

CHAPTER

1

Introduction

1.1 Book objective

Wireless communication is one of the most vibrant areas in the communication field today. While it has been a topic of study since the 1960s, the past decade has seen a surge of research activities in the area. This is due to a confluence of several factors. First, there has been an explosive increase in demand for tetherless connectivity, driven so far mainly by cellular telephony but expected to be soon eclipsed by wireless data applications. Second, the dramatic progress in VLSI technology has enabled small-area and low-power implementation of sophisticated signal processing algorithms and coding techniques. Third, the success of second-generation (2G) digital wireless standards, in particular, the IS-95 Code Division Multiple Access (CDMA) standard, provides a concrete demonstration that good ideas from communication theory can have a significant impact in practice. The research thrust in the past decade has led to a much richer set of perspectives and tools on how to communicate over wireless channels, and the picture is still very much evolving.

There are two fundamental aspects of wireless communication that make the problem challenging and interesting. These aspects are by and large not as significant in wireline communication. First is the phenomenon of *fading*: the time variation of the channel strengths due to the small-scale effect of multipath fading, as well as larger-scale effects such as path loss via distance attenuation and shadowing by obstacles. Second, unlike in the wired world where each transmitter–receiver pair can often be thought of as an isolated point-to-point link, wireless users communicate over the air and there is significant *interference* between them. The interference can be between transmitters communicating with a common receiver (e.g., uplink of a cellular system), between signals from a single transmitter to multiple receivers (e.g., downlink of a cellular system), or between different transmitter–receiver pairs (e.g., interference between users in different cells). How to deal with fading and with interference is central to the design of wireless communication

systems and will be the central theme of this book. Although this book takes a physical-layer perspective, it will be seen that in fact the management of fading and interference has ramifications across multiple layers.

Traditionally the design of wireless systems has focused on increasing the *reliability* of the air interface; in this context, fading and interference are viewed as *nuisances* that are to be countered. Recent focus has shifted more towards increasing the *spectral efficiency*; associated with this shift is a new point of view that fading can be viewed as an *opportunity* to be exploited. The main objective of the book is to provide a unified treatment of wireless communication from both these points of view. In addition to traditional topics such as diversity and interference averaging, a substantial portion of the book will be devoted to more modern topics such as opportunistic and multiple input multiple output (MIMO) communication.

An important component of this book is the *system view* emphasis: the successful implementation of a theoretical concept or a technique requires an understanding of how it interacts with the wireless system as a whole. Unlike the derivation of a concept or a technique, this system view is less malleable to mathematical formulations and is primarily acquired through experience with designing actual wireless systems. We try to help the reader develop some of this intuition by giving numerous examples of how the concepts are applied in actual wireless systems. Five examples of wireless systems are used. The next section gives some sense of the scope of the wireless systems considered in this book.

## 1.2 Wireless systems

Wireless communication, despite the hype of the popular press, is a field that has been around for over a hundred years, starting around 1897 with Marconi's successful demonstrations of wireless telegraphy. By 1901, radio reception across the Atlantic Ocean had been established; thus, rapid progress in technology has also been around for quite a while. In the intervening hundred years, many types of wireless systems have flourished, and often later disappeared. For example, television transmission, in its early days, was broadcast by wireless radio transmitters, which are increasingly being replaced by cable transmission. Similarly, the point-to-point microwave circuits that formed the backbone of the telephone network are being replaced by optical fiber. In the first example, wireless technology became outdated when a wired distribution network was installed; in the second, a new wired technology (optical fiber) replaced the older technology. The opposite type of example is occurring today in telephony, where wireless (cellular) technology is partially replacing the use of the wired telephone network (particularly in parts of the world where the wired network is not well developed). The point of these examples is that there are many situations in which there is a choice

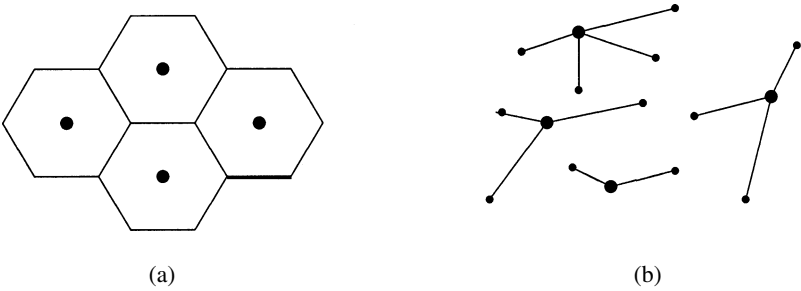
between wireless and wire technologies, and the choice often changes when new technologies become available.

In this book, we will concentrate on cellular networks, both because they are of great current interest and also because the features of many other wireless systems can be easily understood as special cases or simple generalizations of the features of cellular networks. A cellular network consists of a large number of wireless subscribers who have cellular telephones (users), that can be used in cars, in buildings, on the street, or almost anywhere. There are also a number of fixed base-stations, arranged to provide coverage of the subscribers.

The area covered by a base-station, i.e., the area from which incoming calls reach that base-station, is called a cell. One often pictures a cell as a hexagonal region with the base-station in the middle. One then pictures a city or region as being broken up into a hexagonal lattice of cells (see Figure 1.1a). In reality, the base-stations are placed somewhat irregularly, depending on the location of places such as building tops or hill tops that have good communication coverage and that can be leased or bought (see Figure 1.1b). Similarly, mobile users connected to a base-station are chosen by good communication paths rather than geographic distance.

When a user makes a call, it is connected to the base-station to which it appears to have the best path (often but not always the closest base-station). The base-stations in a given area are then connected to a *mobile telephone switching office* (MTSO, also called a *mobile switching center* MSC) by high-speed wire connections or microwave links. The MTSO is connected to the public wired telephone network. Thus an incoming call from a mobile user is first connected to a base-station and from there to the MTSO and then to the wired network. From there the call goes to its destination, which might be an ordinary wire line telephone, or might be another mobile subscriber. Thus, we see that a cellular network is not an independent network, but rather an appendage to the wired network. The MTSO also plays a major role in coordinating which base-station will handle a call to or from a user and when to handoff a user from one base-station to another.

When another user (either wired or wireless) places a call to a given user, the reverse process takes place. First the MTSO for the called subscriber is found,



then the closest base-station is found, and finally the call is set up through the MTSO and the base-station. The wireless link from a base-station to the mobile users is interchangeably called the *downlink* or the *forward channel*, and the link from the users to a base-station is called the *uplink* or a *reverse channel*. There are usually many users connected to a single base-station, and thus, for the downlink channel, the base-station must multiplex together the signals to the various connected users and then broadcast one waveform from which each user can extract its own signal. For the uplink channel, each user connected to a given base-station transmits its own waveform, and the base-station receives the sum of the waveforms from the various users plus noise. The base-station must then separate out the signals from each user and forward these signals to the MTSO.

Older cellular systems, such as the AMPS (advanced mobile phone service) system developed in the USA in the eighties, are analog. That is, a voice waveform is modulated on a carrier and transmitted without being transformed into a digital stream. Different users in the same cell are assigned different modulation frequencies, and adjacent cells use different sets of frequencies. Cells sufficiently far away from each other can reuse the same set of frequencies with little danger of interference.

Second-generation cellular systems are digital. One is the GSM (global system for mobile communication) system, which was standardized in Europe but now used worldwide, another is the TDMA (time-division multiple access) standard developed in the USA (IS-136), and a third is CDMA (code division multiple access) (IS-95). Since these cellular systems, and their standards, were originally developed for telephony, the current data rates and delays in cellular systems are essentially determined by voice requirements. Third-generation cellular systems are designed to handle data and/or voice. While some of the third-generation systems are essentially evolution of second-generation voice systems, others are designed from scratch to cater for the specific characteristics of data. In addition to a requirement for higher rates, data applications have two features that distinguish them from voice:

- Many data applications are extremely bursty; users may remain inactive for long periods of time but have very high demands for short periods of time. Voice applications, in contrast, have a fixed-rate demand over long periods of time.
- Voice has a relatively tight latency requirement of the order of 100 ms. Data applications have a wide range of latency requirements; real-time applications, such as gaming, may have even tighter delay requirements than voice, while many others, such as http file transfers, have a much laxer requirement.

In the book we will see the impact of these features on the appropriate choice of communication techniques.

As mentioned above, there are many kinds of wireless systems other than cellular. First there are the broadcast systems such as AM radio, FM radio, TV and paging systems. All of these are similar to the downlink part of cellular networks, although the data rates, the sizes of the areas covered by each broadcasting node and the frequency ranges are very different. Next, there are wireless LANs (local area networks). These are designed for much higher data rates than cellular systems, but otherwise are similar to a single cell of a cellular system. These are designed to connect laptops and other portable devices in the local area network within an office building or similar environment. There is little mobility expected in such systems and their major function is to allow portability. The major standards for wireless LANs are the IEEE 802.11 family. There are smaller-scale standards like Bluetooth or a more recent one based on ultra-wideband (UWB) communication whose purpose is to reduce cabling in an office and simplify transfers between office and hand-held devices. Finally, there is another type of LAN called an *ad hoc network*. Here, instead of a central node (base-station) through which all traffic flows, the nodes are all alike. The network organizes itself into links between various pairs of nodes and develops routing tables using these links. Here the network layer issues of routing, dissemination of control information, etc. are important concerns, although problems of relaying and distributed cooperation between nodes can be tackled from the physical-layer as well and are active areas of current research.

### 1.3 Book outline

---

The central object of interest is the wireless fading channel. Chapter 2 introduces the multipath fading channel model that we use for the rest of the book. Starting from a continuous-time passband channel, we derive a discrete-time complex baseband model more suitable for analysis and design. Key physical parameters such as coherence time, coherence bandwidth, Doppler spread and delay spread are explained and several statistical models for multipath fading are surveyed. There have been many statistical models proposed in the literature; we will be far from exhaustive here. The goal is to have a small set of example models in our repertoire to evaluate the performance of basic communication techniques we will study.

Chapter 3 introduces many of the issues of communicating over fading channels in the simplest point-to-point context. As a baseline, we start by looking at the problem of detection of uncoded transmission over a narrowband fading channel. We find that the performance is very poor, much worse than over the additive white Gaussian noise (AWGN) channel with the same average signal-to-noise ratio (SNR). This is due to a significant probability that the channel is in *deep fade*. Various *diversity techniques* to mitigate this adverse effect of fading are then studied. Diversity techniques increase

reliability by sending the same information through multiple independently faded paths so that the probability of successful transmission is higher. Some of the techniques studied include:

- interleaving of coded symbols over time to obtain time diversity;
- inter-symbol equalization, multipath combining in spread-spectrum systems and coding over sub-carriers in orthogonal frequency division multiplexing (OFDM) systems to obtain frequency diversity;
- use of multiple transmit and/or receive antennas, via *space-time* coding, to obtain spatial diversity.

In some scenarios, there is an interesting interplay between channel uncertainty and the diversity gain: as the number of diversity branches increases, the performance of the system first improves due to the diversity gain but then subsequently deteriorates as channel uncertainty makes it more difficult to combine signals from the different branches.

In Chapter 4 the focus is shifted from point-to-point communication to studying cellular systems as a whole. Multiple access and inter-cell interference management are the key issues that come to the forefront. We explain how existing digital wireless systems deal with these issues. The concepts of frequency reuse and cell sectorization are discussed, and we contrast narrowband systems such as GSM and IS-136, where users within the same cell are kept orthogonal and frequency is reused only in cells far away, and CDMA systems, such as IS-95, where the signals of users both within the same cell and across different cells are spread across the same spectrum, i.e., frequency reuse factor of 1. Due to the full reuse, CDMA systems have to manage intra-cell and inter-cell interference more efficiently: in addition to the diversity techniques of time-interleaving, multipath combining and soft handoff, *power control* and *interference averaging* are the key interference management mechanisms. All the five techniques strive toward the same system goal: to maintain the channel quality of each user, as measured by the signal-to-interference-and-noise ratio (SINR), as constant as possible. This chapter is concluded with the discussion of a wideband OFDM system, which combines the advantages of both the CDMA and the narrowband systems.

Chapter 5 studies the capacity of wireless channels. This provides a higher level view of the tradeoffs involved in the earlier chapters and also lays the foundation for understanding the more modern developments in the subsequent chapters. The performance over the (non-faded) AWGN channel is used as a baseline for comparison. We introduce the concept of *channel capacity* as the basic performance measure. The capacity of a channel provides the fundamental limit of communication achievable by any scheme. For the fading channel, there are several capacity measures, relevant for different scenarios. Two distinct scenarios provide particular insight: (1) the *slow* fading channel, where the channel stays the same (random value) over the entire time-scale

of communication, and (2) the *fast* fading channel, where the channel varies significantly over the time-scale of communication.

In the slow fading channel, the key event of interest is *outage*: this is the situation when the channel is so poor that no scheme can communicate reliably at a certain target data rate. The largest rate of reliable communication at a certain outage probability is called the outage capacity. In the fast fading channel, in contrast, outage can be avoided due to the ability to average over the time variation of the channel, and one can define a positive capacity at which arbitrarily reliable communication is possible. Using these capacity measures, several resources associated with a fading channel are defined: (1) diversity; (2) number of degrees of freedom; (3) received power. These three resources form a basis for assessing the nature of performance gain by the various communication schemes studied in the rest of the book.

Chapters 6 to 10 cover the more recent developments in the field. In Chapter 6 we revisit the problem of multiple access over fading channels from a more fundamental point of view. Information theory suggests that if both the transmitters and the receiver can track the fading channel, the optimal strategy to maximize the total system throughput is to allow only the user with the best channel to transmit at any time. A similar strategy is also optimal for the downlink. Opportunistic strategies of this type yield a system-wide *multiuser diversity* gain: the more users in the system, the larger the gain, as there is more likely to be a user with a very strong channel. To implement this concept in a real system, three important considerations are: *fairness* of the resource allocation across users; *delay* experienced by the individual user waiting for its channel to become good; and *measurement inaccuracy* and *delay* in feeding back the channel state to the transmitters. We discuss how these issues are addressed in the context of IS-865 (also called HDR or CDMA 2000 1× EV-DO), a third-generation wireless data system.

A wireless system consists of multiple dimensions: time, frequency, space and users. Opportunistic communication maximizes the spectral efficiency by measuring when and where the channel is good and only transmits in those degrees of freedom. In this context, channel fading is *beneficial* in the sense that the fluctuation of the channel across the degrees of freedom ensures that there will be some degrees of freedom in which the channel is very good. This is in sharp contrast to the diversity-based approach in Chapter 3, where channel fluctuation is always detrimental and the design goal is to average out the fading to make the overall channel as constant as possible. Taking this philosophy one step further, we discuss a technique, called *opportunistic beamforming*, in which channel fluctuation can be *induced* in situations when the natural fading has small dynamic range and/or is slow. From the cellular system point of view, this technique also increases the fluctuations of the *interference* imparted on adjacent cells, and presents an opposing philosophy to the notion of interference averaging in CDMA systems.



Chapters 7, 8, 9 and 10 discuss multiple input multiple output (MIMO) communication. It has been known for a while that the uplink with multiple receive antennas at the base-station allow several users to simultaneously communicate to the receiver. The multiple antennas in effect increase the number of degrees of freedom in the system and allow spatial separation of the signals from the different users. It has recently been shown that a similar effect occurs for point-to-point channels with multiple transmit *and* receive antennas, i.e., even when the antennas of the multiple users are co-located. This holds provided that the scattering environment is rich enough to allow the receive antennas to separate out the signal from the different transmit antennas, allowing the *spatial multiplexing* of information. This is yet another example where channel fading is beneficial to communication. Chapter 7 studies the properties of the multipath environment that determine the amount of spatial multiplexing possible and defines an *angular domain* in which such properties are seen most explicitly. We conclude with a class of statistical MIMO channel models, based in the angular domain, which will be used in later chapters to analyze the performance of communication techniques.

Chapter 8 discusses the capacity and capacity-achieving transceiver architectures for MIMO channels, focusing on the fast fading scenario. It is demonstrated that the fast fading capacity increases linearly with the minimum of the number of transmit and receive antennas at all values of SNR. At high SNR, the linear increase is due to the increase in degrees of freedom from spatial multiplexing. At low SNR, the linear increase is due to a power gain from receive beamforming. At intermediate SNR ranges, the linear increase is due to a combination of both these gains. Next, we study the transceiver architectures that achieve the capacity of the fast fading channel. The focus is on the V-BLAST architecture, which multiplexes independent data streams, one onto each of the transmit antennas. A variety of receiver structures are considered: these include the decorrelator and the linear minimum mean square-error (MMSE) receiver. The performance of these receivers can be enhanced by successively canceling the streams as they are decoded; this is known as successive interference cancellation (SIC). It is shown that the MMSE–SIC receiver achieves the capacity of the fast fading MIMO channel.

The V-BLAST architecture is very suboptimal for the slow fading MIMO channel: it does not code across the transmit antennas and thus the diversity gain is limited by that obtained with the receive antenna array. A modification, called D-BLAST, where the data streams are *interleaved* across the transmit antenna array, achieves the outage capacity of the slow fading MIMO channel. The boost of the outage capacity of a MIMO channel as compared to a single antenna channel is due to a combination of both diversity and spatial multiplexing gains. In Chapter 9, we study a fundamental *tradeoff* between the diversity and multiplexing gains that can be simultaneously harnessed over a slow fading MIMO channel. This formulation is then used as a unified framework to assess both the diversity and multiplexing performance



of several schemes that have appeared earlier in the book. This framework is also used to motivate the construction of new tradeoff-optimal space-time codes. In particular, we discuss an approach to design *universal* space-time codes that are tradeoff-optimal.

Finally, Chapter 10 studies the use of multiple transmit and receive antennas in multiuser and cellular systems; this is also called *space-division multiple access* (SDMA). Here, in addition to providing spatial multiplexing and diversity, multiple antennas can also be used to mitigate interference between different users. In the uplink, interference mitigation is done at the base-station via the SIC receiver. In the downlink, interference mitigation is also done at the base-station and this requires *precoding*: we study a precoding scheme, called Costa or dirty-paper precoding, that is the natural analog of the SIC receiver in the uplink. This study allows us to relate the performance of an SIC receiver in the uplink with a corresponding precoding scheme in a *reciprocal* downlink. The ArrayComm system is used as an example of an SDMA cellular system.

## CHAPTER

## 2

## The wireless channel

A good understanding of the wireless channel, its key physical parameters and the modeling issues, lays the foundation for the rest of the book. This is the goal of this chapter.

A defining characteristic of the mobile wireless channel is the variations of the channel strength over time and over frequency. The variations can be roughly divided into two types (Figure 2.1):

- *Large-scale fading*, due to path loss of signal as a function of distance and shadowing by large objects such as buildings and hills. This occurs as the mobile moves through a distance of the order of the cell size, and is typically frequency independent.
- *Small-scale fading*, due to the constructive and destructive interference of the multiple signal paths between the transmitter and receiver. This occurs at the spatial scale of the order of the carrier wavelength, and is frequency dependent.

We will talk about both types of fading in this chapter, but with more emphasis on the latter. Large-scale fading is more relevant to issues such as cell-site planning. Small-scale multipath fading is more relevant to the design of reliable and efficient communication systems – the focus of this book.

We start with the physical modeling of the wireless channel in terms of electromagnetic waves. We then derive an input/output linear time-varying model for the channel, and define some important physical parameters. Finally, we introduce a few statistical models of the channel variation over time and over frequency.

### 2.1 Physical modeling for wireless channels

Wireless channels operate through electromagnetic radiation from the transmitter to the receiver. In principle, one could solve the electromagnetic field equations, in conjunction with the transmitted signal, to find the