Multiuser Wireless Communication Systems

Ashutosh Sabharwal and Behnaam Aazhang

Department of Electrical and Computer Engineering

Rice University

Houston TX 77005

Abstract

Wireless cellular systems have grown dramatically in the last two decades, thanks to several key innovations in communication algorithms and high speed silicon technology. We review fundamental physical layer techniques for the future high speed wireless networks.

Keywords : Wireless communications, channel estimation, multiuser detection, channel coding, multiple antenna, diversity, power control.

1 Introduction

The last two decades have been a witness to the rapid growth and widespread success of wireless connectivity. The success of wireless systems is largely due to breakthroughs in communication theory and progress in the design of low-cost power efficient mobile devices. Beyond the widespread use of voice telephony, new technologies are replacing wires in virtually all modes of communication. For example, in addition to widely recognized outdoor connectivity via cellular wide area networks (WANs), wireless local area networks (LANs) and wireless personal area networks (PANs) have also become popular. Wireless LANs (*e.g.*, IEEE 802.11) provide high speed untethered access inside buildings replacing traditional wired Ethernet, and wireless PANs (*e.g.*, Bluetooth) are replacement for wires between common peripherals like mouse, keyboard, PDAs and printers.

Providing ubiquitous mobile access to a large number of users requires solution to a wide spectrum of scientific and economic issues, ranging from low-power semiconductor design and advanced signal processing algorithms to the design and deployment of large cellular networks. In this paper, we will highlight the challenges in the design of advanced signal processing algorithms for high speed outdoor cellular access. The signal processing algorithms form the core of all wireless systems, and are thus critical for their success. In addition, the techniques and algorithms discussed in this paper form a

basis for most wireless systems, and thus have a wider applicability than outdoor wireless systems. To keep the discussion tractable, we will focus on baseband design for third generation wireless cellular systems (*e.g.*, WCDMA or CDMA2000) based on code division multiple access (CDMA).

Wireless channel is a shared resource, *i.e.*, multiple users in the same geographical locale have to contend for the common spectral resource and in the process interfere with other users. To allow meaningful and resource efficient communication between different users, it is crucial that all participating users agree on a common protocol. The common protocol should enable fair access to the shared resource for all users. The three most commonly used multiple access protocols[1] are time division (TDMA), frequency-division (FDMA) and code-division (CDMA) multiple access. Among the three, direct-sequence CDMA (DS-CDMA) has been adopted as the access technique for all the third generation wireless standards, and thus will be the main focus of this article.

In outdoor cellular systems, the coverage area is divided into smaller regions called *cells*, each capable of supporting a subset of the users subscribing to the cellular system. The cellular structure exploits the fact that electromagnetic signals suffer loss in power with distance, thereby allowing reuse of the same communication channel at another spatially separated location. The reuse of communication channels allows a cellular system to support many more users as compared to a system which treats the whole geographical region as one cell. Each cell is served by a *base station* which is responsible for operations within a cell, primarily serving calls to and from users located in the respective cell. Figure 1 shows the components of a typical cellular system. The size and distribution of the cells [1] are dictated by the coverage area of the base station, subscriber density and projected demand within a geographical region. As mobile users travel from cell to cell, their calls are *handed off* between cells in order to maintain seamless service. The base stations are connected to the *mobile telephone switching office* (MTSO) that serves as a controller to a group of base stations and as an interface with the fixed wired backbone.

Wireless networks, like typical multiple access networks, have a layered architecture [2, 3]. The three main layers of each network are the physical layer, the network layer[2] and the application layer. The physical layer is responsible for actual transport of data between the source and the destination points. The network layer controls the communication session, and the user applications operate in the application layer. Both network and application layer design are critical in wireless networks, and are areas of active research. In this paper, our focus will be on the design of physical layer for wireless networks.

---

[1]We limit our discussion to circuit-switched networks and deterministic multiple access schemes. In packet-switched networks, probabilistic multiple access is used; a good example is contention avoidance/resolution based protocol used in IEEE 802.11, and packet services used in EGPRS and 3G systems.

[2]Network layer consists of several layers which among others include multiple access layer (MAC), data link layer and transport layer.

The rest of the paper is organized as follows. In Section 2, we will briefly discuss the three major challenges in the design of wireless systems and commonly used methods to combat them. Models for wireless channels are discussed in Section 3. In Section 4, we will introduce information theoretic methods to analyze the limits of wireless systems. The core of the paper is in Section 5, which discusses various aspects in the design of a typical transceiver. We conclude in Section 6.

2   Challenges and Design of Wireless Systems

In this section, we highlight the major challenges and techniques employed in wireless system design.

- **Time varying multipath**: Enabling mobility, which is the fundamental premise in designing wireless systems and is the major reason for their success, also presents itself as the most fundamental challenge. Due to mobility of users and their surrounding environment, wireless channels are generally time-varying. Electromagnetic signals transmitted by base-station or mobile users reach the intended receiver via several paths; the multiple paths are caused by reflections from man-made and natural objects (Figure 2). Since the length of the each path may be different, the resultant received signal shows a wide fluctuations in its power profile (Figure 3), thereby complicating the design of spectrally efficient systems.

  To combat time-varying fading, a combination of time, spatial or frequency diversity is commonly used [4]. By using diversity techniques, the receiver obtains multiple copies of the transmitted signal, thereby increasing the chance that at least one of the copies is reliable. To exploit time diversity, error control codes are used in conjunction with an interleaver [4]. Spatial diversity can be obtained by using multiple antennas which are sufficiently separated. Spatial diversity can be tapped by using space-time codes [5] at the transmitter or signal combining [6] at the receiver. Spatial diversity techniques have recently received considerable interest due to their potential to support larger data rates on the same channels compared to current technology. Frequency diversity is analogous to spatial diversity where frequency selectivity due to multipath is used.

- **Shared multiple access**: Unlike wired networks, where new bandwidth is "created" by adding additional physical resources (cables, servers, etc.), users in wireless system have to share limited spectral resources. Although, the available spectrum for commercial wireless system has increased in the last two decades, it is clear that growth in demand will always outpace the available spectrum. Limited growth of resources immediately implies that the requirements of new data rate hungry wireless services can only be sustained by progress in efficiently using the available spectrum. An obvious way of increasing system capacity is to use smaller cells, but using smaller

cells is undesirable due to economic reasons; increased number of base-stations and the required wired backbone are the major reasons for the increased system cost. Further, smaller cells generally lead to increased intercell handoffs and out-of-cell interference, leading to diminishing returns with increasing cell partitioning.

The capacity of cellular systems can also be improved by cell sectorization [7, 8], where each cell is further divided into sectors. Cell sectorization is achieved by using multiple directional antenna [9] at each base-station, thereby reducing the inter-sector interference. Due to the directional antenna response, cell sectorization has also been shown to reduce the delay spread of the received signal leading to power savings [10]. Much like cell splitting, cell sectorization has also its limits too. To achieve smaller sectors using directional antennas requires increasingly large size antennas, which are both expensive and hard to deploy.

Information theoretic results [11] for multiuser systems indicate that the optimal methods to share spectral resources should not attempt to avoid inter-cell and intra-cell interference. The co-channel interference in wireless systems can be suppressed by using multiuser detection [12], leading to increased spectral efficiency [13, 14]. Further improvements in system capacity can be obtained by the use of dynamic resource allocation among users, for example, adaptive channel assignment techniques [15], and dynamic spreading gain and power control [16].

- **Power limitation for mobile users**: Since most of the mobile devices are battery operated, power efficiency is a crucial design parameter in wireless system design. The major consumers of power in wireless handsets are power amplifier used during transmission, silicon based computing units (A/D, D/A and baseband processor) used in reception, and in some cases, the color display.

  Power dissipation in the RF power amplifier can be reduced by using cells with smaller radii, better multiuser signal processing at the base-station, improved coding schemes or receiver diversity. As pointed out earlier, cell splitting is not attractive due to increased system cost with diminishing returns. Advanced signal processing, multiuser channel estimation, and data detection have been shown to greatly reduce the power requirements to achieve a desired performance level [12]. Recent advances in channel coding, namely turbo coding [17], can lead to further reduction in power requirements for the transmitter to achieve a desired performance level. Reduction in power requirements of baseband processing units requires development of hardware-frugal algorithms and low power CMOS circuits. Also, techniques which require more computation at the base-station to cut the complexity of handset are very effective in saving power at the mobile unit.

4

3 Fading Channel Models

In this section, we will describe time-varying wireless channels and the statistical models used to capture their effect on transmitted signals. A detailed discussion of channel models can be found in [4, 18]. A fading multipath channel is generally modeled as a linear system with time-varying impulse response[3] $h(t; \tau)$. The time-varying impulse response is assumed to be a wide-sense stationary random process with respect to the time variable t. Due to time variations of the channel, the transmitted signal is spread in frequency; the frequency spreading is called *Doppler spreading*. The transmitted signal also suffers time spreading due to multipath propagation. Thus, the received signal is spread both in time and frequency.

Two parameters are commonly used to characterize wide-sense stationary channels: *multipath delay spread* and *Doppler spread*. To define the multipath delay and Doppler spread, it is convenient to work with the scattering function $\mathcal{H}(\tau; \lambda)$, which is a measure of average power output[4] of the channel at delay τ and frequency offset λ relative to the carrier. The *delay power spectrum* of the channel is obtained by averaging $\mathcal{H}(\tau; \lambda)$ over λ, *i.e.*,

$$\mathcal{H}_c(\tau) = \int_{-\infty}^{\infty} \mathcal{H}(\tau; \lambda) d\lambda. \tag{1}$$

The multipath delay spread T_m is the maximum delay τ for which delay power spectrum $\mathcal{H}_c(\tau)$ is non-zero. Similarly, the Doppler spread B_d is the maximum value of λ for which the following *Doppler power spectrum* $\mathcal{H}_c(\lambda)$ is non-zero,

$$\mathcal{H}_c(\lambda) = \int_{-\infty}^{\infty} \mathcal{H}(\tau; \lambda) d\tau. \tag{2}$$

The reciprocal of the multipath delay spread is defined as *channel coherence bandwidth*, $B_{coh} = 1/T_m$ and provides an indication of the width of band of frequencies which are similarly affected by the channel. The Doppler spread provides a measure of how fast the channel variations are in time. The reciprocal of Doppler spread is called *channel coherence time* $T_{coh} = 1/B_d$. A large value of T_{coh} represents a slowly fading channel and a small values represents fast fading. If $T_m B_d < 1$, then the channel is said to be *underspread*, else it is *overspread*. In general, if $T_m B_d \ll 1$, then the channel can be accurately measured at the receiver, which can aid in improving the transmission schemes. On the other hand, channel measurement is unreliable for the case of $T_m B_d > 1$.

An appropriate model for a given channel also depends on the transmitted signal bandwidth. If $s(t)$ is the transmitted signal with the Fourier transform $S(f)$, the received baseband signal, with the

---

[3]A linear time-invariant system requires a single-variable transfer function. For a time-varying linear system, two parameters are needed; the parameter t in $h(t; \tau)$ captures the time-variability of the channel.

[4]Under the assumption that all different delayed paths propagating through the channel are uncorrelated.

additive noise, is

$$
\begin{aligned}
z(t) &= \int_{-\infty}^{\infty} h(t;\tau)s(t-\tau)d\tau + \nu(t) \\
&= \int_{-\infty}^{\infty} H(t;f)S(f)e^{j2\pi ft}df + \nu(t),
\end{aligned}
$$

where $H(t;f)$ is the Fourier transform of $h(t;\tau)$ with respect to τ. If the bandwidth W of the transmitted signal $S(f)$ is much smaller than the coherence bandwidth, *i.e.*, $W \ll B_{coh}$, then all the frequency components in $S(f)$ undergo the same attenuation and phase shift during propagation. This implies that within the bandwidth of the signal, the transfer function $H(t;f)$ is constant in f, leading to a *frequency nonselective* or *flat fading*. Thus, the received signal can be rewritten as

$$
\begin{aligned}
z(t) &= H(t;0)\int_{-\infty}^{\infty} S(f)e^{j2\pi ft}df + \nu(t) \\
&= H(t)s(t) + \nu(t), \tag{3}
\end{aligned}
$$

where $H(t) \in \mathbb{C}$ is the complex multiplicative channel. A flat fading channel is said to be *slowly fading* if the symbol time duration of the transmitted signal T_s is much smaller than the coherence time of the channel, $T_s \ll T_{coh}$. The channel is labeled as *fast fading* if $T_s \geq T_{coh}$.

If the signal bandwidth W is much greater than the coherence bandwidth of the channel, then the frequency components of $S(f)$ with frequency separation more than B_{coh} are subjected to different attenuations and phase shifts. Such a channel is called *frequency selective*. In this case, multipath components separated by delay more than $1/W$ are resolvable and the channel impulse response can be written as [4]

$$
h(t;\tau) = \sum_{p=1}^{P} h_p(t)\delta(\tau - p/W). \tag{4}
$$

Since the multipath delay spread is T_m and the time resolution of multipaths is $1/W$, the number of paths L is given by $\lfloor T_m W \rfloor + 1$. In general, the time-varying tap coefficients $h_p(t)$ are modeled as mutually uncorrelated wide-sense stationary processes. The random time-variation of the channel are generally modeled via a probability distribution on the channel coefficients $h_p(t)$. The most commonly used probability distributions are Rayleigh, Ricean and the Nakagami-m [4].

The main purpose of the channel modeling is to characterize the channel in a tractable yet meaningful manner, to allow design and analysis of the communication algorithms. Note that *all* models are approximate representation of the actual channel, and thus development of practical systems requires both theoretical analysis and field testing.

In the sequel, we will consider only slowly fading channels, where $T_s \ll T_{coh}$, *i.e.*, multiple consecutive symbols or equivalently, a block of symbols undergo the same channel distortion. Hence, these channels

are also referred as *block fading channels* [19–22]. As a result of slow time-variation of the channel, the time dependency of the channel will be suppressed, *i.e.*, $h(t)$ will be denoted by h and $h(t; \tau)$ by $h(\tau)$.

4   Capacity of Multiple Access Channels

Developed in the landmark paper by Shannon [23], information theory forms the mathematical foundation for source compression, communication over noisy channels and cryptography. Among other important contributions in [23], the concept of *channel capacity* was developed. It was shown that a noisy channel can be characterized by its capacity, which is the maximum rate at which the information can be transmitted reliably over that channel. Information theoretic methods not only provide the ultimate achievable limits of a communication system, but also provide valuable insight into the design of practical systems.

Typically, a capacity analysis starts by using a simple model of the physical phenomenon. The simplified model captures the basic elements of the problem, such as time-varying fading wireless channel, shared multiple access and power-limited sources. Information theoretic analysis then leads to limits on reliably achievable data rates and provides guidelines to achieve those limits. Although information theoretic techniques are rarely practical, information theory inspired coding, modulation, power control and multiple access methods have led to significant advances in practical systems. Furthermore, the analysis techniques allow performance evaluation of suboptimal but implementation-friendly techniques, thereby providing a useful benchmarking methodology.

In this section, we will provide a brief sampling of results pertaining to time-varying fading wireless channels; the reader is referred to [19] for a detailed review. Our aim is to highlight basic single and multiuser results for fading channels to motivate the algorithms discussed in the sequel. In Section 4.1, we will first introduce two notions of channel capacity, Shannon-theoretic capacity [23] and outage capacity [24]. Capacity of a channel characterizes its performance limits using *any* practical transmitter-receiver pair and is a fundamental notion in evaluating efficacy of practical systems. Single-user fading channels will be analyzed using the two capacity notions, motivating the importance of diversity techniques (like space-time coding and beamforming) and power control. In Section 4.2, the multiuser extensions will be discussed to motivate the use of power controlled CDMA based multiple access.

All results in this section will be given for flat fading channels. The results can be easily extended to frequency selective fading by partitioning the channel into frequency bins of width B_{coh}, and then treating each bin as a separate channel.

4.1 Capacity of Single-user Fading Channels

A channel is deemed *noisy* if it introduces random perturbations in the transmitted signals. In [23], the capacity of a noisy channel was defined as the highest data rate at which reliable communication is possible across that channel. *Communication reliability* is defined as the probability that the receiver will decode the transmitted message correctly; higher reliability means lower errors in decoding messages and vice versa. An information rate is *achievable* if there exists at least one transmission scheme such that any preset level of communication reliability can be achieved. To achieve this (arbitrary level of) reliability, the transmitter can choose any codebook to map information message sequences to channel inputs. If the rate of transmission R is no more than the channel capacity C, then reliable communication is possible by using codebooks which jointly encode increasingly longer input messages. The above notion of channel capacity is commonly referred as *Shannon-theoretic capacity.*

Besides providing a characterization of the channel capacity for a broad class of channels, Shannon [23] also computed the capacity of the following additive white Gaussian noise (AWGN) channel,

$$z(t) = s(t) + \nu(t), \tag{5}$$

as

$$C = W \log_2 \left(1 + \frac{\mathcal{P}_{av}}{\sigma^2} \right) \text{ bits/second.} \tag{6}$$

Note that the AWGN channel in (5) can be considered as a special case of fading channel (3) with $h(t) \equiv 1$. In (6), W represents the channel bandwidth (in Hertz), $\mathcal{P}_{av} = \mathbb{E}_s\{|s(t)|^2\}$ is the average transmitted power over time[5] and σ^2 is the variance of the additive noise $\nu(t)$. The fundamental formula (6) clarifies the role of two important system parameters, the channel bandwidth W and signal to noise ratio (SNR), $\mathcal{P}_{av}/\sigma^2$. The capacity result (6) claims a surprising fact that even for very small amount of power or bandwidth, information can be sent at a non-zero rate with vanishingly few decoding errors. To achieve this reliable communication, the transmitter *encodes* multiple information bits together using a channel code. The encoded bits are then jointly decoded by the receiver to correct errors introduced by the channel (5).

The capacity analysis in [23] forms the basis for deriving capacity of fading channels (3), which we review next. With an average transmitted power constraint, $\mathbb{E}_s\{|s(t)|^2\} \leq \mathcal{P}_{av}$, the Shannon-theoretic capacity of fading channels, with perfect channel information at the receiver, is given by [25]

$$C_{sc}^r = W \mathbb{E}_\gamma \left\{ \log_2 \left(1 + \frac{\mathcal{P}_{av}\gamma(t)}{\sigma^2} \right) \right\}, \tag{7}$$

---

[5]The expectation $\mathbb{E}_s\{|s(t)|^2\}$ represents an average computed over time (assuming that it exists) using the distribution of $s(t)$.

where σ^2 is the variance of the additive i.i.d. Gaussian noise $\nu(t)$ in (3), and $\gamma(t) = |h(t)|^2$ is the received instantaneous power. The expectation in (7) is computed with respect to the probability distribution of the variable $\gamma(t)$. If in addition to perfect channel information at the receiver, the transmitter has knowledge of the instantaneous channel realization, then the transmitter can adapt its transmission strategy based on the channel. The optimal strategy, in this case, turns out to be "water-filling" in time [26]. To water-fill in time, the transmitter waits for the good channel conditions to transmit and does not transmit during poor channel conditions. Thus, the optimal transmission policy is a constant rate Gaussian codebook (see [11] for details on Gaussian codebooks) transmitted using an instantaneous channel SNR dependent power. The optimal transmission power is given by [26]

$$
\mathcal{P}_{sc}(\gamma(t)) = \begin{cases} \mathcal{P}_{av}\left(\frac{1}{\gamma_{sc}} - \frac{1}{\gamma(t)}\right) & , \gamma(t) \geq \gamma_{sc} \\ 0 & , \gamma(t) < \gamma_{sc} \end{cases},
\tag{8}
$$

where the threshold γ_{sc} is found to satisfy the power constraint $\mathbb{E}_{\gamma,s}\left\{\mathcal{P}_{sc}(\gamma(t))|s(t)|^2\right\} \leq \mathcal{P}_{av}$. The achievable capacity is then given by

$$
C_{sc}^{rt} = W\mathbb{E}_{\gamma}\left\{\log_2\left(1 + \frac{\mathcal{P}_{sc}(\gamma(t))\gamma(t)}{\sigma^2}\right)\right\}.
\tag{9}
$$

Note that allocated power in (8) is zero for poor channels whose SNR is less than $\gamma_{sc}(t)$ and increases monotonically as channels conditions improve. Adapting the transmission power based on channel conditions is known as *power control*. Channel state information at the transmitter leads to only modest gains for most fading distributions [26] with a single transmitter and receiver, *i.e.*, C_{sc}^{rt} is only marginally greater than C_{sc}^{r}. But the gains of transmitter information increase dramatically with multiple transmit and receive antennas. Using the extensions of (7) and (9) to multiple antennas [25, 27], a representative example is shown in Figure 4. Thus, building adaptive power control policies is more useful for multiple antenna systems; see [28] for practical methods to achieve a significant portion of this capacity in a practical system. The gain due to channel state information at the transmitter can also be achieved by using imprecise channel information [28–30]. The large gains promised by multiple antenna diversity, with or without channel information at the transmitter, have sparked the rich field of space-time coding [5, 31, 32].

In slow fading channels, achieving Shannon-theoretic capacity requires coding over exceedingly long input blocks. The long codewords are required to average over different fading realizations, which then allow the use of assumed ergodicity[6] of the fading process to prove the capacity theorem. The

---

[6]A stochastic process $h(t)$ is called ergodic if its ensemble averages equal appropriate time averages [33]. The channel capacity theorem proved by Shannon [23] relied on law of large numbers, *i.e.*, the time averages converge to their ensemble averages, which in turn motivated the idea of encoding increasingly long blocks of input messages. Ergodic channels

large delays associated with Shannon-theoretic capacity directly translate into impractical delays in delay sensitive applications like voice and video. Thus, with a delay constraint, the Shannon-theoretic capacity of slowly fading practical channels (more specifically, non-ergodic channels) is zero [24]. In [24], the concept of capacity versus outage was introduced, which captures the effect of delay in slow fading channels. A block of transmitted data, which is assumed to undergo the same fading throughout, is in *outage* if the instantaneous capacity of the channel is less than the rate of transmission. The concept of outage provides a code-independent method (by using asymptotic approximations) to gauge the codeword error probability for practical codes. Assuming that the flat fading channel h is constant for a block of transmitted data, the instantaneous capacity is given by[7] $W \log_2(1 + \mathcal{P}_{av}\gamma(t)/\sigma^2)$. The outage probability, when only the receiver is aware of the channel state, is then given by

$$\Pi_{oc}^r = \mathrm{Prob}\left(W \log_2\left(1 + \frac{\mathcal{P}_{av}\gamma(t)}{\sigma^2}\right) < R\right), \tag{10}$$

where the probability is computed over the distribution of channel $h(t)$. Analogous to the above Shannon-theoretic capacity analysis, the probability of outage can also be computed for different amount of channel state information at the transmitter. With perfect channel state information at the transmitter and receiver, the outage probability is given by

$$\Pi_{oc}^{rt} = \min_{\mathcal{P}_{oc}(\gamma(t))} \mathrm{Prob}\left(W \log_2\left(1 + \frac{\mathcal{P}_{oc}(\gamma(t))\gamma(t)}{\sigma^2}\right) < R\right). \tag{11}$$

The power allocation $\mathcal{P}_{oc}(\gamma(t))$ minimizing the outage is given by [27]

$$\mathcal{P}_{oc}(\gamma(t)) = \begin{cases} \frac{\sigma^2(2^{R/W}-1)}{\gamma(t)} & ,\gamma(t) \geq \gamma_{oc} \\ 0 & ,\gamma(t) < \gamma_{oc} \end{cases} \tag{12}$$

The threshold γ_{oc} is chosen to meet the average power constraint, $\mathbb{E}_{\gamma,s}\{\mathcal{P}_{oc}(\gamma(t))|s(t)|^2\} \leq \mathcal{P}_{av}$. The *outage capacity*, which measures the total number of transmitted bits per unit time not suffering an outage, is given by

$$\begin{aligned} C_{oc}^r &= (1 - \Pi_{out}^r)R \\ C_{oc}^{rt} &= (1 - \Pi_{out}^{rt})R \end{aligned}$$

Due to the extra information at the transmitter, it immediately follows that $\Pi_{out}^{rt} < \Pi_{out}^r$ and hence $C_{oc}^{rt} > C_{oc}^r$. The gain in outage capacity due to transmitter information is much more substantial

---

are the most general channels with dependency across time for which the (strong) law of large numbers holds, thereby allowing a direct extension of capacity theorem [23] to ergodic channels. For a more general capacity theorem without any assumptions on channel structure, see [34].

[7]Assuming that the transmitter is unaware of the instantaneous channel state and receiver has the perfect knowledge of $h(t)$ [25].

compared to Shannon-capacity even for a single antenna system [35]. Similar to the Shannon-capacity, outage capacity increases with the increasing number of transmit and receive antennas [25, 36].

The differences in the objectives of achieving outage capacity versus achieving Shannon-theoretic capacity can be better appreciated by the difference in the optimal power allocation schemes, $\mathcal{P}_{sc}(\gamma(t))$ and $\mathcal{P}_{oc}(\gamma(t))$. In the Shannon-theoretic approach, the transmitter uses more power in the good channel states and less power during poor channel conditions. On the other hand, to minimize outage the transmitter employs *more* power as the channel gets *worse*, which is exactly opposite to the power allocation $\mathcal{P}_{sc}(\gamma(t))$. The difference in power allocation strategies, $\mathcal{P}_{sc}(\gamma(t))$ and $\mathcal{P}_{oc}(\gamma(t))$ can be attributed to optimization goals: Shannon-theoretic capacity maximizes long-term throughput and hence it is not delay-constrained, and outage capacity maximizes short-term throughput with delay constraints.

Irrespective of the capacity notion, the main lesson learnt from information theoretic analysis is that diversity and channel information at the transmitter can potentially lead to large gains in fading channels. The gains promised by above information theoretic results have motivated commonly used methods of space-time coding and power control to combat fading. Readers are referred to [21, 25, 26, 36–38] for detailed results on capacity of single user flat fading channels. In the next section, we will briefly discuss the results for multiple access channels and their impact on the choice of multiple access protocols.

4.2 Multiple User Fading Channels

The primary question of interest in a multiuser analysis is the multiaccess protocol to efficiently share the spectral resources among several power-limited users. An accurate capacity analysis of a complete cellular system is generally intractable. Hence, the information theoretic analysis relies on a series of simplifying assumptions to understand the dominant features of the problem. Our main emphasis will be on uplink communication in a single cell, where multiple users simultaneously communicate with a single receiver, the base-station.

The sampled received baseband signal at the base-station is the linear superposition of K user signals in additive white Gaussian noise, given by

$$y(t) = \sum_{i=1}^{K} h_i(t) s_i(t) + \nu(t), \tag{13}$$

The Gaussian noise $\nu(t)$ is assumed to be zero mean with variance σ^2. The channels for all users $h_i(t)$ are assumed to vary independently of each other and from one coherence interval to another. The fading processes for all users are assumed to be jointly stationary and ergodic. Furthermore, each user is subjected to an average power constraint, $\mathbb{E}_{s_i}\{|s_i(t)|^2\} \leq \mathcal{P}_i$.

11

Equivalent to the capacity of channel in the single-user case, a *capacity region* specifying all the rates which can be simultaneously and reliably achieved are characterized. Thus, the capacity region for K users is a set of rates defined as

$$\mathcal{R} = \{\underline{R} = (R_1, R_2, \ldots, R_K) : \text{ Rates } R_i \text{ can be reliably achieved simultaneously } \}. \tag{14}$$

When the base-station receiver is aware of all the fading realizations of all the users, $\{h_i(t)\}$, then the rate region is described by the following set of inequalities (in the single user case, there is only one inequality, $R \leq C$),

$$\sum_{i \in \mathcal{B}} R_i \leq \mathbb{E}_{\underline{\gamma(t)}} \log_2 \left(1 + \frac{\sum_{i \in \mathcal{B}} \gamma_i(t) \mathcal{P}_{av}}{\sigma^2} \right) \tag{15}$$

where it is assumed that each user has the same average power limit $\mathcal{P}_i = \mathcal{P}_{av}$. In (15), $\mathcal{B}$ represents a subset of $\{1, 2, \ldots, K\}$ and $\gamma_i(t) = |h_i(t)|^2$ is the received power. The expectation of $\mathbb{E}_{\underline{\gamma(t)}}$ is over all the fading states $\{\gamma_i(t)\}_{i \in \mathcal{B}}$. A quantity of interest is the *normalized sum rate* which is the maximum achievable equal rate per user and is obtained by taking $\mathcal{B}$ to be the whole set to yield [39]

$$R_{sum} \quad = \quad \frac{1}{K} \sum_{i=1}^{K} R_i = \mathbb{E}_{\underline{\gamma(t)}} \frac{1}{K} \log_2 \left(1 + \frac{\mathcal{P}_{av} \sum_{i=1}^{K} \gamma_i(t)}{\sigma^2} \right) \tag{16}$$

$$\underset{K \to \infty}{\to} \quad \frac{1}{K} \log_2 \left(1 + \frac{K \mathcal{P}_{av}}{\sigma^2} \right). \tag{17}$$

The asymptotic result (17) shows an interesting phenomenon, that as the number of users increases, the effect of fading is completely mitigated due to the averaging effect of multiple users. The averaging effect due to increasing users is analogous to time or frequency [40] or spatial [25] averaging in single-user channels. In [39], using (16), it was shown that a non-orthogonal multiple access scheme has a higher normalized sum rate R_{sum} than orthogonal schemes like time (frequency) division multiple access[8]. By requiring orthogonality of users, an orthogonal multiple access scheme adds additional constraints on user transmission, which leads to a performance loss compared to optimal non-orthogonal method. Non-orthogonal CDMA is an example of non-orthogonal multiple access scheme. Spread signals, like CDMA signals, occupy more bandwidth than needed and were first conceived to provide robustness against intentional jamming [41]. The capacity versus outage analysis also shows the superiority of CDMA schemes over orthogonal access methods [42].

In [43], a cellular multicell model was introduced to study the effect of multiple cells. The model extends (13) to include inter-cell interference from users in neighbouring cells. The cellular model in [43] was extended to fading channels in [39, 44]. There again, it was concluded that CDMA like wideband

---

[8]In time (frequency) division multiple access, each user transmits in its allocated time (frequency) slot such that no two user share a time (frequency) slot. Thus, the transmission of one user is orthogonal in time (frequency) to any other user.

methods achieve optimal normalized sum rates even in the presence of multicell interference, for several important practical receiver structures. Even though the spread spectrum signals occupy more bandwidth than needed for each signal, multiuser spread spectrum systems are spectrally efficient [13, 14]. Motivated by the success of the second generation CDMA standard, IS-95, currently all third generation wireless systems (CDMA2000 and W-CDMA) use some form of spread spectrum technique. In addition to information theoretic superiority, CDMA based multiple access provides other practical advantages [45]. First, CDMA signals allow finer *diversity combining* due to larger signal bandwidth, thereby providing robustness to multipath fading. In other words, combined with an interleaver, spread spectrum signals naturally exploit both frequency and time diversity. Frequency diversity is not available in bandwidth-efficient TDMA systems. Second, CDMA allows a *frequency reuse* of one in contrast to TDMA/FDMA which require a higher reuse factor. A lower reuse factor immediately implies higher system capacity; a reuse factor of one also simplifies frequency planning. Lastly, CDMA naturally exploits the *traffic activity factor*, the percentage of time during a two-way communication each channel is actually used. Most of the information theoretic analysis completely ignores the data burstiness, a property which is central to higher resource utilization in wired networking [46]; see [47, 48] for insightful reviews.

The CDMA based systems allow communication without the need for a universal clock or equivalently synchronism among different users. The need for synchronism in TDMA requires the use of time guard bands between time slots and hence wastes resources. Finally, in long-code DS-CDMA systems, like the one used in IS-95 standard[9] assigning channels to users is straightforward because each user is given a unique fixed spreading code. In TDMA, time slots are granted adaptively as users handoff from one cell to another, thereby complicating resource management and requiring additional protocol overhead. Also, long-code CDMA leads to the same average performance for all users, and thus a fair resource allocation among users.

Though the area of multiuser information theory is rich and well-studied, we maintain that many fundamental results are yet to be published. For instance, connections with queuing theory [47–49] which is the mathematical basis for networking are far from well understood, but due to the rise of internet, it is more urgent than ever to unify the areas of data networking and wireless communications. Furthermore, with the growth of wireless services beyond voice communication, and advent of newer modes of communication like *ad hoc* networking[10], current information theoretic results should be

---

[9]In long-code CDMA systems, unlike short repeating code CDMA systems, each transmitted bit is encoded with a different spreading code.

[10]In *ad hoc* networking, mobile nodes can communicate with each other without the need for any infrastructure like in cellular systems; IEEE 802.11 and Bluetooth are examples of *ad hoc* networking.

considered as the beginnings of our understanding on the subject of multiuser communications.

5 Typical Architecture of Wireless Transceiver

Most wireless systems transmit signals of finite bandwidth using a high frequency carrier[11]. This immediately leads to the wireless transceiver with three major components, an RF front end which performs the frequency conversion from passband to baseband and vice versa, digital to analog converter (D/A) and analog to digital (A/D) converter, and a baseband processing unit. In this section, we will discuss the signal processing algorithms used in the digital baseband unit. Wherever applicable, we will highlight the differences between the baseband unit at the mobile receiver and that at the base-station.

We briefly note that the hardware receiver design for CDMA systems is generally more challenging than its TDMA counterparts. The design of A/D, D/A converters, and digital baseband processors require special effort. Higher chipping rates in CDMA systems require faster sampling and hence lead to higher computational throughput requirements and increased circuit power dissipation compared to their TDMA counterparts. Fortunately, advances in low-power high-speed complementary metal oxide semiconductor (CMOS) circuits have allowed implementation of sophisticated digital signal processing algorithms, and high speed converters.

5.1 Transmitter

A simplified transmitter for DS-CDMA system is shown in Figure 5. The data obtained from the higher layers is passed through a channel encoder, spread spectrum modulator, digital to analog converter and finally through an RF unit.

5.1.1 Channel Encoding

The source data bits are first encoded using a forward error correction (FEC) code. A FEC code systematically adds redundant bits to the source bits, which are used by the receiver to correct errors in the received signal. Error correction coding is essential to achieve low bit error rates at the receiver and has a strong information theoretic foundation [23]. Following Shannon's work in 1948 [23], error control coding has seen tremendous growth in last fifty years; the readers are referred to [51–54] for recent reviews on state of the art. Several excellent texts [55–58] on channel coding theory are available, hence we will keep our discussion in this section elementary.

The choice of code primarily depends on desired performance level, the specific channel under consideration and the complexity of the resulting receiver. The desired level of performance is based on

---

[11]Carrier-less systems include impulse radio [50].

the type of services to be provided. For instance, loss tolerant services like speech can work with high packet loss probability, while data/email/fax requires a much higher error protection, thereby requiring FEC codes with different amount of error protection capabilities[12]. The complexity of decoding the received packets to correct errors is a major concern in the design of power-limited mobile handsets. Typically, stronger FEC codes are computationally harder to decode and, hence require more battery power for the baseband units; see [59] for a discussion.

The communication channel is a major factor in selection of forward error correcting codes. For example, code design is different for slow and fast fading channels. To illustrate the concept of coding, our discussion will be limited to convolutional codes which are used in both telephone line modems, and both second and third generation digital wireless cellular standards. Further, we will highlight the recent interest in space-time coding by dividing this section into two parts: single-antenna systems and multiple-antennas systems. Our discussion on single-antenna systems will give a quick introduction to convolutional codes with a review of recent coding results for slow and fast fading channels. In the multiple antenna discussion, diversity techniques will be central to our discussion, with an emphasis on spatial and time diversity for wireless systems.

Single antenna systems: The choice of convolutional codes is motivated by their simple optimal decoding structure, systematic construction of strong codes for large block lengths and lower decoding delay compared to block codes. A convolutional code is generated by passing the information sequence through a linear finite-state shift register. In general, the shift register consists of S B-bit stages and m linear algebraic function generators; see Figure 6 [4]. The input data to the encoder, assumed to be binary, is shifted into and along the shift register B bits at a time. The number of output bits for each B input bits is m bits. Consequently, the code rate is defined as $R_c = B/m$. The parameter S is called the *constraint length* of the convolutional code.

To understand the encoding procedure, consider the convolutional encoder for $S = 3$, $B = 1$ and $m = 3$ shown in Figure 7 [4]. All the shift registers are assumed to be in zero state initially. If the first input bit is a 1, the resulting output sequence of 3 bits is $\begin{bmatrix} b[1] & b[2] & b[3] \end{bmatrix} = [1\,1\,1]$. Now if the second input bit is a 0, the next three output bits are $\begin{bmatrix} b[4] & b[5] & b[6] \end{bmatrix} = [0\,0\,1]$ (else the output bits are $[1\,1\,0]$ if the input bit is 1). If the third bit is a 1, the output is $\begin{bmatrix} b[7] & b[8] & b[9] \end{bmatrix} = [1\,0\,0]$. The operation of a non-recursive (Figure 7) convolutional code is similar to that of a finite impulse response (FIR) filter with all the operations done over a finite field; in Figure 7, the finite field consists of only two elements $\{0, 1\}$ with binary addition. The convolutional code has one input and several outputs, equivalent to single-input multiple-output finite impulse response linear system. The equivalent of the impulse response of

---

[12]Some of the networking layers use checksums for error detection and perform error correction by requesting retransmission of packets.

the filter is the *generator polynomial*, which succinctly describes the relation between output and shift register states for a convolutional code. For the example in Figure 7, the generator polynomials are

$$\text{Output } 1 \rightarrow \quad \mathbf{g}_1 = [1\ 0\ 0],$$
$$\text{Output } 2 \rightarrow \quad \mathbf{g}_2 = [1\ 0\ 1],$$
$$\text{Output } 3 \rightarrow \quad \mathbf{g}_3 = [1\ 1\ 1].$$

The generator polynomials of a convolutional code characterize its performance via different metrics, notably minimum distance and distance spectrum [60]. To design any code requires an appropriate metric space, which depends on the channel under consideration. For slowly block fading channels, the Euclidean distance between the codewords is the natural metric [60], while for fast fading channels, Hamming distance is the appropriate metric [61]; see below for further discussion on diversity techniques.

Addition of redundant bits for improving the error probability leads to bandwidth expansion of the transmitted signal by an amount equal to the reciprocal of the code rate. For bandwidth constrained channels, it is desirable to achieve a coding gain with minimal bandwidth expansion. To avoid bandwidth expansion due to channel coding, the number of signal points over the corresponding uncoded system can be increased to compensate for the redundancy introduced by the code. For instance, if we intend to improve the performance of an uncoded system using BPSK modulation, a rate 1/2 code would require doubling the number of signal points to quadrature phase shift keying (QPSK) modulation. However, increasing the number of signals leads to higher probability of error for the same average power. Thus, for the resultant bandwidth efficient scheme to provide gains over the uncoded system, it must be able to overcome the penalty due to increased size of the signal set.

If the modulation (mapping of the bits to channel signals) is treated as an operation independent of channel encoding, very strong convolutional codes are required to offset the signal set expansion loss and provide significant gains over the uncoded system [4]. On the other hand, if the modulation is treated as an integral part of channel encoding, and designed in unison with code to maximize the Euclidean distance between pairs of coded signals, the loss due to signal set expansion is easily overcome. The method of *mapping by set partitioning* [62] provides an effective method for mapping the coded bits into signal points such that the minimum Euclidean distance is maximized. When convolutional codes are used in conjunction with signal set partitioning, the resulting method is known as *trellis coded modulation* (TCM). TCM is a widely used bandwidth efficient coding scheme with a rich associated literature; see [63] for a comprehensive in-depth review.

The fundamental channel coding theorem by Shannon [23] proved the existence of good codes, which can achieve arbitrarily small probability of error, as long as the transmission rate is lower than the channel capacity. The proof in [23] required creating codes which had ever increasing block sizes to

achieve channel capacity. Another key component of the proof in [23] was the choice of codebooks, they were chosen at random. Random codes with large block sizes have no apparent structure to implement a physically tractable decoder. Proven optimality of random codes coupled with the inability to find good structured codes led to a common belief that the structured deterministic codes had a lower capacity than the channel capacity, often called the "practical capacity" [64, 65]. The discovery of *turbo codes* [17] and the rediscovery of *low-density parity check* (LDPC) codes [66] appears to have banished the above "practical capacity" myth. Both turbo and LDPC codes have been shown to operate below the "practical capacity," within a tenth of a decibel of the Shannon capacity. Turbo codes have also been proposed for the third generation wireless standards. The main ingredients of a turbo code are constituent codes (block or convolutional code) and a long interleaver. The long interleaver serves two purposes: lends codewords a "random-like" structure, and leads to long codes which are easily and efficiently decoded using a (sub-optimal yet effective) iterative decoding algorithm. Several extensions of turbo codes are areas of active research, notably, bandwidth efficient turbo codes [67, 68], deterministic interleaver design [69] and space-time turbo codes [70].

We close the discussion on codes for slow fading Gaussian channels, by highlighting that none of the current codes come close to the lower bounds on the performance of codes [71]. Current codes require large block lengths to achieve small probability of decoded message errors, but relatively short block lengths suffice to achieve the same level of performance for "good" codes [71]. Thus, the field of code design, though more than fifty years old, has still significant room to develop.

Multiple antenna systems : The random time-variations in the received signal provide diversity, which can be exploited for improved error performance. Typical forms of diversity include time, frequency and spatial diversity. In Section 4.1, it was noted that diversity is important to improve the outage performance or achievable rates in fading channels. Although only spatial diversity using multiple transmit and receive antennas was studied in Section 4.1, similar benefits are also obtained by using time or frequency diversity or a combination of them. In time and frequency diversity, channel variations in time and across frequency are used to increase reliability of the received signal. In spatial diversity, multiple transmit and/or receive antennas exploit the random spatial time-variations.

The codes designed for Gaussian channels can be used for slowly fading channels if an accurate channel estimator is available and all symbols of a codeword undergo the same channel fading. In the presence of medium to fast fading, where the coherence interval is shorter than a codeword, Hamming distance between the codewords should be maximized [61]. If channel variations are slower than a codeword, an interleaver is commonly used to induce time diversity. For interleaver based schemes to be effective, the interleaver depths should be larger than the coherence interval; this implies that it is useful for fast fading channels or for communications where large delay can be tolerated. For

low-delay application, the interleaver-induced time diversity is not possible. In addition, if the channel is flat-fading (true for narrowband communications), then frequency diversity cannot be used either. Irrespective of the availability of time and frequency diversity, the spatial diversity via multiple antennas is a promising method to achieve higher data rates.

Receiver diversity using multiple receive antennas is a well-understood concept [6] and often used in practice [72]. In contrast, using multiple antennas at the transmitter has gained attention only recently due to discovery of space-time codes [5, 31], motivated by encouraging capacity results [25, 73]. Space-time coding exploits multiple independent channels between different transmit-receive antenna pairs in addition to time diversity (possibly interleaver induced). The work in [5] extended well-founded coding principles to spatial diversity channels, thereby simultaneously achieving coding gain and the highest possible spatial diversity. The space-time codes proposed in [5] have become a performance benchmark for all subsequent research in space-time coding [74–80]. The concept of transmitter diversity can be appreciated using the following elegant *Alamouti scheme* [81] for two transmit antennas.

In a given symbol period, two symbols are simultaneously transmitted from the two antennas. Denote the signal transmitted from antenna 1 as s_1 and from antenna 2 as s_2 (see Figure 8). During the next symbol period, signal $-s_2^*$ is transmitted from antenna 1, and s_1^* is transmitted from antenna 2. Note that the encoding of symbols is done in both space and time. As is evident from Figure 8, the received signal in any symbol interval is a linear combination of the signals transmitted from the two antennas. Thus, the space-time channel is an interference channel. An analogous scenario exploiting frequency diversity would use non-orthogonal carrier frequency to send two symbols in each symbol period. The Alamouti scheme sends orthogonal signals over two time instants from the two antennas, *i.e.*, vector $[s_1 \quad -s_2^*]$ transmitted from antenna 1 over two time symbols is orthogonal to the vector $[s_2 \ s_1^*]$ transmitted from antenna 2. If the channel stays constant over two consecutive symbol periods, then the orthogonality is maintained at the receiver. Since each symbol s_1 and s_2 is transmitted from both the antennas, they travel to the receiver from two different channels, which provides the desired diversity order of two. The orthogonality of the time signals helps resolution of the two symbols at the receiver without affecting the diversity order.

The Alamouti scheme can be extended to more than two transmit antennas using the theory of orthogonal designs [74]. The Alamouti scheme is a rate 1 code and thus requires no bandwidth expansion. But it provides a diversity order of two, which is twice that of any rate 1 single-antenna system. The Alamouti scheme has a very simple optimal receiver structure, thereby making it a prime candidate for practical implementations. In addition to its simplicity, the Alamouti scheme-based systems do not lose in their asymptotic performance. In [79], it was shown that orthogonal transmit diversity schemes are capacity achieving, and thus provided a motivation for the concatenated space-time coding methods

in [79, 80]. The concatenated space-time codes decouple the spatial and temporal diversity to simplify the space-time code design.

All third generation systems have adopted some form of transmit and receive diversity. Multiple antennas at the base-station are relatively easier to implement in comparison to multiple antennas at the mobile handset, due to size limitation. Two cross polarized antennas have been proposed and tested for mobile handsets [82].

5.1.2   Spreading and Modulation

The binary output of the error control encoder is mapped to either ± 1 to obtain the sequence $b_i[k]$, which is multiplied by a spreading sequence, $c_i[n] \in \{-1, 1\}$, of length N; the spreading operation is shown in Figure 9. After spreading the signal, the signal is passed through a digital pulse shaping filter, $\phi[n]$, which is typically a square root raised cosine filter [4]. The pulse shaping filter is chosen to limit the bandwidth of the transmitted signal to the available spectrum, while minimizing the intersymbol interference (ISI) caused by the filter. The digital signal for user i after pulse shaping can be written as

$$s_i[n] = \sum_{k=1}^{G} b_i[k]\psi_i[n - kNL] \tag{18}$$

where L is the number of samples per chip and $\psi_i[n] = \phi[n] \star c_i[n]$ where $\star$ represents linear convolution, and G is the number of bits in the packet.

After converting the digital signal to analog using a D/A converter, the RF upconverter shifts the baseband analog signal to the carrier frequency f_c. The upconverted signal is amplified by a power amplifier and transmitted via an antenna. The transmitted passband signal assumes the following form,

$$\begin{aligned} x_i(t) &= \sqrt{\mathcal{P}_i} e^{-\jmath \omega_c t} \sum_{k=1}^{G} b_i[k]\psi_i(t - kT_s), \\ &= e^{-\jmath \omega_c t} s_i(t) \end{aligned} \tag{19}$$

where T_s is the symbol period and $\mathcal{P}_i$ is the transmitted power. The bits $b_i[k]$ are the output of a suitable channel encoder discussed in Section 5.1.1. Since CDMA signals at the base-station typically have large peak to average power ratios, the operating point of the power amplifier is kept low to avoid amplifier nonlinearities. The amplifier nonlinearities are avoided for several important reasons: (a) RF amplifier efficiency is lower in nonlinear region which increases the power loss and hence total power consumed by the transmitter, (b) the nonlinearity introduces higher spectral components, which can cause increased interference in the neighbouring frequency bands, and (c) algorithm design for resulting nonlinear systems becomes intractable.

As discussed in Section 4.1, multiple antennas at the transmitter and receiver can lead to large gains in fading wireless channels [21, 25, 37]. If multiple transmit antennas are used, the vector transmitted passband signal is given by

$$\mathbf{x}_i(t) = \sqrt{\frac{\mathcal{P}_i}{M}} e^{-\jmath \omega_c t} \sum_{k=1}^{G} \mathbf{b}_i[k] \psi_i(t - kT), \tag{20}$$

where M is the number of transmit antennas. The $M \times 1$ vectors, $\mathbf{x}_i(t)$ and $\mathbf{b}_i[k]$, represent the transmitted vector signal and space-time coded signal, respectively. In (20), we have assumed that the transmitter has no knowledge of the channel and hence uses same average power on each of the transmitter. If the transmitter knows the channel, then the power across different antennas can be adapted to achieve an improved performance [25, 83].

5.2   Base-station Receiver

In cellular systems, the time and spectral resources are divided into different logical *channels*. The generic logical channels are broadcast, control, random access, paging, shared and dedicated channels [84, 85]. All logical channels are physically similar and the distinction is solely made based on the purpose served by each channel. In the sequel, we will restrict our attention to the dedicated and shared channels, since they carry most of the user data and hence impose the biggest computational bottleneck. Implementation details of other channels can be found in [84, 85].

As noted in Section 2, the unknown time-varying multipath is one of the biggest challenges in the design of wireless systems. Optimal transmission schemes which do not require the knowledge of the wireless channel at the receiver can be designed using information theoretic tools [86, and the references therein], but are seldom employed. The primary reason for not using optimal strategies is their high computational complexity, and large latency of the resulting communication method. Hence, suboptimal and computationally efficient solutions are generally employed. The receiver estimates the unknown channel, and then uses the channel estimate to decode the data using a channel decoder.

A simplified illustration of the baseband receiver is shown in Figure 10. The key components of the receiver are multiuser channel estimation, multiuser detection and single-user channel decoding. Most systems also provide feedback from the receiver for power control and automatic repeat request (ARQ) to improve system reliability. The choice of algorithms used in each of the blocks is determined by their computational complexity, desired performance level and the available side information. Mobile units are power and complexity constrained, and have little or no knowledge of the multiple access interference. On the other hand, the base-stations are equipped with higher processing power and detailed information about all in-cell users, thereby allowing more sophisticated processing at the base-stations.

Our discussion will focus on base-station algorithms in the following section, with only bibliographic references to relevant counterparts for the mobile handset.

5.2.1 Received Signal

For each of the active users in a cell, the received signal at the base-station consists of several unknown time-varying parameters. These parameters include propagation delay, amplitude, delay and number of paths, and residual carrier offset. The time-variation in propagation delay is caused as users move closer or away from the base-station. The mobility of the users or the surrounding environment also causes time-variation in the multipath environment. Finally, drift in the local oscillator frequencies of the transmitter and receiver leads to a residual carrier offset at the baseband.

Using the model (4) for the multipath channel impulse response and assuming that the channel coefficients for the ith user $h_{p,i}$ are constant over the observation interval, the received signal for a transmitted signal $x_i(t)$ without additive white noise is given by

$$
\begin{aligned}
z_i(t) &= \sum_{p=1}^{P} h_{p,i} x_i(t - \tau_i - p/W), \\
&= \sqrt{\mathcal{P}_i} e^{-\jmath \omega_c t} \sum_{p=1}^{P} \underbrace{h_{p,i} e^{\jmath \omega_c (\tau_i + p/W)}}_{a_{p,i}} \sum_{k=1}^{G} b_i[k] \psi_i(t - kT - \tau_i - p/W), \\
&= \sqrt{\mathcal{P}_i} e^{-\jmath \omega_c t} \sum_{k=1}^{G} \sum_{p=1}^{P} b_i[k] a_{p,i} \psi_i(t - kT - \tau_i - p/W).
\end{aligned}
\tag{21}
$$

where τ_i is the propagation delay of the received signal. If the number of paths $P = 1$, then it is a flat fading channel else a frequency selective channel. The received signal is amplified and downconverted to baseband. In practice, there is a small difference in the frequencies of the local oscillators at the transmitter and the receiver. The received baseband signal after downconversion (without additive noise) is given by

$$
z_i(t) = \sqrt{\mathcal{P}_i} e^{-\jmath \Delta \omega_i t} \sum_{k=1}^{G} \sum_{p=1}^{P} b_i[k] a_{p,i} \psi_i(t - kT - \tau_i - p/W),
\tag{22}
$$

where $\Delta \omega_i$ represents the residual carrier frequency offset. Assuming that the carrier offset $\Delta \omega_i$ is negligible or is corrected using a multiuser equivalent of digital phase lock loop [4, 87], the sampled baseband (without noise) with L samples per chip can thus be written as

$$
z_i[n] = \sum_{k=1}^{G} \sum_{p=1}^{P} b_i[k] a_{p,i} \psi_i[n - kNL - \tau_i - p].
\tag{23}
$$

In general, the receiver components introduce thermal noise, which is generally modeled as additive noise. For K simultaneously active users, the received baseband signal in the presence of thermal noise at the base-station is

$$z[n] \;\; = \;\; \sum_{i=1}^{K} z_i[n] + \nu[n]. \tag{24}$$

The additive component $\nu[n]$ in (24) is generally modeled as white Gaussian noise. The received signal model in Equations (13) and (24) are similar, both consider a sum of all user signals in additive noise. The main difference is the assumption on the fading statistics, a flat fading model is assumed in (13) compared to a multipath model in (24).

In the sequel, we will focus our discussion on estimation of the unknown channel coefficients and subsequent detection of the data bits, $b_i[k]$ for all users $i = 1, \dots, K$. The development of multiuser channel estimation and data detection is greatly simplified by using linear algebraic methods. We will write the received signal (24) using matrix-vector notation in two different forms. The first form will be used in multiuser channel estimation methods, and the second in multiuser detection.

Channel as unknown: For simplicity, we will assume that all τ_i are multiple of sampling instants, *i.e.*, $\tau_i = l_i$; for the general case, the reader is referred to [88]. Let $u_i[n] = \sum_{k=1}^{G} b_i[k] \psi(n - kNL)$. Then the received signal $z_i[n]$ can be rewritten in matrix-vector notation [89] as

$$\mathbf{z}_i \;\; = \;\; \begin{bmatrix} u_i[1] & 0 & 0 & \cdots & & 0 \\ u_i[2] & u_i[1] & 0 & & & 0 \\ u_i[3] & u_i[2] & u_i[1] & & & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \vdots & u_i[GLN + l_\phi] \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_{1,i} \\ \vdots \\ a_{P,i} \end{bmatrix},$$

$$= \;\; \mathbf{U}_i \mathbf{a}_i, \tag{25}$$

where there are l_i leading zeros in the channel vector $\mathbf{a}_i$ to account for the propagation delay, and l_ϕ is the length of the pulse ϕ (measured in number of samples). The total received signal can thus be written as

$$\mathbf{z} \;\; = \;\; \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_K \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_K \end{bmatrix} + \boldsymbol{\nu}$$

$$= \;\; \mathbf{U}\mathbf{a} + \boldsymbol{\nu}, \tag{26}$$

22

where we recall that N is the spreading gain, L is the number of samples per chip and G is the number of bits in the packet. The above signal model will be used to derive channel estimation algorithms in Section 5.2.2.

Data as unknown: Define $q_i[n] = \sum_{p=1}^{P} a_{p,i}\psi[n-kNP-l_i-p]$; $q_i[n]$ can be understood as the *effective* spreading waveform for the ith user. The waveform $q_i[n]$ is generally longer than one symbol period and hence causes interference between the consecutive symbols. To highlight the presence of intersymbol interference (ISI), we will write the received signal $z_i[n]$ for every symbol duration. For simplicity, we will assume that the length of $q_i[n]$, l_q, is less than two symbol durations, *i.e.*, $l_q < 2NL$. Then the received signal $z_i[n]$ can be written as

$$
\mathbf{z}_i[k] = \begin{bmatrix} 0 & q_i[1] \\ 0 & q_i[2] \\ \vdots & \vdots \\ q_i[NL+1] & q_i[2NL-l_q+1] \\ \vdots & \vdots \\ q_i[l_q] & q_i[NL] \end{bmatrix} \begin{bmatrix} b_i[k-1] \\ b_i[k] \end{bmatrix}
$$
$$
= \mathbf{Q}_i \mathbf{b}_i \tag{27}
$$

The total received signal can be written as

$$
\mathbf{z}[k] = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 & \cdots & \mathbf{Q}_K \end{bmatrix} \begin{bmatrix} \mathbf{b}_1[k] \\ \mathbf{b}_2[k] \\ \vdots \\ \mathbf{b}_K[k] \end{bmatrix} + \boldsymbol{\nu} \tag{28}
$$
$$
= \mathbf{Q}\mathbf{b}[k] + \boldsymbol{\nu} \tag{29}
$$

The above received signal model (29) clearly demonstrates the challenges in multiuser detection. Not only does the receiver have to cancel the multiple access interference, but also the ISI for each user introduced by the multipath channel. The ISI acts to increase the effective multiple access interference experienced by each bit. The multiuser detection methods aim to jointly make all bits decisions $\mathbf{b}[k]$.

In the following section, we will discuss channel estimation, multiuser detection and channel decoding algorithms for DS-CDMA systems.

5.2.2 Multiuser Channel Estimation

Most channel estimation can be divided into two broad classes, training based and blind methods. In each of the classes, a further subdivision[13] is made based on assumptions made regarding the multiple access interference: single-user channel estimation in the presence of multiple access interference or jointly estimating channels for all the users.

Most wireless systems add known symbols periodically to the data packets. The known data symbols are known as *training symbols* and facilitate coarse synchronization, channel estimation and carrier offset recovery. Training based methods simplify estimation of unknown baseband parameters at the cost of throughput loss; symbols used for training could potentially be used to send more information bits. The amount of training depends on the number of simultaneous users, number of transmit antennas [28] and desired reliability of channel estimates. Given the training symbols and assuming perfect carrier offset recovery, multiuser channel estimation can be cast as a linear estimation problem [91], and admits a closed form solution. The work in [91] also discusses extensions to multiple antennas.

A class of blind channel estimation procedures, collectively known as constant modulus algorithms (CMA) were first proposed in [92, 93] using the constant amplitude property of some of the communication signals like BPSK. The CMA algorithms use a nonlinear (nonconvex) cost function to find the channel estimate, and hence can converge to poor estimates. An alternate procedure of blind estimation was proposed in [94, 95], which used the cyclostationarity of the communication signals. Motivated by the method in [94], a single-user blind channel identification method, using only second-order statistics, was proposed in [96]. The blind channel equalization exploits only the second (or higher) order statistics without requiring periodic training symbols, with an assumption that the data symbols are independent and identically distributed. The assumption of i.i.d. data is rarely correct due to channel coding used in almost all systems. Hence, the results based on blind channel estimation should be interpreted with caution. Nonetheless, there is value in exploring blind channel identification methods. Blind estimation can improve the estimates based on training or completely avoid the use of training symbols; the reader is referred to [97, 98] for results on single-user systems.

Single-user channel estimation in the presence of unknown multiple access interference was addressed in [99]. An approximate maximum-likelihood channel estimation for multiple users entering a system was presented in [100]; the estimate-maximize algorithm [101] and the alternating projection algorithm [102] in conjunction with the Gaussian approximation for the multiuser interference were used to obtain a computationally tractable algorithm. Blind multiuser channel estimation has also been addressed in

---

[13]Another possible subdivision can be based on linear and non-linear algorithms. An example of feedback based non-linear algorithm is the decision feedback based equalization [90].

several papers [103, 104], with an assumption of coarse synchronization.

Most of the current work, with a few exceptions [105–108], assume square pulse shaping waveforms leading to closed form optimistic results; see [107] for a detailed discussion. Furthermore, very little attention has been paid to carrier offset recovery in a multiuser system, except for the results in [87]. In this section, we will only discuss channel estimation at the base-station, assuming coarse synchronization and perfect downconversion. For handset channel estimation algorithms, the reader is referred to [107, 109]. Additionally, we restrict our attention to only training based methods; blind techniques are rarely used in wireless systems[14]. The channel model assuming T training symbols for each user can be written as

$$\mathbf{z} = \mathbf{U}\mathbf{a} + \boldsymbol{\nu}, \tag{30}$$

where the size of the vectors $\mathbf{z}$ and $\boldsymbol{\nu}$, and matrix $\mathbf{U}$ is appropriately redefined for an observation length of T symbols, using the definition in (26). The matrix $\mathbf{U}$ depends on the spreading codes, $\phi_i[n]$ and the training symbols $b_i[k]$, all of which are assumed known for all users. Thus, the matrix $\mathbf{U}$ is completely known. The maximum likelihood estimate of the channel coefficients, $\mathbf{a}$, is given by the pseudo-inverse [4, 91],

$$\widehat{\mathbf{a}} = \left(\mathbf{U}^H \mathbf{U}\right)^{-1} \mathbf{U}^H \mathbf{z}. \tag{31}$$

The above solution retains several desirable statistical properties of the maximum likelihood estimates for linear Gaussian problems [110], namely, consistency, unbiasedness and efficiency. Note that there are several leading zeros in $\mathbf{a}$. The variance of the maximum likelihood estimator $\widehat{\mathbf{a}}$ can be reduced by detecting the unknown number of leading (and possibly trailing) zeros in $\mathbf{a}$, which reduces the number of estimated parameters. The above channel estimation procedure can also easily be extended to long code DS-CDMA systems [111]. In practice the additive noise $\boldsymbol{\nu}$ is better modeled as colored Gaussian noise with unknown covariance due to out-of-cell multiuser interference. The maximum likelihood estimate of $\mathbf{a}$ requires estimation of the unknown covariance, thereby leading to more accurate results compared to (31) at the expense of increased computation [89].

Having estimated the channel for all the users, the channel estimates are then used to detect the rest of the information bearing bits in the packet. For bit detection, the received signal representation in (29) is more appropriate, where the matrix $\mathbf{Q}$ is formed using the channel estimates $\widehat{\mathbf{a}}$ and the user signature waveforms $\psi_i[n]$.

[14]A notable exception is high definition television (HDTV) transmission, where no resources are wasted in training symbols, and slow channel time-variation permit the use of blind estimation techniques.

5.2.3 Multiuser Detection

Due to channel induced imperfections and time-varying asynchronism between the users, it is practically impossible to maintain orthogonality between the user signals'. *Multiple access interference* (MAI) is caused by the simultaneous transmission of multiple users, and is the major factor which limits the capacity and performance of DS-CDMA systems. In the second generation CDMA standards, the multiple access interference is treated as part of the background noise and single-user optimal detection strategy is used. The single-user receiver is prone to the *near-far* problem, where a high power user can completely drown the signal of a weak user. To avoid the near-far problem, CDMA based IS-95 standard uses tight power control to ensure that all users have equal received power. Even with the equal received power, the output of the single-user detector is contaminated with MAI and is suppressed by using very strong forward error correction codes.

The MAI is much more structured than white noise and this structure was exploited in [112] to derive the optimal detector which minimizes the probability of error. The optimal detector alleviates the near-far problem which plagues the single-user receiver. The optimal detector, thus, does not require fast power control to achieve a desired level of performance, thereby reducing the system overhead greatly. Further, as the number of users increases, the optimal receiver achieves significant gains over single-user receiver, even with perfect power control. Unfortunately, the optimal receiver is computationally too complex to be implemented for large systems [113]. The computational intractability of multiuser detection has spurred a rich literature on developing low-complexity suboptimal multiuser detectors.

Most of the proposed suboptimal detectors can be classified in one of two categories: linear multiuser detectors and subtractive interference cancellation detectors. Linear multiuser receivers linearly map the soft outputs of single-user receivers to an alternate set of statistics, which can possibly be used for an improved detection. In subtractive interference cancellation, estimates for different user signals' are generated and then removed from the original signal.

To gain insight into different methods for multiuser detection, we will limit the discussion in this section to a simple case of no multipath and no carrier frequency errors. We further assume that the pulse shaping introduces no ISI and all users are synchronous, thereby leading to simplification of (29) as

$$\mathbf{z}[k] = \mathbf{Q}\mathbf{b}[k] + \boldsymbol{\nu}[k], \tag{32}$$

where $\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_K \end{bmatrix}$, $\boldsymbol{q}_i = \begin{bmatrix} q_i[1] & q_i[2] & \cdots & q_i[NP] \end{bmatrix}^T$, and $\mathbf{b}[k] = \begin{bmatrix} b_1[k] & b_2[k] & \cdots & b_K[k] \end{bmatrix}^T$. Note that the above simplification only eliminates ISI not the multiple access interference, which is the primary emphasis of the multiuser detection. We quickly note that all the subsequently discussed multiuser detection methods can be extended to the case of asynchronous and ISI channels. The code

26

matched filter outputs, $\mathbf{y}[k] = \mathbf{Q}^H \mathbf{z}[k]$ can be written as

$$\mathbf{y}[k] = \mathbf{R}\mathbf{b}[k] + \boldsymbol{\nu}[k]. \tag{33}$$

The $K \times K$ matrix $\mathbf{R} = \mathbf{Q}^H \mathbf{Q}$ is the correlation matrix, whose entries are the values proportional to the correlations between all pair of spreading codes. The matrix $\mathbf{R}$ can be split into two parts, $\mathbf{R} = \mathbf{D} + \mathbf{O}$, where $\mathbf{D}$ is a diagonal matrix with $\mathbf{D}_{ii} = \mathcal{P}_i$. Thus (33) can be written as follows

$$\mathbf{y}[k] = \mathbf{D}\mathbf{b}[k] + \mathbf{O}\mathbf{b}[k] + \boldsymbol{\nu}[k]. \tag{34}$$

The matrix $\mathbf{O}$ contains the off-diagonal elements of $\mathbf{R}$, with entries proportional to the cross-correlations between different user codes. The first term in (34), $\mathbf{b}[k]$, is simply the decoupled data of each user and the second term, $\mathbf{O}\mathbf{b}[k]$, represents the MAI.

Matched-filter detector : Also known as single-user optimal receiver, the matched-filter receiver treats the MAI+$\boldsymbol{\nu}[k]$ as white Gaussian noise, and the bit decisions are made by using the matched filter outputs, $\mathbf{y}[k]$. The hard bit decisions are made as

$$\widehat{\mathbf{b}}_{MF}[k] = \mathrm{sign}\left(\mathbf{y}[k]\right), \tag{35}$$

where $\mathrm{sign}(\cdot)$ is a nonlinear decision device and outputs the sign of the input. The matched-filter receiver is extremely simple to implement and requires no knowledge of MAI for its implementation. However the matched-filter receiver suffers from the near-far problem, where a non-orthogonal strong user can completely overwhelm a weaker user; in fading environments, power disparities are commonly encountered and perfect power control is generally impossible.

Maximum a posteriori probability (MAP) detector : As the name suggests, the maximum likelihood detector chooses the most probable sequence of bits to maximize the joint *a posteriori* probability, the probability that particular bits were transmitted having received the current signal, *i.e.* Prob($\mathbf{b}[k]|\mathbf{r}(t)$, for all t). The MAP detector minimizes the probability of error [112]. Under the assumption that all bits are equally likely, the MAP detector is equivalent to the maximum-likelihood detector, which finds the bits $\mathbf{b}[k]$ that maximize the probability Prob$\left(\mathbf{r}(t)|\mathbf{b}[k]\right)$.

For the case of K synchronous users in (32), there are 2^K possible transmitted bit combinations in each received symbol duration. Thus, the computation of the maximum-likelihood bit estimates requires number of operations proportional to 2^K. For large number of users, the number of operations to obtain maximum-likelihood estimates become prohibitive for real-time implementation.

In the general case of asynchronous users, if a block of $M \leq G$ bits per user is used to perform the detection, there are 2^{MK} possible bit decisions, $\{\mathbf{b}[k]\}_{k=1}^{M}$. An exhaustive search over all possible bit

combinations is clearly impractical, even for moderate values of M and K. However, the maximum-likelihood detector can be implemented using the Viterbi algorithm [114]; the Viterbi implementation (see Section 5.2.4 for more details on Viterbi decoding) is similar to maximum likelihood sequence detection for ISI channels [4]. The resulting Viterbi algorithm has a complexity which is linear in block length M and exponential in the number of users, of the order of $M2^K$.

The maximum-likelihood detector requires complete knowledge of all user parameters which not only include the spreading signatures of all users but also their channel parameters. The channel parameters are unknown a priori, and have to be estimated. Despite the huge performance and capacity gains of the maximum-likelihood detector, it remains impractical for real-time systems. The computational intractability of the ML detector has led to several detectors which are amenable to real-time implementation.

Linear detectors : Linear detectors map the matched filter outputs, $\mathbf{y}[k]$, in Equation (33) into another set of statistics to reduce the MAI experienced by each user. Two of the most popularly studied matched filter receivers are the decorrelating detector and minimum mean-squared error (MMSE) detector.

The **decorrelating detector** was proposed by [115, 116] and was analyzed in [117, 118]. The decorrelating detector uses the inverse of the correlation matrix, $\mathbf{R}^{-1}$, to decouple the data of different users. The output of the decorrelating detector before hard decision is given by

$$
\begin{align}
\widehat{\mathbf{b}}_{dec}[k] &= \mathbf{R}^{-1}\mathbf{y}[k], \tag{36}\\
&= \mathbf{b}[k] + \mathbf{R}^{-1}\boldsymbol{\nu}[k], \tag{37}\\
&= \mathbf{b}[k] + \boldsymbol{\nu}_{dec}[k]. \tag{38}
\end{align}
$$

The decorrelating detector completely suppresses the MAI at the expense of reduced signal power[15]. For non-multipath channels and unknown user amplitudes, the decorrelating detector yields optimal maximum likelihood estimates of the bits and the received amplitudes. The decorrelating detector leads to substantial performance improvements over the single-user detector [118] if the background noise is low compared to the MAI. In addition to the noise enhancement problem, the computational complexity of the decorrelating detector can be prohibitive to implement in real-time; however, dedicated application specific integrated circuits (ASIC) can ameliorate the real-time implementation issues. The computational complexity of the decorrelating detector prohibits its use for long-code CDMA systems, since it requires recomputation of $\mathbf{R}^{-1}$ for every bit.

The **MMSE detector** [119] accounts for the background noise and the differences in user powers to suppress the MAI. The detector is designed to minimize the mean-squared error between the actual

---

[15]The decorrelating detector is very similar to the zero-forcing equalizer [4] which is used to completely suppress ISI.

data, $\mathbf{b}$ and the soft estimate of data, $\widehat{\mathbf{b}}_{mmse}$. The MMSE detector hard limits the following transform of the received signal,

$$\widehat{\mathbf{b}}_{mmse} = \left(\mathbf{R} + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{y}[k]. \tag{39}$$

The MMSE detector[16] balances between the suppression of MAI and suppression of background noise. The higher the background noise level, the lesser is the emphasis on suppressing MAI and vice versa. The MMSE detector has been shown to have a better probability of error than the decorrelating detector [12]. It is clear that as the background noise goes to zero, the MMSE detector converges to the decorrelating detector. On the other hand, as the background noise becomes more dominant compared to MAI, the MMSE detector converges to a single-user detector. Unlike the decorrelator and single-user receiver, the MMSE detector requires an estimate of user amplitudes. Further, the complexity of the MMSE detector is similar to that of the decorrelator.

A blind extension of the MMSE detector, which does not require the knowledge of other user codes and parameters, was presented in [120]. The blind MMSE is similar to the commonly used beamformer in antenna array processing [121]. The probability of error performance of the MMSE detector was studied in [122]. The MMSE estimator was extended to multiple data rate systems, used in the third generation standards, in [123, 124].

Subtractive interference cancellation : The basic idea in subtractive interference cancellation is to separately estimate the MAI contribution of each user and use the estimates to cancel a part or all the MAI seen by each user. Such a detector structure can be implemented in multiple stages, where each additional stage is expected to improve the accuracy of the decisions. The bit decisions used to estimate MAI can be hard (after the sign($\cdot$) operation) or soft (before the sign($\cdot$) operation). The nonlinear hard-decision approach uses the bit decisions and the amplitude estimates of each user to estimate the MAI. In the absence of reliable estimates, the hard-decision detectors may perform poorly as compared to their soft-decision counterparts [125, 126].

The **successive interference cancellation** (SIC) detector cancels interference serially. At each stage of the detector, bit decisions are used to regenerate a user signal and cancel out the signal of one additional user from the received signal. After each cancellation, the rest of the users see a reduced interference. The SIC detector is initialized by ranking all the users by their received power. For the following discussion, assume that the subscripts represent the user rank based on their received powers. The received signal corresponding to user 1 is denoted by $z_1[n]$ (cf (32)), and its bit estimate is denoted by $b_1[n]$. The SIC detector includes the following steps:

1. Detect the strongest user bit, $b_1[k]$, using the matched-filter receiver.

---

[16]The MMSE detector is similar to the MMSE linear equalizer used to suppress ISI [4].

2. Generate an estimate, $\widehat{z}_1[n]$, of the user signal based on the bit estimate, $b_1[k]$, and the channel estimate.

3. Subtract $\widehat{z}_1[n]$ from the received signal $z[n]$, yielding a signal with potentially lower multiple access interference.

4. Repeat Steps (1)-(3) for each of the successive users using the "cleaned" version of the signal from the previous stage.

Instead of using the hard bit estimates, $\widehat{b}_i[k]$, soft bit estimates (without the sign operator) can also be used in Step 3. If reliable channel estimates are available, hard decision SIC generally outperforms the soft-decision SIC; the situation may reverse if the channel estimates have poor accuracy [125, 126]. The reasons for canceling the signals in descending order of received signal strength are as follows. First, acquisition of the strongest user is the easiest and has the highest probability of correct detection. Second, the removal of the strongest user greatly facilitates detection of the weaker users. The strongest user sees little or no interference suppression but the weakest user can potentially experience a huge reduction in MAI. Third, SIC is information theoretically optimal, *i.e.*, optimal performance can be achieved using SIC [127].

The SIC detector can improve the performance of the matched-filter receiver with minimal amount of additional hardware, but SIC presents some implementation challenges. First, each stage introduces an additional bit delay, which implies that there is a trade-off between the maximum number of users that are canceled and the maximum tolerable delay [128]. Second, time-variation in the received powers caused by time-varying fading requires frequent reordering of the signals [128]. Again, a trade-off between the precision of the power ordering and the acceptable processing complexity has to be made.

Note that the performance of SIC is dependent on the performance of the single-user matched filter for the strongest users. If the bit estimates of the strongest users are not reliable, then the interference due to the stronger users is quadrupled in power (twice the original amplitude implies four times the original power). Thus, the errors in initial estimates can lead to large interference power for the weaker users, thereby amplifying the near-far effect. So, for SIC to yield improvement over matched filter, a certain minimum performance level of the matched-filter is required.

In contrast to the SIC detector, the **parallel interference cancellation** (PIC) detector [129] estimates and cancels MAI for all the users in parallel. The PIC detector is also implemented in multiple stages:

1. The first stage of the PIC uses a matched-filter receiver to generate bit estimates for all the users, $\widehat{\mathbf{b}}_{MF}[k]$.

2. The signal for the matched filter for user i in the next stage is generated as follows. Using the effective spreading codes and the bit estimates of all but the ith user, the MAI for user i is generated and subtracted from the received signal, $r[n]$.

3. The signal with canceled MAI is then passed to the next stage which hopefully yields better bit estimates.

4. Steps (1)-(3) can be repeated for multiple stages. Each stage uses the data from the previous stage and produces new bit estimates as its output.

The output of $(m+1)^{st}$ stage of the PIC detector can be concisely represented as

$$
\begin{aligned}
\widehat{\mathbf{b}}^{(m+1)}[k] &= \operatorname{sign}\left(\mathbf{y}[k] - \mathbf{O}\widehat{\mathbf{b}}^{(m)}[k]\right) \\
&= \operatorname{sign}\left(\mathbf{Db}[k] + \mathbf{O}(\mathbf{b}[k] - \widehat{\mathbf{b}}^{(m)}[k]) + \boldsymbol{\nu}[k]\right)
\end{aligned}
\tag{40}
$$

The term $\mathbf{O}\widehat{\mathbf{b}}^{(m)}[k]$ is the estimate of MAI after the mth stage. Since soft-decision SIC exploits power variation by canceling in the order of signal strength, it is superior in a non power-controlled system. On the other hand, soft-decision PIC has a better performance in a power-controlled environment. Performance evaluation of soft-decision PIC can be found in [130, 131]. A comparison of the soft-decision PIC and SIC detectors can be found in [130].

The susceptibility of the PIC to the initial bit estimates was discussed in [129]. An improved PIC scheme, which uses a decorrelator in the first stage, was proposed in [132]. The decorrelator based PIC detector provides significant performance gains over the original PIC scheme. Further improvements to PIC detector's performance can be obtained by linearly combining the outputs of different stages of the detector [133].

For long-code systems, multistage detection is best suited for its good performance-complexity trade-off. Multistage detection requires only matrix multiplications in each processing window while other multiuser detectors like the decorrelator and MMSE detector require matrix inversions during each processing window due to the time-varying nature of the spreading codes.

5.2.4 Channel Decoding

Following the multiuser detection, the detected symbols are decoded using a channel decoder to produce an estimate of the transmitted information bits. In this section we will review decoders for FEC coding when the sender uses either one or more than one transmit antenna. For single antenna systems, we will consider Viterbi decoding [134] of convolutional codes and review its lower complexity approximations. For multiple antennas, the ML decoder for the Alamouti scheme is presented along with a discussion on complexity of decoding space-time trellis codes.

Single transmit antenna: The detected bits after the multiuser detection can be treated to be free of multiple access interference and hence a single-user channel decoder can be used. Viterbi decoding for convolutional codes is an application of the dynamic programming principle, and allows efficient hard or soft decision decoding of convolutional codes. Furthermore, Viterbi decoding is amenable to VLSI implementation.

To understand the decoding of a convolutional code, an alternate representation for the encoding process, known as a trellis diagram, is better suited. A convolutional code is a finite state machine, whose next state and output is completely determined by its current state and input. The states of a convolutional code can be depicted using a *trellis diagram*. The trellis diagram for the example code in Figure 7 is given in Figure 11. A close examination of the trellis diagram in Figure 11 reveals that the diagram repeats itself after three stages, which is equal to the constraint length of the code, $S = 3$. In fact, the three outputs are completely determined by the first two states of the system and the input, which explains the four possible states ('00', '01', '10' and '11') in the trellis and the two possible transitions from the current state to the next state based on the input ('0' or '1'). The solid transitions are due to input '0' and the dashed line shows transition due to input '1.' The numbers along the transition describe the output of the decoder due to that transition.

Assume that κ encoded bits were sent using a rate R convolutional code; note that κ can be less than the packet length G if a training sequence is sent in the packet for channel estimation. The maximum a posteriori decoder chooses the information bit sequence which maximizes the posterior probability of the transmitted information symbols given the received noise corrupted signal. To compute the exact estimate of the transmitted information symbols, a total of $2^{\kappa R}$ bits should be considered. In [134], it was shown that due to encoding structure of the convolutional codes, the optimal decoder has a complexity which is linear in the codeword length κ. In the Viterbi algorithm, a metric is associated to each branch of the code trellis. The metric associated with a branch at a particular stage or level i, is the probability of receiving r_i, when the output corresponding to that branch is transmitted. A *path* is defined as a sequence of branches at consecutive levels so that the terminal node of a branch ends in the source node of the next branch. The metric associated with a path is the sum of the metrics associated with the branches in the path. And the metric associated to a node is the minimum metric associated with any path starting from the start node to that node. With the above associations, the MAP codeword corresponds to the path that has the lowest metric from the start node to the final node. If the decoder starts and ends in state 0, with start level labeled 0 and end level labeled κ, then for all $0 < l < \kappa$, the defining equation in the optimization problem is

$$metric(0, \kappa) = \min_{m \in \text{ states}} (metric(0, l_m) + metric(l_m, G)) \tag{41}$$

where $metric(i, j)$ is the minimum metric of any path originating from node i and ending in node j and l_m represent the mth node in level l. With additive Gaussian noise, the metric for each branch is the mean-squared error between the symbol estimate and the received data. Once we know the metric associated with all the nodes in level l, the metric associated with the mth node in level $l + 1$ can be calculated by

$$metric(0, (l+1)_m) = \min_{i \in \text{ states}} \left(metric(0, l_i) + metric(l_i, (l+1)_m) \right). \tag{42}$$

If there is no branch between the node i in state l and node m in state $l + 1$, then the metric associated with that branch is assumed to be infinitely large.

This iterative method of calculating the optimal code reduces the complexity of the decoder to be linear in codeword length κ. However, at every stage of the trellis, the Viterbi algorithm requires computation of the likelihood of each state. The number of states is exponential in the size of the constraint length, S, of the code, thereby making the total complexity of the algorithm of the order of $\kappa 2^{(S+1)}$.

For large constraint lengths, the Viterbi decoding can be impractical for real-time low-power applications. As applications require higher data rates with increasing reliability, higher constraint lengths are desirable. There have been several low-complexity alternatives to Viterbi decoding proposed in the literature: sequential decoding [135], majority logic decoding [136], M-algorithm or list-decoding [137, 138], T-algorithm [139], reduced state sequence detection [140, 141] and maximal weight decoding [59].

As noted in the beginning of this section, most of the channel coding and decoding procedures are designed for single-user AWGN channels or fading channels. In the presence of multiaccess interference, joint multiuser detection and decoding [142–146] can lead to lower error performance at the expense of increased receiver complexity.

Multiple transmit antennas: The information symbols encoded using the Alamouti scheme in Figure 8 admit a simple maximum likelihood decoder. With two transmit and single receive antenna, the sampled received signal in two consecutive time symbols is given by

$$
\begin{aligned}
z[1] &= h_1 s_1 + h_2 s_2 + n_1 \\
z[2] &= -h_1 s_2^* + h_2 s_1^* + n_1
\end{aligned}
$$

where n_1 and n_2 are assumed to be independent instances of circularly symmetric Gaussian noise with zero mean and unit variance. The maximum likelihood detector builds the following two signals,

$$
\begin{aligned}
\widehat{s}_1 &= h_1^* z[1] + h_2 z^*[2] = (|h_1|^2 + |h_2|^2) s_1 + h_1^* n_1 + h_2 n_2^* \\
\widehat{s}_2 &= h_2^* z[1] - h_1 z^*[2] = (|h_1|^2 + |h_2|^2) s_2 - h_0 n_1^* + h_1^* n_1
\end{aligned}
\tag{43}
$$

33

followed by the maximum likelihood detector for each of the symbol s_i, $i = 1, 2$. The combined signals in (43) are equivalent to that obtained from a two-branch receive diversity using maximal ratio combining (MRC) [6]. Thus, the Alamouti scheme provides an order two transmit diversity much like an order two receive diversity using MRC. Note that both the Alamouti and MRC schemes have the same average transmission rate, one symbol per transmission, but the Alamouti scheme requires at least two transmissions to achieve order two diversity, while MRC achieves order two diversity per transmission.

If a space-time trellis code is used, then the decoder is a simple extension of the decoder for the single-antenna case. As the number of antennas is increased to achieve higher data rates, the decoding complexity increases exponentially in the number of transmit antennas [5], thereby requiring power hungry processing at the receiver. Though there is no work on reduced complexity decoders for space-time trellis codes, complexity reduction concepts for single-antenna trellis decoding should apply (see above).

5.3 Power Control

Power control was amply motivated on the capacity grounds in Sections 4.1 and 4.2; in this section, we will only highlight some of the representative research on power control methods and its benefits. Power control is widely used in second and third generation cellular systems. For instance, in IS-95, transmit power is controlled not only to counter the near-far effect, but also to overcome the time-varying fading. By varying the transmit power based on the channel conditions, a fixed received signal to noise ratio (SNR) can be achieved. A SNR guarantee implies a guarantee on the reliability of received information, through the relation between the packet error rate and the received SNR [4].

Information theoretically optimal power control for a multiuser system was discussed in [147–150]. While providing a bound on the achievable capacity, the proposed power control algorithms assume perfect knowledge of the time-varying channel at the transmitter. Hence, the power control policies and the resultant system performance is only a loose bound for the achievable performance. Network capacity analysis with power control errors has appeared in [151, 152, and references therein].

Recently, significant research effort has been devoted to power control algorithms for data traffic, *e.g.*, the work in [153–160]. Most of the above work on power control has been for circuit switched networks, where users are given a certain dedicated channel for their entire session. With the advent of services supporting bursty traffic, like email, web browsing etc., resource allocation for shared channels and packet networks becomes of importance. First steps in these directions can be found in [158, 159, 161]. Lastly, we note that power control can also lead to gain in packet switched networks, like IEEE 802.11 or *ad hoc* networks; preliminary results can be found in [162, 163].

6 Conclusions

If the relentless advances in wireless communications in the past decade are an indicator of things to come, then it is clear that we will witness not only faster ways to communicate, but also newer modes of communication. The fundamental information theoretic bounds hold as long as the assumed communication model holds. The capacity of the channel can be "increased," by introducing new capabilities like multiple antennas and *ad hoc* networking. Thus, it will be safe to conclude that the actual physical limits of wireless communication are still unknown and it is for us to exploit that untapped potential with a mix of creativity and serendipity.

References

[1] R. Steele, J. Whitehead, and W. C. Wong, "System aspects of cellular radio," *IEEE Communications Magazine*, vol. 33, pp. 80–86, Jan. 1995.

[2] U. T. Black, *Mobile and Wireless Networks*. Prentice Hall, 1996.

[3] J. Geier, *Wireless LANs: Implementing Interoperable Networks*. Macmillan Technical Publishing, 1998.

[4] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Info. Th.*, vol. 44, pp. 744–765, Mar. 1998.

[6] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, 1959.

[7] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 1996.

[8] M. G. Jansen and R. Prasad, "Capacity, throughput, and delay analysis of a cellular DS CDMA system with imperfect power control and imperfect sectorization," *IEEE Trans. Veh. Tech.*, vol. 44, pp. 67–75, Feb. 1995.

[9] A. Sabharwal, D. Avidor, and L. Potter, "Sector beam synthesis for cellular systems using phased antenna arrays," *IEEE Trans. Veh. Tech.*, vol. 49, pp. 1784–1792, Sep. 2000.

[10] E. S. Sousa, V. M. Jovanović, and C. Daigneault, "Delay spread measurements for the digital cellular channel in Toronto," *IEEE Trans. Veh. Tech.*, vol. 43, pp. 837–847, Nov. 1994.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[12] S. Verdú, *Multiuser Detection*. Cambridge University Press, 1998.

[13] S. Verdú and S. S. (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Info. Th.*, vol. 45, pp. 622–40, Mar. 1999.

[14] S. S. (Shitz) and S. Verdú, "The impact of frequency flat fading on the spectral efficiency of CDMA," *IEEE Trans. Info. Th.*, vol. 47, pp. 1302–1327, May 2001.

[15] I. Katzela and M. Nagshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Pers. Comm.*, pp. 10–31, June 1996.

[16] S. Jordan, "Resource allocation in wireless networks," *J. High Speed Networks*, vol. 5, no. 1, pp. 23–4, 1996.

[17] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. 1993 Int. Conf. Comm.*, (Geneva, Switzerland), pp. 1064–1070, May 1993.

[18] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Comm. Syst.*, vol. CS-11, pp. 360–393, Dec. 1963.

[19] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communication aspects," *IEEE Trans. Info. Th.*, vol. 44, pp. 2619–2692, Oct. 1998.

[20] E. Malkamäki and H. Leib, "Coded diversity on block-fading channels," *IEEE Trans. Info. Th.*, vol. 45, pp. 771–781, Mar. 1999.

[21] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Info. Th.*, vol. 45, no. 1, pp. 139–157, 1999.

[22] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. Info. Th.*, vol. 46, pp. 189–205, Jan. 2000.

[23] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–423 (Part I), 623–656 (Part II), 1948.

[24] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Info. Th.*, vol. 43, pp. 359–378, May 1994.

[25] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," tech. rep., AT&T Bell Labs, 1995. (Appeared in *Eur. Trans. Tel.*, vol. 10, no. 6, pp. 585–595, 1999).

[26] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Info. Th.*, vol. 43, pp. 1986–92, Nov. 1997.

[27] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Info. Th.*, vol. 45, pp. 1468–1489, July 1999.

[28] A. Sabharwal, E. Erkip, and B. Aazhang, "On side information in multiple antenna block fading channels," in *Proceedings ISITA*, (Honolulu, Hawaii), Nov. 2000.

[29] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE-JSAC*, vol. 16, pp. 1423–1436, Oct. 1998.

[30] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Info. Th.*, vol. 45, pp. 2418–2433, Nov. 1999.

[31] J.-C. Guey, M. Fitz, M. Bell, and W. Y. Kuo, "Signal design for transmitter diversity wireless communication systems over rayleigh fading channels," *IEEE Trans. Comm.*, vol. 46, pp. 527–37, April 1999.

[32] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Comm.*, vol. 47, pp. 199–207, Feb. 1999.

[33] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill International Editions, 1984.

[34] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Info. Th.*, vol. 40, pp. 1147–57, July 1994.

[35] H. Viswanathan, *Capacity of fading channels with feedback and sequential coding of correlated sources*. PhD thesis, Cornell University, August 1997.

[36] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *submitted to IEEE Trans. Info. Th.*, August 1999.

[37] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.

[38] G. Caire and S. Shamai (Shitz), "On the capacity of some channels with channel state information," *IEEE Trans. Info. Th.*, vol. 45, pp. 2007–2019, Sept. 1999.

[39] S. Shamai and A. D. Wyner, "Information theoretic considerations for symmetric, cellular, multiple-access fading channels – Part I," *IEEE Trans. Info. Th.*, vol. 43, pp. 1877–1894, Nov. 1997.

[40] G. W. Wornell, "Spread-signaure CDMA: Efficient multiuser communication in presence of fading," *IEEE Trans. Info. Th.*, vol. 41, pp. 1418–1438, Sept. 1995.

[41] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications - a tutorial," *IEEE Trans. Comm.*, vol. COM-30, pp. 855–884, May 1982.

[42] E. Erkip and B. Aazhang, "Multiple access schemes over multipath fading channels," in *Proc. ISIT*, (Cambridge, MA), August 1998.

[43] A. D. Wyner, "Shannon-theoretic approach to Gaussian cellular multiple access channel," *IEEE Trans. Info. Th.*, vol. 40, pp. 1713–1727, Nov. 1994.

[44] O. Somekh and S. S. (Shitz), "Shannon-theoretic approach to a Gaussian cellular multiple-access channel with fading," in *Proc. 1998 IEEE Int. Symp. Info. Th.*, (Cambridge, MA), p. 393, Aug. 1998.

[45] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*. Addison-Wesley, 1995.

[46] D. Bertsekas and R. Gallager, *Data Networks*. Prentice Hall, 1992.

[47] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Info. Th.*, vol. IT-31, pp. 124–142, Mar. 1985.

[48] A. Ephremides and B. Hajek, "Information theory and communication networks: An unconsummated union," *IEEE Info. Th.*, vol. 44, pp. 2416–2434, Oct. 1998.

[49] I. E. Telatar and R. G. Gallager, "Combining queuing theory with information theory for multi-access," *IEEE J. Sel. Ar. Comm.*, vol. 13, pp. 963–969, Aug. 1995.

[50] M. Win and R. A. Scholz, "Impulse radio: how it works," *IEEE Comm. Lett.*, vol. 2, pp. 36–38, Feb. 1998.

[51] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Info. Th.*, vol. 44, pp. 2260–2299, Oct. 1998.

[52] J. D. J. Costello, J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Trans. Info. Th.*, vol. 44, pp. 2531–2560, Oct. 1998.

[53] A. R. Calderbank, "The art of signalling," *IEEE Trans. Info. Th.*, vol. 44, pp. 2561–2595, Oct. 1998.

[54] I. Blake, C. Heegard, T. Hϕholdt, and V. Wei, "Algebraic-geometry codes," *IEEE Trans. Info. Th.*, vol. 44, pp. 2596–2618, Oct. 1998.

[55] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw Hill, 1968.

[56] G. D. Forney, *Concatenated Codes.* Cambridge, MA: MIT Press, 1966.

[57] J. H. van Lint, *Introduction to Coding Theory.* New York: Springer-Verlag, 1992.

[58] V. S. Pless, W. C. Huffman, and R. A. Brualdi, eds., *Handbook of Coding Theory.* New York: Elsevier, 1998.

[59] S. Das, *Multiuser Information Processing in Wireless Communication.* PhD thesis, Rice University, Houston, TX, Sept 2000.

[60] A. Dholakia, *Introduction to Convolutional Codes with Applications.* Kluwer Academic Publishers, 1994.

[61] D. Divsalar and M. K. Simon, "The design of trellis coded MPSK for fading channels: Performance criteria," *IEEE transactions on Communications*, vol. 36, pp. 1004–1012, September 1988.

[62] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE transactions on Information Theory*, vol. 28, pp. 55–67, January 1982.

[63] G. D. Forney and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Info. Th.*, vol. 44, pp. 2384–2415, Oct. 1998.

[64] J. L. Massey, "Coding and modulation in digital communications," in *Proc. 1974 Int. Zurich Seminar on Digital Comm.*, (Zurich, Switzerland), pp. E2(1)–(4), Mar. 1974.

[65] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* New York: Wiley, 1965.

[66] R. G. Gallager, *Low Density Parity-Check Codes.* Cambridge, MA: MIT Press, 1962.

[67] S. L. Goff, A. Glavieux, and C. Berrou, "Turbo-codes and high spectral efficiency modulation," in *Proc. 1994 Int. Conf. Comm.*, vol. 2, pp. 645–649, May 1994.

[68] W. Liu and S. G. Wilson, "Rotationally-invariant concatenated (turbo) TCM codes," in *Proc. 1999 Asilomar Conf. Sig. Sys. Comp.*, vol. 1, pp. 32–36, Oct. 1999.

[69] O. Y. Takeshita and J. D. J. Costello, "New deterministic interleaver designs for turbo codes," *IEEE Trans. Info. Th.*, vol. 46, pp. 1988–2006, Sept. 2000.

[70] Y. Liu, M. Fitz, and O. Y. Takeshita, "QPSK space-time turbo codes," in *Proc. 2000 Int. Conf. Comm.*, vol. 1, pp. 292–296, June 2000.

[71] V. Tarokh, A. Vardy, and K. Zeger, "Universal bound on the performance of lattice codes," *IEEE Trans. Info. Th.*, vol. 45, pp. 670–681, March 1999.

[72] J. H. Winters, "Smart antennas for wireless systems," *IEEE Pers. Comm.*, vol. 5, pp. 23–27, Feb. 1998.

[73] G. J. F. Jr. and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," in *Wireless Personal Communications*, Kluwer Academic Publishers, March 1998.

[74] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Info. Th.*, vol. 45, pp. 1456–1467, July 1999.

[75] D. M. Ionescu, "New results on space time code design criteria," in *Proc. IEEE Wireless Communications and Networking Conference*, (New Orleans, Lousiana), October 1999.

[76] O. Tirkkonnen and A. Hottinen, "Complex space-time block codes for four Tx antennas," in *Proc. GLOBECOM*, pp. 1005–1009, 2000.

[77] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis coded modulation," *IEEE Comm. Letters*, vol. 4, pp. 20–22, Jan. 2000.

[78] A. R. Hammons and H. E. Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Info. Th.*, vol. 46, pp. 524–542, Mar 2000.

[79] T. Moharemovic and B. Aazhang, "Information theoretic optimality of orthogonal space-time transmission schemes and concatenated code construction," in *Proc. Int. Conf. Comm. (ICC)*, (Alcapulco, Mexico), May 2000.

[80] M. J. Borran, M. Memarzadeh, and B. Aazhang, "Design of coded modulation schemes for orthogonal transmit diversity," *submitted to IEEE Trans. Comm.*, April 2001.

[81] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE journal on Select Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[82] R. Prasad, "Overview of wireless personal communications: Microwave perspectives," *IEEE Comm. Mag.*, pp. 104–108, April 1997.

[83] G. Taricco, E. Biglieri, and G. Caire, "Limiting performance of block-fading channels with multiple antennas," in *Proc. Info. Th. Comm. Workshop*, pp. 27–29, 1999.

[84] http://www.3gpp.org/.

[85] T. Ojanpera and R. Prasad, eds., *Wideband CDMA for Third Generation Mobile Communications.* Artech House Universal Personal Communications Series, 1998.

[86] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainity," *IEEE Trans. Info. Th.*, vol. 44, pp. 2148–2177, Oct. 1998.

[87] K. Li and H. Liu, "Joint channel and carrier offset estimation in CDMA communications," *IEEE Trans. Sig. Proc.*, vol. 47, pp. 1811–22, July 1999.

[88] Z. Pi and U. Mitra, "Blind delay estimation in multi-rate asynchronous DS-CDMA systems," *submitted to IEEE Trans. Comm.*, 2000.

[89] C. Sengupta, *Algorithms and architectures for channel estimation in wireless CDMA communication systems.* PhD thesis, Rice University, Dec. 1998.

[90] Z. Tian, K. L. Bell, and H. L. V. Trees, "A quadratically constrained decision feedback equalizer for DS-CDMA communication systems," in *IEEE Workshop on Sig. Proc. Advances in Wireless Comm.*, pp. 190–193, May 1999.

[91] C. Sengupta, J. Cavallaro, and B. Aazhang, "On multipath channel estimation for DS-CDMA systems with multiple sensors," *IEEE Trans. Comm.*, vol. 49, pp. 543–553, Mar. 2001.

[92] D. N. Godard, "Self-recovering equalization and carrier tracking of two-dimensional data communication systems," *IEEE Trans. Comm.*, vol. 28, pp. 1867–1875, Nov. 1980.

[93] J. R. Treichler and B. G. Agree, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust. Speech Signal Proc.*, vol. 31, pp. 459–472, Apr. 1983.

[94] W. A. Gardner, "A new method for channel identification," *IEEE Trans. Comm.*, vol. 39, pp. 813–817, June 1991.

[95] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE SP Mag.*, pp. 14–36, April 1991.

[96] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Trans. Info. Th.*, vol. 40, pp. 340–349, Mar. 1994.

[97] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Sig. Proc.*, vol. 43, pp. 516–525, Feb. 1995.

[98] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Processing*, vol. 50, no. 1-2, pp. 83–99, 1996.

[99] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communications," *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.

[100] E. Ertin, U. Mitra, and S. Siwamogsatham, "Maximum-likelihood based multipath channel estimation for code-division multiple-access systems," *IEEE Trans. Comm.*, vol. 49, pp. 290–302, Feb. 2001.

[101] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc. Ser. B*, pp. 1–38, 1977.

[102] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Sig. Proc.*, vol. 36, pp. 1553–60, Oct. 1988.

[103] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Sig. Proc.*, vol. 45, pp. 137–147, Jan. 1997.

[104] M. K. Tsatsanis and G. B. Giannakis, "Blind estimation of direct sequence spread spectrum signals in multipath," *IEEE Sig. Proc.*, vol. 45, pp. 1241–52, 1997.

[105] T. Östman and B. Ottersten, "Near far robust time delay estimation for asynchronous DS-CDMA systems with bandlimited pulse shapes," in *Proc. IEEE Vehi. Tech. Conf.*, pp. 1651–54, May 1998.

[106] V. Tripathi, A. Mantravadi, and V. V. Veeravalli, "Channel acquisition for wideband CDMA," *IEEE J. Sel. Areas Comm.*, vol. 18, pp. 1483–1494, Aug. 2000.

[107] E. Aktas and U. Mitra, "Single user sparse channel acquisition for ds/cdma," in *Proc. CISS*, (Princeton, NJ), 2000.

[108] S. Bhashyam, A. Sabharwal, and U. Mitra, "Channel estimation multirate DS-CDMA systems," in *Proc. Asilomar Conf. Sig. Sys. Comp.*, (Pacific Grove, CA), Oct/Nov. 2000.

[109] T. P. Krauss and M. D. Zoltowski, "Blind channel identification on CDMA forward link based on dual antenna receiver at handset and cross-relation," in *Proc. 1999 Asilomar Conf. Sig. Sys. Comm.*, vol. 1, pp. 75–79, Oct. 1999.

[110] C. R. Rao, *Linear Statistical Inference and its Applications*. Wiley, New York, 1973.

[111] S. Bhashyam and B. Aazhang, "Multiuser channel estimation for long code CDMA systems," in *Proc. 2000 Wireless Comm. Networking Conf.*, 2000.

[112] S. Verdú, *Optimum Multiuser Signal Detection*. PhD thesis, University of Illinois at Urbana-Champaign, Aug. 1984.

[113] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, pp. 303–312, 1989.

[114] S. Verdú, "Optimum multiuser asymptotic efficiency," *IEEE Trans. Comm.*, vol. 34, pp. 890–897, Sept. 1986.

[115] K. S. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Trans. Aerospace Elect. Sys.*, vol. AES-15, pp. 181–185, Jan. 1979.

[116] R. Kohno, M. Hatori, and H. Imai, "Cancellation techniques of co-channel interference in asynchronous spread spectrum multiple access systems," *Elect. and Commun. in Japan*, vol. 66-A, no. 5, pp. 20–29, 1983.

[117] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Info. Th.*, vol. 35, no. 1, pp. 123–36, 1989.

[118] R. Lupas and S. Verdú, "Near-far resistance of multi-user detectors in asynchronous channels," *IEEE Trans. Comm.*, vol. 38, pp. 496–508, April 1990.

[119] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE JSAC*, vol. 8, pp. 683–90, May 1990.

[120] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Info. Th.*, vol. 41, pp. 944–960, July 1995.

[121] R. Mailloux, *Phased Array Antenna Handbook*. Artech House, 1994.

[122] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Info. Th.*, vol. 43, pp. 858–871, May 1997.

[123] A. Sabharwal, U. Mitra, and R. Moses, "Cyclic Wiener filtering based multirate DS-CDMA receivers," in *Proc. of IEEE WCNC*, (New Orleans, LA), Sept. 1999.

[124] A. Sabharwal, U. Mitra, and R. Moses, "Low complexity MMSE receivers for multirate DS-CDMA systems," in *Proc. CISS*, (Princeton, NJ), 2000.

[125] H. Y. Wu and A. Duel-Hallen, "Performance comparison of multi-user detectors with channel estimation for flat Rayleigh fading CDMA channel," *Wireless Personal Comm.*, July/Aug. 1996.

[126] S. D. Gray, M. Kocic, and D. Brady, "Multiuser detection in mismatched multiple-access channels," *IEEE Trans. Comm.*, vol. 43, pp. 3080–89, Dec. 1995.

[127] B. Rimoldi and R. Urbanke, "A rate splitting approach to the Gaussian multiple-access channel," *IEEE Trans. Info. Th.*, vol. 42, pp. 364–375, March 1996.

[128] K. I. Pederson, T. E. Kolding, I. Seskar, and J. M. Holzman, "Practical implementation of successive interference cancellation in DS/CDMA systems," in *Proc. IEEE Conference on Universal Personal Commun.*, vol. 1, pp. 321–325, 1996.

[129] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, April 1990.

[130] P. Patel and J. Holzman, "Performance comparison of a DS/CDMA system using a successive interference cancellation (IC) scheme and a parallel IC scheme under fading," in *Proc. ICC*, (New Orleans, LA), pp. 510–514, May 1994.

[131] R. M. Buehrer and B. D. Woerner, "Analysis of adaptive multistage interference cancellation for CDMA using an improved Gaussian approximation," in *Proc. IEEE MILCOM*, (San Diego, CA), pp. 1195–99, Nov. 1995.

[132] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access schemes," *IEEE Trans. Comm.*, vol. 39, pp. 725–736, May 1991.

[133] S. Moshavi, *Multistage linear detectors for DS-CDMA communications.* PhD thesis, City Univ. New York, NY, Jan. 1996.

[134] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Info. Th.*, vol. IT-13, pp. 260–269, April 1967.

[135] J. M. Wozencraft and B. Reiffen, *Sequential Decoding.* Cambridge, MA: MIT Press, 1961.

[136] J. L. Massey, *Threshold Decoding.* Cambridge, MA: MIT Press, 1998.

[137] J. B. Anderson and S. Mohan, "Sequential coding algorithms: A survey and cost analysis," *IEEE Trans. Comm.*, vol. COM-32, pp. 169–176, Feb. 1984.

[138] G. J. Pottie and D. P. Taylor, "A comparison of reduced complexity decoding algorithms for trellis codes," *IEEE J. Sel. Areas Comm.*, vol. 7, pp. 1369–1380, Dec. 1989.

[139] S. T. Simmons, "Breadth-first trellis decoding with adaptive effort," *IEEE Trans. Comm.*, vol. 38, pp. 3–12, Jan. 1990.

[140] M. V. Eyuboglu and S. Qureshi, "Reduced-state sequence estimation for coded modulation on interference channels," *IEEE J. Sel. Areas Comm.*, vol. 35, pp. 944–955, Sept. 1989.

[141] P. R. Chevillat and E. Elephtheriou, "Decoding of trellis-encoded signals in the presence of inter-symbol interference and noise," *IEEE Trans. Comm.*, vol. 37, pp. 669–676, July 1989.

[142] C. Schlegel, P. Alexander, and S. Roy, "Coded asynchronous CDMA and its efficient detection," *IEEE Trans. Info. Th*, vol. 44, pp. 2837–2847, Nov. 1998.

[143] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Comm.*, vol. 47, pp. 1046–61, July 1999.

[144] H. E. Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Sel. Areas Comm.*, vol. 18, pp. 30–41, Jan. 2000.

[145] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding flat-fading channels via mixture Kalman filtering," *IEEE Trans. Info. Th.*, vol. 46, pp. 2079–94, Sept. 2000.

[146] L. Wei and H. Qi, "Near-optimal limited-search detection on ISI-CDMA channels and decoding of long-convolutional codes," *IEEE Trans. Info. Th.*, vol. 46, pp. 1459–1482, July 2000.

[147] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels-Part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Info. Th.*, vol. 44, pp. 2796–2815, Nov. 1998.

[148] S. V. Hanly and D. N. C. Tse, "Multiaccess fading channels-Part II: delay-limited capacities," *IEEE Trans. Info. Th.*, vol. 44, pp. 2816–2831, Nov. 1998.

[149] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. Info. Th.*, vol. 45, pp. 1968–83, Sept. 1999.

[150] S. Hanly and D. Tse, "Power control and capacity of spread-spectrum wireless networks," *Automatica*, vol. 35, pp. 1987–2012, Dec. 1999.

[151] N. Bambos, "Toward power-sensitive network architectures in wireless communications: concepts, issues, and design aspects," *IEEE Personal Comm.*, vol. 5, pp. 50–59, June 1998.

[152] J. Zhang and E. K. P. Chong, "CDMA systems in fading channels: admissibility, network capacity and power control," *IEEE Trans. Info. Th.*, vol. 46, pp. 962–981, May 2000.

[153] J. Wu and R. Kohno, "A wireless multimedia CDMA system based on transmission power control," *IEEE J. Sel. Areas Comm.*, vol. 14, pp. 683–691, May 1996.

[154] J. Jacobsmeyer, "Congestion relief on power-controlled CDMA networks," *IEEE J. Sel. Areas Comm.*, vol. 14, pp. 1758–61, Dec. 1996.

[155] A. Sampath and J. M. Holtzman, "Access control of data in integrated voice/data CDMA systems: benefits and tradeoffs," *IEEE J. Sel. Areas Comm.*, vol. 15, pp. 1511–26, Oct. 1997.

[156] D. Ayyagari and A. Ephremides, "Cellular multicode CDMA capacity for integrated (voice and data) services," *IEEE J. Sel. Areas Comm.*, vol. 17, pp. 928–938, May 1999.

[157] Y. Lu and R. W. Broderson, "Integrating power control, error correction coding, and scheduling for a CDMA downlink system," *IEEE J. Sel. Areas Comm.*, vol. 17, pp. 978–989, May 1999.

[158] D. Kim, "Rate-regulated power control for supporting flexible transmission in future cdma mobile networks," *IEEE J. Sel. Areas Comm.*, vol. 17, pp. 968–977, May 1999.

[159] S. Manji and W. Zhuang, "Power control and capacity analysis for a packetized indoor multimedia DS-CDMA network," *IEEE Trans. Veh. Tech.*, vol. 49, pp. 911–935, May 2000.

[160] D. Goodman and N. Mandayam, "Power control for wireless data," *IEEE Personal Comm.*, vol. 7, pp. 48–54, April 2000.

[161] N. Bambos and S. Kandukuri, "Power controlled multiple access (PCMA) in wireless communication networks," in *Proc. INFOCOM 2000*, vol. 2, pp. 386–395, March 2000.

[162] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Info. Th.*, vol. 46, pp. 388–404, March 2000.

[163] J. Monks, "Power controlled multiple access in *ad hoc* networks," in *Proc. of Multiaccess, Mobility and Teletraffic for Wireless Communications(MMT)*, (Duck Key, FL), December 2000.
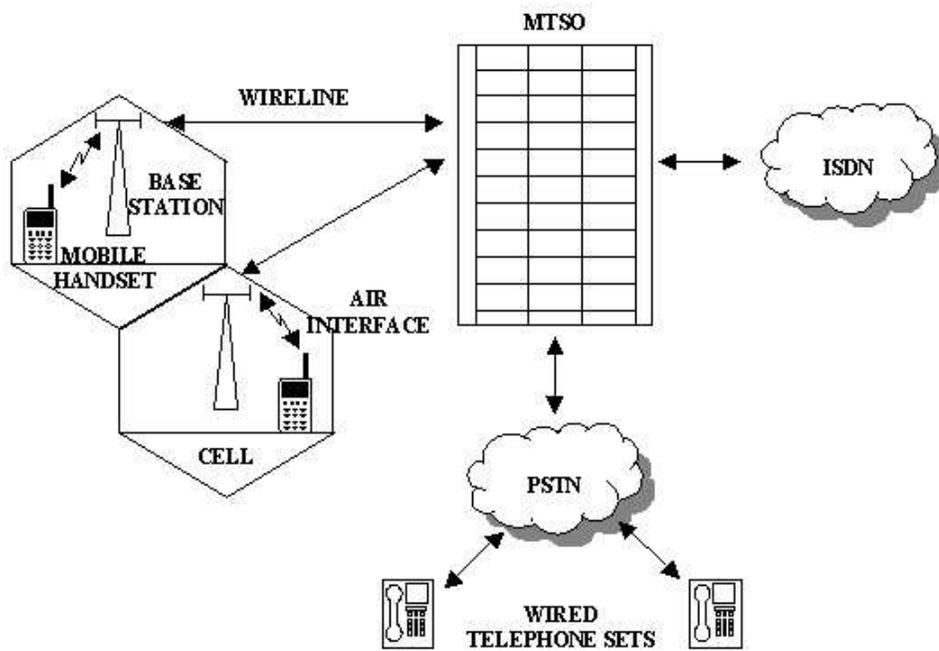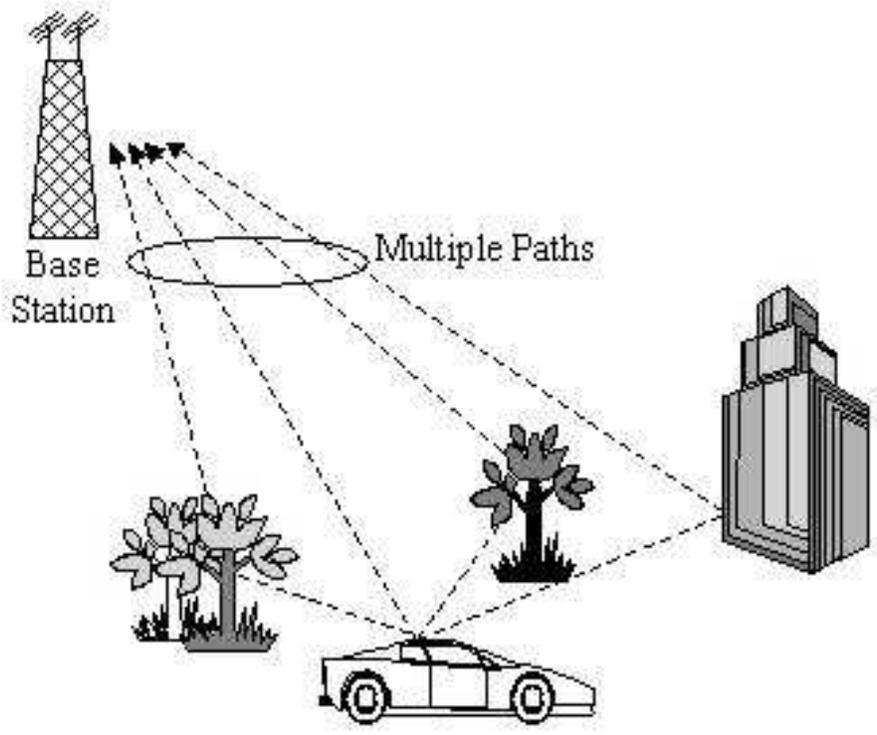
Figure 1: Components of a cellular wireless network.

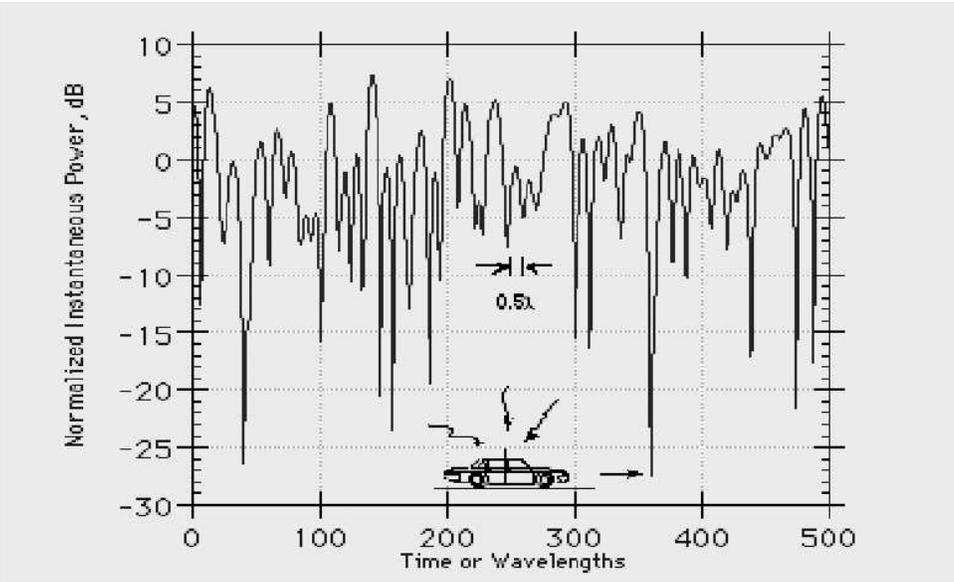Figure 2: Multipath propagation.

Figure 3: Time variations in the received signal power due to multipath and user mobility.
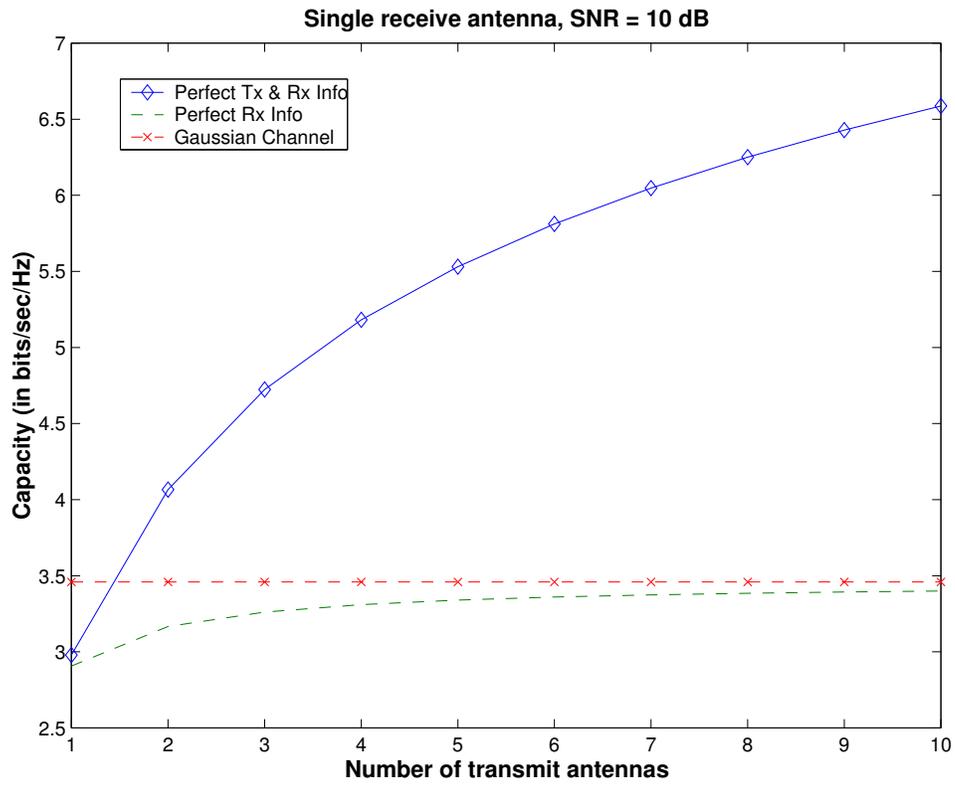
Figure 4: Capacity with multiple transmit antennas and single receive antenna, with different amount of channel state information at the transmitter.
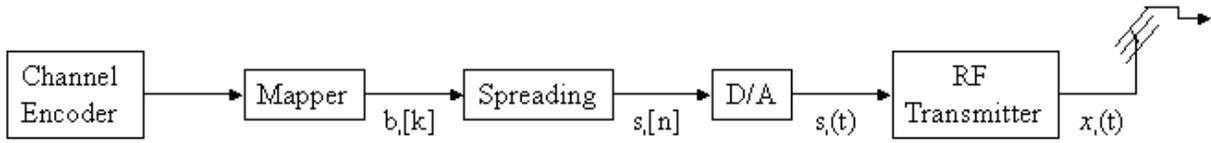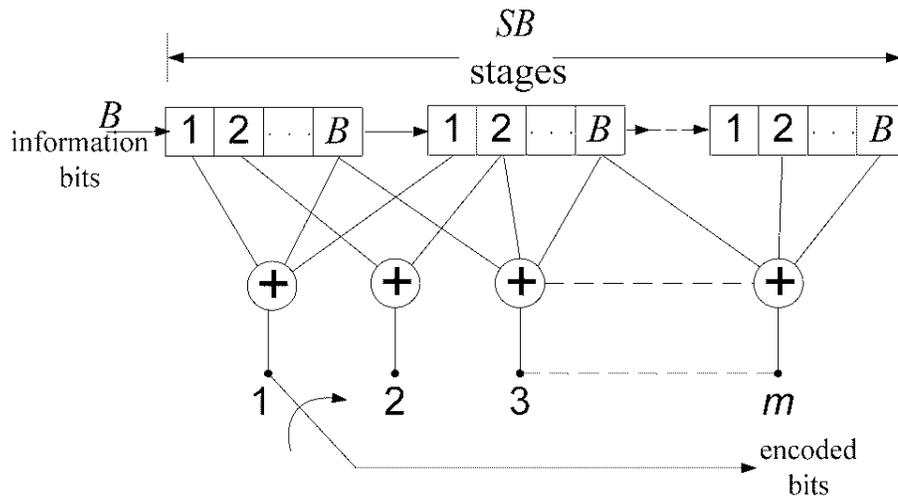
Figure 5: DS-CDMA handset transmitter components.
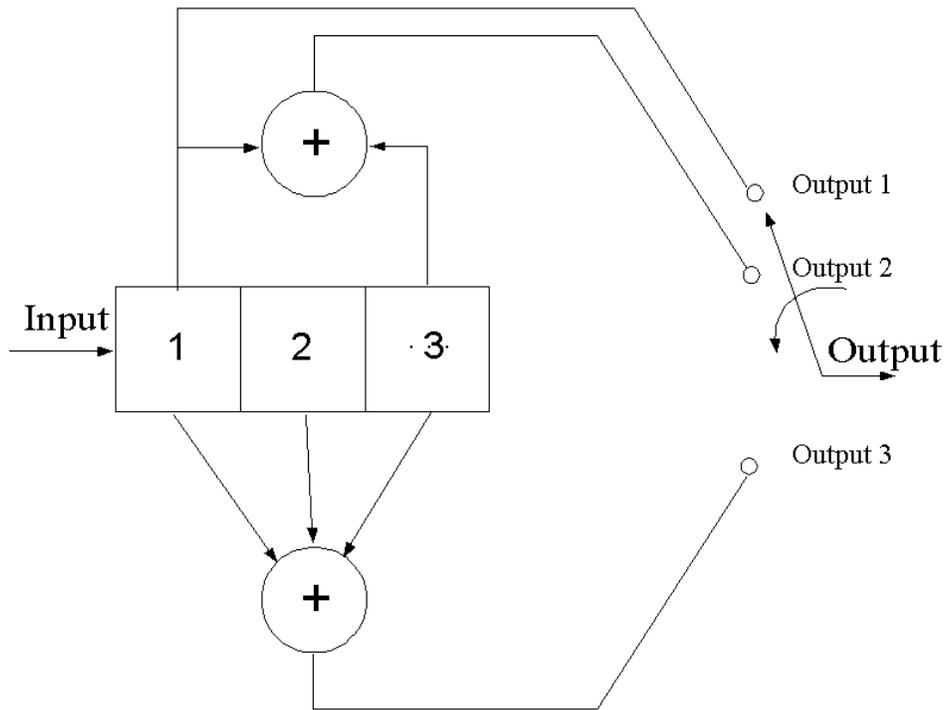

Figure 6: Convolutional encoder.
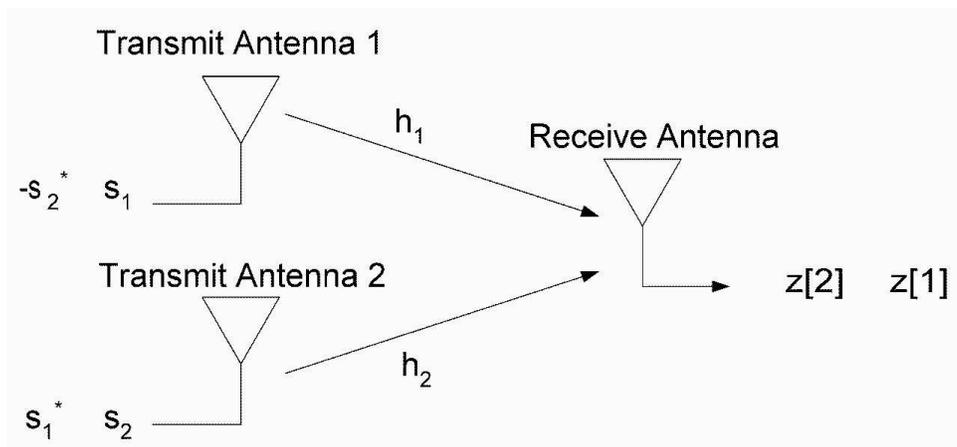
Figure 7: Convolutional encoder for a (3,1) code.

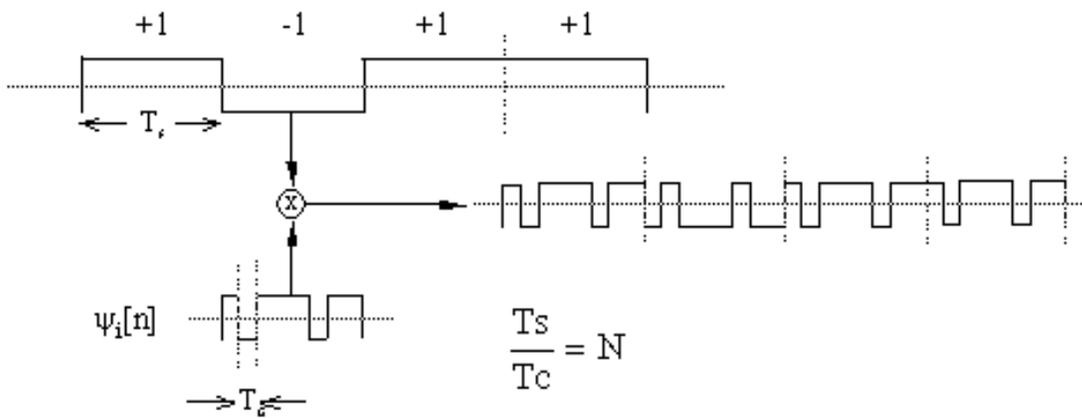
Figure 8: Alamouti encoder for two transmit and one receive antenna.
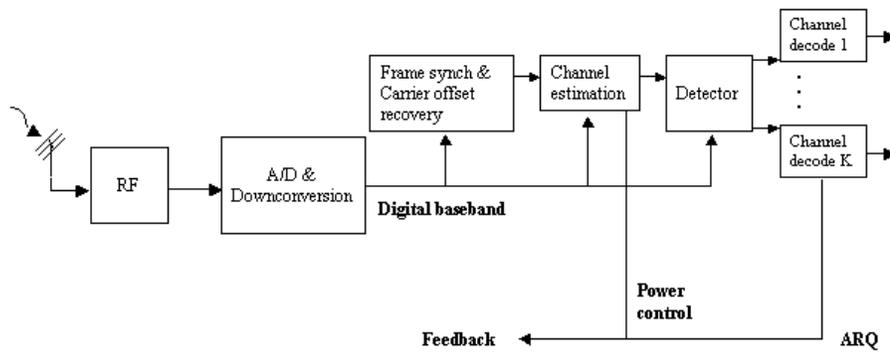
Figure 9: The spreading operation.
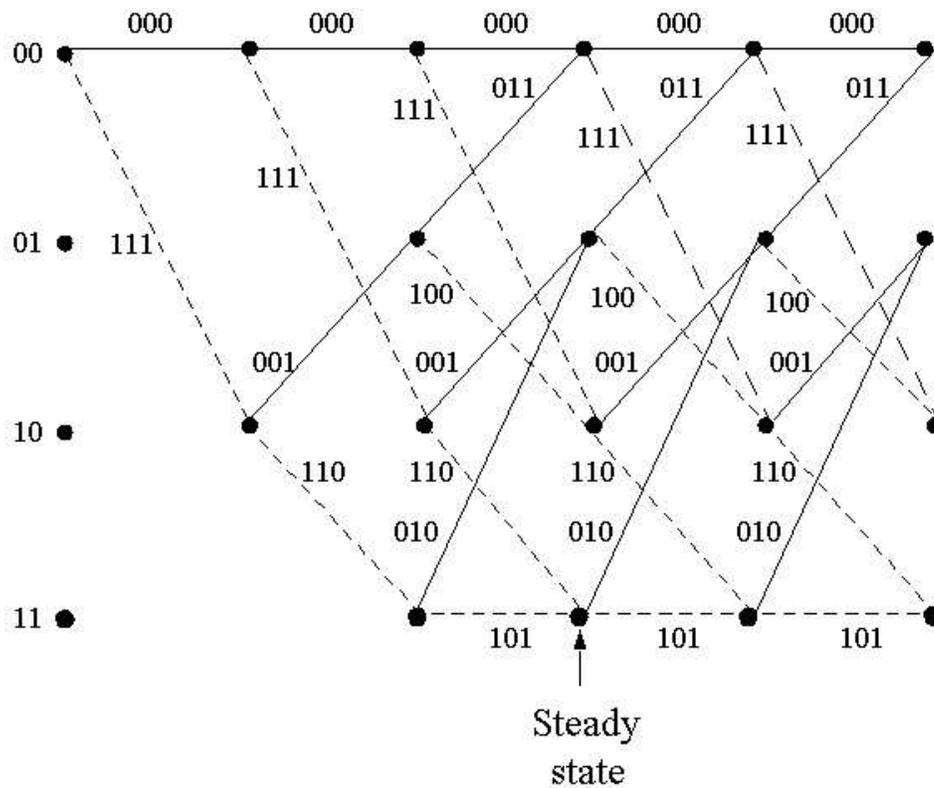
Figure 10: Base-station receiver structure.

Figure 11: Trellis diagram for the example (3,1) convolutional code.