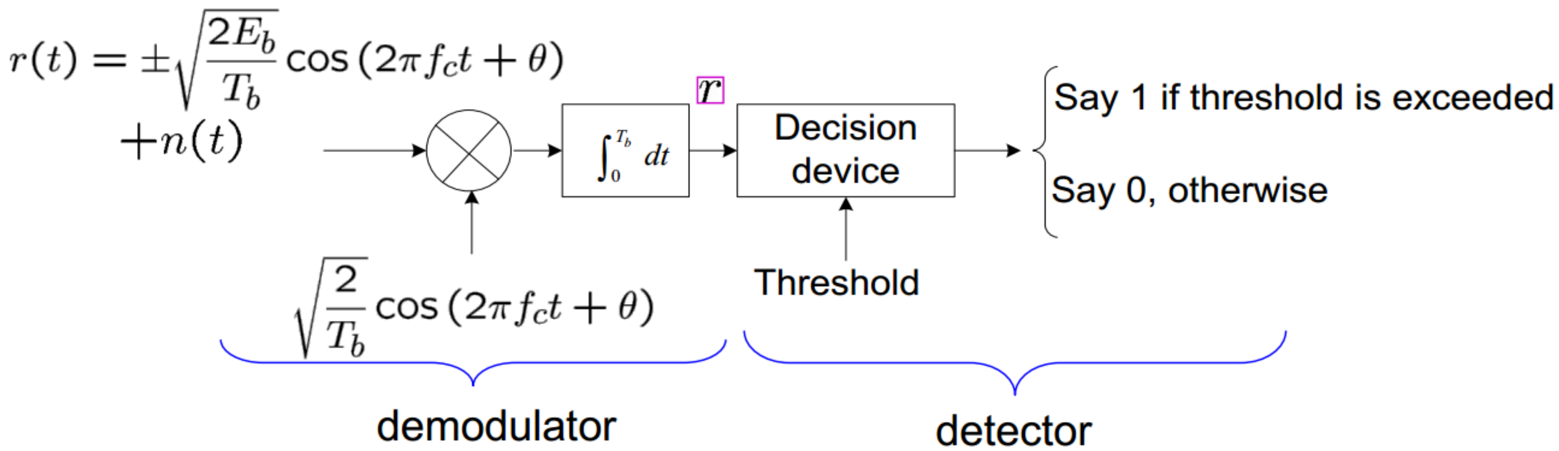
Binary digital modulation

- Binary Phase-Shift Keying (**BPSK**)
  - ➢ Transmitter.



Binary digital modulation

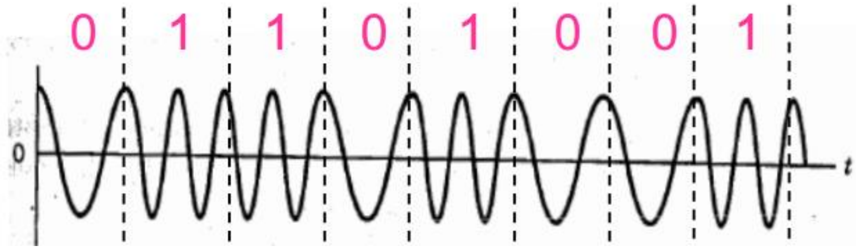
- Binary Phase-Shift Keying (**BPSK**)
  - Receiver.



  - θ is the carrier-phase offset, due to propagation delay or oscillators at the transmitter and receiver are not synchronous
  - The detection is coherent in the sense of phase synchronization and timing synchronization

Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Modulation



"1" $\longrightarrow s_1(t) = \sqrt{\dfrac{2E_b}{T_b}}\cos(2\pi f_1 t)$

"0" $\longrightarrow s_2(t) = \sqrt{\dfrac{2E_b}{T_b}}\cos(2\pi f_2 t) \quad 0 \le t < T_b$

  - E_b : transmitted signal energy per bit

$$\int_0^{T_b} s_1^2(t)dt = \int_0^{T_b} s_2^2(t)dt = E_b$$

  - f_i : transmitted frequency with separation $\Delta f = f_1 - f_0$
  - Δf is selected so that $s_1(t)$ and $s_2(t)$ are orthogonal, i.e.,

$$\int_0^{T_b} s_1(t)s_2(t)dt = 0$$

(Example?)

Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Signal space representation:

$$\phi_1(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_1 t) \qquad 0 \le t < T_b$$

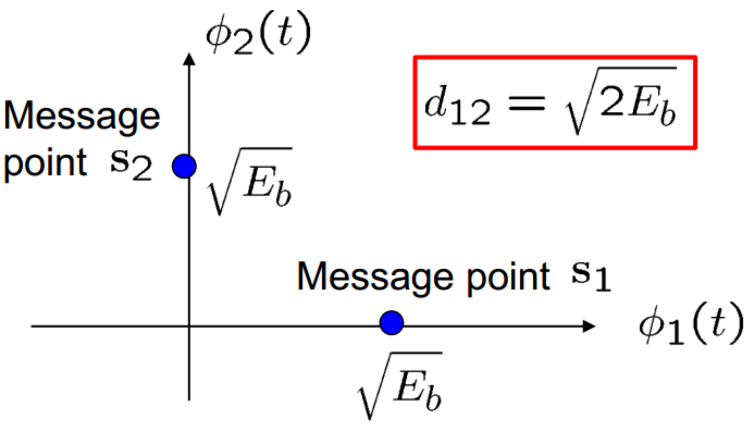
$$\phi_2(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_2 t) \qquad 0 \le t < T_b$$

$$\Rightarrow \quad s_1(t) = \sqrt{E_b}\,\phi_1(t), \qquad s_2(t) = \sqrt{E_b}\,\phi_2(t)$$

$$\mathrm{s}_1 = [\sqrt{E_b} \quad 0]$$

$$\mathrm{s}_2 = [0 \quad \sqrt{E_b}]$$

$$\boxed{d_{12} = \sqrt{2E_b}}$$



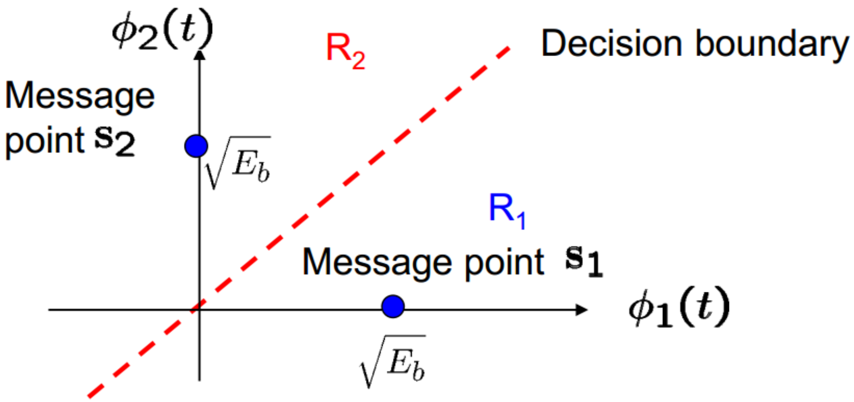
Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Decision regions:

$$\vec{r} = [r_1 \quad r_2]$$

$$r_1 = \int_0^{T_b} r(t)\phi_1(t)dt$$

$$r_2 = \int_0^{T_b} r(t)\phi_2(t)dt$$



1. Guess signal $s_1(t)$ (or binary 1) was transmitted if the received signal point ***r*** falls in region R_1 ($r_2<r_1$)
2. Guess signal $s_2(t)$ (or binary 0) was transmitted otherwise

Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Probability of error analysis.
  - Given that $\mathbf{s}_1$ is transmitted

    $$r_1 = \sqrt{E_b} + n_1 \quad \text{and} \quad r_2 = n_2$$

  - Since the condition $r_2 > r_1$ corresponds to the receiver making a decision in favor of symbol $\mathbf{s}_2$, the conditional probability of error when $\mathbf{s}_1$ is transmitted is given by

    $$P(e|s_1) = P(r_1 < r_2|s_1) = P(\sqrt{E_b} + n_1 < n_2)$$

  - n_1 and n_2 are i.i.d. Gaussian with $n_1, n_2 \in \mathcal{N}(0, N_0/2)$
  - Then $n = n_1 - n_2$ is Gaussian with $n \in \mathcal{N}(0, N_0)$

$$\Rightarrow \quad P(e|s_1) = P(n < -\sqrt{E_b}) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Probability of error analysis.
  - By symmetry, we also have

$$P(e|s_2) = P(r_1 > r_2|s_2) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

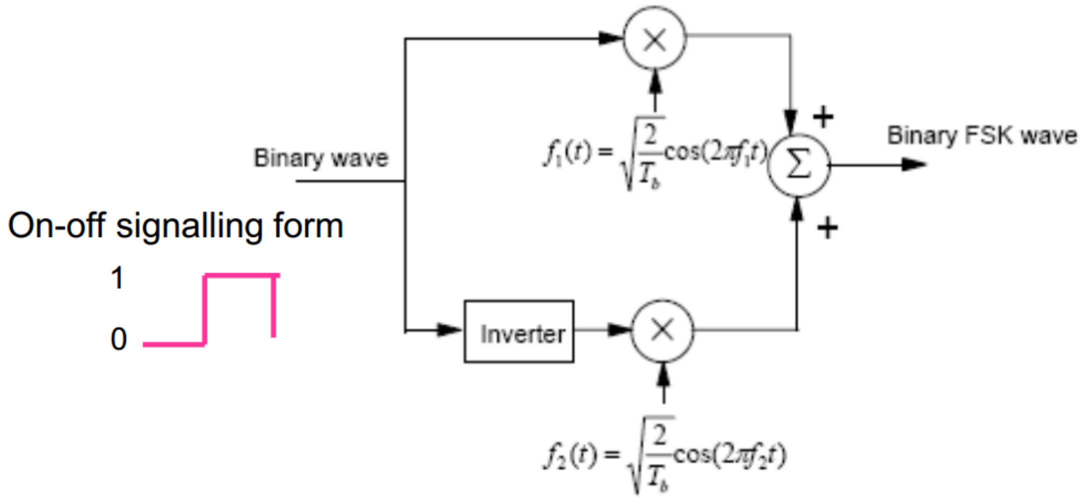
  - Since the two signals are equally likely to be transmitted, the average probability of error for coherent binary FSK is

$$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \quad \Rightarrow \quad \text{3 dB worse than BPSK}$$

To achieve the same P_e, BFSK needs 3dB more transmission power than BPSK

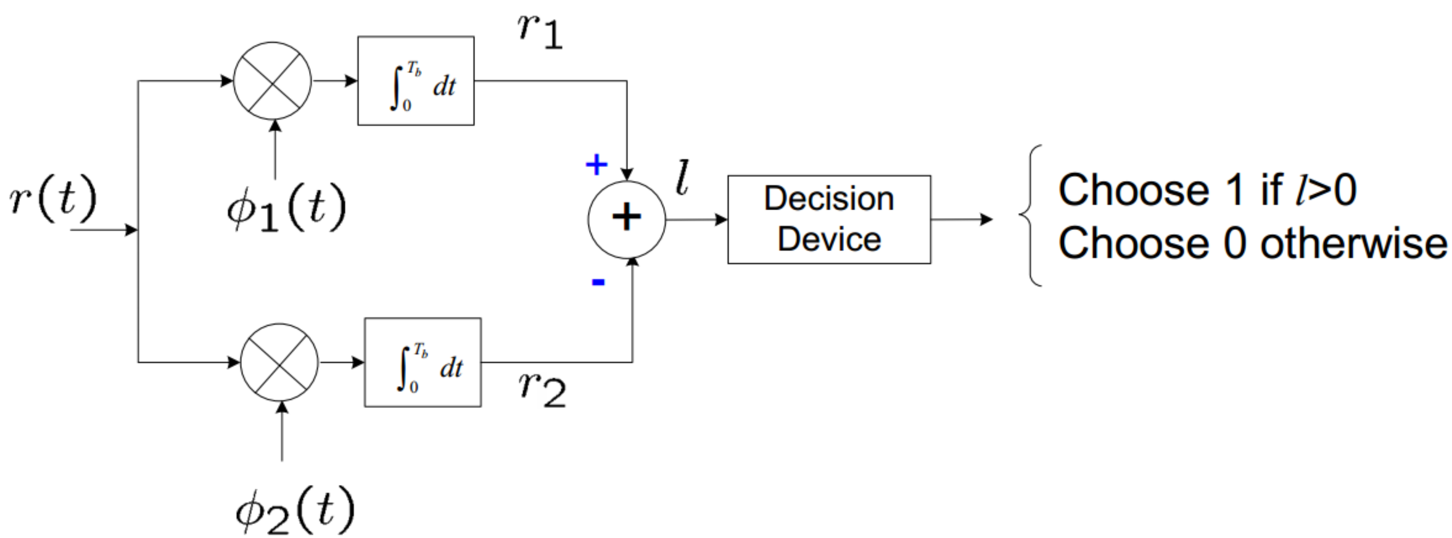
Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Transmitter.



Binary digital modulation

- Binary Frequency-Shift Keying (**BFSK**)
  - Receiver.

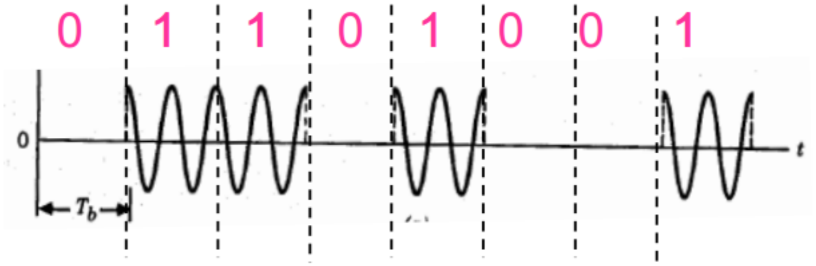


Binary digital modulation

- Binary Amplitude-Shift Keying (**BASK**)
  - Modulation.

"1" $\longrightarrow$ $s_1(t) = \sqrt{\frac{2E}{T_b}} \cos(2\pi f_c t)$

"0" $\longrightarrow$ $s_2(t) = 0 \qquad 0 \le t < T_b$



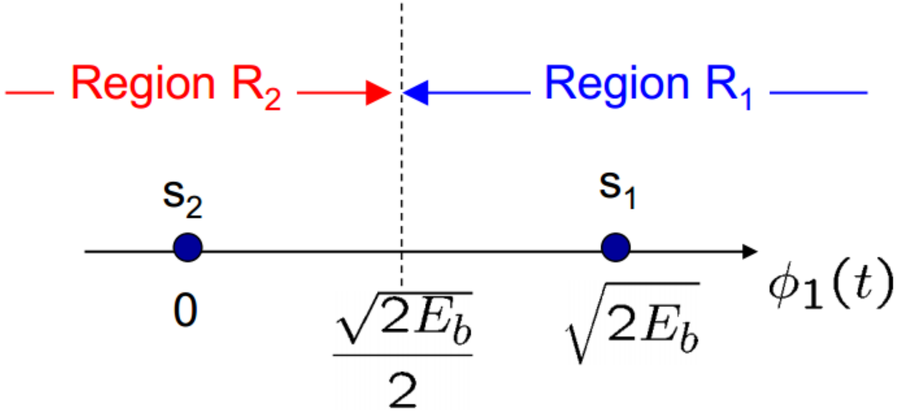
(On-off signaling)

  - Average energy per bit

$$E_b = \frac{E+0}{2} \qquad \text{i.e. } E = 2E_b$$

  - Decision region

$$d_{12} = \sqrt{2E_b}$$



Binary digital modulation

- Binary Amplitude-Shift Keying (**BASK**)
  - Probability of error analysis.
  - Average probability of error

Prove it!

$$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Identical to that of coherent binary FSK

Binary digital modulation

- Comparison

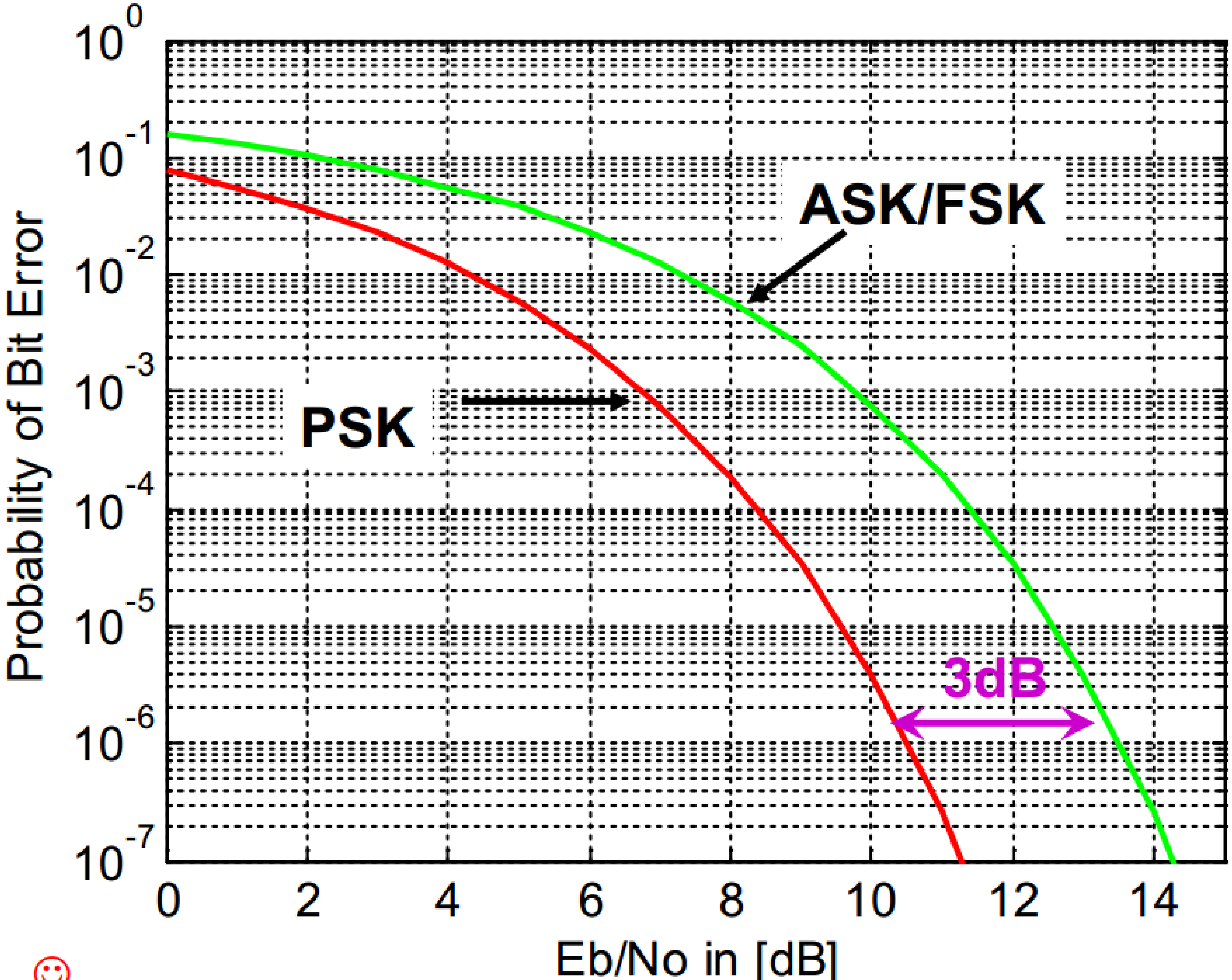
BPSK	BFSK	BASK
$d_{1,2} = 2\sqrt{E_b}$	$d_{1,2} = \sqrt{2E_b}$	$d_{1,2} = \sqrt{2E_b}$
$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$	$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$	$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$

- In general,

$$\boxed{P_e = Q\left(\sqrt{\frac{d_{12}^2}{2N_0}}\right)}$$

Binary digital modulation

- Comparison



Binary digital modulation

- Example
  - Binary data are transmitted over a microwave link at the rate of 10^6 bits/sec and the PSD of the noise at the receiver input is 10^{-10} watts/Hz.
  - Find the average carrier power required to maintain an average probability of error $P_e \leq 10^{-4}$ for coherent binary FSK.
  - What if noncoherent binary FSK?

Binary digital modulation

- Update
  - We have discussed coherent modulation schemes, e.g., BPSK, BFSK, BASK, which need coherent detection assuming that the receiver is able to detect and track the carrier wave's phase
  - In many practical situations, strict phase synchronization is not possible. In these situations, non-coherent reception is required.
  - We now consider non-coherent detection on binary FSK and differential phase-shift keying (DPSK)

Binary digital modulation

- Non-coherent scheme: BFSK
  - Consider a binary FSK system, the two signals are

$$
\begin{aligned}
s_1(t) &= \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_1 t + \theta_1) \\
& \qquad\qquad\qquad\qquad\qquad 0 \le t < T_b \\
s_2(t) &= \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_2 t + \theta_2)
\end{aligned}
$$

  - θ_1, θ_2 : unknown random phases with uniform distribution

$$
p_{\theta_1}(\theta) = p_{\theta_2}(\theta) = \begin{cases} 1/2\pi & \theta \in [0, 2\pi) \\ 0 & \text{else} \end{cases}
$$

Binary digital modulation

- Non-coherent scheme: BFSK
  - Since

$$s_1(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_1 t + \theta_1) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_1 t)\cos(\theta_1) - \sqrt{\frac{2E_b}{T_b}}\sin(2\pi f_1 t)\sin(\theta_1)$$

$$s_2(t) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_2 t + \theta_2) = \sqrt{\frac{2E_b}{T_b}}\cos(2\pi f_2 t)\cos(\theta_2) - \sqrt{\frac{2E_b}{T_b}}\sin(2\pi f_2 t)\sin(\theta_2)$$

  - Choose four basis functions as

$$\phi_{1c}(t) = \sqrt{2/T_b}\cos(2\pi f_1 t) \quad \phi_{1s}(t) = -\sqrt{2/T_b}\sin(2\pi f_1 t)$$

$$\phi_{2c}(t) = \sqrt{2/T_b}\cos(2\pi f_2 t) \quad \phi_{2s}(t) = \sqrt{2/T_b}\sin(2\pi f_2 t)$$

  - Signal space representation

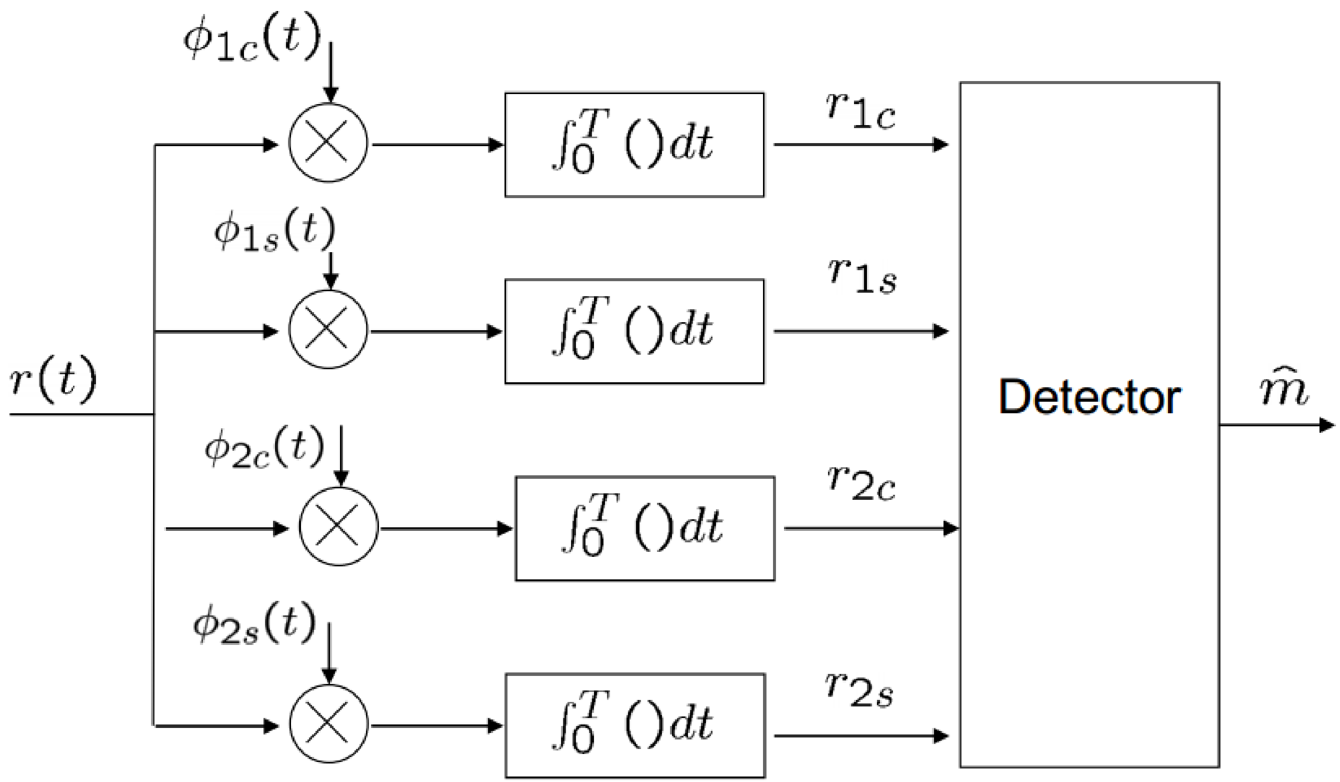
$$\vec{s}_1 = [\ \sqrt{E_b}\cos\theta_1 \quad \sqrt{E_b}\sin\theta_1 \quad 0 \quad 0\ ]$$

$$\vec{s}_2 = [\ 0 \quad 0 \quad \sqrt{E_b}\cos\theta_2 \quad \sqrt{E_b}\sin\theta_2\ ]$$

Binary digital modulation

- Non-coherent scheme: BFSK
  - The vector representation of the received signal

$$\vec{r} = [r_{1c} \quad r_{1s} \quad r_{2c} \quad r_{2s}]$$



Binary digital modulation

- Non-coherent scheme: BFSK
  - Decision rule:

$$f(\vec{r}|\vec{s}_1) \underset{\text{Choose } s_2}{\overset{\text{Choose } s_1}{\gtrless}} f(\vec{r}|\vec{s}_2)$$

ML

  - Conditional pdf

$$f(\vec{r}|\vec{s}_1,\theta_1) = \frac{1}{\pi N_0}\exp\left[-\frac{(r_{1c}-\sqrt{E_b}\cos\theta_1)^2+(r_{1s}-\sqrt{E_b}\sin\theta_1)^2}{N_0}\right] \times \frac{1}{\pi N_0}\exp\left[-\frac{r_{2c}^2+r_{2s}^2}{N_0}\right]$$

  - Similarly

$$f(\vec{r}|\vec{s}_2,\theta_2) = \frac{1}{\pi N_0}\exp\left[-\frac{r_{1c}^2+r_{1s}^2}{N_0}\right] \times \frac{1}{\pi N_0}\exp\left[-\frac{(r_{2c}-\sqrt{E_b}\cos\theta_2)^2+(r_{2s}-\sqrt{E_b}\sin\theta_2)^2}{N_0}\right]$$

Binary digital modulation

- Non-coherent scheme: BFSK
  - For ML decision, we need to evaluate

$$f(\vec{r}|\vec{s}_1) \geq f(\vec{r}|\vec{s}_2)$$

$$\frac{1}{2\pi}\int_0^{2\pi} f(\vec{r}|\vec{s}_1,\theta_1)d\theta_1 \geq \frac{1}{2\pi}\int_0^{2\pi} f(\vec{r}|\vec{s}_2,\theta_2)d\theta_2$$

  - Removing the constant terms

$$\left(\frac{1}{\pi N_0}\right)^2 \exp\left[-\frac{r_{1c}^{\ 2}+r_{1s}^{\ 2}+r_{2c}^{\ 2}+r_{2s}^{\ 2}+E}{N_0}\right]$$

  - We have

$$\frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{1c}\cos(\phi_1)+2\sqrt{E}r_{1s}\sin(\phi_1)}{N_0}\right]d\phi_1$$

$$\geq \frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{2c}\cos(\phi_1)+2\sqrt{E}r_{2s}\sin(\phi_1)}{N_0}\right]d\phi_2$$

Binary digital modulation

- Non-coherent scheme: BFSK
  - By definition

$$\frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{1c}\cos(\phi_1)+2\sqrt{E}r_{1s}\sin(\phi_1)}{N_0}\right]d\phi_1 = I_0\left(\frac{2\sqrt{E(r_{1c}^2+r_{1s}^2)}}{N_0}\right)$$

where $I_0()$ is a modified Bessel function of the zero-th order

  - Thus, the decision rule becomes: choose $\mathbf{s}_1$ if

$$I_0\left(\frac{2\sqrt{E(r_{1c}^2+r_{1s}^2)}}{N_0}\right) \geq I_0\left(\frac{2\sqrt{E(r_{2c}^2+r_{2s}^2)}}{N_0}\right)$$

Binary digital modulation

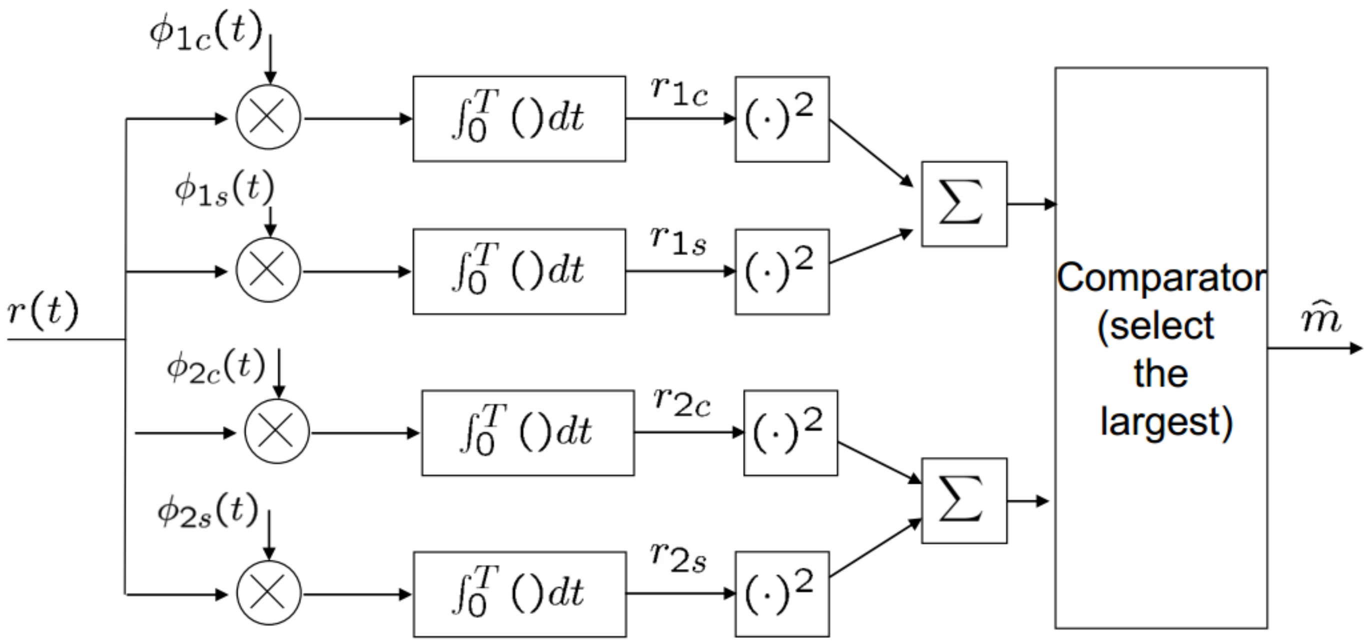
- Non-coherent scheme: BFSK
  - Note that this Bessel function is monotonically increasing. Therefore, we choose **s**1 if

$$\sqrt{r_{1c}^2 + r_{1s}^2} \geq \sqrt{r_{2c}^2 + r_{2s}^2}$$

1. Useful insight: we just compare the energy in the two frequencies and pick the larger (**envelope detector**)
2. Carrier phase is irrelevant in decision making

Binary digital modulation

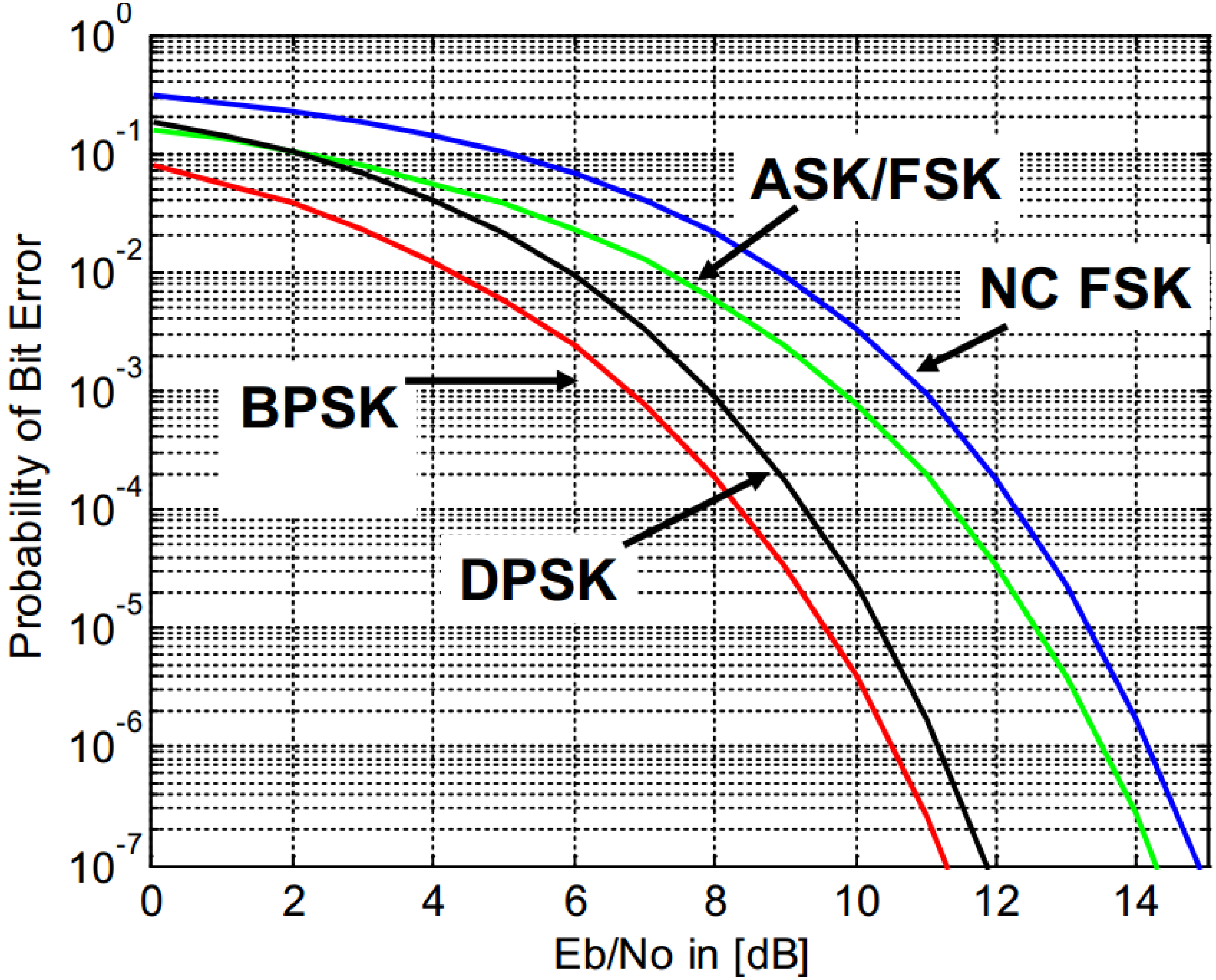
- Non-coherent scheme: BFSK
  - Structure.



$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$$

See Section 9.5.2

- Comparison



Probability of Bit Error
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
2
4
6
8
10
12
14
Eb/No in [dB]
ASK/FSK
NC FSK
BPSK
DPSK

Binary digital modulation

- Differential PSK (**DPSK**)
  - Non-coherent version of PSK
  - Phase synchronization is eliminated using differential encoding

1. Encode the information in phase difference between successive signal transmission.
2. Send “0” , advance the phase of the current signal by 180^o
3. Send “1”, leave the phase unchanged

  - Provided that the unknown phase θ contained in the received wave varies slowly (constant over two bit intervals), the phase difference between waveforms received in two successive bit intervals will be independent of θ

Binary digital modulation

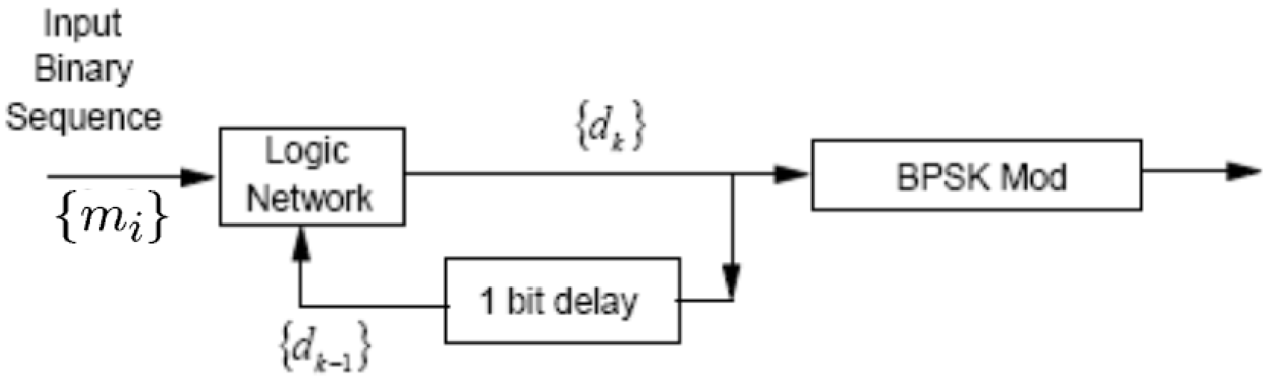
- Differential PSK (**DPSK**)
  - Generate DPSK signals in two steps
    1. Differential encoding of the information binary bits.
    2. Phase shift keying
  - Differential encoding starts with an arbitrary reference bit

Information sequence		1	0	0	1	0	0	1	1	$\{m_i\}$
Differentially encoded sequence	1	1	0	1	1	0	1	1	1	$\{d_i\}$
	Initial bit									
Transmitted Phase	0	0	π	0	0	π	0	0	0	

$$\boxed{d_i = \overline{d_{i-1} \oplus m_i}}$$

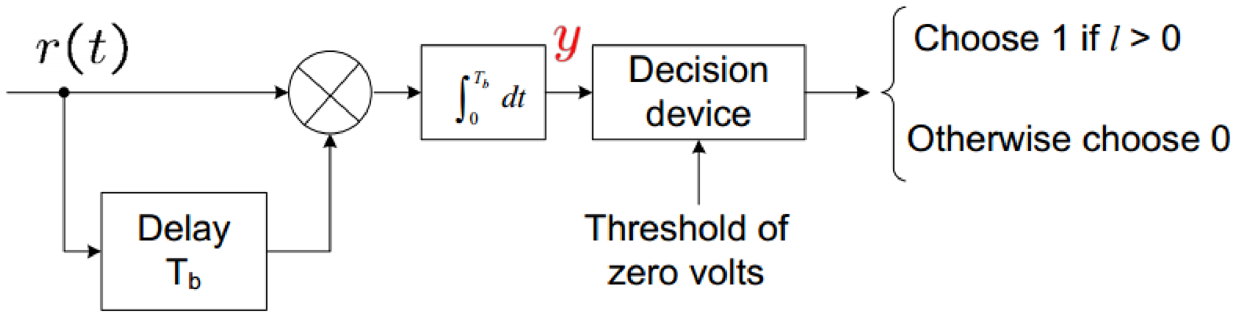
Binary digital modulation

- Differential PSK (**DPSK**)
  - ➢ Structure.



Binary digital modulation

- Differential PSK (**DPSK**)
  - Differential detection.



  - Output of integrator (assume noise free)

$$y = \int_0^{T_b} r(t)r(t-T_b)dt = \int_0^{T_b} \cos(w_c t + \psi_k + \theta)\cos(w_c t + \psi_{k-1} + \theta)dt$$

$$\propto \cos(\psi_k - \psi_{k-1})$$

  - The unknown phase θ becomes irrelevant. The decision becomes: if $\psi_k - \psi_{k-1} = 0$ (bit 1), then y>0; if $\psi_k - \psi_{k-1} = \pi$ (bit 0), then y<0

$$P_e = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$$

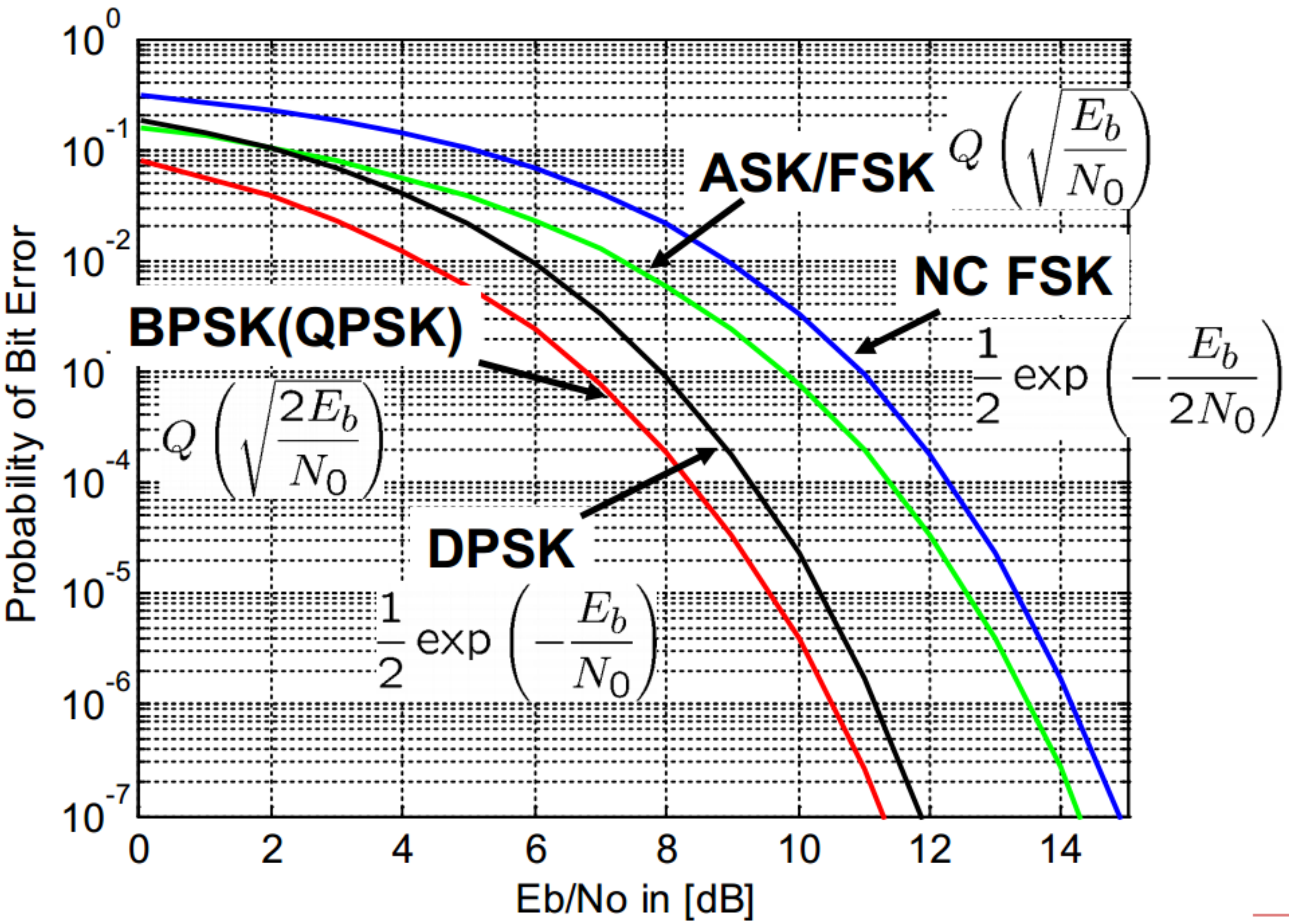
Binary digital modulation

- Comparison

Coherent PSK	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
Coherent ASK	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
Coherent FSK	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
Non-Coherent FSK	$\frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right)$
DPSK	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$

Binary digital modulation

- Comparison



M-ary digital modulation

- Why?

M-ary digital modulation

- In binary data transmission, send only one of two possible signals during each bit interval T_b
- In M-ary data transmission, send one of M possible signals during each signaling interval T
- In almost all applications, $M=2^n$ and $T=nT_b$, where n is an integer
- Each of the M signals is called a **symbol**
- These signals are generated by changing the amplitude, phase, frequency, or combined forms of a carrier in M discrete steps.
- Thus, we have MASK, MPSK, MFSK, and **MQAM**

M-ary digital modulation

- M-ary Phase-shift Keying (**MPSK**)
  - Modulation: The phase of the carrier takes on M possible values

$$\theta_m = 2\pi(m-1)/M, \;\; m = 1, \ldots, M$$

  - Signal set

$$s_m(t) = \sqrt{\frac{2E_s}{T}} \cos\left[2\pi f_c t + \frac{2\pi(m-1)}{M}\right] \qquad \begin{array}{c} m = 1, \ldots, M \\ 0 \le t < T \end{array}$$

  - Es=Energy per symbol, $f_c >> \frac{1}{T}$
  - Basis functions

$$\begin{aligned} \phi_1(t) &= \sqrt{\frac{2}{T}} \cos(2\pi f_c t) \\ \phi_2(t) &= \sqrt{\frac{2}{T}} \sin(2\pi f_c t) \end{aligned} \qquad 0 \le t < T$$

M-ary digital modulation

- M-ary Phase-shift Keying (**MPSK**)
  - Signal space representation.

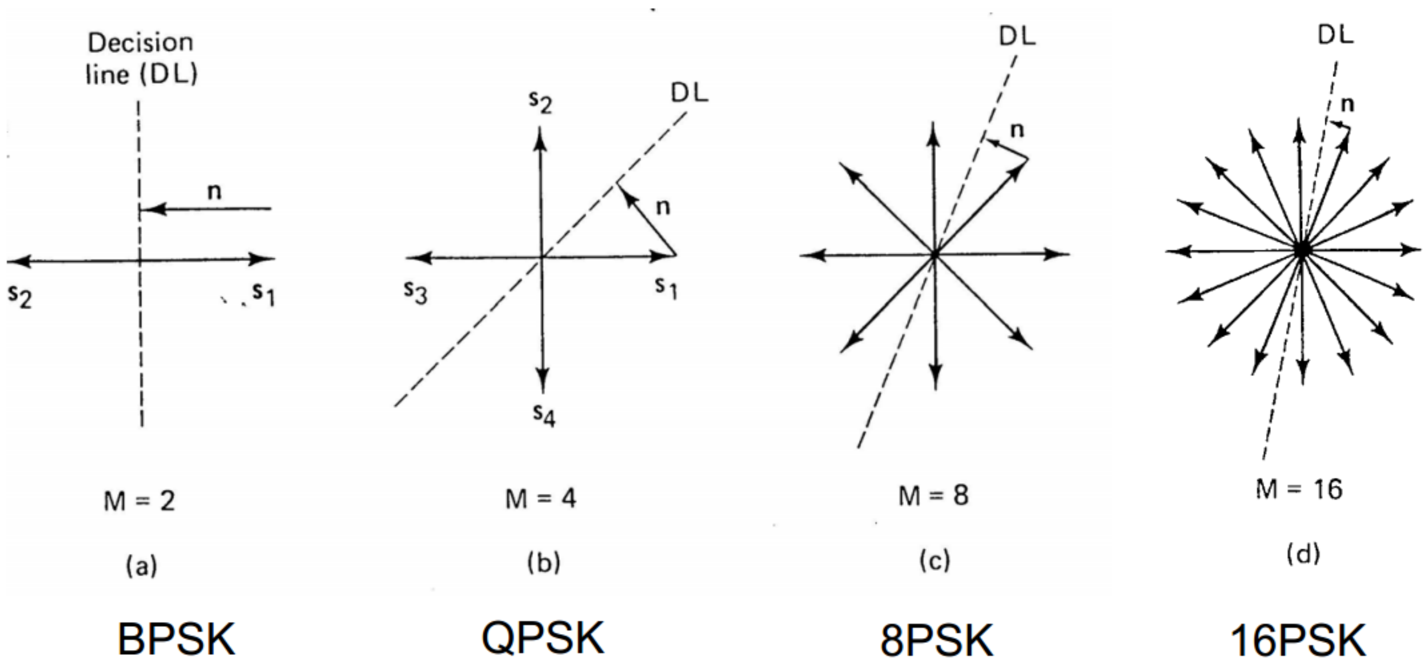
$$\begin{aligned}
s_m(t) &= \sqrt{\frac{2E_s}{T}}\cos\left[2\pi f_c t + \frac{2\pi(m-1)}{M}\right] \\
&= \sqrt{\frac{2E_s}{T}}\cos(2\pi f_c t)\cos\left[\frac{2\pi(m-1)}{M}\right] \\
&\quad -\sqrt{\frac{2E_s}{T}}\sin(2\pi f_c t)\sin\left[\frac{2\pi(m-1)}{M}\right] \\
&= \sqrt{E_s}\cos\left[\frac{2\pi(m-1)}{M}\right]\phi_1(t) - \sqrt{E_s}\sin\left[\frac{2\pi(m-1)}{M}\right]\phi_2(t)
\end{aligned}$$

$$\Rightarrow \quad \mathbf{s}_m = \left[\ \sqrt{E_s}\cos\left(\frac{2\pi(m-1)}{M}\right) \quad \sqrt{E_s}\sin\left(\frac{2\pi(m-1)}{M}\right)\ \right]$$

$$m = 1, \ldots, M$$

M-ary digital modulation

- M-ary Phase-shift Keying (**MPSK**)
  - Signal constellations.



M-ary digital modulation

- M-ary Phase-shift Keying (**MPSK**)
  - Euclidean distance

$$d_{mn} = \left\| \mathbf{s}_m - \mathbf{s}_n \right\| = \sqrt{2E_s \left(1 - \cos \frac{2\pi(m-n)}{M} \right)}$$

  - The minimum Euclidean

$$d_{\min} = \sqrt{2E_s \left(1 - \cos \frac{2\pi}{M} \right)} = 2\sqrt{E_s} \sin \frac{\pi}{M}$$

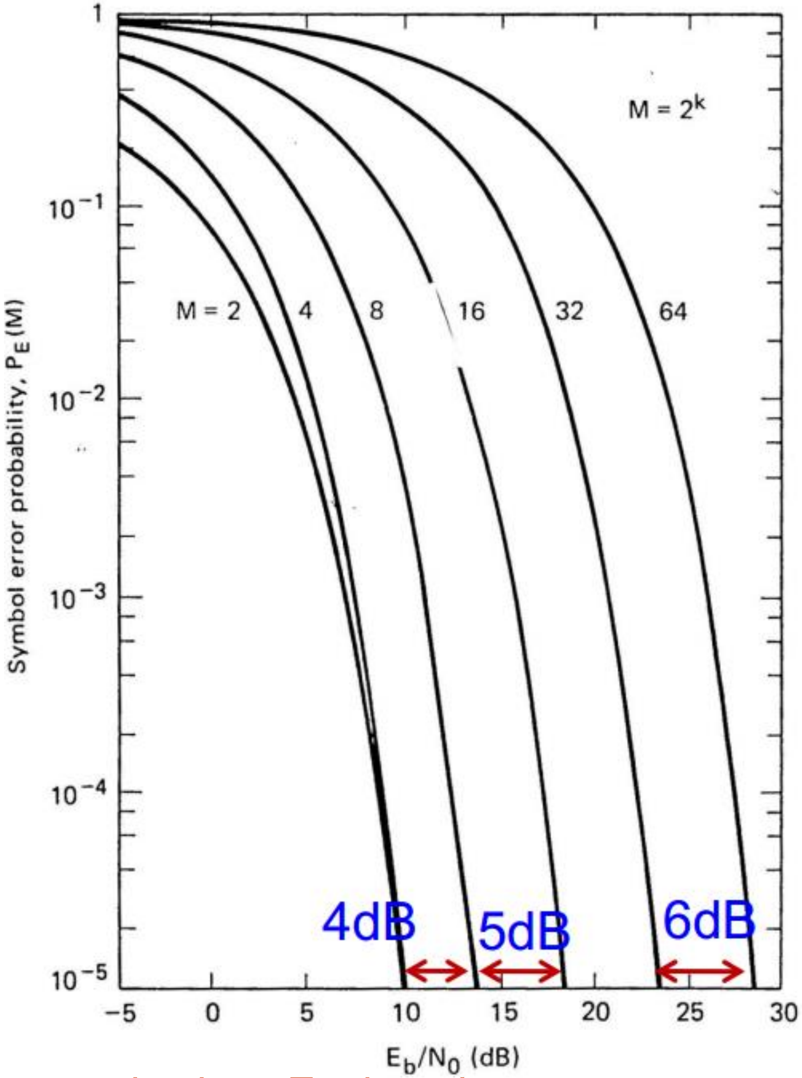
  - $d_{\min}$ plays an important role in determining error performance as discussed previously (union bound)
  - In the case of PSK modulation, the error probability is dominated by the erroneous selection of either one of the two signal points adjacent to the transmitted signal point
  - Consequently, an approximation to the symbol error probability is $P_{MPSK} \approx 2Q\left(\frac{d_{\min}/2}{\sqrt{N_0/2}} \right) = 2Q\left(\sqrt{\frac{2E_s}{N_0}} \sin \frac{\pi}{M} \right)$

M-ary digital modulation

- M-ary Phase-shift Keying (**MPSK**)
  - Exercise: Consider the M=2, 4, 8 PSK signal constellations. All have the same transmitted signal energy Es.
  - Determine the minimum distance $d_{\min}$ between adjacent signal points
  - For M=8, determine by how many dB the transmitted signal energy Es must be increased to achieve the same $d_{\min}$ as M=4.

M-ary digital modulation

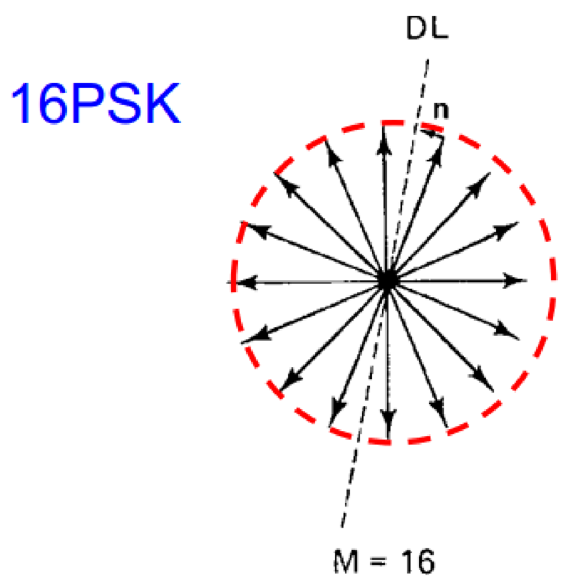
- M-ary Phase-shift Keying (**MPSK**)

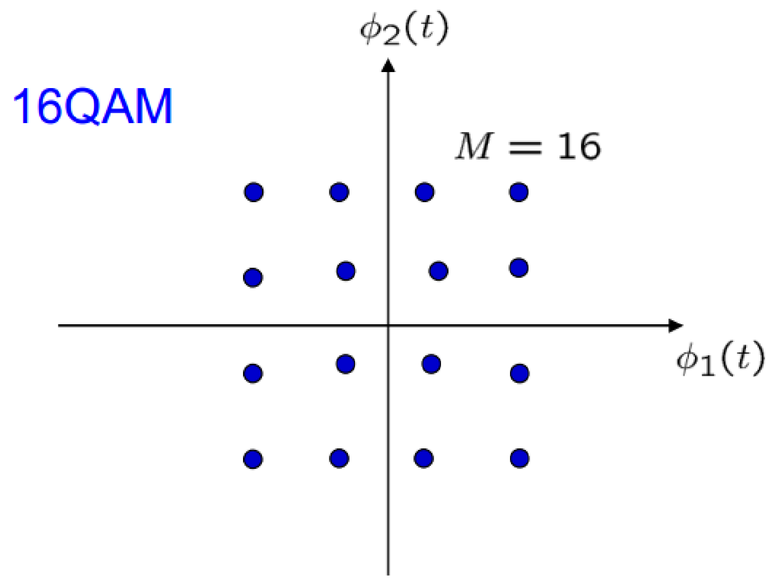


For large M, doubling the number of phases requires an additional 6 dB/bit to achieve the same performance

M-ary digital modulation

- M-ary Quadrature Amplitude Modulation (**MQAM**)
  - In MPSK, in-phase and quadrature components are interrelated in such a way that the envelope is constant (circular constellation)
  - If we relax this constraint, we get M-ary QAM





M-ary digital modulation

- M-ary Quadrature Amplitude Modulation (**MQAM**)
  - Modulation:

$$s_i(t) = \sqrt{\frac{2E_0}{T}} a_i \cos(2\pi f_c t) + \sqrt{\frac{2E_0}{T}} b_i \sin(2\pi f_c t)$$

  - E_0 is the energy of the signal with the lowest amplitude
  - a_i, b_i are a pair of independent integers
  - Basis functions

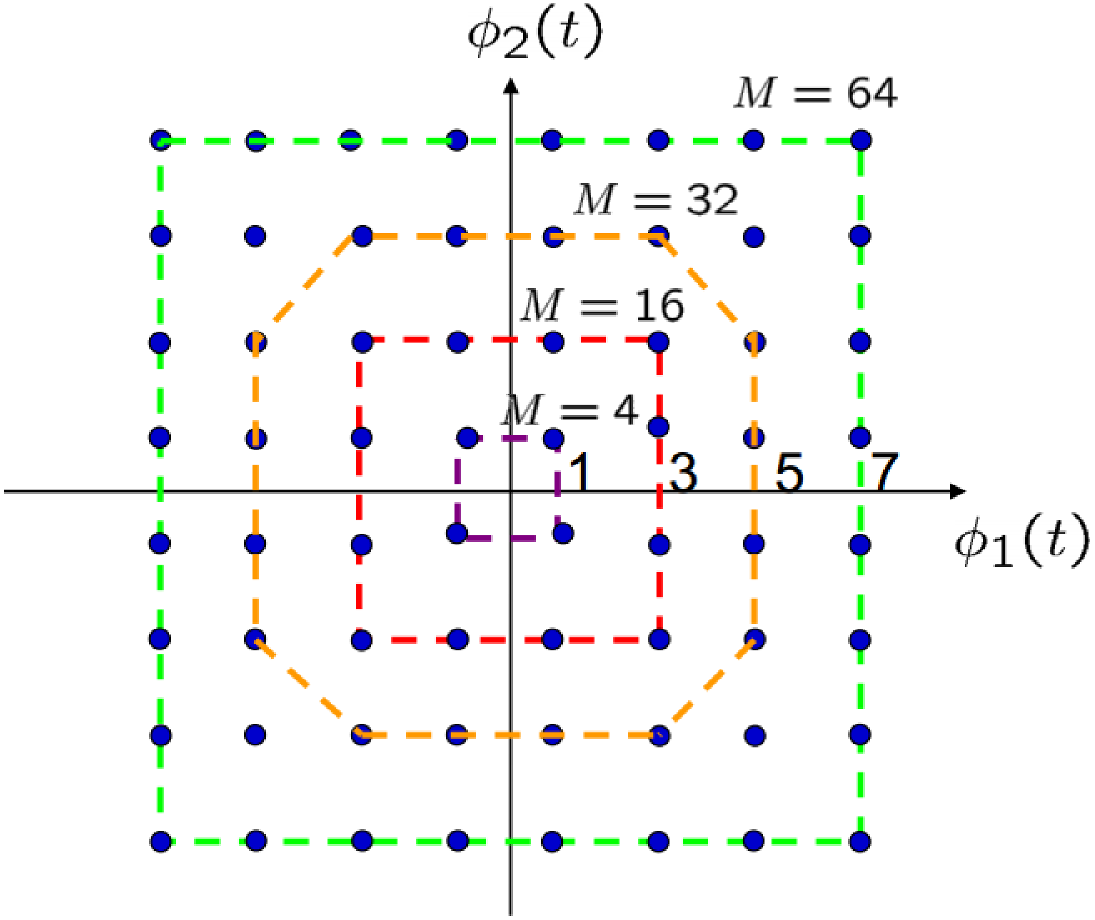
$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t) \quad \phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t) \quad 0 \le t < T$$

  - Signal space representation

$$\vec{s}_i = [\sqrt{E_0} a_i \quad \sqrt{E_0} b_i]$$

M-ary digital modulation

- M-ary Quadrature Amplitude Modulation (**MQAM**)
  - Signal constellation.



M-ary digital modulation

- M-ary Quadrature Amplitude Modulation (**MQAM**)
  - Probability of error analysis.
  - Upper bound of the symbol error probability

$$P_e \le 4Q\left(\sqrt{\frac{3kE_b}{(M-1)N_0}}\right) \qquad \text{(for } M = 2^k \text{)}$$

Think about the increase in Eb required to maintain the same error performance if the number of bits per symbol is increased from k to k+1, where k is large.

M-ary digital modulation

- M-ary Frequency-shift Keying (**MFSK**) (Multitone Signaling)
  - Signal set:

$$s_m(t) = \sqrt{\frac{2E_s}{T}} \cos\{2\pi(f_c + (m-1)\triangle f)t\} \qquad \begin{matrix} m = 1, \ldots, M \\ 0 \le t < T \end{matrix}$$

where $\triangle f = f_m - f_{m-1}$ with $f_m = f_c + m\triangle f$

  - Correlation between two symbols

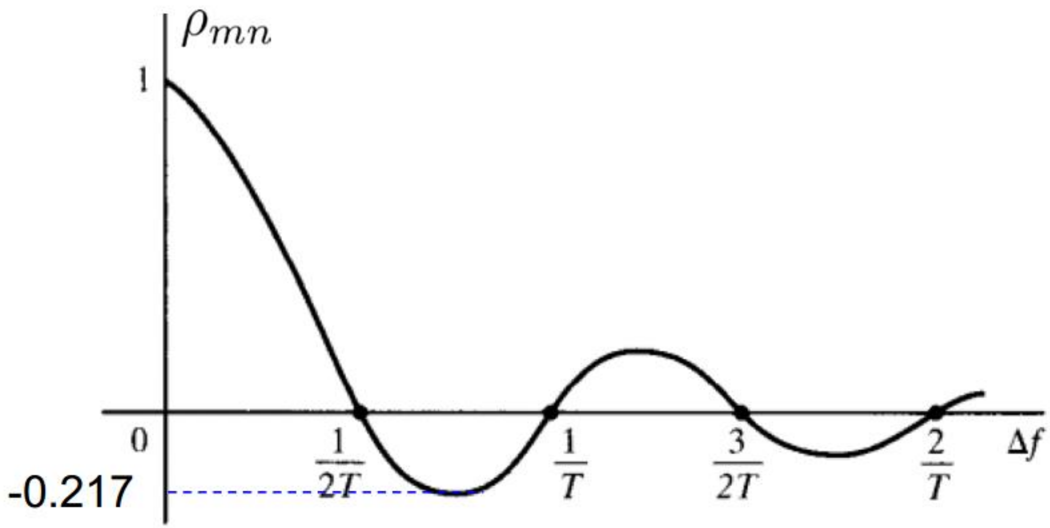
$$\rho_{mn} = \frac{1}{E_s} \int_0^T s_m(t) s_n(t) dt$$

$$= \frac{\sin[2\pi(m-n)\triangle f T]}{2\pi(m-n)\triangle f T}$$

$$= \mathrm{sinc}[2(m-n)\triangle f T]$$

M-ary digital modulation

- M-ary Frequency-shift Keying (**MFSK**) (Multitone Signaling)



➢ For orthogonality, the minimum frequency separation is

$$\Delta f = \frac{1}{2T}$$

M-ary digital modulation

- M-ary Frequency-shift Keying (**MFSK**) (Multitone Signaling)
  - Geometrical representation.

$$\mathbf{s}_0 = \left(\sqrt{E_s}, 0, 0, \cdots, 0\right)$$

$$\mathbf{s}_1 = \left(0, \sqrt{E_s}, 0, \cdots, 0\right)$$

$$\vdots$$

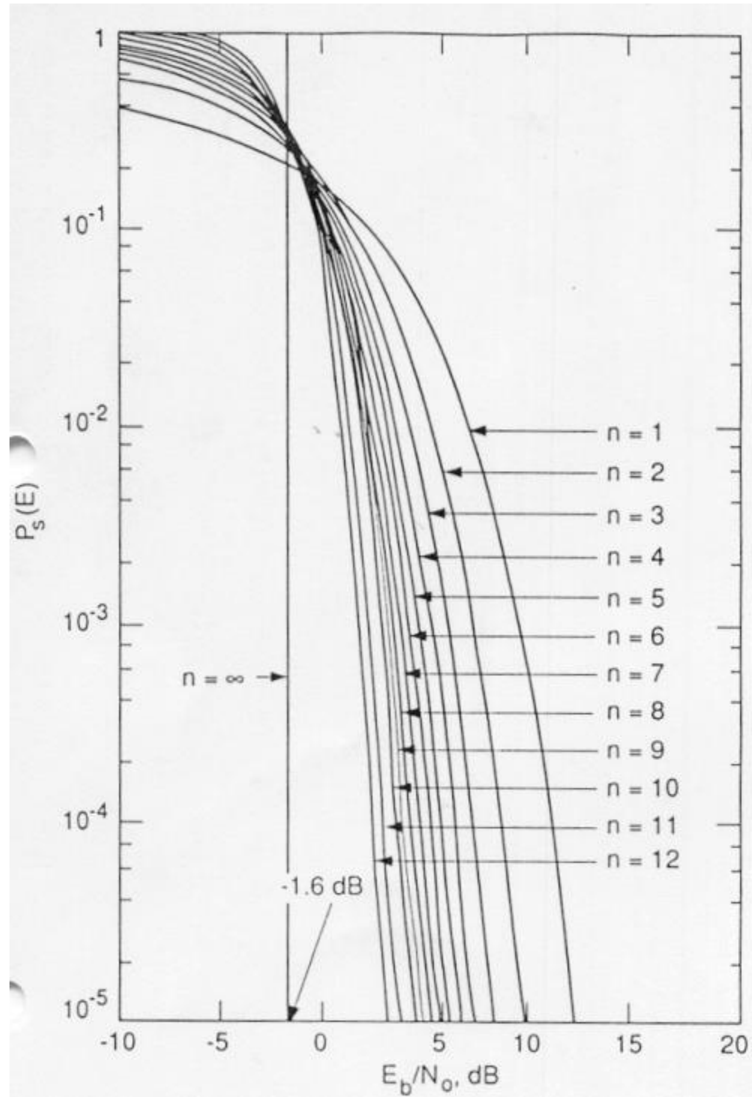
$$\mathbf{s}_{M-1} = \left(0, 0, \cdots, 0, \sqrt{E_s}\right)$$

  - Basis functions.

$$\phi_m = \sqrt{\frac{2}{T}} \cos 2\pi \left(f_c + m\Delta f\right) t$$

M-ary digital modulation

- M-ary Frequency-shift Keying (**MFSK**) (Multitone Signaling)
  - Probability of error.



M-ary digital modulation

- Notes
  - Pe is found by integrating conditional probability of error over the decision region, which is difficult to compute but can be simplified using union bound
  - Pe depends only on the distance profile of the signal constellation

M-ary digital modulation

- **Gray Code**
  - Symbol errors are different from bit errors
  - When a symbol error occurs, all k bits could be in error
  - In general, we can find BER using

$$P_b = \sum_{i=1}^{M} P(\vec{s}_i) \sum_{j=1, j\neq i}^{M} \frac{n_{i,j}}{\log_2 M} P\left(\hat{\vec{s}} = \vec{s}_j | \vec{s}_i\right)$$

n_{ij} the number of different bits between $\mathbf{s}_i$ and $\mathbf{s}_j$

  - Gray coding is a bit-to-symbol mapping, where two adjacent symbols differ in only one bit out of the k bits
  - An error between adjacent symbol pairs results in one and only one bit error

M-ary digital modulation

- **Gray Code**

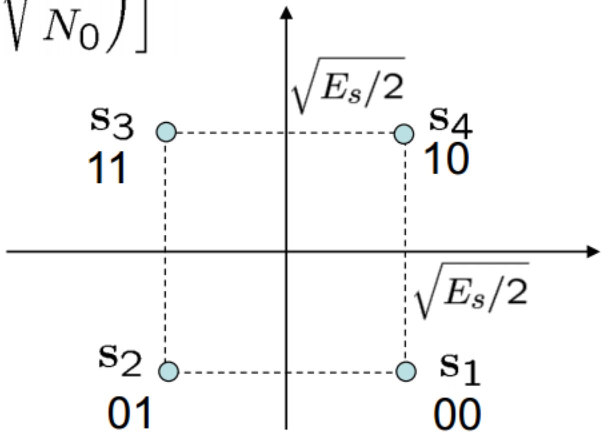
$$P_b = \sum_{i=1}^{M} \frac{1}{4} \sum_{j=1, j\neq i}^{M} \frac{n_{i,j}}{\log_2 M} P\left(\hat{\vec{s}} = \vec{s}_j | \vec{s}_i\right)$$

$$= \frac{1}{2} P\left(\hat{\vec{s}} = \vec{s}_1 | \vec{s}_4\right) + \frac{2}{2} P\left(\hat{\vec{s}} = \vec{s}_2 | \vec{s}_4\right) + \frac{1}{2} P\left(\hat{\vec{s}} = \vec{s}_3 | \vec{s}_4\right)$$

$$= \left[1 - Q\left(\sqrt{\frac{E_s}{N_0}}\right)\right] \cdot Q\left(\sqrt{\frac{E_s}{N_0}}\right) + \left[Q\left(\sqrt{\frac{E_s}{N_0}}\right)\right]^2$$

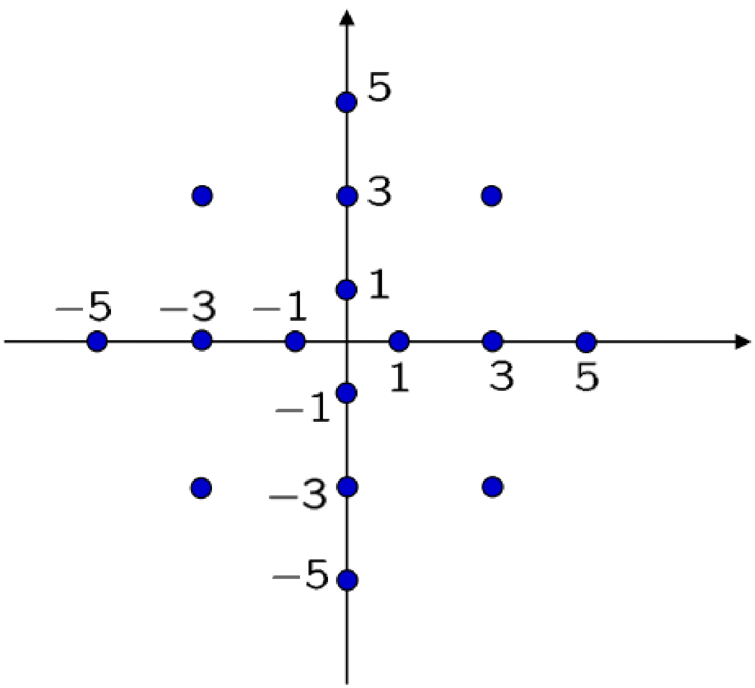
$$= Q\left(\sqrt{\frac{E_s}{N_0}}\right)$$

$$= Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$



M-ary digital modulation

- Example
  - The 16-QAM signal constellation shown right is an international standard for telephone-line modems (called V.29)
  - Determine the optimum decision boundaries for the detector
  - Derive the union bound of the probability of symbol error assuming that the SNR is sufficiently high so that errors only occur between adjacent points
  - Specify a Gray code for this 16-QAM V.29 signal constellation



M-ary digital modulation

- **Gray Code**
  - For MPSK with Gray coding, we know that an error between adjacent symbols will most likely occur. Thus, bit error probability can be approximated by

$$P_b \approx \frac{P_e}{\log_2 M}$$

  - For MFSK, when an error occurs, anyone of the other symbols may result equally likely. Thus, k/2 bits every k bits will on average be in error when there is a symbol error. The bit error rate is approximately half of the symbol error rate

$$P_b \cong \frac{1}{2} P_e$$

Think about why MQAM is more preferrable?

M-ary digital modulation

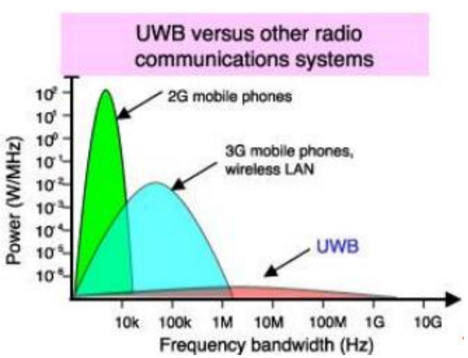
- Channel bandwidth and transmit power are two primary communication resources and have to be used as efficient as possible
  - Power utilization efficiency (energy efficiency): measured by the required E_b/N_0 to achieve a certain bit error probability
  - Spectrum utilization efficiency (bandwidth efficiency): measured by the achievable data rate per unit bandwidth R_b/B
- It is always desired to maximize bandwidth efficiency at a minimal required E_b/N_0

M-ary digital modulation

- Consider for example you are a system engineer in Huawei/ZTE, designing a part of the communication systems. You are required to design a modulation scheme for three systems using MFSK, MPSK or MQAM only. State the modulation level M to be low, medium or high

An ultra-wideband system

- Large amount of bandwidth
- Band overlays with other systems
- Purpose: high data rate



A wireless remote control system

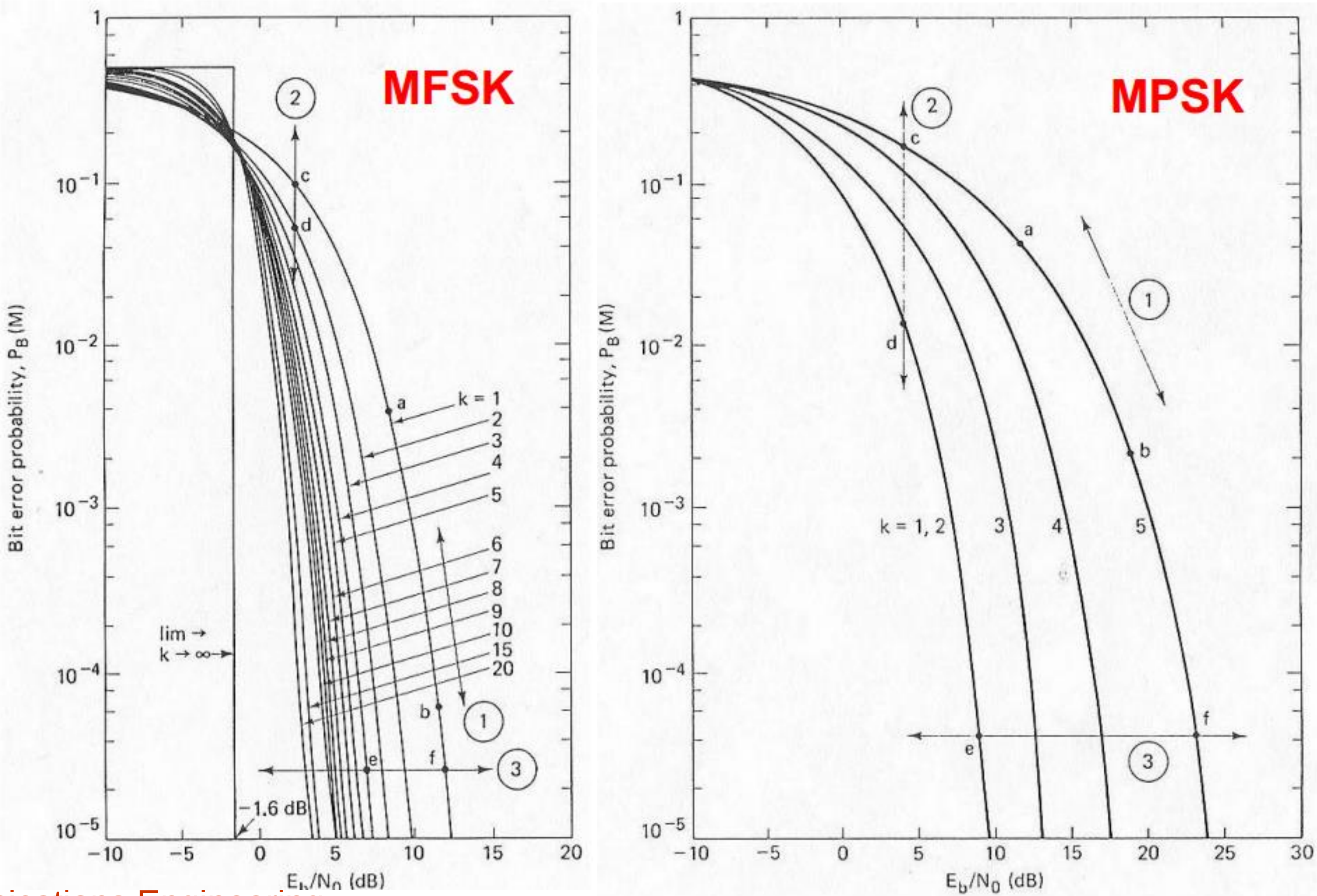
- Use unlicensed band
- Purpose: control devices remotely

A fixed wireless system

- Use licensed band
- Transmitter and receiver fixed with power supply
- Voice and data connections in rural areas

M-ary digital modulation

- Energy efficiency comparison



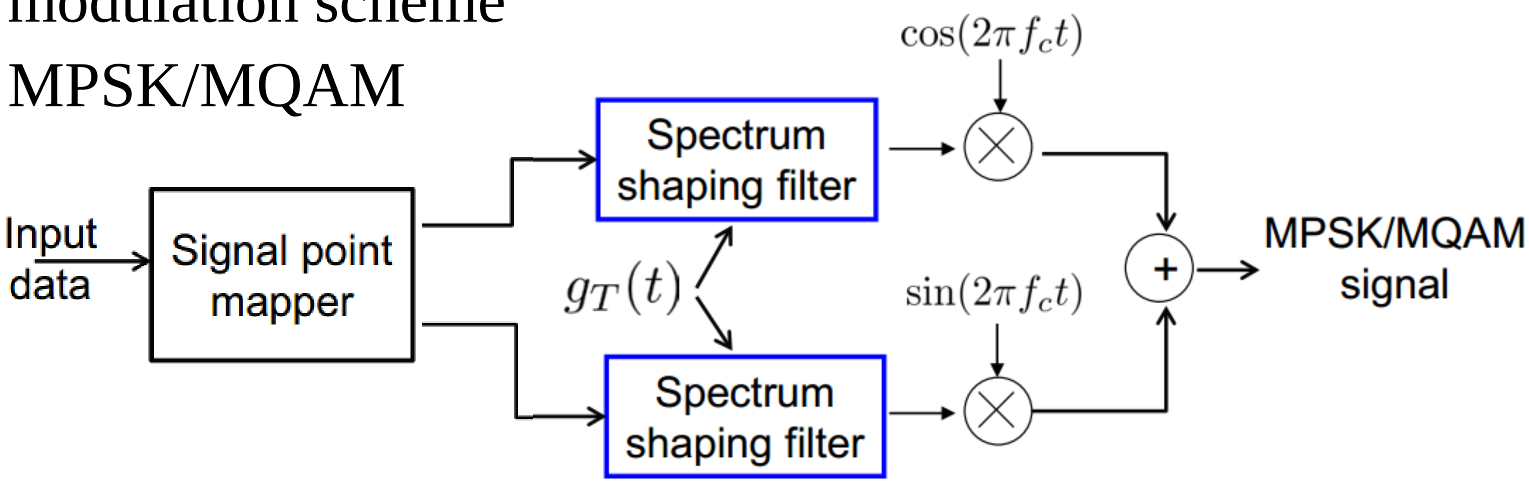
M-ary digital modulation

- Energy efficiency comparison
  - MFSK: At fixed E_b/N_0, increasing M can provide an improvement on P_b; At fixed P_b, increasing M can provide a reduction in the E_b/N_0
  - MPSK: BPSK and QPSK have the same energy efficiency. At fixed E_b/N_0, increasing M degrades P_b; At fxied P_b, increasing M increases the E_b/N_0 requirement

MFSK is more energy efficient than MPSK

M-ary digital modulation

- Bandwidth efficiency comparison
  - To compare bandwidth efficiency, we need to know the power spectral density (power spectra) of a given modulation scheme
  - MPSK/MQAM



  - If $g_T(t)$ is rectangular, the bandwidth of main-lobe is $B = \frac{2}{T_s}$
  - If it has a raised cosine spectrum, the bandwidth is $B = \frac{1+\alpha}{T_s}$

M-ary digital modulation

- Bandwidth efficiency comparison
  - In general, bandwidth required to pass MPSK/MQAM signal is approximately given by $B = \frac{1}{T_s}$.
  - The bit rate is $R_b = \frac{\log_2 M}{T_s}$
  - So the bandwidth efficiency may be expressed as

$$\boxed{\rho = \frac{R_b}{B} = \log_2 M \quad \text{(bits/sec/Hz)}}$$

  - But for MFSK, bandwidth required to transmit MSFK signal is $\boxed{B = \frac{M}{2T}}$
  - Bandwidth efficiency

$$\boxed{\rho = \frac{R_b}{B} = \frac{2\log_2 M}{M} \quad \text{(bits/s/Hz)}}$$

Adjacent frequencies need to be separated by 1/2T to maintain orthogonality

M-ary digital modulation

- Bandwidth efficiency comparison
  - In general, bandwidth required to pass MPSK/MQAM signal is approximately given by $B = \frac{1}{T_s}$.
  - The bit rate is $R_b = \frac{\log_2 M}{T_s}$
  - So the bandwidth efficiency may be expressed as

MPSK/MQAM is more bandwidth efficient than MFSK

signal is $\boxed{B = \frac{M}{2T}}$

  - Bandwidth efficiency

$$\boxed{\rho = \frac{R_b}{B} = \frac{2\log_2 M}{M} \quad \text{(bits/s/Hz)}}$$

Adjacent frequencies need to be separated by 1/2T to maintain orthogonality

M-ary digital modulation

- Fundamental tradeoff: Bandwidth Efficiency vs. Energy Efficiency
  - ➢ To see the ultimate power-bandwidth tradeoff, we need to use Shannon's channel capacity theorem:

Channel capacity is the theoretical upperbound for the maximum rate at which information could be transmitted without error (Shannon 1948)

  - ➢ Specifically, for a bandlimited channel corrupted by AWGN, the maximum achievable rate is given by

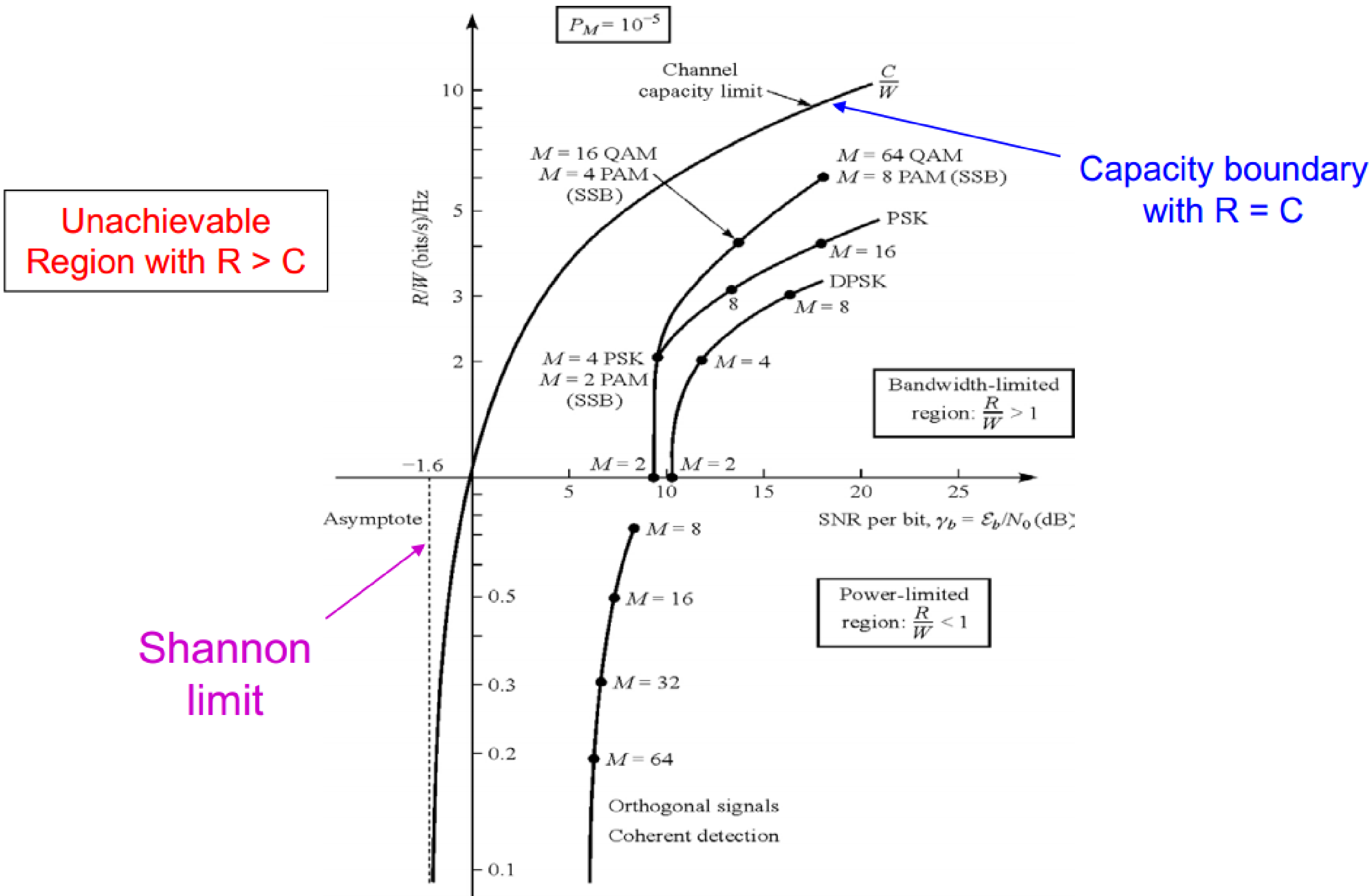
$$R \le C = B\log_2(1+SNR) = B\log_2\left(1+\frac{P_s}{N_0 B}\right)$$

  - ➢ Note that $\frac{E_b}{N_0} = \frac{P_s T}{N_0} = \frac{P_s}{RN_0} = \frac{P_s B}{RN_0 B} = SNR\frac{B}{R}$
  - ➢ Thus,

$$\frac{E_b}{N_0} = \frac{B}{R}(2^{R/B} - 1)$$

M-ary digital modulation

- Fundamental tradeoff: Bandwidth Efficiency vs. Energy Efficiency



M-ary digital modulation

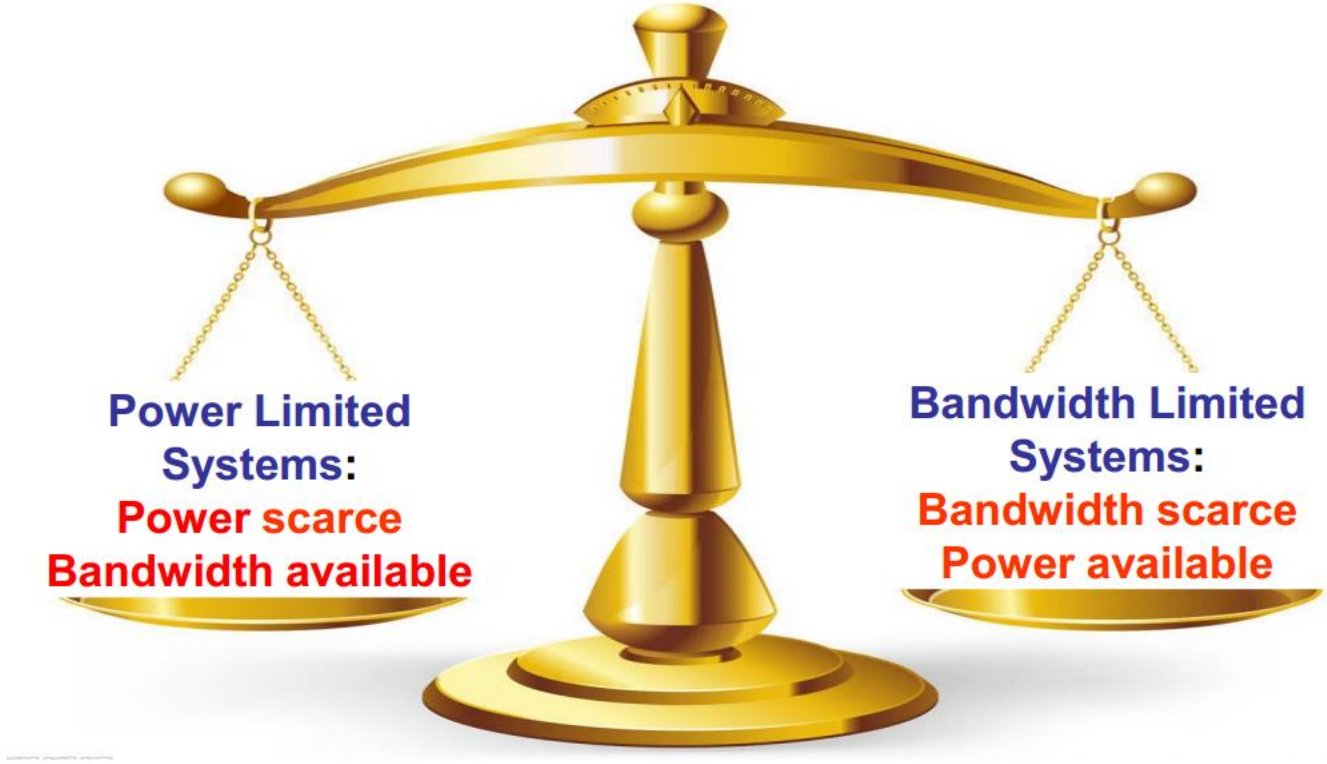
- Fundamental tradeoff: Bandwidth Efficiency vs. Energy Efficiency
  - In the limits as R/B goes to 0, we get

$$\frac{E_b}{N_0} = \ln 2 = 0.693 = -1.59dB$$

  - This value is called the Shannon limit. Received E_b/N_0 must be **>-1.59** dB to ensure reliable communication
  - BPSK and QPSK require the same E_b/N_0 of 9.6 dB to achieve $P_e=10^{-5}$. However, QPSK has a better bandwidth efficiency.
  - MQAM is superior to MPSK
  - MPSK/MQAM increases bandwidth efficiency at the cost of energy efficiency
  - MFSK trades energy efficiency at reduced bandwiidth efficiency

M-ary digital modulation

- Fundamental tradeoff: Bandwidth Efficiency vs. Energy Efficiency
  - ➢ Which modulation to use?

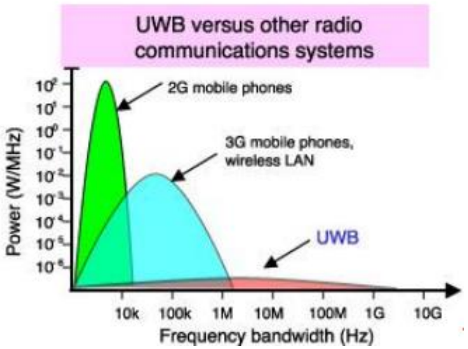


M-ary digital modulation

- Consider for example you are a system engineer in Huawei/ZTE, designing a part of the communication systems. You are required to design a modulation scheme for three systems using MFSK, MPSK or MQAM only. State the modulation level M to be low, medium or high

An ultra-wideband system

- Large amount of bandwidth
- Band overlays with other systems
- Purpose: high data rate



A wireless remote control system

- Use unlicensed band
- Purpose: control devices remotely

A fixed wireless system

- Use licensed band
- Transmitter and receiver fixed with power supply
- Voice and data connections in rural areas

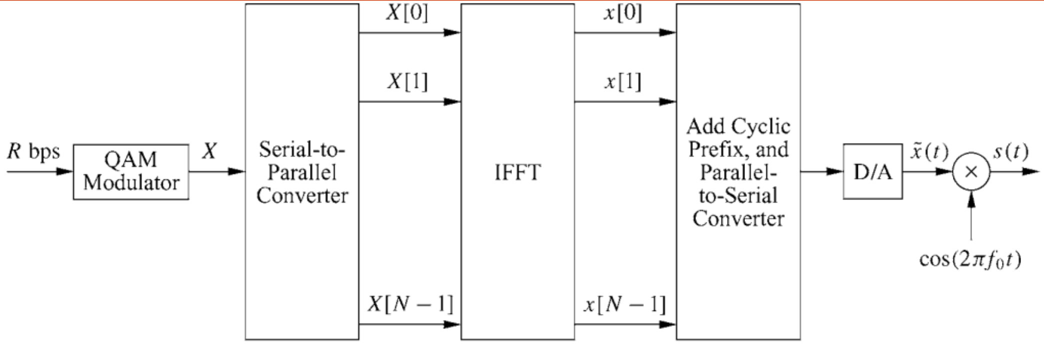
M-ary digital modulation

- Practical applications
  - BPSK: WLAN IEEE 802.11b (1 Mbps)
  - QPSK:
    1. WLAN IEEE 802.11b (2 Mbps, 5.5 Mbps, 11 Mbps)
    2. 3G WCDMA
    3. DVB-T (with OFDM)
  - QAM:
    1. Telephone modem (16-QAM)
    2. Downstream of Cable modem (64-QAM, 256-QAM)
    3. WLAN IEEE 802.11 a/g (16-QAM for 24 Mbps, 36 Mbps; 64-QAM for 38 Mbps and 54 Mbps)
    4. LTE cellular Systems
    5. 5G
  - FSK:
    1. Cordless telephone
    2. Paging system

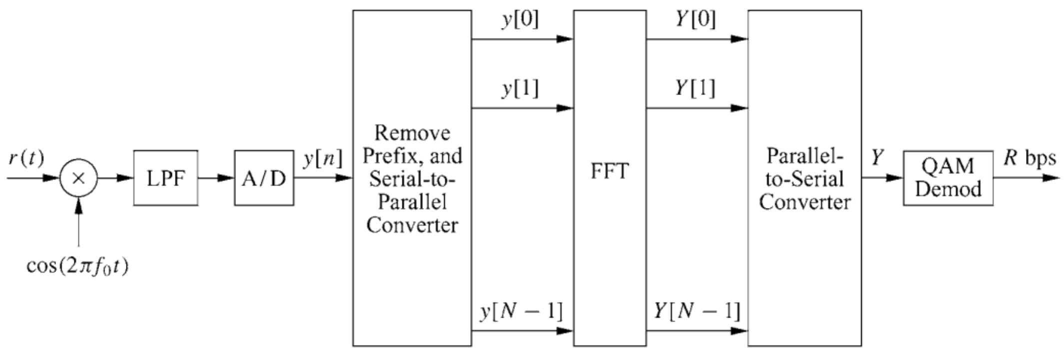
Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- **Multicarrier Communications & OFDM [Matlab]**
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- Synchronization

Multicarrier Communications & OFDM



Transmitter



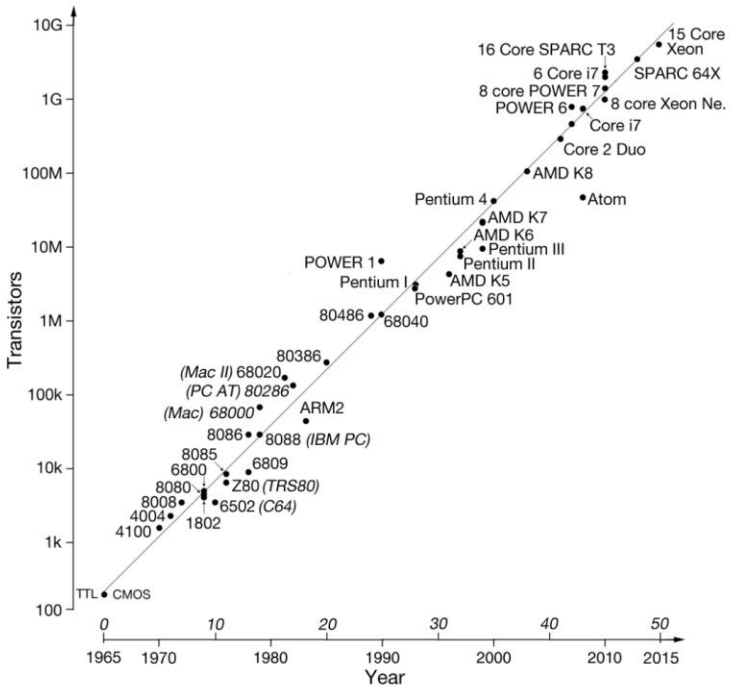
Receiver

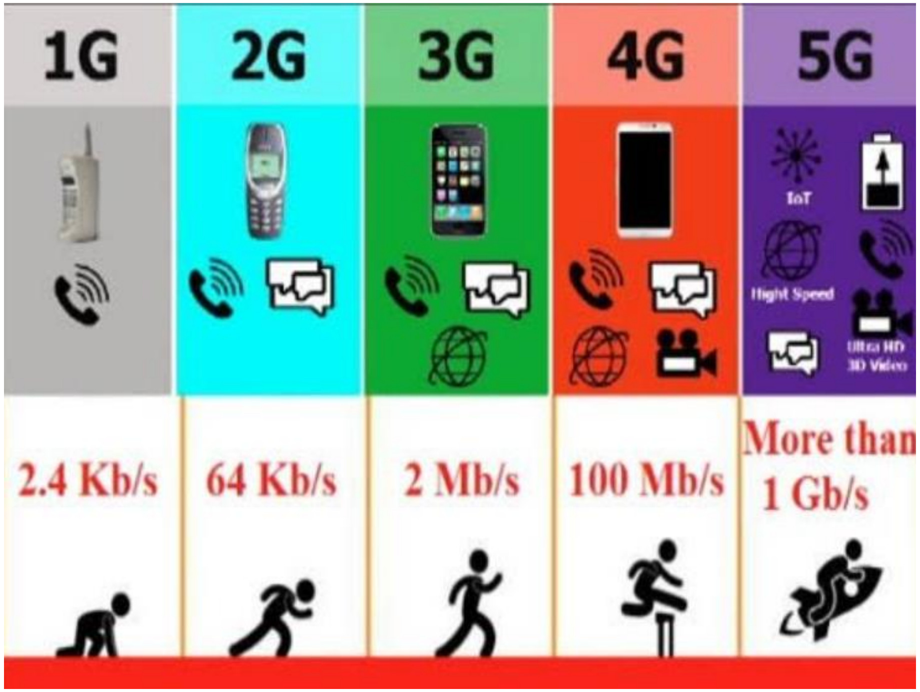
- Motivation
- Fading Channels
- Multicarrier modulation
- OFDM

Chapter 14.1-14.3、
Chapter 11

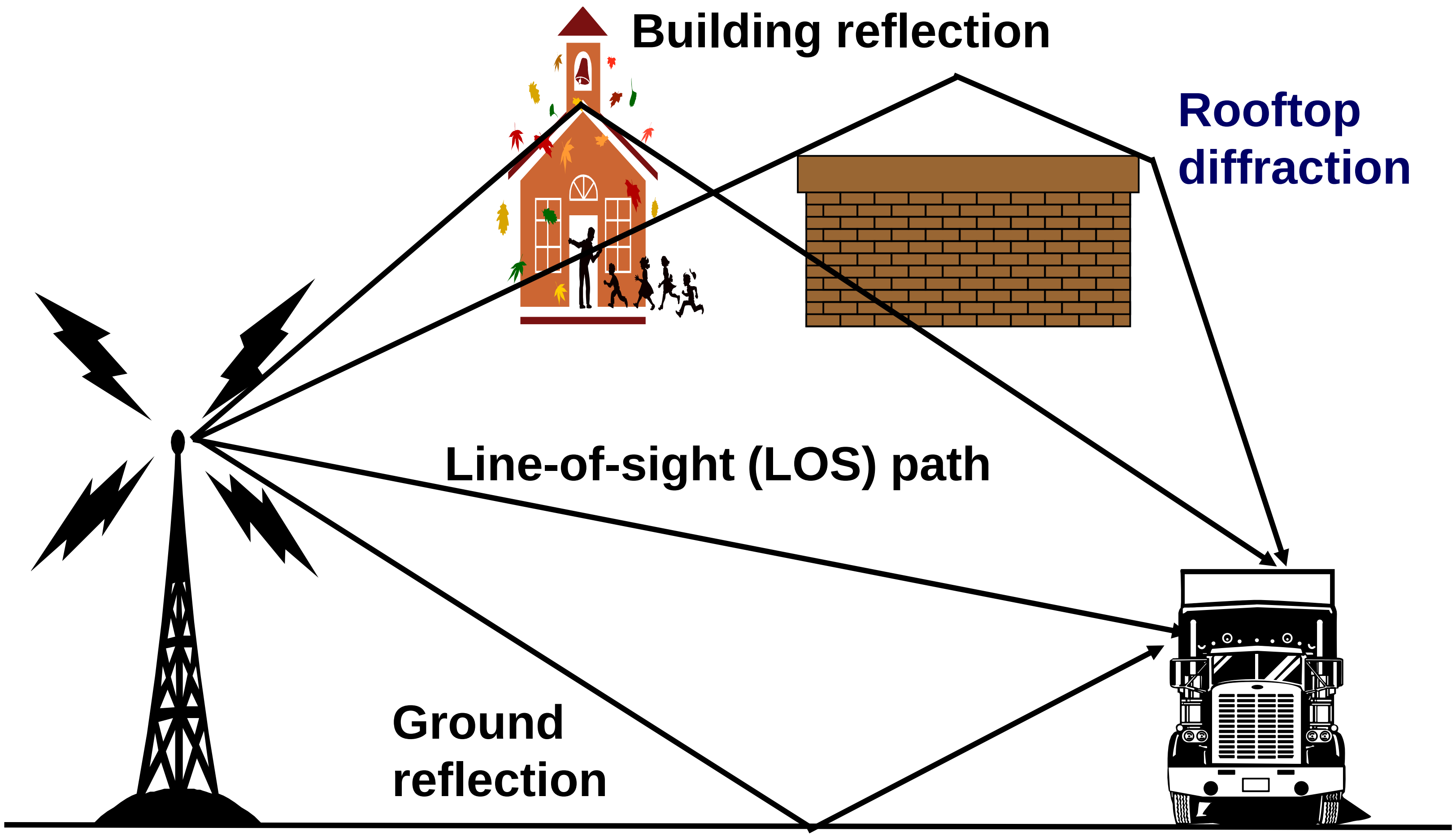
Motivation

- Practical channel impairments.
- Increasing demands high data rate transmission.
- Increasing power of DSP.

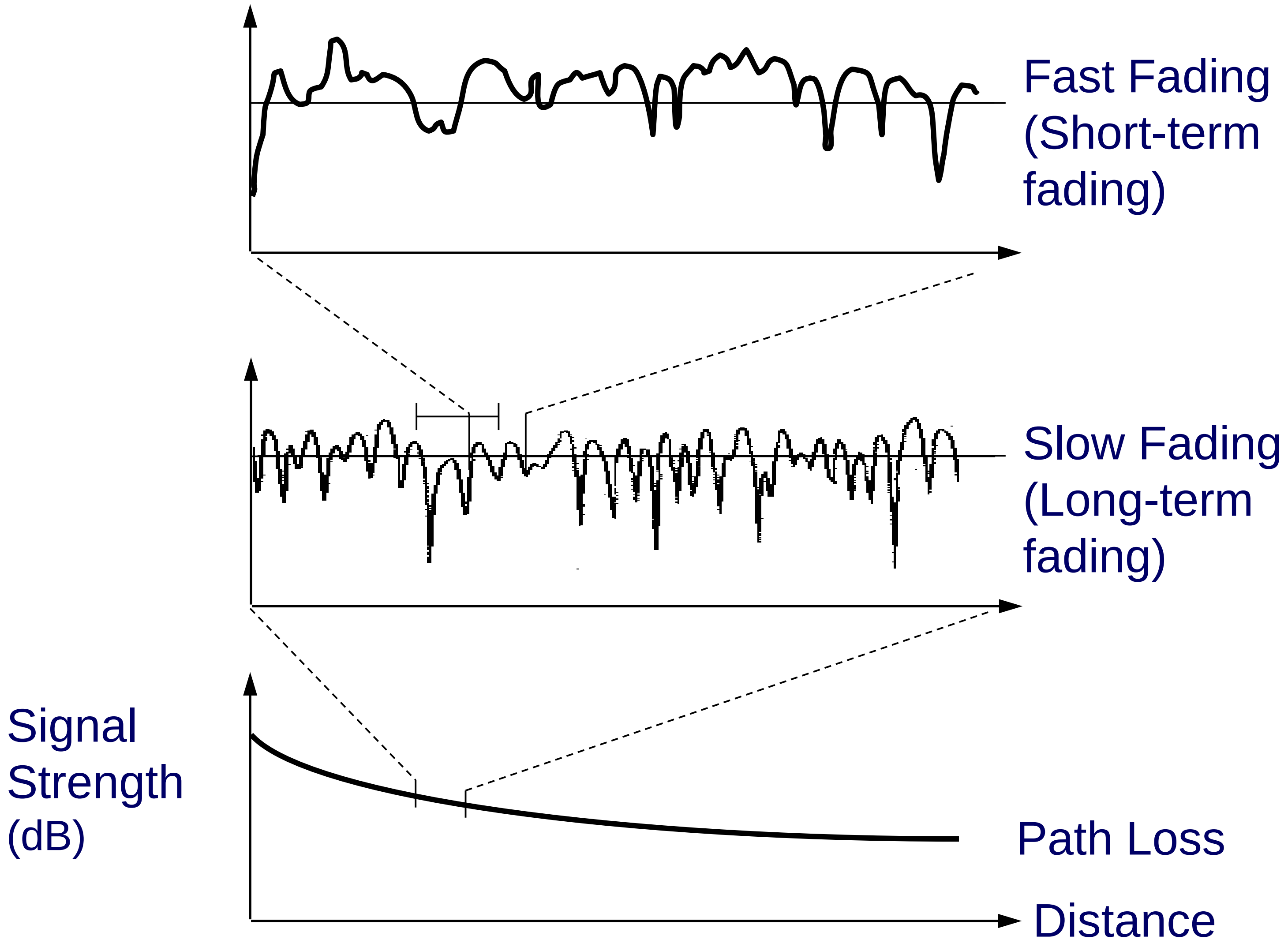




Fading Channel



Fading Channel



Fading Channel

- The received signal strength for the same distance from the transmitter will be different.
- The variation of the signal strength due to locations is often referred to as shadow fading or slow fading.
- Reason:
  - Often time, the fluctuations around the mean value are caused due to the signal being blocked from the receiver by buildings or walls and so on.
  - It is called slow fading because the variations are much slower with distance than another fading phenomenon caused due to multipath.

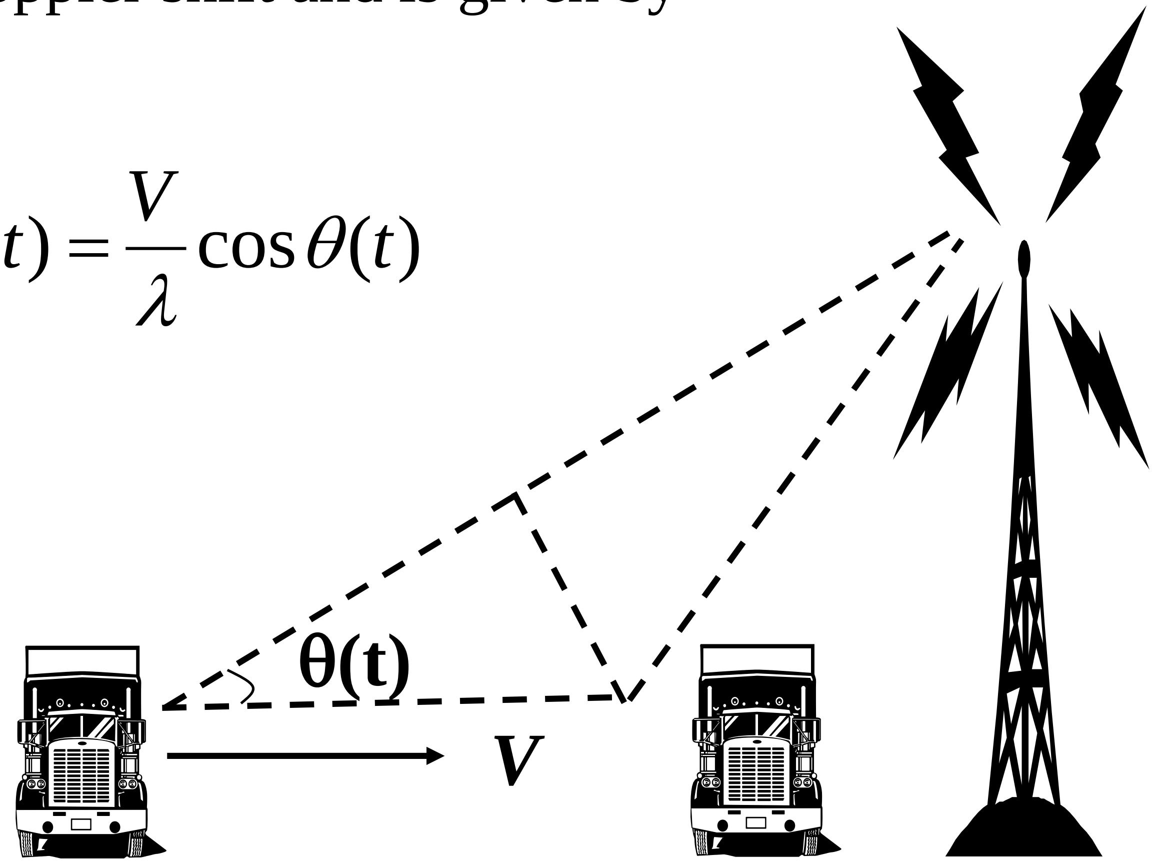
Fading Channel

- *Small-scale fading*: The received signal is rapidly fluctuating due to the mobility of the terminal causing changes in multiple signal components arriving via different paths.
- There are two effects which contribute to the rapid fluctuation of the signal amplitude.
  - **Multipath fading**: caused by the addition of signals arriving via different paths.
  - **Doppler**: caused by the movement of the mobile terminal toward or away from the base station transmitter.
- Small-scale fading results in very high bit error rates. It is not possible to simply increase the transmit power to overcome the problem
  - Error control coding, diversity schemes, directional antennas.

Fading Channel

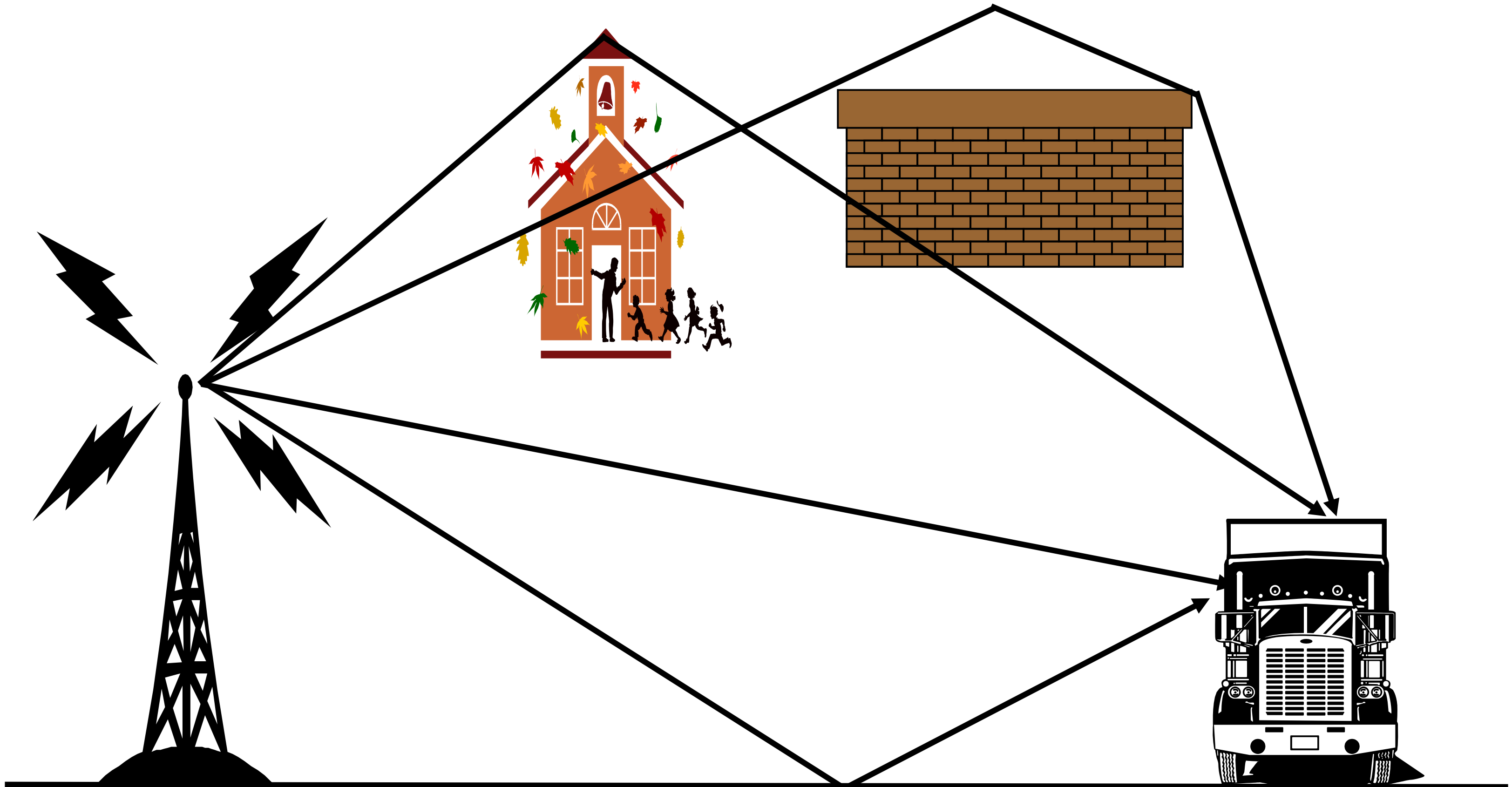
- BS transmits a single frequency *f*, the received signal at the MT at time *t* has a frequency of *f+v(t)*.
- *v(t)* is the Doppler shift and is given by

$$v(t)=\frac{V}{\lambda}\cos\theta(t)$$



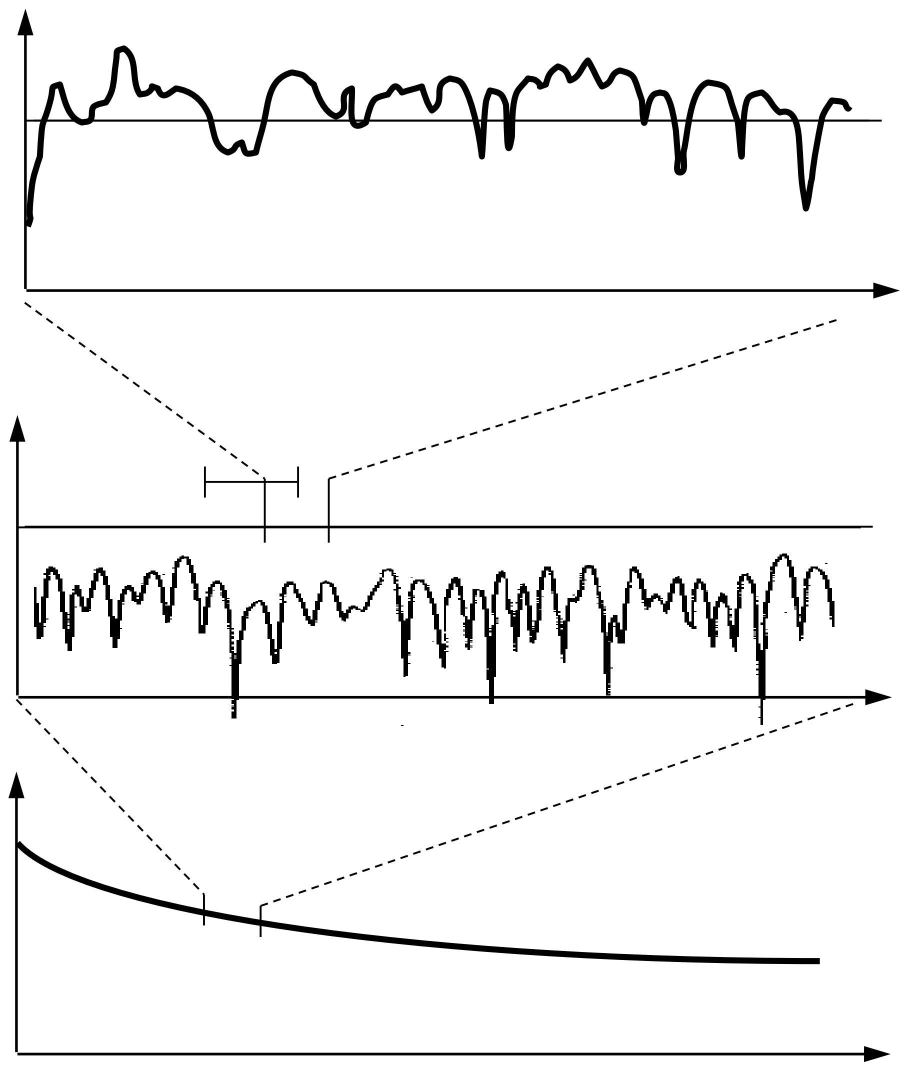
Fading Channel

- Results in fluctuations of the signal amplitude because of the addition of signals arriving with different *phases*.
- This phase difference is caused due to the fact that signals have traveled different distances by traveling along different *paths*.



Fading Channel

- **Narrowband fading**
  - Autocorrelation
  - Power spectral density
- Because of the phases of the arriving paths are changing rapidly, the received signal amplitude undergoes rapid fluctuation that is often modeled as a <u>*random variable*</u>.



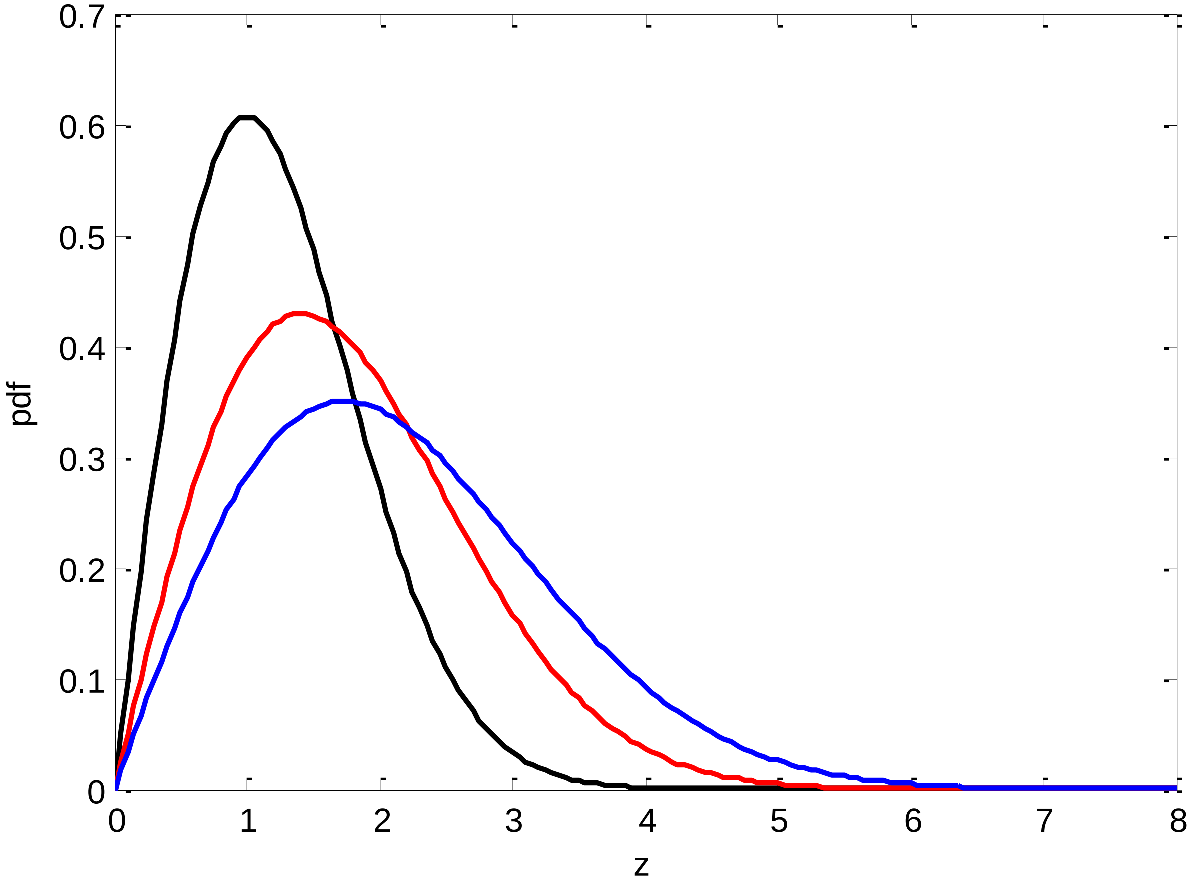
Fading Channel

- Rayleigh distribution (NLOS)
  - Most commonly used distribution for multipath fading (the envelope distribution of received signal) is Rayleigh distribution with pdf

$$f_{ray}(z) = \frac{2z}{P_r}\exp\left(-\frac{z^2}{P_r}\right) = \frac{z}{\sigma^2}\exp\left(-\frac{z^2}{2\sigma^2}\right), z \geq 0$$

  - Assume that all signals suffer nearly the same attenuation, but arrive with different phases.
  - σ^2 is the variance.
  - Middle value r_m of envelope signal within sample range to be satisfied by $P(r \leq r_m) = 0.5.$ We have $r_m = 1.777\,\sigma$.

Fading Channel



The pdf of the envelope variation

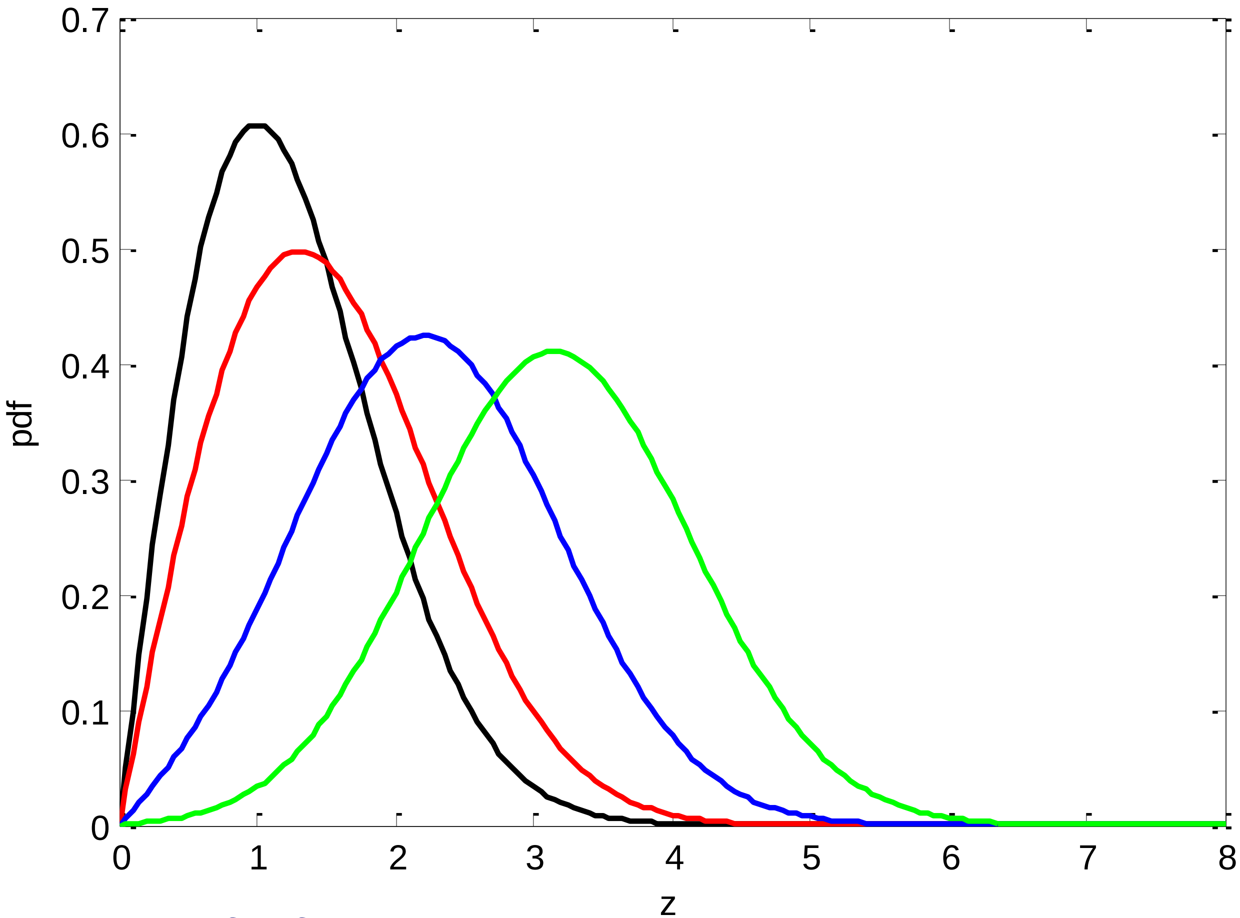
Fading Channel

- Ricean distribution (LOS – transmitter is close)
  - When a strong LOS signal component also exists, the pdf is given by

$$f_{ric}(z) = \frac{z}{\sigma^2} \exp\left(\frac{-(z^2+\alpha^2)}{2\sigma^2} \right) I_0\left(\frac{\alpha z}{\sigma^2} \right), z \geq 0, \alpha \geq 0$$

  - α is a factor that determines how strong the LOS component is relative to the rest of the multipath signals. If α=0, then it becomes Rayleigh distribution.
  - $I_0(x)$ is the zero-order modified Bessel function of the first kind.
  - $K = \frac{\alpha^2}{2\sigma^2}$, $K = 0 \Rightarrow Rayleigh, K = \infty \Rightarrow nofading$

$$f_{ric}(z) = \frac{2z(K+1)}{P_r} \exp\left(-K - \frac{(K+1)z^2}{P_r} \right) I_0\left(2z\sqrt{\frac{K(K+1)}{P_r}} \right), z \geq 0$$



The pdf of the envelope variation

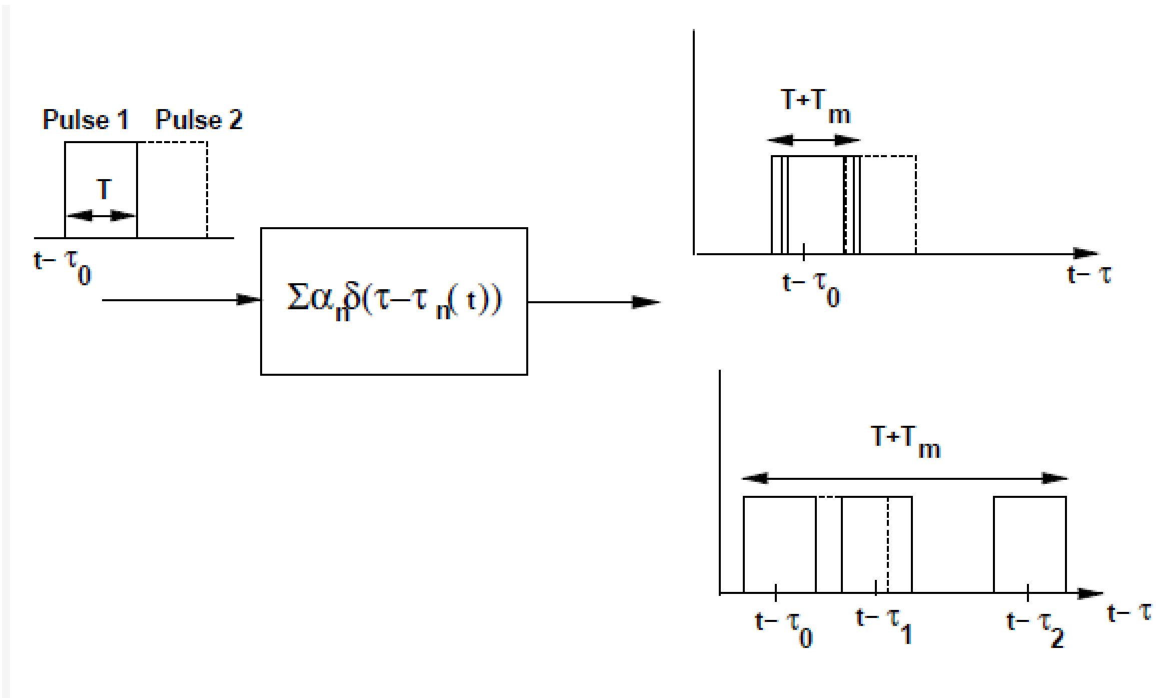
Fading Channel

- Nakagami-m fading (general model)
  - Parameters adjusted to fit a variety of empirical measurements

$$f_{naka}(z) = \frac{2m^m z^{2m-1}}{\Gamma(m) P_r^m} \exp\left(\frac{-mz^2}{P_r} \right) I_0\left(\frac{\alpha z}{\sigma^2} \right), m \geq 0.5$$

  - m=1, Rayleigh fading
  - m = $(K+1)^2/(2K+1)$, approximately Rician fading
  - m= ∞, no fading

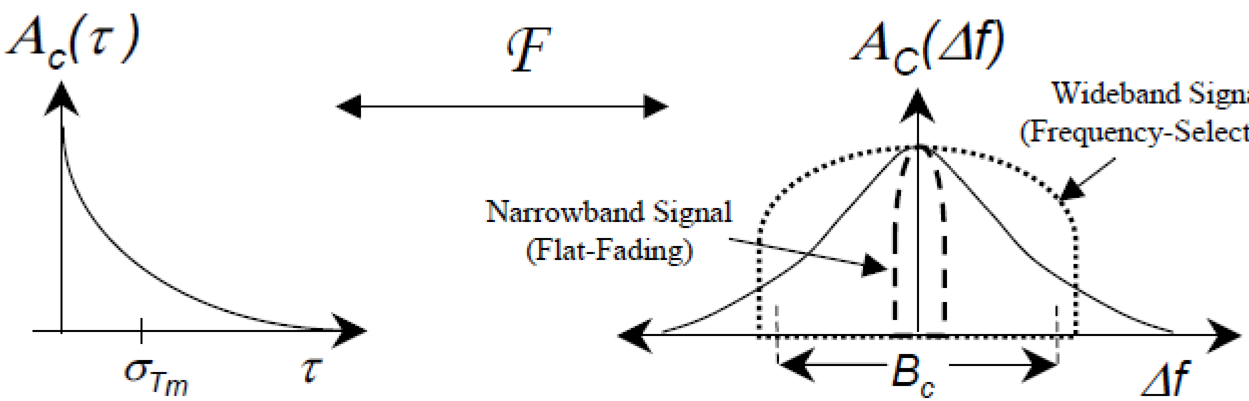
- **Wideband fading**
  - Multipath delay spread



  - Intersymbol interference (ISI)
  - Equalization, Multicarrier modulation, spread spectrum

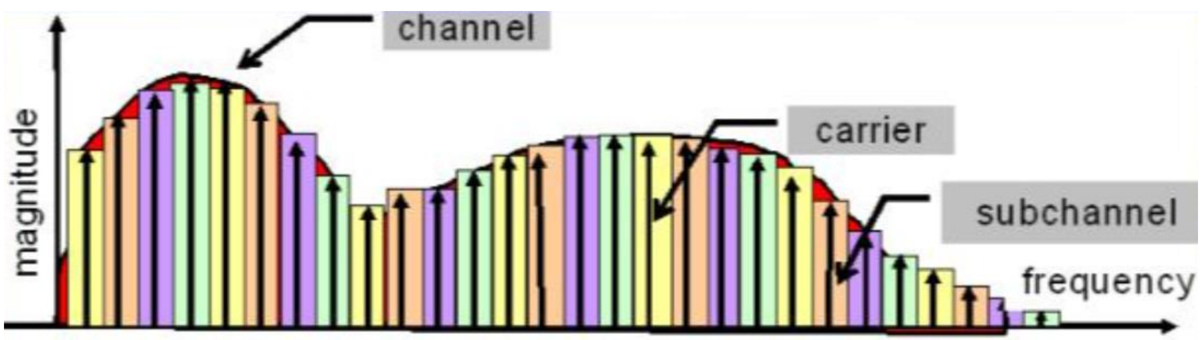
Fading Channel

- **Wideband fading**
  - Deterministic: deterministic scattering function
  - Random: scattering function
  - Important characterizations
    - Power delay profile (PDP)
    - Coherence bandwidth



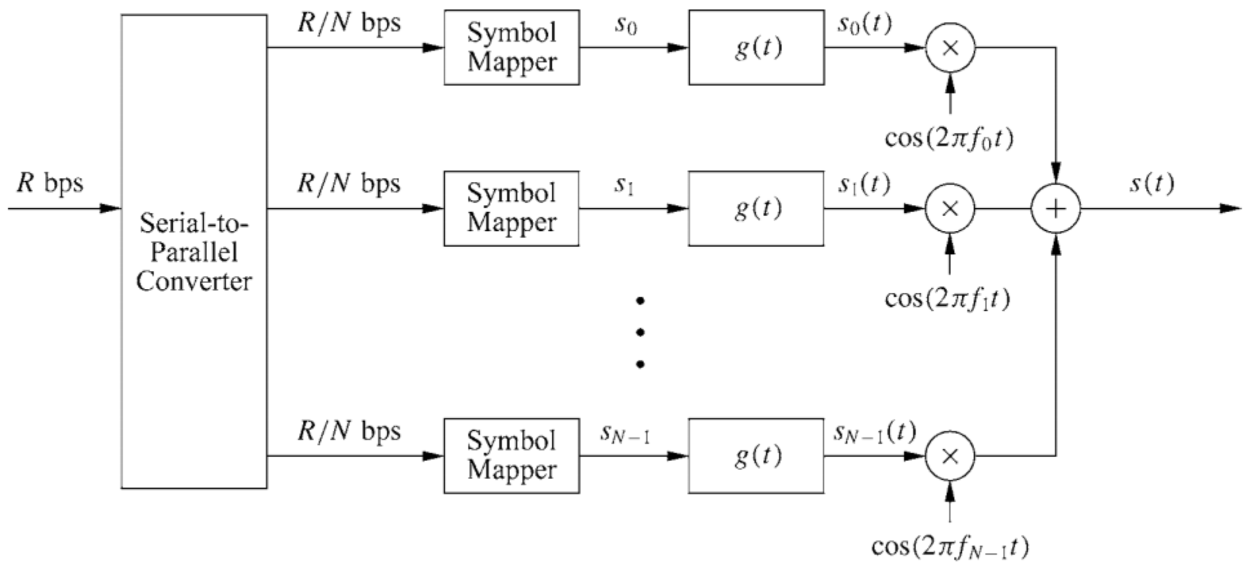
Multicarrier Modulation

- Techniques in which multiple frequencies are used to transmit data
- Divide broadband channel into $K = \frac{W}{\Delta f}$ narrowband subchannels
- Widely used in ADSL, Wireless LAN, and 4G/5G systems

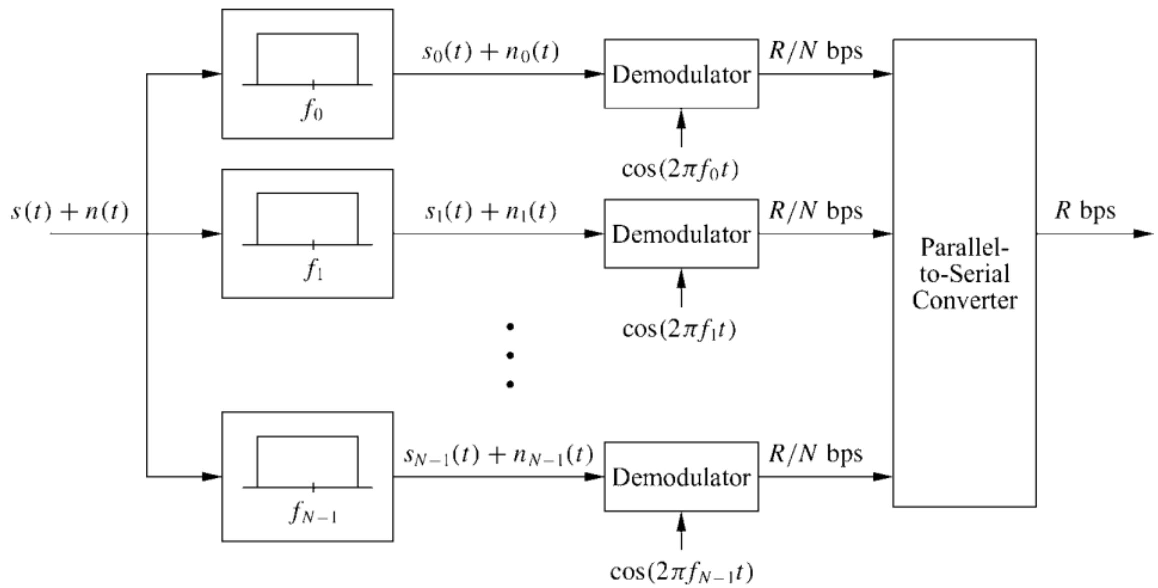


Multicarrier Modulation

Transmitter



Receiver



OFDM

- Review of Discrete Fourier Transform (DFT)
  - DFT of a discrete time sequence $x[n], 0 \le n \le N-1$

$$\mathrm{DFT}\{x[n]\} = X[i] \triangleq \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j2\pi ni/N}, \quad 0 \le i \le N-1.$$

  - x[n] can be recovered from its DFT using IDFT

$$\mathrm{IDFT}\{X[i]\} = x[n] \triangleq \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X[i] e^{j2\pi ni/N}, \quad 0 \le n \le N-1.$$

  - Circular convolution of x[n] and h[n]

$$y[n] = x[n] \circledast h[n] = h[n] \circledast x[n] = \sum_{k} h[k] x[n-k]_N,$$

$[n-k]_N$ denotes [n-k] modulo N, i.e., periodic version of x[n-k] with period N

  - Frequency domain property

$$\mathrm{DFT}\{y[n] = x[n] \circledast h[n]\} = X[i] H[i], \quad 0 \le i \le N-1,$$

Zero-padding for shorter sequence!

OFDM

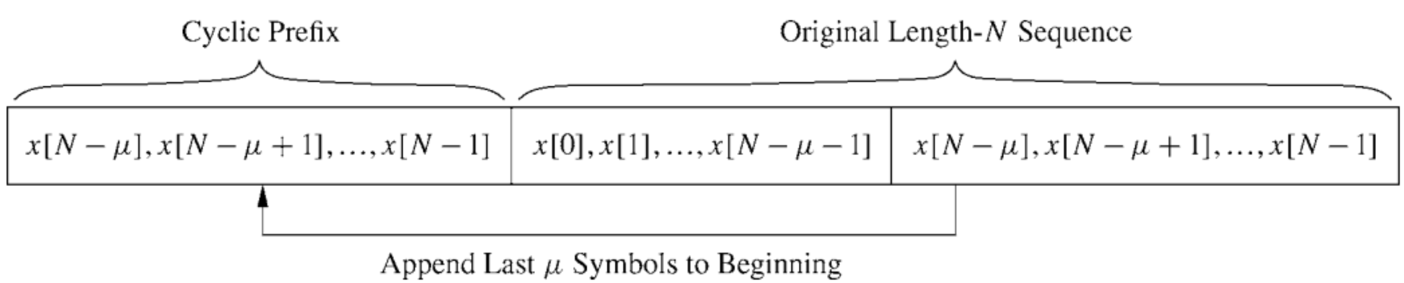
- Cyclic prefix (CP)
  - ➢ Consider input sequence $x[n] = x[0], \ldots, x[N-1]$ passes through a discrete-time channel with finite impulse response (FIR) $h[n] = h[0], \ldots, h[\mu]$, with $\mu + 1 = {}^{T_m}/_{T_s}$
  - ➢ The cyclic prefix for x[n] is defined as

$$\{x[N-\mu], \ldots, x[N-1]\}$$

i.e., the last μ samples are appended to the beginning of the sequence.

  - ➢ This yields a new sequence $\tilde{x}[n],\ -\mu \le n \le N-1$ given by

$$x[N-\mu], \ldots, x[N-1], x[0], \ldots, x[N-1]$$



OFDM

- Cyclic prefix (CP)
  - Suppose $\tilde{x}[n]$ is input to a discrete-time channel with impulse response h[n], the output y[n] is then given by

$$\begin{aligned} y[n] &= \tilde{x}[n] * h[n] \\ &= \sum_{k=0}^{\mu} h[k]\tilde{x}[n-k] \\ &= \sum_{k=0}^{\mu} h[k]x[n-k]_N \\ &= x[n] \circledast h[n], \quad 0 \le n \le N-1. \end{aligned}$$

where $0 \le k \le \mu, \tilde{x}[n-k] = x[n-k]_N$ is applied.

By appending a CP to the channel input, the linear convolution associated with channel output becomes a circular convolution!

OFDM

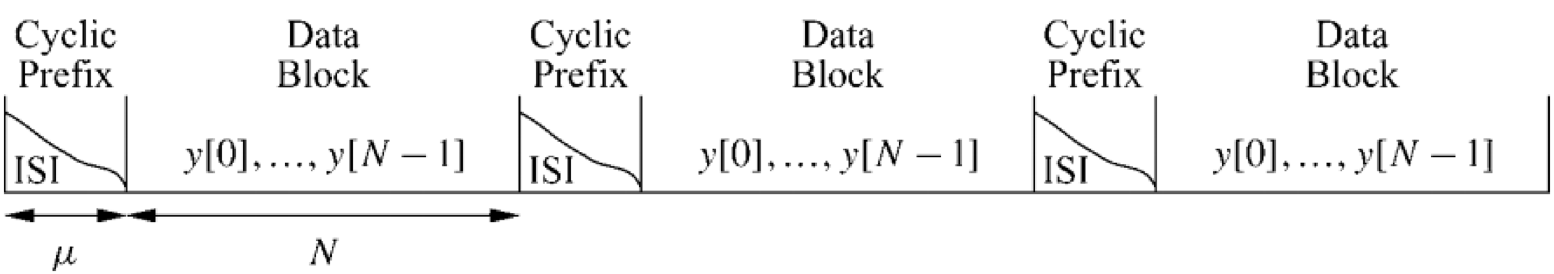
- Cyclic prefix (CP)
  - Taking the DFT of channel output in the absence of noise

$$Y[i] = \mathrm{DFT}\{y[n] = x[n] \circledast h[n]\} = X[i]H[i], \quad 0 \le i \le N-1,$$

  - The input sequence can be recovered from

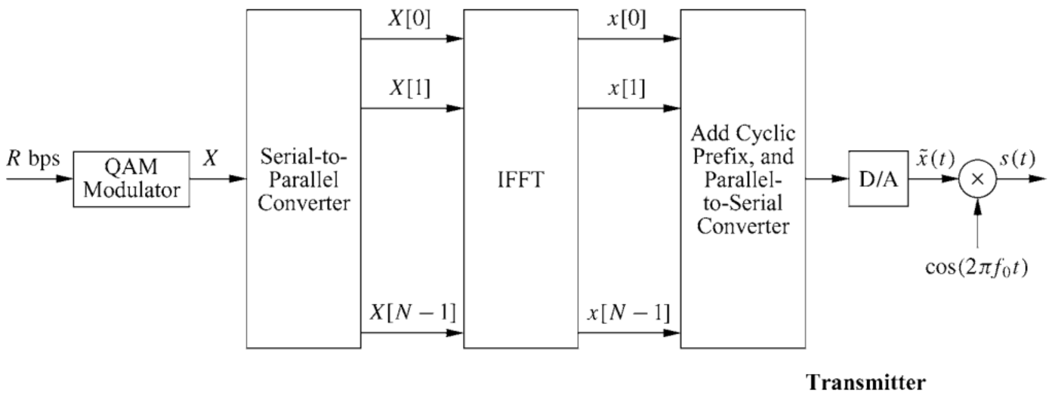
$$x[n] = \mathrm{IDFT}\left\{\frac{Y[i]}{H[i]}\right\} = \mathrm{IDFT}\left\{\frac{\mathrm{DFT}\{y[n]\}}{\mathrm{DFT}\{h[n]\}}\right\}$$

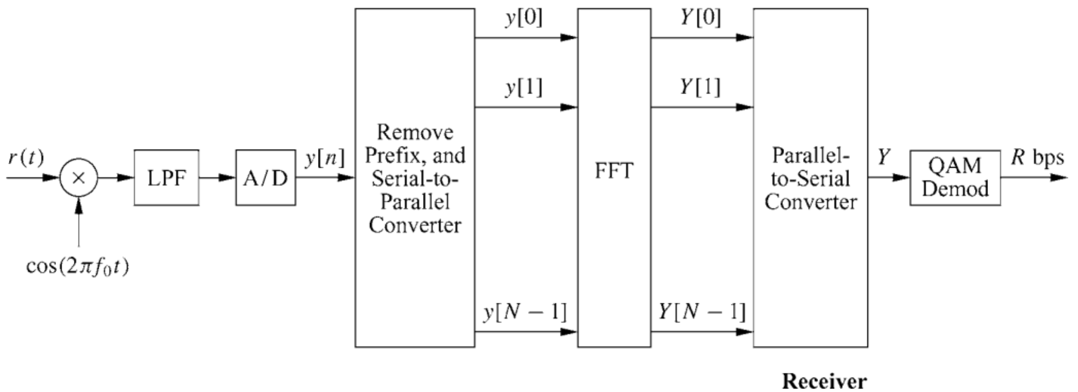
  - The ISI between data blocks of N can be eliminated, at the cost of data rate reduction.



OFDM

- OFDM Implementation



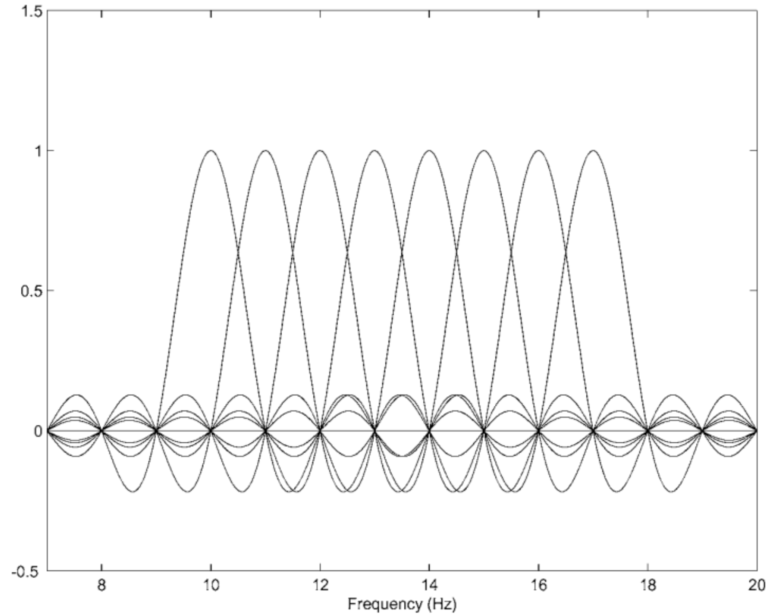


OFDM

- Advantages of OFDM
  - Robustness against multipath fading and intersymbol interference
  - High spectral efficiency
  - Easy equalization for each subcarrier.
  - Flexibility considering adaptive bit and power loading, adaptive modulation and coding, adaptive subcarrier allocation, space-time processing, MIMO, etc.

OFDM

- Challenges of OFDM
  - Time and Frequency offset.



  - High PAPR.

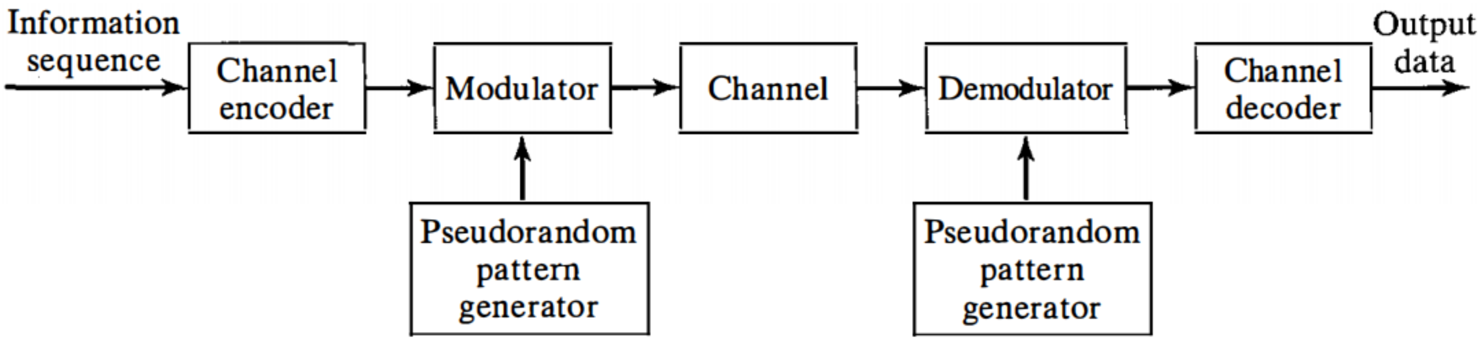
$$\mathrm{PAR} \triangleq \frac{\max_n |x[n]|^2}{\mathbf{E}_n[|x[n]|^2]}.$$

  - Noise enhancement.

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- **Spread Spectrum [Matlab]**
- Channel coding [Matlab]
- Synchronization

Spread Spectrum

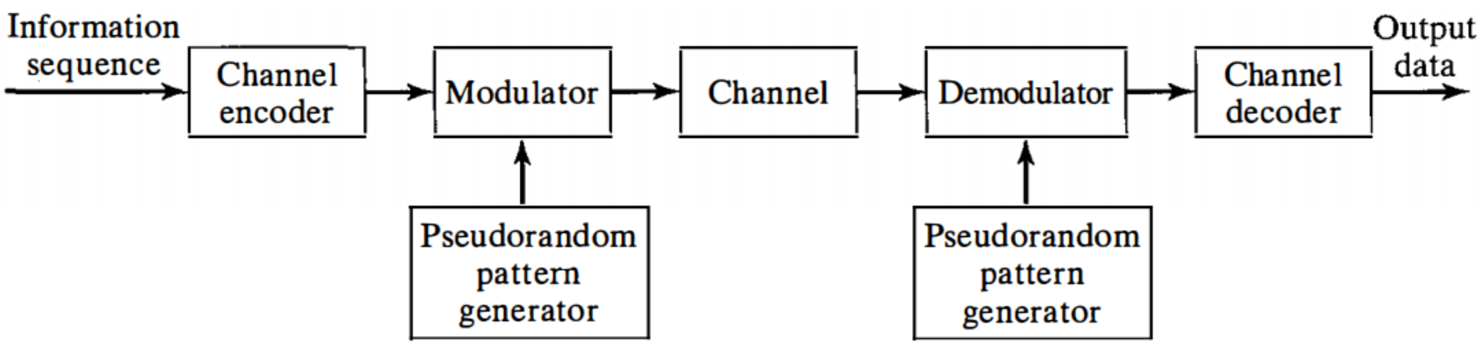


- Introduction
- PN Sequence
- DSSS
- FHSS

Chapter 15

Introduction

- Definition of Spread Spectrum
  - Technique that spreads signal across a given, wider, frequency band in pseudorandom pattern.
  - Two identical pseudorandom sequences.
  - Signals occupy *more* bandwidth than original

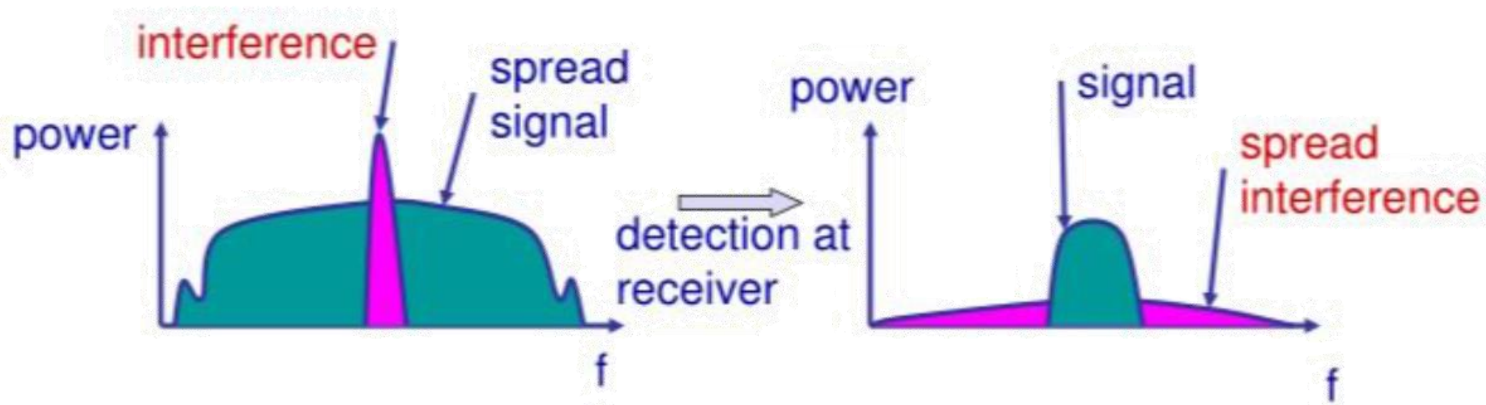


Introduction

- Definition of Spread Spectrum
  - Transmitted signal bandwidth >> information bandwidth).
  - Processing gain

$$\text{Processing Gain } N = \frac{B_{ss}}{B} = 10\log_{10}\left(\frac{B_{ss}}{B}\right)$$

  - Multiplexing gain (How many users can share same frequency band without affecting each other after despreading).



Introduction

- Advantage of Spread Spectrum
  - Immunity from various kinds of noise and multipath distortion
  - Can be used for hiding and encrypting signals
  - Several users can independently use the same higher bandwidth with very little interference, using different codes (CDMA)
  - Resists intentional and non-intentional interference
  - Privacy due to the pseudorandom codes (Millitary)

Introduction

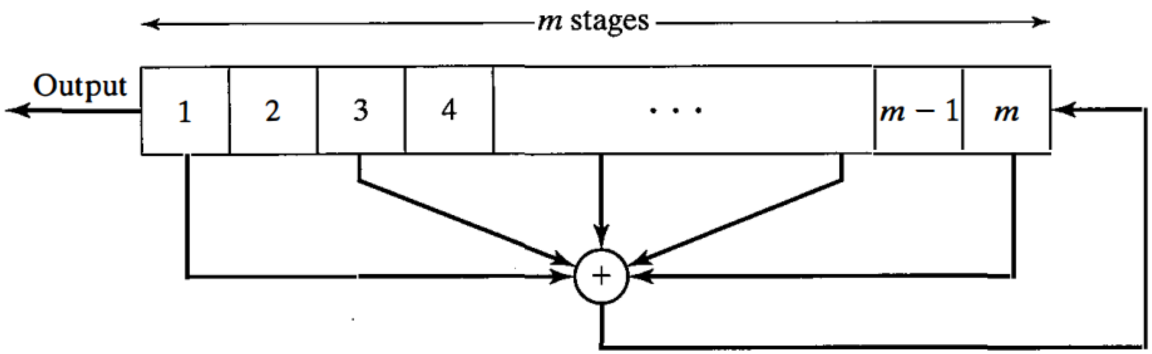
- Disadvantage of Spread Spectrum
  - Bandwidth inefficient if only one user
  - Implementation is more complex

PN Sequences

- PN generator produces periodic sequence that appears to be random
- PN sequences
  - Generated by an algorithm using initial seed
  - Sequences is not statistically random but will pass many test of randomness
  - Sequences referred to as pseudorandom numbers or pseudonoise sequences
  - Unless algorithm and seed are known, the sequence is impractical to predict
- Important properties
  - Randomness
  - Unpredictability

PN Sequences

- Maximum-length shift-register sequence
  - Contains 2^{m-1} ones and $2^{m-1}-1$ zeros
  - For a window of length m slid along outputs for $L=2^m-1$ shifts, each m-tuple appears once, except all zero sequence



PN Sequences

- Maximum-length shift-register sequence
  - Modulo-2 connection

TABLE 15.1 SHIFT-REGISTER CONNECTIONS FOR GENERATING MAXIMUM-LENGTH SEQUENCES

m	Stages Connected to Modulo-2-Adder	*m*	Stages Connected to Modulo-2-Adder	*m*	Stages Connected to Modulo-2 Adder
2	1, 2	13	1, 10, 11, 13	24	1, 18, 23, 24
3	1, 3	14	1, 5, 9, 14	25	1, 23
4	1, 4	15	1, 15	26	1, 21, 25, 26
5	1, 4	16	1, 5, 14, 16	27	1, 23, 26, 27
6	1, 6	17	1, 15	28	1, 26
7	1, 7	18	1, 12	29	1, 28
8	1, 5, 6, 7	19	1, 15, 18, 19	30	1, 8, 29, 30
9	1, 6	20	1, 18	31	1, 29
10	1, 8	21	1, 20	32	1, 11, 31, 32
11	1, 10	22	1, 22	33	1, 21
12	1, 7, 9, 12	23	1, 19	34	1, 8, 33, 34

PN Sequences

- Maximum-length shift-register sequence
  - Sequence contains: 1) One run of ones, length m; 2) One run of zeros, length m-1; 3) One run of ones and one run of zeros, length m-2; 4) Two runs of ones and two runs of zeros, length m-3; 5) 2^{m-3} runs of ones and runs of zeros, length 1
  - Autocorrelation of ± 1 sequence is

$$R_c(m) = \frac{1}{L}\sum_{n=1}^{L} c_n c_{n+m} = \begin{cases} 1, & m = 0 \\ -1/L, & m = 1,2,\ldots,L-1 \end{cases}$$

PN Sequences

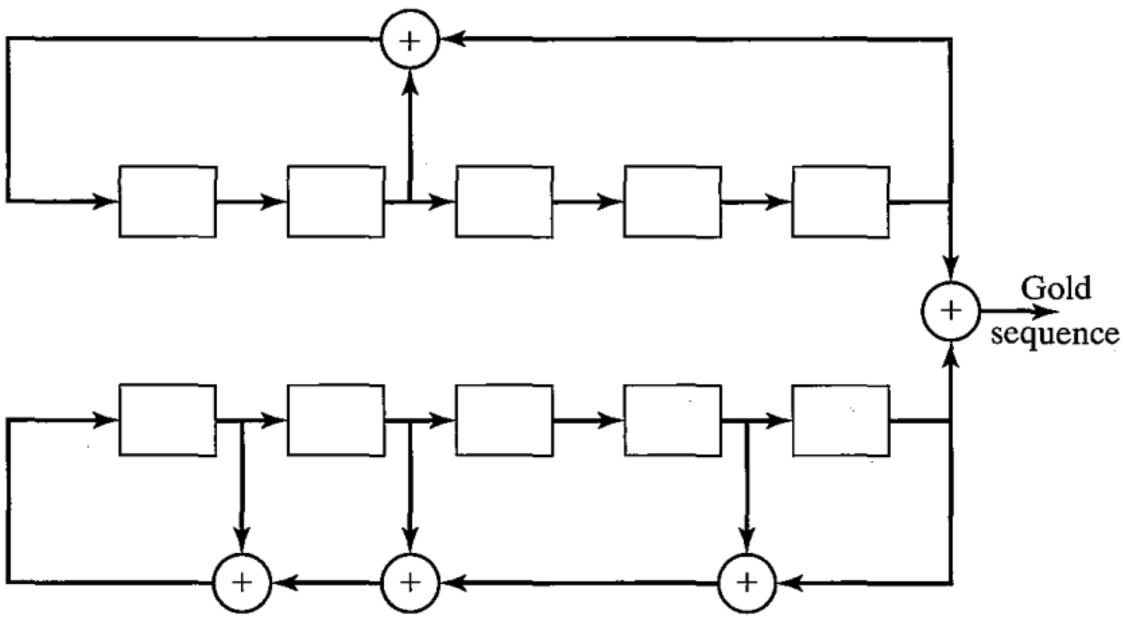
- Maximum-length shift-register sequence
  - Cross correlation compare two sequences from different sources rather than a shift copy with itself

$$R_{xy}(m) = \frac{1}{L}\sum_{n=1}^{L} x_n y_{n+m}$$

  - Cross correlation between m-sequence and noise or two different sequences are low (helpful for filtering out noise or CDMA)

PN Sequences

- Golden Sequence
  - Constructed by the XOR of two m-sequences with the same clocking
  - Well-defined cross correlation properties
  - Only simple circuitry needed to generate large number of unique codes



PN Sequences

- Golden Sequence

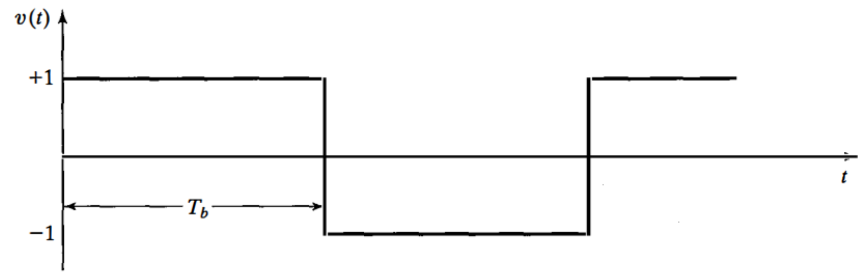
TABLE 15.2 PEAK CROSS CORRELATIONS OF M-SEQUENCES AND GOLD SEQUENCES

		m Sequences				
			Peak Cross		Gold	Sequences
m	$L = 2^{m-1}$	Number	Correlation R_{max}	$R_{max}/R(0)$	R_{max}	$R_{min}/R[0]$
3	7	2	5	0.71	5	0.71
4	15	2	9	0.60	9	0.60
5	31	6	11	0.35	9	0.29
6	63	6	23	0.36	17	0.27
7	127	18	41	0.32	17	0.13
8	255	16	95	0.37	33	0.13
9	511	48	113	0.22	33	0.06
10	1023	60	383	0.37	65	0.06
11	2047	176	287	0.14	65	0.03
12	4095	144	1407	0.34	129	0.03

DSSS

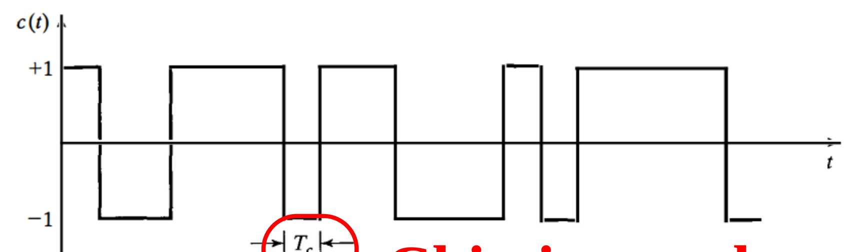
- Each bit in original signal is represented by multiple bits in the transmitted signal

data: $v(t) = \sum_{n=-\infty}^{\infty} a_n g_T(t - nT_b)$



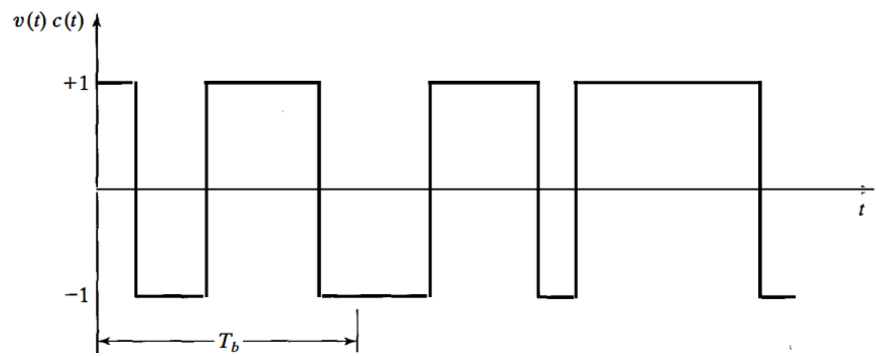
chip

PN signal: $c(t) = \sum_{n=-\infty}^{\infty} c_n p(t - nT_c)$



Chip interval

Product signal:

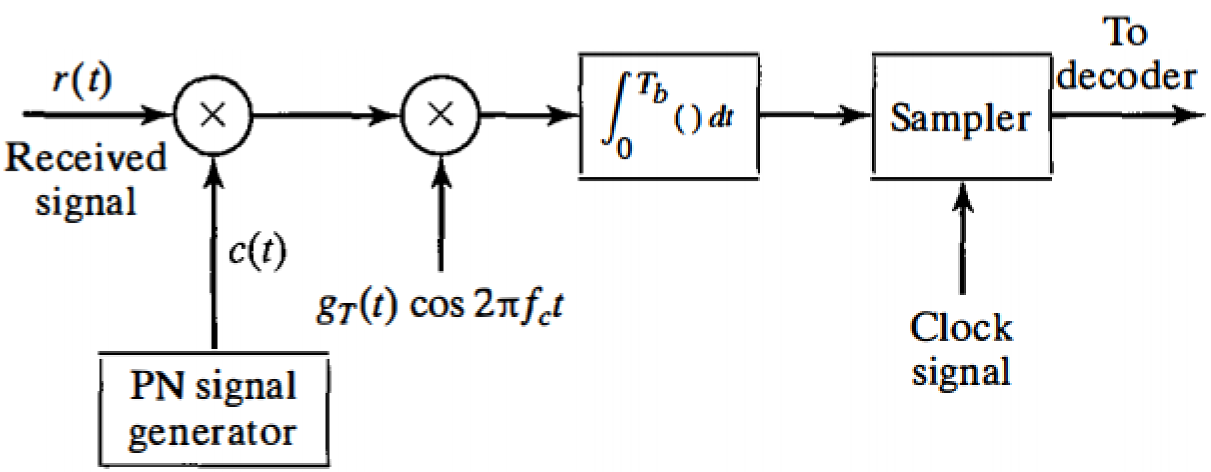


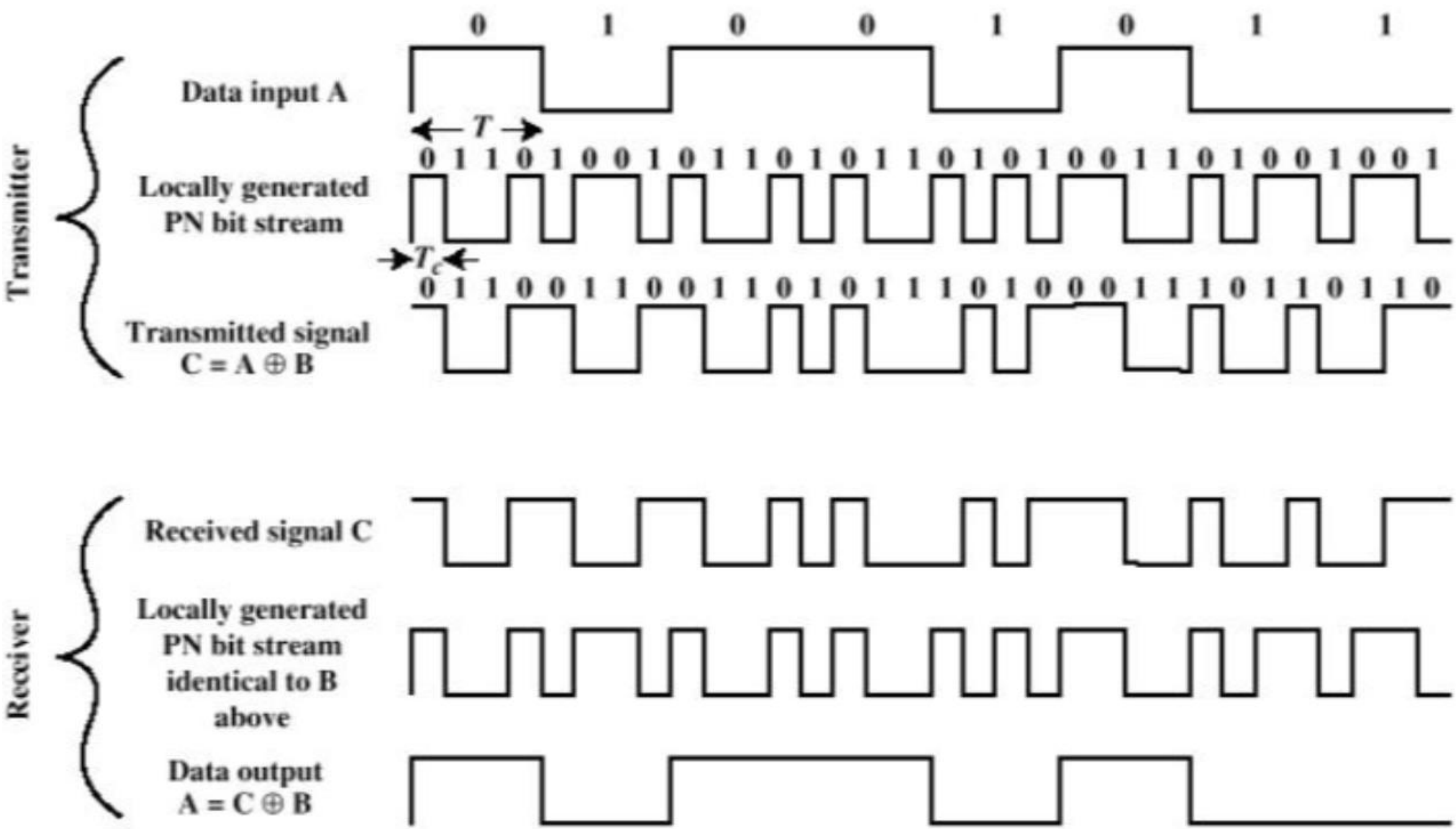
DSSS

- Spreading code spreads signal across a wider frequency band
  - Spread is in direct proportion to number of bits used

$$L_c = \frac{T_b}{T_c}$$

- Demodulator




0 1 0 0 1 0 1 1
Data input A
T
0 1 1 0 1 0 0 1 0 1 1 0 1 0 1 1 0 1 0 1 0 0 1 1 0 1 0 0 1 0 0 1
Locally generated PN bit stream
Tc
0 1 1 0 0 1 1 0 0 1 1 0 1 0 1 1 1 0 1 0 0 0 1 1 1 0 1 1 0 1 1 0
Transmitted signal C = A ⊕ B
Transmitter
Received signal C
Locally generated PN bit stream identical to B above
Data output A = C ⊕ B
Receiver

DSSS

- Performance of DSSS with interference
  - Received signal: $r(t) = A_c v(t)c(t)\cos 2\pi f_c t + i(t)$
  - After dispreading: $r(t)c(t) = A_c v(t)\cos 2\pi f_c t + i(t)c(t)$
- Consider an example with $i(t) = A_I \cos 2\pi f_I t$,

PSD after dispreading: $I_0 \doteq P_I / W$

Total power after demodulator:

$$I_0 R_b = P_I R / W = \frac{P_I}{W/R} = \frac{P_I}{T_b/T_c} = \frac{P_I}{L_c}.$$

$L_c = \frac{T_b}{T_c}$ **Processing gain**

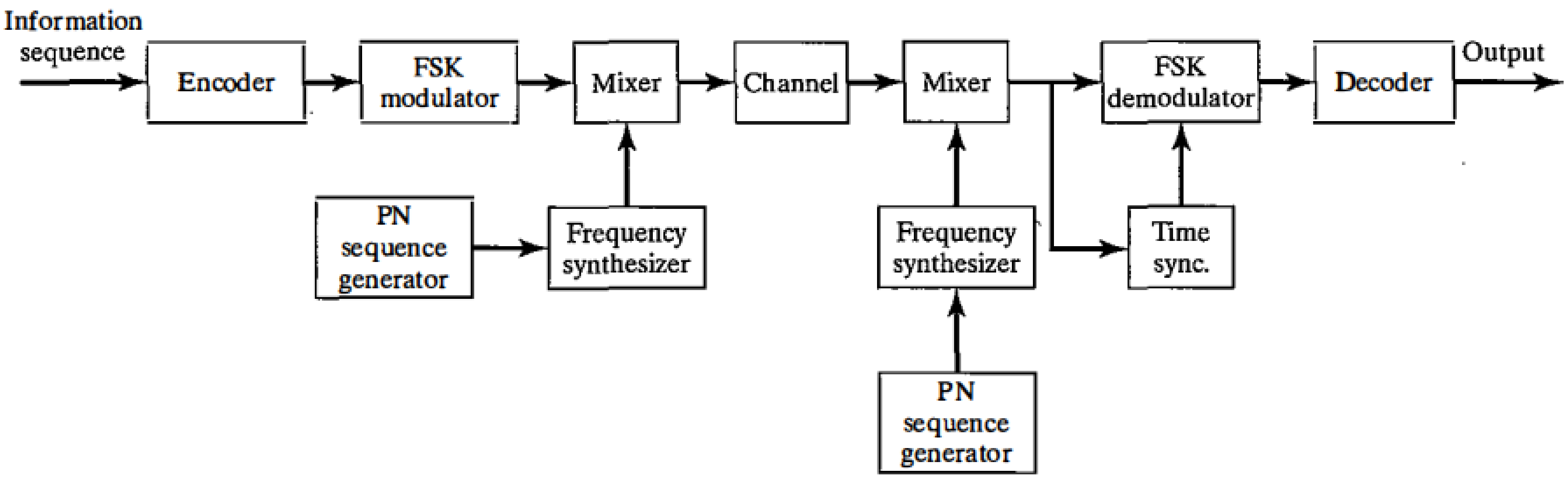
Spread the bandwidth of the interference to larger bandwidth such that the interference power after demodulation is reduced!

FHSS

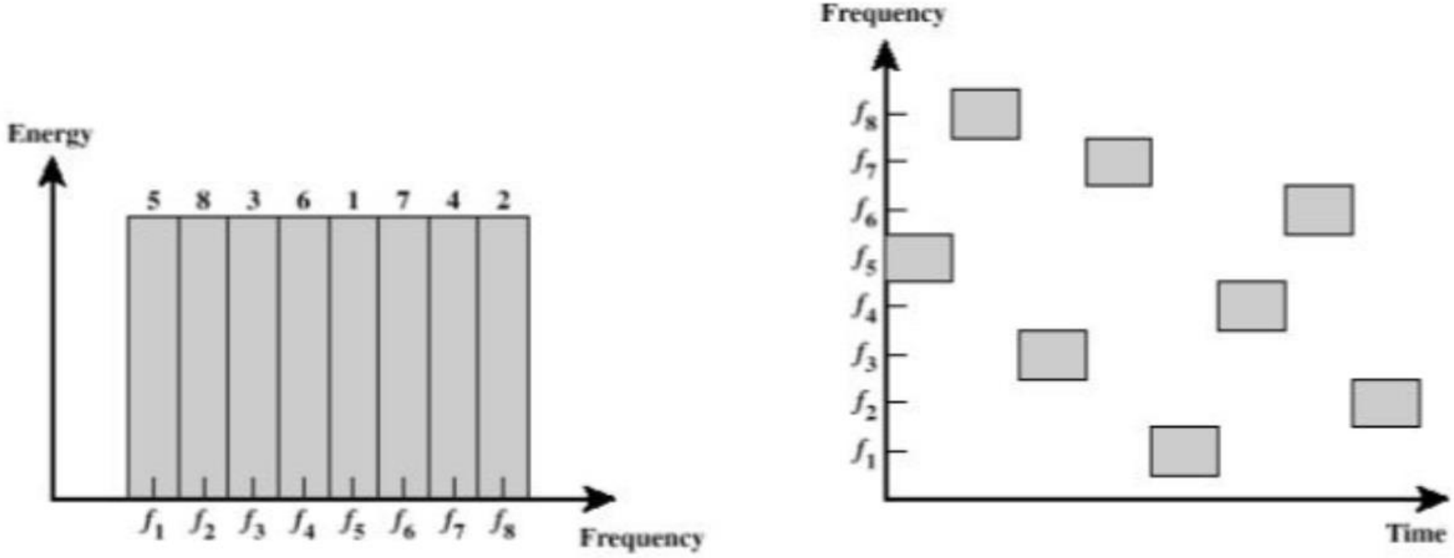
- Signal is broadcast over seemingly random series of radio
  - A number of channels allocated for the FH signal
  - Width of each channel corresponds to bandwidth of input signal
- Signal hops from frequency to frequency at fixed intervals
  - Transmitter operates in one channel at a time
  - Bits are transmitted using some encoding scheme
  - At each successive interval, a new carrier frequency is selected
- Channel sequence dictated by spreading code
- Receiver, hopping between frequencies in synchronization with transmitter, picks up message

FHSS

- Advantages
  - Eavesdroppers hear only unintelligible blips
  - Attempts to jam signal on one frequency succeed only at knocking out a few bits



FHSS



(a) Channel assignment (b) Channel use

Large number of frequencies used

FHSS

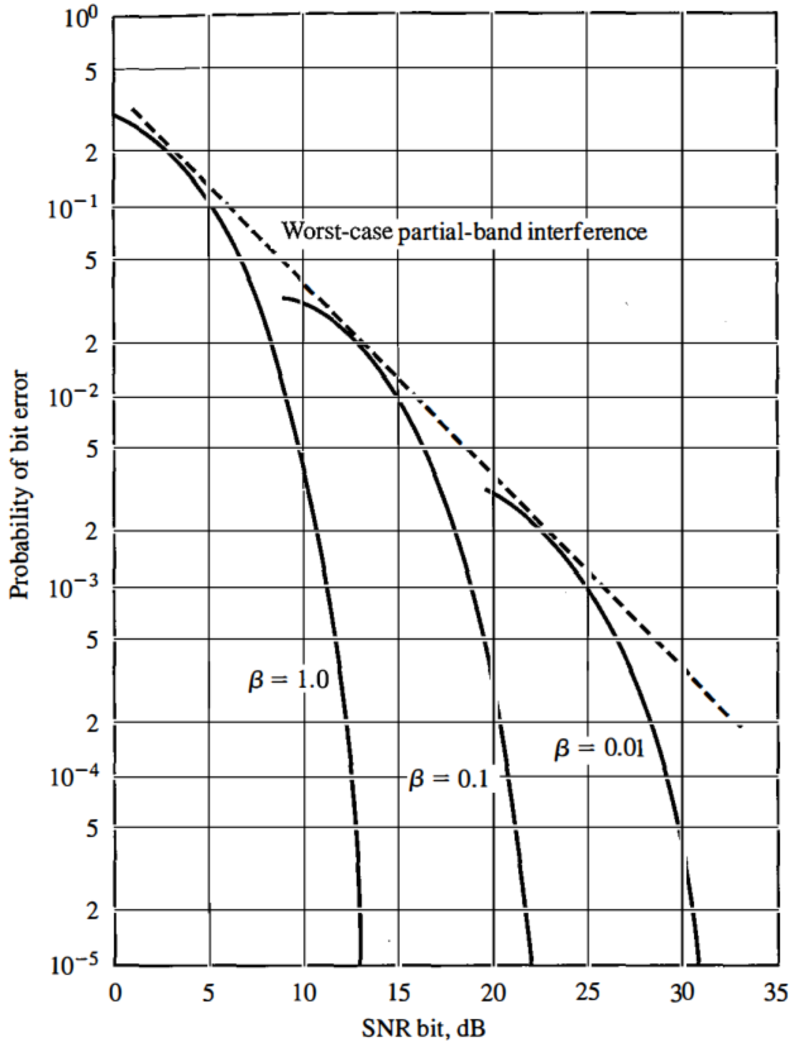
- MFSK signal is translated to a new frequency every Tc seconds by modulating the MFSK signal with the FHSS carrier signal
- For data rate of R:
  - duration of a bit: T = 1/R seconds
  - duration of signal element: Ts = LT seconds
- Tc ≥ Ts - slow-frequency-hop spread spectrum
  - Cheaper to implement but less protection against noise/jamming
  - Popular technique for wireless LANs

FHSS

- Tc ≥ Ts - slow-frequency-hop spread spectrum

SNR: $\rho_b = \frac{\mathcal{E}_b}{I_0} = \frac{W/R}{P_I/P_S}$

$$P_b(\beta) = \frac{\beta}{2} e^{-\beta\rho_b/2}$$
$$= \frac{\beta}{2} \exp\left(-\frac{\beta W/R}{2P_I/P_S}\right)$$



FHSS

- Tc < Ts - fast-frequency-hop spread spectrum
  - More expensive to implement, provides more protection against noise/jamming
  - Fast FH systems are particularly attractive for military communications.

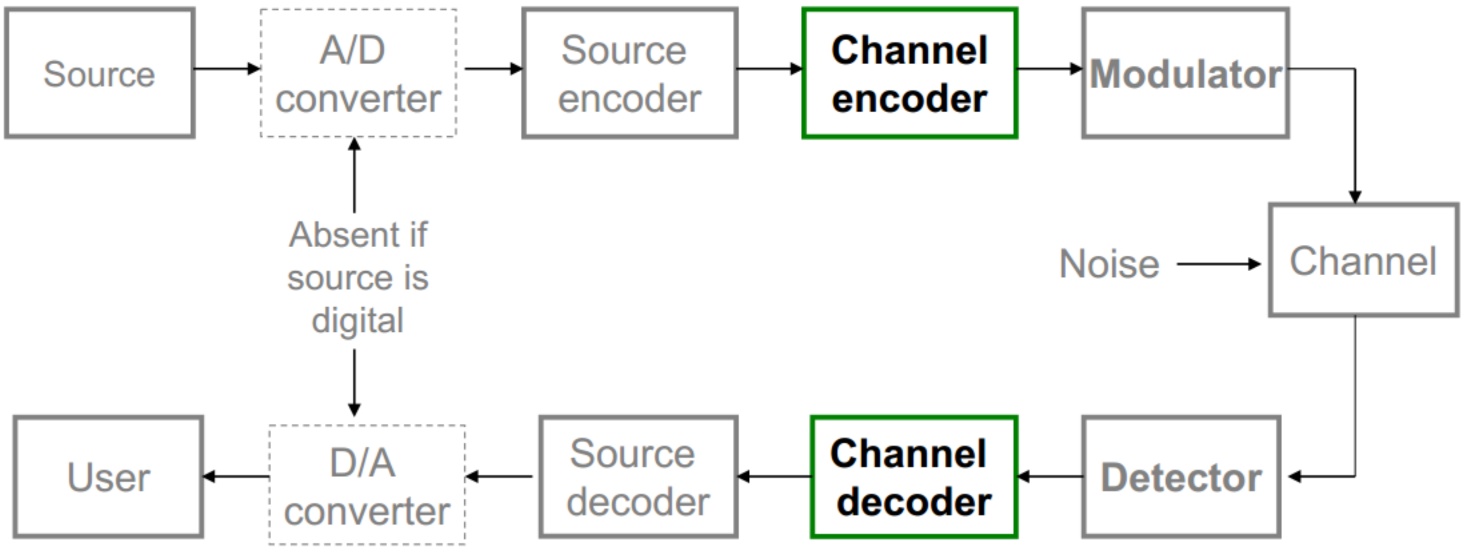
$$P_b = \frac{1}{2^{2N-1}} e^{-\rho_b/2} \sum_{i=0}^{N-1} K_i \left(\frac{\rho_b}{2}\right)^i$$

$$K_i = \frac{1}{i!} \sum_{r=0}^{N-1-i} \binom{2N-1}{r}$$

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- **Channel coding [Matlab]**
- Synchronization

Channel coding



- Linear block code
- Convolutional code

Chapter 13.1-13.3

Channel coding

- Information theory and channel coding
  - Shannon's noisy channel coding theorem tells us that adding controlled redundancy allows transmission at arbitrarily low bit error rate (BER) as long as $R \le C$
  - Error control coding (ECC) uses this controlled redundancy to detect and correct errors
  - ECC depends on the system requirements and the nature of the channel
  - The key in ECC is to find a way to add redundancy to the channel so that the receiver can fully utilize that redundancy to detect and correct the errors, and to reduce the required transmit power (coding gain)

Channel coding

- Information theory and channel coding
  - Consider for example the case that we want to transmit data over a telephone link using a modem under the conditions that link bandwidth = 3 kHz and the modem can operate up to the speed of 3600 bits/sec at an error probability Pe = 8×10^{-4}.
  - Target: transmit the data the rate of 1200 bits/sec at maximum output SNR = 13 dB with a probability of error 1×10^{-4}

Channel coding

- Information theory and channel coding
  - Shannon theorem tells us that channel capacity is

    $$C = B \log_2 \left(1 + \frac{S}{N}\right) = 13,000 \text{ bits/sec}$$

    since B=3000, S/N=13 dB=20
  - Thus, by Shannon's theorem, we can transmit the data with an arbitrarily small error probability
  - Note that without coding Pe = 8×10^{-4}, the target Pe is not met.

Channel coding

- Information theory and channel coding
  - Consider a simple code design with repetition code.
  - Every bit is transmitted 3 times, e.g., when b_k="0" or "1", transmitted codewords are "000" or "111"
  - Based on the received codewords, the decoder attempts to extract the transmitted bits using majority-logic decoding scheme
  - Obviously, the transmitted bits will be recovered correctly as long as no more than one of the bits in the codewords is affected by noise

Tx bits b_k	0	0	0	0	1	1	1	1
Codewords	000	000	000	000	111	111	111	111
Rx bits	000	001	010	100	011	101	110	111
$\hat{b}_k$	0	0	0	0	1	1	1	1

Channel coding

- Information theory and channel coding
  - With this simple error control coding, the probability of error is

$$P_e = P(b_k \neq \hat{b}_k)$$

$$= P\ (2 \text{ or more bits in codeword are in error})$$

$$= \binom{3}{2} q_c^2 (1 - q_c) + \binom{3}{3} q_c^3$$

$$= 3q_c^2 - 2q_c^3$$

$$= 0.0192 \times 10^{-4}$$

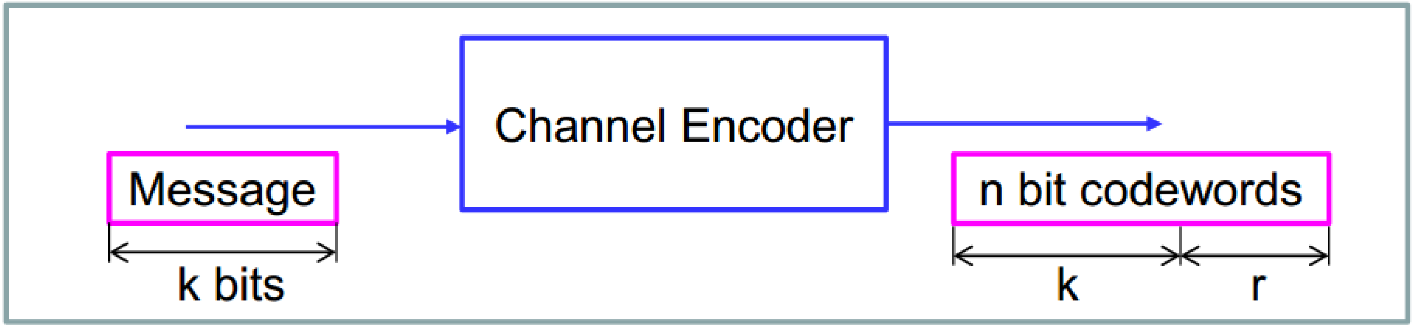
$$\leq \text{Required } P_e \text{ of } 10^{-4}$$

Channel coding

- From the above example, we can see the importance of coding techniques.
- Coding techniques are classified as either block codes or convolutional codes, depending on the presence or absence of memory
- A block code has no memory
  - Information sequence is broken into blocks of length k
  - Each block of k inf. bits is encoded into a block of n coded bits
  - No memory from one block to another block
- A convolutional code has memory
  - A shift register of length k_0L is used
  - Inf. bits enter the shift register k_0 bits at a time and n_0 coded bits are generated
  - These n_0 bits depend not only on the recent k_0 bits, but also on the $k_0(L-1)$ previous bits

Linear block codes

- Block codes
  - An (n,k) block code is a collection of $M=2^k$ codewords of length n
  - Each codeword has a block of k inf. bits followed by a group of r=n-k check bits that are derived from the k inf. bits in the block preceding the check bits



  - The code is said to be linear if any linear combination of 2 codewords is also a codeword, i.e., if c_i and c_j are codewords, then $c_i + c_j$ is also a codeword (addition is module-2)

Linear Block codes

- **Code rate** (rate efficiency) = k/n
- Matrix description:
  - Codeword $\mathbf{c} = (c_1, c_2, ..., c_n)$
  - Message bits $\mathbf{m} = (m_1, m_2, ..., m_k)$
- Each block code can be generated using a **Generator matrix** **G** (dim: kxn)
- Given **G**, then

$$\boxed{\mathbf{c} = \mathbf{m}\mathbf{G}}$$

Codeword message

Linear Block codes

- Generator matrix **G**

$$\mathbf{G} = [\mathbf{I}_k | \mathbf{P}]_{k \times n}$$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & p_{11} & p_{12} & \cdots & p_{1,n-k} \\ 0 & 1 & & 0 & p_{21} & p_{22} & & p_{2,n-k} \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & p_{k,1} & p_{k,2} & \cdots & p_{k,n-k} \end{array}\right]$$

  - $\mathbf{I}_k$ is an identity matrix of order k
  - **P** is a matrix of order kx(n-k), which is selected so that the code will have certain desired properties

Linear Block codes

- Generator matrix **G**
  - The form of **G** implies that the 1^{st} k components of any codeword are precisely the information symbols
  - This form of linear encoding is called systematic encoding
  - Systematic-form codes allow easy implementation and quick look-up features for decoding
  - For linear codes, any code is equivalent to a code in systematic form (given the same performance). Thus, we can restrict our study to only systematic codes

Linear Block codes

- Example
  - Hamming code is a family of (n,k) linear block codes that have the following parameters
    1. Codeword length $n = 2^m - 1, \quad m \geq 3$
    2. # of message bits $k = 2^m - m - 1$
    3. # of parity check bits $n - k = m$
    4. Capable of providing single-error correction capability with $d_{\min} = 3$
  - (7,4) Hamming code with generator matrix

$$\mathbf{G} = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right]$$

Find all codewords

Linear Block codes

- Example
  - (7,4) Hamming code

Message				codeword						
0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1	1	0	1
0	0	1	0	0	0	1	0	1	1	1
0	0	1	1	0	0	1	1	0	1	0
0	1	0	0	0	1	0	0	0	1	1
0	1	0	1	0	1	0	1	1	1	0
0	1	1	0	0	1	1	0	1	0	0
0	1	1	1	0	1	1	1	0	0	1
1	0	0	0	1	0	0	0	1	1	0
1	0	0	1	1	0	0	1	0	1	1
1	0	1	0	1	0	1	0	0	0	1
1	0	1	1	1	0	1	1	1	0	0
1	1	0	0	1	1	0	0	1	0	1
1	1	0	1	1	1	0	1	0	0	0
1	1	1	0	1	1	1	0	0	1	0
1	1	1	1	1	1	1	1	1	1	1

Linear Block codes

- Parity check matrix
  - For each **G**, it is possible to find a corresponding parity check matrix **H** $\mathbf{H} = \left[\mathbf{P}^T \; | \mathbf{I}_{n-k}\right]_{(n-k)\times n}$
  - **H** can be used to verify if a codeword C is generated by **G**
  - Let C be a codeword generated by $\mathbf{G} = [\mathbf{I}_k | \mathbf{P}]_{k\times n}$

$$\Rightarrow \boxed{\mathbf{c}\mathbf{H}^T = \mathbf{m}\mathbf{G}\mathbf{H}^T = 0}$$

  - Think about the parity check matrix of (7,4) Hamming code

Linear Block codes

- Error syndrome
  - Received codeword $\mathbf{r}=\mathbf{c}+\mathbf{e}$, where $\mathbf{e}$=Error vector or Error pattern and it is 1 in every position where data word is in error
  - Example

$$\mathbf{c} = [1 \;\; 0 \;\; 1 \;\; 0]$$

$$\mathbf{r} = [1 \;\; 1 \;\; 0 \;\; 0]$$

$$\mathbf{e} = [0 \;\; 1 \;\; 1 \;\; 0]$$

  - Error syndrome: $\mathbf{s} \triangleq \mathbf{r}\mathbf{H}^T$

Linear Block codes

- Error syndrome
  - Note that

$$
\begin{aligned}
\mathbf{s} = \mathbf{r}\mathbf{H}^T &= (\mathbf{c}+\mathbf{e})\mathbf{H}^T \\
&= \mathbf{c}\mathbf{H}^T + \mathbf{e}\mathbf{H}^T \\
&= \mathbf{e}\mathbf{H}^T
\end{aligned}
$$

  - If **s**=**0**, then **r** = **c** and m is the 1st k bits of **r**
  - If **s**≠**0**, and **s** is the jth row of $\mathbf{H}^T$, then 1 error in jth position of **r**

Linear Block codes

- Error syndrome
  - Consider the (7,4) Hamming code for example

$$\mathbf{H}^T = \left[\mathbf{P}^T | \mathbf{I}_{n-k}\right]^T = \begin{bmatrix} \mathbf{P} \\ \mathbf{I}_{n-k} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ \hdashline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

  - So if $\mathbf{r} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$

$$\Longrightarrow \quad \mathbf{r}\mathbf{H}^T = [0\ 0\ 0]$$

  - But if $\mathbf{r} = [1\ 1\ 1\ 1\ 1\ 1\ 0]$

$$\Longrightarrow \quad \mathbf{r}\mathbf{H}^T = [0\ 0\ 1]$$

= Error syndrome **s**

- Note that s is the last row of $\mathbf{H}^T$
- Also note error took place in the last bit
- Syndrome indicates error position

Linear Block codes

- Cyclic code
  - A code $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{2^k}\}$ is cyclic if

$$(c_1, c_2, \ldots, c_n) \in C \Rightarrow (c_n, c_1, \ldots, c_{n-1}) \in C$$

  - (7,4) Hamming code is cyclic

message	codeword
0001	0001101
1000	1000110
0100	0100011

Linear Block codes

- Important parameters

  Hamming Distance between codewords c_i and c_j:

  $d(c_i, c_j)$ = # of components at which the 2 codewords differ

  Hamming weight of a codeword c_i is

  $w(c_i)$ = # of non-zero components in the codeword

  Minimum Hamming Distance of a code:

  $d_{min} = \min d(c_i, c_j)$ for all $i \neq j$

  Minimum Weight of a code:

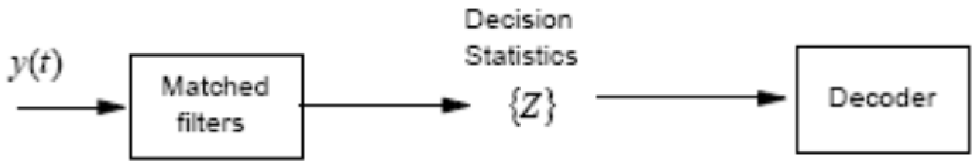
  $w_{min} = \min w(c_i)$ for all $c_i \neq 0$

  Theorem: In any linear code, $d_{\min} = w_{\min}$

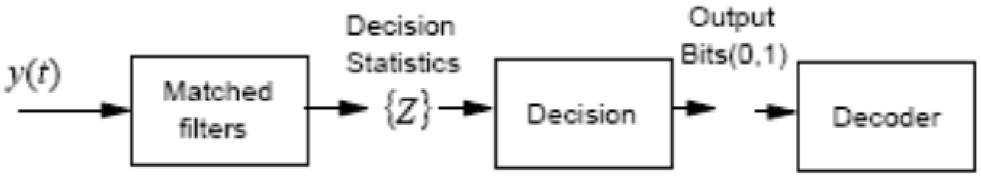
  Exercise: Find d_{min} for (7,4) Hamming code

Linear Block codes

- Soft-decision and hard-decision decoding
  - Soft-decision decoder operates directly on the decision statistics



  - Hard-decision decoder makes "hard" decision (0 or 1) on individual bits



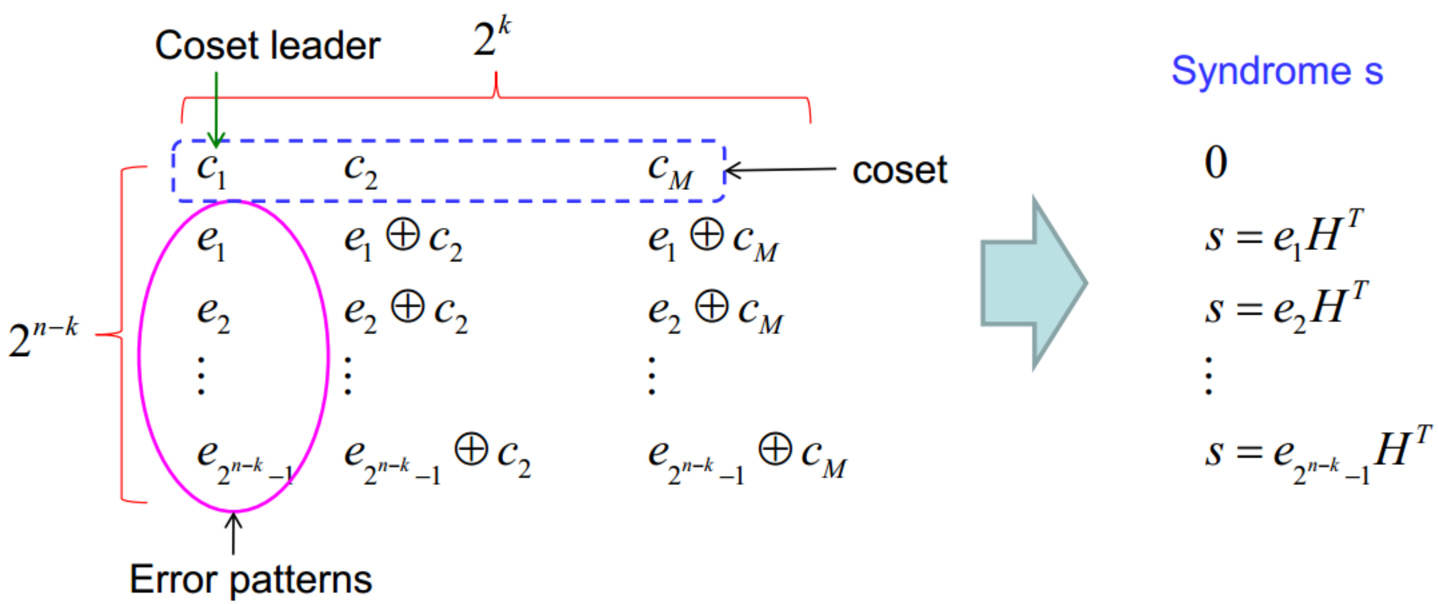
  - Here we only focus on hard decision decoder

Linear Block codes

- Hard-decision decoding
  - Minimum Hamming distance decoding
    1. Given the received codeword **r**, choose **c** which is closest to **r** in terms of Hamming distance
    2. To do so, one can do an exhaustive search (but complexity problem if k is large)
  - Syndrome decoding
    1. Syndrome testing: $\mathbf{r}=\mathbf{c}+\mathbf{e}$ with $\mathbf{s}=\mathbf{r}\mathbf{H}^T$
    2. This implies that the corrupted codeword **r** and the error pattern have the same syndrome
    3. A simplified decoding procedure based on the above observation can be used

Linear Block codes

- Hard-decision decoding
  - Let the codewords be denoted as $\{c_1, c_2, \ldots, c_M\}$ with c_1 being the all-zero codeword
  - A standard array is constructed as

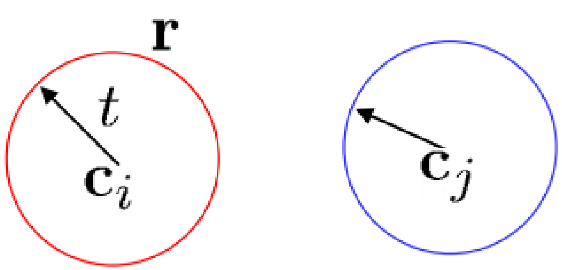


Linear Block codes

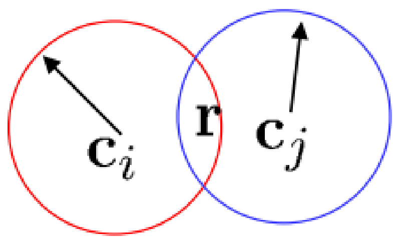
- Hard-decision decoding
  - Hard-decoding procedure
    1. Find the syndrome by **r** using $\mathbf{s}=\mathbf{r}\mathbf{H}^T$
    2. Find the coset corresponding to **s** by using the standard array
    3. Find the coset leader and decode as $\mathbf{c}=\mathbf{r}+\mathbf{e}_j$
  - Try on (7,4) Hamming code

Linear Block codes

- Hard-decision decoding
  - A linear block code with a minimum distance d_{min} can
    1. Detect up to (d_{min}-1) errors in each codeword
    2. Correct up to $t = \lfloor \frac{d_{min}-1}{2} \rfloor$ errors in each codeword
    3. t is known as the error correction capability of the codeword



$$d(\mathbf{c}_i, \mathbf{c}_j) \geq 2t + 1$$



$$d(\mathbf{c}_i, \mathbf{c}_j) < 2t$$

Linear Block codes

- Hard-decision decoding
  - Consider a linear block code (n,k) with an error correcting capability t. The decoder can correct all combination of errors up to and including t errors
  - Assume that the error probability of each individual coded bit is p and that bit errors occur independently since the channel is memoryless
  - If we send n-bit block, the probability of receiving a specific pattern of m errors and (n-m) correct bits is

$$p^m(1-p)^{n-m}$$

  - Total number of distinct patterns of n bits with m errors and (n-m) correct bits is

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Linear Block codes

- Hard-decision decoding
  - Total probability of receiving a pattern with m errors is

$$P(m,n) = \binom{n}{m} \cdot p^m (1-p)^{n-m}$$

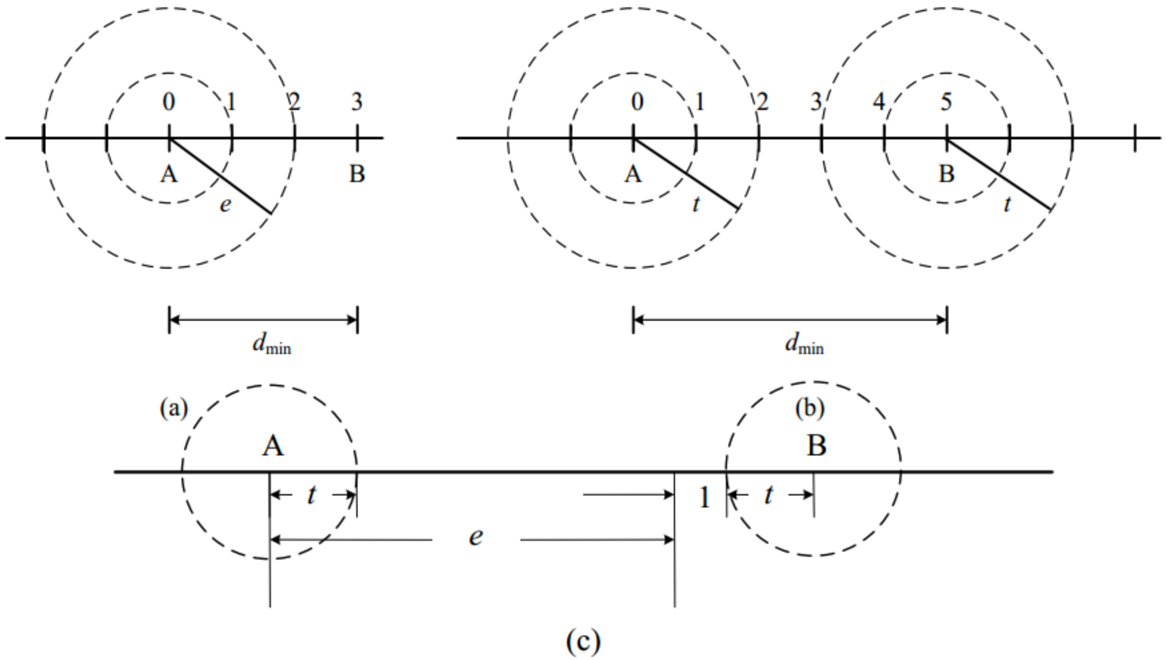
  - Thus, the codeword error probability is upperbounded by

$$P_M \leq \sum_{m=t+1}^{n} \binom{n}{m} p^m (1-p)^{n-m}$$

(with equality for perfect codes)

Linear Block codes

- Hard-decision decoding
  - Key parameters.



  - To detect e bit errors, we have $d_{\min} \geq e+1$
  - To correct t bit errors, we have $d_{\min} \geq 2t+1$

Linear Block codes

- Major classes of block codes
  - Repetition code
  - Hamming code
  - Golay code
  - BCH code
  - Reed-Solomon codes
  - Walsh codes
  - LDPC codes: invented by Robert Gallager in his PhD thesis in 1960, now proved to be capacity approaching and adopted in **5G** standards

Convolutional codes

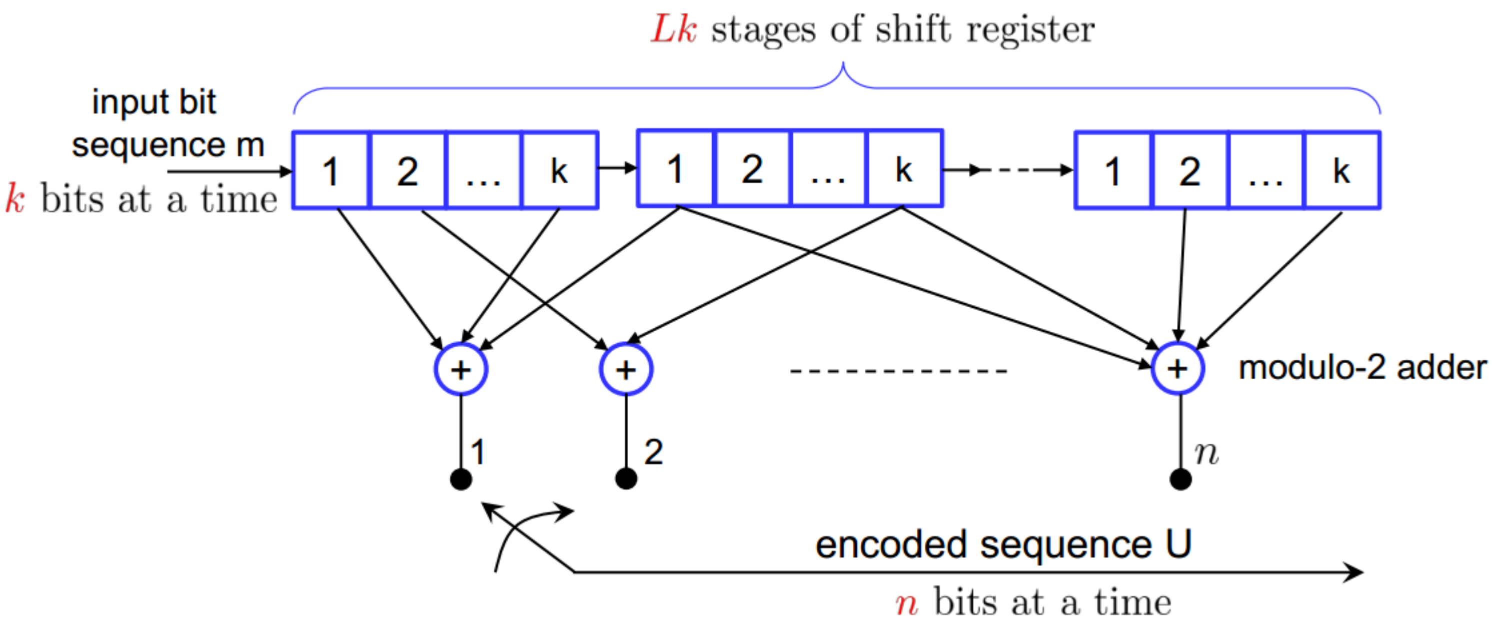
- A convolutional code has memory
  - It is described by 3 integers: n, k, and L
  - Maps k inf. bits into n bits using previous (L-1)k bits
  - The n bits emitted by the encoder are not only a function of the current input k bits, but also a function of the previous (L-1)k bits
  - Code rate = k/n (information bits/coded bits)
  - L is the constraint length and is a measure of the code memory
  - n does not define a block or codeword length

Convolutional codes

- Convolutional encoding
  - A rate k/n convolutional encoder with constraint length L consists of kL-stage shift register and n mod-2 adders
  - At each unit of time
    1. k bits are shifted into the 1^{st} k stages of the register
    2. All bits in the register are shifted k stages to the right
    3. The output of the n adders are sequentially sampled to give the coded bits
    4. There are n coded bits for each input group of k bits or message bits. Hence $R=k/n$ information bits/coded bits is the code rate ($k<n$)

Convolutional codes

- Convolutional encoding
  - ➢ Encoder structure.



  - ➢ Typically, k=1 for binary codes. Hence, consider rate 1/n codes for example.

Convolutional codes

- Convolutional encoding
  - Encoding function: characterizes the relationship between the information sequence m and the output coded sequence U.
  - Four popular methods for representation
    1. Connection pictorial and connection polynomials (usually for encoder)
    2. State diagram
    3. Tree diagram
    4. Trellis diagram

Usually for decoder

Convolutional codes

- Convolutional encoding
  - Connection representation.
  - Specify n connection vectors, $\mathbf{g}_i, i = 1, \ldots, n$ for each of the n mod-2 adders
  - Each vector has kL dimension and describes the connection of the shift register to the mod-2 adders
  - A 1 in the i^{th} position of the connection vector implies shift register is connected
  - A 0 implies no connection exists

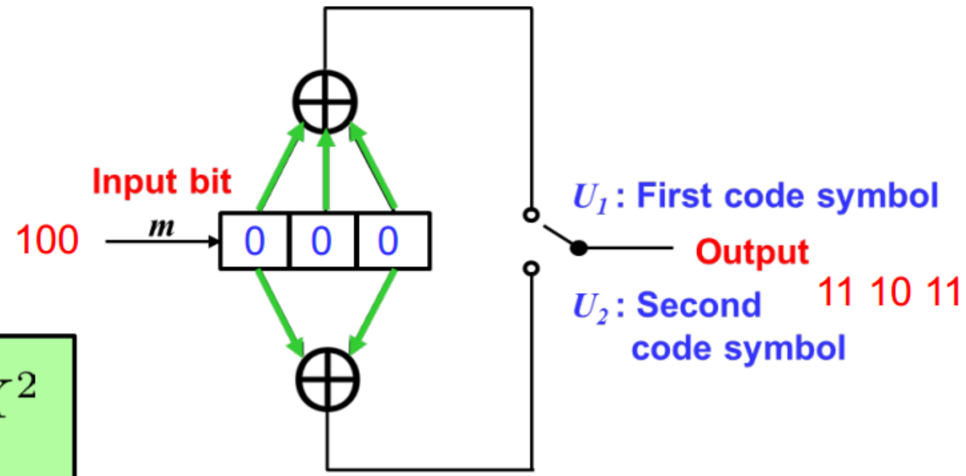
Convolutional codes

- Convolutional encoding
  - Connection representation (L=3, Rate 1/2).

$$g_1 = 1\,1\,1$$
$$g_2 = 1\,0\,1$$

Or

$$g_1(X) = 1 + X + X^2$$
$$g_2(X) = 1 + X^2$$

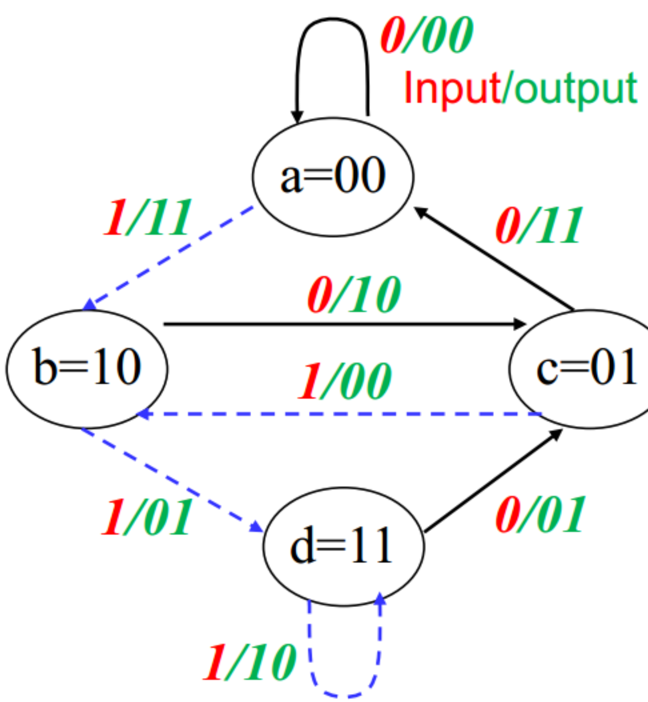


If Initial Register content is 0 0 0 and Input Sequence is 1 0 0. Then Output Sequence is 11 10 11

Convolutional codes

- Convolutional encoding
  - State diagram representation.
  - The contents of the leftmost L-1 stages (or the previous L-1 bits) are considered the current state, 2^{L-1} states
  - Knowledge of the current state and the next input is necessary and sufficient to determine the next output and next state
  - For each state, there are only 2 transitions (to the next state) corresponding to the 2 possible input bits
  - The transitions are represented by paths on which we write the output word associated with the state transition: A solid line path corresponds to an input bit 0, while dashed line for 1

- Convolutional encoding
  - State diagram representation (L=3, Rate 1/2).



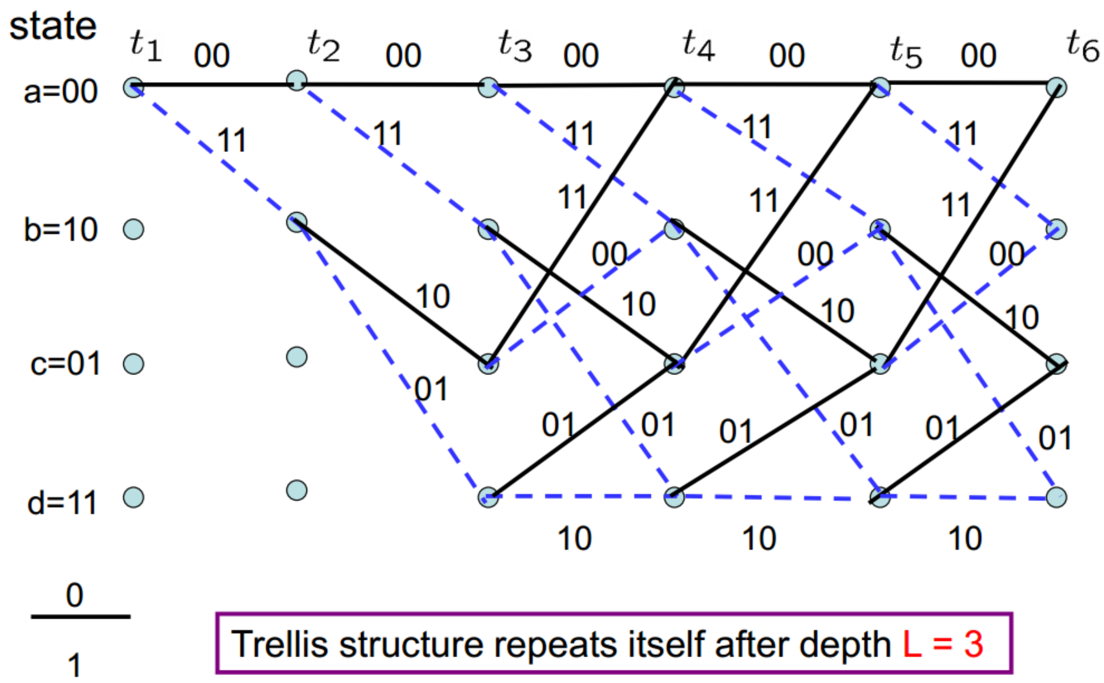
Current State	Input	Next State	Output
00	0	00	***00***
	1	10	***11***
10	0	01	***10***
	1	11	***01***
01	0	00	***11***
	1	10	***00***
11	0	01	***01***
	1	11	***10***

Convolutional codes

- Convolutional encoding
  - State diagram representation.
  - Assume that m=11011 is the input followed by L-1=2 zeros to flush the register. Also assume that the initial register contents are all zero. Find the output sequence U

Convolutional codes

- Convolutional encoding
  - Trellis diagram representation.
  - Trellis diagram is similar to the state diagram, except that it adds the dimension of time.
  - The code is represented by a trellis where each trellis branch describes an output word

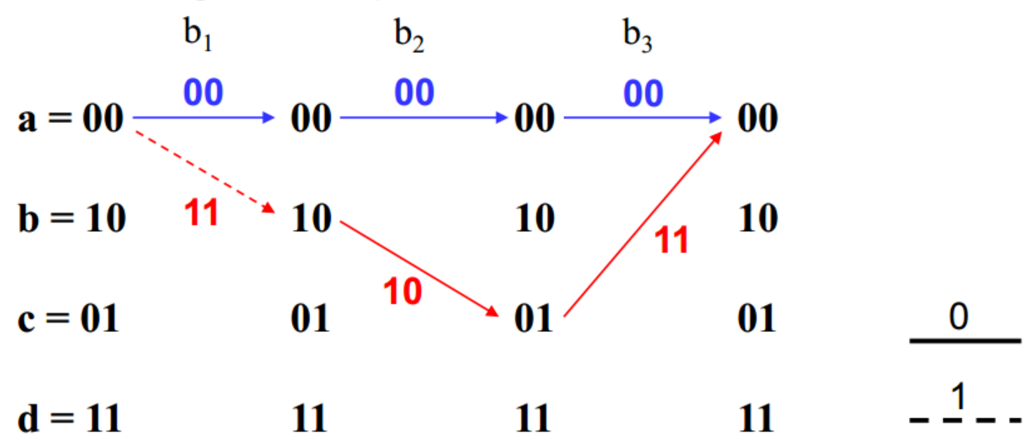


Convolutional codes

- Convolutional encoding
  - Trellis diagram representation.
  - Every input sequence $(m_1, m_2, \ldots)$ corresponds to
    1. a path in the trellis
    2. a state transition sequence $(s_0, s_1, \ldots)$, (assume $s_0 = 0$ is fixed)
    3. an output sequence $((u_1, u_2), (u_3, u_4), \ldots)$
  - For instance, let s_0=00, then

$b_1 b_2 b_3 = 000$ gives output *000000* and states *aaaa*

$b_1 b_2 b_3 = 100$ gives output *111011* and states *abca*



Convolutional codes

- Update
  - We have discussed conv. code with constraint length L and rate 1/n, and the different representations
  - We will talk about decoding of convolutional code with maximum likelihood decoding, Viterbi algorithm, and transfer function

Convolutional codes

- Maximum likelihood decoding
  - Transmit a coded sequence $U^{(m)}$ (corresponds to message sequence m) using a digital modulation scheme (e.g., BPSK or QPSK)
  - Received sequence z
  - Maximum likelihood decoder will
    1. Find the sequence $U^{(j)}$ such that

$$P(\mathbf{Z}|\mathbf{U}^{j}) = \max_{\forall \mathbf{U}^{(m)}} P(\mathbf{Z}|\mathbf{U}^{(m)})$$

    2. Minimize the probability of error if m is equally likely

Convolutional codes

- Maximum likelihood decoding
  - Assume a memoryless channel, i.e., noise components are independent. Then, for a rate 1/n code

$$P(\mathbf{Z}|\mathbf{U}^{(\mathbf{m})}) = \prod_{i=1}^{\infty} P(Z_i|U_i^{(m)}) = \prod_{i=1}^{\infty}\prod_{j=1}^{n} P(z_{ji}|u_{ji}^{(m)})$$

i-th branch of **Z**

  - Then, the problem is to find a path through the trellis such that

$$\max_{\mathbf{U}^{(m)}} \prod_{i=1}^{\infty}\prod_{j=1}^{n} P(z_{ji}|u_{ji}^{(m)})$$

by taking log

$$\max_{\mathbf{U}^{(m)}} \boxed{\sum_{i=1}^{\infty}\sum_{j=1}^{n} \log P(z_{ji}|u_{ji}^{(m)})}$$

Log-likelihood path metric

i-th branch metric

$$= \max_{\mathbf{U}^{(m)}} \sum_{i=1}^{\infty}\sum_{j=1}^{n} LL\left(z_{ji}|u_{ji}^{(m)}\right)$$

Log-likelihood of $z_{ji}|u_{ji}^{(m)}$

Convolutional codes

- Maximum likelihood decoding
  - Log-likelihood.
  - For AWGN channel with soft decision

$z_{ji} = u_{ji} + n_{ji}$ and P($z_{ji}|u_{ji}$) is Gaussian with mean u_{ji} and variance σ^2

Hence

$$\ln p(z_{ji}|u_{ji}) = -\frac{1}{2}\ln\left(2\pi\sigma^2\right) - \frac{(z_{ji}-u_{ji})^2}{2\sigma^2}$$

Note that the objective is to compare which Σ_i ln(p(z|u)) for different **u** is larger, hence, constant and scaling does not affect the results

Then, we let the log-likelihood be $LL(z_{ji}|u_{ji}) = -(z_{ji}-u_{ji})^2$

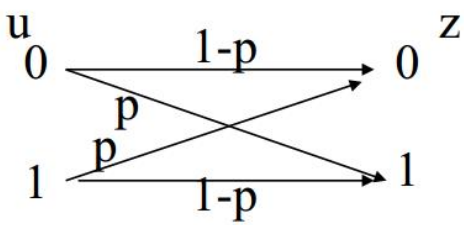
and

$$\boxed{\log P(Z|U^{(m)}) = -\sum_{i=1}^{\infty}\sum_{j=1}^{n}\left(z_{ji}-u_{ji}^{(m)}\right)^2}$$

Thus, soft decision ML decoder is to choose the path whose corresponding sequence is at the minimum Euclidean distance from the received sequence

Convolutional codes

- Maximum likelihood decoding
  - Log-likelihood.
  - For binary symmetric channel (hard decision)



$$p(z|u) = \begin{cases} p & \text{if } z \neq u \\ 1-p & \text{if } z = u \end{cases}$$

$$LL(z_{ji}|u_{ji}) = \ln p(z_{ji}|u_{ji}) = \begin{cases} \ln p & \text{if } z_{ji} \neq u_{ji_i} \\ \ln(1-p) & \text{if } z_{ji} = u_{ji} \end{cases}$$

$$= \begin{cases} \ln p/(1-p) & \text{if } z_{ji} \neq u_{ji} \\ 0 & \text{if } z_{ji} = u_{ji} \end{cases}$$

$$= \begin{cases} -1 & \text{if } z_{ji} \neq u_{ji} \\ 0 & \text{if } z_{ji} = u_{ji} \end{cases}$$ (since p<0.5)

- Thus

$$\boxed{\log P(Z|U^{(m)}) = -d_m}$$

Hamming distance between Z and $U^{(m)}$, i.e. they differ in d_m positions

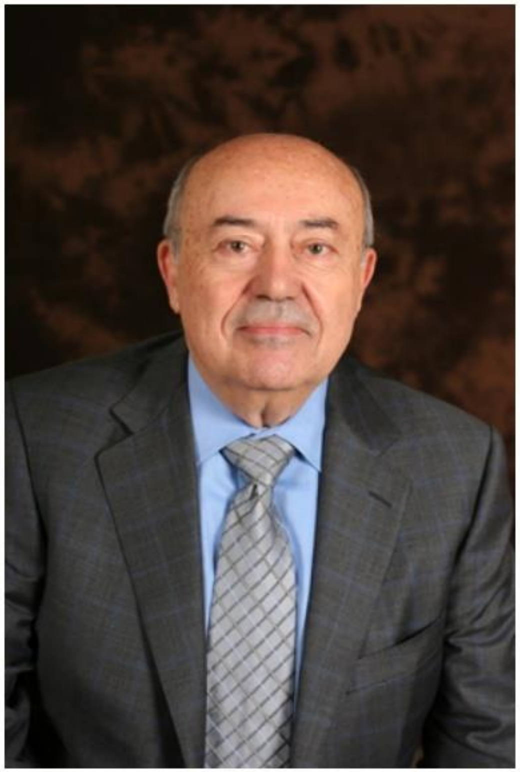
Hard-Decision ML Decoder = Minimum Hamming Distance Decoder

Convolutional codes

- Maximum likelihood decoding
  - Decoding procedure:
    1. Compute, for each branch i, the branch metric using output bits $\{u_{1,i}, u_{2,i}, \ldots, u_{n,i}\}$ associated with that branch and the received symbols $\{z_{1,i}, z_{2,i}, \ldots, z_{n,i}\}$
    2. Compute, for each valid path through the trellis (a valid codeword sequence U(m)), the sum of the branch metrics along that path
    3. The path with the maximum path metric is the decoded path
  - To compare all possible valid paths, we need to do exhaustive search or brute-force, not practical as the # of paths grows exponentially as the path length increases
  - The optimal algorithm for solving this problem is the Viterbi decoding algorithm or Viterbi decoder

Convolutional codes

- Viterbi decoding



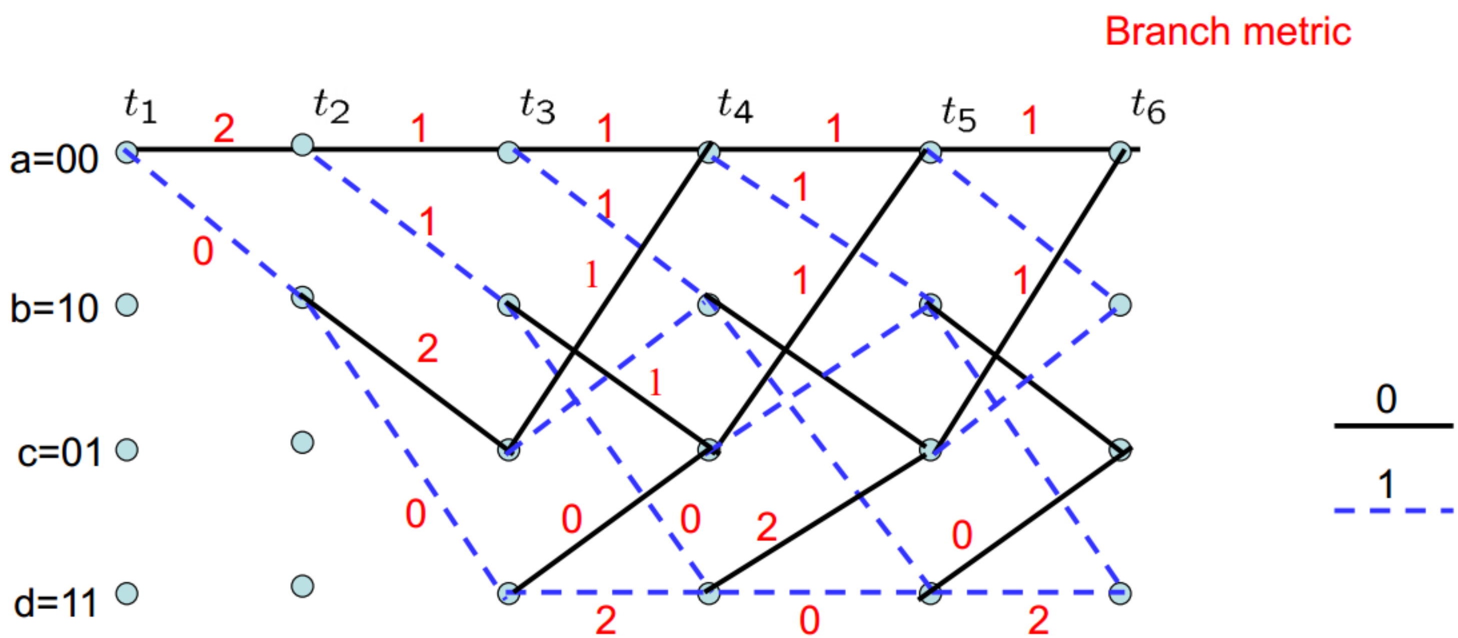
Andrew Viterbi
(1935-)

- BS & MS in MIT
- PhD in University of Southern California
- Invention of Viterbi algorithm in 1967
- Co-founder of Qualcomm Inc. in 1983

Convolutional codes

- Viterbi decoding
  - ➢ Consider R=1/2, L=3 for example.

Input data sequence **m**:	1	1	0	1	1	…
Coded sequence **U**:	11	0 1	01	00	01	…
Received sequence **Z**:	11	0 1	01	10	01	…

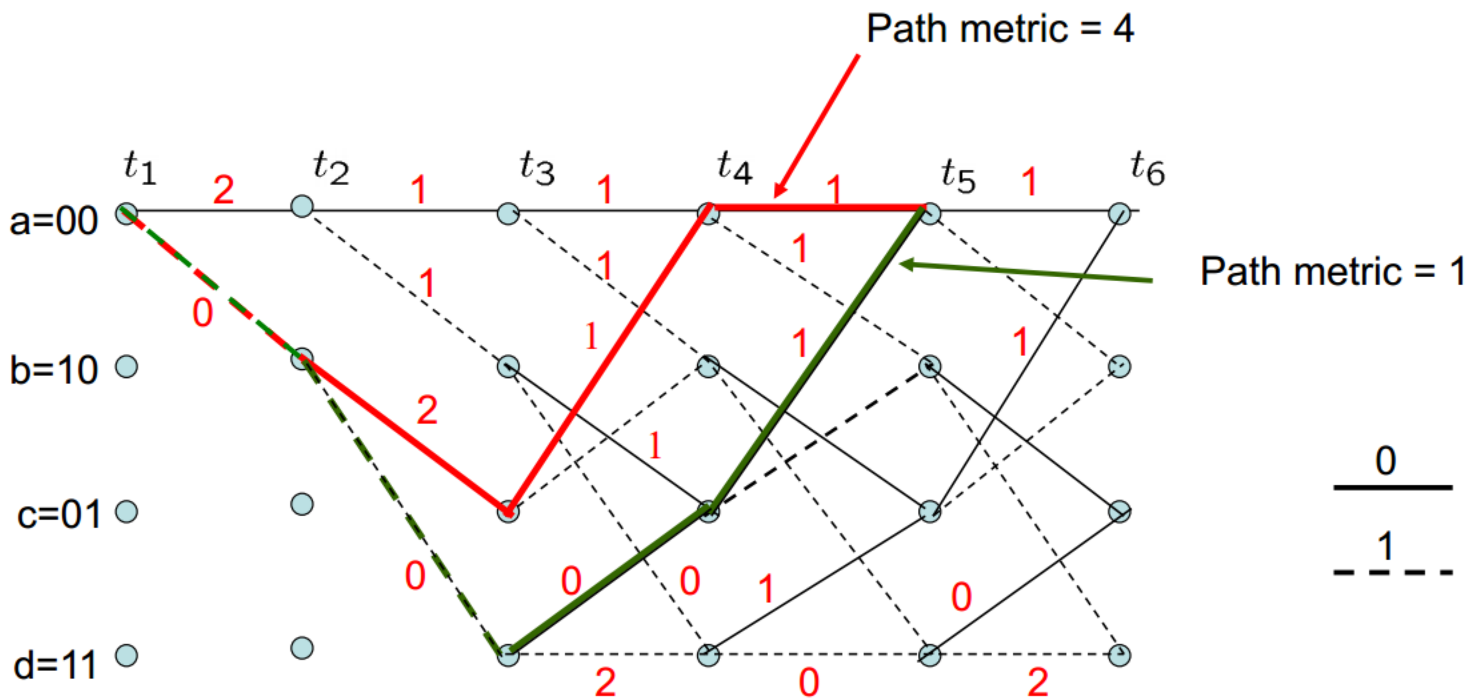


Convolutional codes

- Viterbi decoding
  - Basic idea: If any 2 paths in the trellis merge to a single state, one of them can always be eliminated in the search
  - Let cumulative path metric of a given path at t_i=sum of the branch metrics along that path up to time t_i
  - Consider t_5
    1. The upper path metric is 4, the lower path metric is 1
    2. The upper path metric cannot be path of the optima path since the lower path has a lower metric
    3. This is because future output branches depend on the current state and not the previous state

Convolutional codes

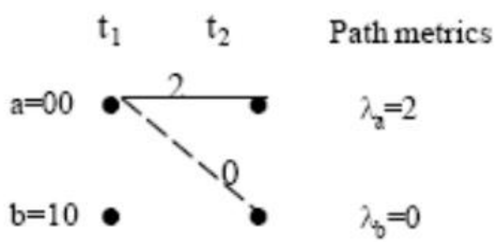
- Viterbi decoding



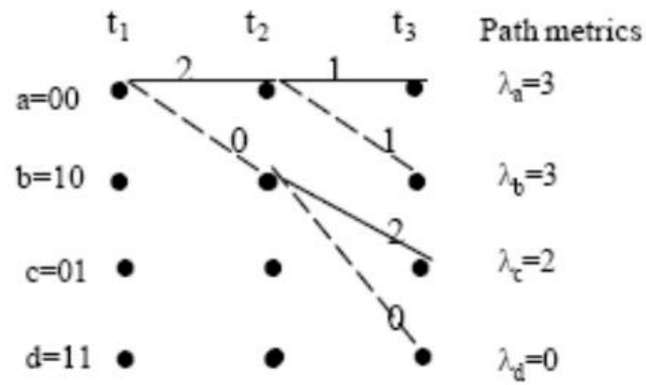
Convolutional codes

- Viterbi decoding
  - At time t_i, there are 2^{L-1} states in the trellis
  - Each state can be entered by means of 2 states
  - Viterbi decoding consists of computing the metric of the 2 paths entering each state and eliminating one of them
  - This is done for each of the 2^{L-1}nodes at time ti
  - The decoder then moves to time t_{i+1} and repeat the process

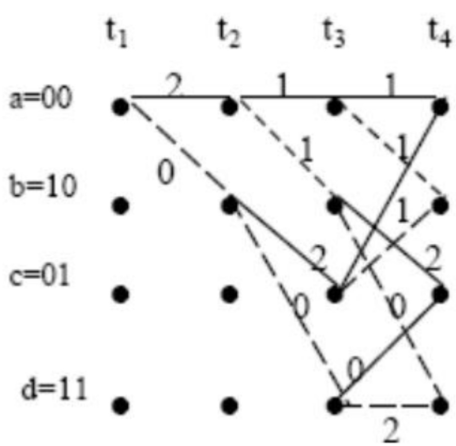
- Viterbi decoding
  - Example.

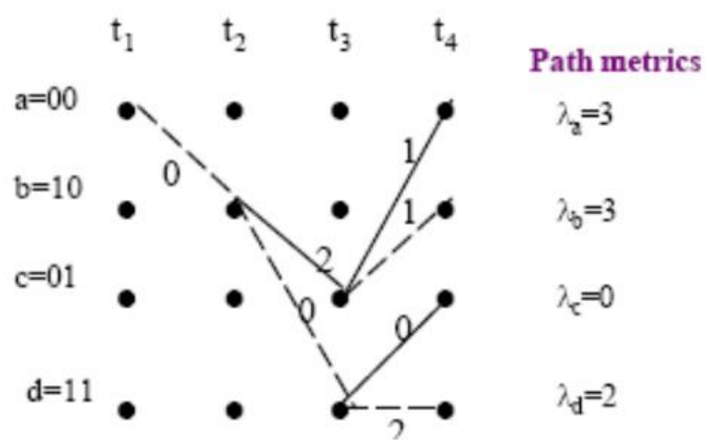


(a)



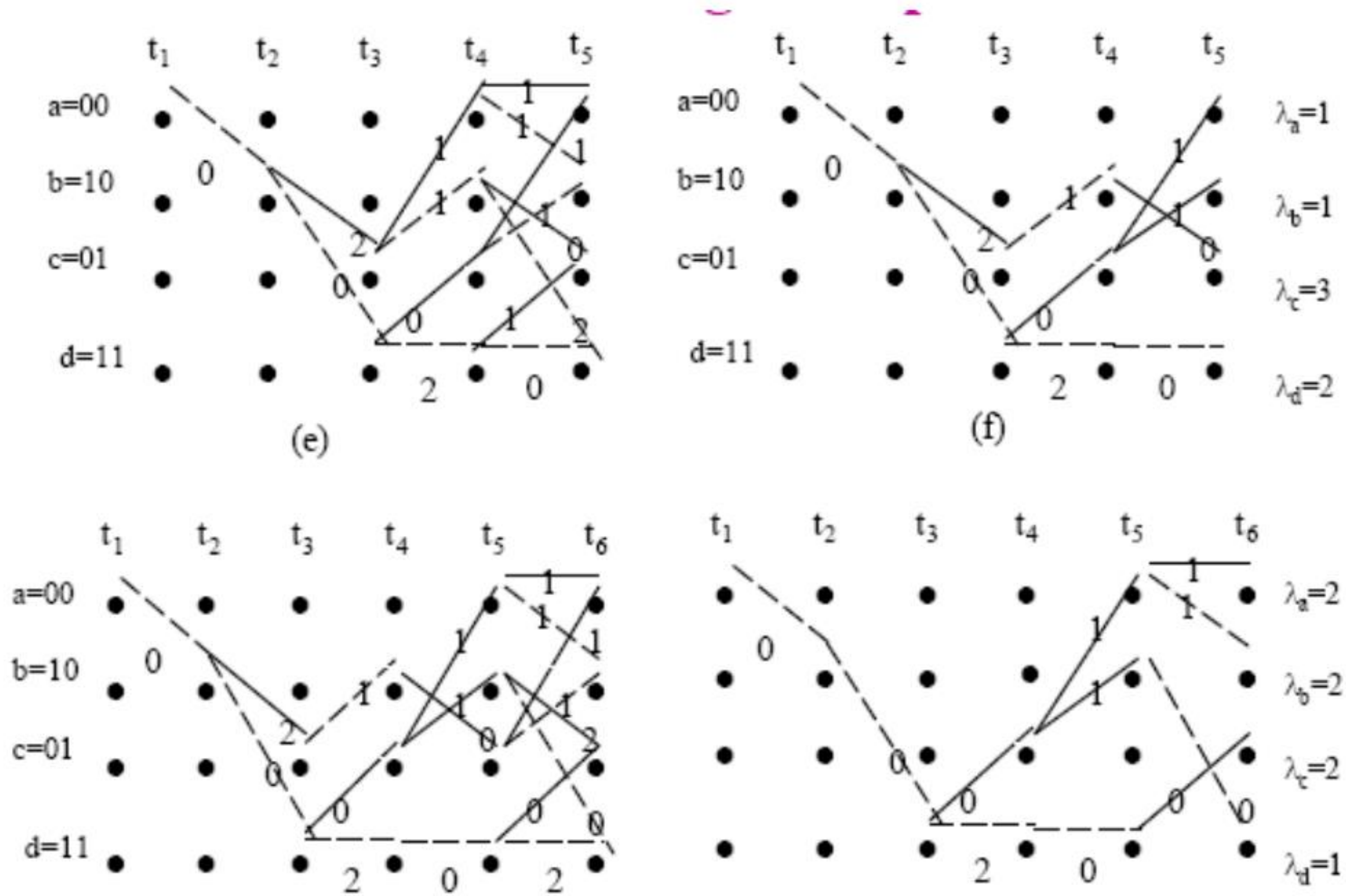
(b)





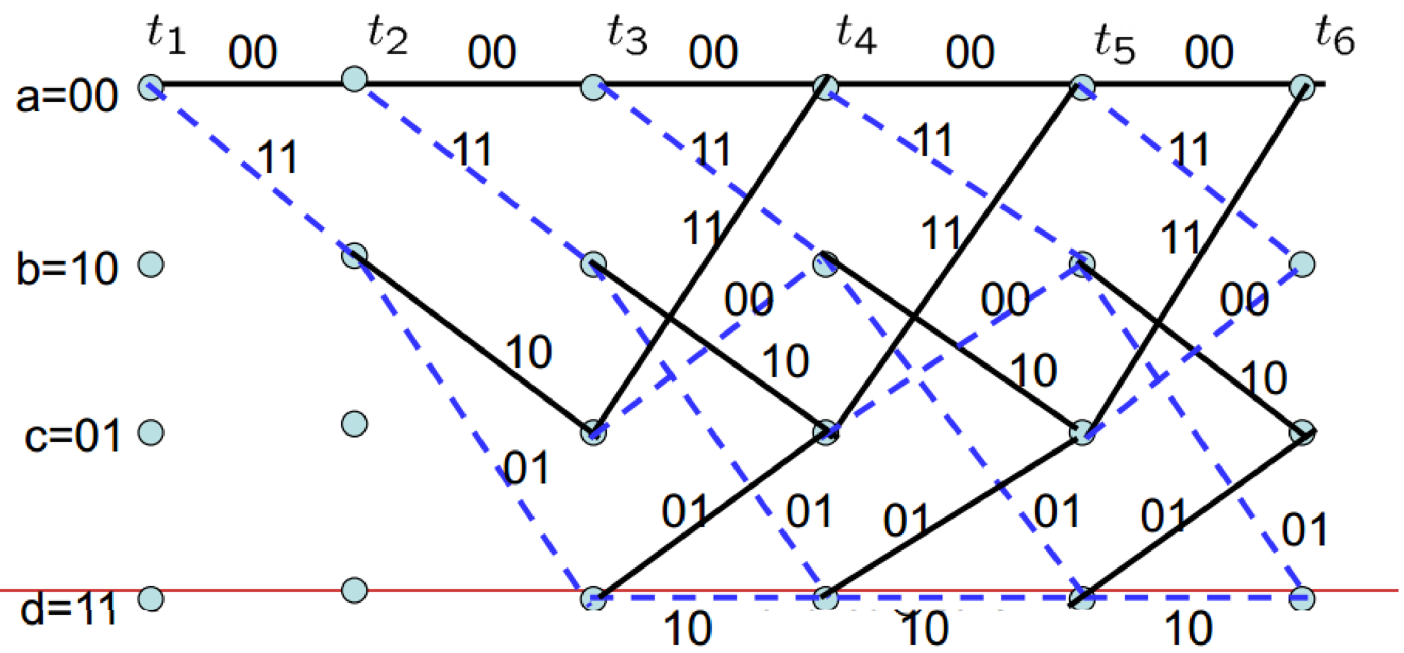
Convolutional codes

- Viterbi decoding
  - ➢ Example.



Convolutional codes

- Viterbi decoding
  - d_{free}=Minimum free distance=Minimum distance of any pair of arbitrarily long paths that diverge and remerge
  - A code can correct any t channel errors where (this is an approximation) $t \leq \lfloor \frac{d_{\text{free}} - 1}{2} \rfloor$

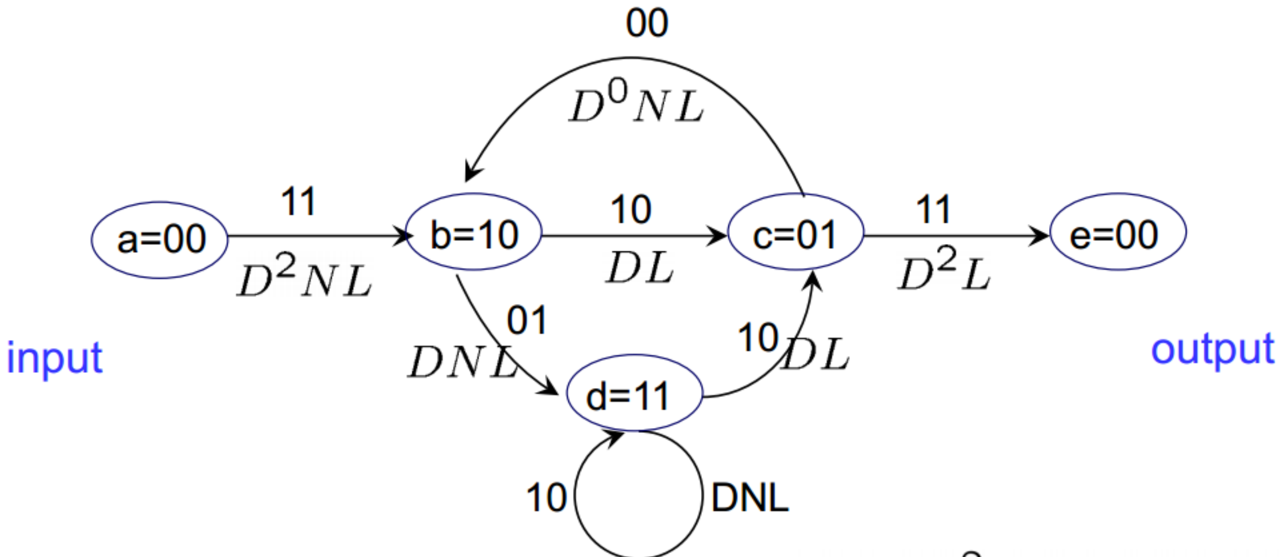


Convolutional codes

- Transfer function
  - The distance properties and the error rate performance of a convolutional code can be obtained from its transfer function
  - Since a convolutional code is linear, the set of Hamming distances of the code sequences generated up to some stages in the trellis, from the all-zero code sequence, is the same as the set of distances of the code sequences with respect to any other code sequence
  - Thus, we assume that the all-zero path is the input to the encoder

Convolutional codes

- Transfer function
  - State diagram labeled according to distance from all-zero path



D^m denote m non-zero output bits

N if the input bit is non-zero

L denote a branch in the path

$$\begin{cases} X_b = D^2NLX_a + LNX_c \\ X_c = DLX_b + DLX_d \\ X_d = DNLX_b + DNLX_d \\ X_e = D^2LX_c \end{cases}$$

Convolutional codes

- Transfer function
  - The transfer function T(D,N,L), also called the wieght enumerating function of the code is

$$T(D,N,L) = \frac{X_e}{X_a}$$

  - By solving the state equations we get

$$T(D,N,L) = \frac{D^5NL^3}{1 - DNL(1+L)}$$

$$= D^5NL^3 + D^6N^2L^4(1+L) + D^7N^3L^5(1+L)^2$$

$$+ \ldots + D^{l+5}N^{l+1}L^{l+3}(1+L)^l + \ldots$$

  - The transfer functions indicates that
    1. There is one path at distance 5 and length 3, which differs 1 bit from the correct all-zeros path
    2. There are 2 paths at distance 6, one of which is of length 4, the other length 5, and both differ in 2 input bits from all-zeros path
    3. $d_{\text{free}} = 5$

Convolutional codes

- Good convolutional codes
  - Good convolutional codes can only be found in general by computer search
  - They are listed in tables and classified by their constraint length, code rate, and their generator polynomials or vectors (typically using octal notation).
  - The error-correction capability of a convolutional code incrases as n increases or as the code rate decreases.
  - Thus, the channel bandwidth and decoder complexity increases.

Convolutional codes

- Good convolutional codes
  - Rate 1/2.

Constraint Length	Generator Polynomials	d_{free}
3	(5,7,7)	8
4	(13,15,17)	10
5	(25,33,37)	12
6	(47,53,75)	13
7	(133,145,175)	15
8	(225,331,367)	16
9	(557,663,711)	18
10	(1117,1365,1633)	20

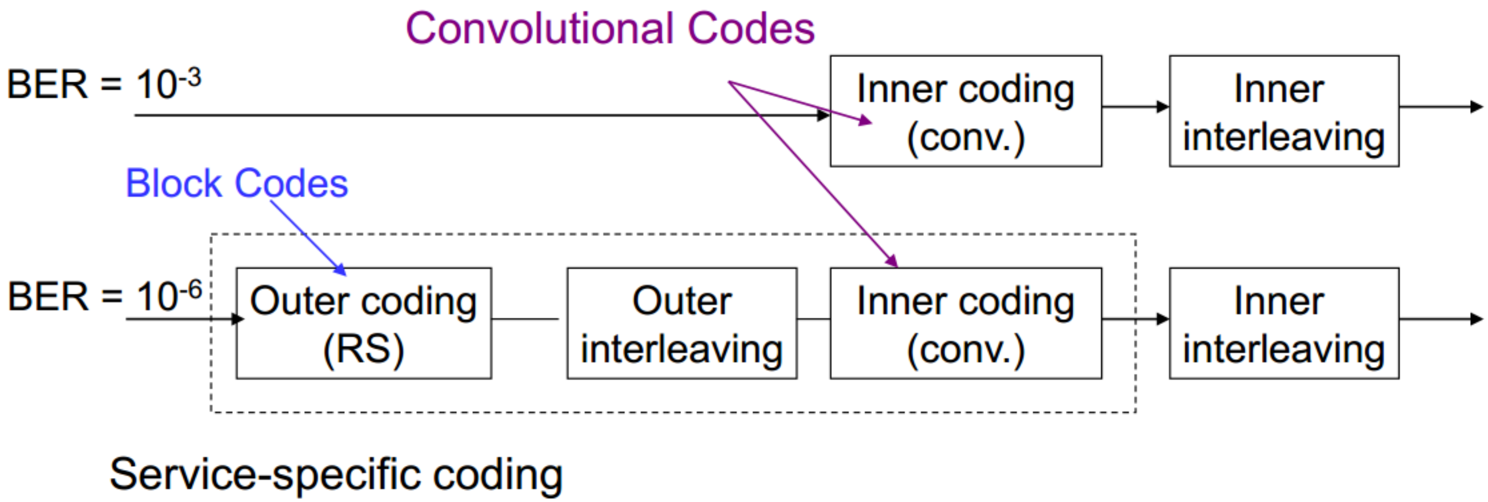
Convolutional codes

- Good convolutional codes
  - Rate 1/3.

Constraint Length	Generator Polynomials	d_{free}
3	(5,7)	5
4	(15,17)	6
5	(23,35)	7
6	(53,75)	8
7	(133,171)	10
8	(247,371)	10
9	(561,753)	12
10	(1167,1545)	12

Convolutional codes

- Channel coding for Wideband CDMA



Convolutional code is rate 1/3 and rate ½,
all with constraint length 9

Convolutional codes

- Channel coding for Wireless LAN (IEEE 802.11a)

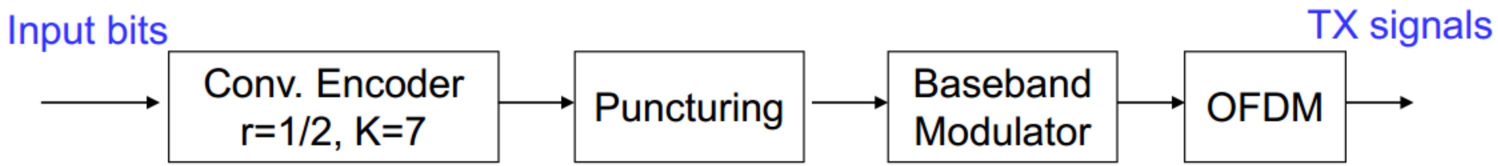


Table 11-3. Encoding details for different OFDM data rates

Speed (Mbps)	Modulation and coding rate (R)	Coded bits per carrier[a]	Coded bits per symbol	Data bits per symbol[b]
6	BPSK, R=1/2	1	48	24
9	BPSK, R=3/4	1	48	36
12	QPSK, R=1/2	2	96	48
18	QPSK, R=3/4	2	96	72
24	16-QAM, R=1/2	4	192	96
36	16-QAM, R=3/4	4	192	144
48	64-QAM, R=2/3	6	288	192
54	64-QAM, R=3/4	6	288	216

Convolutional codes

- Other advanced channel coding
  - Low density parity check codes: Robert Gallager 1960
  - Turbo codes: Berrou et al. 1993
  - Trellis-coded modulation: Ungerboeck 1982
  - Space-time coding: Vahid Tarokh et al. 1998
  - Polar codes: Erdal Arkan 2009

Check the latest coding techniques in 5G standards

Outline

- Introduction
- Digital transmission through baseband channels
- Signal space representation
- Optimal receivers
- Digital modulation techniques [Matlab]
- Multicarrier Communications & OFDM [Matlab]
- Spread Spectrum [Matlab]
- Channel coding [Matlab]
- **Synchronization**

Synchronization

- Synchronization is one of the most critical functions of a communication system with coherent receiver. To some extent, it is the basis of a synchronous communication system.
- Three kinds of synchronization: Carrier synchronization, Symbol/Bit synchronization, and Frame synchronization.
- Carrier synchronization (载波同步): Receiver needs estimate and compensate for frequency and phase differences between a received signal's carrier wave and the receiver's local oscillator for the purpose of coherent demodulation, no matter it is analog or digital communication systems.

Chapter 8.8-8.9

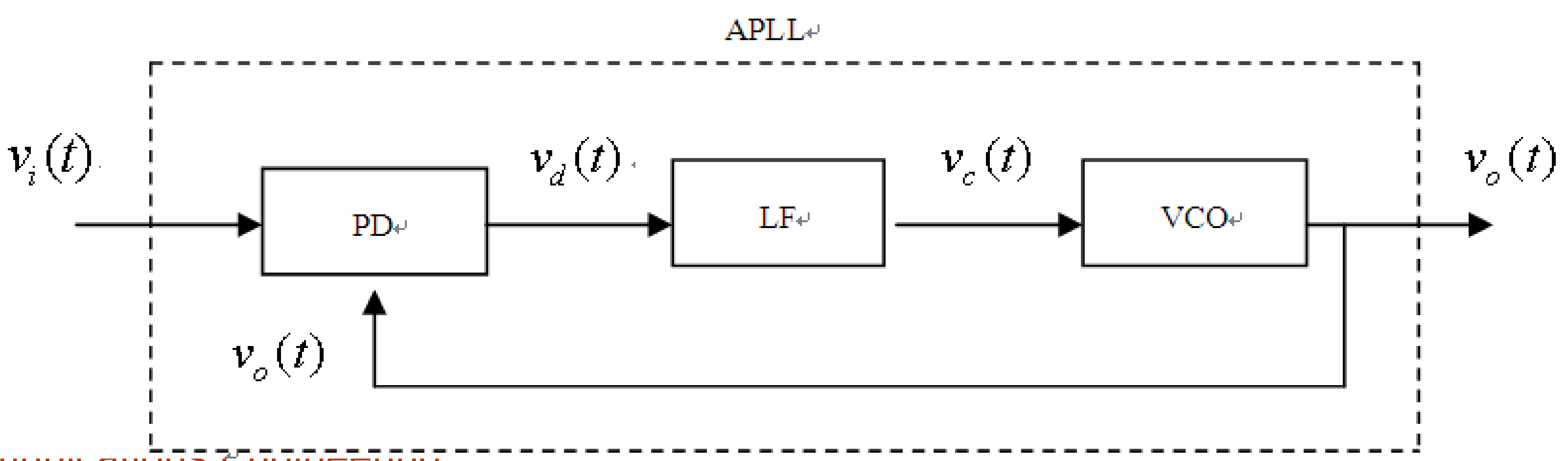
韩声栋等，《通信原理》

Synchronization

- Symbol/Big synchronization (符号/位同步): In digital systems, the output of the receiving filter (i.e. matched filter) must be sampled at the symbol rate and at the precise sampling time instants. Hence, we require a clock signal. The process of extracting such a clock signal at the receiver is called symbol/bit synchronization.
- Frame synchronization (帧同步): In frame-based digital systems, receiver also needs to estimate the starting/stopping time of a data frame. The process of extracting such a clock signal is called frame synchronization.

Synchronization

- Phase-Locked Loop (PLL, 锁相环)
  - ➢ PLL is often used in carrier syn. and symbol syn. It is a closed-loop control system consisting of
  - ➢Phase detector (**PD**): generate the phase difference of $v_i(t)$ and $v_o(t)$.
  - ➢Voltage-controlled oscillator (**VCO**): adjust the oscillator frequency based on this phase difference to eliminate the phase difference. At steady state, the output frequency will be exactly the same with the input



Synchronization

- Phase-Locked Loop (PLL, 锁相环)

$$v_i(t) = v_i \sin[\omega_0 t + \phi(t)]$$

$$v_o(t) = v_o \cos[\omega_0 t + \hat{\phi}(t)]$$

➢A PD contains a multiplier and a low-pass filter. The output of PD is:

$$v_d(t) = \mathrm{K}_d \sin[\phi(t) - \hat{\phi}(t)] = \mathrm{K}_d \sin \phi_e(t)$$

➢LF is also a LPF. The output of the LF is (where F(p) is the transfer function)

$$v_c(t) = F(p) v_d(t)$$

Synchronization

- Phase-Locked Loop (PLL, 锁相环)
  - ➢The output of VCO can be a sinusoid or a periodic impulse train. The differentiation of the output frequency are largely proportional to the input voltage.

$$\frac{d\hat{\phi}(t)}{dt} = \mathrm{K}_v v_c(t)$$

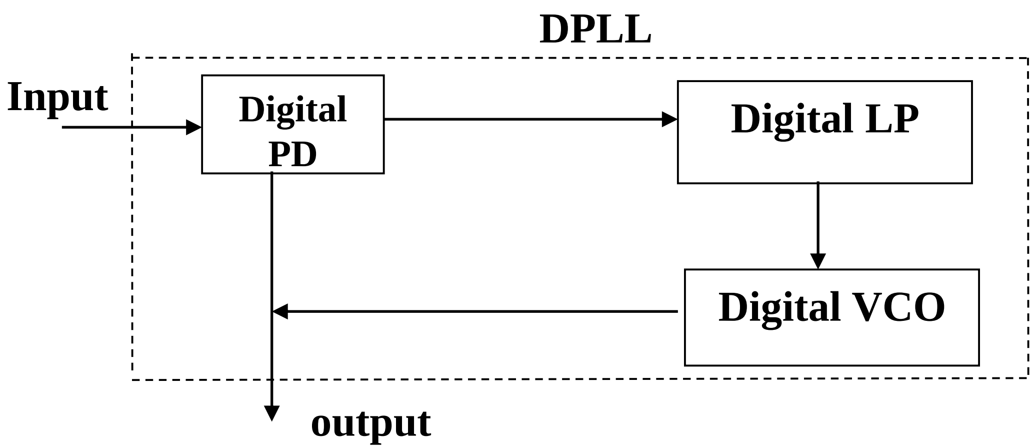
  - ➢If F(p)=1， then

$$\frac{d\hat{\phi}(t)}{dt} = \mathrm{K}\sin\phi_e(t)$$

The first kind of loop!

Synchronization

- Phase-Locked Loop (PLL, 锁相环)
  - ➢Digital PLL.



DPLL

Synchronization

- Phase-Locked Loop (PLL, 锁相环)
  - ➢In a coherence system, a PLL is used for:

    1.PLL can track the input frequency and generate the output signal with small phase difference.

    2.PLL has the character of narrowband filtering which can eliminate the noise introduced by modulation and reduce the additive noise.

    3.Memory PLL can sustain the coherence state for enough time.

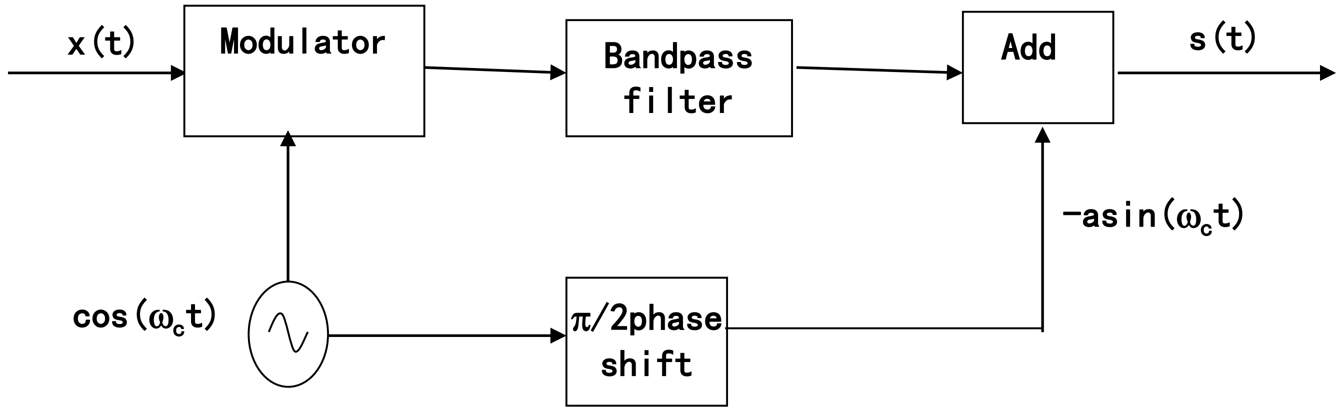
  - ➢CMOS-based integrated PLL has several advantages such as ease of modification, reliable and low power consumption, therefore are widely used in coherence system.

Carrier synchronization

- To extract the carrier:
  - Pilot-tone insertion method: Sending a carrier component at specific spectral-line along with the signal component. Since the inserted carrier component has high frequency stability, it is called pilot (导频).
  - Direct extraction method: Directly extract the synchronization information from the received signal component.

Carrier synchronization

- Pilot-tone insertion method:
  - ➢Insert pilot to the modulated signal

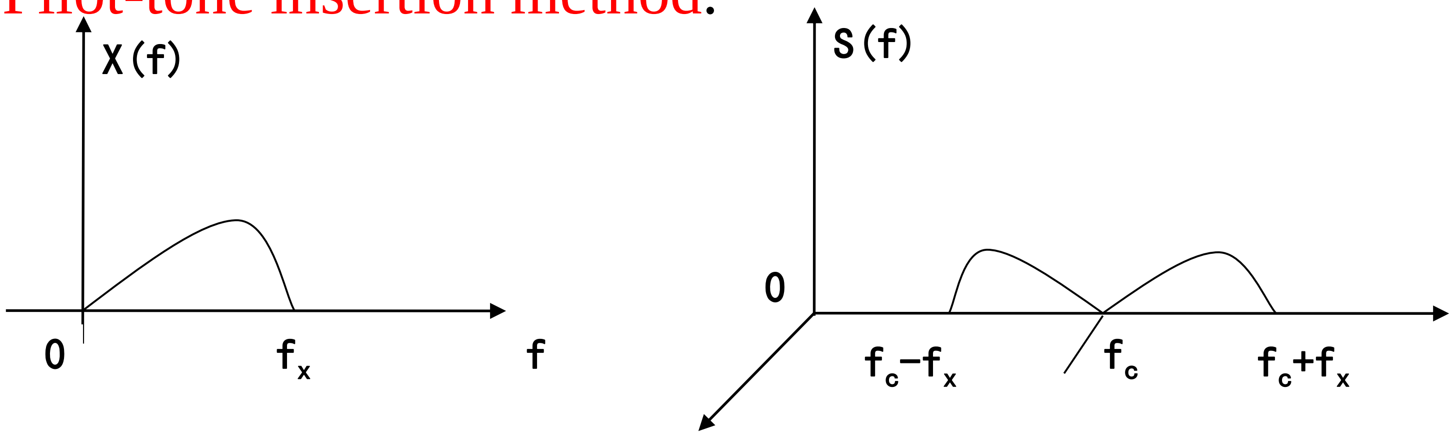


➢The pilot signal is generated by shift the carrier by 90^0 and decrease by several dB, then add to the modulated signal. Assume the modulated signal has 0 DC component, then the pilot is

$$s(t) = f(t)\cos\omega_c t - a\sin\omega_c t$$

Carrier synchronization

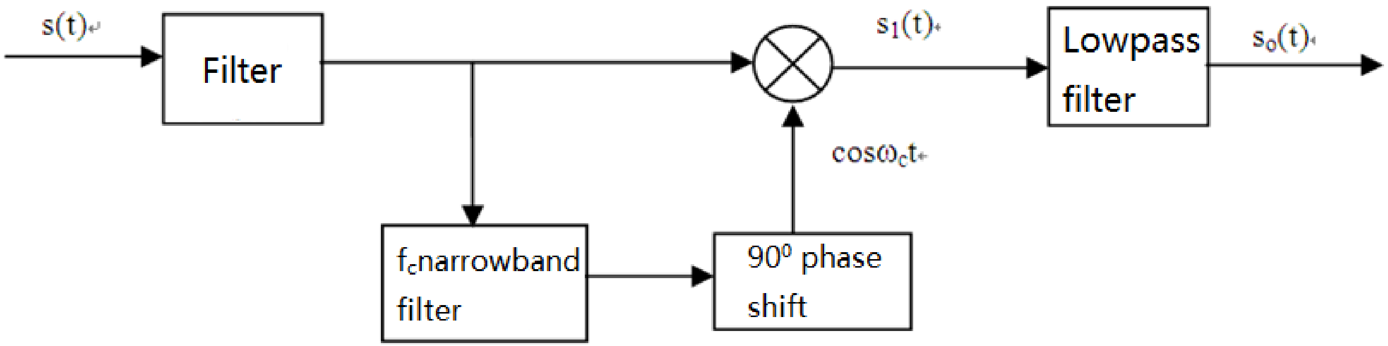
- Pilot-tone insertion method:



➢The receiver uses a narrowband filter with central frequency f_c to extract the pilot $a\sin\omega_c t$ and then the carrier $a\cos\omega_c t$ can be generated by simply shifting 90^0.

Carrier synchronization

- Pilot-tone insertion method:
  - Narrowband filter receiver structure



$$s_1(t) = s(t)\cdot\cos\omega_c t = f(t)\cos^2\omega_c t - a\sin\omega_c t\cos\omega_c t$$

$$= \frac{1}{2}f(t) + \frac{1}{2}f(t)\cos 2\omega_c t - \frac{1}{2}a\sin 2\omega_c t$$

After the LPF $s_0(t) = \frac{1}{2}f(t)$

- DSB, SSB and PSK are all capable of pilot-tone insertion method. VSB can also apply pilot-tone insertion method but with certain modification.

Carrier synchronization

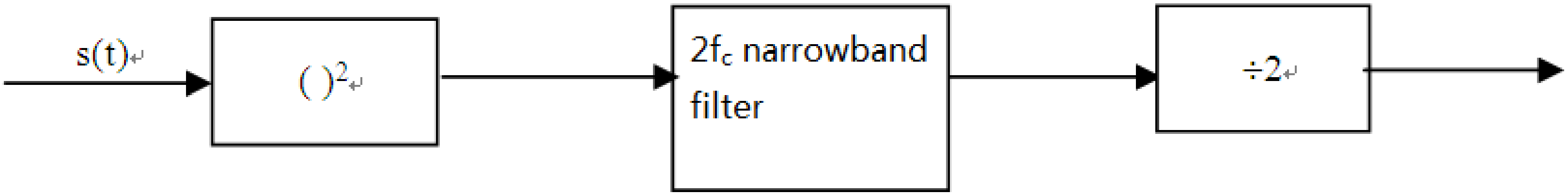
- Pilot-tone insertion method:
  - ➢The drawback of narrowband filter receiver includes:
    1.The pass band is not narrow enough
    2.f_c is fixed, cannot tolerate any frequency drift with respect to the central frequency
    3.Can be replaced by PLL
  - ➢Pilot-tone insertion method is suitable for DSB, SSB, VSB and 2PSK

Carrier synchronization

- Direct extraction method:
  - If the spectrum of the received signal already contains carrier component, then the carrier component can be extracted simply by a narrowband filter or a PLL
  - If the modulated signal suppresses the carrier component, then the carrier component may be extracted by performing nonlinear transformation or using a PLL with specific design.

Carrier synchronization

- Nonlinear transformation based method:
  - Square transformation



Example: a DSB signal $s(t) = f(t)\cos\omega_c t$

If $f(t)$ has 0 DC component, then $s(t)$ *does not have carrier component*

square transformation: $s^2(t) = \frac{1}{2}f^2(t) + \frac{1}{2}f^2(t)\cos 2\omega_c t$

now $f^2(t)$ contains DC component, let it be α, so: $f^2(t) = \alpha + f_m(t)$

then $s^2(t) = \frac{1}{2}\alpha + \frac{1}{2}f_m(t) + \frac{1}{2}\alpha\cos 2\omega_c t + \frac{1}{2}f_m(t)\cos 2\omega_c t$

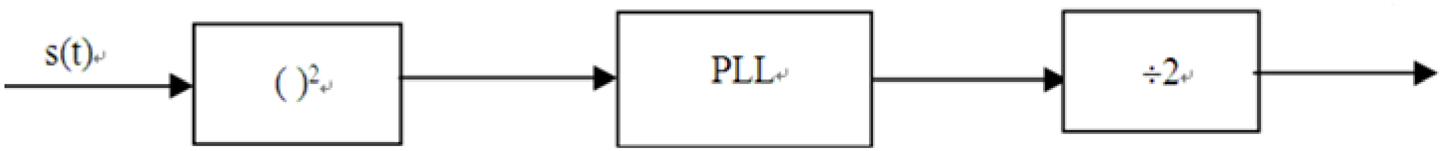
Carrier synchronization

- Nonlinear transformation based method:
  - Square transformation

$$s^2(t) = \frac{1}{2}\alpha + \frac{1}{2}f_m(t) + \frac{1}{2}\alpha\cos 2\omega_c t + \frac{1}{2}f_m(t)\cos 2\omega_c t$$

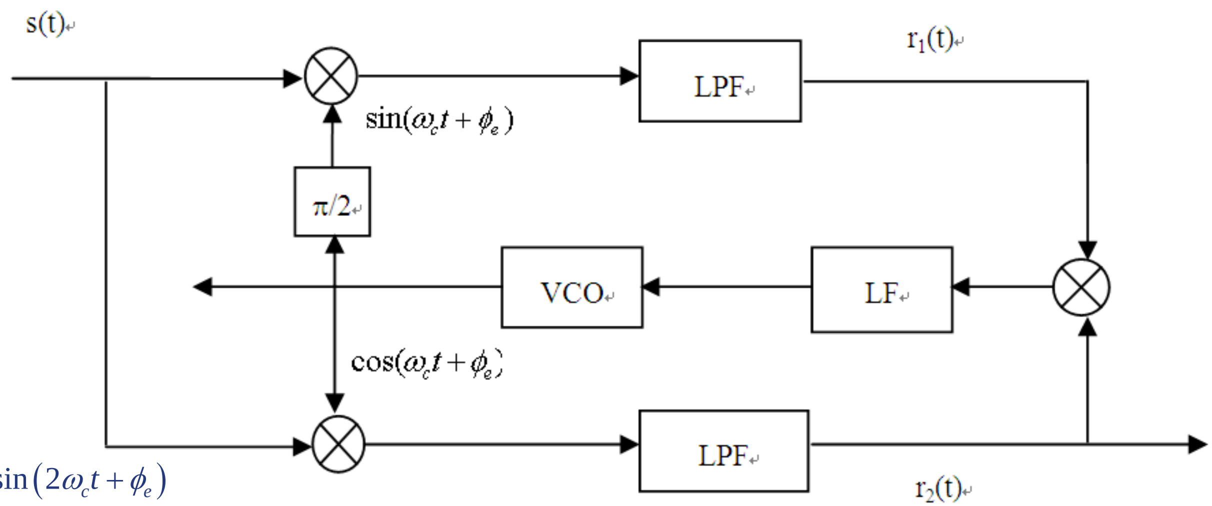
The first term is the DC component. The second term is the low frequency component. The third term is the $2\omega_c$ component. The 4th term is the frequency component symmetrical distributed of $2\omega_c$—modulation noise. After narrowband filtering, only the 3rd term and a small fraction of 4th term left, then the carrier component can be extracted by frequency division.

Since the carrier is extracted by frequency division, its phase may shift by 180°. Besides, modulation noise may cause random phase jitter.



Carrier synchronization

- Nonlinear transformation based method:
  - ➢In-phase orthogonal loop (**Costas loop**)



Let $s(t) = f(t)\cos\omega_c t$

(1) upper branch

$$f(t)\cos\omega_c t\cdot\sin(\omega_c t+\phi_e) = \frac{1}{2}f(t)\sin\phi_e + \frac{1}{2}f(t)\sin(2\omega_c t+\phi_e)$$

ϕ_e is the phase difference between generated carrier and the original carrier

After LPF $r_1(t) = \frac{1}{2}f(t)\sin\phi_e$

When ϕ_e is small, $r_1(t) = \frac{1}{2}f(t)\phi_e$

(2) lower branch

$$r_2(t) = \frac{1}{2}f(t)\cos\phi_e \to \frac{1}{2}f(t)$$

$$(3)\ r_1(t)\cdot r_2(t) \to \frac{1}{4}f^2(t)\phi_e = v_d(t)$$

Contains in-phase branch and orthogonal branch. All parts except LF and VCO are similar with a “phase detector”.

Carrier synchronization

- Nonlinear transformation based method:
  - In-phase orthogonal loop (Costas loop)

    Advantages of Costas loop:

    1.Costas loop works on f_c instead of $2f_c$, so when $\mathrm{f_c}$ is large, Costas loop is easier to realize

    2.The output of in-phase loop $\mathrm{r_2}(t)$ is the signal $\mathrm{f}(t)$

- Performance of carrier synchronization technique

  1) Phase error: steady-state phase error, random phase error

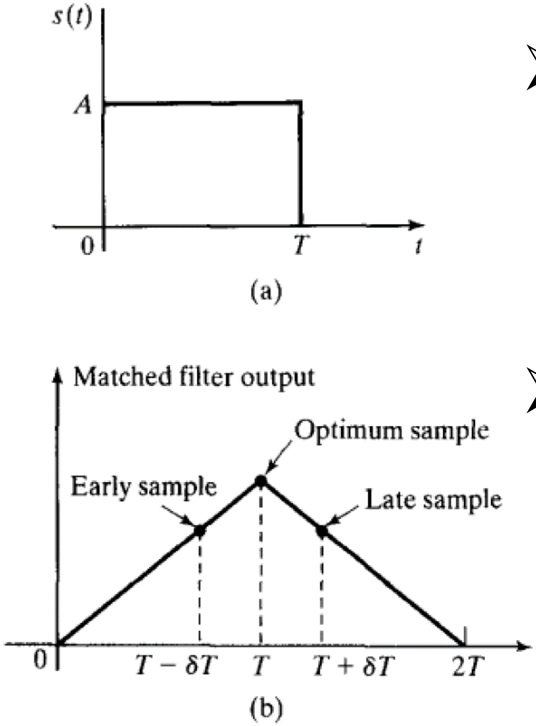
  2) Synchronization build time and hold time

Symbol synchronization

- In a digital communication system, the output of the receiving filter must be sampled periodically at the symbol rate and at the precise sampling time instance.
- To perform this periodic sampling, we need a clock signal at the receiver
- The process of extracting such a clock signal is called symbol synchronization or timing recovery
- One method is for the transmitter to simultaneously transmit the clock frequency along with the information signal. The receive can simply employ a narrowband filter or PLL to extract it. This method requires extra power and bandwidth and hence, but frequently used in telephone transmission systems.
- Another method is to extract the clock signal from the received data signal by using some kind of non-linear transformation.

Symbol synchronization

- Early-late gate synchronization
  - Basic Idea: exploit the symmetry properties of the output signal of matched filter or correlator



  - Due to the symmetry, the values of the correlation function at the early samples $t = T - \delta T$ and the late samples $t = T + \delta T$ are equal.
  - Thus, the proper sampling time is the midpoint between $t = T - \delta T$ and $t = T + \delta T$

Figure 8.48 (a) Rectangular signal pulse and (b) its matched filter output.

Symbol synchronization

- Early-late gate synchronization
  - ➢ Block diagram.

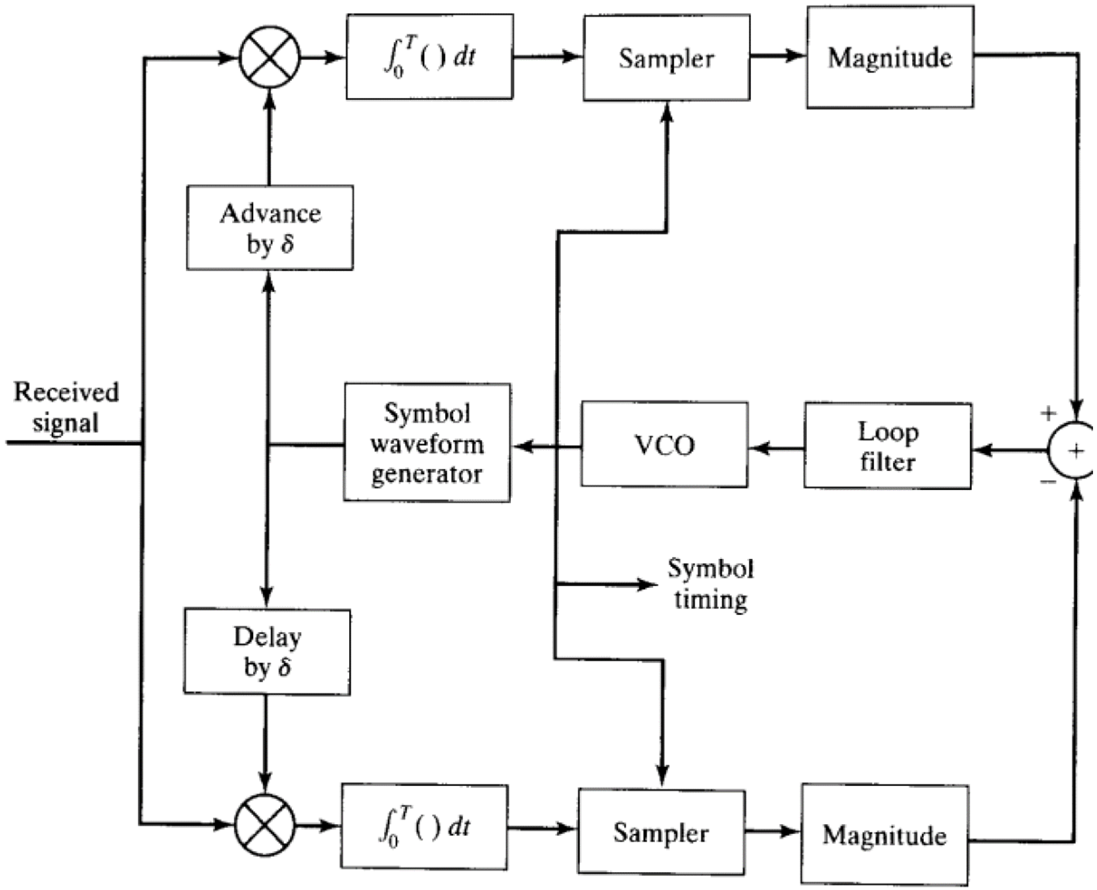
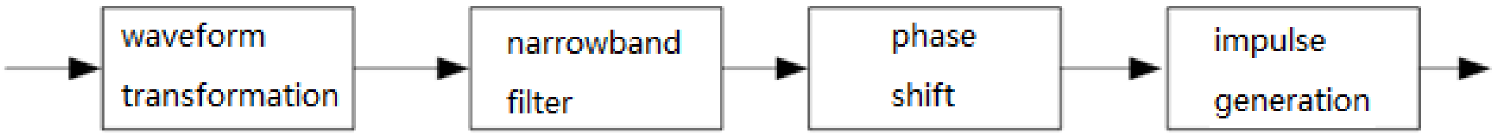


Figure 8.49 Block diagram of early–late gate synchronizer.

Symbol synchronization

- Nonlinear transformation based synchronization

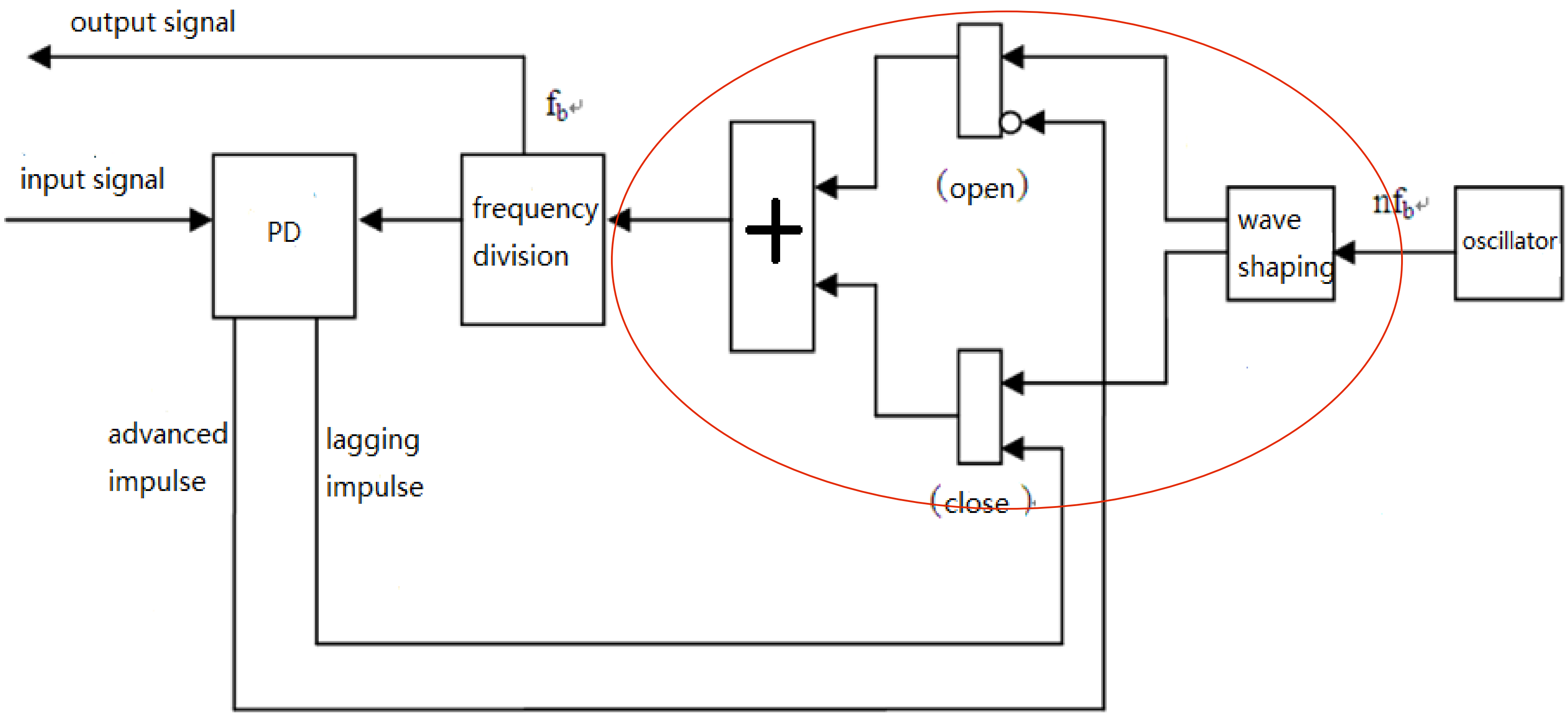


  - Some transformations can add synchronous signal with f=1/T to the original signal. For example, we can transform the signal to return-to-zero waveform. After narrowband filtering and phase shifting, we can generate the clock signal used for synchronization.

$$P_s(f) = f_s P(1-P)\left|G_1(f) - G_2(f)\right|^2 + f_s^2 \sum_{m=-\infty}^{\infty} \left|P G_1(m f_s) + (1-P) G_2(m f_s)\right|^2 \delta(f - m f_s)$$

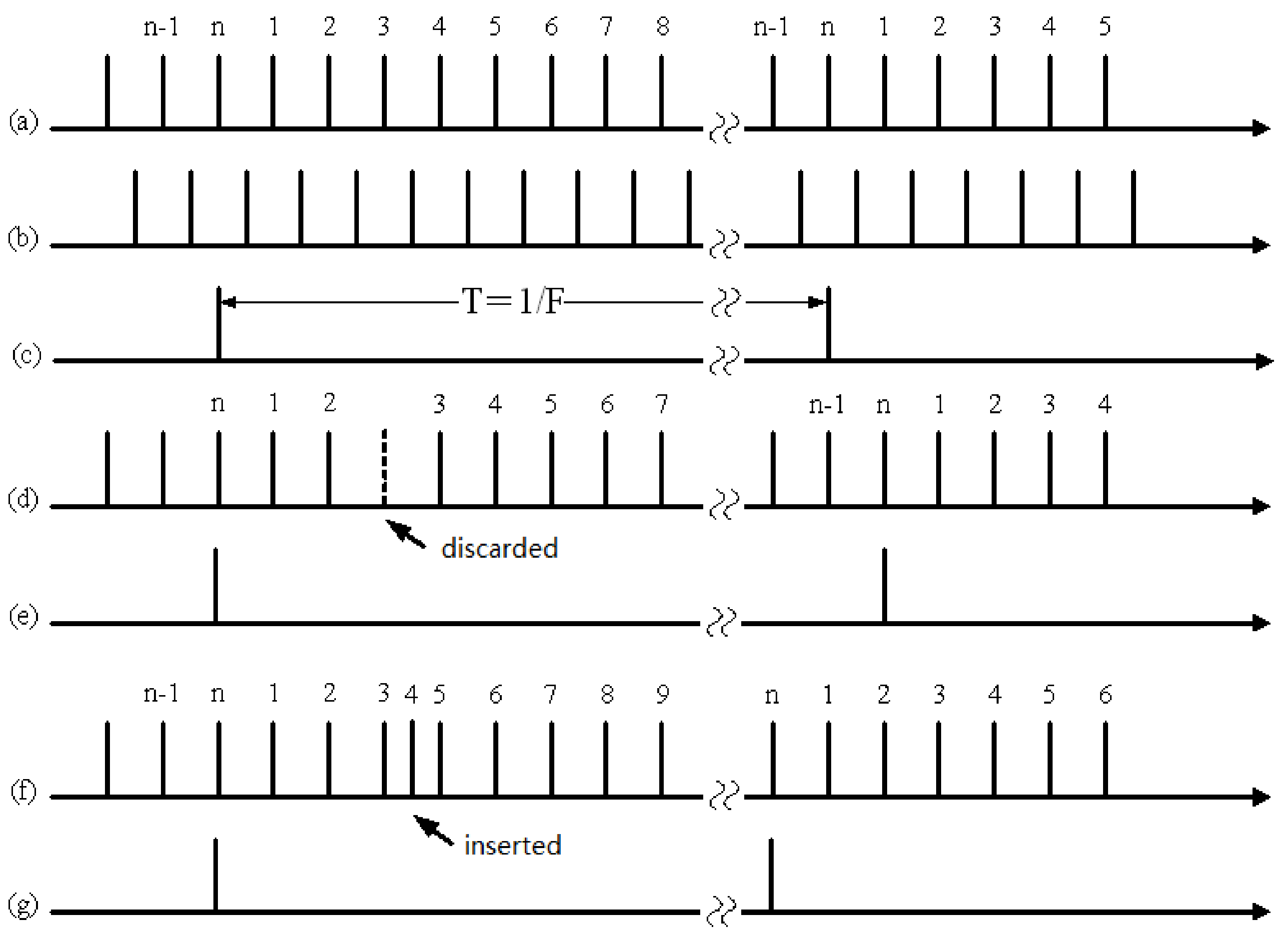
Symbol synchronization

- DPLL



Symbol synchronization

- DPLL



Symbol synchronization

- DPLL
  - Performance.

    1). Phase error

    2). Synchronization build time

    3). Synchronization hold time

    4). Synchronous bandwidth

Frame synchronization

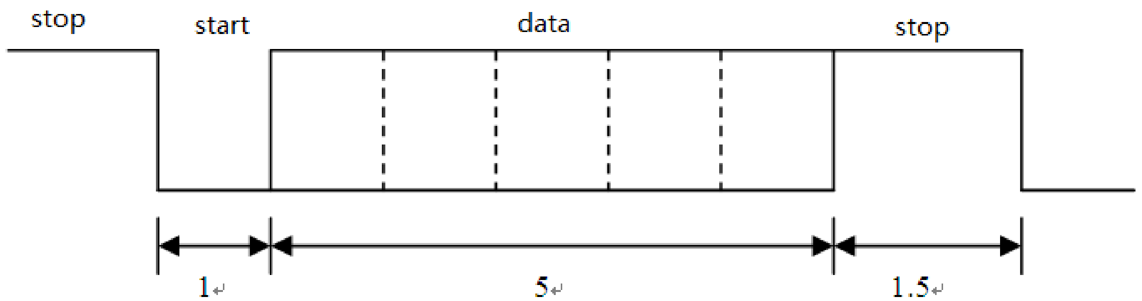
- Recall that carrier and symbol synchronization needs to estimate the phase of synchronous signal which can be realized by using a PLL.
- Frame synchronization is to insert frame alignment signal (distinctive bit sequence) and then detect the alignment symbol.
- Besides adding frame alignment bits, some code such as self-synchronizing code can be synchronized without adding extra bits.
- Here, we only focus on the first method ——inserting frame alignment signal.

Frame synchronization

- Recall that carrier and symbol synchronization needs to estimate the phase of synchronous signal which can be realized by using a PLL.
- Frame synchronization is to insert frame alignment signal (distinctive bit sequence) and then detect the alignment symbol.
- Besides adding frame alignment bits, some code such as self-synchronizing code can be synchronized without adding extra bits.
- Here, we only focus on the first method ——inserting frame alignment signal.
  - Start-stop method
  - Bunched frame alignment signal
  - Distributed frame alignment signal

Frame synchronization

- Start-stop method
  - It is widely used in teleprinter. Each symbol contains 5-8 data bits, a start bit and a stop bit.



start bit: "0", width: T_b

stop bit: "1", width $\geq T_b$

System will keep sending stop bit when it is idle. When "1" $\rightarrow$ "0", the receiver will start to receive a data symbol.

Low transmission efficiency and low timing accuracy

Frame synchronization

- Bunched frame alignment signal

➢ This method inserts synchronous code at a particular place in each frame. The code should have a sharp self-correlation function. The detector should be simple to implement.

➢ Frame synchronous code includes

   1. Barker code
   2. Optimal synchronous code
   3. Pseudo-random code

Frame synchronization

- Bunched frame alignment signal
  - Barker code

(1) Barker code：

A n bits barker code $\{x_1, x_2, x_3 \cdots x_n\}$, x_i=+1 or −1. its self-correlation function satisfies：

$$R_x(j) = \sum_{i=1}^{n-j} x_i x_{i+j} = \begin{cases} n & j = 0 \\ 0 \text{or} \pm 1 & 0 < j < n \\ 0 & j \geq n \end{cases}$$

Barker code is not a periodic sequence. It is proved that when $n < 12100$, we can only find barker code with $n = 2, 3, 4, 5, 7, 11, 13$.

Frame synchronization

- Bunched frame alignment signal
  - Barker code

n	barker code
2	+ +
3	+ + -
4	+ + + - , + + - +
5	+ + + - +
7	+ + + - - + -
11	+ + + - - - + - - + -
13	+ + + + - - + + - + - +

Frame synchronization

- Bunched frame alignment signal
  - Barker code

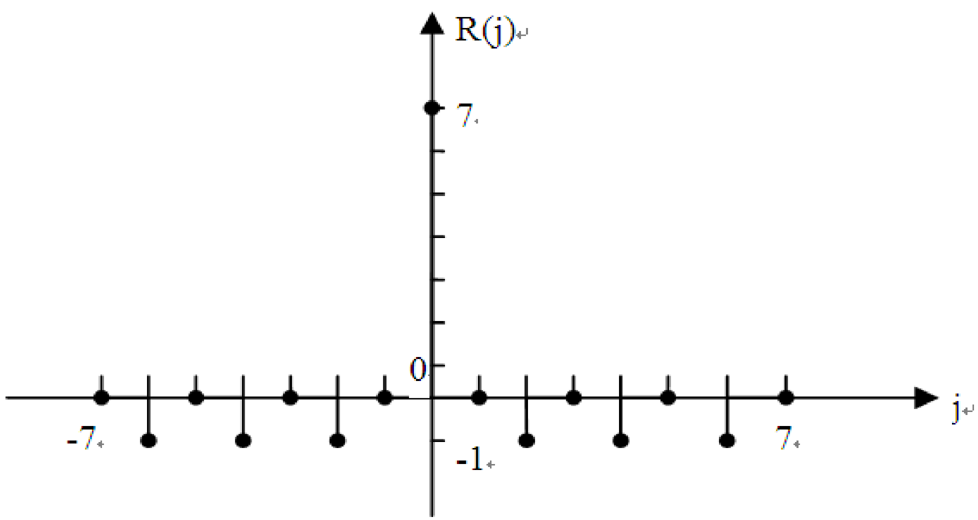
Example: A barker code with $n = 7$, find its self-correlation function

$$j = 0: R_x(0) = x_1x_1 + x_2x_2 + \cdots + x_7x_7 = 7$$

$$j = 1: R_x(1) = x_1x_2 + x_2x_3 + \cdots = 0$$

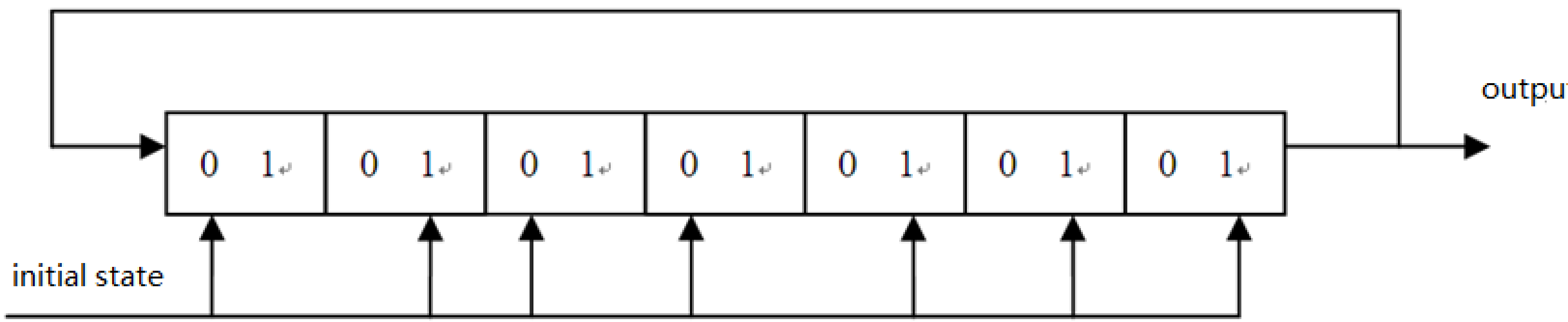
Similarly, we can determine $R_x(j)$.

The result is shown below, we can see it has a sharp peak when $j = 0$.



Frame synchronization

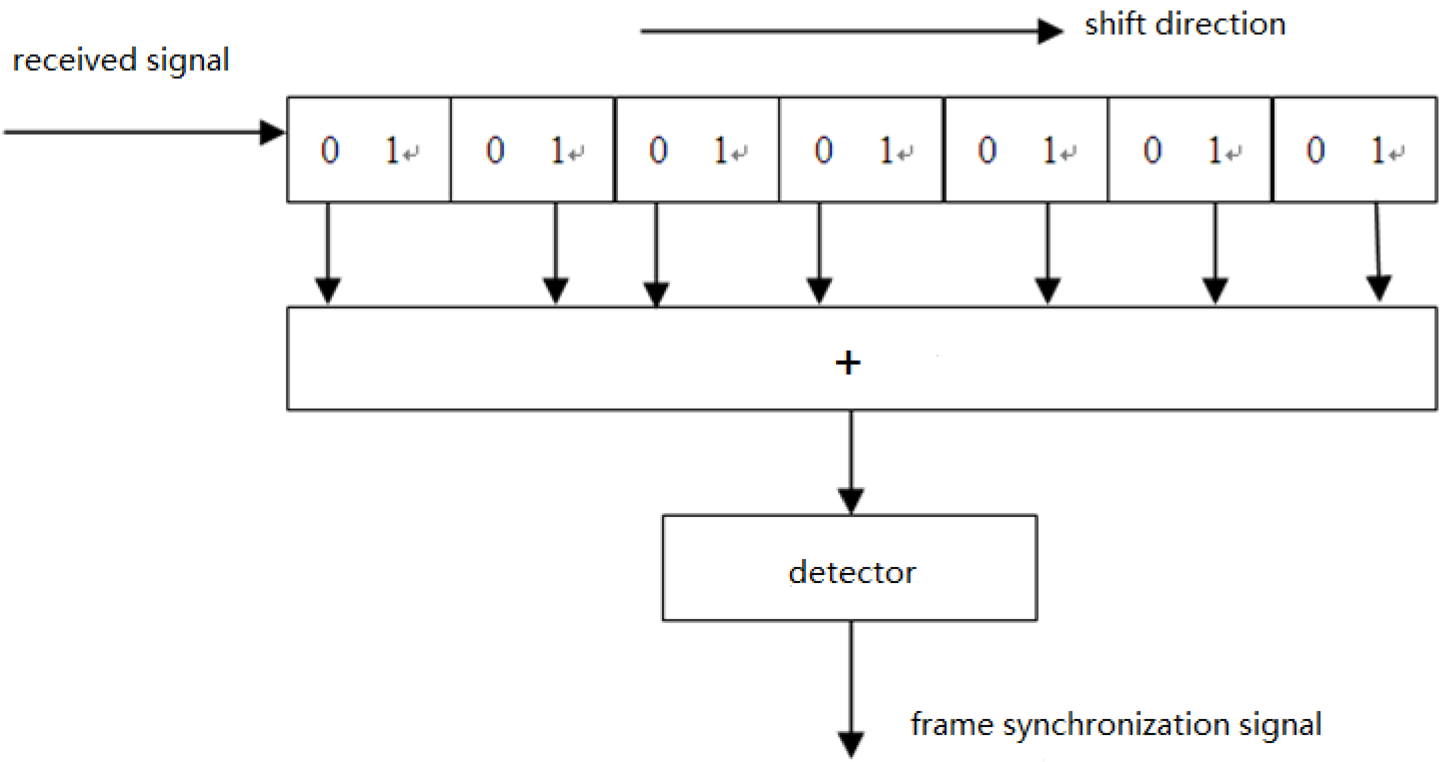
- Bunched frame alignment signal
  - Barker code generator – shift register
  - Example: when n=7, a 7 bits shift register. The initial state is a barker code.



Frame synchronization

- Bunched frame alignment signal

➢ Barker code detector



Frame synchronization

- Bunched frame alignment signal

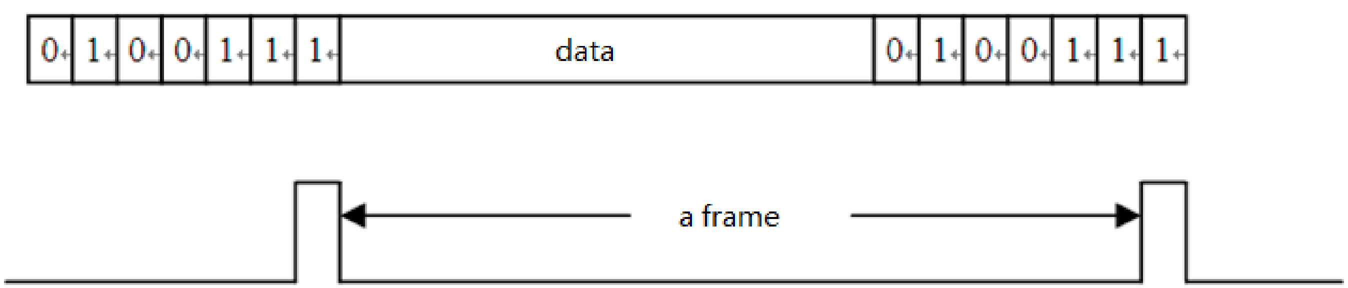
➢ Barker code detector

The barker code detector follows:

$$input:"1"\begin{cases}\text{output "1":? }+1\\ \text{output "0":? }-1\end{cases}$$

$$input:"0"\begin{cases}\text{output "1":? }-1\\ \text{output "0":? }+1\end{cases}$$

If the output connection of the shift register is the same with a barker code, then when the input is a barker code, the output of the shift register is "1111111". The detector will send a synchronous impulse.



Frame synchronization

- Distributed frame alignment signal

➢ The synchronous code is distributed in the data signal. That means between each n bits, a synchronous bit is inserted.

➢ Design criteria of synchronous code:

1. Easy to detect. For example: “11111111”or ”10101010”

2. Easy to separate synchronous code from data code. For example: In some digital telephone system， all ”0” stands for ring, so synchronous code can only use “10101010”

Frame synchronization

- Performance of Bunched frame alignment signal

➢ Probability of missing synchronization P_L

1. Affected by noise, the detector may not be able to detect the synchronous code. The probability of this situation is called probability of missing synchronization P_L.

2. Assume the length of synchronous code is n， bit error rate is Pe. The detector will not be able to detect if more than m bit errors happen, then:

$$P_L = 1 - \sum_{x=0}^{m} C_n^x P_e^x \left(1 - P_e\right)^{n-x}$$

Frame synchronization

- Performance of Bunched frame alignment signal

➤ Probability of false synchronization P_F

Since data code can be arbitrary, it may be the same with synchronous code. The probability of this situation is called probability of false synchronization P_F. P_F equals to the probability of appearance of synchronous code in the data code.

a. In a binary code, assume 0 and 1 appears with the same probability. There are 2^n combinations of a n bit code.

b. Assume when there are more than m bit errors, the data code will also be detected as synchronous code.

Frame synchronization

- Performance of Bunched frame alignment signal

➢ Probability of false synchronization P_F

When $m = 0$, only $1(C_n^0)$ code will be detected as synchronous code;

When $m = 1$, there are C_n^1 codes will be detected as synchronous code;

……

Therefore, the probability of false synchronization is:

$$P_F = \frac{\sum_{x=0}^{m} C_n^x}{2^n} = \left(\frac{1}{2}\right)^n \sum_{x=0}^{\mathrm{m}} C_n^x$$

Frame synchronization

- Performance of Bunched frame alignment signal

➢ Performance

P_L and P_F depends on the length of synchronous code n and the maximum bit error m.

When n↑, P_F↓, P_L↑; when m↑, P_L↓, P_F↑

3. Average build time t_s

Assume both P_L and P_F will not happen, the worst case is we need one frame to build frame synchronization. Assume each frame contains N bits, each bit has a width T_b, then one frame costs NT_b.

Now assume a missing synchronization or a flase synchronization also needs NT_b to rebuild the synchronization, then:

$$t^1{}_s = NT_b\left(1 + P_L + P_F\right)$$

Bisedes, the average build time of using the distributed frame alignment signal is:

$$t^2{}_s = N^2 T_b\left(N >> 1\right)$$

Apparently, $t^1{}_s < t^2{}_s$, so the previous method is more widely used.