

# Capacity Limits of MIMO Channels

Andrea Goldsmith, *Senior Member, IEEE*, Syed Ali Jafar, *Student Member, IEEE*, Nihar Jindal, *Student Member, IEEE*, and Sriram Vishwanath, *Student Member, IEEE*

*Invited Paper*

**Abstract**—We provide an overview of the extensive recent results on the Shannon capacity of single-user and multiuser multiple-input multiple-output (MIMO) channels. Although enormous capacity gains have been predicted for such channels, these predictions are based on somewhat unrealistic assumptions about the underlying time-varying channel model and how well it can be tracked at the receiver, as well as at the transmitter. More realistic assumptions can dramatically impact the potential capacity gains of MIMO techniques. For time-varying MIMO channels there are multiple Shannon theoretic capacity definitions and, for each definition, different correlation models and channel information assumptions that we consider. We first provide a comprehensive summary of ergodic and capacity versus outage results for single-user MIMO channels. These results indicate that the capacity gain obtained from multiple antennas heavily depends on the available channel information at either the receiver or transmitter, the channel signal-to-noise ratio, and the correlation between the channel gains on each antenna element. We then focus attention on the capacity region of the multiple-access channels (MACs) and the largest known achievable rate region for the broadcast channel. In contrast to single-user MIMO channels, capacity results for these multiuser MIMO channels are quite difficult to obtain, even for constant channels. We summarize results for the MIMO broadcast and MAC for channels that are either constant or fading with perfect instantaneous knowledge of the antenna gains at both transmitter(s) and receiver(s). We show that the capacity region of the MIMO multiple access and the largest known achievable rate region (called the dirty-paper region) for the MIMO broadcast channel are intimately related via a duality transformation. This transformation facilitates finding the transmission strategies that achieve a point on the boundary of the MIMO MAC capacity region in terms of the transmission strategies of the MIMO broadcast dirty-paper region and *vice-versa*. Finally, we discuss capacity results for multicell MIMO channels with base station cooperation. The base stations then act as a spatially diverse antenna array and transmission strategies that exploit this structure exhibit significant capacity gains. This section also provides a brief discussion of system level issues associated with MIMO cellular. Open problems in this field abound and are discussed throughout the paper.

**Index Terms**—Antenna correlation, beamforming, broadcast channels (BCs), channel distribution information (CDI), channel state information (CSI), multicell systems, multiple-access channels (MACs), multiple-input multiple-output (MIMO) channels, multiuser systems, Shannon capacity.

Manuscript received November 8, 2002; revised January 31, 2003. This work was supported in part by the Office of Naval Research (ONR) under Grants N00014-99-1-0578 and N00014-02-1-0003. The work of S. Vishwanath was supported by a Stanford Graduate Fellowship.

The authors are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: andrea@wsl.stanford.edu; syed@wsl.stanford.edu; njindal@wsl.stanford.edu; sriram@wsl.stanford.edu).

Digital Object Identifier 10.1109/JSAC.2003.810294

## I. INTRODUCTION

WIRELESS systems continue to strive for ever higher data rates. This goal is particularly challenging for systems that are power, bandwidth, and complexity limited. However, another domain can be exploited to significantly increase channel capacity: the use of multiple transmit and receive antennas. Pioneering work by Winters [81], Foschini [20], and Telatar [69] ignited much interest in this area by predicting remarkable spectral efficiencies for wireless systems with multiple antennas when the channel exhibits rich scattering and its variations can be accurately tracked. This initial promise of exceptional spectral efficiency almost “for free” resulted in an explosion of research activity to characterize the theoretical and practical issues associated with multiple-input multiple-output (MIMO) wireless channels and to extend these concepts to multiuser systems. This tutorial summarizes the segment of this recent work focused on the capacity of MIMO systems for both single-users and multiple users under different assumptions about spatial correlation and channel information available at the transmitter and receiver.

The large spectral efficiencies associated with MIMO channels are based on the premise that a rich scattering environment provides independent transmission paths from each transmit antenna to each receive antenna. Therefore, for single-user systems, a transmission and reception strategy that exploits this structure achieves capacity on approximately  $\min(M, N)$  separate channels, where  $M$  is the number of transmit antennas and  $N$  is the number of receive antennas. Thus, capacity scales linearly with  $\min(M, N)$  relative to a system with just one transmit and one receive antenna. This capacity increase requires a scattering environment such that the matrix of channel gains between transmit and receive antenna pairs has full rank and independent entries and that perfect estimates of these gains are available at the receiver. Perfect estimates of these gains at both the transmitter and receiver provides an increase in the constant multiplier associated with the linear scaling. Much subsequent work has been aimed at characterizing MIMO channel capacity under more realistic assumptions about the underlying channel model and the channel estimates available at the transmitter and receiver. The main question from both a theoretical and practical standpoint is whether the enormous capacity gains initially predicted by Winters, Foschini, and Telatar can be obtained in more realistic operating scenarios and what specific gains result from adding more antennas and/or a feedback link to feed receiver channel information back to the transmitter.

MIMO channel capacity depends heavily on the statistical properties and antenna element correlations of the channel. Recent work has developed both analytical and measurement-based MIMO channel models along with the corresponding capacity calculations for typical indoor and outdoor environments [26]. Antenna correlation varies drastically as a function of the scattering environment, the distance between transmitter and receiver, the antenna configurations, and the Doppler spread [1], [65]. As we shall see, the effect of channel correlation on capacity depends on what is known about the channel at the transmitter and receiver: correlation sometimes increases capacity and sometimes reduces it [16]. Moreover, channels with very low correlation between antennas can still exhibit a “keyhole” effect where the rank of the channel gain matrix is very small, leading to limited capacity gains [12]. Fortunately, this effect is not prevalent in most environments. The impact of channel statistics in the low-power (wideband) regime has interesting properties as well: recent results in this area can be found in [71].

We focus on MIMO channel capacity in the Shannon theoretic sense. The Shannon capacity of a single-user time-invariant channel is defined as the maximum mutual information between the channel input and output. This maximum mutual information is shown by Shannon’s capacity theorem to be the maximum data rate that can be transmitted over the channel with arbitrarily small error probability. When the channel is time-varying channel capacity has multiple definitions, depending on what is known about the channel state or its distribution at the transmitter and/or receiver and whether capacity is measured based on averaging the rate over all channel states/distributions or maintaining a constant fixed or minimum rate. Specifically, when the instantaneous channel gains, called the channel state information (CSI), are known perfectly at both transmitter and receiver, the transmitter can adapt its transmission strategy relative to the instantaneous channel state. In this case, the Shannon (ergodic) capacity is the maximum mutual information averaged over all channel states. This ergodic capacity is typically achieved using an adaptive transmission policy where the power and data rate vary relative to the channel state variations. Other capacity definitions for time-varying channels with perfect transmitter and receiver CSI include outage capacity and minimum-rate capacity. These capacities require a fixed or minimum data rate in all nonoutage channel states, which is needed for applications with delay-constrained data where the data rate cannot depend on channel variations (except in outage states, where no data is transmitted). The average rate associated with outage or minimum rate capacity is typically smaller than ergodic capacity due to the additional constraints associated with these definitions. This tutorial will focus on ergodic capacity in the case of perfect transmitter and receiver CSI.

When only the channel distribution is known at the transmitter (receiver) the transmission (reception) strategy is based on the channel distribution instead of the instantaneous channel state. The channel coefficients are typically assumed to be jointly Gaussian, so the channel distribution is specified by the channel mean and covariance matrices. We will refer to knowledge of the channel distribution as channel distribution

information (CDI). We assume throughout the paper that CDI is always perfect, so there is no mismatch between the CDI at the transmitter or receiver and the true channel distribution. When only the receiver has perfect CSI the transmitter must maintain a fixed-rate transmission strategy optimized with respect to its CDI. In this case, ergodic capacity defines the rate that can be achieved based on averaging over all channel states [69]. Alternatively, the transmitter can send at a rate that cannot be supported by all channel states: in these poor channel states the receiver declares an outage and the transmitted data is lost. In this scenario, each transmission rate has an outage probability associated with it and capacity is measured relative to outage probability<sup>1</sup> (capacity CDF) [20]. An excellent tutorial on fading channel capacity for single antenna channels can be found in [4]. For single-user MIMO channels with perfect transmitter and receiver CSI the ergodic and outage capacity calculations are straightforward since the capacity is known for every channel state. Thus, for single-user MIMO systems the tutorial will focus on capacity results assuming perfect CDI at the transmitter and perfect CSI or CDI at the receiver. Although there has been much recent progress in this area, many open problems remain.

In multiuser channels, capacity becomes a  $K$ -dimensional region defining the set of all rate vectors  $(R_1, \dots, R_K)$  simultaneously achievable by all  $K$  users. The multiple capacity definitions for time-varying channels under different transmitter and receiver CSI and CDI assumptions extend to the capacity region of the multiple-access channel (MAC) and broadcast channel (BC) in the obvious way [28], [48], [49], [70]. However, these MIMO multiuser capacity regions, even for time-invariant channels, are difficult to find. Few capacity results exist for time-varying multiuser MIMO channels, especially under the realistic assumption that the transmitter(s) and/or receiver(s) have CDI only. Therefore, the tutorial focus for MIMO multiuser systems will be on ergodic capacity under perfect CSI at the transmitter and receiver, with a brief discussion of the known results and open problems for other capacity definitions and CSI/CDI assumptions.

Note that the MIMO techniques described herein are applicable to any channel described by a matrix. Matrix channels describe not only multiantenna systems but also channels with crosstalk [85] and wideband channels [72]. While the focus of this tutorial is on memoryless channels (flat-fading), the results can also be extended to channels with memory (ISI) using well-known methods for incorporating the channel delay spread into the channel matrix [59], as will be discussed in the next section.

Many practical MIMO techniques have been developed to capitalize on the theoretical capacity gains predicted by Shannon theory. A major focus of such work is space-time coding: recent work in this area is summarized in [21]. Other techniques for MIMO systems include space-time modulation [30], [33], adaptive modulation and coding [10], space-time

<sup>1</sup>Note that an outage under perfect CSI at the receiver only is different than an outage when both transmitter and receiver have perfect CSI. Under receiver CSI only an outage occurs when the transmitted data cannot be reliably decoded at the receiver, so that data is lost. When both the transmitter and receiver have perfect CSI the channel is not used during outage (no service), so no data is lost.

TABLE I  
TABLE OF ABBREVIATIONS

CSI	Channel State Information
CDI	Channel Distribution Information
CSIT	Transmitter Channel State Information
CSIR	Receiver Channel State Information
CDIT	Transmitter Channel Distribution Information
CDIR	Receiver Channel Distribution Information
ZMSW	Zero Mean Spatially White
CMI	Channel Mean Information
CCI	Channel Covariance Information
DPC	Dirty Paper Coding
MAC	Multiple-Access Channel
BC	Broadcast Channel

equalization [2], [51], space–time signal processing [3], space–time CDMA [14], [34], and space–time OFDM [50], [52], [82]. An overview of the recent advances in these areas and other practical techniques along with their performance can be found in [25].

The remainder of this paper is organized as follows. In Section II, we discuss the capacity of single-user MIMO systems under different assumptions about channel state and distribution information at the transmitter and receiver. This section also describes the optimality of beamforming and training issues. Section III describes the capacity region of the MIMO MAC and the “dirty-paper” achievable region of the MIMO BC, along with a duality connection between these regions. The capacity of multicell systems under dirty paper coding (DPC) and opportunistic beamforming is discussed in Section IV, as well as tradeoffs between capacity, diversity, and sectorization. Section V summarizes these capacity results and describes some remaining open problems and design questions associated with MIMO systems.

A note on notation: We use boldface to denote matrices and vectors and  $\mathbb{E}[\cdot]$  for expectation.  $|\mathbf{S}|$  denotes the determinant and  $\mathbf{S}^{-1}$  the inverse of a square matrix  $\mathbf{S}$ . For any general matrix  $\mathbf{M}$ ,  $\mathbf{M}^\dagger$  denotes the conjugate transpose and  $\text{Tr}(\mathbf{M})$  denotes the trace.  $\mathbf{I}$  denotes the identity matrix and  $\text{diag}(\lambda_i)$  denotes a diagonal matrix with the  $(i, i)$  entry equal to  $\lambda_i$ . For symmetric matrices the notation  $\mathbf{Q} \geq 0$  implies that  $\mathbf{Q}$  is positive semidefinite.

A table of abbreviations used throughout the paper is given in Table I.

## II. SINGLE-USER MIMO

In this section, we focus on the capacity of single-user MIMO channels. While most wireless systems today support multiple users, single-user results are still of much interest for the insight they provide and their application to channelized systems, where users are allocated orthogonal resources (time, frequency bands, etc.). MIMO channel capacity is also much easier to derive for single users than for multiple users. Indeed, single-user

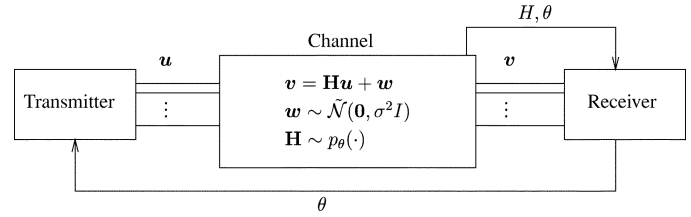


Fig. 1. MIMO channel with perfect CSIR and distribution feedback.

MIMO capacity results are known for many cases, where the corresponding multiuser problems remain unsolved. In particular, very little is known about multiuser capacity without the assumption of perfect channel state information at the transmitter (CSIT) and at the receiver (CSIR). While there remain many open problems in obtaining the single-user capacity under general assumptions of CSI and CDI, for several interesting cases the solution is known. This section will give an overview of known results for single-user MIMO channels with particular focus on special cases of CDI at the transmitter, as well as the receiver. We begin with a description of the channel model and the different CSI and CDI models we consider, along with their motivation.

### A. Channel Model

Consider a transmitter with  $M$  transmit antennas and a receiver with  $N$  receive antennas. The channel can be represented by the  $N \times M$  matrix  $\mathbf{H}$ . The  $N \times 1$  received signal  $\mathbf{v}$  is equal to

$$\mathbf{v} = \mathbf{H}\mathbf{u} + \mathbf{w} \quad (1)$$

where  $\mathbf{u}$  is the  $M \times 1$  transmitted vector and  $\mathbf{w}$  is the  $N \times 1$  additive white circularly symmetric complex Gaussian noise vector, normalized so that its covariance matrix is the identity matrix. The normalization of any nonsingular noise covariance matrix  $\mathbf{K}_w$  to fit the above model is as straightforward as multiplying the received vector  $\mathbf{u}$  with  $\mathbf{K}_w^{-1/2}$  to yield the effective channel  $\mathbf{K}_w^{-1/2}\mathbf{H}$  and a white noise vector.

The CSI is the channel matrix  $\mathbf{H}$ . Thus, with perfect CSIT or CSIR, the channel matrix  $\mathbf{H}$  is assumed to be known perfectly and instantaneously at the transmitter or receiver, respectively. When the transmitter or receiver knows the channel state perfectly, we also assume that it knows the distribution of this state perfectly, since the distribution can be obtained from the state observations.

1) *Perfect CSIR and CDIT*: The perfect CSIR and CDIT model is motivated by the scenario where the channel state can be accurately tracked at the receiver and the statistical channel model at the transmitter is based on CDI fed back from the receiver. This distribution model is typically based on receiver estimates of the channel state and the uncertainty associated with these estimates. Fig. 1 illustrates the underlying communication model in this scenario, where  $\tilde{\mathcal{N}}$  denotes the complex Gaussian distribution.

The salient features of the model are as follows.

- Conditioned on the parameter  $\theta$  that defines the channel distribution, the channel realizations  $H$  at different time instants are independent identically distributed (i.i.d.).

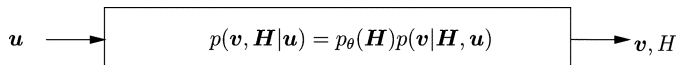


Fig. 2. MIMO channel with perfect CSIR and CDIT ( $\theta$  fixed).

- In a wireless system the channel statistics change over time due to mobility of the transmitter, receiver, and the scattering environment. Thus,  $\theta$  is time-varying.
- The statistical model depends on the time scale of interest. For example, in the short term, the channel coefficients may have a nonzero mean and one set of correlations reflecting the geometry of the particular propagation environment. However, over a long term the channel coefficients may be described as zero-mean and uncorrelated due to the averaging over several propagation environments. For this reason, uncorrelated, zero-mean channel coefficients is a common assumption for the channel distribution in the absence of distribution feedback or when it is not possible to adapt to the short-term channel statistics. However, if the transmitter receives frequent updates of  $\theta$  and it can adapt to these time-varying short-term channel statistics then capacity is increased relative to the transmission strategy associated with just the long-term channel statistics. In other words, adapting the transmission strategy to the short-term channel statistics increases capacity. In the literature adaptation to the short-term channel statistics (the feedback model of Fig. 1) is referred to by many names including mean and covariance feedback, imperfect feedback and partial CSI [38], [40], [42], [45], [46], [56], [66], [76].
- The feedback channel is assumed to be free from noise. This makes the CDIT a deterministic function of the CDIR and allows optimal codes to be constructed directly over the input alphabet [8].
- For each realization of  $\theta$  the conditional average transmit power is constrained as  $\mathbb{E}[\|\mathbf{u}\|^2 | \Theta = \theta] \leq P$ .
- The ergodic capacity  $C$  of the system in Fig. 1 is the capacity  $C(\theta)$  averaged over the different  $\theta$  realizations

$$C = \mathbb{E}_\theta[C(\theta)]$$

where  $C(\theta)$  is the ergodic capacity of the channel shown in Fig. 2. This figure represents a MIMO channel with perfect CSI at the receiver and only CDI about the *constant* distribution  $\theta$  at the transmitter. Channel capacity calculations generally implicitly assume CDI at both the transmitter and receiver except for special channel classes, such as the compound channel or arbitrarily varying channel. This implicit knowledge of  $\theta$  is justified by the fact that the channel coefficients are typically modeled based on their long-term average distribution. Alternatively,  $\theta$  can be obtained by the feedback model of Fig. 1. Thus, motivated by the distribution feedback model of Fig. 1, we will provide capacity results for the system model of Fig. 2 under different distribution ( $\theta$ ) models. For clarity, we explicitly state when CDI is available at either the transmitter or receiver, to contrast with the case where CSI is also available.

Computation of  $C(\theta)$  for general  $p_\theta(\cdot)$  is a hard problem. Almost all research in this area has focused on three special cases for this distribution: zero-mean spatially white channels, spatially white channels with nonzero mean, and zero-mean channels with nonwhite channel covariance. In all three cases, the channel coefficients are modeled as complex jointly Gaussian random variables. Under the zero-mean spatially white (ZMSW) model, the channel mean is zero and the channel covariance is modeled as white, i.e., the channel elements are assumed to be i.i.d. random variables. This model typically captures the long-term average distribution of the channel coefficients averaged over multiple propagation environments. Under the channel mean information (CMI) model, the mean of the channel distribution is nonzero while the covariance is modeled as white with a constant scale factor. This model is motivated by a system where the channel state is measured imperfectly at the transmitter, so the CMI reflects this measurement and the constant factor reflects the estimation error. Under the channel covariance information (CCI) model, the channel is assumed to be varying too rapidly to track its mean, so the mean is set to zero and the information regarding the relative geometry of the propagation paths is captured by a nonwhite covariance matrix. Based on the underlying system model shown in Fig. 1, in the literature the CMI model is also called mean feedback and the CCI model is also called covariance feedback. Mathematically, the three distribution models for  $\mathbf{H}$  can be described as follows:

**Zero-Mean Spatially White (ZMSW):**  $\mathbb{E}[\mathbf{H}] = \mathbf{0}$ ,  $\mathbf{H} = \mathbf{H}^w$ ;  
**Channel Mean Information (CMI):**  $\mathbb{E}[\mathbf{H}] = \bar{\mathbf{H}}$ ,  $\mathbf{H} = \bar{\mathbf{H}} + \sqrt{\alpha}\mathbf{H}^w$ ;  
**Channel Covariance Information (CCI):**  $\mathbb{E}[\mathbf{H}] = \bar{\mathbf{0}}$ ,  $\mathbf{H} = (\mathbf{R}^r)^{1/2}\mathbf{H}^w(\mathbf{R}^t)^{1/2}$ .

Here,  $\mathbf{H}^w$  is an  $N \times M$  matrix of i.i.d. zero mean, unit variance complex circularly symmetric Gaussian random variables. The channel mean  $\bar{\mathbf{H}}$  and  $\alpha$  are constants that may be interpreted as the channel estimate based on the feedback and the variance of the estimation error, respectively.  $\mathbf{R}^r$  and  $\mathbf{R}^t$  are called the receive and transmit fade covariance matrices. Although not completely general, this simple correlation model has been validated through recent field measurements as a sufficiently accurate representation of the fade correlations seen in actual cellular systems [13]. Under CMI the channel mean  $\bar{\mathbf{H}}$  and the variance of the estimation error  $\alpha$  are assumed known and under CCI the transmit and receive covariance matrices  $\mathbf{R}^r$  and  $\mathbf{R}^t$  are assumed known.

2) *CDIT and CDIR:* In highly mobile channels, the assumption of perfect CSI at the receiver can be unrealistic. Thus, we now consider a model where both transmitter and receiver only have information about the channel distribution. Even for a rapidly fluctuating channel where reliable channel estimation is not possible, it might be possible for the receiver to track the short-term distribution of the channel fades, as the channel distribution changes much more slowly than the channel itself. The estimated distribution can be made available to the transmitter through a feedback channel. Fig. 3 illustrates the underlying communication model.

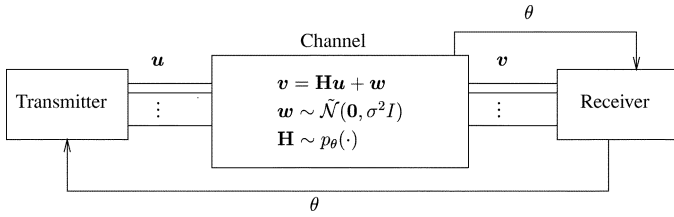
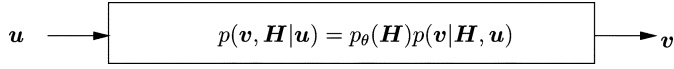


Fig. 3. MIMO channel with CDIR and distribution feedback.

Fig. 4. MIMO channel with CDIT and CDIR ( $\theta$  fixed).

Note that the estimation of the channel statistics at the receiver is captured in the model as a genie that provides the receiver with the correct channel distribution. The feedback channel represents the same information being made available to the transmitter simultaneously. This model is slightly optimistic because in practice the receiver estimates  $\theta$  only from the received signal  $\mathbf{v}$  and therefore will not have a perfect estimate.

As in the previous section, the ergodic capacity turns out to be the expected value (expectation over  $\theta$ ) of the ergodic capacity  $C(\theta)$ , where  $C(\theta)$  is the ergodic capacity of the channel in Fig. 4. In this figure,  $\theta$  is constant and known at both the transmitter and receiver (CDIT and CDIR). As in the previous section, the computation of  $C(\theta)$  is difficult for general  $\theta$ , so we restrict ourselves to the same three channel distribution models described in the previous subsection: the ZMSW, CMI, and CCI models.

Next, we summarize the single-user MIMO capacity results under various assumptions on CSI and CDI.

### B. Constant MIMO Channel Capacity

When the channel is constant and known perfectly at the transmitter and the receiver, the capacity is

$$C = \max_{\mathbf{Q}: \text{Tr}(\mathbf{Q})=P} \log |\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger| \quad (2)$$

where  $\mathbf{Q}$  is the input covariance matrix. Telatar [69] showed that the MIMO channel can be converted to parallel, noninterfering single-input single-output (SISO) channels through a singular value decomposition (SVD) of the channel matrix. The SVD yields  $\min(M, N)$  parallel channels with gains corresponding to the singular values  $\sigma_i^2$  of  $\mathbf{H}$ . Waterfilling the transmit power over these parallel channels leads to the power allocation

$$P_i = \left( \mu - \frac{1}{\sigma_i^2} \right)^+, \quad 1 \leq i \leq \min(M, N) \quad (3)$$

where  $\mu$  is the waterfill level,  $P_i$  is the power in the  $i$ th eigenmode of the channel and  $x^+$  is defined as  $\max(x, 0)$ . The channel capacity is shown to be

$$C = \sum_i^{\min(M, N)} (\log(\mu \sigma_i^2))^+. \quad (4)$$

Although the constant channel model is relatively easy to analyze, wireless channels in practice are not fixed or constant.

Instead, due to the changing propagation environment wireless channels vary over time, assuming values over a continuum. The capacity of fading channels is investigated next.

### C. Fading MIMO Channel Capacity

With slow fading, the channel may remain approximately constant long enough to allow reliable estimation of the channel state at the receiver (perfect CSIR) and timely feedback of this state information to the transmitter (perfect CSIT). However, in systems with moderate to high user mobility, the system designer is inevitably faced with channels that change rapidly. Fading models where only the channel distribution is available to the receiver (CDIR) and/or transmitter (CDIT) are more applicable to such systems. Capacity results under various assumptions regarding CSI and CDI are summarized in this section.

1) *Capacity With Perfect CSIT and Perfect CSIR:* Perfect CSIT and perfect CSIR model a fading channel that changes slow enough to be reliably measured by the receiver and fed back to the transmitter without significant delay. The ergodic capacity of a flat-fading channel with perfect CSIT and CSIR is simply the average of the capacities achieved with each channel realization. The capacity for each channel realization is given by the constant channel capacity expression in the previous section. Thus, the fading MIMO channel capacity assuming perfect channel knowledge at both transmitter and receiver is

$$C = \mathbb{E}_{\mathbf{H}} \left[ \max_{\mathbf{Q}: \text{Tr}(\mathbf{Q})=P} \log |\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger| \right]. \quad (5)$$

2) *Capacity With Perfect CSIR and CDIT:* *ZMSW Model:* Seminal work by Foschini and Gans [22] and Telatar [69] addressed the case of perfect CSIR and a ZMSW channel distribution at the transmitter. Recall that in this case, the channel matrix  $\mathbf{H}$  is assumed to have i.i.d. complex Gaussian entries (i.e.,  $\mathbf{H} \sim \mathbf{H}^w$ ). As described in the introduction, the two relevant capacity definitions in this case are capacity versus outage (capacity CDF) and ergodic capacity. For any given input covariance matrix the input distribution that achieves the ergodic capacity is shown in [22] and [69] to be complex vector Gaussian, mainly because the vector Gaussian distribution maximizes the entropy for a given covariance matrix. This leads to the transmitter optimization problem—i.e., finding the optimum input covariance matrix to maximize ergodic capacity subject to a transmit power (trace of the input covariance matrix) constraint. Mathematically, the problem is to characterize the optimum  $\mathbf{Q}$  to maximize

$$C = \max_{\mathbf{Q}: \text{Tr}(\mathbf{Q})=P} C(\mathbf{Q}) \quad (6)$$

where

$$C(\mathbf{Q}) \triangleq \mathbb{E}_{\mathbf{H}} [\log |\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger|] \quad (7)$$

is the mutual information with the input covariance matrix  $\mathbb{E}[\mathbf{u}\mathbf{u}^\dagger] = \mathbf{Q}$  and the expectation is with respect to the channel matrix  $\mathbf{H}$ . The mutual information  $C(\mathbf{Q})$  is achieved by transmitting independent complex circular Gaussian symbols along the eigenvectors of  $\mathbf{Q}$ . The powers allocated to each eigenvector are given by the eigenvalues of  $\mathbf{Q}$ .

It is shown in [22] and [69] that the optimum input covariance matrix that maximizes ergodic capacity is the scaled identity matrix, i.e., the transmit power is divided equally among all the transmit antennas. Thus, the ergodic capacity is given by

$$C = \mathbb{E}_{\mathbf{H}} \left[ \log \left| \mathbf{I}_N + \frac{P}{M} \mathbf{H} \mathbf{H}^\dagger \right| \right]. \quad (8)$$

An integral form of this expectation involving Laguerre polynomials is derived in [69]. If  $M$  and  $N$  simultaneously become large, capacity is seen to grow *linearly* with  $\min(M, N)$ . Expressions for the growth rate constant can be found in [32] and [69].

Telatar [69] conjectures that the optimal input covariance matrix that maximizes capacity versus outage is a diagonal matrix with the power equally distributed among a *subset* of the transmit antennas. The principal observation is that as the capacity CDF becomes steeper, capacity versus outage increases for low outage probabilities and decreases for high outage probabilities. This is reflected in the fact that the higher the outage probability, the smaller the number of transmit antennas that should be used. As the transmit power is shared equally between more antennas the expectation of  $C$  increases (so the ergodic capacity increases) but the tails of its distribution decay faster. While this improves capacity versus outage for low outage probabilities, the capacity versus outage for high outages is decreased. Usually, we are interested in low outage probabilities<sup>2</sup> and, therefore, the usual intuition for outage capacity is that it increases as the diversity order of the channel increases, i.e., as the capacity CDF becomes steeper. Foschini and Gans [22] also propose a layered architecture to achieve these capacities with scalar codes. This architecture, called Bell Labs Layered Space-Time (BLAST), shows enormous capacity gains over single antenna systems. For example, at 1% outage, 12 dB signal-to-noise ratio (SNR) and with 12 antennas, the spectral efficiency is shown to be 32 b/s/Hz as opposed to the spectral efficiencies of around 1 b/s/Hz achieved in present day single antenna systems. While the channel models in [22] and [69] assume uncorrelated and frequency flat fading, practical channels exhibit both correlated fading, as well as frequency selectivity. The need to estimate the capacity gains of BLAST for practical systems in the presence of channel fade correlations and frequency selective fading sparked off measurement campaigns reported in [24] and [55]. The measured capacities are found to be about 30% smaller than would be anticipated from an idealized model. However, the capacity gains over single antenna systems are still overwhelming.

3) *Capacity With Perfect CSIR and CDIT: CMI and CCI Models:* Recent results indicate that for MIMO channels the capacity improvement resulting from some knowledge of the short-term channel statistics at the transmitter can be substantial. These results have ignited much interest in the capacity of MIMO channels with perfect CSIR and CDIT under general distribution models. In this section, we focus on the cases of CMI and CCI channel distributions, corresponding to distribution feedback of the channel mean or covariance

matrix. Key results on the capacity of such channels have been recently obtained by several authors including Madhoo and Visotsky [76], Trott and Narula [58], [57], Jafar and Goldsmith [42], [40], [38], Jorsweick and Boche [45], [46], and Simon and Moustakas [56], [66].

Mathematically the problem is defined by (6) and (7), with the distribution on  $\mathbf{H}$  determined by the CMI or CCI. The optimum input covariance matrix in general can be a full rank matrix which implies either vector coding across the antenna array or transmission of several scalar codes in parallel with successive interference cancellation at the receiver. Limiting the rank of the input covariance matrix to unity, called *beamforming*, essentially leads to a scalar coded system which has a significantly lower complexity for typical array sizes.

The complexity versus capacity tradeoff is an interesting aspect of capacity results under CDIT. The ability to use scalar codes to achieve capacity under CDIT for different channel distribution models, also called optimality of beamforming, captures this tradeoff and has been the topic of much research in itself. Note that vector coding refers to fully unconstrained signaling schemes for the memoryless MIMO Gaussian channel. Every symbol period, a channel use corresponds to the transmission of a vector symbol comprised of the inputs to each transmit antenna. Ideally, while decoding vector codewords the receiver needs to take into account the dependencies in both space and time dimensions and therefore the complexity of vector decoding grows exponentially in the number of transmit antennas. A lower complexity implementation of the vector coding strategy is also possible in the form of several scalar codewords being transmitted in parallel. It is shown in [38] that without loss of capacity, any input covariance matrix, regardless of its rank, can be treated as several scalar codewords encoded independently at the transmitter and decoded successively at the receiver by subtracting out the contribution from previously decoded codewords at each stage. However, well-known problems associated with successive decoding and interference subtraction, e.g., error propagation, render this approach unsuitable for use in practical systems. It is in this context that the question of optimality of beamforming becomes important. Beamforming transforms the MIMO channel into a single-input single-output (SISO) channel. Thus, well established scalar codec technology can be used to approach capacity and since there is only one beam, interference cancellation is not needed. In the summary given below, we include the results on both the transmitter optimization problem, as well as the optimality of beamforming.

*Multiple-Input Single-Output (MISO) Channels:* We first consider systems that use a single receive antenna and multiple transmit antennas. The channel matrix is rank one. With perfect CSIT and CSIR, for every channel matrix realization it is possible to identify the only nonzero eigenmode of the channel accurately and beamform along that mode. On the other hand, with perfect CSIR and CDIT under the ZMSW model, it was shown by Foschini and Gans [22] and Telatar [69] that the optimal input covariance matrix is a multiple of the identity matrix. Thus, the inability of the transmitter to identify the nonzero channel eigenmode forces a strategy, where the power is equally distributed in all directions.

<sup>2</sup>The capacity for high outage probabilities becomes relevant for schemes that transmit only to the best user. For such schemes, it is shown in [6] that increasing the number of transmit antennas reduces the average sum capacity.

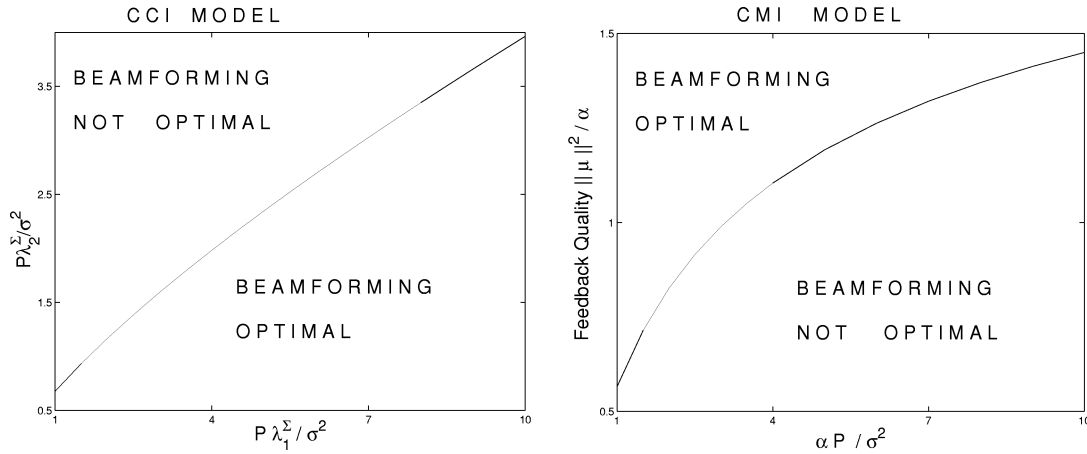


Fig. 5. Plot of necessary and sufficient conditions (9). <Author: Fig. 5 not cited in text>

For a system using a single receive antenna and multiple transmit antennas, the transmitter optimization problem under CSIR and CDIT is solved by Visotsky and Madhoo in [76] for the distribution models of CMI and CCI. For the CMI model ( $\mathbf{H} \sim \tilde{\mathcal{N}}(\bar{\mathbf{H}}, \alpha \mathbf{I})$ ) the principal eigenvector of the optimal input covariance matrix  $\mathbf{Q}^o$  is found to be along the channel mean vector and the eigenvalues corresponding to the remaining eigenvectors are shown to be equal. When beamforming is optimal, all power is allocated to the principal eigenvector. For the CCI model ( $\mathbf{H} \sim \tilde{\mathcal{N}}(\mathbf{0}, \mathbf{R}^t)$ ) the eigenvectors of the optimal input covariance matrix  $\mathbf{Q}^o$  are shown to be along the eigenvectors of the transmit fade covariance matrix and the eigenvalues are in the same order as the corresponding eigenvalues of the transmit fade covariance matrix. Moreover, Visotsky and Madhoo's numerical results indicate that beamforming is close to the optimal strategy when the quality of feedback improves, i.e., when the channel uncertainty  $\alpha$  decreases under CMI or when a stronger channel mode can be identified under CCI. We will discuss quality of feedback in more detail below. Under CMI, Narula and Trott [58] point out that there are cases where the capacity is actually achieved via beamforming. While they do not obtain fully general necessary and sufficient conditions for when beamforming is a capacity achieving strategy, they develop partial answers to the problem for two transmit antennas.

A general condition that is both necessary and sufficient for optimality of beamforming is obtained by Jafar and Goldsmith in [40] for both the CMI and CCI models. The result can be stated as follows.

The ergodic capacity can be achieved with a unit rank matrix if and only if the following condition is true:

$$\mathbb{E}_{w_1} \left[ \frac{1}{1 + P\lambda_1 w_1} \right] \leq \frac{1}{1 + P\lambda_2} \quad (9)$$

where for the CCI model

- 1)  $\lambda_1 > \lambda_2$  are the two largest eigenvalues of the channel fade covariance matrix  $\mathbf{R}^t$ ;
- 2)  $w_1$  is exponential distributed with unit mean, i.e.,  $w_1 \sim e^{-w_1}$ ;

and for the CMI model

- 1)  $\lambda_1 = \lambda_2 = \alpha$ ;
- 2)  $w_1$  has a noncentral chi-squared distribution. More precisely,  $w_1 \sim e^{-\|\mu\|^2/\alpha} I_0 \left( 2\|\mu\| \sqrt{w_1/\alpha} \right)$  where  $I_0(\cdot)$  is the zeroth-order modified Bessel function of the first kind.

Further, for the CCI model the expectation can be evaluated to express (9) explicitly in closed form as

$$\frac{1}{P\lambda_1} e^{1/P\lambda_1} \Gamma(0, \frac{1}{P\lambda_1}) \leq \frac{1}{1 + P\lambda_2}. \quad (10)$$

The optimality conditions are plotted in Fig. 5. For the CCI model the optimality of beamforming depends on the two largest eigenvalues  $\lambda_1, \lambda_2$  of the transmit fade covariance matrix and the transmit power  $P$ . Beamforming is found to be optimal when the two largest eigenvalues of the transmit covariance matrix are sufficiently disparate or the transmit power  $P$  is sufficiently low. Since beamforming corresponds to using the principal eigenmode alone, this is reminiscent of waterpouring solutions where only the deepest level gets all the water when it is sufficiently deeper than the next deepest level and when the quantity of water is small enough. For the CMI model the optimality of beamforming is found to depend on transmit power  $P$  and the *quality of feedback* associated with the mean information, which is defined mathematically as the ratio  $\|\bar{\mathbf{H}}\|^2/\alpha$  of the norm squared of the channel mean vector  $\bar{\mathbf{H}}$  and the channel uncertainty  $\alpha$ . As the transmit power  $P$  is decreased or the quality of feedback improves beamforming becomes optimal. As mentioned earlier, for perfect CSIT (uncertainty  $\alpha \rightarrow 0$  so quality of feedback  $\rightarrow \infty$ ) the optimal input strategy is beamforming, while in the absence of mean feedback (quality of feedback  $\rightarrow 0$  so the CMI model becomes the ZMSW model), as shown by Telatar [69], the optimal input covariance has full rank, i.e., beamforming is necessarily suboptimal. Note that [40], [57], [58], and [76] assume a single receive antenna. Next, we summarize the analogous capacity results for MIMO channels.

**MIMO Channels:** With multiple transmit and receive antennas, capacity with CSIR and CDIT under the CCI model with spatially white fading at the receiver ( $\mathbf{R}^r = \mathbf{I}$ ) is obtained by

Jafar and Goldsmith in [42]. Like the single receive antenna case the capacity achieving input covariance matrix is found to have the eigenvectors of the transmit fade covariance matrix and the eigenvalues are in the same order as the corresponding eigenvalues of the transmit fade covariance matrix. Jafar and Goldsmith also presented in closed form a mathematical condition that is both necessary and sufficient for optimality of beamforming in this case. The same necessary and sufficient condition is also derived independently by Jorsweick and Boche in [45] and Simon and Moustakas in [66]. In [46], Jorsweick and Boche extend these results to incorporate fade correlations at the receiver as well. Their results show that while the receive fade correlation matrix does not affect the eigenvectors of the optimal input covariance matrix, it does affect the eigenvalues. The general condition for optimality of beamforming found by Jorsweick and Boche depends upon the two largest eigenvalues of the transmit covariance matrix and all the eigenvalues of the receive covariance matrix.

Capacity under the CMI model with multiple transmit and receive antennas is solved by Jafar and Goldsmith in [38] when the channel mean has rank one and is extended to general channel means by Moustakas and Simon in [67]. Similar to the MISO case, the principal eigenvector of the optimal input covariance matrix and of the channel mean are the same and the eigenvalues of the remaining eigenvectors are equal. For the case where the channel mean has unit rank, a necessary and sufficient condition for optimality of beamforming is also determined in [38].

These results summarize our discussion of channel capacity with CDIT and perfect CSIR under different channel distribution models. From these results we notice that the benefits of adapting to distribution information regarding CMI or CCI fed back from the receiver to the transmitter are twofold. Not only does the capacity increase with more information about the channel distribution, but this feedback also allows the transmitter to identify the stronger channel modes and achieve this higher capacity with simple scalar codewords.

We conclude this section with a discussion on the growth of capacity with number of antennas. With perfect CSIR and CDIT under the ZMSW channel distribution, it was shown by Foschini and Gans [22] and by Telatar [69] that the channel capacity grows linearly with  $\min(M, N)$ . This linear increase occurs whether the transmitter knows the channel perfectly (perfect CSIT) or only knows its distribution (CDIT). The proportionality constant of this linear increase, called the rate of growth, has also been characterized in [15], [31], [68], [69]. Chuah *et al.* [15] show that with perfect CSIR and CSIT, the rate of growth of capacity with  $\min(M, N)$  is reduced by channel fading correlations at high SNR but is increased at low SNR. They also show that the mutual information under CSIR increases linearly with  $\min(M, N)$  even when a spatially white transmission strategy is used on a correlated fading channel, although the slope is reduced relative to the uncorrelated fading channel. As we will see in the next section, the assumption of perfect CSIR is crucial for the linear growth behavior of capacity with the number of antennas.

In the next section, we explore the capacity when only CDI is available at the transmitter and the receiver.

4) *Capacity With CDIT and CDIR: ZMSW Model:* We saw in the last section that with perfect CSIR, channel capacity grows linearly with the minimum of the number of transmit and receive antennas. However, reliable channel estimation may not be possible for a mobile receiver that experiences rapid fluctuations of the channel coefficients. Since user mobility is the principal driving force for wireless communication systems, the capacity behavior with CDIT and CDIR under the ZMSW distribution model (i.e.,  $\mathbf{H}$  is distributed as  $\mathbf{H}^w$  with no knowledge of  $\mathbf{H}$  at either the receiver or transmitter) is of particular interest. In this section, we summarize some MIMO capacity results in this area.

One of the first papers to address the MIMO capacity with CDIR and CDIT under the ZMSW model is [53] by Marzetta and Hochwald. They model the channel matrix components as i.i.d. complex Gaussian random variables that remain constant for a coherence interval of  $T$  symbol periods after which they change to another independent realization. Capacity is achieved when the  $T \times M$  transmitted signal matrix is equal to the product of two statistically independent matrices: a  $T \times T$  isotropically distributed unitary matrix times a certain  $T \times M$  random matrix that is diagonal, real, and nonnegative. This result enables them to determine capacity for many interesting cases. Marzetta and Hochwald show that, for a fixed number of antennas, as the length of the coherence interval increases, the capacity approaches the capacity obtained as if the receiver knew the propagation coefficients. However, perhaps the most surprising result in [53] is the following: In contrast to the linear growth of capacity with  $\min(M, N)$  under the perfect CSIR assumption, [53] showed that in the absence of CSIT and CSIR, capacity does not increase at all as the number of transmit antennas is increased beyond the length of the coherence interval  $T$ . The MIMO capacity for this model was further explored by Zheng and Tse in [89]. They show that at high SNRs capacity is achieved using no more than  $M^* = \min(M, N, \lfloor T/2 \rfloor)$  transmit antennas. In particular, having more transmit antennas than receive antennas does not provide any capacity increase at high SNR. Zheng and Tse also find that for each 3-dB SNR increase, the capacity gain is  $M^*(1 - M^*/T)$ .

Notice that [53], [89] assume block fading models, i.e., the channel fade coefficients are assumed to be constant for a block of  $T$  symbol durations. Hochwald and Marzetta extend their results to continuous fading in [54] where, within each independent  $T$ -symbol block, the fading coefficients have an arbitrary time correlation. If the correlation vanishes beyond some lag  $\tau$ , called the *correlation time* of the fading, then it is shown in [54] that increasing the number of transmit antennas beyond  $\min(\tau, T)$  antennas does not increase capacity. Lapidath and Moser [47] explored the channel capacity of this CDIT/CDIR model for the ZMSW distribution at high SNR without the block fading assumption. In contrast to the results of Zheng and Tse for block fading, Lapidath and Moser show that without the block fading assumption, the channel capacity grows only double logarithmically in SNR. This result is shown to hold under very general conditions, even allowing for memory and partial receiver side information.

5) *Capacity With CDIT and CDIR: CCI Model:* The results in [53] and [89] seem to leave little hope of achieving the high



capacity gains predicted for MIMO systems when the channel cannot be accurately estimated at the receiver and the channel distribution follows the ZMSW model. However, before resigning ourselves to these less-than-optimistic results we note that these results assume a somewhat pessimistic model for the channel distribution. That is because most channels when averaged over a relatively small area have either a nonzero mean or a nonwhite covariance. Thus, if these distribution parameters can be tracked, the channel distribution corresponds to either the CMI or CCI model.

Recent work by Jafar and Goldsmith [37] addresses the MIMO channel capacity with CDIT and CDIR under the CCI distribution model. The channel matrix components are modeled as spatially correlated complex Gaussian random variables that remain constant for a coherence interval of  $T$  symbol periods after which they change to another independent realization based on the spatial correlation model. The channel correlations are assumed to be known at the transmitter and receiver. As in the case of spatially white fading (ZMSW model), Jafar and Goldsmith show that with the CCI model the capacity is achieved when the  $T \times M$  transmitted signal matrix is equal to the product of a  $T \times T$  isotropically distributed unitary matrix, a statistically independent  $T \times M$  random matrix that is diagonal, real and nonnegative and the matrix of the eigenvectors of the transmit fade covariance matrix  $\mathbf{R}^t$ . It is shown in [37] that the channel capacity is independent of the smallest  $(M - T)^+$  eigenvalues of the transmit fade covariance matrix, as well as the eigenvectors of the transmit and receive fade covariance matrices  $\mathbf{R}^t$  and  $\mathbf{R}^r$ . Also, in contrast to the results for the spatially white fading model where adding more transmit antennas beyond the coherence interval length ( $M > T$ ) does not increase capacity, [37] shows that additional transmit antennas always increase capacity as long as their channel fading coefficients are spatially correlated. Thus, in contrast to the results in favor of independent fades with perfect CSIR, these results indicate that with CCI at the transmitter and the receiver, transmit fade correlations can be beneficial, making the case for minimizing the spacing between transmit antennas when dealing with highly mobile, fast fading channels that cannot be accurately measured. Mathematically, [37] proves that for fast fading channels ( $T = 1$ ), capacity is a Schur-concave function of the vector of eigenvalues of the transmit fade correlation matrix. The maximum possible capacity gain due to transmitter fade correlations is shown to be  $10 \log_{10} M$  db.

6) *Frequency Selective Fading Channels:* While flat fading is a realistic assumption for narrowband systems where the signal bandwidth is smaller than the channel coherence bandwidth, broadband communications involve channels that experience frequency selective fading. Research on the capacity of MIMO systems with frequency selective fading typically takes the approach of dividing the channel bandwidth into parallel flat fading channels and constructing an overall block diagonal channel matrix with the diagonal blocks given by the channel matrices corresponding to each of these subchannels. Under perfect CSIR and CSIT, the total power constraint then leads to the usual closed-form waterfilling solution. Note that the waterfill is done simultaneously over both space and frequency. Even SISO frequency selective fading channels can

be represented by the MIMO system model (1) in this manner [59]. For MIMO systems, the matrix channel model is derived by Bolcskei, Gesbert and Paulraj in [5] based on an analysis of the capacity behavior of OFDM-based MIMO channels in broadband fading environments. Under the assumption of perfect CSIR and CDIT for the ZMSW model, their results show that in the MIMO case, unlike the SISO case, frequency selective fading channels may provide advantages over flat fading channels not only in terms of ergodic capacity but also in terms of capacity versus outage. In other words, MIMO frequency selective fading channels are shown to provide both higher diversity gain and higher multiplexing gain than MIMO flat-fading channels. The measurements in [55] show that frequency selectivity makes the CDF of the capacity steeper and, thus, increases the capacity for a given outage as compared with the flat-frequency case, but the influence on the ergodic capacity is small.

7) *Training for Multiple-Antenna Systems:* The results summarized in the previous sections indicate that CSI plays a crucial role in the capacity of MIMO systems. In particular, the capacity results in the absence of CSIR are strikingly different and often quite pessimistic compared with those that assume perfect CSIR. To recapitulate, with perfect CSIR and CDIT MIMO channel capacity is known to increase linearly with  $\min(M, N)$  when the CDIT assumes the ZMSW or CCI distribution models. However, in fast fading when the channel changes so rapidly that it cannot be estimated reliably at the receiver (CDIR only) the capacity does not increase with the number of transmit antennas at all for  $M > T$  where  $T$  is the channel decorrelation time. Also at high SNR under the ZMSW distribution model, capacity with perfect CSIR and CDIT increases logarithmically with SNR, while the capacity with CDIR and CDIT increases only double logarithmically with SNR. Thus, CSIR is critical for obtaining the high capacity benefits of multiple-antenna wireless links. CSIR is often obtained by sending known training symbols to the receiver. However, with too little training the channel estimates are poor, whereas with too much training there is no time for data transmission before the channel changes. So the key question to ask is how much training is needed in multiple-antenna wireless links. This question itself is the title of the paper [29] by Hassibi and Hochwald where they compute a lower bound on the capacity of a channel that is learned by training and maximize the bound as a function of the receive SNR, fading coherence time, and number of transmitter antennas. When the training and data powers are allowed to vary, the optimal number of training symbols is shown to be equal to the number of transmit antennas—which is also the smallest training interval length that guarantees meaningful estimates of the channel matrix. When the training and data powers are instead required to be equal, the optimal training duration may be longer than the number of antennas. Hassibi and Hochwald also show that training-based schemes can be optimal at high SNR, but are suboptimal at low SNR.

#### D. Open Problems in Single-User MIMO

The results summarized in this section form the basis of our understanding of channel capacity under different CSI and CDI

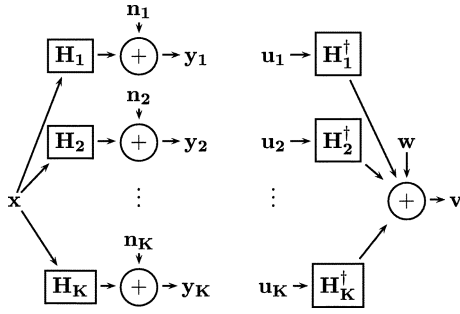


Fig. 6. System models of the (left) MIMO BC and the (right) MIMO MAC channels.

assumptions. These results serve as useful indicators for the benefits of incorporating training and feedback schemes in a MIMO wireless link to obtain CSIR/CDIT and CSIT/CDIT, respectively. However, our knowledge of MIMO capacity with CDI only is still far from complete, even for single-user systems. We conclude this section by pointing out some of the many open problems.

- 1) Combined CCI and CMI: Capacity under CDIT and perfect CSIR is unsolved under a combined CCI and CMI distribution model even with a single receive antenna.
- 2) CCI: With perfect CSIR and CDIT capacity is not known under the CCI model for completely general correlations.
- 3) CDIR: Almost all cases with only CDIR are open problems.
- 4) Outage capacity: Most results for CDI only at either the transmitter or receiver are for ergodic capacity. Capacity versus outage has proven to be less analytically tractable than ergodic capacity and contains an abundance of open problems.

### III. MULTIUSER MIMO

In this section, we consider the two basic multiuser MIMO channel models: the MIMO MAC and the MIMO BC. Since the capacity region of a general MAC has been known for quite a while, there are many results on the MIMO MAC for both constant channels and fading channels with different CSI and CDI assumptions at the transmitters and receivers. The MIMO BC, however, is a relatively new problem for which capacity results have only recently been found. As a result, the field is much less developed, but we summarize the recent results in the area. Interestingly, the MIMO MAC and MIMO BC have been shown to be duals, as we will discuss in Section III-C2.

#### A. System Model

To describe the MAC and BC models, we consider a cellular-type system in which the base station has  $M$  antennas and each of the  $K$  mobiles has  $N$  antennas. The downlink of this system is a MIMO BC and the uplink is a MIMO MAC. We will use  $\mathbf{H}_i$  to denote the *downlink* channel matrix from the base station to user  $i$ . Assuming that the same channel is used on the uplink and downlink, the *uplink* matrix of user  $i$  is  $\mathbf{H}_i^H$ . A picture of the system model is shown in Fig. 6.

In the MAC, let  $\mathbf{u}_k \in \mathbb{C}^{N \times 1}$  be the transmitted signal of user (i.e., mobile)  $k$ . Let  $\mathbf{v} \in \mathbb{C}^{M \times 1}$  denote the received signal and

$\mathbf{w} \in \mathbb{C}^{M \times 1}$  the noise vector where  $\mathbf{w} \sim \tilde{\mathcal{N}}(0, \mathbf{I})$  is circularly symmetric complex Gaussian with identity covariance. The received signal at the base station is then equal to

$$\begin{aligned} \mathbf{v} &= \mathbf{H}_1^H \mathbf{u}_1 + \cdots + \mathbf{H}_K^H \mathbf{u}_K + \mathbf{w} \\ &= \mathbf{H}^H \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_K \end{bmatrix} + \mathbf{w} \quad \text{where} \quad \mathbf{H}^H = [\mathbf{H}_1^H \quad \cdots \quad \mathbf{H}_K^H]. \end{aligned}$$

In the MAC, each user (i.e., mobile) is subject to an individual power constraint of  $P_k$ . The transmit covariance matrix of user  $k$  is defined to be  $\mathbf{Q}_k \triangleq \mathbb{E}[\mathbf{u}_k \mathbf{u}_k^H]$ . The power constraint implies  $\text{Tr}(\mathbf{Q}_k) \leq P_k$  for  $k = 1, \dots, K$ .

In the BC, let  $\mathbf{x} \in \mathbb{C}^{M \times 1}$  denote the transmitted vector signal (from the base station) and let  $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$  be the received signal at receiver (i.e., mobile)  $k$ . The noise at receiver  $k$  is represented by  $\mathbf{n}_k \in \mathbb{C}^{N \times 1}$  and is assumed to be circularly symmetric complex Gaussian noise ( $\mathbf{n}_k \sim \mathcal{N}(0, \mathbf{I})$ ). The received signal of User  $k$  is equal to

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k. \quad (11)$$

The transmit covariance matrix of the input signal is  $\Sigma_x \triangleq \mathbb{E}[\mathbf{x} \mathbf{x}^H]$ . The base station is subject to an average power constraint  $P$ , which implies  $\text{Tr}(\Sigma_x) \leq P$ .

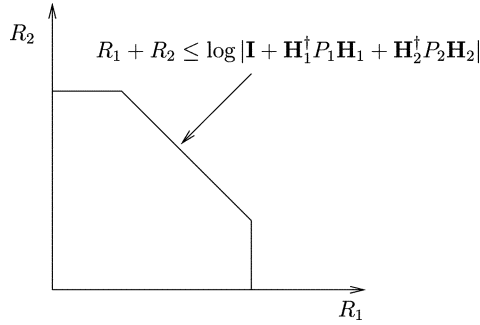
#### B. MIMO Multiple-Access Channel

In this section, we summarize capacity results on the multiple-antenna MAC. We first analyze the constant channel scenario and then consider the fading channel. Since the capacity region of a general MAC is known, the expressions for the capacity of a constant MAC are quite straightforward. For the fading case, one must consider different assumptions about the CSI and CDI available at the transmitter and receiver. We consider three cases: perfect CSIR and CSIT, perfect CSIR and CDIT, and CDIT and CDIR. As above, under CDI, we consider three different distribution models: the ZMSW, CMI, and CCI models.

1) *Constant Channel*: The capacity of any MAC can be written as the convex closure of the union of rate regions corresponding to every product input distribution  $p(u_1) \cdots p(u_K)$  satisfying the user-by-user power constraints [18]. For the Gaussian MIMO MAC, however, it has been shown that it is sufficient to consider only Gaussian inputs and that the convex hull operation is not needed [11], [86]. For any set of powers  $\mathbf{P} = (P_1, \dots, P_K)$ , the capacity of the MIMO MAC is shown in (12), at the bottom of the next page. The  $i$ th user transmits a zero-mean Gaussian with spatial covariance matrix  $\mathbf{Q}_i$ . Each set of covariance matrices  $(\mathbf{Q}_1, \dots, \mathbf{Q}_K)$  corresponds to a  $K$ -dimensional polyhedron (i.e.)

$$\left\{ (R_1, \dots, R_K) : \sum_{i \in S} R_i \leq \frac{1}{2} \log \left| \mathbf{I} + \sum_{i \in S} \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i \right| \quad \forall S \subseteq \{1, \dots, K\} \right\}$$

and the capacity region is equal to the union (over all covariance matrices satisfying the trace constraints) of

Fig. 7. Capacity region of MIMO MAC for  $N = 1$ .

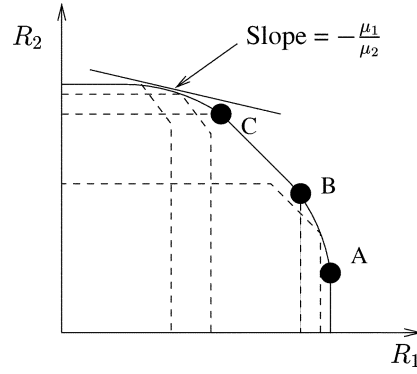
all such polyhedrons. The corner points of each polyhedron can be achieved by *successive decoding*, in which users' signals are successively decoded and subtracted out of the received signal. For the two-user case, each set of covariance matrices corresponds to a pentagon, similar in form to the capacity region of the scalar Gaussian MAC. The corner point where  $R_1 = \log |\mathbf{I} + \mathbf{H}_1^\dagger \mathbf{Q}_1 \mathbf{H}_1|$  and  $R_2 = \log |\mathbf{I} + \mathbf{H}_1^\dagger \mathbf{Q}_1 \mathbf{H}_1 + \mathbf{H}_2^\dagger \mathbf{Q}_2 \mathbf{H}_2| - R_1 = \log |\mathbf{I} + (\mathbf{I} + \mathbf{H}_1^\dagger \mathbf{Q}_1 \mathbf{H}_1)^{-1} \mathbf{H}_2^\dagger \mathbf{Q}_2 \mathbf{H}_2|$  corresponds to decoding user 2 first (i.e., in the presence of interference from user 1) and decoding user 1 last (without interference from user 2). Successive decoding can reduce a complex multiuser detection problem into a series of single-user detection steps [27].

The capacity region of a MIMO MAC for the single transmit antenna case ( $N = 1$ ) is shown in Fig. 7. When  $N = 1$ , the covariance matrix of each transmitter is a scalar equal to the transmitted power. Clearly, each user should transmit at full power. Thus, the capacity region for a  $K$ -user MAC for  $N = 1$  is the set of all rate vectors  $(R_1, \dots, R_K)$  satisfying

$$\sum_{i \in S} R_i \leq \frac{1}{2} \log \left| \mathbf{I} + \sum_{i \in S} \mathbf{H}_i^\dagger P_i \mathbf{H}_i \right| \quad \forall S \subseteq \{1, \dots, K\}. \quad (13)$$

For the two-user case, this reduces to the simple pentagon seen in Fig. 7.

When  $N > 1$ , however, a union must be taken over all covariance matrices. Intuitively, the set of covariance matrices that maximize  $R_1$  are different from the set of covariance matrices that maximize the sum rate. In Fig. 8, a MAC capacity region for  $N > 1$  is shown. Notice that the region is equal to the union of pentagons (each pentagon corresponding to a different set of transmit covariance matrices), a few of which are shown with dashed lines in the figure. The boundary of the capacity region is in general curved, except at the sum rate point, where the boundary is a straight line [86]. Each point on the curved portion of the boundary is achieved by a *different* set of covariance matrices. At point A, user 1 is decoded last and achieves his single-user capacity by choosing  $\mathbf{Q}_1$  as a water-fill of the channel  $\mathbf{H}_1$  (independent of  $\mathbf{H}_2$  or  $\mathbf{Q}_2$ ). User 2 is decoded first, in the presence of interference from user 1, so  $\mathbf{Q}_2$  is chosen as

Fig. 8. Capacity region of MIMO MAC for  $N > 1$ .

a waterfill of the channel  $\mathbf{H}_2$  and the interference from user 1. The sum-rate corner points B and C are the two corner points of the pentagon corresponding to the sum-rate optimal covariance matrices  $\mathbf{Q}_1^{\text{sum}}$  and  $\mathbf{Q}_2^{\text{sum}}$ . At point B user 1 is decoded last, whereas at point C user 2 is decoded last. Thus, points B and C are achieved using the same covariance matrices but different decoding orders.

Next, we focus on characterizing the optimal covariance matrices  $(\mathbf{Q}_1, \dots, \mathbf{Q}_K)$  that achieve different points on the boundary of the MIMO MAC capacity region. Since the MAC capacity region is convex, it is well known from convex theory that the boundary of the capacity region can be fully characterized by maximizing the function  $\mu_1 R_1 + \dots + \mu_K R_K$  over all rate vectors in the capacity region and for all nonnegative priorities  $(\mu_1, \dots, \mu_K)$  such that  $\sum_{i=1}^K \mu_i = 1$ . For a fixed set of priorities  $(\mu_1, \dots, \mu_K)$ , this is equivalent to finding the point on the capacity region boundary that is tangent to a line whose slope is defined by the priorities. See the tangent line in Fig. 8 for an example. The structure of the MAC capacity region implies that all boundary points of the capacity region are corner points of polyhedrons corresponding to different sets of covariance matrices. Furthermore, the corner point should correspond to successive decoding in order of *increasing* priority, i.e., the user with the highest priority should be decoded last and, therefore, sees no interference [70], [73]. Thus, the problem of finding the boundary point on the capacity region associated with priorities  $\mu_1, \dots, \mu_K$  assumed to be in descending order (users can be arbitrarily re-numbered to satisfy this condition) can be written as

$$\begin{aligned} \max_{\mathbf{Q}_1, \dots, \mathbf{Q}_K} \quad & \mu_K \log \left| \mathbf{I} + \sum_{l=1}^K \mathbf{H}_l^\dagger \mathbf{Q}_l \mathbf{H}_l \right| \\ & + \sum_{i=1}^{K-1} (\mu_i - \mu_{i+1}) \log \left| \mathbf{I} + \sum_{l=1}^i \mathbf{H}_l^\dagger \mathbf{Q}_l \mathbf{H}_l \right| \end{aligned}$$

subject to power constraints on the trace of each of the covariance matrices. Note that the covariances that maximize the func-

$$\mathcal{C}_{MAC}(\mathbf{P}; \mathbf{H}^\dagger) \triangleq \bigcup_{\{\mathbf{Q}_i \geq 0, \text{Tr}(\mathbf{Q}_i) \leq P_i \quad \forall i\}} \left\{ (R_1, \dots, R_K) : \sum_{i \in S} R_i \leq \frac{1}{2} \log \left| \mathbf{I} + \sum_{i \in S} \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i \right| \quad \forall S \subseteq \{1, \dots, K\} \right\} \quad (12)$$

tion above are the *optimal* covariances. The most interesting and useful feature of the optimization problem above is that the objective function is concave in the covariance matrices. Thus, efficient convex optimization tools exist that solve this problem numerically [7]. A more efficient numerical technique to find the sum-rate maximizing (i.e.,  $\mu_1 = \dots = \mu_K$ ) covariance matrices, called iterative waterfilling, was developed by Yu *et al.* [86]. This technique is based on the Karush Kuhn Tucker (KKT) optimality conditions for the sum-rate maximizing covariance matrices. These conditions indicate that the sum-rate maximizing covariance matrix of any user in the system should be the single-user water-filling covariance matrix of its own channel with noise equal to the actual noise plus the interference from the other  $K - 1$  transmitters.

2) *Fading Channels*: As in the single-user case, the capacity of the MIMO MAC where the channel is time-varying depends on the definition of capacity and the availability of CSI and CDI at the transmitters and the receiver. The capacity with perfect CSIT and CSIR is very well studied, as is the capacity with perfect CSIR and CDIT under the ZMSW distribution model. However, little is known about the capacity of the MIMO MAC with CDIT at either the transmitter or receiver under the CMI or CCI distribution models. Some results on the optimum distribution for the single antenna case with CDIT and CDIR under the ZMSW distribution can be found in [62].

With perfect CSIR and CSIT the system can be viewed as a set of *parallel* non interfering MIMO MACs (one for each fading state) sharing a common power constraint. Thus, the ergodic capacity region can be obtained as an average of these parallel MIMO MAC capacity regions [87], where the averaging is done with respect to the channel statistics. The iterative waterfilling algorithm of [86] easily extends to this case, with joint space and time waterfilling.

The capacity region of a MAC with perfect CSIR and CDIT under the ZMSW distribution model was found in [23] and [63]. In this case, Gaussian inputs are optimal and the ergodic capacity region is equal to the time average of the capacity obtained at each fading instant with a constant transmit policy (i.e., a constant covariance matrix for each user). Thus, the ergodic capacity region is given by

$$\bigcup_{\{\mathbf{Q}_i \geq 0, \text{Tr}(\mathbf{Q}_i) \leq P_i \quad \forall i\}} \left\{ (R_1, \dots, R_K) : \sum_{i \in S} R_i \leq \frac{1}{2} \mathbb{E}_{\mathbf{H}} \left[ \log \left| \mathbf{I} + \sum_{i \in S} \mathbf{H}_i^\dagger \mathbf{Q}_i \mathbf{H}_i \right| \right] \right. \\ \left. \forall S \subseteq \{1, \dots, K\} \right\}.$$

If the channel matrices  $\mathbf{H}_i$  have i.i.d. complex Gaussian entries and each user has the same power constraint, then the optimal covariances are scaled versions of the identity matrix [69].

There has also been some work on capacity with perfect CSIR and CDIT under the CCI distribution model [41]. In this paper, Jafar and Goldsmith determine the optimal transmit covariance matrices when there is transmit antenna correlation that is known at the transmitters. This topic has yet to be fully investigated.

Asymptotic results on the sum capacity of MIMO MAC channels with the number of receive antennas and the number of transmitters increasing to infinity were obtained by Telatar [69] and by Viswanath *et al.* [80]. MIMO MAC sum capacity with perfect CSIR and CDIT under the ZMSW distribution model (i.e., each transmitter's channel is distributed as  $\mathbf{H}^{\mathbf{W}}$ ) is found to grow *linearly* with  $\min(M, NK)$  [69]. Thus, for systems with large numbers of users, increasing the number of receive antennas at the base station ( $M$ ) while keeping the number of mobile antennas ( $N$ ) constant can lead to linear growth. Sum capacity with perfect CSIR and CSIT also scales linearly with  $\min(M, NK)$ , but perfect CSIT is of decreasing value as the number of receive antennas increases [32], [80]. Furthermore, the limiting distribution of the sum capacity with perfect CSIR and CSIT was found to be Gaussian by Hochwald and Vishwanath [32].

### C. MIMO Broadcast Channel

In this section, we summarize capacity results on the multiple-antenna BC. When the transmitter has only one antenna, the Gaussian broadcast channel is a degraded broadcast channel (i.e., the users can be absolutely ranked by their channel strength), for which the capacity region is known [18]. However, when the transmitter has more than one antenna, the Gaussian broadcast channel is generally nondegraded.<sup>3</sup> The capacity region of general nondegraded broadcast channels is unknown, but the seminal work of Caire and Shamai [9] and subsequent research on this problem have shed a great deal of light on this channel and the sum capacity of the MIMO BC has been found. In subsequent sections, we focus mainly on the constant channel, but we do briefly discuss the fading channel as well which is still an open problem. Note that the BC transmitter (i.e., the base station) has  $M$  antennas and each receiver has  $N$  antennas, as described in Section III-A.

1) *Dirty Paper Coding (DPC) Achievable Rate Region*: An achievable region for the MIMO BC was first obtained for the  $N = 1$  case by Caire and Shamai [9] and later extended to the multiple-receive antenna case by Yu and Cioffi [83] using the idea of DPC [17]. The basic premise of DPC is as follows. If the transmitter (but not the receiver) has perfect, noncausal knowledge of additive Gaussian interference in the channel, then the capacity of the channel is the same as if there was no additive interference, or equivalently as if the receiver also had knowledge of the interference. DPC is a technique that allows noncausally known interference to be "presubtracted" at the transmitter, but in such a way that the transmit power is not increased. A more practical (and more general) technique to perform this presubtraction is the cancelling for known interference technique found by Erez *et al.* in [19].

In the MIMO BC, DPC can be applied at the transmitter when choosing codewords for different receivers. The transmitter first picks a codeword (i.e.,  $\mathbf{x}_1$ ) for receiver 1. The transmitter then chooses a codeword for receiver 2 (i.e.,  $\mathbf{x}_2$ ) with full (noncausal) knowledge of the codeword intended for receiver 1. Therefore,

<sup>3</sup>The multiple-antenna broadcast channel is nondegraded because users receive different strength signals from different transmit antennas. See [18] for a precise definition of degradedness.

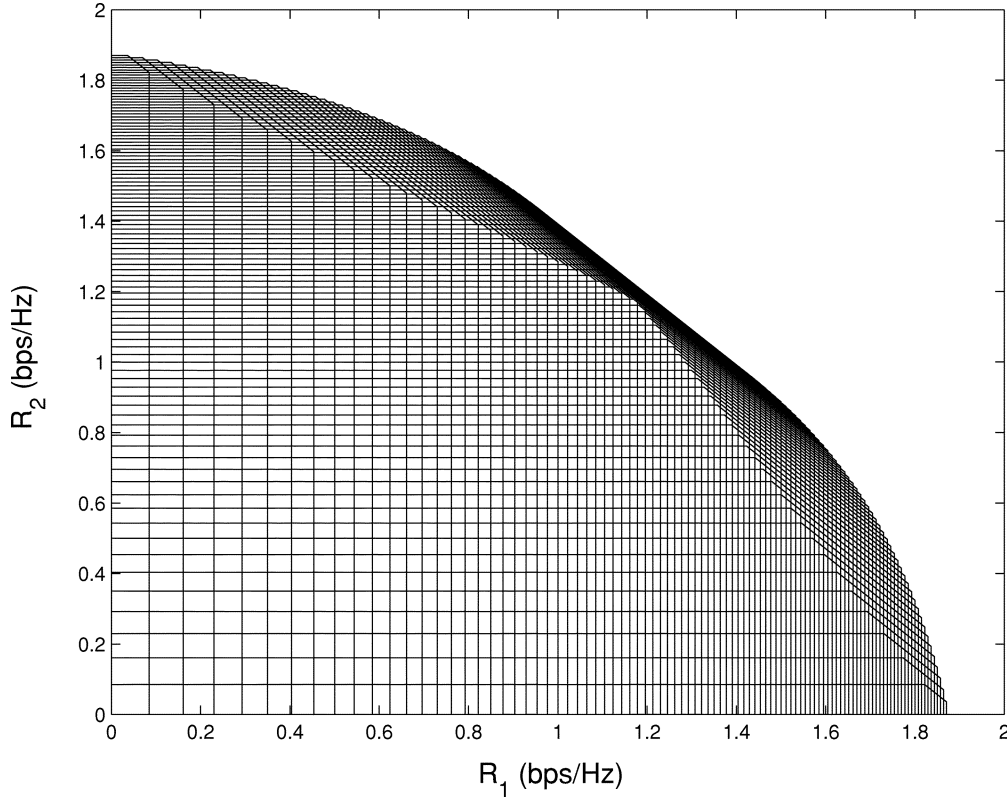


Fig. 9. Dirty paper rate region,  $\mathbf{H}_1 = [1 \ 0.5]$ ,  $\mathbf{H}_2 = [0.5 \ 1]$ ,  $P = 10$ .

the codeword of user 1 can be presubtracted such that receiver 2 does not see the codeword intended for receiver 1 as interference. Similarly, the codeword for receiver 3 is chosen such that receiver 3 does not see the signals intended for receivers 1 and 2 (i.e.  $\mathbf{x}_1 + \mathbf{x}_2$ ) as interference. This process continues for all  $K$  receivers. If user  $\pi(1)$  is encoded first, followed by user  $\pi(2)$ , etc., the following is an achievable rate vector:

$$R_{\pi(i)} = \frac{1}{2} \log \frac{\left| \mathbf{I} + \mathbf{H}_{\pi(i)} \left( \sum_{j \geq i} \boldsymbol{\Sigma}_{\pi(j)} \right) \mathbf{H}_{\pi(i)}^\dagger \right|}{\left| \mathbf{I} + \mathbf{H}_{\pi(i)} \left( \sum_{j > i} \boldsymbol{\Sigma}_{\pi(j)} \right) \mathbf{H}_{\pi(i)}^\dagger \right|}} \quad i=1, \dots, K. \quad (14)$$

The dirty paper region  $\mathcal{C}_{\text{DPC}}(P, \mathbf{H})$  is defined as the convex hull of the union of all such rates vectors over all positive semi-definite covariance matrices  $\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K$  such that  $\text{Tr}(\boldsymbol{\Sigma}_1 + \dots + \boldsymbol{\Sigma}_K) = \text{Tr}(\boldsymbol{\Sigma}_x) \leq P$  and over all permutations  $(\pi(1), \dots, \pi(K))$ :

$$\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \triangleq \text{Co} \left( \bigcup_{\pi, \boldsymbol{\Sigma}_i} \mathbf{R}(\pi, \boldsymbol{\Sigma}_i) \right) \quad (15)$$

where  $\mathbf{R}(\pi, \boldsymbol{\Sigma}_i)$  is given by (14). The transmitted signal is  $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_K$  and the input covariance matrices are of the form  $\boldsymbol{\Sigma}_i = \mathbb{E}[\mathbf{x}_i \mathbf{x}_i^\dagger]$ . From the dirty paper result we find that  $\mathbf{x}_1, \dots, \mathbf{x}_K$  are uncorrelated, which implies  $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_1 + \dots + \boldsymbol{\Sigma}_K$ .

One important feature to notice about the dirty paper rate equations in (14) is that the rate equations are neither a concave nor convex function of the covariance matrices. This makes numerically finding the dirty paper region very difficult, because generally a brute force search over the entire space of covariance matrices that meet the power constraint must be conducted. The dirty paper rate region for a two-user channel with  $M = 2$  and  $N = 1$  is shown in Fig. 9.

Note that DPC and successive decoding (i.e., interference cancellation by the receiver instead of the transmitter) are completely equivalent capacity-wise for scalar channels, but this equivalence does not hold for MIMO channels. It has been shown [36] that the achievable region with successive decoding is contained within the DPC region.

2) *MAC-BC Duality*: In [74], Vishwanath, Jindal, and Goldsmith showed that the dirty paper rate region of the multiantenna BC with power constraint  $P$  is equal to the union of capacity regions of the dual MAC, where the union is taken over all individual power constraints that sum to  $P$

$$\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) = \bigcup_{\mathbf{P}: \sum_{i=1}^K P_i = P} \mathcal{C}_{\text{MAC}}(P_1, \dots, P_K, \mathbf{H}^\dagger). \quad (16)$$

This is the multiple-antenna extension of the previously established duality between the scalar Gaussian broadcast and multiple-access channels [44]. In addition to the relationship between the two rate regions, for any set of covariance matrices in the MAC/BC (and the corresponding rate vector), [74] provides an explicit set of transformations to find covariance matrices in

the BC/MAC that achieve the same rates. The union of MAC capacity regions in (16) is easily seen to be the same expression as in (12) but with the constraint  $\sum_{i=1}^K \text{Tr}(\mathbf{Q}_i) \leq P$  instead of  $\text{Tr}(\mathbf{Q}_i) \leq P_i \ \forall i$  (i.e., a sum constraint instead of individual constraints).

The MAC-BC duality is very useful from a numerical standpoint because the dirty paper region leads to nonconcave rate functions of the covariances, whereas the rates in the dual MAC are *concave* functions of the covariance matrices. Thus, the optimal MAC covariances can be found using standard convex optimization techniques and then transformed to the corresponding optimal BC covariances using the MAC-BC transformations given in [74]. A specialized algorithm to find the optimal MAC covariances can be found in [35]. An algorithm based on the iterative waterfilling algorithm [86] that finds the sum rate optimal covariances is given in [43].

The dirty paper rate region is shown in Fig. 9 for a channel with two-users,  $M = 2$  and  $N = 1$ . Notice that the dirty paper rate region shown in Fig. 9 is actually a union of MAC regions, where each MAC region corresponds to a different set of individual power constraints. Since  $N = 1$ , each of the MAC regions is a pentagon, as discussed in Section III-B1. Similar to the MAC capacity region, the boundary of the DPC region is curved, except at the sum-rate maximizing portion of the boundary. For the  $N = 1$  case, duality also indicates that rank-one covariance matrices (i.e., beamforming) are optimal for DPC. This fact is not obvious from the dirty paper rate equations, but follows from the transformations of [74] which find BC covariances that achieve the same rates as a set of MAC covariance matrices (which are scalars in the  $N = 1$  case).

Duality also allows the MIMO MAC capacity region to be expressed as an intersection of the dual dirty paper BC rate regions [74, Corollary 1]

$$\begin{aligned} C_{\text{MAC}}(P_1, \dots, P_K, \mathbf{H}^\dagger) \\ = \bigcap_{\alpha_i > 0} C_{\text{DPC}}\left(\sum_{i=1}^K \frac{P_i}{\alpha_i}; [\sqrt{\alpha_1} \mathbf{H}_1^T \cdots \sqrt{\alpha_K} \mathbf{H}_K^T]^T\right). \end{aligned} \quad (17)$$

3) *Optimality of DPC*: DPC was first shown to achieve the *sum-rate* capacity of the MIMO BC for the two-user,  $M = 2$ ,  $N = 1$  channel by Caire and Shamai [9]. This was shown by proving that the Sato upper bound [61] on the broadcast channel sum-rate capacity is achievable using DPC. The sum-rate optimality of DPC was extended to the multiuser channel with  $N = 1$  by Viswanath and Tse [79] and to the more general  $N > 1$  case by Vishwanath *et al.* [74] and Yu and Cioffi [84].

It has also recently been conjectured that the DPC rate region is the actual capacity region of the multiple-antenna broadcast channel. Significant progress toward proving this conjecture is made in [75] and [77].

4) *Fading Channels*: Most of the capacity problems for fading MIMO BCs are still open, with the exception of sum-rate capacity with perfect CSIR and CSIT. In this case, as for the MIMO MAC, the MIMO BC can be split into parallel channels with an overall power constraint (see Li and Goldsmith [48] for a treatment of the scalar case).

Asymptotic results for the sum-rate capacity of the MIMO BC for  $N = 1$  under the ZSMW model can be obtained by combining the asymptotic results for the sum-rate capacity of the MIMO MAC with duality [32]. Thus, the role of transmitter side information reduces with the growth in the number of transmit antennas and, hence, the sum capacity of the MIMO BC with  $K$  users and  $M$  transmit antennas tends to the sum capacity of a single-user system with only receiver CSI and  $M$  receive antennas and  $K$  transmit antennas, which is given by  $\mathbb{E}_{\mathbf{H}} \log |\mathbf{I} + P/K\mathbf{H}\mathbf{H}^\dagger|$ . Thus, the asymptotic growth under CSIR and CSIT or CDIT under the ZMSW model is linear as  $\mathcal{C} \min(M, K)$  and the growth rate constant  $\mathcal{C}$  can be found in [32]. As seen for the MIMO MAC, for systems with large numbers of users, increasing the number of transmit antennas at the base station ( $M$ ) while keeping the number of mobile antennas ( $N$ ) constant can lead to linear growth.

#### D. Open Problems in Multiuser MIMO

Multiuser MIMO has been the primary focus of research in recent years, mainly due to the large number of open problems in this area. Some of these are as follows.

- 1) BC with perfect CSIR and CDIT: The broadcast channel capacity is only known when both the transmitter and the receivers have perfect knowledge of the channel.
- 2) CDIT and CDIR: Since perfect CSI is rarely possible, a study of capacity with CDI at both the transmitter(s) and receiver(s) for both MAC and BC is of great practical relevance.
- 3) Non-DPC techniques for BC: DPC is a very powerful capacity-achieving scheme, but it appears quite difficult to implement in practice. Thus, non-DPC multiuser transmissions schemes for the downlink (such as downlink beamforming [60]) are also of practical relevance.

#### IV. MULTICELL MIMO

The MAC and the BC are information theoretic abstractions of the uplink and the downlink of a single cell in a cellular system. However, a cellular system, by definition, consists of many cells. Due to the fundamental nature of wireless propagation, transmissions in a cell are not limited to within that cell. Users and base stations in adjacent cells experience interference from each other. Also, since the base stations are typically not mobile themselves there is the possibility for the base stations to communicate through a high-speed reliable connection, possibly consisting of optical fiber links capable of very high data rates. This opens up the opportunity for base stations to cooperate in the way they process different users' signals. Analysis of the capacity of the cellular network, explicitly taking into account the presence of multiple cells, multiple users and multiple antennas, and the possibilities of cooperation between base stations is inevitably a hard problem and runs into several long-standing unsolved problems in network information theory. However, such an analysis is also of utmost importance because it defines a common benchmark that can be used to gauge the efficiency of any practical scheme, in the same way that the capacity of a single-user link serves as a measure of the performance of practical schemes. There has been some recent

research in this area that extends the single-cell MAC and BC results to multiple cells. In this section, we summarize some of these results.

The key to the extension of single-cell results to multiple-cell systems is the assumption of perfect cooperation between base stations. Conceptually, this allows the multiple base stations to be treated as physically distributed antennas of one composite base station. Specifically, consider a group of  $B$  coordinated cells, each with  $M$  antennas and  $K$  mobiles, each with  $N$  antennas. If we define  $\mathbf{H}_{i,j} \in \mathbb{C}^{N \times M}$  to be the *downlink* channel of user  $i$  from base station  $j$ , then the *composite* downlink channel of user  $i$  is  $\mathbf{F}_i = [\mathbf{H}_{i,1} \cdots \mathbf{H}_{i,B}]$  and the composite uplink channel is  $\mathbf{F}_i^\dagger$ . The received signal of user  $i$  can then be written as  $\mathbf{Y}_i = \mathbf{F}_i \mathbf{W} + \mathbf{n}_i$ , where  $\mathbf{W}$  is the composite transmitted signal defined as  $\mathbf{W}^T = [\mathbf{W}_1^T \cdots \mathbf{W}_B^T]$ . Here, we let  $\mathbf{W}_j$  represent the transmit signal from base  $j$ .

First, let us consider the uplink. As pointed out by Jafar *et al.* [36] the single-cell MIMO MAC capacity region results apply to this system in a straightforward way. Thus, by assuming perfect data cooperation between the base stations, the multiple-cell uplink is easily seen to be equal to the MAC capacity region of the composite channel, defined as  $\mathcal{C}_{\text{MAC}}(\mathbf{F}_1^\dagger, \dots, \mathbf{F}_K^\dagger; P_1, \dots, P_K)$  in (12), where the power constraints of the  $i$ th mobile is  $P_i$ .

On the downlink, since the base stations can cooperate perfectly, DPC can be used over the entire transmitted signal (i.e., across base stations) in a straightforward manner. The application of DPC to a multiple-cell environment with cooperation between base stations is pioneered in recent work by Shamai and Zaidel [64]. For one antenna at each user and each base station, they show that a relatively simple application of DPC can enhance the capacity of the cellular downlink. While capacity computations are not the focus of [64], they do show that their scheme is asymptotically optimal at high SNRs.

The MIMO downlink capacity is explored by Jafar and Goldsmith in [39]. Note that the multicell downlink can be solved in a similar way as the uplink. But this requires perfect data and *power cooperation* between the base stations. If we let  $\mathbf{X}_{i,j}$  represent the transmit vector for User  $i$  from base station  $j$ , the composite transmit vector intended for User  $i$  is  $\mathbf{X}_i^T = [\mathbf{X}_{i,1}^T \cdots \mathbf{X}_{i,B}^T]$ . Thus, the composite covariance of user  $i$  is defined as  $\Sigma_i = \mathbb{E}[\mathbf{X}_i \mathbf{X}_i^T]$ . The covariance matrix of the entire transmitted signal is  $\Sigma_x = \sum_{i=1}^K \Sigma_i$ . Assuming perfect data cooperation between the base stations, DPC can be applied to the composite vectors intended for different users. Thus, the dirty paper region described in Section III-C1, (15), can be achieved in the multicell downlink.

While data cooperation is a justifiable assumption for capacity computations in the sense that it captures the possibility of base stations cooperating among themselves as described earlier in this section, in practice each base station has its own power constraint. The per-base power constraint can be expressed as  $\mathbb{E}[\mathbf{W}_j^\dagger \mathbf{X}_j] \leq P_j$ , where  $P_j$  is the power constraint at base  $j$ . Thus power cooperation, or pooling the transmit power for all the base stations to have one overall transmit power constraint, is not realistic. Note that on the uplink the base stations are only receiving signals and, therefore,

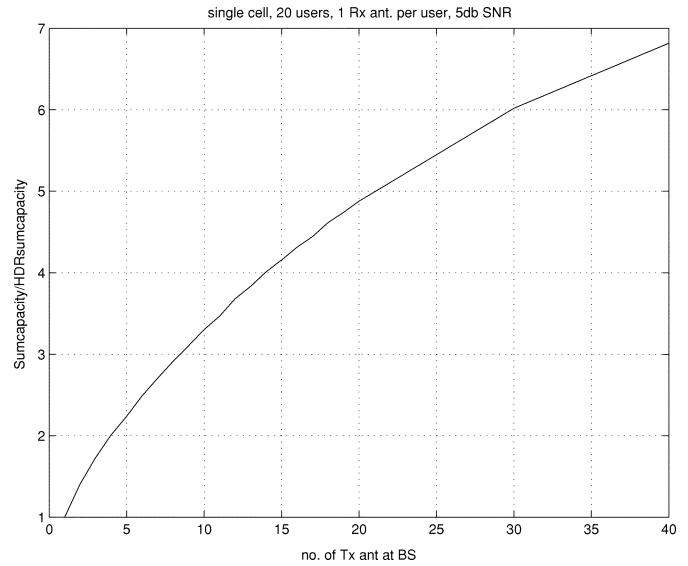


Fig. 10. Optimal sum rate relative to HDR.

no power cooperation is required. The per-base power constraints restrict consideration to covariance matrices such that  $\sum_{i=1}^K \text{Tr}(\mathbb{E}[\mathbf{X}_{i,j} \mathbf{X}_{i,j}^\dagger]) \leq P_j, j = 1, \dots, B$ . This is equivalent to a constraint of  $P_1$  on the sum of the first  $M$  diagonal entries of  $\Sigma_x$ , a constraint of  $P_2$  on the sum on the next  $M$  diagonal entries of  $\Sigma_x$ , etc. These constraints are considerably stricter than a constraint on the trace of  $\Sigma_x$  as in the single-cell case.

Though DPC yields an achievable region, it has not been shown to achieve the capacity region or even the sum-rate capacity with per-base power constraints. Additionally, the MAC-BC duality (Section III-C2) which greatly simplified calculation of the dirty paper region does not apply under per-base power constraints. Thus, even generating numerical results for the multicell downlink is quite challenging.

However, data and power cooperation does give a simple upper bound on the capacity of the network. Based on numerical comparisons between this upper bound and a lower bound on capacity derived in [39], Jafar and Goldsmith find that the simple upper bound with power and data cooperation is also a good measure of the capacity with data cooperation alone.

Note that current wireless systems use the high data rate (HDR) protocol and transmit to only one user at a time on the downlink, where this best user is chosen to maximize the average system data rate. In contrast, DPC allows the base station to transmit to many users simultaneously. This is particularly advantageous when the number of transmit antennas at the base station is much larger than the number of receive antennas at each user—a common scenario in current cellular systems. To illustrate the advantages of DPC over HDR, even for a single cell, the relative gains of optimal DPC over a strategy that serves only the best user at any time are shown in Fig. 10. Note that this single-cell model is equivalent to the multicell system with no cooperation between base stations so that the interference from other cells is treated as noise. With cooperation between base stations the gains are expected to be even more significant as DPC reduces the overall interference by making some users invisible to others.

The capacity results described in this section address just a few out of many interesting questions in the design of a cellular system with multiple antennas. Multiple antennas can be used not only to enhance the capacity of the system but also to drive down the probability of error through diversity combining. Recent work by Zheng and Tse [88] unravels a fundamental diversity versus multiplexing tradeoff in MIMO systems. Also, instead of using isotropic transmit antennas on the downlink and transmitting to many users, it may be simpler to use directional antennas to divide the cell into sectors and transmit to one user within each sector. The relative impact of CDIT and/or CDIR on each of these schemes is not fully understood. Although in this paper we focus on the physical layer, smart schemes to handle CDIT can also be found at higher layers. An interesting example is the idea of opportunistic beamforming [78]. In the absence of CSIT, the transmitter randomly chooses the beamforming weights. With enough users in the system, it becomes very likely that these weights will be nearly optimal for one of the users. In other words, a random beam selected by the transmitter is very likely to be pointed toward a user if there are enough users in the system. Instead of feeding back the channel coefficients to the transmitter the users simply feed back the SNRs they see with the current choice of beamforming weights. This significantly reduces the amount of feedback required. By randomly changing the weights frequently, the scheme also treats all users fairly.

## V. CONCLUSION

We have summarized recent results on the capacity of MIMO channels for both single-user and multiuser systems. The great capacity gains predicted for such systems can be realized in some cases, but realistic assumptions about channel knowledge and the underlying channel model can significantly mitigate these gains. For single-user systems the capacity under perfect CSI at the transmitter and receiver is relatively straightforward and predicts that capacity grows linearly with the number of antennas. Backing off from the perfect CSI assumption makes the capacity calculation much more difficult and the capacity gains are highly dependent on the nature of the CSI/CDI, the channel SNR, and the antenna element correlations. Specifically, assuming perfect CSIR, CSIT provides significant capacity gain at low SNRs but not much at high SNRs. The insight here is that at low SNRs it is important to put power into the appropriate eigenmodes of the system. Interestingly, with perfect CSIR and CSIT, antenna correlations are found to increase capacity at low SNRs and decrease capacity at high SNRs. Finally, under CDIT and CDIR for a zero-mean spatially white channel, at high SNRs capacity grows relative to only the double log of the SNR with the number of antennas as a constant additive term. This rather poor capacity gain would not typically justify adding more antennas. However, at moderate SNRs the growth relative to the number of antennas is less pessimistic.

We also examined the capacity of MIMO broadcast and multiple-access channels. The capacity region of the MIMO MAC is well-known and can be characterized as a convex optimization problem. Duality allows the DPC achievable region for the MIMO BC, a nonconvex region, to be computed from the

MIMO MAC capacity region. These capacity and achievable regions are only known for ergodic capacity under perfect CSIT and CSIR. Relatively little is known about the MIMO MAC and BC regions under more realistic CSI assumptions. A multicell system with base station cooperation can be modeled as a MIMO BC (downlink) or MIMO MAC (uplink), where the antennas associated with each base station are pooled by the system. Exploiting this antenna structure leads to significant capacity gains over HDR transmission strategies.

There are many open problems in this area. For single-user systems the problems are mainly associated with CDI only at either the transmitter or receiver. Most capacity regions associated with multiuser MIMO channels remain unsolved, especially ergodic capacity and capacity versus outage for the MIMO BC under perfect receiver CSI only. There are very few existing results for CDI at either the transmitter or receiver for any multiuser MIMO channel. Finally, the capacity of cellular systems with multiple antennas remains a relatively open area, in part because the single-cell problem is mostly unsolved and in part because the Shannon capacity of a cellular system is not well-defined and depends heavily on frequency assumptions and propagation models. Other fundamental tradeoffs in MIMO cellular designs such as whether antennas should be used for sectorization, capacity gain, or diversity are not well understood. In short, we have only scratched the surface in understanding the fundamental capacity limits of systems with multiple transmitter and receiver antennas, as well as the implications of these limits for practical system designs. This area of research is likely to remain timely, important, and fruitful for many years to come.

## REFERENCES

- [1] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 550–561, Apr. 2002.
- [2] N. Al-Dahir, "Overview and comparison of equalization schemes for space-time-coded signals with application to EDGE," *IEEE Trans. Signal Processing*, vol. 50, pp. 2477–2488, Oct. 2002.
- [3] N. Al-Dahir, C. Fragouli, A. Stamoulis, W. Younis, and R. Calderbank, "Space-time processing for broadband wireless access," *IEEE Commun. Mag.*, vol. 40, pp. 136–142, Sept. 2002.
- [4] E. Biglieri, J. Proakis, and S. S. Shitz, "Fading channels: Information theoretic and communication aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.
- [5] H. Bolcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.
- [6] S. Borst and P. Whiting, "The use of diversity antennas in high-speed wireless systems: Capacity gains, fairness issues, multi-user scheduling," *Bell Labs Tech. Mem.*, 2001.
- [7] S. Boyd and L. Vandenberghe. (2001) Introduction to Convex Optimization With Engineering Applications. [Online]. Available: [www.stanford.edu/~boyd/cvxbook.html](http://www.stanford.edu/~boyd/cvxbook.html)
- [8] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sept. 1999.
- [9] —, "On achievable rates in a multi-antenna broadcast downlink," in *Proc. 38th Annual Allerton Conf. Communications, Control, Computing*, Oct. 2000, pp. 1188–1193.
- [10] S. Catreux, V. Erceg, D. Gesbert, and R. W. Heath, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commun. Mag.*, vol. 40, pp. 108–115, June 2002.
- [11] R. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, May 1993.



- [12] D. Chizhik, G. Foschini, M. Gans, and R. Valenzuela, "Keyholes, correlations and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 361–368, Apr. 2002.
- [13] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple input multiple output measurements and modeling in Manhattan," in *Proc. IEEE Vehicular Technology Conf.*, 2002, pp. 107–110.
- [14] Chong and L. Milstein, "The performance of a space-time spreading CDMA system with channel estimation errors," in *Proc. Int. Communications Conf.*, Apr. 2002, pp. 1793–1797.
- [15] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.
- [16] C.-N. Chuah, D. N. Tse, J. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.
- [17] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.
- [18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [19] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," in *Proc. Int. Symp. Information Theory Applications*, Nov. 2000, pp. 681–684.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] R. G. Gallager, "An inequality on the capacity region of multiaccess fading channels," in *Communication and Cryptography—Two Sides of One Tapestry*. Boston, MA: Kluwer, 1994, pp. 129–139.
- [24] 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.
- [25] D. Gesbert, M. Shafi, D. S. Shiu, P. Smith, and A. Nguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.* Special Issue on MIMO Systems, pt. I, vol. 21, pp. 281–302, Apr. 2003.
- [26] L. Greenstein, J. Andersen, H. Bertoni, S. Kozono, D. Michelson, and W. Tranter, "Channel and propagation models for wireless system design I and II," *IEEE J. Select. Areas Commun.*, vol. 20, Apr./Aug. 2002.
- [27] T. Guess and M. K. Varanasi, "Multiuser decision-feedback receivers for the general Gaussian multiple-access channel," in *Proc. Allerton Conf. Communications, Control, Computing*, Monticello, IL, Oct. 1996, pp. 190–199.
- [28] S. Hanly and D. Tse, "Multiaccess fading channels-Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2816–2831, Nov. 1998.
- [29] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, Apr. 2003.
- [30] —, "Cayley differential unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1485–1503, June 2002.
- [31] B. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, 2002, submitted for publication.
- [32] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in sum rate," in *Proc. 40th Allerton Conf. Communications, Control, Computing*, Monticello, IL, Oct. 2002.
- [33] M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.
- [34] H. Huang, H. Viswanathan, and G. J. Foschini, "Multiple antennas in cellular CDMA systems: Transmission, detection and spectral efficiency," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 383–392, July 2002.
- [35] H. C. Huang, S. Venkatesan, and H. Viswanathan, "Downlink capacity evaluation of cellular networks with known interference cancellation," in *Proc. DIMACS Workshop on Signal Processing Wireless Communications*, DIMACS Center, Rutgers Univ., Oct. 7–9, 2002.
- [36] S. Jafar, G. Foschini, and A. Goldsmith, "Phantomnet: Exploring optimal multicellular multiple antenna systems," in *Proc. Vehicular Technology Conf.*, 2002, pp. 261–265.
- [37] S. Jafar and A. Goldsmith, "Multiple-Antenna Capacity in Correlated Rayleigh Fading With no Side Information. [Online]. Available: <http://wsl.stanford.edu/publications.html>.
- [38] —, "Transmitter optimization and optimality of beamforming for multiple antenna systems with imperfect feedback," *IEEE Trans. Wireless Commun.*, submitted for publication.
- [39] S. A. Jafar and A. Goldsmith, "Transmitter optimization for multiple antenna cellular systems," in *Proc. Int. Symp. Information Theory*, June 2002, p. 50.
- [40] S. A. Jafar and A. J. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. Int. Symp. Information Theory*, June 2001, p. 321.
- [41] —, "Vector mac capacity region with covariance feedback," in *Proc. Int. Symp. Information Theory*, June 2001, p. 321.
- [42] S. A. Jafar, S. Vishwanath, and A. J. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. Int. Conf. Communications*, vol. 7, 2001, pp. 2266–2270.
- [43] N. Jindal, S. Jafar, S. Vishwanath, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 3–6, 2002.
- [44] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," in *Proc. Int. Symp. Inform. Theory*, June 2002, p. 500.
- [45] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE J. Select. Areas Commun.*, submitted for publication.
- [46] —, "Optimal transmission with imperfect channel state information at the transmit antenna array," *Wireless Personal Commun.*, submitted for publication.
- [47] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multi-antenna systems on flat fading channels," *IEEE Trans. Inform. Theory*, submitted for publication.
- [48] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels-Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1083–1102, Apr. 2001.
- [49] —, "Capacity and optimal resource allocation for fading broadcast channels-Part II: Outage capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1103–1127, Mar. 2001.
- [50] Y. Li, J. Winters, and N. Sollenberger, "MIMO-OFDM for wireless communication: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, pp. 1471–1477, Sept. 2002.
- [51] A. Lozano and C. Papadias, "Layered space-time receivers for frequency-selective wireless channels," *IEEE Trans. Commun.*, vol. 50, pp. 65–73, Jan. 2002.
- [52] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 213–225, Apr. 2002.
- [53] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.
- [54] —, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.
- [55] A. F. Molisch, M. Stienbauer, M. Toeltsch, E. Bonek, and R. S. Thoma, "Capacity of MIMO systems based on measured wireless channels," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 561–569, Apr. 2002.
- [56] A. Moustakas and S. Simon, "Optimizing Multi-Transmitter Single-Receiver (MISO) Antenna Systems With Partial Channel Knowledge. [Online]. Available: <http://mars.bell-labs.com>.
- [57] A. Narula, M. Trott, and G. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418–2433, Nov. 1999.
- [58] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
- [59] G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.
- [60] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

- [61] H. Sato, "An outer bound on the capacity region of the broadcast channel," *IEEE Trans. Inform. Theory*, vol. 24, pp. 374–377, May 1978.
- [62] S. Shamai and T. L. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inform. Theory*, vol. 48, pp. 938–942, Apr. 2002.
- [63] 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.
- [64] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE Vehicular Technology Conf.*, May 2001, pp. 1745–1749.
- [65] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
- [66] S. Simon and A. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 406–417, Apr. 2003.
- [67] —, "Optimality of beamforming in multiple transmitter multiple receiver communication systems with partial channel knowledge," in *Proc. DIMACS Workshop Signal Processing Wireless Communications*, DIMACS Center, Rutgers Univ., Oct. 7–9, 2002.
- [68] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. Int. Conf. Communications*, Apr. 2002, pp. 406–410.
- [69] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
- [70] 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.
- [71] A. Tulino, A. Lozano, and S. Verdu, "Capacity of multi-antenna channels in the low power regime," in *Proc. IEEE Information Theory Workshop*, Oct. 2002, pp. 192–195.
- [72] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319–1343, June 2002.
- [73] S. Vishwanath, S. Jafar, and A. Goldsmith, "Optimum power and rate allocation strategies for multiple access fading channels," in *Proc. Vehicular Technology Conf.*, May 2000, pp. 2888–2892.
- [74] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the capacity of multiple input multiple output broadcast channels," in *Proc. Int. Conf. Communications*, Apr. 2002, pp. 1444–1450.
- [75] S. Vishwanath, G. Kramer, S. Shamai(Shitz), S. A. Jafar, and A. Goldsmith, "Outer bounds for multi-antenna broadcast channels," in *Proc. DIMACS Workshop on Signal Processing Wireless Communications*, DIMACS Center, Rutgers Univ., Oct. 7–9, 2002.
- [76] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.
- [77] P. Viswanath and D. Tse, "On the capacity of the multi-antenna broadcast channel," in *Proc. DIMACS Workshop on Signal Processing Wireless Communications*, DIMACS Center, Rutgers Univ., Oct. 7–9, 2002.
- [78] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1277–1294, June 2002.
- [79] 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.
- [80] 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.
- [81] 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.
- [82] Y. Xin and G. Giannakis, "High-rate space-time layered OFDM," *IEEE Commun. Lett.*, pp. 187–189, May 2002.
- [83] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. Global Communications Conf.*, Oct. 2001, pp. 1344–1348.
- [84] W. Yu and J. M. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," in *Proc. Int. Symp. Information Theory*, June 2002, p. 498.
- [85] W. Yu, G. Ginis, and J. Cioffi, "An adaptive multiuser power control algorithm for VDSL," in *Proc. Global Communications Conf.*, Oct. 2001, pp. 394–398.
- [86] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for vector multiple access channels," in *Proc. IEEE Int. Symp. Information Theory*, 2001, p. 322.

- [87] W. Yu, W. Rhee, and J. Cioffi, "Optimal power control in multiple access fading channels with multiple antennas," in *Proc. Int. Conf. Communications*, 2001, pp. 575–579.
- [88] L. Zheng and D. Tse, "Optimal diversity-multiplexing tradeoff in multiple antenna channels," in *Proc. Allerton Conf. Communications, Control, Computing*, Monticello, IL, Oct. 2001, pp. 835–844.
- [89] L. Zheng and D. N. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the noncoherent multi-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.



**Andrea Goldsmith** (S'90–M'93–SM'99) received the B.S., M.S., and Ph.D. degrees in electrical engineering from University of California, Berkeley, in 1986, 1991, and 1994, respectively.

She was an Assistant Professor in the Department of Electrical Engineering, California Institute of Technology (Caltech), Pasadena, from 1994 to 1999. In 1999, she joined the Electrical Engineering Department, Stanford University, Stanford, CA, where she is currently an Associate Professor. Her industry experience includes affiliation with Maxim Technologies, Santa Clara, CA, from 1986 to 1990, where she worked on packet radio and satellite communication systems and with AT&T Bell Laboratories, Holmdel, NJ, from 1991 to 1992, where she worked on microcell modeling and channel estimation. Her research includes work in capacity of wireless channels and networks, wireless information and communication theory, multi-antenna systems, joint source and channel coding, cross-layer wireless network design, communications for distributed control and adaptive resource allocation for cellular systems and *ad-hoc* wireless networks.

Dr. Goldsmith is a Terman Faculty Fellow at Stanford University and a recipient of the Alfred P. Sloan Fellowship, the National Academy of Engineering Gilbreth Lectureship, a National Science Foundation CAREER Development Award, the Office of Naval Research Young Investigator Award, a National Semiconductor Faculty Development Award, an Okawa Foundation Award, and the David Griep Memorial Prize from University of California, Berkeley. She was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 1995 to 2002 and has been an Editor for the IEEE WIRELESS COMMUNICATIONS MAGAZINE since 1995. She is also an Elected Member of Stanford's Faculty Senate and the Board of Governors for the IEEE Information Theory Society.



**Syed Ali Jafar** (S'99) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT), Delhi, in 1997 and the M.S. degree in electrical engineering from California Institute of Technology (Caltech), Pasadena, in 1999. He is a Graduate Research Assistant in the Wireless Systems Lab, Stanford University, Stanford, CA, and is currently working toward the Ph.D. degree in electrical engineering.

He was a Summer Intern in the Wireless Communications Group of Lucent Bell Laboratories, Holmdel, NJ, in 2001 and has two pending patents resulting from that work. He was also an Engineer in the satellite networks division of Hughes Software Systems, India, from 1997 to 1998. His research interests include spread-spectrum systems, multiple antenna systems, and multiuser information theory.



**Nihar Jindal** (S'99) received the B.S. degree in electrical engineering and computer science from University of California, Berkeley, in 1999 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2001, and is currently working toward the Ph.D. degree at the same university.

His industry experience includes summer internships at Intel Corporation, Santa Clara, CA, in 2000 and at Lucent Bell Labs, Holmdel, NJ, in 2002. His research interests include multiple-antenna channels and multiuser information theory and their applications to wireless communication.



**Sriram Vishwanath** (S'99) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT), Madras, in 1998 and the M.S. degree in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 1999. He is a graduate fellow currently working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA.

His research interests include information and coding theory, with a focus on multiple antenna systems. His industry experience includes work at

National Semiconductor Corporation, Santa Clara, CA, in the Summer of 2000 and at the Lucent Bell Labs, Murray Hill, NJ, during the Summer of 2002.