WIRELESS NETWORKS
MULTIUSER DETECTION
IN CROSS-LAYER DESIGN

Information Technology: Transmission, Processing, and Storage

Communication System Design Using DSP Algorithms: With
Laboratory Experiments for the TMS320C6701 and TMS320C6711
Steven A. Tretter

Interference Avoidance Methods for Wireless Systems
Dimitrie C. Popescu and Christopher Rose

MIMO Signals and Systems
Horst J. Bessai

Performance Analysis and Modeling of Digital Transmission Systems
William Turin

Stochastic Image Processing
Chee Sun Won and Robert M. Gray

Wireless Communications Systems and Networks
Mohsen Guizani

Wireless Networks
Multiuser Detection in Cross-Layer Design
Cristina Comaniciu, Narayan B. Mandayam, and H. Vincent Poor

WIRELESS NETWORKS

MULTIUSER DETECTION
IN CROSS-LAYER DESIGN

Cristina Comaniciu
Stevens Institute of Technology
Hoboken, New Jersey

Narayan B. Mandayam
Rutgers University
Piscataway, New Jersey

H. Vincent Poor
Princeton University
Princeton, New Jersey

Printed in the United States of America.

9   8   7   6   5   4   3   2   1

springeronline.com

To our families

Contents

List of Figures

List of Tables

Preface

Wireless networking is undergoing a transformation from what has been primarily a medium for supporting voice traffic between telephones, into what is increasingly becoming a medium for supporting traffic among a variety of digital devices transmitting media of many types (voice, data, images, video, etc.) Wireline networking underwent a similar transformation in the 1990s, which led to an enormous build-up in the capacity of such networks, primarily through the addition of new optical fiber, switches and other infrastructure. Creating a similar build-up in the capacity of wireless networks presents many challenges, including notably the scarcity of two of the principal resources for providing high capacity in wireless networks, namely power and bandwidth. Moreover, the physical nature of wireless communication channels themselves, involving such features as mobility, interference, and fading, adds to the challenge of providing high-quality multimedia communications to large groups of users.

A principal way of enabling the advanced services required of wireless networks is to add intelligence throughout the network in order to exploit increases in processing power afforded by Moore's Law type improvements in microelectronics. One way of doing this is through the introduction of advanced signal processing at the node level of the network, in order to mitigate the impairments of the wireless channel and to exploit the diversity opportunities provided by such channels. Multiuser detection, which addresses issues of optimal signal reception in multiple-access channels, is a major technique in this context. A very extensive research effort has been devoted to the development of multiuser de-

tection algorithms over the past two decades[1]. This research has shown that substantial performance gains can be realized in interference-limited channels through the introduction of advanced signal processing.

Recent research activity in wireless networking has begun to focus on the higher layers of the network, and on the special problems presented at such layers by the particular properties of the wireless physical layer. One of the key issues of this research is cross-layer design, which seeks to enhance the capacity of wireless networks significantly through the joint optimization of multiple layers in the network, primarily the physical (PHY) and medium access control (MAC) layers. Although there are advantages of such design in wireline networks as well, this approach is particularly advantageous for wireless networks due to properties such as mobility and interference that strongly affect performance and design of higher layer protocols. This monograph is concerned with this issue of cross-layer design in wireless networks, and more particularly with the impact of node-level multiuser detection on such design. This is currently a very active research area, and the intention of this work is to provide an introduction to this area, and to present some of the principal methods developed and results obtained to date.

This work is intended for engineers, researchers and students with some prior exposure to the field of communication networks. Although the book is largely self-contained and presents necessary background on wireless networking and multiuser detection, it is not intended to provide a complete treatment of these subjects. However, an extensive bibliography is included to direct the reader to additional details on these subjects as desired.

---

[1] An account of some of this work can be found in the recent book, *Wireless Communication Systems: Advanced Techniques for Signal Reception*, by Xiaodong Wang and H. Vincent Poor (Prentice-Hall: Upper Saddle River, NJ, 2004).

Acknowledgments

Chapter 1

MULTIUSER DETECTION FOR WIRELESS NETWORKS

1. Future Generation Wireless Networks

Future generations of wireless networks will enable heterogeneous services with a variety of data rates that may even reach up to the order of a gigabit per second. One of the strongest motivations for supporting traffic heterogeneity and high speed data rates is the enormous popularity and societal impact of wireline Internet enabled applications. Since the appearance of the desktop computer, two separate evolutionary paths have been emerging: on one hand the laptop and palmtop have become extremely popular as their users enjoy the freedom of being untethered, but on the other hand, the advantages of networking have become increasingly important as users want to maintain connectivity [Goodman, 2000]. Wireless Internet is the answer to merging these seemingly disparate requirements. Indeed, the convergence of computing and wireless communications, in the form of smart phones and similar devices, is the leading trend in these fields. Furthermore, wireless data services are becoming increasingly popular worldwide, with the current reported number of subscribers for third generation (3G) cellular services increasing from 70 million in September 2003, to over 128 million at the end of July 2004 (www.3gtoday.com). Moreover, more than 7 million worldwide subscribers to WiMax wireless broadband services are expected by 2009 (www.wi-fitechnology.com).

For the North American market, WiFi hotspots are becoming widespread, while 3G cellular networks are just now being deployed and are available only for a few regions. A recently emerging trend for commercial data services is to integrate cellular and WiFi, with companies in

1

North America [Brewin, 2004] and Japan leading the way by launching converged WiFi/cellular handhelds and bundled data services.

To support the widespread use of high speed wireless data services for future generation wireless networks, a key element is to reduce the cost of wireless transmission in terms of the actual price per Mbyte, as well as in terms of the amount of required transmission power.

In the following, we will summarize several network solutions that have been proposed to support wireless data services, and we will discuss how the cost of data transmission is influenced by each of these network designs.

1.1   Third Generation (3G) Cellular Networks

The third generation cellular networks currently being deployed are required to provide ubiquitous coverage for heterogeneous applications with varied quality of service (QoS) requirements (Fig. 1.1). This implies that 3G networks must support high data rate traffic in a highly bursty environment.

The wireless technology of choice for implementing 3G systems is code division multiple access (CDMA) due to its soft capacity characterization, which allows a graceful degradation of the network performance as demand increases, and due to its robustness to inter-cell interference which supports the powerful anytime/anywhere principle. Moreover, the nature of the CDMA air interface promotes statistical multiplexing of streams with varied bit error rates and delay requirements.

Both cdma2000 (www.tiaonline.org), developed primarily in North America, and wideband CDMA (WCDMA) [Holma and Toskala, 2002], developed primarily in Europe and Asia (www.3gpp.org), focus on providing high data rates to mobile users. The standard requirements specify a data rate of 384 Kb/s for outdoor devices moving at high speeds, and 2 Mbps for devices moving at pedestrian speeds. However, in reality, the achieved transmission rates depend on the prevalent channel conditions, and consequently, a rate adaptation technique is used. Many times, the high data rates are achieved at the expense of high power consumption and high costs for users. To reduce these transmission costs, 3G networks' capacity enhancements rely primarily on sophisticated resource management techniques, without imposing any improvements in the receiver design. As we will see in this book, multiuser receivers have the potential to increase the network capacity dramatically, thus having a significant impact on the effective price of wireless data.
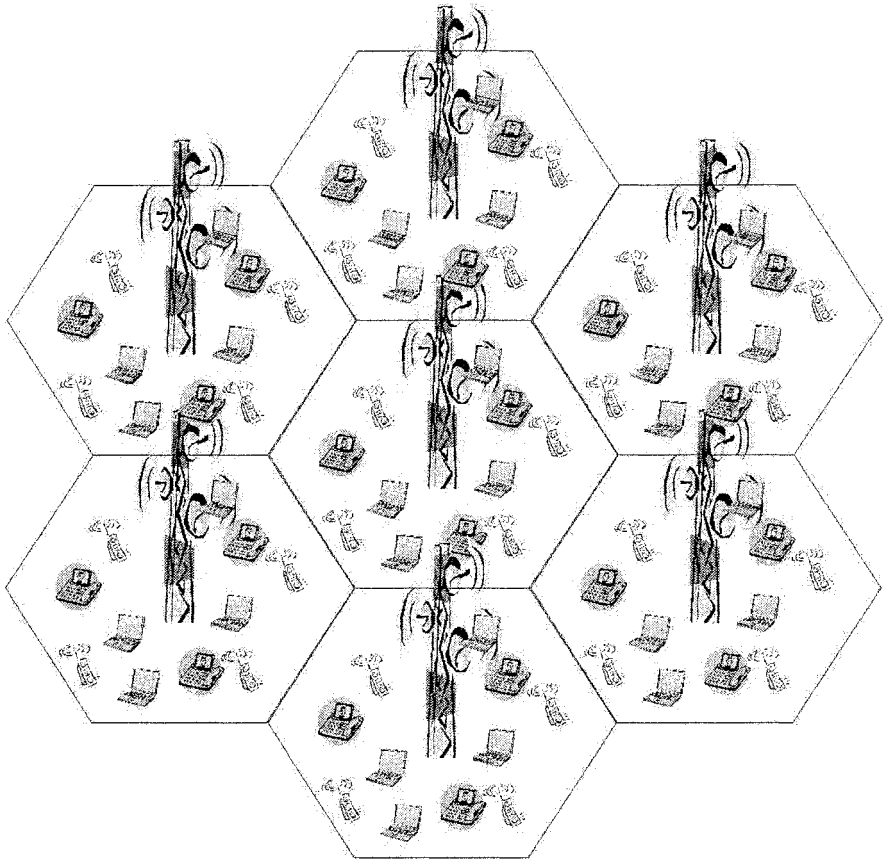
Figure 1.1. Heterogeneous applications and ubiquitous coverage in third generation cellular networks

1.2 Wireless Application Protocol (WAP)

One industry solution to provide low cost wireless Internet access to mobile users with cell phones concentrates on building a "cell phone centric Internet" using WAP (wireless application protocol). WAP is intended to be used for networks of handheld digital wireless devices such as mobile phones, pagers, two-way radios, and smart phones, and is suitable for basic applications such as accessing weather forecasts and

stock quotes, messaging, personal information management, financial services and location based services.

WAP uses high compression for data and improves the cell phone user interface using the WML (Wireless Markup Language) to display text and icons on a cell phone's screen. Despite the advantage of providing an immediate solution for the wireless Internet, it has an inherent, very significant, disadvantage: the "cell phone centric Internet" is not the real World Wide Web; its content is subject to the availability of wireless Internet Web pages for the desired targeted sites. Most of the "cell phone centric Internet" is constructed and managed by the cellular operating companies.

Thus, WAP provides only a partial and interim solution for data wireless networks. While it is useful in a transitional phase, next generation wireless networks must commit to genuine information connectivity.

1.3   Network Costs for Data Transmission

Although at first glance 3G networks seem to be on the right track for providing ubiquitous connectivity, the price per Mbyte may be too high for the successful proliferation of Internet services on such networks. The cost per Mbyte is influenced by the overall cost of the system (C_{system}). For uniform coverage with QoS guarantees, a general system cost formula [Zander, 2001] can be expressed as

$$C_{system} \approx cN_{AP} \approx cK_{user}B_{user}A_{service}f(Q), \qquad (1.1)$$

where N_{AP} is the number of access points (base stations) required to provide services and c is a proportionality constant. The effective bandwidth B_{user} required per user with K_{user} subscribers over a service area $A_{service}$ must be scaled by an overprovisioning factor $f(Q)$ for QoS delivery to high rate data users.

The factor $f(Q)$ can be greatly reduced by efficient access control algorithms relying on statistical traffic multiplexing. It is also immediately apparent from (1.1) that, for a fixed number of users, the system cost is strongly influenced by the effective bandwidths of the users, for a given service coverage area.

It is evident that reducing the effective bandwidth for high rate users will result in a cost reduction for Internet services. While the WAP solution is based on decreasing the bandwidth requirement for the applications (basic applications and higher compression), improvements for third generation cellular technology can be achieved using multiuser receivers for CDMA systems. As we will see in the next section, the capacity improvements achieved by multiuser detectors come at the cost of significant implementation complexity. This complexity has prevented

the use of such receivers in previous cellular systems, which were primarily designed for voice telephony. However, with the emergence of new high speed applications and the rapid increase in the processing speeds of low power, low cost digital signal processing (DSP) devices and integrated circuits, multiuser detection should become an attractive choice for next generation wireless networks.

1.4 Wireless Networks for Unlicensed Bands: WiFi, WiMax, HomeRF, Bluetooth and Infostations

The deployment of wireless data networks in unlicensed bands is ideal for data users who can freely use the spectrum without the need to obtain a license for it. Operating in unlicensed bands can significantly reduce the cost of wireless data, by reducing the implementation price floor related to spectrum acquisition.

In response to different application requirements, several types of networks have emerged in the unlicensed spectrum, such as WiFi, WiMax, HomeRF, Bluetooth, and infostations. In general, all these networks are based either on a star configuration, i.e., there is an access point to which all portable terminals transmit in a single-hop fashion, or they use a peer-to-peer topology that facilitates the deployment of on-the-fly ad hoc networks with multi-hop transmissions[1]. In this section, we discuss some of the key technologies in this category.

WiFi

While high data rate adoption is trailing for 3G cellular networks in North America, the use of wireless local area networks (LANs) for nomadic computing is growing dramatically. Because of this increasing popularity of local network wireless access, hot spot access points are becoming available to users in a variety of commercial areas such as airports, hotel lobbies, coffee shops, book shops, etc.

Wireless LANs are intended for low mobility and stationary users, and have a relatively small coverage area (e.g., a room, a floor, etc.) The name WiFi stands for "wireless fidelity" (similar to HiFi for "high fidelity" in audio systems), and it refers to the fact that wireless LANs were originally targeted primarily at office use requiring high quality transmission.

---

[1] Ad hoc networks will be discussed in more detail in Section 1.5.

Commercially available wireless LANs (WLANs) are based on the IEEE 802.11 family of wireless Ethernet standards, which has several different variants:

- IEEE 802.11a radios transmit at 5 GHz and send data up to 54 Mbps using OFDM (Orthogonal Frequency Division Multiplexing);

- IEEE 802.11b radios transmit at 2.4 GHz and send data up to 11 Mbps using direct sequence spread spectrum modulation;

- IEEE 802.11g is an extension to IEEE 802.11b , with enhanced data rate transmission of up to 54 Mbps within the 2.4 GHz band using OFDM technology.

IEEE 802.11g maintains backward compatibility with IEEE 802.11b at 11 Mbps, while IEEE 802.11a is not interoperable with either IEEE 802.11b or IEEE 802.11g systems.

Wireless LANs can be configured either in a star topology with one access point and portable units transmitting to the access point, or in a peer-to-peer architecture. The latter option is not widely used and appears to have relatively poor performance [Xu and Saadawi, 2001].

Wireless Metropolitan Area Networks

Although IEEE 802.11 based wireless network implementations are very popular for wireless LAN access, a wider area network implementation such as a MAN (Metropolitan Area Network) is difficult to implement with this technology, since IEEE 802.11 has performance limitations for large numbers of users with high bandwidth requirements. In addition, interference is often a significant problem in IEEE 802.11 networks if deployed for large coverage areas, due to the fact that they operate in unlicensed bands.

A solution for wireless MAN implementation is the recently proposed IEEE 802.16 family of standards [IEEE 802.16 Working Group, 2004] which offers a high speed/capacity, low cost, and scalable solution for fiber optic backbone extention. IEEE 802.16 supports point-to-multipoint architectures in the 10-66 GHz range, with data rates up to 120 Mbps. At these frequencies, transmission requires a direct line of sight between the transmitter and receiver. However, non-line-of-site access provisioning at lower frequencies has been proposed in a recent version of the standard: IEEE 802.16a, which also includes support for a mesh architecture, and which operates in both licensed and unlicensed bands between 2GHz and 11GHz, using OFDM.

The IEEE 802.16 [IEEE 802.16 Working Group, 2004] family of standards has a series of very desirable properties such as: support for mul-

tiple services simultaneously with QoS provisioning, bandwidth on demand with spectrum efficient MAC design, and link adaptation (adaptive modulation and coding). The standard also supports the use of adaptive antennas and space-time coding for physical layer performance enhancement.

The technology integrates well with IEEE 802.11 wireless LANs and thus may be used in the future for linking 802.11 hot spots to the Internet via a wireless broadband connection. Moreover, it is a good candidate for home wireless broadband access. At this stage of development, the technology is still too expensive for consumers, but the prices are expected to fall dramatically as major industry players support the new technology. The forum that promotes and supports brodband wireless access networks based on the 802.16 standard is the WiMax forum [WiMax, 2004].

HomeRF

As opposed to the WiFi technology, which was originally oriented towards the corporate user, HomeRF technology aims to provide a cheaper and lower quality (lower data speeds) wireless network technology in the home network environment. Home networks are envisioned to connect PCs, PDAs, laptops, cordless phones, smart appliances, etc., in and around the home. Home networks were promoted by the HomeRF working group which ceased activity in January 2003, after finalizing a standard called Shared Wireless Access Protocol (SWAP).

SWAP is a hybrid standard that supports both voice and data, and interoperates with both the PSTN (Public Switched Telephone Network) and the Internet. The voice support is based on the Digital Enhanced Cordless Telecommunications (DECT) standard, while the data support relies on the IEEE 802.11 wireless Ethernet specification. SWAP supports streaming services (voice and video) via a centralized network controller, as well as ad hoc peer-to-peer transmission for data services. SWAP devices use frequency hopping spread spectrum technology with 50 hops per second and transmit at about 1 Mb/s. Some manufacturers allow for an increase in the transmission speed up to 2 Mb/s when little interference is present. The range of a HomeRF network covers a typical home and backyard (about 75 to 125 feet). The future of such networks is uncertain in view of the increasing popularity of WiFi systems for home use.

Bluetooth

Bluetooth has primarily been proposed as a technology for cable replacement in personal area networks. It is a low cost, low power, short

range wireless link intended as an alternative to IrDA (Infrared Data Association)[Infrared Data Association, 2004], which is based on infrared light pulses and consequently requires direct line of sight between transmitter and receiver.

The Bluetooth standard allows small, inexpensive radio chips (under \$5) to be integrated into many electronic devices (e.g., computers, printers, mobile phones, etc.). Devices that are Bluetooth enabled detect each other independently (without any user intervention) and form a pico-network, within a typical range of 10 meters (using 1 mW of transmit power). The piconet is a star network, with one node acting as master controlling the transmission of the others. The master node synchronizes and schedules the transmissions for all the other nodes. Similar to HomeRF and WiFi, Bluetooth operates in the 2.4 GHz unlicensed band. The physical layer interface is based on frequency hopping spread spectrum, and it supports one data channel at 721 Kb/s and up to three voice channels at 56 Kb/s. Since the standard provides only for low rate transmission and supports only very short range transmissions, Bluetooth is not a technology replacement for either WiFi or HomeRF for wireless LAN implementation.

Infostations

As a hybrid architecture between cellular networks and wireless LANs, the infostation paradigm [Frenkiel and Imielinski, 1996, Frenkiel et al., 2000], abandons the anytime/anywhere requirement, replacing it with a more affordable "many-time/many-where" philosophy, and promises to deliver data inexpensively ("free bits") for high data rate users. This network concept can reduce the cost of providing high-rate data by decreasing the effective bandwidth allocated for high rate users, as a result of using only very good channels in the proximity of access points. Unlike WiFi, HomeRF, and Bluetooth, the infostation concept is not yet implemented in a commercially available system.

The conceptually simple idea behind infostations is based on the well known fact that optimal use of a collection of channels is achieved by waterfilling solutions, in which more power is transmitted on the better channels [Cover and Thomas, 1991], as opposed to transmitting more power when the channel is worse, as is the case for 3G systems. For time-varying fading channels, the optimality of waterfilling in time was verified in [Goldsmith and Varaiya, 1997], and this result can also be extrapolated to channels whose quality variations are due to distance based path loss. Infostations are systems designed to optimize throughput, without the constraint of anytime/anywhere coverage, and thus will have pockets of very high rate coverage and large areas without any ser-

vice. An infostation is a source of information providing low power, very high data rate Internet access to portable devices in a limited surrounding area, similar to a hot spot in a WiFi network.

An example of a potential infostation location is in an airport: an infostation can be located for example at an X-ray machine in an airport security area, so that useful information such as maps and attractions at the flight destination point, or recent e-mails and faxes, can be downloaded to a laptop computer that passes through the machine. Similarly, an infostation can be placed in a jetway corridor, and data generated during the flight can be uploaded, and pertinent local information, such as weather and traffic reports can be downloaded on arrival at an airport after a flight.

The airport example is characteristic of the categories of traffic that can be supported by infostations. Obviously, real time applications cannot be accomodated, and even for delay tolerant services there are several technical challenges that must be overcome. The restricted range of an infostation introduces problems of its own: a portable terminal may be in the range of an infostation for only a few seconds, which may not be enough for completion of a transfer. With very high-speed radios, the bottleneck in information transfer in this architecture would be the organization and transfer of the information from the Internet to the infostation in a timely manner. It is likely that infostations would be located in a cellular service area, which may support the infostation network by providing location updates to the backbone wireline network, which in turn will select the next infostation to receive the requested information for resuming file transfer (Fig. 1.2 [Goodman, 2000]). The cooperation between these two heterogeneous networks, as well as the performance of such two-tier systems [Kishore et al., 2003, Ortigoza-Guerrero and Aghavami, 2000] offer several challenging technical problems still requiring solutions. To help relieve the problems associated with the information transfer it is very likely that local and general interest information would be cached at the infostation site. Examples of location dependent information are local area maps, restaurant locations, traffic and weather reports, etc. General interest information might include stock quotes, electronic news, and popular music recordings.

The implementation of infostations can be built upon the current commercially available short range technologies such as IEEE 802.11 wireless LANs [Crow et al., 1997], the Bluetooth technology [Bhagwat, 2001], or the emerging ultrawideband (UWB) technology [Win and Scholtz, 2000]. The characterization and modeling of the channels for such short range communication scenarios is an active area of research [Domaze-
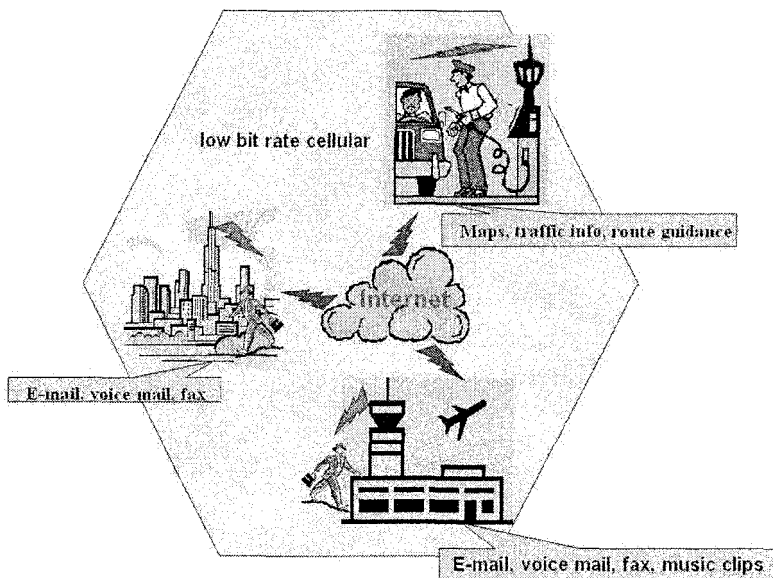
Figure 1.2.   Illustration of the infostation concept

tovic et al., 2002], as are a number of other aspects of the infostation concept.

1.5   Ad Hoc Networks

An even more forward-looking solution for next generation wireless networks, which completely reverses the cellular model, is the ad hoc network architecture.

An ad hoc network is defined as a collection of wireless terminals that self-configure to form a network without relying on a pre-existing infrastructure. Cost reduction in such networks is achieved by lowering the system price floor related to the infrastructure costs (base stations and auctioned spectrum), and also by their inherent multi-hop capacity increase potential. More specifically, ad hoc networks allow for peer-to-peer communication, as well as multihop connections, which have been shown to improve performance in both cellular (multihop routing to the base station)[Jabbari and Zadeh, 2001] and ad hoc network settings. As such, it has been shown that the coverage and capacity of ad hoc networks (measured in bit-meters/sec) increases with the increase in the number of users N. Several studies in the literature have been dedicated to quantify this capacity increase under various scenarios (see for exam-

ple [Shepard, 1995, Gupta and Kumar, 2000, Toumpis and Goldsmith, 2001, Grossglauser and Tse, 2002, Perevalov and Blum, 2003, Comaniciu and Poor, 2004c, Gupta and Kumar, 2003, Bansal and Liu, 2003]. Early work of [Gupta and Kumar, 2000] has shown that a capacity increase in the order of $O(\sqrt{N})$ is achieved for random access, two dimensional fixed ad hoc networks. While this is a rather pessimistic result since the per node throughput will decrease as $O(1/\sqrt{N})$, several papers in the literature have shown that some form of multiuser diversity may be exploited in ad hoc networks to increase capacity. In [Grossglauser and Tse, 2002, Gupta and Das, 2001, Perevalov and Blum, 2003, Bansal and Liu, 2003] mobility of the nodes is exploited to improve the capacity at the expense of very large to moderate transmission delays. In [Comaniciu and Poor, 2004c, Gupta and Kumar, 2003], signal processing based solutions for improved spectral efficiency are used to increase the network performance. As we will discuss in more detail later on in the book, the work in [Comaniciu and Poor, 2004c] shows significant user capacity increase for given network delay constraints in CDMA ad hoc networks using multiuser receivers.

An information theoretic result in [Gupta and Kumar, 2003] shows that a capacity in the order of $O(N)$ may be achieved for certain classes of networks, and gives a constructive example of achieving $O(N)$ capacity for an ad hoc network using a multiple transmit-receive antenna architecture.

The impact of all these studies is that if signal processing[2], combined with smart resource management techniques (e.g. power control, scheduling and routing) can drive the ad hoc network capacity close to $O(N)$, then each new user can support itself and the spectrum becomes essentially free. While this represents only a theoretical performance benchmark, it provides a strong economic motivation for investigation of high capacity ad hoc networks as possible future generation wireless data network solutions.

Although, by definition, ad hoc networks do not require any backbone infrastructure, they may potentially benefit from establishing a node hierarchy, which can improve their performance. However, in contrast with the cellular scenario, such a hierarchy is not a design requirement for ad hoc networks.

The lack of infrastructure in ad hoc networks requires new technologies for mobility management, service discovery and energy efficient in-

---

[2]In this book, our focus is on signal processing in the form of multiuser detection.
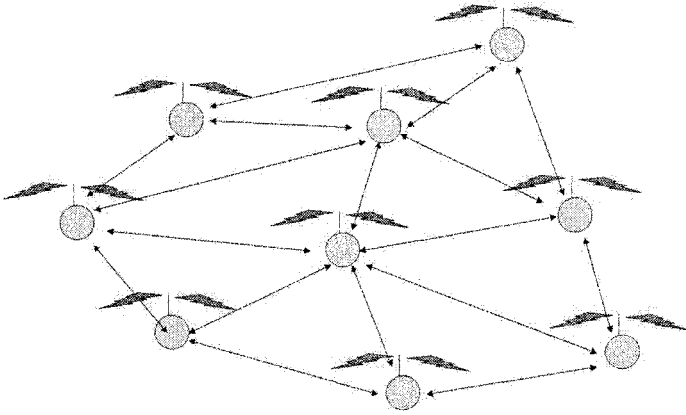
Figure 1.3. Ad hoc network illustration

formation routing, and poses design challenges at all layers of the protocol stack.

Ad hoc networks have the advantage of low cost deployment and they can be easily tailored for specific applications. They are suitable for a large array of applications [Goldsmith and Wicker, 2002] such as data networks, home networks [Lansford et al., 2000], device networks (Bluetooth [Bhagwat, 2001]), sensor networks [Akyildiz et al., 2002], etc. A large range of application-dependent network requirements must be met regardless of the network topology, the link quality at each local node and the node traffic. Moreover, the nodes usually have stringent energy constraints as well. Significant research has been directed towards implementing application-dependent QoS requirements in variable network conditions, and has specifically addressed power control, coding, adaptive techniques at the link layer, scheduling at the MAC (medium access control) layer and energy and delay constrained routing at the network layer. Although most of this research has concentrated on the layered protocol approach and has proposed adaptive and distributed techniques for the particularly considered layer, recent work shows that significant performance improvement can be achieved by considering cross-layer design in ad hoc networks (e.g. [Bertocchi et al., 2003, Cruz and Santhanam, 2003, Goldsmith and Wicker, 2002, Jabbari et al., 2002a]).

The use of a DS-CDMA (direct sequence CDMA) air interface for ad hoc network implementation would have many desirable advantages such as high capacity, low probability of intercept and robust performance in

narrowband interference (particularly attractive for unlicenced bands) and fading. As will become clear in the next section, the use of multiuser receivers may be especially beneficial for CDMA ad hoc networks for which tight power control may be difficult to implement. Further, the use of CDMA transmitter optimization could also alleviate the near/far problem and consequently, significantly increase the network capacity.

1.6   Cross-Layer Design

To summarize the above, several diverse solutions have been proposed for next generation wireless networks. A question that remains to be answered is: should fourth generation (4G) networks be application-specific or should they be designed to be flexible enough so that they will be able to support a large array of applications? Most of the network architectures that we have presented in Section 1.4 are application specific, and provide support for limited mobility. The infostation paradigm extends the mobility support by conceptually implementing a wireless LAN with roaming. Also, mobility extensions for WiMax are under consideration in the IEEE 802.16e version for the MAN standard.

Nevertheless, some of the data networks discussed previously are not suitable for real time applications with mobility (e.g. emergency communications and real-time interactive services such as interactive video and browsing, or voice calls). Most probably, these services will still be deployed in cellular type networks, or maybe in ad hoc networks, or a combination between the two architectures (e.g. [Jabbari et al., 2002b]). While radio resource management remains a key component in such networks, further significant performance gains may be obtained in CDMA based networks by employing multiuser receivers.

In this book, we focus on the current design approaches and state-of-the art analytical tools for wireless CDMA networks that use multiuser detection in cross-layer design; that is, design that simultaneously considers the requirements of multiple network layers.

Cross-layer design has recently captured the interest of the research community due to its possible performance advantages over the traditional layered network design approach. To ensure QoS delivery, adaptability to channel transmission conditions should be implemented at all layers of the protocol stack. A key question that arises is whether this adaptability should be implemented at each layer independently (Fig. 1.4), preserving the classical modular design approach of the Open Systems Interconnect (OSI) model, or the optimization should be jointly implemented over multiple layers of the protocol stack (Fig. 1.5).

This question has stirred some debate over the advantages and disadvantages of cross-layer design. The advantages of using a modular
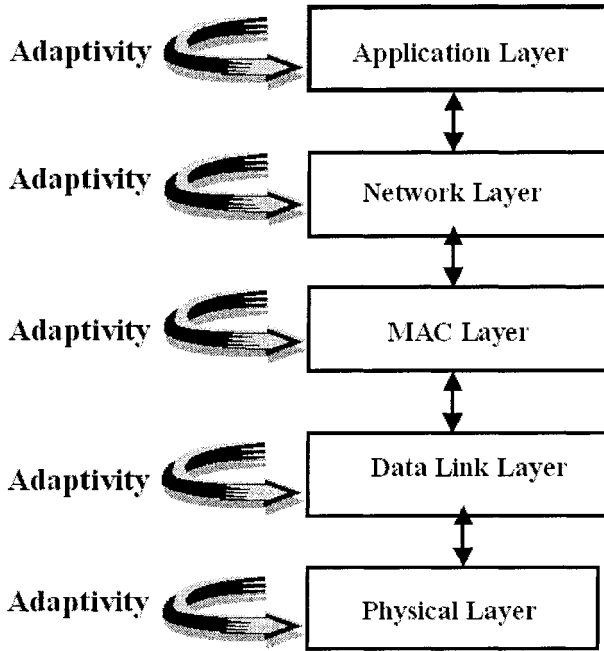
Figure 1.4. Adaptation at local layers in the OSI model

approach are increased flexibility in upgrading certain layers, easy de-
bugging, and low complexity. These are key properties that should be
preserved in a cross-layer design approach, to ensure that the short term
gains in performance and capacity can be transformed into long term
gains [Kawadia and Kumar, 2003], while considering cost, maintainabil-
ity and standardization [Shakkottai et al., 2003].

The advantages of a cross-layer design approach are direct conse-
quences of the nature of the wireless link itself. The wireless link charac-
teristics affect all levels of the network protocol stack, and therefore all
layers must be responsive to changing channel conditions. Furthermore,
tight coupling between protocols at different layers exists.

For example, at the physical layer, receiver filters can be dynamically
adjusted to respond to interference changes; at the link layer, power,
rate and coding can be adapted, again affecting the interference level;
at the MAC layer, adaptive scheduling can be implemented based on
the current level of interference and on the current link quality; adap-
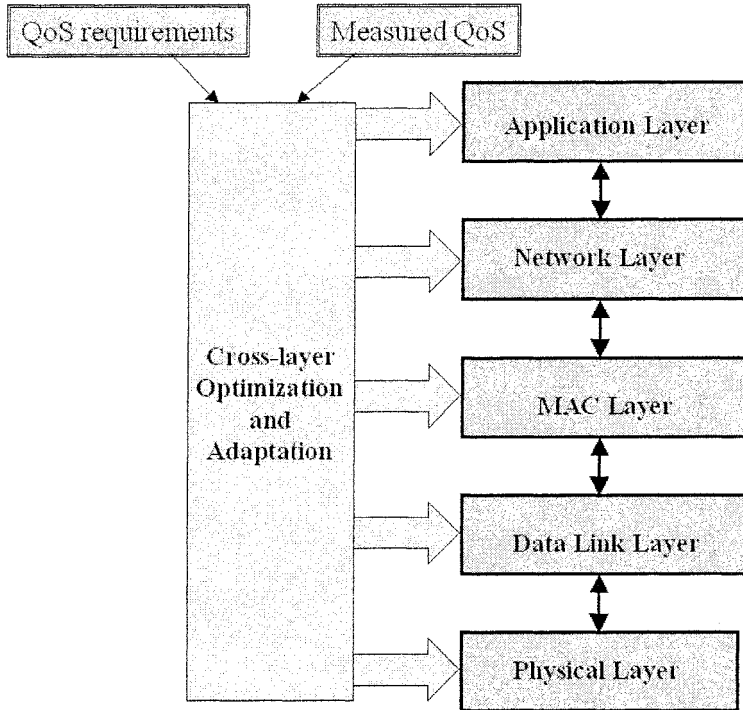
Figure 1.5. Cross-layer adaptation

tive routing (for ad hoc networks) or soft handoff (in cellular systems) can be implemented in response to the current interference level and distribution in the network; at the application layer, soft QoS can be defined, where the application QoS requirements are dynamically adjusted depending again on the current interference levels.

All the above adaptation protocols react to, and have an impact on, the interference level and distribution in the network. As a consequence, for efficient design, the adaptation protocols at each layer should not be independently developed, but rather should be designed in an integrated way, such that the interdependencies between layers can be exploited. Some extensively studied, classical examples of cases in which integration of different adaptation techniques at different layers is crucial for the performance of wireless networks, include the interaction between source and channel coding (e.g. [Aazhang et al., 1998]), and the interaction

between data link layer protocols and the transport control protocol (TCP) (e.g. [Kostic, 2001], [Harris et al., 2001]).

More recently, cross-layer design optimization of resource management algorithms has been proposed for various network scenarios and considering various performance measures (see for example [Alonso and Agusti, 2004, Cruz and Santhanam, 2003, ElBatt and Ephremides, 2004, Jabbari et al., 2002a, Jung and Vaidya, 2002, Radunovic and Boudec, 2002, Radunovic and Boudec, 2004] and the references therein). For ad hoc networks, energy efficient routing implies tight interdependencies among all layers of the protocol stack [Bertocchi et al., 2003, Cruz and Santhanam, 2003, Goldsmith and Wicker, 2002, Jabbari et al., 2002a].

The development of cross-layer protocols enhances the network's ability to adapt: performance information can be exchanged among layers for an optimal response to degrading transmission conditions. The integrated adaptive protocol must still have an hierarchical structure, since network variations take place on different time scales: for example, variations in the achieved link signal-to-interference ratio (SIR) are very fast, on the order of microseconds for high speed mobility, while variations in users' traffic are much slower, on the order of tens to hundreds of seconds [Goldsmith and Wicker, 2002]. The rate of adaptation for a protocol is determined by its location in the protocol stack. However, information exchange between layers and joint optimization may greatly improve the system performance.

Fundamental questions that must be answered in cross-layer design are: what information should be exchanged among layers, and how should such information be factored into each layer's performance adaptation algorithm [Goldsmith and Wicker, 2002]? In this book, we address these questions in the context of integrating the network and physical layer performance in wireless networks using multiuser receivers. We begin, in the following section, with an introduction to basic principles and results for multiuser detection, and with a general discussion of the tradeoffs involved in choosing the "right" receiver for next generation wireless networks.

2.   Introduction to Multiuser Receivers: Pros and Cons

In CDMA systems, the notion of capacity is directly related to the QoS perceived by the users. In general, a certain bit error rate (BER) target is required, which is application specific (e.g. 10^{-3} for voice users, and 10^{-6} or better for data applications). The achieved BER is directly related to the level of interference in the system, which thus dictates the system capacity. It immediately follows that any improvement in

the management and suppression of interference, greatly impacts the system capacity.

Matched filter (MF) receivers, in combination with tight power control and powerful coding have been shown to have reasonably good performance for second generation CDMA systems supporting only voice users. Indeed, under a white Gaussian noise model for the multiple-access interference (MAI), the matched filter receiver is optimal. However, this model is not accurate for wireless data systems, especially when the traffic is characterized by high burstiness as is the case with multimedia traffic. The structure of MAI can be exploited to build better receivers, which leads to the development of multiuser detectors. Better interference management will also certainly increase spectral efficiency, which of course is a desirable feature for all wireless networks.

In what follows, we address the following question: is multiuser detection (MUD) the right solution for future generation wireless networks? In order to answer this question, we start with several more basic questions: Why is multiuser detection superior to conventional, matched filter detection? What is the performance/complexity tradeoff for various MUD schemes? And, do we still need power control if multiuser receivers are used?

2.1 Performance of Matched Filter Receivers

Consider a single cell synchronous DS-CDMA system with K active users. The received signal at the base station in such a system can be expressed as [Verdú, 1998]

$$r(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + n(t), \tag{1.2}$$

where A_k, b_k, $s_k(t)$ are the received signal amplitude, the transmitted symbol and the signature waveform, respectively, of user k, and $n(t)$ is an additive white Gaussian noise (AWGN) process with power spectral density σ^2. For simplicity, we assume throughout that the symbols $\{b_k\}$ take binary ± 1 values, although other cases are readily treated. When the symbols are taken to be random, we assume that they are independent, taking the values ± 1 equiprobably.

For random signature sequences , the signature waveform $s_k(t)$ can be written as

$$s_k(t) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} s_{kj} \ p_{T_c}(t - jT_c + T_c), \tag{1.3}$$

where N is the spreading gain, s_{kj}, $k = 1,\ldots,K$, $j = 1,\ldots,N$ are independent equiprobable ± 1 random variables, T_c is the chip duration and p_{T_c} is the deterministic chip waveform, assumed to have unit energy. The vector $\mathbf{s}_k^T = [s_{k1}, s_{k2}, \ldots, s_{kN}]$ is the signature sequence of user k, and N is the spreading gain. The normalized cross-correlation between two users' signature waveforms over the bit duration T_b can be defined as

$$\rho_{k,\ell} = \int_0^{T_b} s_k(t)s_\ell(t)dt. \tag{1.4}$$

Equation (1.4) is equivalent to

$$\rho_{k,\ell} = \sum_{n=1}^{N} s_{kn}s_{\ell n} = \mathbf{s}_k^T \mathbf{s}_\ell. \tag{1.5}$$

For random signature sequences, it can be shown that $E\{\rho_{k,\ell}\} = 0$, and $E\{\rho_{k,\ell}^2\} = 1/N$ [Verdú, 1998].

Although a synchronous CDMA system is harder to implement in practice for the reverse link of a cellular system than for the forward link (it requires access to a common clock or closed-loop timing control), it is usually the model considered in theoretical analyses. This is due to the fact that the insights gained using a simplified analysis can be, in general, easily extended for a one shot analysis approach for asynchronous systems. In an asynchronous system, because of the time offsets among the reception of users' signals, one must take into account the fact that users transmit a frame or a stream of bits: $\mathbf{b}_k = [b_k[-M],\ldots,b_k[0],\ldots,\ b_k[M]]$. If we consider a one shot approach for detection, then for the symbol $b_k[0]$ of user k (the user of interest), an interfering user ℓ would affect the desired user partly by transmiting bit $b_\ell[-1]$ and partly by transmitting bit $b_\ell[0]$ (Fig. 1.6).

To characterize the influence of user ℓ on user k, equivalent virtual interfering users can be defined, having signature waveforms corresponding to the left $(s_\ell^l(t))$ and right $(s_\ell^r(t))$ signature waveform of user ℓ:

$$s_\ell^l(t) = \begin{cases} \frac{1}{\sqrt{\theta_\ell}}s_\ell(t + T_b - \tau_\ell), \ 0 \leq t \leq \tau_\ell \\ 0, \ \leq \tau_\ell \leq t \leq T_b; \end{cases} ; \tag{1.6}$$

$$s_\ell^r(t) = \begin{cases} \frac{1}{\sqrt{(1-\theta_\ell)}}s_\ell(t - \tau_\ell), \ \leq \tau_\ell \leq t \leq T_b \\ 0, \ 0 \leq t \leq \tau_\ell; \end{cases} , \tag{1.7}$$

where τ_ℓ is the time offset of user j relative to user k, and θ_ℓ is the partial energy of the ℓth interfering signal over the left overlapping bit.
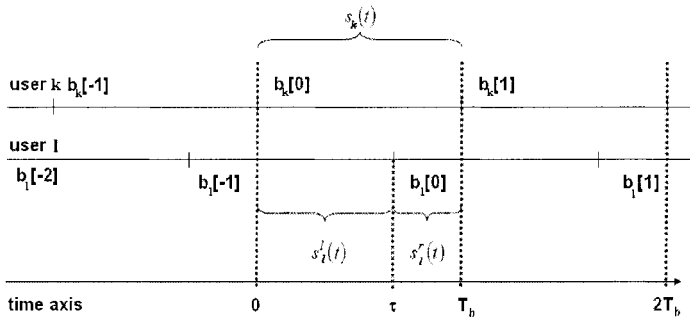
Figure 1.6. Asynchronous CDMA: basic model

Thus, a two-user asynchronous system can be viewed as a three-user synchronous system, and by generalization, a K user asynchronous system is equivalent to a synchronous one with $2K - 1$ users.

Due to this equivalence, many theoretical results developed for synchronous users can be readily adopted for asynchronous systems. With this in mind, in what follows we focus our presentation on synchronous systems, as described by (1.2).

The general multiuser detection problem is to determine the transmitted symbols for all users, given the received signal $r(t)$. In [Verdú, 1986] it was shown that the appropriately sampled outputs of filters matched to the various users' signature waveforms form a sufficient statistic for this decision problem. Given a symbol duration T_b, these matched filter outputs are given by

$$y_k = \int_0^{T_b} r(t)s_k(t)dt, \ k = 1, \ldots, K. \tag{1.8}$$

For the conventional matched filter detector, the decision is made by quantifying these outputs directly as

$$\hat{b}_k = \operatorname{sgn}(y_k), \ k = 1, \ldots, K. \tag{1.9}$$

where $\operatorname{sgn}(\cdot)$ denotes the algebraic sign of its argument.

Equation (1.9) represents the optimal decision for detection in the presence of white Gaussian noise only, under both maximum likelihood (ML) and maximum a posteriori probability (MAP). In a multiuser setting, the matched filter outputs also contain multiple access interference (MAI) components which are not white Gaussian random variables. The output of the matched filter for user k can be expressed as

$$y_k = \underbrace{A_k b_k}_{\text{desired signal}} + \underbrace{\sum_{\ell \neq k}^{K} \rho_{k,\ell} A_\ell b_\ell +}_{MAI} \underbrace{n_k}_{\text{Gaussian noise}} . \qquad (1.10)$$

where $n_k = \int_0^{T_b} n(t) s_k(t) dt$.

To illustrate the performance of the matched filter we consider a simple two-user example ($K = 2$) with $\rho_{1,2} = \rho_{2,1} = \rho$. Without loss of generality, consider the bit error rate for user 1, which can be expressed as

$$P_e^{mf} = P(\hat{b}_1 \neq b_1) = \tfrac{1}{2} \underbrace{P(y_1 < 0 | b_1 = +1)}_{P+} + \tfrac{1}{2} \underbrace{P(y_1 > 0 | b_1 = -1)}_{P-} .$$

By symmetry, $P^+ = P^-$, and thus, $P_e^{mf} = P^+ = P^-$. We derive P^- as follows

$$P^- = \tfrac{1}{2} P(y_1 > 0 | b_1 = -1, b_2 = +1) + \tfrac{1}{2} P(y_1 > 0 | b_1 = -1, b_2 = -1) =$$

$$= \tfrac{1}{2} P(n_1 > A_1 - A_2 \rho) + \tfrac{1}{2} P(n_1 > A_1 + A_2 \rho).$$

Since the signature waveforms are normalized, $n_1 \sim N(0, \sigma^2)$. Hence the probability of error for the conventional matched filter receiver is given by

$$P_e^{mf} = \frac{1}{2} Q\left(\frac{A_1 - A_2 |\rho|}{\sigma}\right) + \frac{1}{2} Q\left(\frac{A_1 + A_2 |\rho|}{\sigma}\right), \qquad (1.11)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$.

Since $Q(.)$ is monotonically decreasing, an upper bound on the probability of error is given by

$$P_e^{mf} \leq Q\left(\frac{A_1 - A_2 |\rho|}{\sigma}\right). \qquad (1.12)$$

The bound is smaller than $1/2$ if $\frac{A_2}{A_1} < \frac{1}{|\rho|}$, which is called the "open eye" condition (the interferer is not dominant). If the interferer is dominant, the conventional receiver exhibits the near/far problem: the error probability is not monotonic with the noise power, and a powerful interferer can completely obscure the reception of a less powerful user.

From the network performance point of view, the achievable power efficiency (the required signal to noise ratio for a given BER target) is of special interest. This can be best illustrated by using power tradeoff region diagrams, which represent the set of required signal to noise ratios

(SNRs) $\{A_1^2/\sigma^2, A_2^2/\sigma^2, \ldots, A_K^2/\sigma^2\}$, such that $\max_k P_k^\sigma \leq \zeta$, where P_k^σ is the error probability of user k, and ζ is a target bit error rate.

In Fig. 1.7 two-user power tradeoff regions are depicted for the matched filter receiver for different values of the cross-correlation coefficient between the users' signature sequences, and for a bit error rate requirement of $\zeta = 10^{-3}$. The upper bound in performance is achieved when the signature sequences are orthogonal, which is equivalent to the case of a single user in additive white Gaussian noise. For fixed cross-correlation, the best performance is obtained for equal powers. Note that, as the cross-correlation increases, the sensitivity to imbalances in the received powers increases as well, and also higher energies are required, even for the case of perfect power control $(A_1 = A_2)$.
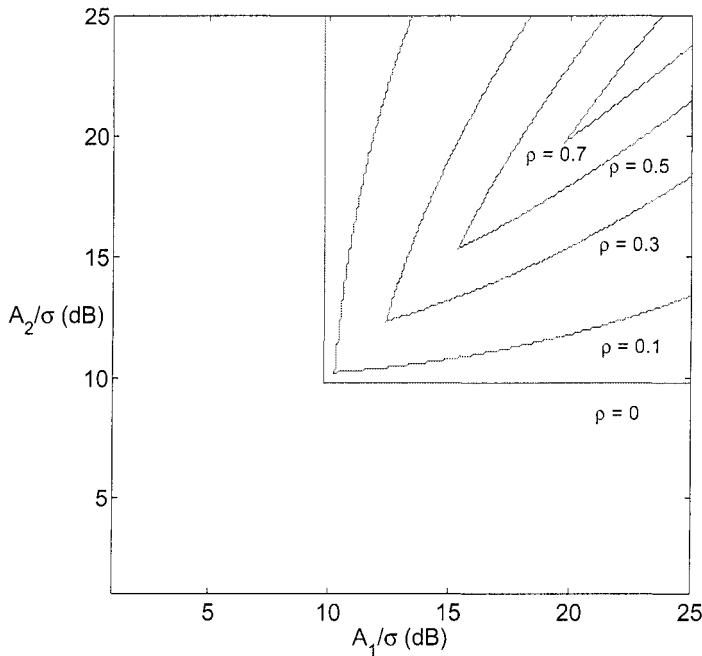
Figure 1.7. Power tradeoff regions for two users employing matched filter receivers

In early work on multiuser detection [Verdú, 1986], the near/far problem was shown to be a consequence of the inability of the matched filter to exploit the structure of the MAI, and not to be associated with CDMA in general. An optimal multiuser detector was proposed based on the maximum likelihood detection of the transmitted symbols.

Before analyzing this optimal receiver, we define some performance measures frequently used to quantify the performance of multiuser detectors.

For a given multiuser detector and background noise level σ, the *effective energy* of user k, $e_k(\sigma)$, is the energy that user k would require to achieve the same bit error rate $P_k(\sigma)$, in an equivalent single user Gaussian channel with the same noise level:

$$P_k(\sigma) = Q\left(\frac{\sqrt{e_k(\sigma)}}{\sigma}\right).\tag{1.13}$$

The *multiuser efficiency* represents the ratio between the effective and actual energies, $e_k(\sigma)/A_k^2$, and quantifies the performance loss due to other users in the channel. The *asymptotic multiuser efficiency* measures the slope with which $P_k(\sigma)$ goes to zero in the high SNR (signal to noise ratio) region:

$$\eta_k = \lim_{\sigma \to 0} e_k(\sigma)/A_k^2.\tag{1.14}$$

The *near/far resistance* represents the minimal multiuser efficiency, minimized over the received energies of all the other users:

$$\bar{\eta}_k = \inf_{A_\ell > 0, \ \ell \neq k} \eta_k.\tag{1.15}$$

2.2   Multiuser Detectors

An optimum detection rule can be based on maximum likelihood detection of the transmitted symbols. Let $\mathbf{y} = [y_1, y_2, \ldots, y_K]^T$ be the vector of matched filter outputs, where y_k is defined as in (1.10). The vector $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \mathbf{R}\mathbf{A}\mathbf{b} + \mathbf{n},\tag{1.16}$$

where $\mathbf{R}$ is the normalized cross-correlation matrix, with 1's on the main diagonal, and entries $\mathbf{R}_{k,\ell} = \rho_{k,\ell}$, $\mathbf{n}$ is a Gaussian noise vector with zero mean and covariance matrix equal to $\sigma^2 \mathbf{R}$, and $\mathbf{b}$ is the vector of information symbols. The matrix $\mathbf{A}$ is a diagonal matrix: $\mathbf{A} = diag[A_1, \ldots, A_K]$.

The likelihood function of $\mathbf{y}$ given $\mathbf{b}$ is given by

$$p(\mathbf{y}|\mathbf{b}) = \exp\left(\frac{-\frac{1}{2}(\mathbf{y} - \mathbf{R}\mathbf{A}\mathbf{b})^T(\sigma^2\mathbf{R})^{-1}(\mathbf{y} - \mathbf{R}\mathbf{A}\mathbf{b})}{(2\pi)^{K/2}\sigma|\mathbf{R}|^{1/2}}\right),$$

where $|\mathbf{R}|$ denotes the determinant of $\mathbf{R}$. The maximum likelihood symbol decisions are thus determined as $\hat{\mathbf{b}} = \arg\max_{\mathbf{b}} \Omega(\mathbf{b})$, with

$$\Omega(\mathbf{b}) = 2\mathbf{b}^T\mathbf{A}\mathbf{y} - \mathbf{b}^T\mathbf{H}\mathbf{b}; \quad \mathbf{H} = \mathbf{A}\mathbf{R}\mathbf{A}. \tag{1.17}$$

The above maximization problem is a combinatorial optimization problem which is known to be NP-hard: its computational complexity increases exponentially with the number of users in the system[3]. This $O(2^K)$ implementation complexity required by the optimal detector makes it impractical for real systems. The optimal detector represents, however, a basis for comparison for other, suboptimal, receivers.

In Fig. 1.8, the power-tradeoff regions for optimal multiuser detection receivers are shown for the same bit error rate probability target of 10^{-3} that was used to illustrate the matched filter case. We notice a very significant performance improvement compared with the matched filter case. Further, an interesting observation is that equal powers are detrimental for the optimum receiver, especially for high cross-correlation values. An intuitive explanation for this is that the receiver can better separate very similar users (with highly correlated sequences) if they are at least received with very different powers.

Although equal power control is not appropriate for this receiver, a minimal transmitted power solution can be achieved by implementing unequal power control. For example, for $\rho = 0.95$, equal power control leads to a requirement of $(18, 18)$ dB for the users, while the minimal power solution requires only $(10, 15)$ dB or $(15, 10)$ dB. It can thus be concluded that power control still helps to improve the system performance, even for the optimal receiver case. The difference from the

---

[3]Note however that, it was shown independentely in [Ulukus and Yates, 1998c] and [Ephremides and Sankaran, 1998] that, for synchronous systems and a specific choice of the signature sequences (i.e., having negative cross-correlations), an optimal multiuser detector can be implemented with polynomial complexity, $O(K^3)$. Also, it has been shown in [Schlegel and Grant, 2000] that, optimal multiuser detection for users with equal cross-correlations has a complexity of $O(K\log(K))$.
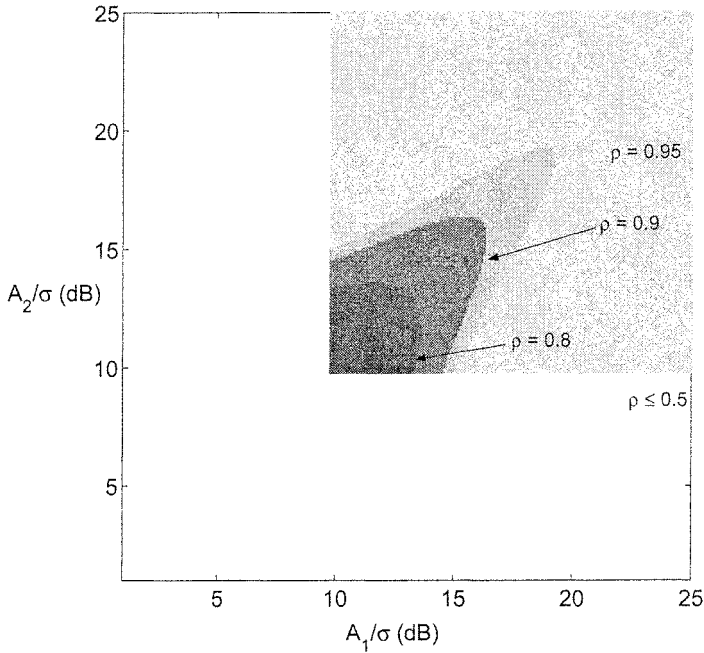
Figure 1.8.   Power tradeoff regions for two users employing optimal multiuser detection

conventional power control for matched filter systems is that, in this case, unequal power targets are used. The fact that the users can be treated interchangeably renders more potential gains for multicell systems: a user further away from the base station may be allocated the lowest received power, thus reducing the interference perceived by the neighboring cells. A more detailed discussion on this topic will follow in Chapter 3.

Numerous suboptimal approaches to multiuser detection have been proposed, to trade off performance and complexity. The most widely studied solutions can be classified into two categories: linear and non-linear multiuser detectors. For linear multiuser receivers, a linear transformation is applied to the vector of matched filter outputs, and a new, better decoupled, set of decision variables is produced, which can then be quantized to produce symbol decisions. The two most important linear receivers are the decorrelating detector [3] and the linear minimum mean-square error (LMMSE) detector [4]. Non-linear detection, also called subtractive detection, is based on estimating the interference and

removing it from the signal before detection. Examples of non-linear receivers are the succesive interference cancellation (SIC)[5, 6] and the parallel interference cancellation (PIC) receivers [7, 8]. This multiuser receiver classification is summarized in Fig. 1.9.

```
                    ┌──────────────┐
                    │  Multiuser   │
                    │  Receivers   │
                    └──────────────┘
      ┌──────────────────────┬──────────────────────┐
┌───────────────────────┐  ┌──────────────────────┐
│       Optimal         │  │     Suboptimal       │
│ Maximum Likelihood    │  │                      │
│     Receiver          │  │                      │
└───────────────────────┘  └──────────────────────┘
              ┌──────────────┬──────────────┐
        ┌──────────┐              ┌──────────────┐
        │  Linear  │              │  Non-linear  │
        └──────────┘              └──────────────┘
     ┌───────┬───────┐        ┌───────────┬───────────┐
┌──────────┐ ┌──────┐   ┌────────────┐ ┌────────────┐
│Decorrelator│ │MMSE│   │  Parallel  │ │Successive  │
│          │ │      │   │Interference│ │Interference│
│          │ │      │   │Cancellation│ │Cancellation│
│          │ │      │   │   (PIC)    │ │   (SIC)    │
└──────────┘ └──────┘   └────────────┘ └────────────┘
```
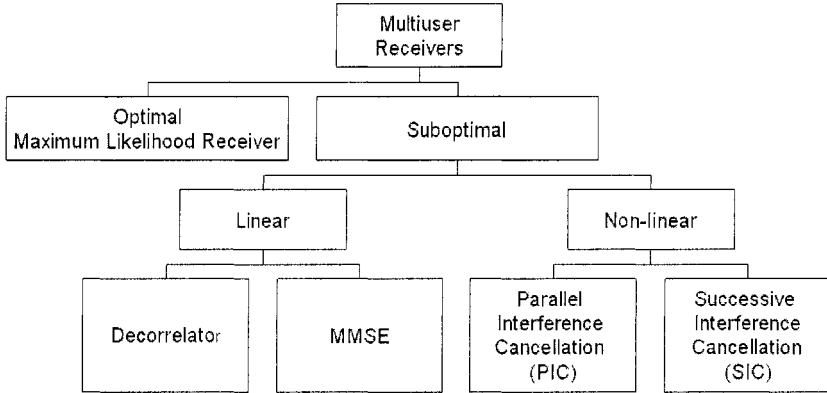
Figure 1.9. A classification of multiuser receivers

In the linear receiver category, the decorrelator [Lupas and Verdú, 1989] completely eliminates the multiple access interference by orthogonalizing the users. Starting from (1.16), if the linear transformation $\mathbf{R}^{-1}$ is applied to the outputs of the matched filters, the resulting decision vector is given by

$$\mathbf{y}^d = \mathbf{R}^{-1}\mathbf{y} = \mathbf{A}\mathbf{b} + \mathbf{R}^{-1}\mathbf{n}. \tag{1.18}$$

From (1.18) it can be immediately inferred that each component of the decision vector $\mathbf{y}^d$ is interference free. On the other hand, the background noise can be enhanced by the transformation $\mathbf{R}^{-1}$. The use of this detector requires that the set of signature sequences be linearly independent. Two advantages of the decorrelator are that it does not require knowledge of the received amplitudes, and it affords a decentralized implementation. Indeed, the output decision variable, y_k^d, for the kth user can be expressed as:

$$y_k^d = \sum_{\ell=1}^{K} \mathbf{R}_{k,\ell}^{+} y_\ell = \sum_{\ell=1}^{K} \mathbf{R}_{k,\ell}^{+} \mathbf{y}^T \mathbf{s}_\ell = \mathbf{y}^T \tilde{\mathbf{s}}_k$$

where $\mathbf{R}_{k,\ell}^+$ is the $(k,\ell)^{th}$ element of matrix $\mathbf{R}^{-1}$, and $\tilde{\mathbf{s}}_k = \sum_{\ell=1}^K \mathbf{R}_{k,\ell}^+ \mathbf{s}_\ell$. Thus, the decorrelator can be implemented as a modified matched filter (see Fig. 1.10) of the form

$$\widehat{\mathbf{b}}_k = \mathrm{sgn}\left\{ \int_0^{T_b} y(t)\tilde{s}_k(t)dt \right\}, \qquad (1.19)$$

where $\tilde{s}_k(t) = \sum_{\ell=1}^K \mathbf{R}_{k,\ell}^+ \tilde{s}_\ell(t)$.
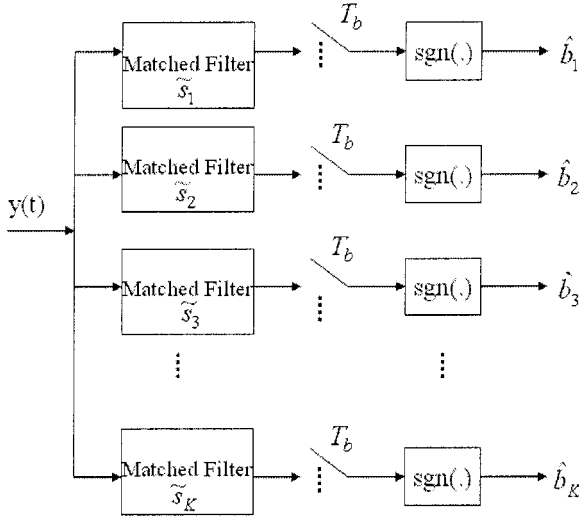
Figure 1.10.   Decorrelator implemented as a modified matched filter receiver

The decorrelating receiver is optimal (in the maximum-likelihood sense) when the amplitudes of the signals are not known. In this case, the decorrelating detector is obtained through joint maximum likelihood estimation of the transmitted symbols and amplitudes. Also, the decorrelating receiver achieves maximal near/far resistance: $\eta_k^d = 1/\mathbf{R}_{k,k}^+$. The probability of error achieved by an arbitrary user k using the decorrelator is given by

$$P_k^d = Q\left(\frac{A_k}{\sigma\sqrt{R_{k,k}^+}} \right). \qquad (1.20)$$

In contrast with the decorrelator, which is optimized to suppress the intereference, and the matched filter, which is optimized for noise suppression, the LMMSE receiver [Madhow and Honig, 1994] takes into

account the relative importance of both interference and background noise. The linear transformation for the LMMSE detector is given by the solution to the problem

$$\min_{\mathbf{C} \in \mathbb{R}^{K \times K}} E\left\{||\mathbf{b} - \mathbf{Cy}||^2\right\} = \min_{\mathbf{C} \in \mathbb{R}^{K \times K}} \text{trace}\{\text{cov}(\mathbf{b} - \mathbf{Cy})\}, \qquad (1.21)$$

where $\text{cov}(\mathbf{b} - \mathbf{Cy}) = E\left\{(\mathbf{b} - \mathbf{Cy})(\mathbf{b} - \mathbf{Cy})^T\right\}$. It is easily seen that $\mathbf{C}^* = A^{-1}[\mathbf{R} + \sigma^2 A^{-2}]^{-1}$ is the solution to (1.21), and thus the LMMSE decisions are given by

$$\widehat{b}_k = \text{sgn}\left(([\mathbf{R} + \sigma^2 \mathbf{A}^{-2}]^{-1}\mathbf{y})_k\right). \qquad (1.22)$$

The covariance matrix $\text{cov}(\mathbf{b} - \mathbf{C}^*\mathbf{y})$ can be expressed as

$$\text{cov}(\mathbf{b} - \mathbf{C}^*\mathbf{y}) = \left[\mathbf{I}_{K \times K} + \sigma^{-2}\mathbf{ARA}\right]^{-1}, \qquad (1.23)$$

and thus the achievable minimum mean square error is

$$MMSE = \text{trace}\left\{\text{cov}(\mathbf{b} - \mathbf{C}^*\mathbf{y})\right\} = \text{trace}\left\{\left[\mathbf{I}_{K \times K} + \sigma^{-2}\mathbf{ARA}\right]^{-1}\right\}. \qquad (1.24)$$

To illustrate the trade off between noise reduction and intererence supression, we analyze the LMMSE receiver for the kth user for two limiting cases: no interferers, and no noise, respectively

- No interference: Suppose $A_k \neq 0$ is fixed and $A_l \to 0$, $\forall l = 1, \ldots, K$, $l \neq k$; then the LMMSE receiver, $\left([\mathbf{R} + \sigma^2 A^{-2}]^{-1}\right)_{(row\ k)} \to \left[\frac{A_k^2}{A_k^2 + \sigma^2}, 0, ..., 0\right]$, becomes the matched filter receiver.

- No noise: Suppose $\sigma \to 0$; then the LMMSE receiver becomes the decorrelator

$$[\mathbf{R} + \sigma^2 A^{-2}]^{-1} \to \mathbf{R}^{-1}.$$

As a consequence of the fact that the LMMSE receiver becomes the decorrelator when $\sigma \to 0$, the LMMSE detector has the same asymptotic efficiency as the decorrelator. The exact error probability for the LMMSE receiver is very difficult to compute. However, a useful approximation for the bit error probability is [Poor and Verdú, 1997]

$$P_k^{mmse} \approx Q(\sqrt{SIR_k}), \qquad (1.25)$$

where SIR_k is the signal-to-interference-plus-noise ratio of user k at the output of the LMMSE transformation. Like (1.20), (1.25) depends on the number of users.

Note that, for a linear detector based on a transformation $\mathbf{Cy}$, we can write $(\mathbf{Cy})_k$ as $\mathbf{c}_k^T\mathbf{r}$, where $\mathbf{c}_k^T$ is an $N \times 1$ vector, and where $\mathbf{r}$ is the received vector:

$$\mathbf{r} = \sum_{k=1}^{K} A_k b_k \mathbf{s}_k + \sigma \mathbf{n}. \tag{1.26}$$

The vector c_k can be interpreted as a set of filter coefficients for detecting user k's data. It is straightforward to show that the LMMSE filter coefficients can be derived as

$$\mathbf{c}_k^* = \frac{A_k}{1 + A_k^2 \mathbf{s}_k^T \Sigma^{-1} \mathbf{s}_k} \Sigma^{-1} \mathbf{s}_k, \tag{1.27}$$

and the corresponding SIR expression becomes

$$SIR_k = A_k^2 \mathbf{s}_k^T \Sigma^{-1} \mathbf{s}_k, \tag{1.28}$$

where $\Sigma = \sigma^2 \mathbf{I}_{K \times K} + \sum_{\ell=1,\ell \neq k}^{K} A_\ell^2 \mathbf{s}_\ell \mathbf{s}_\ell^T$.

We can compare the linear multiuser receivers' performance with that of the matched filter and optimal receiver using the power tradeoff regions. In Fig. 1.11, we compare the achievable power efficiency for both the LMMSE and decorrelating receivers for a fixed BER requirement of 10^{-3}. To compute the power tradeoff regions for the LMMSE receiver, we use the approximate BER formula (1.25).

Analyzing Fig. 1.11 we notice that the decorrelating and LMMSE receivers have fairly close performance characteristics. However, the LMMSE receiver generally outperforms the decorrelator[4]. The difference is pronounced when the correlation coefficient is high, since the decorrelator's noise enhancement is more significant in this case. In terms of sensitivity to power control imbalances, the decorrelator is obviously unaffected by the interferers' powers. However, to achieve a certain target SIR, a minimum transmitted power for the desired user is necessary to surpass the enhanced noise power. Therefore, power control might be useful in systems using decorrelating receivers, in order to preserve the terminal battery and to reduce the interference seen by neighboring cells. The minimal power solution calls for equal received powers for all users. For the LMMSE case, we can also notice (and this will be proved in Chapter 3), that equalizing the received powers of all users can improve the system performance.

The performance of the linear receivers should be compared with that of the optimal receiver (Fig. 1.8) and the matched filter receiver (Fig.

---

[4]This is not uniformly the case however [Moustakides and Poor, 2001].
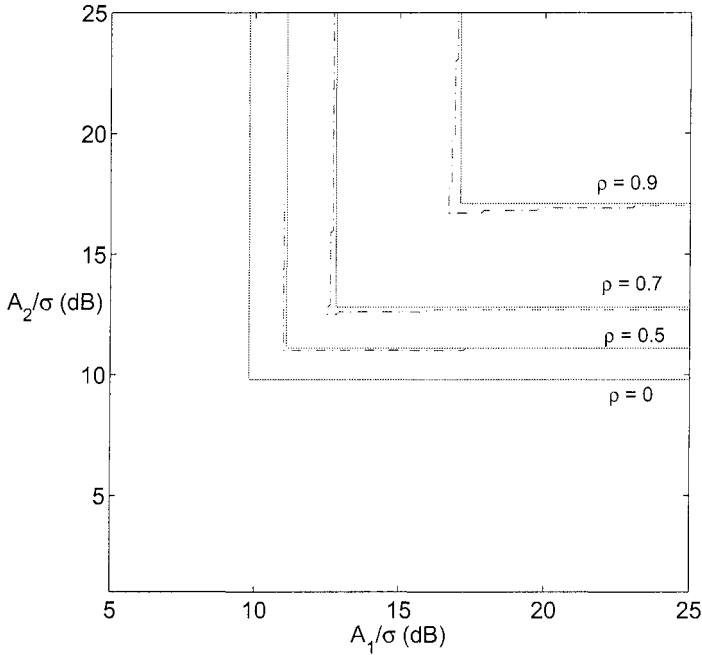
Figure 1.11. Power tradeoff regions for two users employing the decorrelating (solid line) and LMMSE (dash-dot line) receivers.

1.7). Notice that the linear filters and the optimal receiver have very similar performance for the equal received power case, which represents the best case for the suboptimal receivers, but a worst for the optimal receiver. However, a very significant gain in performance is achieved in general compared with the matched filter case.

In terms of complexity, both the decorrelating and LMMSE receivers require computing the inverse of a $K \times K$ matrix related to the correlation matrix of the signature sequences, which can be achieved with $O(K^3)$ complexity. Using a direct implementation, this matrix inverse must be computed each time a user changes activity (enters or exits the system, goes on and off temporarily, etc.). To decrease the computational complexity, rank one updates, and order updates can be used to compute the inverse of the correlation matrix or its Cholesky factorization. Based on these techniques, detector update algorithms having computational complexity of $O(K^2)$ have been proposed [Juntti, 1995].

Direct implementation of linear detectors requires knowledge of the cross-correlation matrix **R**. This can be avoided if the filters are imple-

mented adaptively, with as little knowledge as the timing of the desired user. However, some adaptive implementations require the use of training sequences, which in turn represents a waste of system bandwidth. As an alternative solution, blind adaptive algorithms have been proposed that require only the knowledge of a desired user's signature sequence and its timing information. Note that this is the same amount of information required for a matched filter implementation. This is particularly attractive for ad hoc networks, in which information about network signaling structure is inherently decentralized.

An additional advantage of an adaptive implementation in cellular systems is that a linear adaptive LMMSE detector can supress other-cell interference, in addition to the intra-cell interference. Furthermore, a coherent solution can constructively combine any multipath components falling within the filter's window span [Honig and Tsatsanis, 2000]. A disadvantage is that blind algorithms usually work only for systems with so-called short spreading codes, i.e., systems for which the same spreading code is used during every symbol interval. This type of spreading code is not used in existing 2G and 3G cellular telephony systems.

Because of the above mentioned advantages, considerable work has been done on designing low complexity blind adaptive algorithms with minimal performance loss compared to direct implementation. In practical systems, complexity, tracking and convergence for such algorithms are important performance measures.

Among the most representative blind adaptive algorithms, we mention the BADD (blind adaptive decorrelating detector) [Ulukus and Yates, 1998b], the MOE (minimum output energy) detector [Honig et al., 1995], and subspace tracking methods [Wang and Poor, 1998]. This list is by no means exhaustive; surveys of adaptive multiuser receivers can be found in [Madhow, 2000], [Honig and Tsatsanis, 2000], [Lim and Roy, 1998] and [Wang and Poor, 2004].

The BADD receiver is based on a stochastic convergence to the decorrelating receiver in the mean square sense. It requires knowledge only of the signature sequence of the desired user and the variance of the additive white Gaussian noise. Its complexity is $O(N)$ per iteration (Recall that N is the spreading gain). However, the performance of the detector depends on the adaptation step size, and for a small achievable mean-square error (MSE), the convergence rate is very slow.

The minimum output energy detector is an adaptive implementation of the LMMSE receiver, based on the observation that the mean square error is minimized when the variance at the output of the linear transformation is minimized within an "anchor" constraint on the filter coefficients. The adaptive algorithm uses a constrained optimization approach

and is very sensitive to possible desired-signal signature mismatch that could occur as a result of multipath fading and timing errors. To avoid complete cancellation of the desired signal, one solution proposed in [Honig et al., 1995] is to switch to a decision directed mode before the SIR is compromised. The implementation complexity of the MOE detector depends on the update algorithm. For least-mean-squares (LMS) updates the complexity is $O(N)$ per iteration, while for recursive least squares (RLS) updates, the complexity is $O(N^2)$ per iteration.

A more robust alternative to implementing blind LMMSE, as well as decorrelating receivers, is based on subspace tracking [Wang and Poor, 1998]. Both the LMMSE receiver and the decorrelator can be expressed in terms of signal subspace parameters which can then be tracked using the PASTd [Yang, 1995] or other subspace tracking algorithms. This approach also offers the capability of tracking the rank of the signal subspace, which is equivalent to tracking the number of active users. The complexity of this approach is $O(KN)$ per iteration, and it achieves better performance compared to the MOE receiver. Simulation results show slow convergence with random initialization; however, this can be improved if an SVD (singular value decomposition) is used for initialization.

To summarize the above discussion, although linear detection avoids the exponential complexity of optimal multiuser detection, there are still complexity issues associated with the implementation of linear multiuser receivers. Of particular importance for future wireless networks are the number of floating point operations (flops) and the required information for filter computation and updates, as well as the rate of convergence to the desired filter for the adaptive implementations. We summarize these important properties for the linear receivers discussed above in Tables 1.1 and 1.2

Table 1.1. Linear Receivers: Information Requirements

Req. Inf.	MF	Dec.	LMMSE	Adapt. LMMSE/Dec.	Blind LMMSE/Dec.
Code user	√	√	√	-	√
Code interf.	-	√	√	-	-
Timing user	√	√	√	√	√
Timing interf.	-	√	√	-	-
Rec. amplit.	-	-	√	-	-
Noise level	-	-	√	-	-/√
Training seq.	-	-	-	√	-

Table 1.2.   Linear Receivers: Implementation Complexity

Dec.	LMMSE	LMMSE/Dec. updates	Blind LMMSE/Dec.
$O(K^3)$	$O(K^3)$	$O(K^2)$	$O(N^2)/O(KN)/O(N)$

To achieve further complexity reduction, non-linear interference cancellation receivers have been proposed. Successive interference cancellation (SIC) receivers are among the least complex of multiuser detectors. At each stage of SIC, a single user is detected and removed from the overall received signal, so that the users detected at later stages see reduced multiple access interference. The current user is detected using its matched filter detector. Its signal is then respread and subtracted from the next user's received signal. The disadvantage of this scheme is that it is very sensitive to decision errors in previous stages. If a signal can be correctly reconstructed then it is completely removed from subsequent stages, but if there is an error, the interference is doubled. To improve SIC performance, the "hard" intermediate decisions can be replaced by "soft" decisions. This can be achieved by replacing the sign function in the matched filter receiver by a different nonlinearity. For example, a good option is to use the hyperbolic tangent function scaled with the SIR. This function is very similar to the sign function in the high SIR regime but discounts the effect of unreliable decisions in the low SIR regime. A simplified block diagram for an SIC detector is presented in Fig. 1.12.

An exact evaluation of the bit error rate of the SIC receiver is very difficult. Usually, the performance is measured using simulations, or is computed using approximations. A commonly used approximation assumes that the system performance is equivalent to that of a single user matched filter with a supplementary Gaussian noise source having zero mean and variance $1/N$ for each interferer [Verdú, 1998]. Consequently, the probability of error for user k can be approximated using the following recursive formula:

$$P_k^{SIC} \approx Q \left(\frac{A_k}{\sigma^2 + \frac{1}{N} \sum_{\ell=1}^{k-1} A_\ell^2 + \frac{4}{N} \sum_{\ell=k+1}^{K} A_\ell^2 P_\ell^{SIC}} \right). \qquad (1.29)$$

The expression in (1.29) enables the derivation of power tradeoff regions for appropriate bit error rate targets.
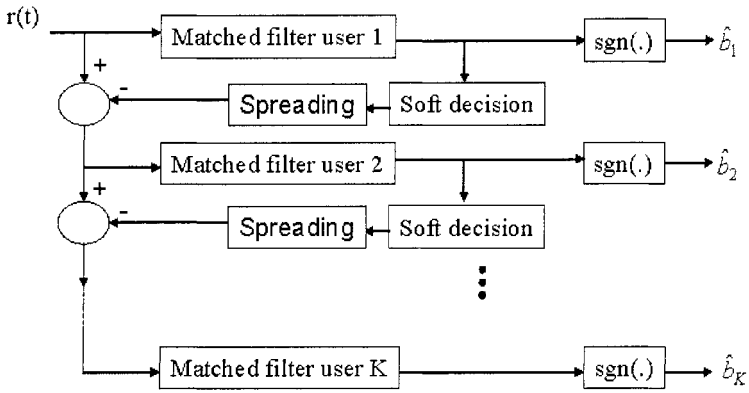
Figure 1.12. SIC block diagram
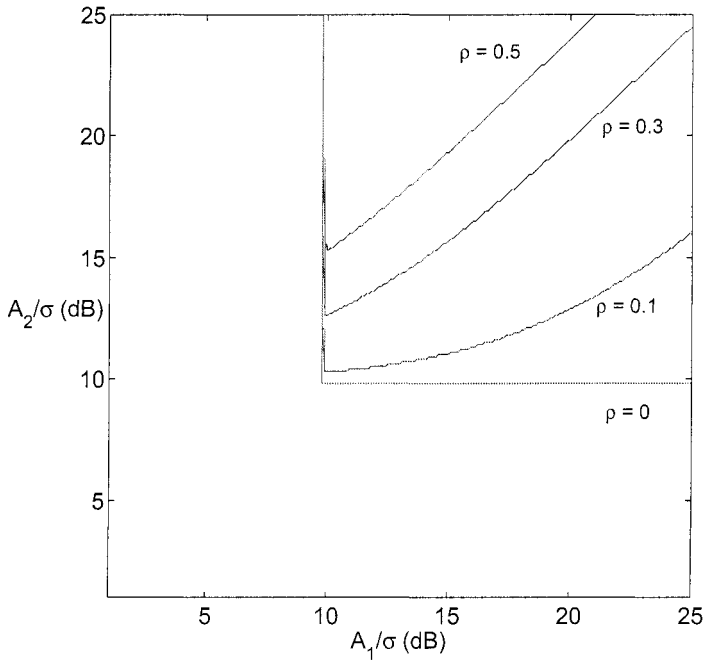
Figure 1.13. Power tradeoff regions for two users employing succesive interference cancellation detector

The power tradeoff regions are shown in Fig. 1.13 for a target bit-error rate of 10^{-3}. We first notice an asymmetric performance for the two users. This is an inherent property of the succesive interference cancellation scheme: while the user decoded first is strongly affected by the power of the interfering user (matched filter performance), the second user might have performance close to the single user system. Note that, in practical systems, the cancellation of the first user is not perfect, since it is strongly affected by the bit error rate and by amplitude and phase estimation errors. A more detailed discussion on this subject will follow in Chapter 3.

Due to the asymmetry in BER performance, equal power control represents the worst case for systems using SIC receivers. Nevertheless, SIC systems can benefit from unequal power control, such that users received with higher powers can be detected first. Although this represents a straightforward solution for the detection order, it completely neglects the effect of correlations between users. A better solution is to detect the users in the decreasing order of their energies at the output of the matched filter receivers [Patel and Holtzman, 1994a]. A more detailed discussion of the detection order for power controlled SIC receivers will follow in Chapter 3. Comparing Figs. 1.13, 1.11 and 1.7, we can see that SIC receivers (especially used in conjuction with unequal power control) outperform the matched filter receiver but have worse performance than that of the linear receivers. On the other hand, the reduction in complexity is substantial compared with their linear conterparts, although there is a penalty in increased detection delay. The implementation complexity and the detection delay for the SIC receiver grow linearly with the number of users. To reduce the detection delay, parallel cancellation (PIC) can be implemented, which detects all the users in parallel, then reconstructs the interference and subtracts it from the useful signal. A comparison between SIC and PIC receivers can be found in [Patel and Holtzman, 1994b].

To improve the performance of PIC, multiple stages can be implemented, which successively refine the estimates of the interfering symbols for progressively better cancellation. However, multistage PIC cannot guarantee improvements in performance with an increase in the number of stages, since incorrect decisions lead to further performance degradation. A solution to this problem has been proposed in [Divsalar et al., 1998]: only a partial cancellation of the MAI is implemented at each stage, with a weighting factor selected according to the level of confidence for the estimators.

Although parallel interference cancellers have lower detection delay than SIC receivers, they have a higher complexity than does SIC. How-

ever, if parallelism is exploited, the time complexity of PIC can be greatly reduced and the number of operations per processor become significantly less than that for SIC.

On a final note, we mention that blind successive interference cancellation receivers have been recently proposed [Samardzija et al., 2002], which require only the knowledge of the desired users' signature sequences, timing and power. As in the case of blind linear receivers, the approach here uses the sample covariance matrix of the received signal vector. Based on a maximum mean energy (MME) criterion, dominant interference components from the received signal are successively removed in a blind manner.

More sophisticated iterative multiuser detectors can be developed by exploiting structure in the symbol vector, improved either through prior information or through error control coding. Such detectors can span the performance gap between linear and optimal multiuser receivers, while maintaining relatively low complexity [Poor, 2004].

2.3   Performance of Blind Receivers

In this section, we discuss the performance of blind and group-blind multiuser receivers. Group-blind multiuser receivers are hybrid receivers built using knowledge of the spreading sequences for only a subset of the users in the system. A group-blind linear multiuser receiver zero-forces the interference caused by the known users, and suppresses the interference caused by the unknown ones using an MMSE criterion.

In what folows, we will summarize the performance analysis results presented in [Zhang and Wang, 2002b] for networks using blind LMMSE receivers. The analysis in [Zhang and Wang, 2002b] assumes that deterministic codes are used, and the receivers are estimated from the received signal samples. The number of received signal samples required for filter estimation is denoted by Υ.

For blind receivers, an estimate of the filter vector $\widehat{\mathbf{c}}_k$ for an arbitrary user k is obtained from the received signals $\{\mathbf{r}[n]\}_{n=1}^{\Upsilon}$ ($\mathbf{r}[n]$ is the nth received signal vector in a sequence of Υ samples used for estimation, with the received vector being defined as in (1.26)), such that the estimation error can be denoted as

$$\Delta\mathbf{c}_k = \widehat{\mathbf{c}}_k - \mathbf{c}_k, \tag{1.30}$$

and it is characterized by a covariance matrix $\mathbf{C}_c$.

The estimation error $\Delta\mathbf{c}_k$, as well as its covariance matrix, depend on the actual implementation of the blind receiver. In [Høst-Madsen and Wang, 2002] and [Zhang and Wang, 2002b], two different implementations are considered: the direct matrix inversion (DMI), and the

subspace approach [Wang and Poor, 1998]. It can be shown that the exact linear MMSE receiver can be computed as

$$\mathbf{c}_k = \mathbf{C}_r^{-1}\mathbf{s}_k \tag{1.31}$$

$$= \mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^T\mathbf{s}_k, \tag{1.32}$$

where $\mathbf{C}_r$ is the covariance matrix of the received vector, and $\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^T$ is the singular value decomposition of the covariance matrix $\mathbf{C}_r$. Based on (1.31), the DMI method replaces the exact covariance matrix of the received vector by

$$\widehat{\mathbf{C}}_r = \frac{1}{\Upsilon}\sum_{n=1}^{\Upsilon}\mathbf{r}[n]\mathbf{r}[n]^T, \tag{1.33}$$

such that

$$\widehat{\mathbf{c}}_k = \widehat{\mathbf{C}}_r^{-1}\mathbf{s}_k. \tag{1.34}$$

For the subspace method, the eigencomponents of the exact covariation matrix ($\mathbf{U}_s$ and $\mathbf{\Lambda}_s$) are replaced in (1.32) by the ones computed using $\widehat{\mathbf{C}}_r$, $\widehat{\mathbf{U}}_s$ and $\widehat{\mathbf{\Lambda}}_s$, such that

$$\widehat{\mathbf{c}}_k = \widehat{\mathbf{U}}_s\widehat{\mathbf{\Lambda}}_s^{-1}\widehat{\mathbf{U}}_s^T\mathbf{s}_k. \tag{1.35}$$

If the number of samples used for filter estimation is large, it has been shown that for a fixed number of users and fixed spreading gain, the output SIR for the blind LMMSE receiver can be approximated as

$$SIR^N(\widehat{\mathbf{c}}_k) = \frac{P_k(\mathbf{c}_k^T\mathbf{s}_k)^2}{\zeta_0 + \frac{N}{\Upsilon}(\zeta_1 + \zeta_2 + \zeta_3)}, \tag{1.36}$$

where

$$\zeta_0 = \sum_{\ell=2}^{K} P_k(\mathbf{c}_k^T\mathbf{s}_\ell)^2 + \sigma^2\|\mathbf{c}_k\|^2,$$

$$\zeta_1 = \frac{K+1}{N}\mathbf{c}_k^T\mathbf{s}_k,$$

$$\zeta_2 = -\frac{2}{N}\sum_{\ell=1}^{K} P_\ell^2(\mathbf{c}_k^T\mathbf{s}_\ell)^2\mathbf{c}_\ell^T\mathbf{s}_\ell,$$

and

$$\zeta_3 = \frac{N-K}{N}\tau\sigma^2,$$

with

$$\tau = \begin{cases} \mathbf{c}_k^T \mathbf{s}_k / \sigma^2, & \text{for DMI} \\ \sigma^2 \mathbf{s}_k^T \mathbf{U}_s \mathbf{\Lambda}_s^{-1} (\mathbf{\Lambda}_s - \sigma^2 \mathbf{I}_{K \times K})^{-2} \mathbf{U}_s^T \mathbf{s}_k, & \text{for subspace.} \end{cases}$$

More information is available for constructing the filter coefficients for the group-blind receiver: the signature sequences of $\tilde{K}$ users are assumed to be known, while those of the remaining $K - \tilde{K}$ users are unknown. It is assumed that $\tilde{\mathbf{S}}$ has full column rank, where the columns of $\tilde{\mathbf{S}}$ are the $\tilde{K}$ known signature sequences. Denoting by $\tilde{\mathbf{e}}_i$ the unit vectors in $\mathbb{R}^{\tilde{K}}$, it can be shown that the filter vector for an arbitrary user k can be computed as

$$\hat{\mathbf{c}}_k = \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s^{-1} \hat{\mathbf{U}}_s^T \tilde{\mathbf{S}} \left(\tilde{\mathbf{S}}^T \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s^{-1} \hat{\mathbf{U}}_s^T \tilde{\mathbf{S}} \right)^{-1} \tilde{\mathbf{e}}_k. \tag{1.37}$$

Using the following partitioning and notation,

$$\mathbf{S}^T \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^T \mathbf{S} = \begin{bmatrix} \mathbf{\Psi}_{11} & \mathbf{\Psi}_{12} \\ \mathbf{\Psi}_{12}^T & \mathbf{\Psi}_{22} \end{bmatrix},$$

$$\mathbf{\Xi} = [\mathbf{P}^{-1}(\mathbf{R} + \sigma^2 \mathbf{P}^{-1})^{-1} \mathbf{P}^{-1} \mathbf{R}^{-1} \mathbf{P}^{-1}]_{1:\tilde{K}, 1:\tilde{K}},$$

(where $\mathbf{R} = \mathbf{S}^T \mathbf{S}$, $\mathbf{P} = \text{diag}\{P_1, \,, ..., P_K\}$, and $\mathbf{\Psi}_{11}$ has dimension $\tilde{K} \times \tilde{K}$), it was shown in [Høst-Madsen and Wang, 2002] that the SIR for the group-blind receiver, can be approximated as

$$SIR^N(\hat{\mathbf{c}}_k) = \frac{P_k (\mathbf{c}_k^T \mathbf{s}_k)^2}{\zeta_0 + \frac{N}{T}(\zeta_1 + \zeta_2 + \zeta_3)}, \tag{1.38}$$

where

$$\zeta_0 = \sum_{j=1}^{K-\tilde{K}} P_{\tilde{K}+j} (\mathbf{c}_k^T \mathbf{s}_{\tilde{K}+j})^2 + \sigma^2 \|\mathbf{c}_k\|^2$$

$$\zeta_1 = \frac{K - \tilde{K}}{N} [\mathbf{\Psi}_{11}^{-1}]_{1,1}$$

$$\zeta_2 = -\frac{2}{N} \sum_{k=1}^{K-\tilde{K}} P_{\tilde{K}+k}^2 [\mathbf{\Psi}_{1,2}^T \mathbf{\Psi}_{11}^{-1}]_{k,1}^2 [\mathbf{\Psi}_{2,2} - \mathbf{\Psi}_{1,2}^T \mathbf{\Psi}_{1,1}^{-1} \mathbf{\Psi}_{1,2}]_{k,k}$$

and

$$\zeta_3 = \frac{N - K}{N} \sigma^2 [\mathbf{\Psi}_{11}^{-1} \mathbf{\Xi} \mathbf{\Psi}_{11}^{-1}]_{1,1}.$$

3.   Multiuser Detection for Next Generation Wireless Networks

As we have previously discussed in Section 1, future generation wireless networks face the challenge of providing low cost, high data rate transmission for a diverse population of users with a wide range of QoS specifications. While the emerging third generation standards have been successful in increasing the data transmission rate, this is still achieved at the expense of increased energy consumption and high bandwidth demands. One possible technique for increasing wireless network efficiency and consequently for decreasing the total system cost for providing high data rate services is to use multiuser receivers. So why have commercial systems not yet adopted this powerful technique? The short answer is implementation complexity. Furthermore, the performance advantages of multiuser receivers are reduced by their sensitivity to imperfections in channel tracking (tracking errors in frequency, phase, and timing).

As we will discuss in later chapters in the book, it has been shown that, although both linear and interference cancelling receivers are affected by imperfections in channel estimation, they still preserve their performance advantages over the matched filter receiver. We will specifically discuss in Chapter 3 the impact of imperfect channel estimation on system capacity for systems using linear receivers and for those using a combination of linear receivers and successive interference cancellation schemes (groupwise successive interference cancellation). The impact of imperfect amplitude and phase estimation for successive interference cancellation receivers will also be examined in Chapter 3.

In terms of implementation complexity, successive interference cancellation receivers are the least costly to implement but also have the disadvantage of long detection delays. For large numbers of users this may be unacceptable, especially for delay sensitive applications. Although linear receivers are more costly to implement, they are superior to interference cancellers in terms of combating the multiple access interference. The tradeoffs between implementation costs and performance must be balanced carefully when choosing a particular receiver over another.

Several studies have focussed on comparisons among these suboptimal receivers, under various conditions. For example, comparisons among SIC, partial PIC, decorrelating and LMMSE detectors are presented in [Buehrer et al., 1996] for various conditions such as perfect power control (equal received powers for all users), single path Rayleigh fading and twopath frequency selective Rayleigh fading. The results in [Buehrer et al., 1996] show that, for perfect power control, the partial PIC has the best performance, while the LMMSE and the decorrela-

tor perform similarly. For Rayleigh fading channels, the linear receivers have a performance advantage, and the performance gap increases for the two-path channel. The performance of PIC and decorrelating receivers is also compared in [Juntti et al., 1997], where a BPSK (binary phase-shift keying) CDMA system operating over a multipath Rayleigh fading channel is considered, together with decision directed channel estimation. Simulation results in this paper indicate that with low to moderate signal-to-noise ratio, PIC outperforms the decorrelator. The decorrelator is only superior to PIC under severe near-far scenarios at high SNR.

Performance comparisons between SIC and PIC receivers have been presented in [Buehrer et al., 1996] and [Patel and Holtzman, 1994b]. The general observation is that perfect power control with equal powers benefits PIC receivers while significantly reducing the performance of SIC. As we will see later in this book, optimal SIC performance is achieved using unequal power control. This result is also suggested by the power-tradeoff regions of the SIC receiver shown in Section 2.

Due to their reduced implementation complexity, interference cancellation receivers are currently preferred for implementation in cellular wireless networks. Several companies, such as Fujitsu, NTT DoCoMo and NEC, have built suboptimal reduced complexity multiuser receivers, based on either successive interference cancellation, or parallel interference cancellation [Ephremides et al., 2000]. Capacity improvements on the order of 100% compared with systems using conventional receivers have been reported, when no intercell interference is present. For the multicell case, the capacity was improved 1.3 times compared with the conventional receiver case.

The performance of PIC receivers can be improved by cascading several detection stages as in [Xu et al., 2002], where a real-time prototype VLSI implementation of the multistage PIC algorithm is presented.

Important progress has also been made in implementing both linear multiuser detectors and interference cancellation receivers using software radio technology based on a combination of DSP and FPGA (field programmable gate array) devices [Seskar and Mandayam, 1999a, Seskar and Mandayam, 1999b]. Software radios can be used to provide reconfigurable radio architectures for diverse QoS guarantees. A combination of receiver flexibility and integrated resource management may offer the widest range of QoS guarantees in wireless networks.

While DSP implementation is currently too power-inefficient to be used in mobiles, rapid advances in DSP and CMOS (Complementary Metal-Oxide Semiconductor) technologies will most likely make possible

the use of multiuser receivers for mobile terminals in next generation wireless networks.

4.   Multi-Rate Multiuser Detection

One of the requirements for the next generation wireless networks is to integrate a wide range of applications, requiring a correspondingly wide range of transmission rates. To implement multirate communications in CDMA systems, several strategies have been proposed, such as

1 Fixed processing gain, variable chip rate [Wu and Geraniotis, 1994], [Wyrwas et al., 1992];

2 Fixed chip rate, variable spreading gain [Ottosson and Svensson, 1995], [Wu and Geraniotis, 1994], [Wyrwas et al., 1992];

3 Multicode (MC) [Chih-Lin and Gitlin, 1995] [Ottosson and Svensson, 1995];

4 Multi-modulation [Ottosson and Svensson, 1995].

The fixed processing gain, variable chip rate strategy introduces significant implementation complexity issues such as the need for synchronization of the receiver to its particular code rate and the need for additional frequency planning due to the unequal bandwidth spreading for different users. For that reason, the other three strategies are preferred, with the most popular choices being the fixed chip rate variable processing gain scheme and the multicode scheme. The two preferred implementations have similar performance, with a slight advantage for the multicode method [Ottosson and Svensson, 1995], [Yao et al., 2004], and outperform the multi-modulation scheme, which suffers from a severe near/far problem [Ottosson and Svensson, 1995]. The multi-modulation access strategy employs an M-ary QAM (quadrature amplitude modulation) scheme and varies the modulation level, M to accommodate multiple bit rates. Different modulation schemes require different transmission powers; hence the occurrence of the near/far problem in such systems.

In the multicode access strategy, all users multiplex their information symbols onto multiple low rate signature waveforms. One advantage of this scheme is that all the users have the same processing gain, and thus it may be easier to construct signature sequences with good crosscorrelation properties. However, this access scheme yields a high peak-to-average power ratio for high rate users, since the sum of many parallel channels gives rise to large amplitude variations. As a consequence, the multicode access strategy requires more costly power amplifiers.

The fixed chip rate, variable spreading gain (VSG) scheme performs similarly to the multicode method for both AWGN (additive white Gaussian noise) and multipath fading channels. For very high rate users only a small spreading gain is used, making this technique more susceptible to intersymbol interference and sensitive to external interference from neighboring cells [Ottosson and Svensson, 1995]. Three different detection strategies for variable spreading gain systems have been proposed: the low rate detector (LRD), the high rate detector (HRD), and groupwise successive interference cancellation (GSIC). Due to similarities between the LRD and HRD, we first focus our discussion on these two schemes. To explain the conceptual differences between the LRD and HRD, we consider a simple two-user system, in which one user transmits with a high rate (user 2), equal to M times the rate of the lower rate user (user 1). As a consequence, in user 1's bit interval of width $T_b^{(1)}$, user 2 transmits M bits, each within a bit interval of width $T_b^{(2)} = T_b^{(1)}/M$ (see Fig. 1.14). In the example shown in Fig. 1.14, user 1 transmits bit 1 while user 2 transmits a succession of bits: $\{1, 1, -1, 1\}$. These bits modulate a spreading sequence of length $N_1 = MN_2$. In the example from Fig. 1.15, $N_2 = 6$, $N_1 = 24$ and $M = 4$. In the LRD case, the detector operates at a lower bit rate and therefore, a decoding delay of M bits is incurred for the high rate user. To implement an LRD, an equivalent system is considered in which each high rate user is equivalent to M virtual low rate users, which have expanded signature sequences. The expanded signature sequences $\tilde{s}_2^{(i)}$ for virtual user i, $i = 1, 2, \ldots, M$, are constructed by zero padding the original signature sequence for transmitted bit i, to form an extended signature sequence of length MN_2. For the considered example:

$$\tilde{\mathbf{s}}_2^{(1)} = [1, 1, -1, 1, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T;$$

$$\tilde{\mathbf{s}}_2^{(2)} = [0, 0, 0, 0, 0, 0, 1, 1, -1, 1, 1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T;$$

$$\tilde{\mathbf{s}}_2^{(3)} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, -1, 1, 1, -1, 0, 0, 0, 0, 0, 0]^T;$$

$$\tilde{\mathbf{s}}_2^{(4)} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, -1, 1, 1, -1]^T.$$

Therefore, the equivalent LRD implementation is a simple multiuser detector implementation for a system with $K = K_1 + MK_2$ virtual users (K_1 is the number of low rate users and K_2 is the number of high rate users).

An alternate implementation of multirate multiuser detectors is the HRD, which operates at the higher transmission rate, and therefore out-
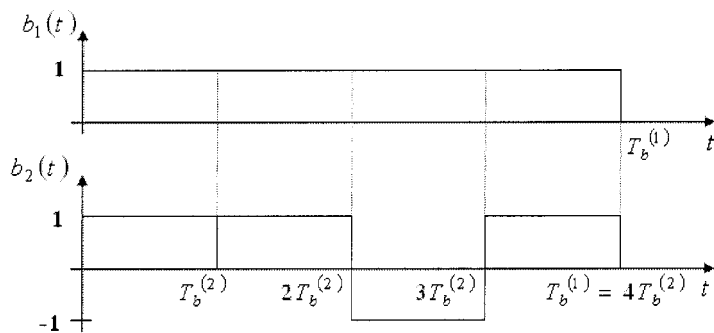
Figure 1.14. Bit transmission for multirate systems
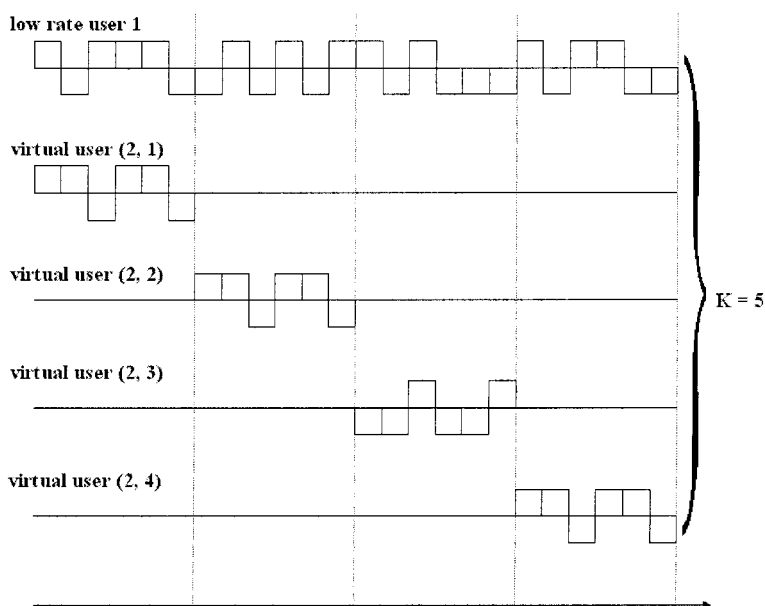
Figure 1.15. Virtual user equivalence in LRD multirate systems

puts bit decisions for the high rate users every $T_b^{(2)}$ time units. The advantage of the HRD is that no bit delays are incurred for the high rate users. As we will see momentarily, this comes at the expense of

worse performance for the low rate users. As we can see from Fig. 1.16, the effective signature sequence for the low rate user is reduced to a truncated signature sequence which varies periodically, with period M. For the considered example, the signature sequence of the low rate user is split into four truncated signature sequences: $\mathbf{s}_1^T = [\mathbf{s}_{1,1}^T, \mathbf{s}_{1,2}^T, \mathbf{s}_{1,3}^T, \mathbf{s}_{1,4}^T]$, with

$$\mathbf{s}_{1,1}^T = [1, 1, 1, 1, 1, -1];$$

$$\mathbf{s}_{1,2}^T = [-1, 1, -1, 1, -1, 1];$$

$$\mathbf{s}_{1,3}^T = [1, -1, 1, -1, -1, -1];$$

$$\mathbf{s}_{1,4}^T = [1, -1, 1, 1, -1, -1].$$

To detect a low rate user, maximal ratio combining is used to combine the outputs of the M subintervals.
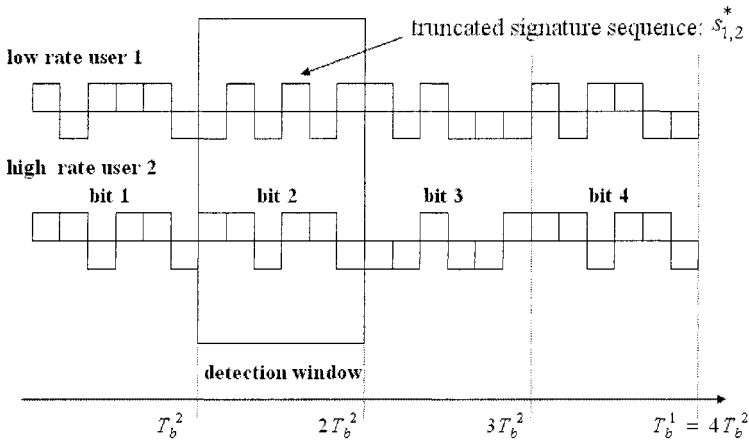
Figure 1.16. HRD for multirate systems

Peformance analyses of LRD and HRD as well as comparisons between the two techniques have been reported for both the decorrelator [Saquib et al., 1999] and the LMMSE [Ge and Ma, 1998] detectors.

As an alternate implementation, groupwise multiuser detection [Wijk et al., 1995] has recently emerged as an appealing solution for multirate multiuser detection. In GSIC systems, users are grouped according to their transmission rates and are detected in groups, while the interfer-

ence among groups is successively cancelled (see Fig. 1.17). In Fig. 1.17 a block diagram of GSIC is presented. Given the received signal $r(t)$, the users in the first detection group are detected and their bits are estimated as $\widehat{\mathbf{b}}_1$. Then, based on the detected bits and the estimated amplitude and phase for all class 1 users, their signals are reconstructed and subtracted from the received signal $r(t)$, such that, ideally, the class 1 users' interference is completely cancelled. This process continues until the last group of users is detected. A natural detection order has been proposed in the literature [Wijting et al., 1999], which considers the detection of the high rate users first. These high rate users are expected to cause more interference due to high power requirements, and in turn, to be less sensitive to the low power users' interference.

Within a group, any type of detector can be implemented, although the simplest, most common choice is to use matched filter receivers. Performance comparisons with various other detectors (e.g. the decorrelator, parallel interference cancellers (PIC) detectors) have been presented in [Juntti, 1998b] and [Juntti, 1998a] using simulations. More recently, performance analysis for power controlled GSIC systems have been presented in [Kim and Bambos, 2001] for a perfect cancellation scenario, and in [Comaniciu and Poor, 2003b] for the imperfect cancellation case. Further discussion regarding the characterization of the capacity of systems using GSIC will follow in Chapter 3.

GSIC systems have the advantages of a relatively simple implementation and good performance, but they have been shown to be quite sensitive to channel estimation errors and they also yield detection delays that increase with an increase in the number of detection groups.
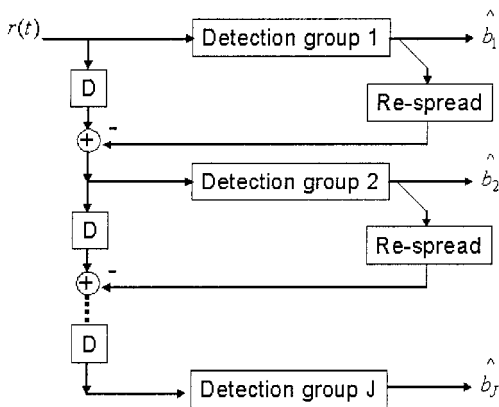
Figure 1.17.   Groupwise successive interference cancellation

5. Information Theoretic Aspects: Spectral Efficiency

While the discussion on the performance of multiuser detection thus far has focussed on performance metrics such as the bit error rate and other measures derived from it, a more fundamental approach is to consider the performance from an information theoretic point of view. In a CDMA system with K users using signature sequences of length N chips, the fundamental figure of merit is the spectral efficiency, which is defined as the total number of bits per chip that can be transmitted reliably. Further, in a CDMA system, approximating its bandwidth as the chip rate, the spectral efficiency can be essentially measured in bits per second per hertz (bits/s/Hz). The most comprehensive treatment of the performance of multiuser detectors in this context is due to [Verdú and Shamai, 1999] where the focus is on systems that use random spreading sequences. The reasoning for this particular choice of signature sequences is that it accurately models second and third generation CDMA cellular systems that use pseudonoise sequences that span several symbol intervals. Additionally, the spectral efficiency obtained by averaging out the random signature sequences serves as a lower bound on the optimal spectral efficiency achievable with deterministically chosen signature sequences.

Before we summarize the results on the spectral efficiency of the various multiuser receivers, it is of interest to consider the following two cases for the purposes of benchmarking the spectral efficiency performance. First, consider a system where there is no spreading imposed and the users are jointly detected in the presence of additive white Gaussian noise with power spectral density $N_o/2$. The maximal spectral efficiency η^{ns} in such a system can be shown to satisfy the following equation

$$\eta^{ns}\left(\frac{E_b}{N_o}\right) = \frac{1}{2}\log\left(1 + 2\eta^{ns}\left(\frac{E_b}{N_o}\right)\frac{E_b}{N_o}\right), \qquad (1.39)$$

where the spectral efficiency is shown as an explicit function of the ratio of the energy per bit E_b to N_o. It can be verified that the solution to the above equation is positive if and only if $\frac{E_b}{N_o} > \log_e 2 = -1.6dB$. If we now consider a system with spreading and further impose that the system be synchronous with orthogonal signature waveforms of chip length N being employed by the K users, then the spectral efficiency η^{orth} can be shown to be

$$\eta^{orth}\left(\frac{E_b}{N_o}\right) = \frac{K}{N}\eta^{ns}\left(\frac{E_b}{N_o}\right), \qquad (1.40)$$

where it is necessary that $K \leq N$. When $K = N$ it is seen that orthogonal multiple access achieves the same spectral efficiency as an unconstrained multiple access system with no spreading for equal rate and equal power users. It is also well known that when $K \geq N$, it is still possible to find spreading codes that incur no loss in spectral efficiency relative to the case of no spreading. These signature sequences are referred to as Welch-bound-equality (WBE) sequences [Rupf and Massey, 1994].

In the case of random spreading sequences, the spectral efficiency is a random variable itself. Using an asymptotic (large N and K with K/N fixed) analysis, [Verdú and Shamai, 1999] shows the convergence of these random spectral efficiencies to deterministic quantities for the cases of the following receivers: the matched filter receiver, the optimal multiuser receiver, the decorrelator and the LMMSE receiver. Letting $\beta = \frac{K}{N}$ denote the number of users per dimension and

$$\mathcal{F}(a,b) = \left(\sqrt{a\left(1 + \sqrt{b}\right)^2 + 1} - \sqrt{a\left(1 - \sqrt{b}\right)^2 + 1} \right)^2 ,$$

the spectral efficiencies in each of the above cases is given as follows.

- For the single-user matched filter receiver, the spectral efficiency, η^{mf}, converges almost surely as $K \to \infty$ to

$$\lim_{K \to \infty} \eta^{mf} = \frac{\beta}{2} \log \left(1 + \frac{SNR}{1 + SNR\ \beta} \right) \qquad (1.41)$$

- For $\beta > 0$, the optimal spectral efficiency, η^{opt}, converges almost surely as $K \to \infty$ to

$$\begin{aligned}
\lim_{K \to \infty} \eta^{opt} &= \frac{\beta}{2} \log \left(1 + SNR - \frac{1}{4}\mathcal{F}(SNR, \beta) \right) \qquad (1.42) \\
&\quad + \frac{1}{2} \log \left(1 + SNR\ \beta - \frac{1}{4}\mathcal{F}(SNR, \beta) \right) \\
&\quad - \frac{\log e}{8\ SNR} \mathcal{F}(SNR, \beta)
\end{aligned}$$

- For $\beta \leq 1$, the spectral efficiency, η^{dec}, of the decorrelator converges in the mean-square sense as $K \to \infty$ to

$$\lim_{K \to \infty} \eta^{dec} = \frac{\beta}{2} \log \left(1 + SNR\left(1 - \beta\right) \right) \qquad (1.43)$$

- For $\beta > 0$, the spectral efficiency, η^{mmse}, of the LMMSE receiver converges in the mean-square sense as $K \to \infty$ to

$$\lim_{K \to \infty} \eta^{mmse} = \frac{\beta}{2} \log \left(1 + SNR - \frac{1}{4} \mathcal{F}(SNR, \beta) \right) \qquad (1.44)$$

In the above expressions, SNR denotes the ratio of the energy per transmitted N chips to the height of the noise power spectral density $N_o/2$, and in each case is given as $SNR = \frac{2N}{K} \frac{E_b}{N_o} \eta$ with η being the appropriate achievable rate for each receiver. Figure 1.18 shows the spectral efficiencies of the various receivers under the large K analysis as a function of the ratio $\frac{K}{N}$ for $\frac{E_b}{N_o} = 10dB$. It is observed that at large values of $\frac{K}{N}$, the optimal receiver does begin to approach the spectral efficiency of the case of no spreading, but the linear multiuser receivers experience a loss in spectral efficiency beyond an optimal value of $\frac{K}{N}$. This leads to the interesting question regarding the optimal trade-off between coding and spreading in a CDMA system. For the optimal receiver and the matched filter reciever, it is seen that the spectral efficiency is maximized by letting $\frac{K}{N} \to \infty$. Thus for these receivers, the coding-spreading trade-off favors coding, and it is best to use these receivers in systems with low rate error-correcting codes with minimal spreading (see also [Hui, 1984]). For large values of K, for the decorrelator, the optimal choice of $\frac{K}{N}$ ranges from 0 to 1 thereby suggesting a larger spreading factor in contrast to the optimal and matched filter receivers. In Fig. 1.19, the spectral efficiency with the optimal choice of $\frac{K}{N}$ is shown for the various receivers as a function of $\frac{E_b}{N_o}$. It is observed that for the decorrelator, the spectral efficiency is higher than that of the matched filter receiver for $\frac{E_b}{N_o} > 5.2dB$ and, unlike the matched filter, the spectral efficiency grows without bound. For the case of the LMMSE receiver, for low values of $\frac{E_b}{N_o}$, the optimal $\frac{K}{N}$ is very large indicating that the spectral efficiency is the same as that of the matched filter receiver. As pointed out in [Verdú and Shamai, 1999], the optimal $\frac{K}{N}$ reaches a value of 1 at $\frac{E_b}{N_o} = 4dB$ and a minimum value of 0.75 at $\frac{E_b}{N_o} = 10dB$.

For low $\frac{K}{N}$ systems such as are encountered in typical cellular CDMA settings, both the decorrelator and the LMMSE receiver provide comparable spectral efficiencies relative to the case of orthogonal signatures. The spectral efficiencies of the above multiuser receivers has also been considered for frequency flat fading channels in [Shamai and Verdú, 2001] in conjunction with power control strategies based on variants of water-pouring optimization. Significant gains in the spectral efficiencies can be realized using optimal as well as suboptimal power control algorithms for the matched filter, the LMMSE receiver and the optimal receiver,

**Spectral Efficiency
(bits per chip)**
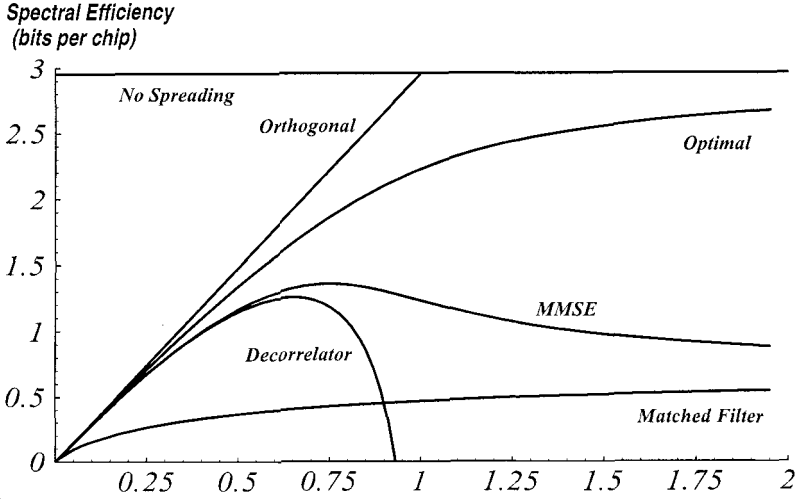
Figure 1.18. Spectral efficiencies for $\frac{E_b}{N_o} = 10dB$ (reprinted with permission from [Verdú and Shamai, 1999])

thereby motivating the need for integrating power control (cross-layer design) with multiuser detection.

6.   Multiuser Detection in Cross-Layer Design: Introductory Remarks and Book Outline

This first chapter of the book has been dedicated to the review of various wireless network architectures, and to the presentation and comparison of different multiuser receivers. While the early research on multiuser detection was focused on showing its superiority in adverse near/far conditions, recently the focus has shifted to studying its performance in power controlled wireless networks, and furthermore, to under-
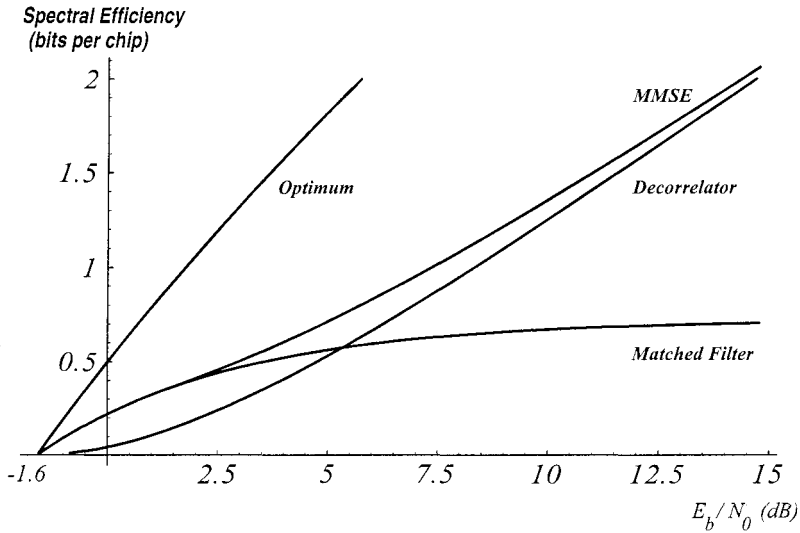
Figure 1.19. Spectral efficiency for optimal K/N (reprinted with permission from [Verdú and Shamai, 1999])

standing its interactions with upper layer protocols in cross-layer design. In Chapter 2 of the book, we briefly present several resource allocation techniques for QoS provisioning in wireless networks and we discuss in more detail their cross-layer interactions with multiuser receivers.

QoS guarantees for various applications can only be achieved if QoS support is provided at all layers of the protocol stack, i.e., network adaptation to changes in the achieved QoS must be hierarchically implemented across all layers. We have already noted that the interactions between different layers can be modeled by an exchange of pertinent information between layers. A certain level of abstraction for the performance of various layers can greatly simplify cross-layer design. Moreover, such abstraction can make it easier to determine what information should be exchanged between layers, and how this information should be used by the adaptation protocols to optimize the overall network performance. Chapter 3 of the book presents such an abstraction of the physical layer, by discussing the user capacity of power controlled wireless networks using multiuser receivers for several implementation

scenarios and different network architectures, including cellular and ad hoc networks. Based on the results in this chapter, integrated admission control is also discussed in detail for both cellular and ad hoc networks in Chapter 4.

Finally, general perspectives on integrated design in wireless networks using multiuser receivers are presented in Chapter 5.

Chapter 2

INTEGRATED RADIO RESOURCE ALLOCATION

1.   Introduction to Radio Resource Allocation

A common characteristic of multiple access wireless communication systems is that their capacity is limited by interference. As a result, admission of a new user into such a system results in more interference to the existing users and a consequent degradation in their signal quality. When the number of active users in the network reaches a certain value, the quality of transmission can become unacceptably low due to interference, implying that additional users cannot be admitted immediately if a certain Quality of Service (QoS) is required. Further, an inherent characteristic of wireless channels is the variation encountered in space, time and frequency due to mobility and to propagation effects encountered by radio waves. The propagation effects are usually classified into two categories: small-scale fluctuations due to scattering, and large-scale fluctuations due to shadowing. The traditional treatment of resource allocation in wireless systems is based on signal quality measurements that are averaged over time scales where small-scale fluctuations become insignificant. As a result, most approaches to understanding and designing resource allocation strategies take into consideration primarily path loss models and occasionally slowly varying large-scale fading effects due to shadowing.

Radio resource management is the collective term used to classify system level strategies for managing the physical layer of wireless networks, including transmitter power control, channel allocation and handoff. It is an important component necessary to sustain any wireless network of multiple users. One lesson of the cellular telephone success story is

51

that effective radio resource management is essential to promote system quality and efficiency, and it will be increasingly important in enabling the wireless data networks of the future. The simplest abstraction of radio resource management is the need to enable the following five requirements necessary for a wireless network:

- **Power Control**: power control is used to provide every user with a transmitter power level necessary to achieve a certain required level of signal quality at the receiver. The measure of signal quality depends on the nature of the wireless application and also on the time scale used in measuring such signal quality.

- **Base Station Assignment** (cellular networks): base station assignment is necessary to provide every mobile user with a receiver or access point to which it can connect. The strategies used to enable this are also often referred to as handoffs and may in general enable connectivity for a given user to more than one access point.

- **Routing** (ad hoc networks): Routing protocols are implemented in ad hoc networks in order to establish multi-hop transmission paths between any source-destination pair of nodes.

- **Channel Assignment**: Channel assignment is used to provide every user with a radio channel on which it can transmit. The channel assigned to transmitter/receiver pairs may be defined by frequency or time slots as in frequency-division multiple-access (FDMA) and time-division multiple-access (TDMA) systems, respectively, or signature sequences as in the case of CDMA systems.

- **Admission Control**: Admission control is necessary to regulate the entry of new users into the network in order to preserve the QoS of existing users and also to guarantee required signal quality for new users.

In all the above cases, the commonly considered metrics of signal quality include the signal-to-interference ratio (SIR), the bit-error rate (BER) and the frame-error rate (FER). The performance of radio resource management strategies is usually characterized by measures such as the probability of dropping existing users, the probability of blocking new users, the delay in handoffs, the frequency of handoffs and the resulting multiuser capacity of the wireless network under various strategies. There have been considerable developments in the area of radio resource management [Zander and Kim, 2001] focussing historically on voice services and more recently on a variety of wireless data applications. These techniques range in character from static to dynamic as well

as from centralized to distributed resource allocation. Earlier approaches to radio resource allocation considered admission control, channel allocation, power control and handoffs [Zander, 1992b, Hong and Rappaport, 1986, Vijayan and Holtzman, 1993, Sivarajan et al., 1989] as problems distinct from each other, while later efforts integrated these problems with a unified perspective [Chuah et al., 1995]. More recent work on resource allocation for wireless data has focussed on using approaches from microeconomics and game theory to model and design efficient networks [Goodman and Mandayam, 2000, Saraydar et al., 2002, Meshkati et al., 2003].

As we will see in the following sections of this chapter, multiuser receivers have a great impact on the design and performance of all resource management techniques. For example, consider the soft handoff problem for CDMA cellular networks. We know that soft handoff is necessary in CDMA cellular systems primarily because of MAI. If either SIC or MUD is used in these networks the effect of MAI is reduced, and as a consequence, the coverage (or range) of each cell is effectively increased, thus reducing the requirements on soft-handoff. Additionally, when we integrate SIC with power control, the cancellation order of users typically suggests ordering of powers in each cell in such a manner that it reduces other-cell interference. This again results in relaxing the requirements for soft-handoff.

Specific examples of integration of different resource management techniques with multiuser detection will be presented shortly. We start our discussion with a brief introduction to power control.

2. Power Control

In a mobile communication network, users are subject to a time varying radio channel that results in fluctuations of their received signals. In addition, a major problem in multiuser systems is the near/far effect, by which a nearby interferer can disrupt the reception (at a fixed receiver) of a highly attenuated desired signal. Both of these effects call for mobile devices to control their transmitter powers so that the received signal quality is acceptable. Additionally, the dynamic range and sensitivity of electronic components such as amplifiers in transceiver circuits requires power control to enable smooth transition in signal strengths for high fidelity operation. Power control algorithms developed in the literature usually require knowledge or estimates of some measure of signal quality such as the SIR, received signal power, or the bit or frame error rate, and sometimes knowledge of the channel gains. All of these quantities vary in time due to the fluctuations inherent in mobile channels. Power control algorithms come in many varieties, and may be classified according to

the nature of implementation as centralized or distributed, synchronous or asynchronous, iterative or non-iterative, and deterministic or stochastic.

The earliest works on power control considered algorithms that were non-iterative, synchronized and centralized [Aein, 1973, Nettleton and Alavi, 1983, Zander, 1992b, Grandhi et al., 1993]. These works identified the power control problem as an eigenvalue problem for nonnegative matrices. The power control problem is typically considered in a static (snapshot) setting where the radio link of every mobile user to its destination receiver (for cellular systems this may be the base station) is completely characterized by a deterministic channel gain.

The simplest and most tractable measure of signal quality is the SIR. For a cellular system, for any user k, the SIR γ_{kj} at base station j is given by

$$\gamma_{kj} = \frac{p_k h_{kj}}{\sum_{\ell=1, \ell \neq k}^{K} p_\ell h_{\ell j} + \sigma_j^2}, \tag{2.1}$$

where h_{kj} is the link gain from user k to base station j, $\mathbf{p} = (p_1, p_2, \ldots, p_K)^T$ is the vector of transmitted powers of the K users in the system and σ_j^2 is the background noise variance seen at receiver j. The QoS requirement for acceptable signal reception is specified by means of a target SIR γ^* as $\gamma \geq \gamma^*$. If it is required that all the K users be received at the base station j with SIRs above the target, then the system of inequalities $\gamma_{kj} \geq \gamma^*, \forall k = 1, \ldots, K$ can be written as a matrix inequality:

$$(\mathbf{I} - \mathbf{F})\mathbf{p} \geq \eta, \tag{2.2}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{F}$ is a normalized link gain matrix such that

$$\mathbf{F}_{k\ell} = \begin{cases} 0, & k = \ell, \\ \gamma^* \frac{h_{\ell j}}{h_{kj}}, & k \neq \ell. \end{cases} \tag{2.3}$$

Here, the vector $\eta = \gamma^*(\frac{\sigma_1^2}{h_{1j}}, \ldots, \frac{\sigma_K^2}{h_{Kj}})$, and the inequality in (2.2) is taken componentwise. The target SIR γ^* is said to be *feasible* if there exists a nonnegative power vector $\mathbf{p} = (p_1, p_2, \ldots, p_K)^T$ such that $\gamma_{kj} \geq \gamma^*$ for all k. Using the theory of nonnegative matrices, it can be shown that the SIR target γ^* is achievable if the Perron-Frobenius (largest) eigenvalue, e_F, of the matrix $\mathbf{F}$ is less than one. The eigenvalue e_F is a measure of system load and resource consumption and it plays a central role in radio resource management issues related to power control and admission control.

For example, in the noiseless case ($\eta = 0$) it can be shown that at all links the maximum achievable target SIR γ^* is equal to the reciprocal of the dominant eigenvalue of the matrix $\mathbf{F}/\gamma^*$. This fact forms the basis

for SIR balancing power control algorithms [Zander, 1992b] that aim to maximize the minimum SIR at all links in the system. Further, when there is noise in the system, the above quantity represents an upper bound on the maximum feasible balanced SIR level.

Centralized power control schemes that solve for a power vector satisfying (2.2) require a centralized mechanism where the link gain matrix $\mathbf{H}$ needs to be known. Distributed power control algorithms are more desirable in that they are more practical in a time varying environment and also avoid the heavy signaling required for centralized control. Several distributed power control algorithms that rely only on local measurements for each mobile user have been proposed, including the works of [Zander, 1992a, Foschini and Miljanic, 1993]. In the simplest abstraction of such algorithms, the power update for user k is given as

$$p_k^{(n+1)} = \frac{\gamma^*}{\gamma_k^{(n)}} p_k^{(n)}, \qquad (2.4)$$

where $\gamma_k^{(n)}$ is the SIR of user k at iteration n and $p_k^{(n)}$ is the corresponding transmit power. The convergence of these algorithms in each specific case can be shown using standard techniques from numerical linear algebra under the assumption that the link gain matrix remains unchanged during the iterations. A unified framework developed in [Yates, 1995] views distributed power control algorithms as iterations of the form

$$\mathbf{p}^{(n+1)} = I(\mathbf{p}^{(n)}), \qquad (2.5)$$

where the iterative mapping $I(\cdot)$ is referred to as the *interference function*. It has been shown that any power control algorithm of the form (2.5) will converge to the unique feasible (if it exists) solution of (2.2) as long as the interference function is *standard*, i.e., if it satisfies the following properties:

- Positivity : $I(\mathbf{p}) > 0$

- Monotonicity : $\mathbf{p} \geq \mathbf{p}' \Rightarrow I(\mathbf{p}) \geq I(\mathbf{p}')$

- Scalability : $\forall \alpha > 1, \ \alpha I(\mathbf{p}) > I(\alpha \mathbf{p})$

Distributed algorithms for power control have also been considered from several other aspects such as integration with base station assignment [Yates and Huang, 1995, Hanly, 1995], using BER as a metric [Kumar et al., 1995] and integration with multiuser receivers [Kumar and Holtzman, 1995, Ulukus and Yates, 1998a]. The topic of integrating power control with multiuser receivers and with admission control will be discussed at length in the remaining sections of this chapter.

The above developments relate primarily to static power control where the link gain matrix is fixed in a snapshot analysis. In reality, these channel gains vary even though signal quality measurements are obtained by averaging over several observations. Analysis of the dynamic behavior of power control can be found in [Ulukus and Yates, 1998d] and [Zander and Kim, 2001]. In practical systems such as 2G and 3G implemetations of CDMA, extensive link information is not readily available for power control updates, and the exchange between the base station and the mobile is limited to qualititative and quantized information. Typically this exchange is limited to 1-bit feedback commands from the receiver to the transmitter to either increase or decrease the transmit power in fixed increments. This scenario can be modelled as a stochastic non-linear feedback control system and an analysis of the same is considered in [Song et al., 2001].

More recent work on power control for wireless data has focussed on using approaches from microeconomics in which the QoS of heterogeneous data users is captured by means of utility functions. Distributed power control algorithms are then analyzed using the framework of game theory [Goodman and Mandayam, 2000, Saraydar et al., 2002, MacKenzie and Wicker, 2001, Xiao et al., 2001, Meshkati et al., 2003]. The issue of cross-layer design in the context of power control interactions with radio link and transport layer data retransmission protocols has been considered in [Song and Mandayam, 2001] using a hierarchical control theoretic approach for modeling and analysis.

3.   Integrated Power Control and Multiuser Detection

As seen in the previous sections, both multiuser detection and power control are primarily techniques that are used to combat the near/far problem while implicitly allowing a graceful degradation of performance with an increase in system load. While multiuser detection may be viewed as a purely physical layer design strategy, power control is often considered to be a system level control technique. In keeping with the cross-layer design paradigm considered in this book, an interesting question to ask is what performance gains may be had if these two techniques are integrated. The first attempts at answering this question were considered in [Kumar and Holtzman, 1995] and independently in [Ulukus and Yates, 1998a]. The basic premise of the integrated approach is to consider a strategy that controls both the transmitter powers and receiver filters of the users in an iterative and distributed manner. Given a QoS requirement specified in terms of a target received SIR for each of the users, the receiver filters are first updated to suppress the inter-

ference optimally, followed by assignment of transmitter powers to each of the users in a manner that causes the least interference to others.

Here, we will present the basic principles for integrating power control and multiuser receivers for a cellular setting, although the integration technique can be applied to any wireless network with fixed access points. Later in the book, we will discuss the integration of power control and multiuser detection in ad hoc wireless networks.

Consider a multicell synchronous CDMA system with K users. Generalizing equation (1.26), the baseband received signal vector, $\mathbf{r}_k \in \mathbb{R}^N$, at the assigned base station of user k can be written as

$$\mathbf{r}_k = \sum_{\ell=1}^{K} \sqrt{p_\ell} \sqrt{h_{k\ell}} b_\ell \mathbf{s}_\ell + \mathbf{n}, \tag{2.6}$$

where $\mathbf{s}_\ell \in \mathbb{R}^N$ is the spreading signature sequence of user ℓ, $h_{k\ell}$ is the link gain of user ℓ to the assigned base station of user k and p_ℓ is the transmit power of user ℓ. $\mathbf{n}$ is a zero mean Gaussian random vector with covariance matrix $\sigma^2 \mathbf{I}$. If a linear receiver filter $\mathbf{c}_k$ is applied for user k at its assigned base station, then the receiver filter output is given by

$$y_k = \mathbf{c}_k^T \mathbf{r}_k = \sum_{\ell=1}^{K} \sqrt{p_\ell} \sqrt{h_{k\ell}} b_\ell (\mathbf{c}_k^T \mathbf{s}_\ell) + \mathbf{c}_k^T \mathbf{n}, \tag{2.7}$$

where $\mathbf{c}_k^T \mathbf{n}$ is a zero mean Gaussian random variable with variance $\sigma^2 \mathbf{c}_k^T \mathbf{c}_k$. The corresponding SIR of user k can be written as

$$SIR_k = \frac{p_k h_{kk} (\mathbf{c}_k^T \mathbf{s}_k)^2}{\sum_{\ell \neq k} h_{k\ell} p_\ell (\mathbf{c}_k^T \mathbf{s}_\ell)^2 + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)^2}, \tag{2.8}$$

The simplest integrated power control and multiuser detection algorithm seeks to find a set of optimal powers $\mathbf{p} = (p_1, p_2, \ldots, p_K)^T$ and a set of receive filters $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K$ such that each user k achieves its QoS requirement of $SIR_k \geq \gamma_k^*$, where γ_k^* is the required SIR target of user k, $k = 1, 2, \ldots, K$. Formally, the integrated power control and multiuser detection problem can be stated as

$$\min_{\mathbf{p}} \sum_{k=1}^{K} p_k \tag{2.9}$$

subject to the constraints that for each user $k = 1, \ldots, K$:

$$p_k \geq \frac{\gamma_k^*}{h_{kk}} \frac{\sum_{\ell \neq k} h_{k\ell} p_\ell (\mathbf{c}_k^T \mathbf{s}_\ell)^2 + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)}{(\mathbf{c}_\ell^T \mathbf{s}_\ell)^2} \tag{2.10}$$

$$p_k \geq 0$$

$$\mathbf{c}_k \in \mathbb{R}^N$$

It can be seen that the solution to the above problem remains the same if the set of constraints in (2.10) are rewritten as

$$p_k \geq \frac{\gamma_k^*}{h_{kk}} \min_{\mathbf{c}_k \in \mathbb{R}^N} \frac{\sum_{\ell \neq k} h_{k\ell} p_\ell (\mathbf{c}_k^T \mathbf{s}_\ell)^2 + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)}{(\mathbf{c}_k^T \mathbf{s}_k)^2} \tag{2.11}$$

$$p_k \geq 0$$

Observing that the optimization in (2.9) is over the power vector only and that the inner optimization in (2.11) is over each user's filter coefficients for a fixed set of powers, a distributed and iterative integrated power control and multiuser detection algorithm can be realized as follows.

For a fixed set of powers, solving the filter optimization in (2.11) for each user, it was shown in [Ulukus and Yates, 1998a] that the filter for user k is given by

$$\mathbf{c}_k^* = \frac{\sqrt{p_k}}{1 + p_k \mathbf{s}_k^T \boldsymbol{\Sigma}_k^{-1} \mathbf{s}_k} \boldsymbol{\Sigma}_k^{-1} \mathbf{s}_k, \tag{2.12}$$

where the $N \times N$ matrix $\boldsymbol{\Sigma}_k$ is given by

$$\boldsymbol{\Sigma}_k = \sum_{l \neq k} p_\ell h_{k\ell} \mathbf{s}_\ell \mathbf{s}_\ell^T + \sigma^2 \mathbf{I}. \tag{2.13}$$

Note that the receiver filter $\mathbf{c}_k^*$ is the same as the LMMSE receiver filter derived in (1.27). Thus, iterating on the filter coefficients for a given set of powers results in a receive filter vector that maximizes the SIR for each user. In order to iterate on the powers, given the set of receivers, the transmit power for each user should be chosen to meet the constraint in (2.11) with equality. Thus, we can chose the set of powers to be

$$p_k = \frac{\gamma_k^*}{h_{kk}} \min_{\mathbf{c}_k \in \mathbb{R}^N} \frac{\sum_{\ell \neq k} h_{k\ell} p_\ell (\mathbf{c}_k^T \mathbf{s}_\ell)^2 + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)}{(\mathbf{c}_k^T \mathbf{s}_k)^2}, \quad k = 1, \ldots, K. \tag{2.14}$$

The power vector updates given a set of filter coefficients can now be written as

$$\mathbf{p}^{(n+1)} = \mathbf{I}(\mathbf{p}^{(n)}), \tag{2.15}$$

where $\mathbf{I}(\mathbf{p}) = [\mathbf{I}_1(\mathbf{p}), \ldots, \mathbf{I}_K(\mathbf{p})]^T$, with

$$\mathbf{I}_k(\mathbf{p}) = \frac{\gamma_k^*}{h_{kk}} \min_{\mathbf{c_k} \in \mathbb{R}^N} \frac{\displaystyle\sum_{\ell \neq k} h_{k\ell} p_\ell (\mathbf{c_k}^T \mathbf{s}_\ell)^2 + \sigma^2 \mathbf{c_k}^T \mathbf{c_k})}{(\mathbf{c_k}^T \mathbf{s}_k)^2}, \quad k = 1, \ldots, K. \tag{2.16}$$

The interference function $\mathbf{I}(\mathbf{p})$ given above is a standard interference function in that it satisfies the conditions of positivity, monotonicity and scalability outlined in Section 2. Thus the power control algorithm in (2.14) can be shown to converge to a unique feasible (if it exists) solution where $\mathbf{p} = \mathbf{I}(\mathbf{p})$. The filter coefficients then converge to the corresponding LMMSE receiver.

In terms of implementing the above iterative power control and multiuser detection algorithm, each user k implements a two-stage iteration as follows. At iteration $n+1$, the LMMSE filter for user k is constructed by using the power vector $\mathbf{p}(n)$ corresponding to the nth iteration. Then, the power vector is updated using the set of new filter coefficients obtained at the $(n+1)^{st}$ iteration. If the SIR targets are feasible, then starting from any initial power vector and filter coefficients, this iterative procedure will converge to the unique minimal power fixed point. It may seem that the value of this algorithm is limited in practice since the implementation for user k requires knowledge of all the other transmitter powers and the link gains to obtain $\boldsymbol{\Sigma}_k$ and hence $\mathbf{c}_k^*$. However, several simple estimation procedures have been proposed in [Ulukus and Yates, 1998a] to estimate $\boldsymbol{\Sigma}_k$, where user k requires only the knowledge of its own link gain h_{kk}. The availability of such link gain information is in fact fairly standard in several practical systems where the downlink information is used to infer the uplink gains and vice versa. An alternative approach to arrive at the same solution for integrated power control and multiuser detection is to use measurements of the mean square error in updating the filter coefficients [Kumar and Holtzman, 1995].

As an illustration of the performance gains obtained by integrating power control with multiuser detection, we present results (adapted with permission from [Ulukus and Yates, 1998a]) for a multicell CDMA system with 25 base stations that are uniformly spaced over a square grid of $5km \times 5km$. K users each using a random signature sequence of length $N = 150$ are independently and uniformly distributed on this grid. The link gains in the system are chosen to be inversely proportional to the fourth power of the distance between the transmitter and the receiver.
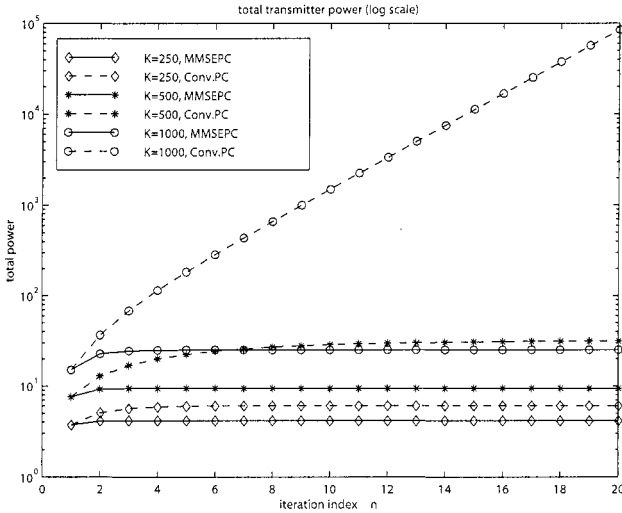
Figure 2.1. Performance gains of integrated power control and multiuser detection
(reprinted with permission from [Ulukus and Yates, 1998a])

(This is a reasonable approximation to path loss in terrestrial cellular
systems.) Figure 2.1 shows the total transmit power in the uplink as a
function of the number of iterations of the integrated power control and
multiuser detection algorithm for the cases of $K = 250, 500$, and 1000
users. The performance of the conventional matched filter receiver with

power control is also shown for reference. Aside from power savings obviously observed, it can also be seen that for the case of $K = 1000$ users, the integrated approach converts an infeasible system (as evidenced by the unbounded increase of the total power for the conventional matched filter receiver) into a feasible one. Thus this approach increases system capacity by allowing individual users to use higher SIR targets (and corresponding data rates). Alternatively, the network capacity can be increased by supporting more users at a fixed target SIR.

While the earliest work [Kumar and Holtzman, 1995, Ulukus and Yates, 1998a] on integrating power control with multiuser detection described above focussed primarily on linear multiuser receivers for single-rate systems, there has been considerable progress in exploring various aspects of such an integrated approach to power control. These range from using reduced complexity [Wang et al., 2001] and pilot based estimation [Almutairi et al., 2000] procedures for various power control parameters to multirate multiuser systems [Saquib et al., 2000, Kim and Bambos, 2001] to integrating power control with non-linear interference suppression techniques [Shum and Cheng, 2000, Varanasi and Das, 2002, Berggren and Slimane, 2002, Andrews and Meng, 2003, Shu and Niu, 2003] and also evaluating such an integrated approach for the forward link [Xiao and Honig, 2002].

An issue of paramount importance in all of the above approaches is the performance of integrated power control algorithms in practice where there is a lack of ideal knowledge of various user parameters. The stochastic power control approach first broached in [Ulukus and Yates, 1998d] for conventional matched filters has been reconsidered notably in the context of multiuser detection for multirate systems in [Saquib et al., 2000] and for multicell systems in [Varanasi and Das, 2002]. The integrated approach in asynchronous multirate systems with linear receivers is similar in philosophy to that described for single-rate systems with the difference being the use of appropriate sliding windows and appropriate sampling intervals to distinguish classes of users with different rates. The work in [Saquib et al., 2000] considers a family of BER objectives and proposes stochastic algorithms for decorrelating receivers in which the feasibility of power control is shown to be related to the condition that the users in the system have non-zero asymptotic efficiencies. Linear and decision-feedback based non-linear multiuser detectors are integrated with stochastic power control for use in a multicell CDMA system in [Varanasi and Das, 2002], where for a feasible system, the power control algorithms are shown to converge in the mean-square sense to the minimal power solution.

As we noted in Chapter 1, non-linear interference suppression schemes such as successive interference cancellation [Patel and Holtzman, 1994a] are being employed in integration with power control in emerging wireless data systems such as 3G and beyond. An interesting issue that arises in this context is the desired ordering of users for cancellation in the SIC scheme. If the total transmit power realized after power control is used as a measure of performance, then in single-rate systems, under identical cancellation errors, the detection order is not important. However, for multirate systems [Shum and Cheng, 2000] showed that under perfect cancellation, ranking users in descending order of link gains minimizes the total transmission power regardless of the different target SIR settings for users with different rates. This result also holds in certain cases when the cancellation is imperfect [Agrawal et al., 2004]. Integrated power control with groupwise serial multiuser detection, which is typically employed in multirate systems using a variable spreading factor, has also been considered in [Kim and Bambos, 2001] where an active link protection algorithm is used along with distributed power control. Optimal power control for groupwise successive interference cancellation systems with LMMSE receivers for in-group detection, and imperfect interference cancellation among groups, has been recently proposed in [Comaniciu and Poor, 2003b]. More recent approaches to the study and design of integrated power control and multiuser detection involve formulating and analyzing distributed algorithms for these in the setting of a noncooperative game [Meshkati et al., 2003].

4.   Access Control, Power Control and Multiuser Detection

We have seen in the previous section that performance gains can be achieved in a cellular system by joint receiver optimization and power control. Another possible dimension to be added is to exploit also the traffic burstiness[1] by appropriately designing the MAC. An integrated access control, power control and multiuser optimization algorithm proposed in [Comaniciu and Mandayam, 2002] adjusts to changes in the interference levels and stucture by both optimizing the physical layer (dynamically adjusting the transmission powers and the receiver filters) and the MAC scheduling. In this algorithm, the MAC layer appropriately schedules the delay insensitive traffic to take advantage of periods of low real-time traffic activity, while physical layer adjustments lead

---

[1]It is a well known fact that taking advantage of traffic burstiness can significantly improve system capacity [Viterbi, 1995].

to increased network capacity. In what follows, this algorithm will be illustrated for a voice and data CDMA system, for which users' QoS requirements are specified in terms of SIR targets and delay constraints. The MAC algorithm guarantees that the voice users' delay requirements are met, by always giving priority to the voice traffic. At the physical layer, the target SIRs are guaranteed if the power control feasibility condition holds, i.e., there exists a positive power vector assignment such that all users can meet their target SIRs.

The design of such an integrated algorithm imposes several challenges:

- Real-time burst activity prediction:

  Real-time changes in traffic activity require frequent updates of linear multiuser receiver filter coefficients, thus significantly increasing the implementation complexity. As a performance/complexity tradeoff, an alternate implementation is considered in which only data users employ multiuser detectors, while voice users (real-time users) employ matched filters. In order to further decrease complexity, only the signature sequences for the data users are considered for the update of the data users' filters, while the voice interference is approximated as background noise. This implementation is termed a partial hybrid system $(H - MMSE^{(p)})$, to differentiate it from the case in which all users make use of multiuser receivers, termed uniform LMMSE $(U - MMSE)$.

  Real-time burst activity prediction is more difficult for the uniform multiuser detector scenario since the access control needs exact knowledge of all users' signature sequences whenever any user changes activity. For the partial hybrid system, only the change in the interference power needs to be known for residual capacity updates, and thus a simple prediction algorithm similar to one proposed in [Comaniciu and Mandayam, 2000] can be used.

- Optimization of filter coefficients:

  Every time the interference pattern changes, the multiuser receivers must be re-optimized. Although not specifically addressed in [Comaniciu and Mandayam, 2002], complexity reduction at this step can be achieved if decorrelating receivers are employed for the partial hybrid system, since the decorrelator filter coefficients do not depend on the white noise level, and thus need not be updated when the real-time interference changes.

- Spectrally efficient data scheduling:

  For both implementation scenarios ($H-MMSE^{(p)}$ and $U-MMSE$), data is scheduled according to the residual capacity determined from the power control feasibility condition, in a round robin fashion.

While delay guarantees for both voice and data users are implemented by the MAC scheduling procedure with the support of the admission control[2], the total capacity available for users is limited by the power control feasiblity condition, which for a heterogeneous (voice and data) network can be derived as follows.

Power Control Feasibility Condition for Heterogeneous Networks

Assuming a network of K_v voice users transmitting at rate R_v, and K_d data users transmitting at a rate $R_d = MR_v$, where $M \geq 1$ is an integer, the general expression for the SIR achieved at the output of a linear multiuser detector with coefficients $\mathbf{c}_k$, for an arbitrary user k is given as:

$$SIR_k =$$

$$\frac{p_k h_k (\mathbf{c}_k^T \mathbf{s}_k)^2}{\displaystyle\sum_{j=K_v+1, j \neq k}^{K_v+K_d'} h_j p_j (\mathbf{c}_k^T \mathbf{s}_j')^2 + \sum_{\ell=1}^{K_v} h_\ell p_\ell (\mathbf{c}_k^T \mathbf{s}^*{}_\ell)^2 + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)}, \qquad (2.17)$$

where K_d' represents the number of virtual data users in the system. For the uniform LMMSE approach, $K_d' = MK_d$, $\mathbf{s}_j' \in \{-1/\sqrt{N_d}, 0, 1/\sqrt{N_d}\}^{N_v}$ with $j = K_v + 1, \ldots K_v + K_d'$, and the receiver filter vector for the kth data user is $\mathbf{c}_k \in \mathbb{R}^{N_v}$. As above, σ^2 denotes the background noise level. For the partial hybrid LMMSE approach, $K_d' = K_d$, $\mathbf{s}_j' \equiv \mathbf{s}_j \in \{-1/\sqrt{N_d}, 1/\sqrt{N_d}\}^{N_d}$, and $\mathbf{c}_k \in \mathbb{R}^{N_d}$. Since the partial hybrid LMMSE system operates at the high rate MR_v, we define $\{\mathbf{s}_l^*\}^{(m)}$, $m = 1, 2, \ldots, M$, to be the truncated voice signature sequence of length N_d, corresponding to the current data bit interval m. For notational simplicity, the superscript m will be suppressed in what follows, so that $\mathbf{s}_\ell^* \in \{-1/\sqrt{N_v}, 1/\sqrt{N_v}\}^{N_d}$, $\ell = 1, \ldots, K_v$, represents the truncated voice signature sequence for the current data bit interval. In reality, the truncated voice signature sequence differs from one data bit

---

[2]The admission control limits the number of users admitted in the system, such that, given the scheduling algorithm, the delay requirements for all users can be met, while maintaining the interference level within a tolerable range.

interval to another, with a periodicity of M data bit intervals. For the uniform LMMSE scenario, $\mathbf{s}_\ell^* \equiv \mathbf{s}_\ell$. For both scenarios, h_j and h_ℓ, $j = K_v + 1, \ldots, K_v + K_d$ and $\ell = 1, \ldots, K_v$, represent the link gains for data and voice, respectively.

Expression (2.17) can be particularized for voice users, so that the filter coefficients represent the desired user's spreading sequence. For random spreading sequences with data users' sequences normalized with $1/\sqrt{N_d}$ and voice sequences normalized with $1/\sqrt{N_v}$, the SIR for voice user k can be expressed as

$$SIR_k = \frac{h_k p_k}{\dfrac{1}{N_v} \displaystyle\sum_{j=1, j\neq k}^{K_v} h_j p_j + \dfrac{1}{N_d} \displaystyle\sum_{\ell=K_v+1}^{K_v+K_d'} h_\ell p_\ell + \sigma^2}. \qquad (2.18)$$

The SIR expression for the partial hybrid LMMSE case can be simplified via the following proposition. The proof is omitted here and can be found in [Comaniciu and Mandayam, 2002].

Proposition 2.1. *Consider a partial hybrid multiuser detector CDMA system, where data and voice users have different transmission rates, reflected in different spreading gains N_d and N_v, respectively ($N_v = M N_d$), the data users employ multiuser receivers built using only knowledge of data signature sequences, and the voice users use conventional receivers.*

Then, for any data filter vector $\mathbf{c}_k$, any truncated voice signature sequence $\{\mathbf{s}_\ell^*\}^{(m)}$, $\ell = 1, \ldots, K_v$, and any data bit decoding interval $m = 1, \ldots, M$, we have

$$\mathbf{E}\left\{ \left(\mathbf{c}_k^T \{\mathbf{s}_\ell^*\}^{(b)} \right)^2 \right\} = \frac{1}{N_v} \mathbf{E}\{(\mathbf{c}_k^T \mathbf{c}_k)\}.$$

As a consequence, the voice interference power for a given filter vector $\mathbf{c}_k \in \mathbb{R}^{N_d}$, can be expressed as $\frac{1}{N_v}(\mathbf{c}_k^T \mathbf{c}_k) \sum_{\ell=1}^{K_v} h_\ell p_\ell$. Hence, the SIR expression for the kth data user for the partial hybrid LMMSE system case becomes

$$SIR_k =$$

$$\frac{p_k h_k (\mathbf{c}_k^T \mathbf{s}_k)^2}{\displaystyle\sum_{j=K_v+1, j\neq k}^{K_v+K_d} h_j p_j (\mathbf{c}_k^T \mathbf{s}_j)^2 + \left(\dfrac{1}{N_v} \displaystyle\sum_{\ell=1}^{K_v} h_\ell p_\ell + \sigma^2 \right) (\mathbf{c}_k^T \mathbf{c}_k)} \qquad (2.19)$$

For both $U-MMSE$ and $H-MMSE^{(p)}$, the LMMSE filter coefficients are computed as in [Verdú, 1998]:

$$\mathbf{c}_k = \frac{\sqrt{p_k}}{1 + p_k \mathbf{s}_k'^T \Sigma_k^{-1} \mathbf{s}_k'} \Sigma_k^{-1} \mathbf{s}_k' \qquad (2.20)$$

where $\Sigma_k = \sigma_e^2 \mathbf{I_{dim}} + \sum_{j=K_v+1, j\neq k}^{K_v+K_d} h_j p_j \mathbf{s}_j' \mathbf{s}_j'^T$, $\mathbf{I_{dim}}$ is the identity matrix of dimension dim, with $dim = N_v \times N_v$ for the uniform LMMSE case, and $dim = N_d \times N_d$ for the partial hybrid LMMSE case, corresponding to the length of the data users' signature sequences (or extended signature sequences for the uniform LMMSE case).

For the uniform LMMSE scenario $\sigma_e^2 \equiv \sigma^2$, whereas for the partial hybrid LMMSE approach $\sigma_e^2 = \sigma^2 + \frac{1}{N_v} \sum_{l=1}^{K_v} h_l p_l$, where σ^2 is the background noise level.

Denote the target SIR for a user k (voice or data) by γ_k^*. Then, the QoS requirements are $SIR_k \geq \gamma_k^*$, $k = 1 \ \ldots \ n$, $n = (K_v + K_d')$.

For the partial hybrid scenario, the system of equations expressing the above conditions can be written as

$$\begin{cases} p_1 = \dfrac{\gamma_1}{h_1} \left[\dfrac{1}{N_v} \displaystyle\sum_{j=2}^{K_v} h_j p_j + \dfrac{1}{N_d} \sum_{j=K_v+1}^{K_v+K_d} h_j p_j + \sigma^2 \right] \\ \\ p_{K_v} = \dfrac{\gamma_{K_v}}{h_{K_v}} \left[\dfrac{1}{N_v} \displaystyle\sum_{j=1}^{K_v-1} h_j p_j + \dfrac{1}{N_d} \sum_{j=K_v+1}^{K_v+K_d} h_j p_j + \sigma^2 \right] \\ \\ p_n = \dfrac{\gamma_n^*}{(\mathbf{c}_n^T \mathbf{s}_n)^2 h_n} \left[\displaystyle\sum_{j=K_v+1, j\neq n}^{K_v+K_d'} (\mathbf{c}_n^T \mathbf{s}_j')^2 h_j p_j + \right. \\ \left. + \dfrac{1}{N_v}(\mathbf{c}_i^T \mathbf{c}_i) \sum_{\ell=1}^{K_v} h_\ell p_\ell + \sigma^2(\mathbf{c}_k^T \mathbf{c}_k) \right] \end{cases} \qquad (2.21)$$

Therefore, the power control feasibility condition is given by

$$(\mathbf{I_{dim}} - (\mathbf{A} - \mathbf{B}))\mathbf{p} = \sigma^2 \mathbf{u} \Leftrightarrow (\mathbf{I_{dim}} - \mathbf{C})\mathbf{p} = \sigma^2 \mathbf{u}, \qquad (2.22)$$

with

$$\mathbf{B} = \text{diag} \left\{ \frac{\gamma_1}{N_v}, \frac{\gamma_2}{N_v}, \ldots, \frac{\gamma_{K_v}}{N_v}, \ \gamma_{K_v+1}, \ldots, \ \gamma_n^* \right\} \qquad (2.23)$$

$$\mathbf{u}^T = \left[\frac{\gamma_1}{h_1}, \ \frac{\gamma_2}{h_2}, \ldots, \ \frac{\gamma_{K_v}}{h_{K_v}}, \ \frac{\gamma^*_{K_v+1}(\mathbf{c}^T_{K_v+1}\mathbf{c}_{K_v+1})}{(\mathbf{c}^T_{K_v+1}\mathbf{s}_{K_v+1})^2 h_{K_v+1}}, \ldots, \ \frac{\gamma^*_n(\mathbf{c}^T_n \mathbf{c}_n)}{(\mathbf{c}^T_n \mathbf{s}_n)^2 h_n} \right]$$

$$\tag{2.24}$$

$$\mathbf{A}^T = [\mathbf{a_1}, \ \mathbf{a_2}, \ldots, \ \mathbf{a_{K_v+1}}, \ \mathbf{a_n}], \tag{2.25}$$

where for $z_{k,j} = h_k/h_j$, the rows of the matrix $\mathbf{A}$ can be written as

$$\mathbf{a}^T_k = \begin{cases} \gamma_k \left[z_{1,k}/N_v, \ldots, \ z_{n,k}/N_v \right], \ \text{for } k = 1, \ \ldots \ K_v \\ \frac{\gamma^*_k}{(\mathbf{c}^T_k\mathbf{s}_k)^2} \left[\frac{1}{N_v}(\mathbf{c}^T_k\mathbf{c}_k)z_{1,k}, \ldots, \ \frac{1}{N_v}(\mathbf{c}^T_k\mathbf{c}_k)z_{K_v,k}, \ (\mathbf{c}^T_k\mathbf{s}_n)^2 z_{n,K_v+1}, \ldots \\ \qquad \ldots, \ (\mathbf{c}^T_k\mathbf{s}_n)^2 z_{n,k} \right], \ \text{for } k = K_v + 1, \ \ldots \ n. \end{cases}$$

$$\tag{2.26}$$

The matrix $\mathbf{C}$ is similar to the matrix $\mathbf{A}$ except that it has zeros on the main diagonal. To find a positive power vector solution, we note that the matrix $(\mathbf{I_{dim}} - \mathbf{C})$ is a nonnegative matrix, and by the Perron-Frobenius theorem [Strang, 1988], it has exactly one positive eigenvalue λ^* for which the corresponding eigenvector is positive (i.e., all components have the same sign). Thus, the equalities $\lambda \mathbf{p} = (\mathbf{I_{dim}} - \mathbf{C})p = \sigma^2 \mathbf{u}$ hold for $\lambda = \lambda^*$. The condition $\lambda^* > 0$ is equivalent to

$$\lambda_{max}(\mathbf{C}) < 1. \tag{2.27}$$

The resulting power vector solution is given by

$$\mathbf{p} = \sigma^2 (\mathbf{I_{dim}} - \mathbf{C})^{-1}\mathbf{u}. \tag{2.28}$$

A similar result can be derived for the uniform LMMSE case. The power control feasibility condition reduces to the same eigenvalue condition (2.27) for a different choice for the matrix $\mathbf{C}$. If identical derivation steps are applied, the system of equations that represents the SIR conditions can be reduced to the same matrix equation (1.16), but with different expressions for $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{u}$:

$$\mathbf{B} = \text{diag}\{\gamma^*_1, \ , \gamma^*_2 \cdots, \ \gamma^*_n\}$$

$$\mathbf{u}^T = \left[\frac{\gamma^*_1}{(\mathbf{c}^T_1\mathbf{s}_1)^2 h_1}, \ \frac{\gamma^*_2}{(\mathbf{c}^T_2\mathbf{s}_2)^2 h_2}, \ \cdots, \ \frac{\gamma^*_n}{(\mathbf{c}^T_n\mathbf{s}_n)^2 h_n} \right]$$

and

$$\mathbf{a}^T_k = \frac{\gamma^*_k}{(\mathbf{c}^T_k\mathbf{s}_k)^2} \left[(\mathbf{c}^T_k\mathbf{s}_1)^2 z_{1,k}, \ (\mathbf{c}^T_i\mathbf{s}_2)^2 z_{2,k}, \ \cdots, \ (\mathbf{c}^T_k\mathbf{s}_n)^2 z_{n,k} \right], \tag{2.29}$$

for $k = 1, \ \ldots, \ n$.

Once we have derived the power control feasibility condition for different network scenarios, the integrated access control algorithm can be summarized as follows.

Integrated Access Control and Detection

At each time slot:

1 *Predict the new voice interference structure and power.*

2 *Re-optimize the filter coefficients according to the new voice interference structure (or interference power) using (2.20).*

3 *Recompute the Perron-Frobenius eigenvalue of the matrix* **C**

4 **If** $\lambda_{max}(C) < 1$:
    Increment the number of data users granted access, and update the powers and filter coefficients;
    If power control is still feasible, schedule more data users for transmission;
   Else
    Decrement the number of data users granted access and update powers and filter coefficients;
    Repeat until power control is feasible.

A flowchart for the access control algorithm is presented in Fig. 2.2. For implementing the access control, the base station (BS) maintains two lists containing IDs for the active and inactive data users respectively. At each time slot, when the voice activity changes may favor the increase of the number of active data users, the data user selected for possible transmission is the one that is at the head of the line in the inactive data users list. Whenever the power control feasibility condition does not hold, data users become successively inactive, starting with the head of the line for the active list, until power control feasibility is satisfied. The head of the line user is the oldest one in that particular list; when a data user changes activity from inactive/active, its ID is attached at the end of active/inactive list.

The implementation of the proposed access control algorithm differs for the uplink and downlink, and also depends on the particular scenario. Implementation issues are summarized in Table 2.1.

For both uplink and downlink scenarios, since the power update for all users is given by (2.28) which depends on the filter coefficients and these in turn depend on the choice of powers, the Perron-Frobenius eigenvalue computation, as well as the power and filter updates must be done iteratively. Simulation results show a typically rapid convergence for the iterative procedure (see Fig. 2.3).

The tradeoffs between the implementation complexity and performance are illustrated in Figs. 2.4 and 2.5. Figure 2.4 illustrates the
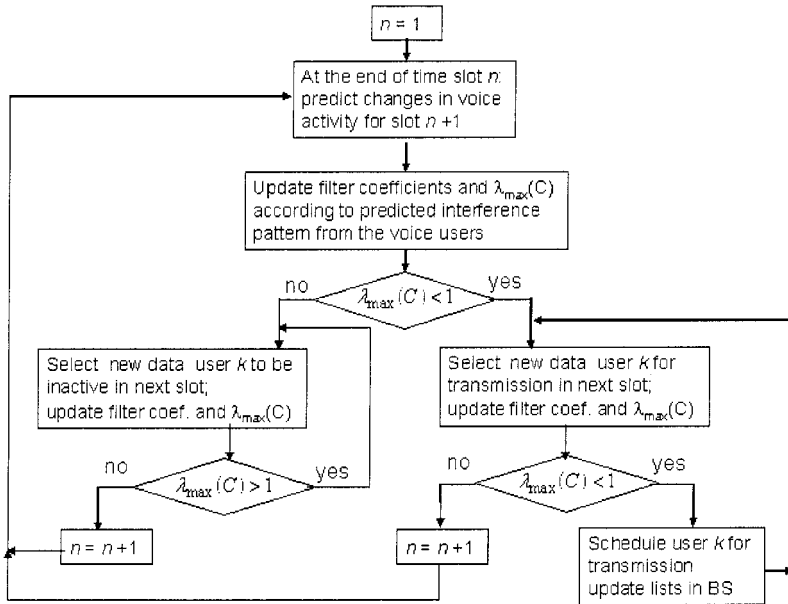
Figure 2.2. Integrated access control and receiver adaptation flowchart

Table 2.1. Implementation Issues Related to Uplink/Downlink

|	Filter updates	Feedback information from base to mobile

Uniform LMMSE uplink	- both voice and data user filters
	    - requires knowledge of
	        active voice users'
	        signature sequences
		change status bit for data user that changed activity

Uniform LMMSE downlink	-both voice and data user filters
	    - requires knowledge of
	        active voice users'
	        signature sequences
		- signature sequences for voice user that changed activity
		- signature sequences for data user that changed activity

Partial Hybrid LMMSE uplink	- only data user filters
	    - requires only knowledge
	    of voice interference power
		change status bit for data user that changed activity

Partial Hybrid LMMSE downlink	- only data user filters
	    - requires only knowledge
	    of voice interference power
		- total received voice power
		- signature sequences for data user that changed activity
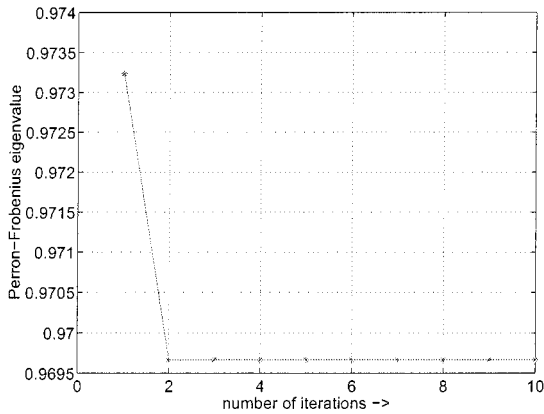
Figure 2.3. Simulated convergence of the Perron-Frobenius eigenvalue for the partial hybrid LMMSE implementation
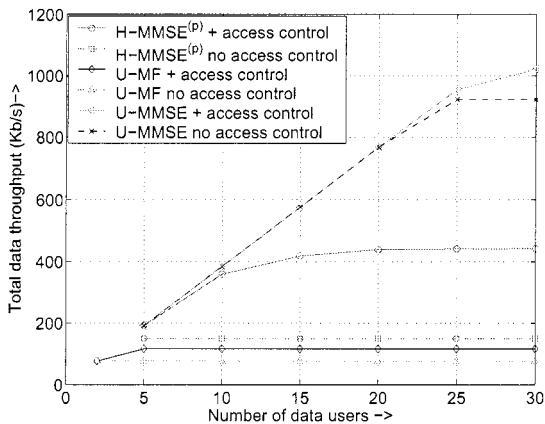

Figure 2.4. Total data throughput capacity

total data throughput that can be accommodated by the system for both LMMSE approaches and also for the uniform MF system. It can be seen that employing access control improves the system capacity regardless of the particular receiver structure. The best performance is obtained, as expected, by the uniform LMMSE approach with access control. However, it can be seen that the partial hybrid system, used in conjunction with access control, has very good performance with substantial complexity reduction.
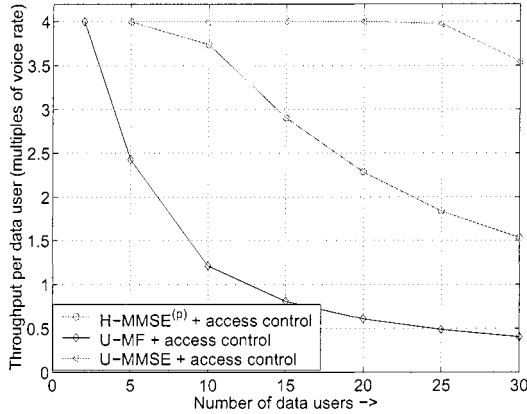
Figure 2.5. Throughput per user for integrated access control and detection

The improved performance achieved by the access control comes with a penalty in the data throughput per user, as seen in Fig. 2.5. This throughput penalty arises as a consequence of the fact that in order to increase the overall system utilization, more data users must be admitted into the system so that the low real-time traffic activity can be fully exploited. However, for the time slots in which the real-time traffic load is high, only a few data users can be scheduled for transmission and thus the average throughput per data user is reduced when the number of admitted data users increases. In Fig. 2.5, the throughputs achieved per data user are represented as multiples of the basic transmission rate R (equal to the voice rate R_v), and are seen to decrease as the number of data users in the system increases.

All numerical values were obtained for $N_v = 128$, $N_d = 32$, $\gamma_v = 5$, $\gamma_d = 10$, $T_s = 0.02s$, $\sigma^2 = 10^{-17}$, and $K_v = 10$. Power vector initializations for received voice and data powers were 20 dB above the noise floor, and perfect voice activity prediction was assumed for maximal gain illustration. The matched filter performance was determined based on results from [Comaniciu and Mandayam, 2000].

5. Traffic-Aided Multiuser Detection

In the previous section, we saw that access control can yield significant gains even for systems using multiuser detectors. The approach in [Comaniciu and Mandayam, 2002] assumes that traffic activity can be predicted, and this information can be used for data scheduling as well as for receiver adaptation. While for the simplified partial hybrid
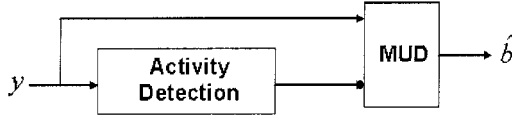
Figure 2.6. Two stage multiuser detector (reprinted with permission from [Chen and Tong, 2001])

approach, only the prediction of the aggregate interference power level is required and an approach similar to that in [Comaniciu and Mandayam, 2000] can be used, the uniform multiuser detection receiver update requires individual activity tracking for all users, which is more difficult to achieve. In this section, we discuss the implementation of a traffic-aided multiuser detection receiver, which dynamically updates the set of active users used for constructing the multiuser detector. This implementation has been proposed in [Chen and Tong, 2001] and it is based on a two stage receiver, as seen in Fig. 2.6.

The second stage multiuser detector is simply a classical multiuser detector when the knowledge of the set of active users is available. In [Chen and Tong, 2001] a decorrelating receiver is considered for stage two. To examine this problem, consider a system in which packets are sent by multiple users in each of a succession of time slots. In order to detect the set of users that are active in a given slot, an activity indicator γ_n^k for the kth user is used for modeling the received signal y_n in slot n:

$$y_n(t) = \sum_{k=1}^{K} \gamma_n^k A_n^{(k)} b_n^{(k)}(t) \mathbf{s}_k + z_n(t), \ t = 1, \dots, T, \qquad (2.30)$$

where $\mathbf{s}_k$, $A_n^{(k)}$ and $b_n^{(k)}$ are respectively, the signature sequence, the amplitude and the symbol for user k in time slot n, K is the total number of users, T is the number of bits in a packet (i.e., the packet length), and $z_n(t)$ is the background noise. The indicator γ_n^k is 1 if user k is active in slot n, and 0 otherwise.

Given (2.30), the user identification problem becomes an estimation problem for the binary random sequence γ_n^k. Chen and Tong use a first order approximation for the traffic burstiness, in which an individual source is modeled as a two-state Markov chain with a state transition matrix

$$a^{(k)} = \begin{bmatrix} 1 - p^{(k)} & p^{(k)} \\ q^{(k)} & 1 - q^{(k)} \end{bmatrix}, \qquad (2.31)$$
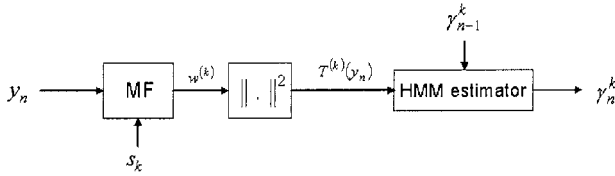
Figure 2.7. State tracker with matched filter receiver (reprinted with permission from [Chen and Tong, 2001])

where $p^{(k)} = P\left(\gamma_n^k = 1 | \gamma_{n-1}^k = 0\right)$, and $q^{(k)} = P\left(\gamma_n^k = 0 | \gamma_{n-1}^k = 1\right)$.

On defining the "hidden" state $\Gamma_n = \left(\gamma_n^{(1)}, \ldots, \gamma_n^{(K)}\right)$, an optimal traffic tracker can be developped by modeling the received vector using a hidden Markov model (HMM). Since, the combined state variable Γ_n is a Markov sequence with state dimensionality 2^K, the complexity grows exponentially with the number of users and therefore suboptimal solutions are preferable. The key to complexity reduction is the decoupling of individual users' state tracking. An important observation is that, although performance loss is incurred by this suboptimal approach, good performance can still be achieved since the required signal to noise ratio for detection of the presence of a particular user is much less that the one required for symbol detection for that user. The presence estimation exploits the fact that γ_n^k remains constant for the entire duration of a packet, which yields a form of diversity, and also the fact that traffic prediction can improve the estimate of γ_n^k.

A simple front end receiver based on matched filtering is proposed in [Chen and Tong, 2001] (see Fig. 2.7). The simple matched filter is selected for the first stage detector, due to its implementation simplicity, and due to the fact that it does not impose any restrictions on the rank of the cross-correlation matrix of all the spreading codes in the system. In general, a full-rank condition might be too restrictive, since the total number of users in the system is typically much higher than the number of active users.

In what follows, we describe in more detail the implementation and performance of the two-stage detector, using a matched filter front end for activity detection. If the received signal is passed through a filter matched to user k's signal, the output is given by

$$
w^{(k)}(t) = \gamma_n^k A_n^{(k)} b_n^{(k)}(t) + \sum_{\ell \in I_n, \ell \neq k} \rho_{k\ell} A_n^{(\ell)} b_n^{(\ell)}(t) + z^{(k)}(t), \ t = 1, 2, \ldots, T,
$$

$$(2.32)$$

where I_n is the set of all active users in time slot n, and $\rho_{k\ell}$ is the correlation coefficient between the signature sequences of users k and ℓ. Assuming that $z_n(t)$ is white and Gaussian with spectral height σ^2, $z^{(k)}(1)$, $z^{(k)}(2)$, ..., $z^{(k)}(T)$ is a sequence of independent zero mean Gaussian random variables, each having variance σ^2.

In order to make the analysis tractable, a Gaussian approximation is made for the second term in (2.32). An enhanced noise variable can be defined as

$$\widehat{z}^{(k)}(t) = \sum_{\ell \in I_n, \ell \neq k} \rho_{k\ell} A_n^{(\ell)} b_n^{(\ell)}(t) + z^{(k)}(t). \qquad (2.33)$$

Assuming zero-mean symbols, the variance of $\widehat{z}^{(k)}(t)$, is given by

$$\sigma_k^2 = \sigma^2 + \sum_{\ell \in I_n, \ell \neq k} \rho_{k\ell}^2 \left[A_n^{(\ell)} \right]^2. \qquad (2.34)$$

The above defined variance cannot be computed due to the fact that the active set of users is not known. Instead a predicted variance can be defined, which takes advantage of the traffic statistics

$$\widehat{\sigma}_k^2 = \sigma^2 + \sum_{\gamma_{n-1}^j = 1, j \neq k} (1 - q^{(j)}) \rho_{kj}^2 \left[A_n^{(j)} \right]^2 + \sum_{\gamma_{n-1}^\ell = 0, \ell \neq k} p^{(\ell)} \rho_{k\ell}^2 \left[A_n^{(\ell)} \right]^2. \qquad (2.35)$$

Given the Gaussian approximation, and combining all the user-k matched filter outputs in one packet into an observation vector $w(k)$, γ_n^k can be estimated using the following hypothesis testing model:

$$\begin{aligned} H_1 &: \mathbf{w}^{(k)} = \widehat{\mathbf{z}}^{(k)} \\ H_2 &: \mathbf{w}^{(k)} = A^{(k)} \mathbf{b}^{(k)} + \widehat{\mathbf{z}}^{(k)}. \end{aligned} \qquad (2.36)$$

In (2.36), $\mathbf{b}^{(k)}$ is the symbol vector of the packet for user k and $\mathbf{z}^{(k)}$ is a Gaussian random vector with zero mean and covariance matrix $\sigma_k^2 \mathbf{I}$. The above model implicitly assumes that the amplitude is constant for all the symbols within a packet. To relax this condition and to deal with the fact that the symbol vector is unknown, a uniformly most powerful invariant (UMPI) test statistic is proposed. Although for the given composite hypothesis testing problem, a uniformly most powerful test does not exist, the invariance principle can be used to determine an optimal UMPI test.

The first step is to determine whether or not the test problem is invariant under a group of transformations. Once this invariant group of transformations is obtained, a maximal invariant (MI) statistic is found,

and the UMPI test is based on this statistic. For the given problem, the test for γ_n^k can be abstracted as the following test concerning the mean $\mathbf{m}$ of a Gaussian random vector $\mathbf{w}^{(k)}$, having covariance matrix equal to $\sigma_k^2 \mathbf{I}$:

$$
\begin{aligned}
H_1 &: \mathbf{m} = 0 \\
H_2 &: \mathbf{m} \neq 0
\end{aligned}
\tag{2.37}
$$

The above hypotheses can be shown to be invariant to any orthogonal transformation, so that an MI statistic is given by

$$
S_{MI} = \mathbf{w}^{(k)^T} \mathbf{w}^{(k)}.
\tag{2.38}
$$

This MI statistic has a chi-square distribution under hypothesis 1 and noncentral chi-square distribution under hypothesis 2. Thus, a UMPI test statistic is the maximum invariant itself: $T^{(k)}(y_n) = \mathbf{w}^{(k)^T} \mathbf{w}^{(k)}$.

Traffic predictability can improve the estimate for γ^k. In particular, given that γ^k is a two-state Markov chain, $T^{(k)}(y_n)$ satisfies an HMM, and an HMM tracker can be used to determine γ^k, given $T^{(k)}(y_n)$ and the state estimate from the previous slot. Chen and Tong compare the performance of this traffic aided estimator with the one-shot UMPI test which does not use any traffic information. The one-shot UMPI decision is determined using a simple thresholding of the MI:

$$
T^{(k)}(y_n) \underset{H_2}{\overset{H_1}{\underset{>}{\lessgtr}}} \eta.
\tag{2.39}
$$

In contrast, the HMM decision uses information about the past states, as follows

$$
\frac{P(\gamma_n^k = 1 | T^{(k)}(y_n) \ldots T^{(k)}(y_1))}{P(\gamma_n^k = 0 | T^{(k)}(y_n) \ldots T^{(k)}(y_1))} \underset{H_2}{\overset{H_1}{\underset{>}{\lessgtr}}} \tau.
\tag{2.40}
$$

The thresholds η and τ are selected so as to trade off the false alarm probability and the probability of a miss. It is shown in [Chen and Tong, 2001] that, under the usual channel conditions, the penalty for a miss detection is more severe than that for a false alarm, when a decorrelating receiver is used in the second stage of the detector. It is also suggested that an optimal threshold might be selected experimentally, due to the high complexity of the analysis.

Some numerical examples are presented in [Chen and Tong, 2001] to illustrate the benefits of using traffic prediction in the receiver design. Particularly, the focus is on comparing the performance of the one-shot UMPI test with the HMM tracker which uses the previous state estimates. For the numerical results, the state transition matrix is chosen to be

$$a^{(k)} = \begin{bmatrix} 0.99 & 0.01 \\ 0.10 & 0.90 \end{bmatrix}. \tag{2.41}$$

When a user is active, it will randomly generate a data packet with BPSK modulation at the nth time slot. No power control is considered and the amplitudes are fixed and equally spaced on a log scale in the range of 20 dB. The packet length is assumed to be 128 bits with error correcting capability for correcting up to 8 bit errors. The total number of users is 20 and the spreading gain is 31. The signature sequences' cross-correlation coefficients are in the range of -0.4839 to 0.3548 with a bell-shaped histogram.

In Fig. 2.8 the detection performance for two approaches are compared using receiver operating characteristic (ROC) curves. Unlike, the traditional ROCs, these are "ergodic ROCs", in the sense that the probability of false alarm and miss detection are computed by averaging over time, instead of using a statistical average based on repeated sampling. The performance of the hidden Markov model tracker is shown to be superior to that of the one-shot approach, especially in the region of interest (probability of detection close to 1).

In Fig. 2.9 the packet error probability is shown as a function of the SNR, for several different approaches, including the HMM tracker, the one-shot detector, the full model (in which no activity detection is performed and all users are considered active at all times), the simple matched filter receiver and the true model (perfect knowledge of activity is assumed). It can be seen that the HMM tracker has excellent performance, very close to that resulting from use of the true model.

6.   Medium Access Control for Multipacket Reception Networks

Traditionally, in the layered model approach, the medium access control is designed without detailed knowledge of the underlying physical layer, considering only simple collision models. In cross-layer design, the performance of the physical layer influences the MAC design, and in turn the access control may also have impact on the physical layer performance, as we have already seen in the previous two sections.

In this section we review results on MAC protocols that exploit the multipacket reception (MPR) capability of CDMA networks with mul-
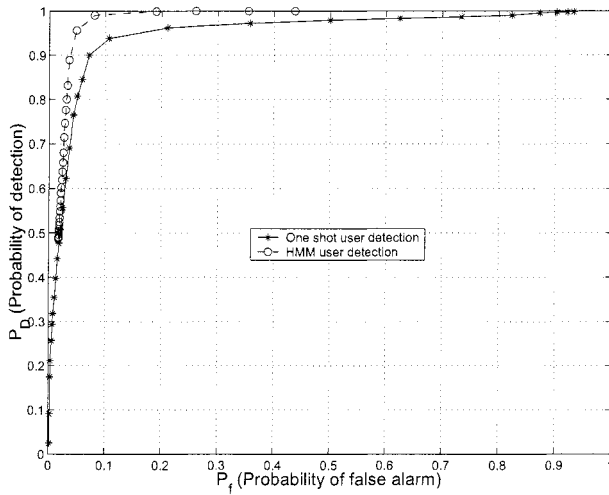
Figure 2.8. Ergodic receiver operating characteristics (ROCs) (reprinted with permission from [Chen and Tong, 2001])
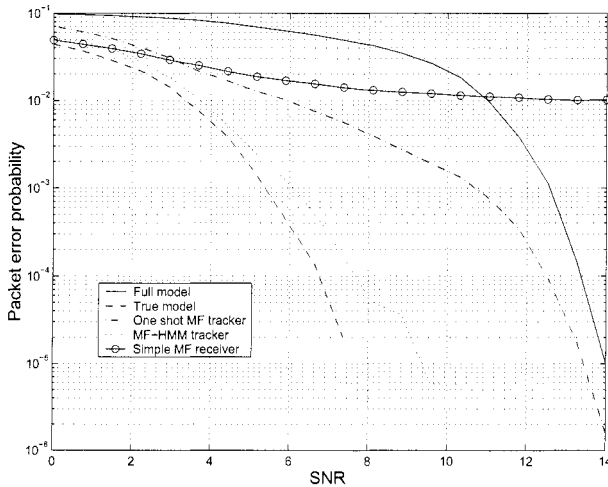

Figure 2.9. Packet error probability (reprinted with permission from [Chen and Tong, 2001])

tiuser receivers. In [Zhao and Tong, 2003, Zhao and Tong, 2004, Mergen and Tong, 2001, Mergen and Tong, 2002] the physical layer performance

is abstracted using an MPR matrix, which describes the MPR capability of a node. For slotted networks, this MPR matrix was proposed in [Ghez et al., 1988, Ghez et al., 1989] and is defined as

$$\mathbf{R} = \begin{pmatrix} R_{1,0} & R_{1,1} & ... & ... \\ R_{2,0} & R_{2,1} & R_{2,2} & ... \\ ... & ... & ... & ... \end{pmatrix}, \tag{2.42}$$

where $R_{i,j}$ is the conditional probability that j packets are correctly received, given that i packets are transmitted.

This MPR matrix is general enough to model the performance of different packet reception scenarios, ranging from the traditional collision channel $\mathbf{R}_0$, to the strongest MPR capability $\mathbf{R}_1$, for which all transmitted packets can be correctly received:

$$\mathbf{R}_0 = \begin{pmatrix} 0 & 1 & ... & ... \\ 1 & 0 & 0 & ... \\ ... & ... & ... & ... \end{pmatrix}. \tag{2.43}$$

$$\mathbf{R}_1 = \begin{pmatrix} 0 & 1 & ... & ... \\ 0 & 0 & 1 & ... \\ ... & ... & ... & ... \end{pmatrix}. \tag{2.44}$$

The focus of [Zhao and Tong, 2003, Zhao and Tong, 2004, Mergen and Tong, 2001, Mergen and Tong, 2002] is not on defining the MPR matrix for given physical layer characteristics, but on building efficient MAC protocols for a given MPR matrix model. The role of the MAC is to regulate the transmission of packets such that the network throughput is maximized. This can be achieved by allowing only an optimal subset of users to access the channel. The difficulty comes from the fact that, while the target number of users to access the channel can be computed, the exact number of users having packets to transmit is a random variable. The solution is to dynamically change the subset of active users, based on throughput performance. While this is a general access problem studied also for classical collision channels, even the simplest access protocol, such as slotted ALOHA [Bertsekas and Gallager, 1992], will need to be adapted to the MPR model [Ghez et al., 1988], and its performance will certainly benefit from better physical layer reception capabilities.

More sophisticated MAC algorithms for general MPR channels for cellular networks were analyzed in [Zhao and Tong, 2003, Zhao and Tong, 2004]. In [Zhao and Tong, 2003] the Multi-Queue Service Room (MQSR) protocol was proposed. In this protocol, the users that are allowed to access the system are first admitted into the so-called service room. These users will then be further scheduled for transmission dynamically, based
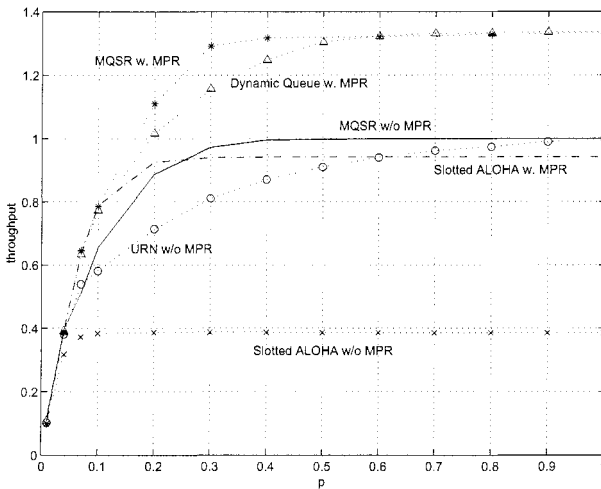
Figure 2.10. Throughput comparisons (reprinted with permission from [Tong et al., 2001])

on resource availability. In order to dynamically adjust the number of transmitting users, the service room is further divided into the access room and the waiting room. Only the users in the access room are actually allowed to transmit. If too many users are currently in the access room, some of them are pushed back into the waiting room. Conversely, more users can be allowed in the access room if more resource become available. This protocol involves a high computational cost, resulting from the need to update the joint distribution for all users' states. A simpler protocol, the dynamic queue protocol, is proposed in [Zhao and Tong, 2004], and is based on splitting the transmission time into time periods. The current time period is used for transmission of packets generated in the previous period. The optimal access set for a time period is determined by minimizing the length of the transmission period, given the probability that a user has a packet, and given the MPR matrix model.

Performance comparisons for the above two protocols with the slotted ALOHA protocol for channels with or wihout MPR capability, and with a simple urn scheme for collision channels are presented in Fig. 2.10 for 10 users in the network and for probability p that a user generates a packet within a given time slot interval. The urn protocol randomly picks a subset of users to access the channel such that the probability that there is one active user in the set is maximized.

From Fig. 2.10, we can make several observations:

1 The MPR capability greatly increases the network throughput, even for simple MAC protocols such as slotted ALOHA;

2 Better MAC protocols are more efficient even for simple collision channels; and

3 The network performance benefits from improvements at both the physical layer and the MAC layer.

We also note that this approach permits a certain design separation at different layers for cellular systems. The only information shared between the layers is the MPR matrix.

Things become more complicated for ad hoc networks, since not all the nodes have the same intended receiver and the access control performance is strongly inter-related with the routing protocol as well. The solution proposed in [Mergen and Tong, 2001, Mergen and Tong, 2002] decomposes the network into independent clusters, each containing a single receiver and its associated transmitters. This reduces the problem essentially to the cellular case, and therefore the previous MAC results can be applied directly. An extension to the dynamic queue protocol for ad hoc networks was proposed in [Mergen and Tong, 2001]. Here, a protocol based on receiver controlled transmissions (RCTs) is analyzed. Receiver controlled transmission is a combination of scheduling and random access. In the scheduling part, disjoint receiver nodes (not interfering with each other) and their neighborhoods are selected for transmission in the next m slots. This is achieved by symmetrically covering the network (a simple Manhattan network model is considered) with various tilings. The nodes in the center of each tile are selected as receivers and then the tilings are shifted so that the network goes through $(2r^2 + 2r + 1)$ states (r is the transmission radius), such that all nodes can be selected as receiver nodes. The network stays in each state for m slots. For each state, each receiver node chooses a subset of transmitters in its neighborhood from which it will receive packets in the next contention period. Since the number of transmitters holding packets is random, a dynamic queue protocol can be applied to this scenario as well. The neighborhood dimension and the transmission period for a given configuration (m) can be optimized as a function of the traffic load in the network. It can be seen that, although a minimum transmission radius is recommended for collisions channels, for a better MPR capability of the nodes, a higher connectivity radius results in higher network throughput. Another approach for access control in ad hoc networks was proposed in [Mergen and Tong, 2002], and is based on randomly

decomposing the network into clusters, based on an *a priori* selection of seeds for all nodes, under the assumption that each node knows the network topology up to two-hop distances and also all the seeds for up to second order neighbors. After decomposing the network, perfect local scheduling/polling is applied for managing the transmissions in a given neighborhood.

A similar approach for designing MAC in ad hoc networks (using cell splitting and MPR channel models) is presented in [Rodoplu and Meng, 2000]. Here it is also assumed that the transmission is time duplexed, and therefore a transmission schedule must be assigned to every cell. The network is divided into blocks, and each block is divided into cells. The assigned duplex schedule is designed such that nodes that fall within different cells can communicate with each other and with cells in adjacent blocks (all blocks have the same duplex schedule). It is assumed that nodes in the same cell will use a separate range of channels, since they would have the same duplex schedule. Signature sequence assignment to improve the network performance is also discussed in [Rodoplu and Meng, 2000]. Two types of interference may affect the network performance: multiple access interference and co-channel interference. To reduce the interference, the sequences are dynamically assigned as the users enter different blocks and different cells within the blocks. Similarly to the cellular concept, they are reused in blocks that are sufficiently far apart.

It it shown in [Rodoplu and Meng, 2000] that the proposed MAC protocol benefits from the MPR capability of multiuser detectors, and outperforms the commonly used 1-persistent CSMA (carrier sense multiple access) MAC protocol.

7. Routing and Multiuser Detection in Ad Hoc Networks

Routing in ad hoc networks has traditionally been studied as a means of providing multi-hop connections from any source node to its selected destination. On the other hand, more recent research treats routing as a resource allocation problem and shows that it influences both the energy consumption in the network, and the interference level, and thus strongly inter-relates with the performance of different protocols at different layers of the protocol stack. A very important QoS measure that is particularly influenced by routing is energy consumption. The concept of "energy aware routing" has recently been proposed for ad hoc networks, and routing protocols that minimize the energy consumption in the network have been developed. In this section, we focus on work related to performance enhancement by integration of routing and mul-

tiuser detection [Cai et al., 2002] and by integration of routing, multiuser detection and power control [Comaniciu and Poor, 2004b].

In [Cai et al., 2002], the authors propose routing protocols that minimize the average transmission power in a synchronous network using blind LMMSE receivers. The proposed routing protocols minimize a routing cost, based on a link cost measure that represents the average power consumption for that particular link. To derive this link cost measure, it is assumed that the nodes transmit with fixed powers and use blind LMMSE receivers and error correcting codes. For data services an ARQ (automatic repeat request) type protocol is also implemented, which allows for retransmissions of incorrectly received packets.

Based on the observation that the output of a linear MMSE receiver is well approximated by a Gaussian random variable [Poor and Verdú, 1997], the bit error rate of an LMMSE receiver can be approximately expressed as

$$P_b \approx Q(\sqrt{SIR}), \tag{2.45}$$

where SIR is determined from (1.36), and $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$.

If data is organized into packets of length L, and a code capable of correcting up to t errors per packet is used, the packet error probability can be expressed as

$$P_e = \sum_{i=t+1}^{L} \binom{L}{i} P_b^i (1 - P_b)^{L-i}. \tag{2.46}$$

Consequently, for a transmission power of P_k for node k, and an average power consumption from transmitting a data packet of αP_k (α is a network parameter), the average transmission power consumption for a transmission from k to a node j can be derived to be

$$P_{av} = \frac{\alpha P_k}{1 - P_e}, \tag{2.47}$$

where P_e is the error probability on the given link. We note that the link costs (2.47) are usually not symmetric, i.e., the link cost for (k, j) is generally different from the cost for (j, k).

For real time services, no retransmission is possible, and therefore the links selected for transmission must be reliable links, i.e., if the required BER is P_{req}, they must meet the condition

$$Q(\sqrt{SIR}) \leq P_{req} + \frac{t}{L}. \tag{2.48}$$

The protocol proposed by [Cai et al., 2002], known as MATPR (minimum average transmission power routing), can be implemented to be

either reactive or proactive. In proactive mode, each terminal maintains a routing table so that a route can be used immediately when a packet needs to be forwarded. In reactive mode, a route discovery process is initiated only on demand, and no routing table is maintained. Faster forwarding is achieved for proactive schemes at the expense of larger overhead for network updates. If the updates are infrequent, the proactive schemes are recommended.

For data services, the main implementation of a proactive protocol can be summarized as follows. Each terminal k should determine the average SIR with which each of its neighbors can be received. Nodes are considered to be neighbors if their received SIRs at the selected node permit demodulation. The packet error rate of each link destined to k is then determined, and this information will be broadcast periodically by k. Also, each node computes link costs (2.47) for all outgoing links, and a local link cost table is established to record the link costs. Periodically all nodes must broadcast their local link cost table to the network, so that every node can have an accurate global link cost table. Based on these cost tables, minimum average transmission power routes can be determined using Dijkstra's algorithm [Bertsekas and Gallager, 1992].

A similar approach is used for real time services, except that only reliable links are considered for the routing process, and a link cost is defined as the transmission power for the transmitting node.

For reactive protocols, each node will determine the neighbors' SIRs and compute the packet error rate of each link destined to it as before, but without broadcasting any information. When a node A attempts to discover a route to B, A broadcasts a route request packet which includes its transmission power and other routing information. If another node receives the packet, it will compute the average transmission power cost using its own transmission power, and also the packet error rate for the corresponding link. Then, the computed link cost and its transmission power are included in the packet and broadcast again. The process is repeated until the route request packet arrives at the destination node, which may receive multiple route request packets, and chooses the route with minimum cost. The destination node will then send a route reply packet to node A including the selected route list. When A receives the reply packet the data communication can begin, using the selected route.

Similarly, a reactive routing protocol can be implemented for real time services with the amendment that all the links selected for the route must be reliable. Thus, if another node receives a packet intended for route discovery, it will include its own information in the packet, only if its link to the source node is reliable.

In [Cai et al., 2002], the performance of the MATPR protocol was compared against the traditional shortest hop distance routing, and it was shown that significant power consumption savings can be achieved with MATPR.

While the results in [Cai et al., 2002] are obtained for fixed transmission powers for all nodes, in [Comaniciu and Poor, 2003c, Comaniciu and Poor, 2004b] it is shown that significant energy gain can be obtained if a joint power control and routing algorithm is implemented in networks using linear receivers.

In [Comaniciu and Poor, 2003c, Comaniciu and Poor, 2004b], joint power control, distributed power control and routing for CDMA ad hoc networks is proposed. Each node k can adjust its transmission power level (P_k), such that all transmissions originating at that node would meet their target SIRs. An ad hoc network consisting of K nodes is considered in the analysis. It is assumed that each node generates traffic to be transmitted towards a randomly chosen destination node. If traffic is relayed by a particular node, the transmissions for different sessions at that node are time multiplexed.

In order to characterize the data QoS measure, a data transmission model similar to that in [Goodman and Mandayam, 2000] is used. Data is transmitted in packets of length L, and a packet received in error is retransmitted until correctly received. Assuming that all the errors can be detected and that a packet is not relayed to its next destination node until it is correctly received from its previous transmission link, the utility per link of an arbitrary terminal k can be measured in the number of information bits correctly received per Joule of energy expended [Goodman and Mandayam, 2000],

$$U_k = \frac{mR\tilde{P}_c(\gamma_k)}{LP_k}, \qquad (2.49)$$

where $\tilde{P}_c(\gamma_k)$ is the approximate probability of correct reception of a packet [Goodman and Mandayam, 2000], γ_k is the target SIR, m is the number of information bits within a packet, R is the transmission rate for terminal k, $R = W/N$, and W is the system bandwidth. As before, N is the spreading gain.

For a particular example of a Gaussian channel with frequency shift keying (FSK) modulation, $\tilde{P}_c(\gamma_k)$ was selected in [Goodman and Mandayam, 2000] to be

$$\tilde{P}_c(\gamma_k) = (1 - 2BER_k)^M, \qquad (2.50)$$

where $BER_k = 0.5\exp(-\gamma_k/2)$. The approximation in (2.50) was used instead of $P_c(\gamma_k) = (1 - BER_k)^M$, which is the probability of correct

reception, in order to avoid the degenerate solution for which maximum utility is obtained for $\mathbf{p} = \mathbf{0}$ (i.e., when the power is turned off). Based on the definition in (2.50), the authors in [Goodman and Mandayam, 2000] showed that maximizing the utility leads to the following condition, which gives an optimal target SIR γ^*:

$$\frac{\partial U_k}{\partial p_k} = 0, \ k = 1, \ 2, \ , \dots, \ K \Rightarrow \gamma^* \frac{\partial \tilde{P}_c(\gamma^*)}{\partial \gamma^*} = \tilde{P}_c(\gamma^*). \tag{2.51}$$

If we define the link QoS measure for terminal k to be the energy consumed for the correct transmission of an information bit, E_b^k:

$$E_b^k = \frac{1}{U_k}, \tag{2.52}$$

then the energy per bit transmission for a particular link can be minimized by using the least amount of power that ensures the target SIR γ^* on that particular link.

Taking into account the above considerations, we express the link QoS requirement for an arbitrary link (k, j), $k, j = 1, 2, ..., K$ as

$$SIR_{(k,j)} \geq \gamma^*, \ \forall (k, j) \in S_a^r, \tag{2.53}$$

where S_a^r is the set of active links for the current routing configuration r, obtained using the routing protocol. The joint optimization problem at the network level can then be formulated as

$$\begin{aligned} & \text{minimize} \sum_{i=1}^{K} P_k \\ & \text{subject to} \\ & \quad SIR_{(k,j)} \geq \gamma^*, \ \forall (k, j) \in S_a^r \\ & \quad P_i \geq 0 \\ & \text{and } r \in \Upsilon, \end{aligned} \tag{2.54}$$

where Υ is the set of all possible routes. From (2.54) we can see that the optimal power allocation depends on the current route selection. On the other hand, for a given power allocation, efficient routing may reduce the interference, thus further decreasing the required energy-per-bit.

Power Control Issues

In a cellular setting, a minimal power transmission solution is achieved when all links achieve their target SIRs with equality. For an ad hoc network, implementation complexity constraints restrict the power control to adapt power levels for each node, and not for each active link. If multiple active transmission links start at node k (Fig. 2.11), then the worst

link must meet the target SIR with equality. If we denote the set of all outgoing links from node k by S_k^*, then the minimal power transmission conditions become

$$\min_{j \in S_k^*} SIR_j = \gamma^*, \ \forall k = 1, 2, ..., K. \tag{2.55}$$
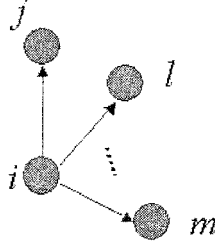

Figure 2.11. Multiple transmissions from node k

The achievable SIR for an arbitrary active link $(k, j) \in S_a^\tau$ can be expressed as

$$SIR_{(k,j)} = \frac{h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2 P_k}{\displaystyle\sum_{\ell=1, \ell \neq k, \ell \neq j}^{K} h_{(\ell,j)}(\mathbf{c}_k^T \mathbf{s}_\ell)^2 P_\ell + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)^2}, \tag{2.56}$$

where $h_{(k,j)}$ is the link gain for link (k, j), $\mathbf{c}_k$ is the filter vector for transmissions from node k, $\mathbf{s}_k$ is the signature sequence for node k, and σ^2 is the background noise level. $\mathbf{c}_k$ can be selected to be any linear receiver, but for the numerical results discussed below, an LMMSE filter is considered, which has the property of maximizing the SIR.

Condition (2.55) can then be expressed as

$$\min_{(k,j) \in S_k^*} \frac{h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2 P_k}{\displaystyle\sum_{\ell=1, \ell \neq k, \ell \neq j}^{K} h_{(\ell,j)}(\mathbf{c}_k^T \mathbf{s}_\ell)^2 P_\ell + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)^2} = \gamma^*. \tag{2.57}$$

From (2.57), the powers can be selected as

$$P_k = \max_{(k,j) \in S_k^*} \frac{\gamma^*}{h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2} \left[\sum_{\ell=1, \ell \neq k, \ell \neq j}^{K} h_{(\ell,j)}(\mathbf{c}_k^T \mathbf{s}_\ell)^2 P_\ell + \sigma^2 (\mathbf{c}_k^T \mathbf{c}_k)^2 \right]$$

$$= \max_{(k,j)} I_{(k,j)}(\mathbf{p}), \tag{2.58}$$

where $\mathbf{p}^T = [P_1, \ P_2, \dots, P_K]$.

Similarly to the cellular case, it can easily be shown that $I_{(k,j)}(\mathbf{p})$ is a standard interference function, and also $T(\mathbf{p}) = \max_{(k,j)} I_{(k,j)}(\mathbf{p})$ is a standard interference function. Thus, for a feasible system, an iterative power control algorithm based on

$$P_k(n+1) = T(\mathbf{p}(n)), \ \forall k = 1, \ 2, \dots, \ K, \tag{2.59}$$

converges to a minimal power solution [Yates, 1995], for both synchronous and asynchronous power updates. Since all the information required for the power updates can be estimated locally, the power control algorithm can be implemented distributively (see also [Ulukus and Yates, 1998a]).

Joint Power Control and Routing

The performance in this setting can be further improved by optimally choosing the routes as well. Finding the optimal routes to minimize the total transmission power over all possible configurations is an NP-hard problem. In [Comaniciu and Poor, 2003c], a suboptimal solution is proposed, based on power control, iterative power control and routing, which is shown to converge rapidly to a local minimum energy solution. Dijkstra's algorithm [Bertsekas and Gallager, 1992] with associated costs for the links is used for finding minimal energy routes. The cost for an arbitrary link (k, j) is determined as

$$Cost(k, j) = \begin{cases} P_k & \text{if } SIR_{(k,j)} \geq \beta^* \\ \infty & \text{if } SIR_{(k,j)} < \beta^* \end{cases} . \tag{2.60}$$

In order to estimate costs for links that are not currently active, the achievable SIRs for all links must be estimated. This requires that each node k updates a routing table which should contain the estimated link gains toward all the other nodes, $h_{(k,j)}$, $j = 1, 2, ..., K$, $j \neq k$, the transmitted powers of all nodes, P_j, $j = 1, 2, ..., K$, and the extended estimated interference at all the other nodes, defined as

$$\tilde{I}(k, j) = \sum_{\ell=1, \ell \neq k, \ell \neq j}^{K} h_{(\ell,j)}(\mathbf{c}_k^T \mathbf{s}_\ell)^2 P_\ell + h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2 P_k, \ j = 1, 2, ..., K, \ j \neq i. \tag{2.61}$$

Hence, the estimated SIR for link (k, j) can be expressed as

$$\widetilde{SIR}_{(k,j)} = \frac{h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2 P_k}{\left(\tilde{I}(k, j) - h_{(k,j)}(\mathbf{c}_k^T \mathbf{s}_k)^2 P_k \right) + \sigma^2(\mathbf{c}_k^T \mathbf{c}_k)} . \tag{2.62}$$

We note that the achievable SIR on any potential link (currently active or not) depends only on the current distribution of nodes, and on the current power assignment, and does not depend on the current assigned routes, and consequently does not change for new route assignments. This property is a result of the fact that multiple sessions are time-multiplexed at a node, and are all transmitted with the same power.
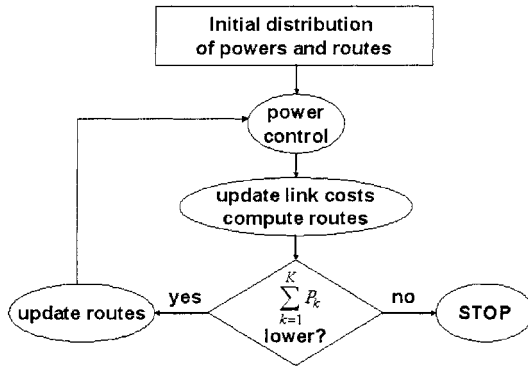

Figure 2.12.   Joint power control and routing algorithm

Starting from an initial distribution of powers and routes, and assuming that the system is feasible for the initial configuration, the joint power control and routing algorithm is summarized in Fig. 2.12. It was proved in [Comaniciu and Poor, 2004b] that the joint power control and routing algorithm converges to a locally minimal transmitted power solution. The achieved local minimum depends on the initial network configuration chosen.

For initialization of the joint protocol, an algorithm similar to that proposed in [Cai et al., 2002] is used: an initial distribution of powers is selected, then routes are determined by assigning link costs equal to the link utility (2.49), without imposing any SIR constraint. This initialization permits us to quantify the energy improvements of the joint optimization versus the initial starting point (with fixed and randomly chosen powers).

As a final observation, we mention that the solution can be improved with little increase in complexity if the algorithm is run several times using different random power initializations, and the best energy solution over all runs is determined.

Simulation results illustrate the performance of the joint protocol for an ad hoc network with 30 nodes, uniformly distributed over a square area of 200×200 meters. The spreading gain was seletected to be $N = 32$, the optimal target SIR was determined to be $\gamma^* \approx 12.5$, and the noise level was set to $\sigma^2 = 10^{-13}$, which corresponds approximately to the thermal noise power for a bandwidth of 1 MHz. Random initial transmission powers were selected, approximately 70 dB above the noise floor.

Figure 2.13 shows the initial distribution of powers, as well as the optimal power distribution after convergence.

(a)

(b)
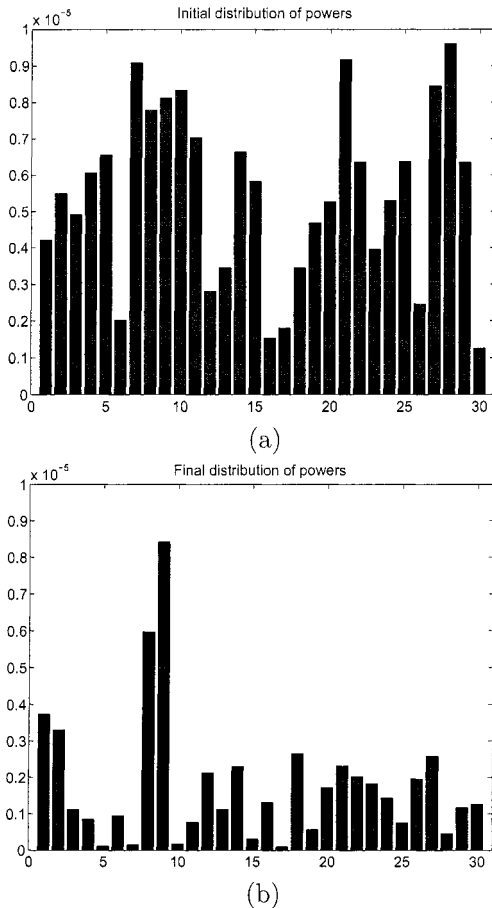
Figure 2.13. Distribution of powers versus node number: (a) initially, (b) after convergence

Figures 2.14 and 2.15 illustrate the performance of the proposed joint optimization algorithm. In Fig. 2.14, it can be seen that the total transmitted power in the network progressively decreases as the proposed algorithm iteratively optimizes power and routes. The values in Fig. 2.14 represent the total transmitted power obtained over a sequence of iterations: [power control, routing, power control, routing, power control].
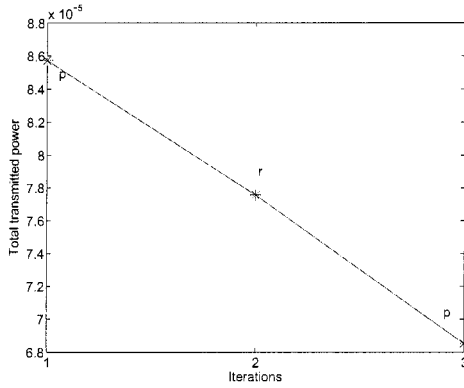
Figure 2.14.   Total transmission power

In Fig. 2.15, the achieved energy-per-bit is compared for the same experiment with the initial energy value (without power control). It can be seen that substantial improvements are achieved by the proposed joint optimization algorithm (approximately one order of magnitude).
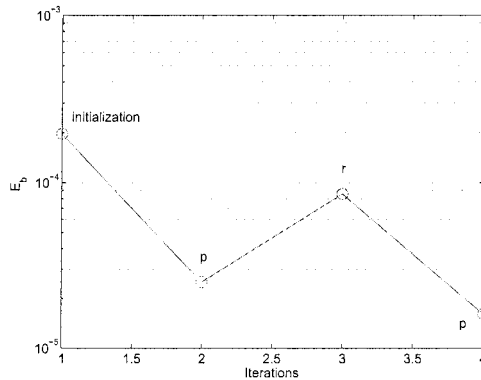
Figure 2.15.   Total energy consumption

As a final observation, we can see that, at the end of each iteration pair [routing, power control], the energy is further minimized. However, after new routes are selected, the powers are not yet optimized, so it is possible that previous routes might have better energy-per-bit performance (for the same power allocation, higher SIRs may improve the energy consumption).

8. Admission Control: General Framework

The role of the admission control is to restrict the number of users in the system so that QoS specifications for all users in the network can be met. In general, network and physical layer QoS requirements may be conflicting. For example, reducing the call blocking probabilities at the network layer results in admitting more users into the network, which consequently increases the level of interference in the system, and may lead to a degradation in the achieved physical layer QoS (SIR and packet access delay). Another way to reduce the call blocking probabilities, while preserving the requested QoS for the physical layer, is to queue incoming call requests, which will lead to an increased call connection delay. All these tradeoffs must be considered carefully when designing call admission control.

Considering a network with J classes of users, the interplay between physical and network layer QoS constraints can be illustrated by using an equivalent queueing system, as in Fig. 2.16 [Comaniciu and Poor, 2003a]. The average connection delays and the blocking probabilities can be derived using a queueing analysis. The service rate for each queue is varied by the admission control such that physical layer QoS requirements (SIR) can be met.

The classical approaches for admission control can be classified into three categories: the complete sharing policy, the threshold policy, and the optimal, state dependent, policy.

The complete sharing policy accepts new users into the network whenever the SIR condition can be met for all users, including the new call requesting connection. No preference is given to different classes of users. As a consequence, this policy cannot control the blocking probability or average call connection delay. The obtained performance is simply characterized by the statistical properties of the traffic. On the other hand, a threshold policy may be designed to accommodate performance constraints for different classes of users. Consider, for example, a network with two classes of users. Each incoming call request is buffered and the queues are served according to the admission control strategy, which means K_1 servers are allocated for class 1, and K_2 servers are allocated for class 2. The network performance is given by the performance of two
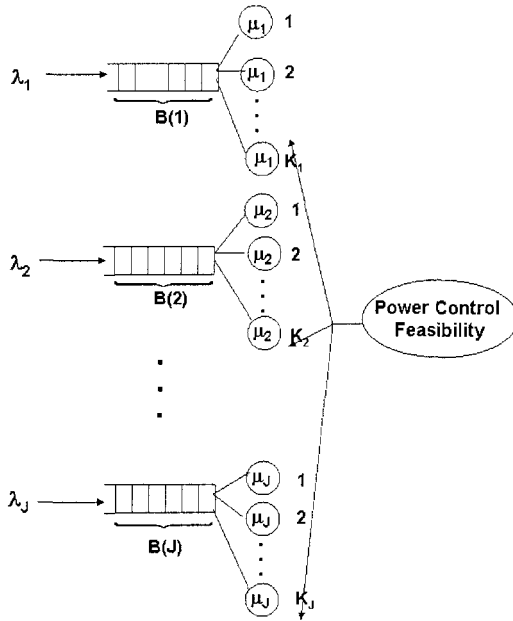
Figure 2.16. Equivalent queueing system (reprinted with permission from [Comaniciu and Poor, 2003a])

$M/M/K_i/B(i)$ queues, $i = 1, 2$, where $B(i)$ is the buffer length for queue i. Note that the threshold can be determined such that blocking probabilities for one of the queues can be met. Also, blocking probabilities and delays cannot be optimized independently in this admission policy, since they are inter-related, and are determined by the performance of the equivalent $M/M/K_i/B(i)$ queue.

The third option for admission control design is to change dynamically the resource allocation for different classes of users depending on the current state of the network. This state must be defined so as to reflect the current QoS in the network. A new call is admitted depending on the current state and on the next state to which the network will transition when the call is admitted. This new state must meet QoS specifications for all users. The admission control policy may be optimized with respect to QoS specifications. More details on how this admission control policy can be implemented in a network using multiuser receivers will follow in Chapter 4.

Regardless of which of these three strategies is used, the admission control has the role of managing the available physical resources among different classes of users, so that network QoS requirements can be met. This pool of resources is given by the physical layer capacity, which is determined such that SIR constraints for all users can be met. This means that the first step in designing admission control is to understand the physical layer performance. In the next chapter, we will discuss the capacity of power controlled networks using multiuser receivers, and then we will resume our discussion of admission control in Chapter 4.

Chapter 3

ASYMPTOTIC CAPACITY FOR WIRELESS NETWORKS WITH MULTIUSER RECEIVERS

Much of the previous work on multiuser detectors has focused on their ability to combat the near/far effect by rejecting the worst case interference. However, as we have seen in the previous chapter, the performance of systems using multiuser receivers can be significantly improved if used in conjunction with power control. Power control is only one of the available resource management tools that provide flexible QoS to different users in a network. QoS requirements can be supported at all layers of the protocol stack, and the system capacity characterizes all resources that are available to users, such that QoS specifications are met. Characterizing the network capacity for a system that uses multiuser receivers is difficult due to the fact that the SIR performance of such systems depends on the particular realization of the signature sequences for the users currently in the system. Moreover, the resource allocation problems encountered at higher networking layers have close interactions with the physical layer for such systems, and thus are more difficult to understand.

In this context, a breakthrough network capacity analysis has been proposed in [Tse and Hanly, 1999], for asymptotically large networks. In this work it was shown that, in a large system (large number of users and large number of degrees of freedom), a decoupling of the interference is possible for various linear receivers, such as the decorrelator and the LMMSE detector. Thus, each interferer can be characterized by a level of effective interference, and will occupy a certain effective bandwidth. Hence, the network capacity can be defined as a sum of these effective bandwidths. The work in [Tse and Hanly, 1999] serves as a foundation for characterizing the network performance for different types of networks (cellular and ad hoc networks), and has been extended to account

for various scenarios (e.g. random and optimal sequences, multiple rate transmission, and combinations of linear receivers with SIC). In this chapter, we will describe this asymptotic approach to characterizing the network capacity for power controlled systems using multiuser receivers.

1. Effective Bandwidths and Capacity for Linear Receivers in Cellular Networks

1.1 General Formulation for Synchronous Networks

We define the network user capacity to be the number of users that can be supported for a given QoS requirement. The analysis in [Tse and Hanly, 1999] assumes random, normalized spreading sequences, modeled as

$$\mathbf{s}_k = \frac{1}{\sqrt{N}}(s_{k1}, s_{k2}, \ldots, s_{kN}), \qquad (3.1)$$

where the s_{kj}'s are iid (independently identically distributed) random variables with zero means and unit variances. While in practice it is common to choose $s_{kj} \in \{-1, 1\}$, the analysis in [Tse and Hanly, 1999] allows for a more general model, with the mild restriction that $E\{s_{kj}^4\} < \infty$. Although the sequences are randomly chosen, it is assumed that they can be acquired by the receiver in a timely manner, i.e., any changes in the sequences occur much more slowly than the time scale required for aquisition.

The system model is restricted to a single power-controlled, synchronous cell, and the QoS measure is the achievable SIR. The capacity analysis has also been extended to consider the asynchronous transmission case [Kiran and Tse, 2000], which we will also discuss later on in this chapter.

We begin our discussion with the case of linear MMSE receivers; then we will compare these results with analogous results for the decorrelator and the matched filter receiver.

The main result in [Tse and Hanly, 1999] is based on properties of the limiting eigenvalue distribution for large matrices with random elements. More specifically, it has been shown in [Silverstein and Bai, 1995] that the empirical distribution of the eigenvalues for random matrices converges to a nonrandom distribution, in the limit, as the matrix dimensions increase without bound. As a consequence, even though for finite systems the achievable SIR is a random variable depending on the current assigned signature sequences (which are random), [Tse and Hanly, 1999] have shown that, in the limit, the SIR converges to a deterministic value.

This result is summarized in the following theorem given in [Tse and Hanly, 1999], for a power controlled, single cell synchronous CDMA system, in which an arbitrary user k is received with power P_k.

Theorem 3.1. *Let $\gamma_k(N)$ be the random SIR of the LMMSE receiver for user k, when the spreading gain in N. Then, $\gamma_k(N)$ converges to γ_k^* in probability as $N \to \infty$ and $\alpha = \frac{K}{N}$ is fixed, where γ_k^* is the unique solution to the equation*

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \alpha E_P \{I(P, P_k, \gamma_k^*)\}}, \tag{3.2}$$

where

$$I(P, P_k, \gamma_k^*) = \frac{P P_k}{P_k + P \gamma_k^*}. \tag{3.3}$$

$E_P\{.\}$ *denotes the expectation with respect to the limiting empirical distribution of the received powers.*

Heuristically, Theorem 3.1 says that, for a large system, the achievable SIR for user k can be expressed as

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \frac{1}{N} \sum_{j=i,\ j \neq i}^{K} I(P_j, P_k, \gamma_k^*)}. \tag{3.4}$$

The proof of Theorem 3.1 is based on applying random matrix results to the covariance matrix of the interference, and can be found in [Tse and Hanly, 1999].

Note that the above result should not be interpreted as stating that the interference is additive accross users, since the term $I(.)$ depends on the SIR, which in turn depends on the interference created by all users in the system. Nevertheless, Theorem 3.1 represents a powerful and simple analytical tool for power controlled networks using LMMSE receivers. In general, when only a verification is needed that user k meets its target SIR (γ_T), it suffices to check that

$$\frac{P_k}{\sigma^2 + \frac{1}{N} \sum_{j=1,\ j \neq k}^{K} I(P_j, P_k, \gamma_T)} \geq \gamma_T. \tag{3.5}$$

The term $I(P_j, P_k, \gamma_T)$ can be interpreted as the effective interference of user j on user k at a target SIR of γ_T.

Although no general explicit solution exists for the achievable SIR in (3.5), for the special case of equal powers, a closed form solution is given by

$$\gamma_k^* = \frac{(1-\alpha)P}{2\sigma^2} - \frac{1}{2} + \sqrt{\frac{(1-\alpha)^2 P^2}{4\sigma^4} + \frac{(1+\alpha)P}{2\sigma^2} + \frac{1}{4}}. \qquad (3.6)$$

The above closed form solution for the achievable SIR has also been obtained independently in [Verdú and Shamai, 1997].

Based on (3.6), we can compare the asymptotic theoretical results for equal power networks, with simulation results for finite networks (finite number of users and finite spreading gain). In Fig. 3.1, we reproduce simulation results from [Tse and Hanly, 1999].
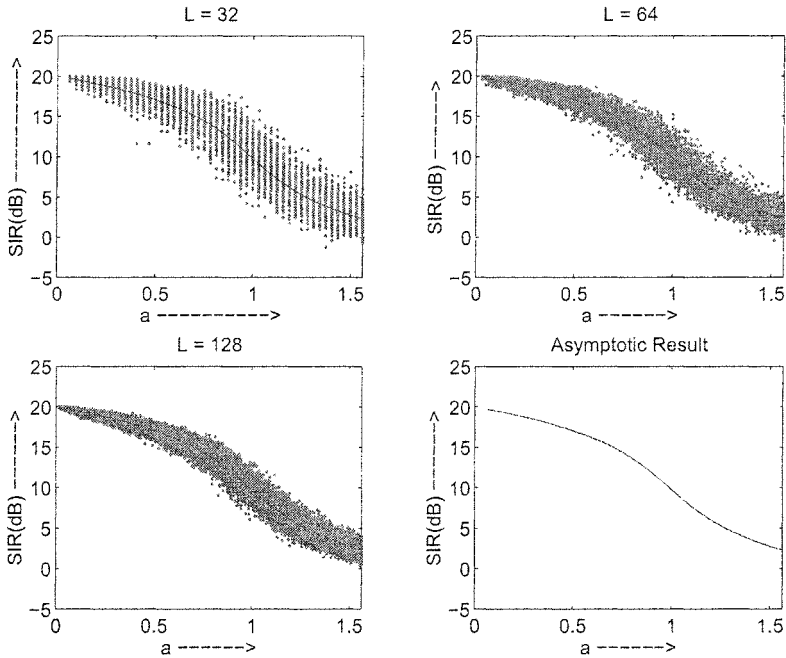
Figure 3.1. Finite network simulations (reprinted with permission from [Tse and Hanly, 1999])

Note that, as the spreading gain increases, the spread around the theoretical values becomes narrower ($\approx 1 - 2$ *dB*). However, for fixed processing gain, the spread is large for a large number of users in the network. As a consequence, we can say that the asymptotic analysis is a very good match for finite networks using large spreading gains and carrying relatively light loads. In general, the asymptotic SIR value is a very good approximation to the achievable mean SIR for finite networks,

but its variance increases with decreasing spreading gain, which makes high-rate networks more susceptible to modelling errors.

A more accurate performance analysis would be based on characterizing the outage probability in the network (i.e., the probability that the achieved SIR is below the target). In order to compute the outage probability, we would need to know the actual SIR distribution at the output of the receiver. This is often very difficult to characterize. In [Honig and Veerakachen, 1996], the SIR distribution was characterized using simulation results, while more recent work [Kim and Honig, 1998] and [Tse and Zeitouni, 2000] assume a Gaussian distribution for the output SIR, based on the central limit theorem, and derive closed form expressions for the SIR variance. It was shown in [Tse and Zeitouni, 2000] that the standard deviation for the SIR distribution decreases as $1/\sqrt{N}$. The same result can be proved for the decorrelator. Simulation results show very good agreement for the standard deviation approximation, but an overly pessimistic result for computing the 1% outage probability, which implies that the Gaussian approximation is not very accurate for predicting the tail of the SIR distribution. A more accurate approximation, which assumes that the SIR has a beta distribution, has been proposed independently for the decorrelator in [Tse and Zeitouni, 2000] and [Müller et al., 1997]. The SIR for the decorrelator exhibits similar convergence properties towards a deterministic constant as that of the LMMSE receiver in the asymptotic case. Asymptotic convergence properties for the decorrelator are summarized in the following theorem from [Tse and Hanly, 1999]:

Theorem 3.2. *Let $\gamma_k(N)$ be the random SIR of the decorrelator receiver for user k, when the spreading gain is N. Then, $\gamma_k(N)$ converges to γ_k^* in probability as $N \to \infty$ with $\alpha = \frac{K}{N}$ fixed, where γ_k^* is given by*

$$\gamma_k^* = \begin{cases} \frac{P_k(1-\alpha)}{\sigma^2}, & \alpha < 1, \\ 0 & \alpha \geq 1. \end{cases} \qquad (3.7)$$

To better characterize the system performance, the notions of effective interference and effective bandwidth were introduced in [Hanly and Tse, 1999]. The effective interference represents the effective level of interference that can be ascribed to a user for its successful demodulation. The effective bandwidth of user k (e_k) can be defined such that all users can meet their SIR requirements if and only if the sum of their effective bandwidths is less than the spreading gain of the system, i.e.,

$$\sum_{k=1}^{K} e_k \leq N. \qquad (3.8)$$

By analyzing the SIR expression in (3.7) we note that the effective interference seen by user k is equal to P_k/γ_k^* and does not depend on the interferers' powers. In Fig. 3.2 a comparison of the effective interference for the linear MMSE, the decorrelator and the matched filter receivers is presented [Hanly and Tse, 1999, Tse and Hanly, 1999].
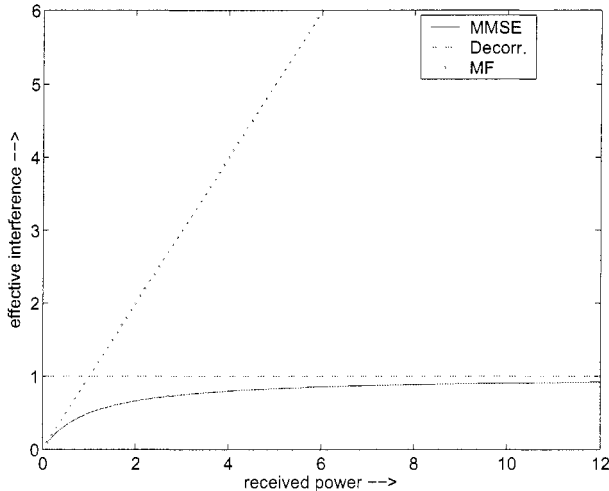

Figure 3.2.   Effective interference for linear receivers

From (3.7), it is straightforward to show that the SIR target γ_k for an arbitrary user k can be met if

$$\alpha \leq 1, \tag{3.9}$$

and

$$P_k \geq \frac{\gamma_k \sigma^2}{1 - \alpha}. \tag{3.10}$$

Consequently, we can see from (3.8) and (3.9) that the effective bandwidth of the decorrelator is limited to $e_{dec} = 1$, irrespective of the powers of the interferers.

It can be conjectured that better capacity performance can be achieved by the linear MMSE receiver, due to its property of maximizing the SIR. In what follows, we confirm this conjecture by presenting capacity results for power controlled networks using linear MMSE receivers. The main result in [Tse and Hanly, 1999] for characterizing the network capacity is based on SIR feasibility condition as follows.

Theorem 3.3. *The SIR requirements (γ^*) for all users in the network can be satisfied when $N \to \infty$ with $\alpha = \frac{K}{N}$ fixed, if and only if*

$$\alpha < \frac{1 + \gamma^*}{\gamma^*}. \tag{3.11}$$

Furthermore, the minimal power solution is achieved when the received powers for all users are

$$P_{mmse}(\gamma^*) = \frac{\gamma^* \sigma^2}{1 - \alpha \frac{\gamma^*}{1 + \gamma^*}}. \tag{3.12}$$

It can readily be seen from (3.8) and (3.11) that the effective bandwith for the LMMSE receiver is $e_{mmse} = \frac{\gamma^*}{1 + \gamma^*}$. For comparison purposes, we also derive the effective bandwidth for the matched filter receiver to be $e_{mf} = \gamma^*$. A comparison of the effective bandwidths for the three linear receivers is given in Fig. 3.3.
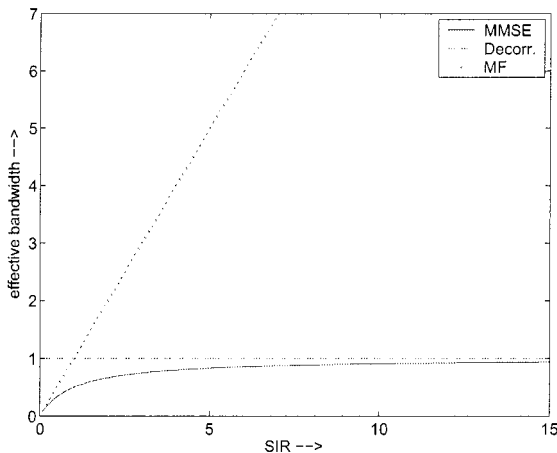
Figure 3.3. Effective bandwiths for linear receivers

General capacity results can be obtained for all three receivers when J classes of users are present in the network, having different SIR requirements, γ_j, $j = 1, \ldots, J$. The simplest case is the matched filter receiver, for which the capacity region is given as

$$\sum_{j=1}^{J} \alpha_j \gamma_j \leq 1, \tag{3.13}$$

where $\alpha_j = \frac{K_j}{N_j}$. If a maximum transmission power $\overline{P}_i$ is imposed for an arbitrary class i, the capacity region becomes

$$\sum_{j=1}^{J} \alpha_j \gamma_j \leq \min_{1 \leq i \leq J} \left[1 - \frac{\gamma_i \sigma^2}{\overline{P}_i} \right]. \tag{3.14}$$

For the decorrelator, (3.9) can be straightforwardly extended for the multi-class case, with or without power constraints as

$$\sum_{j=1}^{J} \alpha_j \leq 1, \tag{3.15}$$

or

$$\sum_{j=1}^{J} \alpha_j \leq \min_{1 \leq i \leq J} \left[1 - \frac{\gamma_i \sigma^2}{\overline{P}_i} \right], \tag{3.16}$$

respectively.

Finally, for the LMMSE receiver, for large networks, $\frac{\gamma_j}{P_j}$ is constant, and therefore the required received power for an arbitrary class i can be specified as

$$P_{mmse}(i) = \frac{\gamma_i \sigma^2}{1 - \sum_{j=1}^{J} \alpha_j \frac{\gamma_j}{1+\gamma_j}}, \quad i = 1, 2, \ldots, J. \tag{3.17}$$

Consequently, the network capacity region becomes

$$\sum_{j=1}^{J} \alpha_j \frac{\gamma_j}{1 + \gamma_j} < 1, \tag{3.18}$$

and if power constraints are imposed:

$$\sum_{j=1}^{J} \alpha_j \frac{\gamma_j}{1 + \gamma_j} < \min_{1 \leq i \leq J} \left[1 - \frac{\gamma_i \sigma^2}{\overline{P}_i} \right]. \tag{3.19}$$

1.2   Partial Hybrid Networks

The capacity advantage of multiuser receivers is evident from Figs. 3.2 and 3.3. However, this advantage comes at the price of higher implementation complexity. This disadvantage becomes more significant for multimedia networks supporting bursty traffic, which requires frequent filter coefficient updates, thus substantially increasing the implementation complexity. One possible approach to overcome this disadvantage is

to use simple matched filter receivers for real-time streams that require on-line filter adaptation and detection and are less sensitive to detection errors (voice traffic is a good example). On the other hand, data users would greatly benefit from using more advanced receivers. Further complexity reduction can be achieved if the real-time users are perceived by the data users' receivers as background noise. This reduces the update frequency for the data users' receivers as well, and it is thus suitable for simple access control design. Some advantages of the partial hybrid receivers have already been discussed in Chapter 2 in the context of MAC design. Such mixed receivers networks (partial hybrid systems) are analyzed in [Comaniciu, 2002], and the corresponding capacity results are presented in the following paragraphs.

To evaluate the asymptotic capacity, the effects of data on voice performance, as well as the effects of voice on data performance should be determined. The bidimensional asymptotic capacity is obtained as a result of the intersection of two QoS requirements for the system: the SIR target guarantee for voice and the SIR target guarantee for data.

It can be shown that, to meet their target SIRs (γ_v^* and γ_d^*, respectively), all voice users and all data users must transmit with the equal powers P_v and P_d, respectively. Denoting as κ the ratio of the data power to the voice users' power, the following feasibility theorems for voice and data, respectively, hold.

Theorem 3.4. *In a CDMA system in which voice users employ matched filters and data users have LMMSE receivers, denoting $\alpha_v = K_v/N$ and $\alpha_d = K_d/N$ as the number of voice users and data users per dimension, and γ_v^* and γ_d^* as the target SIR requirements for voice and data, respectively, a distribution of received powers exists such that the target SIRs for all voice users are met, if and only if*

$$\gamma_v^* \alpha_v + \kappa \gamma_v^* \alpha_d < 1. \tag{3.20}$$

Moreover, if (3.20) holds, the minimum voice power solution is given by

$$P_v^{min} = \inf\{P : SIR_v(P) \geq \gamma_v^*\} = \frac{\gamma_v^* \sigma^2}{1 - \gamma_v^* [\alpha_v + \kappa \alpha_d]}, \tag{3.21}$$

where SIR_v represents the achieved SIR for all voice users.

Theorem 3.5. *In a CDMA system in which voice users employ matched filters and data users employ LMMSE receivers, denoting $\alpha_v = K_v/N$ and $\alpha_d = K_d/N$ as the number of voice users and data users per dimen-*

sion, and γ_v^* and γ_d^* as the target SIR requirements for voice and data, respectively, a distribution of received powers exists such that the target SIRs for all data users are met, if and only if:

$$\alpha_d < \frac{1 + \gamma_d^*}{\gamma_d^*} \left(1 - \frac{1}{\kappa} \gamma_d^* \alpha_v \right). \tag{3.22}$$

Moreover, if (3.22) holds, then the minimum voice power solution is given by:

$$P_d^{min} = \inf\{Q : SIR_d \geq \gamma_d^*\} = \frac{\sigma^2 \gamma_d^*}{1 - \frac{1}{\kappa} \gamma_d^* \alpha_v - \frac{\gamma_d^*}{(1 + \gamma_d^*)\alpha_d}}, \tag{3.23}$$

where SIR_d represents the achieved SIR for all data users.

The proof for Theorem 3.5 relies on the following lemma, which makes the results in [Tse and Hanly, 1999] directly applicable.

Lemma 3.1. *In a partial hybrid multiuser detector CDMA system, where data users employ multiuser receivers built using only knowledge of data signature sequences, and voice users have conventional receivers, we have*
$\mathbf{E}\{(\mathbf{c}^T \mathbf{s})^2\} = \frac{1}{N} \mathbf{E}\{(\mathbf{c}^T \mathbf{c})\}$ *for any data filter vector $\mathbf{c}$ and any voice signature sequence $\mathbf{s}$.* It follows from Lemma 3.1, that such a partial hybrid multiuser detector CDMA system is equivalent from the data users' performance point of view to an all data system employing multiuser detectors, and operating with an enhanced noise power $\Sigma^2 = (\sigma^2 + \frac{K_v}{N} P)$.

Based on the power control feasibility conditions in Theorems 3.4 and 3.5, the hybrid system capacity expression is thus given in the following theorem.

Theorem 3.6. *The bidimensional capacity of a CDMA system having voice users employing matched filter receivers and data users employing partial LMMSE receivers, is expressed as*

$$(\alpha_v, \alpha_d) = \begin{cases} (\alpha_v, \Lambda_d), & \text{voice and data system} \\ \left(\frac{1}{\gamma_v^*}, 0 \right), & \text{voice only system} \\ (0, 1), & \text{data only system} \end{cases} \tag{3.24}$$

where

$$\Lambda_d = min \left\{ \frac{1 + \gamma_d^*}{\gamma_d^*} \left(1 - \frac{1}{\kappa} \gamma_d^* \alpha_v \right), \frac{1}{\kappa} \left(\frac{1}{\gamma_v^*} - \alpha_v \right) \right\}.$$

Recall that the transmission powers for voice and data users are constrained to satisfy $P_d = \kappa P_v$. The system performance will depend on

the particular value of κ chosen. While both voice and data users can be constrained to transmit with minimal powers, this will further decrease the system capacity. If both voice and data users are transmitting with the minimal powers necessary to achieve their target SIRs, and given the constraint that their ratio is equal to κ, the system capacity becomes:

$$(\alpha_v, \alpha_d) = \begin{cases} (\alpha_v, \Lambda_d^*), & \text{voice and data system} \\ \left(\frac{1}{\gamma_v^*}, 0\right), & \text{voice only system} \\ (0, 1), & \text{data only system} \end{cases} \quad (3.25)$$

where

$$\Lambda_d^* = min\left\{ \frac{1}{\kappa \gamma_v^*} \frac{1 + \gamma_d^*}{\gamma_d^*} \left(1 - \kappa \frac{\gamma_v^*}{\gamma_d^*}\right), \frac{1 + \gamma_d^*}{\gamma_d^*} \left(1 - \frac{1}{\kappa}\gamma_d^* \alpha_v\right), \frac{1}{\kappa}\left(\frac{1}{\gamma_v^*} - \alpha_v\right)\right\}.$$

Figures 3.4 (a) and (b) illustrate the system capacity for two cases: without power constraints and with power constraints. It can be seen from these examples that the capacity is reduced if power constraints are enforced.

Similar capacity derivations can be applied for the partial hybrid decorrelator case, based on Lemma 3.1 and the results of [Tse and Hanly, 1999]. Asymptotic capacity results for the partial hybrid decorrelator are summarized in the following paragraphs.

Lemma 3.2. *The random SIR performance of an arbitrary data user in an integrated CDMA system in which real-time traffic employs matched filters and data users employ partial decorrelator receivers, converges as $K_v \to \infty$, $K_d \to \infty$, $L \to \infty$, and $\alpha = (K_v + K_d)/N = \alpha_v + \alpha_d$, to γ_d^* given by*

$$\gamma_d^* = \frac{P_d(1 - \alpha_d)}{\sigma^2 + \alpha_v P_v}. \quad (3.26)$$

Theorem 3.7. *The bidimensional capacity of a CDMA system having voice users employing matched filter receivers and data users employing partial decorrelating receivers, is given by*

$$(\alpha_v, \alpha_d) = \begin{cases} (\alpha_v, \Lambda_d), & \text{voice and data system} \\ \left(\frac{1}{\gamma_v^*}, 0\right), & \text{voice only system} \\ (0, 1), & \text{data only system} \end{cases} \quad (3.27)$$

where

$$\Lambda_d = min\left\{\left(1 - \frac{1}{\kappa}\gamma_d^* \alpha_v\right), \frac{1}{\kappa}\left(\frac{1}{\gamma_v^*} - \alpha_v\right)\right\}.$$
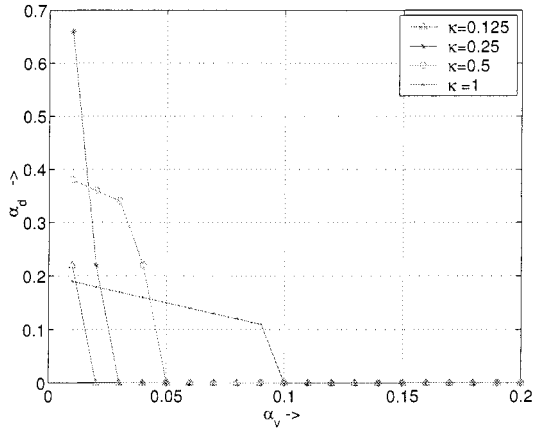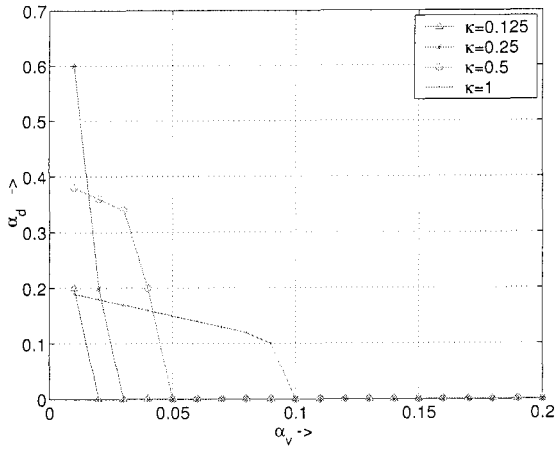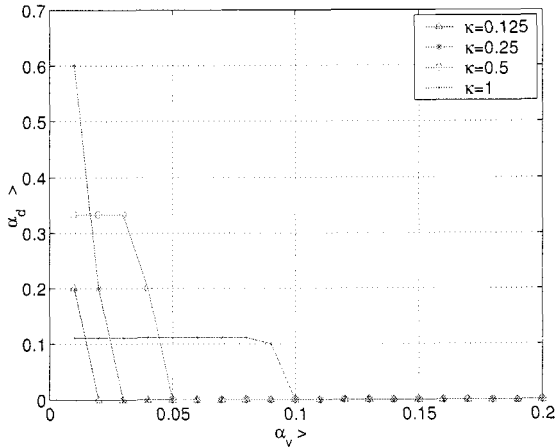
(a)
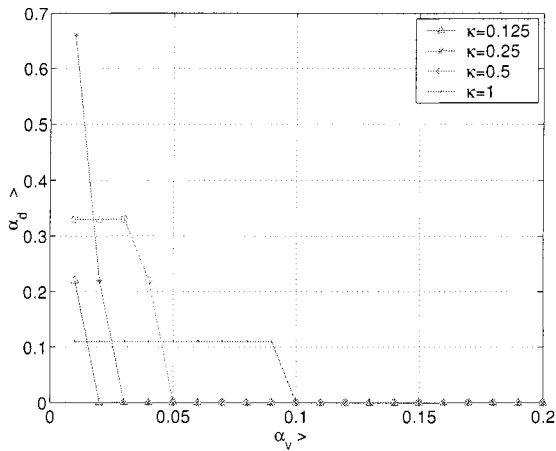
(b)

Figure 3.4. Bidimensional capacity for the $H - MMSE^{(p)}$ system: (a) No power constraints (b) Minimum power transmission for both voice and data and power ratio fixed to κ

The system capacity is illustrated in Fig. 3.5 (a) for different values of κ.

(a)

(b)

Figure 3.5. Bidimensional capacity for the $H-D^{(p)}$ system: (a) No power constraints (b) Minimum power transmission for both voice and data and power ratio fixed to κ

If data power constraints apply (both voice and data users are constrained to transmit with minimal powers), the system capacity becomes:

$$(\alpha_v, \alpha_d) = \begin{cases} (\alpha_v, \Lambda_d^*), & \text{voice and data system} \\ \left(\frac{1}{\gamma_v^*}, 0\right), & \text{voice only system} \\ (0, 1), & \text{data only system} \end{cases} \qquad (3.28)$$
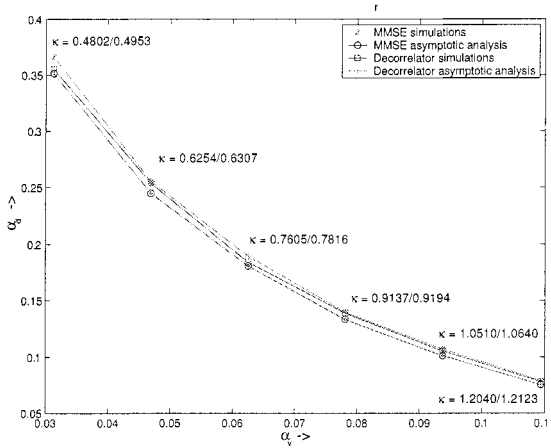
Figure 3.6. Partial hybrid LMMSE and decorrelator: simulations and asymptotic analysis

where

$$
\Lambda_d^* = min \left\{ \left(1 - \kappa \frac{\gamma_v^*}{\gamma_d^*}\right) \frac{1}{\kappa \gamma_v^* \left(1 - \frac{1}{\gamma_d^*}\right)}, \left(1 - \frac{1}{\kappa}\gamma_d^*\alpha_v\right), \frac{1}{\kappa}\left(\frac{1}{\gamma_v^*} - \alpha_v\right) \right\}.
$$

We note that the performance of the partial hybrid systems depends on the particular value of κ chosen. For a higher capacity, κ can be optimized as a function of the number of voice users currently in the system. In Fig. 3.6 we present asymptotic and simulation results to illustrate this for the partial hybrid LMMSE and partial hybrid decorrelator systems. Simulation results are shown for finite systems: finite number of users and finite spreading gains. No power constraints are enforced. A spreading gain of 128 is used for both voice and data users, data SIR targets are 10 and voice SIR targets are 5. The simulation results are averaged over different signature sequence realizations. The ratio κ of data power to voice power is determined from the simulations. The corresponding values are marked on the figure for the LMMSE case and for the decorrelator case, in the order LMMSE/decorrelator.

As a first observation, we can see that the analytical results are very close to the simulation results even for moderate spreading sequence lengths and low numbers of users in the system. As expected, the obtained power ratios from simulations are chosen such that the data powers predominate if more voice users are present in the system, and are lower than voice powers if there are only a few voice users in the sys-

tem. Another interesting observation is that the LMMSE detector and the decorrelator perform identically for the α_v range considered in the simulations. This is due to the fact that, for both cases, the capacity is restricted by the voice power control feasibility condition for the parameter values chosen for simulation. This suggests that, if the partial hybrid decorrelator is used in conjunction with access control, it would give performance similar to that of the partial hybrid LMMSE with a further reduction in complexity: filter coefficients do not depend on changes in the noise level, which includes changes in voice interference power. Recall that a joint access control algorithm for networks using multiuser receivers has been discussed in detail in the previous chapter.

1.3   Optimal Signature Sequences

While the previous sections have considered the case in which normalized, random signature sequences are assigned to users, the performance can potentially be improved by optimizing the signature sequence selection as well. The QoS feasibility condition translates now into finding signature sequences and allocating transmission powers for all users, such that they will meet their target SIRs. In [Viswanath et al., 1999], the capacity of such networks is investigated for optimal signature sequences for all users. Two linear receivers are compared: the linear MMSE receiver, and the matched filter receiver.

A surprising result is demonstrated in [Viswanath et al., 1999]: while the LMMSE receiver performance is asymptotically unchanged for optimal sequences compared to the random assignment case, the matched filter receiver performs as well as the LMMSE receiver when optimal signature sequences are assigned. This result holds for the case when there are no power constraints. If power constraints are imposed, then optimal signature assignment, corresponding to WBE (Welch-bound-equality) sequences, gives better performance for both linear receivers considered. Since the unconstrained power solution for both LMMSE and matched filters using WBE sequences gives identical capacity as the LMMSE with random sequences case, in what follows we present capacity results and minimum power solutions for LMMSE and matched filter (MF) receivers when power constraints are imposed [Viswanath et al., 1999].

Theorem 3.8. *K users, having an SIR requirement of γ and using WBE sequences, are admissible in a network using linear MMSE or matched filter receivers if and only if*

$$K < N \left(1 + \frac{1}{\gamma} - \frac{\sigma^2}{\overline{P_t}} \right), \tag{3.29}$$

where N is the spreading gain, and $\overline{P_t}$ is the maximal transmitted power constraint.

The minimal power solution corresponds to equal received powers for all users:

$$P_k \equiv P_{opt} = \frac{N\gamma\sigma^2}{N(1+\gamma) - K\gamma} \leq \overline{P_t}, \ k = 1, \dots, K. \qquad (3.30)$$

For the power constrained case, again denoting $\alpha = K/N$, we can compare the capacity in (3.29), with an equivalent formula for the case of linear MMSE receivers using random sequences (derived in Section 1 of this chapter): $\alpha < 1 + \frac{1}{\gamma} - (1+\gamma)\frac{G\sigma^2}{\overline{P_t}}$. It can easily be seen that the network with WBE sequences performs better than this random signature case, even for the LMMSE receiver, i.e., the use of WBE sequences leads to a smaller power requirement for the same network capacity, or conversely, for equal received powers it achieves a higher capacity.

More results for multi-class networks, a downlink analysis, as well as detailed derivations and proofs can be found in [Viswanath et al., 1999].

1.4   Multipath Fading Channels

In a fast multipath fading environment, it is very likely that channel estimates will not be perfect, and channel estimation errors will limit the capacity gains obtained using multiuser detectors. A problem that occurs in multipath channels is that the effective signature sequence of a user is characterized by both energy and direction. As a consequence, if the channel is not perfectly known, the effective signature sequence is not known, and it cannot be cancelled by the multiuser receiver. While solutions to this problem have been proposed in the literature, in which the received signature sequence is estimated, and then the receiver is built for the effective set of signature sequences in the network, the estimation errors will clearly adversely impact the performance. Several questions arise naturally in connection with such solutions. In particular, if some measure of the channel estimation errors is available, how can this be factored into the receiver design? And, how would such networks perform?

Answers to these questions for asymptotically large networks using linear multiuser receivers are provided in [Evans and Tse, 2000]. In [Evans and Tse, 2000] the authors assume that any arbitrary user k's channel is not perfectly known; instead, it is statistically characterized by first and second order moments: the average link gain $|\bar{h}_k|^2$, and the estimation error variance ξ_k^2. This paper shows that the variance ξ_k^2 does not depend on the path index if the average power per path is the same for all paths. The model assumes that intersymbol interference can be

neglected (i.e., the delay spread of the channel is small compared to the symbol duration), and that the time delays of the multipath components are known, so that only the path gains need to be estimated.

An implicit assumption for the channel model is that it is conditioned on the slower fading (free space path loss and shadow fading), which does not affect the received power over the time scales of interest. The effects of slow fading are absorbed into the attenuated transmitted power, defined for user k as

$$P_k = z_k P_k^t, \ k = 1, \ 2, \ldots, \ K, \qquad (3.31)$$

where z_k is the path loss due to free space loss and shadow fading, P_k^t is the transmitted power and K is the total number of users in the system.

For simplicity, the analysis in [Evans and Tse, 2000] assumes that the average received powers of all paths of each user are the same, so that for user k, the average received power of path l ($\overline{p}_{k,l}$) can be expressed as

$$\overline{p}_{k,l} = \frac{\overline{p}_k}{L} = P_k |\bar{h}_{kl}|^2, \qquad (3.32)$$

where $\overline{p}_k$ is the total average received power for user k, and L is the total number of paths.

The analysis in [Evans and Tse, 2000] is decoupled into two parts according to the receiver structure, which comprises a data estimator and a channel estimator. The data estimator is a "one-shot" linear receiver which accounts only for information from previously detected symbols through the coupling with the channel estimator. The channel estimator uses training symbols and it is shown that it can provide accurate estimates, as long as its window length is at least equal to the number of resolvable paths.

The main result for an LMMSE network is summarized as follows.

Theorem 3.9. *The SIR achieved by an arbitrary user (say the kth one) in an LMMSE network can be expressed as*

$$SIR_k = \frac{P_k \sum_{l=1}^{L} |\bar{h}_{kl}|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d} = \frac{P_k |\bar{h}_k|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d}, \qquad (3.33)$$

where γ is the unique fixed point in $(0, \infty)$ that satisfies

$$\gamma_d = \left[\sigma^2 + \frac{1}{N} \sum_{j=1, j \neq k}^{K} I_j^*(L, \gamma_d, \xi_j^2, |\bar{h}_j|^2) \right]^{-1}, \qquad (3.34)$$

with $I_j^*(L, \gamma_d, \xi_j^2, |\bar{h}_j|^2) = (L - 1)I(\xi_j^2 P_j, \gamma_d) + I(P_j(\xi_j^2 + |\bar{h}_j|^2), \gamma_d)$, and $I(p, \gamma_d) = \frac{p}{1 + p \gamma_d}$. Here, for $j = 1, \ 2, \ \ldots, \ K$, $|\bar{h}_j|^2$ is the equivalent

estimated average power gain for user j, defined as

$$|\bar{h}_j|^2 = \sum_{l=1}^{L} |\bar{h}_{jl}|^2, \tag{3.35}$$

where $\bar{h}_{j,l}$ is the average link gain for the lth path of user j.

It can be seen that the overall effect of an interferer is equivalent to $L-1$ users with power $\xi_k^2 P_k$ and one with power $\sum_{l=1}^{L} |\bar{h}_{kl}|^2 P_k + \xi_k^2 P_k$. When the channel is known perfectly, the interferer looks like a single interferer with power $P_k \sum_{l=1}^{L} |\bar{h}_{kl}|^2$, whereas as the uncertainty increases, the interferer is perceived as L separate interferers each having reduced power $\bar{p}/L$. However, because of the convexity of the effective interference function, L low power interferers are more damaging than one high power interferer (with the same total power), and this is the reason why significant performance degradation occurs if channel estimates are poor.

Similar capacity results are derived for the post-combining decorrelating receiver. This decorrelating receiver has two stages: the first stage decorrelates the users by considering each path as a separate interferer; the second stage combines the L outputs from different paths for the same user. For this receiver, we have the following.

Theorem 3.10. *The SIR for the decorrelator converges almost surely to the value*

$$SIR_k = \frac{P_k \sum_{l=1}^{L} |\bar{h}_{kl}|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d} = \frac{P_k |\bar{h}_k|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d}, \tag{3.36}$$

where

$$\gamma_d = \frac{1 - \alpha L}{\sigma^2} = \left[\sigma^2 + \alpha L \frac{\sigma^2}{1 - \alpha L} \right], \tag{3.37}$$

and $|\bar{h}_k|^2$ is from (3.35).

For comparisons purposes, results for networks using matched filter receivers are also presented in [Evans and Tse, 2000].

Theorem 3.11. *The SIR for the matched filter receiver converges almost surely to the value*

$$SIR_k = \frac{P_k \sum_{l=1}^{L} |\bar{h}_{kl}|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d} = \frac{P_k |\bar{h}_k|^2 \gamma_d}{1 + P_k \xi_k^2 \gamma_d}, \tag{3.38}$$

where γ_d is the unique fixed point in $(0, \infty)$ that satisfies

$$\gamma_d = \left[\sigma^2 + \frac{1}{N} \sum_{j=1, j \neq k}^{K} I_j^*(L, \gamma_d, \xi_j^2, |\bar{h}_j|^2). \right]^{-1}, \tag{3.39}$$

with $I_j^*(L, \xi_j^2, |\bar{h}_j|^2) = (L-1)I(\xi_j^2 P_j) + I(P_j(\xi_j^2 + |\bar{h}_j|^2))$, *and* $I(p) = p$. *Again,* $|\bar{h}_j|^2$ *is defined as in (3.35).*

As we mentioned previously, the above capacity results rely on average link gain estimates, and on knowledge of channel estimation errors. In [Evans and Tse, 2000], channel estimation is treated as a standard Gaussian estimation problem, in which an LMMSE estimate of the channel is obtained conditioned on knowledge of the transmitted symbols over a training period. The length of the training period is characterixed by the length of the estimation window τ, and characterizes the estimation accuracy. It is shown in [Evans and Tse, 2000] that in the asymptotic case, the channel estimation variance converges almost surely to a deterministic constant ξ^2.

Theorem 3.12. *The minimum square error for any path converges almost surely as* $N \to \infty$ *(with* $\alpha = K/N$ *fixed) to the constant*

$$\xi^2 = \frac{\frac{\bar{p}}{L}}{1 + \bar{p}L\gamma_c}, \tag{3.40}$$

where γ_c *satisfies the equation*

$$\gamma_c = \left[\frac{\sigma^2}{\tau} + \frac{\alpha L}{\tau} \frac{\frac{\bar{p}}{L}}{1 + \frac{\bar{p}}{L}\gamma_c} \right]. \tag{3.41}$$

From (3.41), γ_c can be written in closed form as

$$\gamma_c = \frac{\tau - \alpha L}{2\sigma^2} - \frac{L}{2\bar{p}} + \left[\frac{(\tau - \alpha L)^2}{4\sigma^4} + L\frac{\tau + \alpha}{2\bar{p}\sigma^2} + \frac{L^2}{4\bar{p}^2} \right]^{1/2}. \tag{3.42}$$

We note that, as the window length increases, the contribution of the interference to γ_c becomes negligible and, in the limit, the variance is well approximated by σ^2/τ. Again, we note that as L increases, while τ and the total power per user remain constant, each interferer is perceived by the channel estimator as L low power interferers, and its performance approaches the performance of a matched filter estimator. This corresponds to the case in which only *a priori* statistics of the channel are available.

While Theorem 3.12 was initially proved for long sequences (the signature sequences are independently chosen from symbol to symbol), the proof was extended to the case of repeated sequences as well, under an additional mild assumption that the data symbols have zero mean. Also, it was shown that Theorem 3.12 holds even if the signature sequences

along the different paths are shifted replicas of the same transmitted sequence.

The final result in [Evans and Tse, 2000] couples the results obtained for data and channel estimation in the following theorem for the linear MMSE receiver, when $|\bar{a}_{kl}|^2$ is the received power of path l of user k, and the variance of the received power ξ_k^2 is given by Theorem 3.12.

Theorem 3.13. *The SIR achieved by an arbitrary user (say the kth one) in an LMMSE network can be expressed as:*

$$SIR_k = \frac{\sum_{l=1}^{L}|\bar{a}_{kl}|^2\gamma_d}{1+\xi_k^2\gamma_d}. \tag{3.43}$$

where γ_d is the unique fixed point in $(0,\infty)$ that satisfies

$$\gamma_d = \left[\sigma^2 + \alpha(L-1)\frac{\xi^2}{1+\xi^2\gamma_d} + \alpha\int_{\xi^2}^{\infty}\frac{p}{1+p\gamma_d}g(p)dp\right]^{-1}, \tag{3.44}$$

where $g(p) = \frac{1}{\left(\frac{\bar{p}}{L}-\xi^2\right)^L(L-1)!}(p-\xi^2)^{L-1}\exp\left(-\frac{p-\xi^2}{\frac{\bar{p}}{L}-\xi^2}\right)$, $p \geq \xi^2$.

It follows from Theorem 3.13 that the SIR is aymptotically chi-square distributed with $2L$ degrees of freedom.

A similar coupling result is obtained for the matched filter receiver, which shows that the SIR for the matched filter converges (in probability) to

$$SIR_k = \frac{\sum_{l=1}^{L}|\bar{a}_{kl}|^2\gamma_d}{1+\xi_k^2\gamma_d}, \tag{3.45}$$

with $\gamma_d = [\sigma^2 + \alpha\bar{p}]^{-1}$.

No coupling result is required for the decorrelator, since its SIR is independent of the interferers' received powers.

The above results from [Evans and Tse, 2000] characterize the SIR performance for linear receivers (LMMSE detector, decorrelator and matched filter) in a multipath fading environment for arbitrary transmission powers for the users. The questions to be answered next are: how can the network capacity be optimized with respect to power selection? And what would be the capacity of a multi-class network with different SIR target requirements? These questions were addressed in [Comaniciu and Poor, 2003a] for an LMMSE network with J classes of users, having different target SIRs, γ_j^*, $j = 1, 2, \ldots, J$.

The asymptotic network capacity can be derived building upon the results presented in Theorem 3.9 for the SIR convergence. We can see

that as $K, N \to \infty$, γ_d approaches a constant for all users. If we impose the condition that the achieved SIR should be γ_j^* for all class j users, $j = 1, 2, \ldots, J$ (i.e., the minimal power solution is achieved with equality), and using the assumption that all users in the same class have the same channel characteristics, it follows from (3.33) that all users in the same class j must have equal attenuated transmitted powers:

$$P_j = \frac{\gamma_j^*}{\gamma_d(|\bar{h}_j|^2 - \xi_j^2 \gamma_j^*)} = \frac{\gamma_j^*}{\gamma_d |\bar{h}_j|^2 (1 - \nu_j \gamma_j^*)}, \ j = 1, 2, \ldots, J, \quad (3.46)$$

where $\nu_j = \xi_j^2 / |\bar{h}_j|^2$.

Since the powers must be positive, and $\gamma_d \geq 0$, an immediate result is that $(1 - \nu_j \gamma_j^*) > 0$, and the achievable SIR target must satisfy the condition

$$\gamma_j^* < 1/\nu_j, \ j = 1, 2, \ldots, J. \quad (3.47)$$

Denoting $Q_j = |\bar{h}_j|^2 P_j$ for the class j of users, and imposing that the SIR bound should hold with equality $(SIR_j = \gamma_j^*)$, from (3.33) we can express γ_d as

$$\gamma_d = \frac{\gamma_j^*}{Q_j (1 - \nu_j \gamma_j^*)}, \ j = 1, 2, \ldots, J. \quad (3.48)$$

Thus, for multiple classes of users, the following equality holds:

$$\frac{\gamma_i^*}{Q_i (1 - \nu_i \gamma_i^*)} = \frac{\gamma_j^*}{Q_j (1 - \nu_j \gamma_j^*)}, \ \forall i, j = 1, 2, \ldots, J. \quad (3.49)$$

Expressing the SIR condition for an arbitrary user in class 1, it can be shown that the feasibility condition can be expressed as

$$\left[\sigma^2 + \sum_{j=1}^{J} \alpha_j \left((L-1)\nu_j \gamma_j^* (1 - \nu_1 \gamma_1^*) Q_1 / \gamma_1^* + \right. \right.$$

$$\left. \left. + \frac{Q_1 (1 + \nu_j) \frac{\gamma_j^*}{\gamma_1^*} \frac{1 - \nu_1 \gamma_1^*}{1 - \nu_j \gamma_j^*}}{1 + (1 + \nu_j) \frac{\gamma_j^*}{(1 - \gamma_j^* \nu_j)}} \right) \right]^{-1} = \frac{\gamma_1^*}{Q_1 (1 - \nu_1 \gamma_1^*)}. \quad (3.50)$$

Hence, the transmitting power for user 1 can be determined as

$$P_1^t = \frac{Q_1}{z_1 |\bar{h}_1|^2}, \quad (3.51)$$

where

$$Q_1 = \frac{\gamma_1^* \sigma^2}{(1 - \nu_1 \gamma_1^*) \left(1 - \sum_{j=1}^{J} \alpha_j \left[(L-1)\nu_j \gamma_j^* + \frac{\gamma_j^* (1 + \nu_j)}{1 + \gamma_j^*} \right] \right)}. \quad (3.52)$$

Since the transmitted power must be positive, and considering (3.47) and (3.52), the system capacity is restricted by the power control feasibility condition:

$$\sum_{j=1}^{J} \alpha_j \left[(L-1)\nu_j \gamma_j^* + \frac{\gamma_j^*(1+\nu_j)}{1+\gamma_j^*} \right] < 1. \tag{3.53}$$

The flat fading case can be obtained by setting $L = 1$.

The above derivation can be summarized via the following theorem.

Theorem 3.14. *In an asymptotically large CDMA system ($K_j \to \infty$, $N \to \infty$ with $\alpha_j = K_j/N$ constant, $j = 1, \ldots, J$) operating with linear MMSE receivers in a multipath fading environment with imperfect channel estimation, a minimal received power solution exists such that all users achieve their target SIRs, if and only if*

$$\gamma_j^* < \frac{1}{\nu_j},$$

and

$$\sum_{j=1}^{J} \alpha_j \left[(L-1)\nu_j \gamma_j^* + \frac{\gamma_j^*(1+\nu_j)}{1+\gamma_j^*} \right] < 1.$$

The minimal transmit power solution for a user in class i is given by

$$P_i^t = \frac{\gamma_i^* \sigma^2}{z_i |\bar{h}_i|^2 (1 - \nu_i \gamma_i^*) \left(1 - \sum_{j=1}^{J} \alpha_j \left[(L-1)\nu_j \gamma_j^* + \frac{\gamma_j^*(1+\nu_j)}{1+\gamma_j^*} \right] \right)}.$$

For comparison purposes, the capacity regions for a multi-class network using matched filter receivers are also derived in [Comaniciu and Poor, 2003a], and are specified as a power control feasibility condition:

$$\sum_{j=1}^{J} \alpha_j \epsilon_j \frac{L\nu_j + 1}{1 - \nu_j \epsilon_j} < 1. \tag{3.54}$$

The minimum transmit power solution for a user in class i is shown to be

$$P_i^t = \frac{1}{z_i |\bar{h}_i|^2} \frac{\epsilon_i \sigma^2}{(1 - \nu_i \epsilon_i) \left(1 - \sum_{j=1, \; j \neq i}^{J} \alpha_j \epsilon_j \frac{L\nu_j + 1}{1 - \nu_j \epsilon_j} \right)}.$$

In Fig. 3.7 we present asymptotic physical layer capacity comparisons for LMMSE and matched filter systems, for a network with two classes of users when $L = \{1, 3, 5\}$, and the target SIRs are $\gamma^* = 5$ for both classes.
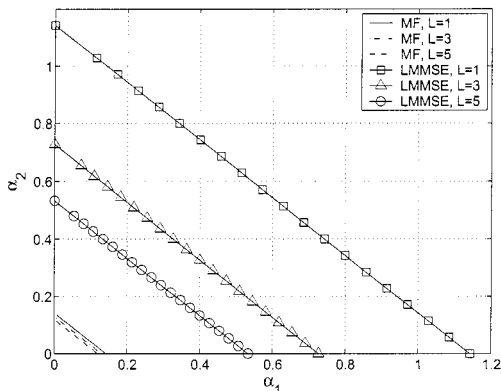
Figure 3.7. Asymptotic capacity comparisons: MF versus LMMSE (reprinted with permission from [Comaniciu and Poor, 2003a])

1.5 Multi-Rate Networks

As we discussed in Section 4 of Chapter 1, QoS support for heterogeneous services requires the implementation of multi-rate receivers. While in Chapter 1 we have focused on describing possible implementations for multi-rate multiuser receivers (LRD - low rate detector, HRD - high rate detector, and GSIC - groupwise successive interference cancellation detector), in this chapter we discuss the asymptotic capacity for power controlled networks using these architectures. This section's material is based mainly on research described in [Yao et al., 2004] for multicode and various LRD and HRD scenarios, and in [Comaniciu and Poor, 2003b], which analyzes the capacity of multi-rate GSIC networks. Asymptotic capacity comparisons among different multi-rate schemes have also been presented in [Biglieri et al., 2000].

For simplicity, the analysis in [Yao et al., 2004] considers a two class network: low rate and high rate, where the rate for the high rate class is M times higher than the rate of the low rate class. It is assumed that all the spreading sequences are randomly chosen and normalized.

We start our discussion with the simplest case, the multicode implementation, for which a high rate users is equivalent to M low rate virtual users, and therefore the results in [Tse and Hanly, 1999] can be straightforwardly applied. It can thus be proven that the SIR of any user k in the network converges to a deterministic value

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \alpha_0 \int I(P, P_k, \gamma_i^*)dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*)dF_1(p)}, \quad (3.55)$$

where the interference function $I(.)$ is defined as in (3.3), $F_0(p)$ and $F_1(p)$ are the distributions of powers for the two classes of users (low rate and high rate, respectively), and $\alpha_0 = \frac{K_0}{N_0}$ and $\alpha_1 = \frac{K_1}{N_1}$ represent the number of users per dimension for the low rate class and the high rate class, respectively.

We note that (3.55) holds for both high rate and low rate users, since there is no difference between a physical and a virtual low rate user. If the SIR requirements for the two classes of users are γ_h^* and γ_l^* for the high and low rate class, respectively, and if maximal power constraints $\overline{P}_h$ and $\overline{P}_l$ are imposed, the user capacity for multicode implementation is given by the following SIR feasibility condition:

$$\alpha_0 \frac{\gamma_l^*}{1 + \gamma_l^*} + M\alpha_1 \frac{\gamma_h^*}{1 + \gamma_h^*} \leq 1 - \max\left\{\frac{\sigma^2 \gamma_l^*}{\overline{P}_l}, \frac{\sigma^2 \gamma_h^*}{\overline{P}_h}\right\}. \tag{3.56}$$

Similar results for the multicode asymptotic capacity region have been derived independently in [Guo and Aazhang, 1999, Guo, 1999]. The multicode implementation is very simple, but has the disadvantage of a high peak to average power ratio requirement. An alternate solution for multi-rate implementation is to use variable spreading gain. As we have seen in Chapter 1, two different implementations may be considered: the LRD (low rate detector) and the HRD (high rate detector). For each scenario, two different signature sequence assignments for the low rate users are discussed in [Yao et al., 2004]: general random codes (GRCs), and random repetition codes (RRCs). The random repetition codes are obtained by repeating the first subinterval sequence (corresponding to a high rate code interval) for each of the subsequent $M - 1$ intervals (the low rate sequence can be considered as a concatenation of M subintervals, each of length equal to the spreading length of high rate users). The random repetition codes have the advantage that they can reduce the implementational complexity and they facilitate the adaptive implementation for the LMMSE receiver.

We summarize the main results of [Yao et al., 2004] as follows.

Theorem 3.15. *(HGRC): For an asymptotically large network, using high rate detectors and general random codes, the output SIR for any user k (low rate or high rate) can be approximated by the solution to the equation*

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \alpha_0 \int I\left(P, P_k, \frac{\gamma_k^*}{M}\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*) dF_1(p)}, \tag{3.57}$$

and the user capacity region of this network is

$$\alpha_0 \frac{\gamma_l^*}{1 + \gamma_l^*/M} + M\alpha_1 \frac{\gamma_k^*}{1 + \gamma_k^*} \leq 1 - \max\left\{ \frac{\sigma^2 \gamma_l^*}{\overline{P}_l}, \frac{\sigma^2 \gamma_k^*}{\overline{P}_k} \right\}. \qquad (3.58)$$

Theorem 3.16. *(LGRC): For an asymptotically large network, using low rate detectors and general random codes, the output SIR for any user k (low rate or high rate) can be approximated by the solution to the equation*

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \alpha_0 \int I\left(P, P_k, \gamma_k^*\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*) dF_1(p)}, \qquad (3.59)$$

and the user capacity region of this network is

$$\alpha_0 \frac{\gamma_l^*}{1 + \gamma_l^*} + M\alpha_1 \frac{\gamma_k^*}{1 + \gamma_k^*} \leq 1 - \max\left\{ \frac{\sigma^2 \gamma_l^*}{\overline{P}_l}, \frac{\sigma^2 \gamma_k^*}{\overline{P}_k} \right\}. \qquad (3.60)$$

Theorem 3.17. *(HRRC): For an asymptotically large network, using high rate detectors and random repetition codes, the output SIR for any high rate user k can be approximated by the solution to the equation*

$$\gamma_k^* = \frac{P_k}{\sigma^2 + \alpha_0 \int I\left(P, P_k, \frac{\gamma_k^*}{M}\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*) dF_1(p)}, \qquad (3.61)$$

while the output SIR for a low rate user is approximated by

$$\gamma_k^* = \frac{P_k}{\sigma^2 + M\alpha_0 \int I\left(P, P_k, \gamma_k^*\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*) dF_1(p)}. \qquad (3.62)$$

The network capacity region is determined by

$$\left\{ (\alpha_0, \alpha_1) \in A(\alpha_0, \alpha_1) \mid \frac{1}{\gamma_h^*} - \frac{M\alpha_1}{1 + \gamma_h^*} - \frac{\alpha_0}{x(\alpha_0, \alpha_1) + \gamma_h^*/M} > \right.$$

$$\left. > \max\left\{ \frac{\sigma^2}{\overline{P}_h}, \frac{\sigma^2}{\overline{P}_l x(\alpha_0, \alpha_1)} \right\} \right\}, \qquad (3.63)$$

where $A(\alpha_0, \alpha_1) = \left\{ (\alpha_0, \alpha_1) \mid \alpha_0 < 1 + \frac{1}{\gamma_l^*}, \ \alpha_1 < 1 + \frac{1}{\gamma_h^*}, \right.$

$$\left. \max\left\{ 0, \frac{M\alpha_1}{\frac{1}{\gamma_l^*} - \frac{M\alpha_0}{1+\gamma_l^*}} - \gamma_l^* \right\} \left(\frac{\alpha_0}{\frac{1}{\gamma_h^*} - \frac{M\alpha_1}{1+\gamma_h^*}} - \frac{\gamma_h^*}{M} \right) < 1 \right\},$$

and where $x(\alpha_0, \alpha_1)$ is the solution to the equation

$$x = \frac{\left(\frac{1}{\gamma_l^*} - \frac{M\alpha_0}{1+\gamma_l^*} - \frac{M\alpha_1}{\frac{1}{x}+\gamma_l^*}\right)}{\left(\frac{1}{\gamma_h^*} - \frac{\alpha_0}{x+\frac{\gamma_h^*}{M}} - \frac{M\alpha_1}{1+\gamma_h^*}\right)}.$$

Theorem 3.18. *(LRRC) : For an asymptotically large network, using low rate detectors and random repetition codes, the output SIR for any high rate user k can be approximated by the solution to the equation*

$$\gamma_k^* = \frac{\gamma_{k,1} + (M-1)\gamma_{k,2} + M\gamma_{k,1}\gamma_{k,2}}{M + (M-1)\gamma_{k,1} + \gamma_{k,2}}, \tag{3.64}$$

where $\gamma_{k,1}$ and $\gamma_{k,2}$ are defined as

$$\gamma_{k,1} = \frac{P_k}{\sigma^2 + M\alpha_0 \int I\left(P, P_k, \gamma_{k,1}\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_{k,1})dF_1(p)}, \tag{3.65}$$

and

$$\gamma_{k,2} = \frac{P_k}{\sigma^2 + M\alpha_1 \int I(P, P_k, \gamma_{k,2})dF_1(p)}. \tag{3.66}$$

The output SIR for a low rate user k can expressed as a solution to

$$\gamma_k^* = \frac{P_k}{\sigma^2 + M\alpha_0 \int I\left(P, P_k, \gamma_k^*\right) dF_0(p) + M\alpha_1 \int I(P, P_k, \gamma_k^*)dF_1(p)}. \tag{3.67}$$

The above results give insights into the performance tradeoffs for different implementations for multi-rate systems. Asymptotically, the multicode network and the LRD for general random codes have the best performance. The bidimensional user capacities for all of the scenarios discussed above are compared in Fig. 3.8 for both the asymptotic limit and finite networks (simulation results).

For these numerical results, the spreading gains were selected to be 128 and 32, respectively, and the target SIRs were chosen to be 1 dB for the low rate users and 10 dB for the high rate users. The power constraints impose that the SNR can be no more than 20 dB (resp. 30 dB) for low (resp. high) rate users. For finite networks, we can see that the MC and LGRC still have the highest capacity, but it is lower than the expected limit. This is due to the fluctuations in the achieved output SIR as discussed also in [Tse and Hanly, 1999]. Determining
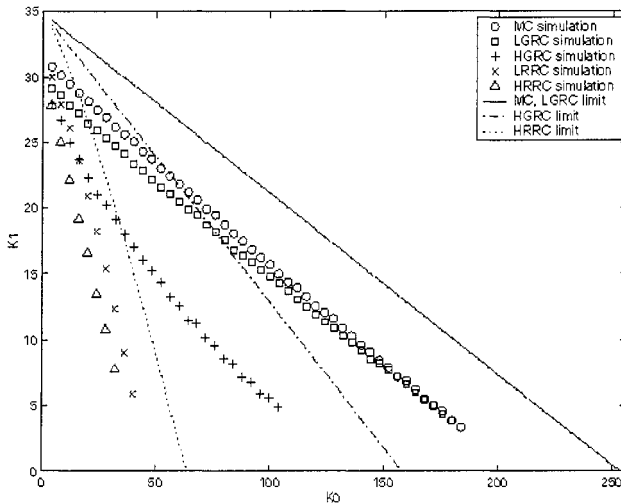
Figure 3.8. Capacity for multi-rate networks (reprinted with permission from [Yao et al., 2004])

the simulated capacity requires that all the users must meet their target SIRs. This calls for capacity overprovisioning, so that we expect that the simulated capacity will always be smaller than the theoretical asymptotic prediction. Following the results in [Tse and Zeitouni, 2000], the SIR can be characterized as being asymptotically Gaussian with mean γ^* and variance $\frac{1}{N}\left[\frac{2\gamma^*(1+\gamma^*)^2}{\frac{\sigma^2}{P}(1+\gamma^*)^2+\alpha} + E\{s_{11}^4\} - 3(\gamma^*)^2\right]$, where s_{11} is the first element of user's 1 signature sequence. It was verified in [Yao et al., 2004] that, by applying the above distribution for the SIR in an MC system in which all users have the same target SIR, very good agreement between the simulation results and the adjusted theoretical limit is obtained. This phenomenon supports the assumption that the mismatch between asymptotic results and finite networks simulations is due to the fact that the output SIR is a random variable for finite networks, rather than being a deterministic constant.

As we have discussed in Chapter 1, an alternative to LRD and HRD implementations is the GSIC system, in which the users are detected in groups, and the interference between groups is cancelled using successive interference cancellation.

In [Comaniciu and Poor, 2003b], optimal power allocation and capacity regions of LMMSE GSIC systems in multipath fading channels are derived for asymptotically large systems. The analysis uses the same model assumptions, and the same notations as in Section 1.4 of this chapter. For GSIC, the impact of channel estimation errors is two-fold: they impact the LMMSE receiver performance within a class of users in the same detection group, and also they are strongly related to the cancellation errors for the successive group interference cancellation.

Consider J classes of users, in which all users in a given class j have the same transmission rates (same spreading gains) and the same SIR targets, γ_j^*. All users in the same class are detected within the same group. The first detected group is selected according to a criterion such as minimal received power. Then, the interference caused by the first group is reconstructed and cancelled from the received signal. This is done successively until the last group of users has been detected. For a given detection order, we denote the groups as $1, 2, \ldots, J$, which represents the detection order.

The imperfect channel estimation yields an imperfect cancellation for group j of users, resulting in a residual interference power $\epsilon_j \sum_{k=1}^{K_j} Q_{j,k}$, where K_j is the number of users in class j, $Q_{j,k} = P_{j,k} |\bar{h}_j|^2$ is the received power of user k from class j, and ϵ_j is the fractional error in canceling the total interference power created by the jth group. This result implicitly assumes that the fractional error for canceling a group j user is the same for all users in class j. Since the target bit-error rates are usually very low, it can be assumed that the cancellation error is mostly determined by the amplitude and phase estimation errors. Similarly to the approach in [Andrews and Meng, 2003] we assume that the cancellation error (ϵ) for the successive interference cancellation, is approximately the same as the total channel estimation standard deviation, ξ. Assuming further that the multipath components are iid, and have estimation error variances of ξ^2, the estimated cancellation error for an L path channel can be approximated by:

$$\epsilon = \sqrt{L\xi^2}. \tag{3.68}$$

Based on the SIR expression for multipath fading channels (3.33), and using the results in [Comaniciu and Poor, 2003a] for multiclass systems using LMMSE receivers, it can be shown that, for a GSIC system, all users within a detection group should have equal received powers. Since each group of users is detected using LMMSE receivers, while the interference created by other groups is perceived as white noise, it can be shown that every group j of users can be approximated as an all LMMSE system with enhanced noise Σ_e^j:

$$\Sigma_e^j = \sigma^2 + \sum_{l<j} K_l \frac{1}{N_l} \epsilon_l Q_l + \sum_{l>j} K_l \frac{1}{N_l} Q_l = \sigma^2 + \sum_{l<j} \epsilon_l \alpha_l Q_l + \sum_{l>j} \alpha_l Q_l, \cdot$$

(3.69)

with α_l defined as $\alpha_l = K_l/N_l$. This equivalence is based on the fact that the receiver filter coefficients for group j users ignore the structure of the interference from other groups, and thus any pair of filter coefficients and signal signature sequences for users in other groups may be considered to be independent.

Using (3.34), (3.48) and (3.69), and after straightforward algebraic manipulation, we can derive the power control feasibility condition such that the target SIR γ_j^* can be met with equality for an arbitrary class j of users:

$$Q_j = \theta_j \sum_{l<j} \epsilon_l \alpha_l Q_l + \alpha_j Q_j \Lambda_j + \theta_j \sum_{l>j} \alpha_l Q_l + \theta_j \sigma^2, \qquad (3.70)$$

where $\theta_j = \gamma_j^*/(1 - \nu_j \gamma_j^*) > 0$, and $\Lambda_j = (L-1)\nu_j \gamma_j^* + (1+\nu_j)\gamma_j^*/(1+\gamma_j^*)$.

Given that target SIRs must be met for all users, the power control feasibility can be expressed as a matrix equation condition

$$(\mathbf{I}_{J\times J} - \mathbf{A})\mathbf{q} = \sigma^2 \mathbf{u}, \qquad (3.71)$$

where $\mathbf{q}^T = [Q_1, Q_2, \ldots, Q_J]$, $\mathbf{u}^T = [\theta_1, \theta_2, \ldots, \theta_J]$, $\mathbf{I}_{J\times J}$ is the identity matrix, and

$$\mathbf{A} = \begin{pmatrix} \alpha_1 \Lambda_1 & \theta_1 \alpha_2 & \ldots & \theta_1 \alpha_J \\ \epsilon_1 \alpha_1 \theta_2 & \alpha_2 \Lambda_2 & \ldots & \theta_2 \alpha_J \\ \ldots & \ldots & \ldots & \ldots \\ \epsilon_1 \alpha_1 \theta_J & \epsilon_2 \alpha_2 \theta_J & \ldots & \alpha_J \Lambda_J \end{pmatrix}. \qquad (3.72)$$

The matrix $\mathbf{A}$ is a nonnegative matrix, but it is not necessarily irreducible, since perfect cancellation of group 1 users results in a reducible matrix. For a nonnegative, irreducible matrix, a positive vector solution to (3.71) exists iff $\rho(\mathbf{A}) < 1$, where $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$. This is usually the practical case since perfect cancellation is difficult to achieve in practice. Nevertheless, this result has been proven to hold under more general circumstances [Agrawal et al., 2004] for a nonnegative matrix and a positive noise vector term in the matrix condition, provided that the power control feasibility condition can be expressed in a standard form, using a standard interference function: $\mathbf{q} = i(\mathbf{q})$. For the considered GSIC system, we can alternatively express the power control feasibility condition as

$$\mathbf{q} = i(\mathbf{q}); \quad i(\mathbf{q}) = \mathbf{A}\mathbf{q} + \sigma^2 \mathbf{u}. \qquad (3.73)$$

It is straightforward to show that function $i(\mathbf{q})$ is a standard interference function by verifying the three properties presented in [Yates, 1995]: positivity, monotonicity and scalability (see Section 2 in Chapter 2). It can be shown that all three properties hold true if $\theta_j > 0$, $\forall j = 1, \ldots, J$, which holds if $\gamma_j^* < 1/\nu_j$, $\forall j = 1, \ldots, J$.

The main capacity result of [Comaniciu and Poor, 2003b] can be summarized in the following theorem.

Theorem 3.19. *In a groupwise successive interference cancellation system with LMMSE receivers within a group, and operating under a multipath fading environment with imperfect channel estimation, a positive power vector solution exists such that all users meet their target SIRs γ_j^*, if and only if*

$$\gamma_j^* < \frac{1}{\nu_j} \ and \ \rho(\mathbf{A}) < 1. \tag{3.74}$$

If (3.74) holds, then the optimal received power allocation for the groups of users is given by

$$\mathbf{q}* = (\mathbf{I}_{J \times J} - \mathbf{A})^{-1} \mathbf{u} \sigma^2. \tag{3.75}$$

It was also proved in [Comaniciu and Poor, 2003b] that the received power requirements for different groups can be derived using a recursive formula. Denoting by Q_j the required received power for detection class j, and using the notation $\Gamma_j = (1 - \alpha_j \Lambda_j)/\theta_j$, it can be shown that

$$Q_{j+1} = \frac{\Gamma_j + \epsilon_j \alpha_j}{\Gamma_{j+1} + \alpha_{j+1}} Q_j, \tag{3.76}$$

or equivalently,

$$Q_j = \Pi_{i=1}^{j-1} \frac{\Gamma_i + \epsilon_i \alpha_i}{\Gamma_{i+1} + \alpha_{i+1}} Q_1, \tag{3.77}$$

with $Q_1 = \dfrac{\sigma^2}{\Gamma_1 - \sum_{j=2}^{J} \alpha_j \Pi_{i=1}^{j-1} \dfrac{\Gamma_i + \epsilon_i \alpha_i}{\Gamma_{i+1} + \alpha_{i+1}}}$.

The total received power requirements for all users, for a given detection order, can then be express as

$$Q_T = \sum_{j=1}^{J} \alpha_j Q_j = Q_1(\alpha_1 + \Gamma_1) - \sigma^2; \tag{3.78}$$

that is,

$$Q_T = \frac{\sigma^2}{\frac{\Gamma_1}{\alpha_1 + \Gamma_1} - \frac{1}{\alpha_1 + \Gamma_1} \sum_{j=2}^{J} \alpha_j \Pi_{i=2}^{j} \frac{\Gamma_{i-1} + \epsilon_{i-1} \alpha_{i-1}}{\Gamma_i + \alpha_i}} - \sigma^2. \tag{3.79}$$

While the above results were derived for a given, arbitrary, detection order, this can be optimized for a minimal received power solution. Using a similar approach to that in [Shu and Niu, 2003], it can be shown that the optimal detection order for GSIC is the same as for SIC systems, i.e., groups must be detected in the ascending order of their cancellation errors.

An interesting observation is that this result is in contrast with the popular recommendation of detecting higher rate users first. Although the analysis in [Comaniciu and Poor, 2003b] considers only the impact of the imperfect amplitude estimation on the cancellation errors, this model can be extended to encompass other effects, such as the performance differences between the asymptotic analysis and the practical finite case. In this case, higher rate users (using lower spreading gains) may have a higher cancellation error due to a higher achieved SIR variance relative to the estimated average SIR for asymptotically large systems [Evans and Tse, 2000].

Capacity regions for the general case of a GSIC system with J groups can be defined in a generic form as

$$C = \left\{ (\alpha_1, \alpha_2, \dots, \alpha_J) \mid \gamma_j^* < 1/\nu_j, \ \forall j = 1, \ \dots, J, \ \rho(\mathbf{A}) < 1 \right\}. \quad (3.80)$$

The computation of the maximal eigenvalue $\rho(\mathbf{A})$ is not very complex since $\mathbf{A}$ is a $J \times J$ matrix, where J is the number of groups, which is usually a small number.

For the particular case of a GSIC system with two detection groups, an explicit dependence between the number of users that can be supported in each class can be obtained as

$$\alpha_1 \Lambda_1 + \alpha_2 \Lambda_2 + \sqrt{\Delta} < 2, \quad (3.81)$$

where $\Delta = (\alpha_1 \Lambda_1 + \alpha_2 \Lambda_2)^2 + 4\alpha_1 \alpha_2 (\theta_1 \theta_2 \epsilon_1 - \Lambda_1 \Lambda_2)$.

A similar capacity expression can be derived for GSIC with matched filter receivers, namely

$$\alpha_1 \Lambda_1^* + \alpha_2 \Lambda_2^* + \sqrt{\Delta^*} < 2, \quad (3.82)$$

with $\Delta^* = (\alpha_1 \Lambda_1^* + \alpha_2 \Lambda_2^*)^2 + 4\alpha_1 \alpha_2 (\theta_1 \theta_2 \epsilon_1 - \Lambda_1^* \Lambda_2^*)$.

In Fig. 3.9 we compare the performance of the matched filter GSIC and the LMMSE GSIC, for different channel estimation errors and for required target SIRs of 10 for both classes. The estimated average link gain $|\bar{h}_j|^2$, $j = 1, 2$ is 1, and the channel path length is $L = 3$. We notice that both implementations are strongly affected by channel estimation errors, but a very substantial performance gap exists in favor of the LMMSE implementation. Also, in Fig. 3.10, the GSIC network capacity is compared with an equivalent MC implementation for $L = 3$.
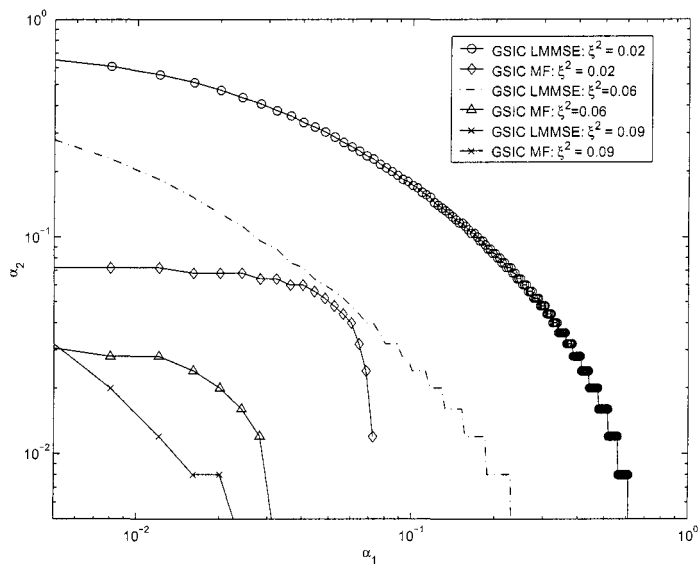
Figure 3.9. Capacity comparisons: GSIC with LMMSE versus GSIC with MF
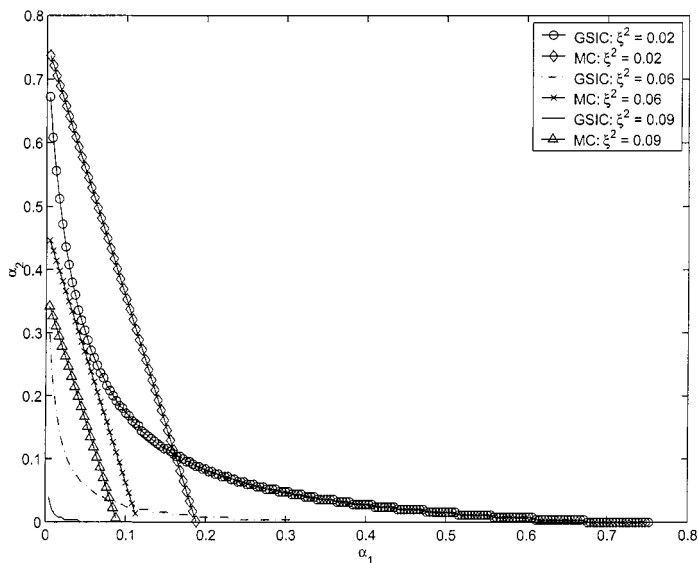
Figure 3.10. Capacity comparisons: GSIC with LMMSE versus MC

1.6   Asynchronous Networks

Although synchronous transmission is often not a practical scenario, the synchronous analysis is usually very useful as a performance benchmark for networks using linear multiuser receivers. In general, a synchronous analysis can be extended to the asynchronous case by considering an equivalent synchronous system with more interferers having lower powers. In [Kiran and Tse, 2000], the authors present a rigorous analysis for the asymptotic performance of asynchronous networks with linear receivers using random spreading codes. They show that the matched filter performance in the asynchronous case is the same as that achieved for the synchronous case. They also show that, if the observation window is infinite, the same is true for the decorrelator and the LMMSE detector. However, for the "one-shot" detection approach, the achieved SIRs for both the decorrelator and the LMMSE detector degrade for the asynchronous case. Although exact capacity values are difficult to derive, in [Kiran and Tse, 2000] the authors present very tight lower bounds on the achievable SIR for both the decorrelator and the LMMSE detector in the asynchronous case, under the simplifying assumption that the nodes are chip-synchronous. They also show by means of simulations that the chip-synchronous scenario provides conservative estimates for a truly asynchronous system.

In an asynchronous system, the various relative delays τ_k are assumed to be random. As a consequence, the SIR performance measure is also a random variable depending on the random spreading sequences and on the random delays. It is assumed that although the delays are random, the receiver has acquired timing information for all users. Also, the underlying assumption of the analysis in [Kiran and Tse, 2000] is that as the system becomes asymptotically large, the empirical distributions of powers and delays converge to fixed distributions $F(P)$ and $G(\tau)$, respectively. For the LMMSE receiver, a tight lower bound for the achieved SIR is proposed in [Kiran and Tse, 2000], under a symmetry assumption for the relative delay distribution function $G(\tau)$: $G(\tau) = 1 - G(1 - \tau)$. It is shown that the achievable SIR, γ_k^*, is lower bounded by $\bar{\gamma}_k$, which (for a given user k) is the solution of the fixed point equation

$$\bar{\gamma}_k = \frac{P_k}{\sigma^2 + \alpha E_P E_\tau \{I(\tau P, P_k, \bar{\gamma}_k) + I((1 - \tau)P, P_k, \bar{\gamma}_k)\}}, \qquad (3.83)$$

with the standard interference function $I(\cdot)$ having the same expression as for the synchronous case previously discussed. We note that for the

asynchronous case, the effective interference created by any user l in the system on user k, is the sum of the corresponding effective interference from the synchronous case for two partial symbols within the observation window.

For equal received powers P and uniform delay distribution, the SIR lower bound, $\bar{\gamma}$, can be expressed as

$$\bar{\gamma} = \frac{P}{\sigma^2 + \frac{2\alpha P}{\bar{\gamma}}\left(1 - \frac{\ln(1+\bar{\gamma})}{\bar{\gamma}}\right)}. \tag{3.84}$$

Based on (3.84), a lower bound on the network capacity can be determined for the LMMSE case, when the maximal transmission power is set to $\overline{P}$:

$$\alpha(\overline{P}, \gamma^*)_{mmse} = \frac{1}{2\left(1 - \frac{\ln(1+\gamma^*)}{\gamma^*}\right)}\left(1 - \frac{\sigma^2\gamma^*}{\overline{P}}\right). \tag{3.85}$$

Further, the user capacity when no power constraints are imposed is given by

$$\alpha(\gamma^*)_{mmse} = \frac{1}{2\left(1 - \frac{\ln(1+\gamma^*)}{\gamma^*}\right)}. \tag{3.86}$$

The capacity expression in (3.85) can be extended straightforwardly to the case with J classes of users, having different SIR requirements γ_j^*, $j = 1, \ldots, J$:

$$\sum_{j=1}^{J} 2\left(1 - \frac{\ln(1 + \gamma_j^*)}{\gamma_j^*}\right)\alpha_{j,mmse} \leq \min_{1 \leq j \leq J}\left[1 - \frac{\gamma_j^*\sigma^2}{\overline{P}_j}\right], \tag{3.87}$$

where $\alpha_{j,mmse} \triangleq K_j/N_j$. From (3.87), we can see that the effective bandwidth for the asynchronous LMMSE network is

$$e_{mmse}^{asy} = 2\left(1 - \frac{\ln(1 + \gamma_j^*)}{\gamma_j^*}\right) \text{ degrees of freedom per user.} \tag{3.88}$$

For the decorrelator, exact SIR convergence results are obtained in [Kiran and Tse, 2000]:

$$\gamma_k^* = \begin{cases} \frac{P_k(1-2\alpha)}{\sigma^2}, & \alpha < 1/2, \\ 0 & \alpha \geq 1/2. \end{cases} \tag{3.89}$$

Consequently, the network capacity and the effective bandwidth for the decorrelator are given by

$$\alpha(\overline{P}, \gamma^*)_{dec} \leq \frac{1}{2} - \frac{\gamma^*\sigma^2}{\overline{P}}, \tag{3.90}$$
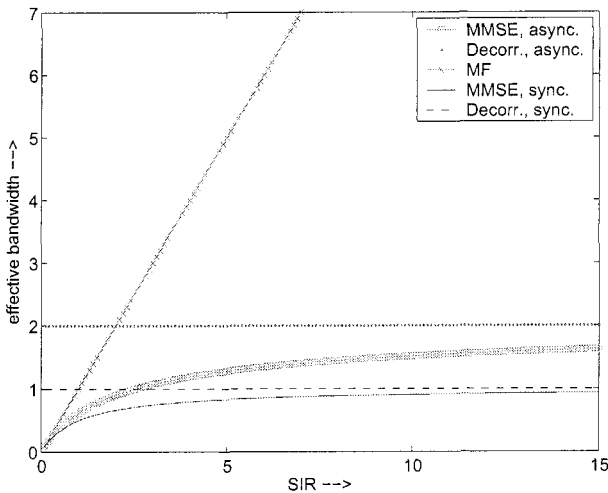
Figure 3.11. Effective bandwidth comparisons

and

$$e_{dec}(\gamma^*) = 2. \tag{3.91}$$

Therefore, the capacity results from the previous subsection can be extended straightforwardly to

$$\alpha_{dec} < 1/2, \tag{3.92}$$

when no power constraints are imposed.

In Fig. 3.11, we compare the effective bandwidths for the three discussed linear receivers - LMMSE detector, decorrelator, and matched filter for both the synchronous and asynchronous cases.

From these results we can see that for small required SIRs, the matched filter receiver has a lower effective bandwidth than the decorrelator, whereas for large required SIRs, the decorrelator outperforms the matched filter. The cross-over between these two regimes occurs at a higher value for the required SIR in the asynchronous system than in the synchronous system. We also note that there is a more significant performance improvement for the LMMSE receiver over the decorrelator in the asynchronous case, compared with the synchronous case. This can be explained by the fact that the LMMSE receiver takes advantage of the fact that the overlap of symbols is only partial and thus yields reduced energy, while the decorrelator loses a full extra degree of freedom by considering an additional interfering symbol per user.

Performance can be improved for both the LMMSE receiver and the decorrelator if a larger detection window is used. More specifically, it has been shown [Kiran and Tse, 2000] that for a detection window length of T symbols, the effective bandwidths for the LMMSE and the decorrelator, respectively, are given by

$$e_{mmse}^{asy}(T, \gamma^*) = \frac{1}{T} \left[(T-1) \left(\frac{\gamma^*}{1+\gamma^*} \right) + 2 \frac{\ln(1+\gamma^*)}{\gamma^*} \right] \qquad (3.93)$$

and

$$e_{dec}^{asy}(T, \gamma^*) = \frac{T+1}{T}. \qquad (3.94)$$

It can easily be seen that as $T \to \infty$, the network performance for the asynchronous case becomes identical to that for the synchronous case.

1.7　Imperfect Power Control

As we have seen up to this point, SIR is a key performance measure in wireless networks using linear multiuser receivers. However, our previous discussions of network performance have assumed that perfect power control is achievable, and thus the network capacity is determined by the power control feasibility condition. A question that remains to be addressed is: what happens in practice when power control loops yield imperfect control of the received powers? For example, how robust is the LMMSE receiver's performance to imperfections in power control and how can the network capacity be characterized under these conditions? These questions have been addressed in [Zhang et al., 2001], which showed that under some mild conditions on the distribution of powers, the output multiple access interference is approximately Gaussian, and thus the bit error rate is strongly related to the achievable SIR[1]. In this case, the network capacity may be determined by imposing an outage condition:

$$P(SIR < \gamma^*) \le \phi, \qquad (3.95)$$

where γ^* is the target SIR and ϕ is the maximum allowable outage probability.

In [Zhang et al., 2001], asymptotic Gaussianity of the MAI as $K, N \to \infty$ with $\alpha = K/N$ fixed, has been proved for synchronous transmission for two different scenarios:

1 The statistics of the output MAI are averaged over the signature sequences:

---

[1] [Poor and Verdú, 1997] also shows similar results in the non-asymptotic case.

For this case, the MAI is shown to be asymptotically Gaussian if the empirical distribution function for the mean received powers $\{\mu_1, \mu_2, \ldots, \mu_K\}$ converges weakly to a distribution function F_μ, and if the second moments of the received powers are bounded.

2 The statistics of the output MAI are conditioned on the signature sequences and powers:

For this case, the conditional distribution is shown to converge weakly to the same Gaussian distribution for almost every realization of the signatures and powers, if the joint empirical distribution function of $\{P_1\mu_1, P_2\mu_2, \ldots, P_K\mu_K\}$ converges weakly to a distribution function $F_{P,\mu}$, and if the received powers P_k are uniformly bounded from above, while the mean powers are bounded from below by a positive number.

Furthermore, it was shown that the variance of the distribution is not affected by the fluctuations in the interferers' powers. The SIR is proved to converge as $K, N \to \infty$, with $\alpha = K/N$ fixed, to the solution to the equation

$$\gamma_k = \frac{P_k}{\sigma^2 + \alpha \int_0^\infty I(\mu, P_k, \gamma_k) dF_\mu(\mu)}. \qquad (3.96)$$

We note that power control is still needed to achieve the desired SIR performance. Imperfect power control leads to a probabilistic guarantee for the SIR as in (3.95). Based on outage probability requirements, in [Zhang and Chong, 2000] the asymptotic capacity expression or a network having LMMSE receivers was derived to be

$$\alpha(\phi) = \frac{F^{-1}(\phi)}{\gamma^*\mu} + 1 - \frac{\sigma^2}{\mu} - \frac{\sigma^2\gamma^*}{F^{-1}(\phi)}, \qquad (3.97)$$

where μ is the mean received power, and σ^2 is the noise power. Similar convergence results for the MAI were proved for the asynchronous case, under the assumption that the system is chip-synchronous and the offsets $\{\tau_1, \tau_2, \ldots, \tau_K\}$ have an empirical distribution which converges weakly to a deterministic distribution function H_τ. The same convergence results as for the synchronous case have been proved for both unconditional and conditional MAI, if the empirical distribution function for the mean received powers $\{\mu_1, \mu_2, \ldots, \mu_K\}$ converges weakly to a distribution function F_μ, and if the received powers P_k are uniformly bounded from above, while the mean powers are bounded from below by a positive number.

1.8   Blind and Group-Blind Multiuser Receivers

Our previous discussion on the performance of networks using linear multiuser receivers assumes that all signature sequences in the network are known by all users and are used to compute the filter coefficients. In practice, this is a major disadvantage for the multiuser receivers' implementation, and has been for a long time an obstacle to the adoption of multiuser detection technology in current commercial networks. As we have already mentioned in Chapter 1, a solution to this problem is to implement blind multiuser receivers. In this section, we discuss the capacity of synchronous, power controlled networks using blind and group blind multiuser receivers. An introduction to the performance of such blind receivers has been presented in Chapter 1.

Building upon results in [Tse and Hanly, 1999], which treats the "exact LMMSE" case (LMMSE receivers built using knowledge of all codes in the network), and [Høst-Madsen and Wang, 2002], which quantifies the SIR estimation error for blind LMMSE receivers, [Zhang and Wang, 2002b] proposes a large system analysis in which the SIR performance for networks using blind receivers is quantified. The analysis assumes that binary random spreading is used, and that the spreading gain N, the number of users K, and the number of received signal samples required for filter estimation, Υ, go to infinity, while the ratios $\alpha_K = K/N$ and $\alpha_\Upsilon = \Upsilon/N$ are fixed. It is also assumed that $\alpha_K < 1$ and α_Υ is "reasonably large".

The main result of [Zhang and Wang, 2002b] reveals a saturation phenomenon when blind receivers are used, such that the achievable SIR is both interference limited and also estimation error limited. In the following paragraphs we will present the main results of [Zhang and Wang, 2002b] on network performance with blind and group-blind multiuser receivers.

Following the analysis in [Høst-Madsen and Wang, 2002], if we assume that the spreading sequences are random, the output SIRs will be random as well. However, it was shown in [Zhang and Wang, 2002b], that the SIR can be well approximated by a deterministic constant for asymptotically large systems. The analysis in [Zhang and Wang, 2002b] assumes that the sequence of powers for all users is bounded above and below by positive numbers and that the distribution of powers converges weakly to a distribution function F_μ. The asymptotic SIR convergence for the blind and group-blind receivers was shown in [Zhang and Wang, 2002b] based on results developed in [Tse and Hanly, 1999], in which it is shown that the SIR achieved per unit power for linear MMSE systems

(γ_0) is the solution to the fixed-point equation:

$$\gamma_0 = \frac{1}{\sigma^2 + \alpha_K \int_0^\infty \frac{\mu}{1+\mu\gamma_0} dF_\mu(\mu)}. \tag{3.98}$$

Asymptotic SIR results for the above blind receivers are summarized in the following theorem [Zhang and Wang, 2002a, Zhang and Wang, 2002b]. (Recall that DMI refers to the direct matrix inversion method of blind multiuser detection.)

Theorem 3.20. *In the asymptotic regime, as $N \to \infty$, with $\alpha_K = \frac{K}{N}$, $\alpha_{\tilde{K}} = \frac{\tilde{K}}{N}$ and $\alpha_\Upsilon = \frac{\Upsilon}{N}$ fixed, for almost every realization of signature sequences, the output SIR achieved by the blind linear MMSE receiver is well approximated as*

$$SIR_k = \begin{cases} \frac{P_k\gamma_0}{1+\frac{1}{\alpha_M}(1+P_k\gamma_0)}, & DMI \\ \frac{P_k\gamma_0}{1+\frac{1}{\alpha_M}[\alpha_K+\delta_k(1-\alpha_K)](1+P_k\gamma_0)}, & subspace \\ \frac{P_k\gamma_0}{1+\frac{1}{\alpha_M}[(\alpha_K-\alpha_{\tilde{K}})+\delta_k(1-\alpha_K)](1+P_k\gamma_0)}, & group\text{-}blind, \end{cases} \tag{3.99}$$

where

$$\delta_k = \frac{\gamma_0 - e_1 + \sigma^2 e_2}{\gamma_0[1 + P_k(e_1 + (1 - \alpha_K)/\sigma^2)]^2},$$

$$e_1 = \lim_{\sigma^2 \to 0} \left(\gamma_0(\sigma^2) - \frac{1-\alpha_K}{\sigma^2} \right),$$

and

$$e_2 = \lim_{\sigma^2 \to 0} -\frac{d\left(\gamma_0(\sigma^2) - \frac{1-\alpha_K}{\sigma^2}\right)}{d\sigma^2}.$$

An important observation is that an SIR saturation phenomenon occurs, such that, when the SNR increases ($P_k/\sigma^2 \to \infty$), the asymptotic SIR for the three receivers approaches

$$\{SIR_k\}_{P_k/\sigma^2 \to \infty} = \begin{cases} \alpha_\Upsilon, & \text{DMI} \\ \frac{\alpha_\Upsilon}{\alpha_K}, & \text{subspace} \\ \frac{\alpha_\Upsilon}{\alpha_K-\alpha_{\tilde{K}}}, & \text{group-blind.} \end{cases} \tag{3.100}$$

The above results indicate that the network capacity for a system using blind receivers is both interference and estimation error limited. This phenomenon is illustrated in Fig. 3.12, in which the dashed lines rep-
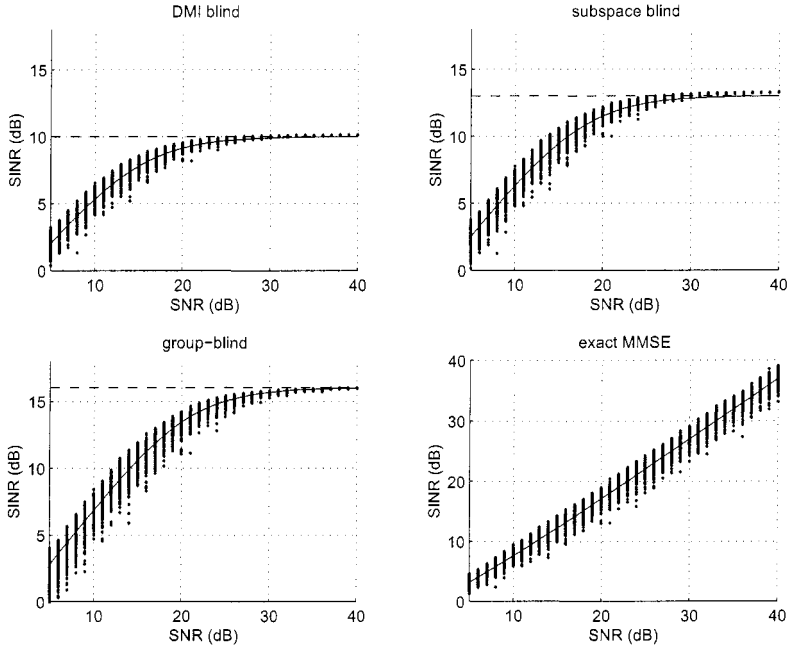
Figure 3.12. Saturation phenomenon for blind LMMSE receivers (reprinted with permission from [Zhang and Wang, 2002b])

resent $\{SIR_k\}_{P_k/\sigma^2 \to \infty}$, computed[2] as in (3.100). The parameter values for this example are $N = 32$, $\alpha_\Upsilon = 10$, and $\alpha_K = 0.5$.

As a final note, we compare the performance of the three blind receiver implementations. It was shown in [Zhang and Wang, 2002b] that if $\delta_k < 1$, then the subspace receiver outperforms the DMI, and if $\delta_k > 1$, DMI performs better. Moreover, we can see that in the high SNR region, the subspace receiver outperforms DMI, whereas in the low SNR region $(P_k \to 0)$, if $\sigma^2 < \frac{e_1}{e_2}$, the subspace receiver is better, otherwise, the DMI is better. As expected, since the group-blind receiver uses more information than the subspace receiver, it will always perform better than the latter one. On the other hand, it performs worse than DMI when $\delta_k \geq 1 + \frac{\alpha_{\bar{K}}}{1 - \alpha_K}$.

---

[2]In this figure, the abbreviation SINR refers to the SIR.

2.   Ad Hoc Networks

Due to the inherent nature of ad hoc network architectures, a major problem in such networks is performance optimization of the air interface, so as to ensure that all users receive their requested QoS, while the network capacity is reasonably high. However, in general, if random access is employed in ad hoc networks, the performance will be strongly affected by collisions, and sophisticated medium access control must be implemented. Even so, pessimistic capacity results for such networks can be obtained. In [Gupta and Kumar, 2000], the authors study the capacity of a fixed ad hoc network in which the nodes' locations are fixed but randomly distributed. They prove that, as the number of nodes (K) per unit area increases, the achievable throughput between any randomly selected source-destination pair is of order $O(1/\sqrt{K})$. A capacity increase for such networks has been reported in [Grossglauser and Tse, 2002], which shows that exploiting mobility can result in a form of multiuser diversity and can improve the system capacity. The authors of [Grossglauser and Tse, 2002] propose a two-hop transmission strategy in which the traffic is first randomly spread (first hop) across as many relay nodes as possible, and then it is delivered (second hop) as soon as any of the relaying nodes is close to the destination. The disadvantage of this scheme is that it involves large delays and therefore it is not suitable for delay sensitive traffic. A capacity increase with mobility has also been noticed in [Gupta and Das, 2001], in which the capacity is empirically determined for a different network model that exploits spatial diversity.

In contrast with the previous approaches to characterize the capacity of ad hoc networks, [Comaniciu and Poor, 2004c] shows that significant capacity gains can be obtained in a CDMA ad hoc network using linear multiuser receivers, even when tight delay and power requirements are enforced. This result is a consequence of the inherent advantages of CDMA (resistance to intereference and fading) and of multiuser receivers (near/far resistance). The capacity analysis in [Comaniciu and Poor, 2004c] builds upon results in [Tse and Hanly, 1999], but considers multi-hop transmissions and delay constraints.

The ad hoc network model considered in the analysis of [Comaniciu and Poor, 2004c] consists of K mobile nodes having a uniform stationary distribution over a square area of dimension $b \times b$. All nodes use random spreading codes and equal transmission powers, P_t, and are assumed to be active at any given time (worst case scenario). As in [Mostofa et al., 2001], a transmitter-oriented protocol is considered, in which each transmiting node has its own signature sequence. Although this implementation yields more complex receivers and longer aquisition times, it

has very good capturing probabilities, allowing multiple packet reception at the same receiver node. To avoid collisions, multiple concurrent transmissions from the same node are not allowed; instead, transmissions from one node to multiple destination nodes are time multiplexed. The analysis considers short range transmissions for the ad hoc network, such that a free space propagation path loss model is suitable. The link gain distribution is then derived using the free space propagation model, and an approximate distribution function for the distance between any two nodes [Miller, 2001], such that the link gain cumulative distribution function is given by

$$F_H(h) = \exp\left(-\frac{C}{h}\right), \ h > 0, \tag{3.101}$$

where $C = \frac{k^2}{4b^2}\lambda^2$, $k = 3.5$ is a constant related to the distance distribution function [Miller, 2001], and λ is the wavelength of the carrier.

Taking the derivative of (3.101) we obtain the probability density function for the link gain:

$$f_H(h) = \frac{C}{h^2} \exp\left(-\frac{C}{h}\right), \ h \geq 0. \tag{3.102}$$

Using (3.102) the mean link gain can be easily computed to be:

$$E_H \approx C\left[E_1(\delta_m^2 C) - E_1(\delta_M^2 C)\right], \tag{3.103}$$

where $E_1(x) = \int_x^\infty \frac{1}{t}\exp(-t)dt$ is the exponential integral.

The traffic can be transmitted directly between any two nodes, or it can be relayed through intermediate nodes. It is assumed that the end-to-end delay can be measured by the number of hops required for a route to be completed. The quality of service requirements for the ad hoc network are the bit error rate (mapped into an SIR requirement), the average source-destination throughput (T_{S-D}), and the transmission delay. Both the throughput and the delay are influenced by the maximal number of hops allowed for a connection and consequently, by the network diameter D. The network diameter is defined as the longest shortest path (measured in number of hops) between any source destination nodes in the network.

Using arguments similar to those in [Gupta and Kumar, 2000], a simplified computation shows that, if the number of hops for a transmission is D, then each node generates $Dl(K)$ traffic for other nodes, where $l(K)$ represents the traffic generation rate for a given node. Thus, the total traffic in the network must meet the stability condition $Dl(K)K \leq W/N$, where W is the system bandwidth, and N is the spreading gain.

This implies that the average source-destination throughput that can be supported by the network must meet the condition

$$T_{S-D} \le \frac{W}{ND}. \qquad (3.104)$$

In [Gupta and Kumar, 2000], it is argued that although (3.104) shows that the throughput decreases with an increase in the number of hops required, this does not account for the fact that if the range of a node increases, more collisions occur and the throughput decreases. In contrast to the model considered in [Gupta and Kumar, 2000], using a CDMA air interface allows multiple packet reception without collisions, and thus, a decrease of the network diameter D is obtained as a result of better physical links, and thus directly translates into a throughput increase for the network.

In terms of SIR requirements, a connection can be established between two nodes if the SIR is greater than or equal to the target SIR γ. The obtained SIR for a particular link is random due to the randomness of the nodes' positions.

The network capacity is defined to be the maximal number of nodes that can be supported such that both the SIR constraints and the delay constraints can be met for any arbitrary source-destination pair of nodes. The delay constraints are mapped into a maximum network diameter constraint D. Using geometric arguments, we will show that this further translates into a link probability constraint p, which at the physical layer represents the probability that a randomly chosen link is feasible. In the following section, we characterize the ad hoc network asymptotic capacity for the case in which the number of nodes and the spreading gain increase without bound, while their ratio is fixed.

2.1   Asymptotic Capacity

Physical Layer Performance

The physical layer capacity is derived under the assumption that a link probability constraint p has been imposed by the network layer, such that delay constraints can be met. In the physical layer, the link probability p is affected by the level of interference in the network and thus it is very sensitive to the choice of receiver. Capacity results for the matched filter, LMMSE detector and decorrelator, are presented in the following for the synchronous transmission case. These results are then extended to the more realistic scenario in which nodes transmit asynchronously.

Matched Filter Receiver - Synchronous Transmission

The SIR condition for an arbitrary node k using a matched filter receiver in a network with random, normalized spreading sequences can be expressed as:

$$SIR_k = \frac{P_t h_k}{\sigma^2 + \frac{1}{N} \sum_{\ell=1, \ \ell \neq k}^{K} P_t h_\ell} = \frac{h_k}{SNR^{-1} + \frac{1}{N} \sum_{\ell=1, \ \ell \neq k}^{K} h_\ell} \geq \gamma,$$
$$(3.105)$$

where we define the $SNR = \frac{P_t}{\sigma^2}$ to be the ratio of the transmit power to the noise power.

Denoting by α the fixed ratio K/N and letting the number of nodes and the spreading gain go to infinity, by using the law of large numbers [Yates and Goodman, 1999], it follows that: $\frac{1}{N} \sum_{\ell=1, \ \ell \neq k}^{K} h_\ell \rightarrow \alpha E_H$, with E_H computed as in (3.103).

The network diameter guarantees require a link probability value equal to p. This translates into a physical layer condition

$$P\left(H \geq \gamma SNR^{-1} + \alpha \gamma E_H\right) = P\left(H \geq T_{MF}\right) = p. \qquad (3.106)$$

Using the notation $T_{MF} = \gamma SNR^{-1} + \alpha \gamma E_H$, the network diameter condition renders an SNR condition

$$\gamma SNR^{-1} + \alpha \gamma E_H = T_{MF} \Rightarrow SNR = \frac{\gamma}{T_{MF} - \alpha \gamma E_H}, \qquad (3.107)$$

where T_{MF} can be derived using (3.101) as follows:

$$p = 1 - F_H(T_{MF}) = 1 - \exp\left(-C\frac{1}{T_{MF}}\right); \qquad (3.108)$$

or

$$T_{MF} = \frac{C}{\log\left(\frac{1}{1-p}\right)}. \qquad (3.109)$$

Equation (3.107) implies that a positive power solution exists if and only if

$$\alpha_{MF} < \frac{T_{MF}}{\gamma E_H} = \frac{\frac{C}{\log\left(\frac{1}{1-p}\right)}}{\gamma E_H}. \qquad (3.110)$$

For ad hoc networks, it is most likely that the mobile nodes are energy limited so that we impose a maximal power transmission limit $\bar{P}_t$. Denoting $SNR_c = \bar{P}_t/\sigma^2$, the ad hoc network capacity becomes:

$$\alpha_{MF} \leq \frac{T_{MF}}{\gamma E_H} - \frac{1}{E_H SNR_c} = \frac{\frac{C}{\log\left(\frac{1}{1-p}\right)}}{\gamma E_H} - \frac{1}{E_H SNR_c}. \qquad (3.111)$$

Decorrelator - Synchronous Transmission

Recall that (see Section 1.1 of this chapter) the SIR of an arbitrary node in an asymptotically large network using decorrelating receivers can be expressed as

$$SIR_d = \begin{cases} \frac{P_t h(1-\alpha)}{\sigma^2}, & \alpha < 1, \\ 0, & \alpha \geq 1 \end{cases}. \qquad (3.112)$$

Thus, if no power constraints are imposed, the network capacity region is

$$\alpha_d < 1. \qquad (3.113)$$

If power constraints are imposed, and $SNR \leq SNR_c$ (SNR_c is the maximal SNR allowed), the physical layer constraint can be expressed as

$$P\left(H \geq \frac{\gamma}{SNR(1-\alpha)}\right) = p. \qquad (3.114)$$

If we define $T_d = \frac{\gamma}{SNR(1-\alpha)}$, the feasibility condition becomes

$$SNR = \frac{\gamma}{T_d(1-\alpha)} \leq SNR_c. \qquad (3.115)$$

Imposing a network constraint on the T_d value, $T_d = \dfrac{C}{\log\left(\frac{1}{1-p}\right)}$, the asymptotic capacity region for a network using decorrelating receivers and having transmission power constraints is given as

$$\alpha_d \leq 1 - \frac{\gamma}{T_d SNR_c} = 1 - \frac{\gamma}{\frac{C}{\log\left(\frac{1}{1-p}\right)} SNR_c}. \qquad (3.116)$$

LMMSE Receiver - Synchronous Transmission

To derive the asymptotic ad hoc network capacity for the case of LMMSE receivers, we express the SIR for an arbitrary node k in a large network with equal transmission powers as follows:

$$SIR_k = \frac{h_k}{SNR^{-1} + \frac{1}{N}\sum_{\ell=1,\ell\neq k}^{K} \frac{h_k h_\ell}{h_k + h_\ell SIR_k}}. \qquad (3.117)$$

Imposing the QoS condition: $SIR_k \geq \gamma$, $\forall k = 1, 2, \ldots, K$, (where γ is the target SIR), we have

$$SIR_k \geq \frac{h_k}{SNR^{-1} + \frac{1}{N}\sum_{\ell=1,\ell\neq k}^{K} \frac{h_k h_\ell}{h_k + h_\ell \gamma}} = \gamma. \qquad (3.118)$$

Again, fixing $\alpha = K/N$, as the number of nodes and the spreading gain increase without bound we can apply the law of large numbers, so that,

$$\frac{1}{N} \sum_{\ell=1, \, \ell \neq k}^{K} \frac{h_k h_\ell}{h_k + h_\ell \gamma} = \alpha \frac{1}{K} \sum_{\ell=1, \, \ell \neq k}^{K} \frac{h_k h_\ell}{h_k + h_\ell \gamma} \to \alpha E\{H|h_k\},$$

where we use the notation $E\{H|h_k\}$ to denote the normalized conditional average interference (normalized to the number of nodes per dimension). It is shown in [Comaniciu and Poor, 2004c] that $E\{H|h_k\}$ can be expressed as:

$$E\{H|h_k\} = C \exp\left(\frac{C\gamma}{h_k}\right) \left[E_1\left(\delta_m^2 C + \frac{C\gamma}{h_k}\right) - E_1\left(\delta_M^2 C + \frac{C\gamma}{h_k}\right) \right].$$

$$(3.119)$$

Thus, the link probability constraint becomes

$$P\left(H \geq \gamma SNR^{-1} + \alpha\gamma E\{H|h\}\right) = p. \qquad (3.120)$$

We define the function $f(h) = h - \gamma SNR^{-1} - \alpha\gamma E\{H|h\}$ and we plot it in Fig. 3.13. We observe that $f(h)$ is a monotonically increasing function of h for the region of interest, and thus we can express the condition (3.120) as

$$P\left(H \geq T_{MMSE}\right) = p. \qquad (3.121)$$
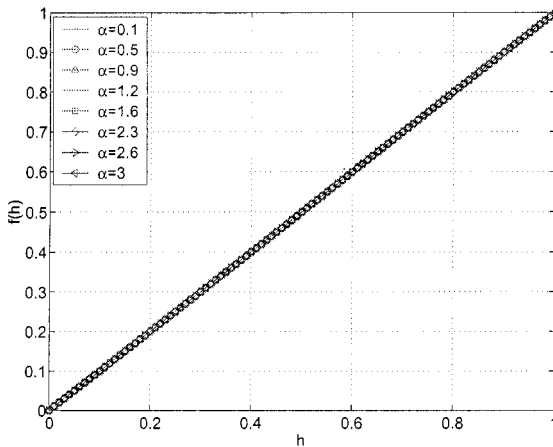
Figure 3.13.   SIR condition monotonicity (all curves are coincident) (reprinted with permission from [Comaniciu and Poor, 2004c])

Equation (3.121) has the same solution as in the previously analyzed cases, and the physical layer constraint becomes

$$SNR = \frac{\gamma}{T_{MMSE} - \alpha\gamma E\{H|h = T_{MMSE}\}}. \tag{3.122}$$

A positive transmitting power solution exists if and only if

$$\alpha_{MMSE} < \frac{T_{MMSE}}{\gamma E\{H|h = T_{MMSE}\}}; \tag{3.123}$$

or equivalently,

$$\alpha_{MMSE} < \frac{\frac{C}{\log\left(\frac{1}{1-p}\right)}}{\gamma C \left(\frac{1}{1-p}\right)^{\gamma} \left[E_1\left(\delta_m^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right) - E_1\left(\delta_M^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right)\right]}. \tag{3.124}$$

If power constraints are imposed, the capacity region becomes

$$\alpha_{MMSE} \leq \frac{T_{MMSE}}{\gamma E\{H|h = T_{MMSE}\}} - \frac{1}{E\{H|h = T_{MMSE}\}SNR_c}; \tag{3.125}$$

or equivalently,

$$\alpha_{MMSE} \leq$$

$$\frac{\frac{C}{\log\left(\frac{1}{1-p}\right)}}{\gamma C \left(\frac{1}{1-p}\right)^{\gamma} \left[E_1\left(\delta_m^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right) - E_1\left(\delta_M^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right)\right]} -$$

$$\frac{1}{C \left(\frac{1}{1-p}\right)^{\gamma} \left[E_1\left(\delta_m^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right) - E_1\left(\delta_M^2 C + \gamma \log\left(\frac{1}{1-p}\right)\right)\right] SNR_c}. \tag{3.126}$$

Figure 3.14 illustrates the physical layer capacity as a function of the link probability constraint p for the three receivers considered, with and without power constraints. For the power-constrained case, a maximal transmission power of $\bar{P}_t = 10^4\sigma^2$ is considered for this example. A target SIR $\gamma = 5$ is imposed. From Fig. 3.14 we can observe that there is a significant capacity advantage if multiuser receivers are used, and conversely, for given capacity requirements, substantial power savings can be achieved by networks using multiuser receivers. As expected, the LMMSE receiver performs the best due to its property of maximizing the SIR. For higher transmission rates and lower delay requirements (translated into a high link probability constraint) using the matched filter is not feasible.
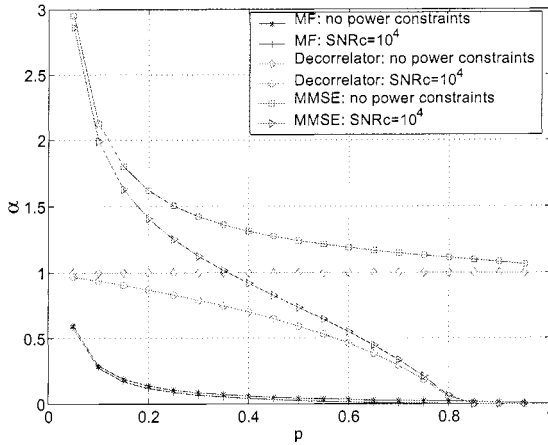
Figure 3.14. Physical layer capacity for given link probability constraint: synchronous transmission (reprinted with permission from [Comaniciu and Poor, 2004c])

Decorrelator - Asynchronous Transmission

Recall that (see Section 1.6 in this chapter), the SIR for the decorrelator can be approximated as

$$SIR_d = \begin{cases} \frac{P_t h(1-2\alpha)}{\sigma^2}, & \alpha < 1/2 \\ 0, & \alpha \geq 1/2 \end{cases} \qquad (3.127)$$

Therefore, the capacity results from the previous subsection can be straightforwardly extended to

$$\alpha_d < 1/2, \qquad (3.128)$$

when no power constraints are imposed.

If power constraints are imposed $(SNR \leq SNR_c)$, we derive the capacity region as

$$\alpha_d \leq \frac{1}{2} - \frac{\gamma}{\frac{2C}{\log\left(\frac{1}{1-p}\right)} SNR_c}. \qquad (3.129)$$

LMMSE Receiver - Asynchronous Transmission

According to our previous discussion in Section 1.6, to characterize the capacity of an asynchronous ad hoc network using LMMSE receivers, we must rely on the lower bound obtained for the achievable SIR:

$$SIR_k = \frac{P_k}{\sigma^2 + \alpha E\{I(\tau P, P_k, SIR_k) + I((1-\tau)P, P_k, SIR_k)\}}, \qquad (3.130)$$

where τ is a random variable that characterizes the delay associated with an arbitrary node (a fraction of a symbol duration), and the expectation is taken with respect to P and τ. Since the received power P can be expressed as $P = P_t h$, for equal transmit powers for all nodes, (3.130) becomes

$$SIR_k = \frac{h_k}{SNR^{-1} + \alpha E\{I(\tau P_t h, P_t h_k, SIR_k) + I((1 - \tau)P_t h, P_t h_k, SIR_k)\}}.$$
$$(3.131)$$

It is straightforward to see that $\alpha E\{I(\tau P_t h, P_t h_k, SIR_k) + I((1 - \tau)P_t h, P_t h_k, SIR_k)\}$ can be expressed as

$$\alpha E\{E\{H|h_k, \tau\}\} =$$

$$\alpha E\left\{ C\tau \exp\left(\frac{C\gamma\tau}{h_k}\right) \left[E_1\left(\delta_m^2 C + \frac{C\gamma\tau}{h_k}\right) - E_1\left(\delta_M^2 C + \frac{C\gamma\tau}{h_k}\right) \right] + \right.$$

$$+ C(1 - \tau) \exp\left(\frac{C\gamma(1 - \tau)}{h_k}\right) \left[E_1\left(\delta_m^2 C + \frac{C\gamma(1 - \tau)}{h_k}\right) - \right.$$

$$\left. \left. - E_1\left(\delta_M^2 C + \frac{C\gamma(1 - \tau)}{h_k}\right) \right] \right\}. \quad (3.132)$$

Using an identical derivation for the network capacity as for the synchronous case, all the capacity formulas hold with $E\{H|h_k\}$ replaced by $E\{E\{H|h_k, \tau\}\}$. In Fig. 3.15, we provide capacity comparisons between networks using LMMSE receivers in the synchronous and the asynchronous cases. (Here we assume that τ is uniformly distributed on $[0, 1]$.)

Network Capacity

The overall network capacity is determined such that both physical layer and network layer QoS requirements can be met. In the previous section we have determined the maximal number of active nodes that can be simultaneously supported by the network, as a function of a link probability constraint, which maps to the transmission delay requirement. In this section, we use geometric arguments to determine the dependence of the link probability on the network diameter constraint (which is a surrogate for the delay constraint).

We consider the asymptotic case, in which we have an infinite number of nodes in the considered square area. The nodes are uniformly distributed, and we ignore the edge effects: the square area can be con-
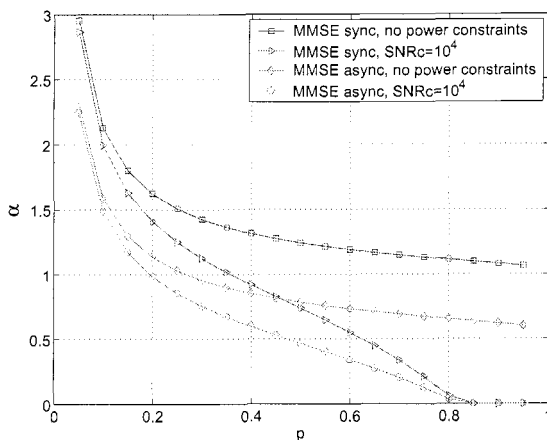
Figure 3.15. Capacity comparisons for ad hoc networks with LMMSE receivers: synchronous versus asynchronous transmission (reprinted with permission from [Comaniciu and Poor, 2004c])

sidered to be a part of a multiple cell layout. It can be seen from Fig. 3.16 that the worst case distance is obtained when the source and destination nodes are on the opposite vertices of the square. We showed in the previous section that the link probability p can be expressed as

$$p = P(H \geq T) = P\left(d_{max} \leq \frac{\lambda}{\sqrt{T}}\right), \qquad (3.133)$$

where the threshold T depends on the particular receiver structure used. This is equivalent to saying that a reliable transmission exists within a radius d_{max} of the transmitting node. Consider now a diameter restriction of $D = 2$, as in Fig. 3.16. In order to be able to transmit from the source node (SN) to the destination node (DN) using only one intermediate node (IN), the minimal value for d_{max} must be

$$d_{max} = \frac{b\sqrt{2}}{2}. \qquad (3.134)$$

For a generic value of D, the constraint becomes

$$d_{max} = \frac{b\sqrt{2}}{D}. \qquad (3.135)$$

From (3.133) and (3.135), we obtain a threshold requirement of

$$T = \frac{\lambda^2 D^2}{2b^2}. \qquad (3.136)$$
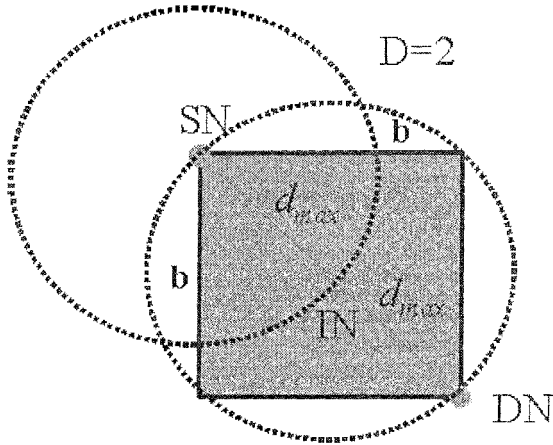
Figure 3.16. Network diameter constraint (reprinted with permission from [Comaniciu and Poor, 2004c])

To determine the link probability constraint we introduce (3.136) into the link probability expression: $p = 1 - \exp(-C/T)$. In Fig. 3.17 we illustrate the mapping between the link probability constraint and the required network diameter.
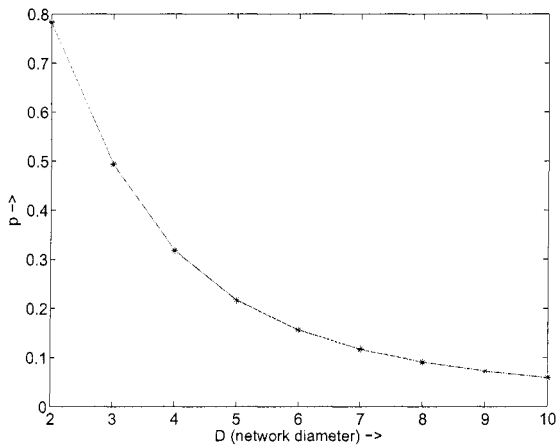
Figure 3.17. Link probability requirement (reprinted with permission from [Comaniciu and Poor, 2004c])

Using the link probability constraint values previously determined, the ad hoc network capacity can be determined for given delay (network diameter) specifications. In Fig. 3.18 an example for the network capacity for a network diameter constraint of $D = 2$ is presented. Fig. 3.18 shows the number of users per dimension that can be supported in an ad hoc network for a given delay constraint, as a function of the maximum transmission power requirement, $SNR_c = \bar{P}_t/\sigma^2$.
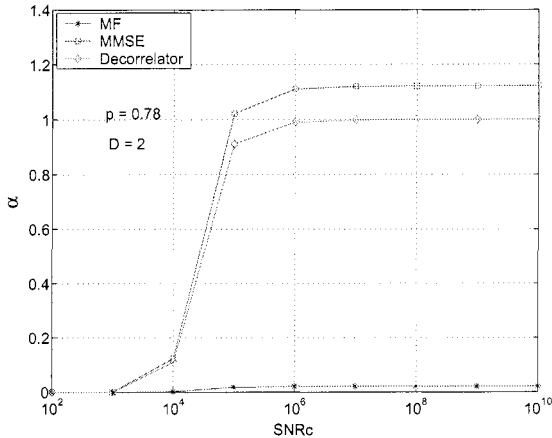
Figure 3.18. Ad hoc network capacity for delay sensitive traffic, $D = 2$ (reprinted with permission from [Comaniciu and Poor, 2004c])

Network Throughput

As we have seen in the previous subsections, ad hoc network capacity is greatly enhanced by using a CDMA access method and separating the users using multiuser detectors. Tight power and delay constraints can thus be met in such networks. We will show now that using multiuser detectors in CDMA ad hoc networks improves also the overall throughput of the network. To see this, we compare the network throughput that can be achieved for our analysis by the LMMSE receiver, with the scenario described in [Gupta and Kumar, 2000], in which random access is used. No delay constraints are enforced, and very similar network models are used for comparison: all nodes are randomly located and independently and uniformly distributed in a unit area (disc in [Gupta and Kumar, 2000], square in our analysis), each node transmits traffic to a randomly chosen destination, all nodes transmit with the same power and the transmission rate is R. Both synchronous and asynchronous

transmission cases are considered for the CDMA network and LMMSE receivers are considered.

For the random access scenario, the order of the average throughput capacity has been shown in [Gupta and Kumar, 2000] to be $l(K) = \theta \left(\dfrac{R}{\sqrt{K \log(K)}} \right)$. For the CDMA network we approximate the network throughput based on (3.104): $l(K) \approx \dfrac{R}{D}$, where $R = W/N$. We compare the network throughput for the Gupta-Kumar analysis (G-K) [Gupta and Kumar, 2000], with both a synchronous and an asynchronous CDMA network using LMMSE receivers. The same numerical values as before are selected for the example plotted in Fig. 3.19, which shows the normalized network throughput as a function of the number of nodes per unit area. The spreading gain is chosen to be $N = 32$. We can see that, although the CDMA ad hoc network capacity also decreases with the increase of the number of nodes per unit area, its capacity is significantly higher than the random access network (G-K). Also, the use of LMMSE receivers yields unreduced throughputs for the network for a fairly large network (approx. 40 nodes per unit area for synchronous transmission). Of course this advantage comes at the price of an increased implementation complexity in acquiring the signature sequences for all users and dynamically adjusting the receivers.
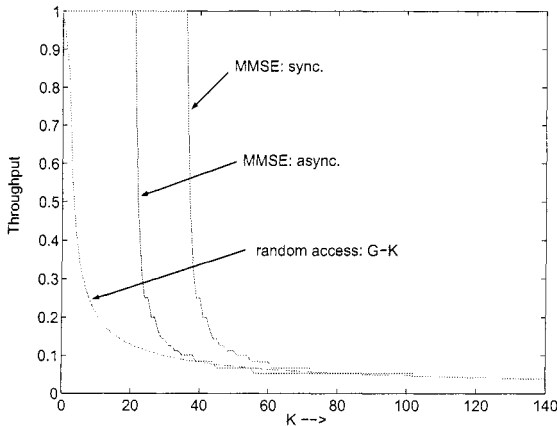
Figure 3.19.   Network throughput comparison (reprinted with permission from [Comaniciu and Poor, 2004c])

2.2   Capacity for Finite Networks: Simulations

Since the above capacity results are asymptotic in nature, a performance validation through simulations is required for practical finite networks. The experiments presented in this section consider unlimited power transmission for the LMMSE case, and maximal power constraints for the decorrelator, $\bar{P}_t = 10^4 \sigma^2$ (the case of the decorrelator with unlimited transmission power is trivial: $\alpha < 1$). For implementation simplicity, all numerical results are obtained for synchronous transmission, and using $b = 6$, $\lambda = 0.1$ m and $\gamma = 5$. The experiments consist of selecting a finite (variable) number of nodes and randomly generating their locations uniformly across a square area. Then, the link gains, and consequently the achieved SIRs are computed for all pair of nodes, using (3.105), (3.112), and (3.117), respectively. We note that the simulations do not consider the SIR formulas' accuracy for finite systems, as this issue has already been discussed earlier in the chapter. In the simulations, a link is considered to be feasible if the computed SIR is greater than or equal to the target SIR, and the network diameter computation uses Dijkstra's algorithm [Bertsekas and Gallager, 1992]. The probabilities associated with a range of network diameters are determined. An infinite network diameter means that the network is disconnected. The link probability p is also determined and compared with the theoretical results.

Some simulation examples are presented in Tables 3.1, 3.2, and 3.3. It can be seen that both the physical layer capacity results, reflected in the achievable link probability p, and the network performance results (i.e., the achieved network diameter) are very close to the asymptotic ones, especially for larger numbers of nodes in the network cell (the considered square area).

2.3   Implications for Admission Control

The results in this chapter provide simple abstract models for the physical layer performance in various scenarios, and therefore, they help to bridge the physical and the network layer design. At higher layers, QoS provisioning will be based on these simple models. Once the physical layer capacity is determined, resources can be managed at the network layer using an admission control policy. Based on the results presented in this chapter, admission control algorithms can be implemented for both cellular and ad hoc networks. This topic is the subject of the following chapter.

Table 3.1. Simulation Results for Ad Hoc Networks with Delay Constraints: MF (reprinted with permission from [Comaniciu and Poor, 2004c])

K/N	p(analysis/sim.)	D (asymptotic/sim.)
44/1024	p=0.5117/0.6107	$D \approx 3$
31/256	p=0.2246/0.3093	$D \approx 5$
144/512	p=0.1037/0.1127	$D \approx 8$
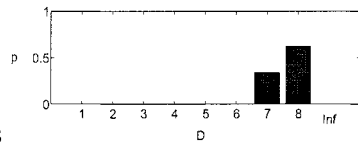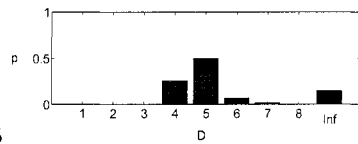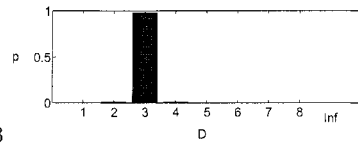
Table 3.2. Simulation Results for Ad Hoc Networks with Delay Constraints: Decorrelator (reprinted with permission from [Comaniciu and Poor, 2004c])

K/N	p (analysis/sim.)	D (asymptotic/sim)
60/512	$p= 0.7773/0.7472$	$D \approx 2$
120/1024	$p=0.7773/0.7510$	$D \approx 2$
28/64	$p=0.6160/0.5670$	$D \approx 3$
92/128	$p=0.3803/0.3392$	$D \approx 4$
96/128	$p=0.3464/0.3074$	$D \approx 4$
100/128	$p=0.3107/0.2764$	$D \approx 4$
57/64	$p=0.1698/0.1515$	$D \approx 7$
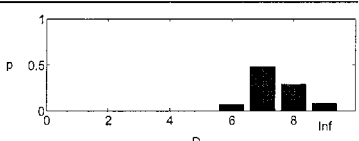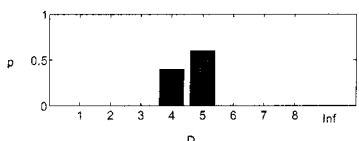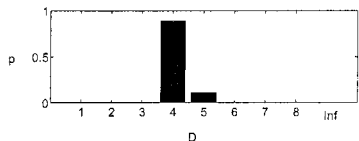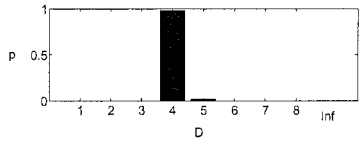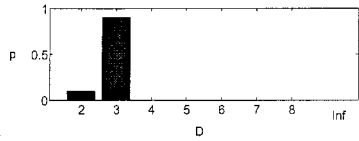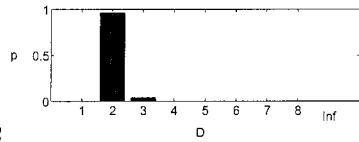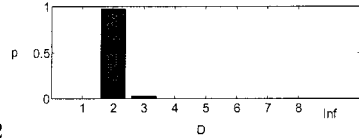
Table 3.3. Simulation Results for Ad Hoc Networks with Delay Constraints: LMMSE (reprinted with permission from [Comaniciu and Poor, 2004c])

K/N	p (analysis/sim.)	D (asymptotic/sim.)
38/32	p=0.6056/0.7491	$D \approx 2$
39/32	p=0.5415/0.4886	$D \approx 3$
42/32	p=0.4024/0.433	$D \approx 3/4$
45/32	p=0.3137/0.3260	$D \approx 4$
46/32	p=0.2913/0.2983	$D \approx 4$
48/32	p=0.2537/0.2590	$D \approx 5$
57/32	p=0.1546/0.1584	$D \approx 7$

Chapter 4

INTEGRATED ADMISSION CONTROL

1.   Cellular Wireless Networks

As we discussed in Chapter 2, the two simplest admission control schemes (complete sharing policy and threshold policy) are based on restricting the maximal number of users into the network using the physical layer capacity as an admission condition. As a consequence, simple admission control algorithms may be derived in a straightforward manner for each of the scenarios analyzed in Chapter 3.

Some earlier work on admission control has adopted very simple models for characterizing the physical layer performance. In [Holma and Laakso, 1999], the system performance improvement using multiuser detectors is quantified by a factor that represents the percentage of intracell interference cancelled, and which is determined experimentally. The capacity characterization in Chapter 3 permits a better understanding of the physical layer performance and impacts the design of upper layer protocols such as admission control.

Based on the SIR convergence results discussed in the previous chapter, an optimal admission control for a multi-class network is proposed in [Singh et al., 2001], which exploits the interplay between the LMMSE receiver performance and the network layer throughput performance (blocking probability). In particular, for Poisson arrivals of new calls, and exponential call durations, an optimal admission policy (which minimizes the blocking probability) is derived using the SMDP (semi-Markov decision process) [Tijms, 1986] theory.

In a semi-Markov decision process, a system of interest is described by a sequence of states, such that the next state of the system depends only on the current state and on an action taken. For call admission

control, the action is selected according to an admission control policy, such that a specified QoS criterion is optimized. SMDPs can essentially be solved by considering an equivalent discrete time average cost Markov decision process, using a process called uniformization [Bertsekas, 1995]. As a consequence, algorithms such as policy iteration, value iteration and linear programming (LP), can be used to provide solutions for the SMDP problem. An advantage of the LP approach is that optimization constraints can easily be added. In [Singh et al., 2001], the admission control policy is determined such that a weighted sum of blocking probabilities for all classes of users is minimized, subject to constraints on blocking probabilities for specific classes of users.

An SMDP is completely characterized [Tijms, 1986] by the following quantities:

- the state space X;

- the action space $\mathbf{A}$;

- $p_{\mathbf{xy}}(\mathbf{a})$ = the probability that at the next decision epoch the system will be in state $\mathbf{y}$, if action $\mathbf{a}$ is selected at the current state $\mathbf{x}$;

- $\tau_{\mathbf{x}}(\mathbf{a})$ (sojourn time)= the expected time until the next decision epoch after action $\mathbf{a}$ is chosen in the present state $\mathbf{x}$:

$$\tau_{\mathbf{x}}(\mathbf{a}) > 0, \forall \mathbf{x} \in X, \ \mathbf{a} \in \mathbf{A_x},$$

where $\mathbf{A_x}$ represents the admissible action space (to be defined below); and

- $c(\mathbf{x}, \mathbf{a})$=the expected costs incurred until the next decision epoch after action $\mathbf{a}$ is chosen in the current state $\mathbf{x}$.

The admission control policy derived in [Singh et al., 2001] optimizes the network layer performance, given SIR constraints at the physical layer (Fig. 4.1). The state space of the SMDP is constructed by selecting all possible configurations of users that meet target SIR requirements at the physical layer, for given fixed transmission powers. To improve the network performance, new call requests that cannot be allowed in the network are queued, using finite length buffers. The possible actions for the admission control are: admit a new or queued user into the network, queue a new call request, or reject a new call request. The network performance measure is the blocking probability, and it is proved in [Singh et al., 2001] that this probability can be expressed as an average cost criterion for the SMDP process. However, the average delay cannot be expressed as an average cost criterion, and thus call connection delays were not considered as a QoS measure in [Singh et al., 2001]. Detailed
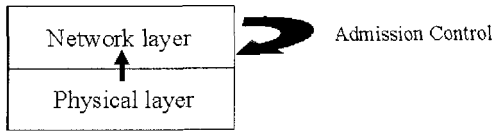
Figure 4.1. Optimization of network layer performance (reprinted with permission from [Comaniciu and Poor, 2003a])
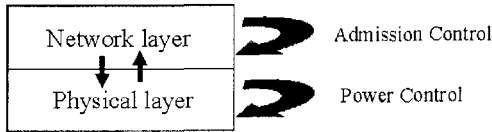
Figure 4.2. Joint optimization across physical and network layers (reprinted with permission from [Comaniciu and Poor, 2003a])

characterization of the SMDP, and complete derivations for its dynamics and cost functions can be found in [Singh et al., 2001].

In this section, we discuss in more detail the work in [Comaniciu and Poor, 2003a], which extends and completes [Singh et al., 2001]. In [Comaniciu and Poor, 2003a], joint optimization across the network and physical layers is proposed (Fig. 4.2). At the physical layer, the QoS requirements are specified in terms of a target SIR, and optimal target powers are dynamically adjusted according to the current number of users in the system. The network QoS is specified in terms of the blocking probabilities and the call connection delays. The network layer guarantees that both the physical layer and the network layer QoS are met by employing admission control.

The analysis in [Comaniciu and Poor, 2003a] models the network as an equivalent queueing system, as described in Chapter 2 (Fig. 2.16). The service rate for each queue is varied by the admission control such that the power control feasibility condition holds and hence all users can meet their target SIRs. A single cell, power controlled synchronous CDMA system is considered, which supports J classes of users, characterized by different target SIRs, γ_j, different blocking probability requirements, P_b^j, and different connection delay constraints, Ξ_j, $j = 1, 2, \ldots, J$. Requests for connections occur with rates λ_j, $j = 1, 2, \ldots, J$ and are Poisson distributed. The call durations are exponentially distributed and the mean duration for class j is μ_j, $j = 1, 2, \ldots, J$.

The admission policy is constructed by accounting for both the physical layer QoS (the physical layer capacity is determined as given in

Theorem 3.14), as well as the network layer QoS (blocking probabilities and call connection delays). At the physical layer, target powers are ajusted according to the current number of users admitted into the system (Theorem 3.14).

At the network layer, the equivalent queueing problem consists of J $M/M/K_j/B(j)$ queues ($B(j)$ represents the length of buffer for queue j, and K_j is the number of class j users), $j = 1, 2, \ldots, J$. The service rates depend on the current number of connections (such that $\alpha_j = K_j/N$ satisfies (3.53)), and also on the current number of calls waiting for connection (according to delay and blocking constraints).

The state space

The state of the jth queue is characterized by the number of users $n_q^j(t)$ in the queue at time t, $t > 0$, and the number of servers $n_s^j(t)$ at time t (equivalent to the number of connections admitted for class j). The state of the system at decision epoch t can be defined as

$$\mathbf{x}(t) = [n_q^1(t), n_s^1(t), n_q^2(t), n_s^2(t), \ldots, n_q^J(t), n_s^J(t)]. \qquad (4.1)$$

Since the arrivals and departures of users are random, $\{\mathbf{x}(t),\ t > 0\}$ represents a finite state stochastic process. The state space X is comprised of all state vectors $\mathbf{x}$, such that SIR constraints can be met:

$$X = \left\{ \mathbf{x} : n_q^i \le B(i),\ i = 1, 2, \ldots, J;\ \sum_{j=1}^{J} \frac{n_s^j}{N} \Lambda_j \le 1 \right\}, \qquad (4.2)$$

where we use the notation $\Lambda_j = (L-1)\nu_j\gamma_j + \frac{\gamma_j(1+\nu_j)}{1+\gamma_j}$.

Decision epochs

Every time a new user arrives and requests a new connection, and any time a departure occurs, the state of the system changes. Since these changes in the system state should affect the admission process, the decision epochs are the set of all arrival and departure instances.

The action space

At each decision epoch, an action $\mathbf{a}$ is chosen that determines how the admission control will perform at the next decision moment. The action vector $\mathbf{a}$ is state dependent, and its components depend on the type of event: arrival or departure. In general, action $\mathbf{a}$ at decision epoch t is defined as

$$\mathbf{a}(t) = [a_1^a(t), a_1^d(t), a_2^a(t), a_2^d(t), \ldots, a_J^a(t), a_J^d(t)], \qquad (4.3)$$

where a_j^a denotes the action for queue j if an arrival occurs and a_j^d denotes the action for queue j if a departure occurs, which are defined as follows:

$$a_j^a = \begin{cases} 0; & \text{maintain the number of servers for queue } j \\ 1; & \text{increase (by 1) the number of servers for queue } j \end{cases}$$

$$a_j^d = \begin{cases} 0; & \text{decrease (by 1) the number of servers for queue } j \\ 1; & \text{maintain the number of servers for queue } j \end{cases}$$

The action space can be defined as the set of all possible actions:

$$\mathbf{A} = \{0, 1\}^{2J}. \tag{4.4}$$

The action space must be restricted for a given state $\mathbf{x} \in X$, such that the selected action will not result in a transition into a state that is not allowed (not in X). Also, the admissible action space $(\mathbf{A_X})$, must be restricted such that $1/\tau_{\mathbf{X}}(\mathbf{a}) > 0$, $\forall \mathbf{x} \in X$, $\mathbf{a} \in \mathbf{A_X}$, i.e, $(a_1^a, a_2^a, \ldots, a_J^a) \neq (0, 0, \ldots, 0)$ if the system is in state $\mathbf{x} = (0, 0, 0, 0, \ldots, 0, 0)$.

Thus, the admissible action space $\mathbf{A_X}$ can be defined as:

$$\mathbf{A_X} = \{\mathbf{a} \in \mathbf{A} : a_j^a = 0 \text{ if } \mathbf{x} + (0, 0, \ldots, \underbrace{0, 1}_{j}, \ldots, 0, 0) \notin X, \text{ and}$$

$$(a_1^a, a_2^a, \ldots, a_J^a) \neq (0, 0, \ldots, 0) \text{ if } \mathbf{x} = (0, 0, 0, 0, \ldots, 0, 0)\}. \tag{4.5}$$

The state dynamics

The state dynamics of an SMDP can be characterized by the transition probabilities of its embedded chain, and the expected sojourn time for each state-action pair [Bertsekas, 1995].

We define the following notation:

- $\delta(x) = \begin{cases} 0; & x = 0 \\ 1; & x > 0 \end{cases}$

- $\mathbf{x}_j = [n_q^j, n_s^j]$ represents the state vector for class j users, such that the state vector for the system can be expressed as $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_J]$.

- $\mathbf{e}_j^q$ represents a vector of dimension $2J$, containing only zeros except for the position $2(j - 1) + 1$, which contains a 1; $\mathbf{x} + \mathbf{e}_j^q$ is equivalent to $\mathbf{x}_j + [1, 0]$ and maps an increase in the queue of class j users by 1.

- $\mathbf{e}_j^s$ represents a vector of dimension $2J$, containing only zeros except for the position $2(j-1)+2$, which contains a 1; $\mathbf{x}+\mathbf{e}_j^s$ is equivalent to $\mathbf{x}_j + [0,1]$ and maps an increase in the number of servers for class j users (number of admitted users) by 1.

Derivations of $p_{\mathbf{xy}}(\mathbf{a})$ and $\tau_{\mathbf{x}}(\mathbf{a})$ rely on the statistical properties of the arrival and departure processes, which are Poisson and mutually independent. It follows that the cumulative process is also Poisson and thus the cumulative event rate is the sum of the rates for all constituent processes. It should be mentioned that arrivals that are blocked do not constitute an event such that the cumulative process includes only the unblocked arrivals, which are also Poisson with rates $\lambda_j(1-P_b^j)$. Hence, the inter-event time $\tau_x(\mathbf{a})$ (the expected sojourn time) can be defined as the inverse of the event rate:

$$\tau_{\mathbf{x}}(\mathbf{a}) =$$

$$= \left[\sum_{j=1}^{J} \lambda_j a_j^a + \sum_{j=1}^{J} \lambda_j(1-a_j^a)\delta(B(j)-n_q^j) + \sum_{j=1}^{J} \mu_j n_s^j \right]^{-1}. \qquad (4.6)$$

Equation (4.6) can be interpreted as follows: the embedded chain always changes state when an arrival occurs unless the arrival is blocked (the queue is full and no new servers are allocated for that particular queue), and also, it always changes state when a departure occurs.

To derive the transition probabilities, the decomposition property of a Poisson process is used: an event of certain type occurs (e.g. arrival class j, departure class i) with a probability equal to the ratio between the rate of that particular type of event and the total cumulative event rate $1/\tau_{\mathbf{x}}(\mathbf{a})$. Hence, the transition probabilities for the embedded Markov chain are determined to be:

$$p_{\mathbf{xy}}(\mathbf{a}) = \begin{cases} \lambda_j a_j^a \tau_{\mathbf{x}}(\mathbf{a}); & \text{if } \mathbf{y} = \mathbf{x} + \mathbf{e}_j^s \\[2mm] \lambda_j(1-a_j^a)\delta(B(j)-n_q^j)\tau_{\mathbf{x}}(\mathbf{a}); & \text{if } \mathbf{y} = \mathbf{x} + \mathbf{e}_j^q \\[2mm] \mu_j n_s^j a_j^d \tau_{\mathbf{x}}(\mathbf{a}); & \text{if } \mathbf{y} = \mathbf{x} - \mathbf{e}_j^q \\[2mm] \mu_j n_s^j(1-a_j^d)\tau_{\mathbf{x}}(\mathbf{a}) + \mu_j n_s^j a_j^d(1-\delta(n_q^j))\tau_{\mathbf{x}}(\mathbf{a}); & \text{if } \mathbf{y} = \mathbf{x} - \mathbf{e}_j^s \\[2mm] 0; & \text{otherwise} \end{cases}$$

$$(4.7)$$

Optimal policy: linear programming approach

For any given state $\mathbf{x} \in X$, an action $\mathbf{a}$ is selected according to a specified policy $\mathbf{R}$. A stationary policy $\mathbf{R}$ is a function that maps the state space into the admissible action space, where the class of admissible policies can be defined as:

$$\mathbf{R}_{\mathbf{X},\mathbf{a}} = \{\mathbf{R} | \mathbf{R} : X \to \mathbf{A}_{\mathbf{X}}, 1/\tau_{\mathbf{X}}(\mathbf{R}) > 0\}. \tag{4.8}$$

According to [Bertsekas, 1995], an average cost criterion for a given policy $\mathbf{R}$ and an initial state $\mathbf{x}_0$ can be associated with the SMDP:

$$J_{\mathbf{R}}(\mathbf{x}_0) = \lim_{T \to \infty} \frac{1}{T} E \left\{ \int_0^T c(\mathbf{x}(t), \mathbf{a}(t)) dt \right\}. \tag{4.9}$$

An optimal policy $\mathbf{R}^*$, that minimizes an average cost criterion $J_{\mathbf{R}}(\mathbf{x}_0)$ for any initial state $\mathbf{x}_0$, exists under the weak unichain assumption [Tijms, 1986].

In (4.9), $c(\mathbf{x}(t), \mathbf{a}(t))$ can be interpreted as the expected cost until the next decision epoch, and will be selected to meet the network layer performance criteria as will be discussed shortly. An optimal policy for the above defined SMDP process can be determined using a linear programming approach. The optimal policy $\mathbf{R}^*(\mathbf{x}) \in \mathbf{R}_{\mathbf{X},\mathbf{a}}$, $\mathbf{x} \in X$, can be obtained using the decision variables $u^*_{\mathbf{x}\mathbf{a}}$, $\mathbf{x} \in X$, $\mathbf{a} \in \mathbf{A}_{\mathbf{X}}$, which are obtained by solving the linear program associated with the SMDP [Tijms, 1986]:

$$\min_{\mathcal{U}} \sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A}_{\mathbf{X}}} \sum_{j=1}^{J} \theta_j c_j(\mathbf{x}, \mathbf{a}) u_{\mathbf{x},\mathbf{a}} \tau_{\mathbf{X}}(\mathbf{a})$$

subject to the constraints

$$\sum_{\mathbf{a} \in \mathbf{A}_{\mathbf{y}}} u_{\mathbf{y}\mathbf{a}} - \sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A}_{\mathbf{X}}} p_{\mathbf{x}\mathbf{y}}(\mathbf{a}) u_{\mathbf{x}\mathbf{a}} = 0, \ \mathbf{y} \in X$$

and

$$\sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A}_{\mathbf{X}}} u_{\mathbf{x}\mathbf{a}} \tau_{\mathbf{X}}(\mathbf{a}) = 1, \tag{4.10}$$

where θ_j represents a weighting of the cost function for class j, and $\mathcal{U}$ is defined as:

$$\mathcal{U} = \{u_{\mathbf{x},\mathbf{a}} : u_{\mathbf{x},\mathbf{a}} \geq 0, \forall \mathbf{a} \in \mathbf{A}_{\mathbf{X}} \text{ and } \forall \mathbf{x} \in X\}.$$

A heuristic explanation [Tijms, 1986] for (4.10) is to interpret $u_{\mathbf{x}\mathbf{a}} \tau_{\mathbf{X}}(\mathbf{a})$ as the steady-state probability of being in state $\mathbf{x}$ and choosing action

a (for an aperiodic Markov chain). Hence, the objective function is to minimize the long run average of the cost function per unit time. The first constraint in (4.10) can be interpreted as a balance equation, and the second constraint requires that the sum of the steady state probabilities should be equal to 1.

After solving (4.10), an optimal policy for the admission control can be constructed as follows [Tijms, 1986]:

- $\forall \mathbf{x}$, choose any $\mathbf{a}^*$, such that $u^*_{\mathbf{x}\mathbf{a}^*} > 0$; then set the optimal policy for state $\mathbf{x}$ to be $\mathbf{R}^*(\mathbf{x}) = \mathbf{a}^*$.

As we will see shortly, for our purpose of meeting QoS requirements in terms of blocking probabilities and average delay constraints, it is important to be able to solve a constrained optimization. The linear programming approach allows us to introduce very easily probabilistic constraints related to an expected cost function $c'_j(\mathbf{x}, \mathbf{a})$ for class j [Tijms, 1986]:

$$\sum_{\mathbf{x} \in X^b_j, \ \mathbf{a} \in \mathbf{A}^b_{\mathbf{X},j}} c'_j(\mathbf{x}, \mathbf{a}) u_{\mathbf{x}\mathbf{a}} \tau_{\mathbf{X}}(\mathbf{a}) \leq C, \qquad (4.11)$$

where C ia a fixed value constraint, and X^b_j, $\mathbf{A}^b_{\mathbf{X},j}$ are the state subspace and the action subspace that result in blocked calls for class j .

When probabilistic constraints are imposed, the optimal policy becomes a randomized policy: in each state, an action $\mathbf{a}$ is chosen randomly according to a probability $\pi^*_{\mathbf{a}}(\mathbf{x}) = u^*_{\mathbf{x},\mathbf{a}} / \sum_{\mathbf{a} \in A_{\mathbf{X}}} u^*_{\mathbf{x},\mathbf{a}}$. The randomized policy can be specified as a matrix $\mathbf{R}^*_{(dim(X),dim(\mathbf{A}_{\mathbf{X}}))}$, with each entry given as $\mathbf{R}^*(i,j) = \pi^*_j(i)$. The $(i,j)^{th}$ entry for matrix $\mathbf{R}^*$ represents the probability that action j is selected when the system is in state i. The matrix $\mathbf{R}^*$ is determined off line, and the admission control randomly chooses actions at each decisions epoch, according to the corresponding probabilities from the matrix $\mathbf{R}^*$.

Cost functions and network QoS

The network layer performance measures are the blocking probabilities, P^j_b, $j = 1, \ldots, J$ (which reflect the network throughput capacity), and the average connection delays W_j, $j = 1, \ldots, J$. The network QoS requirements are specified as

$$\begin{cases} P^j_b \leq \Psi_j, \ j = 1, \ldots, J \\ W_j \leq \Xi_j, \ j = 1, \ldots, J \end{cases} . \qquad (4.12)$$

To determine the optimal admission policy, the cost functions $c(\mathbf{x}, \mathbf{a})$ and $c'(\mathbf{x}, \mathbf{a})$ must be defined.

Blocking probabililty

In [Singh et al., 2001], the authors proved that the blocking probability can be expressed as

$$P_b^j = \lim_{T \to \infty} \frac{1}{T} \mathbf{E} \left\{ \int_0^T (1 - a_j^a(t))(1 - \delta(B(j) - n_j^q(t))) \tau_{\mathbf{X}(t)}(\mathbf{a}(t)) dt \right\}.$$

(4.13)

The expression (4.13) represents the cumulative average blocking probability. We can then obtain an expression for the expected blocking probability until the next decision epoch for class j, when action $\mathbf{a}$ is chosen in current state $\mathbf{x}$:

$$c(\mathbf{x}, \mathbf{a})_j = (1 - a_j^a)(1 - \delta(B(j) - n_j^q)).$$

(4.14)

To minimize a weighted sum of blocking probabilities for all users in the system, (4.14) gives the expected cost in (4.10). Furthermore, blocking probability constraints can be met by selecting $c'(\mathbf{x}, \mathbf{a})_j = (1 - a_j^a)(1 - \delta(B(j) - n_j^q))$, and $C = \Psi_j$ in (4.11).

Average delay

As opposed to the blocking probability, the average connection delay cannot be expressed as an average cost criterion. However, it is shown in [Comaniciu and Poor, 2003a] that a combination of cost functions can be used to ensure that the QoS requirements in (4.12) are met for all users, if the requirements are feasible.

The average connection delay can be expressed using a queueing analysis for the equivalent system in Fig. 2.16. The delay expression for a particular class j is given by Little's theorem as a function of the average number of calls in the jth queue N_q^j, the arrival rate λ_j, and the blocking probability for class j, P_b^j:

$$W_j = \frac{N_q^j}{\lambda_j(1 - P_b^j)}.$$

(4.15)

The delay restrictions imposed in (4.12) for a class j of users can be rewritten as

$$N_q^j \leq \Xi_j \lambda_j (1 - P_b^j).$$

(4.16)

Since $P_b^j \leq \Psi_j$ is also required, (4.16) is guaranteed to be met if

$$N_q^j \leq \Xi_j \lambda_j (1 - \Psi_j).$$

(4.17)

Therefore, the network QoS requirements can be reformulated to be

$$\begin{cases} P_b^j \leq \Psi_j, \ j = 1, \ \dots, J \\ N_q^j \leq \Xi_j \lambda_j (1 - \Psi_j), \ j = 1, \ \dots, J \end{cases} \tag{4.18}$$

The average number of calls in the queue can be expressed as an average cost function by selecting the expected cost until the next decision epoch (for class j) to be

$$c(\mathbf{x}, \mathbf{a})_j = n_q^j. \tag{4.19}$$

Hence, to determine an admission policy that satisfies the restrictions in (4.18), constraints on both the blocking probability and on the average number of queued call requests must be imposed. The last ones can be obtained by selecting $c'(\mathbf{x}, \mathbf{a})_j = n_q^j$ and $C = \Xi_j \lambda_j (1 - \Psi_j)$ in (4.11). The optimal admission policy can be determined as described in the following result.

Proposition 4.1. *An optimal admission control policy can be determined as a solution of a constrained linear programming optimization, such that the network QoS requirements in (4.12) can be guaranteed for all users if the system is feasible. The linear program is formulated as follows:*

$$\min_{\mathcal{U}} \sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A_x}} \sum_{j=1}^{J} \theta_j (1 - a_j^a)(1 - \delta(B(j) - n_j^q)) u_{\mathbf{x}, \mathbf{a}} \tau_{\mathbf{x}}(\mathbf{a})$$

subject to the constraints

$$\sum_{\mathbf{a} \in \mathbf{A_y}} u_{\mathbf{y a}} - \sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A_x}} p_{\mathbf{xy}}(\mathbf{a}) u_{\mathbf{x a}} = 0, \ \mathbf{y} \in X,$$

$$\sum_{\mathbf{x} \in X} \sum_{\mathbf{a} \in \mathbf{A_x}} u_{\mathbf{x a}} \tau_{\mathbf{x}}(\mathbf{a}) = 1,$$

$$\sum_{\mathbf{x} \in X^b, \ \mathbf{a} \in \mathbf{A_x^b}} (1 - a_j^a)(1 - \delta(B(j) - n_j^q)) u_{\mathbf{x a}} \tau_{\mathbf{x}}(\mathbf{a}) \leq \Psi_j,$$

$$j = 1, \dots, J,$$

and

$$\sum_{\mathbf{x} \in X^b, \ \mathbf{a} \in \mathbf{A_x^b}} n_q^j u_{\mathbf{x a}} \tau_{\mathbf{x}}(\mathbf{a}) \leq \Xi_j \lambda_j (1 - \Psi_j), \ j = 1, \dots, J. \tag{4.20}$$

The optimal solution obtained by solving (4.20) minimizes both the blocking probabilities (a weighted sum for all classes) as well as the average delays, subject to the network QoS constraints in (4.12).

Proposition 4.1 gives the optimal admission policy that minimizes blocking probabilities and average delays under certain network QoS constraints. However, the admission policy exists only if the system is feasible, that is, only if for the given arrival rate for each class, the network QoS requirements can be met for the given buffer dimension. As we will see also in the numerical results section, not all buffer configurations result in feasible solutions. The buffers' dimensions are thus parameters of the optimization, being closely related to the blocking probability. In case of infeasibility, the linear programming can be reformulated for different buffer configurations. A numerical example will be discussed shortly.

There is also the scenario in which the arrival rate is too high and no buffer configuration can be found to accommodate the network QoS requirements. Theoretically, a maximal arrival rate per class can be defined, which is the arrival rate that can be supported by the network such that QoS requirements are met. This represents the network capacity. However, this quantity is very hard to determine analytically. As a solution, the LP optimization may be solved for increasingly lower arrival rates, until the system becomes feasible. If $\lambda_j^* < \lambda_j$ is the network capacity for class j, determined using this trial and error procedure, the final blocking probability requirements must be relaxed, and the resulting blocking probability for class j is determined as

$$P_{bf}^j = P_b^j p_b^j, \tag{4.21}$$

where $p_b^j = 1 - \frac{\lambda_j^*}{\lambda_j}$.

In other words, in this situation, the admission control policy will be selected as a solution to an LP formulation with a lower arrival rate λ_j^*, such that it will meet the specified QoS requirements for this arrival rate. The above discussion on the design of such an optimal policy applies directly, with the only difference being that further action is needed to reduce the arrival rate, and the final admission control will be implemented in two steps. To reduce the arrival rate from the initial rate λ_j to λ_j^*, a higher level admission control can be implemented as follows:

Higher-Level Admission Control: *Before requesting a new call connection, each user in class j runs a Bernoulli trial experiment with probability of success $p_{adm}^j = 1 - p_b^j$. In case of success, the request for connec-*

tion is made; otherwise the call is automatically rejected. This ensures that the rate of call connection requests is reduced to $\lambda_j^* = p_{adm}^j \lambda_j$, and the Poisson distribution of the call connection requests is preserved.

After the call connection request has been made, the call is admitted or rejected according to the previously discussed optimal admission policy, based on the SMDP formulation.

We now illustrate the performance of the proposed call admission control using simulations for a two-class system having equal high transmission rates corresponding to an equivalent spreading gain $N_e = 8$ ($N = 128$ and $M = 16$ codes; $N_e = N/M$). Parameter values for the experiments are $\lambda_1 = 1$, $\lambda_2 = 0.5$, $\mu_1 = 0.25$, $\mu_2 = 0.1375$ and $\gamma_1 = \gamma_2 = 10$. It is also assumed that the estimated channel gain is $|\bar{h}|^2 = 1$ and the channel estimation variance is $\xi^2 = 0.05$.

The optimal policy for each experiment is obtained using an LP optimization. For each numerical example, a randomized optimal stationary policy $\mathbf{R}^*$ is obtained, which is then used for simulations. The average call connection delays are obtained from simulations, averaged over 10,000 call requests. The blocking probabilities can be obtained both as a result of the LP optimization, as well as from simulations implemented using the obtained optimal policy.

Both blocking probability and delay constraints are imposed: $\boldsymbol{\Psi} = [0.2, 0.1]$ and $\boldsymbol{\Xi} = [2.5, 0.67]$, respectively. The delay constraints translate into constraints on the average number of users in the queues: $\bar{\mathbf{n}} = [2, 0.3]$. Different buffer configurations are considered, some of which are infeasible: $\mathbf{B} = [1, 1]$, $\mathbf{B} = [1, 2]$, $\mathbf{B} = [1, 3]$, $\mathbf{B} = [2, 3]$, and $\mathbf{B} = [3, 3]$. The network performance is summarized in Table 4.1. The first four columns represent the blocking probabilities and average number of queued calls obtained from the LP optimization. The last four columns represent simulation results. We can see that all four buffer configurations result in admission policies for which the imposed QoS requirements are met. From Table 4.1, we see that when delay constraints are imposed, increasing the length of the buffer for class 2 (the most delay sensitive class) lowers P_b^2, but increases P_b^1. This is a consequence of the fact that class 2 users are more delay sensitive, and by increasing their buffer length (and correspondingly decreasing their blocking probability), their service has to be increased as well, so that the delay constraints can be met. The most delay sensitive class (class 2) is an expensive class, since increasing its share of capacity (lower blocking probability obtained using more buffering) affects the performance of all other classes in the system.

Table 4.1. Numerical Results: Admission Control with Delay and Blocking Probability Constraints (reprinted with permission from [Comaniciu and Poor, 2003a])

B	P_b^1	n1	P_b^2	n2	P_b^1 -sim.	P_b^2 -sim.	del. 1	del. 2
[2,1]	0.1865	0.6865	0.1	0.1880	0.1797	0.1040	0.8349	0.4037
[2,2]	0.2	0.8283	0.0598	0.3	0.1881	0.0585	0.9706	0.6177
[3,1]	0.1645	1.1655	0.1	0.1847	0.1656	0.1022	1.3718	0.4027
[3,2]	0.1855	1.3702	0.0533	0.3	0.1907	0.0573	1.6654	0.6383

On a final note, we compare the performance of the optimal admission policy with two other approaches for call admission control: the complete sharing policy, and the threshold policy. Results for the complete sharing policy are presented in Table 4.2.

Table 4.2. Numerical Results for the Complete Sharing Policy (reprinted with permission from [Comaniciu and Poor, 2003a])

B	P_b^1	P_b^2	delay 1	delay 2
[2,1]	0.1109	0.1892	0.4595	0.6017
[2,2]	0.1352	0.1118	0.5414	1.1026
[3,1]	0.0758	0.1857	0.6289	0.6144
[3,2]	0.0782	0.2019	0.6297	0.6385
[3,3]	0.1146	0.0891	0.9051	1.7118
[5,3]	0.0844	0.1059	1.5627	2.0285
[5,1]	0.0395	0.2234	0.9824	0.7031

For the threshold policy, the resources are partitioned between the two classes such that the blocking probability for the most demanding class (class 2) is met. In order to fairly compare the results, we impose the same blocking probability constraints as the ones considered for the optimal policy: [0.2, 0.1].

According to (3.53), the total number of users that can be accepted into the system for the considered numerical values is $K = K_1 + K_2 = 8$. We wish to find K_1 and K_2 such that $P_b^2 \leq 0.1$. For fixed K_2 and $B(2)$, we have an $M/M/K_2/(B(2) + K_2)$ queue, and the blocking probability can be computed as [Bertsekas and Gallager, 1992]

$$P_b^2 = p_0 \frac{K_2^{K_2} \rho_2^{(B(2)+K_2)}}{K_2!}, \tag{4.22}$$

where $\rho_2 = \frac{\lambda_2}{K_2 \mu_2}$, and p_0 is the probability of an empty queue and no one in service:

$$p_0 = \left[1 + \sum_{k=1}^{K_2-1} \frac{(K_2\rho_2)^k}{k!} + \sum_{k=K_2}^{K_2+B(2)} \frac{(K_2\rho_2)^k}{K_2!} \frac{1}{K_2^{(k-K_2)}} \right]^{-1}. \qquad (4.23)$$

The average connection delay experienced by calls in class 2 can be expressed as

$$W_2 = \frac{\rho_2 P_Q}{\lambda_2(1-\rho_2)}, \qquad (4.24)$$

where P_Q is the probability of queueing, defined as

$$P_Q = p_0 \frac{(K_2\rho)^{K_2}}{K_2!} \sum_{k=K_2}^{K_2+B(2)} \rho_2^{k-K_2}. \qquad (4.25)$$

We note that both the blocking probability and the delay depend on K_2 and $B(2)$; therefore they cannot be optimized independently. If we fix K_2 and $B(2)$ for a given constraint for the blocking probability, the delay is also automatically fixed to the value computed from (4.24).

Table 4.3. Numerical Results for the Threshold Policy (reprinted with permission from [Comaniciu and Poor, 2003a])

B	P_b^1	P_b^2	delay 1	delay 2
[2,1]	0.3325	0.1132	1.3461	0.2340
[2,2]	0.3404	0.0654	1.4325	0.4983
[3,1]	0.3053	0.1154	2.2313	0.2342
[3,2]	0.3040	0.0742	2.1936	0.5359
[3,3]	0.2999	0.0505	2.1757	0.8041
[5,3]	0.2748	0.0540	4.1355	0.8135
[5,1]	0.2722	0.1046	4.0970	0.2282

We represent P_b^2 in Fig. 4.3 as a function of K_2 for three different values of the buffer length. It can be seen that the partition $[K_1, K_2] = [3, 5]$ gives $P_b^2 \approx 0.1$, depending on the designed buffer length. Therefore, this partition is used to obtain simulation results, which are summarized in Table 4.3. We note that the QoS constraints: $\boldsymbol{\Psi} = [0.2, 0.1]$ and $\boldsymbol{\Xi} = [2.5, 0.67]$ cannot be met. Lower blocking probabilities can be obtained if the buffer lengths are increased for both classes, but this comes at the expense of increased delay. We observe that the threshold policy is clearly suboptimal and lacks flexibility in guaranteeing the desired QoS.
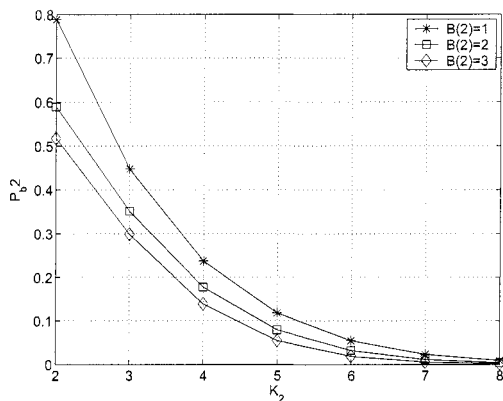
Figure 4.3. Threshold policy: blocking probability for class 2 (reprinted with permission from [Comaniciu and Poor, 2003a])

2. Ad Hoc Networks

Admission control for ad hoc networks usually imposes some hierarchical structure, such that the network is divided into clusters and a cluster-head node is in charge of admission control. Alternatively, a distributed approach based on only local information may be more suitable.

A simple admission control scheme can be designed based on capacity results discussed in Chapter 3. Once we know the maximal number of nodes that can be supported by the network for given QoS specifications (SIR and delay), a complete sharing policy or a threshold policy can be readily determined (recall that for ad hoc networks, the delay is caused by the multi-hop routing). Once admitted, a user is guaranteed to meet its QoS requirements, irrespective of mobility, when a shortest hop routing protocol is used, and for fixed power transmission.

If power control is also implemented, the admission control can be determined as a power control feasibility condition, following the work in [Comaniciu and Poor, 2004b]. If the system is too heavily loaded, the powers will begin to increase without bound (the power control algorithm does not converge), and the new user needs to be dropped out of the network. For this scenario, admission control is effectively integrated with power control and routing.

Integrated admission control, power control and routing is also proposed in [Sankaran and Ephremides, 2002] for multicasting in ad hoc networks, and will be discussed in more detail in the following. In [Sankaran and Ephremides, 2002], circuit switched multicasting for a synchronous, fixed CDMA ad hoc network is analyzed. In this multi-

casting scheme, nodes perform session admission control by specifying reception power bounds for signals. The admission control involves deciding whether or not a node can take part in a new multicast session. If an arbitrary node k decides to take part in the new session, it must specify a minimal ($p_{(k,min)}$) and a maximal ($p_{(k,max)}$) power value that can be received by the node k for that session. The minimal power value ensures that target SIRs are met at the receiving node, whereas the maximal reception power value ensures that the signal for the new session does not cause too much interference to the ongoing multicast sessions. The network connectivity graph for the new multicast session is a function of the minimal and maximal power levels set by the nodes. To set up a new session, three steps are necessary:

1. admitting the new session;

2. building the multicast tree; and

3. implementing power control, such that the power will be in the specified range.

Admission control

The admission control proposed in [Sankaran and Ephremides, 2002] is based only on localized information, i.e., nodes have access only to information from their neighbors. Each node k decides if it can take part of a new session or not, based on its residual capacity. The maximal capacity for node k is determined using (3.14), (3.16), or (3.19) (for the matched filter, decorrelator or LMMSE receiver, respectively), under the assumption that all the processed sessions can be received with the same power $p_{(k,min)}$. In reality, in a multicast session, signals processed at node k usually cannot all be received with the same power, since the transmitting nodes may have multiple destination nodes with links characterized by different link gain coefficients. To overcome this problem, two different versions for the admission control are proposed in [Sankaran and Ephremides, 2002].

Admission control 1

For the first version, the value for $p_{(k,min)}$ is fixed. Choosing its value affects the node's capacity: a higher value means greater capacity, but it is restricted by the capability of all its source nodes to be received with at least that amount of power. One possible approach is to choose the smallest maximal power value that can be received from any of the node's neighbors. Note that it is not necessarily required for the transmitting node to be received at node k with $p_{(k,min)}$; this is necessary only if the

receiving node operates at full capacity. A transmitting node informs its neighbors about its maximal transmission power that may be required for each outgoing session, $P_{(\ell,max)}^s$, which is determined as the maximal power required such that it will be received with at least power $p_{(k,min)}$ at all receiving nodes (k is an arbitrary receiving node index). Since a transmitter may actually be received with a higher power than the minimum required, node k must also select a maximal allowed received power $p_{(k,max)}$. A simple calculation of $p_{(k,max)}$ can be done for the simplest case in which matched filter receivers are used. In this case, the interference depends only on the total interference power and does not depend on the actual number of interferers. If we denote as p_{sum}^k the total power received at node k, including the interference signals, then the node sets its maximal received power to be equal to

$$p_{(k,max)} = \tilde{K}_k p_{(k,min)} - p_{sum}^k, \tag{4.26}$$

where $\tilde{K}_j$ is the approximated user capacity for node j (exact when all users are received with the same power).

A disadvantage of this algorithm is that all nodes assume that every transmitter ℓ transmits with maximal power, and is thus received by an arbitrary node k with at least $p_{(k,min)}$, even though this is not the case for most of the time. This leads to inefficient network utilization and consequently to higher blocking probabilities for new sessions. One way to overcome this problem is to increase the value $p_{(k,min)}$ gradually as the load increases. This leads to the second version of the admission control.

Admission control 2

For this version, a node decides to participate in new sessions by assuming that current transmissions are going to increase their powers by a factor of δ, when the node joins the new session. The factor δ, greatly influences the performance of the admission control and therefore must be chosen carefully. The admission control decision requires several steps:

- If there exists a session which, after the power increase, will violate the maximum received power constraint, node k cannot participate in this new session and sets $p_{(k,max)} = 0$. It is assumed that power increases occur only as a consequence of new session additions.

  Otherwise,

- Node k recomputes $p_{(k,min)}$ and $p_{(k,max)}$ for the scaled powers, and also recomputes its capacity $\tilde{K}_k$. We note that the capacity is com-

puted according to (3.14), (3.16), or (3.19), but considering an enhanced noise power $\sigma_k'^2$, instead of σ_k^2. The enhanced noise power accounts for the interference power at node k (from signals that are not currently processed by the node; see also the discussion on hybrid partial systems in Chapter 3).

- For LMMSE receivers, since the interference function is not additive, node k must determine also the maximal number of neighbors that can transmit for the new session[1]: $N_{k,max} = \lfloor p_{(k,max)}/p_{(k,min)} \rfloor$.

  We note that, for the decorrelator, the SIR performance at one node will depend only on the current SNR, i.e., no constraint needs to be imposed on the maximal number of neighbors that can transmit for the new session, but the enhanced noise plus interference power $\sigma_k'^2$ must be updated for the scaled powers. Based on the updated noise level, a new capacity value $\tilde{K}_k$ is then derived.

A disadvantage of this algorithm is that, even if a node decides to deny a session, due to the scaling algorithm for the powers, an ongoing session for that node can still cause a transmission with a higher power than the maximal allowable power. When this happens, all the nodes that process signals from the node that violated the power constraint do not admit any new session until the transmitter can reduce its power to acceptable levels.

Multicast routing

Once the nodes decide that they can participate in a new session, they will form a set of potential downstream neighbors for that session. A node ℓ is a potential downstream neighbor of k if node k can satisfy the power range requirements for node ℓ. In addition, for LMMSE receivers, the maximum transmitting neighbors condition must also be satisfied. The network connectivity graph for the multicast tree depends on the set of potential downstream neighbors for that particular session. Each multicast tree is characterized by a certain distribution of powers for the nodes, which influences the current interference level and the total energy consumption. We can define the cost of a link (k, ℓ) as the transmission power $P_{k\ell}$ required by node k to be correctly received by node ℓ. In general, a node will have many outgoing links (see an example in Fig. 4.4), and thus the node cost will be the maximal cost over all the outgoing links.

---

[1]It is proved in [Sankaran and Ephremides, 2002] that the SIR constraint will still hold if fewer than the maximal allowable number of neighbor nodes transmit, provided that the total received power satisfies the maximal power constraint.
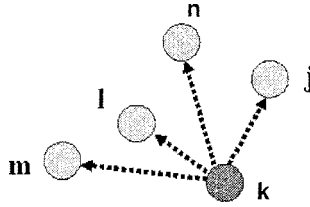
Figure 4.4. Multiple outgoing links in a multicast tree

For the example considered in Fig. 4.4, the cost of node k is $cost_k = \max\{P_{kj}, P_{k\ell}, P_{km}, P_{kn}\}$. The multicast cost of a tree is the sum of costs for all nodes in the tree. While, the goal is to minimize this cost, additional constraints must be met, such as maximal power values specified by each node, as well as a maximal number of transmitting neighbors for the case in which LMMSE receivers are used. The problem of finding an optimal multicast tree is NP-hard, and a suboptimal solution is proposed in [Sankaran and Ephremides, 2002]: the pruned extended tree algorithm (PET). In this algorithm, an extended tree rooted at the source is constructed, and then the multicast tree is obtained by removing nondestination leaf nodes from the tree (pruning). The term "extended tree" means that not all the leaves in the obtained tree will be multicast destinations. Starting with the source, nodes are added to the tree one at a time, until all the multicast destinations have been added. The instantaneous cost of the tree is the cost of the current tree. The rule for addition of new nodes is to add a link that gives the minimal cost addition to the current multicast tree cost. The addition is made only if the above mentioned additional restrictions hold.

After adding a new link, backsweeping may be necessary, i.e, a node parent in the tree is changed if this results in a reduction of the overall cost for the tree. Backsweeping must be done carefully to avoid loops in the tree.

Power control

The last step in the process is to ensure that the sessions will really get the desired QoS, and this implies that power control must be implemented. Similarly to the problem in [Comaniciu and Poor, 2003c], a node may need to satisfy SIR constraints for multiple outgoing links, so that the SIR condition translates to the requirement that the weakest

outgoing link must meet the target SIR. As we discussed in Chapter 2, this power control problem is standard and can be implemented as an iterative power control algorithm algorithm:

$$P(n + 1) = T(\mathbf{p}(n)), \tag{4.27}$$

where $T(\mathbf{p}) = \max_{(k,\ell)} I_{(k,\ell)}(\mathbf{p})$ is a standard interference function, and (k, ℓ) is an arbitrary link. The expressions for $I_{(k,\ell)}$ can be directly determined by imposing the condition that the achievable link SIR must be greater than or equal to a target SIR γ^*. Consequently, for the matched filter we have

$$I_{(k,\ell)}^{MF} = \frac{\gamma^* \left(\frac{1}{N} \sum_{j \neq k} h_{j,\ell} P_j + \sigma^2 \right)}{h_{k,\ell}}; \tag{4.28}$$

and for the LMMSE receiver we have

$$I_{(k,\ell)}^{mmse} = \frac{\gamma^* \left(\frac{1}{N} \sum_{j \neq k} \frac{h_{j,\ell} P_j h_{k,\ell} P_k}{h_{k,\ell} P_k + h_{j,\ell} P_j \gamma^*} + \sigma^2 \right)}{h_{k,\ell}}. \tag{4.29}$$

Here $h_{k,\ell}$ is the link gain from node k to node ℓ.

Performance results

The performance of the proposed admission control (version 2) and the multicast algorithm were illustrated by simulations in [Sankaran and Ephremides, 2002]. We present here an example for 10 simulations, 50 nodes and 100 multicast sessions per simulation. Parameter values used for the experiments are: $p_{min} = \{0.1, 1\}$, maximal transmission power 5, $\delta = 1.3$, $\gamma* = 3$, $N = 511$, and $\sigma^2 = 0.2$. Nodes are placed randomly in a square grid of dimensions (10×10). The arrival process for the multicast sessions is Poisson with mean λ, and the session duration is exponential with mean 1. Three different performance metrics are considered:

- Multicast efficiency:

$$Eff = \frac{1}{X} \sum_{i=1}^{X} \frac{m_i}{n_i}, \tag{4.30}$$

   where X is the number of multicast requests, m_i is the number of destinations reached for the ith session, and n_i is the number of intended destinations for the ith session.

- Blocking probability

$$P_b = \frac{b_X}{X}, \tag{4.31}$$

where b_X is the number of completely blocked multicast requests $(m_i = 0)$.

- Average power consumption

$$P_{av} = \frac{1}{X'} \sum_{i=1}^{X'} \frac{e_i}{T_i t_i}, \qquad (4.32)$$

where X' is the number of multicast request for which at least one destination is reached, e_i is the sum of energies for all nodes participating in session i, T_i is the number of nodes transmitting in session i, and t_i is the duration for session i.

The simulation results are illustrated in Figs. 4.5, 4.6, and 4.7.
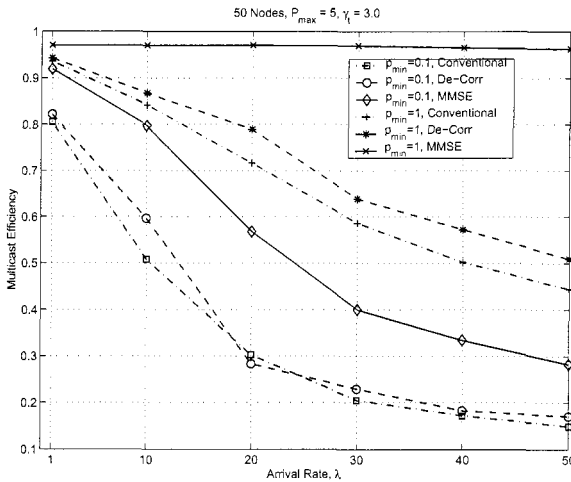
Figure 4.5. Multicast efficiency (reprinted with permission from [Sankaran and Ephremides, 2002])

Performance comparisons among the three receivers reveal that, while the LMMSE receiver performs the best, the performance of the decorrelator is comparable with that of the matched filter.
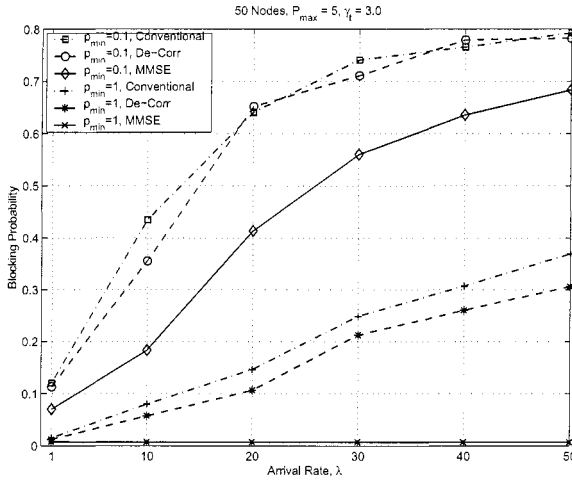
Figure 4.6. Blocking probability(reprinted with permission from [Sankaran and Ephremides, 2002])
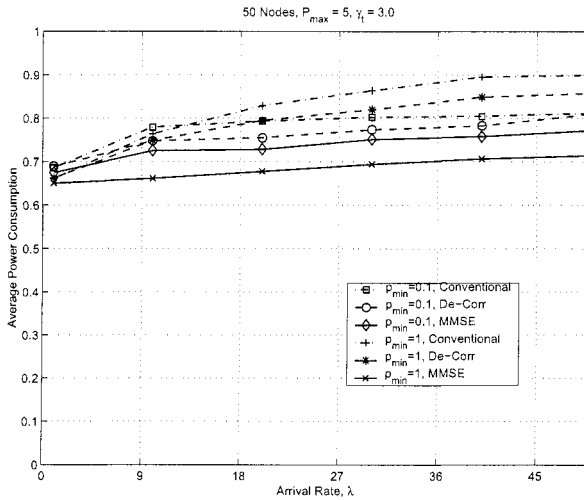
Figure 4.7. Average power consumption (reprinted with permission from [Sankaran and Ephremides, 2002])

Chapter 5

MULTIUSER DETECTION IN CROSS-LAYER DESIGN: PERSPECTIVES

Cross-layer design and multiuser detection are both controversial topics. For multiuser detection, an extensive literature is dedicated to showing that significant gains in spectral efficiency, user capacity and near/far resistance can be achieved. On the other hand, skeptics argue that the complexity cost of multiuser detection makes it impractical for commercial implementation. Despite this, the least complex of the multiuser receiver family (the interference cancellation receivers) are begining to be implemented in third generation cellular systems. Further, there is considerable potential for multiuser receivers to be used in emerging types of networks such as ad hoc networks, which are particularly susceptible to interference and to near/far problems. One of the hurdles to be overcome in implementing multiuser detectors in ad hoc networks is the requirement of continuous adaptation of filter coefficients as a result of changes in the user population, and consequently, the need to continuously learn the signature sequences of neighboring transmitters. We have seen that one possible solution to this problem is to blindly learn the interference subspace and construct the filter coefficients accordingly. For rapidly moving users in an ad hoc network (when the users move quickly in and out of a neighborhood), multiuser detection is not a suitable solution. This type of network is probably most suitable for random access and short bursty transmissions, for which implementing a multiuser receiver is very difficult. This, and related considerations, suggest that the suitability of multiuser detection is application and network dependent. However, as DSP and CMOS technologies continue to evolve, the class of networks for which multiuser detection is suitable will grow.

Another obstacle to the deployment of multiuser receivers has been the fact that the physical layer performance was hard to quantify in

a network context when interacting with upper layer protocols, such as power control, access control and routing. We have seen that some progress has been made in characterizing and optimizing such systems, although there is still considerable room for further research in this area.

Understanding and exploiting the interactions between different layers of the protocol stack is the core of the cross-layer design concept. Several questions about cross-layer design must be answered before these interactions can be successfully exploited. First of all, does cross-layer design mean that we must completely discard the OSI layered model? Do we still need a network architecture? Is cross-layer design suitable for all types of networks and all types of applications? Are the gains obtained from cross-layer design of a short-term nature, or of a long term nature?

We address the above questions by analyzing simple facts related to wireless networks. First of all, wireless networks do not come with fixed links as their wireline counterparts do. This means that the upper layer protocols must rely on a network model that is inherently physical layer dependent. Aside from channel impairments, the reliability of the links that form this network model depends on the level of interference in the system, which in turn may be influenced by upper layer protocols. We have already seen examples in this book that discuss the interdependence among the receiver design, routing, power control, access control and admission control.

On the other hand, the layered architecture has multiple advantages: it has a modular design, it is easily upgradable and suitable for standardization and mass production, which immediately translates into long term gains for this architecture. The success story of the Internet is a very good example supporting the efficiency and suitability of the OSI model. Layered, modular, architecture models have certain advantages that are equally important for wireless networks. A common missconception in cross-layer design is that the layered approach must be completely eliminated, and all layers must be integrated and jointly optimized. While this approach might lead to some overall performance gains in the short term, it is clearly impractical and cannot be considered as a solution for future generation network design. At the other extreme, the isolated design for layers as commonly used in wireline networks might be applied for wireless networks as well. This is also a undesirable since it ignores the interactions between layers and might lead to severe penalties in performance. The solution for cross-layer design should rather be based on a holistic view of wireless networking, which maintains the layered approach, while accounting for the interactions between various protocols at different layers. QoS support for various applications should be implemented at all layers in the protocol

stack and the response to changes in the channel environment should be hierarchically implemented across all layers. Interactions between different protocols can be accounted for by exchanging pertinent information between layers. Abstract models for layers may also greatly simplify the design. We have seen an example in Chapter 3 of how the physical layer performance can be abstracted for upper layers: the capacity of power controlled networks was derived for asymptotically large networks. If such an abstraction is possible, it is much easier to determine what information should be exchanged between layers and how it should be used by the adaptation protocols. In general, the inter-layer coupling is difficult to characterize and this is one of the current key research problems in cross-layer design. Another important issue is that of maintaining the right balance between performance, complexity and scalability for wireless network optimizations.

Note that some inter-layer coupling may occur in all kinds of networks (even wireline), but it is especially strong for wireless architectures, due to the nature of wireless transmission. As a consequence, cross-layer design might be beneficial for all wireless architectures and all types of applications. However, caution should be exercised when implementing cross-layer optimization, since one would expect that the optimal trade-offs among performance, complexity and scalability would be application and network dependent.

As a final remark, in this book we have shown that multiuser detection in cross-layer design has the potential to significantly improve the performance of wireless networks. This topic is fairly new and it opens many new and exciting research problems for all types of wireless networks supporting heterogeneous applications.

List of Acronyms

$H - MMSE^{(p)}$ = Partial Hybrid LMMSE, 63

3G = Third Generation, 1

ARQ = Automatic Repeat Request, 82

AWGN = Additive White Gaussian Noise, 17

BADD = Blind Adaptive Decorrelating Detector, 30

BER = Bit Error Rate, 17

BPSK = Binary Phase Shift Keying, 39

BS = Base Station, 68

CDMA = Code Division Multiple Access, 2

cdma2000 = North American 3G CDMA based standard, 2

CMOS = Complementary Metal-Oxide Semiconductor, 39

CSMA = Carrier Sense Multiple Access, 81

DECT = Digital Enhanced Cordless Telecommunications, 7

DMI = Direct Matrix Inversion, 35

DN = Destination Node, 144

DS-CDMA = Direct Sequence Code Division Multiple Access, 2

DSP = Digital Signal Processing, 5

MAP = Maximum *A Posteriori* Probability, 19

MATPR = Minimum Average Transmission Power Routing, 82

MC = Multicode, 40

MF = Matched Filter, 17

MI = Maximal Invariant, 74

ML = Maximum Likelihood, 19

MME = Maximum Mean Energy, 35

MMSE = Minimum Mean-Square Error, 35

MOE = Minimum Output Energy, 30

MPR = Multipacket Reception, 76

MQSR = Multiqueue Service Room, 78

MSE = Mean-Square Error, 30

MUD = Multiuser Detection, 17

OFDM = Orthogonal Frequency Division Multiplexing, 6

OSI = Open Systems Interconnect, 13

PASTd = Projection Approximation Subspace Tracking with Deflation, 31

PDA = Personal Digital Assistant, 7

PET = Pruned Extended Tree Algorithm, 171

PHY = Physical Layer, xiv

PIC = Parallel Interference Cancellation, 25

PSTN = Public Switched Telephone Network, 7

QAM = Quadrature Amplitude Modulation, 40

QoS = Quality of Service, 2

RCT = Receiver Controlled Transmission, 80

RLS = Recursive Least Squares, 31

References

[Aazhang et al., 1998] Aazhang, B., Jin, L., and Nosratinia, A. (1998). Progressive source-channel coding of images over bursty error channels. In *Proceedings of the IEEE International Conference on Image Processing*, volume 2, pages 127–131, Chicago, IL.

[Aein, 1973] Aein, J. M. (1973). Power balancing in system employing frequency reuse. *COMSAT Technical Review*, 3(2):277–300.

[Agrawal et al., 2004] Agrawal, A., Andrews, J., Cioffi, J., and Meng, T. (2004). Iterative power control for successive interference cancellation. *IEEE Transactions on Wireless Communications,* to appear.

[Akyildiz et al., 2002] Akyildiz, I. F., Cayirci, E., Sankarasubramaniam, Y., and Weilian, S. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114.

[Almutairi et al., 2000] Almutairi, A. F., Miller, S. L., Latchman, H., and Wong, T. F. (2000). Power control algorithm for MMSE receiver based CDMA systems. *IEEE Communications Letters*, 4(11):346–348.

[Alonso and Agusti, 2004] Alonso, L. and Agusti, R. (2004). Automatic rate adaptation and energy-saving mechanisms based on cross-layer information for packet-switched data networks. *IEEE Communications Magazine*, 42(3):S15–S20.

[Andrews and Meng, 2003] Andrews, J. and Meng, T. H. (2003). Optimum power control for successive interference cancellation with imperfect channel estimation. *IEEE Transactions on Wireless Communications*, 2(2):375–383.

[Bansal and Liu, 2003] Bansal, N. and Liu, Z. (2003). Capacity, delay and mobility in wireless ad hoc networks. In *Proceedings of the IEEE INFOCOM*, volume 2, pages 1553–1563, San Francisco, CA.

[Berggren and Slimane, 2002] Berggren, F. and Slimane, S. B. (2002). Power allocation for a simple successive interference cancellation scheme in a multi-rate DS-CDMA system. In *Proceedings of the IEEE International Conference on Communications*, pages 351–355, New York, NY.

183

[Bertocchi et al., 2003] Bertocchi, F., Bergamo, P., Mazzini, G., and Zorzi, M. (2003). MAC and routing solution for energy saving in ad hoc networks: Distributed power control. In *Proceedings of the IEEE Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia*, volume 2, pages 1061–1065, Singapore.

[Bertsekas and Gallager, 1992] Bertsekas, D. and Gallager, R. (1992). *Data Networks*. Prentice Hall, Upper Saddle River, NJ.

[Bertsekas, 1995] Bertsekas, D. P. (1995). *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA.

[Bhagwat, 2001] Bhagwat, P. (2001). Bluetooth: Technology for short-range wireless apps. *IEEE Internet Computing*, 5(3):96–103.

[Biglieri et al., 2000] Biglieri, E., Caire, G., and Taricco, G. (2000). CDMA system design through asymptotic analysis. *IEEE Transactions on Communications*, 48(11):1882–1896.

[Brewin, 2004] Brewin, B. (2004). Wi-Fi and cellular technology move closer to convergence. ComputerWeekly.com.

[Buehrer et al., 1996] Buehrer, R. M., Correal, N. S., and Woerner, B. D. (1996). A comparison of multiuser receivers for cellular CDMA. In *Proceedings of the IEEE Global Telecommunications Conference*, pages 1571–1577, London, UK.

[Cai et al., 2002] Cai, Z., Lu, M., and Wang, X. (2002). Minimum average transmission power routing in CDMA ad hoc networks utilizing the blind multiuser detection. In *Proceedings of the IEEE International Parallel and Distributed Processing Symposium*, pages 428–433, Fort Lauderdale, FL.

[Chen and Tong, 2001] Chen, B. and Tong, L. (2001). Traffic-aided multiuser detection for random-access CDMA networks. *IEEE Transactions on Signal Processing*, 49(7):1570–1580.

[Chih-Lin and Gitlin, 1995] Chih-Lin, I. and Gitlin, R. D. (1995). Multicode CDMA wireless personal communications networks. In *Proceedings of the IEEE International Conference on Communications*, volume 2, pages 1060–1064, Seattle, WA.

[Chuah et al., 1995] Chuah, C. N., Yates, R., and Goodman, D. (1995). Integrated dynamic radio resource management. In *Proceedings of the IEEE 45th Vehicular Technology Conference*, pages 584–588, Chicago, IL.

[Comaniciu, 2002] Comaniciu, C. (2002). *Integrated Access Control and Detection for QoS in Multimedia CDMA Networks*. PhD thesis, Rutgers University, New Brunswick, NJ.

[Comaniciu and Mandayam, 2000] Comaniciu, C. and Mandayam, N. (2000). Delta modulation based prediction for access control in integrated voice/data CDMA systems. *IEEE Journal on Selected Areas in Communication*, 18(1):112–122.

[Comaniciu and Mandayam, 2002] Comaniciu, C. and Mandayam, N. B. (2002). Integrated access control and multiuser detection for multimedia CDMA systems.

In *Proceedings of the Conference on Information Sciences and Systems*, Princeton University, Princeton, NJ.

[Comaniciu and Poor, 2003a] Comaniciu, C. and Poor, H. V. (2003a). Jointly optimal power and admission control for delay sensitive traffic in CDMA networks with LMMSE receivers. *IEEE Transactions on Signal Processing, Special Issue on Signal Processing in Networking*, 51(8):2031 – 2042.

[Comaniciu and Poor, 2003b] Comaniciu, C. and Poor, H. V. (2003b). Multirate groupwise MMSE in multipath fading channels: Optimal power control and asymptotic capacity. In *Proceedings of the Conference on Information Sciences and Systems*, The Johns Hopkins University, Baltimore, MA.

[Comaniciu and Poor, 2003c] Comaniciu, C. and Poor, H. V. (2003c). QoS provisioning for wireless ad hoc data networks. In *Proceedings of the IEEE Conference on Decision and Control*, volume 1, pages 92–97, Maui, HI.

[Comaniciu and Poor, 2004a] Comaniciu, C. and Poor, H. V. (2004a). Capacity regions and optimal power allocation for groupwise multiuser detection. *IEEE Transactions on Wireless Communications*, to appear.

[Comaniciu and Poor, 2004b] Comaniciu, C. and Poor, H. V. (2004b). Joint power control and routing for wireless ad hoc networks. Princeton University, Princeton, NJ. Preprint.

[Comaniciu and Poor, 2004c] Comaniciu, C. and Poor, H. V. (2004c). On the capacity of mobile ad hoc networks with delay constraints. *IEEE Transactions on Wireless Communications*, to appear.

[Cover and Thomas, 1991] Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory*. Wiley-Interscience, New York, NY.

[Crow et al., 1997] Crow, B. P., Kim, L. G., Sakai, P. T., and Widjaja, I. (1997). IEEE 802.11 wireless local area networks. *IEEE Communications Magazine*, 35(9):116–126.

[Cruz and Santhanam, 2003] Cruz, R. L. and Santhanam, A. (2003). Optimal routing, link scheduling, and power control in multi-hop wireless networks. In *Proceedings of the IEEE INFOCOM*, pages 702–711, San Francisco, CA.

[Divsalar et al., 1998] Divsalar, D., Simon, M. K., and Raphaeli, D. (1998). Improved parallel interference cancellation for CDMA. *IEEE Transactions on Communications*, 46(2):258–268.

[Domazetovic et al., 2002] Domazetovic, A., Greenstein, L. J., Mandayam, N. B., and Seskar, I. (2002). A new modeling approach for wireless channels with predictable path geometries. In *Proceedings of the IEEE 56th Vehicular Technology Conference*, volume 1, pages 454–458, Vancouver, BC.

[ElBatt and Ephremides, 2004] ElBatt, T. and Ephremides, A. (2004). Joint scheduling and power control for wireless ad hoc networks. *IEEE Transactions on Wireless Communications*, 3:74–85.

[Ephremides et al., 2000] Ephremides, A., Ihok, T., Pickholtz, R., Iskander, M., Katehi, L., Rao, R., Stark, W., and Winters, J. (2000). *Wireless Technologies and Information Networks*. Technical report, International Technology Research Institute, World Technology Division, Baltimore, MD.

[Ephremides and Sankaran, 1998] Ephremides, A. and Sankaran, C. (1998). Solving a class of optimum multiuser detection problems with polynomial complexity. *IEEE Transactions on Information Theory*, 44(5):1958–1961.

[Evans and Tse, 2000] Evans, J. S. and Tse, D. (2000). Large system performance of linear multiuser receivers in multipath fading channels. *IEEE Transactions on Information Theory*, 46(6):2059–2078.

[Foschini and Miljanic, 1993] Foschini, G. and Miljanic, Z. (1993). A simple distributed autonomous power control algorithm and its convergence. *IEEE Transactions on Vehicular Technology*, 42(4):641–646.

[Frenkiel et al., 2000] Frenkiel, R., Badrinath, B., Borras, J., and Yates, R. (2000). The infostations challenge: Balancing cost and ubiquity in delivering wireless data. *IEEE Personal Communications*, 7(2):66–71.

[Frenkiel and Imielinski, 1996] Frenkiel, R. and Imielinski, T. (1996). *Infostations: The Joy of Many-time, Many-where Communications*. Technical report, WINLAB, Rutgers University, Piscataway, NJ.

[Ge and Ma, 1998] Ge, H. and Ma, J. (1998). Multi-rate LMMSE detectors for asynchronous multi-rate CDMA systems. In *Proceedings of the IEEE International Conference on Communications*, pages 714–718, Atlanta, GA.

[Ghez et al., 1988] Ghez, S., Verdú, S., and Schwartz, S. C. (1988). Stability properties of slotted Aloha with multipacket reception capability. *IEEE Transactions on Automatic Control*, 33(7):640–649.

[Ghez et al., 1989] Ghez, S., Verdú, S., and Schwartz, S. C. (1989). Optimal decentralized control in the random access multipacket channel. *IEEE Transactions on Automatic Control*, 34(11):1153–1163.

[Goldsmith and Varaiya, 1997] Goldsmith, A. and Varaiya, P. (1997). Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, 43(6):1986–1992.

[Goldsmith and Wicker, 2002] Goldsmith, A. J. and Wicker, S. B. (2002). Design challenges for energy-constrained ad hoc wireless networks. *IEEE Wireless Communications*, 9(4):8–27.

[Goodman, 2000] Goodman, D. J. (2000). The wireless Internet: Promises and challenges. *IEEE Computer*, 33(7):36–41.

[Goodman and Mandayam, 2000] Goodman, D. J. and Mandayam, N. B. (2000). Power control for wireless data. *IEEE Personal Communications*, 7(2):1–6.

[Grandhi et al., 1993] Grandhi, S., Vijayan, R., Goodman, D. J., and Zander, J. (1993). Centralized power control for cellular radio systems. *IEEE Transactions on Vehicular Technology*, 42(4):466–468.

[Grossglauser and Tse, 2002] Grossglauser, M. and Tse, D. (2002). Mobility increases the capacity of ad-hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10(4):477–486.

[Guo, 1999] Guo, Y. (1999). *Resource Allocation in Wireless CDMA Multimedia Networks*. PhD thesis, Rice University, Houston, TX.

[Guo and Aazhang, 1999] Guo, Y. and Aazhang, B. (1999). Capacity of multi-class traffic CDMA system with multiuser receiver. In *Proceedings of the IEEE Wireless Communications and Networking Conference*, pages 500–504, New Orleans, LA.

[Gupta and Das, 2001] Gupta, N. and Das, S. R. (2001). A capacity and utilization study of mobile ad hoc networks. In *Proceedings of the 26th Annual IEEE Conference on Local Computer Networks*, pages 576–583, Tampa, FL.

[Gupta and Kumar, 2003] Gupta, P. and Kumar, P. (2003). Towards an information theory of large networks: an achievable rate region. *IEEE Transactions on Information Theory*, 49(8):1877–1894.

[Gupta and Kumar, 2000] Gupta, P. and Kumar, P. R. (2000). The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404.

[Hanly, 1995] Hanly, S. V. (1995). An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity. *IEEE Journal on Selected Areas in Communications*, 13(7):1332–1340.

[Hanly and Tse, 1999] Hanly, S. V. and Tse, D. N. (1999). Power control and capacity of spread spectrum wireless networks. *Automatica*, 35(12):1987–2012.

[Harris et al., 2001] Harris, J. M., Pettinger, J., and Vukovic, I. N. (2001). Modeling RF impact on TCP/RLP in IS-95B. In *Proceedings of the IEEE 53rd Vehicular Technology Conference*, pages 2142–2146, Rhodes, Greece.

[Holma and Laakso, 1999] Holma, H. and Laakso, J. (1999). Uplink admission control and soft capacity with MUD in CDMA. In *Proceedings of the IEEE 49th Vehicular Technology Conference*, volume 1, pages 431–435, Houston, TX.

[Holma and Toskala, 2002] Holma, H. and Toskala, A., editors (2002). *WCDMA for UMTS*. John Wiley & Sons, 2nd edition, New York, NY.

[Hong and Rappaport, 1986] Hong, D. and Rappaport, S. (1986). Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. *IEEE Transactions on Vehicular Technology*, 35(3):77–92.

[Honig et al., 1995] Honig, M., Madhow, U., and Verdú, S. (1995). Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41(4):944–960.

[Honig and Tsatsanis, 2000] Honig, M. and Tsatsanis, M. K. (2000). Adaptive techniques for multiuser CDMA receivers. *IEEE Signal Processing Magazine*, 17(3):49–61.

[Honig and Veerakachen, 1996] Honig, M. L. and Veerakachen, W. (1996). Performance variability of linear multiuser detection for DS-CDMA. In *Proceedings of*

the IEEE 46th Vehicular Technology Conference, volume 1, pages 372–376, Atlanta, GA.

[Høst-Madsen and Wang, 2002] Høst-Madsen, A. and Wang, X. (2002). Performance of blind and group-blind multiuser detectors. *IEEE Transactions on Information Theory*, 48(7):1849–1872.

[Hui, 1984] Hui, J. (1984). Throughput analysis for code divison multiple accessing of the spread spectrum channel. *IEEE Journal on Selected Areas in Communication*, 2(4):482–486.

[IEEE 802.16 Working Group, 2004] IEEE 802.16 Working Group (2004). grouper.ieee.org/groups/802/16.

[Infrared Data Association, 2004] Infrared Data Association (2004). www.irda.org.

[Jabbari et al., 2002a] Jabbari, B., Jean, C. A. S., and Zadeh, A. N. (2002a). Combined routing, channel scheduling, and power control in packet radio ad hoc networks with cellular overlay. In *Proceedings of the IEEE 55th Vehicular Technology Conference*, volume 4, pages 1960–1964, Birmingham, AL.

[Jabbari et al., 2002b] Jabbari, B., Pickholtz, R., Vojcic, B., and Zadeh, A. N. (2002b). Self-organizing packet radio ad hoc networks with overlay. *IEEE Communications Magazine*, 40(6):149–157.

[Jabbari and Zadeh, 2001] Jabbari, B. and Zadeh, A. N. (2001). The capacity modeling of multi-hop cellular packet CDMA networks. In *Proceedings of the IEEE Military Communications Conference*, volume 1, pages 600–604, Washington, DC.

[Jung and Vaidya, 2002] Jung, E.-S. and Vaidya, N. H. (2002). A power control MAC protocol for ad hoc networks. In *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking*, pages 36–47, Atlanta, GA.

[Juntti, 1995] Juntti, M. (1995). Linear multiuser detector update in synchronous dynamic CDMA systems. In *Proceedings of the 6th International Symposium on Personal Indoor and Mobile Radio Communications*, pages 980–984, Toronto, ON.

[Juntti et al., 1998] Juntti, M., Aazhang, B., and Lilleberg, J. (1998). Iterative implementation of linear multiuser detection for dynamic asynchronous CDMA systems,. *IEEE Transactions on Communications*, 46(4):503–508.

[Juntti, 1998a] Juntti, M. J. (1998a). Multiuser detector performance comparisons in multirate CDMA systems. In *Proceedings of the IEEE 48th Vehicular Technology Conference*, volume 1, pages 31–35, Ottawa, ON.

[Juntti, 1998b] Juntti, M. J. (1998b). System concept comparisons for multirate CDMA with multiuser detection. In *Proceedings of the IEEE 48th Vehicular Technology Conference*, volume 1, pages 36–40, Ottawa, ON.

[Juntti et al., 1997] Juntti, M. J., Latvaaho, M., and Heikkila, M. (1997). Performance comparison of PIC and decorrelating multiuser receiver in fading channels. In *Proceedings of the IEEE Global Telecommunications Conference*, pages 609–613, Phoenix, AZ.

[Kawadia and Kumar, 2003] Kawadia, V. and Kumar, P. (2003). A cautionary perspective on cross layer design. University of Illinois at Urbana-Champaign. Preprint.

[Kim and Honig, 1998] Kim, J. B. and Honig, M. L. (1998). Outage probability of multi-code DS-CDMA with linear interference suppression. In *Proceedings of the IEEE Military Communications Conference*, pages 248–252, Boston, MA.

[Kim and Bambos, 2001] Kim, J.-W. and Bambos, N. (2001). Power control for multirate wireless networks with groupwise serial multiuser detection. In *Proceedings of the IEEE Global Telecommunications Conference*, volume 5, pages 3201–3205, San Antonio, TX.

[Kiran and Tse, 2000] Kiran and Tse, D. (2000). Effective bandwidths and effective interference for linear multiuser receivers in asynchronous systems. *IEEE Transactions on Information Theory*, 46(4):1426–1447.

[Kishore et al., 2003] Kishore, S., Greenstein, L. J., Poor, H. V., and Schwartz, S. C. (2003). Capacity in a CDMA macrocell with a hotspot microcell: Exact and approximate analyses. *IEEE Transactions on Wireless Communications*, 2(2):364–374.

[Kostic, 2001] Kostic, Z. (2001). The performance of the WWW traffic due to interactions between TCP and RLP protocols in a cellular system. In *Proceedings of the IEEE 54th Vehicular Technology Conference*, volume 2, pages 1158–1162, Atlantic City, NJ.

[Kumar and Holtzman, 1995] Kumar, P. and Holtzman, J. (1995). Power control for a spread spectrum system with multiuser receivers. In *Proceedings of the 6th International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 955–959, Toronto, ON.

[Kumar et al., 1995] Kumar, P., Yates, R., and Holtzman, J. (1995). Power control based on bit error rate (BER) measurements. In *Proceedings of the IEEE Military Communications Conference*, pages 617–620, San Diego, CA.

[Lansford et al., 2000] Lansford, J., Negus, K. J., and Stephens, A. P. (2000). HomeRF: Wireless networking for the connected home. *IEEE Personal Communications*, 7(1):20–27.

[Lim and Roy, 1998] Lim, T. J. and Roy, S. (1998). Adaptive filters in multiuser (MU) CDMA detection. *Wireless Networks*, 4(4):307–318.

[Lupas and Verdú, 1989] Lupas, R. and Verdú, S. (1989). Multiuser detectors for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 35(1):123–136.

[MacKenzie and Wicker, 2001] MacKenzie, A. B. and Wicker, S. B. (2001). Game theory and design of self-configuring adaptive wireless networks. *IEEE Communications Magazine*, 39(11):126–131.

[Madhow, 2000] Madhow, U. (2000). Multiuser detection: An overview and a new result. In *Proceedings of the 2nd International Workshop on Independent Component Analysis and Blind Signal Separation*, Helsinki, Finland.

[Madhow and Honig, 1994] Madhow, U. and Honig, M. L. (1994). Interference suppression for direct-sequence spread-spectrum CDMA. *IEEE Transactions on Communications*, 42(12):3178–3188.

[Mergen and Tong, 2001] Mergen, G. and Tong, L. (2001). Receiver controlled medium access in multihop ad hoc networks with multipacket reception. In *Proceedings of the IEEE Military Communications Conference*, pages 1014–1018, Washington, DC.

[Mergen and Tong, 2002] Mergen, G. and Tong, L. (2002). Random scheduling medium access for wireless ad hoc networks. In *Proceedings of the IEEE Military Communications Conference*, pages 868–872, Anaheim, CA.

[Meshkati et al., 2003] Meshkati, F., Poor, H. V., Schwartz, S. C., and Mandayam, N. B. (2003). Linear multiuser receivers and power control in wireless data networks: A game-theoretic approach. In *Proceedings of the Conference on Information Sciences and Systems*, The Johns Hopkins University, Baltimore, MA.

[Miller, 2001] Miller, L. E. (2001). Distribution of link distances in a wireless network. *Journal of Research of the National Institute of Standards and Technology*, 106(2):401–412.

[Mostofa et al., 2001] Mostofa, N., Howlader, K., and Woerner, B. D. (2001). System architecture for implementing multiuser detector within an ad-hoc network. In *Proceedings of the IEEE Military Communications Conference*, volume 2, pages 1119–1123, Washington, DC.

[Moustakides and Poor, 2001] Moustakides, G. V. and Poor, H. V. (2001). On the relative error probabilities of linear multiuser detectors. *IEEE Transactions on Information Theory*, 47(1):450–456.

[Müller et al., 1997] Müller, R., Schramm, P., and Huber, J. (1997). Spectral efficiency of CDMA systems with linear interference suppression (English translation. In *Proceedings of the Workshop Kommunikationstechnik*, pages 93–97, Ulm, Germany.

[Nettleton and Alavi, 1983] Nettleton, R. and Alavi, H. (1983). Power control for spread-spectrum cellular mobile radio system. In *Proceedings of the IEEE 33rd Vehicular Technology Conference*, pages 242–246, Toronto, ON.

[Ortigoza-Guerrero and Aghavami, 2000] Ortigoza-Guerrero, L. and Aghavami, A. H. (2000). *Resource Allocation in Hierarchical Cellular Systems*. Artech House, Norwood, MA.

[Ottosson and Svensson, 1995] Ottosson, T. and Svensson, A. (1995). Multi-rate schemes in DS/CDMA systems. In *Proceedings of the IEEE 45th Vehicular Technology Conference*, volume 2, pages 1006–1010, Chicago, IL.

[Patel and Holtzman, 1994a] Patel, P. and Holtzman, J. (1994a). Analysis of a simple successive interference cancellation scheme in a DS/CDMA system. *IEEE Journal on Selected Areas in Communications*, 12(5):796–807.

[Patel and Holtzman, 1994b] Patel, P. and Holtzman, J. (1994b). Performance comparison of a DS/CDMA system using a successive interference cancellation (IC)

scheme and a parallel IC scheme under fading. In *Proceedings of the IEEE International Conference on Communications*, volume 1, pages 510–514, New Orleans, LA.

[Perevalov and Blum, 2003] Perevalov, E. and Blum, R. (2003). Delay limited capacity and ad hoc networks: asymptotically optimal transmission and relaying strategy. In *Proceedings of IEEE INFOCOM*, volume 2, pages 1575–1582, San Francisco, CA.

[Poor, 2004] Poor, H. V. (2004). Iterative multiuser detection. *IEEE Signal Processing Magazine*, 21(1):81–88.

[Poor and Verdú, 1997] Poor, H. V. and Verdú, S. (1997). Probability of error in MMSE multiuser detection. *IEEE Transactions on Information Theory*, 43(5):858–871.

[Radunovic and Boudec, 2002] Radunovic, B. and Boudec, J.-Y. (2002). *Joint Scheduling, Power Control and Routing in Symmetric, One-dimensional, Multihop Wireless Networks*. Technical report, EPFL, Lausanne, Switzerland.

[Radunovic and Boudec, 2004] Radunovic, B. and Boudec, J.-Y. (2004). Optimal power control, scheduling and routing in UWB networks. *IEEE Journal on Selected Areas in Communications*, to appear.

[Rodoplu and Meng, 2000] Rodoplu, V. and Meng, T. H. (2000). Position based CDMA with multiuser detection (P-CDMA/MUD) for wireless ad hoc networks. In *Proceedings of the IEEE 6th International Symposium on Spread Spectrum Techniques and Applications*, volume 1, pages 336–340, Parsippany, NJ.

[Rupf and Massey, 1994] Rupf, M. and Massey, J. (1994). Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 40(4):1226–1266.

[Samardzija et al., 2002] Samardzija, D., Mandayam, N. B., and Seskar, I. (2002). Blind successive interference cancellation for CDMA systems. *IEEE Transactions on Communications*, 50(2):276–290.

[Sankaran and Ephremides, 2002] Sankaran, C. and Ephremides, A. (2002). The use of multiuser detectors for multicasting in wireless ad hoc CDMA networks. *IEEE Transactions on Information Theory*, 48(11):2873–2887.

[Saquib et al., 2000] Saquib, M., Yates, R. D., and Ganti, A. (2000). Power control for an asynchronous multirate decorrelator. *IEEE Transactions on Communications*, 48(5):804–812.

[Saquib et al., 1999] Saquib, M., Yates, R. D., and Mandayam., N. B. (1999). Decorrelating detectors for a dual rate synchronous DS/CDMA channel. *Wireless Personal Communications*, 9(3):197–216.

[Saraydar et al., 2002] Saraydar, C., Mandayam, N. B., and Goodman, D. J. (2002). Efficient power control via pricing in wireless data networks. *IEEE Transactions on Communications*, 50(2):291–303.

[Schlegel and Grant, 2000] Schlegel, C. and Grant, A. (2000). Polynomial complexity optimal detection of certain multiple-access systems. *IEEE Transactions on Information Theory*, 46(6):2246–2248.

[Seskar and Mandayam, 1999a] Seskar, I. and Mandayam, N. B. (1999a). Software defined radio architectures for interference cancellation in DS-CDMA systems. *IEEE Personal Communications*, 6(4):26–34.

[Seskar and Mandayam, 1999b] Seskar, I. and Mandayam, N. B. (1999b). A software radio architecture for linear multiuser detection. *IEEE Journal on Selected Areas in Communications*, 17(5):814–823.

[Shakkottai et al., 2003] Shakkottai, S., Rappaport, T., and Karlsson, P. (2003). Cross layer design for wireless networks. *IEEE Communications Magazine*, 41(10):74–80.

[Shamai and Verdú, 2001] Shamai, S. and Verdú, S. (2001). The impact of frequency-flat fading on the spectral efficiency of CDMA. *IEEE Transactions on Information Theory*, 47(4):1302–1327.

[Shepard, 1995] Shepard, T. (1995). *Decentralized Channel Management in Scalable Multihop Spread-Spectrum Packet Radio Networks*. PhD thesis, Massachusets Institute of Technology, Cambridge, MA. ftp://ftp-pubs.lcs.mit.edu/pub/lcs-pubs/tr.outbox/MIT-LCS-TR-670.ps.gz.

[Shu and Niu, 2003] Shu, T. and Niu, Z. (2003). Capacity optimization by using cancellation-error-ascending decoding order in multimedia CDMA networks with imperfect successive interference cancellation. In *Proceedings of the IEEE International Conference on Communications*, volume 3, pages 2170–2174, Anchorage, AK.

[Shum and Cheng, 2000] Shum, S. M. and Cheng, R. S. (2000). Power control for multirate CDMA systems with interference cancellation. In *Proceedings of the IEEE Global Telecommunications Conference*, volume 2, pages 895–900, San Francisco, CA.

[Silverstein and Bai, 1995] Silverstein, J. W. and Bai, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal on Multivariate Analysis*, 54(2):175–192.

[Singh et al., 2001] Singh, S., Krishnamurthy, V., and Poor, H. V. (2001). Admission control for DS-CDMA systems with fading. In *Proceedings of the IEEE Global Telecommunications Conference*, pages 705–710, San Antonio, TX.

[Sivarajan et al., 1989] Sivarajan, K., McEliece, R., and Ketchum, J. (1989). Channel Assignment in Cellular Radio. In *Proceedings of the IEEE 39th Vehicular Technology Conference*, pages 846–850, San Francisco, CA.

[Song and Mandayam, 2001] Song, L. and Mandayam, N. B. (2001). Hierarchical SIR and rate control on the forward link for CDMA data users under delay and error constraints. *IEEE Journal on Selected Areas in Communications*, 19(10):1871–1882.

[Song et al., 2001] Song, L., Mandayam, N. B., and Gajic, Z. (2001). Analysis of an up/down power control algorithm for the CDMA reverse link under fading. *IEEE Journal on Selected Areas in Communications*, 19(2):277–286.

[Strang, 1988] Strang, G. (1988). *Linear Algebra and its Applications.* Harcourt Brace Jovanovich, San Diego, CA.

[Tijms, 1986] Tijms, H. C. (1986). *Stochastic Modelling and Analysis: A Computational Approach.* John Wiley and Sons, Chichester, UK.

[Tong et al., 2001] Tong, L., Zhao, Q., and Mergen, G. (2001). Multipacket reception in random access wireless networks: From signal processing to optimal medium access control. *IEEE Communication Magazine*, 39(11):108–112.

[Toumpis and Goldsmith, 2001] Toumpis, S. and Goldsmith, A. J. (2001). Capacity regions for wireless ad hoc networks. Stanford University, Stanford, CA. Preprint.

[Tse and Hanly, 1999] Tse, D. and Hanly, S. (1999). Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Transactions on Information Theory*, 45(2):641–657.

[Tse and Zeitouni, 2000] Tse, D. and Zeitouni, O. (2000). Performance of linear multiuser receivers in random environments. *IEEE Transactions on Information Theory*, 46(1):171–188.

[Ulukus and Yates, 1998a] Ulukus, S. and Yates, R. (1998a). Adaptive power control and multiuser interference suppression. *ACM Wireless Networks*, 4(6):489–496.

[Ulukus and Yates, 1998b] Ulukus, S. and Yates, R. (1998b). A blind adaptive decorrelating detector for CDMA systems. *IEEE Journal on Selected Areas in Communications*, 16(8):1530–1541.

[Ulukus and Yates, 1998c] Ulukus, S. and Yates, R. (1998c). Optimum multiuser detection is tractable for synchronous CDMA systems using m-sequences. *IEEE Communications Letters*, 2(4):89–91.

[Ulukus and Yates, 1998d] Ulukus, S. and Yates, R. (1998d). Stochastic power control for cellular radio systems. *IEEE Transactions on Communications*, 46(6):784–798.

[Varanasi and Das, 2002] Varanasi, M. K. and Das, D. (2002). Fast stochastic power control algorithms for nonlinear multiuser receivers. *IEEE Transactions on Communications*, 50(11):1817–1827.

[Verdú, 1986] Verdú, S. (1986). Minimum probability of error for asynchronous Gaussian multiple-access channels. *IEEE Transactions on Information Theory*, 32(1):85–96.

[Verdú, 1998] Verdú, S. (1998). *Multiuser Detection.* Cambridge University Press, Cambridge, UK.

[Verdú and Shamai, 1997] Verdú, S. and Shamai, S. (1997). Multiuser detection with random spreading and error-correction codes: Fundamental limits. In *Proceedings*

of the 35th Annual Allerton Conference on Communication, Control and Computing, University of Illinois, Monticello, IL.

[Verdú and Shamai, 1999] Verdú, S. and Shamai, S. (1999). Spectral efficiency of CDMA with random spreading. *IEEE Transactions on Information Theory*, 45(2):622–639.

[Vijayan and Holtzman, 1993] Vijayan, R. and Holtzman, J. (1993). A model for analyzing handoff algorithms. *IEEE Transactions on Vehicular Technology*, 42(3):351 – 356.

[Viswanath et al., 1999] Viswanath, P., Anantharam, V., and Tse, D. (1999). Optimal sequences, power control and capacity of synchronous CDMA systems with linear MMSE multiuser receivers. *IEEE Transactions on Information Theory*, 45(6):1968–1983.

[Viterbi, 1995] Viterbi, A. J. (1995). *CDMA: Principles of Spread Spectrum Communications*. Addison-Wesley, Reading, MA.

[Wang et al., 2001] Wang, C.-L., Li, M.-H., Wu, K.-M., and Hwang, K.-L. (2001). Adaptive interference suppression with power control for CDMA systems. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, volume 4, pages 286–289, Sydney, Australia.

[Wang and Poor, 1998] Wang, X. and Poor, H. V. (1998). Blind multiuser detection: A subspace approach. *IEEE Transactions on Information Theory*, 44(2):677–690.

[Wang and Poor, 2004] Wang, X. and Poor, H. V. (2004). *Wireless Communication Systems: Advanced Techniques for Signal Reception*. Prentice-Hall: Upper Saddle River, NJ.

[Wijk et al., 1995] Wijk, F., Janssen, G. M. J., and Prasad, R. (1995). Groupwise successive interference cancellation in a DS/CDMA system. In *Proceedings of the 6th International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 2, pages 742–746, Toronto, ON.

[Wijting et al., 1999] Wijting, C. S., Ojanpera, T., Juntti, M., Kansanen, K., and Prasad, R. (1999). Groupwise serial multiuser detectors for multirate DS-CDMA. In *Proceedings of the 49th IEEE Vehicular Technology Conference*, volume 1, pages 836–840, Houston, TX.

[WiMax, 2004] WiMax (2004). www.wimaxforum.org/home.

[Win and Scholtz, 2000] Win, M. Z. and Scholtz, R. A. (2000). Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple access communications. *IEEE Transactions on Communications*, 48(4):679–689.

[Wu and Geraniotis, 1994] Wu, T. and Geraniotis, E. (1994). CDMA with multiple chip rates for multimedia communications. In *Proceedings of the Conference on Information Sciences and Systems*, pages 992–997, Princeton University, Princeton, NJ.

[Wyrwas et al., 1992] Wyrwas, R., Miller, M. J., Anjaria, R., and Zhang, W. (1992). Multiple access options for multimedia wireless systems. In *Proceedings of the 3rd*

Workshop on Third Generation Wireless Information Networks, pages 289–294, WINLAB, Rutgers University, Piscataway, NJ.

[Xiao et al., 2001] Xiao, M., Shroff, N. B., and Chong, E. (2001). Utility-based power control in cellular wireless systems. In *Proceedings of the IEEE INFOCOM*, pages 412–421. Anchorage, AK.

[Xiao and Honig, 2002] Xiao, W. and Honig, M. L. (2002). Forward-link performance of satellite CDMA with linear interference suppression and one-step power control. *IEEE Transactions on Wireless Communications*, 1(4):600–610.

[Xu et al., 2002] Xu, G., Rajagopal, S., Cavallaro, J. R., and Aazhang, B. (2002). VLSI implementation of the multistage detector for next generation wideband CDMA receivers. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 30(1-3):21–33.

[Xu and Saadawi, 2001] Xu, S. and Saadawi, T. (2001). Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks? *IEEE Communications Magazine*, 39(6):130–137.

[Xue et al., 1998] Xue, G., Weng, J., and T. Le-Ngoc, S. T. (1998). *Multiuser Detection Techniques: An Overview*. Technical report, Department of Electrical and Computer Engineering, Concordia University, Montreal, PQ.

[Yang, 1995] Yang, B. (1995). Projection approximation subspace tracking. *IEEE Transactions on Signal Processing*, 44(1):95–107.

[Yao et al., 2004] Yao, Y., Poor, H. V., and Sun, F. (2004). User capacity for synchronous multirate CDMA systems with linear MMSE receivers. *IEEE Transactions on Information Theory*, 50(11).

[Yates, 1995] Yates, R. (1995). A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1348.

[Yates and Huang, 1995] Yates, R. and Huang, C. (1995). Integrated power control and base station assignment. *IEEE Transactions on Vehicular Technology*, 44(3):638–644.

[Yates and Goodman, 1999] Yates, R. D. and Goodman, D. (1999). *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers*. John Wiley & Sons, New York, NY.

[Zander, 1992a] Zander, J. (1992a). Distributed cochannel control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(3):305–311.

[Zander, 1992b] Zander, J. (1992b). Performance of optimum transmitter power control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41:57–62.

[Zander, 2001] Zander, J. (2001). Affordable QoS in future wireless networks: Myth or reality. In *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages C–39– C–43, San Diego, CA.

[Zander and Kim, 2001] Zander, J. and Kim, S. (2001). *Radio Resource Management for Wireless Networks.* Artech House, Norwood, MA, first edition.

[Zhang and Chong, 2000] Zhang, J. and Chong, E. K. P. (2000). Power control for spread-spectrum networks in fading channels. In *Proceedings of the IEEE 6th International Symposium on Spread-Spectrum Techniques and Applications*, pages 790–794, Parsippany, NJ.

[Zhang et al., 2001] Zhang, J., Chong, E. K. P., and Tse, D. N. C. (2001). Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems. *IEEE Transactions on Information Theory*, 47(3):1128–1144.

[Zhang and Wang, 2002a] Zhang, J. and Wang, X. (2002a). Large-system analysis of blind and group-blind multiuser receivers. In *Proceedings of the IEEE International Symposium on Information Theory*, page 190, Lausanne, Switzerland.

[Zhang and Wang, 2002b] Zhang, J. and Wang, X. (2002b). Large-system performance analysis of blind and group-blind multiuser receivers. *IEEE Transactions on Information Theory*, 48(9):2507–2523.

[Zhao and Tong, 2003] Zhao, Q. and Tong, L. (2003). A multi-queue service room MAC protocol for wireless networks with multipacket reception. *IEEE/ACM Transactions on Networking*, 11(1):125–137.

[Zhao and Tong, 2004] Zhao, Q. and Tong, L. (2004). A dynamic queue protocol for multiaccess wireless networks with multipacket reception. *IEEE Transactions on Wireless Communications,* to appear.

About the Authors

Cristina Comaniciu received the M.S. degree in electronics from the Polytechnic University of Bucharest in 1993, and the Ph.D. degree in electrical and computer engineering from Rutgers University in December 2001. From 1998 to 2001 she was with the Wireless Information Network Laboratory (WINLAB) at Rutgers University, working on integrated access control and detection algorithms for multimedia CDMA systems. From 2002 to 2003 she was affiliated with the Electrical Engineering Department at Princeton University as a research associate. Since August 2003, she has been an assistant professor of Electrical and Computer Engineering at Stevens Institute of Technology. Her research interests include cross-layer design for wireless networks, radio resource management for cellular and ad hoc networks, admission/access control for multimedia wireless systems, and multiuser detection.

Narayan Mandayam is a Professor of Electrical and Computer Engineering (ECE) and also the Associate Director of the Wireless Information Network Laboratory (WINLAB) at Rutgers University. He received the Ph.D. degree in ECE from Rice University in 1994 and the B. Tech degree from the Indian Institute of Technology (IIT), Kharagpur in 1989. His research interests are primarily in the area of wireless networks with emphasis on radio resource management and signal processing. He is the author or co-author of more than 100 publications in these areas. He has served as an editor for the journals *IEEE Communications Letters* and *IEEE Transactions on Wireless Communications*. Among his honors are the National Science Foundation CAREER award and the Institute Silver Medal from IIT, Kharagpur.

H. Vincent Poor is the George Van Ness Lothrop Professor in Engineering at Princeton University. He received the Ph.D. degree in EECS from Princeton in 1977, and prior to joining the Princeton faculty in 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. His research interests are primarily in the areas of wireless networks, statistical signal processing, and related fields. He is the author or co-author of more than 500 publications in these areas, including nine books. Dr. Poor is a member of the National Academy of Engineering, and is a Fellow of the IEEE and other organizations. He is a past President of the IEEE Information Theory Society, and is currently Editor-in-Chief of the *IEEE Transactions on Information Theory*. Among his honors are the IEEE Graduate Teaching Award, the NSF Director's Award for Distinguished Teaching Scholars, and a Guggenheim Fellowship.

Index

199