1

Introduction

1.1 General remarks

The explosion in demand for wireless services experienced over the past 20 years has put significant pressure on system designers to increase the capacity of the systems being deployed. While the spectral resource is very scarce and practically exhausted, the biggest possibilities are predicted to be in the areas of spectral reuse by unlicensed users or in exploiting the spatial dimension of the wireless channels. The former approach is now under intense development and is known as the cognitive radio approach (Haykin 2005). The latter approach is as old as communication systems themselves and is known mostly through the receive diversity techniques, well studied in both Western (Middleton 1960; Simon and Alouini 2000) and former USSR literature (Fink 1970; Klovski 1982a). These techniques are mainly used to improve the signal to noise ratio in the receiver in fading environments. In order to exploit the additional (spatial) dimension of the wireless channel, a number of technologies were suggested in early 1990s, including smart antennas. The development of this antenna technology mainly focused on the development of the estimation of the angle of arrival, optimal beamforming, and space-time signal processing. However, these techniques offered only a limited increase in the channel capacity.

In recent years, however, development of the multiple-input multiple-output (MIMO) system has emerged as the most potent technique for increasing the capacity of wireless channels. This technique exploits sampling in the spatial dimension on both sides of the communication links, combined in such a way that they either create virtual, multiple, parallel spatial data pipes to increase capacity linearly with the number of pipes and/or to add diversity to improve the quality of the links. A large number of research articles and monographs have treated different aspects of this subject (Correia, Ed. 2007; Paulraj et al. 2003; Tse and Viswanath 2005;

Wireless Multi-Antenna Channels: Modeling and Simulation, First Edition.
Serguei L. Primak and Valeri Kontorovich.

Verdu 1998), including channel modeling, modulation, diversity-multiplexing trade-off, and so on. Since initial papers by Teletar (Teletar 1999), MIMO technologies have been included in many existing standards of 4G communications. Today MIMO technology appears to be the natural candidate for most large-scale commercial wireless products.

Most of the researchers are focusing on investigation of MIMO systems under the important assumption that fading is Rayleigh, channel state information is perfectly known, and the scattering is reach. Such results provide good limiting estimates for capacity, performance, and delay. However, it often provides over optimistic results. The main contribution of this book lies in addressing the following issues:

- Suggestion of the generalized Gaussian model of MIMO wireless channels.
- Investigation of performance of different coding schemes in generalized Gaussian channels.
- Suggestion of an efficient simulator of MIMO wireless channels based on a geometric-based modeling paradigm.
- In-depth studying of the effect of channel estimation on performance in MIMO systems.
- Introduction of the multitaper approach to channel estimation.
- Investigation of the second-order statistics of MIMO channel capacity.

The book is organized as follows. In this chapter we briefly discuss models for signals used in this text. Chapter 2 describes a novel, four-parametric model of a SISO wireless channel and extends the concept to MIMO configuration. We also consider channels with a fluctuating number of scatterers and other deviations from Gaussian models. Chapter 3 expands on the modeling of MIMO channels based on scattering geometry and explores different geometry characteristics that effect the channel model. It also describes narrowband MIMO channel models while Chapter 4 is dedicated to wideband models. Chapter 5 treats topics related to the investigation of the capacity of the MIMO channel under different geometrical conditions, treats time variation of the capacity, and capacity of sparse channels. Chapter 6 deals with the methodology of MIMO channel prediction, while Chapter 7 deals with effects of errors on different aspects of communication system performance. Finally, Chapter 8 deals with the investigation of space-time code performance in generalized Gaussian MIMO channels.

Finally we would like to express our gratitude to a number of people and agencies that were instrumental in supporting the research that has resulted in most of the content of this book. First of all, we would like to express our admiration for our late teacher and colleague Prod. Daniil (Dani) Klovski. His investigation of diversity in the time-space communication channel in the late 1960s to the early 1980s laid a solid foundation of smart antenna techniques in the former USSR while also inspiring us to offer a generalized Gaussian model of MIMO channels.

We are also deeply indebted to our graduate students, now independent researchers themselves. Our deepest thanks to Drs. Vanja Subotic, Khaled Almustafa, Dan Dechene, and A. F. Ramos-Alarcón for their diligent work in developing most of the ideas presented in this manuscript. We would like to thank Drs. Tricia Willink and Karim Baddour from the Communications Research Centre Canada (CRC Canada) for numerous discussions, especially on topics related to MIMO channel estimation. Our research and graduate students have been financially supported through a number of research grants, provided by NSERC Canada, CONACYT Mexico, and CRC Canada. We are grateful to the University of Western Ontario and CINVESTAV-IPN, Mexico for creating excellent working conditions and the opportunity to spend a few months both in Mexico and Canada jointly working on the manuscript. We also would like to thank our colleagues at the University of Agder, Prof. Matthias Patzold, Drs. Gulzaib Rafiq, Dmitri Umanski, and others for a number of useful discussions and suggestions. And last, but not least, we would like to thank wonderful staff at Wiley for their patience and indispensable help in preparing the manuscript.

1.2 Signals, interference, and types of parallel channels

Here we consider the problem of transmission of digital information over a set of parallel channels. The discrete message is chosen from an alphabet $\mathcal{A} = \{x_k\}$ of size $m = \dim(\mathcal{A})$. Once a symbol, say x from the alphabet $\mathcal{A}$ is chosen it is encoded to a signal waveform $z_k(t)$. A unique waveform corresponds to each symbol of the alphabet. In general one would select n symbols from the alphabet at a single moment of time to transmit them over n channels simultaneously. This can be accomplished by using mn different signals $z_{rk}(t)$, $r = 1, \cdots, m$, and $k = 1, \cdots, n$. We consider coherent systems where the duration of the symbols at every channel is fixed to be T and the start of the symbols in each channel coincide.

The received signals $z'_k(t)$ at the output of each of the parallel channels have a statistical relation to the transmitted signal $z_k(t)$, albeit one that does not coincide with it due to noise and interference. The received signals are processed by a decision-making block. We will focus on synchronous detection. This means that the decision-making algorithm observes a symbol over time period T and then decides which symbol has been transmitted.

Parallel channels are assumed to be linear and the signals $z_{rk}(t)$ narrowband. This means that most of the energy of the signals $z_{rk}(t)$ is concentrated in a frequency band F_{rk} that is much smaller than the carrier frequency f_{rk}. In this case the following representation is valid

$$z_{rk}(t) = E_{rk} \cos\left[2\pi f_{rk} t + \Phi_{rk}(t)\right], \; 0 \leq t < T \tag{1.1}$$

In this case the envelope

$$E_{rk}(t) = \sqrt{z_{rk}^2 + \tilde{z}_{rk}^2} \tag{1.2}$$

and the initial phase

$$\Phi_{rk}(t) = \text{atan}\frac{z_{rk}}{\tilde{z}_{rk}} - 2\pi f_{rk}t \tag{1.3}$$

are slowly varying factions with respect to $\cos(2\pi f_{rk}t)$. Here $\tilde{z}_{rk}$ is the Hilbert transform (Proakis 1997).

$$z_{rk}(t) = E_{rk}\sin\left[2\pi f_{rk}t + \Phi_{rk}(t)\right], 0 \le t < T \tag{1.4}$$

The low pass equivalent of this signal is thus

$$s_{rk} = E_{rk}\exp(j\Phi_{rk}) \tag{1.5}$$

The received signal can then be written as

$$z'_{rk} = \mu_k(t)z_{rk}\left[t - \tau_{rk}(t)\right] + \xi_k(t) \tag{1.6}$$

where $\mu_k(t)$ is a coefficient describing the attenuation of the signal transmitted through the k-th channel, τ_k is the delay associated with this channel, and $\xi_k(t)$ is the associated additive noise component. For the majority of realistic channels with variable parameters the time delay can be written as a sum

$$\tau_k(t) = \tau_{k\,\text{av}} - \Delta\tau_k(t) \tag{1.7}$$

where $\tau_{k\,\text{av}} = E\{\tau_k(t)\}$ is the average delay and $\Delta\tau_k(t)$ is the random fluctuation of the delay. The former can be associated with the overall propagation delay due to the finite distance between the transmit and the receive antennas, while the latter can be associated with variation in the channel and mutual position between the receiver and the transmitter. In order to have an effect on the performance of the system the fluctuating component should have a root-mean-square value comparable to the duration of the bit interval.

Let $\theta_k(t) = 2\pi f_{rk}\Delta\tau_k(t)$ be a random excessive phase associated with the k-th channel and r-th frequency. Then Equation (1.6) can be rewritten as

$$z'_{rk} = \mu_k(t)z_{rk}\left[t - \tau_{k\,\text{av}}, \theta_k(t)\right] + \xi_k(t) \tag{1.8}$$

where $z(t-\tau, \theta)$ represents signal $z(t)$ delayed by τ and the phase of its carrier shifted by θ.

During a one-bit detection the receiver processes segments of the signal $z'_k(t)$ of duration T one by one. It is quite clear that without loss of generality a constant delay $\tau_{k\,\text{av}}$ can be eliminated from consideration. Thus, the receiver observes a signal

$$z'_k(t) = \mu_k(t)z_{rk}[t, \theta_k(t)] + \xi_k(t),\ (l-1)T \le t < lT \tag{1.9}$$

where l indicates the order of the transmitted symbols.

For the narrowband process z'_k we can write its expansion in terms of in-phase and quadrature components (equation 2.44)

$$z'_k(t) = \mu_k(t)\cos\theta_k(t)z_{rk}(t) + \mu_k(t)\sin\theta_k(t)\tilde{z}_{rk}(t) + \xi_k(t),\ (l-1)T \le t < lT \tag{1.10}$$

Here $\mu_k(t)$ and $\theta_k(t)$ represent the magnitude and the phase of the transmission coefficient of the k-th channel.

In practice it is more efficient to represent fading either in terms of in-phase and quadrature components

$$\mu_{ck}(t) = \mu_k(t)\cos\theta_k(t) \tag{1.11}$$

$$\mu_{sk}(t) = \mu_k(t)\sin\theta_k(t) \tag{1.12}$$

or in its phasor (complex low-pass equivalent) form

$$\mu_k(t) = \mu_k(t)\exp[-j\theta_k(t)] \tag{1.13}$$

In this case

$$z'_k = \mu_{ck}(t)z_{rk}(t) + \mu_{sk}(t)\tilde{z}_{rk}(t) + \xi_k(t),\ (l-1)T \le t < lT \tag{1.14}$$

If $\mu_k(t)$ and $\theta_k(t)$ are random functions of time, they can be considered as multiplicative noise.

Without going into detail about the statistical properties of the random channel transmission coefficient it is important to provide a general classification of the digital communication systems utilizing such channels. The main purpose of a system with multiple channels is to increase the capacity of the channels. This can be achieved when the additive noise and the multiplicative noise are decorrelated.

Digital systems with multiple channels can be categorized as follows

- Channels with distinct media of propagation corresponding to different channels;
- Channels that share a media with the same physical properties.

The former category may include channels where the information is simultaneously transmitted over wired and wireless links. Such systems would mostly be used to provide redundancy and reliability of communication systems since the amount of information is the same regardless of number of channels used.

Systems belonging to the latter category are widely used in the diversity reception and recently in so-called MIMO systems. The improvement in the capacity of such systems is to a large degree defined by the number of channels used, the physical location of the antennas, and the signal processing techniques.

Depending on the method of forming multiple channels one can distinguish the following groups of parallel channels

1. systems with parallel channels formed on the transmitting side;
2. systems with parallel channels formed on the receiving side;
3. systems with parallel channels formed on both the receiving and transmitting side.

In the systems belonging to the first group the information about each bit/symbol is transmitted into the channel by means of multiple signals that are formed by the transmitter during the modulation stage. In this case the total number of different signals is mn where n is the number of parallel channels. Each bit/symbol can be transmitted simultaneously over all channels or sequentially over the same time interval T. Systems with frequency and time diversity are two well-known representatives of this class. In the case of frequency diversity the same signal is transmitted over different frequencies. The total signal in the receive antenna is thus

$$z'(t) = \sum_{k=1}^{n} \left[\mu_{ck}(t) z_{rk}(t) + \mu_{sk}(t) \tilde{z}_{rk}(t)\right] + \xi(t) \tag{1.15}$$

where $\mu_{ck}(t) + j * \mu_{sk}$ is a complex transmission coefficient for the signal $z_{rk}(t)$.

If the signals $z_{rk}(t)$ and $z_{rp}(t)$ are orthogonal for $k \neq p$ in the strong sense, then instead of a single received signal in Equation (1.15) one can consider a set of parallel channels with output signals defined by Equation (1.14) where $\mu_{ck}(t)$ and $\mu_{sk}(t)$ are in-phase and quadrature components of the transmission coefficient for the signal $z_{rk}(t)$ and $\xi_k(t)$ is the additive noise acting within the bandwidth allocated to the signal $z_{rk}(t)$. In general, quantities $\mu_{ck}(t)$, $\mu_{sk}(t)$, $\xi_k(t)$, and $\mu_{cp}(t)$, $\mu_{sp}(t)$, $\xi_p(t)$ are represented by correlated processes for all $k \neq p$. An optimal choice of signals $z_{rk}(t)$ will be such as to provide nearly complete decorrelation between the fading components and the additive noise of signals transmitted over different channels. For example, if all $z_{rk}(t)$ have non-overlapping spectra, than the additive noise $\xi_k(t)$ and $\xi_p(t)$ will be uncorrelated for $k \neq p$. Furthermore, if the propagation medium has a selectivity property,[1] one can choose carrier frequencies f_k in such a way that $\mu_{ck}(t)$, $\mu_{sk}(t)$ and $\mu_{cp}(t)$, $\mu_{sp}(t)$ are also almost completely decorrelated.

The simplest method of formation of two parallel channels is frequency shift keying (FSK). Indeed, FSK can be considered as two inverse amplitude modulations with different carriers.

[1] The majority of situations which involve the reflection of the transmitted signal from multiple obstacles will have this property.