NETWORK COORDINATION FOR SPECTRALLY EFFICIENT COMMUNICATIONS IN WIRELESS NETWORKS

BY M. KEMAL KARAKAYALI

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Prof. Roy D. Yates and Dr. Gerard J. Foschini

and approved by

———————————————————

———————————————————

———————————————————

———————————————————

———————————————————

New Brunswick, New Jersey

January, 2007

ABSTRACT OF THE DISSERTATION

Network Coordination for Spectrally Efficient Communications in Wireless Networks

by M. Kemal Karakayali

Dissertation Director: Prof. Roy D. Yates and Dr.

Gerard J. Foschini

In conventional cellular systems, each base station transmits signals intended for users within its cell coverage. Depending on the users' channel conditions, interference caused by the neighboring cell transmissions can sharply degrade the received signal quality. Thus, the downlink capacity of cellular wireless networks is limited by inter-cell interference. Fortunately, since the base stations can be connected via a high-speed backbone, there is an opportunity to coordinate the base antenna transmissions so as to minimize the inter-cell interference, and hence to increase the downlink system capacity. In this thesis, we study various aspects of network coordination in cellular downlink systems.

In the first part of the study, we describe various coordination techniques, and conduct their performance analysis. The performance of each technique is given in terms of the max-min fair rate achievable subject to per-base power constraints.

We compare the performance of coordinated transmissions to that of conventional cellular networks without coordination. It is shown that the coordinated base station transmissions can help to eliminate inter-cell interference, and result in a great capacity improvement on the downlink cellular networks.

In the second part of the study, we consider coordinated networks with multiple antennas. The great advantage of using multiple antennas is that, without increasing power or bandwidth, the capacity of a point-to-point link scales linearly with the minimum of the number of transmit or receive antennas deployed. The gain, in terms of the marginal increase in rate when an additional antenna is deployed, is especially large when the signal-to-noise ratio is high. We show that, without coordination, the link qualities can be very poor because of inter-cell interference. In this case, the network does not benefit significantly from multiple antennas. When the coordination is employed, the inter-cell interference is mitigated so that the links can operate in the high signal-to-noise ratio regime. This enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas.

In the final part of the study, we investigate linear beamforming design with per-antenna power constraints. We show that the standard beamforming techniques used mostly in the sum-power constrained systems are suboptimal when there are per-antenna power constrains. We formulate convex optimization problems finding the optimum zero-forcing beamforming vectors. We observe that optimizing the antenna outputs based on the per-antenna constraints may improve the rate considerably when the number of transmit antennas is much larger the number of receive antennas. The network coordination techniques assume the existence of a high-speed backhaul enabling communications between the base stations. We conclude the thesis by considering the design of a mesh architecture providing such a backhaul support in a cellular network.

Acknowledgements

I would like to thank my advisors, Prof. Roy Yates and Dr. Gerard Foschini for their invaluable technical advice and personal support throughout my PhD studies. Both Prof. Roy Yates and Dr. Gerard Foschini have been truly inspiring role models. I am indebted to both of them for their willingness and generosity in sharing their experiences with me. Their positive and caring attitude made my graduate school life an enjoyable educational experience.

I would like to extend my thanks to Prof. Narayan Mandayam, Prof. Predrag Spasojevic and Dr. Reinaldo Valenzuela for being in my thesis committee. Their comments and suggestions improved the quality of the thesis. Most of the work in this thesis is a result of collaborative efforts with Bell Labs. In particular, I would like to thank members of the Wireless Communications Research Department and the Network Performance and Reliability Department for their contributions in Chapters 2-4 and Chapter 5 respectively. I learned wireless communications at WINLAB, and I am thankful to all WINLAB colleagues for making a friendly and fruitful research environment.

Finally, I would like to thank all my friends, here at Rutgers and elsewhere, for making my life enjoyable. My deepest thanks go to my family for the lifelong encouragement, support and love.

Table of Contents

vii

List of Figures

Chapter 1

Introduction

As the demand for wireless applications continues to grow, future wireless systems are engineered to provide high-speed broadband services with quality of service (QoS) support for a wide range of applications, including voice and multimedia data. The design of such advanced network architectures is a challenging task. First of all, radio resources such as power and bandwidth are often scarce. Therefore, efficient resource allocation and network optimization are vital. Secondly, the wireless channel has its unique impairments such as fading and multi-path. Also, the mobiles sharing a common communication medium interfere with each other. The future wireless system designs are expected to address all these challenges, and networks with high spectral efficiency is a major design goal.

Traditionally, channel fading and multi-path are thought of as impairments that have to be dealt with. A breakthrough technique which dramatically changed this point of view is the use of multiple antennas at the transmitters and/or receivers [1–3]. The advantage of multiple input-multiple output (MIMO) antenna arrangements is that, when the wireless channel provides a rich scattering environment, multiple independent signal paths can be obtained between two communicating units employing multiple antennas. In this case, the link capacity can scale linearly with the number of antennas deployed, even without increasing transmit power or bandwidth.
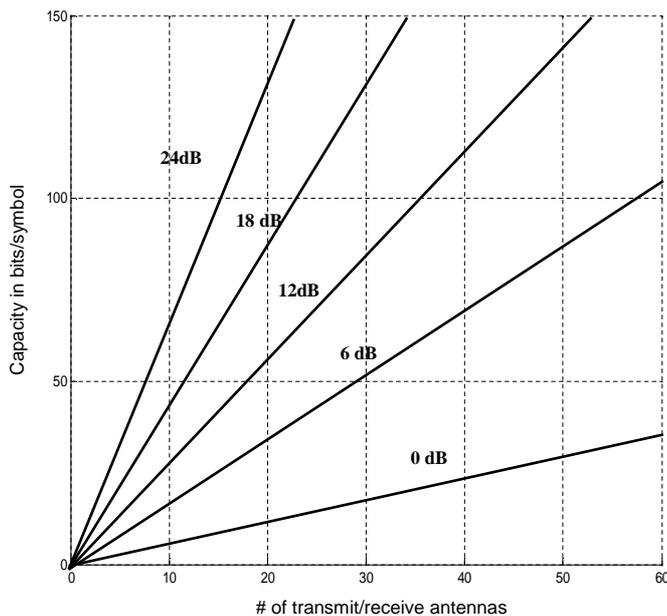
Figure 1.1: MIMO capacity gains in a point-to-point link.

The tremendous capacity improvement predicted by the multiple antennas attracted a huge attention within the research community [4–23]. In addition to the experimental verification and the measurement of these gains [5], multiple antennas have been successfully employed in a number of commercial products, especially in short-range wireless local area networks (WLANs); see for example [24]. However, the deployment of multiple antennas in cellular networks has not been successful so far. One important reason is that cellular networks suffer from inter-cell interference. Therefore, the link qualities in cellular networks can be relatively poor compared to short-range radio links. On the other hand, the multiple antenna gain, in terms of marginal increase in rate when an additional antenna is deployed, is especially large when the signal-to-noise ratio (SNR) is high [6]. This fact can be observed in Figure 1.1 illustrating capacity gains in

a point-to-point link. The figure shows that the slope of the capacity vs. the number of antennas curve is small when the received SNR is low. For example, at 0 dB SNR, adding one more antenna to both the transmitter and the receiver contributes an additional ≈ 0.65 bits/symbol to the spectral efficiency, while the capacity improves by about 6.5 bits/symbol when the SNR is 24 dB. In realistic cellular environments, the received SNR is typically around 18 dB when the mobile is close to the cell border. On the other hand, when the inter-cell interference is present, the SINR (signal-to-noise-plus-interference ratio) may drop to 0 dB. In this case, the improvement due to multiple antennas would be very limited if the interference is not successfully mitigated, as in the conventional cellular networks.

In this dissertation, we propose novel techniques that enable efficient use of multiple antennas in cellular networks. In particular, we study *network coordination* as a means to provide spectrally efficient communications in cellular downlink systems. When the network coordination is employed, all base antennas act together as a single network antenna array, and each mobile may receive useful signals from several nearby base stations. Furthermore, the antenna outputs are chosen in ways to minimize the out-of-cell interference, and hence to increase the downlink system capacity. When the out-of-cell interference is mitigated, the links can operate in the high SNR regime. This enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas.

1.1 Related Work

The pioneering multiple antenna studies are due to [1–3]. The initial multiple antenna research mainly focused on point-to-point links. In particular, [1,2] showed that the link capacity scales linearly with the minimum of the number of transmit and receive antennas. This result requires that the wireless channel provides enough scattering so that each antenna pair would experience an independent channel fading and the matrix channel between the transmitter and the receiver becomes full-rank. Also, the channel information must be available at the receiver. The availability of the channel information at the transmitter enables optimal water-filling power allocation on the eigen-modes of the channel [2]. A number of efficient coding/modulation techniques that help to realize the predicted capacity gains can be found in [6–10].

Multi-user multiple antenna problems are studied in the context of, first, multi-access channels [11–15], and then broadcast channels [16–21]. The capacity region of the multiple antenna multi-access channel follows easily from the capacity region of its scalar counterpart [25]. Basically, successive interference cancelation extends to the multiple antenna system [11]. In this case, each corner point on the boundary of the capacity region corresponds to a particular user order, and each user selects its optimal single-user transmit covariance assuming the existence of interference due to un-canceled interferers in the user order. The multiple antenna broadcast channel problem is much more difficult because it is a nondegraded broadcast channel [25]. Recently, there has been progress towards characterizing its capacity region. Namely, the dirty paper coding [26, 27], which is a technique to cancel the interference causally known to the transmitter without any transmit power penalty, achieves the boundary of the broadcast capacity

region [19, 20]. The multi-access and the broadcast capacity regions are related through a duality relationship [18]. The usefulness of this relationship is that relatively simple multi-access problem can be solved to find optimum transmission policies achieving a particular point on the dual broadcast capacity region. Note that a sum-power constraint is assumed in the above broadcast channel problems. Moreover, the above results are derived for single-cell systems.

While multiple antennas have been studied extensively in the context of point-to-point links and single-cell systems, the effectiveness of multiple antennas in multi-cell systems is not well-understood. In fact, some of the above results apply to multi-cell systems. For example, an uplink multi-cell system with base station cooperation is essentially equivalent to a single cell system with an increased receive antenna count. However, the downlink multi-cell model is fundamentally different from the single cell model. This is mainly due to the fact that each base station/antenna has its own power constraint in a multi-cell system. In this case, the results and the tools, such as the uplink-downlink duality developed for broadcast channels with sum-power constraints, can not be used easily for the multi-cell downlink model. Therefore, new tools and approaches are necessary to analyze the multiple antenna multi-cell downlink networks.

In this thesis, the objective is to contribute to the understanding of the multiple antennas in cellular networks. To achieve this goal, we consider coordinated cellular networks where the base stations can communicate and cooperate via a high-speed backbone. We study the performance limits of such networks with and without multiple antennas. We investigate the impact of such cooperation on the effectiveness of multiple antennas. To the best of our knowledge, this thesis is the first study proposing network coordination as a means to improve

the performance of multiple antennas in cellular networks. For earlier studies on base station cooperation, see [28] for an uplink multi-cell system. We should note that it is hard to derive analytical expressions in multi-cell formulations since the results depend heavily on the network topology. Thus, [28] considers a limited circular network topology to investigate the effect of joint-decoding on the uplink multi-cell network. For downlink networks, [29] studied the base station cooperation. Considering single antenna base stations and mobiles, and assuming a sum-power constraint relaxation, simple analytical expressions are derived for the capacity of such networks.

1.2 Overview of Dissertation

The thesis study starts with the analysis of the downlink cellular networks with single antenna base stations and mobiles. In Chapter 2, we examine coordinated networks in which a high-speed backbone enables all network antennas to operate as a single network-wide antenna array. We compare the max-min rate performance of practical coordination techniques to that of a conventional network with power control. We show that the coordinated base antenna transmissions provide large capacity improvements over the conventional cellular networks.

In Chapter 2, we also study the theoretical performance limits of coordination. Our system can be modeled as a non-degraded Gaussian broadcast channel for which the optimal scheme achieving the boundary of the capacity region involves the dirty paper coding [19, 27]. However, characterizing the max-min rate point on the capacity region is nontrivial as the dirty paper rate functions are in general non-concave in transmit covariances [18]. In this case, the optimal policy may

involve time-sharing, and one may need to specify time-sharing combinations as well as a user ordering for the dirty paper encoding. This becomes computationally infeasible for large systems as one needs to search over all possible user combinations and encoding orders. Our approach is to develop tight upper bounds on the max-min rate. We show that the rate achievable by a particular form of zero-forcing dirty paper coding scheme combined with a heuristic user ordering is very close to the sharpest upper bound obtained. The established bounds not only help to avoid a huge computational burden, but also they give insights into the ultimate performance of the system.

Also in Chapter 2, we discuss the impact of sectorization on the downlink performance. We compare sectorization with coordination, and then consider a coordinated network architecture with sectorization. While the sectorization by itself helps to reduce inter-cell interference, employing network coordination with sectorization gives substantial additional improvements in system capacity. In this chapter, we also discuss some implementation issues such as availability of channel information. In particular, we evaluate the system performance when only a partial channel information is available at the transmitters.

In Chapter 3, we study the impact of network coordination on multiple antenna systems. We show that the coordinated transmissions are especially effective when the base stations and the mobiles are equipped with multiple antennas. Note that deploying multiple antennas in a cellular network is a challenging task due to its complexity of implementation and the cost of infrastructure upgrades in the current cellular architecture. In order to justify the use of multiple antennas in a cellular network, the predicted multiple antenna gains must be realized

in practice. In this chapter, we show how to design a coordinated network employing multiple antennas. Our results indicate that the interference mitigation capability of network coordination enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas.

In Chapters 2 and 3, we study coordination techniques that require simple, practical linear beamforming techniques. While these beamforming methods provide significant capacity improvements over the conventional cellular networks, they are not claimed to be the optimal linear coordination techniques. The objective in Chapter 4 is to find the optimum beamforming vectors. More specifically, in all problem formulations in the thesis, we consider per-antenna power constraints for which well-known beamforming techniques such as the pseudo-inverse zero-forcing is suboptimal. We show that finding the optimum beamforming vectors requires solving a convex optimization problem. We will see that optimizing the antenna outputs based on the per-antenna constraints may improve the rate considerably when the number of transmit antennas is much larger the number of receive antennas. The reason is that more transmit antennas will give more degrees of freedom to optimize the antenna outputs. When the number of transmit and receive antennas are close to each other, there are not much room left to exploit in the signal space.

The network coordination techniques assume the existence of a high-speed backhaul enabling communications between the base stations. In Chapter 5, we study the design of such a backhaul in a cellular network. In particular, we consider a mesh backhaul network consisting of fixed base stations (mesh routers) connected by wireless links. Some of the mesh routers are assumed to have connections to the wired network, and therefore can function as gateways. The

network has multi-hop capability where the traffic entering the mesh backhaul through the gateway routers can be carried over multiple wireless links towards the destination mesh routers. The objective in this chapter is to study the performance limits of such networks. Assuming the use of an OFDMA air-interface for the mesh backhaul network, we formulate a cross-layer optimization problem that involves power control, channel allocation, link scheduling and routing. We show that when the radio resources are optimized carefully, OFDM transmissions may provide tone-diversity advantage in the form of efficient bandwidth utilization by choosing better channels for transmissions and scheduling, or in the form of improved routing performance by providing more path options to route the traffic. Our numerical results indicate that OFDMA-based mesh architecture provides an efficient backhaul solution in cellular networks.

Chapter 2

Network Coordination for Single Antenna Base Stations and Mobiles

In conventional cellular systems, each base station transmits signals intended for users within its cell coverage. Depending on the users' channel conditions, interference caused by the neighboring cell transmissions can sharply degrade the received signal quality. Thus, the downlink capacity of cellular wireless networks is limited by inter-cell interference. Fortunately, since the base stations can be connected via a high-speed backbone, there is an opportunity to coordinate the base antenna transmissions so as to minimize the inter-cell interference, and hence to increase the downlink system capacity. In this chapter, we study various aspects of network coordination in cellular downlink systems.

Figure 2.1 shows the basic idea of the network coordination. On the top figure, the base antenna transmissions are not coordinated, and therefore neighboring base transmissions are received as inter-cell interference (each color represents a signal useful for a given mobile). The objective of network coordination is to enable cooperation between the base stations so that useful signals, as opposed to the interference, can be received from the neighboring base antennas. In the following, we will describe the transmission techniques achieving this objective.
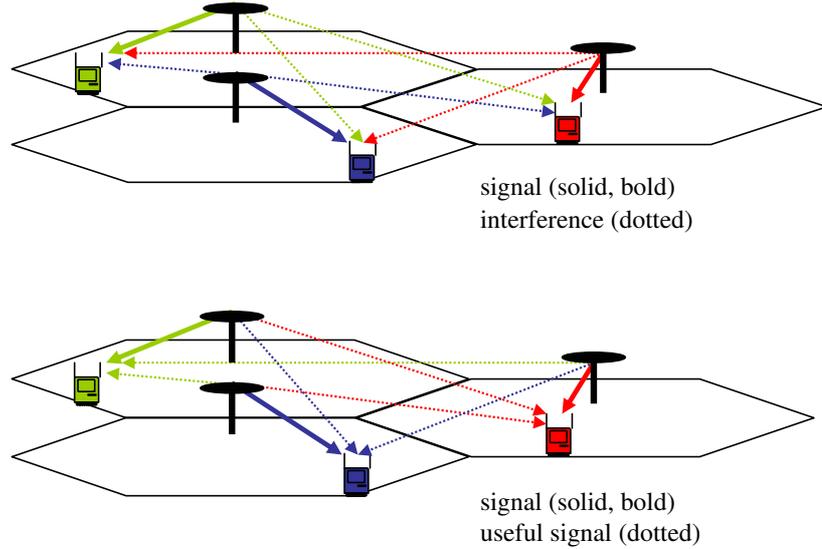
Figure 2.1: Conventional Cellular Networks (top) vs. Coordinated Networks (bottom).

2.1 System Model

In general, the system model for a downlink network with M single antenna base stations and N single antenna mobiles is given by

$$\mathbf{Y} = \mathbf{Hx} + \mathbf{n} \tag{2.1}$$

where $\mathbf{H} = [h_{ij}]_{N \times M}$ denotes the channel matrix with h_{ij} being the complex channel gain between mobile i and base station j, $\mathbf{x} = [x_1, x_2, \ldots, x_M]^T$ denotes the complex antenna outputs, and $\mathbf{n} = [n_1, n_2, \ldots, n_N]^T$ denotes an additive white noise vector with covariance $\sigma^2 \mathbf{I}$. When coordination is employed, all M base stations can act together, and each mobile may receive useful signals from all base stations. Denoting the vector of data symbols by $\mathbf{d} = [d_1, d_2, \ldots, d_N]^T$ where d_i is the ith mobile's complex data symbol, a linear spatial pre-filter matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ is used to map the data symbols to the antenna outputs, i.e. $\mathbf{x} = \mathbf{Ad}$.

Thus, in the case of coordination, the antenna output at the jth base station is a linear combination of N data symbols, i.e., $x_j = \sum_{i=1}^{N} A_{ji} d_i$. For the conventional cellular transmissions, each base station simply transmits the data symbol for the mobile in its own cell coverage, and the linear pre-filter matrix is not necessary. In our analysis, we assume that each base is loaded at most with one user, i.e., $N \leq M$. For example, the system model (2.1) may correspond to the set of mobiles using a particular orthogonal dimension, i.e., a time slot for TDMA, a tone for OFDM, an orthogonal spreading code for CDMA etc., in each base station, and some of the mobiles may be in outage due to their undesirable channel and interference conditions.

In the following sections, we will present different transmission techniques, and explain the methodology for their performance evaluations. The metric for comparison will be the max-min rate achievable subject to per-base power constraints. The max-min rate objective is motivated by the fairness concern, i.e., the need to guarantee a quality of service (QoS) for a large number of users spread over many cells.

2.2   Conventional Cellular Network

The baseline for comparison with the coordinated networks is the conventional cellular systems without network coordination. In this case, each base station transmits signals intended for the user within its cell coverage, and neighboring base transmissions cause inter-cell interference. Given the model (2.1), the antenna output at the ith base antenna is the data symbol for its associated mobile,

i.e., $x_i = d_i$. Denoting the power of the ith data symbol by p_i, the signal-to-interference-plus-noise ratio (SINR) for mobile i is given by

$$\rho_i = \frac{p_i |h_{ii}|^2}{\sum_{j \neq i} p_j |h_{ij}|^2 + \sigma^2}. \tag{2.2}$$

and the corresponding Shannon rate is given by $\log_2(1 + \rho_i)$ bits/symbol/Hz. The following optimization problem formulates the max-min rate problem for an uncoordinated network subject to per-base power constraints:

$$\max \quad r \tag{2.3}$$

$$s.t. \quad \log_2\left(1 + \frac{p_i |h_{ii}|^2}{\sum_{j \neq i} p_j |h_{ij}|^2 + \sigma^2}\right) \geq r \tag{2.3a}$$

$$0 \leq p_i \leq p_{\max}, \quad i = 1, \ldots, N \tag{2.3b}$$

$$r \in \Re^+, p_i \in \Re^+ \tag{2.3c}$$

where $p_{\max}$ is the maximum transmit power of a base station antenna, and r can be interpreted as the minimum rate to be maximized, i.e., $r = \min_i \log_2(1 + \rho_i)$. The max-min rate can be found in an iterative fashion by solving a series of linear feasibility problems. One can initially start with a small target rate r, and can solve a simple linear system of equations to find if there exists a feasible power allocations (either as a centralized LP or using distributed power control algorithms). The target rate is improved further as long as the power constraints are satisfied. We note that channel phase knowledge is not needed for uncoordinated transmissions.

2.3 Coordinated Network by Zero-Forcing Transmission

A simple form of coordination is obtained by zero-forcing transmission where each mobile's signal vectors are projected away from other users. This can be achieved when a pseudo-inverse pre-filter matrix

$$\mathbf{A} = \mathbf{H}^{\dagger}(\mathbf{H}\mathbf{H}^{\dagger})^{-1} \tag{2.4}$$

is used to map the data symbols to the antenna outputs, i.e.,

$$\mathbf{x} = \mathbf{H}^{\dagger}(\mathbf{H}\mathbf{H}^{\dagger})^{-1}\mathbf{d}. \tag{2.5}$$

Notice that each column of the pseudo-inverse matrix is an M-dimensional weight vector mapping a data symbol to M network antennas. In this case, the received signal is given by

$$\mathbf{Y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{H}^{\dagger}(\mathbf{H}\mathbf{H}^{\dagger})^{-1}\mathbf{d} + \mathbf{n} = \mathbf{d} + \mathbf{n} \tag{2.6}$$

and the ith mobile receives $y_i = d_i + n_i$. Thus, the channel has been diagonalized; all network antennas in range can help the transmission of each message, but the message is received only by the intended user with no interference. Given the received signal (2.6), mobile i can achieve the rate $\log_2(1 + p_i/\sigma^2)$ bits/symbol/Hz where

$$p_i = E[|d_i|^2] \tag{2.7}$$

is the symbol power.

To formulate the max-min rate optimization problem, one has to specify per-base power constraints. Notice that each base antenna is subject to an average power constraint given by $E[|x_m|^2] \leq p_{\max}, \quad m = 1, \ldots, M$. These constraints can be transformed into a set of linear constraints in terms of the power of the data symbols p_i, $i = 1, \ldots, N$. Note that base antenna powers are on the diagonals of the following transmit covariance matrix

$$E[\mathbf{x}\mathbf{x}^\dagger] = \mathbf{A}E[\mathbf{d}\mathbf{d}^\dagger]\mathbf{A}^\dagger = \mathbf{A} \begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_N \end{bmatrix} \mathbf{A}^\dagger, \tag{2.8}$$

where $\mathbf{x} = \mathbf{A}\mathbf{d}$, and assuming independent data symbols, the constraints on the diagonals can be expressed in matrix form as

$$\begin{bmatrix} |A_{11}|^2 & \ldots & |A_{1N}|^2 \\ \vdots & & \vdots \\ |A_{M1}|^2 & \ldots & |A_{MN}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq p_{\max}\mathbf{1} \tag{2.9}$$

where $\mathbf{1} = [1, 1, \ldots, 1]^T$ is an M-dimensional column vector of 1s. The problem of maximizing the minimum rate subject to per-base power constraints can be

written follows:

$$\max \quad r \tag{2.10}$$

$$s.t. \quad \log_2\left(1 + \frac{p_i}{\sigma^2}\right) \geq r, \; i = 1, \ldots, N \tag{2.10a}$$

$$\begin{bmatrix} |A_{11}|^2 & \ldots & |A_{1N}|^2 \\ \vdots & & \vdots \\ |A_{M1}|^2 & \ldots & |A_{MN}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq p_{\max}\mathbf{1} \tag{2.10b}$$

$$r \in \Re^+, p_i \in \Re^+ \tag{2.10c}$$

In the above problem definition, r can be interpreted as the minimum rate to be maximized, i.e., $r = \min_i \log_2\left(1 + p_i/\sigma^2\right)$. Notice that each constraint in (2.10a) defines a convex set, which is the region underneath a concave logarithm function intersected with the region above the hyperplane defined by r. Moreover, the power constraints are linear, and therefore the max-min rate problem becomes a convex optimization problem [30]. Note that the max-min rate cannot be increased any further when any of the M constraints in (2.10b) becomes active, i.e., when $p_{\max}$ is attained at one of the base stations.

The above convex optimization formulation holds for any pre-filter matrix **A**. For the particular choice of the pseudo-inverse matrix, the max-min rate can actually be obtained in closed-form by enforcing users to have equal rate assignments, i.e., by solving the maximum common rate problem. Notice that any mobile achieving a larger rate than the max-min rate can give up some of its transmit power to make its rate equal to the max-min rate, and this would

actually help the power constraints. Under the equal rate constraint, each user's received power would be the same p. In this case, the covariance matrix simply becomes $E[\mathbf{x}\mathbf{x}^\dagger] = p\mathbf{A}\mathbf{A}^\dagger$. Similarly, the base antenna power constraints on the diagonals of the covariance matrix can be written as

$$p[\mathbf{A}\mathbf{A}^\dagger]_{(j,j)} = p\sum_{i=1}^{N}|A_{ji}|^2 \leq p_{max}, \quad j = 1, \ldots, M, \tag{2.11}$$

where $[.]_{(j,j)}$ denotes the jth diagonal element of the matrix. Maximizing the common rate is equivalent to maximizing the common received power p, which occurs at $p^* = p_{\max}/\max_j \sum_{i=1}^{N}|A_{ji}|^2$. It follows that the maximum common rate is given by $r^* = \log_2(1 + p^*/\sigma^2)$, which is also equivalent to the max-min rate.

We mention that an MMSE pre-filter is also possible [31, 32], which should have a performance beyond that of the zero-forcing scheme. However, as we will see in the numerical examples section, the typical cellular network setup that we are interested in will have a relatively high SNR, and therefore we expect the performance improvement due to MMSE pre-filter to be small. Also, complete channel knowledge (including phase and magnitude) is needed for the zero-forcing technique, as well as for the zero-forcing dirty paper coding technique of the next section.

2.4 Coordinated Network by Zero-Forcing Dirty Paper Coding

Zero-forcing comes with a penalty in the sense that a mobile's transmissions are constrained to a smaller subspace after projecting away from the other mobiles'

channels. An improved form of coordination is obtained when a limited form of zero-forcing is combined with dirty paper coding [20, 29]. Dirty paper coding can be employed when the interference is known causally at the transmitter, which is possible for downlink transmissions. First, mobiles are indexed according to some order $\pi = [\pi(1), \pi(2), \ldots, \pi(N)]$. By dirty paper encoding, a mobile can be made invisible interference-wise to other mobiles with higher indices in the user ordering [27]. When dirty paper coding is combined with a limited form of zero forcing, the visible interference is nulled out due to the zero forcing.

The combined zero-forcing dirty paper coding scheme assumes a linear pre-filter matrix $\mathbf{A}$ obtained through LQ decomposition of the channel matrix $\mathbf{H}$, and is given by $\mathbf{A} = \mathbf{Q}^\dagger$ where $\mathbf{H} = \mathbf{LQ}$, $\mathbf{L} \in \mathbb{C}^{\mathbb{N} \times \mathbb{N}}$ is lower triangular, and $\mathbf{Q} \in \mathbb{C}^{\mathbb{N} \times \mathbb{M}}$ is a unitary matrix with $\mathbf{QQ}^\dagger = \mathbf{I}$ [20, 29]. We should note that while this scheme is shown to be sum-rate optimal at high SNR regime subject to a sum power constraint, the same scheme is not claimed to be optimal for the max-min rate problem subject to per-base power constraints. The corresponding system model (2.1) is given by

$$\mathbf{Y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HAd} + \mathbf{n} = \mathbf{LQQ}^\dagger\mathbf{d} + \mathbf{n} = \mathbf{Ld} + \mathbf{n}. \tag{2.12}$$

Therefore the ith user receives the signal

$$y_i = L_{ii}d_i + \sum_{j<i} L_{ij}d_j + n_i. \tag{2.13}$$

The particular choice of spatial filter matrix $\mathbf{A} = \mathbf{Q}^\dagger$ nulls the interference from users with indices $j > i$, and the remaining part of the interference due to

users with indices $j < i$ (causally known to the transmitter) is taken care of by the dirty paper coding [27]. In this case, the ith user experiences a single user channel, and can achieve the following rate

$$r_i = \log_2 \left(1 + \frac{|L_{ii}|^2 p_i}{\sigma^2} \right) \tag{2.14}$$

where $p_i = E[|d_i|^2]$. Each base station antenna is subject to an average power constraint given by

$$E[|x_m|^2] \leq P_{\max}, \quad m = 1, \ldots, M. \tag{2.15}$$

Next, the base antenna power constraints are transformed into a set of linear constraints in terms of power of the data symbols p_i, $i = 1, \ldots, N$. Note that base antenna powers are on the diagonals of the following transmit covariance matrix

$$E[\mathbf{xx}^\dagger] = \mathbf{Q}^\dagger E[\mathbf{dd}^\dagger]\mathbf{Q} = \mathbf{Q}^\dagger \begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_N \end{bmatrix} \mathbf{Q}. \tag{2.16}$$

The constraints on the diagonals can be expressed in matrix form as

$$\begin{bmatrix} |Q_{11}|^2 & \cdots & |Q_{N1}|^2 \\ \vdots & & \vdots \\ |Q_{1M}|^2 & \cdots & |Q_{NM}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq p_{\max} \mathbf{1} \tag{2.17}$$

where $\mathbf{1} = [1, 1, \ldots, 1]^T$ is an N-dimensional column vector of 1s. The problem

Input : $\mathbf{h}_i = [h_{i1}, h_{i2}, \ldots, h_{iM}]$, $i = 1, 2, \ldots, N$ (channel responses).
Output : $\pi = [\pi(1), \pi(2), \ldots, \pi(N)]$, ($\pi(i)$ projects away from $\pi(j)$ where $i > j$).
Initialization : $\mathbf{S} = \{1, 2, \ldots, N\}$.
for i = 1 : N
   $\pi(i) = \arg\min_{i \in S} |\mathbf{h}_i|^2$,
   $\mathbf{S} = \mathbf{S} - \pi(i)$,
   $\mathbf{e}_{\pi(i)} = \frac{\mathbf{h}_{\pi(i)}}{|\mathbf{h}_{\pi(i)}|}$,
   $\mathbf{h}_j = \mathbf{h}_j - (\mathbf{h}_j^\dagger \mathbf{e}_{\pi(i)})\mathbf{e}_{\pi(i)}$, $\forall j \in \mathbf{S}$.
end

Figure 2.2: The heuristic user-ordering algorithm.

of maximizing the minimum rate subject to per-base power constraints can be written follows:

$$\max \quad r \tag{2.18}$$

$$s.t. \quad \log_2\left(1 + \frac{|L_{ii}|^2 p_i}{\sigma^2}\right) \geq r, \ i = 1, \ldots, N \tag{2.18a}$$

$$\begin{bmatrix} |Q_{11}|^2 & \cdots & |Q_{N1}|^2 \\ \vdots & & \vdots \\ |Q_{1M}|^2 & \cdots & |Q_{NM}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq p_{\max}\mathbf{1} \tag{2.18b}$$

$$r \in \Re^+, p_i \in \Re^+ \tag{2.18c}$$

In the above problem definition, r can be interpreted as the minimum rate to be maximized, i.e. $r = \min_i \log_2\left(1 + |L_{ii}|^2 p_i/\sigma^2\right)$. Similar to the zero-forcing scheme, the base antenna power constraints are linear, the rate functions are concave in symbol powers and therefore (2.18a) defines a convex set. As in the zero-forcing, the max-min rate optimization becomes a convex programming problem [30].

The above combined zero-forcing dirty paper coding scheme assumes a user ordering. Notice that when a particular user i's transmission becomes invisible to some others due to dirty paper encoding, those users who do not receive interference from user i do not have to project away from this particular user. Moreover, dirty paper encoding does not induce any power penalty [27], and therefore comes for free (except the complexity of its practical implementation). On the other hand, zero-forcing incurs a penalty in the sense that a user projecting its signals away from others would give up some part of its signal space. Since the objective is fairness, we propose a heuristic user ordering based on the idea that the price should be paid by the users with good channels. In this case, a user with a good channel "owes a favor" to users with relatively bad channels, and therefore projects away form those disadvantageous users. In turn, the users with bad channels "pay the favor back" by being invisible via the dirty paper encoding. The heuristic user ordering algorithm is shown in Figure 2.2. We will see in the numerical examples section that the algorithm performs quite well in our cellular network setup.

2.5   Performance Limits of Network Coordination

Thus far, we have considered coordination techniques that involve practical linear filtering operations, and thus are analytically tractable. In this section, our objective is to study the ultimate performance limits of network coordination, and see how far the techniques of the previous sections are from the ultimate limits.

The system model described in (2.1) is an example of a non-degraded Gaussian broadcast channel with M geographically separated but perfectly cooperating

transmitters, and N non-cooperating receivers. Finding the capacity region for such channels has been a major research challenge for a long time. Recently, it has been shown that subject to sum power constraint $E[|\mathbf{x}|^2] \leq p_{\max}$, the optimal scheme achieving the boundary of the capacity region involves dirty paper coding [19,20], where the transmitter encodes N data symbols in an order given by $[\pi(1), \ldots, \pi(N)]$ such that a user with index $\pi(i)$ does not suffer any interference from users with lower indices, i.e., $\pi(j)$ with $j < i$. Moreover, the optimal policy may involve time-sharing, as the dirty paper rate functions are in general non-concave in transmit covariances [18].

Notice that, unique to our multiple base coordinated network model, we have a set of constraints and assumptions that adds a new dimension to the classical broadcast channel problems. First, instead of a sum-power constraint, we are concerned with per-base (or per-antenna) power constraints as the transmit power cannot be transferred from one base station to the other. Second, we have a large number of users spread over many cells, and therefore we would like to guarantee a quality of service (QoS) for each user. Hence, our objective is to maximize the minimum rate in the network, instead of a more common sum-rate objective. The classical approach to solve this problem would be to characterize the capacity region with per-base power constraints, and then determine the policy maximizing the minimum of all rates. The difficulty is that even with the duality results [18], it is computationally hard to derive policies, i.e., an encoding order and corresponding transmit covariances etc., achieving a specific point (max-min point in our case) on the broadcast capacity region. Moreover, when the optimal policy involves time-sharing, one may need to specify time-sharing combinations as well. This becomes computationally infeasible for large

systems as one needs to search over all possible user combinations and encoding orders. Therefore, our approach is to develop tight upper bounds on the max-min rate. These bounds not only help to avoid a huge computational burden, but also they give insights into the ultimate performance of the system.

2.5.1 Upper Bounds on the Max-Min Rate of a Coordinated Network

In this section, we derive upper bounds on the max-min rate achievable in a coordinated network. From the first upper bound to the last, each upper bound improves the tightness of the previous upper bounds at the expense of an increased computational complexity. The common context of all the derivations is a configuration that has no user suffering interference. Although, at the first sight this may seem to be too drastic a departure, we will establish its usefulness. Recall that the optimum transmission scheme involves some form of dirty paper coding, with which users causing significant interference can be made invisible to others. Because of the no-interference assumption, each user enjoys a single user channel with M transmit antennas. However, multiple users are coupled through power sharing, i.e., each of them competes for a portion of the total available power.

The interference-free assumption will enable us to work with concave rate functions in user covariances so that the standard convex optimization techniques can be easily employed. On the other hand, for any transmission policy, the set of rates calculated based on no interference will be an upper bound on the actual user rates. Importantly, the bounds are valid even when the optimal policy involves

time sharing. The reason is that, without interference, we are left with single user channels in which it is better not to do time sharing due to concavity of the logarithm function. In other words, time sharing rates, which are already upper bounded by ignoring the interference, will be further bounded when we restrict our attention to simpler transmission policies. Also, when users do not interfere with each other, the issue of dirty paper encoding order becomes irrelevant. Therefore, the bounds are valid for all encoding orders when the optimal transmission policy involves the dirty paper coding. In short, by being reasonably optimistic, we greatly simplify a complex problem as we do not have to deal with all permutations of dirty paper encoding order, or time-sharing combinations.

Single-User Bound

Our first upper bound is based on a single-user argument that the network does its best for one of the users by granting him exclusive use of all network resources. In other words, the network operator lets all M base station antennas serve the user at full power, and produce a coherently reinforced signal at the user's antenna. Let us consider user i, and assume that the antenna at the mth base transmits a voltage $v_{im} = \sqrt{p_{im}}e^{j\phi_{im}}$ to this user, where p_{im} is the transmit power and ϕ_{im} is the phase. In this case, the transmit voltage vector is given by $\mathbf{v}_i = [v_{i1}, v_{i2}, \ldots, v_{iM}]^T$. Given the ith user's complex vector channel $\mathbf{h}_i = [h_{i1}, h_{i2}, \ldots, h_{iM}]^T$, the maximum received signal power is obtained by having phase matched transmit voltages, and coherent voltage addition at the receiver side. In this case, the ith user achieves the single user rate given by

$$r_i = \log_2\left(1 + \frac{\left(\sqrt{p_{i1}}|h_{i1}| + \cdots + \sqrt{p_{iM}}|h_{iM}|\right)^2}{\sigma^2}\right). \tag{2.19}$$

Given that all base antennas transmit at full power, i.e. $p_{im} = p_{\max}$ for $m = 1, \ldots, M$, (2.19) can be written as $r_i = \log_2(1 + p_{\max}\|\mathbf{h}_i\|_1^2/\sigma^2)$, where $\|\mathbf{h}_i\|_1$ denotes L_1 norm of the vector channel $\mathbf{h}_i$. The first upper bound is based on the fact that max-min rate of N users cannot exceed the single user Shannon rate of any user (2.19). Therefore, denoting the max-min rate by r, it follows that

$$r \leq \min_{i \in \{1, \ldots, N\}} \log_2\left(1 + \frac{p_{\max}\|\mathbf{h}_i\|_1^2}{\sigma^2}\right) \tag{2.20}$$

Multi-User Bound with Sum-Power Constraint

We can improve the previous upper bound by considering presence of multiple users in the system. Similar to the first bound, we assume that users do not suffer any interference, but they are coupled through power sharing. Given the ith user's transmit voltage vector $\mathbf{v}_i = [v_{i1}, v_{i2}, \ldots, v_{iM}]^T$, and the channel $\mathbf{h}_i = [h_{i1}, h_{i2}, \ldots, h_{iM}]^T$, the received signal power is $|\mathbf{v}_i \cdot \mathbf{h}_i|^2$, which is by coherent voltage addition. In this case, each user achieves the following rate

$$r_i = \log_2\left(1 + \frac{|\mathbf{v}_i \cdot \mathbf{h}_i|^2}{I_i + \sigma^2}\right) \quad i = 1, \ldots, N \tag{2.21}$$

where I_i denotes the total interference power user i would actually suffer. An upper bound on the max-min rate r can be written as

$$r \leq r_i = \log_2\left(1 + \frac{|\mathbf{v}_i \cdot \mathbf{h}_i|^2}{I_i + \sigma^2}\right) \overset{(a)}{\leq} \log_2\left(1 + \frac{|\mathbf{v}_i \cdot \mathbf{h}_i|^2}{\sigma^2}\right)$$

$$\overset{(b)}{\leq} \log_2\left(1 + \frac{\|\mathbf{v}_i\|^2\|\mathbf{h}_i\|^2}{\sigma^2}\right), \tag{2.22}$$

where (a) is obtained by ignoring the interference term I_i, and (b) is due to the Cauchy-Schwartz inequality. In the inequality (2.22), the magnitude squared voltage term $\|\mathbf{v}_i\|^2$ represents the total power allocated to the ith user across all base stations; we denote it by $P_i = \|\mathbf{v}_i\|^2$. It follows from (b) that

$$\frac{(2^r - 1)\sigma^2}{|\mathbf{h}_i\|^2} \leq P_i. \tag{2.23}$$

The sum-power of all users must be smaller than the total available power in the network implying

$$\sum_{i=1}^{N} P_i \leq M p_{\max}. \tag{2.24}$$

This is a relaxation to per-base power constraints, which will be analyzed in the following subsection, and will lead to a better upper bound. It follows that

$$(2^r - 1) \sum_{i=1}^{N} \frac{\sigma^2}{\|\mathbf{h}_i\|^2} \leq \sum_{i=1}^{N} P_i \leq M p_{\max} \tag{2.25}$$

The second upper bound follows easily from (2.25):

$$r \leq \log_2 \left(1 + \frac{M p_{\max}}{\sum_{i=1}^{N} \frac{\sigma^2}{\|\mathbf{h}_i\|^2}} \right). \tag{2.26}$$

Multi-User Bound with Per-Base Power Constraint

A tighter upper bound is obtained by considering per-base power constraints, instead of the previous section's constraint relaxation in the form of a sum-power constraint. Note in the previous section that P_i represents the sum of contributions of each antenna power for user i, i.e., $P_i = \sum_{m=1}^{M} p_{im}$. Here, we will choose,

for each user, transmit powers at each base so as to satisfy per base power constraints given by

$$\sum_{i=1}^{N} p_{im} \leq p_{\max}, \ m = 1, \ldots, M \tag{2.27}$$

Consider the following optimization problem:

$$\max \ r^{'} \tag{2.28}$$

$$s.t. \ (2^{r^{'}} - 1)\frac{\sigma^2}{\|\mathbf{h}_i\|^2} - \sum_{m=1}^{M} p_{im} \leq 0, \ i = 1, \ldots, N \tag{2.28a}$$

$$\sum_{i=1}^{N} p_{im} \leq p_{\max}, \ m = 1, \ldots, M \tag{2.28b}$$

$$r^{'} \in \Re^{+}, p_{im} \in \Re^{+} \tag{2.28c}$$

where the constraint (2.28a) expresses the bound (2.22b) using the definition of P_i in this section. The solution to the optimization problem, say r^*, is an upper bound on the max-min rate, i.e., $r \leq r^*$. The argument is that, for $r^{'} = r$, the optimum power, and phase allocations maximizing the minimum rate, i.e., achieving r, would satisfy (2.22b), and therefore (2.28a) as well. Moreover, the optimum power allocations are feasible, and therefore the constraint (2.28b) is also satisfied. Thus, we would at least obtain r by solving the optimization problem (2.28), and possibly overshoot r due to the constraint relaxation (2.28a). The optimization problem can easily be solved as a convex program, or as a linear program (LP) by defining an auxiliary variable $t = 2^{r^{'}} - 1$, and maximizing t subject to the constraints (2.28a)-(2.28c). Since t is an increasing function of $r^{'}$, maximizing t would maximize $r^{'}$ as well. If the solution occurs at t^*, then r^* simply follows as $r^* = \log_2 (1 + t^*)$.

Multiple User Bound

In this section, we solve the max-min rate problem subject to per-base power constraints when each user's rate is calculated as if no interference is present in the system. Note that the received signal power is maximized when the transmitted signals are in-phase with the channel, and add up coherently at the receiver, i.e., $|\mathbf{v}_i \cdot \mathbf{h}_i|^2 \leq \left(\sqrt{p_{i1}}|h_{i1}| + \sqrt{p_{i2}}|h_{i2}| + \cdots + \sqrt{p_{iM}}|h_{iM}|\right)^2$. It follows that the optimum power allocations achieving r would satisfy

$$r \leq \log_2\left(1 + \frac{\left(\sqrt{p_{i1}}|h_{i1}| + \cdots + \sqrt{p_{iM}}|h_{iM}|\right)^2}{\sigma^2}\right) \tag{2.29}$$

We obtain our last upper bound by solving the following optimization problem:

$$\max \quad r' \tag{2.30}$$

$$s.t. \quad \log_2\left(1 + \frac{\left(\sqrt{p_{i1}}|h_{i1}| + \cdots + \sqrt{p_{iM}}|h_{iM}|\right)^2}{\sigma^2}\right) \geq r' \tag{2.30a}$$

$$\sum_{i=1}^{N} p_{im} \leq p_{\max}, \ m = 1, \ldots, M \tag{2.30b}$$

$$r' \in \Re^+, p_{im} \in \Re^+, i = 1, \ldots, N. \tag{2.30c}$$

The optimization problem can be solved using convex optimization techniques. We find it useful to make the following change of variables to better observe the convex structure of the problem. First, we rewrite (2.30a) as

$$\sqrt{(2^{r'} - 1)\sigma^2} - \left(\sqrt{p_{i1}}|h_{i1}| + \cdots + \sqrt{p_{iM}}|h_{iM}|\right) \leq 0. \tag{2.31}$$

Then, we define an auxiliary variable $t = \sqrt{2^{r'} - 1}$, and maximize t within

the constraints (2.30a)-(2.30c). Since t is an increasing function of r', maximizing t would maximize r' as well. Also, we define $y_{im} = \sqrt{p_{im}}$, and solve the following optimization problem:

$$\max \quad t \tag{2.32}$$

$$s.t. \quad t\sigma - (y_{i1}|h_{i1}| + \cdots + y_{iM}|h_{iM}|) \leq 0, \tag{2.32a}$$

$$\sum_{i=1}^{N} y_{im}^2 \leq p_{\max}, \ m = 1, \ldots, M, \tag{2.32b}$$

$$t \in \Re^+, y_{im} \in \Re^+, i = 1, \ldots, N. \tag{2.32c}$$

The first set of constraints (2.32a) is linear in optimization variables, and therefore defines a convex region. The second set of constraints (2.32b) defines an intersection of spheres, which is also convex. Thus, the problem is a convex optimization problem with linear and non-linear constraints. The solution of such a convex optimization problem is well-known [30].

2.6   Coordination vs. Sectorization

Another effective method for downlink interference mitigation is the use of sectorized antennas at the base station. Sectorization can be interpreted as spatial/fixed beamforming where directional narrow-beam antennas are used to separate interfering users spatially. On the other hand, the coordination provides user-specific beamforming, and therefore is more flexible than sectorization. For example, in our numerical examples, we will present results for a 36 base network with 6 sectors in each cell, and one user per antenna. In this case, when the mobiles are

uniformly distributed across the cell area, each sector might have only one user, and the sectorization can effectively mitigate the intra-cell interference. On the other hand, if there are more than one user in a sector area, sectorized antennas cannot separate these users while the coordinated antennas can effectively separate them through zero-forcing or dirty paper coding. We will see in our numerical examples that while the sectorization by itself helps to reduce inter-cell interference, employing network coordination with sectorization gives substantial additional improvements in system capacity.

2.7 Limited Coordination: Partial Channel Information

The coordinated transmission methods require the channel information at the base station. This could be achieved in practice by using channel estimates on the uplink in the time division duplexing (TDD) mode, or by a channel feedback from the mobiles in the frequency division duplexing mode (FDD). However, there will be channel estimation errors and feedback delays, and therefore availability of an accurate channel information is a challenge. Here, we consider a limited coordination scenario where only partial information is available about the channel matrix $\mathbf{H}$. In particular, we assume that, for each mobile antenna, only the channels of those base antennas with strong enough links can be reliably estimated. Figure 2.4 shows a sample scenario where only the channel information for the two base station antennas with the strongest links is available for each mobile antenna. Here, the partial channel matrix is denoted by $\mathbf{H}_p$. As a case study, we will consider the zero-forcing coordination where the precoding matrix is chosen to be $\mathbf{A} = \mathbf{H}_p^{\dagger}(\mathbf{H}_p \mathbf{H}_p^{\dagger})^{-1}$ based on the partial channel matrix. In this case, the

channel output becomes

$$\mathbf{Y} = \mathbf{Hx} + \mathbf{n} = \mathbf{HH_p^\dagger}(\mathbf{H_p H_p^\dagger})^{-1}\mathbf{d} + \mathbf{n}. \tag{2.33}$$

Thus, the received signal for the ith mobile antenna includes both the data signal and the interference. Notice that, in the signal model (2.33), the matrix $\mathbf{HH_p^\dagger}(\mathbf{H_p H_p^\dagger})^{-1}$ defines an effective channel matrix whose input is the data vector $\mathbf{d}$. In this case, denoting the effective channel matrix by $\mathbf{W} = \mathbf{HH_p^\dagger}(\mathbf{H_p H_p^\dagger})^{-1}$, and the power of the ith data symbol d_i by p_i, the SINR for mobile i is given by

$$\rho_i = \frac{p_i |w_{ii}|^2}{\sum_{j \neq i} p_j |w_{ij}|^2 + \sigma^2}. \tag{2.34}$$

where w_{ij} is the (i,j)th entry of the effective channel matrix $\mathbf{W}$, and the corresponding Shannon rate is given by $\log_2(1 + \rho_i)$ bits/symbol/Hz. In this case, the following optimization problem formulates the max-min rate problem for the partial coordination case:

$$\max \quad r \tag{2.35}$$

$$s.t. \quad \log_2\left(1 + \frac{p_i |w_{ii}|^2}{\sum_{j \neq i} p_j |w_{ij}|^2 + \sigma^2}\right) \geq r \tag{2.35a}$$

$$\begin{bmatrix} |A_{11}|^2 & \cdots & |A_{1N}|^2 \\ \vdots & & \vdots \\ |A_{M1}|^2 & \cdots & |A_{MN}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq p_{\max}\mathbf{1} \tag{2.35b}$$

$$r \in \Re^+, p_i \in \Re^+, i = 1, 2, \ldots, N \tag{2.35c}$$
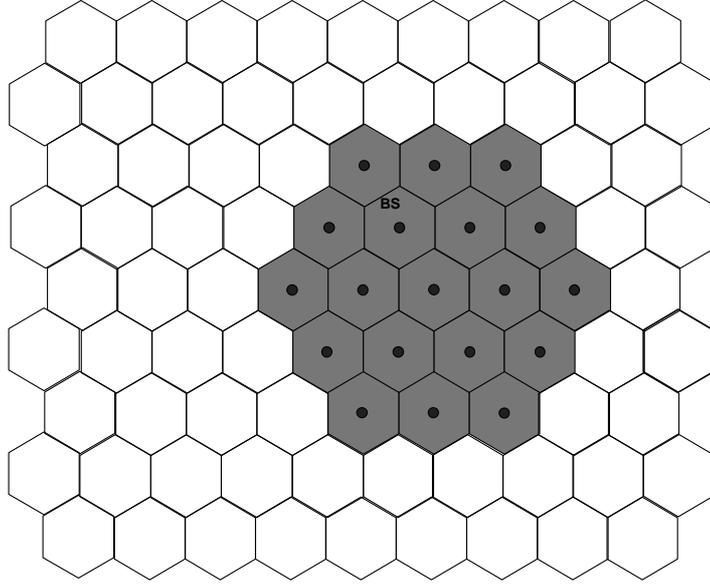
Figure 2.3: Layout for a 64 cells cellular network with a base at the center of each hexagon. A cell with two concentric hexagonal rings of surrounding cells is highlighted.

where the power constraint (2.35b) follows from (2.10b) with the precoding matrix given by $\mathbf{A} = \mathbf{H}_p^\dagger (\mathbf{H}_p \mathbf{H}_p^\dagger)^{-1}$ based on the partial channel.

As in Section 2.2, the max-min rate can be found in an iterative fashion by solving a series of linear feasibility problems. Notice that the power constraints (2.35b) are linear. In this case, one can start with a small target rate r, and can solve a linear system of equations to find if there exists a feasible power allocations. The target rate is improved further as long as the power constraints are satisfied.
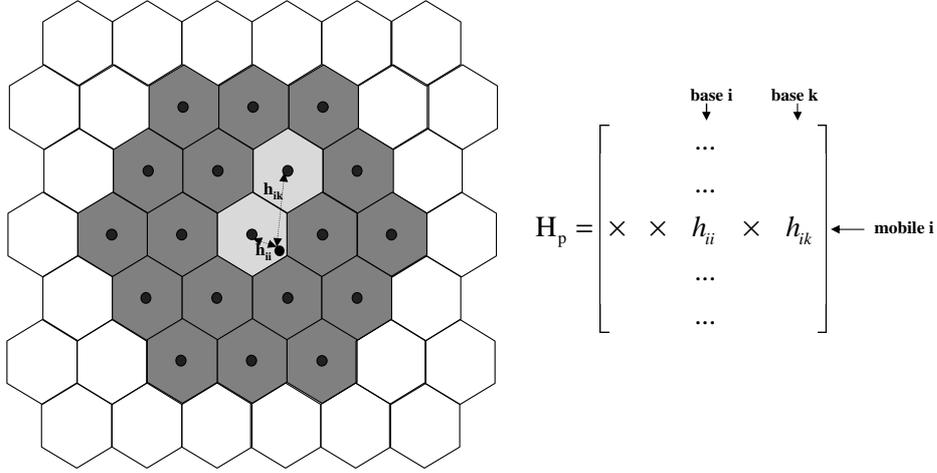
Figure 2.4: For each mobile antenna, only the channel information for the two base station antennas with the strongest links is available.

2.8 Performance Evaluation

In this section, we will present results for a 64 base cellular network. We assume that each base station is located at the center of an hexagonal cell shown in Figure 2.3. A flat torus is formed to avoid the boundary effects. From the base antenna i to the mobile antenna j at a distance of d meters, the channel propagation characteristic is a triple product, i.e., $h_{ij} \propto \alpha_{ij} s_{ij}^{1/2} d^{-\epsilon/2}$ where in addition to path-loss with a propagation exponent of $\epsilon = 3.8$, the channels are affected by log-normal shadowing s_{ij} with 0 mean and 8 dB standard deviation, and 0 mean unit variance complex Gaussian component α_{ij} (Rayleigh piece). We assume a maximum base power of 10 W, a mean power loss of 134 dB at the
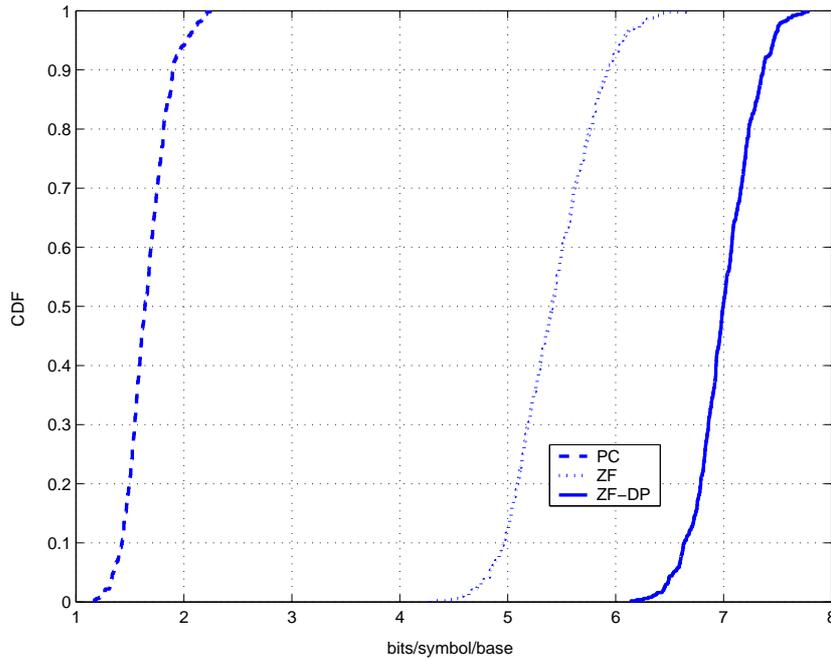
Figure 2.5: Empirical max-min rate CDFs for a 64 base network with 18 dB reference SNR at the cell border.

reference distance of 1.6 km from the base (distance from the base station to any corner of the hexagon), a receiver noise figure of 5 dB, a vertical antenna gain of 10.3 dBi, a channel bandwidth of 5 MHz, and a receiver temperature of 300°K. Thus, accounting only for path loss and ignoring shadowing and Raleigh fading, SNR at the reference distance is 18 dB. However, because of inter-cell interference, the SINR can be much smaller than 18 dB (can even be negative) in our cellular network setup. In Figure 2.3, two rings around a cell are highlighted to emphasize that a mobile in the center cell can receive significant signal only from those colored cells because of the exponential decay in the signal power [35, 36]. In other words, those colored cells may potentially interfere with the mobile in the center cell, while at the same time they can be coordinated to help overcome the interference.

Figure 2.5-2.7 show empirical CDFs of the max-min rate for conventional
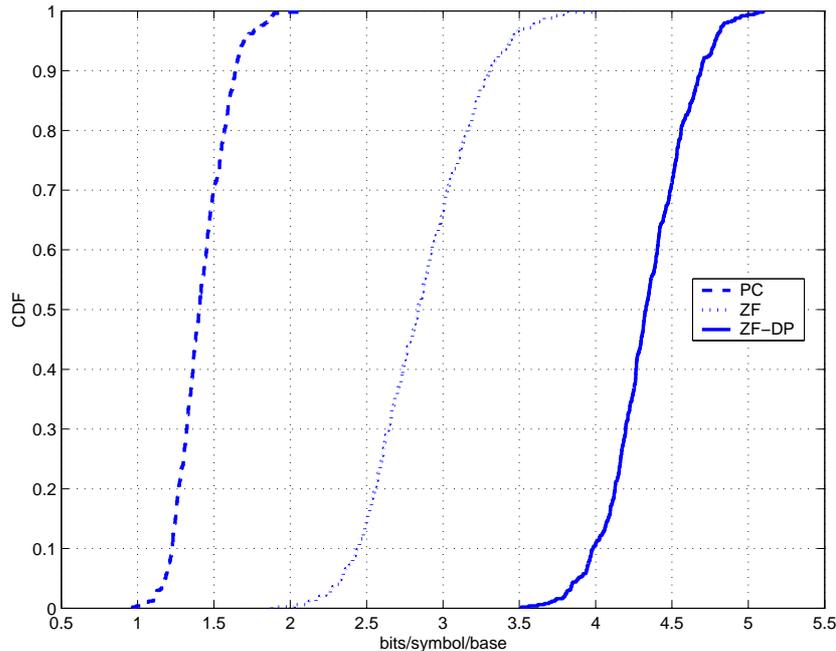
Figure 2.6: Empirical max-min rate CDFs for a 64 base network with 9 dB reference SNR at the cell border.

power control, and coordinated (zero forcing and zero-forcing dirty paper coding) transmission techniques for different reference SNR values. The CDFs are obtained using max-min rates of 500 network instances. At each network instance, a set of uniformly distributed mobile locations and random channel realizations (including path-loss, shadow and Rayleigh fading) are generated. Each base is loaded with one user, and 10% of the users are allowed to be in outage. Many heuristic methods are possible to assign mobiles to the outage state. For example, one may discard those mobiles which have the weakest links to their associated bases, or have the smallest SINRs when all base antennas transmit at full power. In our experiments, we used an iterative user discard rule in which the mobile that makes the power constraint to become active (causing the max-min rate not to be improved any further) is discarded at each iteration. Figure 2.5 shows that, based on the median points of the rate CDFs, the coordination improves the max-min
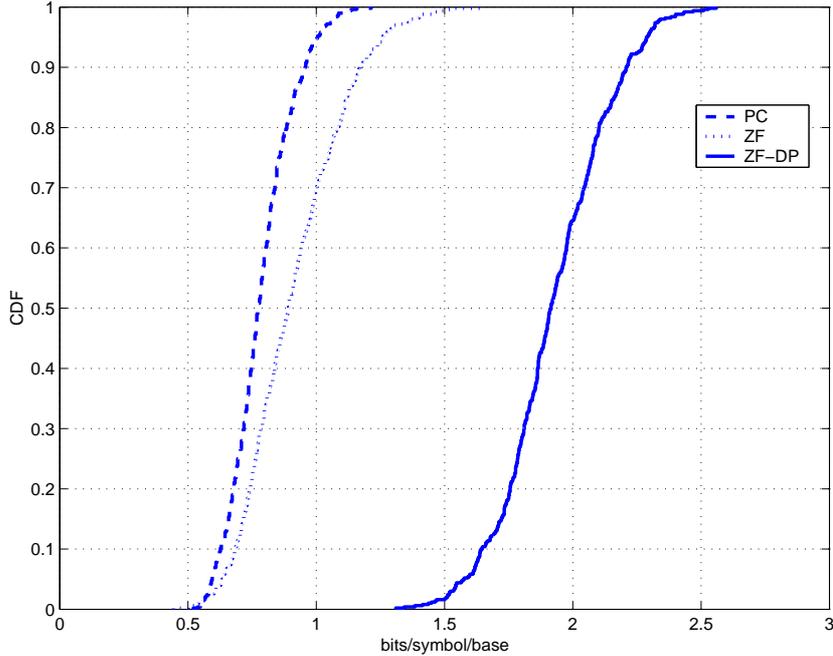
Figure 2.7: Empirical max-min rate CDFs for a 64 base network with 0 dB reference SNR at the cell border.

rate by about a factor-of-3 when the zero-forcing coordination is employed, and by about a factor-of-5 when the combined zero-forcing dirty paper coding scheme is used. The reported numbers in Figure 2.5 are based on a typical cellular wireless channel, and therefore should be interpreted as the potential gains in practical wireless sytems. To show the dependence of our results to the reference SNR, we also plotted the rate CDFs for 0 dB and 9 dB reference SNRs in Figure 2.6 and Figure 2.7 respectively. From Figure 2.7, we observe that the zero-forcing, which performs well at high SNRs, performs far from optimal at very low SNRs. On the other hand, the combined zero-forcing dirty paper coding scheme has an advantage over the conventional scheme at all SNR ranges.

In Figure 2.8, we compare the max-min rate of our best coordination technique with the upper bounds. The figure shows that the rate achievable by the linear zero-forcing beamforming combined with dirty paper coding is close (≈ 1
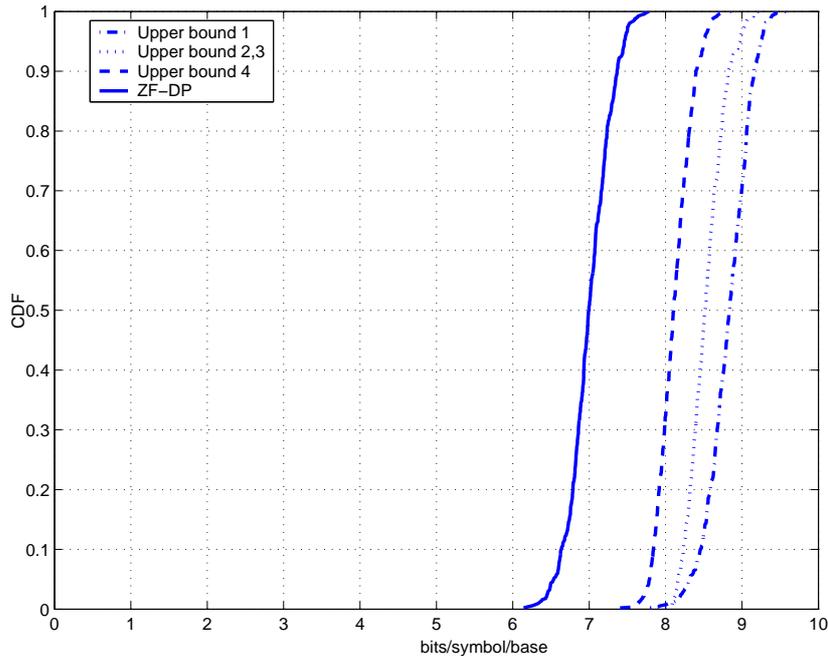
Figure 2.8: Performance limits of network coordination: Upper bounds for a 64 base network with 18 dB reference SNR at the cell border.

bit/symbol/base) to the sharpest upper bound obtained. Hence, the coordination techniques presented achieve the desired performance.

In Figure 2.9, the sectorization is compared with the coordination. Here, we have a 36 base network with 6 omni or sector antennas in each cell, and one mobile per antenna. The reference SNR at the cell border is 18 dB, and sectorized antennas have an additional $10 \log_{10} 6 \approx 7.8$ dB power gain. The figure shows that, without coordination, 6 mobiles in a cell interfere with each other significantly when omni antennas are used, and therefore the max-min rate is very small. The use of sectorized antennas improves the max-min rate as it helps reducing the intra-cell interference, and provides additional power gain. Combining the zero-forcing coordination with sectorization gives substantial additional improvements in system capacity. Notice that, the power gain of the sectorized antennas adds an additional ≈ 3 bits/symbol/base to the max-min rate at high
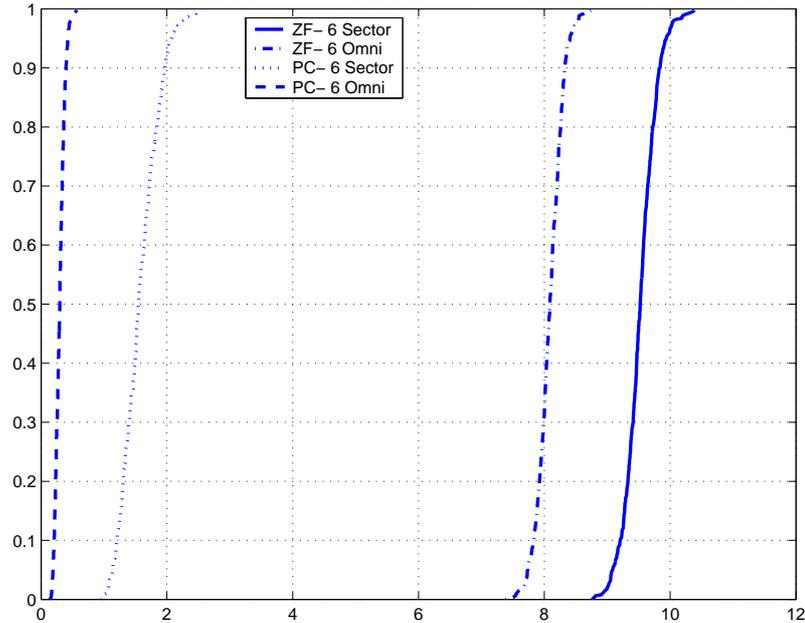
Figure 2.9: Empirical max-min rate CDFs for a 36 base network with 6 sectorized or omni transmit antennas, and with 18 dB reference SNR at the cell border.

SNRs. Without the power gain, the coordinated sectorized antennas perform worse than the coordinated omni antennas. The reason is that the omni antennas provide mobile-specific beamforming instead of fixed/spatial beamforming, and have more degrees of freedom for coordination.

In Figure 2.10, the effect of partial channel information is presented in the context of zero-forcing coordination. We observe that more than half of the throughput achievable with the full channel information can be obtained when only the channels of $3 - 4$ base antennas with the strongest links are available.

2.9 Chapter Summary and Conclusion

Coordinating base antenna transmissions to mitigate inter-cell interference is a promising idea suggesting large capacity improvements over the conventional cellular networks. Practical concerns regarding the coordination are the need for
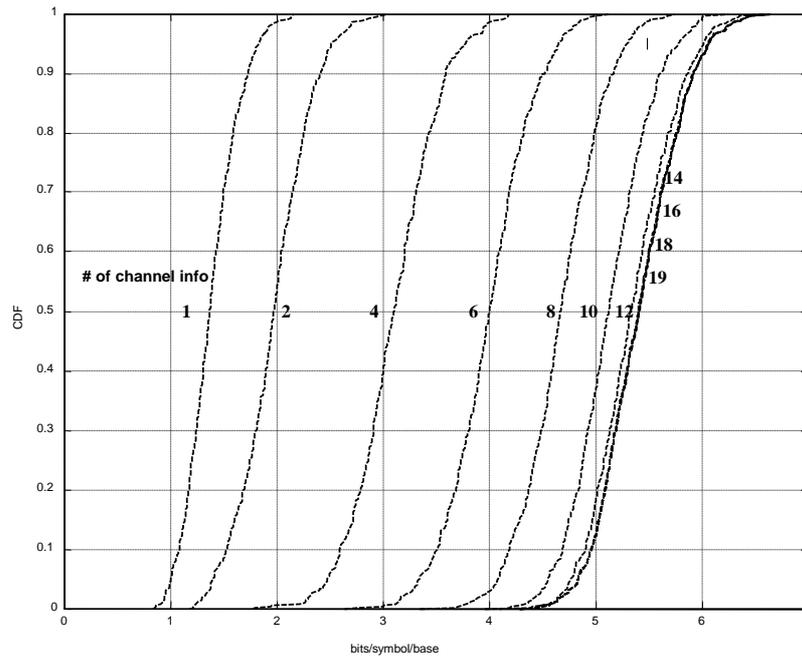
Figure 2.10: The effect of limited channel information: Empirical max-min rate CDFs for a 64 base network with ZF transmission, and with 18 dB reference SNR at the cell border.

channel knowledge, and a backbone enabling communication between the base stations, both not far from reach technically. Thus, the network coordination should be given serious consideration for future wireless networks.

Chapter 3

Multiple Antenna Network Coordination

In the previous chapter, we presented network coordination as a means to provide high spectral efficiency in cellular downlink systems with single antenna units. Here, we will study the impact of network coordination on multiple antenna systems. We will show that the coordinated transmissions are especially effective when the base stations and the mobiles are equipped with multiple antennas. Without coordination, the link qualities can be very poor because of inter-cell interference. In this case, the network does not benefit significantly from multiple antennas since the improvement in rate due to the antenna units is small at low SNRs. When the coordination is employed, inter-cell interference is mitigated so that the links can operate in the high SNR regime. This enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas.

3.1   System Model

We consider a cellular network with M base stations, each equipped with t transmit antennas, and N mobiles, each with r receive antennas. All M base stations can act together, and each user may receive signals from up to tM base antennas. As in Chapter 2, each base is loaded at most with one user, and also the total

number of transmit antennas is assumed to be larger than the total number of receive antennas, i.e., $tM \geq rN$. Taking all the base antennas as input and all the mobile antennas as output, we have a MIMO network. The received signal model for the kth mobile is as follows:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \qquad k = 1, 2, \ldots, N \qquad (3.1)$$

where $\mathbf{y}_k \in \mathbb{C}^{r \times 1}$ is the received signal, $\mathbf{H}_k = [h_{ij}]_{r \times tM}$ denotes the kth user's channel matrix with h_{ij} being the complex channel gain between the ith receive antenna and the jth transmit antenna, $\mathbf{x} \in \mathbb{C}^{tM \times 1}$ denotes the complex antenna outputs (without subscript k since it is composed of signals for all N users), and $\mathbf{n}_k \in \mathbb{C}^{r \times 1}$ denotes the white noise vector with covariance $\sigma^2 \mathbf{I}_r$. To simplify our analysis, we redefine the vectors in (3.1) to be in normalized form, meaning that each vector has been divided by the standard deviation of the additive noise component, σ. Then, the components of $\mathbf{n}_k$ have unit variance. Also, the N vectors $\{\mathbf{n}_k\}_{k=1}^{N}$ are i.i.d.

The fact that each user has r receive antennas can be exploited in the spatial domain by transmitting up to r independent symbol streams simultaneously for each user. Moreover, since all base antennas are coordinated, the complex antenna output vector $\mathbf{x}$ is composed of signals for all N users. Therefore, $\mathbf{x}$ can be written as follows:

$$\mathbf{x} = \sum_{j=1}^{r} b_{1j} \mathbf{w}_{1j} + \sum_{j=1}^{r} b_{2j} \mathbf{w}_{2j} + \cdots + \sum_{j=1}^{r} b_{Nj} \mathbf{w}_{Nj} \qquad (3.2)$$

where b_{ij} denotes the jth symbol of mobile i. In the context of coordinated multiple antenna transmissions, $\mathbf{w}_{ij} = [w_{ij}^1, w_{ij}^2, \ldots, w_{ij}^{tM}]^T$ denotes the complex unit norm antenna weight vector that is multiplied by b_{ij}. The selection of appropriate

antenna weight vectors, the mathematical problem, and the solution associated with each transmission method will be given in the next three sections.

3.2 Multiple Antenna Network Coordination by Zero-Forcing

For the multiple antenna zero-forcing coordination, the antenna weight vectors are selected so that each user's data does not interfere with other users' data. On the other hand, a user's own data symbols can interfere with each other. Thus, each normalized zero-forcing weight vector $\mathbf{w}_{ij}$ satisfies

$$\mathbf{H}_k \mathbf{w}_{ij} = 0, \quad \|\mathbf{w}_{ij}\|^2 = 1, \quad i \neq k, \quad j = 1, \ldots, r. \tag{3.3}$$

In other words, each unit norm weight vector $\mathbf{w}_{ij}$ has to be orthogonal to the subspace spanned by other users' channels. We note that each row of the channel matrix $\mathbf{H}_k$ corresponds to the channel seen by one of the receive antennas of user k. Let us denote the mth row of the matrix $\mathbf{H}_k$ by $\mathbf{h}_{km}$ for $m = 1, \ldots, r$. The channel vector $\mathbf{h}_{km}$ can be expressed as a sum of two vectors $\mathbf{h}_{km} = \mathbf{q}_{km} + \mathbf{q}'_{km}$ where $\mathbf{q}'_{km}$ denotes the part of the vector $\mathbf{h}_{km}$ in the subspace spanned by other users' channels. Similarly, we write $\mathbf{H}_k = \mathbf{Q}_k + \mathbf{Q}'_k$ where $\mathbf{q}_{km}$ and $\mathbf{q}'_{km}$ are the mth rows of the matrix $\mathbf{Q}_k$ and $\mathbf{Q}'_k$ respectively. Notice that the row spaces of $\mathbf{Q}_k$ and $\mathbf{Q}'_k$ are orthogonal spaces. The zero-forcing weight vectors are selected in such a way that user k's transmissions do not cause interference to other users, and therefore are confined into the subspace spanned by the vectors $\mathbf{q}_{km}$ for $m = 1, \ldots, r$ only, or equivalently in the row space of $\mathbf{Q}_k$. In order to find the basis for the row space, we use the singular value decomposition theorem, and

write $\mathbf{Q}_k$ as

$$\mathbf{Q}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^\dagger \qquad (3.4)$$

where $\mathbf{U}_k \in \mathbb{C}^{r \times r}$ and $\mathbf{V}_k \in \mathbb{C}^{tM \times tM}$ are unitary, and $\mathbf{S}_k \in \mathbb{C}^{r \times tM}$ is a zero matrix except for the square roots of r nonzero eigenvalues $(r < tM)$ of the matrix $\mathbf{Q}_k \mathbf{Q}_k^\dagger$ on the diagonals. We denote each diagonal by $\lambda_{kj}^{1/2}$ for $j = 1, \ldots, r$. By the singular value decomposition theorem, the first r columns of $\mathbf{V}_k$ are the bases for the row space of $\mathbf{Q}_k$, and therefore are selected to be the user k's zero-forcing weight vectors $\mathbf{w}_{kj}$ for $j = 1, \ldots, r$. Using (3.1)-(3.2), and with the particular selection of zero-forcing weight vectors, the received signal for user k is given by

$$
\begin{aligned}
\mathbf{y}_k &= \mathbf{H}_k \mathbf{x} + \mathbf{n}_k & (3.5) \\
&= \mathbf{H}_k \left(\sum_{j=1}^{r} b_{1j} \mathbf{w}_{1j} + \sum_{j=1}^{r} b_{2j} \mathbf{w}_{2j} + \cdots + \sum_{j=1}^{r} b_{Nj} \mathbf{w}_{Nj} \right) + \mathbf{n}_k & (3.6) \\
&= \mathbf{H}_k \left(\sum_{j=1}^{r} b_{kj} \mathbf{w}_{kj} \right) + \mathbf{n}_k & (3.7) \\
&= (\mathbf{Q}_k + \mathbf{Q}_k') \left(\sum_{j=1}^{r} b_{kj} \mathbf{w}_{kj} \right) + \mathbf{n}_k & (3.8) \\
&= \mathbf{Q}_k \left(\sum_{j=1}^{r} b_{kj} \mathbf{w}_{kj} \right) + \mathbf{n}_k & (3.9) \\
&= \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^\dagger \left(\sum_{j=1}^{r} b_{kj} \mathbf{w}_{kj} \right) + \mathbf{n}_k & (3.10) \\
&= \mathbf{U}_k \begin{bmatrix} \lambda_{k1}^{1/2} b_{k1} \\ \lambda_{k2}^{1/2} b_{k2} \\ \vdots \\ \lambda_{kr}^{1/2} b_{kr} \end{bmatrix} + \mathbf{n}_k & (3.11)
\end{aligned}
$$

where (3.7)-(3.9) is due to the fact that each user's zero-forcing weight vectors are orthogonal to the subspace spanned by other users' channels, and (3.11) follows from the fact that user k's weight vectors are selected to be the first r columns of the unitary matrix $\mathbf{V}_k$, and $\mathbf{S}_k$ is a diagonal matrix with the square roots of r nonzero eigenvalues on the diagonals. User k recovers its symbols by match-filtering the received signal with $\mathbf{U}_k^\dagger$:

$$\tilde{\mathbf{y}}_k = \mathbf{U}_k^\dagger \mathbf{y}_k = \begin{bmatrix} \lambda_{k1}^{1/2} b_{k1} \\ \lambda_{k2}^{1/2} b_{k2} \\ \vdots \\ \lambda_{kr}^{1/2} b_{kr} \end{bmatrix} + \tilde{\mathbf{n}} \tag{3.12}$$

where the noise vector $\tilde{\mathbf{n}}_k = \mathbf{U}_k^\dagger \mathbf{n}_k$ remains white with covariance $\sigma^2 \mathbf{I}_r$ due to the unitary transformation. It follows that the rate achievable in the parallel Gaussian channels (3.12) is given by

$$R_k = \sum_{j=1}^{r} \log_2 \left(1 + \lambda_{kj} E[|b_{kj}|^2]\right). \tag{3.13}$$

Given the antenna weight vector $\mathbf{w}_{kj} = [w_{kj}^1, w_{kj}^2, \ldots, w_{kj}^{tM}]^T$, the symbol b_{kj} with power $P_{kj} = E[|b_{kj}|^2]$ contributes power $|w_{kj}^q|^2 P_{kj}$ to the total transmit power at antenna q. The total transmit power at antenna q is the sum of the contributions of the symbols of all users, i.e.,

$$P_q = \sum_{k=1}^{N} \sum_{j=1}^{r} |w_{kj}^q|^2 P_{kj}. \tag{3.14}$$

Now that we have described how to calculate the achievable rates, and the individual power constraints at each base antenna, we give the formal problem statement. The objective is to maximize the minimum rate in the network subject to per-base (or per-antenna) power constraints. The problem can be defined as

$$\max \quad r_0 \tag{3.15}$$

$$\text{s.t.} \quad \sum_{j=1}^{r} \log_2 \left(1 + \lambda_{kj} P_{kj} \right) \geq r_0, \ k = 1, \ldots, N \tag{3.15a}$$

$$\sum_{k=1}^{N} \sum_{j=1}^{r} |w_{kj}^q|^2 P_{kj} \leq P_{\max}, \ q = 1, \ldots, tM \tag{3.15b}$$

$$r_0 \in \Re^+, P_{kj} \in \Re^+ \quad \forall \, k, j. \tag{3.15c}$$

Note that r_0 can be interpreted as the minimum rate to be maximized, i.e., $r_0 = \min_k \sum_{j=1}^{r} \log_2 \left(1 + \lambda_{kj} P_{kj} \right)$. Also, notice that the constraint in (3.15b) defines a per-antenna power constraint. However, other types of power constraints can easily be included in the above optimization problem. For example, if we need a constraint on the total base power, we need to sum up the contribution of each base antenna given on the left side of the inequality in (3.15b) and constrain the sum. In our numerical examples, we will use a per-base power constraint. The problem (3.15) is a convex optimization problem since the constraints in (3.15b) are linear, and the constraints in (3.15a) define a convex region. The logarithm function is concave in the power assignments, and the sum of concave functions is also concave. Each constraint in (3.15a) defines a convex set, which is the region underneath a concave function intersected with the region above the hyperplane defined by r_0. This is because the intersection of convex sets is another convex set. Therefore, standard convex optimization techniques can be used to solve the

problem [30].

3.3 Multiple Antenna Network Coordination by Zero-Forcing Dirty Paper Coding

Similar to the single antenna network coordination schemes, an improved form of multiple antenna coordination is achieved when zero forcing and dirty paper coding work together to get rid of the interference. In this section, we will use a reduced form of zero-forcing to get rid of the interference that the dirty paper coding does not remove. In the multiple antenna zero forcing coordination, each user's interference was nulled out by selecting the weight vectors to be orthogonal to the other user's channels. In the dirty paper coding approach, when the interference is known causally at the transmitter, N users' codewords can be chosen such that given an ordering of users $[\pi(1), \ldots, \pi(N)]$, where π is a permutation, a user with index $\pi(i)$ does not suffer any interference from users with lower indexes, i.e., $\pi(k)$ with $k < i$. When dirty paper coding is combined with the reduced form of zero forcing, the interference still present from dirty paper coding will be nulled out due to the zero forcing weight vectors. In this case, the weight vectors for the data symbols of user $\pi(k)$ with $k > i$ have to be orthogonal to the user $\pi(i)$'s channel. Thus, each zero forcing weight vector $\mathbf{w}_{\pi(k)j}$ must satisfy

$$\mathbf{H}_{\pi(i)}\mathbf{w}_{\pi(k)j} = 0, \quad \|\mathbf{w}_{\pi(k)j}\|^2 = 1, \quad k > i, \quad j = 1, \ldots, r. \qquad (3.16)$$

The weight vectors satisfying the above conditions can be found in a similar way to the zero-forcing case. First, we write the channel vector $\mathbf{h}_{\pi(k)m}$ as a sum of

two vectors $\mathbf{h}_{\pi(k)m} = \mathbf{q}_{\pi(k)m} + \mathbf{q}'_{\pi(k)m}$ where $\mathbf{q}'_{\pi(k)m}$ denotes the part of the vector $\mathbf{h}_{\pi(k)m}$ in the subspace spanned by channels of users $\pi(i)$ with $i < k$. Similarly, we write $\mathbf{H}_{\pi(k)} = \mathbf{Q}_{\pi(k)} + \mathbf{Q}'_{\pi(k)}$. The zero-forcing weight vectors are selected in such a way that user $\pi(k)$'s transmissions are confined into the subspace spanned by the vectors $\mathbf{q}_{\pi(k)m}$ for $m = 1, \ldots, r$ only, or equivalently in the row space of $\mathbf{Q}_{\pi(k)}$. In order to find the bases for the row space, we use the singular value decomposition theorem, and write $\mathbf{Q}_{\pi(k)} = \mathbf{U}_{\pi(k)}\mathbf{S}_{\pi(k)}\mathbf{V}^{\dagger}_{\pi(k)}$. The first r columns of $\mathbf{V}_{\pi(k)}$ are selected to be the user $\pi(k)$'s zero-forcing weight vectors $\mathbf{w}_{\pi(k)j}$ for $j = 1, \ldots, r$. In this case, the received signal for user $\pi(i)$ is given by

$$
\begin{aligned}
\mathbf{y}_{\pi(i)} &= \mathbf{H}_{\pi(i)}\mathbf{x} + \mathbf{n}_{\pi(i)} & (3.17)\\
&= \mathbf{H}_{\pi(i)}\left(\sum_{j=1}^{r} b_{1j}\mathbf{w}_{1j} + \sum_{j=1}^{r} b_{2j}\mathbf{w}_{2j} + \cdots + \sum_{j=1}^{r} b_{Nj}\mathbf{w}_{Nj}\right) + \mathbf{n}_{\pi(i)} & (3.18)\\
&= \mathbf{H}_{\pi(i)}\left(\sum_{j=1}^{r} b_{\pi(i)j}\mathbf{w}_{\pi(i)j}\right) + \mathbf{H}_{\pi(i)}\left(\sum_{k>i}\sum_{j=1}^{r} b_{\pi(k)j}\mathbf{w}_{\pi(k)j}\right) + \mathbf{n}_{\pi(i)} & (3.19)\\
&= \mathbf{H}_{\pi(i)}\left(\sum_{j=1}^{r} b_{\pi(i)j}\mathbf{w}_{\pi(i)j}\right) + \mathbf{n}_{\pi(i)} & (3.20)\\
&= \left(\mathbf{Q}_{\pi(i)} + \mathbf{Q}'_{\pi(i)}\right)\left(\sum_{j=1}^{r} b_{\pi(i)j}\mathbf{w}_{\pi(i)j}\right) + \mathbf{n}_{\pi(i)} & (3.21)\\
&= \mathbf{Q}_{\pi(i)}\left(\sum_{j=1}^{r} b_{\pi(i)j}\mathbf{w}_{\pi(i)j}\right) + \mathbf{n}_{\pi(i)} & (3.22)\\
&= \mathbf{U}_{\pi(i)}\mathbf{S}_{\pi(i)}\mathbf{V}^{\dagger}_{\pi(i)}\left(\sum_{j=1}^{r} b_{\pi(i)j}\mathbf{w}_{\pi(i)j}\right) + \mathbf{n}_{\pi(i)} & (3.23)\\
&= \mathbf{U}_{\pi(i)}\begin{bmatrix} \lambda^{1/2}_{\pi(i)1}b_{\pi(i)1} \\ \lambda^{1/2}_{\pi(i)2}b_{\pi(i)2} \\ \vdots \\ \lambda^{1/2}_{\pi(i)r}b_{\pi(i)r} \end{bmatrix} + \mathbf{n}_{\pi(i)} & (3.24)
\end{aligned}
$$

Input : $\mathbf{H}_k = [h_{ij}]_{r \times tM}$, $k = 1, 2, \ldots, N$ (channel response matrices).
Output : $\pi = [\pi(1), \pi(2), \ldots, \pi(N)]$, ($\pi(i)$ projects away from $\pi(j)$ where $i > j$).
Initialization : $\mathbf{S} = \{1, 2, \ldots, N\}$.
for k = 1 : N
    $\pi(k) = \arg\min_{k \in S} \log_2 \left[\det(\mathbf{I}_r + P_{max}\mathbf{H}_k\mathbf{H}_k^{\dagger}) \right],$
    $\mathbf{S} = \mathbf{S} - \pi(k),$
    $\mathbf{e}_{\pi(k),i} = \frac{\mathbf{h}_{\pi(k),i}}{|\mathbf{h}_{\pi(k),i}|}, i = 1, \ldots, r,$ ($\mathbf{h}_{\pi(k),i}$: ith column of $\mathbf{H}_{\pi(k)}$)
    $\mathbf{h}_{j,l} = \mathbf{h}_{j,l} - (\mathbf{h}_{j,l}^{\dagger}\mathbf{e}_{\pi(k),i})\mathbf{e}_{\pi(k),i}, \forall j \in \mathbf{S}, i, l = 1, \ldots, r.$
end

Figure 3.1: The heuristic user-ordering algorithm for multiple antenna networks.

where (3.19) follows from the fact that a user with index $\pi(i)$ does not suffer any interference from users $\pi(k)$ with $k < i$ due to the dirty paper coding. The interference that dirty paper coding cannot get rid of, i.e., the term inside the second parenthesis in (3.19), is nulled out by appropriate zero forcing weight vectors (3.16), and accordingly (3.20) represents the received signal after both dirty paper coding and zero-forcing. Equation (3.22) is due to the fact that the weight vector $\mathbf{w}_{\pi(i)j}$ is a basis in the row space of the matrix $\mathbf{Q}_{\pi(i)}$, which is orthogonal to the row space of $\mathbf{Q}_{\pi(i)}'$. The remaining steps are due to the singular value decomposition theorem.

Given the received signal (3.24), the calculation of the achievable rates, and the power constraints at each base follow the same steps as in the previous section. Similarly, the problem statement is the same as (3.15), except that the eigenvalues and the weight vectors have to be determined based on the zero-forcing dirty paper coding scheme described in this section.

The above multiple antenna network coordination scheme assumes a particular user ordering. In this case, a similar version of the heuristic we mentioned in the context of single antenna systems can be used. Namely, the user with the smallest single-user log-det capacity is a disadvantaged user. Since the objective is fairness,

all other users have to project away from this user. In turn, the disadvantaged user "pays the favor back" by being invisible via the dirty paper encoding. In the subspace that does not cause interference to this disadvantageous user, the next user with the smallest log-det capacity is chosen to be the second user in the ordering. These steps are followed in the same way for the remaining mobiles. The heuristic user ordering algorithm is shown in Figure 3.1. We will see in the numerical examples section that the algorithm performs quite well in our cellular network setup.

3.4 Multi-antenna Cellular Networks with Power Control

Our baseline for comparison with the coordinated schemes is a conventional multiple antenna cellular network with power control. As an extension of the single antenna power control scheme, we will assume a simple form of power control where each antenna of a base station transmits with the same power P_k, and the power level at each base is controlled. We also assume that $\mathbf{x}_k$ is Gaussian with covariance $P_k \mathbf{I}_t$. In this case, user k's rate is given by $\log_2\left[\det(\mathbf{I}_r + P_k \mathbf{H}_k \mathbf{H}_k^\dagger \mathbf{R}^{-1})\right]$ where $\mathbf{R} = \sum_{i=1, i \neq k}^{N} P_i \mathbf{H}_i \mathbf{H}_i^\dagger + \mathbf{I}_r$. To maximize the minimum rate in the network, we solve the following problem

$$\max \quad r \tag{3.25}$$

$$\text{s.t.} \quad \log_2\left[\det(\mathbf{I}_r + P_k \mathbf{H}_k \mathbf{H}_k^\dagger (\sum_{i=1, i \neq k}^{N} P_i \mathbf{H}_i \mathbf{H}_i^\dagger + \mathbf{I}_r)^{-1})\right] \geq r, \ k = 1, \dots, N \tag{3.25a}$$

$$0 \leq P_k \leq P_{\max}, r \in \Re^+. \tag{3.25b}$$

The way we approach this problem is also analogous to the power control problems in single-antenna systems [37, 38]. First, we start with a very low, easily achievable target rate r, and initialize all power assignments to zero. Next, we determine the power level for each user to achieve the target rate given the interference of other users. In the second iteration, we increase the target rate, and given the set of power assignments from the previous iteration and corresponding interference levels, we calculate the new set of power levels to achieve the new target rate. We follow the same procedure, with higher target rates, until any of the base station power constraints is violated. By the very nature of this procedure, the successive iterations generate a bounded sequence of increasing rates, and therefore the method converges.

3.5   Performance Evaluation

In this section, we will compare the performance of coordinated multiple antenna networks to that of conventional cellular networks with inter-cell interference. In Chapter 2, we showed that the coordination is successful in eliminating inter-cell interference. Our results in this section indicate that the coordination does more than mitigate the inter-cell interference. Namely, the coordination improves the effectiveness of multiple antennas in cellular networks.

Our basic experimental setup is the same as in Chapter 2. In particular, each base is located at the center of one of the 64 hexagonal cells whose sample layout is shown in Figure 2.3. We denote by (t, r) a configuration in which each base station has t transmit antennas and each mobile has r receive antennas. Three antenna configurations will be considered: (1,1), (2,2) and (4,4) where the base
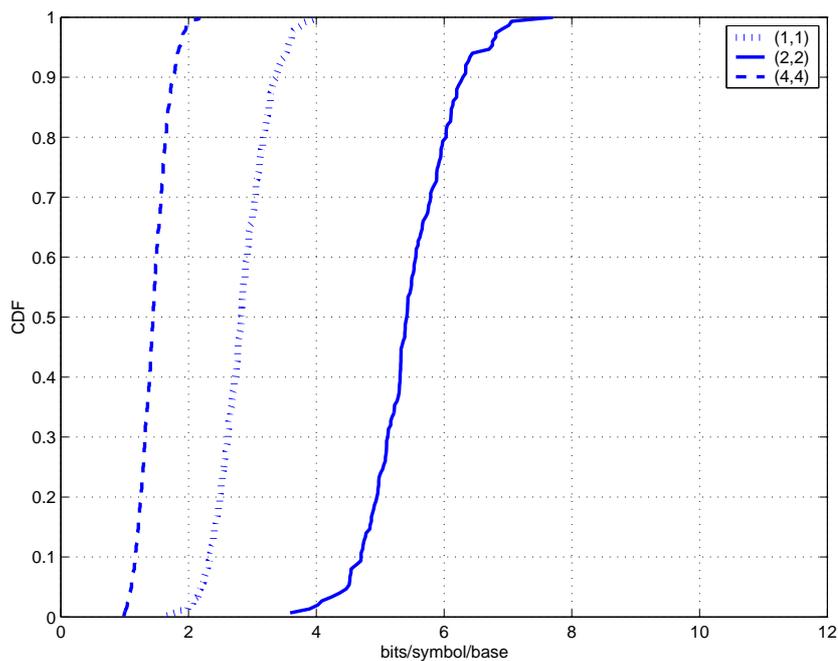
Figure 3.2: Empirical max-min rate CDFs: Multiple antenna power control results for a 64 base network.
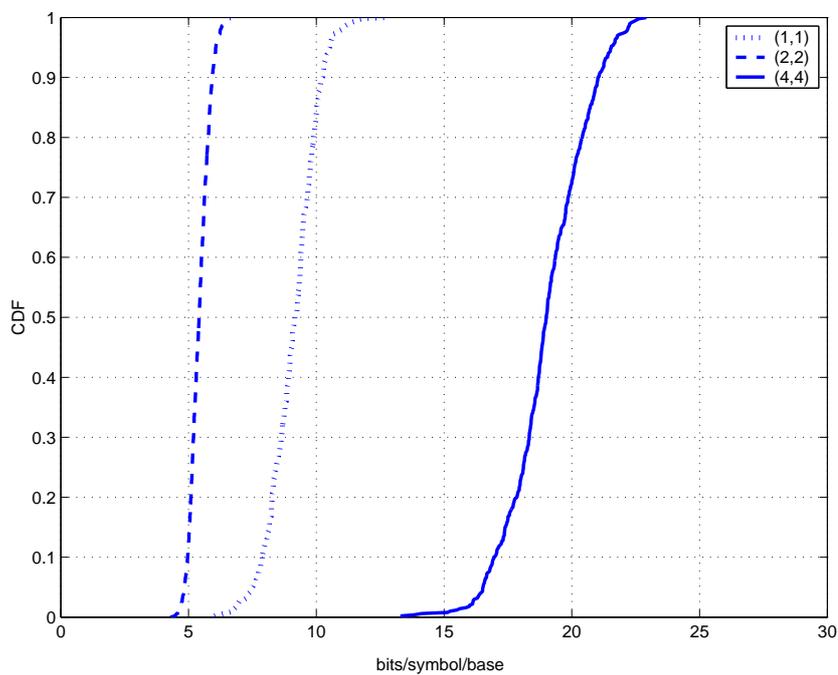
Figure 3.3: Empirical max-min rate CDFs: Multiple antenna zero-forcing results for a 64 base network.
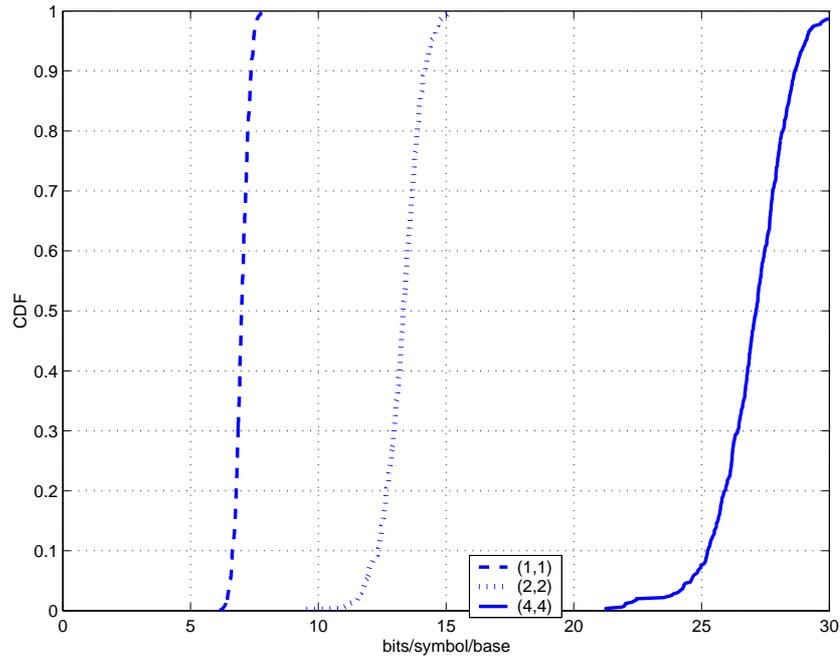
Figure 3.4: Empirical max-min rate CDFs: Multiple antenna zero-forcing dirty paper coding results for a 64 base network.

stations and the mobiles are equipped with 1, 2 and 4 antennas respectively. Each base is loaded with one mobile, and 10% of the mobiles are allowed to be in outage. We assume path-loss with a 3.8 propagation exponent, log-normal shadowing with 0 mean and 8 dB standard deviation, and Rayleigh fading for each transmit/receive antenna pair with 0 mean, unit variance complex Gaussian component. Given a maximum base power of 10 W, a mean power loss of 134 dB at the reference distance of 1.6 km from the base, a receiver noise figure of 5 dB, a vertical antenna gain of 10.3 dBi, a channel bandwidth of 5 MHz, and a receiver temperature of 300^oK, the SNR at the reference distance is 18 dB (considering one transmit/receieve antenna pair), accounting only for path loss and ignoring shadowing and Raleigh fading. However, because of inter-cell interference, the SINR can fall below 0 dB in our network setup.

We have seen in the previous chapter that, for cellular networks with single
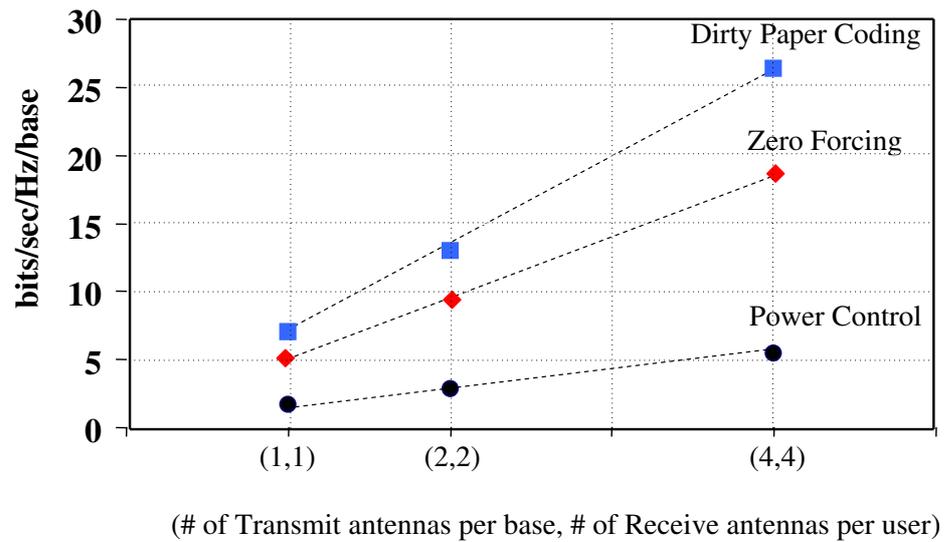
Figure 3.5: Spectral efficiency v.s. # of antennas.
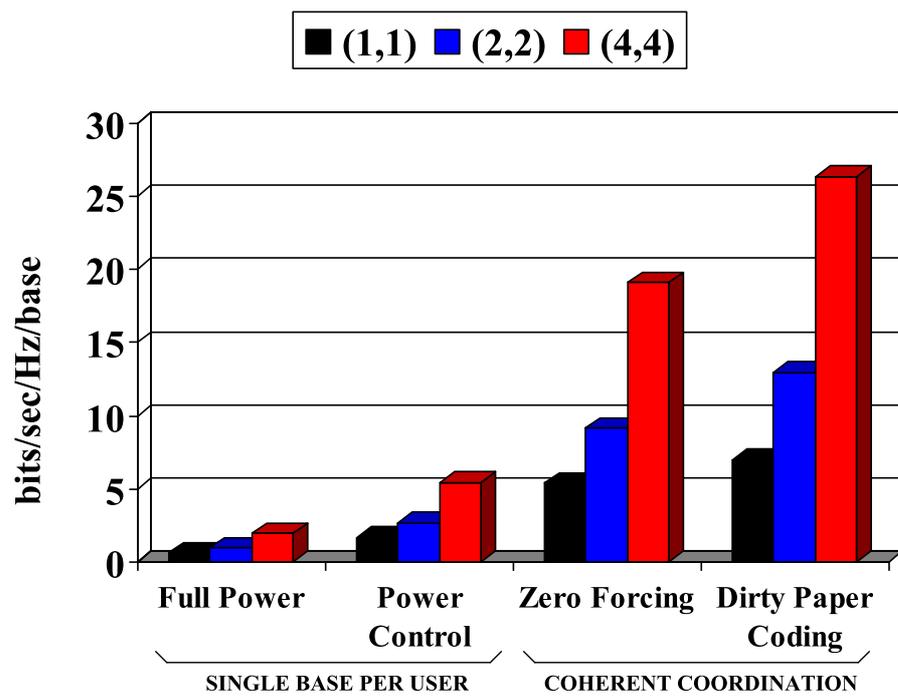
Figure 3.6: Summary of the multiple antenna results. The results for a conventional network with each base antenna transmitting at the maximum power limit is also presented.

antenna terminals, the coordination improves the max-min rate of the network by about a factor-of-3 when the zero-forcing coordination is employed, and by about a factor-of-5 when the combined zero-forcing dirty paper coding scheme is used. Corresponding multiple antenna spectral efficiency results obtained by averaging over many network instances are shown in Figure 3.2-3.6 for all three antenna arrangements. The figure shows that the use of multiple antennas improves (linearly) the spectral efficiency of the network for all transmission methods. However, the gain in terms of the marginal increase in rate when an additional antenna is deployed is small without coordination. We observe in Figure 3.5 that the slope of the rate versus the number of antennas curve is small for the uncoordinated transmissions with power control, while it is significantly improved when the coordination is employed. More than a factor-of-10 improvement in spectral efficiency is reported due to both network coordination and $(4, 4)$ antenna arrangements compared to the baseline of uncoordinated transmissions with single antenna terminals. The results show that the removal of "out of cell" interference greatly enhances the spectral efficiency improvement associated with increasing the number of antennas.

3.6   Chapter Summary and Conclusion

In this chapter, we have considered multiple antenna cellular networks with and without network coordination. We have seen that the network suffers from inter-cell interference when the coordination is not employed, and this would reduce the effectiveness of multiple antennas due to relatively low link SNRs. On the other hand, employing network coordination improves the link qualities, and enables

the cellular network to realize the predicted capacity improvement due to the multiple antennas. This suggests that if the multiple antennas are to be deployed in a practical interference-limited cellular downlink system, it is necessary to employ some form of interference mitigation technique to improve link SNRs. The network coordination is an effective way of achieving this objective.

Chapter 4

Beamforming Design With Per-Antenna Power Constraints

Thus far, we have considered coordination techniques that require simple, practical linear beamforming techniques. For the single antenna zero-forcing method, each mobile's data symbol is multiplied by a beamforming vector given by the column of the pseudo-inverse matrix. Similarly, for the combined zero-forcing dirty paper coding scheme with single antenna terminals, the beamforming vectors are given by the columns of the unitary matrix obtained by the LQ decompositions of the channel matrix. Both derivations are extended to the multiple antenna systems as well. While these linear beamforming techniques provide significant capacity improvements over the conventional cellular networks, they are not claimed to be the optimal linear coordination techniques. The objective in this chapter is to find the optimal beamforming vectors for the coordinated transmissions.

Notice that, in our problem formulations, each base station antenna has a separate power constraint. The optimal system design with per-antenna power constraints is different than the design based on a sum-power constraint. This fact can be illustrated by a simple example where M transmit antennas serve one mobile with a single antenna. Let's assume that the mobile's i.i.d. vector channel is denoted by $\mathbf{h} = [h_1, h_2, \ldots, h_M]^T$, and the antenna outputs are denoted by $\mathbf{x} = [x_1, x_2, \ldots, x_M]^T$. When the system is sum-power constrained, the optimal

transmission method is the transmit beamforming where the antenna outputs are aligned with the mobile's channel, and the beamforming vector is scaled to the maximum transmit power P. Thus, the optimum antenna output becomes $\mathbf{x}_{opt} = P\mathbf{h}^{\dagger}/|\mathbf{h}|$. When each antenna has a separate power constraint, say $P^{'} = P/M$ for each antenna, the optimum transmission method is to transmit full power at each antenna, and phase-match the output voltages to the channels, i.e., $\mathbf{x}_{opt} = P^{'}[h_1^{\dagger}/|h_1|, h_2^{\dagger}/|h_2|, \ldots, h_M^{\dagger}/|h_M|]^T$. The problem with the transmit beamforming in this case is that it gives more weights to the antennas whose channels are good, and in the case of a per-antenna power constraint, this makes the antenna power to drain quickly.

So far, our beamforming design ignores the existence of a per-antenna power constraint, and therefore is similar to the transmit beamforming design. For example, when the pseudo-inverse matrix is used for the zero-forcing, each mobile's beamforming vector becomes aligned with the part of the mobile's channel which does not deposit interference on the others. Similarly, for the multiple antenna zero-forcing system, beamforming vectors are chosen to be the eigenvectors of the part of the mobile's channel matrix which is orthogonal to the others. Notice that, despite ignoring the existence of per-antenna power constraints, these beamforming designs performed quite well as their performance gets very close to the upper bounds obtained in Chapter 2. The reason is that, as we will see in the following sections, optimizing the antenna outputs based on per-antenna power constraints may improve the rate considerably when the number of transmit antennas is much larger the number of receive antennas, i.e., $tM \gg rN$ in the multiple antenna system and $tM \gg N$ in the single antenna system. This is because more transmit antennas will give more degrees of freedom to optimize the

antenna outputs. When the number of transmit and receive antennas are close to each other, there are not much room left to exploit in the signal space (consider the zero-forcing design where each mobile's transmission has to be projected away from the others), which is the case in the previous two chapters.

4.1 System Model

We consider a broadcast system with a transmitter equipped with M antennas, and N mobiles, each with a single receive antenna. The received signal model for the kth mobile is as follows:

$$y_k = \mathbf{h}_k^\dagger \mathbf{x} + n_k \qquad k = 1, 2, \ldots, N \tag{4.1}$$

where y_k is the received signal, $\mathbf{h}_k = [h_{k1}, h_{k2}, \ldots, h_{kM}]^\dagger$ denotes the kth mobile's complex channel vector, $\mathbf{x} = [x_1, x_2, \ldots, x_M]^\dagger$ denotes the antenna outputs (without subscript k since it is composed of signals for all N mobiles), and n_k denotes the complex Gaussian noise with variance σ^2. In the above model, the M transmit antennas can be thought of as multiple transmit antennas of a base station in a single cell system, or alternatively as a network-wide transmit antenna array in a multi-cell system when M base stations are connected via a high-speed backbone enabling coordinated base transmissions.

We assume that the transmitter employs linear beamforming where the vector $\mathbf{w}_k = [w_{k1}, w_{k2}, \ldots, w_{kM}]^\dagger$ is used to map the kth mobile's data symbol b_k to the antenna outputs. The antenna output $\mathbf{x}$ is composed of signals for all N mobiles,

and therefore is given by

$$\mathbf{x} = \sum_{k=1}^{N} b_k \mathbf{w}_k. \tag{4.2}$$

As a case study, we are interested in the zero-forcing beamforming where each mobile's transmissions do not cause interference to others. Our analysis can simply be extended to the zero-forcing dirty paper coding system design as well. In the zero-forcing case, the beamforming vector $\mathbf{w}_k$ is chosen to be orthogonal to all other mobiles' channels:

$$\mathbf{h}_j^{\dagger} \mathbf{w}_k = 0, \quad j \neq k. \tag{4.3}$$

Given the requirement (4.3), and using (4.2), the kth mobile's received signal with zero-forcing beamforming becomes $y_k = b_k \mathbf{h}_k^{\dagger} \mathbf{w}_k + n_k$. Assuming that the data symbol b_k is zero-mean Gaussian with unit variance, mobile k's achievable rate is given by

$$r_k = \log_2 \left(1 + \frac{|\mathbf{h}_k^{\dagger} \mathbf{w}_k|^2}{\sigma^2} \right). \tag{4.4}$$

Finally, each transmit antenna has a power constraint given by $E[|x_m|^2] \leq P_{max}$, $m = 1, \ldots, M$. Since the data symbols for different mobiles are i.i.d., this constraint can be written as $\sum_{k=1}^{N} |w_{km}|^2 \leq P_{max}$, $m = 1, \ldots, M$.

4.2   Problem Formulation and Solution

The objective is to design a zero-forcing beamforming system so as to maximize the minimum rate in the network subject to per-antenna power constraints. The

optimization problem can be written as

$$\max_{\mathbf{w}_k} \quad r_0 \tag{4.5}$$

$$\text{s.t.} \quad \log_2\left(1 + \frac{|\mathbf{h}_k^\dagger \mathbf{w}_k|^2}{\sigma^2}\right) \geq r_0, \ k = 1, \ldots, N, \tag{4.5a}$$

$$\sum_{k=1}^{N} |\mathbf{e}_m^\dagger \mathbf{w}_k|^2 \leq P_{max}, \ m = 1, \ldots, M, \tag{4.5b}$$

$$\mathbf{h}_j^\dagger \mathbf{w}_k = 0 \quad \forall j \neq k, \ \mathbf{w}_k \in \mathbb{C}^{M \times 1}, \ r_0 \in \Re^+. \tag{4.5c}$$

where $\mathbf{e}_m$ is a standard basis vector which has 1 for its mth component, and 0 for every other component. In the above formulation, r_0 can be interpreted as the minimum rate to be maximized, i.e., $r_0 = \min_k \log_2(1 + |\mathbf{h}_k^\dagger \mathbf{w}_k|^2/\sigma^2)$. First, notice that the zero-forcing constraints in (4.5c) are simply linear. Moreover, the per-antenna power constraints (4.5b) define a convex set, since the function $f(\mathbf{w}_k) = |\mathbf{e}_m^\dagger \mathbf{w}_k|^2$ is convex in $\mathbf{w}_k$, the sum of convex functions is convex, and α-sublevel set of a convex function defined by $\{\mathbf{w}_k \in \mathbf{dom}\, f \,|\, f(\mathbf{w}_k) \leq \alpha\}$ is a convex set [30]. The rate constraints (4.5a) can be written as $|\mathbf{h}_k^\dagger \mathbf{w}_k| \geq \sqrt{(2^{r_0} - 1)\sigma^2}$, $k = 1, \ldots, K$, which defines a non-convex set since it is the region outside of a convex set defined by $|\mathbf{h}_k^\dagger \mathbf{w}_k| \leq \sqrt{(2^{r_0} - 1)\sigma^2}$. Therefore, the problem above is not a convex optimization problem, for which a tractable analytical solution would be possible.

Fortunately, the problem can be simplified significantly by observing that the optimum beamforming vectors are invariant to phase-shifts, i.e., if $\mathbf{w}_k^*$ is an optimum beamforming vector for the problem (4.5), then the vector $e^{j\theta}\mathbf{w}_k^*$ is also an optimum beamforming vector [33, 34]. This is because of the fact for any feasible $\mathbf{w}_k$, the vector $e^{j\theta}\mathbf{w}_k$ would also satisfy the constraints (4.5a)-(4.5c). In

this case, without loss of generality, we can search for the optimum beamforming solutions among the vectors that would result in positive and real projections $\mathbf{h}_k^\dagger \mathbf{w}_k$. Thus, an equivalent optimization problem can be written as

$$\max_{\mathbf{w}_k} \quad t \tag{4.6}$$

$$\text{s.t.} \quad (\mathbf{h}_k^\dagger \mathbf{w}_k) \geq t, \ k = 1, \ldots, N, \tag{4.6a}$$

$$\text{Im}(\mathbf{h}_k^\dagger \mathbf{w}_k) = 0 \quad \forall k, \tag{4.6b}$$

$$\sum_{k=1}^{K} |\mathbf{e}_m^\dagger \mathbf{w}_k|^2 \leq P_{max}, \ m = 1, \ldots, M, \tag{4.6c}$$

$$\mathbf{h}_j^\dagger \mathbf{w}_k = 0 \quad \forall j \neq k, \ \mathbf{w}_k \in \mathbb{C}^{M \times 1}, t \in \Re^+. \tag{4.6d}$$

where $t = \sqrt{(2^{r_0} - 1)\sigma^2}$, and the optimum beamforming vectors maximizing t would also maximize r_0. Similar to (4.5), the constraints (4.6c) and (4.6d) are convex and linear respectively, and (4.6a) and (4.6b) are linear constraints. Since the objective function is also linear, the problem (4.6) is a convex optimization problem.

In the next section, we will consider the conventional zero-forcing technique based on the Moore-Penrose pseudo-inverse, and investigate why and when this technique can be an optimal/suboptimal solution to the above problem (4.6).

4.2.1 Optimality/Suboptimality of the Moore-Penrose Zero-Forcing

The simplest, and probably the most widely used, form of zero-forcing involves channel inversion when the channels are invertible, or the Moore-Penrose pseudo-inverse operation otherwise. In this section, we will study the optimality of these simplest forms of zero-forcing. We will show that the zero-forcing based on the conventional pseudo-inverse operation is a suboptimal zero-forcing technique when there are per-antenna power constraints.

First, we briefly revisit the optimal max-min rate solution under the assumption of zero-forcing by channel inversion, or by the pseudo-inverse operation. The system model is written in matrix form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{4.7}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_N]^\dagger$ is the vector of the received signals, $\mathbf{H} \in \mathbb{C}^{N \times M}$ is the channel matrix with the kth row given by $\mathbf{h}_k^T$, $\mathbf{x}$ is the complex antenna outputs, and $\mathbf{n} = [n_1, n_2, \ldots, n_N]^\dagger$ denotes an additive white noise vector with covariance $\sigma^2 \mathbf{I}$. Denoting the vector of data symbols by $\mathbf{b} = [b_1, b_2, \ldots, b_N]^\dagger$, the linear beamforming operation is performed by a pre-filter matrix $\mathbf{W} \in \mathbb{C}^{M \times N}$, which maps the data symbols to the antenna outputs, i.e., $\mathbf{x} = \mathbf{W}\mathbf{b}$. It is clear from (4.2) that the kth column vector of $\mathbf{W}$ is the beamforming vector $\mathbf{w}_k$.

When there are as many users as the number of antennas, $\mathbf{H}$ is a square matrix, and there can be only one pre-filter matrix achieving zero-forcing transmission, which is the inverse channel given by $\mathbf{W} = \mathbf{H}^{-1}$. When there are more antennas

than users, the pre-filter matrix takes the form of a pseudo-inverse matrix given by $\mathbf{W} = \mathbf{H}^\dagger(\mathbf{HH}^\dagger)^{-1}$. Geometrically speaking, in both cases, each user's channel is projected away from all other users' channels, and the remaining piece of the channel gives the beamforming direction which would not cause interference to others. When the above pre-filter matrices are used, each user receives its own data symbol corrupted by the noise, i.e., $y_k = b_k + n_k$. Denoting the power of b_k by $p_k = E[|b_k|^2]$, the following convex problem results in the max-min rate solution for the given pre-filter matrix $\mathbf{W}$:

$$\max_{p_k} \quad r_0 \tag{4.8}$$

$$s.t. \quad \log_2\left(1 + \frac{p_k}{\sigma^2}\right) \geq r_0, \ k = 1,\ldots,N, \tag{4.8a}$$

$$\begin{bmatrix} |W_{11}|^2 & \cdots & |W_{1N}|^2 \\ \vdots & & \vdots \\ |W_{M1}|^2 & \cdots & |W_{MN}|^2 \end{bmatrix} \begin{bmatrix} p_1 \\ \vdots \\ \vdots \\ p_N \end{bmatrix} \leq P_{max}\mathbf{1}, \tag{4.8b}$$

$$r_0, p_k \in \Re^+ \tag{4.8c}$$

In the above problem, (4.8b) represents the per-antenna power constraints, and is obtained using the fact that the antenna powers are on the diagonals of the following transmit covariance matrix

$$E[\mathbf{xx}^\dagger] = \mathbf{W}E[\mathbf{bb}^\dagger]\mathbf{W}^\dagger = \mathbf{W} \begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_N \end{bmatrix} \mathbf{W}^\dagger, \tag{4.9}$$
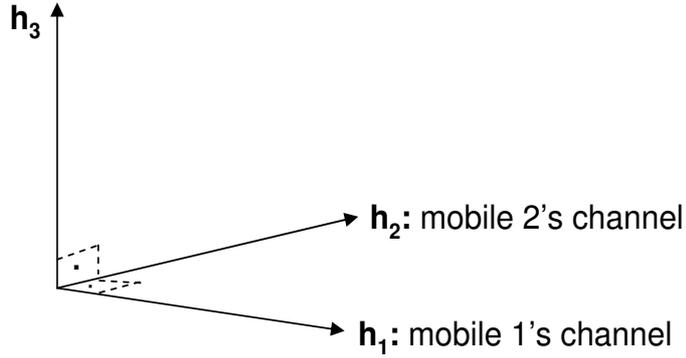
Figure 4.1: 3 transmit antennas and 2 mobiles with orthogonal channels.

and the constraints on the diagonals can be expressed in the matrix form (4.8b).

We will show in the case of 2 users that, when there is a sum-power constraint instead of a per-antenna power constraint (4.5b), the columns of the Moore-Penrose pseudo-inverse matrix defines the optimum beamforming directions, and the above problem (4.8) would result in the max-min rate solution for the original problem (4.5). On the other hand, the same zero-forcing form is suboptimal when there are per-antenna power constraints.

Proposition 1. *The beamforming based on the Moore-Penrose pseudo-inverse matrix* $\mathbf{W} = \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ *is a suboptimal form of zero-forcing when there are per-antenna power constraints.*

Proof. We prove the statement by an example. Consider a broadcast system with 3 transmit antennas and 2 single antenna mobiles with channels $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^{3\times1}$. Assume that $\mathbf{h}_1$ and $\mathbf{h}_2$ are orthogonal, $\mathbf{h}_1^\dagger\mathbf{h}_2 = 0$. In three dimensional space, there is a third vector $\mathbf{h}_3 \in \mathbb{C}^{3\times1}$, which is orthogonal to both $\mathbf{h}_1$ and $\mathbf{h}_2$. These three vectors are shown in Figure 4.1. The matrix $\mathbf{H} \in \mathbb{C}^{2\times3}$ representing the channel between the two mobile antennas and the transmit antenna array has $\mathbf{h}_1^T$

and $\mathbf{h}_2^T$ on its rows. The zero-forcing beamforming directions are on the column vectors of the pseudo-inverse channel $\mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$, and are given by $\mathbf{h}_1/|\mathbf{h}_1|^2$ and $\mathbf{h}_2/|\mathbf{h}_2|^2$ respectively. In this case, the signals are transmitted in the directions of $\mathbf{h}_1$ and $\mathbf{h}_2$ intended for the first and the second mobile respectively. These particular beamforming selections make sense since they provide interference-free transmission because $\mathbf{h}_1^\dagger \mathbf{h}_2 = 0$, and any signal transmitted in the direction of $\mathbf{h}_3$ cannot add up rate for neither mobiles. However, we will show that the third orthogonal dimension $\mathbf{h}_3$ is in fact useful.

Let us write the antenna outputs as

$$\mathbf{x} = b_1\sqrt{P_1}\mathbf{w}_1 + b_2\sqrt{P_2}\mathbf{w}_2 + b_3\sqrt{P_3}\mathbf{w}_3 \tag{4.10}$$

where P_k is the transmit power associated with zero-mean unit-variance data symbol b_k, and $\mathbf{w}_k = \mathbf{h}_k/|\mathbf{h}_k|$ is the unit-power beamforming vector. If no signal is transmitted in the direction of $\mathbf{h}_3$, then $P_3 = 0$ which is the case with the pseudo-inverse zero-forcing. If b_3 is chosen to be an independent data symbol, then the total transmit power at the mth antenna is $\sum_{k=1}^{3} P_k|\mathbf{e}_m^\dagger \mathbf{w}_k|^2$, and clearly P_3 is wasted as b_3 cannot be received by neither mobiles ($\mathbf{w}_3$ is orthogonal to both $\mathbf{h}_1$ and $\mathbf{h}_2$). The only remaining option is to transmit either b_1 or b_2 (or a linear combination of them) in the third orthogonal dimension. For example, if $b_3 = b_1$, then the total transmit power at the mth antenna is given by $P_2|\mathbf{e}_m^\dagger \mathbf{w}_2|^2 + |\mathbf{e}_m^\dagger(\sqrt{P_1}\mathbf{w}_1 + \sqrt{P_3}\mathbf{w}_3)|^2$. This shows that, although transmitting b_1 in the direction of $\mathbf{h}_3$ cannot add up rate for mobile 1, it affects the power constraint, and may even help to save some power by coherently nulling out transmit voltage at an antenna.

For example, assume two orthogonal mobile channel vectors $\mathbf{h}_1 = (0, 2, 0)$ and $\mathbf{h}_2 = (0.8, 0, 0.6)$. In this case, the third orthogonal dimension is given by[1] $\mathbf{h}_3 = (0.6, 0, -0.8)$. When the pseudo-inverse zero-forcing is used, the unit-norm beamforming vectors are given by $\mathbf{w}_1 = (0, 1, 0)$, $\mathbf{w}_2 = (0.8, 0, 0.6)$, and

$$\mathbf{x}_{zf} = b_1 \sqrt{P_1} \mathbf{w}_1 + b_2 \sqrt{P_2} \mathbf{w}_2. \qquad (4.11)$$

Assuming that $P_{max} = 1$ W, the max-min received power is $P_{max}/(0.8)^2 \cong 1.56$ W, which is achieved when $P_1 \cong 0.39$ W, $P_2 \cong 1.56$ W. Notice that the power constraint is active only for the first transmit antenna. Instead, the following signal can be transmitted

$$\mathbf{x}' = b_1 \sqrt{P_1'} \mathbf{w}_1 + b_2 \sqrt{P_2'}(1.4\mathbf{w}_2 - 0.2\mathbf{w}_3), \qquad (4.12)$$

where $\mathbf{w}_3 = (0.6, 0, -0.8)$, $P_1' = 0.49$ W and $P_2' = 1$ W. In this case, the third orthogonal dimension is used to send the second mobile's data, which helps to increase the max-min rate received power to 1.96 W. This improvement can be explained in two ways. First, the first and the third antennas transmit at full power in this case, which improves the second mobile's received power. Second, more power is transmitted in the direction of the second mobile's channel $\mathbf{h}_2$, and the beamforming direction $\mathbf{w}_2$. This would not be feasible without $\mathbf{w}_3$, since the power constraint at the first antenna would be violated. Sending signal on $\mathbf{w}_3$ has the effect of nulling out part of the transmit voltage on the first antenna, and

---

[1]This particular numerical example is given for its simplicity, and is not intended to represent realistic channel realizations. Also, the channel values include both the magnitude and the phase components, and therefore can be negative depending on its phase.

therefore easing the power constraint. □

Theorem 1. *The beamforming based on the Moore-Penrose pseudo-inverse matrix* $\mathbf{W} = \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ *is the optimal form of downlink zero-forcing beamforming when the system is sum-power constrained.*

Proof. We will prove the theorem in the case of 2 users with single antenna receivers. The same result can be extended to multiple users as well. We have to show that each mobile's beamforming vector obtained by the pseudo-inverse zero-forcing corresponds to the part of the mobile's channel orthogonal to the other mobiles' channels. The optimality of the pseudo-inverse zero-forcing follows from the fact that, subject to a sum power constraint, the transmit beamforming is optimal, and therefore there is no need to transmit any signal in any direction that would not contribute to the data rate.

First, notice that the (i, j)th entry of the matrix $\mathbf{H}\mathbf{H}^\dagger$ is given by $\mathbf{h}_i^T \mathbf{h}_j$, which is the projection of the ith mobile's channel onto the jth mobile's channel. For the 2 user case, its inverse matrix is given by

$$(\mathbf{H}\mathbf{H}^\dagger)^{-1} = \frac{1}{|\mathbf{h}_1|^2|\mathbf{h}_2|^2 - (\mathbf{h}_1^T\mathbf{h}_2)^2} \begin{pmatrix} |\mathbf{h}_2|^2 & -\mathbf{h}_1^T\mathbf{h}_2 \\ -\mathbf{h}_1^T\mathbf{h}_2 & |\mathbf{h}_1|^2 \end{pmatrix} \qquad (4.13)$$

The first column of the matrix $\mathbf{W} = \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ is the beamforming vector $\mathbf{w}_1$ used to transmit the data symbol b_1 (similarly, $\mathbf{w}_2$ for b_2). Since the kth column of $\mathbf{H}^\dagger$ is $\mathbf{h}_k$, it follows that

$$\mathbf{w}_1 = \frac{|\mathbf{h}_2|^2}{|\mathbf{h}_1|^2|\mathbf{h}_2|^2 - (\mathbf{h}_1^T\mathbf{h}_2)^2} \left(\mathbf{h}_1 - \frac{\mathbf{h}_1^T\mathbf{h}_2}{|\mathbf{h}_2|} \frac{\mathbf{h}_2}{|\mathbf{h}_2|} \right). \qquad (4.14)$$

The term inside the parenthesis in (4.14) is the part of $\mathbf{h}_1$ that is orthogonal to $\mathbf{h}_2$. In M-dimensional space, there is a subspace defined by $M - 2$ basis vectors, which are orthogonal to both $\mathbf{w}_1$ and $\mathbf{h}_2$, and therefore can be used to transmit the data symbol b_1 without violating zero-forcing constraints (without causing interference to the second mobile). Thus, the optimal beamforming vector for the first mobile can be in the following form

$$\mathbf{w}_{opt} = \alpha_1 \mathbf{w}_1 + \sum_{\forall \mathbf{e}_k \perp \mathbf{w}_1, \mathbf{h}_2} \alpha_k \mathbf{e}_k \tag{4.15}$$

where $\mathbf{e}_k$ is a basis vector in the subspace orthogonal to $\mathbf{w}_1$ and $\mathbf{h}_2$, and for some $\alpha_1, \alpha_k \in \Re$. Notice that only the signal transmitted in the direction of $\mathbf{w}_1$ can be received by the first mobile. Moreover, when the sum-power is a concern, any signal transmitted in the direction of $\mathbf{e}_k$ would be a waste since $E[|b_1 \mathbf{w}_{opt}|^2] = p_1 \alpha_1^2 |\mathbf{w}_1|^2 + p_1 \sum_{\forall \mathbf{e}_k \perp \mathbf{w}_1, \mathbf{h}_2} \alpha_k^2 |\mathbf{e}_k|^2$ where $p_1 = E[|b_1|^2]$, and the second term in the sum does not contribute to the received power. Thus, the signal intended for the first mobile must be transmitted in the direction of $\mathbf{w}_1$, which is the beamforming direction chosen by the pseudo-inverse zero-forcing. The same analysis applies to the second mobile as well. □

4.3 Performance Evaluation

In this section, we will compare the performance of the optimum zero-forcing design with that of the zero-forcing based on the Moore-Penrose pseudo-inverse when there are per-antenna transmit power constraints. We will see that the pseudo-inverse zero-forcing performs increasingly worse compared to the optimum design as the number of transmit antennas increases.

Our basic experimental setup is as follows. We consider a downlink network where the transmitter is equipped with multiple antennas, and each mobile has a single receive antenna. This model is valid for a single cell system with multiple transmit antennas at the base station, or a multi-cell system with single or multiple antenna base stations. Each transmit antenna is assumed to have a separate power constraint. We evaluate the performance improvement due to the optimum beamforming design in a single-cell system. The total base transmit power is constrained to 10 Watts, and in the case of t transmit antennas, each antenna power will be constrained to $10/t$ Watts. We assume path-loss with a 3.8 propagation exponent, log-normal shadowing with 0 mean and 8 dB standard deviation, and Rayleigh fading for each transmit/receive antenna pair with 0 mean, unit variance complex Gaussian component. Given a mean power loss of 134 dB at the reference distance of 1.6 km from the base, a receiver noise figure of 5 dB, a vertical antenna gain of 10.3 dBi, a channel bandwidth of 5 MHz, and a receiver temperature of 300ºK, the SNR at the reference distance is 18 dB (considering one transmit/receieve antenna pair), accounting only for path loss and ignoring shadowing and Raleigh fading.

Figure 4.2 compares the spectral efficiency (averaged over many random network realizations) of the optimum zero-forcing and the pseudo-inverse zero-forcing in the case of a 2 user system. When there are only 2 transmit antennas, there is only one way of achieving the zero-forcing transmission, which is given by the channel inversion. In this case, the optimum zero-forcing design and the pseudo-inverse design have the same spectral efficiency. On the other hand, the optimum design becomes increasingly better compared to the pseudo-inverse design when

Figure 4.2: Optimum zero-forcing beamforming vs. Moore-Penrose pseudo-inverse zero-forcing: 2 users with single antenna receivers.

the number of transmit antennas increases, as it effectively uses available signal space to optimize the antenna outputs, and to reduce the transmit power at particular antennas with limited transmit power, while the pseudo-inverse zero-forcing does not have the same capability. Notice that, according to Figure 4.2, the optimum zero-forcing design requires only 8 transmit antennas to achieve a target spectral efficiency of 8 bits/second/Hz, while the pseudo-inverse design requires 16 transmit antennas for the same spectral efficiency. Figures 4.3 and 4.4 also support similar conclusions in the case of 4 users and 8 users respectively.

Figure 4.3: Optimum zero-forcing beamforming vs. Moore-Penrose pseudo-inverse zero-forcing: 4 users with single antenna receivers.

4.4 Chapter Summary and Conclusion

In this chapter, we studied the optimum zero-forcing beamforming design with per-antenna power constraints. We observe that the optimum beamforming vectors are invariant to the phase-shifts. This observation helps us to formulate convex optimization problems finding the optimum vectors. Our numerical results indicate that optimizing the antenna outputs based on the per-antenna constraints improves the rate when the number of transmit antennas is larger the number of receive antennas. In this case, the additional degrees of freedom given by more transmit antennas helps the per-antenna power constraints. In particular, the additional degrees is used to null out part of the transmit voltages at particular antennas with relatively high transmit powers. Our analysis can be

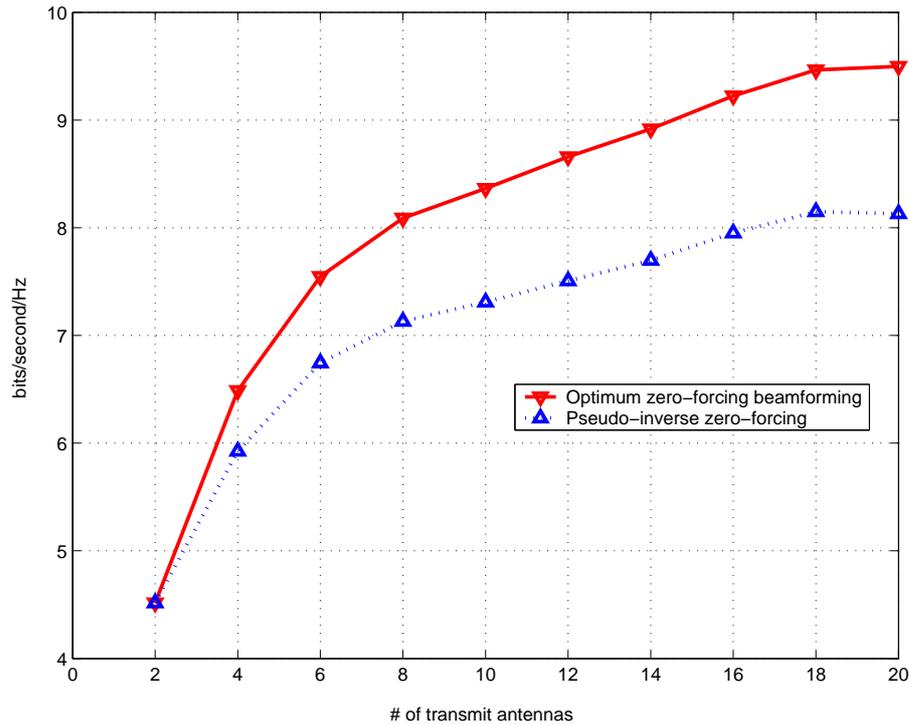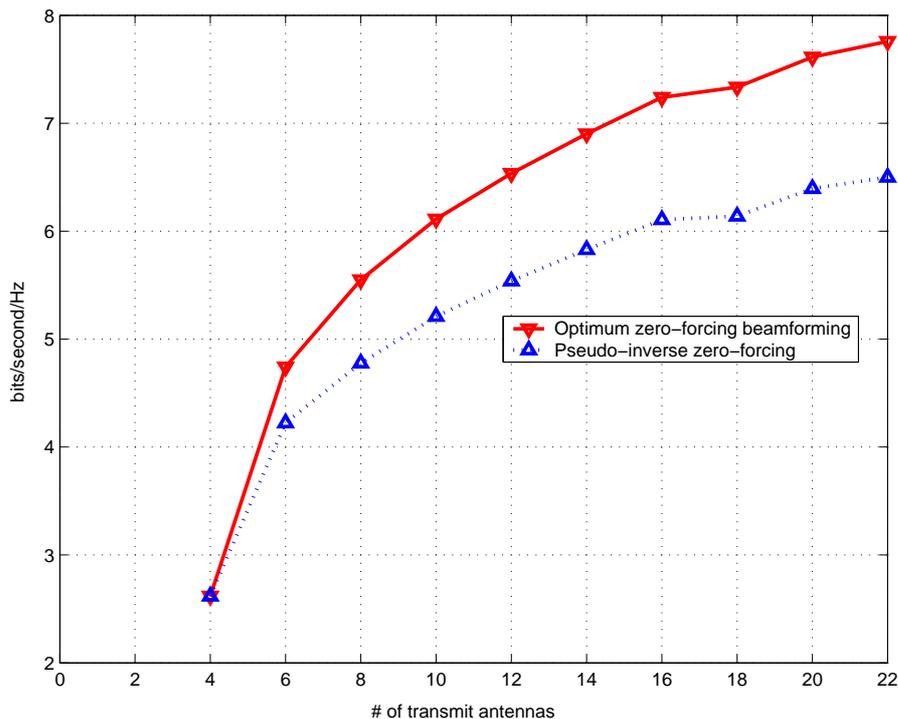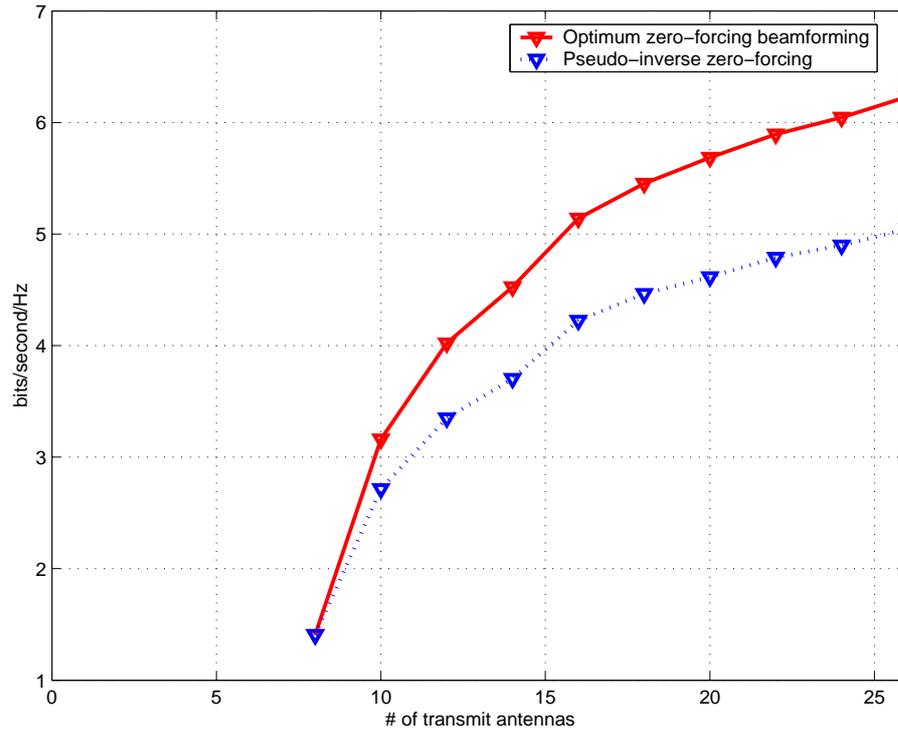Figure 4.4: Optimum zero-forcing beamforming vs. Moore-Penrose pseudo-inverse zero-forcing: 8 users with single antenna receivers.

extended to improve the performance of the zero-forcing scheme in the combined zero-forcing dirty paper coding technique of the previous two chapters. Also, we plan to extend our analysis to downlink networks with multiple antenna receivers.

Chapter 5

Cellular Backhaul Design to Enable Communication and Coordination Between the Base Stations

The techniques of the previous chapters assume the existence of a high-speed backhaul enabling communications between the base stations. In this chapter, we study the design of such a backhaul in a cellular network. More specifically, we consider a mesh backhaul network consisting of fixed base stations (mesh routers) connected by wireless links. Some of the mesh routers have connections to the wired network, and therefore can function as gateways. The network has multi-hop capability where the traffic entering the mesh backhaul through the gateway routers can be carried over multiple wireless links towards the destination mesh routers. The objective in this work is to study the capacity of such networks. Assuming the use of an OFDMA air-interface for the mesh backhaul network, we formulate a cross-layer optimization problem that involves power control, channel allocation, link scheduling and routing. Our results indicate that OFDMA-based mesh architecture provides an efficient backhaul solution in cellular networks.

5.1 Introduction

Next generation wireless networks are designed to provide broadband services with quality of service (QoS) support for various applications including voice and multimedia data. The successful deployment of these all-IP networks requires advanced network architectures with high spectral efficiency, ubiquitous coverage and cross-layer optimized system design. To achieve these goals, a service provider may deploy a large number of access points (base stations) in a given service area. On the other hand, the cost of connecting these base stations by a wired backhaul is too high, mainly because of the cable deployment cost. In this case, a wireless backhaul is necessary to enable communications between the base stations. As an efficient backhaul architecture, wireless mesh network is a promising system design providing interoperability among different wireless technologies including Wi-Fi and 3G, and its all-IP network design enables extended network coverage through multi-hop routing. In this study, we consider a mesh network backhaul architecture for a cellular system, and investigate how a mesh backhaul can improve the cellular network coverage and the capacity.

In our network architecture, we assume the use of an OFDMA air-interface for the mesh backhaul. In this case, the wideband channel available in each base-to-base link is divided into narrowband subchannels, called tones, for efficient use of spectrum in multipath environments. Our objective is to optimize the radio resources, such as transmit power and tone allocations, and study efficient routing and scheduling techniques so as to maximize the end-user throughput. As we will see, this problem is relatively complex because of the non-linear constraints involved, and the size of a typical network; hence a large number of parameters involved (flow control, link scheduling, tone and power allocations for every link

in the network). However, the base locations are fixed, and therefore the channels are relatively static and time-invariant. This fact can be used to perform network optimization in an off-line fashion.

Wireless mesh networks have been studied extensively in the literature. A survey paper summarizing the recent state of art, and open research issues can be found in [39]. A mesh architecture in a cellular network context has been considered in [40]. Given a set of scheduling modes (sets of simultaneously active links), and assuming constant transmit power, a linear programming framework is used to optimize the throughput. Joint routing and scheduling schemes have been studied to characterize the capacity of mesh networks in [41–43]. Assuming constant link capacities and no power control, and by modeling the link scheduling feasibility as an explicit constraint in the optimization problem, the authors develop efficient algorithms to find the set of achievable rates in multi-hop wireless networks. We will use the same scheduling constraint in our problem formulation to determine the set of simultaneously active OFDM tones. Joint power control, scheduling and/or routing in multi-hop adhoc networks is considered in [44–47]. In [44], average power consumption of a link is calculated by taking into account of the fraction of time the link is scheduled to be active. We will use a similar model to define the mesh node power constraints in an OFDMA network. Among the joint power control, scheduling and/or routing problems studied in the literature, [44, 46] are closest to our work. The analysis in both [44, 46] apply to single carrier systems. In the case of a single-carrier system, when a link between two nodes with bandwidth W is active, the whole bandwidth W is used for transmission. In the case of multi-carrier systems, the link between the same two nodes can be partially utilized by using a fraction of the bandwidth W (preferably part

of the spectrum with favorable channel conditions) for transmission. This enables more efficient use of spectrum in the mesh backhaul. Moreover, our scheduling model is more constrained than the model in [46] in the sense that we impose per-tone level half-duplex and multi-access/broadcast constraints as we will explain in more detail in Section 5.2.3. Comparison of different routing metrics and protocols in multi-hop wireless networks has been studied in [48, 49]. Capacity limits of adhoc networks and their scaling laws are investigated in [50, 51].

Our main contribution in this study is the analysis of the use of an OFDMA type air-interface [52] in wireless mesh networks. To the best of our knowledge, this is the first study conducting a detailed analysis of the use of OFDMA in a mesh backhaul. By providing multiple orthogonal channels in each link, OFDMA provides multi-radio advantage with a single radio in mesh networks. Moreover, when the radio resources are optimized carefully, OFDM transmissions may provide tone-diversity advantage in the form of efficient bandwidth utilization by choosing better channels for transmissions and scheduling, or in the form of improved routing performance by providing more path options to route the traffic. To study the performance limits of OFDMA mesh backhaul, and to find efficient algorithms achieving these performance limits, we formulate a cross-layer optimization problem involving power control, OFDMA resource allocation, link scheduling and routing. Using our framework, we quantify the throughput of a mesh backhaul in realistic cellular environments.

The outline of the study is as follows. In Section 5.2, we describe the system model, and explain transmission constraints in the physical layer, scheduling constraints in the link layer, and flow constraints in the routing layer. Cross-layer
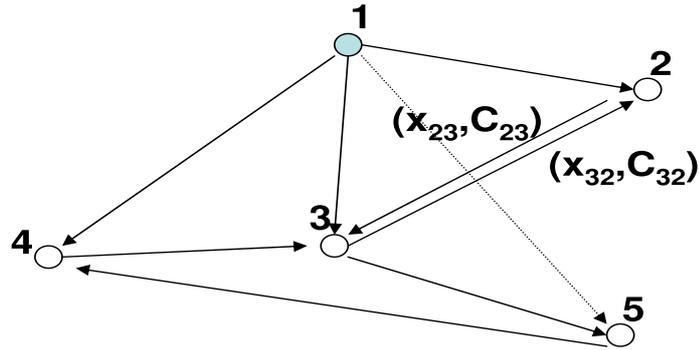
Figure 5.1: A sample network topology.

throughput optimization problem is defined in Section 5.3. The problem of identifying individual routing paths achieving the optimal throughput is studied in Section 5.4. Simulation setup, and the numerical results are given in Section 5.5. We conclude the study, and give future research directions in Section 5.6.

5.2   System Model

We consider the backhaul of a mesh network consisting of fixed base station routers connected by wireless links. Some of the mesh routers have connections to the wired network, and therefore can function as gateways. The network has multi-hop capability where the traffic entering the network through the gateway routers can be carried over multiple mesh-to-mesh wireless links towards the destination mesh client. The objective in this work is to study the capacity of such networks.

5.2.1 Mesh Network Topology

We consider the downlink of a mesh network, and aim to maximize the throughput received by each mesh node. We use a directed network flow diagram to model the mesh topology. A network flow can be mapped onto a directed graph $G = (M, A)$ consisting of a set M of nodes and a set A of links. Figure 5.1 shows a sample directed graph corresponding to a simple 5 node mesh network. In the figure, C_{ij} denotes the link capacity, and $x_{ij} \leq C_{ij}$ denotes the amount of flow on link (i, j). The set of nodes and the set of links are given by $M = \{1, 2, 3, 4, 5\}$ and $A = \{(1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (3, 2), (4, 3), (3, 5), (5, 4)\}$ respectively.

In our mesh network setup, base station routers correspond to the set M of nodes, and the wireless links connecting the mesh routers correspond to the set A of links. Each base station router i with gateway functionality is a source node with a total supply of T_i bits/sec. In this case, T_i is the total amount of traffic entering the network through node i that has to be distributed to users located in the service areas of the mesh routers. Moreover, a symmetric traffic-user distribution is assumed, which means that the total demand at each sink router is assumed to be the same. In this case, the destination nodes have uniform demands of T bits/sec. These sink nodes have the capability to route traffic for the other nodes, while the net incoming flow into each sink is greater than its net outgoing flow. Let us assume that the set of all nodes M is divided into two subsets $M = M_{so} \bigcup M_{si}$, where M_{so} contains the source nodes, and M_{si} contains the sink nodes. In Figure 5.1, $M_{so} = \{1\}$ and $M_{si} = \{2, 3, 4, 5\}$. The network may also have a number of base stations with routing capabilities only. In this case, the net traffic flow would be zero for these routers.

5.2.2 PHY Model and Transmission Constraints

We assume the use of a multi-carrier (OFDM) technology for the mesh backhaul network. Given that the total bandwidth available on each link is B Hz, OFDM transmissions divide the bandwidth into L frequency bands (tones) of size B/L Hz each. Typically, the wireless channel is frequency-selective across the wide bandwidth B because of the multi-path propagation. However, when the channel is divided into many narrowband pieces, it is possible to have a flat fading channel on each tone. In our network flow model, the existence of multiple non-interfering (orthogonal) OFDM tones between base station routers can be accounted for by having L distinct links between node pairs. Let us use a superscript f to denote each of these links, i.e., x_{ij}^f denotes the flow on tone f in use between nodes i and j.

The capacity of each link is defined as the maximum rate that can be transmitted over the link reliably. In this case, the maximum rate achievable on a particular tone f can be expressed as

$$C_{ij}^f = W \log_2(1 + \rho_{ij}^f) \tag{5.1}$$

in bits/sec where $W = B/L$ Hz, and ρ_{ij}^f is the received signal-to-interference-plus-noise ratio (SINR) on this tone. Here, the SINR is given by

$$\rho_{ij}^f = \frac{p_{ij}^f h_{ij}^f}{I_{ij}^f + \sigma^2} \tag{5.2}$$

where p_{ij}^f is the tone transmit power, h_{ij}^f is the channel fading loss (that may

include path-loss, shadow fading and Rayleigh fading), I_{ij}^f represents the total interference received by the node j on this particular tone, and σ^2 is due to the white noise. Any rate x_{ij}^f below C_{ij}^f is feasible. Notice that the factors contributing to I_{ij}^f are, first, the interference caused by the nodes $l \neq i$ trying to communicate with node j using the same tone f simultaneously, and second, the interference caused by other pairs of nodes communicating using the tone f. We will deal with the first part of the interference by designing appropriate MAC scheduling schemes. In particular, our MAC design will guarantee that multiple nodes will not transmit to a particular node using the same tone. We should emphasize that the mesh backhaul is a fixed wireless network for which the channels are relatively static and time-invariant, and therefore can be estimated pretty accurately. In this case, timing offsets and synchronization information necessary for the tone orthogonality are available. More specifically, a node receiving data from multiple nodes simultaneously can have orthogonal reception with appropriate timing offset adjustments. The second part of the interference is taken care of by the use of directional antennas between the base stations. This can be achieved by employing antenna arrays at the base stations which can beamform the data in a particular direction by adjusting the antenna weights. Moreover, frequency domain beamforming [53] can be employed so that a base station can transmit different groups of subcarriers to different base stations simultaneously.

Each node i has a total power budget P that has to be shared among all transmitting links/tones of node i. It is well-known that the water-filling policy maximizes the total rate achievable on a number of parallel/orthogonal channels subject to a total power constraint [25]. When the water-filling policy is employed, the tones with favorable channel conditions are allocated more power

compared to the ones with deep channel fades, and determining the power allocations in this case is a relatively easy task. However, in our network setup, power control is a much more challenging task. The reason is that, in a network with many source-destination pairs, a cross-layer design approach is necessary in which power control, MAC scheduling (which links/tones can be active simultaneously), and routing/flow control decisions have to be optimized jointly. For example, the set of link rates determined by the water-filling policy with the objective to maximize the sum-rate out of a particular node may not be the set of rates desired by the routing/flow control unit, i.e., the flow control unit may wish to route the traffic through a link with a relatively bad channel condition to avoid congestion in the network, or to provide a minimum service rate to a mesh router with unfavorable channel conditions, while the water-filling policy tends to avoid allocating power on such links. Moreover, there is a scheduling unit that determines which links/tones can be active simultaneously (for example, in the case of an RTS/CTS protocol in 802.11 type of MAC). In a network with many source-destination pairs, there might be many active link combinations (transmission modes). In this case, the power control becomes a difficult task since one needs to specify how the total power is distributed among these transmission modes as well as among active links/tones in each transmission mode.

5.2.3 MAC Model and Scheduling Constraints

We assume a half-duplex operation per-tone level, i.e., a node cannot receive from and transmit to other nodes using the same tone simultaneously. Moreover, a node cannot use the same tone to transmit to multiple nodes (no multicasting, broadcasting per tone level), and the link/tone transmissions are scheduled so that

multiple nodes do not transmit to the same node using the same tone. However, each node can communicate with multiple nodes simultaneously using different OFDM tones.[1]

Notice that the above conditions are related to edge-coloring of a graph. Namely, corresponding to each OFDM tone, there is a network graph with each edge representing the particular OFDM tone channel between two nodes. Edge-coloring is interested in finding the minimum number of colors needed to color the edges of a graph such that the same color are not incident on the same node [54]. In this case, the set of edges with the same color is a transmission mode describing a set of links permitted to be active simultaneously using the particular tone, and the minimum number of colors needed is the same as the number of transmission modes. In [54], Shannon derived an upper bound on the minimum number of colors needed. Using this result, it is shown [41, 55] that if the following condition is satisfied, there exists a feasible link scheduling policy satisfying the above MAC conditions:

$$\sum_{(i,j)\in A_{out}(i)} \frac{x_{ij}^f}{C_{ij}^f} + \sum_{(j,i)\in A_{in}(i)} \frac{x_{ji}^f}{C_{ji}^f} \leq \beta, \quad i \in M, \ f = 1, 2, \ldots, L, \qquad (5.3)$$

where $A_{out}(i)$ and $A_{in}(i)$ are the set of outgoing and incoming links connected to node i respectively, x_{ij}^f/C_{ij}^f is the fraction of time that tone f on the link (i,j) is active, and $\beta = 2/3$ defines a sufficient condition for a feasible schedule,

---

[1]Other types of MAC structures are also possible. For example, multiple nodes may be allowed to communicate with a destination node using the same tone simultaneously. Also, the destination node may wish to employ multi-user detection or interference cancelation at the physical layer when decoding multiple signals on the same frequency band. Our MAC structure is relatively simple, and does not require advanced receiver/transmitter architectures. Other MAC structures, in particular multi-node reception with multi-user detection and interference cancelation, is one of our future work items.

while $\beta = 1$ defines a necessary condition. Thus, when the above condition is satisfied, it is possible to have a net flow of x_{ij}^f from the base router i to the base router j using the fth OFDM tone, and there exists a scheduling policy specifying the set of simultaneously active tones that results in the set of rates in (5.3) without violating the MAC conditions explained above. Efficient heuristic graph-coloring/scheduling algorithms are available in the literature [56].

We note that providing a link with capacity C_{ij}^f requires a transmit power of $p_{ij}^f = (2^{C_{ij}^f/W} - 1)\sigma^2/h_{ij}^f$ watts, where the interference term in (5.2) from other nodes communicating using the same tone is taken care of by a combination of the MAC scheduling policy avoiding link conflicts and the use of directional antennas between base routers. Since the link is active x_{ij}^f/C_{ij}^f percent of the time, the average power $\bar{p}_{ij}^f$ consumed in this link is given by

$$\bar{p}_{ij}^f = \frac{x_{ij}^f}{C_{ij}^f} \frac{(2^{C_{ij}^f/W} - 1)\sigma^2}{h_{ij}^f} \tag{5.4}$$

In this case, the average node power constraints can expressed as:

$$\sum_{(i,j)\in A_{out}(i)} \sum_{\forall f} \frac{x_{ij}^f}{C_{ij}^f} \frac{(2^{C_{ij}^f/W} - 1)\sigma^2}{h_{ij}^f} \le P, \quad i \in M. \tag{5.5}$$

5.2.4 Routing and Flow Constraints

The network has multi-hop capability where the incoming traffic from the base routers connected to the wired backbone can be carried over multiple wireless links towards the destination mesh routers. In this case, for each source node i,

the following node flow constraint have to be satisfied:

$$\sum_{(i,j)\in A_{out}(i)} \sum_{\forall f} x_{ij}^f = T_i, \quad i \in M_{so} \tag{5.6}$$

where T_i is the total traffic to be distributed to the mesh clients through node i. Since we consider the downlink of the mesh network, the source nodes which have wired connections to the internet will not receive traffic from the wireless mesh backhaul, and will only route the traffic coming from the wired backbone. This condition, which also encourages forward routing, requires

$$\sum_{(j,i)\in A_{in}(i)} \sum_{\forall f} x_{ji}^f = 0, \quad i \in M_{so}. \tag{5.7}$$

Our objective is to maximize the incoming traffic to each sink router. Using the fact that the total demand at each sink router is T bits/sec (symmetric traffic-user distribution in the network), and together with the fact that the sink routers can help each other by multi-hop routing, the flow constraint for the sink nodes can be written as

$$\sum_{(j,i)\in A_{in}(i)} \sum_{f} x_{ji}^f - \sum_{(i,j)\in A_{out}(i)} \sum_{f} x_{ij}^f = T, \quad i \in M_{si}. \tag{5.8}$$

Finally, the total traffic from the source mesh nodes must be equal to the sum of the traffic received by the destination mesh nodes. It follows that

$$\sum_{i\in M_{so}} T_i = \sum_{i\in M_{si}} T. \tag{5.9}$$

5.3 Problem Definition and Approach

The problem is to maximize the throughput received by each destination router subject to the node power constraints, the scheduling constraints, and the flow constraints. The problem formulation is as follows:

$$\max \quad T \tag{5.10}$$

$$s.t. \quad \sum_{(j,i) \in A_{in}(i)} \sum_{f} x_{ji}^{f} - \sum_{(i,j) \in A_{out}(i)} \sum_{f} x_{ij}^{f} = T, \quad i \in M_{si}, \tag{5.10a}$$

$$\sum_{(i,j) \in A_{out}(i)} \sum_{f} x_{ij}^{f} = T_i, \quad i \in M_{so}, \tag{5.10b}$$

$$\sum_{(j,i) \in A_{in}(i)} \sum_{f} x_{ji}^{f} = 0, \quad i \in M_{so}, \tag{5.10c}$$

$$\sum_{i \in M_{so}} T_i = \sum_{i \in M_{si}} T, \tag{5.10d}$$

$$\sum_{(i,j) \in A_{out}(i)} \sum_{f} \frac{x_{ij}^{f}}{C_{ij}^{f}} \frac{(2^{C_{ij}^{f}/W} - 1)\sigma^2}{h_{ij}^{f}} \leq P, \quad i \in M, \tag{5.10e}$$

$$\sum_{(i,j) \in A_{out}(i)} \frac{x_{ij}^{f}}{C_{ij}^{f}} + \sum_{(j,i) \in A_{in}(i)} \frac{x_{ji}^{f}}{C_{ji}^{f}} \leq \frac{2}{3}, \quad i \in M, \tag{5.10f}$$

$$T, \ T_i, \ x_{ij}^{f}, C_{ij}^{f} \geq 0, \quad (i,j) \in A, \ f = 1, 2, \ldots, L. \tag{5.10g}$$

Thus, we are searching for an optimal flow through the mesh network maximizing the incoming traffic T to each sink node, while each flow x_{ij}^{f} meets the source and the sink net flow constraints (5.10a)-(5.10d), the node power constraints (5.10e), and the link scheduling constraints (5.10f). Notice that the constraint (5.10f) implies $x_{ij}^{f} \leq C_{ij}^{f}$. The solution of the above problem results in the set

of link rates between each source-destination pairs as well as the set of OFDM tones allocated, and the power assignments achieving these rates. In the above problem formulation, the objective function and the constraints (5.10a)-(5.10d) are linear, while (5.10e) and (5.10f) are non-linear constraints.

With its current form above, (5.10) is a non-linear optimization problem with non-convex constraints. In this case, approximate solutions can be obtained by heuristic approaches [44]. Our approach is to decompose the problem into two smaller but tractable pieces, while allowing partial coupling between the two steps. In the first step, we determine the link capacities by employing the water-filling policy [25] for power allocation. When the water-filling is employed at each node with the objective to maximize the node's output throughput, only a subset of the links/tones with favorable channel conditions may be selected at each node. Although this seems to be a locally efficient solution from a node's perspective, it can cause connectivity problems in the network as some of the mesh routers will not be receiving traffic because of its neighbors' greedy link selection preferences. One way to deal with this problem is to couple the water-filling power allocation with the scheduling and the routing decisions by distributing power efficiently to each link so that the network becomes fully-connected. Then, the scheduling and the routing unit determine which one of the links/tones will be active at any given time instance under the constraints on the nodes' average transmit power (5.10e). When the link capacities are determined, the remaining (scheduling/routing) problem becomes a linear program, as (5.10e)-(5.10f) becomes linear constraints. The two step hierarchical solution is summarized in Figure 5.2. We will see in our numerical examples that our hierarchical approach results in promising capacity enhancements due to multi-hop routing in the mesh

- **Step 1**: Given sets of L channels, determine the link capacities as follows:

  Input : P (initial power level), $\mathbf{h} = \{h^1, h^2, \ldots, h^L\}$ (channel responses), σ^2

  Output : $\mathbf{p} = \{p^1, p^2, \ldots, p^L\}$ (power allocations), $\mathbf{C} = \{C^1, C^2, \ldots, C^L\}$ (tone capacities)

  Initialization : $\mathbf{p} = \mathbf{0}$, $\mathbf{C} = \mathbf{0}$

  Find λ^* such that $\sum_{f=1}^{L} \max\{0, \lambda^* - \frac{\sigma^2}{h^f}\} = P$,

  $p^f = \max\{0, \lambda^* - \frac{\sigma^2}{h^f}\}$, $\quad C^f = \log_2(1 + \frac{p^f h^f}{\sigma^2})$, $\forall f$

  end

- **Step 2**: Solve the remaining Linear Program (LP).

Figure 5.2: The outline of the two steps hierarchical solution.

backhaul.

We should note that while we use our framework to obtain the maximum rate that can be delivered to all mesh routers in the backhaul, other optimization objectives, such as sum-rate maximization, can be studied using the same framework with simple changes in the objective and the constraint functions.

5.3.1 Link Capacity Assignments

In this section, we will outline different link capacity assignment techniques in a mesh network (Step 1). Among the three techniques given below, the link-based water-filling method, in which all the links in the network are allocated a non-zero capacity based on an efficient water-filling algorithm, will provide a good compromise between network connectivity and the throughput.

Spatial Water-filling

Consider the set of outgoing links from a node $i \in M$ denoted by $A_{out}(i)$. For example, in Figure 5.1, $A_{out}(1) = \{(1, 2), (1, 3), (1, 4), (1, 5)\}$. Moreover, L OFDM

tones are available for use between each of these node pairs in the set $A_{out}(i)$. The spatial water-filling distributes available node transmit power across the links in the set $A_{out}(i)$. On the other hand, each tone can only be active on one of the links in the set due to the scheduling constraints given in Section 5.2.3, and therefore L links have to be selected for L tones. The following theorem states that, in order to maximize the total rate out of a particular node, each tone must be assigned to the link which has the strongest channel on that particular tone.

Theorem 1. *Given that L orthogonal channels are available for use in N distinct links, and assuming that each orthogonal channel can be used in at most one of the N links, the optimal sum-rate maximizing strategy subject to a sum-power constraint is to pick, for each orthogonal channel, the link with the largest gain on that channel, and water-fill across the selected L link/channel pairs.*

Proof. A power allocation vector $\mathbf{p} = [p_1, p_2, \ldots, p_L]$ is feasible if $\sum_{l=1}^{L} p_l \leq P$, where P is the sum-power constraint. Consider two links h and g with channel gains h_l and g_l on the lth orthogonal channel. Also, assume $h_l > g_l$. For *any* feasible power allocation vector $\mathbf{p}$, the sum-rate improves if the channel l is assigned to link h instead of link g since $\log_2(1 + p_l h_l/\sigma^2) > \log_2(1 + p_l g_l/\sigma^2)$. Thus, the channel l must be assigned to the link with the largest gain among N links. Since the previous statement is true for any feasible power vector $\mathbf{p}$, one can first assign L orthogonal channels to the corresponding L strongest links. Given these link/channel assignments, the water-filling maximizes the sum-rate subject to a sum-power constraint [25]. □

In the spatial water-filling, each node makes a *greedy* attempt by selecting the links with the strongest channels to maximize its own throughput. When all the

nodes in the network make such local decisions, some of the mesh routers with unfavorable channel conditions may get disconnected from the mesh backhaul. The following two methods perform better than the spatial water-filling in terms of the network connectivity.

Random Link Assignments with Water-filling

In the *random link assignment* case, each tone will be assigned to a link which is randomly selected in the set $A_{out}(i)$, and the total node transmit power will be water-filled across these randomly selected link/tone pairs. In a large mesh backhaul with many routers, the random link selection will increase the probability of each mesh router to get connected to some other mesh routers. Thus, random link assignments increase the likelihood of a fully connected network, while it does not necessarily maximize each node's throughput.

Link-based Water-filling

A third strategy, *link-based water-filling*, will outperform previous two strategies, and is intended to compromise between these two approaches. In the link-based water-filling, each node pair will be considered separately, and all the links/pairs will have non-zero capacities (by water-filling power on each link). This ensures the connectivity of the network. On the other hand, tones with good channel conditions will be favored by the water-filling policy. In other words, per-link level, capacity allocations are performed efficiently. When the water-filling is performed on each link, the power level is assumed to be the node power constraint P. The rationale is that assuming N links are connected to the node i, the total node power would be NP if all the links/tones would be used at the same time.

However, only one link is active for each tone at any given time out of the N links. Roughly speaking, this would mean that the total power utilization would be $1/N$th of NP on the average. We note that the task of determining the transmission modes, i.e., the set of simultaneously active links/tones, is performed by the scheduling and the routing unit in this case (see [56] for efficient heuristic graph-coloring/scheduling algorithms). Moreover, the average node power constraint is explicitly expressed in (5.10e), and the remaining linear program will be solved under this constraint.

5.4   Identifying Individual Routing Paths

In the previous section, we were interested in finding the optimum flow assignments so as to maximize the destination throughput. While the resulting solution specifies how much traffic should be carried over each link, it does not identify which users' traffic is carried over the links. In other words, the solution does not give routing paths. In this section, we are interested in finding the routing paths for each destination mesh router.

Assume that the optimum link flow assignment for the fth OFDM tone on the link (i, j) is $x_{ij}^f = a_{ij}^f$. There are N destination routers, and a_{ij}^f may contain traffic for all of these destinations. Denoting the part of a_{ij}^f belonging to the destination node d by $a_{ij}^f(d)$, it follows that $a_{ij}^f = \sum_{d=1}^{N} a_{ij}^f(d)$. Also assume that the optimum values of T and T_i, $i \in M_{so}$, occurs a T^* and T_i^* respectively. In this case, the net flow constraints for the source nodes can be written in terms of

individual user traffic as

$$\sum_{(i,j)\in A_{out}(i)} \sum_{f} \left(\sum_{d=1}^{N} a_{ij}^{f}(d) \right) = T_i^*, \quad i \in M_{so}, \tag{5.11}$$

$$\sum_{(j,i)\in A_{in}(i)} \sum_{f} \left(\sum_{d=1}^{N} a_{ji}^{f}(d) \right) = 0, \quad i \in M_{so}. \tag{5.12}$$

For the sink routers, we need to modify the net flow constraints (5.8) to account for the fact that the only traffic that can be routed by a sink router is the traffic belonging to the others. The following constraints say that the sink routers do not route their own traffic:

$$\sum_{(j,i)\in A_{in}(i)} \sum_{f} a_{ji}^{f}(i) = T^*, \quad i \in M_{si}, \tag{5.13}$$

$$\sum_{(i,j)\in A_{out}(i)} \sum_{f} a_{ij}^{f}(i) = 0, \quad i \in M_{si}. \tag{5.14}$$

The fact that the sink routers help each other by routing each other's traffic is expressed in the following constraint:

$$\sum_{(j,i)\in A_{in}(i)} \sum_{f} a_{ji}^{f}(d) = \sum_{(i,j)\in A_{out}(i)} \sum_{\forall f} a_{ij}^{f}(d), \quad i \neq d, \; i,d \in M_{si}. \tag{5.15}$$

where the left side of (5.15) denotes the traffic belonging to the destination d, and to be routed by the node i, and the right side denotes the traffic routed. The problem of identifying individual routing paths can be solved as a linear feasibility problem subject to the constraints (5.11)-(5.15). Basically, we solve the following

linear program:

$$\text{find} \quad \mathbf{a} = [a_{ij}^f(.)]_{|A| \times L \times N} \quad (\text{or} \min 0 \times \mathbf{a}) \quad (5.16)$$

$$\text{subject to} \quad (5.11) - (5.15)$$

$$\mathbf{a} \geq 0$$

Notice that we may have more unknowns than equations as we have $|A| \times L \times N$ variables (total number of links is $|A|$, L tones are available on each link, N users may share each flow), while the number of equations is limited by the number of nodes in the network. In this case, there may be multiple solutions satisfying the constraints (5.11)-(5.15), which means that multiple alternative routing paths may exist with the same destination throughput. All of these routes are the same in terms of their power consumption and the link scheduling feasibility, as they all satisfy the constraints (5.10a)-(5.10f). The differences might exist in the number of hops required for different routing paths. Finding the paths with minimum number of hops is one of our future research directions.

5.5 Performance Evaluation

In this section, we will evaluate the capacity improvements in the mesh backhaul due to multi-hop routing and the use of OFDMA-based air-interface providing tone diversity in power allocation, scheduling and routing. Our baseline for comparison will be a mesh backhaul without multi-hop routing capability. Four network topologies will be considered: 19 cells cellular network with 2 rings of mesh routers around a source node, 37 cells cellular network with 3 rings of
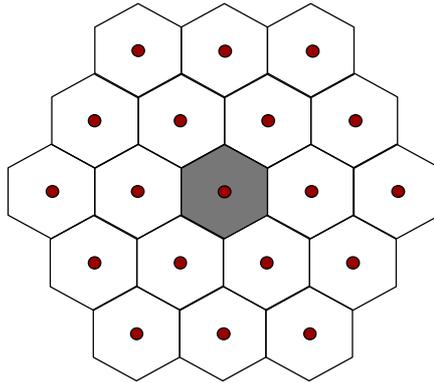
Figure 5.3: Simulation setup, 19 cells cellular network (2 rings around the source node).

mesh routers around a source node, and 37 cells cellular networks with 2 and 3 source nodes. Each base stations is located in an hexagonal cell where the distance from the center to any of the corners of the hexagon is 0.5 mile. Each base station antenna can radiate 10 Watts of transmit power on the average, in a 10MHz bandwidth with 1024 OFDM tones available for transmissions. A simple distance path-loss model is used in which the path loss PL at a distance of D mile is given by $PL = 32.6 + 35 \log_{10}(D)$ dB. An exponential power-delay profile is assumed with 0.5 μs RMS delay spread, and each channel tap experiences an independent Rayleigh fading with zero mean, unit variance complex Gaussian component. Also, log-normal shadow fading is assumed with 8 dB standard deviation. Given the RMS delay spread of the channels, the coherence bandwidth is about 2 MHz, and therefore 5 groups of independently fading tones can be seen in a 10 MHz bandwidth. To simply the numerical analysis, we will assume that the tones are grouped into 16 sub-bands, each with an independent fading realization as explained above.

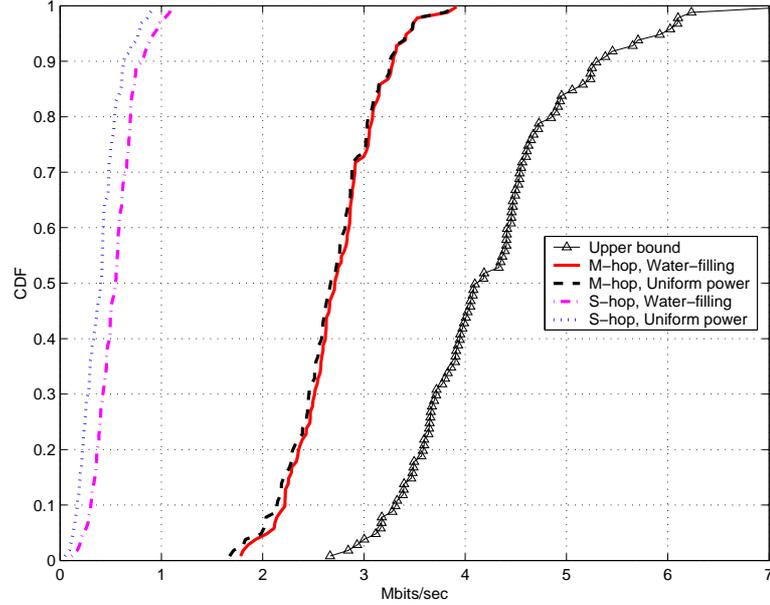For the baseline single-hop scheme, the source node transmits to only one base

Figure 5.4: Results for 19 cells cellular network (2 rings around the source node). Multi-hop throughput is based on the link-based capacity assignment method.

station at a time, and the transmissions are time-scheduled to provide uniform demands of all base stations. Denoting the rate that will be delivered to all bases by R, and the capacity of each single-hop link between the source node and destination node i by R_i, the fraction of time that has to be allocated to the base station i is given by $t_i = R/R_i$. Since $\sum_i t_i = 1$, it follows that $R = 1/\sum_i (1/R_i)$. For the single-hop multiple sources case, each destination node can be assigned to the source node with the strongest link, and the sources might time-schedule transmissions for their assigned bases. For the multi-hop mesh network, among the three link capacity assignment methods mentioned in Section 5.3.1, we will use the throughput of the link-based water-filling technique for comparison with the single-hop schemes. For both single-hop and multi-hop scenarios, two types of power allocation techniques will be considered: the water-filling power allocation, and the uniform power allocation where there is no power adaption with respect to the frequency selective channels.
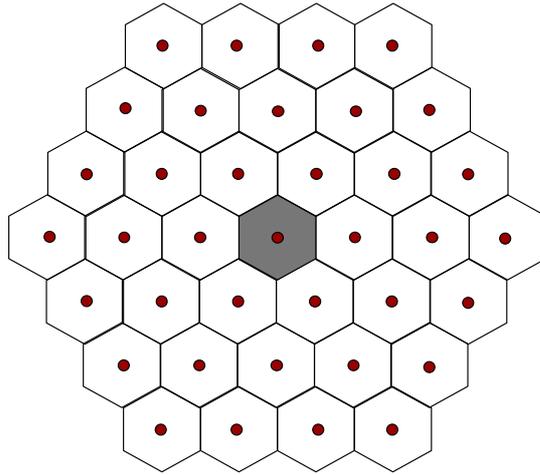
Figure 5.5: Simulation setup, 37 cells cellular network (3 rings around the source node).

We also compare the single-hop and the multi-hop results with a simple upper bound on the mesh throughput. The upper bound is obtained by the fact that the total rate delivered to all mesh routers cannot be larger than the maximum sum-rate out of the source node (sum of the source rates in the case of multiple sources). The maximum source sum-rate follows from Theorem 1. In this case, each OFDM tone is used on an outgoing link from the source node with the largest channel gain, and the total source transmit power is then water-filled across the selected tone/link pairs. The maximum mesh throughput delivered to each mesh router is the source sum-rate divided by the number of mesh routers.

Figures 5.3-5.10 show different network topologies, and the corresponding rate CDFs obtained over 100 random realizations of the channels. For all network topologies, we observe a significant capacity improvement in the mesh backhaul due to multi-hop routing. Based on the median points of the rate CDFs, multi-hop routing with water-filling power allocation improves the rate by about a factor of 5 compared to the single-hop transmissions with water-filling in 19 cells network
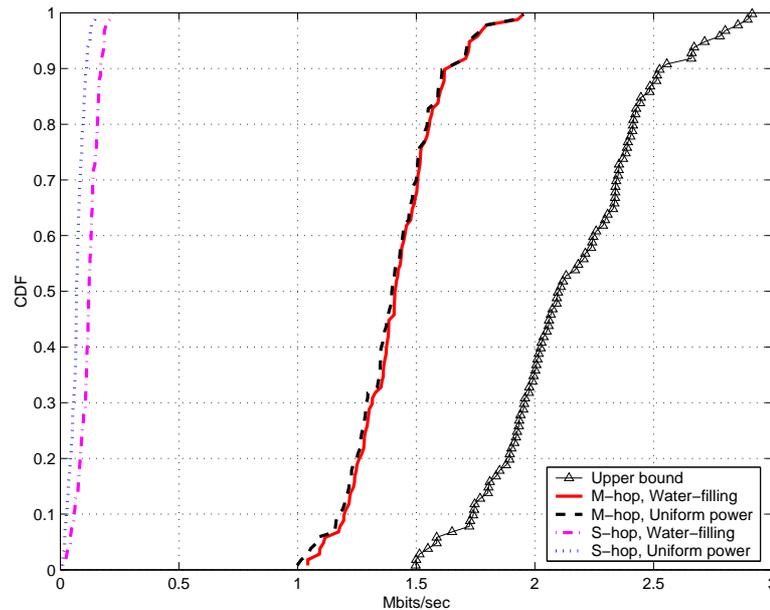
Figure 5.6: Results for 37 cells cellular network (3 rings around the source node). Multi-hop throughput is based on the link-based capacity assignment method.

topology; by about a factor of 10 in 37 cells topology with one source, and by about a factor 7 and 4 in 37 cells topology with 2 and 3 sources respectively. We observe that as transmission ranges for the single-hop schemes get larger, the single-hopping tends to be more suboptimal, and therefore the gain due to the multi-hop routing increases. Moreover, the improvement in rate due to the water-filling type power allocation as opposed to the uniform power allocation is relatively larger in the single-hop scenarios than in the multi-hop scenarios. This is due to the fact that, in the multi-hop case, most of the transmissions occur between the neighboring base stations. These links are relatively high SNR links because of relatively shorter transmission distances (compared to single-hop transmissions to 2 or 3 rings away), and hence small path-loss values. In this case, the water-filling does not improve the rate much as the power level is significantly higher than the noise level (the noise level on all frequency ranges looks flat compared to the high power level), and therefore the uniform power
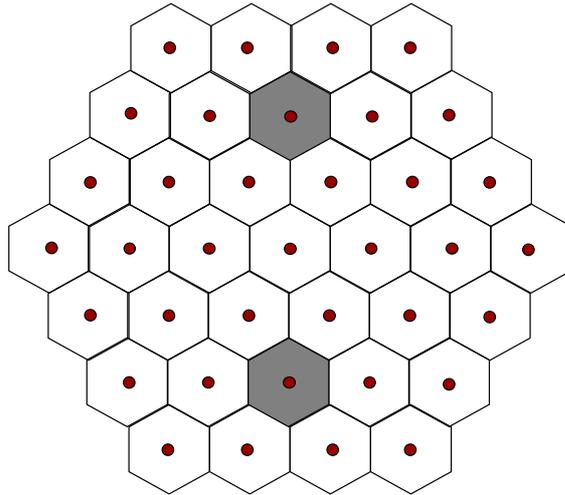
Figure 5.7: Simulation setup, 37 cells cellular network with 2 source nodes.

allocation becomes as good as the water-filling power allocation. We should note that the two-step heuristic is a simple approach to a complicated problem, yet it provides significant improvement over the single-hop throughput, and performs well when it is compared to a simple upper bound[2].

Figure 5.11 compares the throughput of an OFDMA-based mesh backhaul with a single carrier mesh network. The network topology and the channels are the same in both cases. The only difference is that the bandwidth of each link is divided into L frequency bands in the OFDMA-based architecture, while the whole bandwidth is used in the single carrier system when the link is selected to be active. Furthermore, while we use the link-based water-filling heuristic in the OFDMA-based system, the link capacities are calculated optimally in the single carrier system (by water-filling across the frequency selective channels).

---

[2]The reported multi-hop throughput results are based on $\beta = 2/3$ in (5.10f), which defines a sufficient condition for the link scheduling feasibility. Further throughput improvement is possible with the two-step hierarchical framework in some cases where only the necessary condition ($\beta = 1$) is satisfied.
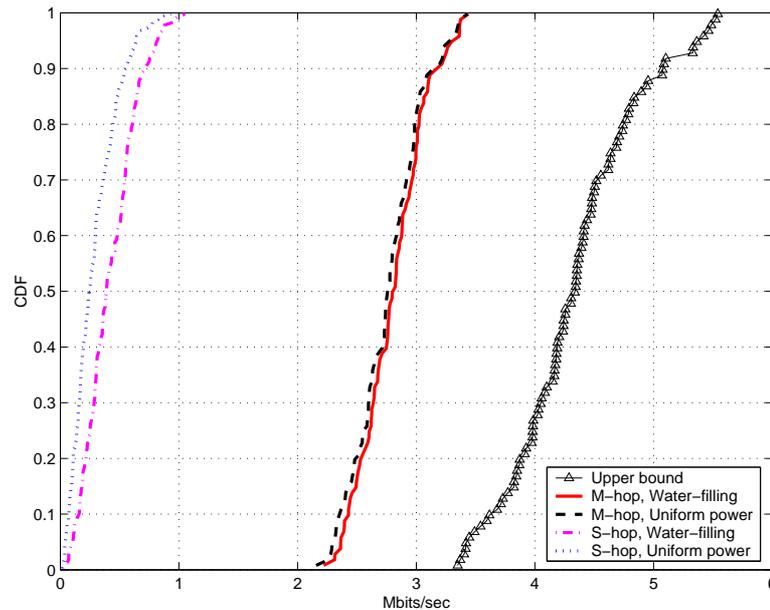
Figure 5.8: Results for 37 cells cellular network with 2 source nodes. Multi-hop throughput is based on the link-based capacity assignment method.

The figure shows that the OFDMA-based mesh backhaul has a throughput advantage over the single carrier mesh backhaul. This relative gain is due to the fact that OFDM transmissions provide tone-diversity advantage in the form of efficient bandwidth utilization by choosing better channels for transmissions and scheduling, and in the form of improved routing performance by providing more path options to route the traffic.

In Figure 5.12, we compare the three link assignment strategies. We observe that the link-based water-filling is the best strategy among the three as it effectively distributes the link power across the frequency bands by the water-filling policy, while ensuring the connectivity of the network. The spatial-water filling suffers from the connectivity problems as the nodes' local greedy link selections make some of the mesh routers get disconnected. On the other hand, when the channel delay spread increases and the coherence bandwidth gets smaller, each
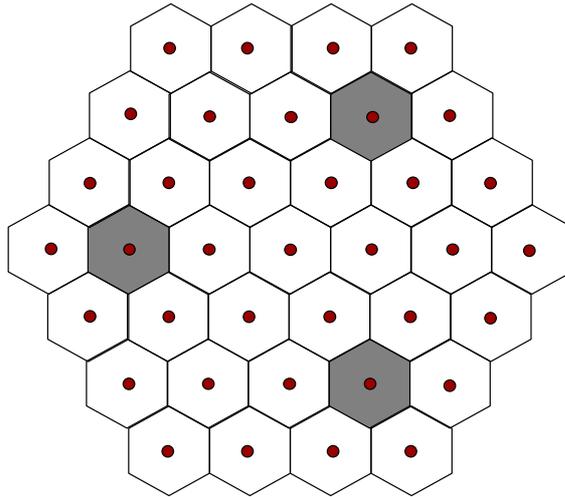
Figure 5.9: Simulation setup, 37 cells cellular network with 3 source nodes.

link would have more groups of tones with independent fades. This increased randomness in the channel provides the disconnected mesh routers with more link opportunities to get connected to the mesh backhaul.

5.6   Chapter Summary and Conclusion

Wireless mesh network is a promising network architecture providing cost-efficient backhaul solution in wireless systems. In this study, we considered an OFDMA-based mesh backhaul in a cellular base station network. In this architecture, the base stations can help each other by multi-hop routing, which not only helps avoiding congestion in the network, but also enables efficient radio resource utilization by having shorter transmission paths. Moreover, when an OFDMA air-interface is used, multiple orthogonal channels are generated on each link, which gives additional channel-diversity advantage in power allocation, scheduling and routing. In short, OFDMA-based mesh architecture provides an efficient backhaul solution in cellular networks.
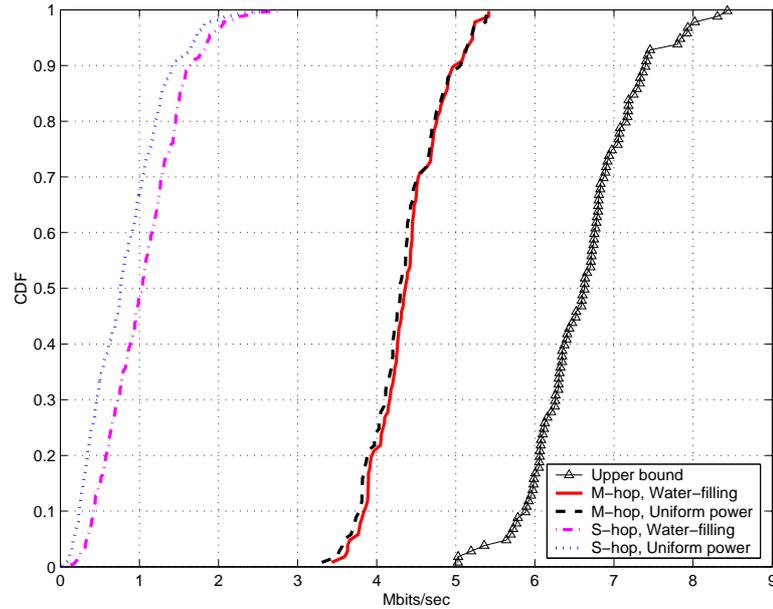
Figure 5.10: Results for 37 cells cellular network with 3 source nodes. Multi-hop throughput is based on the link-based capacity assignment method.

We should note that our framework can be used for other optimization objectives as well. For example, one of our future works includes the use of the framework to determine the best placements of mesh routers in a backhaul network. Also, we plan to have a dynamic optimization framework extension that accounts for traffic variations in the network. More accurate modeling of the interference, alternative MAC schemes, and the routing techniques that require minimum number of hops are also under investigation.
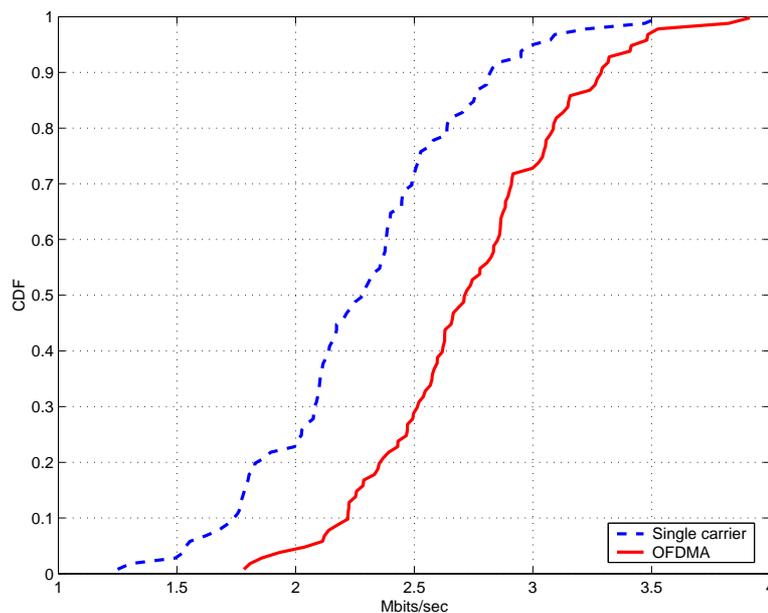
Figure 5.11: Single carrier vs OFDMA. 19 cells cellular network (2 rings around the source node).
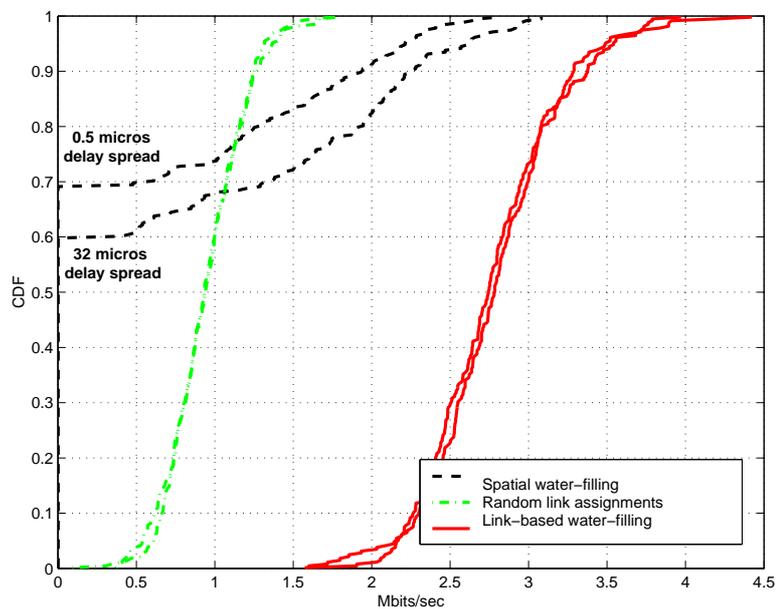

Figure 5.12: Spatial water-filling, random link assignments, link-based water-filling. 19 cells cellular network (2 rings around the source node). Results are presented for 0.5μs and 32μs delay spreads.

Chapter 6

Conclusion

Our results indicate that the network coordination is a promising technique suggesting large capacity improvements over the conventional cellular networks. Without network coordination, the downlink system capacity is limited by inter-cell interference. Furthermore, since the links often operate in the low SINR regime, the use of multiple antennas does not improve the system capacity significantly. On the other hand, when network coordination is employed, inter-cell interference is successfully eliminated. Then, the links can operate in the high SNR regime. This enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas. This strongly suggests that if the multiple antennas are to be deployed in a practical interference-limited cellular downlink system, it is very advisable to consider employing some form of coordination to eliminate inter-cell interference.

The coordination requires a high-speed backbone enabling information (data, control/syncronization and channel state) exchange between the base stations. Moreover, the coherent methods need channel information at the base stations, which could be achieved in practice by using channel estimates on the uplink in the time division duplexing (TDD) mode, or by a channel feedback from the mobiles in the frequency division duplexing mode (FDD). Also, the timing/phase synchronization is essential to achieve coherent combining of the signals from

multiple base stations. Impairments due to Doppler are another challenge to be addressed in practice. As processing power required for advanced signal processing techniques improves, implementation of complex transmission techniques such as dirty paper coding will be closer to reality.

As a future research direction, it would be interesting to have distributed, or local, implementation of network coordination in cellular networks. To achieve this, one has to identify essential part of the information (channel state, synchronization etc.) that has to be shared among the base stations to realize significant portion of the promised gains. For example, it might be enough for each base station to have the channel information of a few neighboring base stations. Also, it is important to quantify the amount of backhaul resources required to implement network coordination in a practical system. We note that network coordination is an example of cooperation in cellular systems, and the base stations can actually cooperate in many different ways. As an example, the base stations can take advantage of the bursty nature of the data transmissions by sharing on-off information to avoid interference during the bursty data arrivals. Also, spectrum allocations can be performed in a coordinated way to avoid overlapping bursty transmissions.

The base station coordination is a futuristic system design that may potentially offer a significant capacity improvement in cellular networks. Our study establishes the initial foundations of the design, and quantifies the theoretical performance gains to encourage further studies striving towards these gains in practice.

References

[1] G. J. Foschini and M. J. Gans, On limits of wireless communications in a fading environment when using multiple antennas, Wireless Personal Commun.: Kluwer Academic Press, no. 6, pp. 311-335, 1998.

[2] E. Telatar, Capacity of multi-antenna Gaussian channels, Eur. Trans. Telecomm. ETT, vol. 10, no. 6, pp. 585-596, Nov. 1999.

[3] J.Winters, On the capacity of radio communication systems with diversity in a Rayleigh fading environment, IEEE J. Select. Areas Commun., vol. 5, pp. 871-878, June 1987.

[4] A. Goldsmith, S. Jafar, N. Jindal and S. Vishwanath, Capacity limits of the MIMO Channel, IEEE Transactions on Information Theory, Vol. 21, No. 5, June 2003, pp 684-702

[5] M. J. Gans, N. Amitay, Y. S. Yeh, H. Xu, T. Damen, R. A. Valenzuela, T. Sizer, R. Storz, D. Taylor, W. M. MacDonald, C. Tran, and A. Adamiecki, Outdoor BLAST measurement system at 2.44 GHz: Calibration and initial results, IEEE J. Select. Areas Commun., vol. 20, pp. 570-581, Apr. 2002.

[6] G. J. Foschini, D. Chizhik, M. Gans, C. Papadias, and R. A. Valenzuela, Analysis and performance of some basic spacetime architectures, IEEE J. Select. Areas Commun., Special Issue on MIMO Systems, pt. I, vol. 21, pp. 303-320, Apr. 2003.

[7] G. J. Foschini, Layered space-time architecture for wireless communication in fading environments when using multi-element antennas, Bell Labs Tech. J., pp. 41-59, 1996.

[8] V. Tarokh, N. Seshardri, and A. R. Calderbank, Spacetime codes for high data rate wireless communications: Performance analysis and code construction, IEEE Trans. Inform. Theory, vol. 44, pp. 744-765, Mar. 1998.

[9] G. Rayleigh and J. M. Cioffi, Spatio-temporal coding for wireless channels, IEEE Trans. Commun., vol. 44, pp. 357-366, Mar. 1998.

[10] S. Alamouti, A Simple Transmitter Diversity Scheme for Wireless Communication, IEEE Journal on Selected Areas in Communications, Vol. 17, No 10, October 1998, pp 1451-1458.

[11] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, Iterative Water-Filling for Gaussian Vector Multiple-Access Channels, IEEE Trans. Inform. Theory, vol. 50, pp 145-152, Jan. 2004.

[12] R. Cheng and S. Verdu, Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling, IEEE Trans. Inform. Theory, vol. 39, pp. 773-785, May 1993.

[13] D. Tse and S. Hanly, Multiaccess fading channels-Part I: Polymatroid structure, optimal resource allocation and throughput capacities, IEEE Trans. Inform. Theory, vol. 44, pp. 2796-2815, Nov. 1998.

[14] P. Viswanath, D. N. Tse, and V. Anantharam, Asymptotically optimal water-filling in vector multiple-access channels, IEEE Trans. Inform. Theory, vol. 47, pp. 241-267, Jan. 2001.

[15] S. Serbetli and A. Yener, "Transceiver Optimization for Multiuser MIMO Systems", IEEE Transactions on Signal Processing, January 2004, 52(1), pp. 214-226.

[16] Wei Yu, and John Cioffi: Sum Capacity of Gaussian Vector Broadcast Channels, IEEE Transactions on Information Theory, vol. 50, no. 9, pp.1875-1892, September 2004

[17] P. Viswanath and D. N. Tse, Sum capacity of the multiple antenna Gaussian broadcast channel, in Proc. Int. Symp. Information Theory, June 2002, p. 497.

[18] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, Achievable Rates, and Sum-rate Capacity of Gaussian MIMO Broadcast Channels. *IEEE Transactions on Information Theory*, 49:2658-2668, October 2003.

[19] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz). The Capacity Region of the Gaussian MIMO Broadcast Channel. In *38th Annual Conference on Information Sciences and Systems (CISS 2004)*, March 17-19 2004.

[20] G. Caire and S. Shamai. On the Achievable Throughput of a Multiantenna Gaussian Broadcast Channel. *IEEE Transactions on Information Theory*, 49(7):1691-1706, July 2003.

[21] D. Samardzija, D. Chizhik and N. Mandayam, "Downlink Multiple Antenna Transmitter Optimization on Spatially and Temporally Correlated Channels with Delayed Channel State Information", Conf. on Info. Sciences and Systems (CISS), Princeton, 2004.

[22] D. Samardzija and N. Mandayam, "Pilot Assisted Estimation of MIMO Fading Channel Response and Achievable Data Rates", IEEE Transactions on Signal Processing, Special Issue on MIMO, Vol. 51, pp. 2882-2890, November 2003.

[23] D. Samardzija and N. Mandayam, "Unquantized and Uncoded Channel State Information Feedback on Wireless Channels", Conf. on Info. Sciences and Systems (CISS), Princeton, 2004.

[24] http://www.atheros.com/pt/whitepapers/MIMO_WLAN_Perf_whitepaper.pdf.

[25] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: Wiley, 1991.

[26] S. I. Gel'fand and M. S. Pinsker, Coding for channel with random parameters, Problems of Control and Information Theory, vol. 9, no. 1, pp. 19-31, 1980.

[27] M. Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory*, 29(3):439-441, May 1983.

[28] S. Shamai and A. D. Wyner, Information-theoretic considerations for symmetric, cellular, multiple-access fading channels, IEEE Trans. Inform. Theory, vol. 43, pp. 1877-1991, Nov. 1997.

[29] B.M. Zaidel and S. Shamai. Enhancing the Cellular Downlink Capacity via Co-processing at the Transmitting End. In *Vehicular Technology Conference*, pages 1745-1749, May 6-9 2001.

[30] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[31] W. Yu and T. Lan, "Transmitter Optimization for the Multi-Antenna Downlink with Per-Antenna Power Constraints", to appear in IEEE Transactions on Signal Processing, 2006.

[32] W. Yu and T. Lan,"Downlink beamforming with per-antenna power constraints", Sixth IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), New York City, USA. June 2005.

[33] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, Robust adaptive beamforming using worstcase performance optimization: A solution to the signal mismatch problem, IEEE Trans. Signal Process., vol. 51, no. 2, pp. 313-324, Feb. 2003.

[34] D. P. Palomar, A. Pascual-Iserte, J. M. Cioffi, and M. A. Lagunas, Convex Optimization Theory Applied to Joint Transmitter-Receiver Design in MIMO Channels, in Space-Time Processing for MIMO Communications, Chapter 8, pp. 269-318, John Wiley & Sons, ISBN 0-470-01002-9, April 2005.

[35] EURO-COST231, Urban Transmission Loss Models for Mobile Radio in the 900 and 1800 MHz Bands: The Hague, Revision 2, Sept. 1991.

[36] M. Hata, Empirical Formula for Propagation Loss in Land Mobile Radio Services, IEEE Trans. Veh. Technology, Vol. VT 29, Aug. 1980

[37] G. J. Foschini and Z. Miljanic, A Simple Distributed Autonomous Power Control Algorithm and Its Convergence, IEEE Trans. on Vehicular Tech., Vol. 42, No. 4, Nov. 1993, pp 641-646

[38] R. Yates, A Framework for Uplink Power Control in Cellular Radio Systems, IEEE Journal on Selected Areas In Communications, Vol. 13, No. 7, Sept. 1995, pp 1341-1347

[39] I. F. Akyildiz, X. Wang, and W. Wang, Wireless Mesh Networks: A Survey, Computer Networks Journal (Elsevier), vol. 47, no. 4, pp. 445-487, Mar. 2005.

[40] S. Mukherjee and H. Viswanathan, Throughput Range Tradeoff of Wireless Mesh Backhaul Networks, IEEE Journal on Selected Areas in Communications, Vol. 24, No. 3. (2006), pp. 593-602.

[41] M. S. Kodialam and T. Nandagopal, Characterizing achievable rates in multi-hop wireless mesh networks with orthogonal channels, IEEE/ACM Trans. Netw. 13(4): 868-880 (2005).

[42] M. S. Kodialam and T. Nandagopal, Characterizing the capacity region in multi-radio multi-channel wireless mesh networks, MOBICOM 2005: 73-87.

[43] M. Alicherry, R. Bhatia and L. Li, Joint Channel Assignment and Routing for Throughput Optimization in Multi-radio Wireless Mesh Networks, Proc. ACM International Conference on Mobile Computing and Networking (MobiCom), Cologne, Germany, August 2005.

[44] R. Bhatia and M. S. Kodialam, On Power Efficient Communication over Multi-hop Wireless Networks: Joint Routing, Scheduling and Power Control, INFOCOM 2004.

[45] T. ElBatt and A. Ephremides, Joint scheduling and power control for wireless ad hoc networks, IEEE Trans. Wireless Commun., vol. 3, no. 1, pp. 74-85, Jan. 2004.

[46] R. L. Cruz and A. V. Santhanam, Optimal routing, link scheduling and power control in multi-hop wireless networks, in Proc. INFOCOM, Mar./Apr. 2003, pp. 702-711.

[47] L. Xiao, M. Johansson and S. Boyd, Simultaneous routing and resource allocation via dualdecomposition, IEEE Transactions on Communications, vol. 52, pp. 1136-1144, July 2004.

[48] R. Draves, J. Padhye, and B. Zill, Routing in Multi-radio, Multi-hop Wireless Mesh Networks, ACM MobiCom, Philadelphia, PA, September 2004.

[49] R. Draves, J. Padhye, and B. Zill, Comparison of Routing Metrics for Static Multi-Hop Wireless Networks, ACM SIGCOMM, Portland, OR, August 2004.

[50] P. Gupta and P. R. Kumar, The Capacity of Wireless Networks, IEEE Transactions on Information Theory, 46(2):388-404, March 2000.

[51] S. Toumpis and A. J. Goldsmith, Capacity regions for wireless ad hoc networks, IEEE Trans. Wireless Commun., vol. 2, no. 4, pp. 736-748, Jul. 2003.

[52] L.J. Cimini, Jr., Analysis and simulation of a digital mobile radio channel using orthogonal frequency division multiplexing, IEEE Trans. Commun., vol.33, no.7, pp.665-765, July 1985.

[53] K.-W. Wong, R.S.-K. Cheng, K. Ben Letaief and R.D. Murch, "Adaptive antennas at the mobile and base stations in an OFDM/TDMA system", IEEE Trans. Comm. 49 (1), pp.195-206, Jan. 2001.

[54] C. E. Shannon, A theorem on coloring the lines of a network, J. Math. Phys., vol. 28, pp. 148-151, 1949.

[55] B. Hajek and G. Sasaki, Link scheduling in polynomial time, IEEE Trans. Inf. Theory, vol. 34, no. 5, pp. 910-917, Sep. 1988.

[56] T. Nishizeki and K. Kashiwagi, On the 1.1 edge-coloring of multi-graphs, in SIAM J. Discrete Math., vol. 3, no. 3, 1990, pp. 391-410.