Modelling of Mobile Fading Channels with Fading Mitigation Techniques

Lei Shang (B.Eng., M.Eng.)

School of Electrical and Computer Engineering
Science, Engineering and Technology (SET) Portfolio
RMIT University, Melbourne, Australia

September 2006

Declaration

The work contained in this thesis has not been previously submitted for a degree or diploma at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signed: ..... Date: .....

Acknowledgements

I am indebted to my supervisor Professor Zahir M. Hussain for his invaluable advice during the course of my PhD study. I am lucky to be the tutor of Professor Zahir M. Hussain's two Digital signal Processing (DSP) courses for the past four years, which was an opportunity to study and enjoy the various inspirational DSP topics.

I would like to thank Dr. Peter O'Shea and Professor Richard Harris for leading me during my Master's studies at RMIT University into the exciting fields of Digital Signal Processing and digital communications. I am very grateful to Professor Richard Harris, who had been providing a friendly research environment at the Center for Advanced Technology in Telecommunications (CATT), RMIT. Thanks to all people at CATT for the shared advice. Special thanks to Dr. Horace King and Dr. Irena Atov for their advice and encouragement.

The thesis is dedicated to my wife, Christy Yahong Liu, for her support and patience during the past four years of my PhD studies.

Lei Shang

Melbourne
February 2006

Contents

Abstract

This thesis aims to contribute to the developments of wireless communication systems. The work generally consists of three parts: the first part is a discussion on general digital communication systems, the second part focuses on wireless channel modelling and fading mitigation techniques, and in the third part we discuss the possible application of advanced digital signal processing, especially time-frequency representation and blind source separation, to wireless communication systems.

The first part considers general digital communication systems which will be incorporated in later parts. Today's wireless communication system is a subbranch of a general digital communication system that employs various techniques of A/D (Analog to Digital) conversion, source coding, error correction, coding, modulation, and synchronization, signal detection in noise, channel estimation, and equalization. We study and develop the digital communication algorithms to enhance the performance of wireless communication systems.

In the Second Part we focus on wireless channel modelling and fading mitigation techniques. A modified Jakes' method is developed for Rayleigh fading channels. We investigate the level-crossing rate (LCR), the average duration of fades (ADF), the probability density function (PDF), the cumulative distribution function (CDF) and the autocorrelation functions (ACF) of this model. The simulated results are verified against the analytical Clarke's channel model. We also construct frequency-selective geometrical-based hyperbolically distributed scatterers (GBHDS) for a macro-cell mobile environment with the proper statistical characteristics. The modified Clarke's model and the GBHDS model may be readily expanded to a MIMO channel model thus we study the MIMO fading channel, specifically we model the MIMO channel in the angular domain. A detailed analysis of Gauss-Markov approximation of the fading channel is also given. Two fading mitigation techniques are investigated: Orthogonal Frequency Division Multiplexing (OFDM) and spatial diversity.

In the Third Part, we devote ourselves to the exciting fields of Time-Frequency Analysis and Blind Source Separation and investigate the application of these powerful Digital Signal Processing (DSP) tools to improve the performance of wireless communication systems.

Keywords

Multipath interference, channel modelling, fading, channel capacity, channel tracking, channel estimation, Doppler spread, coherence time, delay spread, detection, space-time geometrical based hyperbolically distributed scatterers (GBHDS), autocorrelation function (ACF), level-crossing rate (LCR), average duration of fades (ADF), diversity, error correction coding, OFDM (Orthogonal Frequency Division Multiplexing), MIMO (Multi-Input and Multi-Output), additive Gaussian noise, signal-to-noise-ratio (SNR), Sigma-Delta quantization, blind source separation, time-frequency distribution (TFD).

Acronyms

ADC	analog-to-digital converter
ADF	average duration of fades
AWGN	additive white Gaussian noise
BPSK	binary phase-shift-keying (BPSK)
BSS	blind source separation
CDMA	code division multiple access
CWD	Choi-Williams distribution
DFT/IDFT	discrete Fourier transform/inverse discrete Fourier transform
FIR	finite impulse response
GBHDS	Geometrical-Based Hyperbolically Distributed Scatterers
FT	Fourier transform
HT	Hilbert transformer
ICA	independent component analysis
IF	instantaneous frequency
i.i.d.	independent and identically distributed
ISI	inter-symbol interference
ISR	interference-to-signal ratio
LCR	level-crossing rate
LFM	linear frequency modulated signal
MIMO	multiple input multiple output
MMSE	minimum mean square error
OFDM	orthogonal frequency division multiplexing
SIMO	single input multiple output
SVD	singular value decomposition
LTV	linear time-varying
pdf	probability density function
SNR	signal-to-noise ratio
TFD	time-frequency distribution
WLAN	wireless LAN
WSS	wide-sense stationary
WVD	Winger-Ville distribution

Publications

Below are the publications in conjunction with the author during his PhD candidacy.

1. Lei Shang and Zahir M. Hussain, "A Frequency-Selective Mobile Radio Channel with Hyperbolically Distributed Scatterers", *IEEE TENCON*, November 21-24, Melbourne, Australia 2005.

2. Lei Shang and Zahir M. Hussain, "A Spectrum Analysis of Single-Bit Sigma-Delta Modulator Using Linear Frequency Modulation," *IEEE TENCON*, November 21-24, Melbourne, Australia 2005.

3. Lei Shang and Zahir M. Hussain, "Level-Crossing Rate and Average Duration of Fades for Mobile Radio Channel with Hyperbolically Distributed Scatterers," IEEE APCC 2005, Perth, Australia, 2005.

4. Lei Shang, Zahir M. Hussain, Richard Harris, "An Efficient Mobile Rayleigh Fading Channel Simulator: A Comparison with Clarke's Channel Model," *IEEE TENCON* 2004.

5. Lei Shang, Zahir M. Hussain, Richard Harris, "Performance of Quadratic Time-Frequency Distributions in Blind Source Separation of Speech Signals", IEEE International Symposium on Spread Spectrum Techniques and Applications, Sydney, Australia, August 2004.

6. Lei Shang, Zahir M. Hussain, Richard Harris, "Blind Source Separation of FM and Doppler signals Using T-distributions," WSEAS Transactions on Computers, vol. 3, no. 1, January 2004.

7. Lei Shang, Zahir M. Hussain, Richard Harris, "Performance of Quadratic Time-Frequency Distributions in Blind Source Separation," Australian Telecommunication Networks and Applications Conference, Melbourne, Australia, December 2003.

List of Figures

Chapter 1

Introduction

1.1 Motivation

The huge surge of wireless applications has inspired enormous amount of research activities in physical-layer wireless communication theory. The fundamental challenge in wireless communication is to combat *fading* and *interference* which otherwise may not exist in wired communication systems. In addition to the efforts of eliminating fading and interference, increasing of *spectral efficiency* is another goal to achieve.

The enormous challenges in wireless communications, such as complex and highly dynamic channels, multiple-access/ cochannel interference, practical constraints on portability, power consumption and complexity, have provided a rich research field for signal processing. Research in signal processing such as filter design, time-frequency representation, adaptive antenna array and higher order statistics, servers as a springboard to the development of novel methods in digital communications. In this thesis we aim to understand the various aspects of wireless communications in an effort to contribute to the research community.

1.2 Mobile Channel Modelling

The design of spectrally efficient wireless communication systems requires a detailed understanding of the radio propagation environment. The characteristics of the radio channel have to be captured accurately in a channel model to assess the relative performance of the various communication architectures. Efficiency is another concern for computer simulation of mobile communication systems. We investigate the classical mobile channel models and extend them to wideband channel models for future mobile communication systems which exploit the directional properties of the mobile radio channel by using antenna arrays.

In a system with multiple transmit and receive antennas (MIMO channel), the spectral efficiency can be greatly increased from that of the conventional single antenna channels, without increasing the total transmitted power. Various fading mitigation techniques are also investigated in an effort to validate the channel models.

1.3 Blind Source Separation and Time-Frequency Representation

We investigate the blind source separation techniques and time-frequency representation in the context of MIMO wireless communication as these topics share some common ground. In blind source separation (BSS), multiple observations acquired by an array of sensors are processed in order to recover the initial multiple source signals. The mixture is either convolutive or linear. In today's digital cellular system where transmitted signal is interfered not only by its echoes (multipath interference), but also by other users in the neighborhood (cochannel interference), we face the source separation problem similar to that in a cocktail party.

The communications community has recognized the importance of blind signal processing techniques, partly due to the fact that wireless communications field experienced explosive growth and demand for high data-rate services has been increasing. In conventional techniques, channel estimation and equalization rely on training signals. This approach lowers the throughput, for example: in GSM, about 20% of the symbols are used for training. See [84], equalization of linear FIR channels with training is commonly used in practice especially with wired line transmissions over telephone lines, cable television, and asymmetric digital subscriber loops (ADSL). Blind approaches show more promise for future wireless and mobile communications, high-frequency modems, digital audio and video broadcasting systems where rapid channel variations render the overhead for training prohibitive.

1.4 Scope of the Thesis

The wireless communication is a web of interlocking concepts. The concepts can be structured roughly into three levels according to [92]:

- Channel characteristics and modelling;
- Communication concepts and techniques;

- Application of these concepts in a system context.

In this thesis we aim to apply signal processing techniques to solve problems in $\Sigma\Delta$ modulation, channel modelling, blind source separation, and time frequency representation in the background of wireless communications.

1.5 Contributions of the Thesis

- A spectral analysis of $\Sigma\Delta$ modulation system and an investigation of validity of the white noise linear model
- A modified Smith method with easier implementation and greater mathematical tractability
- A verification of Geometrical-Based Hyperbolically Distributed Scatterers (GBHDS) model using statistical characteristics ACF,LCR, AFD.
- A frequency-selective GBHDS model is constructed
- A time-frequency viewpoint of diversity for spatial multiplexing/ blind source separation/channel estimation
- An information-theoretic analysis of spatial diversity to mitigate fading
- A numerical solution to ϵ-outage capacity with L-fold receive diversity
- A Gauss-Markov approximation to GBHDS model
- A MIMO fading channel is modelled with beamforming

1.6 Thesis Organization

- Chapter 2: a general discussion of the components of a digital communication system: source coding, channel coding, modulation, and detection in AWGN.
- Chapter 3: modelling of wireless communication channels, focusing on channel capacity, modelling, statistical properties.
- Chapter 4: fading mitigation techniques, mainly with OFDM and diversity techniques.
- Chapter 5: blind source separation and time-frequency representation.
- Chapter 6: Our conclusions and future directions.

Chapter 2

The Digital Communication System: An Overview

2.1 Components of the Digital Communication System

In his 1948 landmark paper, Shannon states that: "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point [82]". Figure 2.1 illustrates functional diagram and the basic elements of a digital communication system. The source encoder converts the messages produced by the source into a sequence of binary digits. These binary digits, with redundancy possibly removed by the source encoder, forms a finite set of bits as the message. This message is then passed to the channel encoder to add in some redundancy to fight the detrimental effects in the transmission of the signal through the channel. As most of the physical channels are capable of transmitting electrical signals (continuous waveforms), the binary digits from the output of the channel encoder have to be mapped into waveforms that are suitable for the physical transmission media. The communication channel will generally corrupt the signal waveforms in a random manner and therefore after the demodulation of the corrupted signal waveforms, even with the help of the channel decoder, there is still a possibility of errors when reproducing the message. The role of a communication system is therefore to optimally reproduce at the receiver the message sent at the transmitter. The components of the digital communication system are studied in this section to serve as a foothold for further chapters.

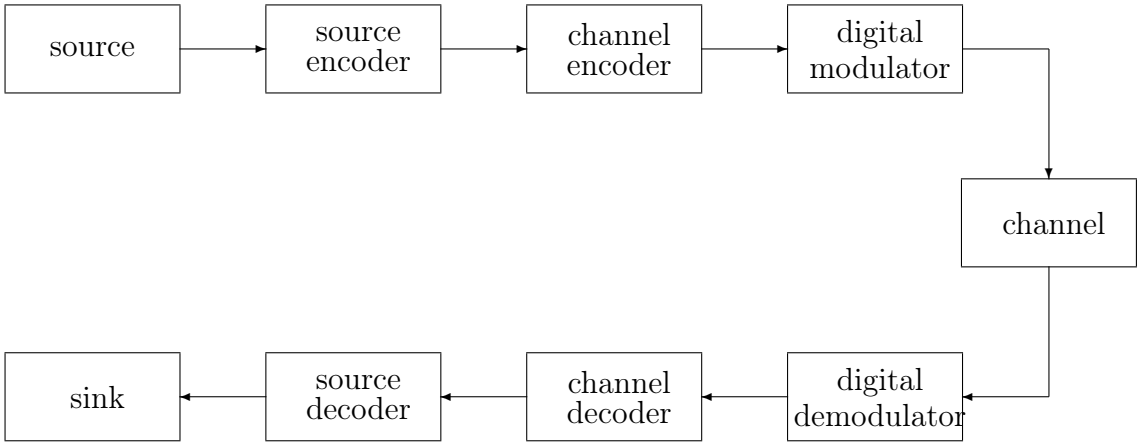


Figure 2.1: Basic elements of a digital communication system.

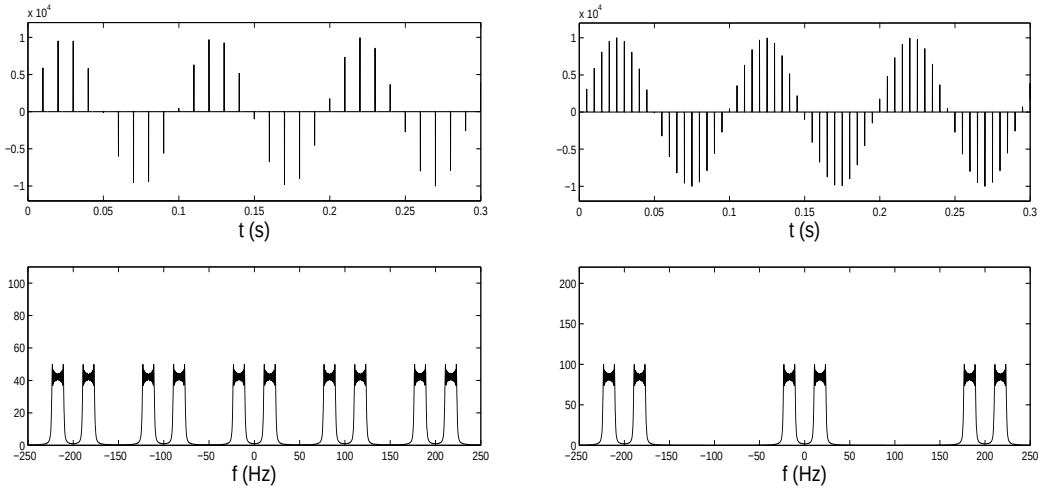
2.2 Quantization: Sigma-Delta modulation

A digital communication system is designed to transmit information in digital form. The output of the information source therefore must be converted to a format that can be transmitted digitally. As shown in Fig. 2.1, this conversion of the source output to a digital form is performed by source encoder. The discussion of mathematical models to describe information sources and coding for discrete sources can be found in [64]. Here we mainly focus on a coding technique for analog sources, the $\Sigma\Delta$ modulation.

2.2.1 Introduction

Assume the signal source in Fig. 2.1 is an analog source which emits a message waveform $x(t)$ that is a sample function of a stochastic process $X(t)$ which can be represented by a sequence of uniform samples taken at the Nyquist rate. The samples are then quantized in amplitude and each of these discrete amplitudes can be mapped to a sequence of binary digits. This process is named analog-to-digital (A/D) conversion.

Sigma-Delta ($\Sigma\Delta$) modulation based analog-to-digital (A/D) conversion technology is a cost effective alternative for high-resolution applications. The fundamental technique in a $\Sigma\Delta$ system is oversampling. Oversampling eases analog filter design, and also generates a spectrum with quantization noise pushed to-



(a) Sampling frequency $f_s = 100$ Hz (b) Sampling frequency is $f_s = 200$ Hz

Figure 2.2: Effects of increasing sampling frequency

wards higher frequencies which are inaudible in audio applications. Over the last few years $\Sigma\Delta$ analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs) have become widely available, particularly for low-frequency applications such as high-fidelity audio and speech processing, metering applications, and voice-band data telecommunications [27] [83] [87].

Fig. 2.2 demonstrates that increasing the sampling frequency in the time domain will effectively increase the distance between the shifted replicas in the frequency domain and consequently a lower order of reconstruction filter may be applied.

Digital $\Sigma\Delta$ systems are easy for implementation and analysis. Analysis of $\Sigma\Delta$ modulation in the z-domain conventionally involves the assumption of a linear model in which the quantizer is modelled as Additive White Gaussian Noise (AWGN) source. In some circumstances this white noise assumption is not valid [43]. The input signal applied normally is a sinusoid or a DC. We present here an analysis by applying a Linear FM (LFM) signal as the input signal. The LFM possesses all possible frequencies with the same magnitude in the bandwidth of interest. All signals used in our analysis will be discrete sequences.

The oversampling process in $\Sigma\Delta$ modulators improves the resolution of a Nyquist rate data converter. This improvement is achieved by sampling the input signal at a significantly faster rate than the Nyquist rate. The ratio between the sampling rate and two times the signal bandwidth is defined as the

oversampling ratio (OSR) [43]

$$\text{OSR} = \frac{f_s}{2f_b} \tag{2.1}$$

where f_s is the sampling rate and f_b is the signal bandwidth. A system with OSR = 1 is generally called a Nyquist-rate system, while a system with OSR $\gg 1$ is called an oversampling system. In this analysis of $\Sigma\Delta$ systems, the bandwidth of the LFM applied signal is approximately 2 Hz and the sampling rate is 100 Hz, hence OSR = 25.

A discrete-time LFM is given by [54]

$$x[n] = \cos(\omega_o n^2) \tag{2.2}$$

whose instantaneous frequency (IF) is $2\omega_o n$.

The z-transform of a sequence $x[n]$ is defined as [54]

$$X(z) = \mathcal{Z}(x[n]) = \sum_{n=-\infty}^{\infty} x[n] z^{-n}. \tag{2.3}$$

The z-transform evaluated on the unit circle corresponds to the discrete Fourier transform (DFT). As for practical digital Δ or $\Sigma\Delta$ systems, all signals are finite-duration sequences, it is convenient to find their spectra by DFT. We plot the signals in the continuous-time domain for demonstration purposes.

2.2.2 Spectral Analysis of the Delta Modulator

We have studied the characteristics of Delta and Sigma-Delta systems using a frequency-domain approach [55]. Consider the 1-bit digital Δ modulator (encoder) and demodulator (decoder) shown in Fig. 2.3. A sinusoid of normalized frequency is applied to the system as a test signal. Fig. 2.4 shows the signals in the time domain and their corresponding spectra evaluated at the upper half of the unit circle by z-transform for the configuration shown in Fig. 2.3. Using this oversampled $\Sigma\Delta$ technique, the sinusoid is represented as a single-bit binary data stream $\in \{-1, 1\}$. For the plots of magnitude and phase in Fig. 2.4, the value of 1 in the frequency axis corresponds to $f_s/2$ which in this case, OSR = 50. We see in this trivial scenario that, as long as no slope-overload distortion occurs, the output of the quantizer $y[n]$, which is a single-bit binary data stream $\in \{-1, 1\}$, preserves the spectrum of the original signal in the bandwidth of interest. The spectrum obtained by DFT of this binary data stream $\in \{-1, 1\}$ is approximately equal to the spectrum of the input sequence $x[n]$ over baseband, and contains noise outside of baseband.

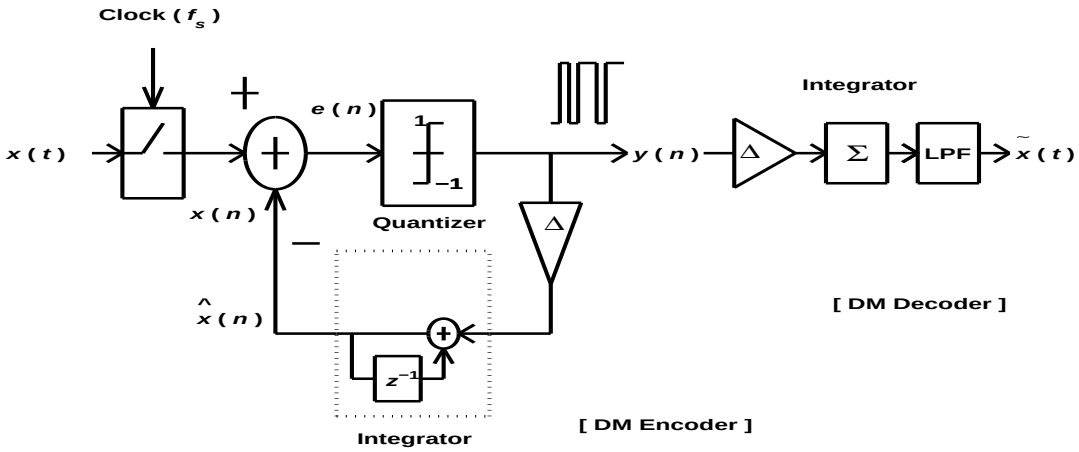


Figure 2.3: First-order digital Δ modulator and demodulator.

Now we use LFM as the input signal. The quantization process is shown in Fig. 2.5. We see that the spectrum of $y[n]$ is distorted as compared to the spectrum of $x[n]$. The Δ modulator (encoder) is sensitive to the sampling frequency and the quantization step. An integrator is necessary in the demodulator (decoder) to reconstruct the signal to $\hat{x}[n]$, as shown in Fig. 2.3, i.e, a step amplifier and an integrator are needed in the demodulator before low-pass filtering (LPF). We also notice that the phase is also "matched" in the bandwidth of interest.

Our design of the encoder focuses on matching the spectrum of $y[n]$ to the spectrum of the input signal $x[n]$. We first try an adaptive delta modulator (DM).

For the Δ modulator shown in Fig. 2.3, if the step is small and the input signal has a steep slope, the Δ modulator may loose tracking and a slope overload occurs. A large step size will incur a large granular noise. To avoid this problem, adaptive *Delta* modulator should be used. Fig. 2.6 shows a first-order digital adaptive Δ system, and the adaptive step size is given as [64]:

$$\Delta_n = \Delta_{n-1} K^{y(n)y(n-1)}. \tag{2.4}$$

where $K > 1$ is constant that can be arranged to minimize error. The quantization process of the LFM is shown in Fig. 2.7. The spectrum of $y[n]$ is getting closer to the spectrum of $x[n]$ as compared to non-adaptive Δ system over the baseband, but the distortion is large, thus an integrator is still necessary at the decoder.

The above analysis shows that the Δ modulator is sensitive to the rate of change of the signal.

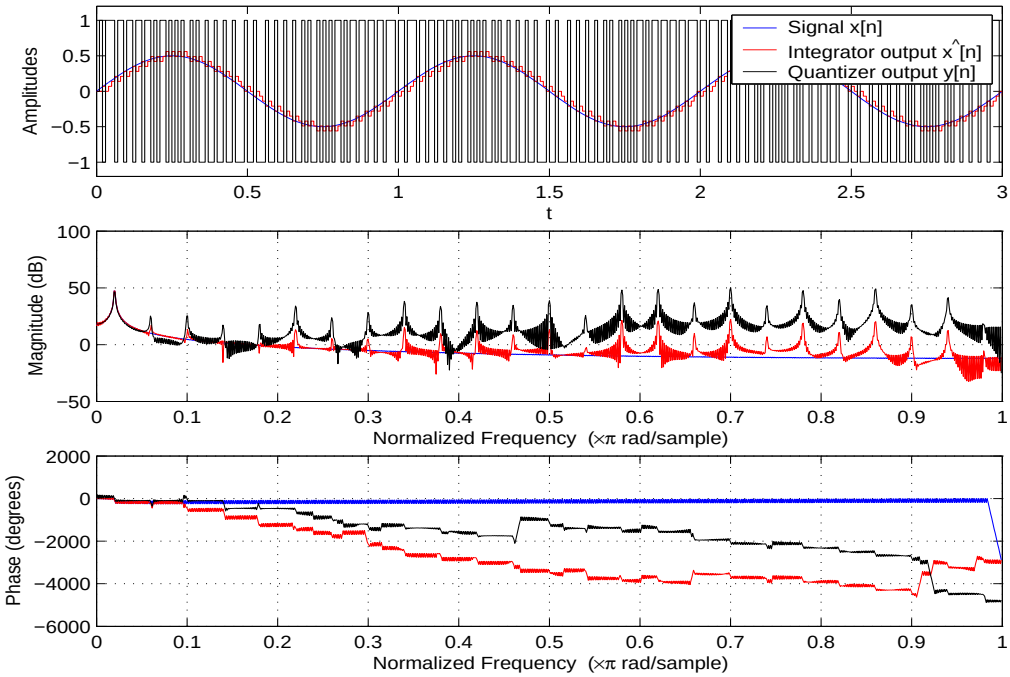


Figure 2.4: Quantizing a sinusoid by the Δ modulator: $x[n]$ is the input signal, $\hat{x}[n]$ is its estimate, and $y[n]$ is the output single-bit stream. The corresponding spectra are also shown.

2.2.3 Spectral Analysis of the Sigma-Delta Modulator

Delta modulation requires two integrators for the modulation and demodulation processes as shown in Fig. 2.3. Since integration is a linear operation, the second integrator can be moved before the modulator and becomes a preprocessing integrator. This will then be the configuration of $\Sigma\Delta$ modulator with two integrators. Again, based on the linear property of integration, the two integrators in Fig. 2.3 can be combined into a single integrator. A general configuration of a 1-bit $\Sigma\Delta$ modulator is shown in Fig. 2.8.

We again implement a digital version of Fig. 2.8 and show that a $\Sigma\Delta$ modulation system is a better solution for spectral matching. $\Sigma\Delta$ modulators encode the integral of the signal and thus their performance is insensitive to the rate of change of the signal.

Fig. 2.9 shows that the spectrum of the LFM $x[n]$ is squashed to lower frequencies by the preprocessing integrator and thus we have a situation similar to that in Fig. 2.3 and Fig. 2.4. The spectrum of the encoder output $y[n]$ matches the spectrum of the input of the encoder $x[n]$ over the bandwidth of interest, and hence only a LPF is needed in the decoder. The $\Sigma\Delta$ modulator is

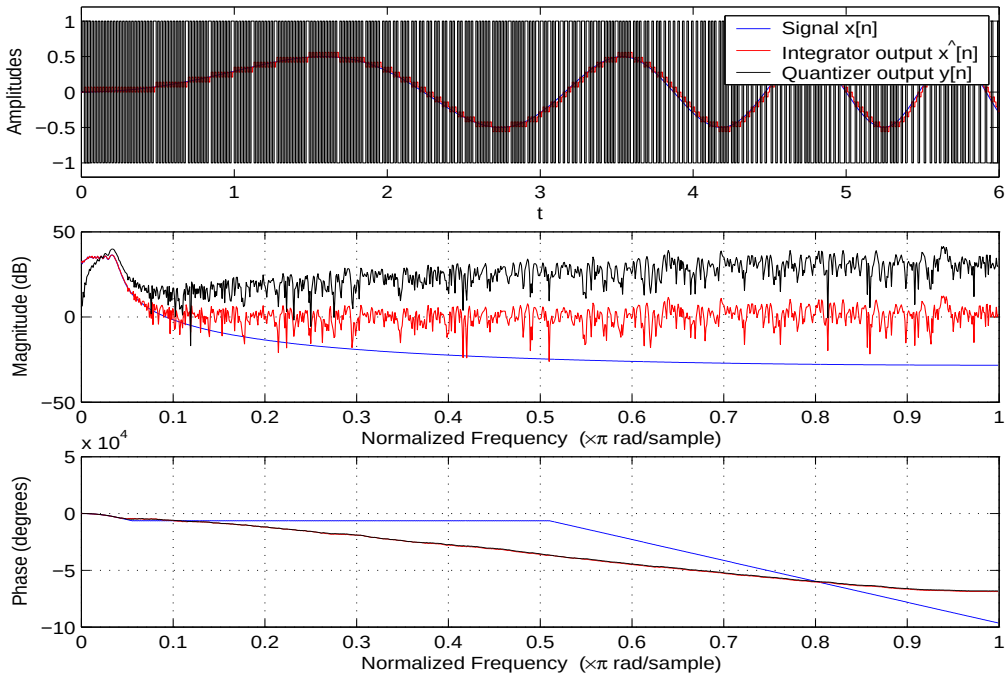


Figure 2.5: Quantizing an LFM signal by the Δ modulator: $x[n]$ is the input signal, $\hat{x}[n]$ is its estimate, and $y[n]$ is the output single-bit stream. The corresponding spectra are also shown.

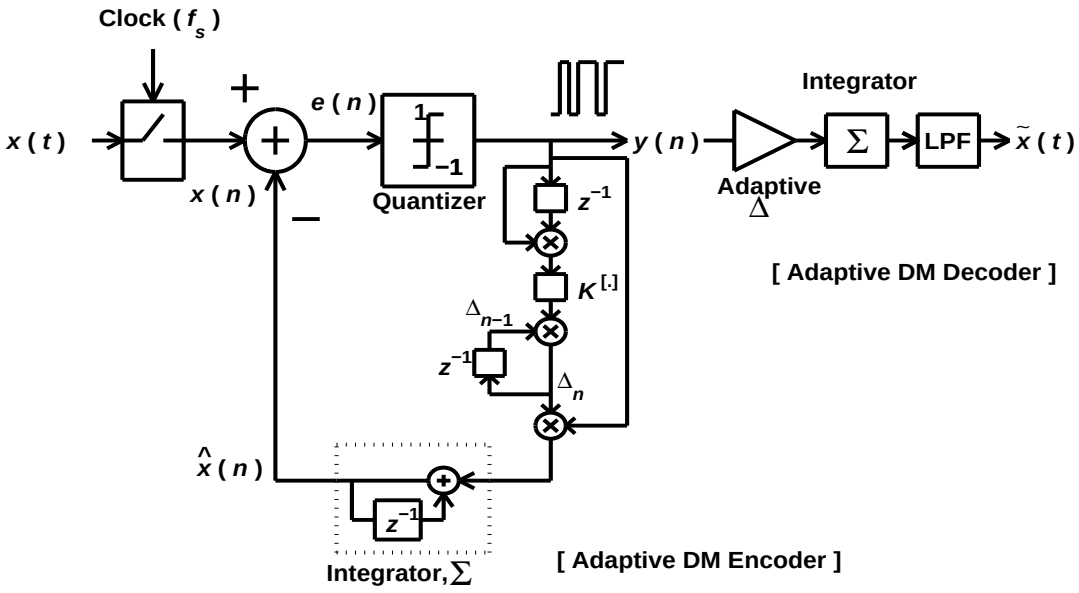


Figure 2.6: First-order digital adaptive Δ modulator and demodulator.

less sensitive to the quantization step and the sampling rate.

It should be also interesting to investigate the spectra of the $\Sigma\Delta$ modulator with one integrator as in Fig. 2.8. As shown in Fig. 2.10, there is no "spectrum squash" observed and the bandwidth of the output of the integrator is almost the same as that of the input signal $x[n]$. As in the case of two-integrator $\Sigma\Delta$, the spectrum of the input signal $x[n]$ matches that of the quantizer output $y[n]$ inside the bandwidth of interest, thus the decoder only needs a LPF to recover the input signal.

2.2.4 Noise Shaping of the Sigma-Delta Modulator

The noise shaping ability of an oversampling $\Sigma\Delta$ modulator allows the input signal of interest (baseband) to pass essentially unfiltered through the modulator but high-pass filters the quantization noise. For the ease of analysis of the important characteristics such as noise shaping, the nonlinearity of the quantizer in the $\Sigma\Delta$ modulator is approximated by an analytical linear model. There are arguments about the validity of the white noise linear model [43]. We investigate under what condition the linear model is valid. In the literature, Fig. 2.8 is approximated by a linear model to make the analysis tractable, i.e., the quantizer is linearized by using an input-independent additive white noise

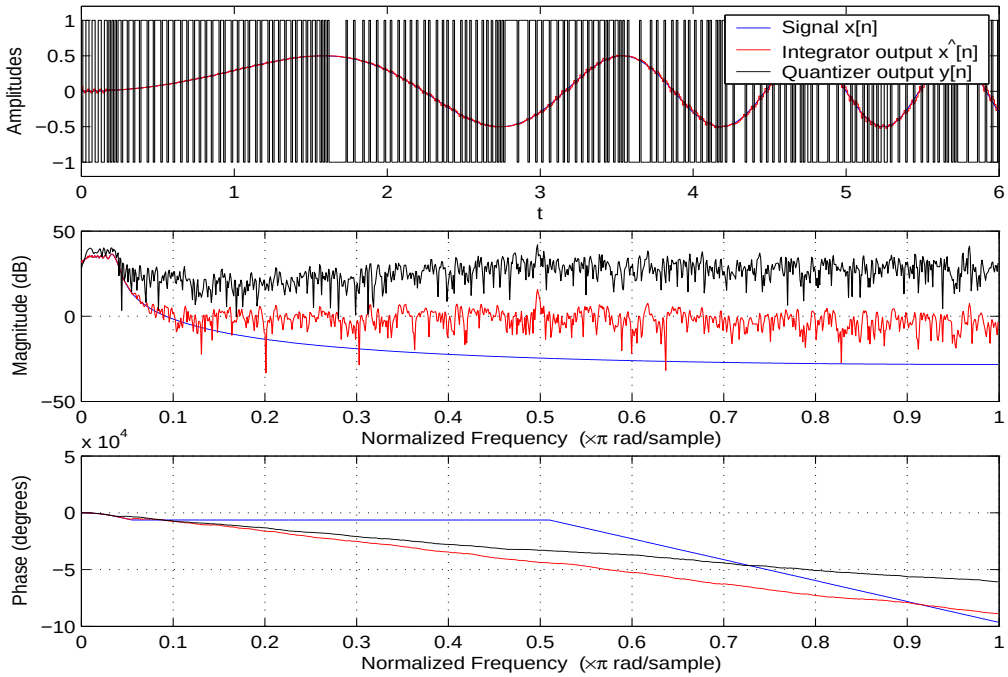


Figure 2.7: Quantizing an LFM by adaptive Δ modulator.

model, and the modulator output is given by [43] [4]:

$$Y(z) = X(z)z^{-1} + E(z)(1 - z^{-1}) \tag{2.5}$$

where $X(z)$, $Y(z)$, and $E(z)$ are the z-transforms of the input, the output, and the quantization error, respectively. If we let $H_x = z^{-1}$ and $H_e(z) = (1 - z^{-1})$, the output is just a delayed version of the signal plus quantization noise that has been shaped by a first-order z-domain differentiator or a high-pass filter. This process is known as the "noise shaping". The corresponding time-domain version of the modulator output is

$$y[n] = x[n-1] + e[n] - e[n-1] \tag{2.6}$$

where the $e[n] - e[n-1]$ term is the first-order difference of $e[n]$.

The transfer function $H_e(z) = (1 - z^{-1})$ is also called *noise transfer function* (NTF) $N(z) = (1 - z^{-1})$, and the magnitude of the NTF can be found out by letting $z = e^{\mathrm{j}2\pi f/f_s}$ and we have

$$|N(f)| = 2\sin(\pi f/f_s). \tag{2.7}$$

From Eq. (2.5) and Eq. (2.6), we obtain $\mathcal{Z}(e[n] - e[n-1]) = \mathcal{Z}(y[n] - x[n-1]) = E(z)(1 - z^{-1})$. This provides a way to evaluate the spectrum of

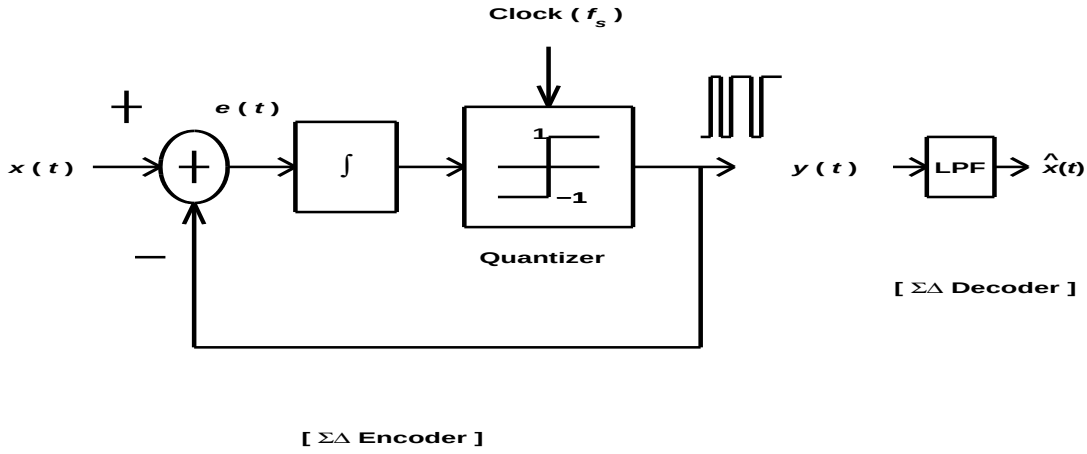


Figure 2.8: First-order digital 1-bit $\Sigma\Delta$ modulator and demodulator.

$E(z)(1 - z^{-1})$ from the difference of the input $x[n-1]$ and the output $y[n]$. Again, we use the LFM as the input signal $x[n]$. Fig. 2.11 shows a comparison of the magnitude spectra for $N(f)$ and simulated power spectral density (PSD) of $E(z)(1 - z^{-1})$ for the $\Sigma\Delta$ modulator shown in Fig. 2.8. Apparently, the match between the two curves in shape means $E(z) =$ constant, and hence $e[n]$ is a white noise process. In our simulation with an LFM input, increasing the duration of the LFM means a better match of the two curves (up to a scaling factor). The PSD of $(y[n] - x[n-1])$ in Fig. 2.11 is obtained with a sample size of 2^{12} and OSR $= 25$. The *noise shaping* is observed as the noise over the band of interest is significantly attenuated and is high-pass filtered outside the band of interest . In other words, the noise introduced by the quantization is pushed to higher frequencies which can be easily filtered out by a low pass filter (LPF) in DAC. Simulation also shows that with a higher OSR, the noise over the band of interest will be further attenuated. We conclude that at least in this case the linear model with the quantization error modelled as a white noise source is a valid assumption.

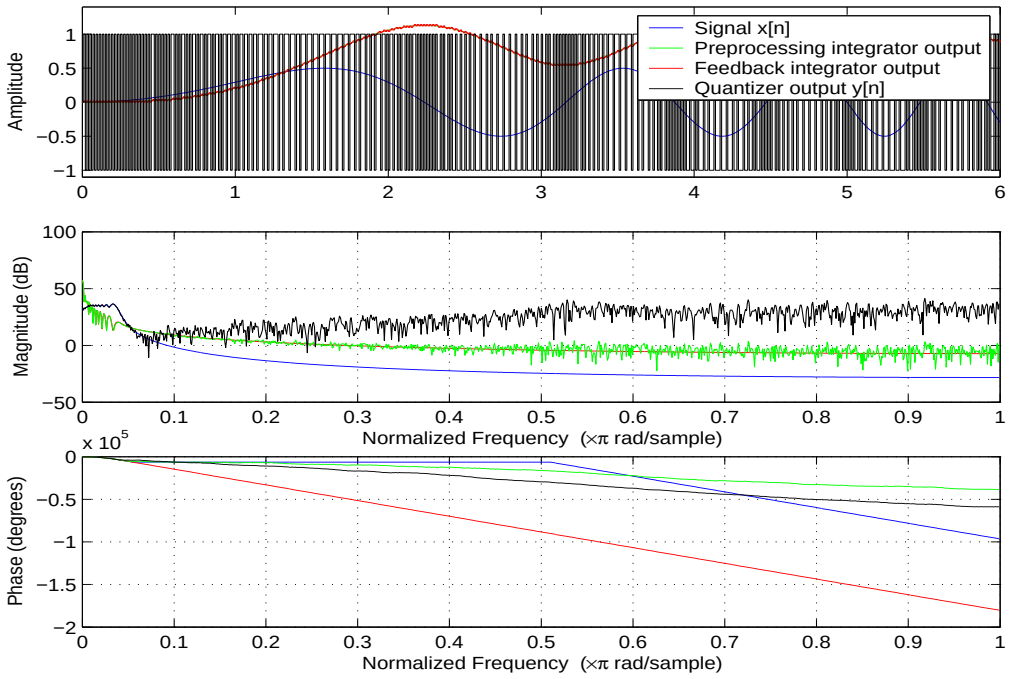


Figure 2.9: Quantizing an LFM by the $\Sigma\Delta$ modulator with two integrators: $x[n]$, preprocessing integrator output, feedback integrator output, $y[n]$ and their corresponding spectra.

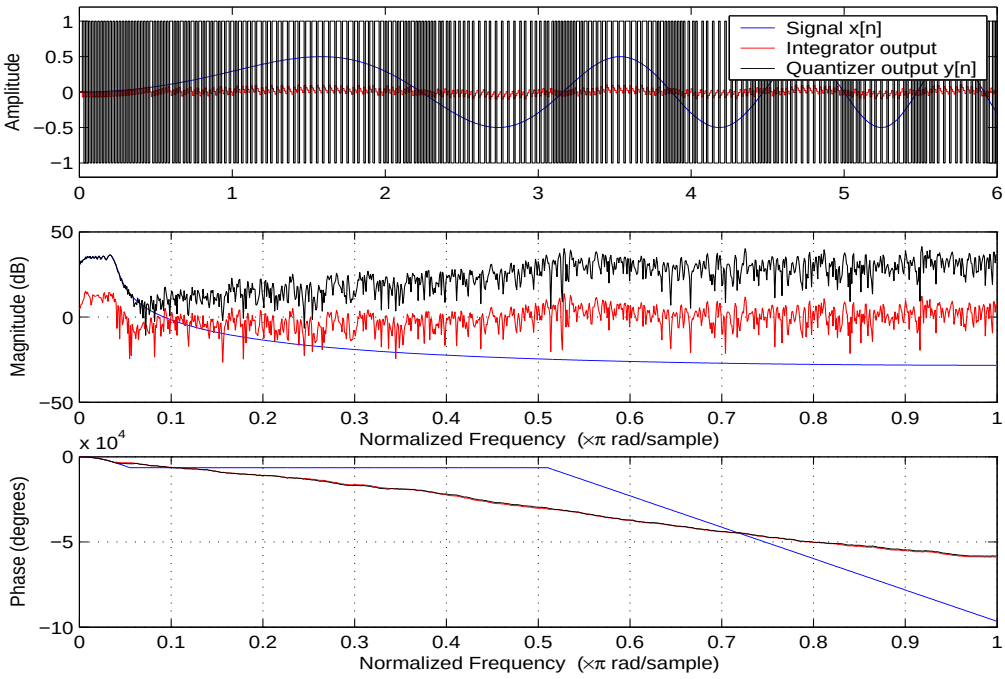


Figure 2.10: Quantizing an LFM by the $\Sigma\Delta$ modulator with one integrators: $x[n]$, Integrator output, $y[n]$ and their corresponding spectra.

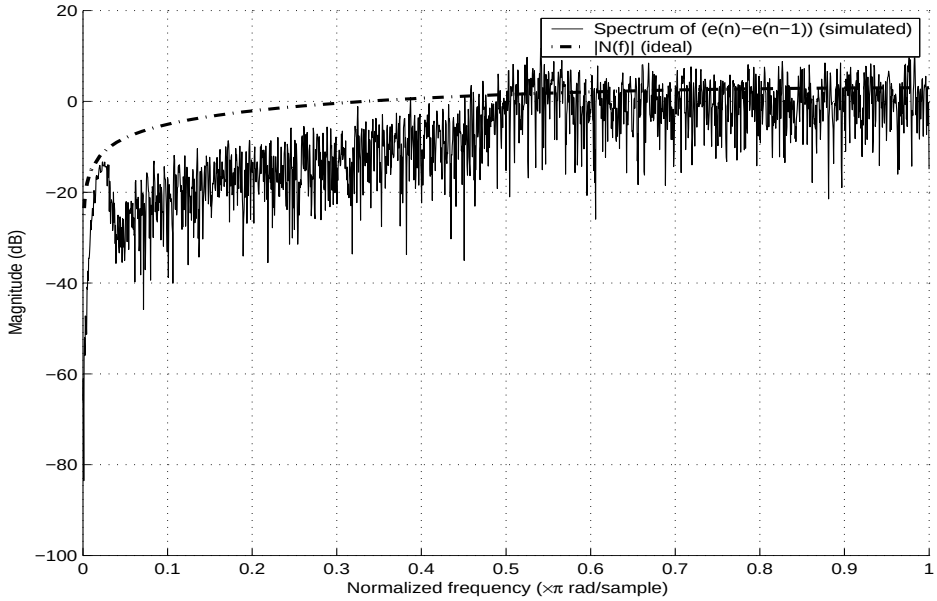


Figure 2.11: Noise shaping of the first-order $\Sigma\Delta$ modulator. Compare between $N(f)$ and the power spectrum density of $(y[n] - x[n-1])$ or $(e[n] - e[n-1])$. The dotted line shows the first-order noise-shaping characteristic as predicted by Eq. 2.5.

2.3 A Geometric View of Digital Modulation

Binary phase-shift keying (BPSK) is the simplest form of the digital modulation schemes, which can be represented in one-dimensional space. The concepts of signal spaces here are discussed in the case of two dimensions. Extending to higher-dimensional signal spaces is straightforward. Let $\varphi_1(t)$ and $\varphi_2(t)$ forms a orthonormal set. That is,

$$\int_{-\infty}^{\infty} \varphi_i(t)^2 dt = 1, \quad i = 1, 2$$

and

$$\int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt = 0, \quad i, j = 1, 2, i \neq j.$$

Define the inner product

$$< \varphi_i(t), \varphi_j(t) >= \int_{-\infty}^{\infty} \varphi_i(t)\varphi_j(t)dt, \quad i, j = 1, 2,$$

and $\{\varphi_1(t), \varphi_2(t)\}$ form an orthonormal set if

$$< \varphi_i(t), \varphi_j(t) >= \delta_{ij} \quad i, j = 1, 2,$$

where $\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$.

The signal $s(t)$ as a function of time can be obtained as the linear combination of the two functions $\varphi_1(t)$ and $\varphi_2(t)$:

$$s(t) = x_1\varphi_1(t) + y_1\varphi_2(t).$$

Projection of $s(t)$ on $\varphi_1(t)$ and $\varphi_2(t)$ will result in x_1 and y_1 respectively. That is, a signal $s(t)$ can be represented by a point (x_1, y_1) in the two-dimensional space.

A transmitted QAM signal waveforms may be expressed as

$$s_m(t) = A_m g_T(t)\cos(2\pi f_c t + \theta_m), \quad m = 1, 2, ..., M, \quad 0 \leq t \leq T \tag{2.8}$$

where f_c is the carrier frequency and $g_T(t)$ is a baseband pulse shape, which determines the spectral characteristics of the transmitted signal.

Expand s_m as

$$\begin{aligned} s_m(t) &= A_m g_T(t)\cos(2\pi f_c t + \theta_n) \\ &= A_m g_T(t)\cos 2\pi f_c t \cos\theta_n - A_m g_T(t)\sin 2\pi f_c t \sin\theta_n \\ &= \sqrt{\frac{\mathcal{E}_g}{2}} A_m \cos\theta_n \sqrt{\frac{2}{\mathcal{E}_g}} g_T(t)\cos 2\pi f_c t - \sqrt{\frac{\mathcal{E}_g}{2}} \mathrm{A_m} \sin\theta_n \sqrt{\frac{2}{\mathcal{E}_g}} g_T(t)\sin 2\pi f_c t \end{aligned} \tag{2.9}$$

where $\mathcal{E}_g$ is the energy of the pulse $g_T(t)$. Note that the orthonormal basis functions are $\varphi_1(t) = \sqrt{\frac{2}{\mathcal{E}_g}} g_T(t)\cos 2\pi f_c t$, and $\varphi_2 = -\sqrt{\frac{2}{\mathcal{E}_g}} g_T(t)\sin 2\pi f_c t$. And the QAM signals can be represented geometrically as two-dimensional vectors

$$\mathbf{s}_m = (\sqrt{\frac{\mathcal{E}_g}{2}} A_m \cos\theta_n, \sqrt{\frac{\mathcal{E}_g}{2}} A_m \sin\theta_n) \tag{2.10}$$

Fig. 2.12 illustrates a 16-QAM signal constellation.

In general, for signal $x(t)$ and $y(t)$, we define

- Energy of $x(t)$:

$$\mathcal{E}_x = \int_{-\infty}^{\infty} |x(t)|^2 dt \tag{2.11}$$

- Length of $x(t)$:

$$\|x\| = \sqrt{\mathcal{E}_x} = \sqrt{\int_{-\infty}^{\infty} |x(t)|^2 dt} \tag{2.12}$$

- Distance between $x(t)$ and $y(t)$:

$$D(x(t), y(t)) = \|x(t) - y(t)\| = \sqrt{\int_{-\infty}^{\infty} |x(t) - y(t)|^2 dt} \tag{2.13}$$

The vector space representation of the signal provides a convenient way to compute the above quantities. It is easy to prove relation like $\|x(t)\| = \|\mathbf{x}\|$. For the signal constellations shown in Fig. 2.12, the average energy/symbol is given by $\mathcal{E}_{\mathrm{av}} = \frac{1}{M}\sum_{i=1}^{M} \|\mathbf{s}_i\|^2$. The Euclidean distance between any pair of signal points is

$$d_{mn} = \sqrt{\|\mathbf{s}_m - \mathbf{s}_n\|} \tag{2.14}$$

To further demonstrate the understanding of a digital communication system and the concept of the geometric representation of signal waveforms, we simulate

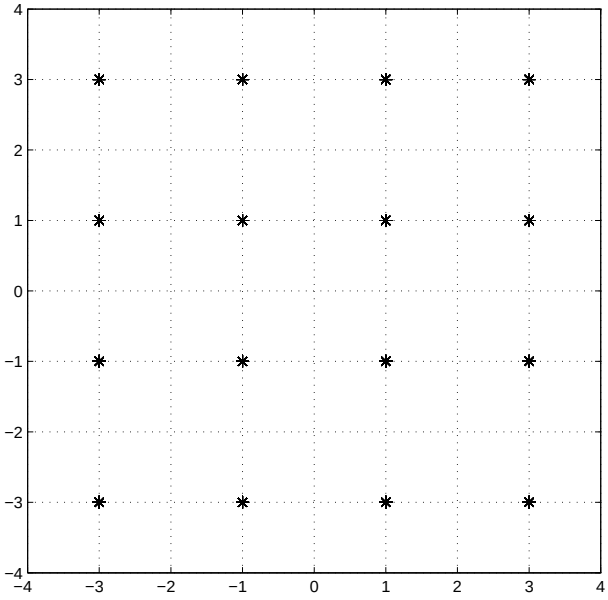


Figure 2.12: M = 16-QAM signal constellation.

a 16-QAM system with ideal transmission channel, i.e., the channel is noiseless and no Inter-Symbol-Interference (ISI) is introduced by the channel. The pulse $g_T(t)$ having the raised cosine spectrum is

$$g_T(t) = \text{sinc}(t/T)\frac{\cos(\pi\alpha t/T)}{1 - 4\alpha^2 t^2/T^2} \tag{2.15}$$

Fig. 2.13 illustrates the pulse for $\alpha = 0.5$ which is used in simulation of the 16-QAM digital transmission system. In practice, the overall raised cosine spectral characteristic is split evenly between the transmitting filter and the receiving filter. The raised cosine pulse is a commonly used pulse which satisfies Nyquist criterion. As the physical application is generally band-limited, the Nyquist criterion indicates a maximum symbol rate for avoiding ISI for a given bandwidth.

Figure 2.14 shows that the symbols modulated baseband signal in both time domain and frequency domain. The randomness of the signal is due to the random nature of the symbols. The symbols are modulated to the maximum symbol rate for a given bandwidth. That is, assume the constraint of $|f| < W$, the maximum symbol rate that can be achieved with zero ISI is $1/T = 2W$.

As the practical communication channels are generally passband in nature, the baseband signal is modulated to the channel center frequency f_c for the transmission over the channel. The passband signal is illustrated in Fig. 2.15.

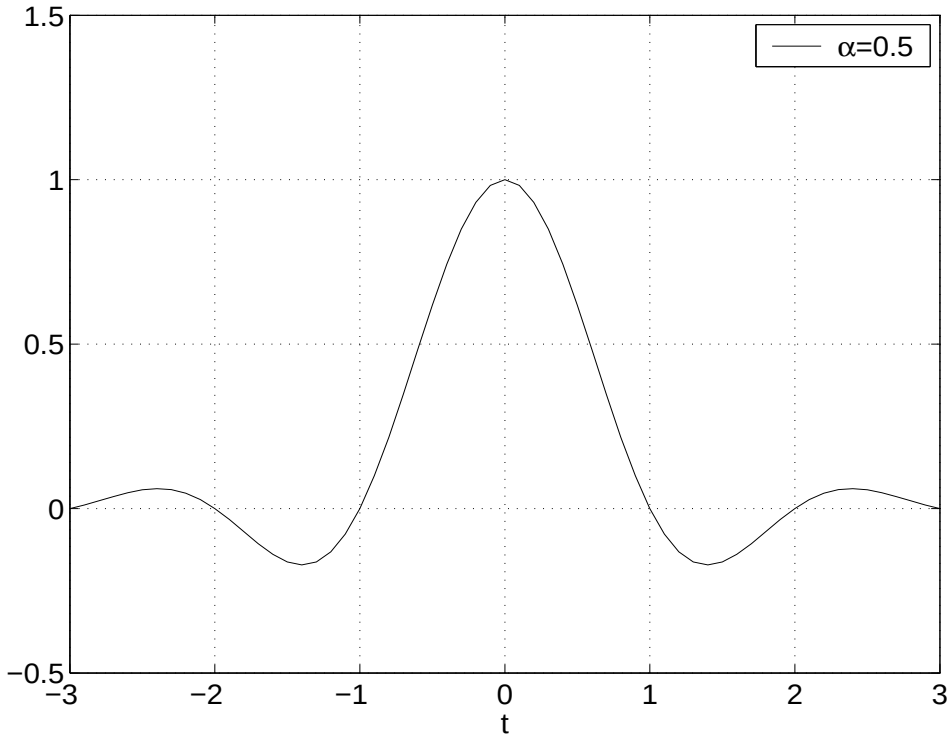


Figure 2.13: Pulse having a raised cosine spectrum with the rolloff factor $\alpha = 0.5$.

At the receiver, the passband is down-converted to the baseband and sampled to recover the constellation as that of Fig. 2.12. The spectrum of the down-converted signal and the low passed filtered signal is shown in Fig. 2.16.

In practice, the received signal is corrupted by ISI and noise introduced by the channel and other factors such as multi-user sharing the same transmission, in addition to constraints such as power consumption, therefore much more complicated techniques are employed. Advanced Shannon-limit approaching coding algorithms are also used to improve the system performance [25] [26] [72]. In the case of wireless communication, emerging and promising techniques such as Orthogonal Frequency Division Multiplexing (OFDM), space-time coding are employed to combat the highly dynamic channel [68].

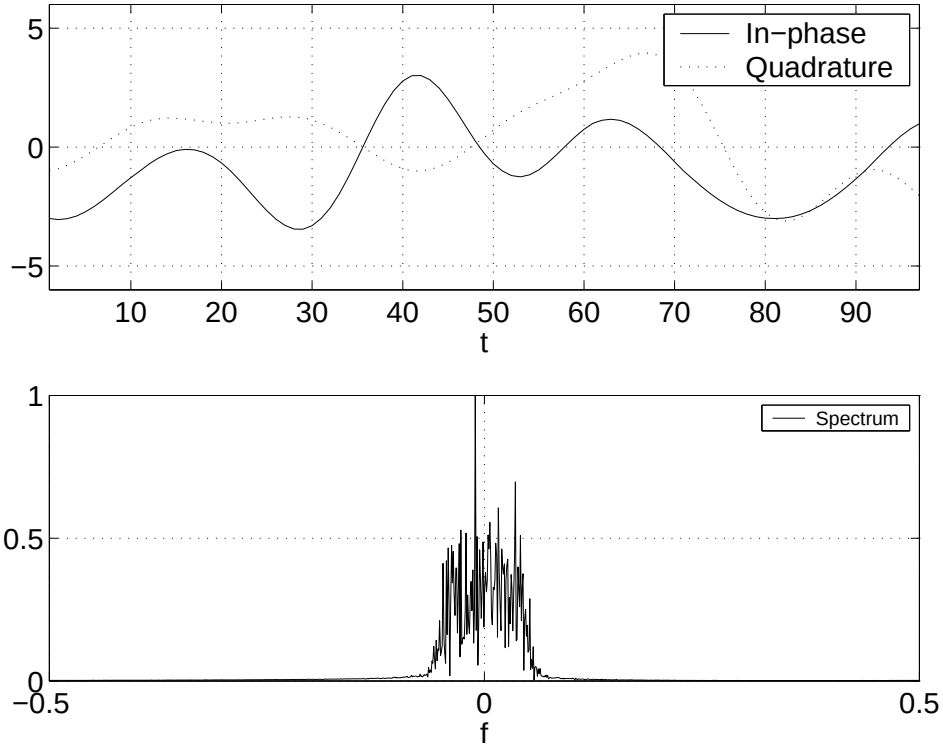


Figure 2.14: The baseband signal in time domain (top) and in frequency domain (bottom).

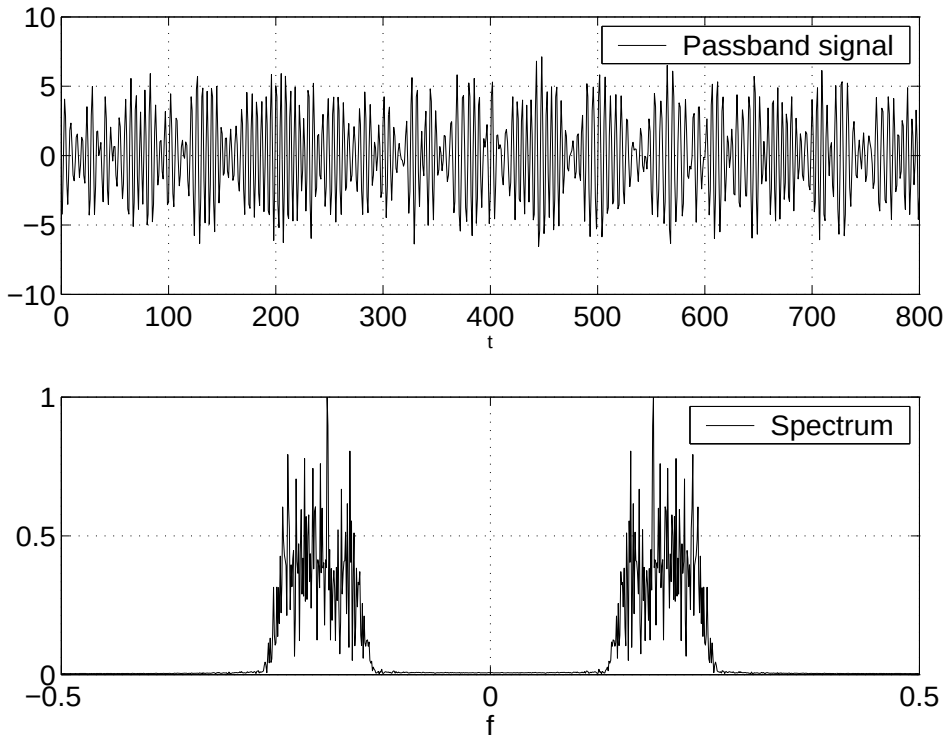


Figure 2.15: The passband signal in time domain (top) and in frequency (bottom).

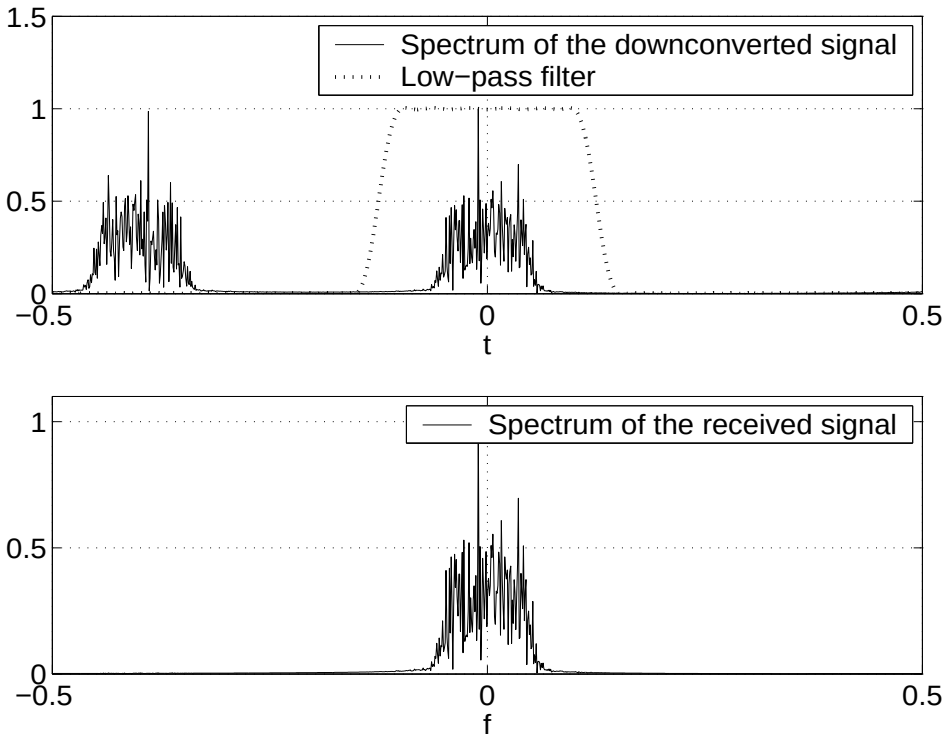


Figure 2.16: The IQ down-converted signal (top) and the received baseband signal after low-pass filtering (bottom).

2.4 The Promise of Coding

In the previous section we discussed a digital communication system in the noise free environment. In practice, $s(t)$ may be disturbed by interference and noise. Assuming that $s(t)$ is a set of two waveforms, we consider the case of $s(t)$ corrupted by the additive white Gaussian noise corrupting

$$r(t) = s(t) + n(t), \tag{2.16}$$

where the channel is modelled a white Gaussian noise process, having the properties that

$$E[n(t)] = 0 \qquad \forall t,$$

$$R_n(\tau) = E[n(t)n(t-\tau)] = \frac{N_0}{2}\delta(\tau),$$

where $n(t)$ denotes the sample function of the additive white Gaussian noise (AWGN) process with power-spectral density $\frac{N_0}{2}$ W/Hz. At the receiver, the signal demodulator and the detector convert the received waveform $r(t)$ into

$$\mathbf{R} = \mathbf{s} + \mathbf{N} \tag{2.17}$$

all are n-dimensional vectors, with $\mathbf{N} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$, and $\mathbf{S} \in \{\mathbf{a}, \mathbf{b}\}$ are two transmitted vectors. Recovering the transmitted messages at the receiver is now a classic detection problem. See [62] for a strict treatment of detection and estimation theory.

The general relation for the probability of decision error is [49][63]

$$P_{\mathrm{e}} = Q\left(\frac{d}{2\sigma}\right) \tag{2.18}$$

where $d = \sqrt{\|\mathbf{a} - \mathbf{b}\|}$ is the Euclidean distance between vectors, and $Q(\cdot)$ is the complementary cumulative distribution function of an $N(0,1)$ random variable.

With the variance for the channel is expressed as $\sigma^2 = \frac{N_0}{2}$ by optimal demodulation, P_{e} can be also expressed as

$$P_{\mathrm{e}} = Q\left(\frac{d}{\sqrt{2N_0}}\right) \tag{2.19}$$

In the special case of BPSK signalling, the transmitted signal is $S \in \{-\sqrt{\mathcal{E}_{\mathrm{b}}}, \sqrt{\mathcal{E}_{\mathrm{b}}}\}$, the Euclidean distance is therefore $d = 2\sqrt{\mathcal{E}_{\mathrm{b}}}$ and for the noise we have $\sigma^2 = \frac{N_0}{2}$. From (2.19) we calculate the probability of error as

$$P_{\mathrm{e}} = Q\left(\sqrt{\frac{2\mathcal{E}_{\mathrm{b}}}{N_0}}\right) \tag{2.20}$$

The main purpose of using coding in communication systems is to increase the Euclidean distance d and consequently reduce P_{e} with the power constraint.

A significant result that emerges from information theory is that if the entropy of the source is less than the capacity of the channel, then error-free communication over the channel can be achieved, with the help of channel coding [82]. We demonstrate the promise of error correction coding using repetition code and Hamming code. Although the repetition code improves nothing on the performance, the study of it reveals the essential ideas of coding.

Assume a communication of BPSK over BSC (Binary Symmetric Channel) with crossover probability p. We first investigate (n,1) repetition coding. The probability of decoding error using *hard-decision* decoder with majority vote strategy for a repetition code of length n is

$$P_{\mathrm{e}}^{n} = \sum_{i=t+1}^{n} \binom{n}{i} p^{i}(1-p)^{n-i} + \text{terms of higher degree in } p. \tag{2.21}$$

This shows that arbitrarily small probability of error can be obtained as n increasing with p a constant. The price paid for this arbitrarily small probability of error is the transmission rate approaching zero.

The AWGN channel can be roughly viewed as an equivalent BSC. With power constraints, the energy available for each coded bit, which is denoted as $\mathcal{E}_{\mathrm{c}}$, is $\mathcal{E}_{\mathrm{c}} = \mathcal{E}_{\mathrm{b}}/n$. The crossover probability for the equivalent BSC is therefore

$$p = Q(\sqrt{2\mathcal{E}_{\mathrm{c}}/N_0}) = Q(\sqrt{2\mathcal{E}_{\mathrm{b}}/nN_0})$$

Since $Q(x)$ is a decreasing function, the effective crossover probability p is higher as a result of using the repetition. Replacing the crossover probability p in (2.21) with different n shows that the coded performance by the repetition coding is worse than the uncoded performance, and the performance gets worse with increasing n. We conclude that with power constraint, repetition coding by hard-decision decoding will only make the performance worse than uncoded transmission.

We now show that *soft-decision* decoding is better than *hard-decision* decoding, and with the *soft-decision* decoding, the probability of error for the $(n, 1)$ repetition code is exactly the same as uncoded transmission-still no performance improvement.

Denote the output of BSC by

$$\mathbf{r} = \mathbf{c} + \mathbf{n},$$

where $\mathbf{r}$ is the received vector, $\mathbf{c}$ is the coded message by the repetition encoding. The likelihood function is

$$p(\mathbf{r}|\mathbf{c}) = \prod_{i=1}^{n} p(r_i|c_i), \tag{2.22}$$

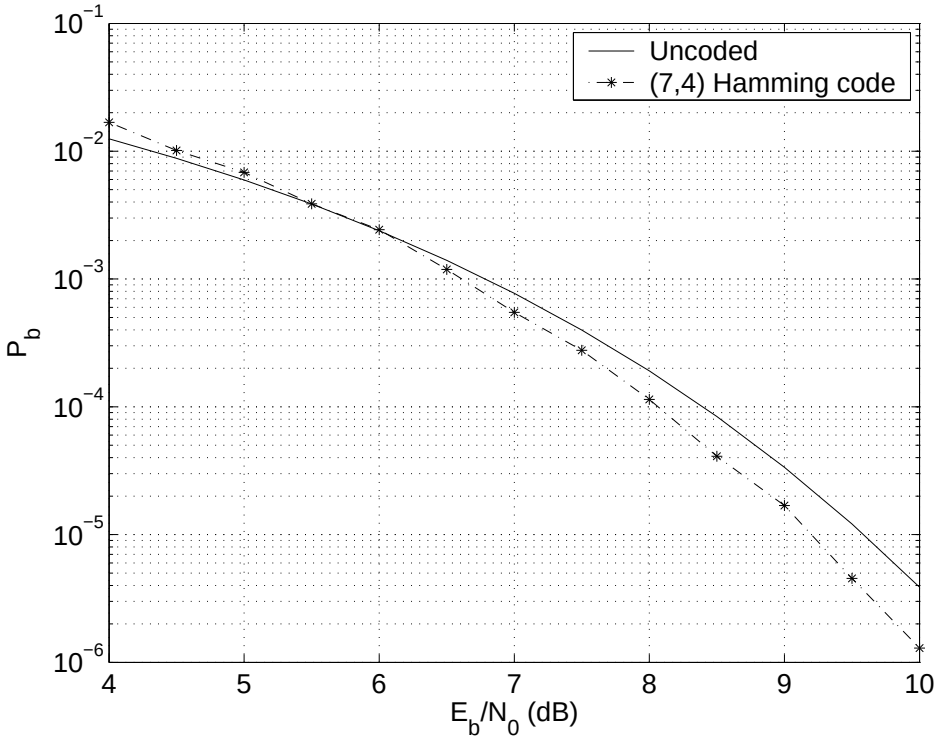


Figure 2.17: Performance of (7,4) Hamming code.

so that the log likelihood ratio can be derived as

$$\begin{aligned}
\Lambda(\mathbf{r}) &= \log\frac{p(\mathbf{r}|m=1)}{p(\mathbf{r}|m=0)} \\
&= \log\frac{\prod_{i=1}^{n} p(r_i|m=1)}{\prod_{i=1}^{n} p(r_i|m=0)} \\
&= \log\frac{\prod_{i=1}^{n} p(r_1|s_1=\sqrt{\mathcal{E}_\mathrm{c}})}{\prod_{i=1}^{n} p(r_1|s_1=-\sqrt{\mathcal{E}_\mathrm{c}})} \\
&= \log\frac{\prod_{i=1}^{n} \exp(-\frac{1}{2\sigma^2}(r_i-\sqrt{\mathcal{E}_\mathrm{c}})^2)}{\prod_{i=1}^{n} \exp(-\frac{1}{2\sigma^2}(r_1-\sqrt{\mathcal{E}_\mathrm{c}})^2)} \\
&= \frac{2\sqrt{\mathcal{E}_\mathrm{c}}}{\sigma^2}\sum_{i=1}^{n} r_i.
\end{aligned} \tag{2.23}$$

The soft-decoder thus decides $\hat{m}$ as

$$\hat{m} = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} r_i > 0 \\ 0 & \text{if } \sum_{i=1}^{n} r_i < 0. \end{cases}$$

In this case, the Euclidean distance $d = 2n\sqrt{\mathcal{E}_c}$ and $\sigma^2 = n\frac{N_0}{2}$. From (2.19), we calculate the probability of error for the $(n, 1)$ repetition code with soft-decision decoding is

$$P_e = Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right). \tag{2.24}$$

This is the same as (2.20). We conclude that soft-decision decoding is superior to hard-input decoding in terms of probability of error but is not effective as an error correction code.

We simulate the Hamming codes which are much better than repetition codes. The simulation result is given in Fig. 2.17. As shown in the figure, (7,4) Hamming code will give a coding gain of 0.45 dB, i.e., the code is able to overcome the 2.4 dB of loss due to rate (less energy per bit due to power constraint), and add another 0.45 dB of improvement. According to Shannon, there is at least 9.6 dB of coding gain possible.

The price paid to increase the performance is a bandwidth expansion due to a narrow transmission pulse is needed. This can be intuitively understood from the discussion of the previous section. For example, the bandwidth expansion ratio for the (7,4) Hamming code is

$$\frac{W_{\text{coded}}}{W_{\text{uncoded}}} = \frac{n}{k} = \frac{7}{4} = 1.75.$$

2.5 Conclusion

In this chapter we have studied some of the fundamental notions in a digital communication system. First, we considered the quantization process, including sigma-delta data conversion, where a study based on frequency domain analysis is introduced. A linear FM (LFM) signal is applied as a test signal at the input of digital Σ and $\Sigma\Delta$ modulators to reveal the relationship of the spectra at different stages. By squashing the spectrum of the signal using the $\Sigma\Delta$ modulator, a simpler decoder and a better performance is achieved for the $\Sigma\Delta$ modulator. We see that for the same parameters such as the sampling frequency and quantization step, the $\Sigma\Delta$ modulator outperforms the adaptive Δ modulator. The noise shaping ability and the validity of the linear model of the $\Sigma\Delta$ modulator are also investigated. Similar principle may be applied to other configurations of the $\Sigma\Delta$ modulator in future works.

We also presented a geometric view of the main digital modulation techniques. Error performance is also studied with AWGN channels. This performance will be the basis for evaluating the chosen channel model and the communication system.

Chapter 3

Modelling of the Wireless Channels

3.1 General Description of the Wireless Channel

The communication system presented in the previous chapters is simulated using an ideal channel model. In wireless communication, the channel will introduce degradation or interference. All relevant components of a mobile radio system, from digital modulation techniques over to channel coding through to network aspects, are determined by the propagation characteristics of the channel. Therefore, a precise knowledge of the mobile radio channels is crucial for the development, evaluation and test of current and future mobile radio communication systems [56] [3].

Sources of degradation from transmission channel can be categorized as [101]:

1. Additive Gaussian noise
2. Bandlimiting of the signal by the channel
3. Non-Gaussian noise, such as impulse noise due to lightning discharges or switches
4. Radio frequency interference due to other transmitters
5. Multiple transmission paths, termed multipath, due to stratifications in the transmission medium or objects that reflect or scatter the propagating signal.

The basic propagation mechanisms of the mobile radio channel such as reflection, diffraction and scattering impairments cause the signal at the receiver to distort or fade significantly as compared to AWGN channels.

Fading, or small-scale fading, is caused by interference between two or more versions of the transmitted signal which arrives at the receiver at slightly different times. These waves, called *multipath waves*, combine at the receiver antenna to give a resultant signal which can vary widely in amplitude and phase, depending on the distribution of the intensity and relative propagation time of the waves and the bandwidth of the transmitted signal.

In the case of BPSK, the received direct-path signal may be described as

$$s_d(t) = Ad(t)\cos 2\pi f_c t \tag{3.1}$$

where f_c is the carrier frequency.

We consider a two-ray multipath model and the received signal is [101]

$$y(t) = s_d(t) + \beta s_d(t - \tau_m) + n(t) \tag{3.2}$$

where $s_d(t)$ is the received direct-path signal, β is the attenuation of the multipath component, and τ_m is its delay.

To find out the transfer function of the channel, $H_c(f)$, we take Fourier transform for $y(t)$ and $s_d(t)$

$$H_c(f) = \frac{\mathcal{F}[y(t)]}{\mathcal{F}[s_d(t)]} \tag{3.3}$$

For the Fourier pair $s_d(t) \leftrightarrow S_d(f)$, we apply the time delay property $x(t - t_0) \leftrightarrow X(f)e^{-j\omega t_0}$ to (3.2), we have

$$y(t) \leftrightarrow S_d(f) + \beta S_d(f)e^{-\mathrm{j}2\pi f\tau_m} = S_d(f)(1 + \beta e^{-\mathrm{j}2\pi f\tau_m})$$

and therefore the transfer function is

$$H_c(f) = 1 + \beta e^{-\mathrm{j}2\pi f\tau_m} \tag{3.4}$$

We demonstrate here with a simple two-ray multipath model that the wireless channel is in generally frequency-selective. Theoretically, the channel effect can be reversed by a filter, referred to as an equalizer with the transfer function

$$H_{\mathrm{eq}}(f) = \frac{1}{H_c(f)} = \frac{1}{1 + \beta e^{-\mathrm{j}2\pi f\tau_m}} \tag{3.5}$$

In practice, channels are generally waveform channels that accept continuous-time waveforms as their inputs and output continuous-time waveforms. Practical channels are always band-limited, and by using the sampling theorem, a waveform channel can be equivalently represented by a discrete-time channel. This equivalence facilitates analysis and algorithm development because standard digital signal processing may apply.

The equivalent discrete-time channel may be modelled as a discrete-time FIR filter. If a discrete channel does not have memory, it is called a *discrete-memoryless channel*. A special case of a discrete-memoryless channel is the *binary-symmetric channel* (BSC) which we have encountered in the previous chapter.

The countermeasures to fading channels are *equalization*, *diversity* and *channel Coding*.

Equalization by equalizer such as Eq. (3.5) compensates for *inter-symbol interference* (ISI) created by multipath within time dispersive channels. *Channel Coding* such as the $(7, 4)$ Hamming code improves the small-scale link performance by adding redundant data bits.
Diversity is used to reduce the depth and duration of the fades. Diversity can be further categorized as

- spatial diversity: multiple antennas
- time diversity: CDMA RAKE receiver
- frequency diversity: OFDM
- antenna polarization

As with an equalizer, diversity improves the quality of a wireless communications link without altering the common air interface, and without increasing the transmitted power or bandwidth. However, while equalization is used to counter the effects of time dispersion (ISI), diversity is considered more effective in wireless communication to combat the depth and duration of the fades experienced by a receiver in a local area which are due to motion [70].

3.2 Modelling Philosophy

The channel filter taps are modelled statistically as random process. It should be noted that probabilistic models for the channel filter taps are going to be far less believable than the models for additive noise. On the other hand, despite the inaccuracy of the channel models, we still need such models for system design and performance evaluation. The reason is that even with a highly oversimplified model, we can compare different algorithms for system designs. As demonstrated in Chapter 2, a simple BSC will enable us to evaluate different coding schemes in the example of $(7, 4)$ Hamming code.

In typical wireless applications, communication occurs in a passband with a certain bandwidth. The use of a baseband equivalent of the communication system simplifies the analysis because the analysis is independent of carrier

frequencies and channel frequency bands. Another advantage is that the only the slowly varying parts of the signal need to be processed and therefore is computationally less expensive.

3.3 Clarke's Model with a Continuum of Paths

The Clarke's model (or sometimes called *Jakes'* model) which is detailed in [17] is a popular statistical model for flat fading channel. In this section we derive the time-varying impulse response of this channel model. We incorporate the input-output relationship of this model to the general linear time-varying filter channel model and derive the time-varying channel impulse response for the Clarke's model with a continuum of paths. The mathematical structure and physical interpretation of this derivation are insightful and will be very helpful for further studies.

Physical channels such as wireless channels may be characterized by a time-variant channel impulse response $h(\tau, t)$, where $h(\tau, t)$ is the response of the channel at time t due to an impulse applied at time $t - \tau$. For an input signal $x(t)$, the channel output signal in the noiseless condition is

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t)x(t - \tau)d\tau \tag{3.6}$$

In the Clarke's model, the transmitter is fixed, and the scatterers are assumed uniformly located around the mobile receiver which is moving at speed v. In the discrete model, there are K paths, the ith path arriving at an angle $\theta_i := 2\pi i/K, i = 0, ..., K-1$, with respect to the direction of motion, where K is assumed to be large. The path coming at angle θ with respect to the direction of motion of the mobile has a delay of $\tau_\theta(t)$ and a time-invariant gain a_θ, and the input-output relationship is given by:

$$y(t) = \sum_{i=0}^{K-1} a_{\theta_i} x(t - \tau_{\theta_i}(t)) \tag{3.7}$$

The continuous version of (3.7) is

$$y(t) = \int_0^{2\pi} a_\theta x(t - \tau_\theta(t))d\theta \tag{3.8}$$

Eq. (3.8) can be reorganized as

$$\begin{aligned} y(t) &= \int_0^{2\pi} a_\theta x(t-\tau_\theta(t))d\theta \\ &= \int_0^{2\pi} a_\theta \Big(\int_{-\infty}^{\infty} \delta(\tau-\tau_\theta(t)) x(t-\tau_\theta(t))d\tau \Big) d\theta \\ &= \int_0^{2\pi} a_\theta \Big(\int_{-\infty}^{\infty} \delta(\tau-\tau_\theta(t)) x(t-\tau)d\tau \Big) d\theta \\ &= \int_{-\infty}^{\infty} \Big(\int_0^{2\pi} a_\theta \delta(\tau-\tau_\theta(t))d\theta \Big) x(t-\tau)d\tau \end{aligned} \tag{3.9}$$

Comparing the last equality with (3.6), we obtain the impulse response as

$$h(\tau,t) = \int_0^{2\pi} a_\theta \delta(\tau-\tau_\theta(t))d\theta \tag{3.10}$$

Alternatively, this result can be obtained intuitively and consequently more insights into the channel can be gained. Let the input be $x(t) = \delta(t-t')$, we then have $x(t-\tau_\theta(t)) = \delta(t-\tau_\theta(t)-t')$. Plug this into (3.8), we have

$$\begin{aligned} y(t) &= \int_0^{2\pi} a_\theta x(t-\tau_\theta(t))d\theta \\ &= \int_0^{2\pi} a_\theta \delta(t-\tau_\theta(t)-t')d\theta \\ &\quad (\text{Let } t' = t-\tau) \\ &= \int_0^{2\pi} a_\theta \delta(\tau-\tau_\theta(t))d\theta \end{aligned} \tag{3.11}$$

which is the same as (3.10). This shows that $h(\tau,t)$ can be interpreted as the response of the channel to an impulse that occurs at time $t-\tau$.

3.4 Single Tap Discrete-Time Baseband Channel Model

In this section we derive the single tap discrete-time baseband channel model. A wireless channel bandlimited in $[f_c - W/2, f_c + W/2]$ can be modelled by a discrete-time baseband model in terms of channel filter taps as [92]

$$y[m] = \sum_\ell h_\ell[m]x[m-\ell] + w[m], \tag{3.12}$$

where

$$h_\ell[m] = \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)} \mathrm{sinc}[\ell - \tau_i(m/W)W], \tag{3.13}$$

and $a_i(m/W)$ and $\tau_i(m/W)$ are sampled attenuation and delay respectively. The noise $w[m]$ is distributed as $\mathcal{CN}(0, N_0)$ and i.i.d. over time. We now show that in the case that the delay spread of the channel T_d satisfies $T_\mathrm{d} \ll 1/W$, the discrete-time baseband model can be approximately represented by a single tap model.

From $T_\mathrm{d} \ll 1/W$ we know for all $\tau_i(m/W)$, we have $\tau_i(m/W) \ll 1/W$, thus

$$0 \le \tau_i(m/W)W \ll 1, \quad \forall i$$

That is, $\tau_i(m/W)W$ is approximately zero for all index i. Therefore (3.13) can be expressed as

$$\begin{aligned} h_\ell[m] &= \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)} \mathrm{sinc}[\ell - \tau_i(m/W)W] \\ &= \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)} \delta[\ell] \\ &= \delta[\ell] \cdot \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)} \\ &= \begin{cases} \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)}, & \ell = 0 \\ 0, & \text{else.} \end{cases} \end{aligned} \tag{3.14}$$

Plug this result into (3.12), we have the single tap model as

$$\begin{aligned} y[m] &= h_0[m]x[m] + w[m] \\ &= h[m]x[m] + w[m], \end{aligned} \tag{3.15}$$

where $h_0[m] = h[m] = \sum_i a_i(m/W) e^{-\mathrm{j}2\pi f_c \tau_i(m/W)}$ is the sum of large number of small independent symmetric random variables, and by Central Limit Theorem (CLT), is modelled as a zero-mean complex Gaussian random variable.

3.5 Capacity of Fading Channels

The motivation to investigate the capacity of fading channels is many-sided. Harnessing information-theoretic tools to the investigation of fading channels, in the widest sense of this notion, has not only resulted in an enhanced understanding of the potential and limitations of those channels, but in fact Information Theory provided in numerous occasions the right guidance to the specific design of efficient communications systems [9].

We have shown in Chapter 2 that by an intelligent design of coding such as $(7, 4)$ Hamming code, one can increase the performance of communication over AWGN channel by making the error probability small without reducing the data rate. Claude Shannon's information theory indicates that in fact the error probability can be made as small as possible as long as the transmission rate is below the *capacity* of the channel.

The AWGN channel, which is a special case of (3.15) with $h[m] = 1$, can be represented as

$$y[m] = x[m] + w[m], \tag{3.16}$$

where $w[m]$ is $\mathcal{CN}(0, N_0)$ and is i.i.d over time. The capacity of this channel is given by [82] [19] [92]

$$C = W\log\left(1 + \frac{P}{N_0 W}\right) \text{ bits/sec} \tag{3.17}$$

where W is the channel capacity, P is the signal power and N_0 is the noise power spectrum. Reliable communication is possible if the transmission rate R is less than C, that is,

$$R < W\log\left(1 + \frac{P}{N_0 W}\right) \tag{3.18}$$

If $\mathcal{E}_\text{b}$ is the energy per bit, then $\mathcal{E}_\text{b} = \frac{P}{C}$. By substituting in (3.17), we obtain

$$\frac{C}{W} = \log\left(1 + \frac{C}{W}\frac{\mathcal{E}_\text{b}}{N_0}\right) \tag{3.19}$$

Consequently,

$$\frac{\mathcal{E}_\text{b}}{N_0} = \frac{2^{C/W} - 1}{C/W} \tag{3.20}$$

This relation is plotted in Fig. 3.1. In the region below the curve in Fig. 3.1, reliable communication is possible and in the region above the curve, reliable communication is not possible. We can also calculate the absolute minimum value of $\mathcal{E}_\text{b}/N_0$ for reliable communication

$$\frac{\mathcal{E}_\text{b}}{N_0} = \ln 2 = 0.693 \sim -1.6 \text{ dB} \tag{3.21}$$

as C/W tends to 0. The curve is thus the performance of an optimal system.

We investigate the optimal performance achievable on a given channel and the techniques to achieve such optimal performance. The AWGN channel is used as a building block to study the capacity of wireless fading channels. For a

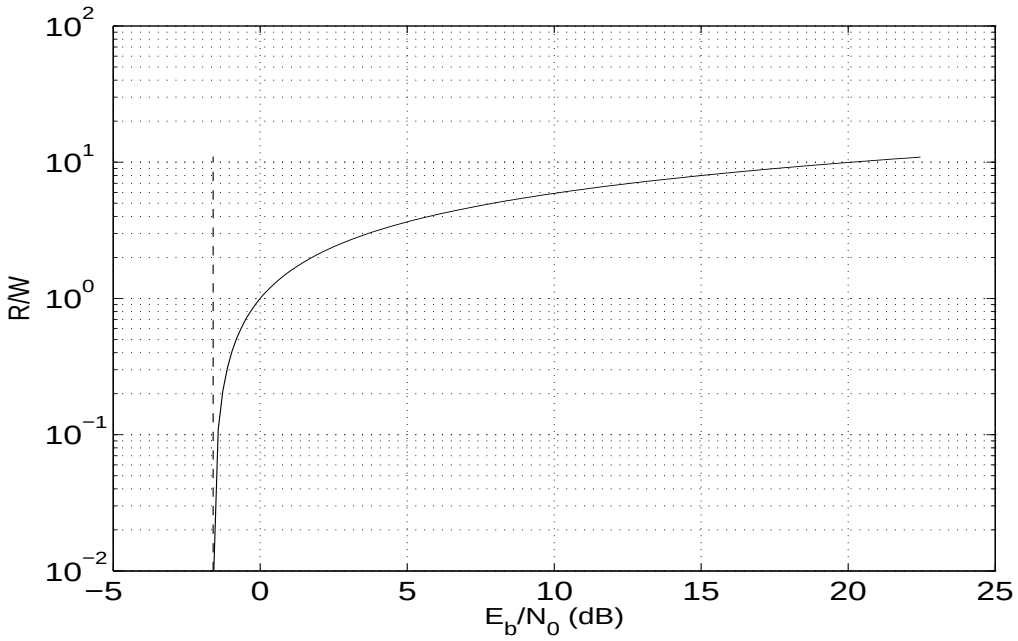


Figure 3.1: Spectral bit rate versus SNR/bit in an optimal system.

given wireless fading channel, we investigate the optimal performance achievable and techniques to achieve such optimal performance just like the the scenario of AWGN channel. Unlike the well established information theory for AWGN channel, there is no single definition of capacity for fading channels that is applicable in all scenarios. We first investigate the simplest scenario, the slowly flat fading channel. In this scenario, the channel gain is random but remains constant in a quasi-static period, i.e., $h[m] = h$ for all m. In this case, the flat fading channel input-output:

$$y[m] = h[m]x[m] + w[m] \tag{3.22}$$

is reduced to a *slow flat fading* input-output:

$$y[m] = hx[m] + w[m], \tag{3.23}$$

with the fading process $\{h[m]\}$ reduced to a random variable $h \sim \mathcal{CN}(0, \sigma^2)$.

The distribution of $r = |h|$ is Rayleigh,

$$f(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \qquad r \geq 0, \tag{3.24}$$

and $x = |h|^2$ is exponentially distributed,

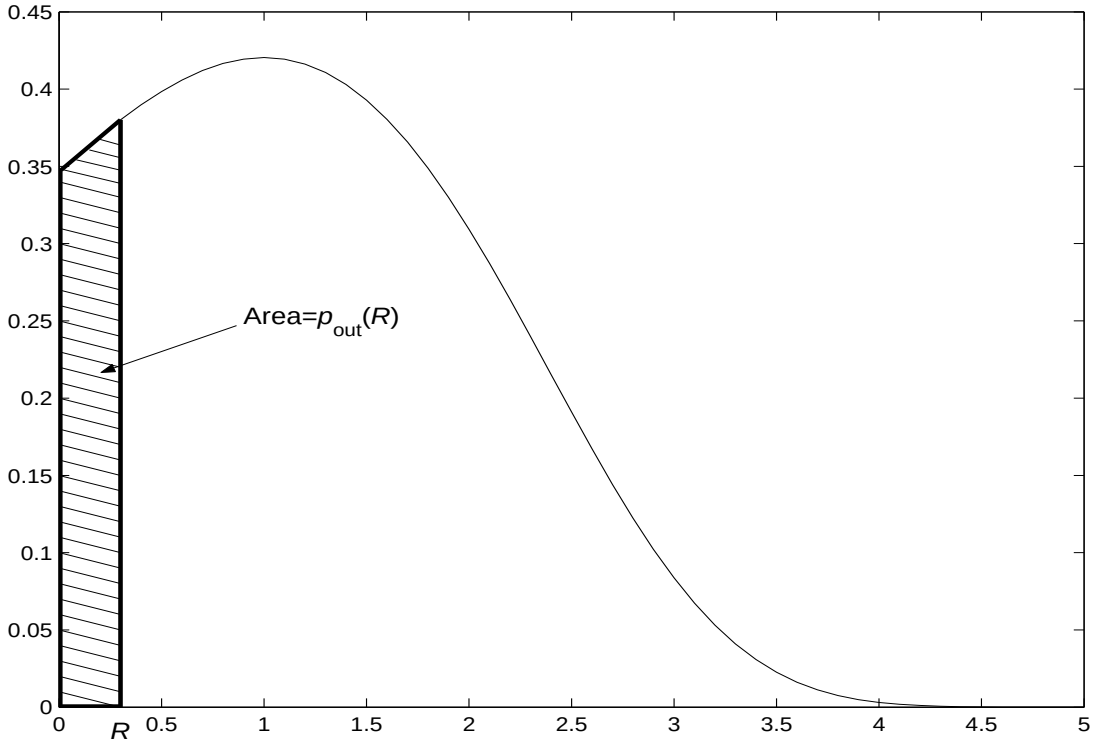


Figure 3.2: Density of $\log(1+|h|^2\mathsf{SNR})$, for Rayleigh fading and $\mathrm{SNR} = 0$ dB. For any target rate R, there is a non-zero outage probability.

$$f(x) = \frac{1}{2\sigma^2} e^{-x/2\sigma^2}, \qquad x \geq 0, \tag{3.25}$$

The capacity of the channel can now be proceeded based on the well-established results for AWGN channel. Let the symbol rate be W Hz, the power constraint is P joules/symbol, and $\mathbb{E}[|h|^2] = 1$ is assumed for normalization, i.e., $h \sim \mathcal{CN}(0,1)$. The noise process $\{w[m]\}$ is i.i.d. $\mathcal{CN}(0, N_0)$. Hence $\mathsf{SNR} := P/N_0$ is the average received SNR.

Conditional on a realization of the channel h, Eq. (3.23) is a scenario of AWGN channel with signal-to-noise ration $|h|^2\mathsf{SNR}$. The maximum rate of reliable communication supported by this channel is $\log(1+|h|^2\mathsf{SNR})$ bits/s/Hz. The rate of communication R can be chosen to be less than the capacity and an arbitrarily small error probability is achieved. As $x = |h|^2$ is random, so is $\log(1+|h|^2\mathsf{SNR})$. The density of this quantity is plotted in Fig. 3.2.

Let $y = \log(1+|h|^2\mathsf{SNR})$ be any possible value of random variable Y. With

$h \sim \mathcal{CN}(0,1)$, the distribution of $x = |h|^2$ is Eq. (3.25) with $\sigma^2 = 1$:

$$f(x) = \frac{1}{2} e^{-x/2}, \qquad x \geq 0. \tag{3.26}$$

The distribution of random variable Y can then be derived as

$$f_Y(y) = \frac{2^y \ln 2}{2\mathsf{SNR}} \exp\{-\frac{2^y - 1}{2\mathsf{SNR}}\}, \qquad y \geq 0. \tag{3.27}$$

Given a targeted transmission rate R, the outage probability is defined as

$$p_{\text{out}}(R) := \mathbb{P}\{\log(1 + |h|^2 \mathsf{SNR} < R\}. \tag{3.28}$$

For the slow fading channel with $h \sim \mathcal{CN}(0,1)$, $p_{\text{out}}(R)$ can be obtained as

$$\begin{aligned} p_{\text{out}}(R) &= \mathbb{P}\{\log(1 + |h|^2 \mathsf{SNR} < R\} \\ &= 1 - \exp\left(-\frac{2^R - 1}{\mathsf{SNR}}\right) \\ &\approx \frac{(2^R - 1)}{\mathsf{SNR}} \quad \text{at high SNR.} \end{aligned} \tag{3.29}$$

The ϵ-outage capacity C_ϵ is defined as the largest rate of transmission R such that the outage probability $p_{\text{out}}(R)$ is less than ϵ. From this definition we can find C_ϵ by solving

$$p_{\text{out}}(R) := \mathbb{P}\{\log(1 + |h|^2 \mathsf{SNR} < R\} = \epsilon. \tag{3.30}$$

The ϵ-outage probability for the slow fading channel with $h \sim \mathcal{CN}(0,1)$ can then be obtained by solving

$$\begin{aligned} p_{\text{out}}(R) &= \mathbb{P}\{\log(1 + |h|^2 \mathsf{SNR} < R\} = \epsilon \\ &= 1 - \exp\left(-\frac{2^R - 1}{\mathsf{SNR}}\right) = \epsilon. \end{aligned} \tag{3.31}$$

The obtained *ϵ-outage capacity* C_ϵ is

$$C_\epsilon = \log(1 - \mathsf{SNR} \cdot \mathsf{ln}(1 - \epsilon)) \tag{3.32}$$

Note the above result is obtained by assuming $h \sim \mathcal{CN}(0,1)$. Let F be the complementary cumulative distribution function of $|h|^2$, i.e., $F(x) = \mathbb{P}\{|h|^2 > x\} = e^{-x}$. We then have $F^{-1}(1-\epsilon) = -\ln(1-\epsilon)$, and Eq. (3.32) can be expressed as

$$C_\epsilon = \log(1 + F^{-1}(1 - \epsilon)) \text{ bits/s/Hz}. \tag{3.33}$$

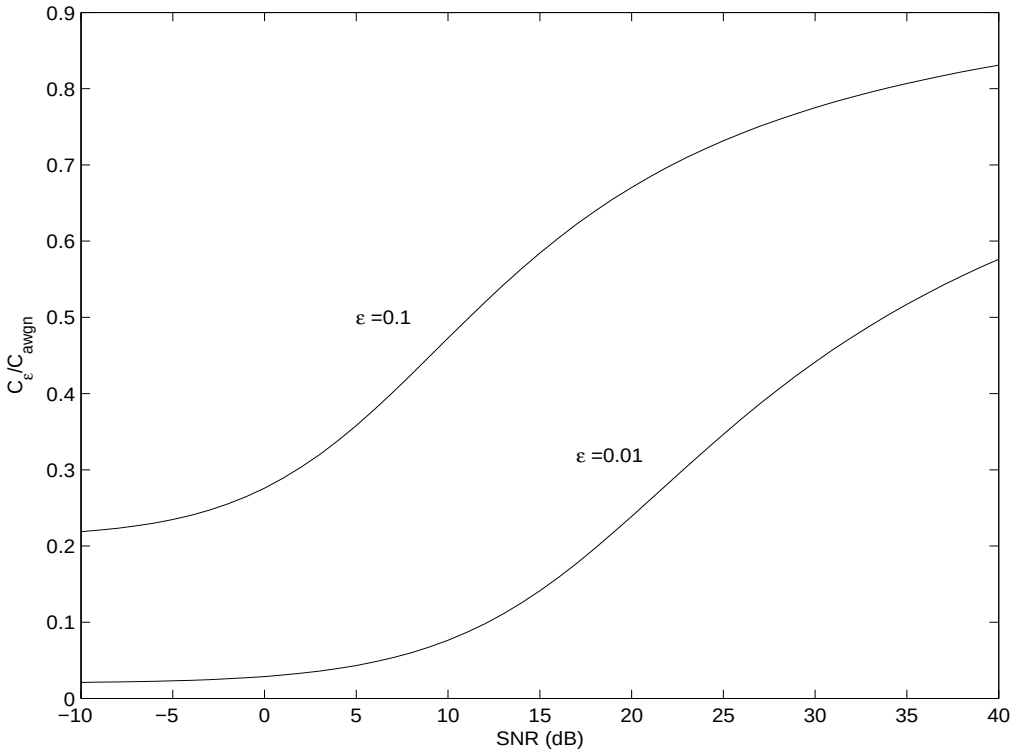


Figure 3.3: ϵ-outage capacity as a fraction of AWGN capacity under Rayleigh fading, for $\epsilon = 0.1$ and $\epsilon = 0.01$.

To obtain some intuitions, Fig. 3.3 plots the ϵ-outage capacity as a function of SNR for the Rayleigh channel. At an outage probability of 0.01, the outage capacity is only 1% of the AWGN capacity. Low and high SNR have different impact on the ϵ-outage capacity.

It should be noted although (3.33) is obtained by assuming a Rayleigh fading channel, it is in fact the general result which can be derived with the distribution of the channel gain h unknown.

3.6 Jakes' Model: from Continuous to Discrete

In this section, we show that the various forms of Jakes' model either continuous or discrete, are essentially the same and can be used interchangeably.

Based on the sampling theory, we now prove that the continuous form of Jakes' model given in [64] is equivalent to its discrete form given in [92].

The continuous-time baseband input-output relation is described as [64]

$$r_l(t) = \sum_{n=1}^{L} c_n(t) s_l\left(t - \frac{n}{W}\right) \tag{3.34}$$

where $L = \lfloor T_{\mathrm{d}} W \rfloor + 1$ is the number of channel taps truncated for the multipath spread T_{d}. The impulse response for the channel can be derived as

$$c(\tau; t) = \sum_{n=1}^{L} c_n(t) \delta(\tau - n/W) \tag{3.35}$$

Let $C(f; t)$ denote the time-variant transfer of $c(\tau; t)$. If $c(\tau; t)$ is modelled as a complex-valued zero-mean Gaussian random process in the t variable, it follow that $C(f; t)$ also has the same statistics. In the case of Jakes' model, the autocorrelation of the time-variant transfer function $C(f; t)$ is given as

$$\begin{aligned} \phi_C(\Delta t) &= \frac{1}{2} E[C^*(f; t) C(f; t + \Delta t)] \\ &= J_0(2\pi f_{\mathrm{D}} \Delta t) \end{aligned} \tag{3.36}$$

where $J(\cdot)$ is the zero-order Bessel function of the first kind and $f_{\mathrm{D}} = v f_{\mathrm{c}}/c$ is the Doppler frequency, where v is the vehicle speed in meters per second (m/s), f_{c} is the carrier frequency, and c is the speed of light (3×10^8 m/s). The Fourier transform of this autocorrelation function gives the Doppler power spectrum:

$$\begin{aligned} S_c(\lambda) &= \int_{-\infty}^{\infty} \phi_C(\Delta t) e^{-\mathrm{j}2\pi\lambda\Delta t} d\Delta t \\ &= J_0(2\pi f_{\mathrm{D}} \Delta t) e^{-\mathrm{j}2\pi\lambda\Delta t} d\Delta t \\ &= \begin{cases} \frac{1}{\pi f_{\mathrm{D}} \sqrt{1 - (f/f_{\mathrm{D}})^2}} & (|f| \le f_{\mathrm{D}}) \\ 0 & (|f| \ge f_{\mathrm{D}}) \end{cases} \end{aligned} \tag{3.37}$$

Sample the autocorrelation function $\phi_C(\Delta t)$ by the sampling frequency $f_s = 1/T_s = W$, where W is the bandwidth of $[f_{\mathrm{c}} - W/2, f_{\mathrm{c}} + W/2]$ around the center frequency f_{c}, we can readily obtain the discrete counterpart of the above Fourier pair based on the sampling theorem. The discrete autocorrelation function $R_0[n]$ is therefore

$$\begin{aligned} R_0[n] &= 2a^2 \pi J_0(n\pi D_{\mathrm{s}}/W) \\ &= 2a^2 \pi J_0(2n\pi f_{\mathrm{D}}/W) \\ &= 2a^2 \pi J_0(2n\pi f_{\mathrm{D}} T_s) \end{aligned} \tag{3.38}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind:

$$J_0(x) := \frac{1}{\pi}\int_0^{\pi} e^{jx\cos\theta}d\theta. \tag{3.39}$$

and $D_\mathrm{s} = 2f_\mathrm{c}v/c$ is the Doppler spread and $f_\mathrm{D} = f_\mathrm{c}v/c$ is the Doppler frequency. And the power spectral density $S(f)$, normalized by W, is given by

$$S(f) = \begin{cases} \frac{4a^2W}{D_\mathrm{s}\sqrt{1-(2fW/D_\mathrm{s})^2}} & -D_\mathrm{s}/(2W) \le f \le +D_\mathrm{s}/(2W) \\ 0 & \text{else.} \end{cases} \tag{3.40}$$

Eq. (3.38) and (3.40) are in fact the autocorrelation function and power spectral density for the tap gain process $\{h_0[m]\}$ in (3.15) respectively.

Define the coherence time T_c to be the value of n/W such that $R_0[n] = 0.05R_0[0]$, then

$$T_\mathrm{c} = \frac{J_0^{-1}(0.05)}{\pi D_\mathrm{s}} \tag{3.41}$$

Clearly, this relation shows that a slowly changing channel has a large coherence time or, equivalently, a small Doppler spread.

3.7 An Efficient Mobile Rayleigh Fading Channel Simulator

3.7.1 Introduction

We have shown that the sequence of random variables $\{h_\ell[m]\}$ (discrete-time filter taps) and $c_n(t)$ (continuous-time filter taps) are generally modelled as complex-valued random process. We now proceed to develop efficient simulation models for these channel taps of the fading channels. To be general, we denote the complex random process under investigation as $\mu(t) = \mu_1(t) + j\mu_2(t)$.

Simulation models for fading channels are extremely important for the development, performance analysis, and test of modern wireless communication systems. The designed fading channel simulator should fit the desired statistical behavior with high precision before analyzing the performance of a new mobile communication system. Accuracy, efficiency, flexibility and ease of implementation are the challenging requirements in designing the simulation models [57] [22].

The characteristic quantities describing the statistics of mobile fading channels are the probability density function (PDF), cumulative distribution function (CDF), the autocorrelation function (ACF), the level-crossing rate (LCR)

and the average duration of fades (ADF). The *level-crossing rate* (LCR) and *average duration of fades* (ADF) are useful for designing error control codes and diversity schemes to be used in mobile communication systems, since it becomes possible to relate the time rate of change of the received signal to the signal level and velocity of the mobile [70].

John I. Smith demonstrated a simple computer program in [86]. We modified this method and developed a new method which promises a greater mathematical tractability. Verification of this new method against the analytic fading channel model is thus necessary and a detailed analysis of the statistical characteristics of this new method is given.

The aim of this section is to design a fading channel simulator based on Smith's method and to analyze the LCR and ADF of this simulation model for Rayleigh fading channels. The results obtained will serve as the foothold for the further investigation of the recently proposed parabolic channel model [78], which will be analyzed in the following sections.

3.7.2 Description of the Analytical Model

The detailed derivation of *Jakes Power spectral density* or *Clarke power spectral density* can be found in [17]. This derivation is briefly reviewed here and some initial simulation results are given.

To derive the Jakes power spectral density, we follow these assumptions [56]:

1. The propagation of the electromagnetic waves takes place in the two-dimensional(horizontal) plane, and the receiver is located in the center of an isotropic scattering area.

2. The angles of arrival α of the waves arriving the receiving antenna are uniformly distributed in the interval $[-\pi, \pi)$.

3. The antenna radiation pattern of the receiving antenna is circular-symmetrical (omnidirectional antenna).

The probability density function of the angles of arrival α is thus given by

$$p_\alpha = \begin{cases} \frac{1}{2\pi}, & \alpha \in [-\pi, \pi), \\ 0, & \text{elsewhere.} \end{cases} \tag{3.42}$$

The Doppler frequencies can then be defined by

$$f = f(\alpha) := f_\mathrm{D} \cos(\alpha), \tag{3.43}$$

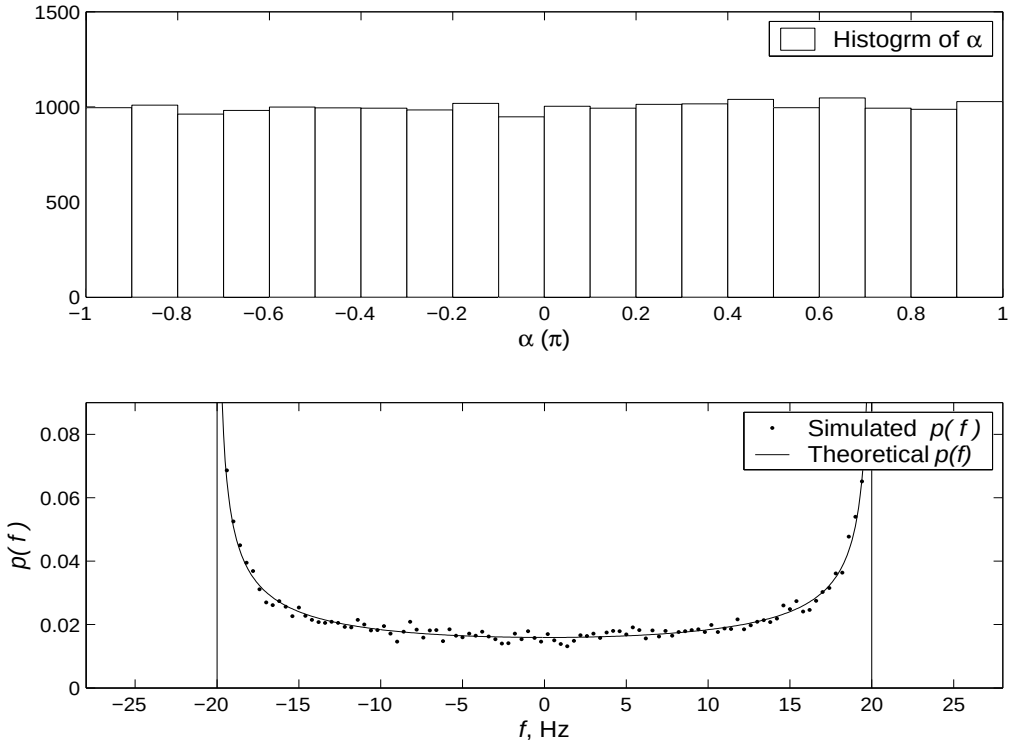


Figure 3.4: Histogram of the angles of arrival α and the pdf of the Doppler frequencies f.

where f_{D} is the Doppler frequency. Obviously f is also a random variable. The probability density function of the Doppler frequencies f, denoted by $p_f(f)$, can be given by [17]

$$p_f(f) = \begin{cases} \frac{1}{\pi f_{\mathrm{D}} \sqrt{1-(f/f_{\mathrm{D}})^2}}, & |f| \leq f_{\mathrm{D}}, \\ 0, & |f| > f_{\mathrm{D}}. \end{cases} \tag{3.44}$$

This function of random variables is simulated and shown in Fig. 3.4 with $f_{\mathrm{D}} = 20$ Hz.

The power spectral density $S_{\mu\mu}(f)$ of the scattered components $\mu(t) = \mu_1(t) + j\mu_2(t)$, received at the receiving antenna, is obviously proportional to $p_f(f)$ of the Doppler frequencies. The following relation holds:

$$S_{uu}(f)df \sim p_f(f)df, \tag{3.45}$$

The relation between the probability density function and the power spectrum density is thus established. Further investigation which takes Eq. (3.44) into account will lead into the *Jakes Power spectral density* or *Clarke power spectral*

density as that of (3.37) (up to a scaling factor):

$$S_{uu}(f) = \begin{cases} \frac{2\sigma_0^2}{\pi f_{\mathrm{D}}\sqrt{1-(f/f_{\mathrm{D}})^2}}, & |f| \le f_{\mathrm{D}}, \\ 0, & |f| > f_{\mathrm{D}}. \end{cases} \tag{3.46}$$

where $\int_{-\infty}^{\infty} S_{uu}(f)df = 2\sigma_0^2$, and $2\sigma_0^2$ is the power of the scattered components $\mu(t) = \mu_1(t) + j\mu_2(t)$.

The autocorrelation function (ACF) $r_{\mu\mu}(\tau)$ of the scattered component $\mu(t) = \mu_1(t) + j\mu_2(t)$ can be obtained by taking the inverse Fourier transform of the Jakes power spectral density of Eq. (3.46):

$$r_{\mu\mu}(\tau) = 2\sigma_0^2 J_0(2\pi f_{\mathrm{D}}\tau) \tag{3.47}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind.

3.7.3 Frequency Zero-Appending for Time Interpolation

Zero-padding is a well-known topic in signal analysis, normally associated with upsampling or unifying signals' lengths in circular convolution [66].

We use zero-padding in the frequency domain for the purpose of increasing the resolution in the time domain. This approach is defined by simply appending a number of zeros to both ends of the spectrum, hence we use the name *zero-appending*. Based on the duality property of the Fourier Transform, we can state that zero-appending in the frequency domain corresponds to ideal interpolation in the time domain. We demonstrate this theorem using a simple example. The sequence $x(n)$ is obtained by sampling a sinc function $x(t) = \mathrm{sinc}(t/T)$ with $T = 1$ s, $T_s = 0.5$ s, and the total time duration is from $-L = 5$ s to $L = 5$ s, $N = 2L/T_s$, therefore $N = 20$, that is, both $x(n)$ and its DFT $X(k)$ are 20 point sequence. We exploit again the time-domain/frequency-domain duality of DFT, that is, the time duration of the signal in time domain will determine the frequency resolution in frequency domain, while the frequency span in frequency domain will determine the time resolution in time domain. We append a total of $M - N$ zeros to the spectrum, divided into two halves, one half on each end of the spectrum of $X(k)$. An IDFT is then performed on $X(k)$ to provide a finer interpolation in the time domain. Fig. 3.5 shows that a total of $M - N = 30$ zeros are padded in the frequency domain to provide sufficient detail to yield a good picture of $x(n)$.

3.7.4 Design and Analysis of the Rayleigh Fading Channel

Smith demonstrated a computer simulator of Rayleigh fading channel in [86]. A detailed presentation of this method can also be found in [70]. This method

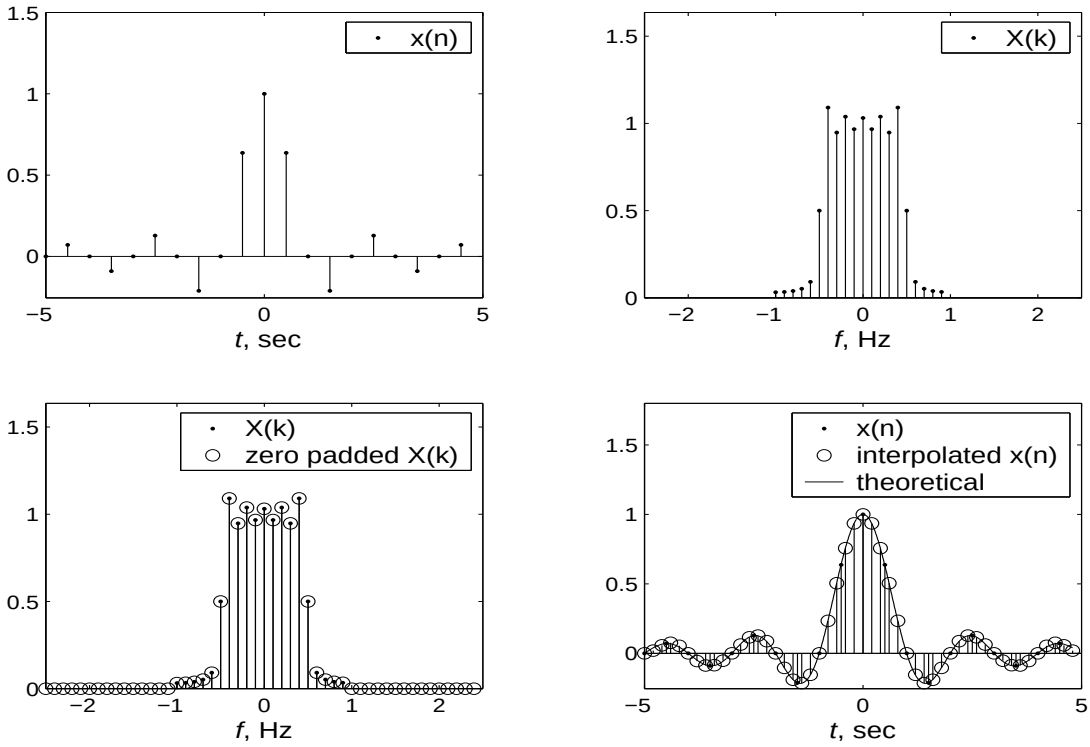


Figure 3.5: Zero-appending $X(k)$ to provide finer interpolation for a time sequence $x(n)$ with duration $2L = 10$ s and sampling period $T_s = 0.5$ s. After zero-appending, time duration remains unchanged, but the new sampling period is $T_s = 0.2$ s, thus a better resolution is achieved for $x(n)$.

uses a complex Gaussian random number generator to produce a line spectrum with complex weights in the positive frequency band. The maximum frequency component of the line spectrum is f_{D}. The negative frequency components are constructed by simply conjugating the complex Gaussian values obtained for the positive frequencies. The random valued line spectrum is then multiplied by a discrete frequency presentation of $\sqrt{S_{uu}(f)}$ having the same number of points as the noise source. An IFFT is then performed on the resulting frequency domain signal to get two time series. The square root of the summed squared of the two time series is the Rayleigh fading signal with proper Doppler spread and time correlation. Note that this method requires a truncation of the spectrum because Equation (3.46) approaches infinity at the baseband edge. We present here a new method which avoids this truncation and gives better resolution by zero-appending. Moreover, we will show that this method promises greater mathematical tractability.

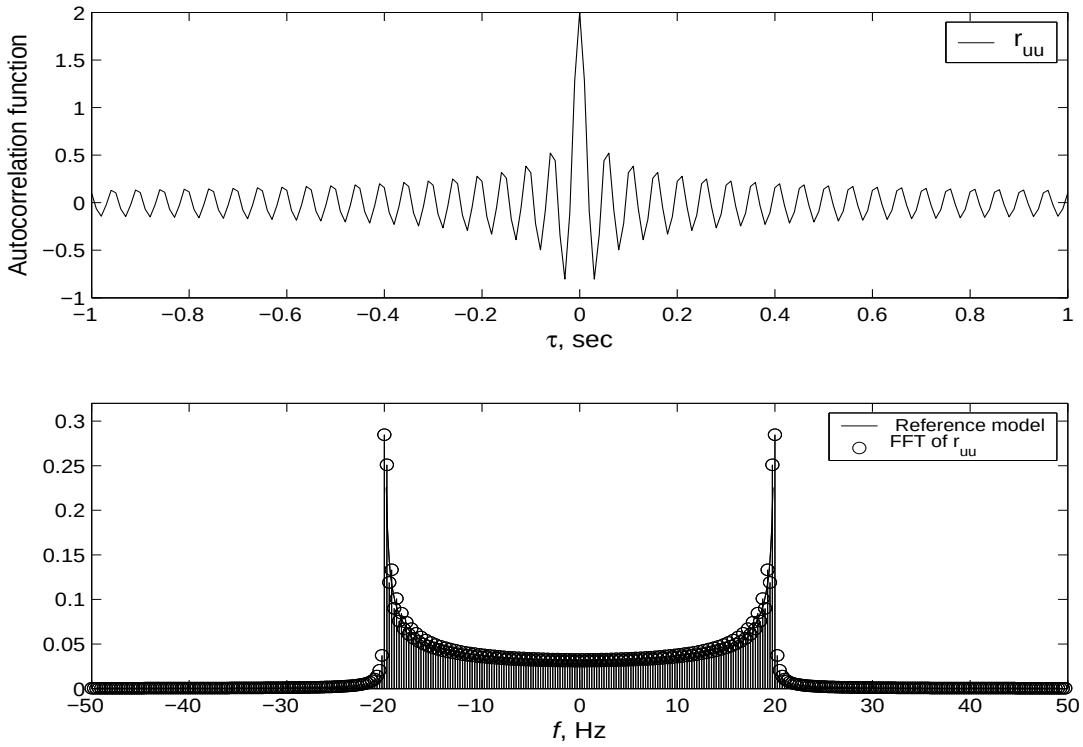


Figure 3.6: The ACFs of the reference model and its FFT ($f_{\mathrm{D}} = 20$ Hz, $\sigma_0^2 = 1$, $K = 5$, $\Delta\tau = 1/(Kf_{\mathrm{D}}) = 0.01$ s, $L = 2$ s).

To implement this simulator, the following steps are used:

1. In Eq. (3.47), specify the power σ_0^2, maximal Doppler spectral f_{D}, and let $\Delta\tau = 1/(Kf_{\mathrm{D}})$, here $K \geq 1$ is a constant. Define the range of τ from $-L$ to L. Compute the N point $r_{\mu\mu}$.

2. To perform an FFT on $r_{\mu\mu}$ to get N point $S_{\mu\mu}$.

3. Generate N complex Gaussian random variables, normalize the total power to 1.

4. Multiple the N complex Gaussian random variable by the obtained Doppler spectrum $\sqrt{S_{\mu\mu}}$ from step 2.

5. Perform an IFFT on the resulting frequency domain signal to obtain the N point scattered component $u = u_1 + ju_2$. The envelope of obtained u is the simulated Rayleigh fading signal with the proper Doppler spread.

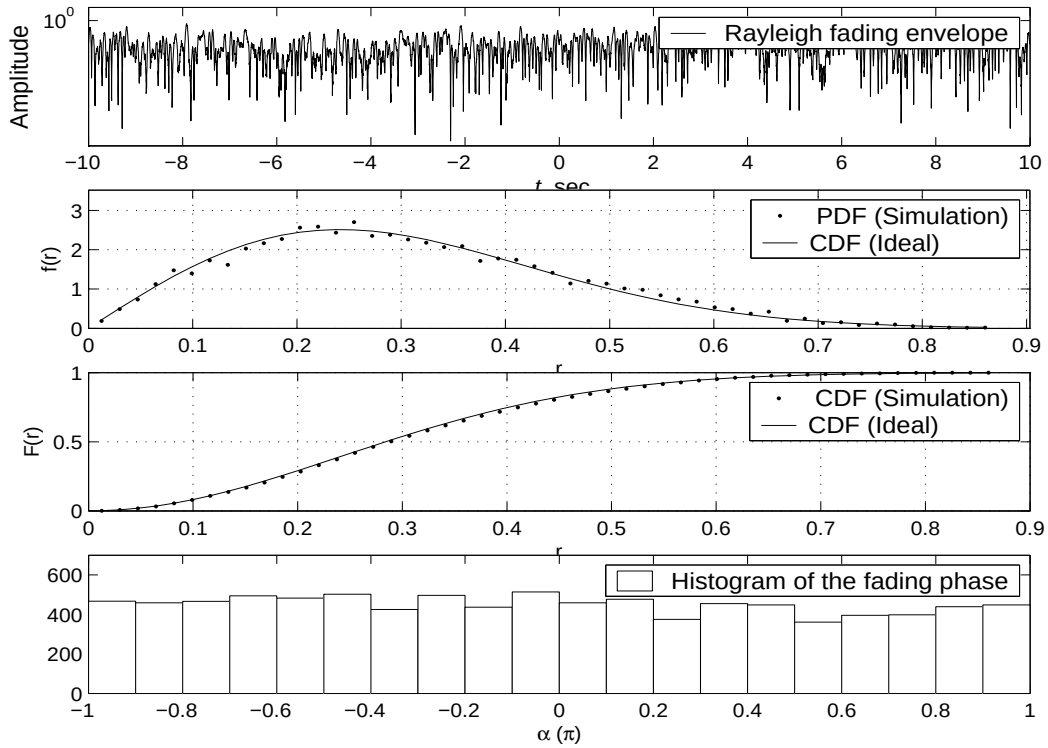


Figure 3.7: The simulated fading envelop, its PDF and CDF, and the histogram of the fading phase ($f_{\rm D} = 20$ Hz, $\sigma_0^2 = 1$, $K = 10$, $\Delta t = 1/(Kf_{\rm D}) = 0.005$ s, $L = 2$ s).

Fig. 3.6 shows the ACF and its FFT with $f_{\rm D} = 20$ Hz, $\sigma_0^2 = 1$, $K = 5$, $\Delta\tau = 1/(Kf_{\rm D}) = 0.01$ s, $L = 2$ s. Based on the discussion in the previous section, the parameters of the fading process u can be easily calculated. The variance of u, $\sigma_u^2 = (2L)\sigma_0^2$, the duration of the fading signal $2L = 4$ s, and the time resolution of the fading signal $\Delta t = \Delta\tau = 0.01$ s.

We now decide that the resolution is not satisfactory with $K = 5$ and let $K = 10$. The time resolution is thus doubled, $\Delta t = 0.005$ s. The envelope of the simulated $\mu(t)$ is shown in Fig. 3.7. The PDF of the envelope is Rayleigh distributed, and the phase of $\mu(t)$ is uniformly distributed. The plot also shows the CDF of the envelope. The ACF of $\mu(t)$ is plotted in Fig. 3.8. Theoretical values are also plotted for the PDF, CDF and the ACF for the purpose of verification.

The advantages of this method are obvious. First, it is easy to implement. Second, the truncation of the spectrum is avoided. Third, a better resolution can be achieved by increasing the constant K, which means effectively appending more zeros to the Doppler spectrum. Moreover, the variance and the ACF

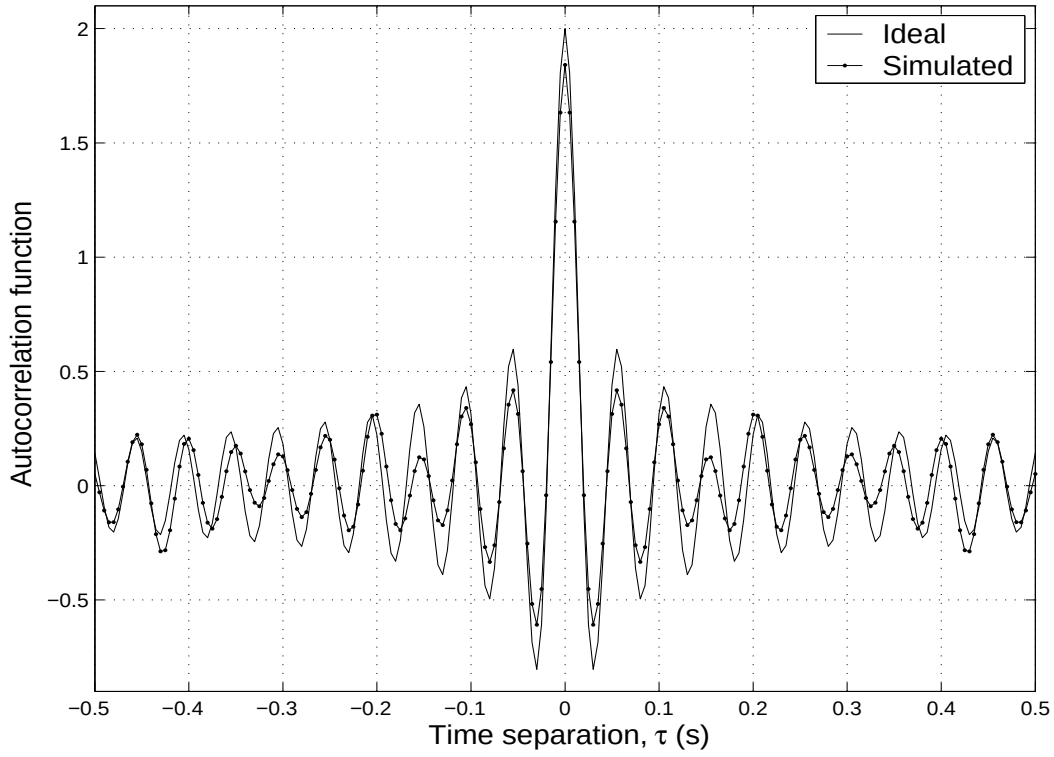


Figure 3.8: The ACFs of the reference model and the simulation model ($f_{\mathrm{D}} = 20$ Hz, $\sigma_0^2 = 1$, $K = 10$, $\Delta t = 1/(Kf_{\mathrm{D}}) = 0.005$ s, $L = 2$ s).

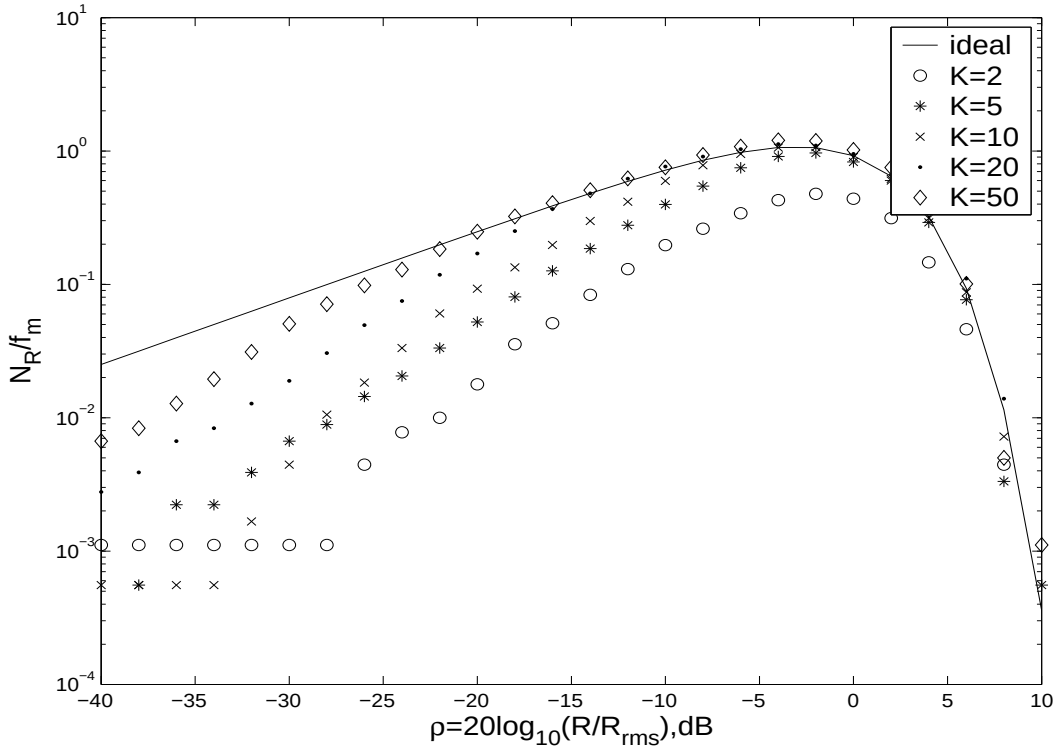


Figure 3.9: Normalized level crossing rates of the envelopes. Simulation values obtained at $f_{\mathrm{D}} = 90$ Hz, $L = 10$ s for $K = 2, 5, 10, 20, 50$.

of the simulated u are simply related to the power σ_0^2 and the theoretical ACF specified in step 1: both are simply scaled down by the time duration $2L$ defined in step 1. Based on the above discussion and the previous section, it is obvious that all the properties of the simulated Rayleigh fading signal can be specified in step 1, e.g., power, Doppler spread, time resolution, and time duration.

Apart from the probability density function, the cumulative distribution function, the autocorrelation function, and other statistical quantities that characterize the mobile fading channels like *level crossing rate* (LCR) and *average duration of fades* (ADF), are all important for the design of the channel simulator and mobile communication systems.

The level crossing rate, N_R, is defined as the expected rate at which the envelope crosses a specified signal level, R, in the positive direction. In general, it is given by [41]

$$N_R = \int_0^\infty \dot{r} p(R, \dot{r}) d\dot{r} = \sqrt{2\pi} f_{\mathrm{D}} \rho e^{-\rho^2}, \tag{3.48}$$

where the dot indicates the time derivative and $p(R, \dot{r})$ is the joint density

function of r and $\dot{r}$ at $r = R$. And $\rho = R/R_{rms}$ is the value of the specified level R, normalized to the local rms amplitude of the fading envelope. We simulate the N_R for different K at $f_{\mathrm{D}} = 90$ Hz, $L = 10$ s and plot the results to compare with the ideal situation. From Fig. 3.9, the effect of zero-appending is significant. In this example, curves with $K > 10$ give satisfactory results which are close enough to the ideal curve.

The *average duration of fades*, $\bar{\tau}$, is defined as the average period of time for which the received signal is below a specified level R. Let τ_i be the duration of the ith fade, then the average duration of fade for a total time interval of length T is $\bar{\tau} = \sum \tau_i/(N_R T)$. Like level crossing rate N_R, the average duration of fade can be also expressed as a function of ρ and f_{D} as [41]

$$\bar{\tau} = e^{\rho^2 - 1}/(\rho f_{\mathrm{D}} \sqrt{2\pi}) \tag{3.49}$$

Fig. 3.10 shows the simulation of the average duration of fade for different K at $f_{\mathrm{D}} = 90$ Hz, $L = 10$ s and the theoretical curve. Like the level crossing rate, the effect of zero-appending is significant. Again, curves with $K > 10$ give satisfactory results which are close enough to the ideal curve. Simulation also shows that K should not be arbitrarily large with the fixed time duration of $2L$. In the example, overshoot is observed when $K > 50$ for both the level crossing rate and the average duration of fades with $L = 10$ s.

3.7.5 Conclusions

In this section, we present a method to accurately and efficiently simulate Rayleigh fading channels. The significance of this method is its ease of implementation and greater mathematical tractability of the model parameters. Some important statistics especially the LCR and ADF of the model are investigated for the verification against the analytical model. Simulation results show that the channel simulator accurately reproduces all of the important statistical properties, such as the probability density function (PDF), autocorrelation (ACF), level crossing rate (LCR) and the average duration of fades (ADF).

3.8 A Frequency-Selective Mobile Radio Channel with Hyperbolically Distributed Scatterers

In the previous section, we have shown an efficient way to simulate the classical Jakes or Clarke's model. Classical models provide information about signal power level distributions and Doppler shifts of the received signals. Modern

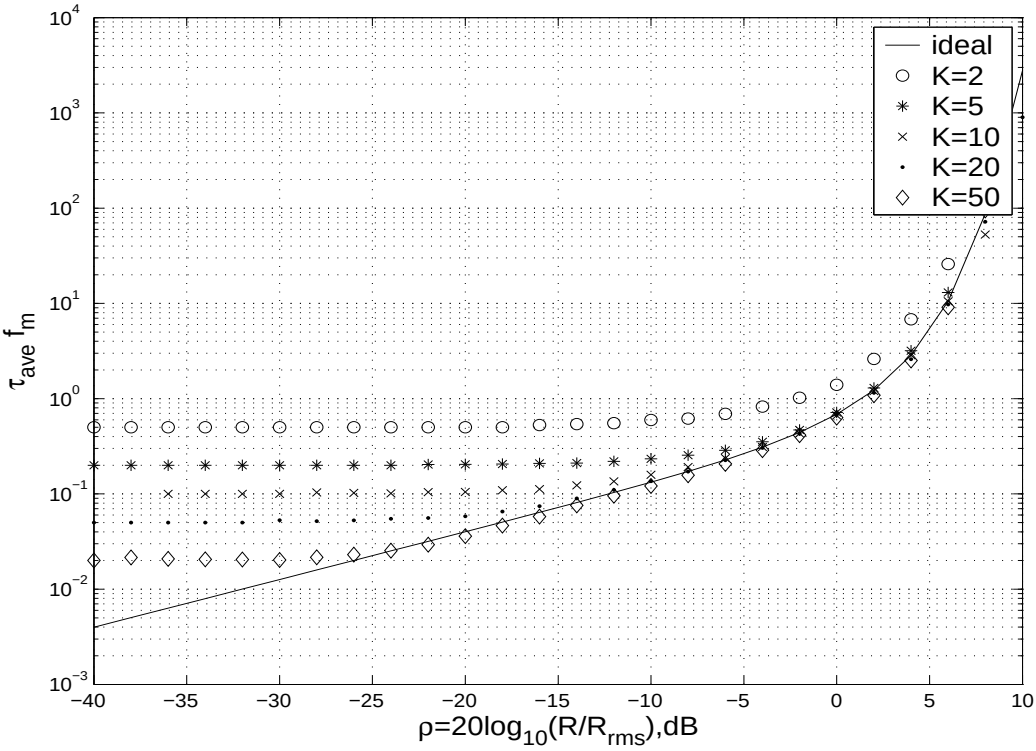


Figure 3.10: Normalized durations of fade of the envelopes. Simulation values obtained at $f_{\mathrm{D}} = 90$ Hz, $L = 10$ s for $K = 2, 5, 10, 20, 50$.

spatial channel models build on classical understanding of fading and Doppler spread and incorporate additional information such as angle of arrival (AOA), time delay spread and array geometries. In this section we investigate a spatial channel model and construct a frequency-selective geometrical and time-variant wireless channel model with hyperbolically distributed scatterers for a macrocell mobile environment. We verify this model against the statistical properties such as the level-crossing rate (LCR), the average duration of fades (ADF), the probability density function (PDF), the cumulative distribution function (CDF), and the autocorrelation functions (ACF) of this model. Simulation results are verified against the analytical Wide Sense Stationary Uncorrelated Scattering (WSSUS) channel model.

3.8.1 Introduction

In [78], a space-time geometrical based hyperbolically distributed scatterers (GBHDS) model for a macrocell mobile environment was proposed. The com-

bination of stochastic and geometrical assumptions results in a mathematically tractable and computationally efficient channel model. This model provides the power of each path, the time-of-arrival (TOA), and the direction-of-arrival (DOA) of the multipath component as well as the fading effect. The model enables the simulation of downlink beamforming as well as space diversity concepts and handles both spatially narrowband and wideband signals.

Verification of the statistical properties for this newly proposed GBHDS model against the analytic fading channel model will be given using the same methods presented in the previous section. The aim of this work is to extend previous results and to construct a frequency-selective GBHDS mobile channel based on previous work [78] [79] [80] [81].

3.8.2 The GBHDS Channel Model

In this section we provide a general description for the space-time geometrical-based hyperbolically distributed scatterers (GBHDS) model [78]. This model combined a scalar stochastic fading model for the local scatterers with the geometrical hyperbolic model proposed in [79, 80] for the distribution of the dominant scatterers. The model in [78] assumes that the scatterers are arranged circularly around the mobile, with the distances between 1) the mobile and the local scatterers and 2) the local and dominant scatterers, both being distributed hyperbolically according to an inverse-cosh-squared distribution. This model provides directional information as well as concerning with mobility issue. Fig. 3.11 shows the geometry for the GBHDS model. The angle of departure ψ_{l_k} is uniformly distributed in the interval [0,2π]. The angle θ_{l_k} is the direction of arrival at the base station, while D denotes the distance between the base station and the mobile station. The mobile is located at the origin. This model has the following assumptions [78]

- The scatterers were arranged circularly around the mobile, with the distance between the mobile and the local scatterers r_{l_k} and the distance between the local and dominant scatterers R_{l_k} are both distributed hyperbolically.
- Signals received at the base station are plane waves propagating along the horizon (there is no vertical component to the signal propagation).
- Scatterers are omnidirectional re-radiating elements.
- The scatterers have identical scattering coefficients.
- The macrocell antenna heights are relatively high and there is no signal scattering from locations near the base station.

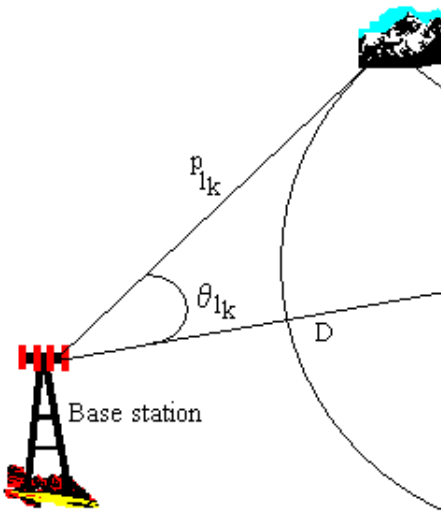


Figure 3.11: Geometry of the space-time hyperbolic model.

The probability density functions (pdf) of the distances r_{l_k} and R_{l_k} for the k^{th} user in multiuser environment are given by [78]

$$f_{r_{l_k}}(r_{l_k}) = \frac{a_1}{\tanh(a_1 R_{ls})\cosh^{\,2}(a_1\, r_{l_k})} \qquad 0 \le r_{l_k} \le R_{ls} \tag{3.50}$$

and

$$f_{R_{l_k}}(R_{l_k}) = \frac{a_2}{\tanh(a_2 R_{ds})\cosh^{\,2}(a_2\, R_{l_k})} \qquad 0 \le R_{l_k} \le R_{ds} \tag{3.51}$$

and

where R_{ls} is the radius of the circle enclosing the local scatterers, and R_{ds} is the radius of the dominant scatterers circle. The applicable values of a_1 and a_2 lie in the interval (0,1). From the spatial probability density functions of the scatterers in (3.50) and (3.51) we can determine the the DOA, TOA, and the signal amplitude. A comprehensive study of these models (at theoretical and simulation levels) as well as their validation with practical data have been considered. They proved to be more realistic than other models in the literature when tested against practical data [80, 81].

3.8.3 Statistical Properties of the GBHDS channel model

To verify the statistical properties of the GBHDS channel, we first plot the envelope of the Rayleigh process $\mu(t)$, shown in Fig. 3.12. The elementary properties of $\mu(t)$ such as the the PDF, CDF and the phase of the Rayleigh process $\mu(t)$ are plotted in Fig. 3.13. We simulated $\mu(kT)$ with a sampling interval of $T = 0.5 \times 10^{-4}$s and sample number of 8000. The remaining model parameters were maximum doppler frequency $f_{\mathrm{D}} = 80$ Hz, the carrier frequency $f_{\mathrm{c}} = 900$ MHz. The autocorrelation (ACF) is also plotted in Fig. 3.14 for the purpose of verification.

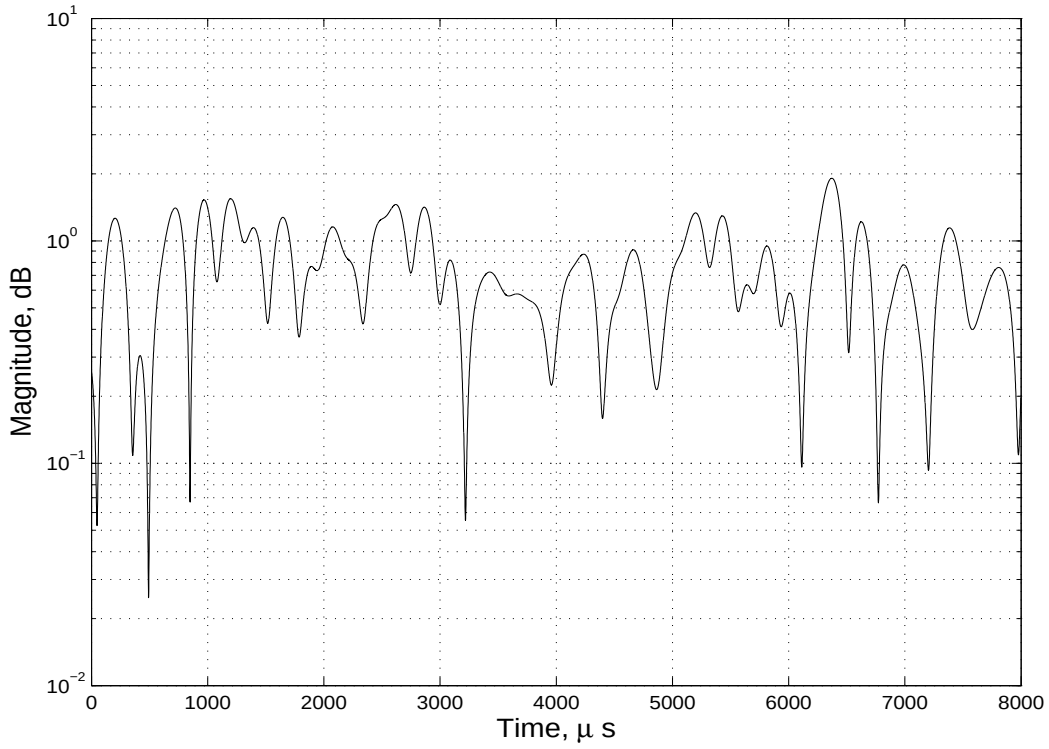


Figure 3.12: The Rayleigh envelope of the simulated GBHDS channel.

The level crossing rate is plotted in Fig. 3.15 which shows that the GBHDS model give satisfactory results which are close enough to the ideal curve.

The average duration of fades is plotted in Fig. 3.16 which shows the simulated average duration of fade for the above mentioned parameters with the theoretical curve. Again, the simulated curve gives satisfactory results which are close enough to the ideal curve.

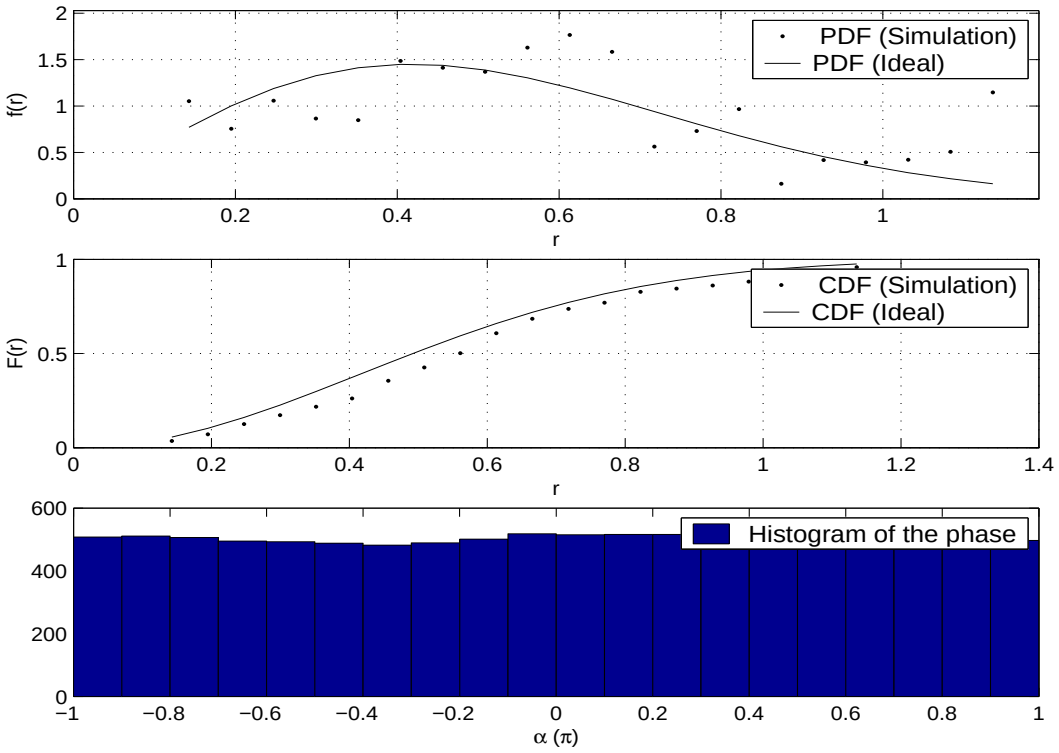


Figure 3.13: The PDF and CDF of $\mu(t)$, and the histogram of the fading phase.

3.8.4 Description of the Analytical Model

For digital communication systems, propagation delay differences from multipath cannot be ignored when the transmission data rate becomes higher or symbol period becomes shorter. A multipath channel with differential path delays introduces amplitude and phase distortion is classified as *frequency-selective fading channel*.

The baseband input signal $s_l(t)$ and the output signal $r_l(t)$ of a time-variant frequency-selective fading channel can be expressed as [64] [89]

$$r_l(t) = \sum_{n=-\infty}^{\infty} c_n(t) s_l(t - n/W) \tag{3.52}$$

where W is the bandwidth covered by the band-pass signal. Therefor, the time-variant frequency-selective channel can be modelled or represented as a tapped delay line with tap spacing $1/W$ and tap weight coefficients $\{c_n(t)\}$, Hence, with the truncation by multipath spread T_{d}, the time-varying impulse response

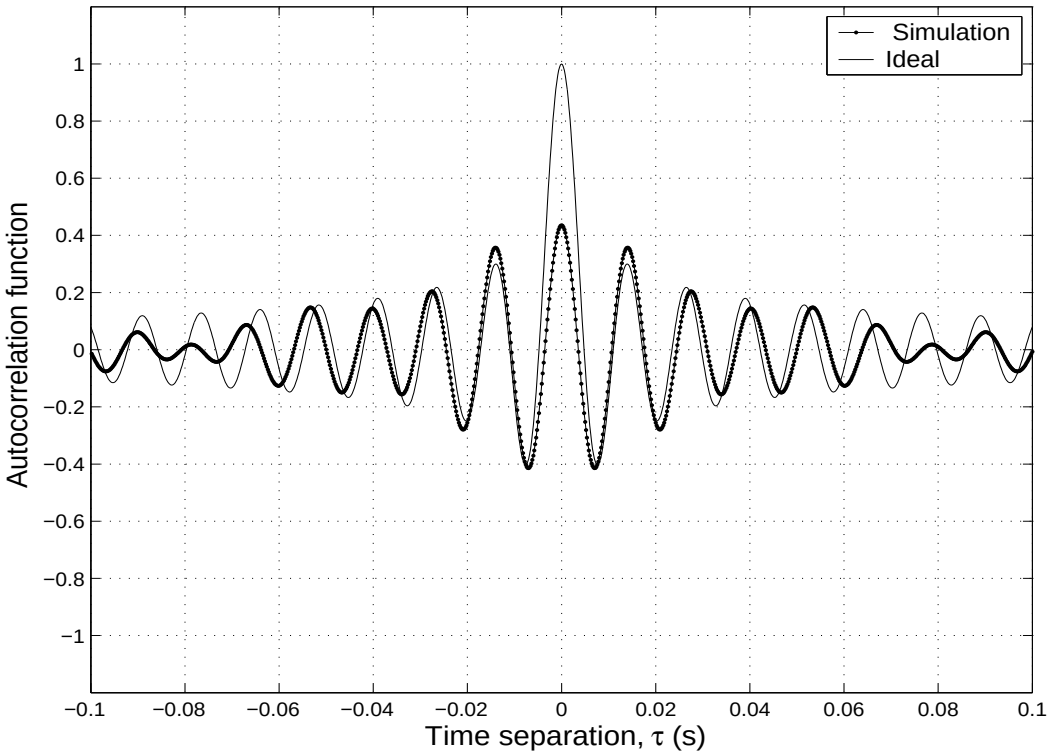


Figure 3.14: The ACF of the reference model and the simulation model.

for the channel can be deduced from Eq. (3.52) as

$$c(\tau;t) = \sum_{n=1}^{L} c_n(t)\delta(\tau - n/W) \tag{3.53}$$

where $L = \lfloor T_{\mathrm{d}} W \rfloor + 1$ is the number of taps, and $c_n(t)$ and $\tau_n = n/W$ are the complex gains and path delays associated with the taps.

The time-variant tap weights $\{c_n(t)\}$ are complex-valued stationary random processes with Doppler power spectrum $S_c(\lambda)$. There are four types of $S_c(\lambda)$ specified by COST207 for different path delays [56] [89]. One of the four types, *Jakes* or *Classical*, has been investigated in a previous section. The Jakes power spectral density only occurs in the case of very short propagation delays (≤ 0.5 μs). The power spectral density of the taps for longer delays may be modelled as *GAUS1* or *GAUS2* (Sum of two Gaussian functions). The power spectral density for shortest path is modelled as *Rice*.

We further assume a Wide Sense Stationary Uncorrelated Scattering (WSSUS) channel introduced by [6]. Due to the simplicity of WSSUS channel models, they are of great practical importance and are nowadays almost exclusively

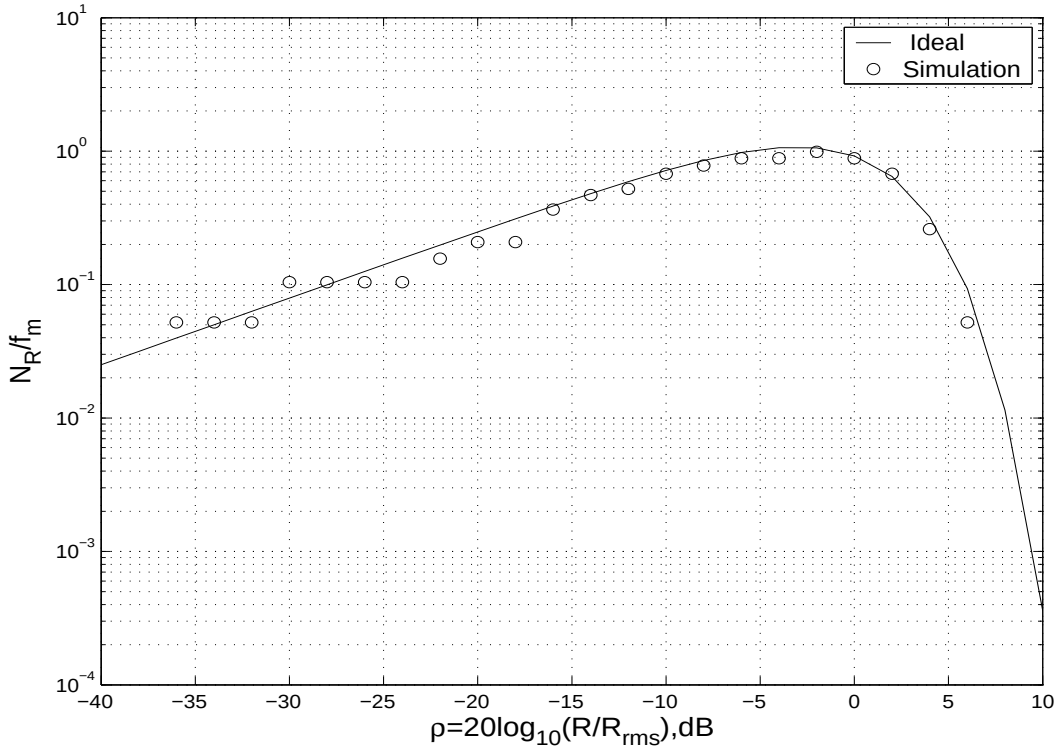


Figure 3.15: Normalized level crossing rates of the envelopes.

employed for modelling frequency-selective mobile radio channels [56]. For the WSSUS channel, the *time-frequency correlation function* $\phi_C(\Delta f; \Delta t)$ or the *scattering function* $S(\tau; \lambda)$ solely characterizes its statistical properties. Another common assumption is separability of scattering function. We investigate a mobile radio channel model with separable scattering function to facilitate the analysis and simulation and to develop some intuitions.

Assume a separable scattering function defined as:

$$S(\tau; \lambda) = \frac{1}{\sigma_0^2} S_c(\lambda) \phi_C(\tau) \tag{3.54}$$

where $\sigma_0^2 = \phi_C(0; 0)$. Then the multipath intensity profile is exponential,

$$\phi_C(\tau) = \begin{cases} \frac{\sigma_0^2}{\tau_0} e^{-\frac{\tau}{\tau_0}}, & \tau \geq 0, \\ 0, & \tau < 0. \end{cases} \tag{3.55}$$

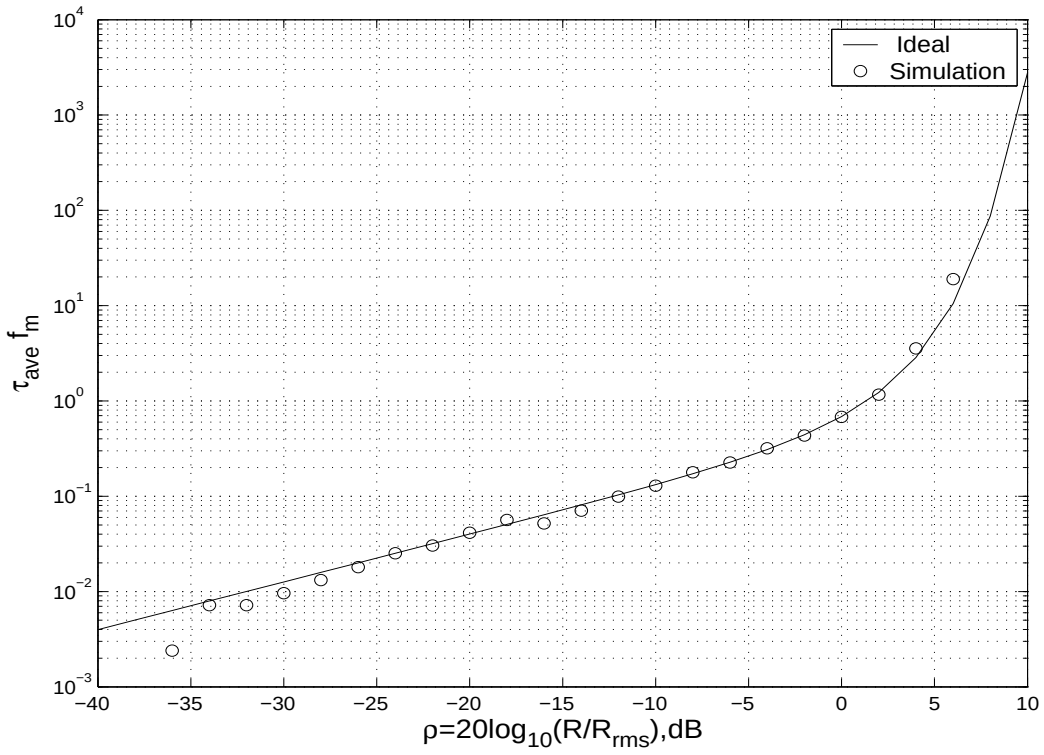


Figure 3.16: Normalized durations of fade of the envelopes.

and the Doppler power spectrum is given by

$$S_c(\lambda) = \begin{cases} \frac{\sigma_0^2}{\pi f_{\rm D}\sqrt{1-(f/f_{\rm D})^2}}, & |f| \le f_{\rm D}, \\ 0, & |f| > f_{\rm D}. \end{cases} \tag{3.56}$$

The associated time-frequency correlation function $\phi_C(\Delta f; \Delta t)$ can be determined by the double Fourier transform of $S_c(\tau; \lambda)$

$$\begin{aligned} \phi_C(\Delta f; \Delta t) &= \int\int S(\tau;\lambda) e^{-{\rm j}2\pi\Delta f\tau} e^{{\rm j}2\pi\Delta t\lambda} d\tau d\lambda \\ &= \frac{\sigma_0^2}{1+{\rm j}2\pi\Delta f\tau_0} \cdot J_0(2\pi f_{\rm D}\Delta t) \end{aligned} \tag{3.57}$$

For the purpose of demonstration, we let $\sigma_0^2 = 0.5$, $\tau_0 = 0.05$ s, and $f_{\rm D} = 5$ Hz, and plot the absolute value of time-frequency correlation function $|\phi_C(\Delta f; \Delta t)|$ and the scattering function $S_c(\tau; \lambda)$ as shown in Fig. 3.17 and Fig. 3.18, respectively. The significance of scattering function or time-frequency correlation

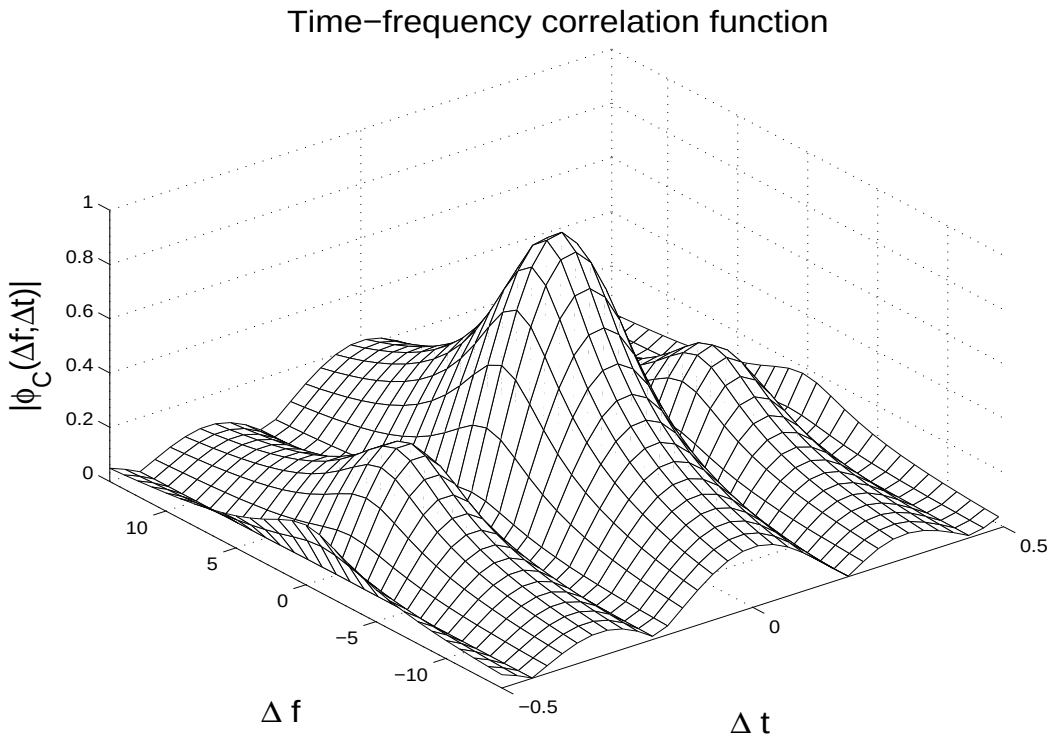


Figure 3.17: Time-frequency correlation function $|\phi_C(\Delta f; \Delta t)|$ with $\sigma_0^2 = 0.5$, $\tau_0 = 0.05$ s, and $f_{\rm D} = 5$ Hz.

function is that important parameters such as the multipath spread $T_{\rm d}$ can be derived.

The above discussion will serve as the basic principle of our simulation for the frequency-selective mobile channels. In simulation, we take non-ideality into account. For example, we may use the multipath intensify profile (or *power delay profile*) of the typical urban defined (discrete in value and approximately exponential) by COST207 instead of Eq. (3.55), and the Doppler power spectrum may be a combination of the four types by COST207 mentioned previously.

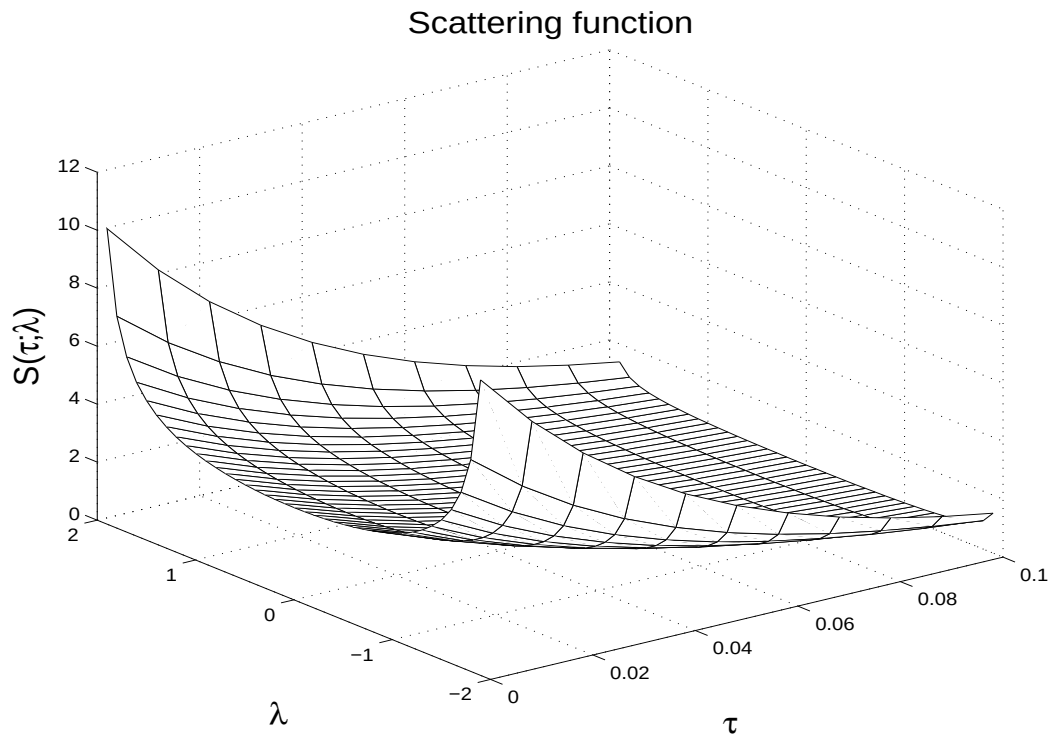


Figure 3.18: Scattering function $S(\tau;\lambda)$ with $\sigma_0^2 = 0.5$, $\tau_0 = 0.05$ s, and $f_\mathrm{D} = 5$ Hz.

3.8.5 A Frequency-Selective GBHDS Channel Model

Let $T = 1/W$ be the transmission symbol period. If

$$T_{\mathrm{d}} \ll T \tag{3.58}$$

holds, then the frequency-selectiveness or time dispersion as an effect of propagation delay differences can be ignored.

Based on the discussion in the previous sections, we develop a frequency-selective GBHDS mobile channel model. We simulate the Typical Urban (TU) (non-hilly) channel defined by COST207 [89] with all the time-variant taps $\{c_n(t)\}$ experience Doppler spectrum generated from GBHDS model. The simulation can be readily extended to a combination of taps with different types of Doppler spectra.

The power delay profile simulated is shown as in the following table:

Delay μs	Fractional Power	Doppler Category
0.0	0.092	Jakes
0.1	0.115	Jakes
0.3	0.231	Jakes
0.5	0.127	Jakes
0.8	0.115	Jakes
1.1	0.074	Jakes
1.3	0.046	Jakes
1.7	0.074	Jakes
2.3	0.051	Jakes
3.1	0.032	Jakes
3.2	0.018	Jakes
5.0	0.025	Jakes

The simulation parameters for the GBHDS model are given as follows: the distance between mobile and scatterer is 1000 m, the light seed is 3×10^8 m/s, the carrier frequency is 900 MHz, the maximum Doppler frequency is 90 Hz.

The statistical properties of the time-variant tap weights $\{c_n(t)\}$ verified are the probability density function, the cumulative distribution function, the autocorrelation function, the *level crossing rate* (LCR) and *average duration of fades* (ADF).

Fig. 3.19 shows the input BPSK signal and the output BPSK signal with $T_{\mathrm{d}} = 5$ μs and $T = 20$ μs. In this case of $T_{\mathrm{d}} \ll T$, the time dispersion caused by the frequency-selective GBHDS mobile channel is negligible. As shown in

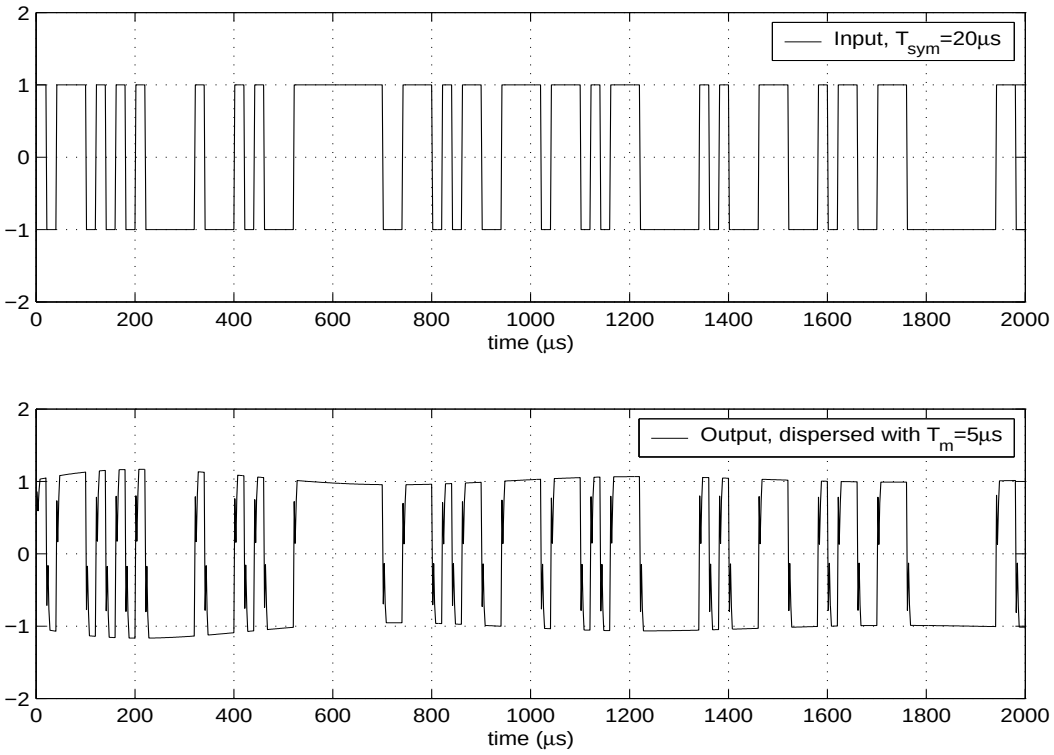


Figure 3.19: BPSK of symbol with period $T = 20$ μs, input (top), and output (bottom), dispersed by channel with multipath spread $T_{\rm d} = 5$ μs.

Fig. 3.20, with $T_{\rm d} = 5$ μs and $T = 4$ μs, the time dispersion by the frequency-selective GBHDS mobile channel can no longer be ignored as the channel introduces amplitude and phase distortion into the signal.

3.8.6 Conclusion

In this section, we simulate a frequency-selective GBHDS channel. The simulated model is verified against analytical model. Important statistical characteristics such as amplitude and phase probability density function (pdf), also the higher-order statistics such as the level crossing rate (LCR) and the average duration of fades (ADF) are verified when constructing the frequency-selective GBHDS channel. The results provide more support to the GBHDS model and will be useful for simulating other practical channels, such as MIMO channels, and space-time-selective mobile fading channels.

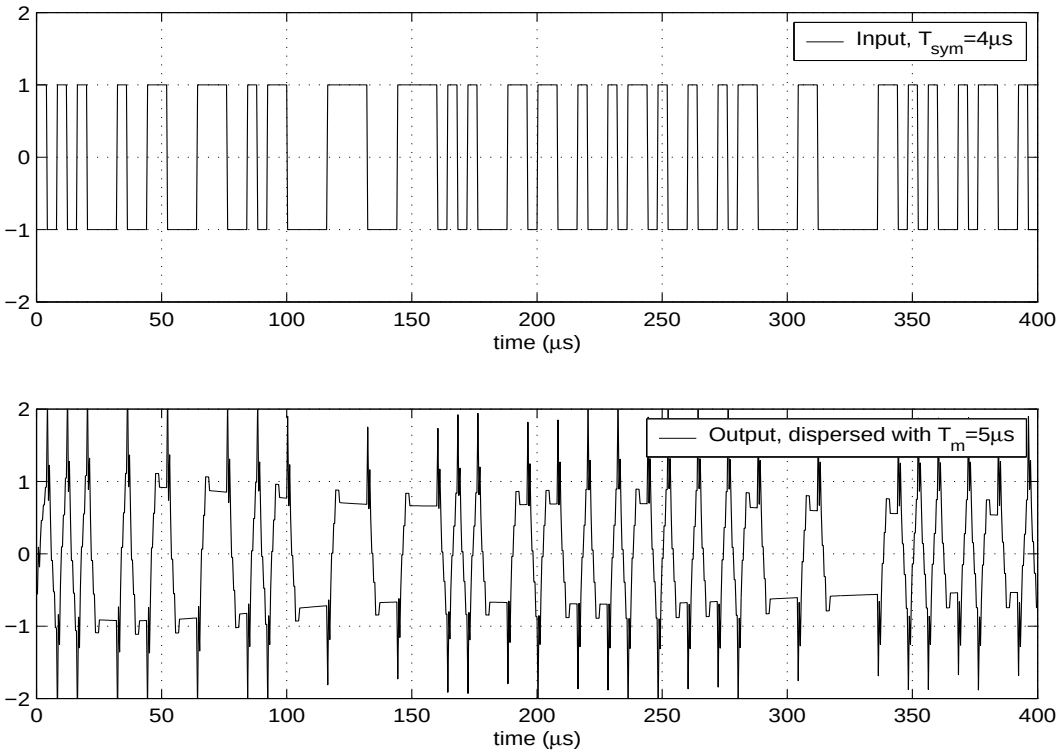


Figure 3.20: BPSK of symbol period $T = 4$ μs, input (top), and output (bottom), dispersed by channel with multipath spread $T_{\mathrm{d}} = 5$ μs.

3.9 Gauss-Markov Channel Model

3.9.1 Introduction

With an appropriate channel model, it is possible to design simple, effective receivers. One such modelling approach is to approximate the fading process experienced in the mobile wireless channel as a Gauss-Markov, and hence discrete, random process. In doing so, well-known decoding techniques such as the Viterbi algorithm (VA) can immediately be applied to decode the transmitted information. The Gauss-Markov approximation of the fading channel is also employed for channel estimation/tracking algorithms and evaluation of multiuser diversity algorithms in the context of cellular networks [53] [73] [74].

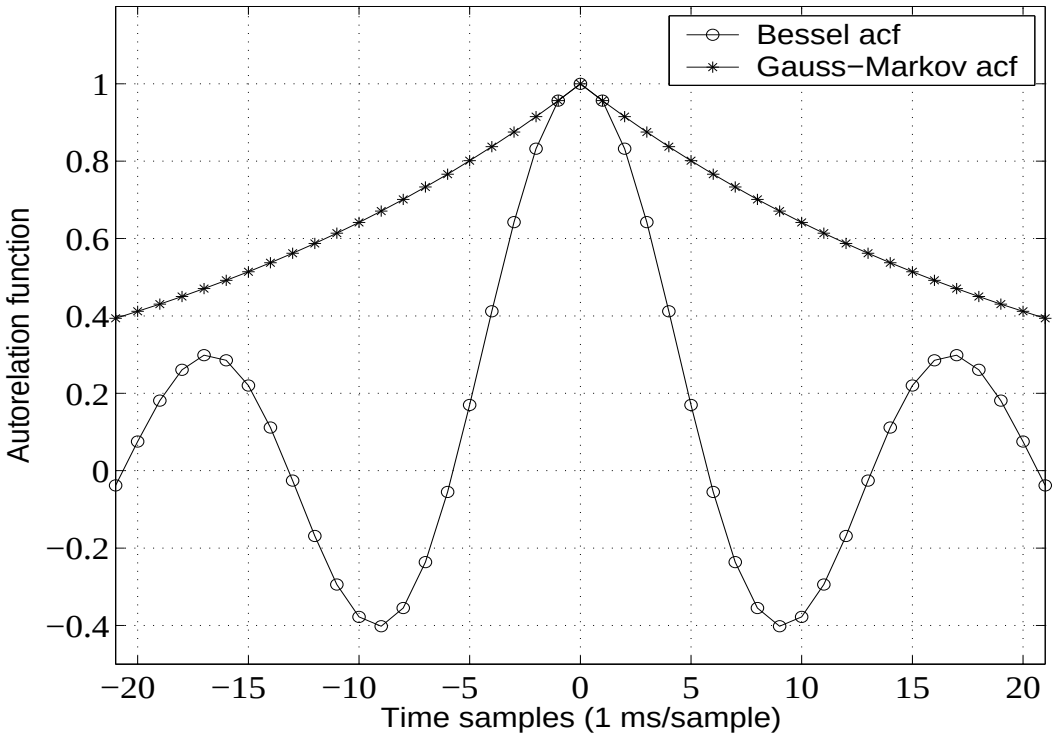


Figure 3.21: A Gauss-Markov approximation of Bessel autocorrelation with $f_{\mathrm{D}} = 66.67$ Hz and $f_s = 1000$ Hz.

3.9.2 A Gauss-Markov Approximation of the Rayleigh Fading Channel

We have seen from the previous discussion that the transmitted sequence $x[m]$ passes through a discrete time frequency non-selective Rayleigh fading channel (multipath spread $T_{\mathrm{m}} \ll 1/W$) with AWGN, the received discrete-time baseband signal $y[m]$ is

$$y[m] = h[m]x[m] + w[m] \tag{3.59}$$

where $h[m] \sim \mathcal{CN}(0,1)$ is the Rayleigh distributed fading coefficient and $w[m] \sim \mathcal{CN}(0,1)$ is a complex white Gaussian noise sample. From (3.38), the autocorrelation function of the sequence $\{h[m]\}$, now regarded as a random process, is

$$R[n] = J_0(2n\pi f_{\mathrm{D}} T_s) \tag{3.60}$$

Exact modelling of the time evolution of $h[m]$ with an autoregressive moving-average (ARMA) model is impossible because the autocorrelation functions are non-rational [53]. Accurate but large-order AR models for the fading channel are

possible but the first few correlation terms are more important for the design of the receiver. Low-order AR models, even a simple Gauss-Markov model, matching the Bessel autocorrelation of (3.38) well for small lags, can capture most of the channel dynamics and lead to effective tracking algorithms. A first-order (Gauss-Markov) approximation for the variation of $h[m]$ is to assume that $h[m]$ varies according to the auto-regressive model:

$$h[m+1] = \sqrt{1-\delta}h[m] + \sqrt{\delta}w[m+1], \quad m \geq 0, \tag{3.61}$$

with $\{w[m]\}$ a sequence of i.i.d. $\mathcal{CN}(0,1)$ random variables independent of $h[0] \sim \mathcal{CN}(0,1)$. The coherence time of the channel is controlled by the parameter δ. The autocorrelation function of the process in (3.61) can be obtained as

$$\tilde{R}[n] = (\sqrt{1-\delta})^{|n|} \tag{3.62}$$

A detailed derivation of (3.62) can be found in the Appendix.

Using (3.60) and (3.62), we establish the following relation by letting $R[1] = \tilde{R}[1]$

$$\delta = 1 - [J_0(2\pi f_{\mathrm{D}} T_s)]^2 \tag{3.63}$$

Fig. 3.21 compares the Bessel autocorrelation with its Gauss-Markov approximated autocorrelation for different T_s (or $f_s = 1/T_s$) with $f_{\mathrm{D}} = 66.67$ Hz. It can be observed that the Gauss-Markov model matches the Bessel model well for small lags and the coherence time T_{c} is inversely proportional to the Doppler spread D_{s}

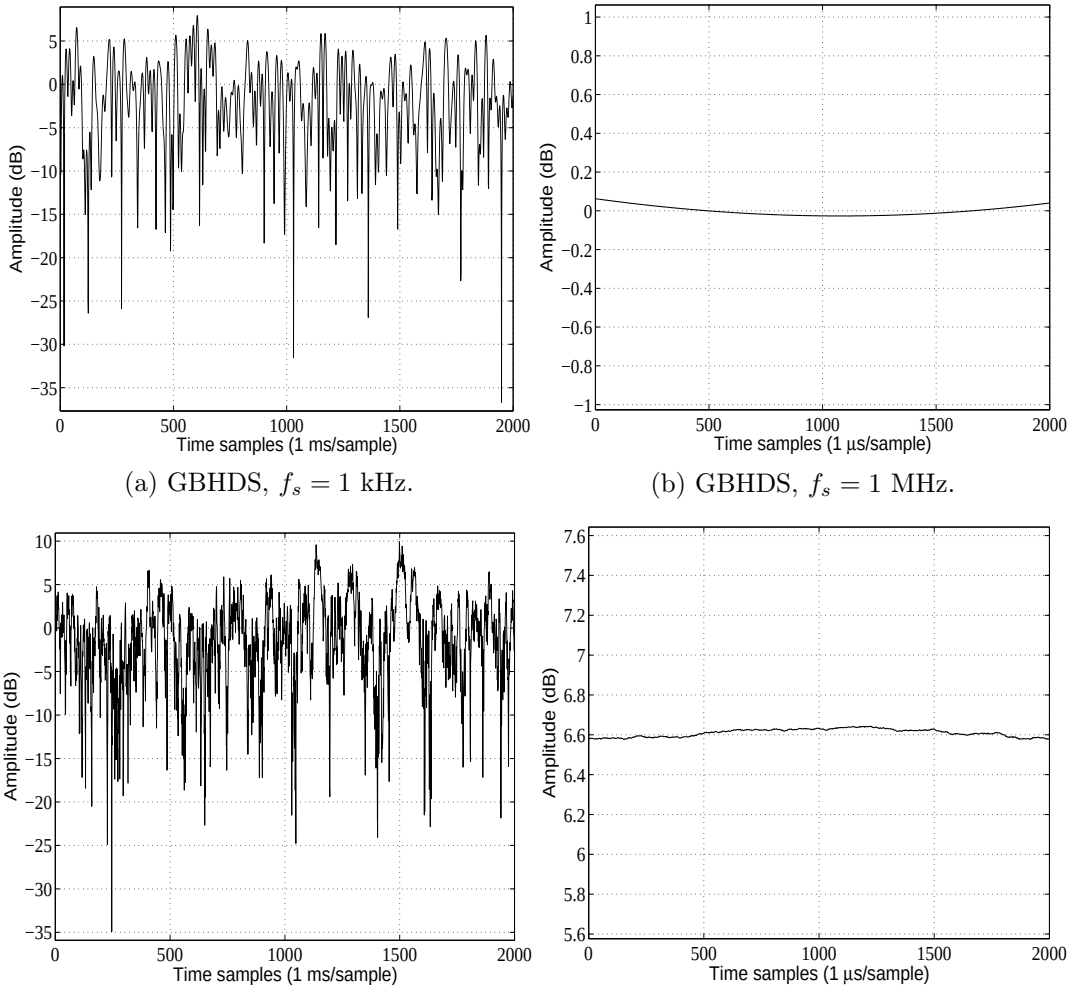
$$T_{\mathrm{c}} = \frac{1}{4D_{\mathrm{s}}}, \tag{3.64}$$

which is approximately 2 ms here. To further verify the result, a comparison plot is given in Fig. 3.22 for the amplitude of the fading sequences generated from GBHDS model and the Gauss-Markov approximation. Resemblance of time evolution can be observed. Further study may be extended to higher order statistics such as level-crossing rate (LCR) and average duration of fades (ADF).

3.10 Conclusions

In this chapter we presented fundamentals of channel modelling and developed an efficient channel simulator suitable for mobile communication. We analyzed the sadistical characteristics of the fading channel model and its capacity. We had results which proved that the diversity is an efficient way to improve the capacity and general performance of wireless communication systems.

Based on these studies, we constructed a frequency-selective channel model with hyperbolically-distributed scatterers for fading channel. We verified this model against the theoretical statistical properties.



(a) GBHDS, $f_s = 1$ kHz.

(b) GBHDS, $f_s = 1$ MHz.

(c) Gauss-Markov, $f_s = 1$ kHz.

(d) Gauss-Markov, $f_s = 1$ MHz.

Figure 3.22: Amplitude plots of fading sequences generated by GBHDS model and the Gauss-Markov approxmiation with Doppler frequency $f_{\mathrm{D}} = 66.67$ Hz.

We also developed a Gauss-Markov approximation to the proposed model for development of channel estimation/tracking algorithms and evaluation of multiuser diversity algorithms.

Chapter 4

Fading Mitigation Techniques

4.1 Introduction

An efficient way to combat multipath fading and dramatically improve the performance over fading channels is diversity. Diversity techniques can be generally identified as:

1. Frequency diversity, one example is Orthogonal Frequency Division Multiplexing (OFDM); Another example is code division multiple access (CDMA), which uses *Rake* receiver to extract frequency diversity.

2. Time diversity, such as coding and interleaving;

3. Spatial diversity, can be obtained by multiple transmit antennas or receive antennas (MIMO).

MIMO systems employing various techniques promise to boost the capacity of the communication and at the same time simplify the structure of the transmitter/receiver. For example, space-time coding exploits spatial diversity and results in a simple receiver structure [2]. Due to its many attractive features, this signalling technique has been adopted in several wireless standards for code division multiple access (CDMA) communications such as WCDMA and CDMA2000. Wireless LAN standards, with OFDM to combat fading channels, also employ spatial diversity at the mobile terminal and access point.

4.2 Orthogonal Frequency Division Multiplexing (OFDM)

Orthogonal Frequency Division Multiplexing (OFDM) is a promising technique for achieving high data rate and combating multipath fading in wireless com-

munications. OFDM systems are multi-carrier systems. In OFDM systems, transmit precoding is performed to convert the ISI channel into a set of non-interfering, orthonormal sub-carriers, each experiencing narrowband flat fading.

4.2.1 Inter-Symbol Interference (ISI) Elimination

In the previous chapter, we have established and simulated the discrete-time baseband model for the mobile fading channel. The input/output relation is

$$y[m] = \sum_{\ell} h_\ell[m]x[m-\ell] + w[m]. \tag{4.1}$$

which is an LTV system with the sequence of $\{h_\ell[m]\}$ for any given ℓ modelled as a complex circular symmetric process independent across the taps. Also, the simulation shows that if the delay spread T_d is greater than the symbol period T, ISI occurs. The basic idea behind OFDM is to divide the available bandwidth into subcarriers and consequently the symbol period T becomes longer so that ISI is negligible. In other words, the frequency-selective channel is decoupled into flat fading subchannels. With the truncation of the multipath delay spread T_d, the channel has a finite number of taps which can be defined as $L = T_\mathrm{d}W$. Further assuming the transmission happens within the coherence time T_c, we can rewrite the channel model in (4.1) as an LTI system

$$y[m] = \sum_{\ell=0}^{L-1} h_\ell x[m-\ell] + w[m]. \tag{4.2}$$

The OFDM scheme performs several operations to convert the ISI channel into orthogonal subchannels: cyclic prefix, IDFT/DFT and eigenvalue decomposition. Consider a block of input symbols of length N:

$$\mathbf{x} = [x[0], x[1], ..., x[N-1]]^t,$$

and denote the output of the channel (4.2) as

$$\mathbf{y} = [y[0], y[1], ..., y[N-1]]^t.$$

Also the N i.i.d $\mathcal{CN}(0, N_0)$ random variables of $w[m]$ form a standard circular symmetric Gaussian random vector $\mathbf{w}$.

If we prefix the block $\mathbf{x}$ cyclicly to create a new input block of length $N+L-1$ as

$$\tilde{\mathbf{x}} = [x[N-L+1], x[N-L+2], ..., x[N-1], x[0], x[1], ..., x[N-1]]^t,$$

i.e., a *cyclic prefix* of length $L-1$ symbols is inserted. To remove the ISI, the length of the cyclic prefix has to be greater than the channel tap length L. In wireless LAN (WLAN) standard such as IEEE 802.11a, the duration of the cyclic prefix is referred as *guard interval*, therefore the guard interval has to last longer than the multipath spread T_{d}. For analysis, we generally assume they are equal.

Eq. (4.2) can be expressed as

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \mathbf{w} \tag{4.3}$$

where

$$\tilde{\mathbf{H}} = \begin{bmatrix} h_0 & 0 & . & 0 & h_{L-1} & h_{L-2} & . & h_1 \\ h_1 & h_0 & 0 & . & 0 & h_{L-1} & . & h_2 \\ . & . & . & . & . & . & . & . \\ 0 & . & 0 & h_{L-1} & h_{L-2} & . & h_1 & h_0 \end{bmatrix} \tag{4.4}$$

is a circulant matrix of dimension $N \times N$. By adding cyclic prefix a Toeplitz channel matrix is converted to a circulant channel matrix. Circulant matrices have many interesting and useful properties. The eigenvalues of such matrices can easily be found exactly [30]. Investigation of the circulant channel matrix $\tilde{\mathbf{H}}$ will further simplify the problem. Define the operator matrix $\mathbf{U}$, which is the DFT matrix with the entries given by

$$u_{kn} = \frac{1}{\sqrt{N}} \exp\left(\frac{-\mathrm{j}2\pi kn}{N}\right), \quad k, n = 0, 1, ..., N-1, \tag{4.5}$$

and the IDFT matrix $\mathbf{U}^*$ is the complex conjugate transpose of $\mathbf{U}$. The DFT matrix $\mathbf{U}$ is unitary, i.e.,

$$\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}, \tag{4.6}$$

The circulant matrix $\tilde{\mathbf{H}}$ can be eigen-decomposed as

$$\tilde{\mathbf{H}} = \mathbf{U}^*\mathbf{\Lambda}\mathbf{U}. \tag{4.7}$$

Denote the first column in (4.4) the channel vector $\mathbf{h} = [h_0, h_1, ..., h_{L-1}, 0, ..., 0]^t$, we have the eigenvalues in the diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}(\lambda_0, ..., \lambda_{N-1})$ calculated as

$$\lambda_m = \sum_{k=0}^{N-1} h_\ell \exp\left(\frac{-\mathrm{j}2\pi km}{N}\right), \quad m = 0, 1, ..., N-1. \tag{4.8}$$

We notice the similarity between (4.5) and (4.8). We come to the conclusion that

$$\begin{aligned} \tilde{\mathbf{H}} &= \mathbf{U}^*\mathbf{\Lambda}\mathbf{U} \\ &= \mathbf{U}^* \cdot \mathrm{diag}(\sqrt{N}\mathbf{U}\mathbf{h}) \cdot \mathbf{U} \end{aligned} \tag{4.9}$$

that is, the eigenvalues of $\tilde{\mathbf{H}}$ are the DFT coefficients of the channel $\mathbf{h}$. Eq. (4.3) can thus be written as

$$\begin{aligned} \mathbf{y} &= \tilde{\mathbf{H}}\mathbf{x} + \mathbf{w} \\ &= \mathbf{U}^*\mathbf{\Lambda}\mathbf{U}\mathbf{x} + \mathbf{w} \end{aligned} \tag{4.10}$$

At the receiver, the output vector vector $\mathbf{y}$ is applied the DFT matrix $\mathbf{U}$ to obtain the vector $\tilde{\mathbf{y}}$, and at the transmitter, the vector $\mathbf{x}$ is obtained by applying IDFT $\mathbf{U}^*$ to the actual data symbols $\tilde{\mathbf{d}} = [\tilde{d}_0, \tilde{d}_1, ..., \tilde{d}_{N-1}]^t$. The final output vector $\tilde{\mathbf{y}}$ and the actual data vector $\tilde{\mathbf{d}}$ are related through

$$\tilde{\mathbf{y}} = \mathbf{\Lambda}\tilde{\mathbf{d}} + \tilde{\mathbf{w}} \tag{4.11}$$

where $\tilde{\mathbf{w}} = \mathbf{U}\mathbf{w}$ has the same distribution as $\mathbf{w}$, i.e., a vector of i.i.d. $\mathcal{CN}(0, N_0)$ random variables. The above operations effectively eliminate ISI and convert the frequency-selective channel into narrowband *parallel* sub-channels. The capacity is therefore the familiar form:

$$C = \max_{P_n} \sum_{n=0}^{N-1} \log\left(1 + \frac{|\lambda_n|^2 P_n}{N_0}\right) \quad \text{bits/OFDM symbol,} \tag{4.12}$$

subject to the power constraint $\sum_n P \leq NP$.

We see that while the channel knowledge is not necessary at the transmitter, it is needed at the receiver for the estimation of the data symbol vector $\tilde{\mathbf{d}}$ in Gaussian noise $\tilde{\mathbf{w}}$. In the case that channel information is available at the transmitter, *waterfilling* strategy can be applied to achieve the maximum rate of reliable communication.

The drawback of adding cyclic prefix is introducing overhead. To overcome this a large block size N is preferable. The block size N is constrained by the coherence time T_c of the channel and can not be arbitrary large if the channel is frequency-selective slowly time-varying as in (4.2). For a slowly time-varying channel, the coherence time T_c is much larger than the delay spread T_d, the coherence block length $T_\mathrm{c}W$ is thus much larger than the multipath length $L = T_\mathrm{d}W$. To approximate an LTV channel by (4.2) while keeping the overhead small, the block length N is constrained as:

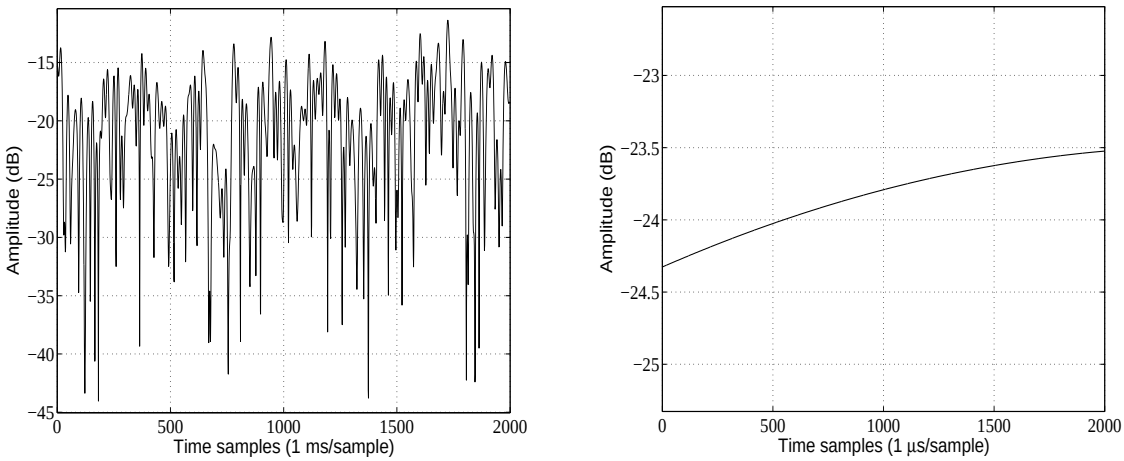
$$T_\mathrm{d}W \ll N \ll T_\mathrm{c}W.$$

For example, the indoor wireless channels in WLAN standards are well characterized by frequency-selective slowly time-varying channels. For reference, a list of key parameters of the OFDM standards regarding our discussion here is listed in Table 4.1.

Table 4.1: Key Parameters of the OFDM Standards

Number of Subcarriers	52
Number of Pilot Tones	4
OFDM Symbol Duration	4 μs
Guard Interval	800 ηs
Subcarrier Spacing	312.5 kHz
Signal Bandwidth	16.66 MHz

In the case of a macrocell environment discussed in the previous chapter, the OFDM block size N may be chosen by calculating the coherence time T_{c} of the channel. This can be illustrated by plotting the amplitude of Rayleigh fading sequence generated by the GBHDS model. It can be observed that the Rayleigh fading sequence in Fig. 4.1 undergoes different variations. The simulation parameters are typical macrocell channel with carrier frequency $f_{\mathrm{c}} =$ 900 MHz, velocity of mobile $v = 64$ km/h, Doppler shift $f_{\mathrm{D}} = 50$ Hz, the distance between transmitter and receiver $d = 1$ km.



(a) $f_s = 1$ kHz. (b) $f_s = 1$ MHz.

Figure 4.1: Amplitude of a Rayleigh fading sequence.

In the example, the time interval needed for the fading sequence to undergo a significant change as large as 30 dB is a few milli-seconds. To be specific, the coherence time is inversely proportional to the Doppler spread D_{s} [92]

$$T_{\mathrm{c}} = \frac{1}{4D_{\mathrm{s}}},$$

and T_{c} is 2.5 ms for $D_{\mathrm{s}} = 100$ Hz. With sampling frequency $f_s = 1$ MHz, the variation of the fading sequence is significantly slow. This indicates that if the

OFDM symbol duration is properly chosen, for example 4 μs in OFDM WLAN standards, the assumption of a frequency-selective slowly time-varying channel is valid.

4.2.2 Implementation of OFDM Systems

One of the key advantages of OFDM systems is that it can be implemented with low cost by applying a discrete-time signal to an ideal D/A converter. Again we assume the transmission is passband-limited to W, the symbol duration for the information sequence is therefore $T = 1/W$. Consider the following set of orthonormal pulses:

$$g_n(t) = \frac{1}{\sqrt{T_{\rm os}}} e^{{\rm j}2\pi nt/T_{\rm os}} w(t), \quad \text{for } n = 0, ..., N-1, \tag{4.13}$$

where $w(t) = u(t) - u(t - T_{\rm os})$ is a rectangular pulse window over $[0, T_{\rm os})$ and $T_{\rm os} = NT$ is the duration of one OFDM symbol. To show the pulses in (4.13) are orthonormal, the Fourier transform of $g_n(t)$, denoted as $G_n(f)$, is calculated and plotted in Fig. 4.2. The transmission bandwidth W is divided into N orthogonal subcarriers or tones, and each tone occupies a bandwidth of W/N. The data symbols are grouped into a block $\tilde{\mathbf{d}} = [\tilde{d}_0, \tilde{d}_1, ..., \tilde{d}_{N-1}]^t$. Instead of transmitting at the rate of W, each information symbol in the block is modulated on a subcarrier in (4.13) and transmitted at a much slow rate of W/N. As the N information symbols in the block are transmitted simultaneously, the overall transmission rate remains unchanged as W. A one-shot OFDM transmitter sends the baseband signal:

$$x(t) = \sum_{n=0}^{N-1} \tilde{d}_n g_n(t). \tag{4.14}$$

The passband signal $x_p(t)$ with carrier frequency $f_{\rm c}$ can be conveniently expressed by:

$$\begin{aligned} x_p(t) &= \Re\{e^{{\rm j}2\pi f_{\rm c} t} x(t)\} \\ &= \Re\left\{ \frac{1}{\sqrt{T_{\rm os}}} \sum_{n=0}^{N-1} \tilde{d}_n \exp\left({\rm j}2\pi \left(f_{\rm c} + \frac{n}{T_{\rm os}} \right) t \right) \right\}, \quad t \in [0, T_{\rm os}]. \end{aligned} \tag{4.15}$$

As the pulses in (4.13) are not bandlimited due to the rectangular window, in practical, $g_n(t)$ is approximated by [5]

$$\tilde{g}_n(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp\left(\frac{{\rm j}2\pi nk}{N} \right) p(t - kT_{\rm os}/N), \quad \text{for } n = 0, ..., N-1, \tag{4.16}$$

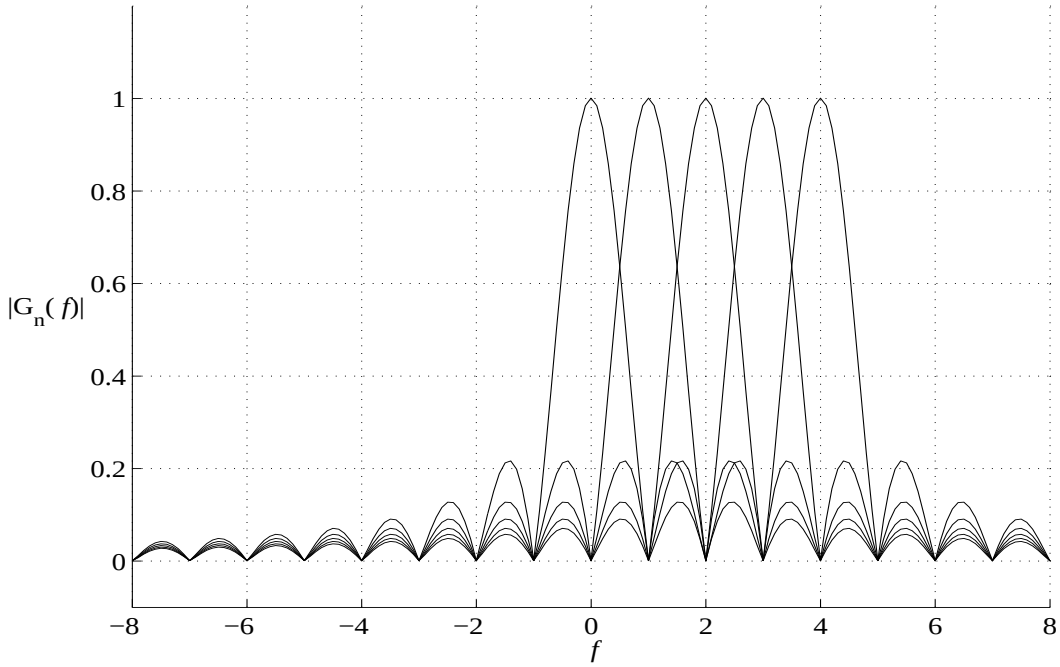


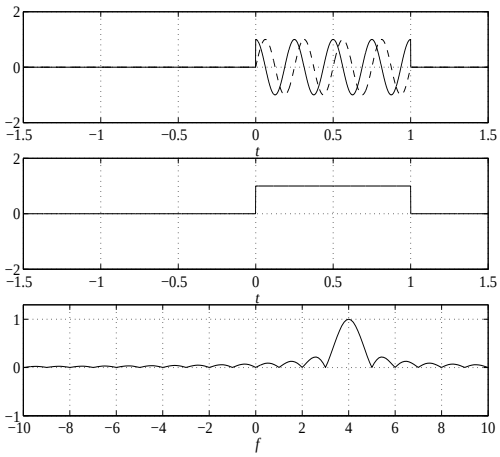
Figure 4.2: $|G_n(f)|$, spectrum of $g_n(t)$ for n=0,1,2,3,4. The OFDM symbol duration T_{os} is normalized to one.

where $p(t)$ is an ideal unit-energy reconstruction filter with a sampling period T:

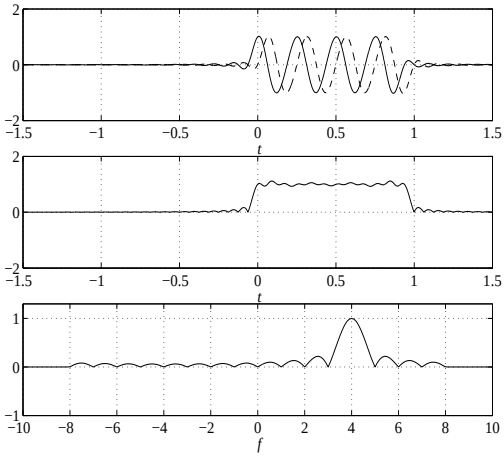
$$p(t) = \frac{1}{\sqrt{T}}\text{sinc}\left(\frac{t}{T}\right) \tag{4.17}$$

where the sinc function is defined as $\text{sinc} = \sin(\pi t)/(\pi t)$.

The approximation is demonstrated in Fig. 4.3 with $n = 4$, $N = 16$ and T_{os} normalized to 1. It can be also shown that $\tilde{g}_n(t)$ is bandlimited to $|f| < N/(2T_{\text{os}})$ with n taking values from 0 to $N - 1$.



(a) $\mathrm{Re}\{g_4(t)\}$, $\mathrm{Im}\{g_4(t)\}$, $|g_4(t)|$ and $|G_4(f)|$.



(b) $\mathrm{Re}\{\tilde{g}_4(t)\}$, $\mathrm{Im}\{\tilde{g}_4(t)\}$, $|\tilde{g}_4(t)|$ and $|\tilde{G}_4(f)|$.

Figure 4.3: Approximating $g_4(t)$ by $\tilde{g}_4(t)$, $N = 16$, $n = 4$ and $T_{\mathrm{os}} = 1$.

With (4.16), (4.14) can be expressed as:

$$\begin{aligned} x(t) &= \sum_{n=0}^{N-1} \tilde{d}_n \tilde{g}_n(t) \\ &= \sum_{n=0}^{N-1} \tilde{d}_n \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp\left(\frac{\mathrm{j}2\pi nk}{N}\right) p(t - kT_{\mathrm{os}}/N) \\ &= \sum_{k=0}^{N-1} \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \tilde{d}_n \exp\left(\frac{\mathrm{j}2\pi nk}{N}\right) \right\} p(t - kT_{\mathrm{os}}/N) \\ &= \sum_{k=0}^{N-1} x[k] p(t - kT_{\mathrm{os}}/N). \end{aligned} \tag{4.18}$$

where $x[k]$ is the kth inverse DFT coefficient of the vector $\tilde{\mathbf{d}}$. This result indicates a simple implementation scheme of an OFDM signal of (4.13). First we perform an IDFT of the data symbols and then modulate the obtained coefficients to unit-energy pulses shown in Fig. 4.4. Another observation is that the symbols $x[k]$ propagate through the channel serially at the symbol rate of $1/T$ whereas effectively the data symbols $\tilde{d}_n$ transmit though narrowband *parallel* subchannels at the overall rate of $N(1/T_{\mathrm{os}}) = 1/T$.

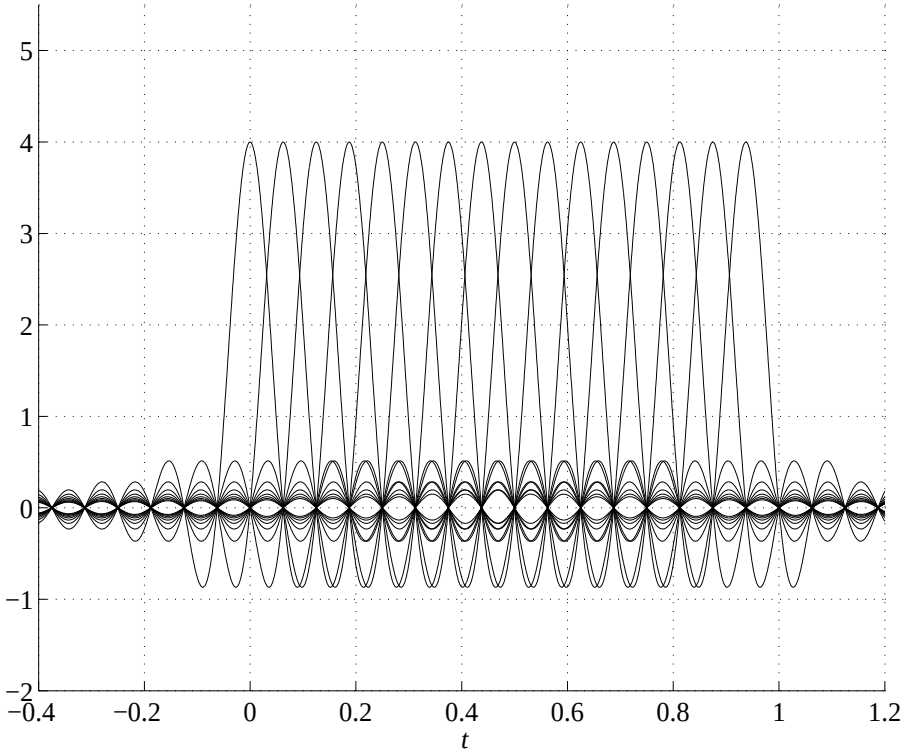


Figure 4.4: Unit-energy pulses $p(t - kT_{\text{os}})$, $k = 0, ..., 15$, with T_{os} normalized to 1.

4.3 Spatial Diversity

Spatial diversity or antenna diversity can be obtained by placing multiple antennas at the transmitter and/or the receiver. A system that employs multiple antennas at both transmitter and receiver is now termed as a MIMO system. Channels with multiple transmit and multiple receive antennas are MIMO channels. We will see that diversity is an efficient way to combat fading channels and in addition to providing diversity (improvement of the reliability), MIMO channels also provide additional *degrees of freedom* (increase of the capacity) for communication. In other words, the benefits of MIMO communication are twofold: a multiplexing gain and a diversity gain. Investigation of the tradeoff between these two types of gains leads to optimal space-time coding schemes [99] [100].

There are three categories of MIMO techniques based on diversity-multiplexing tradeoff [98] :

1. This kind of schemes improves the power efficiency by maximizing spatial diversity, such as space-time block codes (STBC);

2. A layered approach is used to increase capacity, such as the vertical-Bell Laboratories layered space-time (V-BLAST);

3. This type needs the channel knowledge at the transmitter to decompose the channel matrix using singular value decomposition (SVD) and uses these decomposed unitary matrices as pre- and post-filters at the transmitter and receiver to achieve capacity gain.

4.3.1 Selection Combining and Maximal-Ratio Combining: A Simulation Study

Fundamentally, diversity techniques improve the system performance over a fading channel by changing its statistical characteristics. We have shown that the input/output relation over a slow flat Rayleigh fading channel can be expressed as

$$y[m] = hx[m] + w[m], \tag{4.19}$$

where $h \sim \mathcal{CN}(0, 1)$.

Assuming a single input multiple output (SIMO) channel with a selection combiner at the receiver using L receive antennas, and the received signal of the ℓth diversity branch is defined by

$$y_\ell[m] = h_\ell x[m] + w_\ell[m], \quad \ell = 1, 2, ..., L. \tag{4.20}$$

and by the design, all h_ℓ are i.i.d $\mathcal{CN}(0, 1)$. Let $\alpha_\ell = |h_\ell|$. Under the assumption of Rayleigh fading α_ℓ is Rayleigh distributed and α_ℓ^2 is exponentially distributed.

Let γ_ℓ denote the *instantaneous* signal-to-noise ratio measured at the output of the kth receiver during the transmission of a given symbol. We may write

$$\gamma_\ell = \frac{\mathcal{E}_\mathrm{b}}{N_0}\alpha_\ell^2 \quad \ell = 1, 2, ..., L \tag{4.21}$$

where $\mathcal{E}_\mathrm{b}$ is the energy per bit and N_0 is the one-sided noise spectral density. Since all the α_ℓ^2 have the same mean, γ_av, we may express the probability density functions of the random variables Γ_ℓ pertaining to the individual branches as

$$f_{\Gamma_\ell}(\gamma_\ell) = \frac{1}{\gamma_\mathrm{av}}\exp\left(-\frac{\gamma_\ell}{\gamma_\mathrm{av}}\right) \quad \gamma_\ell \geq 0 \quad \text{and } \ell = 1, 2, ..., L. \tag{4.22}$$

By selection combining, we have a random variable Γ_sc described by the value

$$\gamma_\mathrm{sc} = \max\{\gamma_1, \gamma_2, ..., \gamma_L\}. \tag{4.23}$$

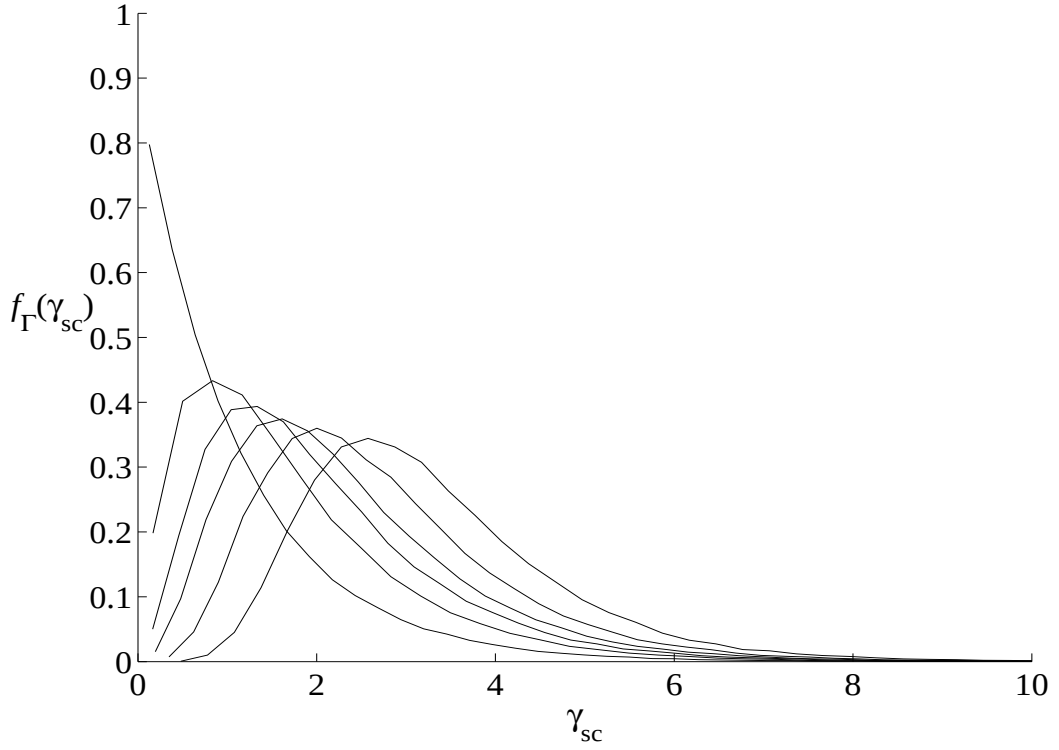


Figure 4.5: Simulation of selection combiner, the probability density function of $\Gamma_{\rm sc}$ for different number L of receive antennas ($L = 1, 2, 3, 4, 6, 10$) and $\mathcal{E}_{\rm b}/N_0$ is normalized to 1.

The probability of density of $\Gamma_{\rm sc}$ can be derived as [64] [34]

$$f_\Gamma(\gamma_{\rm sc}) = \frac{L}{\gamma_{\rm av}} \exp\left(-\frac{\gamma_{\rm sc}}{\gamma_{\rm av}}\right)\left[1 - \exp\left(-\frac{\gamma_{\rm sc}}{\gamma_{\rm av}}\right)\right]^{(L-1)} \quad \gamma_{\rm sc} \geq 0 \tag{4.24}$$

Fig. 4.5 plots the simulation result of $f_\Gamma(\gamma_{\rm sc})$ for $L = 1, 2, 3, 4, 6, 10$ with $\mathcal{E}_{\rm b}/N_0$ normalized to 1. From the figure, the following observations can be made:

1. As expected, when single antenna, $L = 1$, $f_\Gamma(\gamma_{\rm sc})$ is exponentially distributed. As the number of diversity branches, L, is increased, the probability density function $f_\Gamma(\gamma_{\rm sc})$ moves progressively to the right.

2. $f_\Gamma(\gamma_{\rm sc})$ approaches Gaussian as L is increased.

In other words, a slow flat Rayleigh fading channel is modified into a Gaussian channel through the use of selection combining, provided that the number

L of diversity channels is sufficiently large. This modification is favorable and is a "digital communication theorist's dream" [34] because the classic theory of communication can now be applied.

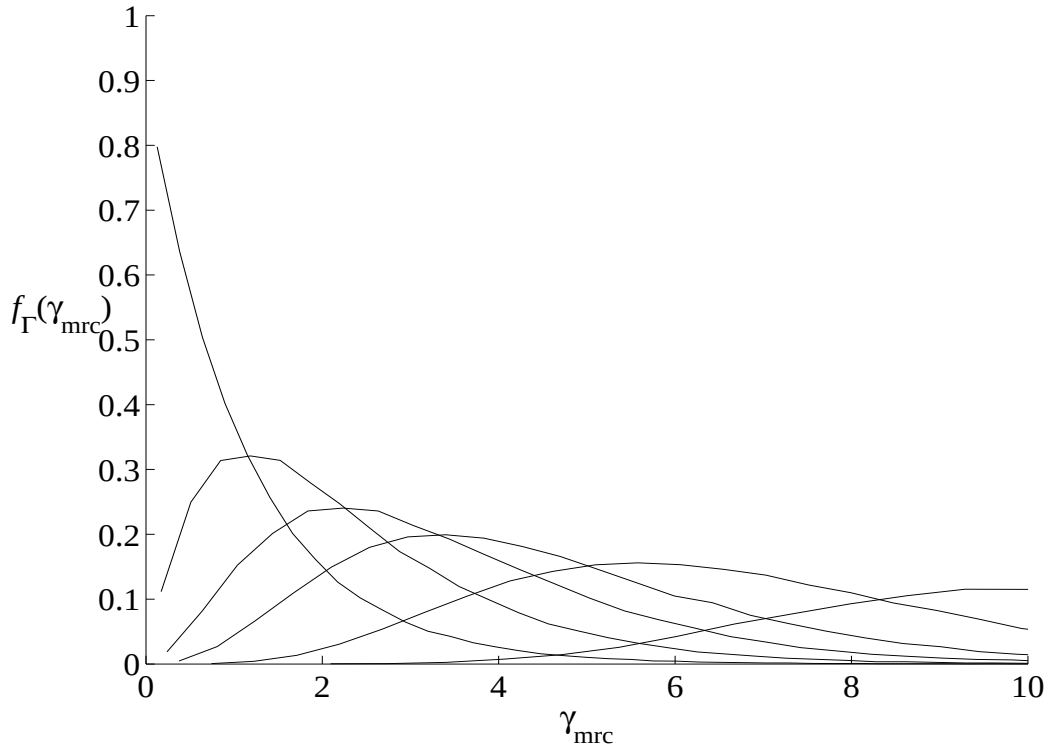


Figure 4.6: Simulation of maximal-ratio combiner, the probability density function of Γ_{mrc} for different number L of receive antennas ($L = 1, 2, 3, 4, 6, 10$) and $\mathcal{E}_{\text{b}}/N_0$ is normalized to 1.

The maximal-ratio combiner produces an instantaneous output signal-to-noise ratio that is the sum of the instantaneous signal-to-noise ratios of the individual branches

$$\gamma_{\text{mrc}} = \sum_{\ell=1}^{L} \gamma_\ell \tag{4.25}$$

that is, γ_{mrc} is equal to the sum of L exponentially distributed random variables for a slow flat Rayleigh fading channel. Unlike selection combiner chooses the maximum value and discards other information, it takes consideration of available information from all the branches. The distribution of γ_{mrc} is (*chi-square*

with $2L$ degrees of freedom)

$$f_\Gamma(\gamma_{\rm mrc}) = \frac{1}{(L-1)!}\frac{\gamma_{\rm mrc}^{L-1}}{\gamma_{\rm av}^{L}}\exp\left(-\frac{\gamma_{\rm mrc}}{\gamma_{\rm av}}\right) \quad \gamma_{\rm mrc} \geq 0 \tag{4.26}$$

Fig. 4.6 plots the simulation result of $f_\Gamma(\gamma_{\rm mrc})$ for $L = 1, 2, 3, 4, 6, 10$ with $\mathcal{E}_{\rm b}/N_0$ normalized to 1. Similar observation can be made to those for the selection combiner except that the probability density function is scaled differently by L. A practical example of maximal-ratio combiner is Rake receiver employed by CDMA system which combines the signals from different delay paths. Further analysis shows that the probability of error of system employing maxima ratio combiner decreases inversely with the Lth power of the SNR [64].

4.3.2 Capacity of Fading Channels: the Diversity Gain

We have shown diversity increase the system performance by changing the channel statistics. In this section we investigate the same problem from a more fundamental point of view. We extend the results of ϵ-outage capacity for signal antenna flat fading channel in (3.22) to multiple receive antennas. Again we consider a SIMO channel with one transmit antenna and L receive antennas:

$$y_\ell[m] = h_\ell x[m] + w_\ell[m] \quad \ell = 1, 2, ..., L, \tag{4.27}$$

where h_ℓ is i.i.d. $\mathcal{CN}(0,1)$ but remains constant for all time, and $w_\ell[m]$ is $\mathcal{CN}(0, N_0)$ additive Gaussian noise independent across antennas. Assuming a perfect channel estimation at the receiver, i.e., the channel vector $\mathbf{h} = [h_1, ..., h_L]^t$ is known to the receiver, (4.27) can be converted to an AWGN channel with sufficient statistic for detecting $x[m]$

$$\tilde{y}[m] = \|\mathbf{h}\|^2 x[m] + \mathbf{h}^*\mathbf{w}[m], \tag{4.28}$$

where the random variable $X = \|\mathbf{h}\|^2$ is distributed as Chi-square with $2L$ degrees of freedom

$$f_X(x) = \frac{1}{(L-1)!}x^{L-1}e^{-x}, \quad x \geq 0 \tag{4.29}$$

In a similar fashion to (3.29), the outage probability can be obtained as

$$\begin{aligned} p_{\rm out}(R) &= \mathbb{P}\{\log(1 + \|\mathbf{h}\|^2\mathsf{SNR} < R\} \\ &= \mathbb{P}\{\|\mathbf{h}\|^2 < \frac{2^R - 1}{\mathsf{SNR}}\} \\ &\approx \frac{(2^R-1)^L}{L!\mathsf{SNR}^L} \quad \text{at high SNR.} \end{aligned} \tag{4.30}$$

Obviously, (3.29) is the special case of (4.30) with $L = 1$. We see a gain from diversity: the outage probability now decays like $1/\mathsf{SNR}^L$.

An alternative way to see this performance gain from diversity is the ϵ, which can be found as (cf. (3.33))

$$C_\epsilon = \log(1 + F^{-1}(1-\epsilon)) \text{ bits/s/Hz},$$

with F the complementary cumulative distribution function of the random variable $X = \|\mathbf{h}\|^2$, i.e., $F(x) = \mathbb{P}\{\|\mathbf{h}\|^2 > x\}$. The cumulative distribution function (CDF) of X with distribution of (4.29) is

$$F_X(x) = 1 - e^{-x} \sum_{k=0}^{L-1} \frac{1}{k!} x^k, \quad x \geq 0, \tag{4.31}$$

so that the complementary CDF of X is

$$F(x) = e^{-x} \sum_{k=0}^{L-1} \frac{1}{k!} x^k, \quad x \geq 0. \tag{4.32}$$

This result can be verified again the result obtained by the single antenna system.

To find out C_ϵ, we have to solve this equation for different L:

$$e^{-x} \sum_{k=0}^{L-1} \frac{1}{k!} x^k = 1 - \epsilon, \tag{4.33}$$

which can be solved numerically by Newton's method [71]. The ϵ-outage capacity, normalized by AWGN channel capacity $C_{\text{awgn}} = \log(1 + L\mathsf{SNR})$, is plotted in Fig. 4.7. The improvement on outage capacity by diversity is significant. At $\mathsf{SNR} = -10$ dB, the outage capacity is increased from 1% ($L = 1$) to more than 50% ($L = 5$).

We have investigated the performance gain by L-fold receive diversity. The transmit diversity gain can be investigated in the same manner.

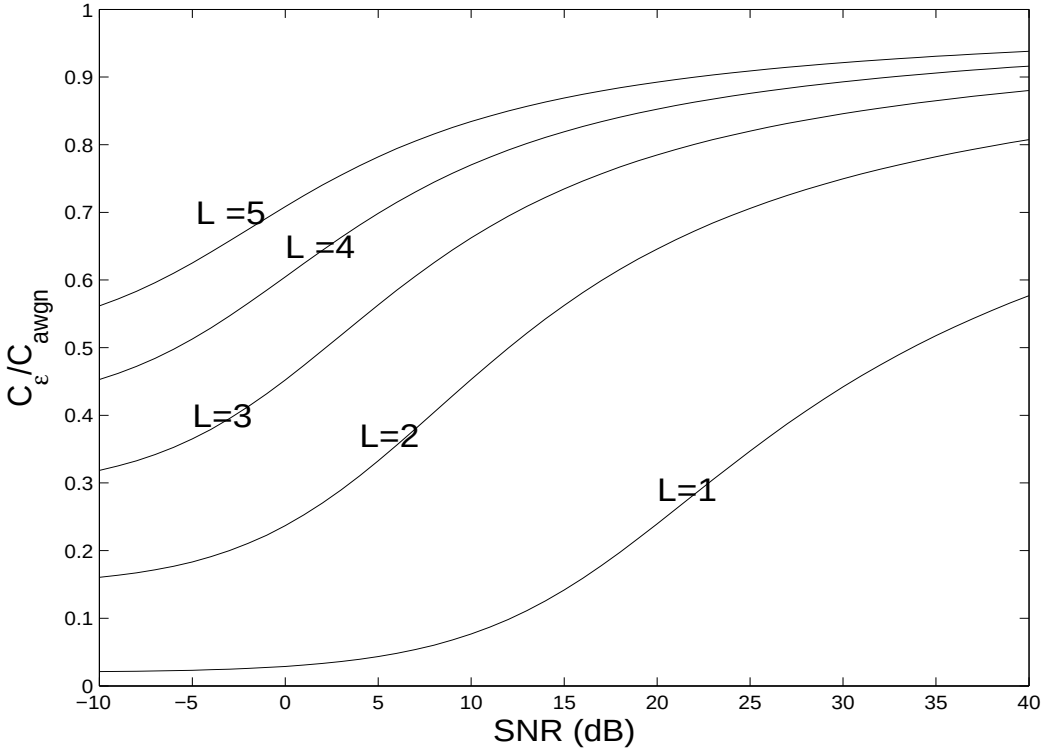


Figure 4.7: ϵ-outage capacity with L-fold receive diversity, as a fraction of the AWGN capacity $\log(1 + L\mathsf{SNR})$, for $\epsilon = 0.01$ and different L.

4.3.3 MIMO Channel Modelling

In the previous section, we have seen that multiple transmit or multiple receive antennas can provide a *diversity* gain or a *power* gain. With a MIMO channel having both multiple transmit and multiple receive antennas, a *degree-of-freedom* gain can be produced. This is achieved by spatially multiplexing several data streams onto the MIMO channel and consequently the capacity is increased. In this section, we first analyze the capacity of the MIMO channel via singular value decomposition and find the analogy between MIMO channel and OFDM parallel channels. We then model the MIMO channel in the *angular domain* which can be extended to a statistical model.

A narrowband time-invariant wireless channel with n_{t} transmit and n_{r} receive antennas is described by an n_{r} by n_{t} deterministic matrix $\mathbf{H}$. For wideband system, the multipath characteristic of the environment causes the MIMO channel to be frequency-selective. Assuming that the transmission occurs within the coherence time T_{c}, we consider the time-invariant frequency-selective MIMO channel:

$$\mathbf{y}[m] = \sum_{\ell=0}^{L-1} \mathbf{H}_\ell \mathbf{x}[m-\ell] + \mathbf{w}[m]. \tag{4.34}$$

where $\mathbf{x}[m] \in \mathcal{C}^{n_{\mathrm{t}}}$, $\mathbf{y}[m] \in \mathcal{C}^{n_{\mathrm{r}}}$, and $\mathbf{w}[m] \sim \mathcal{CN}(0, N_0\mathbf{I}_{n_{\mathrm{r}}})$ denote the transmitted signal, received signal and white Gaussian noise respectively at the symbol time m. The channel matrix $\mathbf{H}_\ell \in \mathcal{C}^{n_{\mathrm{r}} \times n_{\mathrm{t}}}$. An appropriate OFDM transmission and reception scheme can transform this frequency-selective MIMO channel into a set of parallel frequency-flat MIMO channels

$$\tilde{\mathbf{y}}_n = \tilde{\mathbf{H}}_n \tilde{\mathbf{x}}_n + \tilde{\mathbf{w}}_n. \quad n = 0, ..., N-1 \tag{4.35}$$

This combination of MIMO and OFDM boosts capacity and reduces receiver complexity. Consequently, the analysis can be focused on the following time-invariant frequency-flat MIMO channel model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{4.36}$$

where time index is dropped for simplicity. The MIMO channel can be converted into equivalent parallel channels by singular value decomposition (SVD). Let the SVD be given by $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$, then $\mathbf{U}$ and $\mathbf{V}$ are unitary and $\mathbf{\Lambda} \in \mathcal{R}^{n_{\mathrm{r}} \times n_{\mathrm{t}}}$ is a rectangular matrix whose diagonal elements are non-negative real numbers and whose off-diagonal elements are zero, i.e. $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_{n_{\min}}, 0, ..., 0)$, where $n_{\min} := \min(n_{\mathrm{t}}, n_{\mathrm{r}})$. The MIMO channel (4.36) can then be rewritten as a parallel Gaussian channel:

$$\tilde{\mathbf{y}} = \Lambda\tilde{\mathbf{x}} + \tilde{\mathbf{w}} \tag{4.37}$$

where $\tilde{\mathbf{y}} = \mathbf{U}^*\mathbf{y}$, $\tilde{\mathbf{x}} = \mathbf{V}^*\mathbf{x}$, and $\tilde{\mathbf{w}} = \mathbf{U}^*\mathbf{w}$. Since $\mathbf{U}$ is unitary, $\tilde{\mathbf{w}}$ has the same distribution as $\mathbf{w}$, and $\|\tilde{\mathbf{x}}\|^2 = \|\mathbf{x}\|^2$. Thus the pre- and post-possessing will not change the signal-to-noise ratio. The capacity of the parallel Gaussian channel (4.37) is

$$C = \max_{P_i} \sum_{i=1}^{n_{\min}} \log\left(1 + \frac{\lambda_i^2 P_i^*}{N_0}\right) \quad \text{bits/s/Hz}, \tag{4.38}$$

where $P_1^*, ..., P_{n_{\min}}^*$ are the waterfilling power allocations:

$$P_i^* = \left(\mu - \frac{N_0}{\lambda_i^2}\right)^+, \tag{4.39}$$

with μ chosen to satisfy the total power constraint $\sum_i P_i^* = P$. This SVD decomposition structure is analogous to the decomposition of a circulant channel matrix in the OFDM system that we have discussed previously except that SVD decomposition is channel realization dependent and that channel knowledge at transmitter is necessary.

We have seen that a MIMO channel provides $n_{\min}$ spatial degrees of freedom. Mathematically, the rank and the condition number $(\max_i \lambda_i / \min_i \lambda_i)$ of $\mathbf{H}$ determines the performance of a MIMO channel. A good MIMO channel model should therefore capture the essence of the spatial degree of freedom and reflect the condition of the physical environment. We investigate the MIMO model with angular domain representation shown in [76] [92].

Consider the time-invariant frequency-flat MIMO channel in (4.36). The n_{t} transmit and n_{r} receive antennas are placed in uniform linear arrays of normalized lengths L_{t} and L_{r}, respectively. The normalization is by the wavelength λ_{c} of the passband transmitted signal, i.e., the physical length of the transmit antenna array is $L_{\mathrm{t}}\lambda_{\mathrm{c}}$ and the physical length of the receive antenna array is $L_{\mathrm{r}}\lambda_{\mathrm{c}}$.

Suppose that there is an arbitrary number of physical paths between the transmitter and the receiver; the ith path has an attenuation of a_i, makes an angle of $\phi_{\mathrm{t}i}$ ($\Omega_{\mathrm{t}i} := \cos\phi_{\mathrm{t}i}$) with the transmit antenna array and an angle of $\phi_{\mathrm{r}i}$ ($\Omega_{\mathrm{r}i} := \cos\phi_{\mathrm{r}i}$) with the receive antenna array. The channel matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \sum_i a_i^{\mathrm{b}} \mathbf{e}_{\mathrm{r}}(\Omega_{\mathrm{r}i}) \mathbf{e}_{\mathrm{r}}(\Omega_{\mathrm{t}i})^* \tag{4.40}$$

where,

$$a_i^{\mathrm{b}} = a_i \sqrt{n_{\mathrm{t}} n_{\mathrm{r}}} \exp\left(-\frac{\mathrm{j}2\pi d^{(i)}}{\lambda_{\mathrm{c}}}\right),$$

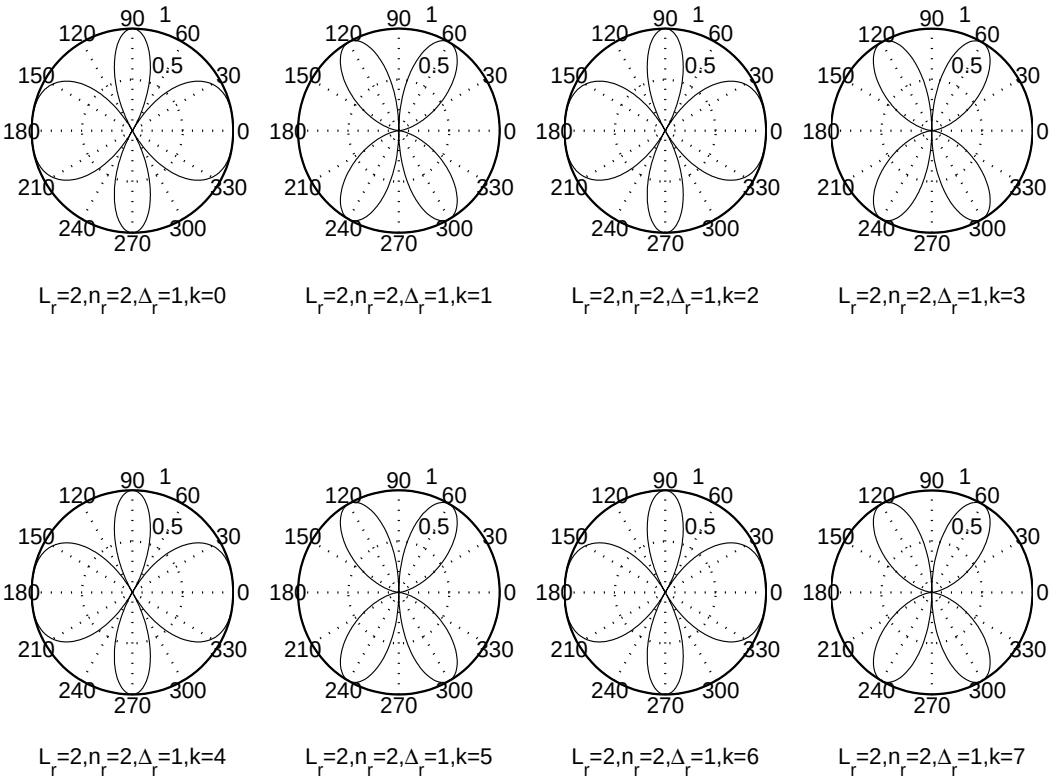


Figure 4.8: $L_{\mathrm{r}} = 2, n_{\mathrm{r}} = 2$, different Δ_{r} and k.

$$\mathbf{e}_{\mathrm{r}}(\Omega) := \frac{1}{\sqrt{n_{\mathrm{r}}}} \begin{bmatrix} 1 \\ \exp(-\mathrm{j}2\pi\Delta_{\mathrm{r}}\Omega) \\ \vdots \\ \exp(-\mathrm{j}2\pi(n_{\mathrm{r}}-1)\Delta_{\mathrm{r}}\Omega) \end{bmatrix}, \tag{4.41}$$

$$\mathbf{e}_{\mathrm{r}}(\Omega) := \frac{1}{\sqrt{n_{\mathrm{r}}}} \begin{bmatrix} 1 \\ \exp(-\mathrm{j}2\pi\Delta_{\mathrm{r}}\Omega) \\ \vdots \\ \exp(-\mathrm{j}2\pi(n_{\mathrm{r}}-1)\Delta_{\mathrm{r}}\Omega) \end{bmatrix}. \tag{4.42}$$

Also, $d^{(i)}$ is the distance between transmit antenna 1 and receive antenna 1 along path i. The vector $\mathbf{e}_{\mathrm{t}}(\Omega)$ and $\mathbf{e}_{\mathrm{r}}(\Omega)$ are, respectively, the transmitted and received unit spatial signatures along the direction Ω. Examples of the beamforming patterns of the receive angular basis vectors $\mathbf{e}_{\mathrm{r}}(k/L_{\mathrm{r}})$ is shown in Fig. 4.8, 4.9, and 4.10 with $k = 0, ..., 7$ for comparison.

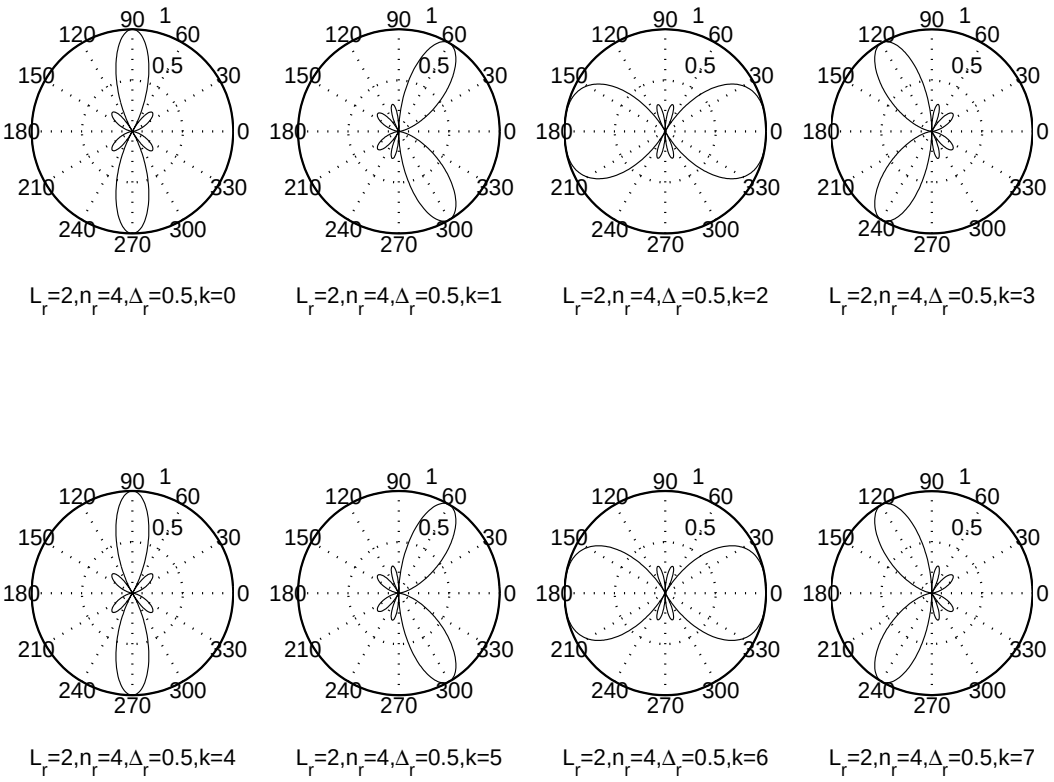


Figure 4.9: $L_{\mathrm{r}} = 2, n_{\mathrm{r}} = 4$, different Δ_{r} and k.

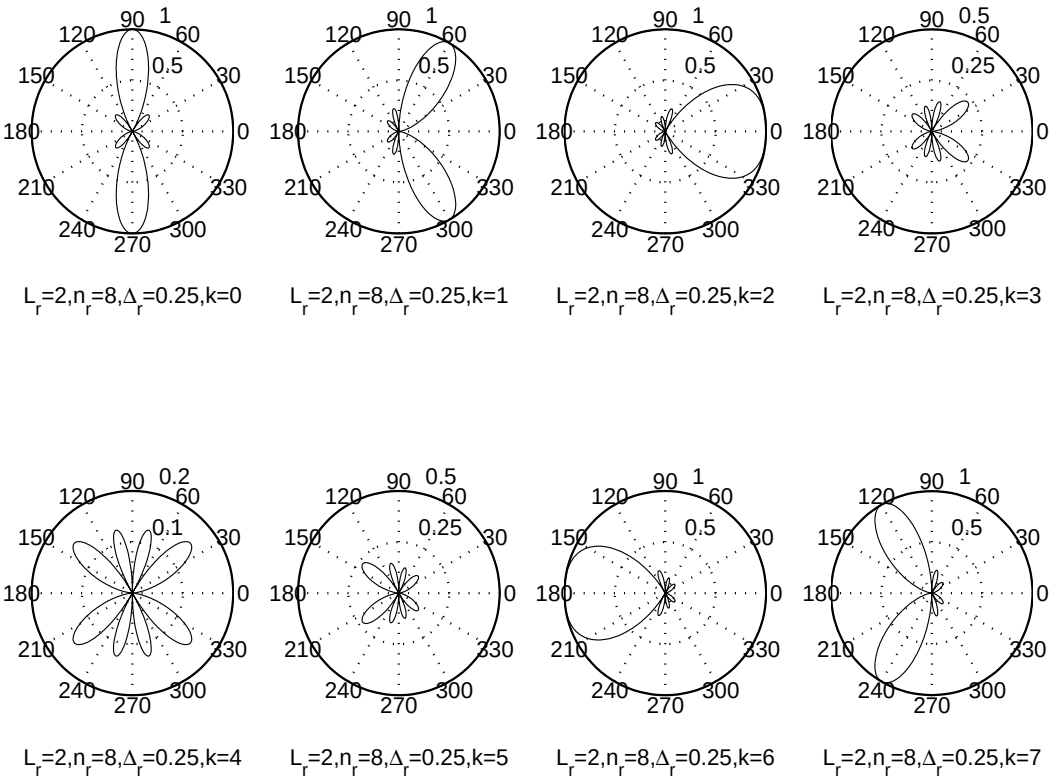


Figure 4.10: $L_{\mathrm{r}} = 2, n_{\mathrm{r}} = 8$, different Δ_{r} and k.

Let $\mathbf{x}$ and $\mathbf{x^a}$ be the n_t-dimensional vector of transmitted signals from the antenna array and its angular domain representation respectively, they are related by

$$\mathbf{x} = \mathbf{U}_\mathrm{t}\mathbf{x^a}, \qquad \mathbf{x^a} = \mathbf{U}_\mathrm{t}^*\mathbf{x}, \tag{4.43}$$

where the entry of $\mathbf{U}_\mathrm{t}$ is

$$\frac{1}{\sqrt{n_\mathrm{t}}}\exp\left(\frac{-\mathrm{j}2\pi kl}{n_\mathrm{t}}\right) \qquad k, l = 0, ..., n_\mathrm{t} - 1. \tag{4.44}$$

Similarly define the transformation for the received signal

$$\mathbf{y} = \mathbf{U}_\mathrm{r}\mathbf{y^a}, \qquad \mathbf{y^a} = \mathbf{U}_\mathrm{r}^*\mathbf{y}, \tag{4.45}$$

where, the entry of $\mathbf{U}_\mathrm{r}$ is

$$\frac{1}{\sqrt{n_\mathrm{r}}}\exp\left(\frac{-\mathrm{j}2\pi kl}{n_\mathrm{r}}\right) \qquad k, l = 0, ..., n_\mathrm{r} - 1. \tag{4.46}$$

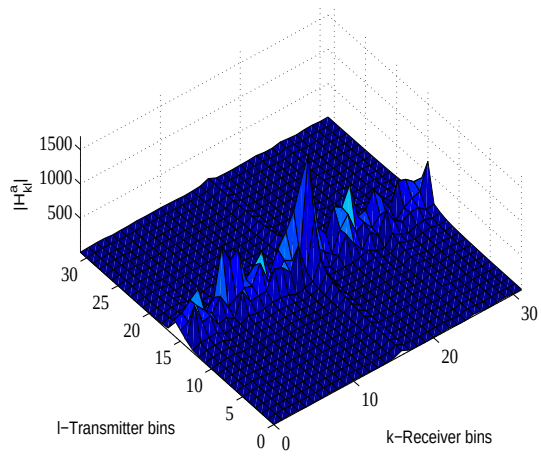
With the above definition, the MIMO channel described in (4.36) can be equivalently represented by

$$\begin{aligned} \mathbf{y^a} &= \mathbf{U}_\mathrm{r}^*\mathbf{H}\mathbf{U}_\mathrm{t}\mathbf{x^a} + \mathbf{U}_\mathrm{r}^*\mathbf{w}, \\ &= \mathbf{U^a}\mathbf{x^a} + \mathbf{w^a}. \end{aligned} \tag{4.47}$$

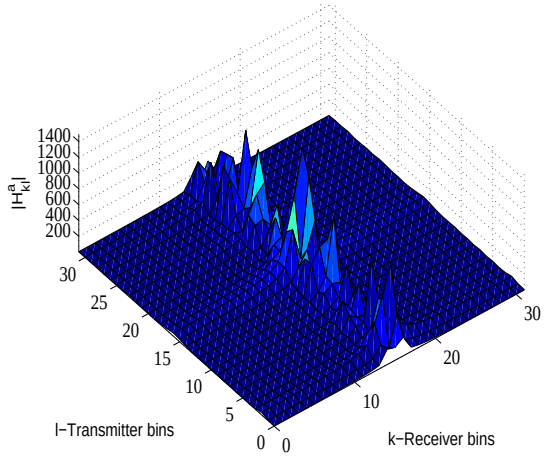
We hence have an angular domain representation of the matrix $\mathbf{H}$:

$$\mathbf{H^a} = \mathbf{U}_\mathrm{r}^*\mathbf{H}\mathbf{U}_\mathrm{t}. \tag{4.48}$$

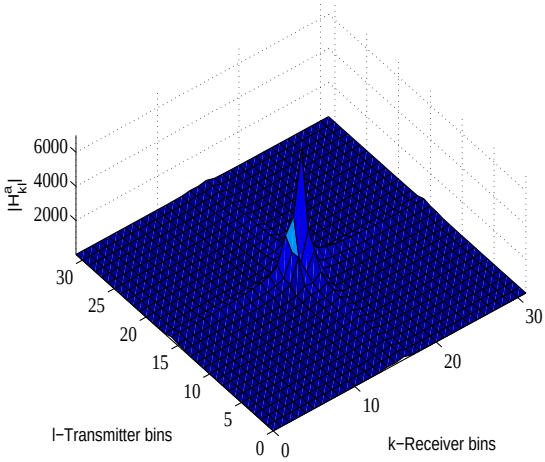
Some simulation results of $\mathbf{H^a}$ are shown in Fig. 4.11. It can be observed that how angular spread contributes to $\mathbf{H^a}$. And we can also conclude that in a rich scattering environment, i.e., there exists a large number of physical paths, the Central Limit Theorem can be invoked and the entries of $\mathbf{H^a}$ can be modelled as complex circular symmetric Gaussian variables. Further discussion about $\mathbf{H^a}$ can be extended to the correlation of the entries, degree of freedom etc.



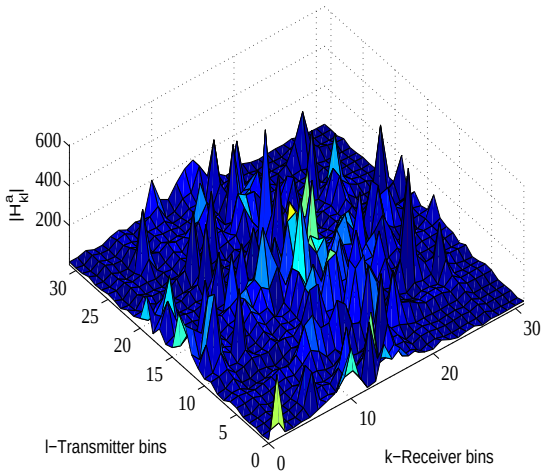
(a) 60° spread at transmitter, 360° spread at receiver.



(b) 360° spread at transmitter, 60° spread at receiver.



(c) 60° spread at transmitter, 60° spread at receiver.



(d) 360° spread at transmitter, 360° spread at receiver.

Figure 4.11: Some examples of $\mathbf{H^a}$.

4.4 Conclusions

In this chapter we presented fading mitigation techniques, including OFDM, MIMO Beamforming. We analyzed the channel capacity gain obtained by these techniques. We studies and constructed a MIMO channel model based on angular domain representation.

Chapter 5

Time-Frequency Signal Analysis and Blind Source Separation

5.1 Introduction to Blind Source Separation

Blind Source Separation (BSS) or blind signal separation is very closely related to the method called *Independent Component Analysis* (ICA). A *source* means here an original signal, i.e., independent component, like the speaker in the cocktail-party problem. *Blind* means that we know very little, if anything, of the mixing matrix, and make very weak assumptions on the source signals. ICA is one method, perhaps the most widely used, for performing blind source separation. ICA is a very general-purpose statistical technique in which observed random data are expressed as a linear transform of components that are statistically independent from each other.

The motivation to investigate blind source sedation is that communication schemes provide several possibilities for applying ICA and BSS in a meaningful way. The time-invariant frequency-selective channel described by (4.34) is termed as convolutive mixtures in the domain of ICA/BSS. Mathematically the CDMA signal model can be cast in the form of a noisy matrix ICA model and ICA/BSS provides a low complexity minimization approach for estimating fading channels [1] [39].

A classical BSS method is based on maximization of non-Gaussianity. Kurtosis is used to measure non-Gaussianity. The fundamental idea is that since a sum of even two independent random variables is more Gaussian than the original variables. The whitening makes the variables closer to Gaussian. The whitening transformation is a subspace method. The subspace method is mathematically clean and elegant. The use of eigenvectors to characterize the signal and noise portions of the observations is one of the cornerstones of modern statistical signal processing.

An alternative approach for BSS is to use time frequency analysis. The advantage of time frequency based method is its ability to recover the non-stationary signals. Moreover, the effects of spreading the noise power while localizing the source energy in the time frequency (t-f) plane amounts to increasing the signal to noise ratio (SNR) and hence improved performance.

5.2 Gaussian Variables are Forbidden

A n-dimensional random vector $\mathbf{x}$ is said to be Gaussian if the probability density function of $\mathbf{x}$ has the form

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}(\det \mathbf{C}_{\mathbf{x}})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^t \mathbf{C}_{\mathbf{x}}^{-1}(\mathbf{x}-\mathbf{m}_{\mathbf{x}})\right), \tag{5.1}$$

where $\mathbf{m}_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\}$ is the mean vector and $\mathbf{C}_{\mathbf{x}} = \mathbb{E}\{(\mathbf{x}-\mathbf{m}_{\mathbf{x}})(\mathbf{x}-\mathbf{m}_{\mathbf{x}})^t\}$ is the covariance matrix.

Consider an example of two-dimensional random vector $\mathbf{x} = (x_1, x_2)^t$. Assume both x_1 and x_2 are Gaussian random variables with zero-mean and variance is 1. We have $\mathbf{m}_{\mathbf{x}} = 0$, and

$$\begin{aligned}\mathbf{C}_{\mathbf{x}} &= \mathbb{E}\{\mathbf{x}\mathbf{x}^t\} \\ &= \mathbb{E}\begin{bmatrix} x_1^2 & x_1x_2 \\ x_2x_1 & x_2^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\end{aligned}$$

therefore, we have

$$\det \mathbf{C}_{\mathbf{x}} = \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1,$$

and $\mathbf{C}_{\mathbf{x}}^{-1} = \mathbf{C}_{\mathbf{x}}$. Substituting all the results into (5.1), we have

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right).$$

With the above result, we now discuss the problem why Gaussian variables are forbidden in ICA/BSS.

Let $\mathbf{x}$ and $\mathbf{y}$ are n-dimensional random vectors that are related by the vector mapping

$$\mathbf{y} = \mathbf{g}(\mathbf{x}), \tag{5.2}$$

for which the inverse mapping is

$$\mathbf{x} = \mathbf{g}^{-1}(\mathbf{y}). \tag{5.3}$$

The density $p_{\mathbf{y}}(\mathbf{y})$ of $\mathbf{y}$ can be obtained from the density $p_{\mathbf{x}}(\mathbf{x})$ of $\mathbf{x}$ as follows:

$$p_{\mathbf{y}}(\mathbf{y}) = \frac{1}{|\det J\mathbf{g}(\mathbf{g}^{-1}(\mathbf{y}))|} p_{\mathbf{x}}(\mathbf{g}^{-1}(\mathbf{y})) \tag{5.4}$$

where $J\mathbf{g}$ is the *Jacobian matrix*. In the special case of $\mathbf{y} = \mathbf{A}\mathbf{x}$, and $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$, so that (5.4) simplifies to

$$p_{\mathbf{y}}(\mathbf{y}) = \frac{1}{|\det \mathbf{A}|} p_{\mathbf{x}}(\mathbf{A}^{-1}\mathbf{y}). \tag{5.5}$$

Let $\mathbf{x}$ be a two-dimensional vector, and $\mathbf{A}$ a 2×2 real-valued matrix, and $\mathbf{V}$ the *whitening matrix* we then have $\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{s} = \tilde{\mathbf{A}}\mathbf{s}$, therefore $\mathbf{s} = \mathbf{A}^{-1}\mathbf{z}$.

Assume that the joint distribution of two independent components, s_1 and s_2, is Gaussian, i.e.,

$$p_{\mathbf{s}}(\mathbf{s}) = p(s_1, s_2) = \frac{1}{2\pi} \exp\left(-\frac{s_1^2 + s_2^2}{2}\right) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{s}\|^2}{2}\right). \tag{5.6}$$

From (5.5), we can find out the pdf of $\mathbf{z}$:

$$\begin{aligned} p_{\mathbf{z}}(\mathbf{z}) &= \frac{1}{\left|\det \tilde{\mathbf{A}}\right|} p_{\mathbf{s}}(\tilde{\mathbf{A}}^{-1}\mathbf{z}) \\ &= \frac{1}{\left|\det \tilde{\mathbf{A}}\right|} \frac{1}{2\pi} \exp\left(-\frac{\|\tilde{\mathbf{A}}^{-1}\mathbf{z}\|^2}{2}\right). \end{aligned} \tag{5.7}$$

Sine $\tilde{\mathbf{A}}$ is orthogonal, it has the following properties, $\tilde{\mathbf{A}}^{-1} = \tilde{\mathbf{A}}^t$, $\|\tilde{\mathbf{A}}\mathbf{z}\|^2 = \|\mathbf{z}\|^2$, and $\left|\det \tilde{\mathbf{A}}\right| = 1$. And finally we have:

$$p_{\mathbf{z}}(\mathbf{z}) = p(z_1, z_2) = \frac{1}{2\pi} \exp\left(-\frac{\|\mathbf{z}\|^2}{2}\right). \tag{5.8}$$

Compare (5.6) and (5.8), we come to the conclusion that the original and mixed distributions are identical. Therefore, there is no way we could infer the mixing matrix from the mixtures.

5.3 Gauss-Markov Process and Whitening Procedure

5.3.1 Whitening in Blind Source Separation

The aim of Blind Source Separation (BSS) is to separate sources without a prior knowledge of the mixture structure. Blind source separation promises wide applications in signal processing and communication systems [39][32]. The first

step (or preprocessing) is whitening. We explore the properties of Gauss-Markov process to investigate the whitening procedures. We propose a Gauss-Markov modelling of the signal source and visualize the whitening procedure geometrically. An information-theoretic analysis and a computer simulation are developed to give intuitive and qualitative insights of the whitening in BSS and enhance the understanding of the interrelation of Gaussian processes, and estimation theory. Blind Source Separation (BSS) is used to recover the original waveforms of sources without a prior knowledge of the mixture's structure. To define BSS, assume that we observe m linear mixtures $x_1, ..., x_m$ of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n, \qquad \forall j. \tag{5.9}$$

where each mixture x_j as well as each independent component s_k is a random variable. This model can be expressed as a compact vector-matrix notation

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{5.10}$$

where $\mathbf{x} = (x_1, x_2, ..., x_m)^t$, $\mathbf{s} = (s_1, s_2, ..., s_n)^t$, and $\mathbf{A}$ being the matrix with elements a_{ij}.

In BSS, we estimate both $\mathbf{A}$ and $\mathbf{s}$ based on the observed random vector $\mathbf{x}$. The term "blind" means we have very little knowledge about the mixing matrix $\mathbf{A}$, and the only assumption made here is that the components of $\mathbf{s}$ are statistically independent.

The BSS of the random vector $\mathbf{x} = (x_1, x_2, ..., x_m)^t$ is obtained by finding an $n \times m$, full-rank, linear transformation (unmixing) matrix $\mathbf{W}$ such that the output signal vector $\mathbf{y} = (y_1, y_2, ..., y_n)^t$, defined as

$$\mathbf{y} = \mathbf{W}\mathbf{x}, \tag{5.11}$$

which contains the components that are the estimates of x_j. The most important preprocessing for BSS is *whitening*. The BSS problem is greatly simplified if the observed mixture vectors $\mathbf{x}$ are first whitened.

The Central Limit Theorem states that the sum of finite-variance independent sources, regardless of their initial distribution, will tend at the limit to form a Gaussian distribution. The histogram of a clean speech signal is decidedly more peaked than Gaussian. A speech signal recorded with a microphone far across a room from the source does appear to be Gaussian. The multipath reflections of the room have summed at the microphone to produce an almost Gaussian signal. Equation (5.9) describes the situation where several microphones in a room picking up the speeches of several human speakers. Here, the Central Limit Theorem "Gaussianization" is due both to the convolution

of distributions from the direct-signal speech signals and also the multipath reflections [32]. We conclude that the observed components x_j in Equation (5.9) are jointly Gaussian with certain correlation. This gives the possibility of modelling the mixed signal source $\mathbf{x}$ using Gauss-Markov processes to investigate the whitening procedure in BSS.

A zero-mean random vector $\mathbf{z} = (z_1, z_2, ..., z_n)^t$ is said to be *white* (or *sphered*) if its elements z_i are *uncorrelated* and have *unit variances*:

$$\mathbb{E}\{z_i z_j\} = \delta_{ij}, \tag{5.12}$$

where $\delta_{ij} = 1$ when $i = j$ and zero otherwise. In terms of the covariance matrix, this means that $\mathbb{E}\{\mathbf{z}\mathbf{z}^t\} = \mathbf{I}$, with $\mathbf{I}$ being the identity matrix. The whitening procedure can be stated as follows: given a random vector $\mathbf{x}$ with n elements defined in (5.9), find a linear transformation $\mathbf{V}$ into another vector $\mathbf{z}$ such that

$$\mathbf{z} = \mathbf{V}\mathbf{x}$$

is white.

After whitening, the blind separation task usually becomes somewhat easier, because the subsequent unmixing matrix $\mathbf{W}$ can be constrained to be an orthogonal matrix for real-valued signals and a unitary matrix for complex-valued signals and weights, that is, $\mathbf{W}\mathbf{W}^t = \mathbf{I}$. Since Equation (5.11) now becomes $\mathbf{y} = \mathbf{W}\mathbf{z} = \mathbf{W}\mathbf{V}\mathbf{x}$, the auto-correlation matrix $\mathbf{R}_\mathbf{y}$ of $\mathbf{y}$ will satisfy

$$\mathbf{R}_\mathbf{y} = \mathbb{E}\{\mathbf{y}\mathbf{y}^t\} = \mathbb{E}\{\mathbf{W}\mathbf{z}\mathbf{z}^t\mathbf{W}^t\} = \mathbf{W}\mathbf{W}^t = \mathbf{I}. \tag{5.13}$$

The whitening transform is given by

$$\mathbf{V} = \mathbf{D}^{-1/2}\mathbf{E}^t, \tag{5.14}$$

where $\mathbf{E} = (\mathbf{e_1}, ..., \mathbf{e}_n)$ is the matrix whose columns are the unit-norm eigenvectors of the covariance matrix $\mathbf{C}_\mathbf{x} = \mathbb{E}\{\mathbf{x}\mathbf{x}^t\}$, $\mathbf{D} = \mathrm{diag}(d_1, ..., d_n)$ is the diagonal matrix of the eigenvalues of $\mathbf{C}_\mathbf{x}$, and $\mathbf{D}^{-1/2}$ is computed by a simple component-wise operation as $\mathbf{D}^{-1/2} = \mathrm{diag}(d_1^{-1/2}, ..., d_n^{-1/2})$. Since $\mathbf{C}_\mathbf{x}$ can be expressed as $\mathbf{C}_\mathbf{x} = \mathbf{E}\mathbf{D}\mathbf{E}^t$, it can be readily verified that $\mathbf{C}_\mathbf{z} = \mathbf{I}$.

An immediate observation is that further repeating of the whitening transform on $\mathbf{z}$ will not go anywhere but another whitened vector. Geometrically, the distribution of the whitened vector $\mathbf{z}$ is invariant to rotations.

Another observation here is that if we remove $\mathbf{D}^{-1/2}$ in the whitening matrix $\mathbf{V}$ defined in (5.14) and let $\mathbf{V} = \mathbf{E}^t$, we will have $\mathbf{C}_\mathbf{z} = \mathbf{D}$ instead of $\mathbf{C}_\mathbf{z} = \mathbf{I}$. This means that the elements of $\mathbf{z}$ are now uncorrelated, but not normalized to unit variances. The intuition we obtained is that the matrix of eigenvectors is rotating $\mathbf{C}_\mathbf{x}$, while the matrix of eigenvalues is scaling $\mathbf{C}_\mathbf{x}$, which is duly implied by the terminology of eigen*vector* and eigen*value*.

5.3.2 Gauss-Markov Modelling of the Signal Source

Gauss-Markov processes are stochastic processes that satisfy the requirement for both Gaussian processes and Markov processes. A Gauss-Markov process can readily be generated by passing a Gaussian process through a linear shift-invariant system. The output and the input of the system will form a random vector with a non-diagonal covariance matrix [91].

To design a random vector of source signals with predefined covariance, let us consider a linear shift-invariant system represented by the difference equation

$$y[n] + a_1 y[n-1] + \cdots + a_P y[n-P] = b_0 x[n] + \cdots + b_Q x[n-Q] \quad (5.15)$$

The mean of the output has the specific form [91]

$$m_y = \frac{\sum_{j=0}^{Q} b_j}{1 + \sum_{i=1}^{P} a_i} m_x \quad (5.16)$$

The covariance functions for (5.15) satisfy the difference equation

$$C_{yx}[l] + a_1 C_{yx}[l-1] + \cdots + a_P C_{yx}[l-P] = b_0 C_x[l] + \cdots + b_Q C_x[l-Q] \quad (5.17)$$

In our design, we consider the following causal first order linear system to generate a Markov process

$$y[n] = -a_1 y[n-1] + b_o x[n] \quad (5.18)$$

where the sequence of $\{x[n]\}$ is a Gaussian process with mean m_x and variance σ_x^2. To be general, we assume a_1 is complex.

To solve (5.16) we obtain the mean

$$m_y = b_0 \frac{m_x}{1 + a_1} \quad (5.19)$$

And from (5.17), the covariance is in the form

$$C_{yx}[l] = C_{xy}^*[-l] = b_0 \sigma_x^2 (-a_1^*)^l u[l] \quad (5.20)$$

where * is the complex conjugate, and $u[l]$ is the unit step function.

Also form (5.17), the autocovariance function of the sequence of $\{y[n]\}$ has this specific form

$$\begin{aligned} C_y[l] &= b_0^2 \frac{\sigma_x^2}{1-|a_1|^2}(-a_1^*)^{-l} \qquad l \leq 0 \\ C_y[l] &= b_0^2 \frac{\sigma_x^2}{1-|a_1|^2}(-a_1^*)^{l} \qquad l > 0. \end{aligned} \tag{5.21}$$

Let $l = 0$, we solve (5.20) and (5.21) to obtain b_0 and a_1 as

$$b_0 = \frac{C_{yx}[0]}{\sigma_x^2} = \frac{C_{xy}^*[0]}{\sigma_x^2} \tag{5.22}$$

$$a_1 = \pm\sqrt{1 - \frac{b_0^2\sigma_x^2}{C_y[0]}} \tag{5.23}$$

Note that in our design we take the negative value of a_1 for the ease of plotting.

Assume that a_1 is real-valued, an inequality can be derived from equation (5.22) and equation (5.23) as follows

$$-1 \leq \frac{C_{xy}[0]}{\sqrt{C_x[0]}\sqrt{C_y[0]}} \leq 1 \tag{5.24}$$

Note that (5.24) is defining the important property of the correlation coefficient (or normalized covariance), that is,

$$-1 \leq \rho = \frac{\mu_{xy}}{\sigma_x\sigma_y} \leq 1 \tag{5.25}$$

where $\sigma_x^2 = C_x[0]$ is the variance of $x[n]$, $\sigma_y^2 = C_y[0]$ is the variance of $y[n]$, and $\mu_{xy} = C_{xy}[0]$ is the covariance of $x[n]$ and $y[n]$.

The obtained results will enable us to model Gauss-Markov signal source with a predefined Hermitian symmetric matrix with elements satisfying the constraint defined by (5.25).

5.3.3 Whitening the Gauss-Markov Source

The discussion in the previous sections suggests a possibility for us to model the signal sources as a Gauss-Markov process during the whitening procedure.

If a Gaussian process is passed through a linear time-invariant (LTI) system, the output of the system is also a Gaussian process. The output random process is actually known to be both a Gaussian random process and a Markov process.

We conclude that the input and the output of the LTI system described in the previous section are jointly Gaussian.

It is convenient to group random variables of the same distribution into a random vector. The density function for a Gaussian random vector can be expressed as

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C_x}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m_x})^t \mathbf{C_x}^{-1}(\mathbf{x}-\mathbf{m_x})\right) \tag{5.26}$$

where $\mathbf{x} = (x_1, x_2, ..., x_N)^t$, and all the components of $\mathbf{x}$ are real-valued random variables.

The correlation of the vector components can be visualized in a geometrical structure, e.g., a three-dimensional plot or a contour plot. We define the contour as

$$(\mathbf{x}-\mathbf{m_x})^t \mathbf{C_x}^{-1}(\mathbf{x}-\mathbf{m_x}) = c \tag{5.27}$$

where c is a positive constant.

A linear transformation of the random vector $\mathbf{x}$ discussed in the previous section will result in a random vector $\mathbf{z}$, whose pdf can be also described as in (5.26) and thus the correlation of the components of $\mathbf{z}$ can be also visualized by a contour plot defined as in (5.27).

The Gauss-Markov signal sources we designed in the previous section provide us with a vector of source signals with a predefined Hermitian symmetric matrix as its covariance matrix. The contour plots or the three-dimensional plots for both $\mathbf{x}$ and $\mathbf{z}$ will thus enable us to visualize the whitening procedure.

5.3.4 Numerical Simulation

We present simulation results for the analysis described in the previous sections. Let the input and output of an LTI system hold the relation defined by (5.18) be

$$x_2[n] = -a_1 x_2[n-1] + b_o x_1[n] \tag{5.28}$$

where $\{x_1[n]\}$ is a wide-sense stationary (WSS) Gaussian process with zero mean and variance $\sigma_1^2 = 0.1$. Let the variance of $\{x_2[n]\}$ be $\sigma_2^2 = .2$, and the correlation coefficient be $\rho_1 = 0.4$.

Note here for the convenience of analysis we use the notation $x_1[n]$ and $x_2[n]$ instead of $x[n]$ and $y[n]$ and drop the time index n due to stationarity to form a vector $\mathbf{x} = (x_1, x_2)^t$. As $x_1[n]$ and $x_2[n]$ are Gaussian random variables, the covariance matrix $\mathbf{C_x}$ of the observed signal vector $\mathbf{x} = (x_1, x_2)^t$ can then be expressed as

$$\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} .1 & -0.0566 \\ -0.0566 & .2 \end{bmatrix}$$

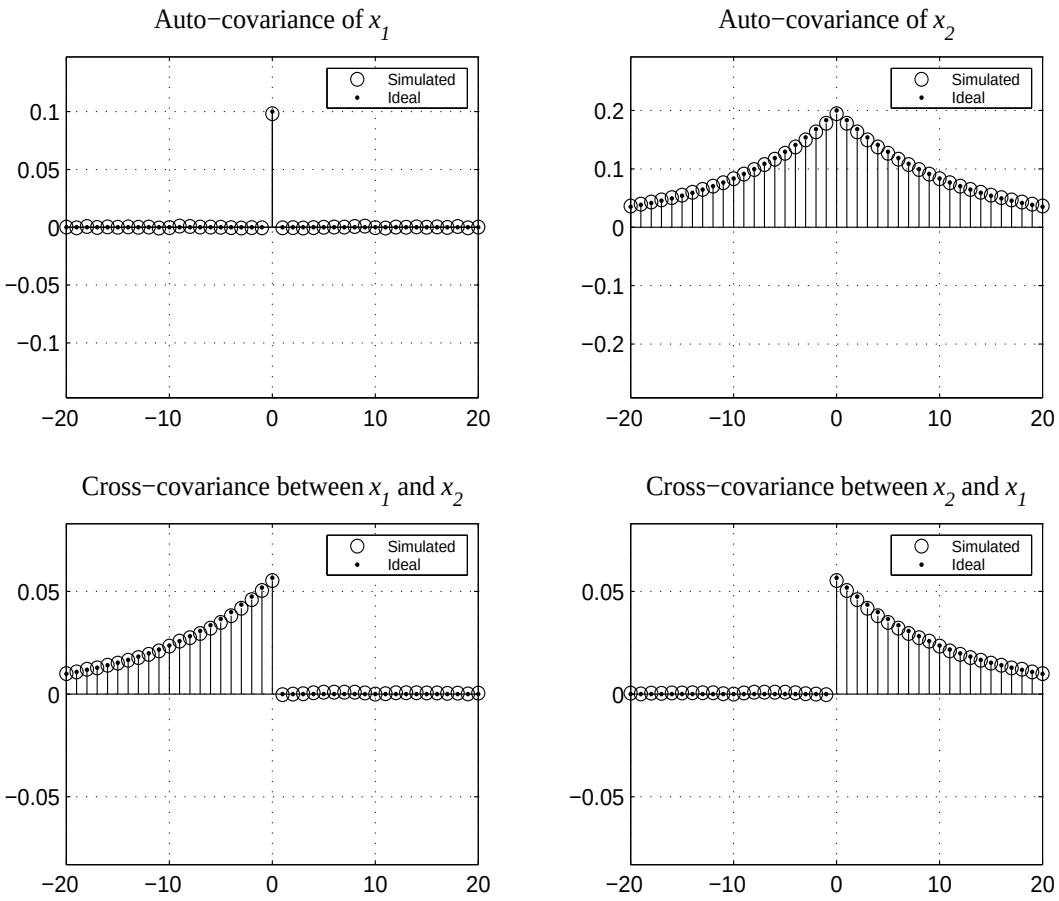


Figure 5.1: Auto-covariance and Cross-covariance between $\{x_1[n]\}$ and $\{x_2[n]\}$.

From (5.23) and (5.22), we calculate $a_1 = -0.9165$ and $b_0 = 0.5657$. The Gauss-Markov process $\{x_2[n]\}$ is then generated by (5.28). To verify the results, we calculate the auto-covariance function and the cross-covariance function of $\mathbf{x} = (x_1, x_2)^t$ and compare with the theoretical results shown in (5.20) and (5.21). The results are plotted in Fig. 5.1. The marginal pdf of $p_{x_1}(x_1)$ and $p_{x_2}(x_2)$ as well as the conditional pdf of $p_{x_1|x_2}(x_1|x_2 = 0.5)$ and $p_{x_1|x_2}(x_1|x_2 = -0.5)$ are plotted in Fig. 5.2. The three-dimensional plot and the contour plot for the joint pdf of x_1 and x_2 are plotted in Fig. 5.3. Using sufficient large sample size, the simulation results match the theoretical curves perfectly.

The fact that the ellipse is tilted with respect to the original coordinates shows that there is a correlation between the vector components x_1 and x_2 of $\mathbf{x}$. A positive slope for the major axis of the ellipse indicates a positive correlation and a negative slope indicates a negative correlation. The principal

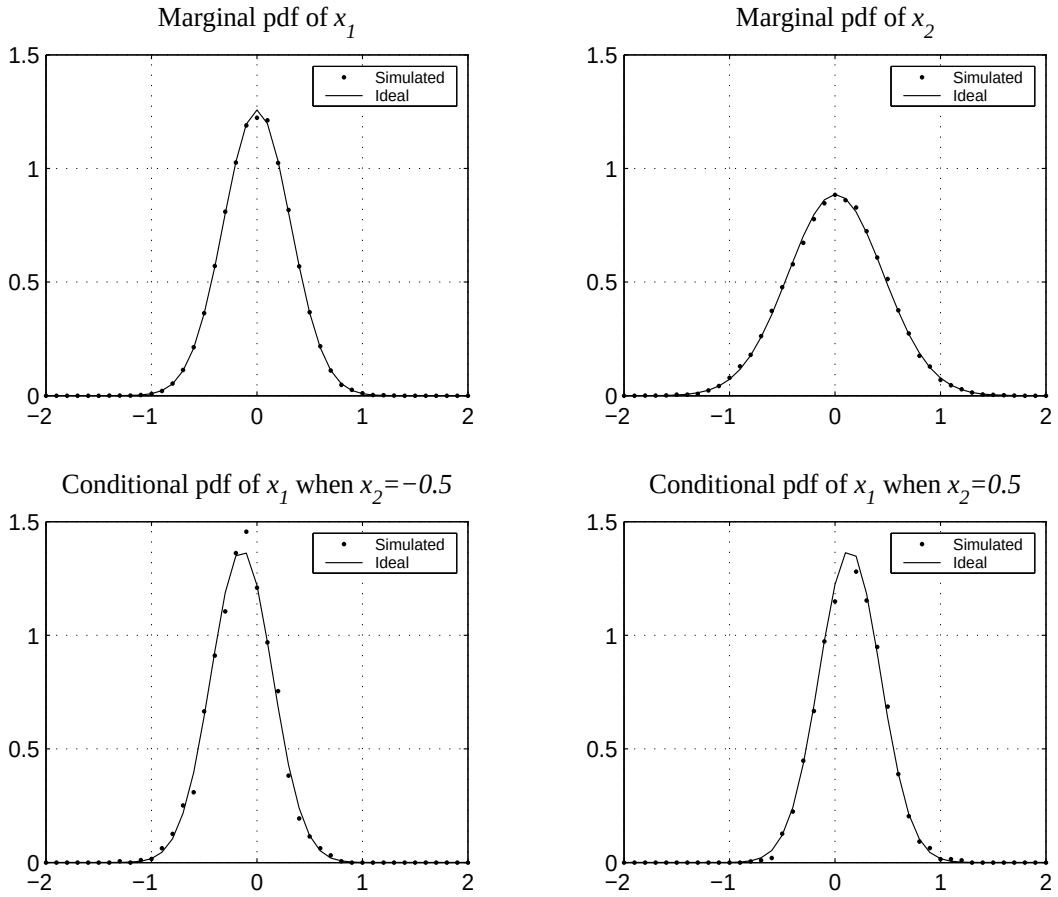


Figure 5.2: Marginal pdf and conditional pdf of x_1 and x_2 of $\mathbf{x}$, where $\mathbf{x} = (x_1, x_2)^t$.

axes of the ellipse are aligned with the eigenvectors, and their sizes are given by $2\sqrt{d_j c}, j = 1, 2$, which is demonstrated in Fig. 5.3 and Fig. 5.4. It can be also observed that Fig. 5.4 and Fig. 5.3 are related by the correlation coefficients as $\rho_2 = -\rho_1$. The ellipses are symmetrical about the vertical axis, and the tilted angles have the relation of $\theta_1 + \theta_2 = 180°$.

We now investigate the whitening procedure. As we discussed in Section II, if the transformation matrix $\mathbf{V} = \mathbf{E}^t$ is applied, the random vector $\mathbf{x} = (x_1, x_2)^t$ is rotated to the random vector $\mathbf{z} = (z_1, z_2)^t$ with its components uncorrelated. Fig. 5.5 shows the geometrical structure of the resulting random vector $\mathbf{z}$. The axes of the ellipse are parallel to the coordinate axes, and hence the components of $\mathbf{z}$ are uncorrelated. If the transformation matrix $\mathbf{V} = \mathbf{D}^{-1/2}\mathbf{E}^t$ is applied, the random vector $\mathbf{x}$ will not only be rotated by the eigenvector matrix $\mathbf{V}$, but also scaled by the eigenvalue matrix $\mathbf{D}$. Fig. 5.6 shows the joint pdf and the

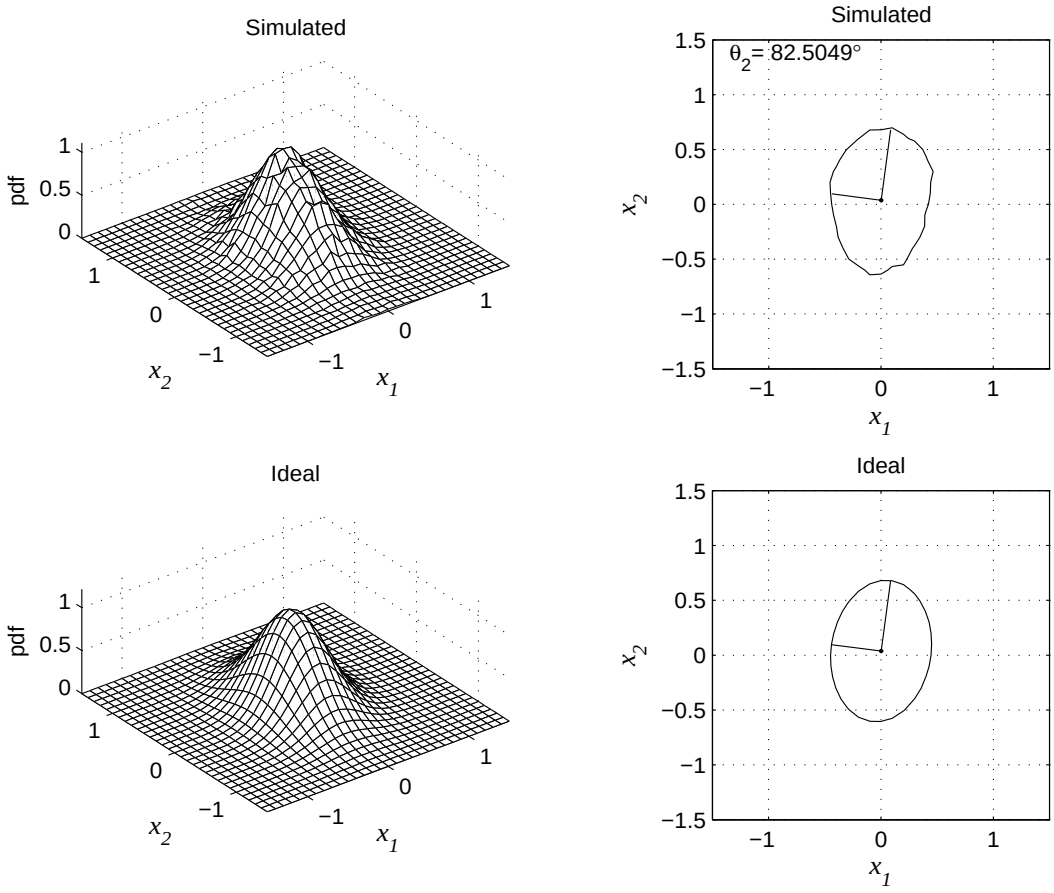


Figure 5.3: The three-dimensional plot and the contour plot of $f_{\mathbf{x}}(\mathbf{x})$ with the correlation coefficient $\rho_1 = 0.4$.

contour plot of the resulting random vector $\mathbf{z}$ which is transformed by $\mathbf{D}^{-1/2}\mathbf{E}^t$. The components of $\mathbf{z}$ are jointly Gaussian of unit variance with zero correlation. The geometric interpretation of Fig. 5.6 is that the transformation involves a rotation and a scaling for the covariance of the random vector $\mathbf{x}$. The contour plot shows a circle instead of an ellipse. It can be noted that in this situation the probability density of $\mathbf{z}$ is less concentrated and thus a higher number of data samples are needed to produce the three-dimensional plot and the contour plot which are close enough to the theoretical curves.

It is thought-provoking to investigate the invariance of the whitened vector $\mathbf{z}$. As shown in Fig. 5.6, the direction of eigenvectors are far from ideal with limited sample size. Repeating whitening transformation recursively will result in a perfect identity matrix of $\mathbf{C}_{\mathbf{z}}$ as shown in Fig. 5.7.

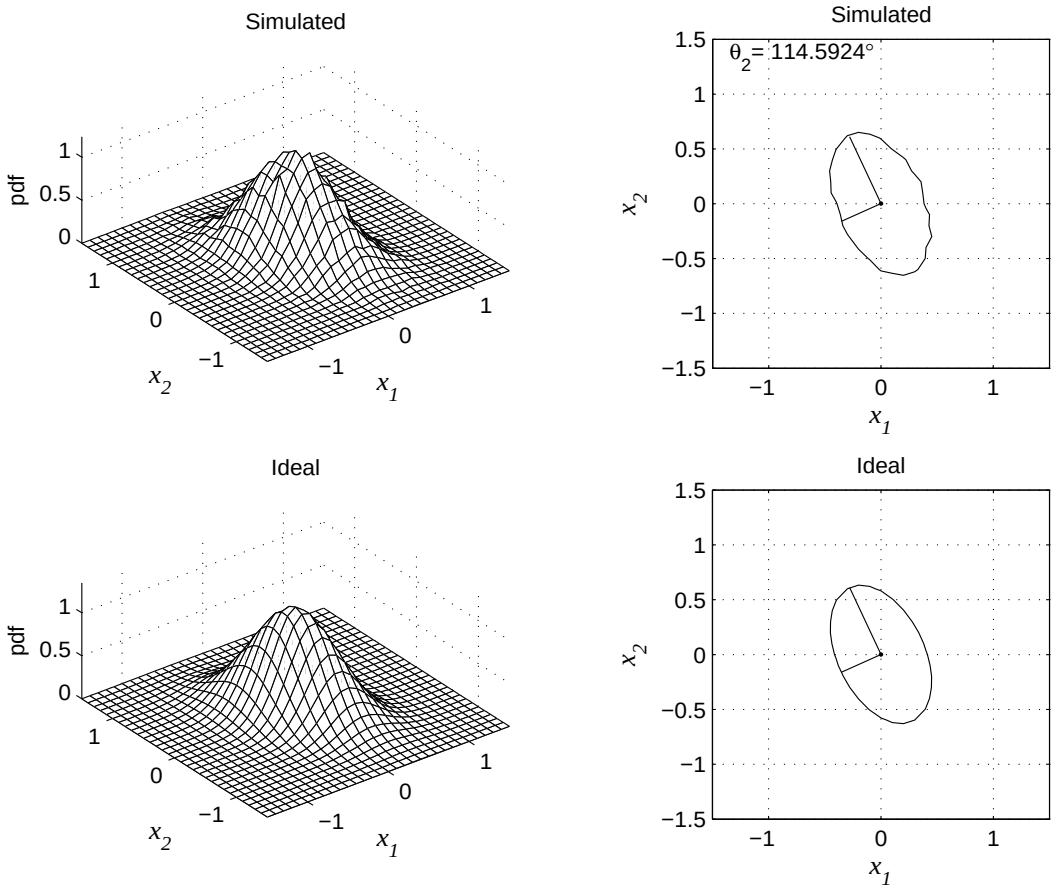


Figure 5.4: The three-dimensional plot and the contour plot of $f_{\mathbf{x}}(\mathbf{x})$ with the correlation coefficient $\rho_2 = -\rho_1 = -0.4$.

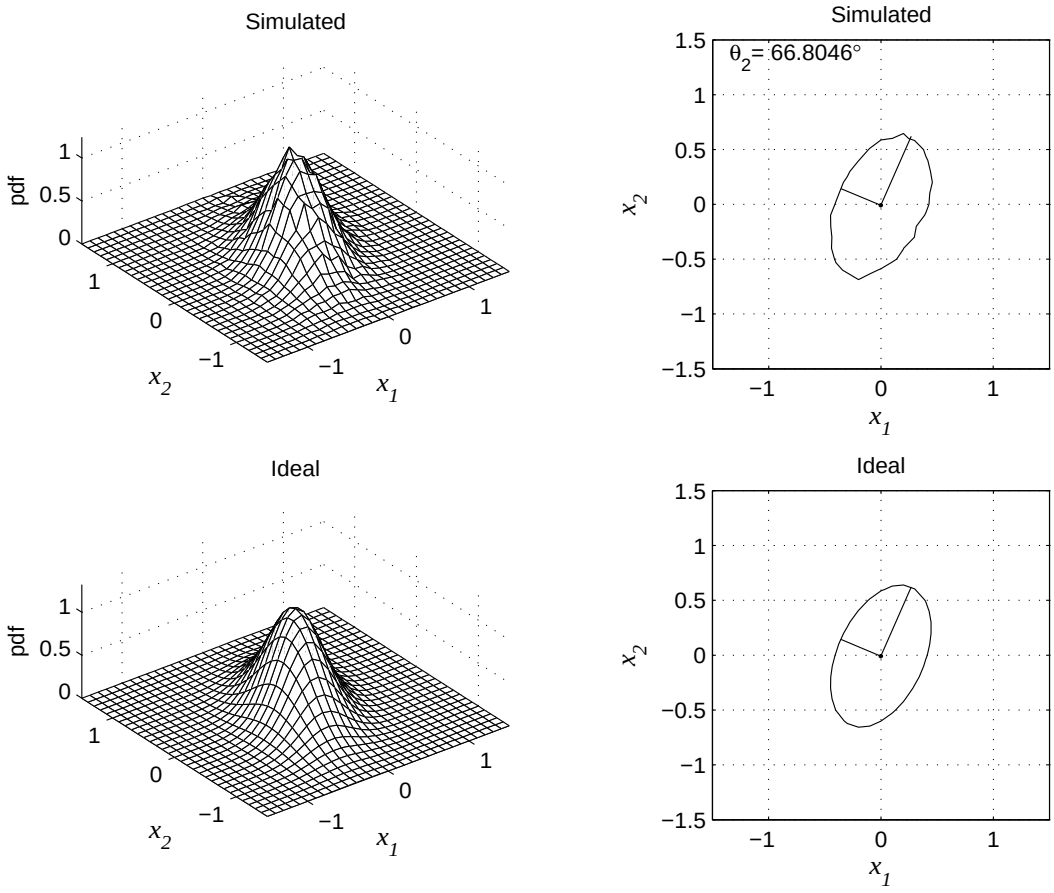


Figure 5.5: Diagonalizing: the transformation matrix $\mathbf{V} = \mathbf{E}^t$, the covariance matrix of $\mathbf{z} = (z_1, z_2)^t$ is the diagonal matrix, i.e., $\mathbf{C}_z = \mathbf{D}$ (uncorrelated).

5.3.5 An Information-Theoretic Analysis of the Whitening Transformation

Assuming wide-sense stationary (WSS) of the random processes we discussed in the previous sections, we present an information-theoretic viewpoint of the whitening transformation.

The mutual information between two random variables X and Y with joint density $f(x, y)$ is defined as [19]

$$I(X;Y) = \int f(x,y) \log \frac{f(x,y)}{f(x)f(y)} dxdy. \tag{5.29}$$

where X is with a density $f(x)$ and Y is with a density $f(y)$.

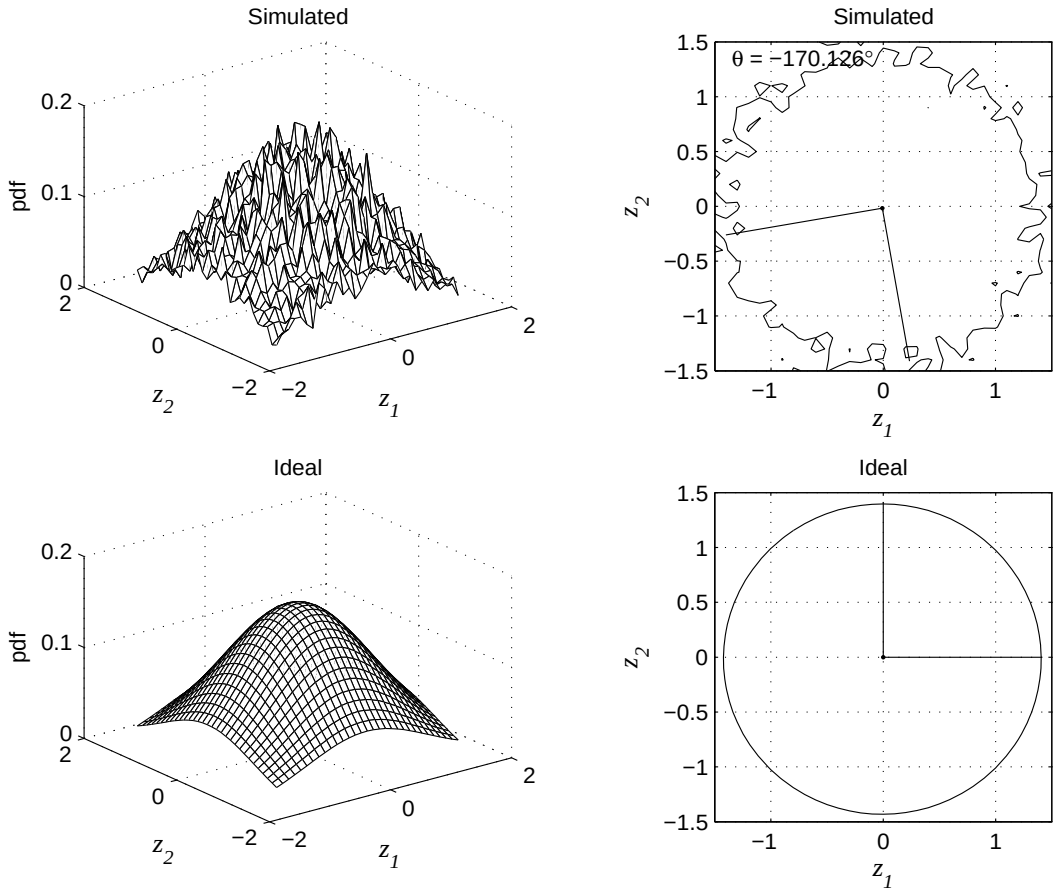


Figure 5.6: Whitening: the transformation matrix $\mathbf{V} = \mathbf{D}^{-1/2}\mathbf{E}^t$, the covariance matrix of $\mathbf{z} = (z_1, z_2)^t$ is the identity matrix, i.e., $\mathbf{C_z} = \mathbf{I}$ (whitened or sphered).

For the correlated Gaussian random variables defined by

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2 \left(0, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

and the mutual information $I(X;Y)$ can be derived from Equation (5.29) as

$$I(X;Y) = \frac{1}{2} \log \frac{1}{1-\rho^2}. \tag{5.30}$$

Other less significant quantities such as the differential entropy $h(X)$ and $h(Y)$, joint entropy $h(X,Y)$, conditional entropy $h(X|Y)$ and $h(Y|X)$ are derived as follows

$$h(X) = \frac{1}{2} \log(2\pi e \sigma_1^2) \tag{5.31}$$

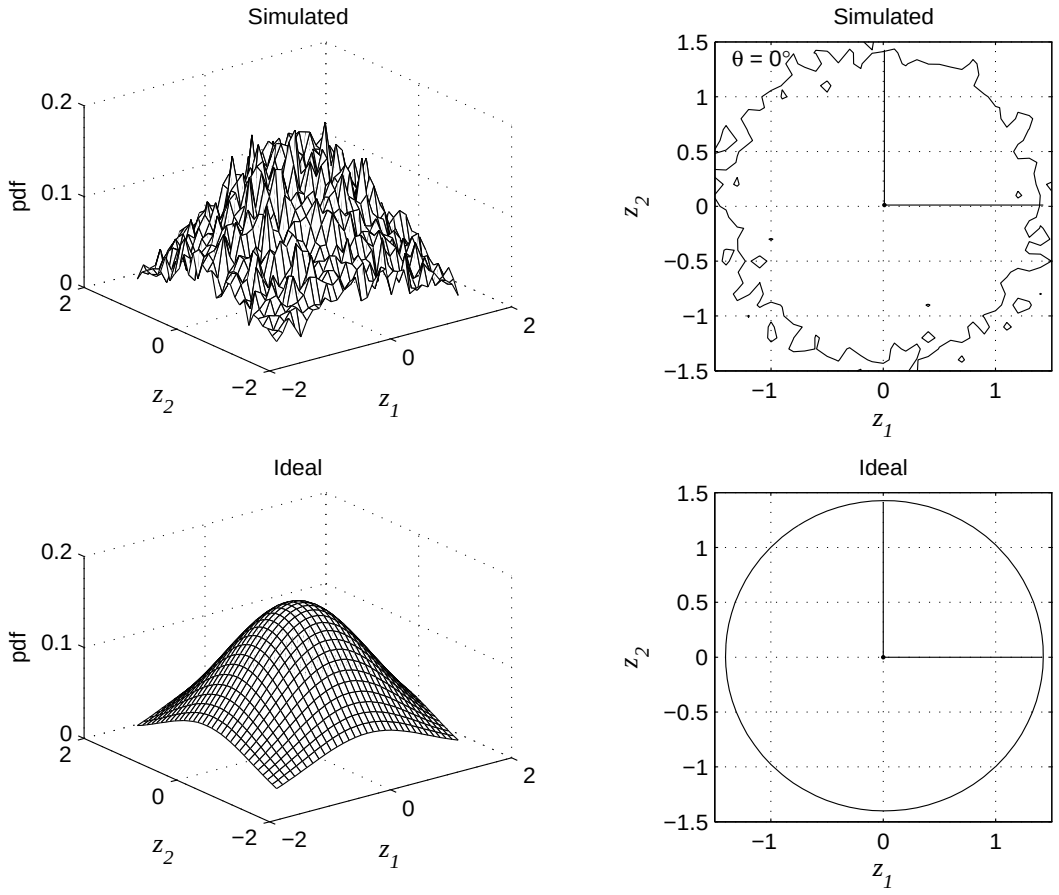


Figure 5.7: Repeating the whitening transformation recursively.

$$h(Y) = \frac{1}{2}\log(2\pi e\sigma_2^2) \tag{5.32}$$

$$h(X,Y) = \frac{1}{2}\log[(2\pi e)^2\sigma_1^2\sigma_2^2(1-\rho^2)] \tag{5.33}$$

$$h(X|Y) = \frac{1}{2}\log[2\pi e\sigma_1^2(1-\rho^2)] \tag{5.34}$$

$$h(Y|X) = \frac{1}{2}\log[2\pi e\sigma_2^2(1-\rho^2)] \tag{5.35}$$

Assuming normalized variances for X and Y, the above quantities are plotted in Fig. 5.6. Equation (5.30) shows that the mutual information between the two correlated Gaussian random variables is independent of the variances σ_1 and σ_2 and is determined by the correlation coefficient ρ only. Moreover, the mutual information $I(X;Y)$ reaches zero when $\rho = 0$. This important result

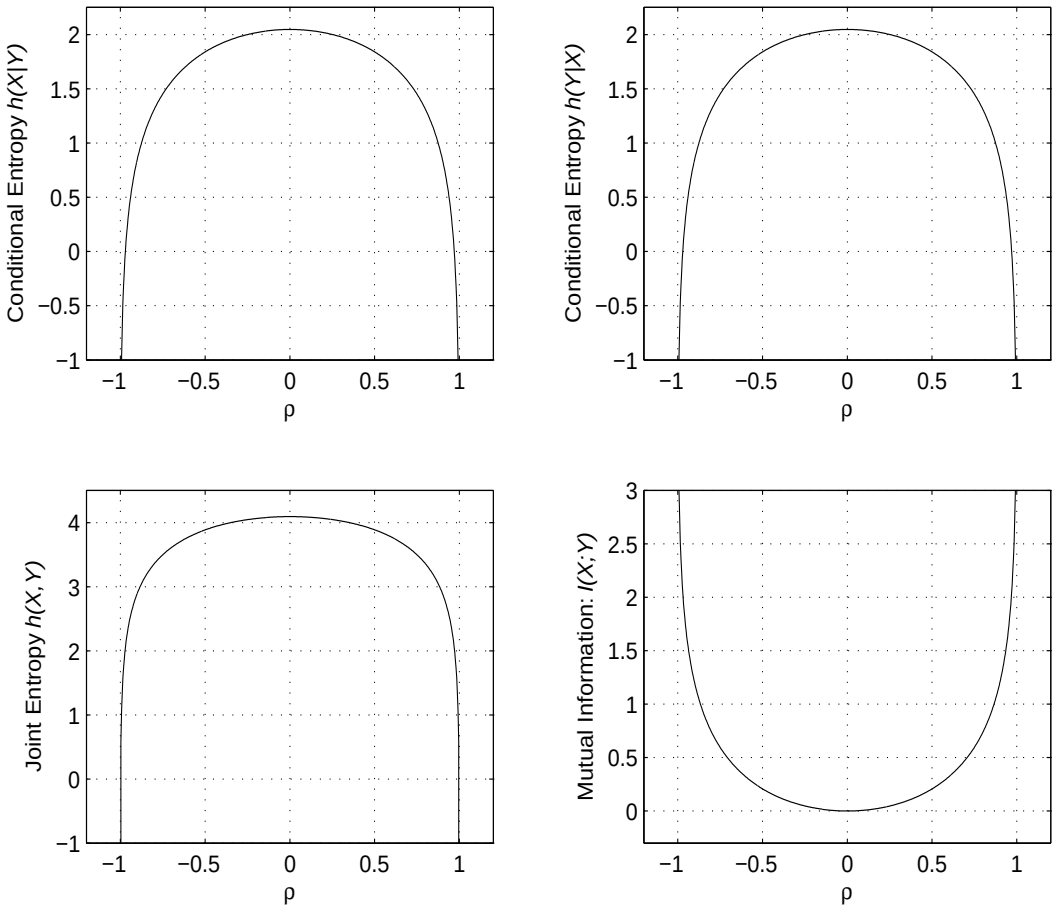


Figure 5.8: Conditional entropy, joint entropy and mutual information.

shows that the whitening transformation minimizes the mutual information of the received signal sources. As we know, mutual information minimization is the implicit or explicit purpose of all the BSS methods.

5.3.6 Conclusions

A Gauss-Markov modelling of the signal source with predefined covariance matrix is presented. The generated Gauss-Markov sources introduce the freedom of arbitrary statistical properties up to the constraint of defining the correlation coefficient. The freedom is exploited to interpret the whitening in BSS geometrically. The visualization of the whitening procedure offers more mathematical tractability for the analysis of BSS algorithms. We also demonstrated that whitening transformation is a way of mutual information minimization.

5.4 Method based on Time-Frequency Signal Representations

5.4.1 Introduction

Blind source separation based on time-frequency distributions (TFD's) allows the separation of Gaussian sources with identical spectral shape but different *t-f* localization properties. As TFD's spread noise power while localizing the source energy in the *t-f* domain, the time-frequency BSS is characterized by robustness of separation and improved overall performance. However, quadratic TFD's, which are the most important class of TFD's suitable for this approach, differ widely in resolution and their ability to reduce the cross-terms that disturb the signal interpretation.

Blind Source Separation (BSS) is used to recover the original waveforms of sources without a prior knowledge of the mixture's structure [8]. The mixture can be convolutive or instantaneously linear. The latter corresponds to a linear memoryless channel. This means that BSS algorithms can be applied to the estimation of communication channels [39].

A time-frequency based BSS approach was presented in [8]. In this section, we investigate the performance of the above approach using the optimal smoothing kernels designed in [37]. A subclass of Cohen's Class of time-frequency distributions, referred to as the T-distributions, was presented in [37] and shown to be suitable for efficient amplitude and instantaneous frequency (IF) estimation of mono- and multi-component FM signals. The exponential and hyperbolic time-only kernels have proven to be efficient in reducing cross-terms while retaining high resolution, with a compromise between these two requirements depending on the selected parameters. We apply these kernels to time-frequency based blind source separation of speech signals, linear FMs and non-linear FMs to investigate possible performance improvements based on optimal kernel selection. The simulation results will show that the T-distributions have advantages over Wigner-Ville and Choi-William distributions in terms of noise performance, stability, and wider optimal dynamic range.

5.4.2 Method Description

We assume that the multidimensional observations of the signal $\mathbf{s}(t)$ are given by

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{5.36}$$

where $\mathbf{x}(t) = [x_1(t), ...x_n(t)]^t$ is the observed noisy instantaneous linear mixture of source signals $\mathbf{s}(t) = [s_1(t), ...s_n(t)]^t$, and $\mathbf{n}(t)$ is the additive noise. The

$m \times n$ matrix $\mathbf{A}$ is the mixing matrix. The assumptions here are that the source signal vector $\mathbf{s}(t)$ is a non-stationary multivariate process with its components mutually uncorrelated, with their cross-correlation equal to zero. The additive noise $\mathbf{n}(t)$ is a stationary, temporally white, zero-mean, complex random process, independent of the source signals.

In addition to the above assumptions, we also take advantage of the indeterminacies in the sources by normalizing the source signals to unit power, this implies that the covariance matrix $\mathbf{R}_s$ of $\mathbf{s}(t)$ is the identity matrix, that is, $\mathbf{R}_s = \mathbf{I}$, so that the covariance matrix $\mathbf{R}_y$ of $\mathbf{y}(t)$ is

$$\mathbf{R}_y \stackrel{\text{def}}{=} \lim_{T\to\infty} \sum_{t=1}^{T} \mathbf{y}(t)\mathbf{y}^*(t) = \mathbf{A}\mathbf{A}^H \tag{5.37}$$

where * denotes the conjugate transpose of a matrix.

The method is a two-step process involving whitening and diagonalizing schemes as presented below:

Whitening: to transform the mixing matrix $\mathbf{A}$ into a unitary matrix.

A whitening matrix $\mathbf{W}$ is applied to the observed $\mathbf{x}(t)$ so that

$$\mathbf{z}(t) \stackrel{\text{def}}{=} \mathbf{W}\mathbf{x}(t) = \mathbf{W}(\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)) = \mathbf{U}\mathbf{s}(t) + \mathbf{W}\mathbf{n}(t)) \tag{5.38}$$

The matrix $\mathbf{U}$ is a unitary matrix, and the matrix $\mathbf{W}$ can be estimated by the following implementation:

1. Estimate the autocorrelation matrix $\hat{\mathbf{R}}$ from data samples. Denote by $\lambda_1, ...\lambda_n$ the n largest eigenvalues and $h_1, ..., h_n$ the corresponding eigenvectors.

2. Under the white noise assumption, an estimate of the noise variance is the average of the smallest eigenvalues of $\hat{\mathbf{R}}$.

We then have the estimation of $\mathbf{W}$ as

$$\hat{\mathbf{W}}(t) = [(\lambda_1 - \hat{\sigma}^2)^{-\frac{1}{2}} h_1, ...(\lambda_n - \hat{\sigma}^2)^{-\frac{1}{2}} h_n]^H. \tag{5.39}$$

Diagonalizing: to jointly diagonalize a set of data-STFD (spatial *t-f* distribution) matrices and retrieve the unitary matrix $\mathbf{U}$.

The STFD matrix is given by $[\mathbf{D}_{xx}(t,f)]_{ij} = D_{x_i y_j}(t,f)$ for $i,j = 1, ...n$, where $D_{x_i y_j}(t,f)$ is the discrete-time form of the Cohen's class of TFD's, given by [18]

$$D_{x_i y_j}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l) x_i(t+m+l) \times x_j^*(t+m-l) e^{-j4\pi fl} \quad (5.40)$$

The matrix $\mathbf{U}$ can then be estimated by the following implementation:

1. Form K matrices by computing the STFD of $z(t)$ for a fixed set of (t_i, f_i) points, $i = 1, ..., K$, corresponding to signal auto-terms.

2. A unitary matrix is then obtained as a joint diagonalizer of the set

$$\{D_{zz}(t_i, f_i) | i = 1, ...K\}.$$

The obtained $\hat{\mathbf{U}}$ and $\hat{\mathbf{W}}$ from the above two steps can then be used to estimate the source signal $s(t)$ and mixing matrix $\mathbf{A}$ as: $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^* \hat{\mathbf{W}} \mathbf{x}(t)$, and $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#} \hat{\mathbf{U}}$, where the superscript $^{\#}$ denotes the Moore-Penrose pseudoinverse.

The performance index is the interference-to-signal ratio (ISR), defined as [8]:

$$\mathcal{I}_{pq} = E|(\hat{\mathbf{A}}^{\#} \mathbf{A})_{pq}|^2. \quad (5.41)$$

I_{pq} measures the ratio of the power of the interference of the q^{th} source to the power of the p^{th} source signal estimated as in (5.41), and the global rejection level is defined as

$$\mathcal{I}_{perf} \stackrel{def}{=} \sum_{q \neq p} I_{pq}. \quad (5.42)$$

5.4.3 Selected Time-Frequency Distributions

As mentioned in the previous section, a TFD that provides a good reduction of the cross-terms is needed to make sure that the off-diagonal elements of the TFD matrix of the sources are negligible and so that a diagonal structure can be maintained.

The time-frequency distribution of the analytic signal $z(t)$ associated with the original real signal $s(t)$ can be expressed as follows [37]

$$\rho(t,f) = \underset{\tau \to f}{\mathcal{F}} [G(t,\tau) \underset{(t)}{*} K_z(t,\tau)] \quad (5.43)$$

where $K_z(t,\tau) = z(t+\frac{\tau}{2}) z(t-\frac{\tau}{2})$ is the instantaneous autocorrelation product, $\mathcal{F}$ is the Fourier transform, and $\underset{(t)}{*}$ denotes time convolution. The time-lag kernel $G(t,\tau)$ completely characterizes the corresponding TFD.

Wigner-Ville distribution (WVD) and Choi-Williams distribution (CWD) are the most famous TFDs of Cohen's Class [18]. Time-only kernels (kernels of the T-distributions) presented in [37] were shown to be more efficient than their two-dimensional counterparts in Cohen's Class in terms of *t-f* resolution and cross-terms reduction, thus potentially they are more efficient in time-frequency based blind source separation. In this work we consider the above TFDs with the following kernels:

1. The constant kernel $G(t,\tau) = 1$, which corresponds to the Wigner-Ville distribution (WVD).

2. The exponential kernel of the Choi-Williams distribution(CWD), defined by
$$G(t,\tau) = \sqrt{\alpha/4\pi\tau^2}e^{-\alpha t^2/4\tau^2} \tag{5.44}$$
where α is a real parameter.

3. The exponential time-only kernel, defined by
$$G(t,\tau) = G_\alpha(t) = \sqrt{\alpha/\pi}e^{-\alpha t^2} \tag{5.45}$$
where α is a real parameter and $\sqrt{\alpha/\pi}$ is the normalization factor.

4. The hyperbolic time-only kernel, defined by
$$G(t,\tau) = G_\alpha(t) = k_\alpha/\cosh^{2\alpha}(t) \tag{5.46}$$
where α is a real positive number and the normalization factor $k_\alpha = \int_{-\infty}^{\infty} 1/\cosh^{2\alpha} dt = \Gamma(2\alpha)/2^{2\alpha-1}\Gamma^2(\alpha)$, Γ stands for the gamma function.

5.4.4 Numerical Simulation

To investigate the potential performance improvements of the T-distributions, we carried out numerical simulations as shown in the following examples. Eight TFD matrices were considered. The corresponding *t-f* points are those of the highest power in the *t-f* domain.

Example 1: The source signals are two speech signals, with sample size $N = 512$. The 3×2 complex mixing matrix is arbitrarily chosen using the Matlab® expression: $\mathbf{A} = \text{rand}(3,2) + j*\text{rand}(3,2)$. Here the Matlab® function rand(3,2) produces 3-by-2 matrix with uniformly distributed random entries with values range between 0 and 1.

One of the T-distributions, the hyperbolic distribution is used to separate these two signals. Fig. 5.9 and Fig. 5.10 illustrate the successful separation in the time domain and time frequency distribution respectively. Note there are three mixed signals because the number of sensors is three. For convenience of illustration, the third one is not plotted here. Notice the reverse order of the estimated signals. This permutation (more obvious with $n > 2$) is caused by the inability to determine the ordering and the phases of the mixture matrix $\mathbf{A}$, hence any permutation of the estimated sources is also a satisfactory solution. We can also notice the phase shift of the recovered signals because the mixing matrix $\mathbf{A}$ is a complex matrix.

Example 2: Because of the non-linearity with the frequency components in speech signals, we use two non-linear FMs to investigate the performance of different kernels. The source signals are two non-linear FM (parabolic law) signals with total length $N = 512$ and instantaneous frequencies given according to:

$s(t) = e^{\mathrm{j}2\pi(A_0 t + \frac{A_1}{2}t^2 + \frac{A_2}{3}t^3)}$.

The 3×2 complex mixing matrix is arbitrarily chosen using Matlab® expression: $\mathbf{A} = \mathrm{rand}(3,2) + j * \mathrm{rand}(3,2)$. This time the exponential distribution is used to separate these two signals.

The separation process is similar to that shown in Fig. 5.9 and Fig. 5.10. Fig. 5.11 shows the successful separation in time frequency distribution.

All the four above mentioned distributions are used to separate these two non-linear FMs to compare the performance. Fig. 5.12 shows the mean rejection level versus different values of α for the above-mentioned four kernels.

These two non-linear FMs are close enough together that the Wigner-Ville distribution has failed to separate them because of the large cross-terms.

It can be concluded that the two T-distributions provide a wider range of optimal values. The simulation also shows that the T-distributions are more stable and less signal-dependent in comparison to Wigner-Ville and Choi-Williams distributions. Fig. 5.12 also demonstrates the fact that the two T-distributions converge to the Wigner-Ville case as α becomes large. This is in accord with the fact that as α increases, the T-distributions converge to WVD [38].

The optimal parameter value α_{opt} (in terms of the mean rejection level) for each TFD is then obtained from Fig. 5.12. For Choi-Williams distribution, $\alpha_{\mathrm{opt}} = 0.05$, for the hyperbolic T-distribution $\alpha_{\mathrm{opt}} = 0.15$, and for the exponential T-distributions, $\alpha_{\mathrm{opt}} = 0.3$. We then obtain the noise performance of the different TFDs as shown in Fig. 5.13. The mean rejection levels are evaluated here over 100 Monte-Carlo runs. The performance improvement using time-only kernels is clearly observed in Fig. 5.13.

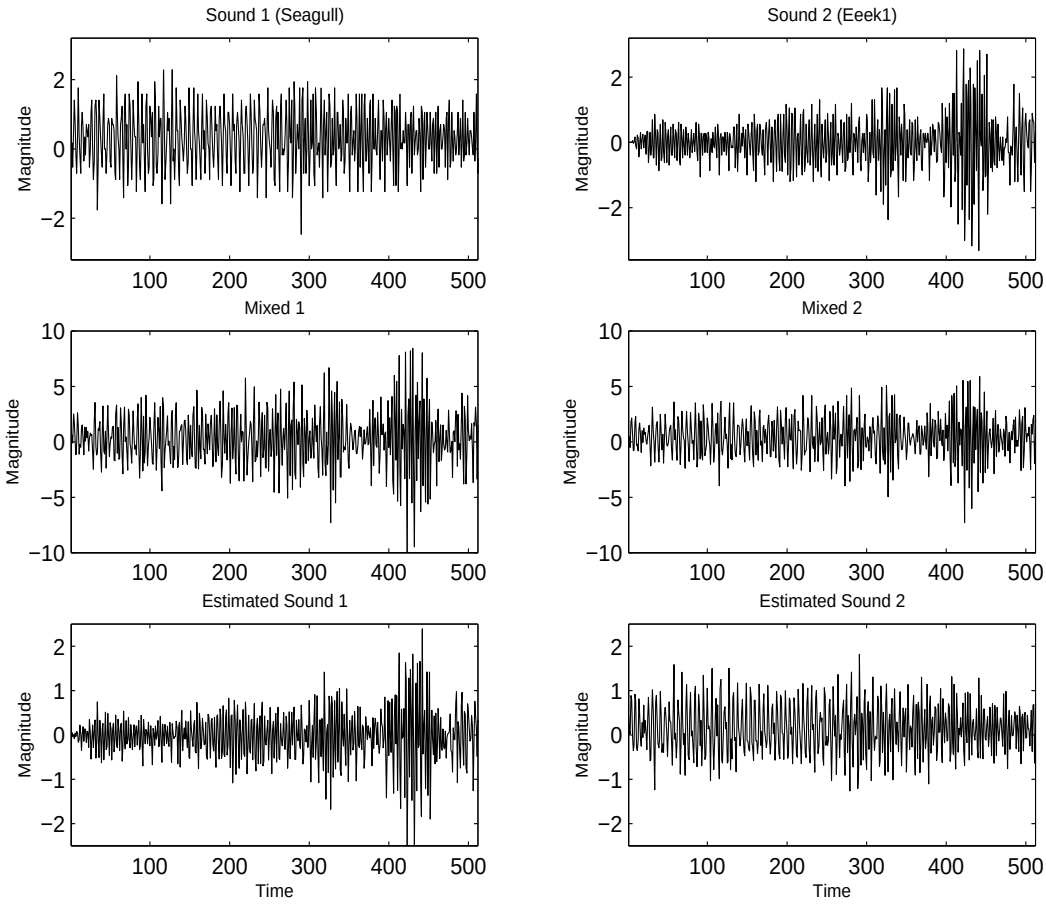


Figure 5.9: Time-domain signals for the blind source separation process using the hyperbolic T-distribution, with SNR = 50 dB. First row: two test sound signals. Middle row: signals resulting from mixing the above sound signals using a randomly-generated mixing matrix. Last row: blindly separated two sound signals (compare with the first row).

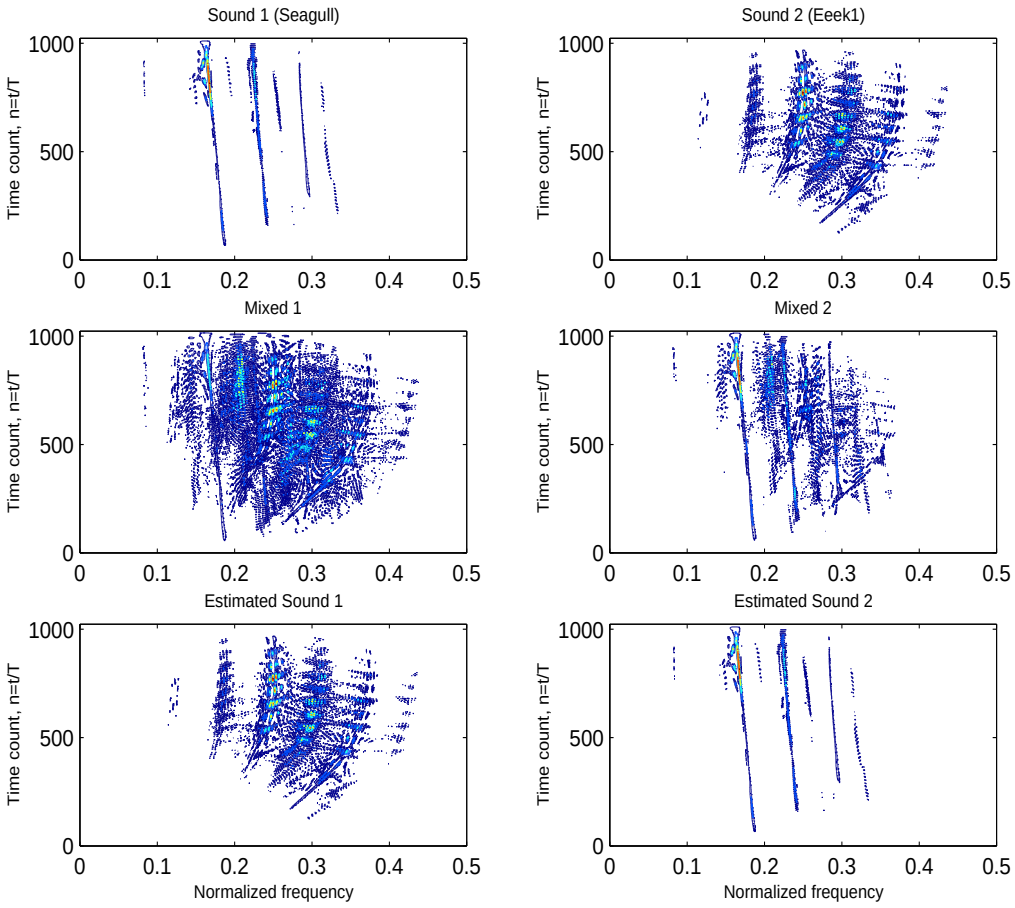


Figure 5.10: Time-frequency distribution of the signals shown in Fig. (5.9) for the blind source separation process of sound signals using the hyperbolic T-distribution. The mean rejection level (MRL) obtained is -23.8 dB.

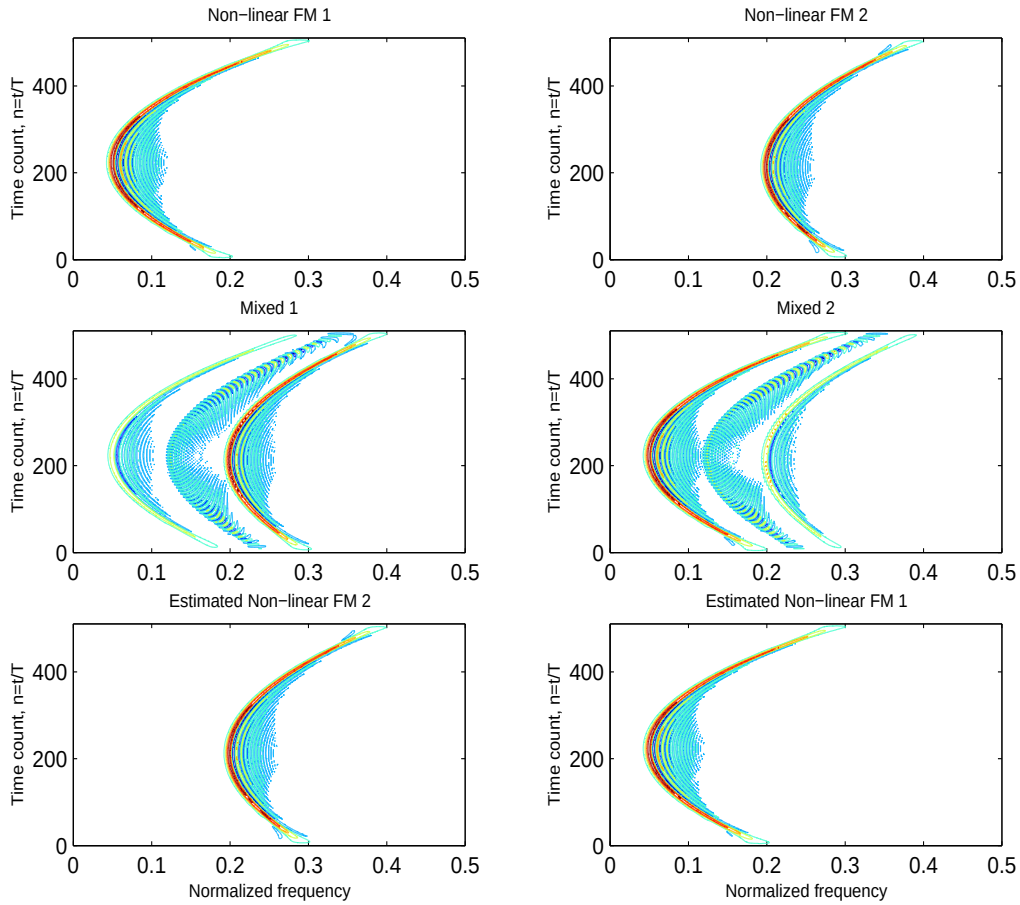


Figure 5.11: Time-frequency distribution of the signals for the blind source separation non-linear FMs using exponential distribution. The mean rejection level (MRL) obtained is -40.3 dB.

Example 3: In this example, we validate the effects of sample size to the rejection level mentioned in [8]. One of the source signals is a non-linear FM and the other is a linear FM, with sample size $N = 1024$. Again, the mixing matrix is a 3×2 complex matrix, arbitrarily chosen using the Matlab® expression: $\mathbf{A} = \text{rand}(3, 2) + j * \text{rand}(3, 2)$.

Fig. 5.13 shows the separation process by Choi-Williams distribution. Fig. 5.14 shows that a larger number of *t-f* points improves the performance.

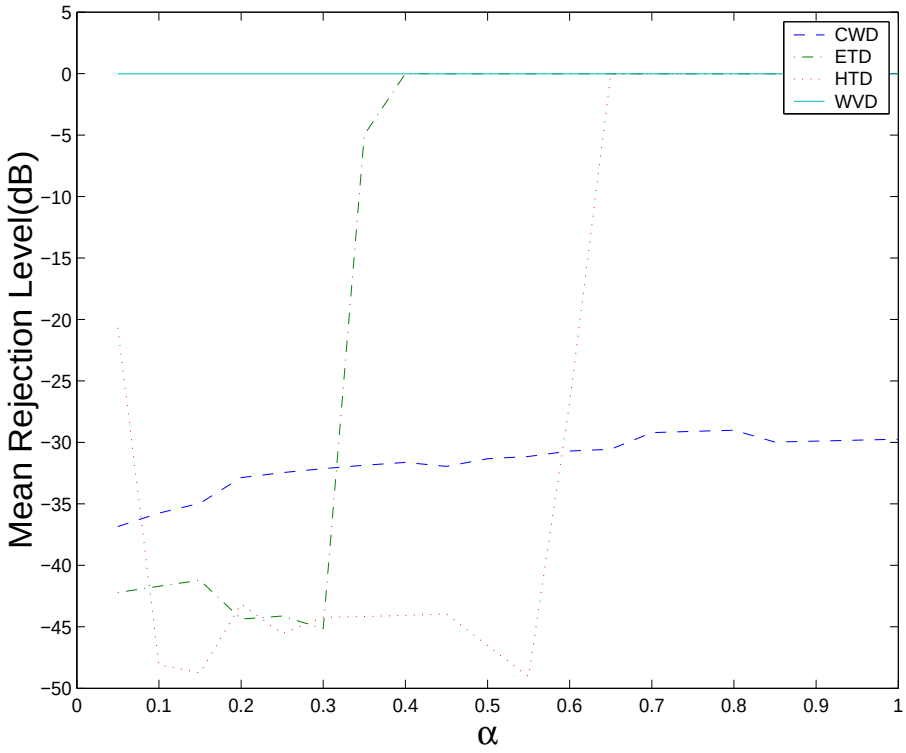


Figure 5.12: Performance of different TFD's versus their parameter α in terms of the mean rejection level. Notice that the T-distributions converge in the limit to WVD.

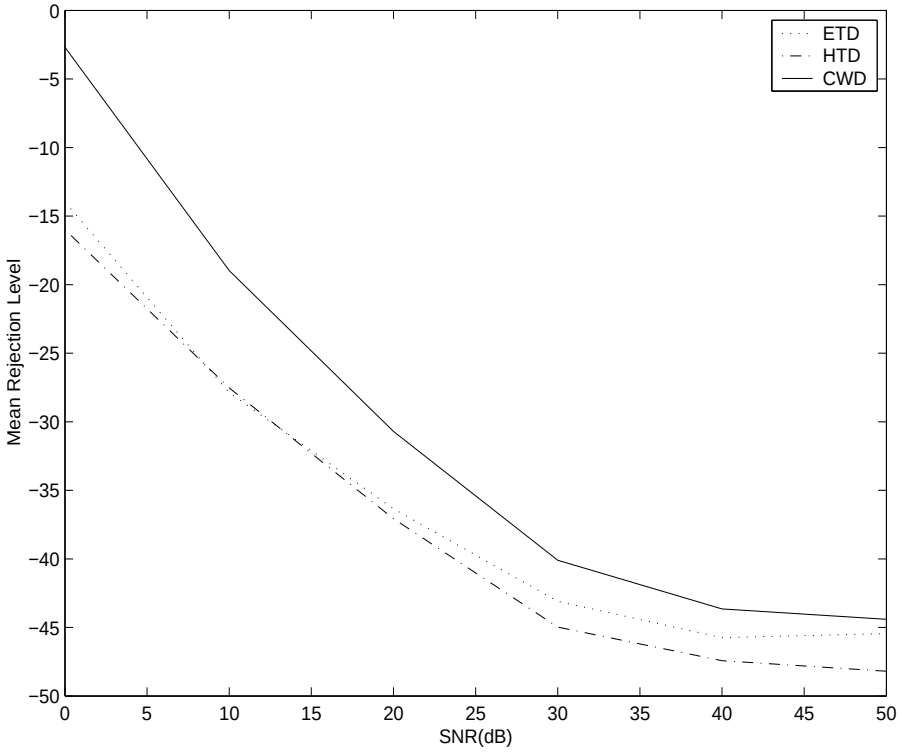


Figure 5.13: Performance of different TFDs versus SNR in terms of the mean rejection level.

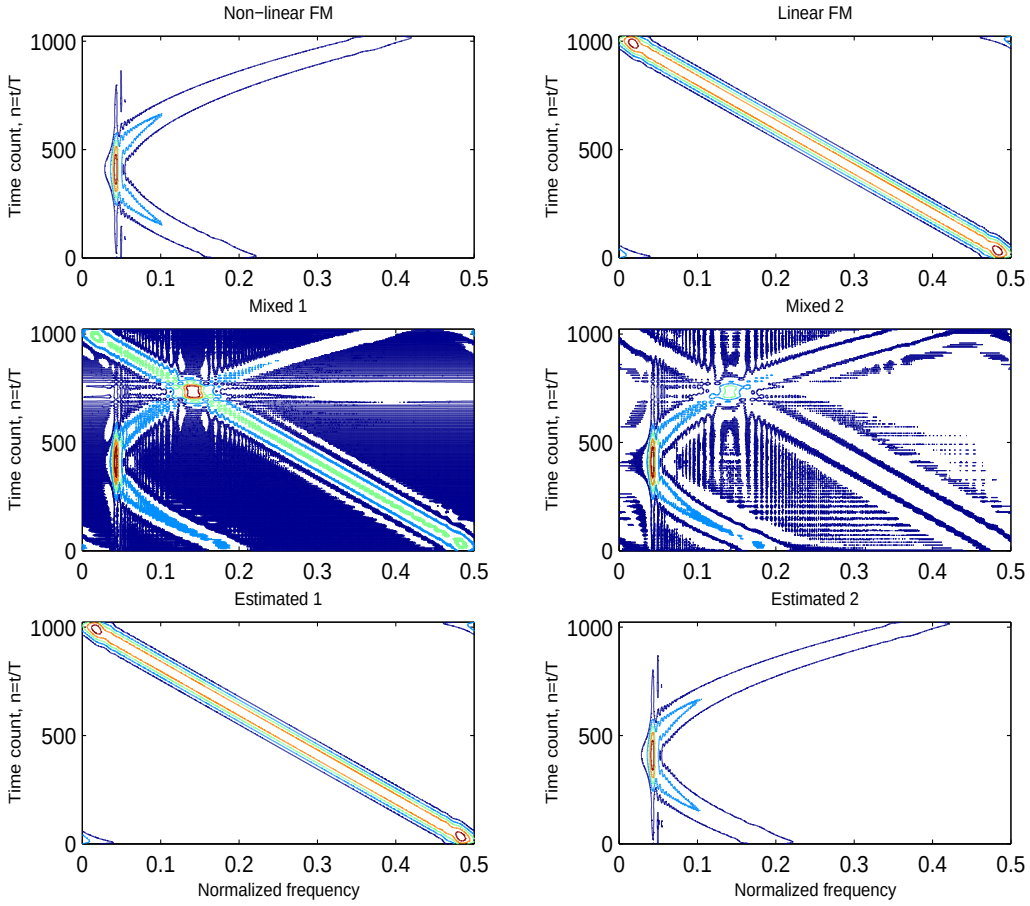


Figure 5.14: Time-frequency distribution of the signals used for the blind source separation process (one is a linear FM signal and the other is a non-linear FM signal). Choi-Williams distribution has been used with $\alpha = 11$. The mean reject level (MRL) obtained is -28.9dB.

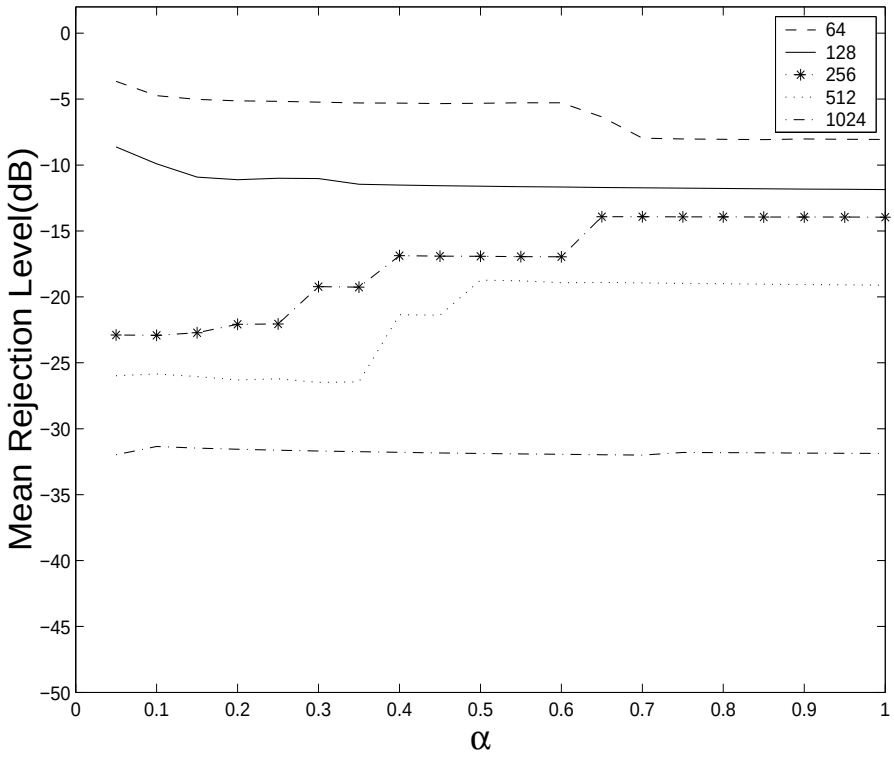


Figure 5.15: Performance validation versus α with different sample sizes for signals in Fig. 5.14.

5.4.5 Conclusions

A comparative study has been done using some well-known and recently-proposed TFDs.

It is shown that the choice of the time-frequency distribution (TFD) has a direct impact on the performance of a recently proposed *t-f* based blind source separation (BSS). It is also true that the above *t-f* based method is highly signal-dependent.

It is found that the T-distributions (TFDs with time-only kernels) have some advantages over Choi-Williams and Wigner-Ville distributions in the above *t-f* based BSS. The T-distributions have wider and more stable optimal ranges (less signal-dependent); also better noise performance than WVD or CWD for the separation of non-linear FMs.

The optimal value of α is signal-dependent and is obtained by sweeping through all possible values by computer simulations. For the T-distributions the optimal values are easier to obtain than for Choi-Williams distribution.

5.5 Applications to Wireless Communications

ICA/BSS is closely related to *blind deconvolution* or *blind equalization* in communication system design. Although our discussion on BSS and TFRs so far is concentrated on solving instantaneous mixture of speech signal processing, the concepts and knowledge gained can be readily extended to wireless communication systems. We now show that ICA/BSS and MIMO systems are intimately related.

The time-invariant frequency-selective channel described by (4.34) is termed as convolutive mixtures in the domain of ICA/BSS. Consider the MIMO channel (4.34) without noise:

$$\mathbf{y}[m] = \sum_{\ell=0}^{L-1} \mathbf{H}_\ell \mathbf{x}[m-\ell], \tag{5.47}$$

or, equivalently,

$$y_i[m] = \sum_{j=1}^{n_\mathrm{t}} \sum_{\ell=0}^{L-1} h_{\ell,ij} x_j[m-\ell] \quad \text{for } i = 1, \ldots, n_\mathrm{r} \tag{5.48}$$

where $h_{\ell,ij}$ is the ith row and jth column entry of $\mathbf{H}_\ell$, $\mathbf{x}[m] = [x_1[m], \ldots, x_{n_\mathrm{t}}[m]]^t$ is the n_t inputs, $\mathbf{y}[m] = [y_1[m], \ldots, y_{n_\mathrm{r}}[m]]^t$ is the n_r outputs. Rearrange the input/output relation to an instantaneous mixture:

$$\tilde{\mathbf{y}}[m] = \tilde{\mathbf{H}} \tilde{\mathbf{x}}[m] \tag{5.49}$$

where

$$\tilde{\mathbf{x}}[m] = [x_1[m], \ldots, x_1[m-(L+L')], \ldots, x_{n_\mathrm{t}}[m-(L+L')]]^t \tag{5.50}$$

$$\tilde{\mathbf{y}}[m] = [y_1[m], \ldots, y_1[m-L'+1], \ldots, y_{n_\mathrm{r}}[m-L'+1]]^t \tag{5.51}$$

and,

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}_{11} & \cdots & \tilde{\mathbf{H}}_{1n_\mathrm{t}} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{H}}_{n_\mathrm{r}} & \cdots & \tilde{\mathbf{H}}_{n_\mathrm{r} n_\mathrm{t}} \end{bmatrix} \tag{5.52}$$

with

$$\tilde{\mathbf{H}}_{ij} = \begin{bmatrix} h_{0,ij} & \cdots & h_{L-1,ij} & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \cdots & h_{0,ij} & \cdots & h_{L-1,ij} \end{bmatrix} \tag{5.53}$$

The relation of (5.49) is the classic matrix representation of the sum of a bank of filters and can be also interpreted as a standard linear ICA model. One can

now apply ordinary ICA methods to this standard linear ICA model. Note that the number of receive antennas has to be greater than the number of transmit antennas, i.e., $n_{\mathrm{r}} > n_{\mathrm{t}}$. Also $\tilde{\mathbf{H}}$ is a $[n_{\mathrm{r}}L' \times n_{\mathrm{t}}(L + L' - 1)]$ matrix and $\tilde{\mathbf{H}}_{ij}$ are $[L' \times (L + L' - 1)]$ matrices. L' is chosen such that $n_{\mathrm{r}}L' \geq n_{\mathrm{t}}(L + L' - 1)$.

A possible application in this case is blind estimation of the channel matrix $\mathbf{H}_\ell$. As we have discussed in the previous section, to achieve the blind estimation of $\mathbf{H}_\ell$ the sources $\mathbf{x}[m]$ are assumed to have different structures and localization properties in the time-frequency domain and $\tilde{\mathbf{H}}$ is full column rank.

A more meaningful application of ICA/BSS is CDMA systems. In the similar manner discussed above, the CDMA signal model can be cast in the form of a noisy matrix ICA model. ICA/BSS will then provide a low complexity minimization approach for estimating fading channels or estimating the desired user's symbol process.

Assume there are K users in the CDMA cell. The data of the kth user are encoded into two BPSK sequences $\{a_k^{\mathrm{I}}[m]\}$ and $\{a_k^{\mathrm{Q}}[m]\}$, which can be assume to have equal amplitude for all m. Each sequence is modulated by a pseudonoise sequence or chip sequence, so that the transmitted complex sequence is

$$x_k[m] = a_k^{\mathrm{I}}[m]s_k^{\mathrm{I}}[m] + ja_k^{\mathrm{Q}}[m]s_k^{\mathrm{Q}}[m], \quad m = 1, 2, \ldots, \tag{5.54}$$

where $s_k^{\mathrm{I}}[m]$ and $s_k^{\mathrm{Q}}[m]$ are chip sequence taking values ± 1. The length of the chip sequence $s_k^{\mathrm{I}}[m]$ or $s_k^{\mathrm{Q}}[m]$ is denoted by C.

We can only discuss the in-phase part of $x_k[m]$, let

$$x_k[m] = a_k[m]s_k[m], \quad m = 1, 2, \ldots, \tag{5.55}$$

and define the chip sequence

$$s_k[m] = \begin{cases} \pm 1, & m = 1, 2, \ldots, C, \\ 0 & \text{elsewhere.} \end{cases} \tag{5.56}$$

The ith symbol of the kth user is denoted by b_{ki}, and we have

$$a_k[m] = b_{ki}, \quad m = 1, 2, \ldots, \quad \text{and} \quad i = (m \mod C) + 1 \tag{5.57}$$

Note that the rate for the sequence $x_k[m]$ is *chip rate*, m is therefore called *chip time*. The date rate for the symbols b_{ki} is much slower than the chip rate. For example, in IS-95, the chip rate is 1.2288 MHz and the data rate is 9.6 kbits/s or less, and the the length of the chip sequence is $C = 1228.8/9.6 = 128$.

The transmitted sequence of user k goes through a discrete-time baseband equivalent multipath channel $h^{(k)}$ and is superimposed at the receiver:

$$y[m] = \sum_{k=1}^{K} \left(\sum_{\ell} h_\ell^{(k)}[m] x_k[m - \ell] \right) + w[m]. \tag{5.58}$$

Assume a downlink transmission of N data symbols and the coherence time of the channel is greater than the data symbol period but smaller than the transmission time of N data symbols, the signal at the receiver expressed in (5.58) becomes

$$y[m] = \sum_{i=0}^{N-1} \sum_{k=1}^{K} b_{ki} \sum_{\ell=0}^{L-1} h_{\ell i} s_k[m - iC - d_\ell] + w[m]. \tag{5.59}$$

where $L =$ is the number of the multipaths, and $d_\ell \in \{0, \ldots, \lfloor (C-1)/2 \rfloor\}$ is the time delay for ℓth path.

Define the $C-$ length vectors as

$$\mathbf{y}_i = (y[iC+1], y[iC+2], \ldots, y[(i+1)C])^t, \tag{5.60}$$

we can rearrange (5.59) in vector form [39]

$$\mathbf{y}_i = \sum_{k=1}^{K} \sum_{\ell=0}^{L-1} [h_{\ell i} b_{ki} \underline{\mathbf{g}}_{k\ell} + h_{\ell,i+1} b_{k,i+1} \bar{\mathbf{g}}_{k\ell}] + \mathbf{w}_i, \tag{5.61}$$

where $\mathbf{w}_i$ denotes the noise vector consisting of C samples of noise. The vector $\underline{\mathbf{g}}_{k\ell}$ denotes the "early" part of the chip sequence, is given by

$$\underline{\mathbf{g}}_{k\ell} = [s_k[C - d_\ell + 1], \ldots, s_k[C], \mathbf{0}^t_{(C-d_\ell)}]^t \tag{5.62}$$

and $\bar{\mathbf{g}}_{k\ell}$ the "late" part, is given by

$$\bar{\mathbf{g}}_{k\ell} = [\mathbf{0}^t_{(d_\ell)}, s_k[1], \ldots, s_k[C - d_\ell]]^t \tag{5.63}$$

where $\mathbf{0}^t_{(C-d_\ell)}$ is row vector having $(C - d_\ell)$ zeros as its elements, and $\mathbf{0}^t_{(d_\ell)}$ is row vector having d_ℓ zeros as its elements. The reason to use $\underline{\mathbf{g}}_{k\ell}$ and $\bar{\mathbf{g}}_{k\ell}$ is to reflect the fact that multipath will generally affect two adjacent data symbols $b_{k,i}$ and $b_{k,i+1}$, which can be observed from (5.59).

$$\mathbf{Y} = [\mathbf{y}_0, \mathbf{y}_2, \ldots, \mathbf{y}_{N-1}], \tag{5.64}$$

$$\mathbf{Y} = \mathbf{GF} + \mathbf{W} \tag{5.65}$$

where the $C \times 2KL$ matrix $\mathbf{G}$ contains all the KL early and late parts of the chip sequence

$$\mathbf{G} = [\underline{\mathbf{g}}_{11}, \bar{\mathbf{g}}_{11}, \ldots, \underline{\mathbf{g}}_{K,L-1}, \bar{\mathbf{g}}_{K,L-1}] \tag{5.66}$$

and the $2KL \times N$ matrix $\mathbf{F} = [\mathbf{f}_0 \ldots \mathbf{f}_{N-1}]$ contains the data symbols and fading terms

$$\mathbf{f}_i = [h_{0i} b_{0i}, h_{0,i+1} b_{0,i+1}, \ldots, h_{L-1,i} b_{Ki}, h_{L-1,i} b_{Ki}]^t \tag{5.67}$$

The vector $\mathbf{f}_i$ represents the $2KL$ data symbols and fading terms of all the users and paths corresponding to the ith pair of early and late part of the chip sequence. Also $\mathbf{W}$ is the $C \times 2KL$ noise matrix.

It is reasonable to assume all the multipaths are independent of each other. Same assumption can be made to each user. Also each user's subsequent transmitted symbols are assumed to be independent. Hence every product of $h_\ell b_i$ can be regarded as an independent source signal. Based on the discussion from the previous sections, the expression of (5.65) is a linear ICA/BSS model with $\mathbf{F}$ the matrix of source signals, $\mathbf{Y}$ the observed data matrix, and $\mathbf{G}$ the unknown mixing matrix. Hence classical ICA/BSS method or the Time-frequency method investigated in the previous can be applied to estimate the fading channels or to estimate the desired user's symbol.

5.6 Conclusions

In this chapter we studied two important tools, independent component analysis (ICA) or blind source separation (BSS) and time-frequency representation with the aim to improve the performance of wireless communication systems. The blind source separation using time-frequency representation is investigated. Compare to classical methods based on second order statistics, time-frequency analysis based method exploits the non stationarity of the signal. Different time-frequency distributions are compared to separate speech signals. We also show that ICA/BSS and communication schemes such as MIMO and CDMA are intimately related. Mathematically, a MIMO or a CDMA scheme can be formulated into an ICA/BSS problem, hence classical ICA/BSS methods or time-frequency representation based methods can be applied to improve the performance of the communication system.

It should be also noted that time-frequency representation alone can be used to combat fading channels in OFDM or CDMA systems [75]. The study in this chapter opens up a gate to apply ICA/BSS and TFR to wireless communication systems.

Chapter 6

Conclusions and Future Directions

Growing interest in high data rate wireless communications attracted much research interest in new techniques such as OFDM, MIMO, Blind Source Separation (BSS) and Time-Frequency Representation (TFR) in the past few years. The central topic is to combat the hostile wireless channel in an effort to achieve higher spectral efficiency and more reliable communication. Many fundamental problems have to be solved. This thesis is a part of our effort in that direction. The knowledge we acquired and the concepts we developed in this thesis open up a family of new problems that help us design effective communication schemes.

The knowledge acquired and the concepts developed in this thesis can be summarized as:

1. An information-theoretic understanding of the wireless communication systems

2. A complete analysis and simulation of mobile fading channels

3. Several fading mitigation techniques are studied, such as channel estimation, CDMA, OFDM, MIMO, Time-frequency representation.

Future Directions

1. In Section 2.2, a linear FM (LFM) signal is applied as a test signal at the input of digital Σ and $\Sigma\Delta$ modulators to reveal the relationship of the spectra at different stages. The noise shaping ability and the validity of the linear model of the $\Sigma\Delta$ modulator are also investigated. Similar principle may be applied to other configurations of the $\Sigma\Delta$ modulator in future works.

2. In Section 3.7, the Rayleigh fading simulator may be extended to multiple cross-correlated Rayleigh fading channels according to the method described in [95].

3. In Section 4.3.1, we simulated the selection combiner and the maximal-ratio combiner over slow flat Rayleigh fading channel and obtained their corresponding probability density functions. The simulation may be extended to investigate the performance of different modulation schemes with diversity, see [64].

4. In Section 3.9, a Gauss-Markov approximation of the fading channel is discussed. As an application of the Gauss-Markov channel model, we may use it for joint channel estimation and Low-density-parity-check (LDPC) decoding in flat Rayleigh fading shown in [53] and verify the statement in the literature that a first-order Markov process is sufficient to represent channel fading.

5. In Section 4.2, we discussed the principle of OFDM, and in Section 4.3.3 we investigated the capacity of MIMO channels by SVD decomposition. We have seen that the combination of MIMO and OFDM boosts capacity and reduces receiver complexity. To further improve the performance of a MIMO-OFDM system, we might adopt LDPC code to provide the channel coding gain as shown in [31].

6. In Chapter 5, we studied the BSS method based on entropy maximization principle. This method is used to estimate the Rayleigh fading channel in CDMA downlink transmission and shows performance improvement over MMSE based channel estimation scheme in [39]. Other methods such as time frequency based BSS might be used to further improve the performance of estimation.

7. In Chapter 5, we studied time-frequency representation in the context of Blind Source Separation (BSS). The powerful tool of time-frequency representation in fact can be used to characterize the random LTV mobile fading channel investigated in Section 3.8. The insightful results offered by the time-frequency perspective are very useful in CDMA or OFDM transceiver design [46] [75].

Appendix A

A.1 Gaussian and Rayleigh Distributions

Assume two independent and identically distributed (i.i.d) Gaussian random variables $X_1 \sim N(0, \sigma^2)$ and $X_2 \sim N(0, \sigma^2)$. Let $R = \sqrt{X_1^2 + X_2^2}$ and $\Theta = \text{actan}(\frac{X_2}{X_1})$, i.e., (X_1, X_2) denotes the coordinates of a point in the plane, then R is Rayleigh distributed

$$f_R(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, & r \geq 0 \\ 0, & r < 0 \end{cases} \tag{A.1}$$

and Θ is uniformly distributed

$$f_\Theta(\theta) = \frac{1}{2\pi}, \qquad 0 \leq \theta \leq 2\pi \tag{A.2}$$

A simulation is illustrated in Fig. A.1 for a demonstration. The detailed derivation can be found in [63] or [64].

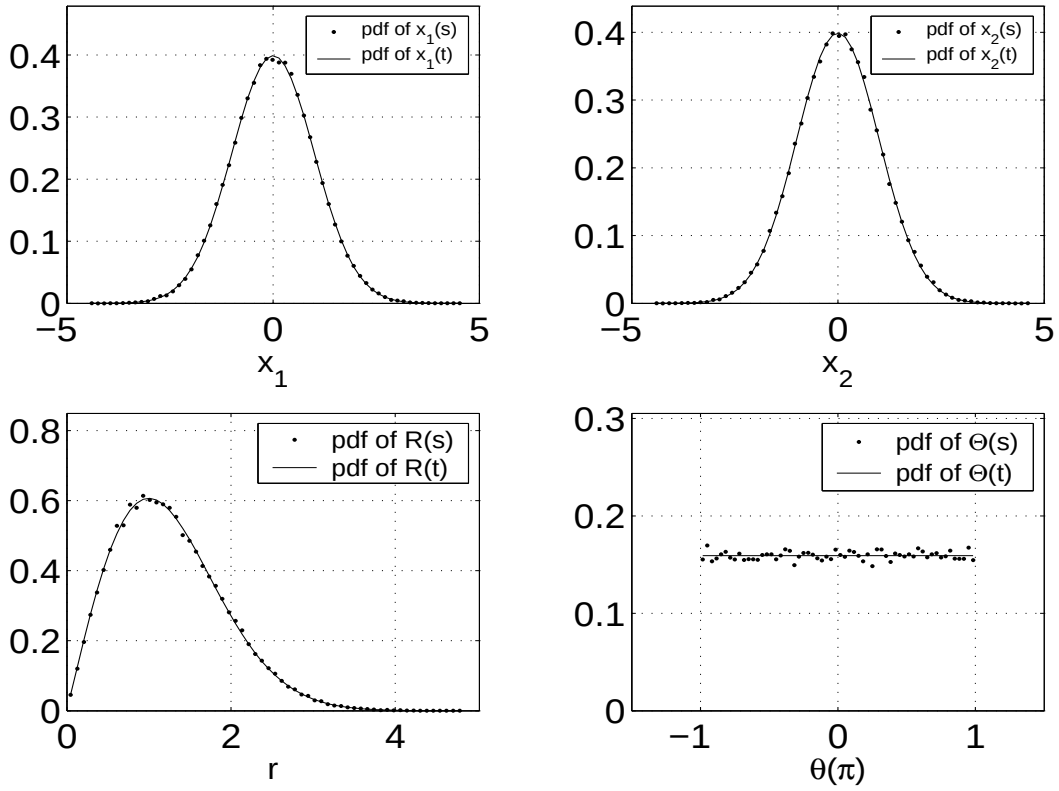


Figure A.1: The histogram and pdf of X_1, X_2, R and Θ. With enough samples, the simulation results match the theory well.

A.2 Central Limit Theorem (CLT) and Approximation of Integrals by Riemann Sums

In signal processing and digital communications, we constantly use Central Limit Theorem which states that the probability distribution function of the sum of the large number of random variables approaches a Gaussian distribution.

We consider the sum of Y of N independent random variables $X_1, X_2, \ldots, X_N$. The density function of $Y = X_1 + X_2 + \cdots + X_N$ is the $(N-1)$-fold convolution of the N individual density functions [60]:

$$f_Y(y) = f_{X_N}(x_N) * f_{X_{N-1}}(x_{N-1}) * \cdots * f_{X_1}(x_1) \tag{A.3}$$

Integrals are involved to find out $f_Y(y)$. Computer simulation does not have the ability to evaluate integrals but the values of the integral can be approximated to any level of accuracy desired. The simplest method of computing an approximation of the integral is to use Riemann sums. If $f(x)$ is continuous in $[a, b]$, and let $a = x_0 < x_1 < \cdots < x_n = b$, $\Delta x_k = x_k - x_{k-1}$, $x_{k-1} \leq \xi_k \leq x_k$, $d = \max \Delta x_k$, then [67]

$$\int_a^b f(x)dx = \lim_{d\to 0} \sum_{k=1}^{n} f(\xi_k)\Delta x_k = \lim_{n\to\infty} \frac{b-a}{n} \sum_{k=1}^{n} f\left(a + \frac{k}{n}(b-a)\right) \tag{A.4}$$

We give an example to verify CLT by using approximation of integrals by Riemann sums in computer simulation. Assuming the N random variables $X_i, i = 1, 2, \ldots, N$, are all independent and distributed uniformly over $[-1/2, 1/2]$. Calculate the $(N-1)$-fold convolution using Riemann sums, the resulting distribution $f_Y(y)$ is approaching Gaussian $\mathcal{N}(0, N\frac{1}{12})$. This is illustrated in Fig. A.2, where $N = 10$. All the convolution results are shown in the figure.

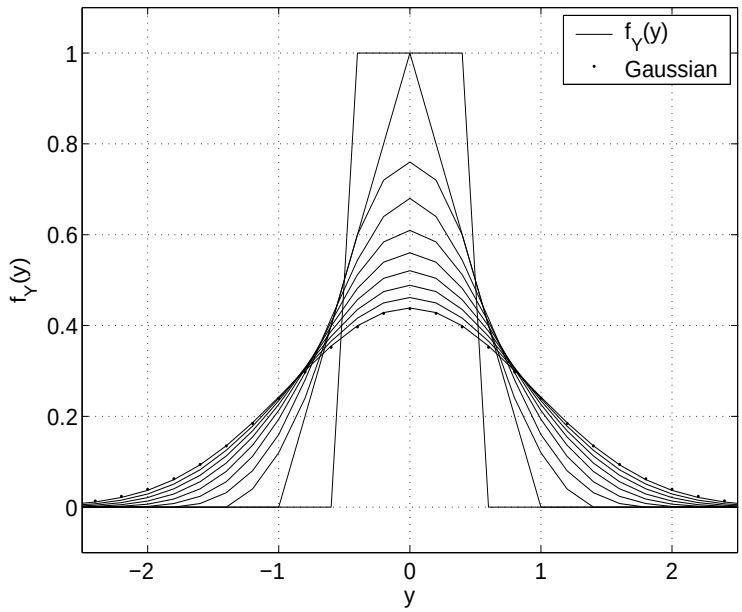


Figure A.2: Sum of several independent uniformly distributed random variables. Here $N = 10$, a 9-fold convolution of the 10 individual density functions gives good approach of Gaussian.

A.3 The Autocorrelation Function of Gauss-Markov Time-Varying Channel Model

In this section we derive the autocorrelation function of the channel process of (3.61).

To avoid notation confusion, we adopt the notation in [54] and rewrite (3.61) as

$$y[m+1] = \sqrt{1-\delta}y[m] + \sqrt{\delta}x[m+1], \quad m \geq 0, \tag{A.5}$$

First we notice both the input and the output are zero mean stationary processes as $\{x[m]\}$ is a sequence of i.i.d $\mathcal{CN}(0,1)$ r.v.. This result is helpful because the two stationary random processes $\{x[m]\}$ and $\{y[m]\}$ will have the following relations for autocorrelation $\phi_{xx}[m]$, $\phi_{yy}[m]$ and autocovariance $\gamma_{xx}[m]$, $\gamma_{yy}[m]$ [54]

$$\phi_{xx}[m] = \gamma_{xx}[m], \tag{A.6}$$

$$\phi_{yy}[m] = \gamma_{yy}[m]. \tag{A.7}$$

This provides convenience to find the autocorrelation function of $y[m]$ because in general the z-transform can be used to represent the covariance function but not a correlation function. When a signal has nonzero average value, its correlation function will contain an additive constant component that does not have a z-transform representation. Let $\Gamma_{xx}(z)$ and $\Gamma_{yy}(z)$ be the z-transform of $\gamma_{xx}[m]$ and $\gamma_{yy}[m]$, respectively. We have this relation

$$\Gamma_{yy}(z) = H(z)H^*(1/z^*)\Gamma_{xx}(z), \tag{A.8}$$

where $H(z)$ is the transfer function to describe the linear time-invariant discrete-time system in (A.5).

The transfer function $H(z)$ of the system with difference equation (A.5) is

$$H(z) = \frac{\sqrt{\delta}}{1-\sqrt{1-\delta}z^{-1}} \tag{A.9}$$

Since $\{x[m]\}$ is a sequence of i.i.d $\mathcal{CN}(0,1)$ r.v., we have $\Gamma_{xx}(z) = 1$. And $\Gamma_{yy}(z)$ can then be expressed as

$$\Gamma_{yy}(z) = \frac{\sqrt{\delta}}{(1-\sqrt{1-\delta}z^{-1})(1-\sqrt{1-\delta}z)}. \tag{A.10}$$

Since $\Phi_{yy}[m]$ is an autocorrelation function, it is two-sided sequence (symmetric), we can determine the ROC here is $\sqrt{1-\delta} < |z| < 1/\sqrt{1-\delta}$. The autocorrelation $\phi_{yy}[m]$ can be obtained as

$$\begin{aligned}\phi_{yy}[m] &= \gamma_{yy}[m] \\ &= (\sqrt{1-\delta})^m u[m] - [-(1/\sqrt{1-\delta})^m]u[-m-1] \\ &= (\sqrt{1-\delta})^{|m|}\end{aligned} \tag{A.11}$$

Also note that the unit circle is in the ROC of $\sqrt{1-\delta} < |z| < 1/\sqrt{1-\delta}$, thus Fourier transform exists for $\phi_y[m]$, which is the power spectrum density of the sequence $\{y[m]\}$.

A.4 Kullback Leibler Distance and Entropy Maximizer

We have seen the importance of *circular symmetric complex Gaussian* random variable in the analysis of OFDM scheme and MIMO channel capacity. To demonstrate the understanding of circular symmetric complex Gaussian random variable, we give the detailed proof of the following lemma: circular symmetric complex Gaussian random vectors are entropy maximizers. The proof involves a serial steps of definitions and theorems.

As the derivation involves *expected value rule* and matrix transformation of random vectors, it is convenient to use the following notation in order to avoid notation ambiguity:

- Deterministic scalar: x.
- Deterministic vector: $\boldsymbol{x}$.
- Random variable: X.
- Random vector: $\boldsymbol{X}$.
- Deterministic matrix: H.
- Random Matrix: $\mathbf{H}$.

And we use superscript ^{t} to stand for transpose of a vector or a matrix, superscript * for complex conjugate-transpose of a vector or a matrix.

Remark. In this thesis and in the literature such as [92] and [39], a moderate level of notation ambiguity is allowed for convenience. For example, notation such as $\mathbf{x}$ is used for both random vector and deterministic vector.

Definition A.4.1. The *differential entropy* $h(X)$ of a continuous random variable X with a density $f(x)$ is defined as

$$h(X) = -\int_S f(x) \log f(x) dx, \tag{A.12}$$

where S is the support set of the random variable.

The definition of differential entropy of a signal random variable can be extended to a random vector.

Definition A.4.2. The *differential entropy* of a set $X_1, X_2, \ldots, X_n$ of random variables with density $f(x_1, x_2, \ldots, x_n)$ is defined as

$$h(X_1, X_2, \ldots, X_n) = -\int f(x_1, x_2, \ldots, x_n) \log f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n. \tag{A.13}$$

or,

$$h(f) = -\mathbb{E}_f[\log f(\boldsymbol{X})] \tag{A.14}$$

Definition A.4.3. The Kullback Leibler distance $D(f \parallel g)$ between two densities f and g is defined by

$$D(f \parallel g) = \int f \log \frac{f}{g}. \tag{A.15}$$

Theorem A.4.1.

$$D(f \parallel g) \geq 0. \tag{A.16}$$

with equality iff $f = g$ almost everywhere (a.e.).

Proof. see [19]. □

Theorem A.4.2. *Let $\boldsymbol{X} \in \mathbb{C}^n$ be a circular symmetric Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance* Q*. The the probability density function $\boldsymbol{X}$ is given by*

$$f(\boldsymbol{x}) = \det(\pi \mathrm{Q})^{-1} \exp(-(\boldsymbol{x} - \boldsymbol{\mu})^* \mathrm{Q}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})). \tag{A.17}$$

Proof. see [90]. □

Lemma A.4.1. *Suppose the complex random vector $\boldsymbol{X} \in \mathbb{C}^n$ is zero-mean with covariance matrix $\mathbb{E}[\boldsymbol{X}\boldsymbol{X}^\dagger] = \mathrm{Q}$. Then the entropy of $\boldsymbol{X}$ satisfies $h(\boldsymbol{X}) \leq \log\det(\pi e \mathrm{Q})$ with equality if and only if $\boldsymbol{X}$ is a circular symmetric Gaussian with $\mathbb{E}[\boldsymbol{X}\boldsymbol{X}^\dagger] = \mathrm{Q}$.*

Proof. Let

$$f(\boldsymbol{x}) = \det(\pi \mathrm{Q})^{-1} \exp(-\boldsymbol{x}^* \mathrm{Q}^{-1} \boldsymbol{x}), \tag{A.18}$$

and let p be *any* density function satisfying $\mathbb{E}_p[\boldsymbol{X}\boldsymbol{X}^*] = \mathbb{E}_f[\boldsymbol{X}\boldsymbol{X}^*] = \mathrm{Q}$.

We have the following relation:

$$\begin{aligned} -\mathbb{E}_f[\log f(\boldsymbol{X})] &= \log\det(\pi \mathrm{Q}) + (\log e)\mathbb{E}_f[\boldsymbol{X}^* \mathrm{Q}^{-1} \boldsymbol{X}] \\ &= \log\det(\pi \mathrm{Q}) + (\log e)\mathrm{tr}(\mathbb{E}_f[\boldsymbol{X}\boldsymbol{X}^*]\mathrm{Q}^{-1}) \\ &= \log\det(\pi \mathrm{Q}) + (\log e)\mathrm{tr}\ \mathbf{I} \end{aligned}$$

and,

$$\begin{aligned} -\mathbb{E}_p[\log f(\boldsymbol{X})] &= \log\det(\pi \mathrm{Q}) + (\log e)\mathbb{E}_p[\boldsymbol{X}^* \mathrm{Q}^{-1} \boldsymbol{X}] \\ &= \log\det(\pi \mathrm{Q}) + (\log e)\mathrm{tr}(\mathbb{E}_p[\boldsymbol{X}\boldsymbol{X}^*]\mathrm{Q}^{-1}) \\ &= \log\det(\pi \mathrm{Q}) + (\log e)\mathrm{tr}\ \mathbf{I} \end{aligned}$$

hence, we have this equality:

$$\mathbb{E}_f[\log f(\boldsymbol{X})] = \mathbb{E}_p[\log f(\boldsymbol{X})]$$

Then, from the definition of differential entropy, we have

$$\begin{aligned} h(p) - h(f) &= -\mathbb{E}_p[\log p(\boldsymbol{X})] + \mathbb{E}_f[\log f(\boldsymbol{X})] \\ &= -\mathbb{E}_p[\log p(\boldsymbol{X})] + \mathbb{E}_p[\log f(\boldsymbol{X})] \\ &= -\mathbb{E}_p\left[\log \frac{p(\boldsymbol{X})}{f(\boldsymbol{X})}\right] \\ &= -D(p \parallel f) \\ &\leq 0, \end{aligned}$$

with equality only if $p = f$. Thus $h(p) \leq h(f)$. From the above derivation, it is easy to obtain the upper bound of $h(\boldsymbol{X}) = \log\det(\pi e \mathrm{Q})$ when $\boldsymbol{X}$ is circular symmetric complex Gaussian. □

A.5 The Mean Vector

In this section, we demonstrate how to derive the mean value of a component in a random vector. Let $\boldsymbol{X}$ be a n-dimensional *real random vector*,

$$\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^t \tag{A.19}$$

and $p_{\boldsymbol{X}}(\boldsymbol{x})$ is the multivariate pdf (joint pdf) of $\boldsymbol{X}$.

The mean vector $\boldsymbol{m} = (m_1, m_2, \ldots, m_n)^t$ of $\boldsymbol{X}$ is defined as the expectation of $\boldsymbol{X}$:

$$\begin{aligned} \boldsymbol{m} &= \mathbb{E}[\boldsymbol{X}] \\ &= \int_{-\infty}^{\infty} \boldsymbol{x} p_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \boldsymbol{x} p_{\boldsymbol{X}}(\boldsymbol{x}) dx_n \ldots dx_2 dx_1 \end{aligned} \tag{A.20}$$

Note that the definition of the vector integral is the set of scalar integrals over all of the components of the random vector.

Denote the marginal density of the ith component X_i of $\boldsymbol{X}$ by $p_{X_i}(x_i)$, which can be expressed as:

$$p_{X_i}(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\boldsymbol{X}}(\boldsymbol{x}) dx_n \ldots dx_2 dx_1, \tag{A.21}$$

that is, $p_{X_i}(x_i)$ is obtained by integrating all the other random variables in $\boldsymbol{x}$ ($x_1, x_2, \ldots, x_n$, except x_i).

By combining (A.20) and (A.21), each component m_i of the n-vector $\boldsymbol{m}$ can be obtained as:

$$\begin{aligned} m_i &= \mathbb{E}[X_i] \\ &= \int_{-\infty}^{\infty} x_i p_{X_i}(x_i) dx_i \\ &= \int_{-\infty}^{\infty} x_i p_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \end{aligned} \tag{A.22}$$

Eq. (A.22) can be demonstrated by using m_1 (without losing generality) as:

$$\begin{aligned} m_1 &= \mathbb{E}[X_1] \\ &= \int_{-\infty}^{\infty} x_1 p_{X_1}(x_1) dx_1 \\ &= \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\boldsymbol{X}}(\boldsymbol{x}) dx_n \ldots dx_3 dx_2 \right) dx_1 \quad \text{(apply (A.21))} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 p_{\boldsymbol{X}}(\boldsymbol{x}) dx_n \ldots dx_2 dx_1 \\ &= \int_{-\infty}^{\infty} x_1 p_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}. \end{aligned}$$

Bibliography

[1] K. Abde-Meraim, A. Belouchrani and A. R. Leyman, "Blind source separation using time-frequency distributions," Chapter in *Time Frequency Signal Analysis and Processing, a Comprehensive Reference*, B.Boashash, Ed., Elsevier, pp. 323-333, 2003.

[2] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communication," *IEEE Journal on Selected Areas in Communication*, 16(8), pp.1451-1458, Oct. 1998.

[3] J. B. Andersen, T. S. Rappaport, and S. Yoshida, "Propagation Measurements and Models for Wireless Communications Channels", *IEEE Communications Magazine (special issue about wireless personal communication),* 33(1), pp.42-49, Jan. 1995.

[4] P. M. Aziz, H. V. Sorensen and J. V. Der Spiegel, "An Overview of Sigma-Delta Converters," *IEEE Signal Processing Magzine,* Jan, 1996.

[5] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, 3rd ed, 2004.

[6] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, 11(4), pp. 360-393, Dec. 1963.

[7] A. Belouchrani, K. Abed-Meraim, Jean-François Cardoso and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, 45(2), pp. 434-444, Feb. 1997.

[8] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representations," *IEEE Trans. Signal Processing*, 46(11), pp. 2888-2897, Nov. 1998.

[9] E. Biglieri, J. Proakis, et al., "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory,* 44(6), pp. 2619-2692, Oct. 1998.

[10] B. Boashash (Editor), *Time Frequency Signal Analysis and Processing, A Comprehensive Reference*, Elsevier, 2003.

[11] H. Bölcskei, "Principles of MIMO-OFDM wireless systems," Chapter in *Signal Processing for Mobile Communications Handbook*, M.Ibnkahla, Ed., CRC Press, pp. 12.1 - 12.22, 2004.

[12] H. Bölcskei and A. J. Paulraj, "Multiple-input multiple-output (MIMO) wireless systems," Chapter in *"The Communications Handbook"*, 2nd edition, J. Gibson, Ed., CRC Press, pp. 90.1-90.14, 2002.

[13] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.,* 50(2), pp. 225-234, Feb. 2002.

[14] Jean-Francois Cardoso and A. Souloumiac, "Blind Beamforming For Non Gaussian Signals," *IEE-Proceedings-F,* 140(6), pp. 362-370, Dec. 1993.

[15] S. Chandran (Editor), *Adaptive Antenna Arrays*, Springer, 2004.

[16] Chong-Yung Chi, Ching-Yung Chen, Chii-Horng Chen, and Chih-Chun Feng, "Batch Processing Algorithms for Blind Equalization Using Higher-Order Statistics", *IEEE Signal Processing Magazine*, Jan, 2003.

[17] R. H. Clarke, "A Statistical Theory of Mobile-Radio Reception," *Bell Systems Technical Journal*, vol. 47, pp. 957-1000, 1968.

[18] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[19] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.

[20] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. Dissertation, MIT, 1989.

[21] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport and J. H. Reed, "Overview of Spatial Channel Models for Antenna Array Communication Systems", *IEEE Personal Communications*, pp. 10-22, Feb. 1993.

[22] S. A. Fechtel, "A Novel Approach to Modelling and Efficient Simulation of Frequency-Selective Fading Radio Channle", *IEEE Joural on Selected Areas in Communications*, 11(3), pp. 422-431, Apr. 1993.

[23] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-elemenent antennas," *Bell Systems Technical Journal*, 1(2), pp. 41-59, 1996.

[24] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications,* 6(3), pp. 311-335, Mar. 1998.

[25] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.

[26] ——, *Low-Density Parity-Check Codes*, Cambridge, MA:M.I.T. Press, 1963.

[27] I. Galton, "Delta-Sigma Data Conversion in Wireless Transceivers," *IEEE Trans. Microwave Theory and Techniques*, 50(1), pp. 302-315, Jan. 2002.

[28] D. Gesbert *et al*, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journal on Selected Areas in Comm.,* 21(3), pp. 281-302, Apr. 2003.

[29] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Fundamental Capacity of MIMO Channels," *IEEE Journal on Selected Areas in Communications, Special Issue on MIMO Systems*, 2003.

[30] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*, available online at http://www-ee.stanford.edu/ gray/toeplitz.html.

[31] J. Ha, A. N. Mody, J. Sung, J. R. Barry, S. W. McLaughlin and G. L. Stuber, "LDPC Coded OFDM with Alamouti/SVD Diversity Technique", *Wireless Personal Communications*, 23(1), pp. 183-194, Oct. 2002.

[32] S. Haykin and B. V. Veen, *Signals and Systems*, John Wiley & Sons, 2nd ed, 2003

[33] S. Haykin, *Communication Systems*, John Wiley & Sons, 4th ed, 2001.

[34] S. Haykin and M. Moher, *Modern Wireless Communications*, Prentice Hall, 2005.

[35] J. Heiskala and J. Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*, Sams Publishing, 2002.

[36] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, reprinted with corrections, (first published in 1985), 1990 .

[37] Z. M. Hussain and B. Boashash, "Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals," *The Sixth International Symposium on Signal Processing and Its Applications (ISSPA'2001)*, pp. 13-16 Aug. 2001.

[38] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using time-frequency distributions," *IEEE Transactions on Signal Processing,* 50(8), pp. 1866-1876, 2002.

[39] A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Willey & Sons, 2001.

[40] M. Ibnkahla (Editor), *Signal Processing for Mobile Communications Handbook*, John Wiley & Sons, 4th ed, 2001.

[41] W. C. Jakes, Jr, *Microwave Mobile Communications*, Wiley, reprinted, (first published in 1974), 1994.

[42] M. Jankiraman, *Space-time Codes and MIMO Systems*, Artech House, 2004.

[43] M. Kozak and I. Kale, *Oversampled Delta-Sigma Modulators, Analysis, Applications and Novel topologies*, Kluwer Academic Publishers, 2003.

[44] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, 1969.

[45] D. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge, 2003.

[46] G. Matz and F. Hlawatsch, "Time-frequency characterization of random time-varying channels," Chapter in *Time Frequency Signal Analysis and Processing, a Comprehensive Reference*, B.Boashash, Ed., Elsevier, pp. 410-419, 2003.

[47] R. J. McEliece, *The theory of information and coding*, Cambridge, Student edition, 2004.

[48] M. H. Hayes, *Statistical Digital Signal Processing And Modeling*, John Wille& Sons, 1996.

[49] T. K. Moon, *Error correction coding : Mathematical methods and algorithms*, John Willey & Sons, 2005.

[50] T. K. Moon and W. C. Stirling *Mathematical Metods and Algorithms for Signal Processing*, Prentice Hall, 2000.

[51] B. Muquet, Marc de Courville and P. Duhamel, "Subspace-Based Blind and Semi-Blind Channel Estimation for OFDM Systems," *IEEE Transactions on Signal Processing,* 50(7), pp. 1699-1712, 2002.

[52] C. L. Nikias and J. M. Mendal, *Signal Processing with Higher-Order Spectra*, IEEE Signal Processing Magazine, 1993.

[53] H. Niu, M. Shen, J. A. Ritcey and H. Liu, "Iterative Channel Estimation and LDPC Decoding over Flat-Fading Channel: A Factor Graph Approach," *The 2003 Conference on Information Sciences and Systems*, The Johns Hopkins University, March 12-14, 2003.

[54] A. V. Oppenheim, R. W. Schafer and J. R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, 2nd ed, 1999

[55] L. Shang and Z. M. Hussain, "A Spectrum Analysis of Single-Bit Sigma-Delta Modulator Using Linear Frequency Modulation," *IEEE TENCON*, Melbourne, Australia, Nov. 21-24, 2005.

[56] M. Pätzold, *Mobile Fading Channels: Modelling, Analysis and Simulation*, John Wiley & Sons, 1st ed, 2002.

[57] M. Pätzold and F.Laue, "Level-Crossing Rate and Average Duration of Fades of Deterministic Simulation Models for Rice Fading Channels," *IEEE Trans. Vehicular Technology*, 48(4), pp. 1121-1129, July, 1999.

[58] M. Pätzold, "Modelling and Simulation of Spatio-Temporal Wideband Mobile Radio Channels," in *Signal Adaptive Antenna Arrays, Trends and Applications*, S. Chandran, Ed., Springer, pp. 293-313, 2004.

[59] A. J. Paulraj, D. A. Gore, et al., "An Overview of MIMO Communications-A Key to Gigabit Wireless," *Proceedings of the IEEE*, 92(2), pp. 198-218, Feb. 2004.

[60] P. Z. Peebles, JR., *Probability, Random Variables And Random Signal Principles*, McGraw-Hill, 2001.

[61] H. V. Poor and G. W. Wornell, *Wireless Communications, Signal Processing Perspectives*, Prentice Hall, 1998.

[62] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 2nd ed, 1998.

[63] J. G. Proakis and M. Salehi, *Communication System Engineering*, Prentice Hall, 1st ed, 1994.

[64] J. G. Proakis, *Digital Communications*, Prentice Hall, 4th ed, 2001.

[65] J. G. Proakis and M. Salehi, *Communication System Engineering*, Prentice Hall, 2nd ed, 2002.

[66] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing, Principles, Algorithms, and Applications*, Prentice Hall, 3rd ed, 1996.

[67] L. Rade and B. Westergren, *Mathematics Handbook for Science and Engineering*, Springer, 4th ed, 1998.

[68] Q. M. Rahman and M. Ibnkahla, "Signal Processing for Future Mobile Communications Systems: Challenges and Perspectives," Chapter in *Signal Processing for Mobile Communications Handbook*, M. Ibnkahla, Ed., CRC Press, pp. 1.2-1.38, 2004.

[69] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, 46(3), pp. 357-366, 1998.

[70] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, 2nd ed, 2002.

[71] G. W. Recktenwald, *Numerical Methods with Matlab, implementation and application*, Prentice Hall, 2000.

[72] T. J. Richardson and R. Urbanke, *Modern Coding Theory*, to be published,

[73] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, 2nd ed, 2004.

[74] S. Sanghavi and B. Hajek, "Adaptive Induced Fluctuations for Multiuser Diversity," *IEEE ISIT 2002*, pp.450, Jul. 2002.

[75] A. M. Sayeed, "Communication over multipath fading channels: a time-frequency perspective," Chapter in *Time Frequency Signal Analysis and Processing, a Comprehensive Reference*, B.Boashash, Ed., Elsevier, pp. 549-557, 2003.

[76] ——, "Deconstructing multi-antenna fading channels," *IEEE Transactions on Signal Processing,* 50(10), pp. 2563-2579, 2002.

[77] L. L. Scharf, *Statistical Signal Processing: detection, estimation, and time series analysis*, Addison-Wesley, 1991.

[78] S. S. Mahmoud, Z. M. Hussain, and P. O'Shea, "Space-time model for mobile radio channel with hyperbolically distributed scatterers," *IEEE Antennas and Wireless Propagation Letters*, 1(12), pp. 211-214, 2002.

[79] ——, "Geometrical model for mobile radio channel with hyperbolically distributed scatterers," *The 8th IEEE International Conference on Communications Systems*, vol. 1, pp. 17-20, Singapore, Nov. 2002.

[80] ——, "A Geometrical-Based Channel Model with Hyperbolically Distributed Scatterers for a Macrocell Mobile Environment with Antenna Array," *Multimedia Cyberscape Journal*, vol. 2, pp.1-10, 2004.

[81] ——, "Spatial and Temporal Statistics for the Geometrical-Based Hyperbolic Macrocell Channel Model," *Submitted to IEEE Transactions on Vehicular Technology*, 2005.

[82] C. E. Shannon, "A mathematical theory of communication", *Bell Sys. Tech. Journal*, 27: 379-423 and 623-656, 1948.

[83] Z. Shi, "Sigma-Delta ADC and DAC for Digital Wireless Communication," *IEEE Radio Frequency Integrated Circuits Symposium*, 1999.

[84] B. Sklar, A. Hero, et al., "Highlights of Statistical Signal and Array Processing", *IEEE Signal Processing Magazine*, June, 1998.

[85] B. Sklar, *"How I Learned to Love the Trellis"*, *IEEE Signal Processing Magazine*, May, 2003.

[86] J. I. Smith, "A Computer Generated Multipath Fading Simulation for Mobile Radio," *IEEE Trans. Vehicular Technology*, 24(3), pp. 39-40, Aug. 1975.

[87] R. W. Stewart and E. Pfann, "Oversampling and Sigma-Delta Strategies for Data Conversion," *IEE Electronics & Communication Engineering Journal*, pp. 37-47, Feb. 1998.

[88] G. Strang, *Linear Algebra and Its Applications*, Academic Press, New York, 3rd ed. 1988.

[89] G. L. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publishers, 2001.

[90] E. Telatar, "Capacity of the multiple antenna Gaussian channel," *European Transactions on Telecommunications*, 10(6), pp. 585-595, Nov. 1999.

[91] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, 1992.

[92] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[93] D. Tse, P. Viswanath and L. Zheng, "Diversity-multiplexing tradeoff in multiple access channels," *IEEE Transactions on Information Theory*, 50(9), pp. 1959-1974, 2004.

[94] A. M. Tulino and S. Verdu, *Random matrices and wireless communication*, Fundations and Trends in Communications and Information Theory, 1(1), 2004.

[95] C. X. Wang and M. Pätzold, "Efficient simulation of multiple cross-correlated Rayleigh fading channels" *Proceedings of the* 14^{th} *PIMRC*, 2003.

[96] S. G. Wilson, *Digital Modulation and Coding*, Prentice Hall, 1996.

[97] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, John Wiley & Sons, 1965.

[98] H. Yang, "A road to future broadband wireless access: MIMO-OFDM-Based air interface," *IEEE Communications Magazine*, 43(1), pp. 53-60, Jan. 2005.

[99] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple antenna channels," *IEEE Transactions on Information Theory*, 49(5), pp.1073-1096, May 2003.

[100] L. Zheng, *Diversity-Multiplexing Tradeoff: A Comprehensive View of Multiple Antenna Systems*, Ph.D. Dissertation, University of California at Berkeley, 2002,

[101] R. E. Ziemer and W. H. Tranter, *Pinciples of Communications, Systems, Modulation and Noise*, John Willey & Sons, 5th ed, 2002.