Table of Contents

Course Catalog 4
Textbook 4
References 4
Books 4
Instructor 4
Prerequisites 4
Topics Covered 4
Evaluation 5
I. Introduction to Communications 6
I.1. Building Blocks of Communication Systems 6
I.1.A. Transmitter 6
I.1.B. Receiver 6
I.1.C. Channel 7
I.2. Wireless Communications Evolution 7
I.2.A. Brief History 7
I.2.B. 0G (Pre Cellular) 8
I.2.C. 1G 8
I.2.D. 2G 8
I.2.E. 3G 9
I.2.F. 4G 9
I.2.G. 5G 9
I.3. Multiple Access Techniques 10
I.3.A. Time Division Multiple Access (TDMA) 10
I.3.B. Frequency Division Multiple Access (FDMA) 10
I.3.C. Code Division Multiple Access (CDMA) 11
II. Communication over Bandlimited Channels 12
II.1. Preliminaries 12
II.1.A. The Nyquist Criterion (Zero ISI) 16
II.1.B. Design of Band-Limited Signals with Controlled ISI – Partial Response Signals 21
II.1.C. Data Detection for Controlled ISI 26

II.1.D. Signal Design for Channels with Distortion 30

II.2. Optimum Receiver for Channels with ISI and AWGN 30

II.2.A. Optimum Maximum-Likelihood Receiver 30

II.2.B. A Discrete-Time Model for a Channel with ISI 30

II.2.C. Maximum-Likelihood Sequence Estimation (MLSE) for the Discrete-Time White Noise Filter Model 33

II.2.D. Performance of MLSE for Channels with ISI 33

II.3. Linear Equalization 33

II.3.A. Peak Distortion Criterion 34

II.3.B. Mean-Square-Error (MSE) Criterion 37

II.3.C. Performance Characteristics of the MSE Equalizer 40

II.3.D. Probability of error performance of linear MSE equalizer 40

II.4. Fractionally Spaced Equalizers 40

II.5. Baseband and Passband Linear Equalizers 40

II.6. Decision-Feedback Equalization 40

II.7. Reduced Complexity ML Detectors 40

II.8. Iterative Equalization and Decoding—Turbo Equalization 40

III. Wireless Communication Channels 41

III.1. Propagation Effects 41

III.1.A. Large-Scale Effects 41

III.1.B. Small-Scale Effects 43

III.1.C. Multipath Effects 43

III.1.D. Multipath Mitigation 44

III.2. Fading Channel Characterization 44

III.2.A. Fading Classifications 44

III.2.B. Time Varying Mulipath Channel 45

III.2.C. Channel Impulse Response 46

III.2.D. Channel Correlation Functions 47

III.2.E. Multipath Definitions 49

III.3. Time-Dispersive vs. Frequency-Dispersive Fading 50

III.3.A. Time Dispersion: Frequency-Selective Fading Channel 50

III.3.B. Frequency Dispersion: Time-Selective Fading Channel 52

III.4. Statistical Models 53

III.4.A. Small-Scale Fading Parameters 53
III.4.B. Probability Densities 54
III.5. System Performance Measures 55
III.5.A. Average Signal-to-Noise Ratio 55
III.5.B. Outage Probability 56
III.5.C. Average Bit Error Probability 57
III.5.D. Amount of Fading (AF) 58
III.5.E. Average Outage Duration (AOD) 58
III.6. Performance Analysis: Slowly Fading Frequency Non-Selective Channels 59
III.7. Diversity and Combining 63
III.7.A. Types of diversity: 63
III.7.B. Types of Combining 63
III.7.C. Binary PSK- MRC 63
III.7.D. MRC – Asymptotic P(e) 64
III.8. Frequency Selective Channels 65
III.8.A. Tapped-Delay-Line Channel Model 65
III.8.B. Rake Receiver 67
III.8.C. Performance of Rake Demodulator 69
IV. Spread Spectrum Communication Systems 73
IV.1. Preliminaries 74
IV.1.A. Spreading Codes 74
IV.1.B. Maximal-Length PN Codes 75
IV.1.C. PN Code Examples 75
IV.1.D. LFSR Tap Connections 77
IV.2. Direct Sequence Spread Spectrum 77
IV.2.A. DS Signal Generation – Method 1 78
IV.2.B. DS Signal Generation – Method 2 78
IV.3. Frequency Hopping Spread Spectrum 78
IV.4. Model of Spread Spectrum Digital Communication System 79
IV.5. Direct Sequence Spread Spectrum Signals 80
IV.5.A. Error Rate Performance of the Decoder 82

SYLLABUS

Course Catalog

3 Credit hours (3 h lectures). Digital signaling over channels with intersymbol interference and AWGN. Wireless multipath channel models: time and frequency dispersive channels, level crossing and average fade duration. Diversity concepts: modeling and error probability performance evaluation. Spread spectrum in digital transmission over multipath fading channels, performance analysis and fading mitigation techniques.

Textbook

Several books and journal articles.

References

BOOKS

1. John G. Proakis and Masoud Salehi, *Digital Communications*, 5th ed., McGraw-Hill, 2008
2. Marvin K. Simon and Mohamed-Slim Alouini, *Digital Communication over Fading Channels*, 2nd ed., Wiley, 2005
3. Andrea Goldsmith, *Wireless Communications*, Cambridge University Press, 2005

Instructor

Instructor: Dr. Mohammad M. Banat

Email Address: banat@just.edu.jo

Prerequisites

- Background in linear algebra, signal analysis, random processes and DSP
- Graduate standing in digital communications

Topics Covered

Week	Topics
1-2	Introduction
3-6	Communications over Bandlimited Channels
7-12	Communication over Fading Channels
13-14	Spread Spectrum Techniques
15-16	Miscellaneous Topics

Evaluation

Assessment Tool	Expected Due Date	Weight
Mid-Term Exam	22 April 2021	25%
Term Project Report	6 May 2021	15%
Term Project Presentation	20 May 2021	10%
Final Exam		50%

I. Introduction to Communications

I.1. Building Blocks of Communication Systems

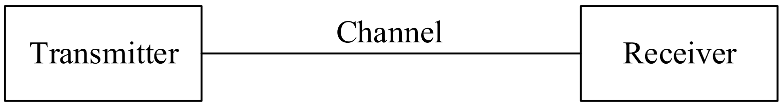


Figure I.1

Transmitter:	Transmits data	Channel:	Transmission Medium	Receiver:	Receives Data

I.1.A. Transmitter

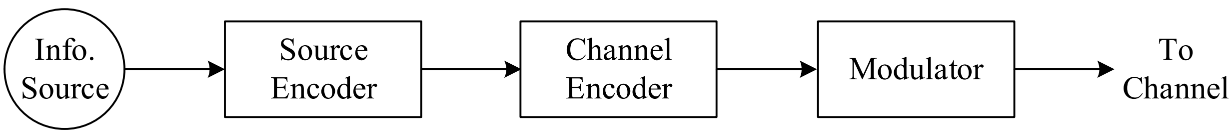


Figure I.2

Information Source:	Generates data to be transmitted, *usually in the form of symbols from a finite alphabet.*
Source Encoder:	Removes redundancy from the source symbol stream, *in order to reduce the required transmission bit rate.*
Channel Encoder:	Adds controlled redundancy to the source encoder symbol stream, *in order to improve system error rate performance.*
Modulator:	Maps the symbol stream into a finite set of signal waveforms. *Each different symbol is assigned a different waveform that is transmitted during the duration of the symbol.*

I.1.B. Receiver

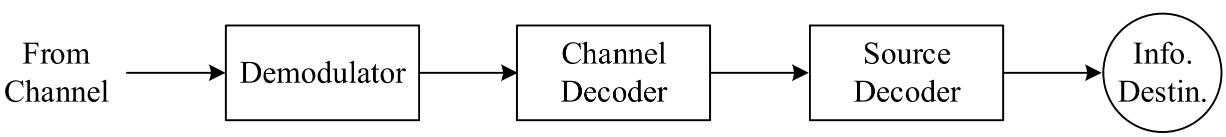


Figure I.3

Demodulator:	Re-maps the received waveforms into an encoded symbol stream. *Some symbol errors occur due to channel impairments. Receiver performance is usually measured by the error rate.*
Channel Decoder:	Removes the redundancy added by the channel encoder. *Some decoding errors occur due to demodulator errors.*
Source Decoder:	*Recovers part (lossy) or all (lossless) of the redundancy removed by the source encoder.*
Information Destination:	Restores original form of transmitted information. Restoration is usually imperfect due to demodulation/decoding errors.

I.1.C. CHANNEL

Channel Types

Linear	Nonlinear
Time-Invariant	Time-Varying
Non-Distorting	Distorting
Deterministic	Stochastic
Narrowband	Wideband

Channel Impairments

Attenuation	Noise
Distortion	Nonlinearities
Multipath	Fading
Interference	Jamming

Transmission Types

Guided Transmission	Unguided Transmission
Wires	Radio/Wireless
Cables	Lightwave
Waveguides	
Optical Fibers	

I.2. Wireless Communications Evolution

I.2.A. BRIEF HISTORY

Radio	**By Guglielmo Marconi in 1896**
Cellular Concept	**1947**
Advanced Mobile Phone Service (AMPS) (Analog, FDMA, Voice Only)	**Early 1980s**

(2G) Digital Telephony (TDMA, Voice and Text Data)	**PCD: 1991 in Japan** **GSM: 1992 in Europe** **IS-54: 1993 in NA**
CDMA Cellular Systems	**IS-95: 1995 in NA**
GPRS	**General Packet Radio Service** **2.5G – 1997: 140 kbps**
EDGE	**Enhanced Data for GSM Evolution (2.75G – 1999: 384 kbps)** **Enhanced Circuit-Switched Data – ECSD** **Enhanced General Packet Radio Service (EGPRS)**
3G	**Started in 2001 in Japan**
4G	**~2014**
5G	**~2020**

I.2.B. 0G (Pre Cellular)

This refers to systems that preceded modern cellular mobile communication systems. These mobile telephone systems can be distinguished from earlier closed radio telephone systems in that they were available as a commercial service that was part of the public switched telephone network (PSTN), with their own telephone numbers, rather than part of a closed network such as a police radio or taxi dispatch system.

I.2.C. 1G

1G systems were analog telecommunications systems, introduced in the late 1970s. Here the voice channel used frequency modulation. 1G systems used frequency division multiple access (FDMA) techniques. The major disadvantages of 1st generation wireless systems are poor voice quality, poor battery quality and large phone size.

I.2.D. 2G

The 2G systems were digital and were oriented to voice with low speed data services. 2G used GSM technology, where GSM stands for global system for mobile communication. It is a circuit switched, connection based technology, where the end systems are devoted to the entire call period. Therefore, it causes low efficiency in usage of bandwidth and resources. Generally GSM enabled systems don't support high data rates and they are generally unable to handle complex data like video. Next comes 2.5G. 2.5G is not an officially defined term, rather it was invented for marketing purposes. Same applies to 2.75G.

I.2.E. 3G

3G systems have the capability to handle complex data like video. They also support high data rates. Generally 3G wireless systems use Code Division Multiple Access Technique (CDMA). The 3G technology adds multimedia facilities to 2G phones by allowing video, audio, and graphics applications. Apart from that, 3G provided increased bandwidth, 384 kbps when the device holder is walking, 128 kbps in a car and 2 Mbps in a fixed application.

Compared to 2G, 3G offered the following new and improved features:

- More Reliability
- Longer Battery Life
- Cost Effectiveness
- Support for Multimedia
- Higher Capacity
- More Secure Internet
- Global Standards

I.2.F. 4G

4G systems have been lunched in many countries. In 2009 IMT-A specified the requirements for 4G standards. A 4G system is intended to provide a comprehensive and secure solution to laptop and mobile devices. Internet access, gaming services and streamed multimedia are provided to users. Technologies like coded orthogonal frequency division multiplexing (COFDM), multiple input multiple output (MIMO) and link adaptation are used in 4th generation wireless system.

4G offers the following new and improved features:

- 100 Mbps for high mobility communication (such as from trains and cars) and 1 Gbps for low mobility communication (such as pedestrians and stationary users).
- Applications include mobile web access, IP telephony, gaming services, high-definition mobile TV, video conferencing and 3D TV.
- Mobile WiMAX Release 2 (also known as WirelessMAN-Advanced or IEEE 802.16m) and LTE Advanced (LTE-A) are IMT-Advanced compliant backwards compatible versions of WiMAX and LTE.
- 4G systems do not support circuit-switched communication services. Instead, they provide all-internet protocol (IP) packet-switched services.

I.2.G. 5G

Research is going on in developing 5th generation wireless (5G) standards. It is expected that, it will fulfill the entire requirement that has not been fulfilled by 4G. 5G technology has changed the means to use cell phones within very high bandwidth. Users never encountered ever before such a high value technology. All kinds of advanced features will be included to make 5G technology most powerful and in huge demand in the near future. 5G systems use MIMO.

5G offers the following new and improved features:

- 5G has better coverage and higher data rates at the edges of the cells.
- Around 1 Gbps data rate is easily possible.
- Better security.
- Better energy efficiency and spectral efficiency.
- D2D communications.

I.3. Multiple Access Techniques

I.3.A. Time Division Multiple Access (TDMA)

This multiple access technique divides the time axis into different time slots, each of length T_{slot} according to Figure I.4. Each data packet or burst is assigned to a certain time slot, even though a user can occupy several slots. Users are assigned the time slots in a round robin manner, with a cycle ending in N_{slot} time slots. Guard intervals of length ΔT are inserted between successive slots in order to avoid interference between them. Within these intervals, no information is transmitted so that they represent redundancy and reduce the spectral efficiency of the communication system.

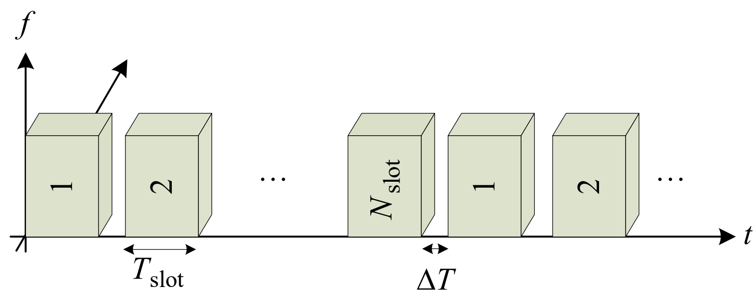


Figure I.4: TDMA

I.3.B. Frequency Division Multiple Access (FDMA)

In this multiple access technique, the frequency axis is divided into N_f subbands each of width B as illustrated in Figure I.5. The data streams are now distributed over different frequency bands. In mobile environments, the signal bandwidth is spread by the Doppler effect, so that neighboring bands can interfere. Thus, gaps of an appropriate width Δf are required to combat this effect at the expense of a reduced spectral efficiency.

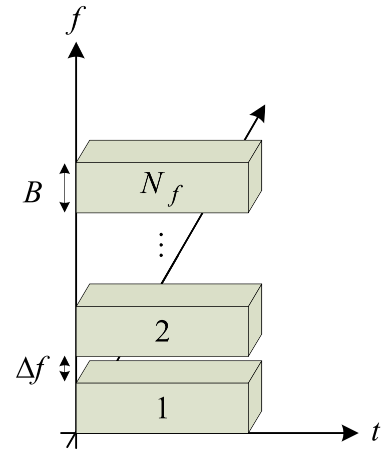


Figure I.5: FDMA

I.3.C. Code Division Multiple Access (CDMA)

In contrast to both the preceding schemes, CDMA allows simultaneous access to the channel in the same frequency range. The basic principle is to spectrally spread the data streams with specific sequences called spreading codes (Spread Spectrum technique). The signals can be distinguished by assigning them individual spreading codes. This opens a third dimension, as can be seen in Figure I.6. An intuitive choice would lead to orthogonal codes, ensuring a parallel transmission of different user signals. However, the transmission channel generally destroys the orthogonality and multiuser interference (MUI) becomes a limiting factor concerning spectral efficiency.

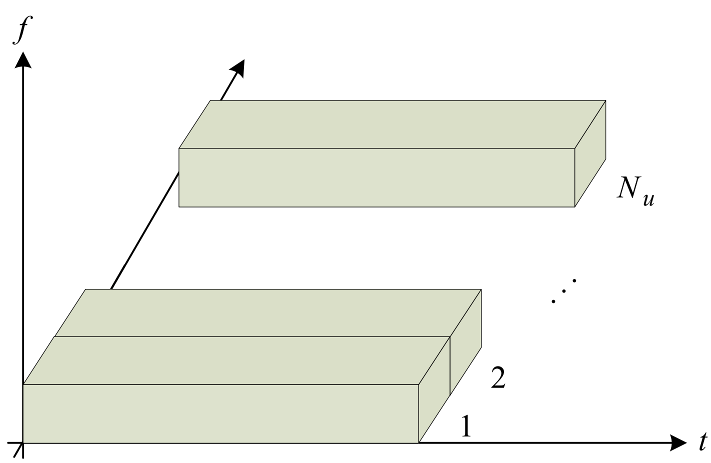


Figure I.6: CDMA

***

II. Communication over Bandlimited Channels

II.1. Preliminaries

Lowpass representation is used for channel and signals models. Channel is modeled as a linear time-invariant filter whose lowpass equivalent impulse response is denoted by $c(t)$. Lowpass equivalent channel frequency response is denoted by $C(f)$. $c(t)$ and $C(f)$ constitute a Fourier transform pair, i.e.,

$$c(t) \rightleftharpoons C(f) \tag{II.1}$$

Lowpass channel bandwidth is equal to W. This means that

$$C(f) = 0, |f| > W \tag{II.2}$$

Note that the bandpass channel bandwidth is equal to $2W$. Transmitted symbols are denoted by I_n, where n represents discrete time (multiples of symbol period T). $\{I_n\}$ are chosen from a signal space constellation, like a PAM or a QAM constellation. For certain constellations, e.g., 4-QAM, $\{I_n\}$ are complex valued.

A pulse shape function $g_T(t)$ is used at the transmitter. The pulse shape function has a duration T. The lowpass equivalent of the transmitted signal is given by:

$$s(t) = \sum_{n=-\infty}^{\infty} I_n g_T(t - nT) \tag{II.3}$$

The bandpass transmitted signal is given by:

$$\tilde{s}(t) = \mathrm{Re}\left\{ s(t) e^{j2\pi f_c t} \right\} \tag{II.4}$$

Let the received signal be given by $\tilde{r}(t)$. The lowpass equivalent of $\tilde{r}(t)$ is given by:

$$\begin{aligned} r(t) &= \mathrm{LP}\{\tilde{r}(t)\} \\ &= s(t) * c(t) + z(t) \\ &= \int_{-\infty}^{\infty} s(\tau) c(t-\tau) d\tau + z(t) \\ &= r_s(t) + z(t) \end{aligned} \tag{II.5}$$

where $z(t)$ is zero-mean AWGN, with $N_0/2$ PSD. The signal part of $r(t)$ is the inverse Fourier transform of

$$R_s(f) = S(f) C(f) \tag{II.6}$$

As a complex quantity, $C(f)$ can be written in the form

$$C(f) = |C(f)| e^{j\theta(f)} \tag{II.7}$$

Envelope Delay:

$$\tau(f) = -\frac{1}{2\pi} \frac{d}{df} \theta(f) \tag{II.8}$$

Non-Distorting (Ideal) Channel:

- Magnitude $|C(f)|$ is constant $\forall |f| \le W$
- Phase $\theta(f)$ is a linear function of f $\forall f \le W$, meaning that $\tau(f)$ is constant $\forall f \le W$.

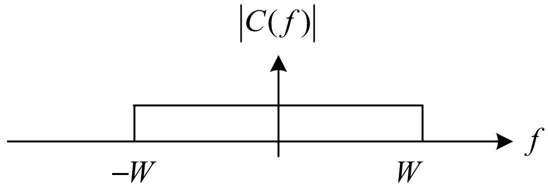


Figure II.1: Amplitude response of a non-distorting channel

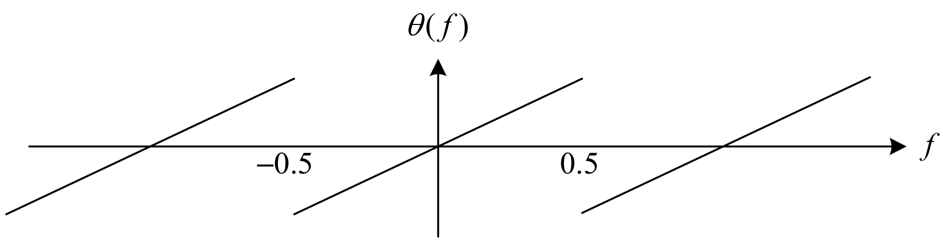


Figure II.2: Phase response of a non-distorting channel

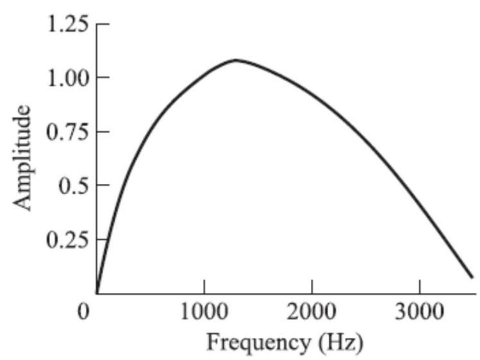


Figure II.3: Amplitude response of a distorting telephone channel

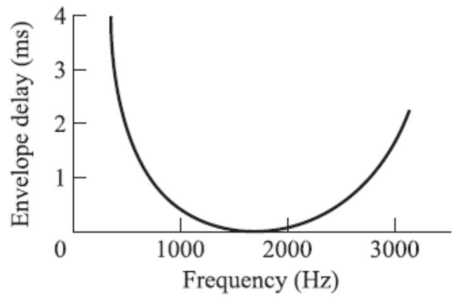


Figure II.4: Envelope delay as function of frequency over a distorting telephone channel

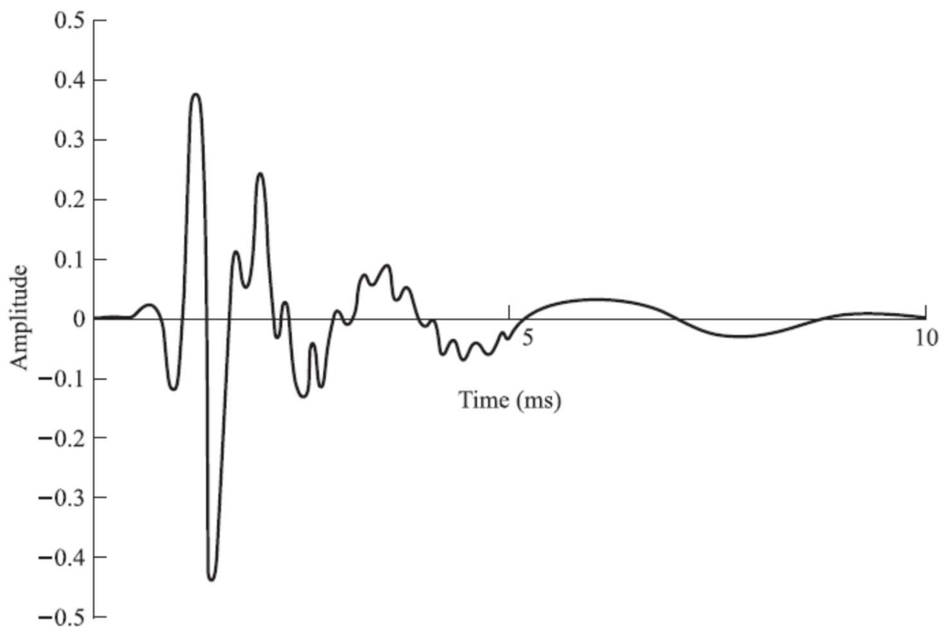


Figure II.5: Impulse response of a distorting telephone channel

The LP received signal has the form

$$r(t) = \left(\sum_{n=-\infty}^{\infty} I_n g_T(t - nT) \right) * c(t) + z(t) \tag{II.9}$$

This can be simplified as

$$r(t) = \sum_{n=-\infty}^{\infty} I_n h(t - nT) + z(t) \tag{II.10}$$

where

$$h(t) = g_T(t) * c(t) \tag{II.11}$$

At the receiver, the received signal is processed by the receive filter, as shown in Figure II.6.

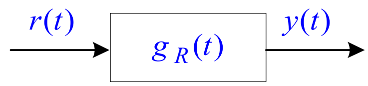


Figure II.6: Receive Filter

$$\begin{aligned} y(t) &= r(t) * g_R(t) \\ &= \left(\sum_{n=-\infty}^{\infty} I_n h(t-nT) + z(t) \right) * g_R(t) \\ &= \sum_{n=-\infty}^{\infty} I_n x(t-nT) + \xi(t) \end{aligned} \tag{II.12}$$

where

$$\begin{aligned} x(t) &= h(t) * g_R(t) \\ &= g_T(t) * c(t) * g_R(t) \end{aligned} \tag{II.13}$$

$$\xi(t) = z(t) * g_R(t) \tag{II.14}$$

Let $y(t)$ be sampled at $kT + t_0$, then

$$\begin{aligned} y(kT+t_0) &= \sum_{n=-\infty}^{\infty} I_n x(kT + t_0 - nT) + \xi(kT+t_0) \\ y_k &= \sum_{n=-\infty}^{\infty} I_n x_{k-n} + \xi_k \\ &= I_n * x_n + \xi_k \end{aligned} \tag{II.15}$$

Note that the desired symbol at $kT + t_0$ is I_k. Taking the term involving I_k out of the summation in (II.15) yields

$$\begin{aligned} y_k &= x_0 I_k + \sum_{\substack{n=-\infty \\ n \neq k}}^{\infty} I_n x_{k-n} + \xi_k \\ &= x_0 \left(I_k + \frac{1}{x_0} \sum_{\substack{n=-\infty \\ n \neq k}}^{\infty} I_n x_{k-n} \right) + \xi_k \end{aligned} \tag{II.16}$$

Letting $x_0 = 1$,

$$y_k = \underbrace{I_k}_{\text{Desired Symbol}} + \underbrace{\sum_{\substack{n=-\infty \\ n\neq k}}^{\infty} I_n x_{k-n}}_{\text{Intersymbol Interference (ISI)}} + \underbrace{\xi_k}_{\text{Additive Noise}} \tag{II.17}$$

II.1.A. The Nyquist Criterion (Zero ISI)

Let $C(f)=1, |f|\le W$. Then,

$$h(t)=g_T(t) \tag{II.18}$$

In this case, the received signal is an undistorted version of the transmitted signal plus AWGN. Therefore, the receive filter should be a matched filter, i.e.,

$$g_R(t)=g_T^*(T-t) \tag{II.19}$$

In the frequency domain, this is equivalent to

$$G_R(f)=G_T^*(f)e^{-j2\pi fT} \tag{II.20}$$

As for the overall channel seen by the transmitted signal (transmit filter, channel, receive filter), we have

$$X(f)=|G_R|^2 e^{-j2\pi fT} \tag{II.21}$$

To achieve the zero-ISI situation, we should have

$$x_k=\begin{cases}1, & k=0\\ 0, & k\neq 0\end{cases} \tag{II.22}$$

Note that $x_k=x(kT)=x(t)\big|_{t=kT}$. The inverse Fourier transform can be used to express $x(kT)$ in the form

$$x_k=\int_{-\infty}^{\infty} X(f)e^{j2\pi fkT}df \tag{II.23}$$

This can be rewritten in the form

$$x_k=\sum_{m=-\infty}^{\infty}\int_{(2m-1)/2T}^{(2m+1)/2T} X(f)e^{j2\pi fkT}df \tag{II.24}$$

Using a change of variables,

$$x_k = \sum_{m=-\infty}^{\infty} \int_{-1/2T}^{1/2T} X\left(f + \frac{m}{T}\right) e^{j2\pi fkT} df \tag{II.25}$$

This can be written in the form

$$x_k = \int_{-1/2T}^{1/2T} \sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right) e^{j2\pi fkT} df \tag{II.26}$$

Let's define

$$B(f) = \sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right) \tag{II.27}$$

Substituting (II.27) into (II.26),

$$x_k = \int_{-1/2T}^{1/2T} B(f)\, e^{j2\pi fkT} df \tag{II.28}$$

Obviously, $B(f)$ is a periodic function in the variable f with period $1/T$, and, therefore, it can be expanded in terms of its Fourier series coefficients $\{b_n\}$ as

$$B(f) = \sum_{k=-\infty}^{\infty} b_k e^{j2\pi kfT} \tag{II.29}$$

where

$$b_k = T \int_{-1/2T}^{1/2T} B(f) e^{-j2\pi kfT} df \tag{II.30}$$

Comparing (II.30) and (II.28),

$$\begin{aligned} b_k &= Tx_{-k} \\ &= Tx(-kT) \\ &= T\,x(t)\big|_{t=-kT} \end{aligned} \tag{II.31}$$

Therefore, for no ISI we should have

$$b_k = \begin{cases} T, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{II.32}$$

Substituting (II.32) into (II.29) yields

$$B(f) = T \tag{II.33}$$

Equivalently,

$$\sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right) = T \tag{II.34}$$

Consider the following general example of $B(f)$

$$B(f) = \sum_{m=-\infty}^{\infty} X\left(f + \frac{m}{T}\right)$$

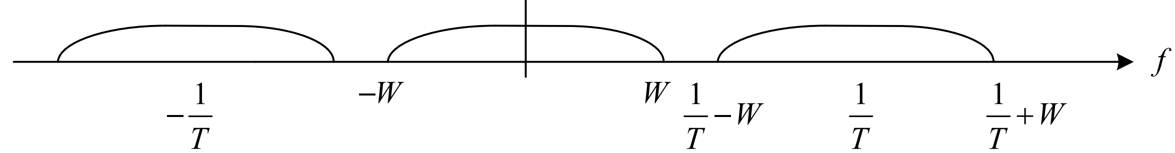


Figure II.7: General example of $B(f)$

- Case 1: $\frac{1}{T} > 2W \Rightarrow \frac{1}{T} - W > W \Rightarrow$ Because of the zero non-overlapping region in $B(f)$, there is no way $B(f) = T$ can be achieved. ISI cannot be avoided no matter how $x(t)$ is designed.
- Case 2: $\frac{1}{T} = 2W \Rightarrow$ Nyquist Rate. If

$$X(f) = \begin{cases} T, & |f| < W \\ 0, & \text{otherwise} \end{cases} \tag{II.35}$$

then $B(f) = T$ is satisfied. The resulting $x(t)$ is called the Nyquist pulse shape, and is given by:

$$x(t) = \frac{\sin(\pi t/T)}{\pi t/T} = \operatorname{sinc}\left(\frac{\pi t}{T}\right) \tag{II.36}$$

This means that the smallest value of T for which transmission with zero ISI is possible is $T = 1/2W$, and for this value, $x(t)$ has to be a sinc function. This shape is non-causal and therefore non-realizable. To make the sinc shape realizable, usually a delayed version of it, i.e., $\operatorname{sinc}\left(\pi(t-t_0)/T\right)$, is used and t_0 is chosen such that for $t<0$, we have $\operatorname{sinc}\left(\pi(t-t_0)/T\right) \approx 0$. Of course, with this choice of $x(t)$, the sampling time must also be shifted to $mT + t_0$.

A second difficulty with this pulse shape is that its rate of convergence to zero is slow. The tails of $x(t)$ decay as $1/t$; consequently, a small mistiming error in sampling the output of the matched filter at the demodulator results in an infinite series of lSI components. Such a series is not absolutely summable because of the $1/t$ rate of decay of the pulse, and, hence, the sum of the resulting lSI does not converge.

- Case 3: $\frac{1}{T} < 2W \Rightarrow \frac{1}{T} - W < W$ $B(f)$ consists of overlapping replications of $X(f)$ separated by $1/T$. In this case, there exist numerous choices for $X(f)$ that satisfy $B(f) = T$.

A particular pulse spectrum, for the $T > 1/2W$ case, that has desirable spectral properties and has been widely used in practice is the raised cosine spectrum. The raised cosine frequency characteristic is given as

$$X_{rc}(f) = \begin{cases} T, & 0 \le |f| \le \frac{1-\beta}{2T} \\ \frac{T}{2}\left\{1 + \cos\left[\frac{\pi \mathrm{T}}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right]\right\}, & \frac{1-\beta}{2T} \le |f| \le \frac{1+\beta}{2T} \\ 0, & |f| > \frac{1+\beta}{2T} \end{cases} \tag{II.37}$$

where β, called the roll-off factor, takes values in the range $0 \le \beta \le 1$. The bandwidth occupied by the signal beyond the Nyquist frequency $1/2T$ is called the excess bandwidth and is usually expressed as a percentage of the Nyquist frequency. For example. When $\beta = 0.5$ the excess bandwidth is 50%, and when $\beta = 1$ the excess bandwidth is 100%. The pulse having the raised cosine spectrum is given by

$$\begin{aligned} x_{rc}(t) &= \frac{\sin(\pi t/T)}{\pi t/T} \frac{\cos(\pi \beta t/T)}{1 - 4\beta^2 t^2/T^2} \\ &= \mathrm{sinc}(\pi t/T) \frac{\cos(\pi \beta t/T)}{1 - 4\beta^2 t^2/T^2} \end{aligned} \tag{II.38}$$

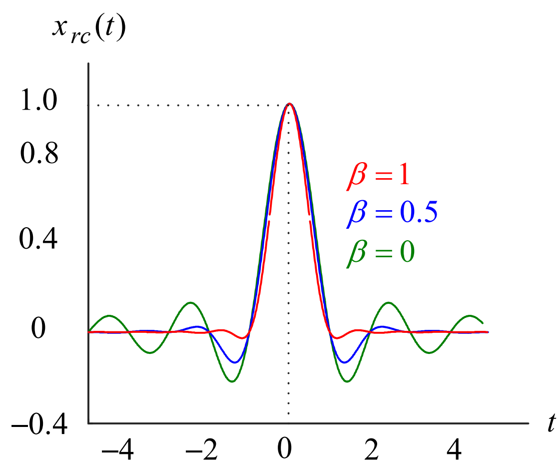


Figure II.8: Raised cosine pulse in the time domain

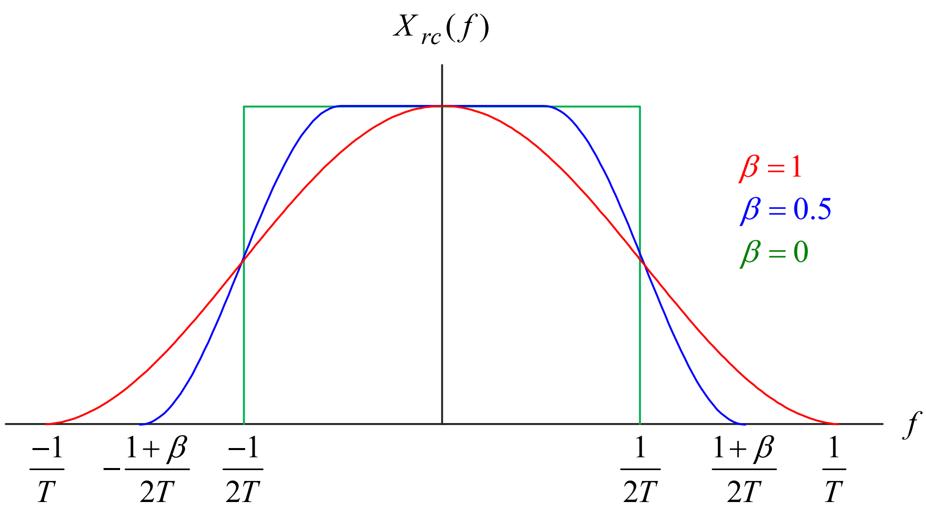


Figure II.9: Raised cosine pulse in the frequency domain

$$X_{rc}(f) = \begin{cases} T, & 0 \le |f| \le \frac{1-\beta}{2T} \\ \frac{T}{2}\left\{1+\cos\left[\frac{\pi T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right]\right\}, & \frac{1-\beta}{2T} \le |f| \le \frac{1+\beta}{2T} \\ 0, & |f| > \frac{1+\beta}{2T} \end{cases} \tag{II.39}$$

Note that for $\beta = 0$, the pulse reduces to $x(t) = \text{sinc}(\pi t/T)$, and the symbol rate $1/T = 2W$. When $\beta = 1$, the symbol rate is $1/T = W$.

In general, the tails of $x(t)$ decay as $1/t^3$ for $\beta > 0$. Consequently, a mistiming error in sampling leads to a series of lSI components that converges to a finite value.

II.1-Preliminaries

Due to the smooth characteristics of the raised cosine spectrum, it is possible to design practical filters for the transmitter and the receiver that approximate the overall desired frequency response.

In the special case where the channel is ideal, i.e., $C(f)=1$, $|f|\le W$, we have

$$X_{rc}(f)=G_T(f)G_R(f) \tag{II.40}$$

In this case, if the receiver filter is matched to the transmitter filter ($G_R(f)=G_T^*(f)$), we have

$$X_{rc}(f)=\left|G_T(f)\right|^2 \tag{II.41}$$

We let

$$G_T(f)=\sqrt{\left|X_{rc}(f)\right|}e^{-j2\pi f t_0} \tag{II.42}$$

where t_0 is some nominal delay that is required to ensure physical realizability of the filter.

II.1.B. Design of Band-Limited Signals with Controlled ISI – Partial Response Signals

It is necessary to reduce the symbol rate $1/T$ below the Nyquist rate of $2W$ symbols/s to realize practical transmitting and receiving filters. Suppose we choose to relax the condition of zero ISI and, thus, achieve a symbol transmission rate of $2W$ symbols/s. By allowing for a controlled amount of ISI, we can achieve this symbol rate. Suppose that we design the band-limited signal to have controlled ISI at one time instant. This means that we allow one additional nonzero value in the samples $\{x(nT)\}$. The ISI that we introduce is deterministic or "controlled" and, hence, it can be taken into account at the receiver, as discussed below.

Duobinary Pulse

One special case that leads to (approximately), physically realizable transmitting and receiving filters is specified by the samples:

$$x(nT)=\begin{cases}1, & n=0,1\\ 0, & \text{otherwise}\end{cases} \tag{II.43}$$

Then

$$b_n=\begin{cases}T, & n=0,-1\\ 0, & \text{otherwise}\end{cases} \tag{II.44}$$

Therefore,

$$B(f)=T+Te^{-j2\pi fT} \tag{II.45}$$

For $T=1/2W \Rightarrow 1/T=2W$,

$$X(f) = \begin{cases} \frac{1}{2W}\left(1+e^{-j\pi f/W}\right), & |f|<W \\ 0, & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{1}{W} e^{-j\pi f/2W} \cos\left(\frac{\pi f}{2W}\right), & |f|<W \\ 0, & \text{otherwise} \end{cases} \quad \text{(II.46)}$$

Therefore,

$$x(t) = \text{sinc}(2\pi W t) + \text{sinc}\left[2\pi\left(Wt - \frac{1}{2}\right)\right] \quad \text{(II.47)}$$

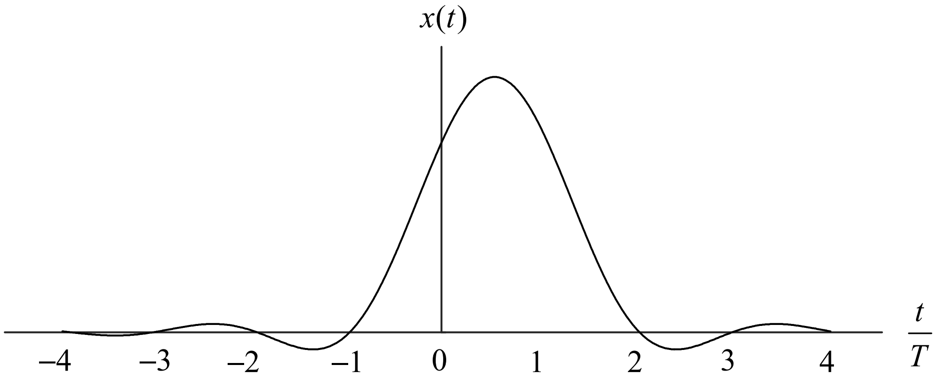


Figure II.10: Duobinary pulse

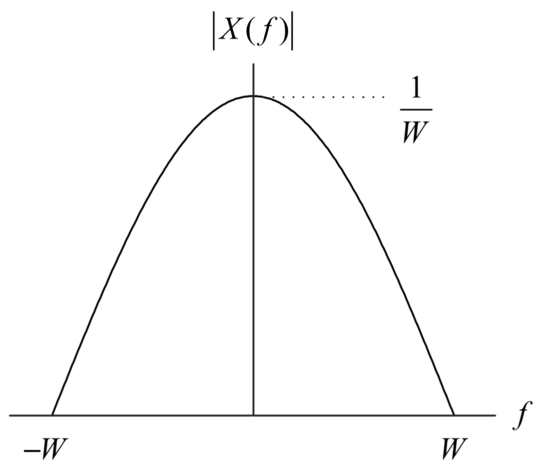


Figure II.11: Duobinary pulse in the frequency domain

The spectrum decays to zero smoothly, which means that physically realizable filters can be designed that approximate this spectrum very closely.

Modified Duobinary Pulse

Another special case that leads to (approximately) physically realizable transmitting and receiving filters is specified by the samples:

$$x(nT) = x\left(\frac{n}{2W}\right) = \begin{cases} 1, & n = -1 \\ -1, & n = 1 \\ 0, & \text{otherwise} \end{cases} \tag{II.48}$$

$$x(t) = \operatorname{sinc}\left(\frac{\pi(t+T)}{T}\right) - \operatorname{sinc}\left(\frac{\pi(t-T)}{T}\right) \tag{II.49}$$

$$\begin{aligned} X(f) &= \begin{cases} \frac{1}{2W}\left(e^{j\pi f/W} - e^{-j\pi f/W}\right), & |f| \le W \\ 0, & |f| > W \end{cases} \\ &= \begin{cases} \frac{j}{W}\sin\left(\frac{\pi f}{W}\right), & |f| \le W \\ 0, & |f| > W \end{cases} \end{aligned} \tag{II.50}$$

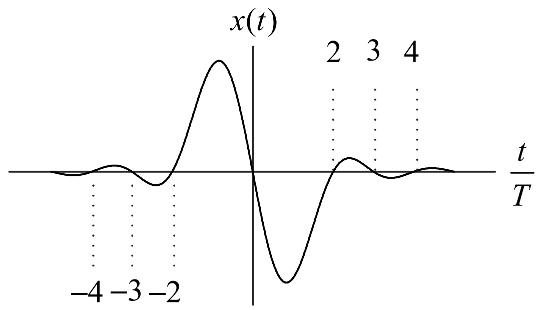


Figure II.12: Modified duobinary pulse

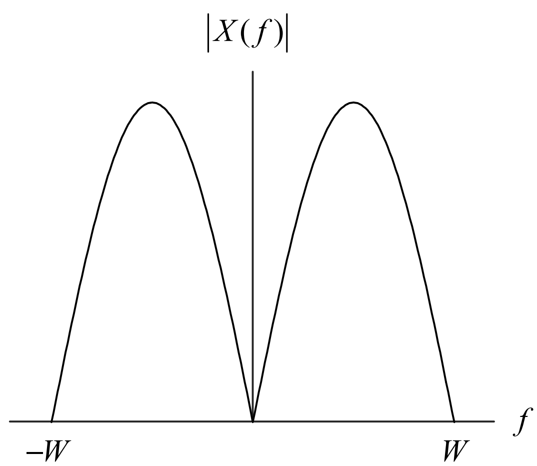


Figure II.13: Modified duobinary pulse in the frequency domain

This signal has a zero at $f=0$, making it suitable for transmission over a channel that does not pass DC.

Partial Response Signals

Consider the class of bandlimited signals pulses that have the form

$$x(t)=\sum_{n=-\infty}^{\infty} x(nT)\operatorname{sinc}\left[2\pi W(t-nT)\right] \quad \text{(II.51)}$$

$$x(t)=\sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right)\operatorname{sinc}\left[2\pi W\left(t-\frac{n}{2W}\right)\right] \quad \text{(II.52)}$$

The operation in (II.51) is equivalent to the recovery of a baseband waveform $x(t)$ with bandwidth W from samples $x_n=x(nT)=x(n/2W)$ that have been generated by sampling $x(t)$ at the Nyquist rate of $2W$ samples per second. The recovery process is done by passing the samples through a lowpass filter with bandwidth W. Equation (II.51) is known as the interpolation formula.

The waveform in (II.51) has the Fourier transform

$$X(f)=\begin{cases}\dfrac{1}{2W}\displaystyle\sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right)e^{-j\pi nf/W}, & |f|\le W\\ 0, & \text{otherwise}\end{cases} \quad \text{(II.53)}$$

Depending on the values of $\{x_n\}$, the Nyquist pulse, the duobinary and the modified duobinary can be seen as special cases of (II.51). Note that the raised cosine pulse is not generally associated with sampling at the Nyquist rate.

One can obtain interesting and physically realizable filter characteristics by selecting different values for the samples $\{x(n/2W)\}$ such that there are more than two nonzero samples. This is called partial response signaling. Note that as we select more nonzero samples, the problem of unraveling the controlled ISI becomes more cumbersome and impractical.

For example, x_n may be designed such that it can be non-zero for a finite $N>1$,

$$x_n=x\left(\frac{n}{2W}\right), \quad n=0,1,\cdots,N-1 \quad \text{(II.54)}$$

Let (II.17) be rewritten in the form

$$y_k=\beta_k+\xi_k \quad \text{(II.55)}$$

where

$$\beta_n = I_n + \sum_{\substack{l=-\infty \\ l\neq n}}^{\infty} I_l x_{n-l} \tag{II.56}$$

Note that y_k contains purposely introduced (controlled) ISI. Based on (II.54),

$$\beta_n = \sum_{l=0}^{N} x_l I_{n-l} \tag{II.57}$$

This is equivalent to passing the discrete-time sequence $\{I_n\}$ through a finite impulse response (FIR) discrete-time filter with an impulse response sequence $\{x_n\}_{n=0}^{N-1}$.

The resulting signal pulses allow us to transmit information symbols at the Nyquist rate of $2W$ symbols/s.

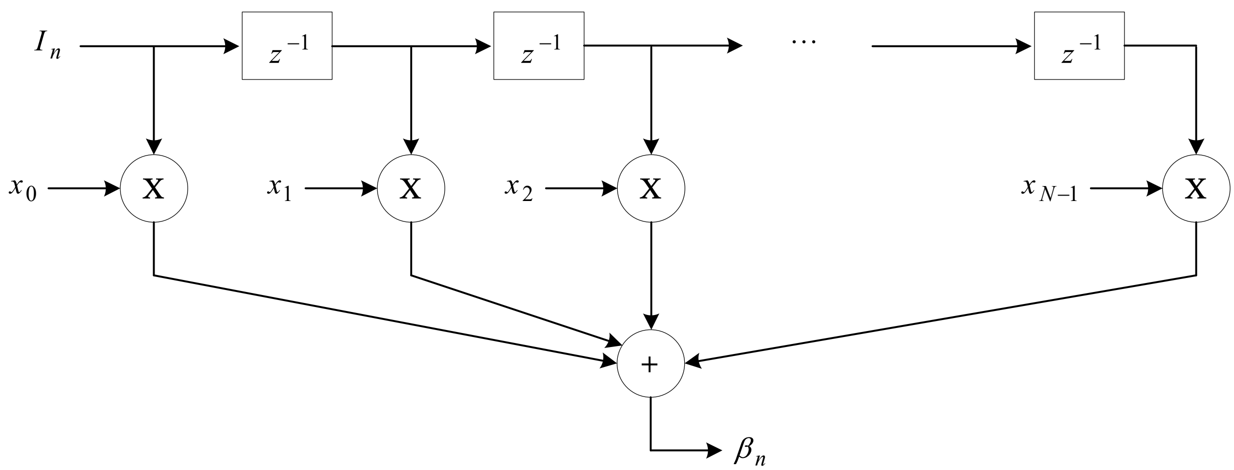


Figure II.14: Partial response – controlled ISI

The sequence of symbols $\{\beta_n\}$ is correlated as a consequence of the filtering performed on the sequence $\{I_n\}$. Let $\{I_n\}$ have equally probable values. Let $\{I_n\}$ be uncorrelated.

$$\begin{aligned}\phi_I(m) &= \mathrm{E}\left[I_n I_{n+m}\right] \\ &= \begin{cases}\mathrm{E}\left[I_n^2\right], & m=0 \\ \mathrm{E}\left[I_n I_{n+m}\right] = \mathrm{E}\left[I_n\right]\mathrm{E}\left[I_{n+m}\right], & m\neq 0\end{cases}\end{aligned} \tag{II.58}$$

Let $I_n = \pm 1$ (BPAM).

$$\begin{aligned}\mathrm{E}\left[I_n\right] &= (-1)(0.5) + (1)(0.5) \\ &= 0\end{aligned} \tag{II.59}$$

$$\phi_I(m) = \begin{cases} \mathrm{E}\left[I_n^2\right], & m = 0 \\ 0, & m \neq 0 \end{cases} \tag{II.60}$$

Because $I_n = \pm 1$,

$$\begin{aligned} \phi_I(m) &= \begin{cases} 1 & m = 0 \\ 0, & m \neq 0 \end{cases} \\ &= \delta_m \end{aligned} \tag{II.61}$$

The autocorrelation function of $\{\beta_n\}$ is

$$\begin{aligned} \phi_\beta(m) &= \mathrm{E}\left[\beta_n \beta_{n+m}\right] \\ &= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} x_k x_l \, \mathrm{E}\left[I_{n-k} I_{n+m-l}\right] \end{aligned} \tag{II.62}$$

When the input sequence is zero-mean and uncorrelated, and with the normalization $\mathrm{E}\left[I_n^2\right] = 1$,

$$\mathrm{E}\left[I_{n-k} I_{n+m-l}\right] = \delta_{m+k-l} \tag{II.63}$$

Substituting (II.63) into (II.62),

$$\phi_\beta(m) = \sum_{k=0}^{N-1-|m|} x_k x_{k+|m|}, \quad m = 0, \pm 1, \cdots, \pm (N-1) \tag{II.64}$$

The power spectral Density of β_n is then,

$$\begin{aligned} \Phi_\beta(f) &= \sum_{m=-(N-1)}^{N-1} \phi_\beta(m) e^{-j2\pi fmT} \\ &= \left| \sum_{m=0}^{N-1} x_m e^{-j2\pi fmT} \right|^2 \end{aligned} \tag{II.65}$$

where $T = 1/2W$ and $|f| \leq W$.

II.1.C. Data Detection for Controlled ISI

We consider two methods of symbol detection:

- Symbol-by-symbol detection is relatively easy to implement.
- Maximum-likelihood sequence minimizes the probability of error but is a little more complex to implement.

In particular, we consider the detection of the duobinary and the modified duobinary partial response signals.

We assume that the desired spectral characteristic $X(f)$ for the partial response signal is split evenly between the transmitting and receiving filters:

$$\begin{aligned}|G_T(f)| &= |G_R(f)| \\ &= |X(f)|^{\frac{1}{2}}\end{aligned} \tag{II.66}$$

Symbol-by-Symbol Suboptimum Detection

For the duobinary signal pulse

$$\begin{aligned}y_m &= \beta_m + \xi_m \\ &= I_m + I_{m-1} + \xi_m\end{aligned} \tag{II.67}$$

Let us ignore the noise for the moment and consider the binary case where $I_m = \pm 1$ with equal probability. Then β_m takes on one of three possible values, namely, $\beta_m = -2, 0, 2$ with corresponding probabilities 1/4, 1/2, 1/4. If I_{m-1} is the detected symbol from the $(m-1)^{\text{th}}$ signaling interval, its effect on β_m, the received signal in the m^{th} signaling interval, can be eliminated by subtraction, thus allowing I_m to be detected. This process can be repeated sequentially for every received symbol.

The major problem with this procedure is that errors arising from the additive noise tend to propagate. Error propagation can be avoided by precoding the data at the transmitter instead of eliminating the controlled ISI by subtraction at the receiver. The precoding is performed on the binary data sequence prior to modulation. From the data sequence $\{D_m\}$ of 1s and 0s that is to be transmitted, a new sequence $\{P_m\}$, called the precoded sequence, is generated. For the duobinary signal, the precoded sequence is defined as

$$P_m = D_m \ominus P_{m-1} \tag{II.68}$$

Note that this means that

$$D_m = P_m \oplus P_{m-1} \tag{II.69}$$

Let the source bits be encoded into

$$I_m = 2P_m - 1 \tag{II.70}$$

The noise-free samples at the output of the receiving filter are given by

$$\begin{aligned}\beta_m &= I_m + I_{m-1} \\ &= (2P_m - 1) + (2P_{m-1} - 1) \\ &= 2(P_m + P_{m-1} - 1)\end{aligned} \tag{II.71}$$

$$P_m + P_{m-1} = \frac{1}{2}\beta_m + 1 \tag{II.72}$$

Note that from (II.68),

$$\begin{aligned}D_m &= P_m \oplus P_{m-1} \\ &= \left(\frac{1}{2}\beta_m + 1\right) \quad \text{mod } 2\end{aligned} \tag{II.73}$$

This means that

$$D_m = \begin{cases} 0, & \beta_m = \pm 2 \\ 1, & \beta_m = 0 \end{cases} \tag{II.74}$$

Data Sequence D_n		1	1	1	0	1	0	0	1	0	0	0	1	1	0	1
Precoded Sequence P_n	0	1	0	1	1	0	0	0	1	1	1	1	0	1	1	0
Transmitted Sequence I_n	-1	+1	-1	+1	+1	-1	-1	-1	+1	+1	+1	+1	-1	+1	+1	-1
Received Sequence β_n		0	0	0	2	0	-2	-2	0	2	2	2	0	0	2	0
Decoded Sequence D_n		1	1	1	0	1	0	0	1	0	0	0	1	1	0	1

In the presence of additive noise, $y_m = \beta_m + \xi_m$ is compared with the two thresholds set at +1 and -1. The data sequence is obtained according to the detection rule

$$D_m = \begin{cases} 1, & |y_m| < 1 \\ 0, & |y_m| \geq 1 \end{cases} \tag{II.75}$$

Symbol-by-Symbol Suboptimum Detection for M-ary PAM

The extension from binary PAM to multilevel PAM signaling using the duobinary pulses is straightforward. In this case the M-level amplitude sequence $\{I_m\}$ results in a (noise-free) sequence

$$\beta_m = I_m + I_{m-1} \tag{II.76}$$

which has $2M-1$ possible equally spaced levels. The amplitude levels are determined from the relation

$$I_m = 2P_m - (M-1) \tag{II.77}$$

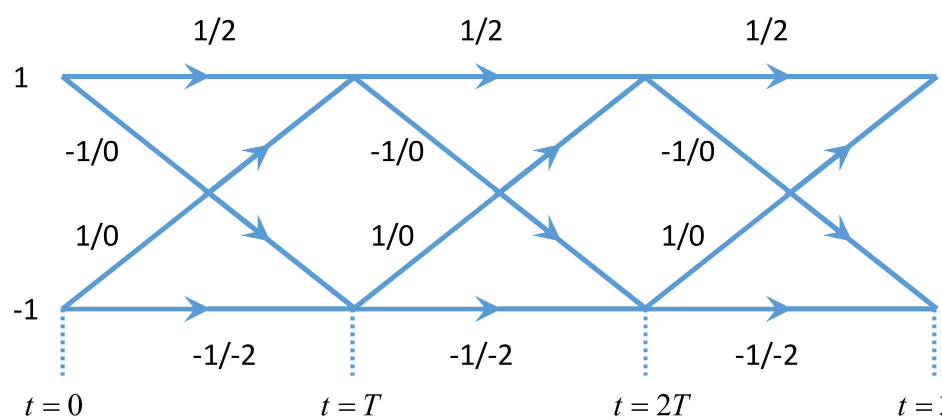
where $\{P_m\}$ is the precoded sequence that is obtained from an M-level data sequence $\{D_m\}$ according to the relation

$$P_m = D_m \ominus P_{m-1} \quad (\text{mod } M) \tag{II.78}$$

Symbol-by-Symbol Suboptimum Detection for Binary PAM – Error Probability

Maximum-Likelihood Sequence Detection

It is clear from the above discussion that partial-response waveforms are signal waveforms with memory. This memory is conveniently represented by a trellis. Following is the trellis for binary data transmission with the duobinary partial response signaling.



This trellis contains two states, corresponding to the two possible input values $I_m = \pm 1$. Each branch in the trellis is labeled by two numbers. The first number on the left is the new data bit, i.e., $I_{m+1} = \pm 1$.This number determines the transition to the new state. The number on the right is the received signal level.

The duobinary signal has a memory of length $L = 1$. Hence, for binary modulation the trellis has $S_t = 2$ states. In general, for M-ary modulation, the number of trellis states is M^L.

The optimum maximum-likelihood (ML) sequence detector selects the most probable path through the trellis upon observing the received data sequence $\{y_m\}$ at the sampling instants $t = mT, m = 1, 2, \cdots$. In general, each node in the trellis will have M incoming paths and M corresponding metrics. One out of the M incoming paths is selected as the most probable, based on the values of the metrics and the other $M-1$ paths and their metrics are discarded. The surviving path at each node is then extended to M new paths, one for each of the M possible input symbols, and the search process continues. This is basically the Viterbi algorithm for performing the trellis search.

II.1.D. Signal Design for Channels with Distortion

II.2. Optimum Receiver for Channels with ISI and AWGN

II.2.A. Optimum Maximum-Likelihood Receiver

II.2.B. A Discrete-Time Model for a Channel with ISI

In dealing with band-limited channels that result in ISI, it is convenient to develop an equivalent discrete-time model for the analog (continuous-time) system. The transmitter sends discrete-time symbols at a rate of $1/T$ symbols/s. The sampled output of the matched filter at the receiver is also a discrete-time signal with samples occurring at a rate of $1/T$ per second.

The cascade of the analog filter at the transmitter with impulse response $g_T(t)$, the channel with impulse response $c(t)$, the matched filter at the receiver with impulse response $g_R(t) = g_T^*(-t)$, and the sampler can be represented by an equivalent discrete-time transversal filter having tap gain coefficients $\{x_k\}$.

Consequently, we have an equivalent discrete-time transversal filter that spans a time interval of $2LT$ seconds. The input of the discrete-time filter is the sequence of information symbols $\{I_k\}$ and its output is the discrete-time sequence $\{y_k\}$.

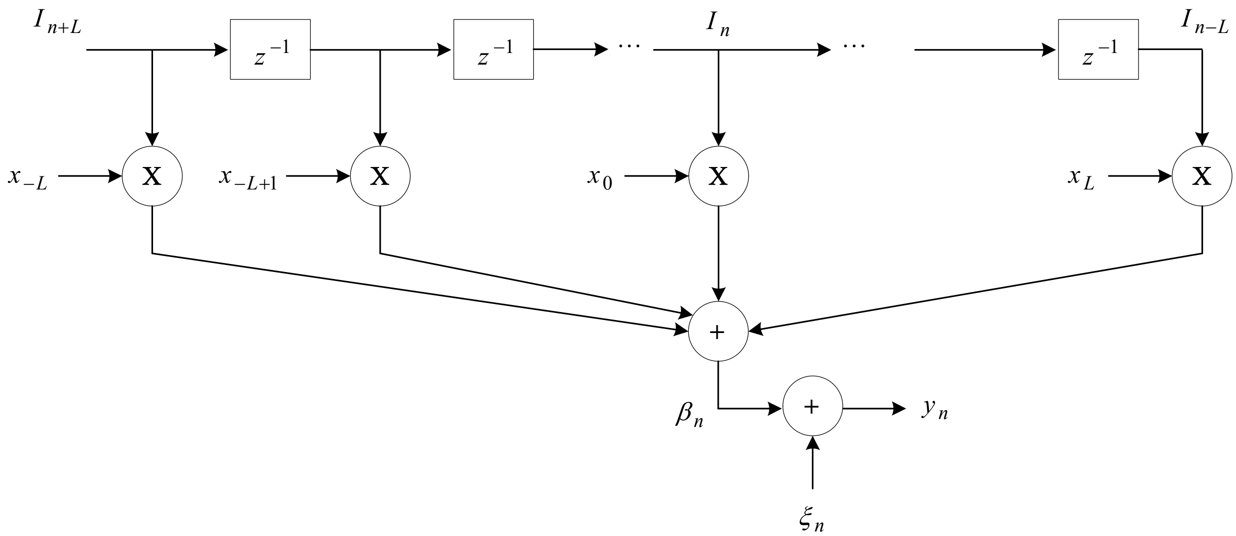


Figure II.15: Discrete-time channel model

The major difficulty with this discrete-time model occurs in the evaluation of performance of the various equalization or estimation techniques. The difficulty is caused by the correlations in the noise sequence $\{\xi_k\}$ at the output of the matched filter. The set of noise variables $\{\xi_k\}$ is a Gaussian-distributed sequence with zero-mean and autocorrelation function

$$\mathrm{E}\left[\xi_k^*\xi_j\right]=\begin{cases}2N_0x_{j-k}, & |k-j|\le L\\ 0, & \text{otherwise}\end{cases} \tag{II.79}$$

Hence, the noise sequence is correlated unless $x_k=0$ for $k\neq 0$. Since it is more convenient to deal with the white noise sequence when calculating the error rate performance, it is desirable to whiten the noise sequence by further filtering the sequence $\{y_k\}$.

Let $X(z)$ denote the (two-sided) z transform of $\{x_k\}$, i.e.,

$$X(z)=\sum_{l=-L}^{L}x_l z^{-l} \tag{II.80}$$

Note that $x_k=x_{-k}^*\Rightarrow X(z)=X^*\left(1/z^*\right)$. The $2L$ roots of $X(z)$ have the symmetry that if ρ is a root, $1/\rho^*$ is also a root. Hence, $X(z)$ can be factored and expressed as

$$X(z)=F(z)F^*\left(\frac{1}{z^*}\right) \tag{II.81}$$

Where $F(z)$ is a polynomial of degree L having the roots $\rho_1, \rho_2, \cdots, \rho_L$ and $F^*\left(1/z^*\right)$ is a polynomial of degree L having the roots $1/\rho_1^*, 1/\rho_2^*, \cdots, 1/\rho_L^*$. Assuming that there are no roots on the unit circle, an appropriate noise-whitening filter has a z transform $1/F^*\left(1/z^*\right)$.

Consequently, passage of the sequence $\{y_k\}$ through the digital filter $1/F^*\left(1/z^*\right)$ results in an output sequence $\{v_k\}$ that can be expressed as

$$v_n = \sum_{l=0}^{L} f_l I_{n-l} + \eta_n \qquad \text{(II.82)}$$

Note that $\{\eta_k\}$ is AWGN. The cascade of the matched filter, the sampler, and the noise-whitening filter is called the whitened matched filter (WMF). It is convenient to normalize the energy of $F(z)$ to unity

$$\sum_{l=0}^{L} \left|f_l\right|^2 = 1 \qquad \text{(II.83)}$$

In summary, the cascade of the transmitting filter impulse response $g_T(t)$, the channel impulse response $c(t)$, the receive filter impulse response $g_R(t)$, the sampler, and the discrete-time noise-whitening filter $1/F^*\left(1/z^*\right)$ can be represented as an equivalent discrete-time transversal filter having the set $\{f_l\}$ as its tap coefficients. The additive noise sequence $\{\eta_k\}$ corrupting the output of the discrete-time transversal filter is a white Gaussian noise sequence having zero-mean and variance N_0. We refer to this model as the equivalent discrete-time white noise filter model.

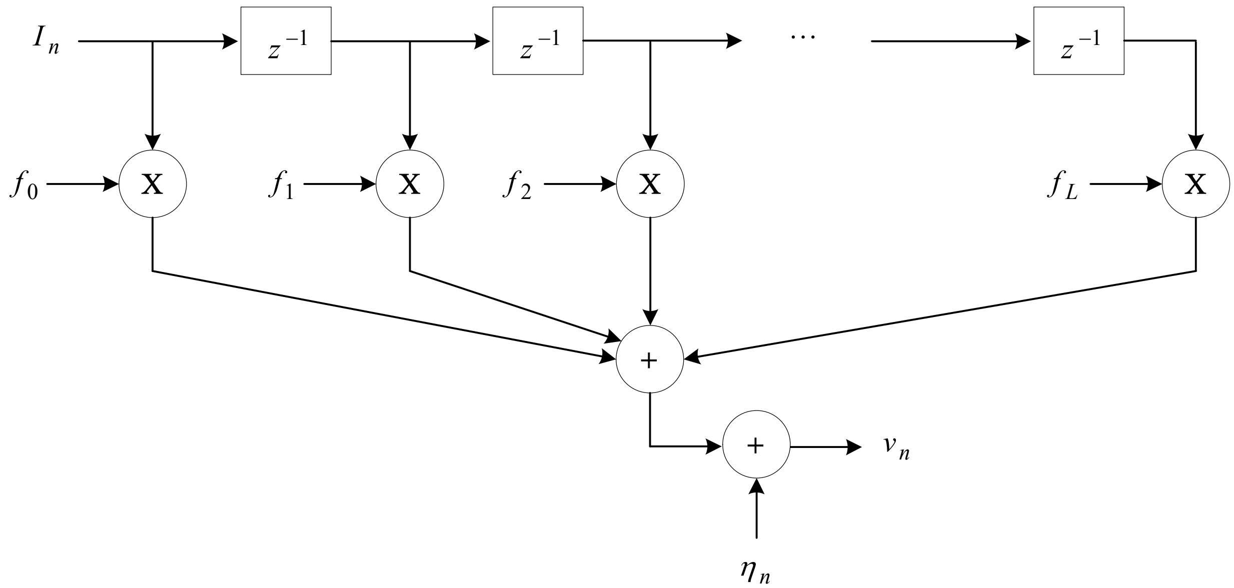


Figure II.16: Equivalent discrete-time white noise channel model

When the channel impulse response is changing slowly with time, the matched filter at the receiver becomes a time-variable filter. In this case, the time variations of the channel/matched-filter pair result in a discrete-time filter with time-variable coefficients. As a consequence, we have time-variable intersymbol interference effects, which can be modeled by the filter above, where the tap coefficients are slowly varying with time.

The discrete-time white noise linear filter model for the intersymbol interference effects that arise in high-speed digital transmission over nonideal band-limited channels will be used throughout our discussion of compensation techniques for the interference. In general, the compensation methods are called equalization techniques or equalization algorithms.

II.2.C. Maximum-Likelihood Sequence Estimation (MLSE) for the Discrete-Time White Noise Filter Model

II.2.D. Performance of MLSE for Channels with ISI

II.3. Linear Equalization

The MLSE for a channel with ISI has a computational complexity that grows exponentially with the length of the channel time dispersion. If the size of the symbol alphabet is M and the number of interfering symbols contributing to ISI is L, the Viterbi algorithm computes M^{L+1} metrics for each new received symbol.

Next, we describe suboptimum channel equalization approaches to compensate for the ISI. One approach employs a linear transversal filter. This filter structure has a computational complexity that is a linear function of the channel dispersion length L.

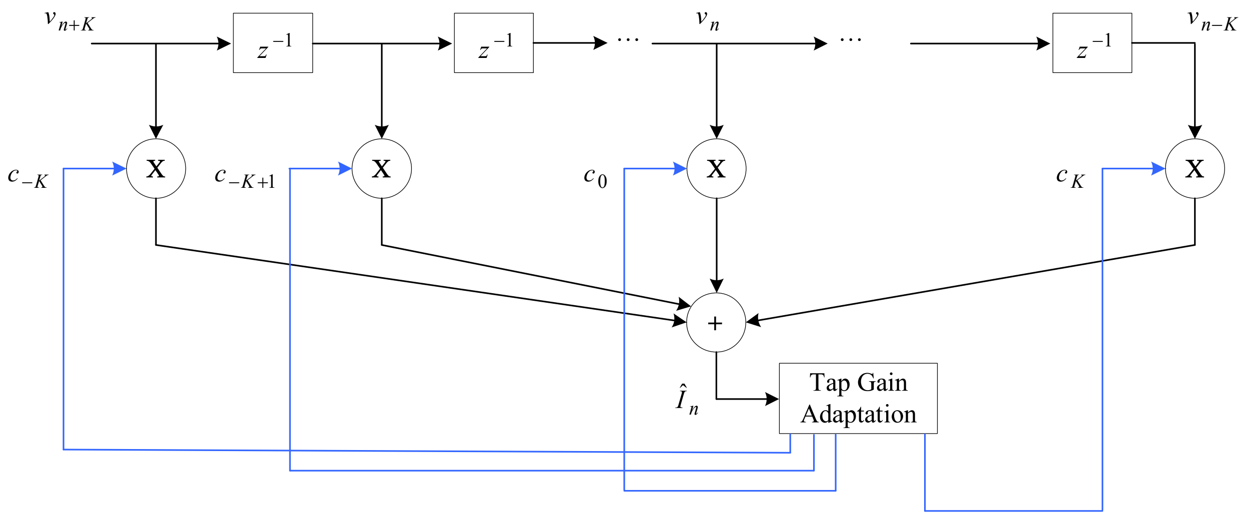


Figure II.17: Linear equalizer

The filter input is the sequence $\{v_k\}$ and its output is an estimate of the information sequence $\{I_k\}$. The estimate of the k^{th} symbol may be expressed as

$$\hat{I}_k = \sum_{j=-K}^{K} c_j v_{k-j} \tag{II.84}$$

where $\{c_j\}$ are the $2K+1$ complex-valued tap weight coefficients of the filter. The estimate $\hat{I}_k$ is quantized to the nearest (in distance) information symbol to form the decision $\tilde{I}_k$. If $\tilde{I}_k$ is not identical to the transmitted information symbol I_k, an error has been made.

The most meaningful measure of performance for a digital communication system is the average probability of error. Therefore, it is desirable to choose the filter coefficients to minimize this performance measure. However, the probability of error is a highly non-linear function of $\{c_j\}$.

Consequently, the probability of error as a performance index for optimizing the tap weight coefficients of the equalizer is computationally complex. Two criteria have found widespread use in optimizing the equalizer coefficients $\{c_j\}$. One is the peak distortion criterion and the other is the mean-square-error criterion.

II.3.A. Peak Distortion Criterion

The peak distortion is simply defined as the worst-case intersymbol interference at the output of the equalizer. The minimization of this performance index is called the peak distortion criterion.

Infinite-Length Equalizer

We observe that the cascade of the discrete-time linear filter model having an impulse response $\{f_n\}$ and an equalizer having an impulse response $\{c_n\}$ can be represented by a single equivalent filter having the impulse response

$$\begin{aligned} q_n &= \sum_{j=-\infty}^{\infty} c_j f_{n-j} \\ &= c_n * f_n \end{aligned} \tag{II.85}$$

The equalizer is assumed to have an infinite number of taps. Its output at the k^{th} sampling instant can be expressed in the form

$$\hat{I}_k = \underbrace{q_0 I_k}_{\text{Desired Symbol}} + \underbrace{\sum_{n \neq k} I_n q_{k-n}}_{\text{Intersymbol Interference}} + \underbrace{\sum_{j=-\infty}^{\infty} c_j \eta_{k-j}}_{\text{Noise}} \tag{II.86}$$

Let $q_0 = 1$. The peak value of intersymbol interference, which is called the peak distortion, is

$$\begin{aligned} \mathcal{D}(c) &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} |q_n| \\ &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left| \sum_{j=-\infty}^{\infty} c_j f_{n-j} \right| \end{aligned} \tag{II.87}$$

Note that $\mathcal{D}(c)$ is a function of the equalizer tap weights.

With an equalizer having an infinite number of taps, it is possible to select the tap weights so that $\mathcal{D}(c) = 0$, i.e., $q_n = 0$ for all $n \neq 0$. That is, the intersymbol interference can be completely eliminated. The values of the tap weights for accomplishing this goal are determined from the condition

$$\begin{aligned} q_n &= \sum_{j=-\infty}^{\infty} c_j f_{n-j} = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases} \\ &= \delta_n \end{aligned} \tag{II.88}$$

By taking the z transform of the last equation we obtain

$$\begin{aligned} Q(z) &= C(z)F(z) \\ &= 1 \end{aligned} \tag{II.89}$$

which means

$$C(z) = \frac{1}{F(z)} \tag{II.90}$$

Therefore, complete elimination of the intersymbol interference requires the use of an inverse filter to $F(z)$. We call such a filter a zero-forcing filter.

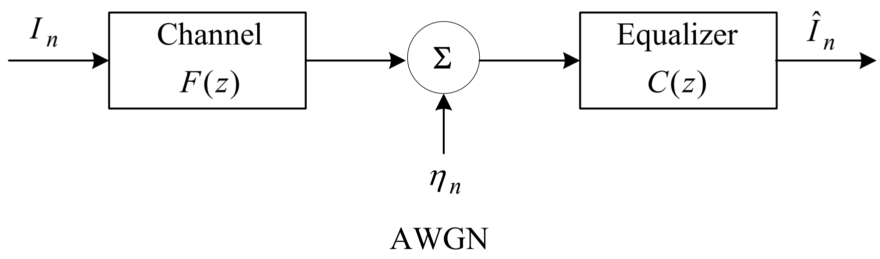


Figure II.18: Zero Forcing Equalizer

The cascade of the noise-whitening filter having the transfer function $1/F^*(1/z^*)$ and the zero-forcing equalizer having the transfer function $1/F(z)$ results in an equivalent zero-forcing equalizer having the transfer function

$$\begin{aligned} C'(z) &= \frac{1}{F(z)F^*\left(1/z^*\right)} \\ &= \frac{1}{X(z)} \end{aligned} \tag{II.91}$$

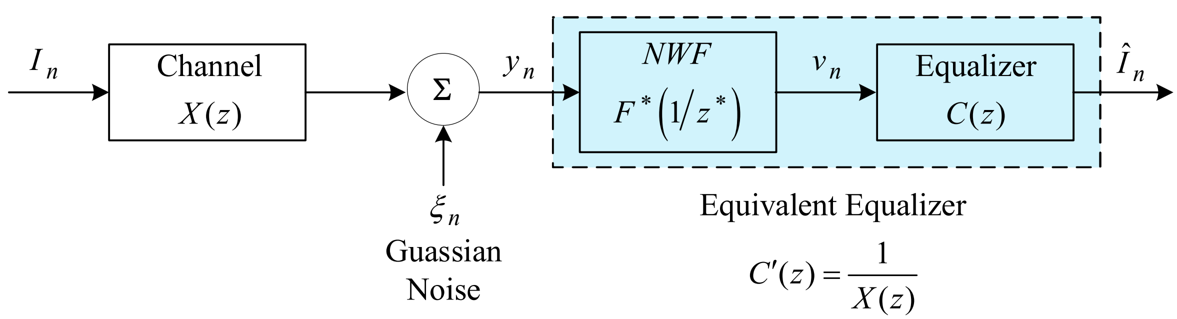


Figure II.19: Zero Forcing Equalizer – Separate Whitening

The performance of the infinite-tap equalizer that completely eliminates the intersymbol interference can be expressed in terms of the SNR at its output.

Performance of the Infinite-Tap Zero-Forcing Equalizer

Finite-Length Equalizer (Peak Distortion)

Let the equalizer have $2K+1$ taps, i.e., $c_j = 0$ for $|j| > K$.

Since $c_j = 0$ for $|j| > K$, the convolution of $\{f_n\}$ with $\{c_n\}$ is zero outside the range $-K \le n \le K + L - 1$.

With q_0 normalized to unity, the peak distortion is

$$\begin{aligned}\mathcal{D}(c) &= \sum_{\substack{n=-K \\ n\neq 0}}^{K+L-1} |q_n| \\ &= \sum_{\substack{n=-K \\ n\neq 0}}^{K+L-1} \left| \sum_j c_j f_{n-j} \right|\end{aligned} \tag{II.92}$$

It is generally impossible to completely eliminate the intersymbol interference at the output of the equalizer. The peak distortion has been shown by Lucky (1965) to be a convex function of the coefficients $\{c_j\}$. That is, it possesses a global minimum and no local minima. Its minimization can be carried out numerically using, for example, the method of steepest descent.

II.3.B. Mean-Square-Error (MSE) Criterion

In the MSE criterion, the tap weight coefficients $\{c_j\}$ of the equalizer are adjusted to minimize the mean square value of the error

$$\varepsilon_k = I_k - \hat{I}_k \tag{II.93}$$

When the information symbols $\{I_k\}$ are complex-valued, the performance index for the MSE criterion, denoted by J, is defined as

$$\begin{aligned}J &= \mathrm{E}\left[|\varepsilon_k|^2 \right] \\ &= \mathrm{E}\left[\left| I_k - \hat{I}_k \right|^2 \right]\end{aligned} \tag{II.94}$$

Note that J is a quadratic function of the equalizer coefficients $\{c_j\}$.

Infinite-Length Equalizer

In this case the estimate $\hat{I}_k$ is expressed as

$$\hat{I}_k = \sum_{j=-\infty}^{\infty} c_j v_{k-j} \tag{II.95}$$

Substitution into the expression for J and expansion of the result yields a quadratic function of the coefficients $\{c_j\}$. This function can be easily minimized with respect to the $\{c_j\}$ to yield a set (infinite in number) of linear equations for the $\{c_j\}$. Alternatively, the set of linear equations can be obtained by invoking the orthogonality principle in mean square estimation. That is, we select the coefficients $\{c_j\}$ to render the error ε_k orthogonal to the signal sequence $\{v_{k-l}^*\}$ for $-\infty < l < \infty$. Thus,

$$\mathrm{E}\left[\varepsilon_k v_{k-l}^*\right] = 0, \quad -\infty < l < \infty \tag{II.96}$$

$$\mathrm{E}\left[\left(I_k - \sum_{j=-\infty}^{\infty} c_j v_{k-j}\right) v_{k-l}^*\right] = 0 \tag{II.97}$$

$$\sum_{j=-\infty}^{\infty} c_j \, \mathrm{E}\left[v_{k-j} v_{k-l}^*\right] = \mathrm{E}\left[I_k v_{k-l}^*\right], \quad -\infty < l < \infty \tag{II.98}$$

Note that

$$\begin{aligned}\mathrm{E}\left[v_{k-j} v_{k-l}^*\right] &= \sum_{n=0}^{L} f_n^* f_{n+l-j} + N_0 \delta_{lj} \\ &= \begin{cases} x_{l-j} + N_0 \delta_{lj}, & |l-j| \le L \\ 0, & \text{otherwise} \end{cases}\end{aligned} \tag{II.99}$$

Note also that

$$\mathrm{E}\left[I_k v_{k-l}^*\right] = \begin{cases} f_{-l}^*, & -L \le l \le 0 \\ 0, & \text{otherwise} \end{cases} \tag{II.100}$$

Now, if we substitute (II.99) and (II.100) into (II.98) and take the z transform of both sides of the resulting equation, we obtain

$$C(z)\left[F(z)F^*\left(\frac{1}{z^*}\right) + N_0\right] = F^*\left(\frac{1}{z^*}\right) \tag{II.101}$$

Therefore, the transfer function of the equalizer based on the MSE criterion is

$$C(z) = \frac{F^*\left(\frac{1}{z^*}\right)}{F(z)F^*\left(\frac{1}{z^*}\right) + N_0} = \frac{F^*\left(\frac{1}{z^*}\right)}{X(z) + N_0} \tag{II.102}$$

When the noise-whitening filter is incorporated into $C(z)$, we obtain an equivalent equalizer having the transfer function

$$\begin{aligned} C'(z) &= \frac{1}{F(z)F^*\left(\frac{1}{z^*}\right) + N_0} \\ &= \frac{1}{X(z) + N_0} \end{aligned} \tag{II.103}$$

We observe that the only difference between this expression for $C'(z)$ and the one based on the peak distortion criterion is the noise spectral density factor N_0. A measure of the residual intersymbol interference and additive noise is obtained by evaluating the minimum value of J, denoted by $J_{\min}$,

$$J_{\min} = 1 - \sum_{j=-\infty}^{\infty} c_j f_{-j} \tag{II.104}$$

Finite-Length Equalizer

The output of the equalizer in the k^{th} signaling interval is

$$\hat{I}_k = \sum_{j=-K}^{K} c_j v_{k-j} \tag{II.105}$$

The MSE for the equalizer having $2K+1$ taps, denoted by $J(K)$, is

$$\begin{aligned} J(K) &= \mathrm{E}\left[\left|I_k - \hat{I}_k\right|^2\right] \\ &= \mathrm{E}\left[\left|I_k - \sum_{j=-K}^{K} c_j v_{k-j}\right|^2\right] \end{aligned} \tag{II.106}$$

Minimization of $J(K)$ with respect to the tap weights $\{c_j\}$ or, equivalently, forcing the error ε_k to be orthogonal to the signal samples v_{j-l}^*, $|l| \le K$, yields the following set of simultaneous equations

$$\sum_{j=-K}^{K} c_j \Gamma_{lj} = \xi_l, \quad l = -K, \cdots, -1, 0, 1, \cdots, K \tag{II.107}$$

where

$$\Gamma_{lj} = \begin{cases} x_{l-j} + N_0 \delta_{lj}, & |l-j| \le L \\ 0, & \text{otherwise} \end{cases} \tag{II.108}$$

$$\xi_l = \begin{cases} f_{-l}^*, & -L \le l \le 0 \\ 0, & \text{otherwise} \end{cases} \tag{II.109}$$

It is convenient to express the set of linear equations in matrix form. Thus,

$$\sum_{j=-K}^{K} c_j \Gamma_{lj} = \xi_l, \quad l = -K, \cdots, -1, 0, 1, \cdots, K \tag{II.110}$$

$$\Gamma \underline{c} = \underline{\xi} \tag{II.111}$$

where $\underline{c}$ denotes the column vector of $2K+1$ tap weight coefficients, Γ denotes the $(2K+1)\times(2K+1)$ Hermitian covariance matrix with elements Γ_{ij} and $\underline{\xi}$ is a $(2K+1)$-dimensional column vector with elements ξ_i.

$$\underline{c}_{\text{opt}} = \Gamma^{-1} \underline{\xi} \tag{II.112}$$

$$\begin{aligned} J_{\min}(K) &= 1 - \sum_{j=-K}^{0} c_j f_{-j} \\ &= 1 - \underline{\xi}^H \Gamma^{-1} \underline{\xi} \end{aligned} \tag{II.113}$$

II.3.C. Performance Characteristics of the MSE Equalizer

II.3.D. Probability of error performance of linear MSE equalizer

II.4. Fractionally Spaced Equalizers

II.5. Baseband and Passband Linear Equalizers

II.6. Decision-Feedback Equalization

II.7. Reduced Complexity ML Detectors

II.8. Iterative Equalization and Decoding—Turbo Equalization

***

III. Wireless Communication Channels

There are two fundamental aspects of wireless communications that make the channel characterization problem challenging and interesting. These aspects are by and large not as significant in wireline communication.

- First is the phenomenon of fading: the time variation of the channel strength due to the small-scale effects of multipath fading, as well as large-scale effects such as path loss via distance attenuation and shadowing by obstacles.
- Second, unlike in the wired world where each transmitter–receiver pair can often be thought of as an isolated point-to-point link, wireless users communicate over the air and there is significant interference between them. The interference can be between transmitters communicating with a common receiver (e.g., uplink of a cellular system), between signals from a single transmitter to multiple receivers (e.g., downlink of a cellular system), or between different transmitter–receiver pairs (e.g., interference between users in different cells).

Traditionally the design of wireless systems has focused on increasing the reliability of the air interface; in this context, fading and interference are viewed as nuisances that are to be countered. Recent focus has shifted more towards increasing the spectral efficiency; associated with this shift is a new point of view that fading can be viewed as an opportunity to be exploited.

III.1. Propagation Effects

III.1.A. Large-Scale Effects

- **Usually in seconds**
- **Result from:**
  - **Attenuation**
  - **Slow shadow fading**
- **Result in slow signal variations**
- **Log-normally distributed**

The free-space propagation model is used for predicting the received signal strength in the line-of-sight (LOS) environment where there is no obstacle between the transmitter and receiver. Let d denote the distance in meters. When non-isotropic antennas are used with a transmit gain of G_t and a receive gain of G_r, the received power at distance d, $P_r(d)$, is expressed by the well-known Friis equation, given as

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{III.1}$$

where P_t represents the transmit power (watts), λ is the wavelength of radiation (m), and L is the system loss factor which is independent of the propagation environment. The system loss factor represents overall attenuation or loss in the actual system hardware, including transmission line, filter, and antennas. In general, $L > 1$, but $L = 1$ if we assume that there is no loss in the system hardware. The free-space path loss PL_F without any system loss can be directly derived from (III.1) with $L = 1$ as

$$PL_F(d)_{dB} = 10\log\left(\frac{P_t}{P_r}\right) \tag{III.2}$$

Without antenna gains ($G_t = G_r = 1$), (III.2) reduces to

$$PL_F(d)_{dB} = 20\log\left(\frac{4\pi d}{\lambda}\right) \tag{III.3}$$

A more general form of the path loss model can be constructed by modifying the free-space path loss with the path loss exponent n that varies with the environment. This is known as the log-distance path loss model, in which the path loss at distance is given as

$$PL_{LD}(d)_{dB} = PL_F(d_0) + 10n\log\left(\frac{d}{d_0}\right) \tag{III.4}$$

where d_0 is a reference distance within which the path loss inherits the characteristics of free-space loss in (III.2). The reference distance d_0 must be properly determined for different propagation environments. For example, d_0 is typically set as 1 km for a cellular system with a large coverage (e.g., a cellular system with a cell radius greater than 10 km). However, it could be 100 m or 1 m, respectively, for a macro-cellular system with a cell radius of 1 km or a microcellular system with an extremely small radius. As shown in Table III.1, the path loss exponent can vary from 2 to 6, depending on the propagation environment. Note that $n = 2$ corresponds to the free space. Moreover, n tends to increase as there are more obstructions.

Table III.1

Environment	Path Loss Exponent
Free space	2
Urban area cellular radio	2.7-3.5
Shadowed urban cellular radio	3-5
In-building line-of-sight	1.6-1.8
Obstructed in-building	4-6
Obstructed in-factories	2-3

Different paths may have different path losses since the surrounding environments may vary with the location of the receiver. However, the aforementioned path loss models do not take this situation into account. A shadowing model is useful when dealing with a more realistic situation. Let X_σ denote a Gaussian random variable with a zero mean and a standard deviation σ.Then, the log-normal shadowing model is given as

$$\begin{aligned} PL(d)_{dB} &= PL_F(d_0) + 10n\log\left(\frac{d}{d_0}\right) + X_\sigma \\ &= \overline{PL}(d) + X_\sigma \end{aligned} \tag{III.5}$$

In other words, this particular model allows the receiver at the same distance to have a different path loss, which varies with the random shadowing effect.

III.1.B. Small-Scale Effects

- **Usually in milliseconds**
- **Result from:**
  - **Reflections**
  - **Scattering**
- **Result in rapid signal variations**
- **Complex Gaussian distributed**

Rapid variation of the received signal level in the short term as the user terminal moves a short distance. It is due to the effect of multiple signal paths, which cause interference when they arrive subsequently in the receive antenna with varying phases (i.e., constructive interference with the same phase and destructive interference with a different phase). Small-scale fading depends on the existence of multiple paths, speed of mobility, speed of surrounding objects, and transmission bandwidth of signal.

III.1.C. Multipath Effects

Objects may exist in the signal transmission path

- **Buildings**
- **Trees, Hills**
- **Vehicles, …**

These objects affect signal through

- **Attenuation**
- **Reflection**
- **Refraction**
- **Diffraction**
- **Scattering, …**

Several instances of the signal arrive at the receiver with different delays

- ***MULTIPATH***

Different paths can often be assumed to fade independently

Severe multipath can be experienced in urban areas

Indoor propagation has a more complex multipath structure than that of terrestrial mobile radio

- **Building structure**
- **Room layout**
- **Type of construction materials, …**

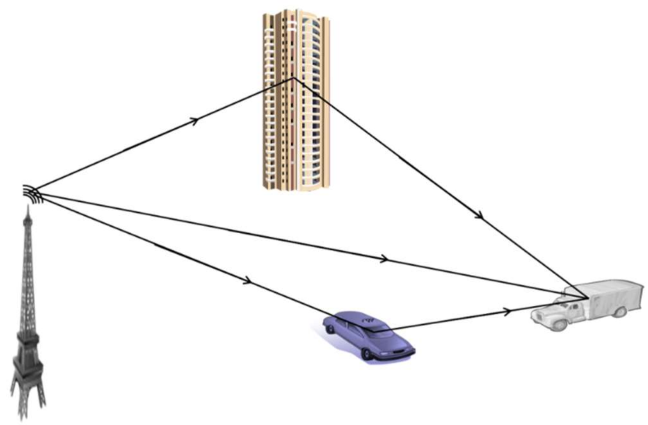


Figure III.1: Multipath

III.1.D. Multipath Mitigation

- **Multicarrier Modulation**
- **Spread Spectrum**
- **MIMO**
- **Channel Coding**

III.2. Fading Channel Characterization

III.2.A. Fading Classifications

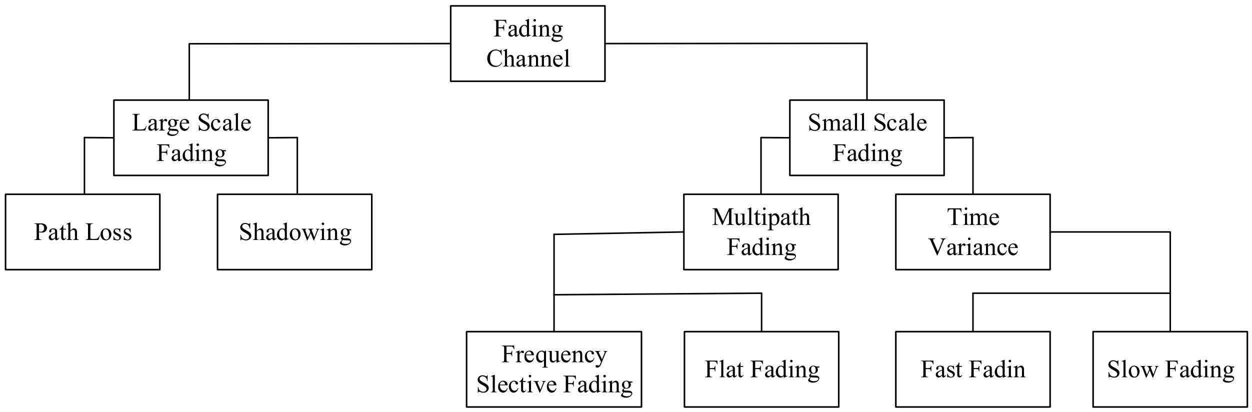


Figure III.2: Classification of fading effects

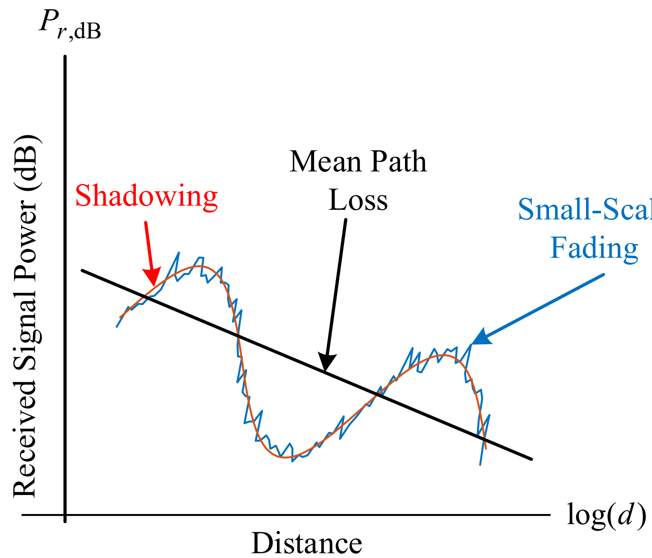


Figure III.3: Large-scale fading vs. small-scale fading

III.2.B. TIME VARYING MULIPATH CHANNEL

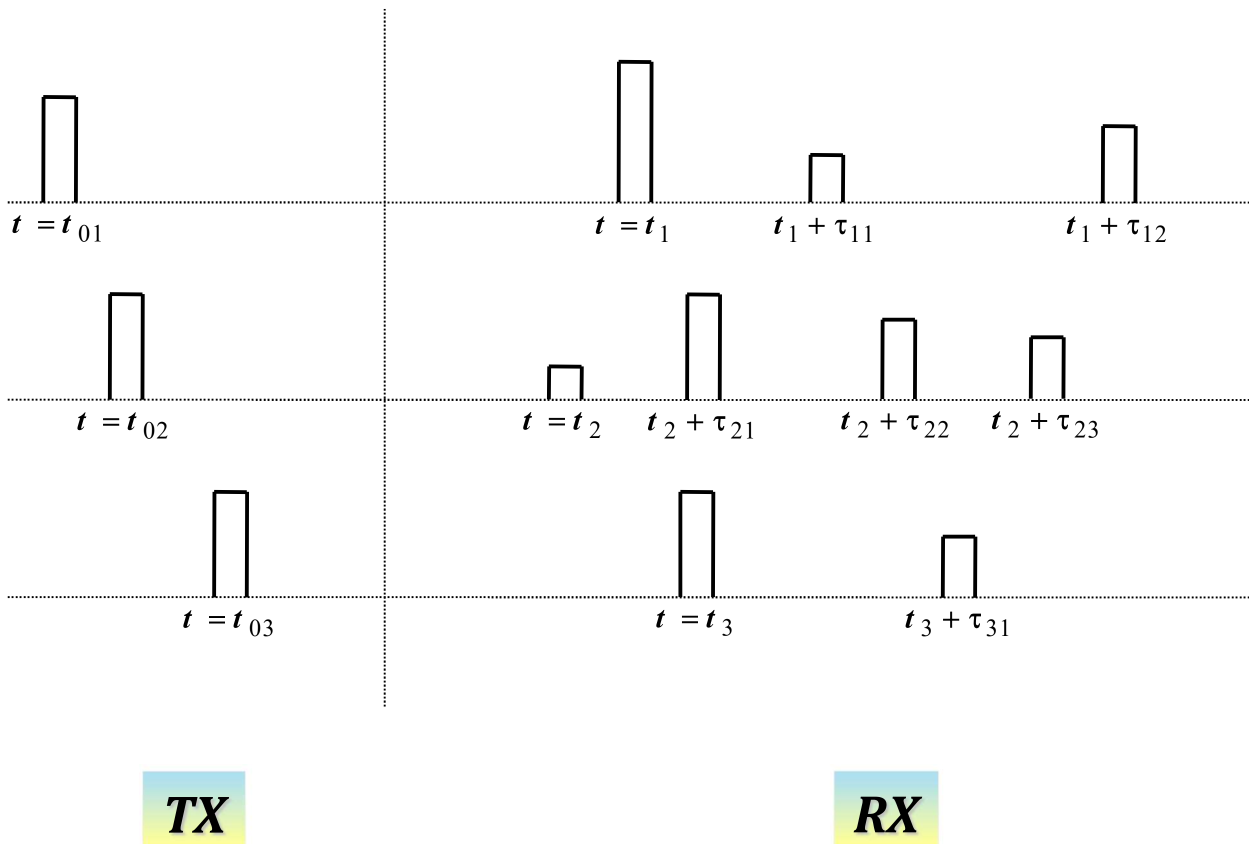


Figure III.4: Multipath transmission

III.2.C. CHANNEL IMPULSE RESPONSE

Input Signal

$$s(t) = \mathrm{Re}\left\{\tilde{s}(t)e^{j2\pi f_c t}\right\} \tag{III.6}$$

Received Signal

$$x(t) = \sum_n \alpha_n(t) s\left(t - \tau_n(t)\right) \tag{III.7}$$

Lowpass Received Signal

$$\tilde{x}(t) = \sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} \tilde{s}\left(t - \tau_n(t)\right) \tag{III.8}$$

Discrete Multipath Components

$$c(\tau, t) = \sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} \delta\left(t - \tau_n(t)\right) \tag{III.9}$$

Continuous Multipath Components

$$x(t) = \int_{-\infty}^{\infty} \alpha(\tau, t) s(t - \tau) d\tau \tag{III.10}$$

$$c(\tau, t) = \alpha(\tau, t) e^{-j2\pi f_c \tau} \tag{III.11}$$

- No Fixed Scatterers or Reflectors:

$c(\tau, t) = \alpha(\tau, t) e^{-j\theta(\tau)}$ is zero-mean complex gaussian distributed .

Envelope $|\alpha(\tau, t)|$ is Rayleigh distributed .

Angle $\theta(\tau)$ is uniformly distributed .

- Fixed Scatterers or Reflectors:

$c(\tau, t) = \alpha(\tau, t) e^{-j\theta(\tau)}$ is non-zero-mean complex gaussian distributed .

Envelope $|\alpha(\tau, t)|$ is Ricean distributed .

Angle $\theta(\tau)$ is uniformly distributed .

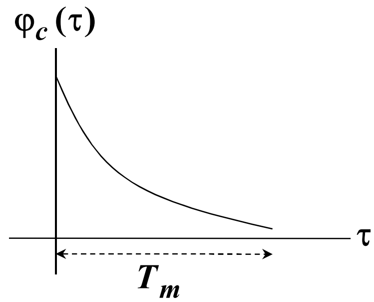
III.2.D. Channel Correlation Functions

Multipath Intensity Profile (MIP) - Also called "Delay Power Spectrum"

$$\varphi_c(\tau_1,\tau_2,\Delta t)=\frac{1}{2}E\left[c^*(\tau_1,t)c(\tau_2,t+\Delta t)\right] \quad \text{(III.12)}$$

- Assume uncorrelated scattering and fix time variable.

$$\varphi_c(\tau_1,\tau_2,\Delta t)\rightarrow\varphi_c(\tau) \quad \text{(III.13)}$$



T_m: Multipath Spread .

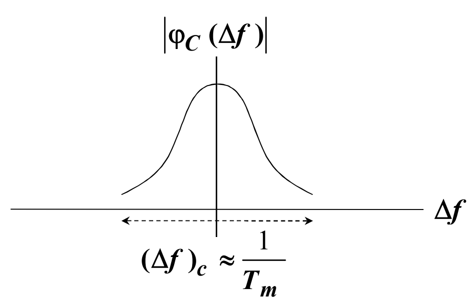
Spaced Time Spaced Frequency Correlation Function

$$\varphi_C(\Delta f,\Delta t)=\int_{-\infty}^{\infty}\varphi_c(\tau,\Delta t)e^{-j2\pi\Delta f\tau}d\tau \quad \text{(III.14)}$$

Assume uncorrelated scattering and fix time variable. Then,

$$\varphi_C(\Delta f,\Delta t)\rightarrow\varphi_C(\Delta f) \quad \text{(III.15)}$$

Spaced Frequency Correlation Function



$(\Delta f)_c$: Coherence Bandwidth .

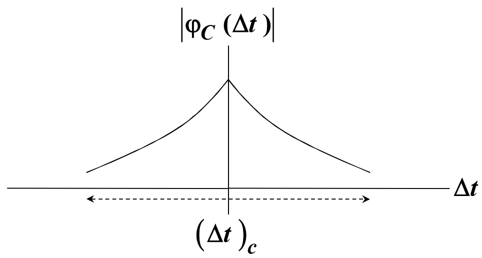
Doppler Power Spectrum

$$S_C(\Delta f,\lambda)=\int_{-\infty}^{\infty}\varphi_C(\Delta f,\Delta t)e^{-j2\pi\lambda\Delta t}d\Delta t \quad \text{(III.16)}$$

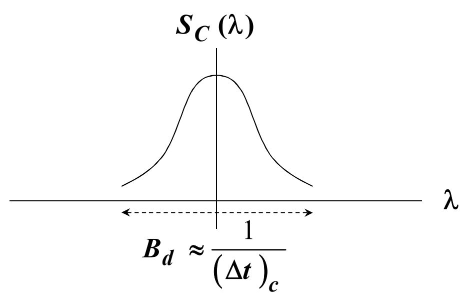
$$\Delta f=0$$
$$\varphi_C(\Delta f,\Delta t)\to\varphi_C(\Delta t)\text{: Spaced Time Correlation Function} \quad \text{(III.17)}$$
$$S_C(\Delta f,\lambda)\to S_C(\lambda)$$

$$S_C(\lambda)=\int_{-\infty}^{\infty}\varphi_C(\Delta t)e^{-j2\pi\lambda\Delta t}d\Delta t\text{: Doppler Power Spectrum} \quad \text{(III.18)}$$

Coherence Time and Doppler Spread



where $(\Delta t)_c$ is called the coherence time.



where B_d is called the Doppler shift.

III.2.E. MULTIPATH DEFINITIONS

Multipath Intensity Profile	Signal intensity (power) as a function of multipath time delay
Multipath Spread	Time delay between arrival of first and last multipath components
Coherence Bandwidth	Range of frequencies over which the channel response is highly correlated
Coherence Time	Time interval over which the channel response if highly correlated
Doppler Spread	Difference in Doppler shifts

Frequency Selectivity

If coherence bandwidth is small compared to signal bandwidth, the channel is said to be ***Frequency-Selective***, otherwise, channel is said to be ***Frequency Non-Selective*** or ***Frequency Flat***

Usually happens when date rate is high

Signal is severely distorted

Distinct multipaths can be resolved

ISI can take place

Solutions: Equalization and Multicarrier Modulation

Time Selectivity

If coherence time is small compared to signal duration, the channel is said to be *Time-Selective*, *Fast Fading* or otherwise, channel is said to be *Time Non-Selective* or *Slowly Fading*

Usually happens when user is highly mobile

Signal is severely distorted

Conventional equalization too complex to realize. Channel tracking is usually required

Solution: Spread Spectrum

III.3. Time-Dispersive vs. Frequency-Dispersive Fading

As a mobile terminal moves, the specific type of fading for the corresponding receiver depends on both the transmission scheme and channel characteristics. The transmission scheme is specified with signal parameters such as signal bandwidth and symbol period. Meanwhile, wireless channels can be characterized by two different channel parameters, multipath delay spread and Doppler spread, each of which causes time dispersion and frequency dispersion, respectively. Depending on the extent of time dispersion or frequency dispersion, frequency-selective fading or time-selective fading is induced, respectively.

III.3.A. TIME DISPERSION: FREQUENCY-SELECTIVE FADING CHANNEL

For a given channel frequency response, frequency selectivity is generally governed by signal bandwidth. Due to multipath-induced time dispersion, channel response varies with frequency. Here, the transmitted signal is subject to frequency-non-selective (narrowband) fading when the signal bandwidth is narrow enough such that it may be transmitted over the flat response. On the other hand, the signal is subject to frequency-selective fading when signal bandwidth is wide enough such that it may be filtered out by the finite channel bandwidth.

As shown in Figure III.5, the received signal undergoes frequency-non-selective fading (narrowband) as long as the bandwidth of the wireless channel is wider than that of the signal, while maintaining a constant amplitude and linear phase response within the passband. Constant amplitude scaling over the entire signal bandwidth induces flat fading, which is another term to refer to frequency-non-selective fading. Here, a narrower bandwidth implies that symbol period is greater than delay spread of the multipath channel. As long as this is the case, the current symbol does not affect the subsequent symbol as much over the next symbol period, implying that intersymbol interference (ISI) is not significant. Even while amplitude is slowly time-varying in the frequency-non-selective fading channel, it is often referred to as a narrowband channel, since the signal bandwidth is much narrower than the channel bandwidth.

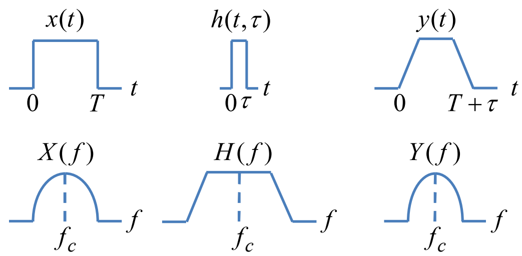


Figure III.5: Pulse transmission in Frequency non-selective fading

To summarize the observation above, a transmit signal is subject to frequency-non-selective fading under the following conditions:

$$B_s \ll (\Delta f)_c \text{ and } T_s \gg T_m \tag{III.19}$$

where B_s is the signal bandwidth and T_s is the symbol duration.

The signal undergoes frequency-selective fading when the wireless channel has a constant amplitude and linear phase response only within a channel bandwidth narrower than the signal bandwidth. In this case, the channel impulse response has a larger delay spread than a symbol period of the transmit signal. Due to the short symbol duration as compared to the multipath delay spread, multiple-delayed copies of the transmit signal are significantly overlapped with the subsequent symbol, incurring inter-symbol interference (ISI). The term frequency selective channel is used simply because the channel amplitude frequency response varies with the frequency, as opposed to the frequency-flat nature of the frequency non-selective fading channel.

As illustrated in Figure III.6, the occurrence of ISI is obvious in the time domain since channel delay spread is much greater than the symbol period. This implies that signal bandwidth is greater than coherence bandwidth and thus, the received signal will have a different amplitude in the frequency response (i.e., undergo frequency-selective fading).

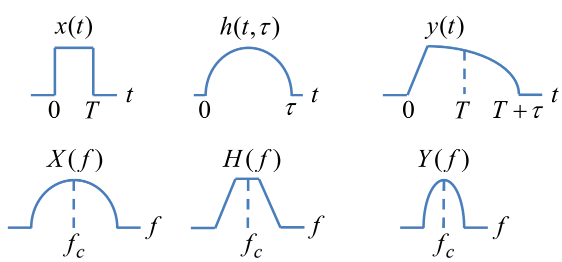


Figure III.6: Pulse transmission in Frequency-selective fading

To summarize the observation above, transmit signal is subject to frequency-selective fading under the following conditions:

$$B_s > (\Delta f)_c \text{ and } T_s < T_m \tag{III.20}$$

A channel is typically classified as frequency-selective when $T_m > 0.1T_s$.

III.3.B. Frequency Dispersion: Time-Selective Fading Channel

Depending on the extent of the Doppler spread, the received signal undergoes fast or slow fading. In a fast fading channel, the coherence time is smaller than the symbol period and thus, a channel impulse response quickly varies within the symbol period. Variation in the time domain is closely related to movement of the transmitter or receiver, which incurs a spread in the frequency domain, known as a Doppler shift.

Therefore, $T_s > (\Delta t)_c$ implies that $B_s < B_d$. The transmit signal is subject to fast fading under the following conditions:

$$T_s > (\Delta t)_c \text{ and } B_s < B_d \tag{III.21}$$

On the other hand, consider the case where the channel impulse response varies slowly as compared to variation in the baseband transmit signal. In this case, we can assume that the channel does not change over the duration of one or more symbols and thus, it is referred to as a static channel. This implies that the Doppler spread is much smaller than the bandwidth of the baseband transmit signal. In conclusion, transmit signal is subject to slow fading under the following conditions:

$$T_s \ll (\Delta t)_c \text{ and } B_s \gg B_d \tag{III.22}$$

In slow fading, a particular fade level will affect many successive symbols, which may lead to burst errors, whereas in fast fading the fading decorrelates from symbol to symbol. In this latter case and when the communication receiver decisions are based on an observation of the received signal over two or more symbol times (such as differentially coherent or coded communications), it becomes necessary to consider the variation of the fading channel from one symbol interval to the next. Various autocorrelation models and their corresponding power spectral densities are tabulated in Table III.2, in which for convenience the variance of the fast-fading process is normalized to unity.

Table III.2: Correlation of Various Types of Fading Processes of Practical Interest

Type of Fading Spectrum	Fading Autocorrelation Function
Rectangular	$\frac{\sin 2\pi f_d \tau}{2\pi f_d \tau}$

Gaussian	$e^{-\pi^2 f_d^2 \tau^2}$
Land Mobile	$J_0(2\pi f_d \tau)$
First Order Butterworth	$e^{-2\pi\lvert f_d \tau \rvert}$
Second Order Butterworth	$e^{-\frac{\pi\lvert f_d \tau\rvert}{\sqrt{2}}}\left[\cos\left(\frac{\pi f_d \tau}{\sqrt{2}}\right)+\sin\left(\frac{\pi\lvert f_d \tau\rvert}{\sqrt{2}}\right)\right]$

III.4. Statistical Models

III.4.A. SMALL-SCALE FADING PARAMETERS

Characteristics of a multipath fading channel are often specified by a power delay profile (PDP), by means of which different multiple signal paths are characterized by their relative delay and average power. The relative delay is an excess delay with respect to the reference time while average power for each path is normalized by that of the first path.

Table III.3: Power Delay Profile Example (ITU-R Pedestrian A Model)

Path	Relative Delay (ns)	Average Power (dB)
1	0	0
2	110	-9.7
3	190	-19.2
4	410	-22.8

Several instances of the signal, received via the multipaths, interfere with one another destructively or constructively. Multipath fading imposes random amplitude and phase variations onto the transmitted waveform. This degrades system performance by inducing reception errors.

Mean excess delay and RMS delay spread are useful channel parameters that provide a reference of comparison among the different multipath fading channels. They are also bases of a general guideline to designing a wireless transmission system.

Let τ_l denote the channel delay of the lth path while a_l and $P(\tau_l)$ denote the amplitude and power of the same path, respectively. The mean excess delay $\bar{\tau}$ is given by the first moment of the PDP as

$$\bar{\tau} = \frac{\sum_{l=1}^{L} a_l^2 \tau_l}{\sum_{l=1}^{L} a_l^2} = \frac{\sum_{l=1}^{L} \tau_l P(\tau_l)}{\sum_{l=1}^{L} P(\tau_l)} \tag{III.23}$$

The RMS delay spread τ_{rms} is given by the square root of the second moment (mean square value) of the PDP as

$$\tau_{rms} = \sqrt{\overline{\tau^2}} \tag{III.24}$$

where

$$\overline{\tau^2} = \frac{\sum_{l=1}^{L} a_l^2 \tau_l^2}{\sum_{l=1}^{L} a_l^2} = \frac{\sum_{l=1}^{L} \tau_l^2 P(\tau_l)}{\sum_{l=1}^{L} P(\tau_l)} \tag{III.25}$$

The standard deviation of the delay spread σ_τ is given by the square root of the second central moment (the variance σ_τ^2) of the PDP as

$$\sigma_\tau = \sqrt{\sigma_\tau^2} = \sqrt{\overline{\tau^2} - (\bar{\tau})^2} \tag{III.26}$$

In general, coherence bandwidth is inversely-proportional to the standard deviation of the delay spread, that is

$$(\Delta f)_c \approx \frac{1}{\sigma_\tau} \tag{III.27}$$

Sometimes, the coherence bandwidth is defined as a bandwidth with correlation of 0.9 or above, then it is assumed that

$$(\Delta f)_c \approx \frac{1}{50\sigma_\tau} \tag{III.28}$$

If the coherence bandwidth is defined as a bandwidth with correlation of 0.5 or above, then it is assumed that

$$(\Delta f)_c \approx \frac{1}{5\sigma_\tau} \tag{III.29}$$

III.4.B. PROBABILITY DENSITIES

- **Large number of multipaths with no LOS → Envelope is Rayleigh**

$$f_R(r) = \frac{2r}{\Omega} e^{-r^2/\Omega}$$
$$\Omega = \mathrm{E}\left[R^2\right]$$

- **Large number of multipaths with LOS ➔ Envelope is Ricean**

- **More flexible distribution (2 parameters)**
  - **Envelope is Nakagami with a fading severity parameter *m***

$m \geq 0.5$	
m=0.5	Most severe fading
$m = 1$	Rayleigh
$m = 0.5$	Single-sided Gaussian
$m \to \infty$	No fading

III.5. System Performance Measures

III.5.A. AVERAGE SIGNAL-TO-NOISE RATIO

Traditionally, the term "noise" in signal-to-noise ratio refers to the ever-present thermal noise at the input of the receiver. In the context of a communication system subject to fading, the more appropriate performance measure is the average SNR, where the term "average" refers to statistical averaging over the probability distribution of the fading.

In simple mathematical terms, if γ denotes the instantaneous SNR [a random variable (RV)] at the receiver output that includes the effect of fading, then the average SNR is given by

$$\overline{\gamma} = \int_0^\infty \lambda f_\gamma(\lambda) d\lambda \qquad \text{(III.30)}$$

Let's rewrite (III.30) in terms of the moment generating function (MGF) associated with γ:

$$M_\gamma(s) = \int_0^\infty f_\gamma(\lambda) e^{s\lambda} d\lambda \qquad \text{(III.31)}$$

Taking the first derivative of (III.31) with respect to s and evaluating the result at $s = 0$, we immediately see from (III.30) that

$$\overline{\gamma} = \frac{d}{ds} M(s) \bigg|_{s=0} \qquad \text{(III.32)}$$

That is, the ability to evaluate the MGF of the instantaneous SNR (perhaps in closed form) allows immediate evaluation of the average SNR via a simple mathematical operation, namely, differentiation.

In some important cases (e.g., diversity), the SNR can be expressed as a sum (or some other combination) of several individual SNRs (e.g., on diversity branches). For example, we could have

$$\gamma = \sum_{l=1}^{L} \gamma_l \tag{III.33}$$

It is often reasonable in practice to assume that the individual SNRs are statistically independent. In such instances, the MGF of the overall SNR can be expressed as a product of the individual MGFs, i.e.,

$$M_\gamma(s) = \prod_{l=1}^{L} M_{\gamma_l}(s) \tag{III.34}$$

Generally, finding the MGF in (III.34) is significantly easier that finding the joint PDF, which in the simplest case of independent individual SNRs, requires an L-fold convolution operation.

III.5.B. OUTAGE PROBABILITY

The outage probability is defined as the probability that the instantaneous error probability exceeds a specified value. Equivalently, the outage probability is the probability that the output SNR γ falls below a certain specified threshold γ_{th}. Mathematically speaking, we have

$$P_{\text{out}} = \int_0^{\gamma_{\text{th}}} f_\gamma(\lambda) d\lambda \tag{III.35}$$

This is the cumulative distribution function (CDF) of γ, evaluated at $\gamma = \gamma_{th}$. Mathematically,

$$P_{out} = F_\gamma(\gamma_{th}) \tag{III.36}$$

It is well-known that the PDF and the CDF are related by

$$f_\gamma(\gamma) = \frac{d}{d\gamma} F_\gamma(\gamma) \tag{III.37}$$

Since $F_\gamma(0) = 0$, then the Laplace transforms of these two functions are related by

$$\begin{aligned} \mathcal{L}\left\{F_\gamma(s)\right\} &= \frac{1}{s}\mathcal{L}\left\{f_\gamma(s)\right\} \\ \hat{F}(s) &= \frac{1}{s}\hat{f}_\gamma(s) \\ &= \frac{M_\gamma(-s)}{s} \end{aligned} \tag{III.38}$$

Using (III.36),

$$P_{out} = \mathcal{L}^{-1}\left\{\frac{M(-s)}{s}\right\}_{\gamma=\gamma_{th}} \tag{III.39}$$

III.5.C. AVERAGE BIT ERROR PROBABILITY

The average bit error probability (BEP) is much more difficult to compute than the average SNR and the outage probability. However, it is much more revealing about the nature of the system behavior.

The primary reason for the difficulty in evaluating average BEP is the fact that the conditional (on the fading) BEP is, in general, a nonlinear function of the instantaneous SNR. The nature of the nonlinearity is a function of the modulation/detection scheme employed by the system. Nevertheless, we will see that an MGF-based approach is still quite useful in simplifying the BEP analysis.

Suppose that the conditional BEP is of the form

$$P_b(E\,|\,\gamma) = C_1 e^{-a_1\gamma} \tag{III.40}$$

Equation (III.40) can represent the cases of differentially coherent detection of binary phase-shift-keying (PSK) signal and noncoherent detection of binary orthogonal frequency-shift-keying (FSK) signals. Then, the average BEP can be written as

$$\begin{aligned} P_b(E) &= \int_0^\infty P_b(E\,|\,\gamma) f_\gamma(\gamma)\,d\gamma \\ &= \int_0^\infty C_1 e^{-a_1\gamma} f_\gamma(\gamma)\,d\gamma \\ &= C_1 M_\gamma(-a_1) \end{aligned} \tag{III.41}$$

Note that $M_\gamma(s)$, and hence $P_b(E)$, depends only on the assumed fading channel model.

Suppose next that the relationship between $P_b(E\,|\,\gamma)$ and γ can be expressed as an integral whose integrand has an exponential dependence on γ in the form

$$P_b(E\,|\,\gamma) = \int_{\xi_1}^{\xi_2} C_2 h(\xi) e^{-a_2 g(\xi)\gamma}\,d\xi \tag{III.42}$$

where $h(\xi)$ and $g(\xi)$ are arbitrary functions of the integration variable, and ξ_1 and ξ_2 are finite. Note that (III.40) becomes a special case of (III.42) when $h(\xi)$ is allowed to be a Dirac delta function that is located in the interval $\xi_1 \le \xi \le \xi_2$. The $P_b(E\,|\,\gamma)$ form in (III.42) can be encountered, for example, in the analysis of coherent detection of PSK and noncoherent detection of QPSK. Averaging over the fading yields

$$P_b(E) = C_2 \int_{\xi_1}^{\xi_2} h(\xi) M_\gamma(-a_2 g(\xi)) d\xi \tag{III.43}$$

Integrals of the form in (III.43) can, in many cases, be obtained in closed form. The above approach for evaluating average error probability is referred to as the unified MGF-based approach.

Not every fading channel communication problem fits the above description. Thus, alternative, but still simple and accurate, techniques are desirable for evaluating system error probability in such circumstances.

Assignment III.1

Read the remainder of section 1.1.3 Average Bit Error Probability (BEP) of Alouini's book.

III.5.D. Amount of Fading (AF)

This performance measure is most appropriate in systems with diversity and combining. Let γ_t denote the total instantaneous SNR at the combiner output, we define AF by

$$\begin{aligned} \mathrm{AF} &= \frac{\sigma_{\gamma_t}^2}{\bar{\gamma}_t^2} \\ &= \frac{\overline{\gamma_t^2} - \bar{\gamma}_t^2}{\bar{\gamma}_t^2} \end{aligned} \tag{III.44}$$

which can be expressed in terms of the MGF of γ_t by

$$\mathrm{AF} = \frac{\left.\frac{d^2}{ds^2} M_{\gamma_t}(s)\right|_{s=0} - \left(\left.\frac{d}{ds} M_{\gamma_t}(s)\right|_{s=0}\right)^2}{\left(\left.\frac{d}{ds} M_{\gamma_t}(s)\right|_{s=0}\right)^2} \tag{III.45}$$

Because the AF is computed at the output of the combiner, its evaluation will reflect the behavior of the particular diversity combining technique as well as the statistics of the fading channel and thus, as mentioned above, is a measure of the performance of the entire system.

III.5.E. Average Outage Duration (AOD)

In certain communication system applications such as adaptive transmission schemes, the performance metrics discussed above do not provide enough information for the overall system design and configuration. In that case, in addition to these performance measures, the frequency of outages and the average outage duration (AOD) (also known as the "average fade duration") are important performance criteria for the proper selection of the transmission symbol rate, interleaver depth, packet length, and/or time slot duration.

In purely noise-limited systems, an outage is declared whenever the output SNR γ falls below a predetermined threshold γ_{th}.

The AOD, $T(\gamma_{th})$, in seconds, is a measure of how long, on the average, the system remains in the outage state. Mathematically speaking, the AOD is given by

$$T(\gamma_{th}) = \frac{P_{out}}{N(\gamma_{th})} \tag{III.46}$$

where $N(\gamma_{th})$ is the frequency of outages or equivalently the average level crossing rate (LCR), which is given by

$$N(\gamma_{th}) = \int_0^{\infty} \dot{\gamma} f_{\gamma,\dot{\gamma}}(\gamma_{th},\dot{\gamma}) d\dot{\gamma} \tag{III.47}$$

where $f_{\gamma,\dot{\gamma}}(\gamma,\dot{\gamma})$ is the joint PDF of γ and its time derivative $\dot{\gamma}$.

III.6. Performance Analysis: Slowly Fading Frequency Non-Selective Channels

$$\tilde{x}(t) = \alpha e^{-j\phi}\tilde{s}(t) + z(t), \quad 0 \le t < T \tag{III.48}$$

$$\begin{aligned} P(e/\gamma_b) &= g(\gamma_b) \\ \gamma_b &= \frac{\alpha^2 E_b}{N_0} \end{aligned} \tag{III.49}$$

$$P(e) = \int_0^{\infty} P(e/\gamma_b) f_{\gamma_b}(\gamma_b) d\gamma_b \tag{III.50}$$

PSK

$$P(e/\gamma_b) = Q\left(\sqrt{2\gamma_b}\right) \tag{III.51}$$

$$\begin{aligned} f_{\gamma_b}(\gamma_b) &= \frac{1}{\bar{\gamma}_b} e^{-\gamma_b/\bar{\gamma}_b} \\ \bar{\gamma}_b &= \frac{E_b}{N_0} \mathrm{E}\left[\alpha^2\right] \end{aligned} \tag{III.52}$$

$$P(e) = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}_b}{1+\bar{\gamma}_b}}\right) \tag{III.53}$$

FSK

$$P(e/\gamma_b) = Q\left(\sqrt{\gamma_b}\right) \tag{III.54}$$

$$P(e) = \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\gamma}_b}{2+\overline{\gamma}_b}}\right) \tag{III.55}$$

DPSK

$$P(e/\gamma_b) = \frac{1}{2}e^{-\gamma_b} \tag{III.56}$$

$$P(e) = \frac{1}{2\left(1+\overline{\gamma}_b\right)} \tag{III.57}$$

NCFSK

$$P(e/\gamma_b) = \frac{1}{2}e^{-\gamma_b/2} \tag{III.58}$$

$$P(e) = \frac{1}{2+\overline{\gamma}_b} \tag{III.59}$$

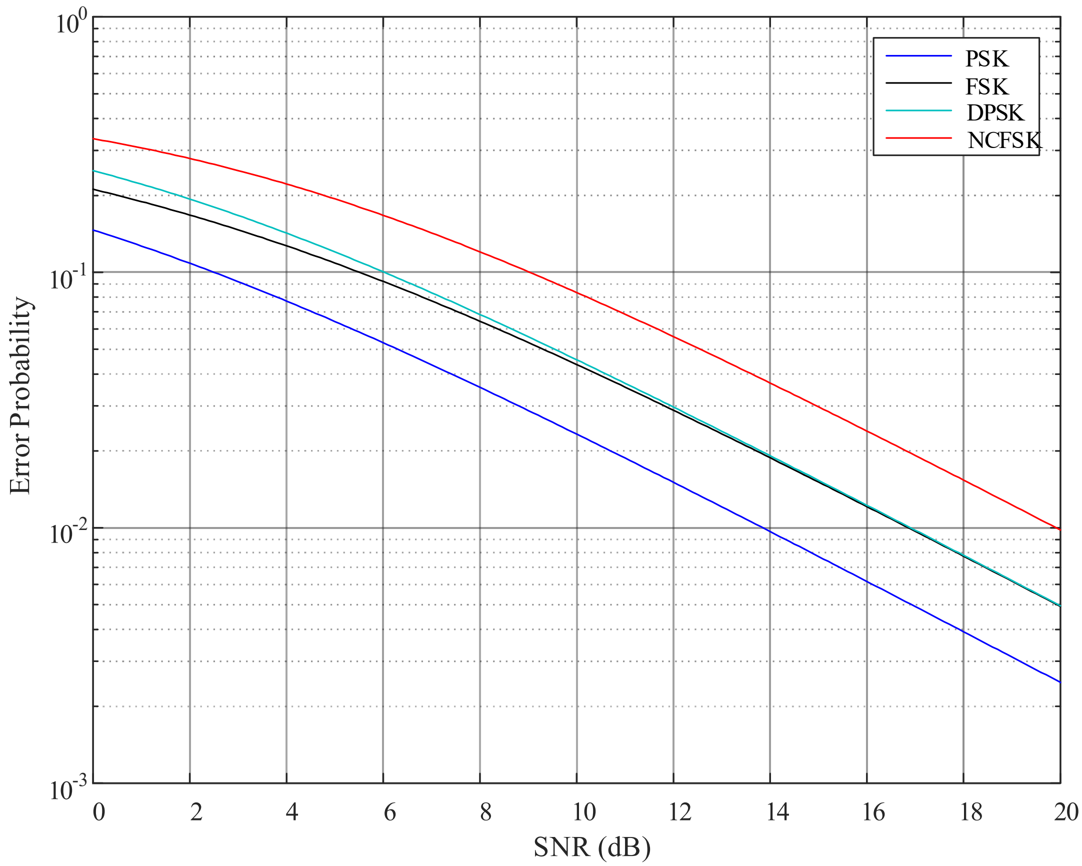


Figure III.7: Error probability under Rayleigh fading

Asymptotic P(e)

$$\bar{\gamma}_b >> 1 \tag{III.60}$$

$$x << 1 \rightarrow$$

$$\sqrt{\frac{1}{1+x}} \approx 1 - \frac{x}{2} \tag{III.61}$$

$$\sqrt{1+x} \approx 1 + \frac{x}{2}$$

$$P(e) \approx \begin{cases} \dfrac{1}{4\overline{\gamma}_b} & PSK \\ \dfrac{1}{2\overline{\gamma}_b} & FSK \\ \dfrac{1}{2\overline{\gamma}_b} & DPSK \\ \dfrac{1}{\overline{\gamma}_b} & NCFSK \end{cases} \tag{III.62}$$

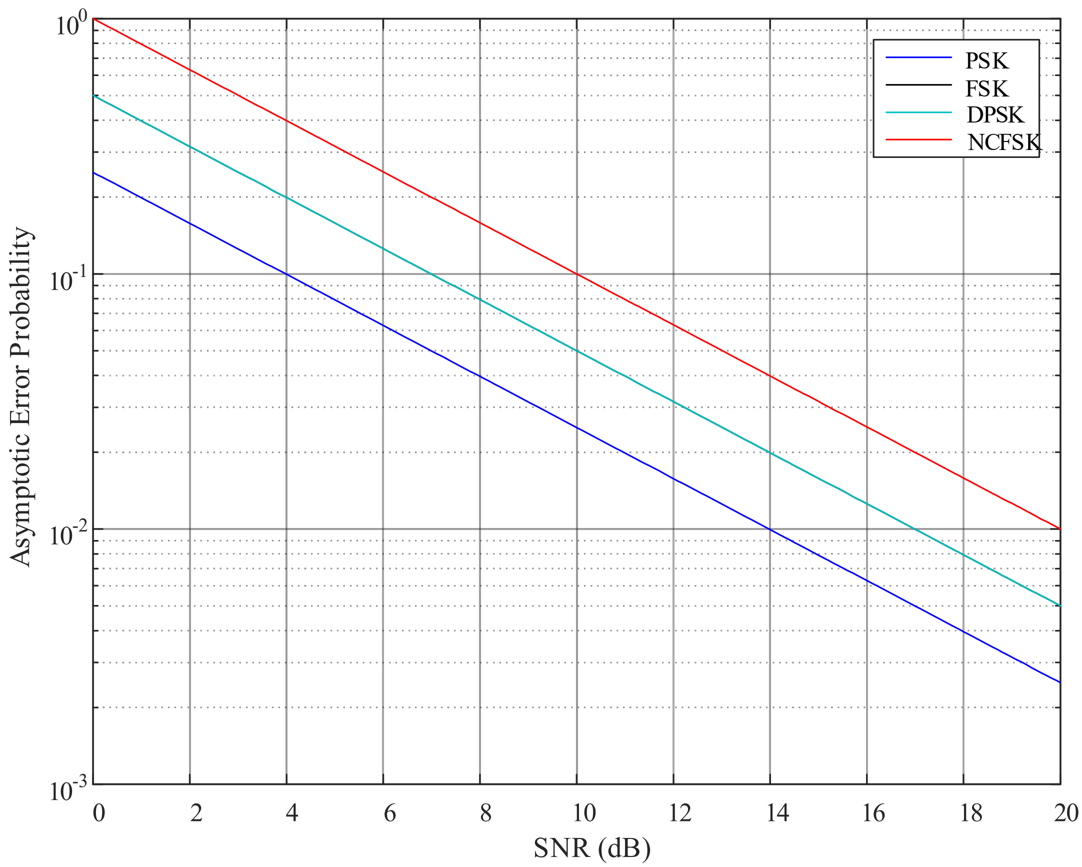
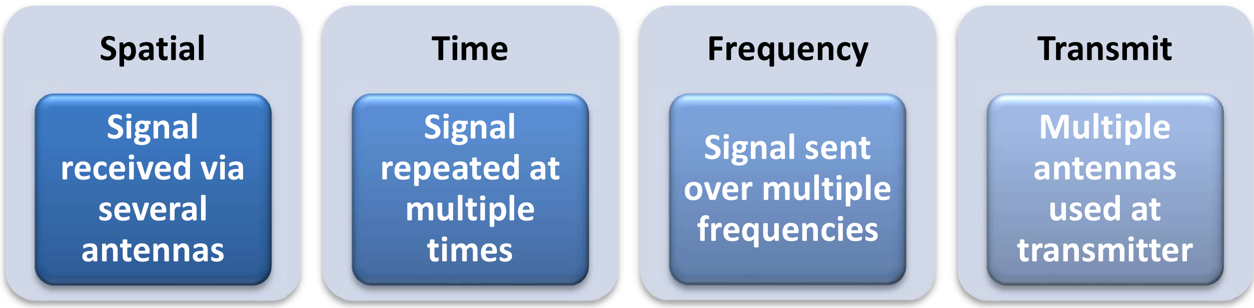


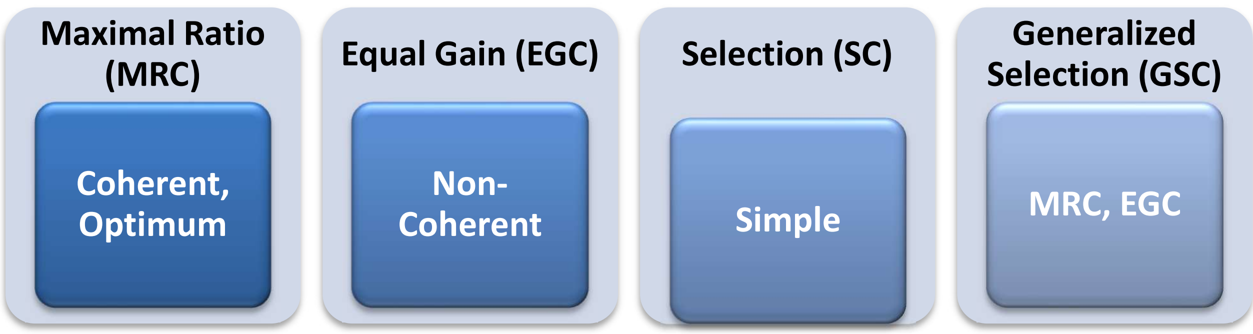
Figure III.8: Asymptotic error probability under Rayleigh fading

III.7. Diversity and Combining

III.7.A. TYPES OF DIVERSITY:



III.7.B. TYPES OF COMBINING



III.7.C. BINARY PSK- MRC

$$P(e) = P(e/1) = P(U < 0)$$
$$U = \sum_{l=1}^{L} U_l \qquad \text{(III.63)}$$

$$U_l = \mathrm{Re}\left\{ \int_0^T \tilde{r}_l(t)\alpha_l e^{j\phi_l} \sqrt{\frac{E_b}{T}}\, dt \right\} \qquad \text{(III.64)}$$

$$\tilde{r}_l(t) = \alpha_l e^{-j\phi_l} \sqrt{\frac{E_b}{T}} + z_l(t), \quad l = 1, 2, \cdots, L \qquad \text{(III.65)}$$

$$\begin{aligned} U_l &= \alpha_l^2 E_b + \alpha_l \sqrt{\frac{E_b}{T}}\, \mathrm{Re}\left\{ \int_0^T e^{j\phi_l} z_l(t)\, dt \right\} \\ &= \alpha_l^2 E_b + \alpha_l N_l \end{aligned} \qquad \text{(III.66)}$$

Conditioned on $\{\alpha_l\}_{l=1}^{L}$, U is Gaussian.

$$\mu_l = E_b \alpha_l^2 \tag{III.67}$$

$$\sigma_l^2 = E_b N_0 \alpha_l^2 \tag{III.68}$$

$$\mu_U = E_b \sum_{l=1}^{L} \alpha_l^2 \tag{III.69}$$

$$\sigma_U^2 = E_b N_0 \sum_{l=1}^{L} \alpha_l^2 \tag{III.70}$$

$$P(e/\gamma_b) = Q\left(\sqrt{2\gamma_b}\right) \tag{III.71}$$

$$\begin{aligned} \gamma_b &= \sum_{l=1}^{L} \gamma_l \\ \gamma_l &= \frac{E_b}{N_0} \alpha_l^2 \end{aligned} \tag{III.72}$$

$$P(e) = \int_0^\infty P(e/\gamma_b) f_{\gamma_b}(\gamma_b) d\gamma_b \tag{III.73}$$

Assuming $\{\alpha_l\}_{l=1}^{L}$ are independent and identically distributed, let

$$\overline{\gamma}_c = \frac{E_b}{N_0} \mathrm{E}\left[\alpha_l^2\right] \tag{III.74}$$

Then,

$$f_{\gamma_b}(\gamma_b) = \frac{1}{(L-1)!\overline{\gamma}_c^L} \gamma_b^{L-1} e^{-\gamma_b/\overline{\gamma}_c} \tag{III.75}$$

$$P(e) = \left(\frac{1-\mu}{2}\right)^L \sum_{l=0}^{L-1} \binom{L-1+l}{l} \left(\frac{1+\mu}{2}\right)^l \tag{III.76}$$

$$\mu = \sqrt{\frac{\overline{\gamma}_c}{1+\overline{\gamma}_c}} \tag{III.77}$$

III.7.D. MRC – Asymptotic P(e)

Assume $\overline{\gamma}_c \gg 1$.

$$P_{PSK}(e) = \left(\frac{1}{4\bar{\gamma}_c}\right)^L \binom{2L-1}{L} \tag{III.78}$$

$$P_{FSK}(e) = \left(\frac{1}{2\bar{\gamma}_c}\right)^L \binom{2L-1}{L} \tag{III.79}$$

$$P_{DPSK}(e) = \left(\frac{1}{2\bar{\gamma}_c}\right)^L \binom{2L-1}{L} \tag{III.80}$$

$$P_{NCFSK}(e) = \left(\frac{1}{\bar{\gamma}_c}\right)^L \binom{2L-1}{L} \tag{III.81}$$

III.8. Frequency Selective Channels

III.8.A. Tapped-Delay-Line Channel Model

Let the bandpass signal bandwidth be $W >> (\Delta f)_c$. The lowpass equivalent bandwidth is

$$W_s = \frac{W}{2} \tag{III.82}$$

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} \tilde{s}\left(\frac{n}{W}\right) \frac{\sin\left[\pi W\left(t - n/W\right)\right]}{\pi W\left(t - n/W\right)} \tag{III.83}$$

$$\tilde{S}(f) = \begin{cases} \dfrac{1}{W} \displaystyle\sum_{n=-\infty}^{\infty} \tilde{s}\left(\frac{n}{W}\right) e^{-j2\pi fn/W}, & |f| \le \dfrac{W}{2} \\ 0, & \text{otherwise} \end{cases} \tag{III.84}$$

The noiseless received signal from a frequency-selective channel is given by

$$\tilde{r}(t) = \int_{-\infty}^{\infty} \tilde{S}(f) C(f,t) e^{j2\pi ft} df \tag{III.85}$$

Substitution for $\tilde{S}(f)$ from (III.84) into (III.85) yields

$$\begin{aligned} \tilde{r}(t) &= \frac{1}{W} \sum_{n=-\infty}^{\infty} \tilde{s}\left(\frac{n}{W}\right) c\left(t - n/W, t\right) \\ &= \frac{1}{W} \sum_{n=-\infty}^{\infty} \tilde{s}\left(t - n/W\right) c\left(n/W, t\right) \end{aligned} \tag{III.86}$$

It is convenient to define a set of time-variable channel coefficients as

$$c_n(t) = \frac{1}{W} c\left(\frac{n}{W}, t\right) \qquad \text{(III.87)}$$

Then (III.86) expressed in terms of these channel coefficients becomes

$$\tilde{r}(t) = \sum_{n=-\infty}^{\infty} c_n(t)\tilde{s}\left(t - n/W\right) \qquad \text{(III.88)}$$

The form for the received signal in (III.88) implies that the time-variant frequency-selective channel can be modeled or represented as a tapped delay line with tap spacing $1/W$ and tap weight coefficients $\{c_n(t)\}$. We deduce from Equation (III.88) that the low-pass impulse response for the channel is

$$c(\tau, t) = \sum_{n=-\infty}^{\infty} c_n(t)\delta\left(\tau - n/W\right) \qquad \text{(III.89)}$$

While the corresponding time-variant transfer function is

$$C(f, t) = \sum_{n=-\infty}^{\infty} c_n(t) e^{-j2\pi f n/W} \qquad \text{(III.90)}$$

Thus, with an equivalent low-pass-signal having a bandwidth $\frac{W}{2}$, where $W \gg (\Delta f)_c$, we achieve a resolution of $1/W$ in the multipath delay profile. Since the total multipath delay spread is T_m, for all practical purposes the tapped delay line model for the channel can be truncated at $L = \lfloor T_m W \rfloor + 1 = \lceil T_m W \rceil$ taps. Then the noiseless received signal can be expressed in the form

$$\tilde{r}(t) = \sum_{n=0}^{L} c_n(t)\tilde{s}\left(t - \frac{n}{W}\right) \qquad \text{(III.91)}$$

$$\tilde{r}(k/W) = \sum_{n=-\infty}^{\infty} c_n(k/W)\tilde{s}\left(k/W - n/W\right) \qquad \text{(III.92)}$$

$$\tilde{r}(k) = \sum_{n=-\infty}^{\infty} c_n(k)\tilde{s}\left(k - n\right) \qquad \text{(III.93)}$$

$\rightarrow$

Delay Line with Tap Spacing = $1/W$

$$\tilde{r}(k) = \sum_{n=0}^{L} c_n(k)\tilde{s}\left(k-n\right) \tag{III.94}$$

$$\tilde{r}(t) = \sum_{n=0}^{L} c_n(t)\tilde{s}\left(t-n/W\right) + z(t) \tag{III.95}$$

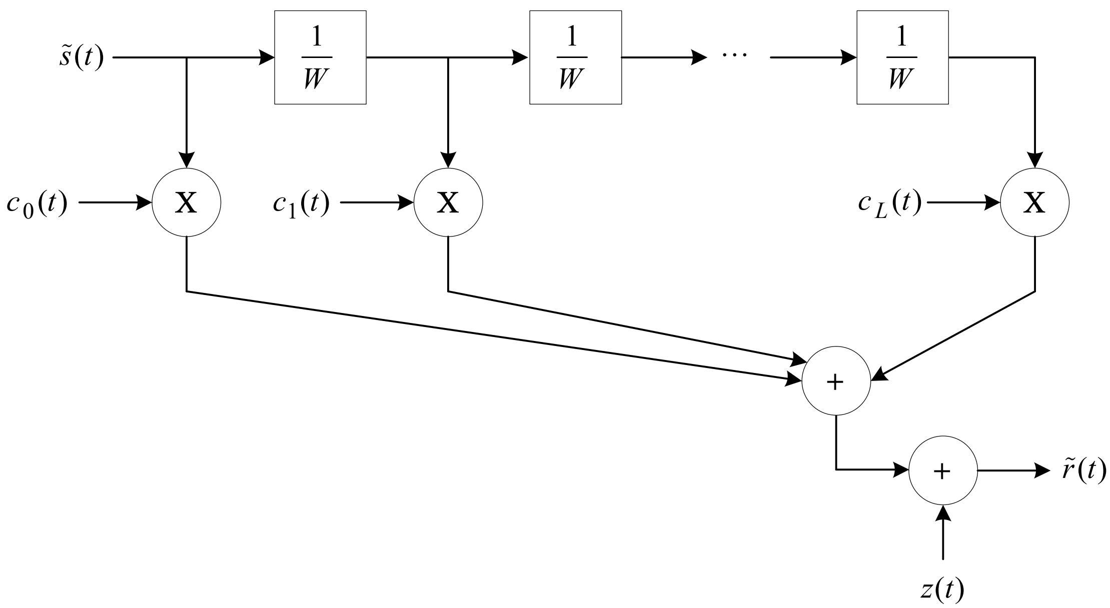


Figure III.9: Tapped delay line frequency selective channel model

In accordance with the statistical characterization of the channel presented earlier, the time-variant tap weights $\{c_n(t)\}$ are complex-valued stationary random processes. In the special case of Rayleigh fading, the magnitudes $|c_n(t)| = \alpha_n(t)$ are Rayleigh-distributed and the phases $\phi_n(t)$ are uniformly distributed. Since the $\{c_n(t)\}$ represent the tap weights corresponding to the L different delays $\tau_n = n/W$, $n = 0,1,\cdots,L$, the uncorrelated scattering assumption implies that the $\{c_n(t)\}$ are mutually uncorrelated. When the $\{c_n(t)\}$ are Gaussian random processes, they are statistically independent.

III.8.B. Rake Receiver

It is apparent that the tapped delay line model with statistically independent tap weights provides us with L replicas of the same transmitted signal at the receiver. Hence, a receiver that processes the received signal in an optimum manner will achieve the performance of an equivalent L^{th}-order diversity communication system.

Let us consider binary signaling over the channel. We have two equal-energy signals $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$, which are either antipodal or orthogonal. Their time duration is selected to satisfy the condition $T \gg T_m$. Thus, we may neglect any intersymbol interference due to multipath. Since the bandwidth of the signal exceeds the coherenc bandwidth of the channel, the received signal is expressed as

$$\begin{aligned}\tilde{r}(t) &= \sum_{l=1}^{L} c_l(t)\tilde{s}_i\left(t-\frac{l}{W}\right)+z(t), \quad 0 \le t \le T, \quad i=1,2 \\ &= v_i(t)+z(t)\end{aligned} \tag{III.96}$$

where $z(t)$ is a complex-valued zero-mean white Gaussian noise process. Assume for the moment that the channel tap weights are known. Then the optimum demodulator consists of two filters matched to $v_1(t)$ and $v_2(t)$. The demodulator output is sampled at the symbol rate and the samples are passed to a decision circuit that selects the signal corresponding to the largest output. An equivalent optimum demodulator employs cross correlation instead of matched filtering. In either case, the decision variables for coherent detection of the binary signals can be expressed as

$$\begin{aligned}U_m &= \mathrm{Re}\left\{\int_0^T \tilde{r}(t)v_m^*(t)dt\right\} \\ &= \mathrm{Re}\left\{\sum_{l=1}^{L}\int_0^T \tilde{r}(t)c_l^*(t)\tilde{s}_m^*\left(t-\frac{l}{W}\right)dt\right\}, \quad m=1,2 \\ &= \mathrm{Re}\left\{\sum_{l=1}^{L}\int_0^T c_l^*(t)\tilde{s}_m^*\left(t-\frac{l}{W}\right)\left[v_m(t)+z(t)\right]dt\right\}, \quad m=1,2\end{aligned} \tag{III.97}$$

Figure III.10 below illustrates the operations involved in the computation of the decision variables. In this realization of the optimum receiver, the two reference signals are delayed and correlated with the received signal $\tilde{r}(t)$.

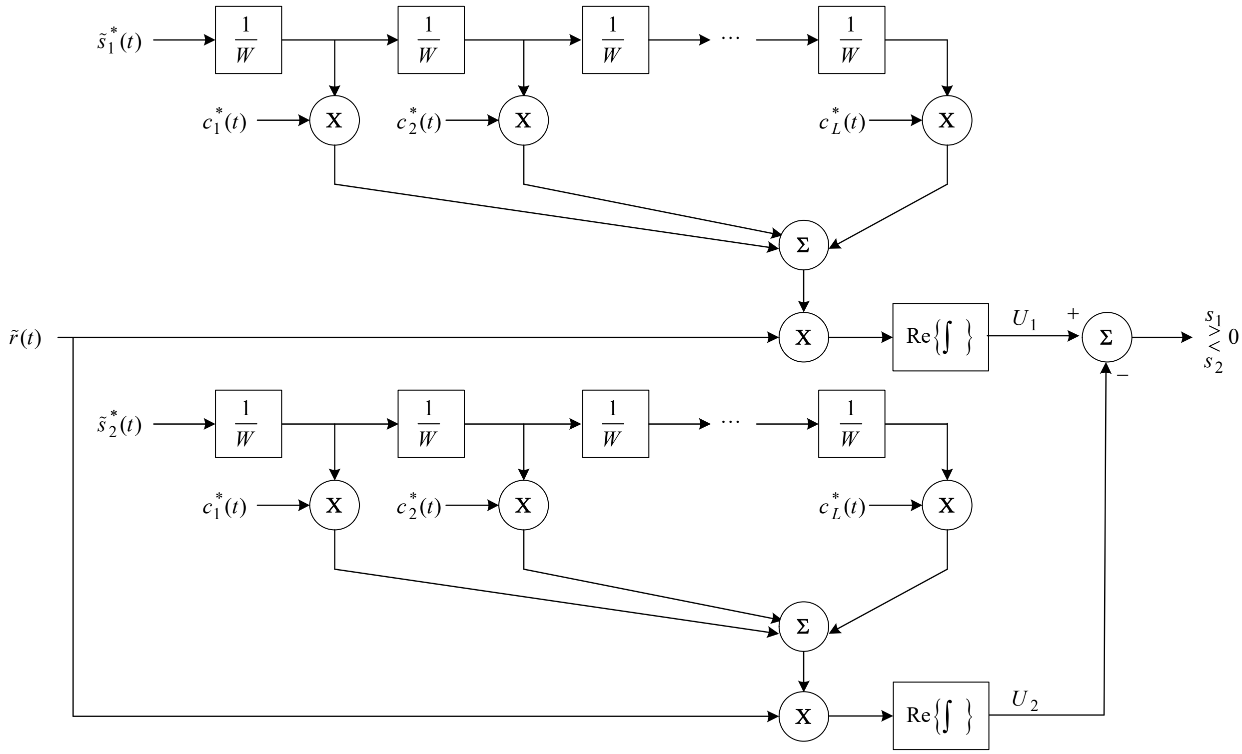


Figure III.10: Rake receiver

Assignment III.2

Sketch a single delay line rake receiver that is equivalent to the one in Figure III.10.

In effect, the tapped delay line demodulator attempts to collect the signal energy from all the received signal paths that fall within the span of the delay line and carry the same information. Its action is somewhat analogous to an ordinary garden rake and, consequently, the name "Rake demodulator" has been coined for this demodulator structure. The taps on the Rake demodulator are often called "Rake fingers."

III.8.C. Performance of Rake Demodulator

We shall now evaluate the performance of the rake demodulator under the condition that the fading is sufficiently slow to allow us to estimate $c_k(t)$ perfectly (without noise). Furthermore, within any one signaling interval, $c_k(t)$ is treated as a constant and denoted as c_k.

The decision variables in (III.97) may be expressed in the form

$$U_m = \mathrm{Re}\left\{ \sum_{l=1}^{L} c_l^* \int_0^T \tilde{r}(t)\tilde{s}_m^*\left(t - \frac{l}{W}\right) dt \right\}, \quad m = 1, 2 \tag{III.98}$$

Without loss of generality, suppose the transmitted signal is $\tilde{s}_1(t)$; then the received signal is

$$\tilde{r}(t)=\sum_{l=1}^{L} c_l \tilde{s}_1\left(t-\frac{l}{W}\right)+z(t), \quad 0 \le t \le T \tag{III.99}$$

Substitution of (III.99) into (III.98) yields

$$\begin{aligned} U_{1,2} &= \operatorname{Re}\left\{\sum_{l=1}^{L} c_l^* \sum_{n=1}^{L} c_n \int_0^T \tilde{s}_1\left(t-\frac{n}{W}\right)\tilde{s}_{1,2}^*\left(t-\frac{l}{W}\right)dt\right\} \\ &+\operatorname{Re}\left\{\sum_{l=1}^{L} c_l^* \int_0^T z(t)\tilde{s}_{1,2}^*\left(t-\frac{l}{W}\right)dt\right\} \end{aligned} \tag{III.100}$$

The wideband signals $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ can be assumed to have the property

$$\int_0^T \tilde{s}_i\left(t-\frac{n}{W}\right)\tilde{s}_i^*\left(t-\frac{k}{W}\right)dt \approx 0, \quad k \ne n, i=1,2 \tag{III.101}$$

Then (III.100) simplifies to

$$\begin{aligned} U_{1,2} &= \operatorname{Re}\left\{\sum_{l=1}^{L} |c_l|^2 \int_0^T \tilde{s}_1\left(t-\frac{l}{W}\right)\tilde{s}_{1,2}^*\left(t-\frac{l}{W}\right)dt\right\} \\ &+\operatorname{Re}\left\{\sum_{l=1}^{L} c_l^* \int_0^T z(t)\tilde{s}_{1,2}^*\left(t-\frac{l}{W}\right)dt\right\} \end{aligned} \tag{III.102}$$

When the binary signals are antipodal, a single decision variable suffices. In this case, (III.102) reduces to

$$U_1=\operatorname{Re}\left\{2\mathcal{E}\sum_{l=1}^{L}\alpha_l^2+\sum_{l=1}^{L}\alpha_l N_l\right\} \tag{III.103}$$

where $\alpha_l=|c_l|$ and

$$N_l=e^{-j\phi_l}\int_0^T z(t)\tilde{s}_1^*\left(t-\frac{l}{W}\right)dt \tag{III.104}$$

As shown by (III.103), the Rake demodulator with perfect (noiseless) estimates of the channel tap weights is equivalent to a maximal ratio combiner in a system with L^{th}-order diversity. Thus, when all the tap weights have the same mean-square value, i.e., $\mathrm{E}\left[\alpha_l^2\right]$ is the same for all l, the error rate performance of the Rake demodulator is given by (III.76) for the case of PSK. On the other hand, when the mean-square values $\mathrm{E}\left[\alpha_l^2\right]$ are not identical for all l, the derivation of the error rate performance must be repeated since (III.76) no longer applies.

We shall derive the probability of error for binary antipodal and orthogonal signals under the condition that the mean-square values of $\{\alpha_l\}$ are distinct. We begin with the conditional error probability

$$P(e/\gamma_b) = Q\left(\sqrt{(1-\rho)\gamma_b}\right) \tag{III.105}$$

where $\rho = -1$ for antipodal signals, $\rho = 0$ for orthogonal signals, and

$$\begin{aligned} \gamma_b &= \sum_{l=1}^{L} \gamma_l \\ \gamma_l &= \frac{E_b}{N_0}\alpha_l^2 \end{aligned} \tag{III.106}$$

$$f_{\gamma_b}(\gamma_b) = \sum_{l=1}^{L} \frac{\pi_l}{\overline{\gamma}_l} e^{-\gamma_b/\overline{\gamma}_l}, \quad \gamma_b > 0 \tag{III.107}$$

$$\begin{aligned} \pi_l &= \sum_{\substack{i=1 \\ i\neq l}}^{L} \frac{\overline{\gamma}_l}{\overline{\gamma}_l - \overline{\gamma}_i} \\ \overline{\gamma}_l &= E\left[\gamma_l\right] \end{aligned} \tag{III.108}$$

$$P(e) = \frac{1}{2}\sum_{l=1}^{L} \pi_l \left(1 - \sqrt{\frac{(1-\rho)\overline{\gamma}_l}{2+(1-\rho)\overline{\gamma}_l}}\right) \tag{III.109}$$

$$\begin{aligned} &\overline{\gamma}_l >> 1 \rightarrow \\ &P(e) \approx \binom{2l-1}{L} \prod_{l=1}^{L} \frac{1}{2(1-\rho)\overline{\gamma}_l} \end{aligned} \tag{III.110}$$

Experimental measurements indicate that the mobile radio channel is well characterized by an exponentially decaying PDP for indoor office buildings and congested urban areas:

$$\Omega_l = \mathrm{E}\left[\alpha_l^2\right] \tag{III.111}$$

$$\Omega_l = \Omega_1 e^{-\tau_l/\tau_{\max}} \tag{III.112}$$

Delays are often assumed to be equally spaced ($\tau_{l+1}-\tau_l$ is constant and equal to the symbol time T) and, with this assumption, we get the equally spaced exponential profile given by

$$\Omega_l = \Omega_1 e^{-(l-1)\delta} \tag{III.113}$$

where the parameter δ is the power decay factor that reflects the rate at which the average fading power decays.

***

IV. Spread Spectrum Communication Systems

Spread spectrum signals used for the transmission of digital information are distinguished by the characteristic that their bandwidth W is much greater than the information rate R in bits/s. That is, the bandwidth expansion factor $B_e = W/R$ for a spread spectrum signal is much greater than unity. The large redundancy inherent in spread spectrum signals is required to overcome the severe levels of interference that are encountered in the transmission of digital information over some channels. Since coded waveforms are also characterized by a bandwidth expansion factor greater than unity and since coding is an efficient method for introducing redundancy, it follows that coding is an important element in the design of spread spectrum signals and systems.

A second important element employed in the design of spread spectrum signals is pseudo-randomness, which makes the signals appear similar to random noise and difficult to demodulate by receivers other than the intended ones.

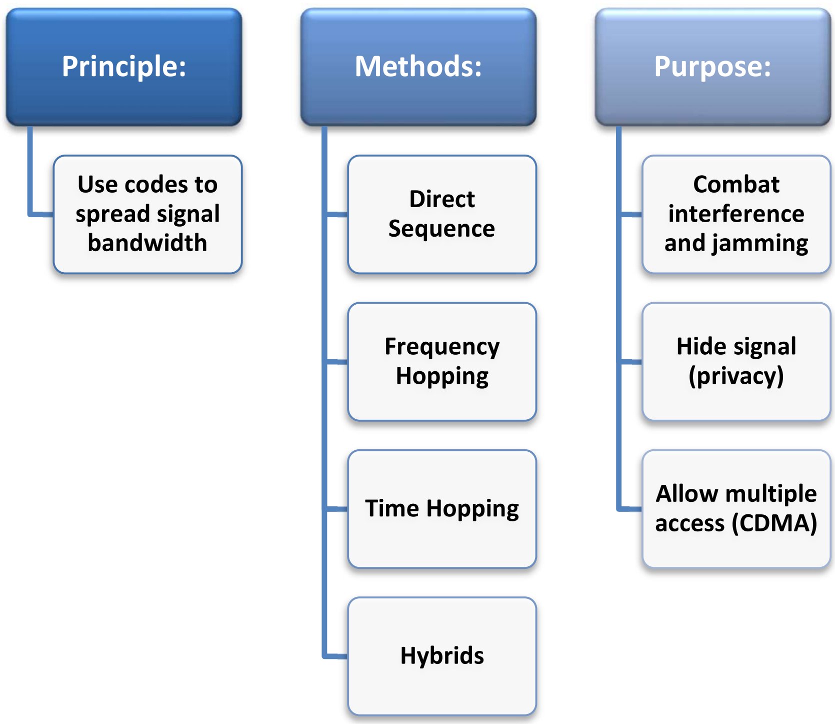
Spread spectrum signals are used for

- Combating or suppressing the detrimental effects of interference due to jamming, interference arising from other users of the channel, and self-interference due to multipath propagation.
- Hiding a signal by transmitting it at low power spectral density and, thus, making it difficult for an unintended listener to detect in the presence of background noise.
- Achieving message privacy in the presence of other listeners.
- In combating intentional interference (jamming), it is important to the communicators that the jammer who is trying to disrupt the communication does not have prior knowledge of the signal characteristics except for the overall channel bandwidth and the type of modulation (PSK, FSK, etc.) being used.
- The transmitter introduces an element of unpredictability or randomness (pseudo-randomness) in each of the transmitted coded signal waveforms that is known to the intended receiver but not to the jammer. As a consequence, the jammer must synthesize and transmit an interfering signal without knowledge of the pseudorandom pattern.
- Interference from the other users arises in multiple-access communication systems in which a number of users share a common channel bandwidth. At any given time, a subset of these users may transmit information simultaneously over the common channel to corresponding receivers. Assuming that all the users employ the same code for the encoding and decoding of their respective information sequences, the transmitted signals in this common spectrum may be distinguished from one another by superimposing a different pseudo-random pattern, also called a code, in each transmitted signal. Thus, a particular receiver can recover the transmitted information intended for it by knowing the pseudo-random pattern, i.e., the key, used by the corresponding transmitter. This type of communication technique, which allows multiple users to simultaneously use a common channel for transmission of information, is called code division multiple access (CDMA).
- Resolvable multipath components resulting from time-dispersive propagation through a channel may be viewed as a form of self-interference. This type of interference may be

suppressed by the introduction of a pseudo-random pattern in the transmitted signal, as will be described later.

- A message may be hidden in the background noise by spreading its bandwidth with coding and transmitting the resultant signal at a low average power spectral density. Because of its low PSD level, the transmitted signal is said to be “covert.” It has a low probability of being intercepted (detected) by a casual listener and, hence, is also called a low-probability-of-intercept (LPI) signal.
- Finally, message privacy may be obtained by superimposing a pseudo-random pattern on a transmitted message. The message can be demodulated by the intended receivers, who know the pseudo-random pattern or key used at the transmitter, but not by any other receivers who do not have knowledge of the key.

IV.1. Preliminaries



IV.1.A. Spreading Codes

IV.1-Preliminaries

IV.1.B. Maximal-Length PN Codes (Linear Feedback Shift Register (LFSR) Codes)

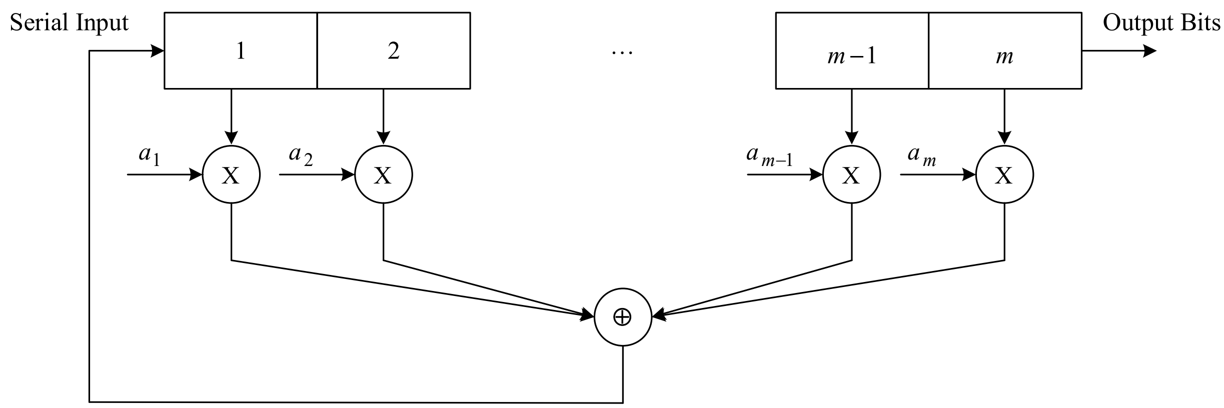


Figure IV.1: Linear feedback shift register (LFSR)

Maximal-length shift register codes are a class of cyclic codes equivalent to the maximum-length codes known as duals of Hamming codes. These are a class of cyclic codes with

$$(n,k)=\left(2^m-1,m\right) \qquad \text{(IV.1)}$$

where m is a positive integer. Note that, with the exception of the all-zero codeword, all the codewords generated by the shift register are different cyclic shifts of a single codeword. The output sequence is periodic with length $n=2^m-1$.

IV.1.C. PN Code Examples

Table IV.1: Maximal-length code for $m=3$.

Information Bits			Codewords						
0	0	0	0	0	0	0	0	0	0

0	0	1	0	0	1	1	1	0	1
1	0	0							
1	1	0							
1	1	1							
0	1	1							
1	0	1							
0	1	0							
0	1	0	0	1	0	0	1	1	1
0	1	1	0	1	1	1	0	1	0
1	0	0	1	0	0	1	1	1	0
1	0	1	1	0	1	0	0	1	1
1	1	0	1	1	0	1	0	0	1
1	1	1	1	1	1	0	1	0	0

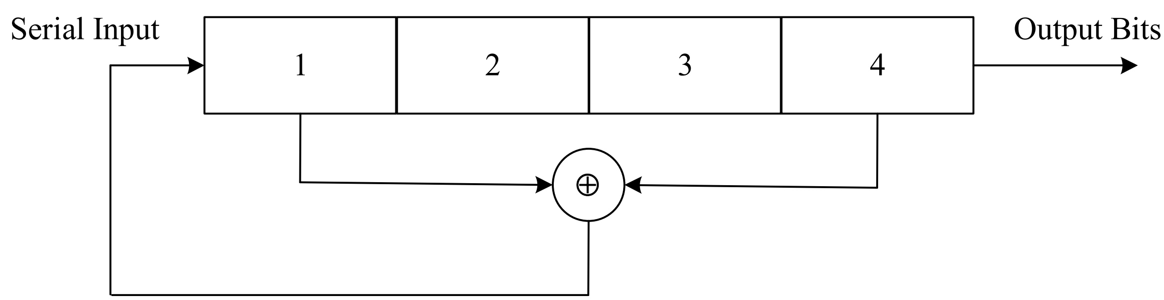


Figure IV.2: (1,4) connection for $m = 3$

Code Period = 15

Number of “1”s = 8

Number of “0”s = 7

Maximum Correlation= 15

Minimum Correlation = -1

Maximum Number of Users = 15

IV.1.D. LFSR Tap Connections

Table IV.2: LFSR tap connections

m	Tap Connections
2	[2,1]
3	[3,1]
4	[4,1] [4,3]
5	[5,2] [5,3] / [5,4,3,2] [5,3,2,1] / [5,4,2,1] [5,4,3,1]
6	[6,1] / [6,5,2,1] / [6,5,3,2]
7	[7,1] / [7,3] / [7,3,2,1] / [7,4,3,2] / [7,6,4,2] / [7,6,3,1] / [7,6,5,2] / [7,6,5,4,2,1] / [7,5,4,3,2,1]
8	[8,4,3,2] / [8,6,5,3] / [8,6,5,2] / [8,5,3,1] / [8,6,5,1] / [8,7,6,1] / [8,7,6,5,2,1] / [8,6,4,3,2,1]

Note that in any set of connections in Table IV.2 in the form $[i_1 = m, i_2, i_3, \cdots]$, the coefficients $a_{i_1}, a_{i_2}, a_{i_3}, \cdots$ in Figure IV.1 are one, while all other coefficients are zero. For example when $m = 5$ and the connection [5,2], we have $a_5 = a_2 = 1$ and $a_4 = a_3 = a_1 = 0$. Note also that if $[m, i_1, i_2, \cdots]$ is a maximal length connection, then $[m, m - i_1, m - i_2, \cdots]$ is a maximal length connection as well.

IV.2. Direct Sequence Spread Spectrum

Direct Sequence is a spread spectrum technique that enables multiple access. The transmitted data sequence is spread across the spectrum after being encoded by spreading codes. Each spreading code is assigned uniquely to one of the users at a higher rate than the symbol rate of the information data.

Capacity of a DS-CDMA system is limited by multiple access interference (MAI).

IV.2.A. DS Signal Generation – Method 1

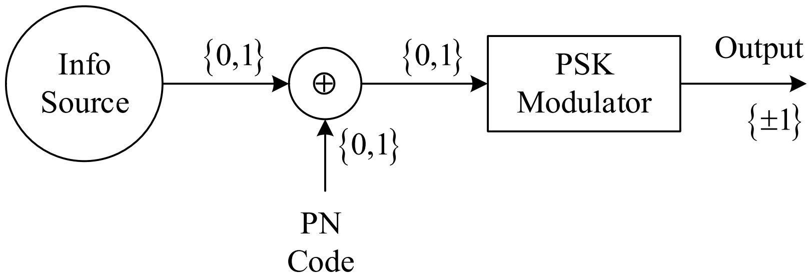


Figure IV.3: DS signal generation using modulo-2 addition

IV.2.B. DS Signal Generation – Method 2

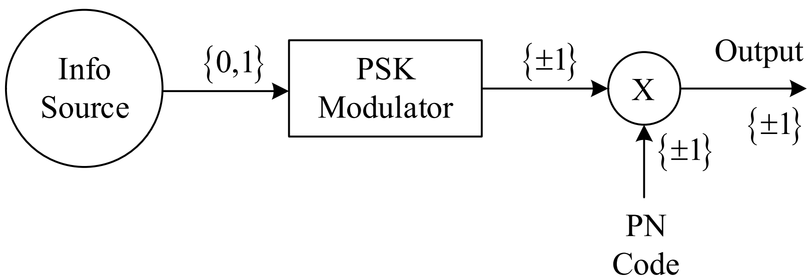


Figure IV.4: DS signal generation using product modulation

IV.3. Frequency Hopping Spread Spectrum

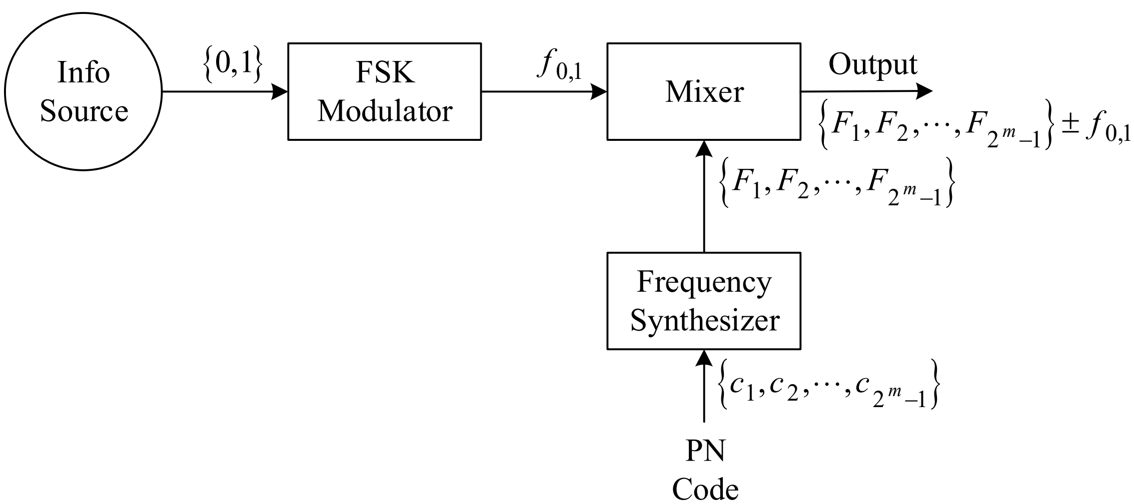


Figure IV.5: FH signal generation

IV.4. Model of Spread Spectrum Digital Communication System

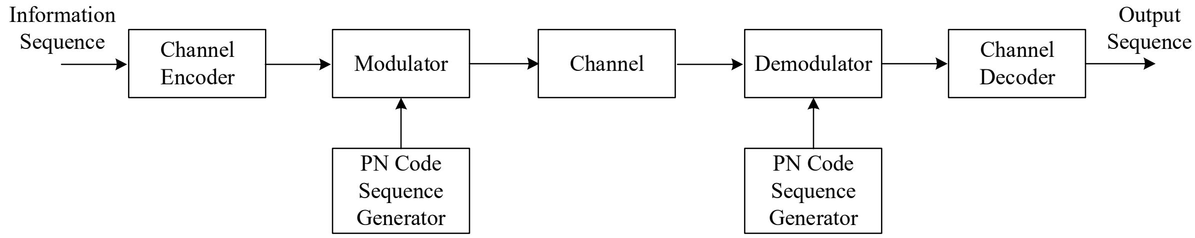


Figure IV.6: Spread spectrum system

(10111001➔ +-+++--+)(10111001+-+++--+) ➔ 7

We have two identical pseudorandom code sequence generators, one that interfaces with the modulator at the transmitting end and a second that interfaces with the demodulator at the receiving end. The generators generate a pseudorandom or pseudonoise (PN) binary-valued sequence which is impressed on the transmitted signal at the modulator and removed from the received signal at the demodulator.

Synchronization of the PN sequence generated at the receiver with the PN sequence contained in the incoming received signal is required in order to demodulate the received signal. Initially, prior to the transmission of information, synchronization may be achieved by transmitting a fixed pseudorandom bit pattern that the receiver will recognize in the presence of interference with a high probability. After time synchronization of the generators is established, the transmission of information may commence.

Interference is introduced in the transmission of the information-bearing signal through the channel. The characteristics of the interference depend to a large extent on its origin. It may be categorized as being either broadband or narrowband relative to the bandwidth of the information-bearing signal and as either continuous or pulsed (discontinuous) in time. For example, an interfering signal may consist of one or more sinusoids in the bandwidth used to transmit the information. The frequencies of the sinusoids may remain fixed or they may change with time according to some rule. As a second example, the interference generated in CDMA by other users of the channel may be either broadband or narrowband, depending on the type of spread spectrum signal that is employed to achieve multiple access. If it is broadband, it may be characterized as an equivalent additive white Gaussian noise.

The PN sequence generated at the modulator is used in conjunction with the PSK modulation to shift the phase of the PSK signal pseudorandomly as described earlier. The resulting modulated signal is called a direct sequence (DS) or pseudo-noise (PN) spread spectrum signal.

When used in conjunction with binary or M -ary ($M > 2$) FSK, the pseudorandom sequence selects the frequency of the transmitted signal pseudorandomly. The resulting signal is called a frequency-hopped (FH) spread spectrum signal. Although a number of other types of spread spectrum signals exist, the emphasis of our treatment will be on DS and FH spread spectrum signals.

IV.5. Direct Sequence Spread Spectrum Signals

We assume that the information rate at the input to the encoder is R bits/s and the available spread spectrum bandwidth is W Hz. The modulation is assumed to be binary PSK. In order to utilize the entire available channel bandwidth, the phase of the carrier is shifted pseudorandomly according to the pattern from the PN generator at a rate W times/s. The reciprocal of W, denoted by T_c, defines the duration of a pulse, which is called a chip; T_c is called the chip interval. The pulse is the basic element in a DS spread spectrum signal.

If we define $T_b = 1/R$ to be the duration of a rectangular pulse corresponding to the transmission time of an information bit, the bandwidth expansion factor W/R may be expressed as

$$\begin{aligned} B_e &= \frac{W}{R} \\ &= \frac{T_b}{T_c} \end{aligned} \tag{IV.2}$$

In practical systems, the ratio T_b/T_c is an integer,

$$L_c = \frac{T_b}{T_c} \tag{IV.3}$$

which is the number of chips per information bit.

Suppose that the encoder takes k information bits at a time and generates a binary linear (n,k) block code. The time duration available for transmitting the n code elements is kT_b seconds. The number of chips that occur in this time interval is kL_c. Hence, we may select the block length of the code as $n = kL_c$. If the encoder generates a binary convolutional code of rate k/n, the number of chips in the time interval kT_b is also $n = kL_c$. Therefore, the following discussion applies to both block codes and convolutional codes. We note that the code rate is

$$\begin{aligned} R_c &= \frac{k}{n} \\ &= \frac{1}{L_c} \end{aligned} \tag{IV.4}$$

One method for impressing the PN sequence on the transmitted signal is to alter directly the coded bits by modulo-2 addition with the PN sequence.

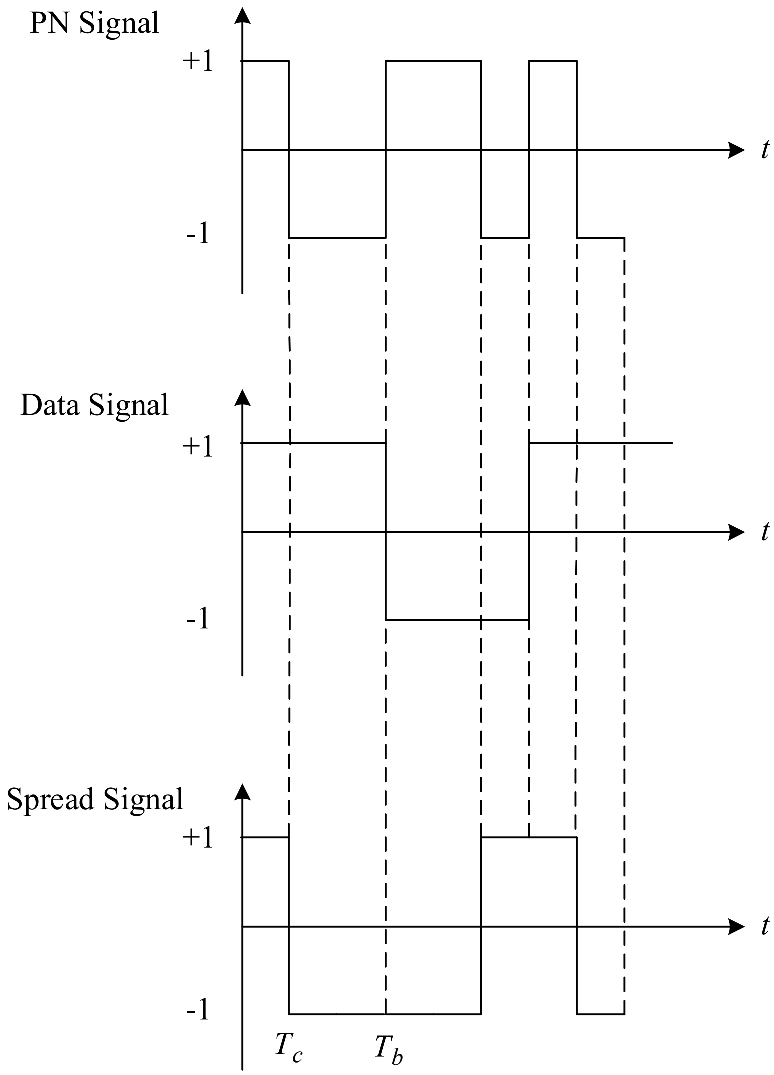


Figure IV.7: Direct sequence spread spectrum waveform

If b_i represents the i th bit of the PN sequence and c_i is the corresponding bit from the encoder, the modulo-2 sum is

$$a_i = b_i \oplus c_i \tag{IV.5}$$

The sequence $\{a_i\}$ is mapped into a binary PSK signal of the form $s(t) = \pm \mathrm{Re}\left\{g(t)e^{j2\pi f_c t}\right\}$ according to the convention

$$g_i(t) = \begin{cases} g(t - iT_c), & a_i = 0 \\ -g(t - iT_c), & a_i = 1 \end{cases} \tag{IV.6}$$

where $g(t)$ represents a pulse of duration T_c seconds and arbitrary shape.

The modulo-2 addition of the coded sequence $\{c_i\}$ and the sequence $\{b_i\}$ from the PN generator may also be represented as a multiplication of two waveforms (as in Figure IV.7). Suppose that the elements of the coded sequence are mapped into a binary PSK signal according to the relation

$$c_i(t) = (2c_i - 1)g(t - iT_c) \tag{IV.7}$$

Similarly, let

$$p_i(t) = (2b_i - 1)p(t - iT_c) \tag{IV.8}$$

where $p(t)$ is a rectangular pulse of duration T_c. Then the equivalent low-pass transmitted signal corresponding to the i^{th} coded bit is

$$\begin{aligned} g_i(t) &= p_i(t)c_i(t) \\ &= (2b_i - 1)(2c_i - 1)g(t - iT_c) \end{aligned} \tag{IV.9}$$

Although it is easier to implement modulo-2 addition followed by PSK modulation instead of waveform multiplication, it is convenient, for purposes of demodulation, to consider the transmitted signal in the multiplicative form given by (IV.9).

The received equivalent low-pass signal for the i th code element is

$$\begin{aligned} r_i(t) &= p_i(t)c_i(t) + z(t) \qquad\qquad iT_c \le t < (i+1)T_c \\ &= (2b_i - 1)(2c_i - 1)g(t - iT_c) + z(t) \end{aligned} \tag{IV.10}$$

IV.5.A. Error Rate Performance of the Decoder

Let the unquantized output of the demodulator be denoted by y_j, $1 \le j \le n$. First we consider a linear binary (n,k) block code. Note that y_i is given by

$$\begin{aligned} y_i &= \mathrm{Re}\left\{(2b_i - 1)\int_{iT_c}^{(i+1)T_c} g^*(t - iT_c)r_i(t)dt\right\} \\ &= \mathrm{Re}\left\{(2c_i - 1)\int_{iT_c}^{(i+1)T_c} g^*(t - iT_c)g(t - iT_c)dt + (2b_i - 1)\int_{iT_c}^{(i+1)T_c} g^*(t - iT_c)z(t)dt\right\} \\ &= \mathrm{Re}\left\{(2c_i - 1)\int_0^{T_c} g^*(t)g(t)dt + (2b_i - 1)\int_0^{T_c} g^*(t)z(t + iT_c)dt\right\} \end{aligned} \tag{IV.11}$$

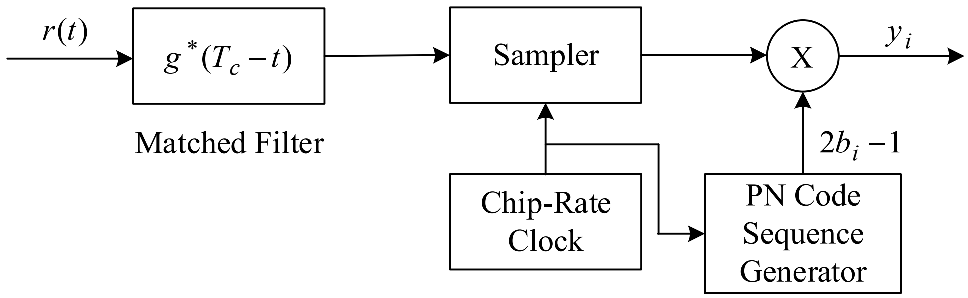


Figure IV.8: Matched filter DS demodulator

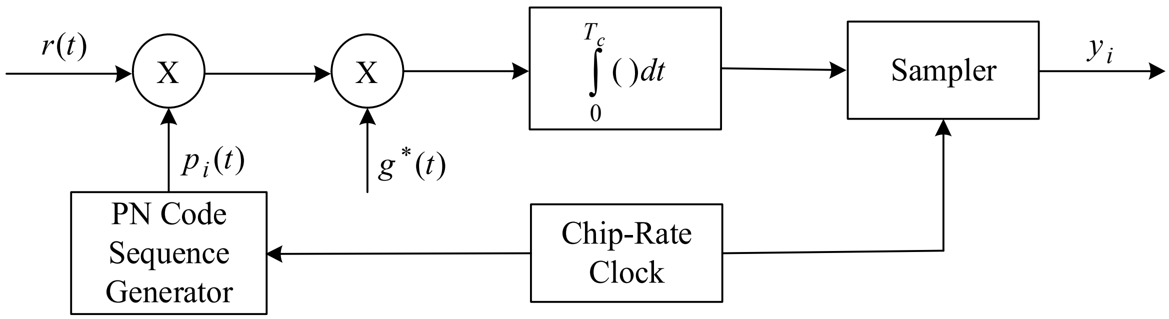


Figure IV.9: Despreading correlator DS demodulator

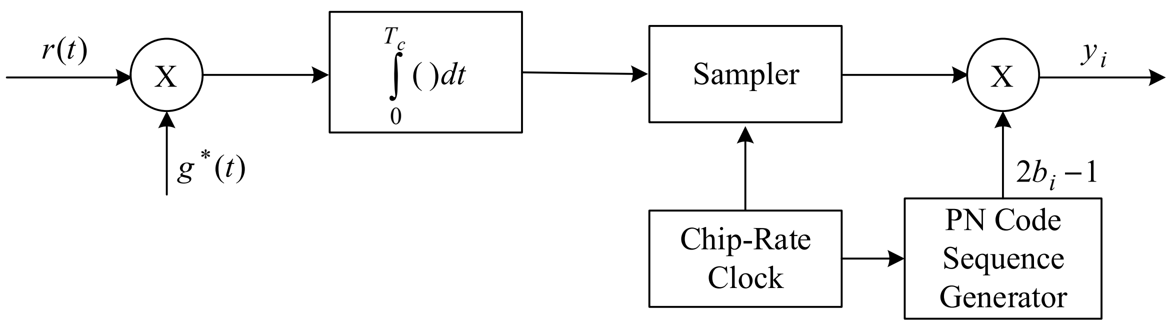


Figure IV.10: Correlator despreading DS demodulator

Without loss of generality, we assume that the all-zero code word is transmitted. In this case, $2c_i - 1 = -1$, and

$$\begin{aligned} y_i &= -\int_0^{T_c} g^*(t)g(t)dt + (2b_i - 1)\operatorname{Re}\left\{\int_0^{T_c} g^*(t)z(t + iT_c)dt\right\} \\ &= -2E_c + (2b_i - 1)v_i \end{aligned} \quad \text{(IV.12)}$$

A decoder that employs soft-decision decoding computes the correlation metrics

$$CM_i = \sum_{j=1}^{n} (2c_{ij} - 1)y_j, \quad i = 1, 2, \cdots, 2^k \quad \text{(IV.13)}$$

The correlation metric corresponding to the all-zero code word is

$$\begin{aligned} CM_1 &= 2n\mathcal{E}_c + \sum_{j=1}^{n} (2c_{1j} - 1)(2b_j - 1)v_j \\ &= 2n\mathcal{E}_c + \sum_{j=1}^{n} (1 - 2b_j)v_j \end{aligned} \tag{IV.14}$$

where v_j, $1 \le j \le n$, is the additive noise and interference term corrupting the j^{th} coded bit and $\mathcal{E}_c$ is the chip energy. v_j is defined as

$$v_j = \mathrm{Re}\left\{ \int_0^{T_c} g^*(t) z\left[t + (j-1)T_c\right] dt \right\} \tag{IV.15}$$

Similarly, the correlation metric corresponding to code word $\underline{c}_m$ having weight w_m is

$$CM_m = 2\mathcal{E}_c \left(1 - \frac{2w_m}{n}\right) + \sum_{j=1}^{n} (2c_{mj} - 1)(2b_j - 1)v_j \tag{IV.16}$$

The difference between CM_1 and CM_m is

$$\begin{aligned} D &= CM_1 - CM_m \\ &= 4\mathcal{E}_c w_m - 2\sum_{j=1}^{n} c_{mj} \left(2b_j - 1\right) v_j \end{aligned} \tag{IV.17}$$

Since the codeword $\underline{c}_m$ has weight w_m, there are w_m nonzero components in the summation of noise terms above. We assume that the minimum distance of the code is sufficiently large that we can invoke the central limit theorem for the summation of noise components. This assumption is valid for DS spread spectrum signals that have a bandwidth expansion of 10 or more. Typically, the bandwidth expansion factor in a spread spectrum signal is of the order of 10 to 100 and sometimes higher. Thus, the summation of noise components is modeled as a Gaussian random variable. Since $\mathrm{E}\left[2b_j - 1\right] = 0$ and $\mathrm{E}\left[v_j\right] = 0$, the mean of the second term in the equation above is zero. The variance is

$$\sigma_m^2 = 4\sum_{j=1}^{n}\sum_{i=1}^{n} c_{mi} c_{mj}\, \mathrm{E}\left[\left(2b_i - 1\right)\left(2b_j - 1\right)\right] \mathrm{E}\left[v_i v_j\right] \tag{IV.18}$$

The sequence of binary digits from the PN generator are assumed to be uncorrelated, hence

$$\mathrm{E}\left[\left(2b_i - 1\right)\left(2b_j - 1\right)\right] = \delta_{ij} \tag{IV.19}$$

$$\sigma_m^2 = 4w_m \mathrm{E}\left[v^2 \right] \tag{IV.20}$$