# FEEDBACK AND COOPERATION IN INTERFERENCE-LIMITED WIRELESS NETWORKS

by

David Gesbert

A proposal presented to the
UNIVERSITY OF NICE
In Partial Fulfillment of the
Requirements for the
HABILITATION A DIRIGER DES RECHERCHES

April 2013

# Chapter 0

# Table of Contents

# Chapter 1

# Introduction and structure of document

This report is a supporting document towards my defense of the French Habilitation à diriger des recherches (HDR). Unlike the PhD thesis, it is not intended as a research monograph but rather a multi-faceted document aimed at illustrating the various dimensions of responsilities linked with my experience as a researcher, then as associate professor, in my 16 years following the PhD. In the writing we favor conciseness and the coherence among selected topics over exhaustiveness.

There are two main parts in this document, pointing respectively at the research contributions per se, then at the key elements of my CV. In the first part, we present a short write-up attempting to put into perspective the contributions brought to the area of optimization of wireless networks with a focus on multiple antenna (MIMO) communications. We introduce the general philosophy underpinning the undertaken research avenues, emphasizing a form of continuity (albeit with clear evolutions) behind the various directions followed over time. Since it is difficult to account in detail for all the research results obtained since the PhD (March 1997) in a legible manner, we make choices as to which problem are commented in more details and which are covered superficially. In the interest of the reader's time, we give a very short summary of earlier research contributions (for instance those dealing with point to point wireless optimization) and cover in a bit more details the more recent ones because of the greater relevance that these exhibit to some of the current hot topics in the area of communications research. This is particularly true when it comes to the study of multi-terminal cooperative networks.

In the second part, we provide an extended CV, cover teaching and research supervision responsibilities since PhD, as well as project management experience. We then give the complete list of publications, organized in book, book chapters, journal papers, conference papers, and patents. To conclude the document, three publications are selected and reproduced *in extenso*. They are selected among other publications for their tutorial level, and their value in highlighting the three major milestones in the evolution of multiple antenna systems's studies, namely (i) point-to-point MIMO (2003), (ii) point-to-multipoint (multiuser) MIMO (2007), and (iii) multipoint-to-multipoint (multicell) MIMO systems (2010).

# Chapter 2

# Summary of research contributions

## 2.1 Introduction to the research

In this part of the document, we describe the research directions in high level terms, trying to depict relations between addressed questions and most importantly their evolution over time. We try and show their relevance to the broad problem of wireless communications engineering, emphasizing common methodologies and assumptions. There is an unavoidable evolution in the types of research scenarios considered in the last 16 years, which is the result of personal choices and tastes. The most striking feature of this evolution lies in the fact that our early research interests are rooted in signal processing, focussing on algorithm design at *one* transmitter or *one* receiver while today's focus has a clear added networking tone.

Asserting that such evolutions are dictacted by pure independent will would be pretentious. The shift towards communication theoretic questions within a networking context also reflects the influence of the research community. Many of today's hot (or perceived to be) topics combine classical physical layer design together with *some* communications protocol issues. On can see there an effect of the growing convergence between wireless commmunications research and internet related questions. This is also reflected in how both public and industrial research funding is currently allocated.

In what follows we try and illustrate this evolution through a couple of key examples taken from our contributions, described with varying levels of details. In the face of having to choose what to bring forward, we opt to emphasize recent work (say the last 6 years) over earlier one. We also favor the presentation of concepts over the mathematical models, as the latter can be found in the publications.

The research contributions are selected as follows: We first illustrate a couple of past results pertaining to point to point link studies. These are covered very superfically. Multi-user studies are also summarized with a few key references given to the reader. We then turn to so-called interference networks, featuring multiple contending point-to-multipoint links, a scenario fulfilled by many modern wireless systems such as cellular, WLAN, or cognitive radio networks. In this, a subset of the accomplished research work is selected and examined in more detail, and specific approaches are described with the emphasis placed on the notions of information feedback and transmitter cooperation.

Finally the final document concludes by suggesting a couple of promising open problems for the way forward.

### 2.1.1 General interest and methodology

We are generally interested in the optimization of communications networks, and more specifically wireless networks, through the exploitation of tools from communications theory, information theory, and signal processing. By optimization, we understand the improvement of a variety of network's key performance indicators, such as the average throughput, the maximum throughput, fairness indicators, or in some cases, more abstract information theory related performance metrics such as the number of degree of freedom (DoF) or multplexing gain (MG) which indicate the scaling factor with which the network's throughput increases with the logarithm of the signal to noise ratio (SNR) in the high SNR region. The emphasis, to a large extent, has been on original scheme constructions at the physical or the link layer by exploiting systems's and signals' strucutural properties whenever possible. Some work was dedicated to performance analysis in some cases, in particular when such type of study serves to illustrate a new standpoint on system design or some interesting trade-off. An example is that of scaling law analysis for interference limited networks with multi-user diversity scheduling where our results demonstrate how powerful a suitable resource allocation can be in combatting interference. In retrospect, it is quite striking to see how questions of the type "how much information or knowleddge does it take to achieve a certain decoding or estimation task" or "how can data be communicated or extracted with *less* prior information" have remained a stable point of interest throughout the years (including PhD thesis on *blind* estimation!).

## 2.2 Point-to-point MIMO systems

### 2.2.1 Blind estimation for multi-channel systems

Our research is initially rooted in the area of receiver design for point-to-point wireless links, using considerably abstracted and simplified models for the communication protocols and focussing on statistical signal processing to improve receiver performance. During the PhD, several new techniques [1–4] are proposed which aim at the blind estimation or equalization of the propagation channel based on the sole observation of the channel output's second order statistics (i.e. without the use of pilot sequences). It is interesting to note that the common enabler behind the new blind approaches is the availability of multiple channel outputs corresponding to the same transmitted sequence. A typical instance of such channels is obtained in a scenario where the receiver is equipped with multiple antennas. Hence the work on blind second-order estimation can be cast quite naturally in a more general framework of studies emerging in the mid 90s under the acronym of MIMO (multiple input multiple output).

### 2.2.2 Study of point-to-point MIMO systems

In the late 90s, a transition takes place in the focus of research, as we move from blind estimation problems to transmitter and receiver design for MIMO systems and MIMO channel capacity studies. In the early 2000s, MIMO systems are a very promising field of research with huge potential, yet with a number of crucial unresolved questions remain. At this stage it is unknown to what extent the formidable gains promised by Shannon capacity analysis are realizable in practice for MIMO. Proper channel modeling is lacking and the most well known MIMO transmission methods are either too complex or too simple and not robust enough. At this stage, our research primarily focusses on deriving good insightful channel models for point to point MIMO systems, which was eventually achieved for both the flat-fading [5] and the OFDM-based frequency selective case [6]. The models we obtain there attracted some attention in their ability to explain quantitatively the role played by certain key parameters in increasing or reducing the MIMO capacity, such as the antenna spacing, the position of the surrounding scatterers, and the line of sight component. In particular the results in [5] predict the possibility of so-called pin-hole MIMO channels which have large size yet small rank, a finding confirmed in parallel by an experimental study made by Bell Labs researchers.

#### Space-time code design with limited feedback

In a second phase of studies, our effort is shifted on the problem of algorithm design for MIMO systems in the point to point case. In the early 2000s, much focus by the community is on the design of diversity oriented schemes, i.e. space time codes. In parallel, yet quite independently, the importance of feedback in improving wireless system performance is gaining recognition. Although timidly exploited in 3G networks, the trend towards building more powerful feedback links into newer types of systems is rapidly developing. Today, feedback is exploited for several purposes, from power control, to link adaptation, to resource allocation and scheduling, to multiple antenna transmit precoding and cooperative transmission. Yet, several years ago, feedback was still in its infancy. Among the interesting issues related to the combination of feedback and single user MIMO, one may ask the following question: How can space time codes, which are by design open loop (blind) diversity methods benefit from even a small amount of feedback? This specififc problem was addressed in [7] where we show a few bits on instantaneous feedback can substantially improve the classical space time block code (Alamouti type) performance. Then a more general class of feedback was investigated by which the receiver is allowed to send back long-term statistical information about the fading channel such as Ricean K factor transmit and receive spatial covariance matrices. There again, one shows how such information can be exploited by modifying the structure of the space time code. The idea, initially brought forward by previous authors, consists in precoding the space time code matrix with a long term beamforming matrix which is a function of second order channel statistics. A series of results are obtained for various classes of fading channels in [8–10].

## 2.3 Multi-terminal networks

### 2.3.1 Multi-user channels with limited feedback

Although feedback can improve the performance of point to point radio links, its role is nowhere as crucial as in the multi-terminal case. In [11] we highlight the importance of feedback in enabling efficient communications with multiple terminals simultaneously using the spatial dimention offered by MIMO. We build on prior expertise in exploiting feedback for single user links in order to explore feedback-based transmission schemes for the downlink of multi-user MIMO systems (a.k.a MIMO broadcast channels in the information theoretic terminology).

Since over-the-air feedback generates overhead and consumes some of the scarce spectral resources, we address the more specific problem of *limited feedback* scheme design. the concept of *limited* feedback encompasses a number of subscenarios ranging from situations where the feedback is limited by degradations of its quality which are hard to control by the system designer, to situations where the feedback quality is restricted *on purpose* to avoid an excessive comsumption of uplink resources. Designer-independent feedback quality degradation can originate from channel noise corrupting the fedback data or CSI aging caused by various sources of propagation, processing, decoding, and framing delays in the uplink channel. Designer-dependent reduced feedback schemes often rely on the use of coarse quantization schemes for channel representation (digital feedback) or the user of limited transmit power levels while emitting channel estimates (analog feedback). Our work addresses both situations and even combinations of these. A couple of examples of results are shown below.

The philosophy of our work within the area of reduced feedback can be summarized as follows: Although a complete feedback protocol would entail describing (quantizing) the complete channel vectors for all users with a large number of bits, reduced feedback schemes aim at extracting most of the feedback-related gains by exploiting a suitable *reduced dimensioned* representation of the same channels. Several results are obtained following this principle. They address mainly the problems of reduced feedback-based scheduling or reduced feedback-based MIMO spatial precoding.

#### Scheduling with reduced feedback

In the context of multi-user diversity scheduling, we propose in [12] (with the key idea originally presented in [13]) a method referred to as *selective multiuser diversity*, whereby only terminals which detect a high enough SNR, are allowed to engage in the feedback process towards the base station. The idea follows the intuition that only terminals with good channel conditions are likely to be selected by the scheduler, while others tend to waste feedback resources. There one shows that 90 percent of the capacity of a complete feedback based system can be preserved while just 10 percent of the feedback overhead is created.

**Precoding with reduced feedback**

In most MIMO-enabled systems, feedback is both used for the user selection as well as to bring vector channel state information to the transmitter so as to compute the precoding coefficients (multi-user beamforming). In [14] we show how the trade-off between multi-user diversity and multiplexing gain can be exploited to reduce feedback overhead. We proposed a method where a small portion of feedback is allocated to all users to describe the channel with just enough quality, based on which the base station performs user selection, while the rest of the feedback resource is allocated to just the users which were selected in the first round so as to describe their channel with enough accuracy for spatial precoding. This two stage feedback splitting approach is then optimized analytically.

In [15] limited feedback multi-user precoding is proposed based on the concept of random beamforming (RB). RB schemes are known to operate well with very low (just SNR) feedback which is used simply for scheduling purposes, but do not work well if the number of users is low because randomly designed beams are likely to miss all users. A new scheme is presented which exploits the feedback in the beamforming design stage, so as to obtain a family of beamforming schemes ranging from purely random to channel-dependant.

**Multiuser precoding with delayed feedback**

In most practical scenarios, perfect channel state information at transmitter (CSIT) may not be available due to the time-varying nature of wireless channels as well as the limited resource for channel estimation. Note that this is regardless of how many bits are affordable to use for channel quantization and feeback. Assuming, in the best case, that perfect estimation at the terminal takes place and infinite quantization is used, the CSIT will be available for precoding purposes only after a certain delay. The delay value typically depends on the framing structure (e.g. waiting time between an uplink feedback slot and the next downlink data slot) and pre-processing time at the base station. On the other hand, the *timeliness* of the delay (i.e. the relevance of the aged CSIT to the actual channel realization) depends on the ratio of the delay value to the fading coherence period. When the delay value approaches or exceeds the coherence period, an interesting question is whether delayed CSIT can still be exploited to precode the downlink data in any meaningful manner. We considered this problem in the context of the two-user MISO broadcast channel, where the transmitter equipped with $m$ antennas wishes to send two private messages to two receivers each with a single antenna. For the case of perfect CSIT, the optimal degrees of freedom (DoF) of this channel is two and achieved by linear strategies such as zero-forcing (ZF) beamforming. When the transmitter suffers from constant inaccuracy of channel estimation, it has been shown by Lapidoth et al. that the degrees of freedom per user is upper-bounded by $2/3$. It is also well known that the full multiplexing gain can be maintained under imperfect CSIT if the error in CSIT decreases as $O(P)$ or faster as $P$ grows [16].

Moreover, for the case of the temporally correlated fading channel such that the transmitter can predict the current state with error decaying as $O(P^{-\alpha})$ for some constant $\alpha$ [0; 1], ZF can only achieve a fraction of the optimal degrees of freedom [16]. This result somehow reveals the bottleneck of a family of precoding schemes relying only on instantaneous CSIT as the temporal correlation decreases. Three years ago, a breakthrough has been made in order to overcome such problem. In [17], Maddah-Ali and Tse showed a surprising result that even completely outdated CSIT can be very useful in terms of degree of freedom, as long as it is accurate enough. For a system with $m \geq 2$ antennas and two users, the proposed scheme in [17], hereafter called MAT, achieves the multiplexing gain of $2/3$ per user, irrespectively of the temporal correlation. This work shifts the paradigm of broadcast precoding from space-only to space-time alignment. The role of perfect delayed CSIT can be re-interpreted as a feedback of the past signal/interference heard by the receivers. This side information enables the transmitter to perform retrospective alignment in the space and time domain, as demonstrated in different multiuser network systems (see e.g. [18]). Despite its DoF optimality, the MAT scheme is designed assuming the worst case scenario where the delayed channel feedback provides no information about the current one (i.e. the delay value exceeds the fading coherence period). This assumption is over pessimistic as most practical channels exhibit some form of temporal correlation. Therefore a fundamental question arises as to whether a better way of exploiting both delayed CSIT and current (imperfect since predicted from the past one) CSIT exists. Studying the achievable DoF under such CSIT assumption is of practical and theoretical interest.

We proposed a novel strategy that combined the ZF precoding, based on the imperfect current state information, and the MAT alignment, based on the perfect past state information. The main role of current CSIT is to reduce, via spatial precoding, the overheard interference power in the original MAT alignment. This power reduction then enables, via source compression/quantization, to save the resources related to the transmission of the overheard interference. The overall scheme was shown to attain a DoF of $(2 + \alpha)/3$ where $\alpha$ is an index in $[0, 1]$ which indicates the quality of the channel prediction obtained from ther delayed feedback: i.e. $\alpha = 0$ indicates that no meaningful prediction can be realized because the delay exceeds the coherence time, while $\alpha = 1$ is the other extreme where quasi perfect prediction can be carried out. This DoF was also shown to be the best achievable one for the two user MISO channel scenario with delayed feedback. The initial result, published in [19], was lated generalized to encompass the interference channel scenario, then finally the MIMO scenario [19].

## 2.4 Interference and multicell networks

The transition from point to point radio link towards multi-user channels is extended quite naturally by adding a dimension featuring multiple parallel point-to-multipoint links. Such a situation can be found most commonly in cellular network with full spatial reuse of the spectrum, including 3G, LTE, LTE-A or WiFi networks or other networks where otherwise competing service providers share the same spectral resource (spectrum sharing, cognitive radios). In this model, multiple transmitters (e.g. base stations) share spectral resources while trying to communicate messages to or from their users. As the SNR increases, such network are typically inter-cell interference limited. Hence, the bulk of our work from 2009 addresses the problem of interference in multi-cell networks. Several avenues are considered for this task which are briefly commented below, with some examples of publications. It is worth noting that a common viewpoint exists behind most of these contributions. First, a central notion is that of a *cooperative communication* approach, by which otherwise interfering transmitters agree on jointly optimizing their transmission parameters for the benefit of the served terminals by virtue of a reduced interference level. The other transversal concept lies in the taking into account of a limited feedback overhead (hard or loose) constraint, similar to the adopted philosophy in the context of single-cell multi-user channels. In the context of multi-transmitter cooperation, the notion of feedback can be seen in the wide sense of the terms, encompassing any type of prior information (data, channel state,..) that is exchanged between (i) terminals and base stations, over-the-air, and (ii) the transmitters (e.g. the base stations) themselves over a pre-established signalling link.

Two fundamental question which underpin our research are as follows:

- how much feedback is necessary to tackle interference in cooperative networks?
- how can *local* information be exploited at each cooperative node for the benefit of the overall network (a.k.a distributed schemes)?

This is a vast and challenging problem however, for which our contributions are merely opening leads. In the below, we differentiate two kinds of approaches to cooperative multicell communications with limited feedback. The first tackles interference through radio resource management (RRM) methods such as power control or user scheduling. The other follows a signal-processing based cooperative communication framework, essentially based on coordinated (or jointly optimized) multiple antenna processing. The first approach is described briefly, while the second is covered is some more depth.

### 2.4.1 Interference coordination through cooperative RRM

Transmitters typically control a number of parameters at the two lowest layers which have a direct impact on the generated interference in neighboring cells. Interference coordination refers to a general class of interference control schemes which exploit the coordinated selection of such transmission parameters at the various mutually interfering base stations so as to minimize interference while maximizing their own user's received signal quality. For instance a coordination gain can be obtained by the jointly optimized choices of (i) power control levels, (ii) time or frequency slot assignment, (iii) code, in CDMA, (iv) user scheduling, and by generalization (this will appear more clearly in the next section): (v) multiple antenna precoding coefficients.

In [20] we present an overview of such RRM-based coordination methods with an emphasis on power control based coordination. When it comes to centralized multicell power control, in [21] we establish a striking result indicating that in certain classes of networks (two cell networks, or general networks with either low SNR or high SNR) then *binary* multicell power control is sum-rate optimal. From the initial centralized scheme, several distributed schemes can be derived [20].

In [22] we consider interference coordination using coordinated *user scheduling*. By resorting to ordered statistics tools, we establish fundamental limits to scaling laws of throughput versus the numner of users. We show that coordinated scheduling is a powerful tool towards reducing interference while allowing the use of simple distributed scheduling rules, such as the max SINR rule. The most important result indicates that coordinated scheduling can reduce interference to the point that the coordinated network can exhibit the *same* capacity scaling (with SNR) as a system without any interference. Interestingly, while coordinated scheduling can *reduce* interference, it cannot fully eliminate it, especially with a fine number of users to exploit. The exact limits are analyzed in [23]. As a result, our study indicates that a stronger formm of cooperation between interfering transmitters ought to be considered as well for practical nertwork scenarios. A solution to this problem can be found in the form of multiple antenna based cooperative communications, considered below.

### 2.4.2 Interference coordination with multi-antenna techniques

The role played by multiple antenna combining in mitigating interference by means of zero forcing (ZF) (or related criteria) beamforming is well established. Over the last few years, the combination of multiple-antenna approaches together with the concept of cooperation among interfering wireless devices was explored, showing strong promise. In particular, TX-based cooperation allows for avoidance of the interference before it even takes place (e.g. Network MIMO), or helps to shape it in a way which makes it easier for the receivers (RX) to suppress it. TX cooperation methods can be categorized depending on whether the data messages intended at the users must be known at several TXs simultaneously or not. For systems not allowing such an exchange (e.g. due to privacy regulations or low backhaul capabilities), interference alignment (IA) has been shown to be instrumental [24]. In contrast, when user data message exchange is made possible by a specific backhaul routing architecture, multi-cell, a.k.a. "network" MIMO, methods offer the best theoretical benefits [25].

#### When cooperation buys antennas

A distinct advantage of TX cooperation over conventional approaches relying on per-user interference rejection, lies in the reduced number of antennas needed at each RX to ZF residual interference. This gain is further amplified when user data messages exchange among TXs is made possible. For instance, in the case of three interfering two-antenna TXs, relying on RX based interference rejection alone requires three antennas to ZF the interference at each RX, while just two are needed when coordination is enabled via IA [26]. Further, if the three user messages are exchanged among the TXs, thus enabling network MIMO precoding, then just one antenna per TX and RX is sufficient to preserve interference-free transmission.

**Network MIMO with user grouping**

Network MIMO (Coordinated Multipoint Transmission - CoMP in the 3GPP terminology) includes strong forms of multicell cooperation, including the possibility of jointly combining all antennas belonging to several base station together, so as to mimick a large virtual MIMO system. A comprehensive tutorial was published regarding the theory of such systems in [27]. On the downlink of cellular networks, this implies the existence of a high capacity backhaul by which all user packets can be shared across the base stations, along with some of the channel state information. The concept of *grouping* of users and base stations into cooperation clusters can be used to limit the size and the complexity of such systems. In our earlier work in this area we examined some solutions based on dynamic cell grouping, where only a finite of neighboring base stations are allowed to cooperate together, the clusters are optimized based on graph theoretic algorithms [28].

### 2.4.3 Network MIMO with reduced user data sharing

As an alternative to clustering (which poses some edge effects issues), we also investigated how just the right amount of information can be exchanged between base stations so that the cooperation gains are balanced with the overhead linked to user data and CSI sharing. In [29] we formulate a rate optimization problem under the constraint of a finite backhaul. We show that packet data sharing is not always beneficial but becomes more so as the inteference is stronger. We adopt a rate splitting formalism to find the amount of traffic that ought to be shared between two interfering base stations. In doing so, we bring forward a family of theoretical MIMO communications scenario bridging the MIMO interference channels and the Network MIMO channels.

**MIMO cooperation with limited CIST feedback and exchange**

The benefits of multiple antenna transmit cooperation go at the expense of requiring channel state information (CSI) at the TXs. Indeed, whether one considers cooperation with or without user's data sharing, the TXs should in principle acquire the complete CSI pertaining to every TX and RX pair in the network. This is also the case for *wide-sense* distributed schemes (e.g. [30]) where the computation of precoders typically relies on iterative techniques where each iteration involve the acquisition of local feedback. Yet, as local feedback is updated over the iterations, this approach implicitly allows each TX to collect information about the precoders and channels of other TXs, hence amounting to an iterative global CSI acquisition at all TXs. Hence such *wide-sense* distributed schemes can be opposed to *strict sense* distributed schemes where transmitters have to make a one-shot decision about how to communicate-cooperate based on the sole local information they have initially acquired.

At first glance, it would seem that CSI feedback and sharing requirements in transmitter cooperation schemes grow unbounded as the network grows in size. Since over the air feedback and backhaul exchange links are always rate and latency limited, this means the practical application of TX cooperation in dense large networks may be very difficult, even if one assummes the user data packets can be routed for free to (i.e. shared across) all cooperating transmitters.

In a series of recent results, we challenge the common view that interfering TXs engaging in a cooperative scheme *can* or *should* share global (network-wide) CSI. Instead, we formulate the problem of a suitable CSI *dissemination* (or allocation) policy across transmitting devices while maintaining performance close to the full CSIT sharing scenario. We report below a couple of findings revealing how the need for CSIT sharing can be alleviated by exploiting specific antenna configurations or decay property of signal strength versus distance, hence making TX cooperation distributed and scalable. We use interference alignment and network MIMO respectively as our driving scenarios.

More specifically, for the cooperation scenario without user data message sharing where alignment of interference is sought, we show how perfect alignment is possible in certain antenna topologies without knowledge of all the channel elements at some TXs. For the network MIMO scenario, this is not the case and we illustrate instead how power decay versus distance can be exploited to substantially reduce the CSI sharing requirements while fulfilling optimal asymptotic rate performance conditions. A common trait behind the findings is that different cooperating TXs can (and often must) live with their own individual partial version of the global CSIT. Hence, CSIT representation quality is bound to be non-uniform across TXs. Consequently, we discuss briefly the problem of multiple-antenna precoding with *TX-dependent* CSI.

## 2.5 Key notions in spatial transmitter cooperation

We consider fast fading multiple-antenna wireless networks where the transmission can be mathematically represented by writing $\mathbf{y}_i$, the received signal at RX $i$, as

$$\mathbf{y}_i = \mathbf{H}_{ii}\mathbf{x}_i + \sum_{j \neq i} \mathbf{H}_{ij}\mathbf{x}_j \tag{2.1}$$

where $\mathbf{x}_i$ is the signal emitted by TX $i$ and $\mathbf{H}_{ij}$ is a matrix containing the channel elements between TX $j$ and RX $i$. The transmitted symbols $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_K]^{\mathrm{T}}$ are then obtained from the user's data symbols $\boldsymbol{s} = [s_1, \dots, s_K]^{\mathrm{T}}$ by multiplication with a precoder $\mathbf{T}$, i.e., $\mathbf{x} = \mathbf{T}\boldsymbol{s}$. If the user's data symbols are not shared between the TXs, the precoder $\mathbf{T}$ is restricted to a particular block-diagonal structure, while it can otherwise take any form. The received filter $\boldsymbol{g}_i^{\mathrm{H}}$ is then applied to the received signal $\boldsymbol{y}_i$ to obtain an estimate of the transmitted data symbol.

Here, we briefly discuss the leading techniques for MIMO based cooperation with or without user data message exchange. We point out commonly made assumptions in terms of CSIT sharing and feedback design.

### 2.5.1 Interference alignment for Interference Channels

When the user's data symbols are not shared between the TXs, the setting is referred to as a *Interference Channel (IC)* in the communication theoretic literature. In MIMO ICs, a method called *Interference Alignment (IA)* has been recently developed and shown to achieve the maximal number of degrees-of-freedom (DoF), or pre-log factor, in many cases [24, 26]. As a consequence, IA has attracted a lot of interest in the community. Here, we take the DoF as our key performance metric, such that we focus on IA schemes.

IA is said to be *feasible* if the antenna configuration (i.e. the distribution of antenna elements at the TXs and the RXs) yields enough optimization variables to allow for the interference-free transmission of all user's data symbols, which means fulfilling [26]

$$\forall i, \forall j \neq i,\ \ \boldsymbol{g}_i^{\mathrm{H}} \mathbf{H}_{ij} \boldsymbol{t}_j = 0. \tag{2.2}$$

Intuitively, IA consists in letting the TXs coordinate among themselves to beamform their signals such that the interferences received at each of the RXs are confined in a subspace of reduced dimensions, which can then be suppressed by linear filtering at the RXs with a smaller number of antennas.

### 2.5.2 Precoding in the Network MIMO

When the user's data symbols are shared between the TXs, the TXs form a *distributed antenna array* and a joint precoder can be applied at the transmit side [25]. Consequently, this setting becomes similar to the single TX multi-user MIMO downlink channel and the interference between the TXs can be completely canceled, e.g., by applying a global ZF precoder $\mathbf{T} \propto \mathbf{H}^{-1}$.

### 2.5.3 Limited feedback versus limited sharing

The limited feedback capabilities have been recognized as a major obstacle for the practical use of the precoding schemes described above. Consequently, a large literature has focused on this problem and both efficient feedback schemes and robust transmission schemes have been derived, for Network MIMO [31, 32] and IA [33, 34].

Yet, all these works assume that the imperfect channel estimates obtained via limited feedback are *perfectly* shared between all the transmit antennas. This is a meaningful assumption when the TXs are colocated but less realistic otherwise, as we shall now see.

#### CSIT sharing issues

One obstacle to the sharing of global CSIT follows from the fact that the amount of CSI which has to be exchanged increases very quickly with the number of TXs. In fact, each TX needs to obtain the CSI relative to the full multi-user channel, which consists of $(NK)^2$ scalars in a $K$-user setting with $N$ antennas at each node.

In addition, acquiring the CSI at a particular TX can be realized either by a direct broadcast of the CSI to all the listening TXs or by an over-the-air feedback to the *home* base station alone, followed by an exchange of the local CSIs over the backhaul, as it is currently advocated by 3GPP LTE-A standards [35]. Note that exchange over the backhaul can involve further quantization loss and may lead to a different CSI-aging at each TX, due to protocol latency. Either case, the channel estimates available at the various TXs will not be exactly the same. This leads to a form of CSI discrepancy which is inherent to the cooperation among non-colocated TXs.

In order to capture multiple-antenna precoding scenarios whereby different TXs obtain an imperfect *and* imperfectly shared estimate of the overall multi-user channel, we denote by $\mathbf{H}^{(j)}$ the network-wide channel matrix estimate available at TX $j$. Consequently, the precoding schemes have to be modified to take into account that each TX will compute its precoder based on its own channel estimate. Thus, TX $j$ transmits $\mathbf{x}_j = \mathbf{T}^{(j)}\boldsymbol{s}$ based on the knowledge of $\mathbf{H}^{(j)}$ only.

The fundamental questions which arise are: $(i)$ how complete and accurate should the estimate $\mathbf{H}^{(j)}$ be for each $j$ while operating under reasonable CSI overhead constraints? and $(ii)$ how should precoders be designed given the likely discrepancies between various channel estimates? Although these questions prove to be difficult and to a large extent remain open, we shed some light on the problem for two key scenarios in the following sections.

## 2.6 Aligning interference with incomplete CSIT

Let us first consider an IC, i.e., without user's data sharing. Feasibility studies for IA are typically carried out under the assumption of full CSIT. Yet, one can show that IA feasibility and the CSIT model are in fact tightly coupled notions. Assume for instance that all the RXs were given a generous number of antennas equaling or exceeding the number of TXs, it is well known that the interference could be suppressed at the RXs alone and no precoding, and hence no CSIT, is necessary. This example suggests the existence of a trade-off between the number of antennas and the CSI sharing requirements. Thus, it is possible to design IA algorithms using less CSIT that conventionally thought, without performance degradations by exploiting the availability of extra-antennas at a subset of devices. More specifically, the problem of finding the minimal CSIT allocation which preserves IA feasibility can be formulated. The minimality refers to the size of a CSIT allocation, defined as the total number of scalars sent through the multi-user feedback channel.

We differentiate between antenna configurations where IA is feasible and the number of antennas at the TXs and the RXs provide just enough optimization variables to satisfy alignment conditions, denoted as *tightly-feasible*, and the ones where extra antennas are available, denoted as *super-feasible*. Furthermore, we call a CSIT allocation *strictly incomplete* if at least one TX does not have the complete multi-user CSI. With such concepts in place, the following lesson can be drawn.

### 2.6.1 Tightly-feasible ICs

A strictly incomplete CSIT allocation implies that some TXs compute their precoders in order to fulfill IA inside a smaller IC formed by a subset of RXs and a subset of TXs. Most of the time, this creates additional constraints for the optimization of the other precoders which makes IA unfeasible. Yet, it can be shown that IA feasibility can be preserved under the following condition [36].

*Lesson* 1. In a tightly-feasible IC, there exists a strictly incomplete CSIT allocation preserving IA feasibility if there exists a tightly-feasible sub-IC strictly included in the full IC.

Exploiting this result, a CSIT allocation algorithm is derived in [36] along with an algorithm which achieves IA based on this incomplete CSIT allocation. In a few words, it consists in giving to each TX the CSI relative to the smallest tightly-feasible sub-IC to which it belongs. The precoders are then designed to align interference inside this smaller sub-IC, thence requiring only a part of the total CSIT, while IA feasibility is preserved. We will see in the simulations results presented in the following that the reduction in the CSIT size is significant. In fact, the reduction of the CSIT allocation feeds on the heterogeneity of the antenna configuration such that the more heterogeneous the antenna configuration is, the larger is the saving brought by using the minimal CSIT allocation. This is particularly appealing in regards to the future networks where mobile units and base-stations from different generations with different number of antennas are likely to co-exist.

### 2.6.2 Super-feasible ICs

In super-feasible settings, the additional antennas can be used to reduce the size of the minimal CSIT allocation. Yet, how to exploit optimally these additional antennas to reduce the feedback size is a very intricate problem. Still, a low complexity heuristic CSIT allocation can be derived [36]. The main idea behind the algorithm is to let some TXs or RXs ZF less interference dimensions such that small tightly-feasible settings are formed inside the original setting.

The effective CSIT reduction is illustrated in Figure 2.1 for a 3-user IC. The results are averaged over $1000$ random distributions of the antennas across the TXs and the RXs. If $12$ antennas are distributed between the TXs and the RXs, the setting is tightly-feasible and the previous CSIT allocation policy for tightly-feasible settings is used. With more than $12$ antennas, the algorithm exploits every additional antenna to reduce the size of the CSIT allocation.

When the setting is tightly-feasible, the reduction in feedback size requires neither a DoF reduction nor any additional antenna and comes in fact “for free”: It simply results from exploiting the heterogeneity in the antenna numbers at the TXs and the RXs.

## 2.7 A CSIT allocation policy for Network MIMO

When the user’s data symbols are jointly precoded at the TXs, complete CSIT allocation, in the sense defined above, is needed in all the practically relevant scenarios. So in this case, a

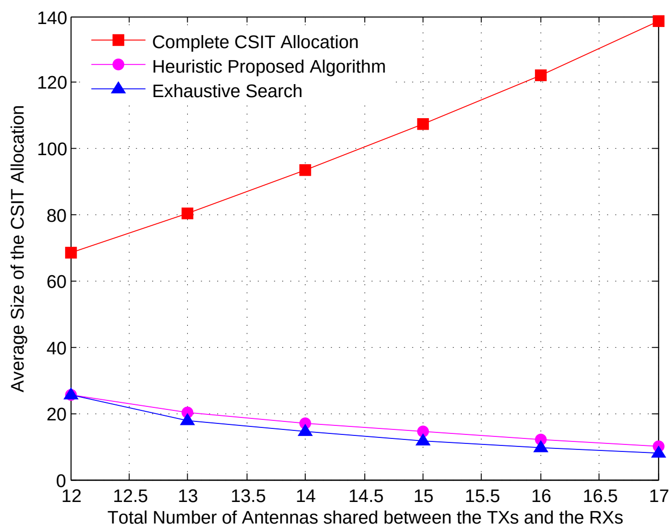



Figure 2.1: Average CSIT allocation size in terms of the number of antennas randomly distributed across the TXs and the RXs for $K = 3$ users.

different notion of reduced CSIT sharing must be advocated. The essential ingredient of this approach is the classical intuition that a TX should have a more accurate estimate for channels creating the strongest interference, i.e. originating from devices in the close neighborhood. This means that the fact that interference decays with pathloss can be exploited in principle to reduce the CSIT sharing requirements. This concept was recently introduced in [37]. A mathematical tool known as generalized degree-of-freedom comes handy to capture the effect of path loss on the multiplexing gain of cooperating networks with partially shared CSIT. Additionally, a simplified model referred to as Wyner model is used in this context to aid analytical tractability and provide first insights into this problem.

### 2.7.1 Generalized degrees of freedom

TX cooperation methods are often evaluated through the prism of DoF performance. Unfortunately the DoF is essentially pathloss-independent, such that a DoF analysis fails to properly capture the behavior of a large (extended) network MIMO. An extension of the notion of DoF, introduced in [38] as the *generalized DoF*, offers a much better grip over the problem as it can better take pathloss models into account. Upon defining the *interference level* $\gamma$ as $\gamma \triangleq \log(\text{INR})/\log(\text{SNR})$ with SNR denoting the signal-to-noise ratio and INR the interference-to-noise ratio, it is possible to define the *generalized DoF* as the DoF obtained when the SNR and the INR tend both to infinity for *a given interference level* $\gamma$.

For ease of exposition, we consider scenarios where all the TXs and RXs have a single antennas. The CSI is distributed, meaning that each TX has its own channel estimate based on which it computes its transmit coefficient without further exchange of information with the other TXs. We denote the estimate at TX $j$ by $\mathbf{H}^{(j)}$ and its $i$-th row, which corresponds to the channel from all TXs to RX $i$, by $\boldsymbol{h}_i^{(j)}$. We consider a digital quantization with a number of bits quantization $\boldsymbol{h}_i^{(j)}$ equal to $B_i^{(j)}$. Therefore, TX $j$ computes its own version of the precoding matrix $\mathbf{T}^{(j)}$ based on its own estimate $\mathbf{H}^{(j)}$. It then transmits $\mathrm{x}_j = \boldsymbol{e}_j^{\mathrm{T}}\mathbf{T}^{(j)}\boldsymbol{s}$.

### 2.7.2 The one-dimensional Wyner model

In the simple 1D *Wyner model* [39], the TXs are regularly placed along a line and solely the direct neighboring TXs emit non-zero interference. The channel is thus represented by a tridiagonal matrix. Furthermore, we assume that the interference from the direct neighbors are attenuated with a coefficient $\mu = P^{\gamma-1}$, according to the generalized DoF model.

Our objective is to evaluate how small $B_i^{(j)}$ can be while guaranteeing the same DoF as a system with perfect CSIT. Obviously, sharing the most accurate CSIT to all the TXs is a possible solution, yet, the size of the CSI required *at each TX* grows then unbounded with the number of users $K$, making this solution both inefficient and unpractical. In contrast, a much more efficient CSI sharing policy achieving the maximal generalized DoF, denoted as *distance-based*, is summarized below [37].

### 2.7.3 Distance-based CSIT allocation

We are interested in a CSIT allocation strategy, referred below as "distance-based", whereby each TX receives a number of CSI scalars which remains *bounded* as the number of users $K$ increases.

The distance-based CSIT allocation is obtained by setting for all $i, j$,

$$B_i^{(j)\,\mathrm{Dist}} = \lceil ([1+(\gamma-1)|i-j|]^+ + 2[\gamma + (\gamma-1)|i-j|]^+) \log_2(P) \rceil \tag{2.3}$$

where $[\bullet]^+$ is equal to zero if the argument is negative and to the identity function otherwise, and $\lceil \bullet \rceil$ is the ceiling operator. It can be shown that the CSIT allocation $\{B_i^{(j)\,\mathrm{Dist}}\}_{i,j}$ allows to achieve the maximal generalized DoF, i.e., those achieved in a system with perfect feedback [37]. The proof is based on the off-diagonal exponential decay of the inverse of the tridiagonal channel matrix.

Setting $\gamma = 1$ (no significant pathloss attenuation) in the previous equation, a conventional CSIT allocation is obtained where all channels are described with the same number of bits at all TXs (uniform allocation). For $\gamma < 1$ however, the number of bits allocated decreases with the distance $|i-j|$ between the considered TX and the index of the quantized channel, until no bit at all is used for quantizing the channel if the distance $|i-j|$ between the RX and the TX is larger than $\lceil 1/(1-\gamma) \rceil$. Crucially, this solution allocates to each TX a total number of bits which no longer grows with $K$ as only the CSI relative to a *neighborhood* is shared at each TX.

The different CSIT allocations are compared in Figure 2.2 for $K = 15$ users and $\gamma = 0.5$. The conventional CSIT allocation consists in providing the best quality to all the TXs, while the other strategies have the same size as the distance-based CSIT allocation, but the feedback bits are respectively shared uniformly and according to a conventional clustering of size 3. It can be seen from (2.3) that the ratio between the size of the distance-based CSIT allocation and the conventional CSIT allocation is independent of the SNR. Here, the distance-based CSIT allocation represents only $10\%$ of the size of the conventional CSIT allocation.

Hence, the distance-based CSIT allocation achieves the maximal number of generalized DoF with only a small share of the total CSIT, and outperforms the other schemes of comparison. Additionally, the user's data sharing can also be reduced to a neighborhood without loss of performance. Consequently, this scheme can be seen as an alternative to clustering in which the hard boundaries of the clusters are replaced by a smooth decrease of the level of cooperation.

## 2.8 Distributed precoding in the Network MIMO channel

### 2.8.1 Defining distributed precoding

In an effort to come up with scalable cooperative transmissions methods and reduce feedback overhead, one must be ready to consider the scenarios where the transmitters are not provided with the full multi-user CSI but rather an incomplete, imperfect, and importantly *individual* version of it. On the downlink, the base stations seek to jointly combine the users signals by using

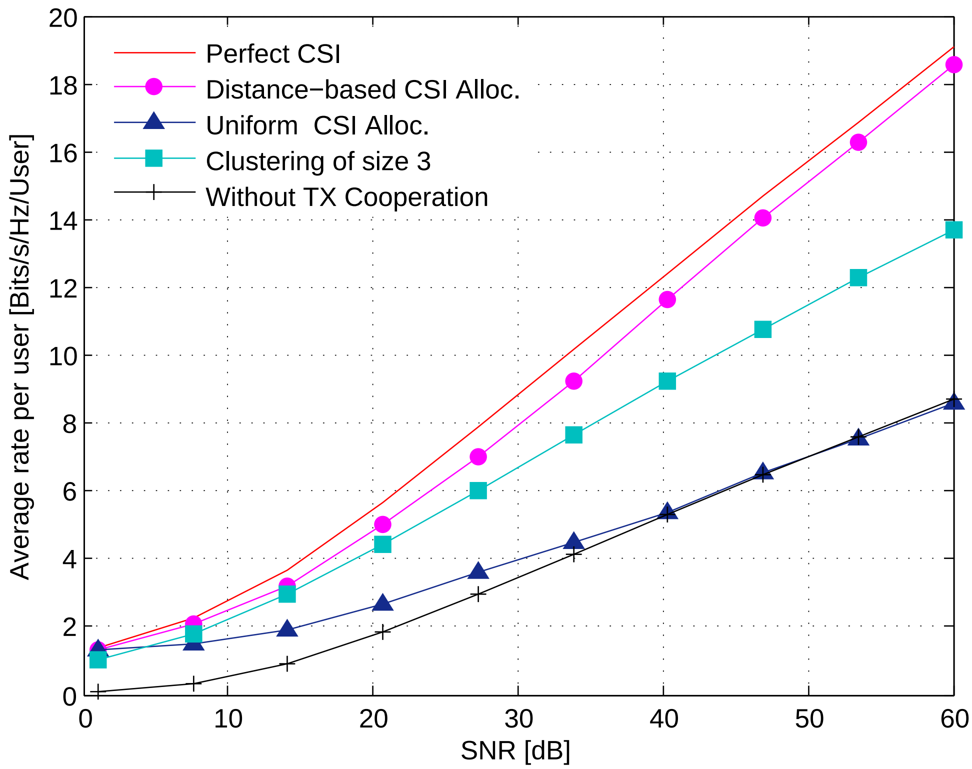



Figure 2.2: Average rate per user in terms of the SNR. The distance-based CSIT allocation, the uniform allocation, and the clustering one have all a size equal to $10\%$ of the size of the conventional CSIT allocation.

precoding coefficients which must be computed on the basis of this individual CSI estimates. This scenario, coined Distributed CSI Network MIMO precoding was initially introduced to the community in our papers [40–42]. Precoding over the D-CSI Network MIMO channel poses new challenges due to the distributed nature of the information structure. We established connections with a field known in the optimization theory field as *Team Decision* theory. In this field, a set of network agents seek to cooperate in order to maximize a common objective. There, performance loss with respect to a centralized optimization does not come from conflict or competition between the agents but from the fact that they are unable to share their views of the system state. In our wireless scenario, cooperating base stations would like to jointly invert a common multi-user channel matrix, unfortunately each base station possesses just an incomple and imperfect channel matrix estimate. For instance, a base has accurate information about terminals in its neighborhood but weak knowledge about more distant users. The fundamental questions which arise in this context are:

- Can one use classical precoding methods (zero-forcing, MMSE, etc..)
- What are optimal (robust) methods? What is their gain?
- What are low complexity distributed precoding techniques?
- Under a total feedback and information exchange overhead constraint, how should individual user CSI be disseminated throughout the network?

Below we depict some leads to start answering the first and second questions. One key lesson is the important role played by the *consistency* of information shared by the cooperating nodes over the *accuracy*.

### 2.8.2 Initial results on distributed cooperative precoding

As was clear from the previous sections, an efficient CSI dissemination policy naturally leads to a significant reduction of the CSI sharing requirements. As a consequence, the CSIT is represented non-uniformly across the TXs. Since non-uniform sharing is the best strategy in order to maximize performance under a given feedback overhead constraint, it is a natural consequence that some user's channels will be coarsely described at certain TXs and more accurately at others. Hence, a distributed CSI network MIMO is likely to occur in such cases.

Interestingly, the problem of designing precoders that can accommodate such a peculiar CSIT scenario is by and large open. In particular, new robust precoding schemes should be developed, as conventional precoders are designed under the assumption that the same imperfect CSIT is shared *perfectly* among the TXs.

Let us consider the model of distributed CSI described in Section 2.7 where TX $j$ receives its own channel estimate $\mathbf{H}^{(j)}$ with the $i$-th row, denoted by $\boldsymbol{h}_i^{(j)}$, obtained using $B_i^{(j)}$ quantizing bits. Assume a network where each TX has roughly the same average pathloss to each RX. The DoF which can be achieved with limited feedback is studied in [32] for the single TX MIMO downlink channel. We can extend this to the setting of non-uniform CSI so as to gain insight into the design of efficient precoders. In this case, *CSI scaling* coefficients $\alpha_i^{(j)}$ are introduced and defined by the limit of $B_i^{(j)}/((K-1)\log_2(P))$ when $P$ goes to infinity.

ZF is widely used and well known to achieve the maximal DoF in the MIMO downlink channel with perfect CSIT [32]. One may wonder how conventional ZF performs in the presence of CSI discrepancies brought by imperfect sharing. The answer is strikingly pessimistic: The sum DoF achieved can be shown to be equal to just $K \min_{i,j} \alpha_i^{(j)}$ [40]. Intuitively, this can be restated as follows.

*Lesson* 2. In the network MIMO with distributed CSI, the worst channel estimate across the TXs and the users limits the DoF achieved by each user using ZF precoding.

This is in strong contrast to the single TX case studied in [32] where the quality of the feedback of user $i$ relative to $\boldsymbol{h}_i$ solely impacts the DoF of user $i$. It shows clearly the disastrous impact of the CSI non-uniformity, since one inaccurate estimate at one TX degrades the performances of *all* the users. One may also wonder whether conventional-type robust precoders [31] can offer a better response. The answer is negative, unfortunately. Instead, a novel precoder design is needed that is tailored to the non-uniform CSIT sharing model.

Preliminary results to this end [40] suggest that it is possible to dramatically improve the DoF in certain scenarios. For instance, in the two-TX network, a scheme referred to as *Active-Passive (AP) ZF*, consisting in letting the TX with degraded CSIT arbitrarily fix its transmit coefficient while the other TX compensates to zero-out the interference, can be shown to recover the optimal DoF.

## 2.9 Open Problems

Transmitter cooperation method form a promising answer to the interference problem in wireless networks. New techniques for feedback limitation and for the CSIT sharing in wireless networks have been derived and their potential to reduce signaling overhead has been shown. We have presented some insights into a new problem which presents serious challenges, but also research opportunities for the future. Firstly, IA algorithms with incomplete CSIT are based on a DoF-preserving criterion only, i.e., on the performance at asymptotically high SNR. The impact of the incomplete CSIT on the performance at finite SNR should then be investigated to obtain practical solutions. Similarly, robust precoding schemes for the MIMO network with distributed CSI have so far considered DoF only as a metric. By and large, precoding over network MIMO with distributed CSI remains a challenging problem. Designing robust schemes at finite SNR is both interesting and, unfortunately, particularly difficult. It is important to note the difference with the classical and much easier task of designing robust schemes for the single decision-maker scenario (for instance for the MIMO broadcast channel). Here the cooperating nodes must make a transmission decision while guessing the strategy used by the others, despite not sharing the same view of the system state. Although this problem may in some case be recast in the form of a single decision making problem with a Bayesian estimation framework, the algorithms one derives with such approaches are typically very complex, hence lack applicability. Simple, elegant, closed-form expressions that are function of the local channel knoweledge and the knowledge of the statistics for non local channels are yet to be obtained.

# Bibliography

[1] D. Gesbert and P. Duhamel, "Unbiased blind adaptive channel identification and equalization," *Signal Processing, IEEE Transactions on*, vol. 48, no. 1, pp. 148–158, 2000.

[2] D. Gesbert, J. Sorelius, P. Stoica, and A. Paulraj, "Blind multiuser mmse detector for cdma signals in isi channels," *Communications Letters, IEEE*, vol. 3, no. 8, pp. 233–235, 1999.

[3] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *Signal Processing, IEEE Transactions on*, vol. 45, no. 9, pp. 2307–2317, 1997.

[4] D. Gesbert, A.-J. Van der Veen, and A. Paulraj, "On the equivalence of blind equalizers based on mre and subspace intersections," *Signal Processing, IEEE Transactions on*, vol. 47, no. 3, pp. 856–859, 1999.

[5] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor mimo wireless channels: Models and performance prediction," *Communications, IEEE Transactions on*, vol. 50, no. 12, pp. 1926–1934, 2002.

[6] H. Böcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of wireless systems employing ofdm-based spatial multiplexing," 2002.

[7] J. Akhtar and D. Gesbert, "Extending orthogonal block codes with partial feedback," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 6, pp. 1959–1962, 2004.

[8] A. Hjorungnes and D. Gesbert, "Precoded orthogonal space–time block codes over correlated ricean mimo channels," *Signal Processing, IEEE Transactions on*, vol. 55, no. 2, pp. 779–783, 2007.

[9] ——, "Precoding of orthogonal space-time block codes in arbitrarily correlated mimo channels: Iterative and closed-form solutions," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 3, pp. 1072–1082, 2007.

[10] A. Hjorungnes, D. Gesbert, and J. Akhtar, "Precoding of space-time block coded signals for joint transmit-receive correlated mimo channels," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 3, pp. 492–497, 2006.

[11] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the mimo paradigm," *Signal Processing Magazine, IEEE*, vol. 24, no. 5, pp. 36–46, 2007.

[12] V. Hassel, D. Gesbert, M.-S. Alouini, and G. E. Oien, "A threshold-based channel state feedback algorithm for modern cellular systems," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 7, pp. 2422–2426, 2007.

[13] D. Gesbert and M. S. Alouini, "How much feedback is multi-user diversity really worth?" *Proc. IEEE ICC*, vol. 4, pp. 234–238, 2004.

[14] R. Zakhour and D. Gesbert, "A two-stage approach to feedback design in multi-user mimo channels with limited channel state information," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*. IEEE, 2007, pp. 1–5.

[15] M. Kountouris, D. Gesbert, and T. Salzer, "Enhanced multiuser random beamforming: Dealing with the not so large number of users case," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 8, pp. 1536–1545, 2008.

[16] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[17] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, 2010, pp. 1188–1195.

[18] S. Jafar, *Interference Alignment*, ser. Foundations and trends in communications and information theory. Now Publishers Incorporated, 2011. [Online]. Available: http://books.google.fr/books?id=GfwB7ItK4esC

[19] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally-correlated mimo networks with delayed csit," *arXiv preprint arXiv:1211.3322*, 2012.

[20] D. Gesbert, S. G. Kiani, A. Gjendemsj *et al.*, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2393–2409, 2007.

[21] A. Gjendemsj, D. Gesbert, G. E. Oien, and S. G. Kiani, "Binary power control for sum rate maximization over multiple interfering links," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3164–3173, 2008.

[22] D. Gesbert and M. Kountouris, "Rate scaling laws in multicell networks under distributed power control and user scheduling," *Information Theory, IEEE Transactions on*, vol. 57, no. 1, pp. 234–244, 2011.

[23] P. de Kerret and D. Gesbert, "The asymptotic limits of interference in multicell networks with channel aware scheduling," in *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*. IEEE, 2011, pp. 466–470.

[24] V. R. Cadambe and S. A. Jafar, “Interference alignment and degrees of freedom of the K-user interference channel,” *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[25] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, “Network coordination for spectrally efficient communications in cellular systems,” *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[26] C. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, “On feasibility of interference alignment in MIMO interference networks,” *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[27] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, “Multi-cell mimo cooperative networks: A new look at interference,” *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1380–1408, 2010.

[28] A. Papadogiannis, H. J. Bang, D. Gesbert, and E. Hardouin, “Efficient selective feedback design for multicell cooperative networks,” *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 1, pp. 196–205, 2011.

[29] R. Zakhour and D. Gesbert, “Optimized data sharing in multicell mimo with finite backhaul capacity,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 6102–6111, 2011.

[30] K. Gomadam, V. R. Cadambe, and S. A. Jafar, “A distributed numerical approach to interference alignment and applications to wireless interference networks,” *IEEE Trans. Inf. Theo.*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[31] M. B. Shenouda and T. N. Davidson, “Robust linear precoding for uncertain MISO broadcast channels,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2006.

[32] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[33] H. Bolcskei and I. J. Thukral, "Interference alignment with limited feedback," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2009.

[34] O. E. Ayach and R. W. Heath, "Interference alignment with analog channel state feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 626–636, Feb. 2012.

[35] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS long term evolution: From theory to practice*, 2nd ed. Wiley, 2011.

[36] P. de Kerret and D. Gesbert, "Interference alignment with incomplete CSIT sharing," submitted to IEEE Trans. Wireless Commun., Nov. 2012. [Online]. Available: http://arxiv.org/pdf/1211.5380v1.pdf

[37] ——, "CSI feedback allocation in multicell MIMO channels," in *Proc. International Conference on Communications (ICC)*, 2012.

[38] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[39] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[40] P. de Kerret and D. Gesbert, "The multiplexing gain of the network mimo channel with distributed csi," *IEEE Transactions on Information Theory*, Nov. 2011.

[41] R. Zakhour and D. Gesbert, "Team decision for the cooperative MIMO channel with imperfect CSIT sharing," in *Proc. Information Theory and Applications Workshop (ITA)*, 2010.

[42] P. de Kerret and D. Gesbert, “The multiplexing gain of a two-cell MIMO channel with unequal CSI,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2011.

Chapter 3

# Extended CV

**David Gesbert** **gesbert@eurecom.fr**
Born Sept. 30th, 1969, France

---

***Employment***

***Since Jan. 2006 Professeur Classe 2, Dept. of Mobile Communications, EURECOM, Sophia Antipolis, France***

***Since 2011, Head of Mobile Communications Dept., EURECOM***
***Since 2009 Leader of Communications Theory Research Group.***

*2003-2006 Maitre de Conferences (Assistant Professor), Dept. of Mobile Communications, EURECOM, Sophia Antipolis, France*

*2002-2006, Adjunct Professor, Dept. of Informatics, Univ. of Oslo, Norway,*

*October 1998-October 2001 Co-founder, Iospan Wireless, Inc., San Jose California, (today part of INTEL)*
*Iospan was an R&D startup in the area of broadband wireless internet systems, spun off of my former Stanford University research group. Technology featured MIMO-OFDM and influenced early 802.16 standardization.*

*March 1997 – October 98, Postdoc researcher, Stanford University, Information Systems Laboratory, Stanford, California.*

*June 1993 – March 1997, Research Scientist, French Telecom Research Center (CNET), and Ecole Nationale Superieure des Telecommunications, Paris, France.*

***PhD Supervisor***

***Supervised 10 completed PhD students since 2002, + 7 ongoing.***

***Projects***

*Since 2004, leader at EURECOM for 5 EU projects, 2 ANR projects, 4 CIFRE projects, 3 industry projects*

***Teaching***

*Teaches two full courses at Eurecom: Advanced Topics in Wireless, Information Theory.*
*Since 2005, Academic leader for Engineering Specialization in "Mobile Communications".*

***Education***

***Ph.D. Electrical Engineering**, Ecole Nationale Superieure des Telecommunications (ENST), Paris, France, 1997. Specialty: Signal and Communications. Highest honors.*

***M.Sc. Electrical Engineering** degree from Telecom SudParis (INT), Evry, France, 1993. Electrical Engineering.*

***Publications***

*About 200+ publications in IEEE journals, and international peer-reviewed conferences in the areas of communication theory, signal processing, and wireless networking.*

**Impact**

*H-index (according to Google Scholar): 42*
*Paper citations (according to Google Scholar): 11000+*
.

**Honors**

*Recipient of IEEE Fellow (2011).*
*Recipient Best Paper Award 2012, IEEE Signal Processing Magazine*
*Co-Author of paper receiving Best Paper Award (Young Author) 2006, IEEE Signal Processing Society.*
*Recipient of IEEE Best Tutorial Paper of the Year 2004, Communications Society.*
*Recipient of ACM MSWiM 2004 Conference Best Paper Award*

**Patents**

*Co-author of 18 patents applications, all in the area of communication systems.*

**Past Committees**

*Elected member of the IEEE Signal Processing for Communications and Networking technical committee (2004-2009).*
*Member of CNRS French Network Expert Committee (2004-2009).*

**Expertise**

*Expert reviewer for Hong Kong Research Council, Irish Science Foundation, Swedish Foundation for Strategic Research, French Research Agency, Saudi Arabia Research Council, Qatar Research Council.*

**Editor**

*Guest Editor, IEEE Journal Selected Topics in Signal Processing, "Massive MIMO systems" 2013*

*Lead Guest Editor, IEEE Journal Areas in Communications, Special Issue on "Cooperative communications in MIMO cellular networks", 2010.*

*Guest Editor, IEEE Journal Areas in Communications, Special Issue on "Communications with limited feedback", 2007.*

*Guest Editor EURASIP Journal on Applied Signal Processing, Special Issue on "MIMO systems with partial feedback", 2007.*

*Guest Editor IEEE Journal on Wireless Communications, Special Issue on "Recent advances in multiple-antenna communications systems", 2006.*

*Guest Editor IEEE Journal Selected Areas in Communications, Special Issue on "MIMO wireless systems" 2003.*

*Guest Editor EURASIP Journal on Applied Signal Processing, Special Issue on "System integration oriented transceiver design for wireless networks beyond 3G", 2004.*

**Organizer/ Chair**

*General Chair, IEEE Communication Theory Workshop (CTW), Phuket, Thailand, 2013*

*Technical Chair, IEEE Communications Theory Symposium, International Conference on Communications 2013, Budapest, Hungary.*

*Technical Program Co-Chair, Communication Theory Symposium of the IEEE Personal Indoor Mobile Radio Communications Conference, 2008, Cannes, France.*

*Co-Chair, RAWNET'08 Workshop on Resource Allocation in Wireless Networks, April 2008, Berlin, Germany.*

*Co-organizer and Technical Program Committee Chairman, IEEE Workshop on Signal Processing Advances in Wireless Communications, June 2006.*

*Technical Chair, "Fifth Annual Smart Antenna Workshop", July 1998, Stanford University. Leading scientific meeting in smart antennas for mobile communications, with more than 200 international attendees.*

*Member of Technical Program Committee for numerous IEEE Conferences since 2001 and several special sessions.*

***Keynotes/ Panels***

*Keynote speaker at Sixth IEEE Workshop on Advanced Information Processing for Wireless Communication Systems*

*Keynote speaker at IEEE Workshop on Heterogeneous Networks (HETnet) 2011, Kyoto, Japan*

*Keynote speaker at IEEE International Symposium on Wireless Communications Systems, Bristol, England, 2010.*

*Keynote speaker at IEEE International Symposium on Wireless Communications Systems 2007, Trondheim, Norway.*

*Tutorial presenter at IEEE PIMRC Conference, Athens, Sept. 2007*

*Tutorial presenter at the IEEE ICASSP Conference, 2005 Toulouse, France.*

*Invited panelist at CROWNCOM 2010, ICC2009, ICC'03, WCNC'07.*

## Book

" Space-Time Wireless Systems: From Array Processing to MIMO Communications ", by H. Boelcskei, D. Gesbert, C. Papadias, A. J. van der Veen, Editors. Cambridge University Press ISBN 052185105x. Spring 2006.

## Chapters in Books

1. D. Gesbert, R. Zakhour, "Downlink Decentralized Multi-cell Transmission", Chapter in book *Coordinated Multi-point in Mobile Communications: From Theory to Practice*, Cambridge Univ. Press, 2011.

2. T. Salzer, D. Gesbert, F. Tosato, C. van Rensburg, "Multiple Antenna techniques in LTE and LTE Advanced", Chapter in book *UMTS Long Term Evolution: From Theory to Practice*, Second Edition covering LTE-Advanced", pub. J. Wiley, to appear 2010.

3. D. Gesbert, F. Tosato, C. van Rensburg, "Multiple Antenna techniques in LTE", Chapter in book "UMTS Long Term Evolution: From Theory to Practice", pub. J. Wiley, 2009. (An extract of the chapter is only available for download for copyright issues with Wiley, sorry).

4. D. Gesbert, J. Akhtar, `` Transmitting over ill-conditioned MIMO channels: From spatial to constellation multiplexing ". In *Smart Antennas: State of the Art* . S. Kaiser et al. Editors, EURASIP BOOK SERIES, 2005.

5. I. Stojanovic, M. Airy, D. Gesbert, H. Saran, `` Performance of TCP/IP Over Next Generation Broadband Wireless Access Networks ". In *Wireless IP and Building the Mobile Internet* . S. Dixit, R. Prasad Eds, Artech House Pubs, 2003.

6. A. Paulraj, D. Gesbert and C. Papadias,`` Antenna Arrays for mobile communications ", in *The Encyclopedia of Electrical and Electronics Engineering* , J. Webster Ed., John Wiley and Sons, New York, 1999.

7. D. Gesbert, B. C. Ng and A. Paulraj, `` Blind space-time receivers for CDMA communications ", in *Spread-spectrum: Developments for the new Millenium* , Kluwer Acad. Publishers, sept. 1998.

## Journal papers

### Tutorial/Magazine papers

1. P. de Kerret, D. Gesbert, "CSI Sharing Strategies for transmitter cooperation in wireless networks", in IEEE Wireless Communications Magazine, Feb. 2013.

2. D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference", in IEEE Journal on Selected Areas in Communications, Dec. 2010.

3. D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, M. Andrews, "An overview of limited feedback in wireless communications systems" (Tutorial Paper), IEEE Journal on Selected Areas in Communications, Oct. 2008.

4. D. Gesbert, M. Kountouris, R. Heath, C-B. Chae, T. Salzer, "From Single User to Multiuser Communications: Shifting the MIMO paradigm", IEEE Signal Processing Magazine, Sept. 2007. **Recipient of the 2012 Signal Processing Society Signal Magazine Best Paper Award.**

5. D. Gesbert, G. Oien, S. Kiani, A. Gjendemsjo, "Adaptation, Coordination and Distributed Resource Allocation in Interference-Limited Wireless Networks", Proceeding of the IEEE, December 2007.

6. D. Gesbert, M. Shafi, D. Shiu, P. Smith, " From theory to practice: An overview of space-time coded MIMO wireless systems ", *IEEE Journal on Selected Areas on Communications* (JSAC). April 2003, special issue on MIMO systems. (**Recipient of the 2004 IEEE Best Tutorial Paper Award by IEEE Comm. Society**).

## Regular transactions papers

1. P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the K-User time correlated broadcast channel with delayed CSIT", submitted to arXiv, [online] http://arxiv.org/abs/1301.2138.

2. P. de Kerret and D. Gesbert, "Spatial CSIT Allocation Policies for Network MIMO Channels", submitted to Transactions on Information Theory, Jan. 2013. Also available under http://arxiv.org/abs/1302.5376

3. Paul de Kerret and David Gesbert, "Interference alignment with incomplete CSIT sharing", submitted to Transactions on Wireless Communications, Nov. 2012. Also available under arxiv http://arxiv.org/pdf/1211.5380.pdf

4. X. Yi, S. Yang, D. Gesbert, M. Kobayashi ''The Degree of Freedom Region of Temporally Correlated MIMO Networks with Delayed CSIT", to appear in IEEE Trans. on Information Theory, October 2012. Available under arxiv.org/abs/1211.3322

5. B. Godana, D. Gesbert "Egoistic vs. Altruistic beamforming in the presence of feedback and backhaul delays", submitted to IEEE Trans. on Signal Processing, October 2012.

6. X. Yi, D. Gesbert, "Precoding Methods for MISO Broadcast Channel with Delayed CSIT", to appear in IEEE Trans. Wireless Communications, 2013. Also available under arxiv at http://arxiv.org/abs/1207.2103

7. S. Yang, M. Kobayashi, D. Gesbert, X. Yi, "On the Degree of Freedom Region of Time Correlated MISO Broadcast Channels with Delayed CSIT", to appear in IEEE Transactions on Information Theory, 2013. Also available under ArXiv.

8. H. Yin. D. Gesbert, M. Filippou, Y. Liu "A Coordinated Approach to Channel Estimation in Large-scale Multiple-antenna Systems", to appear in IEEE Journal on Selected Areas in

Communications, Special Issue on Large Scale Antenna Systems. 2013. Available under arxiv.org/abs/1203.5924.

9. P. de Kerret, D. Gesbert, "The Degrees of Freedom of the Network MIMO Channel With Distributed CSI", in the IEEE Transactions on Information Theory, Nov. 2012.
10. R. Zakhour, D. Gesbert, "Optimized data sharing in multicell MIMO with finite backhaul capacity", in IEEE Transactions on Signal Processing, Dec. 2011.
11. D. Gesbert, M. Kountouris, "Rate Scaling Laws in Multicell Networks under Distributed Power Control and User Scheduling", in IEEE Trans. On Information Theory, Jan. 2011.
12. H. J. Bang, D. Gesbert, P. Orten, "On the rate gap between multi and single cell processing under opportunistic scheduling", in IEEE Trans. on Signal Processing, Jan. 2012.
13. R. Zakhour, D. Gesbert, "Distributed multicell MIMO precoding using the layered virtual SINR framework". In the IEEE Transactions on Wireless Communications, August 2010.
14. U. Salim, D. Gesbert, D. Slock "Combining training and quantized feedback in Multi-Antenna Reciprocal Channels", IEEE Trans. on Signal Processing, March 2012.
15. Z. Ho, D. Gesbert, "Balancing Egoism and Altruism on the MIMO interference channel", in preparation (available under arxiv also).
16. E. Bjornson, R. Zakhour, D. Gesbert, B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI", IEEE Trans. on Signal Processing, Aug. 2010.
17. A. Papadogiannis, H. Bang, D. Gesbert, E. Hardouin, "Efficient Selective Feedback Design for Multicell Cooperative Networks" in IEEE Trans. Vehicular Technologies, October 2010.
18. A. Papadogiannis, E. Hardouin, and D. Gesbert, " Decentralising Multi-Cell Cooperative Processing on the Downlink: a Novel Robust Framework " in EURASIP Journal on Wireless Communications and Networking, Special Issue on Broadband Wireless Access, August 2009.
19. Nadia Fawaz, Keyvan Zarifi, Merouane Debbah, David Gesbert, " Asymptotic Capacity and Optimal Precoding in MIMO Multi-Hop Relay Networks", in IEEE Trans. on Information Theory, April 2011.
20. K. Tourki, D. Gesbert, L. Deneire, "Cooperative diversity using per-user power control in the MAC channel" in Elsewier Journal on Physical Communication, 2010.
21. F. Kaltenberger, M. Kountouris, D. Gesbert, R. Knopp, "On the trade-off between feedback and capacity in measured MU-MIMO channels", in IEEE Transactions on Wireless Communications, September 2009.
22. S. G. Kiani, D. Gesbert, A Gjendemsjø, G. E. Øien "Distributed Power Allocation for Interfering Wireless Links Based on Channel Information Partitioning", in IEEE Transactions on Wireless Communications. 2009, vol. 8, n°6, pp. 3004-3015.

23. M. Kountouris, D. Gesbert, T. Salzer, “Enhanced multiuser random beamforming: Dealing with the not-so-large number of users case”, IEEE Journal on Selected Areas in Communications, Special Issue on Exploiting Limited Feedback in Tomorrow’s Wireless Communications Networks, Oct 2008.

24. N. Fawaz, D. Gesbert, M. Debbah, “When network coding and dirty paper coding meet in a cooperative adhoc network”, in IEEE Transaction on Wireless Communications, May 2008. Volume: 7, Issue: 5, Pages: 1862-1867. Longer version available on “arxiv” data base.

25. M. Kountouris, T. Salzer, D. Gesbert, ’’Scheduling for multiuser MIMO downlink channels with ranking-based feedback”,in EURASIP Journal on Wireless Communications and Networking (Special Issue on MIMO systems with limited feedback), Vol. 2008. Article ID 854120.

26. H. Skjevling, D. Gesbert, A. Hjorungnes ’’Low complexity distributed multi-base transmission and scheduling’’, in EURASIP Journal on Advances in Signal Processing (Special Issue on “Distributed space time systems”), Volume 2008 (2008), Article ID 741593.

27. A. Gjendemsjø, H. Holm, G. E. Øien, M.-S. Alouini, D. Gesbert, K. J. Hole, P. Orten “Rate and Power Allocation for Discrete-Rate Link Adaptation,” Eurasip Journal on Wireless Communications and Networking, Volume 2008 (2008), Article ID 394124.

28. Gjendemsjoe, D. Gesbert, G. Oien, S. Kiani, "Binary power control for sum rate maximization over multiple interfering links", IEEE Transactions on Wireless Communications, August 2008.

29. A. Roumy, D. Gesbert, "Optimal matching in wireless sensor networks", in IEEE Journal on Selected Topics in Signal Processing, Special issue on Convex Optimization Methods in Signal Processing, Dec. 2007.

30. H. Bang, T. Ekman, D. Gesbert, "Channel Predictive Proportional Fair Scheduling", IEEE Transactions on Wireless Communications, Feb. 2008.

31. V. Hassel, G. Oien, D. Gesbert, "Throughput Guarantees for Wireless Networks with Opportunistic Scheduling: A Comparative Study". IEEE Transactions on Wireless Communications, December 2007.

32. A. Hjoerungnes, D. Gesbert, " Precoded Orthogonal Space-Time Block Codes over Correlated Ricean MIMO Channels " IEEE Transactions on Signal Processing, Jan. 2007.

33. A. Hjoerungnes, D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results", accepted in IEEE Transactions on Signal Processing. Full (extended) version also available on-line at the following: "Introduction to Complex-valued matrix differentiation" (internal report).

34. V. Hassel, M.-S. Alouini, G. Oeien, D. Gesbert "Rate Optimal Multiuser Scheduling with Reduced Feedback Load and Analysis of Delay Effects", in *EURASIP Journal on Wireless Communications and Networking – Special Issue on Radio Resource Management in 3G+ Systems*, vol. 2006, Article ID 36424, 7 pages, 2006..

35. S. Kiani, D. Gesbert “Optimal and distributed scheduling for multicell capacity maximization ", IEEE Transaction of Wireless Communications, Jan. 2008.

36. H. Skjevling, D. Gesbert, A. Hjorungnes ” Precoded Distributed Space-time block Codes in cooperative diversity-based downlink”, in IEEE Trans. on Wireless Communications, Dec. 2007.

37. A. Hjorungnes, D. Gesbert " Precoding of orthogonal space-time block codes in arbitrarily correlated MIMO channels: Iterative and closed-form solutions , " IEEE Transactions on Wireless Communications March 2007.

38. V. Hassel, D. Gesbert, M. Slim-Alouini, G. Oien, " A threshold based channel state feedback algorithm for modern cellular systems ", IEEE Trans. on Wireless Communications, July 2007.

39. M. Kobayashi, G. Caire, D. Gesbert, " Transmit diversity vs. Opportunistic beamforming in packet data downlink transmission, " IEEE Trans. on Wireless Communications, Jan. 2007.

40. A. Hjorungnes, D. Gesbert, J. Akhtar " Precoding of space-time block coded signals for joint transmit receive correlated MIMO channels ," IEEE Transactions on Wireless Communications, Feb. 2006

41. T. Dahl, S. Silva, N.Christophersen, D. Gesbert, "Intrinsic Subspace Convergence in TDD MIMO Communications", in IEEE Trans. Signal Processing, June 2007.

42. J. Akhtar, D. Gesbert, "Spatial multiplexing over correlated MIMO channels with a closed form precoder ," in *IEEE Transactions on Wireless Communications, 2004.*

43. J. Akhtar, D. Gesbert, " Extending Orthogonal Block Codes With Partial Feedback ," *IEEE Transactions on Wireless Communications,* Nov. 2004.

44. D. Gesbert, "Robust Linear MIMO Receivers: A minimum Error-rate Approach ", *IEEE Transactions on Signal Processing. S* pecial issue on MIMO systems, Nov. 2003.

45. T. Dahl, N. Christophersen, D. Gesbert, "Blind MIMO eigenmode transmission based on the algebraic power method" , *in IEEE Trans. on Signal Processing, Sept. 2004.* ***Recipient of the BEST PAPER AWARD by SP SOCIETY (YOUNG AUTHOR)*** .

46. D. Gesbert, J. Akhtar, " Breaking the barriers of Shannon's capacity: An overview of MIMO wireless system ", Telektronikk Telenor Journal. Jan. 2002.

47. S. Catreux, V. Erceg, D. Gesbert, R. Heath, " Adaptive modulation and MIMO coding for broadband wireless data networks ", *IEEE Communications Magazine*. June 2002.

48. R. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, A. Paulraj, " Performance of multi-antenna signal strategies in the presence of polarization diversity ", *IEEE Trans. on Signal Processing* Oct. 2002.

49. D. Gesbert, L. Haumonte, H. Bolcskei, R. Krishnamoorthy, A.Paulraj, `` Performance of Non-Line-of-Sight Broadband Wireless Access Networks ", *IEEE Communications Magazine, April 2002.*

50. D. Gesbert, H. Bolcskei, D. Gore, A. Paulraj, `` Outdoor MIMO wireless channels: Models and performance prediction ", *IEEE Trans. on Communications.* Dec. 2002.

51. D. Gesbert and P. Duhamel,`` Unbiased blind adaptive channel identification and equalization ", *IEEE Trans. on Signal Processing* , January 2000.

52. K. Sheikh, D. Gesbert, D. Gore, A. Paulraj,``Smart Antennas for Broadband Wireless Access Networks, IEEE Communications Magazine, Nov.1999.

53. H. Bolcskei, D. Gesbert, A. Paulraj,`` On the capacity of wireless systems employing OFDM-based spatial multiplexing ", IEEE Trans.Communications, February 2002.

54. D. Gesbert, J. Sorelius, P. Stoica, A. Paulraj,``Blind multi-user MMSE detectior for CDMA signals in ISI channels", IEEE Communications Letters, August 1999.

55. D. Gesbert, P. Duhamel and S. Mayrargue,``On-line blind multichannel equalization based on mutually referenced filters', IEEE Trans. on Signal Processing, September 1997,vol. 45, pages 2307-2317.

56. D. Gesbert and A. J. van der Veen,``On the equivalence of blind equalizers based on MRE and subspace intersections", IEEE Trans. Signal Processing, March 1999, Vol. 47.

### Conference papers

1. P. de Kerret, M. Guilaud, D. Gesbert, "Degrees of Freedom of Certain Interference Alignment Schemes with Distributed CSI" submitted to IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2013.

2. X. Yi, D. Gesbert, S. Yang, M. Kobayashi, "Degrees of Freedom of Time-Correlated Broadcast Channels with Delayed CSIT: The MIMO Case", submitted to IEEE International Symposium on Information Theory, 2013.

3. P. de Kerret, X. Yi, D. Gesbert, "On the Degrees of Freedom of the K-User Time-Correlated Broadcast Channel with Delayed CSIT", submitted to IEEE International Symposium on Information Theory, 2013.

4. R. Gangula, D. Gesbert, J. Lindblom, E.G. Larsson "On the Value of Spectrum Sharing among Operators in Multicell Networks", in Proc. IEEE VTC-Spring 2013, Dresden, Germany.

5. H. Yin, D. Gesbert, M. Filippou "Decontaminating Pilots in Massive MIMO Systems", in proc. of IEEE International Conference on Communications 2013.

6. M.C. Filippou, D. Gesbert, G.A. Ropokis "Optimal Combining of Instantaneous and Statistical CSI in the SIMO Interference Channel" in Proc. of Vehicular Technology Conference (VTC-Spring), Dresden, Germany, 2013

7. X. Yi, P. de Kerret, D. Gesbert "The DoF of Network MIMO with Backhaul Delays", in proc. of IEEE International Conference on Communications 2013.

8. P. de Kerret, M. Guillaud, D. Gesbert "Degrees of Freedom of Interference Alignment with Distributed CSI", submitted to IEEE Allerton Conference on Communications, Control, and Computing.

9. S. Yang, M. Kobayashi, D. Gesbert, X. Yi "Degrees of Freedom of MISO Broadcast Channel with Perfect Delayed and Imperfect Current CSIT", in Proc. of the IEEE Information Theory Workshop, 2012.

10. P. de Kerret, "MIMO Interference Alignment Algorithms With Hierarchical CSIT", in Proc. of the IEEE International Symposium on Wireless Communications Systems" Paris (ISWCS), 2012, Invited paper.

11. M. Filippou, D. Gesbert, H. Yin, "Decontaminating Pilots in Cognitive Massive MIMO Systems", in Proc. Of the IEEE International Symposium on Wireless Communications Systems (ISWCS), Paris, 2012 (Invited Paper).

12. X. Yi, D. Gesbert, S. Yang, M. Kobayashi, "On the DoF of the multiple-antenna time correlated interference channel with delayed CSIT" to appear in Proc. Asilomar Conference on Signals and Systems (Invited Paper), 2012. (available upon request).

13. M. Kobayashi, S. Yang, D. Gesbert, X. Yi, "On the Degrees of Freedom of time correlated MISO broadcast channel with delayed CSIT" in Proc. IEEE Intern. Symposium on Information Theory, 2012. (ArXiv version available).

14. X. Yi, D. Gesbert, "Precoding on the Broadcast MIMO Channel with Delayed CSIT: The Finite SNR Case", in Proceedings of IEEE ICASSP 2012 Conference, Kyoto Japan.

15. P. de Kerret, D. Gesbert "CSI Feedback Allocation in Multicell MIMO Channels", in Proceedings of IEEE International Conference on Communicationce, Ottawa, Canada, 2012.

16. P. de Kerret, D. Gesbert, "Sparse Precoding in Multicell MIMO Systems", in Proceedings of IEEE Wireless Communication and Networking Conference (WCNC), 2012.

17. E. Yılmaz, D. Gesbert and R. Knopp , "Interference Relay Channel in 4G Wireless Networks" , in Proceedings of IEEE Wireless Communication and Networking Conference (WCNC), 2012.

18. P. de Kerret, R. Gangula, D. Gesbert, "A Practical Precoding Scheme for Multicell MIMO Channels with Partial User's Data Sharing" , 4G Wireless Workshop at the WCNC2012 Conference, Paris, 2012.

19. E. Kapiridis, D. Gesbert, M. Haardt, K.M. Ho, E. Jorswieck, E. Larsson, J. Li, J. Lindblom, C. Scheunert, M. Schubert, N. Vucic, "Transmit Beamforming for Inter Operator Spectrum Sharing', In proceedings of FUNEMS 2011 (Future Networks and Mobile Summit).

20. R. Gangula, P. de Kerret, D. Gesbert, M. Al-Odeh "Optimized Data Symbol Sharing in Multiple Antenna Interference Channels », Invited paper, Asilomar Conference on Signals and Systems, 2011.

21. M. Guillaud, D. Gesbert, "Interference Alignement in Partially Connected Interfering Multiple-Access and Broadcast Channels", in proceedings of IEEE Global Communications Conference (GLOBECOM), 2011.

22. P. de Kerret, D. Gesbert, "The Multiplexing Gain of a Two-cell MIMO Channel with Unequal CSI" IEEE International Symposium on Information Theory (ISIT), 2011, St. Petersburg, Russia.

23. P. de Kerret, D. Gesbert, “The Asymptotic Limits of Interference in Multicell Networks with Channel Aware Scheduling”, IEEE SPAWC 2011, San Francisco, USA.

24. M. Guillaud, D. Gesbert, “Interference Alignement in the partially connected K-user MIMO interference channel”, invited paper, 19th European Signal Processing Conference, Barcelona, Spain, 2011.

25. Arun Singh, Kiran Gowda David Gesbert and Petros Elia, “Diversity-Multiplexing Tradeoff for the non-separated two-way DF Channel.” in the IEEE SPAWC 2011, San Francisco***. Best Student Paper Award***

26. E. Yilmaz, R. Knopp, D. Gesbert, “Error Exponents for Multi-Source Multi-Relay Parallel Relay Networks with Limited Backhaul Capacity”, in Proceedings of the IEEE International Conference on Communications (ICC), 2011, Kyoto, Japan.

27. K. M. (Zuleita) Ho, D. Gesbert, E. Jorswieck, R. Mochaourab, “Beamforming on the MISO interference channel with multi-user decoding capability”, (Invited paper) Proc. of the Asilomar Conference, Ca., Nov. 2010.

28. E. Yilmaz, R. Knopp, F. Kaltenberger, D. Gesbert, “Low-complexity Multiple-relay Strategies for Improving Uplink Coverage in 4G Wireless Networks” in Proceedings of the Asilomar Conference, Ca., Nov. 2010.

29. R. Mochaourab, E. Jorswieck, K. M. (Zuleita) Ho, D. Gesbert, “Bargaining and beamforming in interference channels”, (Invited paper) Proc. of the Asilomar Conference, Ca., Nov. 2010.

30. E. Yilmaz, R. Knopp, D. Gesbert, “Error exponents for bakhaul constrained parallel relay networks”, in proceedings of IEEE PIMRC 2010, Istanbul, Turkey.

31. M. Kaynia, A. Goldsmith, D. Gesbert, G. Oien “On the usage of antennas in MIMO and MISO interference channels” In proceedings of the IEEE SPAWC Workshop, Marrakhesh, Marocco, 2010.

32. R. Zakhour, D. Gesbert , “Team decision for the cooperative MIMO channel with imperfect CSIT sharing”, Invited Paper, The Information Theory and Applications (ITA) Workshop, San Diego CA., February 2010.

33. R. Zakhour, D. Gesbert “On the value of data sharing in constrained-backhaul network MIMO”, In proc. of the International Zurich Seminar on Communications, Zurich, March 2010.

34. K.-M. (Zuleita) Ho, D. Gesbert, “Balancing Egoism and Altruism on the Interference Channel: The MIMO case”, In proc. of the IEEE ICC Conference, Cape Town, South Africa, 2010.

35. E. Yilmaz, R. Zakhour, D. Gesbert, R. Knopp, “Multi-pair Two-way Relay Channel with Multiple Antenna Relay Station”, In proc. of the IEEE ICC Conference, Cape Town, South Africa, 2010.

36. K.-M. (Zuleita) Ho, M. Kaynia, D. Gesbert, “Distributed power control and beamforming on MIMO interference channels”, In proc. of the European Wireless Conference, Lucca, Italy, 2010.

37. U. Salim, D. Gesbert, D. Slock, Z. Beyastas, “Hybrid pilot/quantization-based feedback in multi-antenna TDD systems”, Proc. of IEEE Globecom Conference, Hawai, 2009.

38. E. Bjornson, R. Zakhour, D. Gesbert, B. Ottersten, “Distributed multicell and multiantenna precoding: characterization and performance evaluation”, Proc. of IEEE Globecom Conference, Hawai, 2009.

39. R. Zakhour, Z. Ho, D. Gesbert, “Distributed Beamforming Coordination in Multicellular MIMO Systems”, in Proc. of IEEE VTC 2009, Barcelona, Spain.

40. V. Corvino, D. Gesbert, R. Verdone, “A novel distributed interference mitigation technique using power planning”, in proceeding of the IEEE WCNC conference 2009.

41. Erhan YILMAZ, Raymond Knopp and David Gesbert, “Relaying with Wireless Infrastructure Links in Cellular Networks”, Proc. of the 2009 IEEE Winter School on Information Theory.

42. R. Zakhour, D. Gesbert “Coordination on the MISO Interference Channel using the Virtual SINR Framework” submitted to the International ITG/IEEE Workshop on Smart Antennas (WSA’09), Berlin, Germany, 2009.

43. M. Kaynia, G. Oien, N. Jindal, D. Gesbert, “Comparative Performance Evaluation of MAC Protocols in Ad Hoc Networks with Bandwidth Partitioning”, in Proc. of IEEE International Symposium on Personal, Indoor, Mobile Radio Communications (PIMRC), Cannes, 2008.

44. A. Papadogiannis, E. Hardouin, and D. Gesbert, "A framework for decentralising multi-cell cooperative processing on the downlink," in Proc. IEEE Global Communications Conference (GLOBECOM 2008), November 30-December 4, 2008, New Orleans, USA.

45. A. Papadogiannis, D. Gesbert, “Downlink overhead reduction for multicell cooperative processing enabled wireless networks”, in Proc. of IEEE International Symposium on Personal, Indoor, Mobile Radio Communications (PIMRC), 2008.

46. Z. Ho, D. Gesbert, “Spectrum Sharing in Multiple Antenna Channels: A distributed Cooperative Game Theoretic Approach”, in Proc. IEEE International Symposium on Personal, Indoor, Mobile Radio Communications (PIMRC), 2008.

47. A. Papadogiannis, E. Hardouin, A. Saadani, D. Gesbert, and P. Layec, "A Novel Framework for the Utilisation of Dynamic Relays in Cellular Networks", in Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers (Asilomar 2008), Pacific Grove, USA, October 2008.

48. F. Kaltenberger, M. Kountouris, D. Gesbert, R. Knopp, "Performance of Multi-user MIMO Precoding with Limited Feedback over Measured Channels," in IEEE Global Communications Conference (GLOBECOM 2008), New Orleans, USA, 30.11.-4.12.2008.

49. F. Kaltenberger, M. Kountouris, R. Knopp, D. Gesbert, "Correlation and capacity of measured multi-user MIMO channels," in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2008), Cannes, France, 15.-18.9.2008.

50. M. Kountouris, D. Gesbert, T. Salzer, "Distributed Transmit Mode Selection for MISO Broadcast Channels with Limited Feedback: Switching from SDMA to TDMA" in Proc. IEEE

Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Recife, Brazil, 2008.

51. F. Kaltenberger, M. Kountouris, L. Cardoso, R. Knopp, D. Gesbert, “Capacity of Linear Multiuser Precoding Schemes with Measured Channel Data”, in Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Recife, Brazil 2008.

52. F. Kaltenberger, L. Cardoso, M. Kountouris, R. Knopp, D. Gesbert, "Real-time Multi-user MIMO Channel Sounding and Capacity Evaluations," in COST 2100 4th Management Comitee Meeting, Wroclav, Poland, 6.-8.2.2008.

53. N. Fawaz, K. Zarifi, M. Debbah, D. Gesbert, “Asymptotic Capacity and Optimal Precoding Strategy of Multi-Level Precode and Forward in Correlated Channels”, in Proc. IEEE Information Theory Workshop, ITW 2008, May 2008.

54. R. Zakhour, D. Gesbert, “Adaptive Feedback Rate Control in MIMO Broadcast Systems” in Proc. of the IEEE Information Theory Workshop 2008.

55. E. Yilmaz, D. Gesbert, R. Knopp, ’’Parallel Relay Networks with Phase Fading“, Proceedings IEEE Globecom, 2008.

56. E. Yilmaz, D. Gesbert, R. Knopp, ’’Some system aspects for compressive relaying“, Proceedings IEEE Globecom, 2008.

57. R. Zakhour, D. Gesbert, “Adaptive Feedback Rate Control in MIMO Broadcast Systems with User Scheduling” Proceeding of the 2008 Information Theory and Applications (ITA) Workshop, San Diego, Ca. 2008.

58. Z. Beyaztas, A. Pandharipande, D. Gesbert, ‘’Optimal Power Allocation in a hierarchical spectrum sharing system’’, Proc. of the IEEE ICC 2008 (Workshop on Cognitive Radio), Beijing, China, May 2008.

59. A. Papadogiannis, D. Gesbert, E. Hardoin, “A Dynamic Clustering Approach in Wireless Networks with Multi-Cell Cooperative Processing”, In the proc. of the International Conference on Communications, Beijing, 2008.

60. N. Fawaz, Z. Beyaztas, D. Gesbert, M. Debbah, ’’Improving Adhoc network capacity and connectivity using dynamic blind beamforming“, Proc. of VTC Conference Spring 2008

61. S. Kiani, D. Gesbert, “Capacity maximizing power allocation for interfering wireless links: A distributed approach”, in Proc. of IEEE Globecom Conference, Washington DC, 2007.

62. R. Zakhour, D. Gesbert, “A two-stage approach to feedback design in MU-MIMO channels with limited channel state information”, In Proceedings of the IEEE PIMRC Conference 2007 (Invited).

63. R. de Lacerda, L. Sampaio, R. Knopp, M. Debbah, D. Gesbert, “EMOS Platform: Real-time capacity estimation of MIMO channel in the UMTS-TDD band”, to appear in the Proceedings of the IEEE IWCS Conference 2007.

64. N. Fawaz, M. Debbah, D. Gesbert "Capacity and positioning in dense scattering environments", Proceedings of IEEE SPAWC Conference, Helsinki, Finland, 2007.

65. H. Skjevling, D. Gesbert, A. Hjoerungnes, "A low complexity distributed multibase transmission scheme for improving the sum capacity of wireless networks", IEEE SPAWC Conference, Helsinki, Finland, 2007 (Invited paper).

66. A. Gjendemsjoe, G. Oien, D. Gesbert, "Binary power control for multicell capacity maximization", Proceedings of IEEE SPAWC Conference, Helsinki, Finland, 2007.

67. A. Roumy, D. Gesbert, "Optimal matching in wireless sensor networks ", Proceedings of IEEE International Symposium on Information Theory, 2007.

68. K. Tourki, D. Gesbert, L. Deneire, "Cooperative diversity using per-user power control in the multiuser MAC channel ", Proceedings of IEEE International Symposium on Information Theory, 2007.

69. M. Kountouris, R. de Francisco, D. Gesbert, D. Slock, T. Salzer, "Efficient Metric for Scheduling in MIMO Broadcast Channels with Limited Feedback" In Proceedings of IEEE ICASSP 2007, Hawaii, USA.

70. D. Gesbert, M. Kountouris, "Resource allocation in multicell wireless networks: Some capacity scaling laws", submitted to the IEEE WiOpt (Workshop on Resource Allocation in Wireless Networks RAWNET), 2007.

71. R. de Francisco, M. Kountouris, D. Slock, D. Gesbert, "Orthogonal Linear Beamforming in MIMO Broadcast Channels ", in Proc. IEEE Wireless Communications and Networking Conference, 2007.

72. S. Kiani, G. Oien, D. Gesbert, "Maximizing Multicell Capacity Using Distributed Power Allocation and Scheduling ", in Proc. IEEE Wireless Communications and Networking Conference (WCNC), 2007.

73. M. Kountouris, R. de Francisco, D. Gesbert, D. Slock, T. Salzer, "Multiuser Diversity: Multiplexing trade-off in MIMO broadcast channels with limited feedback ", in Proc. 40th Asilomar Conference on Signals Systems and Computers, 2006.

74. A. Hjorungnes, D. Gesbert, "Hessian of scalar function of complex-valued matrices: A systematic computational approach ", in Proc. IEEE International Symposium on Signal Processing and Its Applications (ISSPA'07), Sharjah, United Arab Emirates, IEEE, February 2007.

75. S. G. Kiani, D. Gesbert, J. E. Kirkebø, A. Gjendemsjø, and G. E. Øien, "A Simple Greedy Scheme for Multicell Capacity Maximization ", in Proc. IEEE International Telecommunications Symposium (ITS '06) , Fortaleza, Brazil, September 2006.

76. J. E. Kirkebø, D. Gesbert, S. G. Kiani, "Probabilistic Access Functions for Multicell Wireless Schemes", in Proc. IEEE International Telecommunications Symposium (ITS '06) , Fortaleza, Brazil, September 2006.

77. V. Hassel, G. Oien, D. Gesbert, "Throughput Guarantees for Wireless Networks with Opportunistic Scheduling ", Proceedings of IEEE Globecom Conference, San Francisco, 2006.

78. J.-E. Kirkeboe, D. Gesbert, S. Kiani "Maximizing the Capacity of Wireless Networks using Multi-Cell Access Schemes ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, Cannes, France, 2006.

79. M. Kountouris, R. De Francisco, D. Gesbert, D. Slock "Low Complexity Scheduling and Beamforming for Multiuser MIMO Systems ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, Cannes, France, 2006.

80. V. Hassel, G. Oien, D. Gesbert, "Throughput Guarantees for Opportunistics Scheduling Algorithms: A Comparative Study ", Proceedings of IEEE International Telecommunications Symposium, Fortaleza, Brazil, 2006.

81. M. Kountouris, D. Gesbert, Lars Pittman, "Transmit Correlation-Aided Opportunistic Beamforming and Scheduling Using reciprocal channel information in multiuser MIMO scheduling ", Proc. EUSIPCO Conference, Florence, Italy, 2006 (Invited paper).

82. D. Gesbert, A. Hjorungnes, H. Skjevling, "Cooperative spatial multiplexing with hybrid channel knowledge ", Proc. International Zurich Seminar on Broadband Communications, Feb. 2006.

83. S. Kiani, D. Gesbert, "Maximizing the capacity of large networks: Optimal and distributed solutions ", In Proceedings of IEEE International Symposium on Information Theory, 2006.

84. H. Skjevling, D. Gesbert, A. Hjorungnes, "Receiver-Enhanced Cooperative Spatial Multiplexing with Hybrid Channel Knowledge ", Proc. International Conf. on Acoustics Speech and Signal Processing,Toulouse, France, 2006.

85. D. Gesbert, L. Pittman, M. Kountouris "Transmit Correlation-aided Scheduling in Multiuser MIMO Networks ", Proc. International Conf. on Acoustics Speech and Signal Processing,Toulouse, France, 2006.

86. A. Gjendemsjø, D. Gesbert, G. Oien, S. Kiani, " Optimal Power Allocation and Scheduling for Two-Cell Capacity Maximization ", IEEE WiOpt (Workshop on Resource Allocation in Wireless Networks), 2006.

87. H. Skjevling, D. Gesbert, A. Hjorungnes " Precoding for distributed space-time codes in cooperative diversity-based downlink ", Proc. IEEE Intern. Conf. on Communications, 2006.

88. M. Kountouris, D. Gesbert, " Robust Multi-User Opportunistic Beamforming for Sparse Networks ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC) 2005, New York.

89. H. Bang, T. EKman, D. Gesbert, " A channel predictive proportional fair scheduling algorithm ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC) 2005, New York.

90. V. Hassel, M.-S. Alouini, G. E. Øien, and D. Gesbert, "Rate-optimal multiuser scheduling with reduced feedback load and analysis of delay effects ," in Proc. EUSIPCO-2005, Antalya, Turkey, September 2005.

91. M. Kobayashi, G. Caire, D. Gesbert " Opportunistic vs. Space time coding in a queued downlink ", Proceedings of IST Summit, 2005, Dresden, Germany.

92. A. Forenza, M. Airy, M. Kountouris, R. Heath, D. Gesbert, S. Shakkottai, " Performance of the MIMO Downlink Channel with Multi-Mode Adaptation and Scheduling ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC) 2005, New York.

93. M. Koutouris, D. Gesbert, " Memory based opportunistic multi-user beamforming ", In Proceedings of IEEE International Symposium on Information Theory, (ISIT) 2005, Australia.

94. M. Kobayashi, G. Caire, D. Gesbert, " Impact of multiple transmit antennas in a queued SDMA/TDMA downlink" ", Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC) 2005, New York.

95. A. Hjorungnes, D. Gesbert, "Precoding of orthogonal space-time block codes over (Non)-kronecker correlated Ricean MIMO channels ," In Proceedings of IEEE Vehicular Technology Conference (VTC), Stockholm, Sweden, 2005.

96. M. Koutouris, A. Pandharipande, H. Kim, D. Gesbert, "QoS-based user scheduling for multiuser MIMO systems", Vehicular Technology Conference (VTC) 2005, In proceedings.

97. V. Hassel, M. Slim-Alouini, D. Gesbert, G. Oien, " Exploiting multiuser diversity using multiple feedback thresholds ", Vehicular Technology Conference (VTC) 2005, In proceedings.

98. L. Yang, M. Slim-Alouini, D. Gesbert, " Further Results on Selective Multi-User Diversity ", Proceedings of the Seventh ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, October 2004. **Won the 2004 MSWiM Best Paper Award**.

99. Vegard Hassel, Mohamed-Slim Alouini, David Gesbert and Geir E. Øien,"Minimizing Feedback Load for Nested Scheduling Algorithms", Proceedings of COST 289 Workshop (Spectrum and power efficient broadband communications), 2005

100. M. Kobayashi, G. Caire, D. Gesbert, " Antenna diversity versus multi-user diversity: Quantifying the trade-offs ", Proceedings of IEEE International Symposium on Information Theory and its Applications (ISITA), October 2004.

101. J. Akhtar, D. Gesbert, A. Hjorungnes " Linear closed-form precoding of MIMO multiplexing systems in the presence of transmit correlation and Ricean channel ", In Proceedings of IEEE Globecom Conference, Dallas, TX, Nov. 2004.

102. A. Hjorungnes, D. Gesbert, " Minimum exact-SER precoding of orthogonal space-time block codes for correlated MIMO channels ," In Proceedings of IEEE Globecom Conference, Dallas, TX, Nov. 2004.

103. A. Hjorungnes, D. Gesbert, "Exact-SER precoding of orthogonal space-time block coded correlated MIMO channels: An iterative approach," Proceesings of the Nordic Signal Processing Conference (NORSIG), Finland, 2004.

104. A. Hjorungnes, J. Akhtar, D. Gesbert, "Precoding for space-time block codes in (non-)kronecker correlated MIMO channels", Invited paper, Proceedings of European Signal Processing Conference (Eusipco), Vienna, Austria, 2004.

105. D. Gesbert, " Multipath: Curse or blessing? A system performance analysis of MIMO wireless systems ", Invited paper, Proceedings of International Zurich Seminar on Communications (IZS), Zurich, Switzerland, 2004.

106. D. Gesbert, M. Slim Alouini, " How much feedback is multi-user diversity really worth? ", In Proceedings of IEEE Intern. Conf. On Communications (ICC), 2004.

107. D. Gesbert, M. Slim Alouini, " Selective Multi-user Diversity " Proc of IEEE Intern. Symposium on Signal Proc. And Info. Techn.(ISSPIT), Darmstadt, Ger. 2003.

108. J. Akhtar, D. Gesbert," A Closed-Form Precoder for Spatial Multiplexing over Correlated MIMO Channels ", Proc. of IEEE Global Communications Conference, 2003.

109. H. Skjevling, D. Gesbert, N. Christophersen, " Combining Space Time Block Codes and Multiplexing in Correlated MIMO Channels: An antenna assignement strategy " Proc. of Nordic Signal Processing Conference (NORSIG) 2003.

110. D. Gesbert," Minimum-error Linear Receivers for Ill-conditioned MIMO Channels ", In Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications, Roma, Italy, 2003.

111. H. Holm, G. Øien, M. Slim-Alouini, D. Gesbert, K. Hole," Optimal design of adaptive coded modulation schemes for maximum average spectral efficiency ". In Proceedings of IEEE Workshop on Signal Prcessing Advances in Wireless Communications, Roma, Italy, 2003.

112. J. Akhtar, D. Gesbert, “Partial Feedback Based Space Time Block Coding ,’’, Proceedings IEEE Vehicular Technology Conference, 2003.

113. D. Gesbert, T. Ekman, N. Christophersen, "Capacity limits of dense palm-sized MIMO arrays' ', in Proceedings of Globecom Conference 2002, Taipei, Taiwan.

114. T. Dahl, N. Christophersen, D. Gesbert, "BIMA Blind MIMOAlgorithm", In proceedings of IEEE International Conference on Signal Processing (ICASSP), 2002.

115. T. Dahl, N. Christophersen, D. Gesbert, "Identification of parallel MIMO channels in wireless communications using power iterations", In proceedings of IEEE Nordic Signal Processing Conference (NORSIG), 2002.

116. D. Gesbert, L. Haumonte, R. Krishnamoorthy, A. Paulraj,``Performance of Next-Generation Fixed Wireless Access Networks", Proc. of IEEE Radio and Wireless Conference , Boston, August 2001.

117. I. Stojanovic, M. Airy, D. Gesbert, H. Saran, "Performance of TCP/IP Over Next Generation Broadband Wireless Access Networks". In Proc. IEEE Symposium on Wireless Personal Multimedia Communications , Aalborg, DK, Sept. 2001.

118. H. Bolcskei, R. Nabar, V. Erceg, D. Gesbert, A. Paulraj, ``Performance of Spatial Multiplexing in the presence of polarization diversity", Proc.of IEEE International Acoustics Speech and Signal Processing (ICASSP) Conference , Salt Lake City, May 2001.

119. D. Gesbert, H. Bolcskei, D. Gore, A. Paulraj, `` MIMO wireless channels: Capacity and performance prediction ", Proceedings of IEEE Globecom Conference , Nov. 2000.

120. D. Gesbert, H. Bolcskei, D. Gore, A. Paulraj, "Performance evaluation for scattering MIMO channel models", Proc. 34th Asilomar conference on signals systems and computers , Nov 2000.

121. H. Bolcskei, D. Gesbert, A. Paulraj, ``On the capacity of OFDM-based multi-antenna systems", Proc.IEEE ICASSP-2000, Istanbul, Turkey,June 2000, submitted.

122. Geert Leus, D. Gesbert,``Recursive blind source separation for BPSK signals", Proc. of 2nd IEEE Workshop on Signal Processing Advances on Wireless Communications (SPAWC), May 1999.

123. Joakim Sorelius, D. Gesbert,``Blind chip and symbol rate CDMA receivers", Proc. of 2nd IEEE Workshop on Signal Processing Advances on Wireless Communications (SPAWC), May 1999.

124. C. Papadias, D. Gesbert, A. Paulraj, "Direct second order blind equalization of polyphase channels based on a decorrelation criterion" IEEE Intern. Conference on Acoustics, Speech and Signal Processing, May 1999.

125. D. Gesbert, J. Sorelius and A. Paulraj,``Blind CDMA receivers using mixed-rate constraints", Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers , 1998.

126. D. Gesbert, J. Sorelius and A. Paulraj,``Blind multi-user MMSE detection of CDMA signals", Proc. of ICASSP, 1998.

127. D. Gesbert and A. Paulraj,``Blind multi-user linear detection of CDMA signals in frequency selective channels", Proc. of the International Communications Conference, 1998.

128. B. Ng, D. Gesbert and A. Paulraj, ``A Semi-blind approach to structured channel equalization", Proc. ICASSP, 1998.

129. A. J. van der Veen and D. Gesbert,``On the equivalence ofblind equalizers based on MRE and subspace intersections", IEEE DigitalSignal Processing Workshop, , 1998.

130. D. Gesbert, J. Sorelius, P. Stoica, and A. Paulraj,``Multi-user MMSE detection of CDMA signals: Blind and semi-blind algorithms", IEEE Digital Signal Processing Workshop, 1998.

131. C. Papadias, D. Gesbert, and A. Paulraj,``Robust Second-Order Blind Equalization of Polyphase Channels", Proc. European Signal Processing Conference , 1998.

132. M. Ibrahim, A. Belouchrani, D. Gesbert, K. Abed-Meraim, "Parametric estimation and suppression of non-stationary interference in DS-spread spectrum communication", 32nd Asilomar Conference on Signal Systems and Computers, 1998.

133. D. Gesbert and P. Duhamel,``Robust Blind Channel Identification and Equalization based on multistep predictors", Proc. ICASSP, 1997.

134. D. Gesbert and P. Duhamel,``Unimodal blind adaptive channel equalization: An RLS implementation of the Mutually Referenced Equalizers", Proc. First IEEE Workshop on Signal Processing Advances for Wireless Communications, 1997.

135. D. Gesbert, A. Paulraj and P. Duhamel,``Blind Joint Multiuser Detection Using Second-Order Statistics and Structure Information", Proc. 40th Midwest Symposium on Circuits and Systems, 1997.

136. D. Gesbert, C. Papadias and A. Paulraj,``Blind Equalization of Polyphase FIR Channels: A Whitening Approach", Proc. 31th Asilomar Conf. on Signals, Systems, and Computers, 1997.

137. D. Gesbert and P. Duhamel,``Robust blind joint data/channel estimation based on bilinear optimization", Proc. 8th IEEE Signal Processing Workshop on Statistical Signal Array Processing, 1996.

138. D. Gesbert, P. Duhamel and S. Mayrargue,``Blind multichannel adaptive MMSE equalization with controlled delay", Proc. 8th IEEE Signal Processing Workshop on Statistical Signal ArrayProcessing, 1996.

139. D. Gesbert, P. Duhamel and S. Mayrargue,``Blind least-squares criteria for joint data/channel estimation", IEEE Digital Signal Processing Workshop, 1996.

140. K. Abed Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S.Mayrargue, E. Moulines and D. Slock,``Prediction error methods for time-domain blind identification of multichannel FIR filters", Proc. ICASSP ,1995.

141. D. Gesbert, P. Duhamel and M. Unser,``Regression lineaire en presence de bruit: Une solution adaptative non biaisee", Proc. GRETSI Conference , Juan-les-Pins, France, 1995.

142. D. Gesbert, P. Duhamel and S. Mayrargue,``A bias removal technique for the prediction-based blind adaptive multichannel deconvolution", Proc. 29th Asilomar Conf. on Signals, Systems, and Computers, 1995.

143. D. Gesbert, P. Duhamel and S. Mayrargue,``Subspace-based adaptive algorithms for the blind equalization of multichannel FIR filters", Proc. European Signal Processing Conference, 1994.

## Patents

Granted or filed

- Patent 1 U.S. Patent No.6,067,290 Issued: May 23, 2000, SPATIAL MULTIPLEXING IN A CELLULARNETWORK Filed: July 30, 1999
- Patent 2 Extensions SPATIAL MULTIPLEXING IN A CELLULAR NETWORK. Number: 6,067,290. Assigned May 2000.
- Patent 3 DATA ROUTING FOR SPATIAL MULTIPLEXING IN A CELLULAR NETWORK Filed: March 3, 2000

- Patent 4 SUBSCRIBER UNITINCORPORATING SPATIAL MULTIPLEXING Filed: April 7, 2000
- Patent 5 SUBSCRIBER UNITIN A HYBRID LINK INCORPORATING SPATIAL MULTIPLEXING Filed: May 3, 2000
- Patent 6 WIRELESS COMMUNICATION SYSTEMS USING MULTIPLE ANTENNAS AND ADAPTIVE CONTROL FOR MAXIMIZING A COMMUNICATION PARAMETER. Number: 6,351,499. Assigned Feb 2002.
- Patent 7 METHOD AND WIRELESS COMMUNICATION SYSTEMS USING MULTIPLE ANTENNAS AND ADAPTIVE CONTROL FOR MAXIMIZING A COMMUNICATION PARAMETER Filed: December 15, 2000
- Patent 8 WIRELESS COMMUNICATION SYSTEM USING JOINED TRANSMIT AND RECEIVE PROCESSING. Number: 6,377,819. Assigned May 2002.
- Patent 9 SPATIAL SEPARATION AND MULTI-POLARIZATION OF ANTENNAE IN A WIRELESS CELLULAR NETWORK Filed: July 21, 2000
- Patent 10 AN APPARATUS AND METHOD FOR OPTIMIZING DATA TRANSFER CAPACITY OF A MULTIPLE BASE TRANSCEIVER STATION CELLULAR WIRELESS NETWORK SYSTEM Filed: September 28, 2000
- Patent 11 MODE LOOKUP TABLES FOR DATA TRANSMISSION IN WIRELESS COMMUNICATION CHANNELS BASED ON STATISTICAL PARAMETERS Filed: December 5, 2000
- Patent 12 WIRELESS COMMUNICATION SYSTEMS WITH ADAPTIVE CHANNELIZATION AND LINK ADAPTATION, Filed 6/5/01
- Patent 13 A SYSTEM AND METHOD OF CLASSIFYING REMOTE USERS ACCORDING TO LINK QUALITY, AND SCHEDULING WIRELESS TRANSMISSION OF INFORMATION TO THE TO THE USERS BASED UPON THE CLASSIFICATIONS, Filed July 24, 2001
- Patent 14 A SYSTEM AND METHOD FOR COORDINATION OF MULTI-CELL USERS IN A WIRELESS COMMUNICATION NETWORK. Filed: in filing process.
- Patent 15 SYSTEM AND METHOD OF DYNAMICALLY OPTIMIZING A TRANSMISSION MODE OF WIRELESSLY TRANSMITTED INFORMATION- Filed April 2002.
- Patent 16 AN ITERATIVE BEAMFORMING ALGORITHM and SYSTEM – Filed Dec. 2008
- Patent 17 “A SYSTEM FOR FRACTIONAL FREQUENCY REUSE AND POWER CONTROL” – Filed Feb. 2009.
- Patent 18 “METHOD AND TRANSMITTER FOR ITERATIVELY MODIFYING BEAMFORMING VECTOR” –Filed March 2010.

# Chapter 4

# PhD supervision

## 4.1 Current PhD students

- Paul de Kerret (information and decision theoretic MIMO cooperation)
- Xinping Yi (information theory in delayed feedback networks)
- Miltiades Filippou (cognitive radio networks)
- Haifan Yin (Cooperative cellular networks)
- Rajeev Gangula (cooperation in energy harvesting wireless systems)
- Sandeep Kottath (limited feedback schemes in 4G)
- Qianrui Li (Distributed signal processing in networks)

## 4.2 Alumni PhD students

- Zuleita (Ka-Ming) Ho (postdoc, TU Dresden, Germany)
- Nadia Fawaz (postdoc MIT, USA, Technicolor USA) (won Best PhD prize under French Defense Labs) -co-supervised with Merouane Debbah

- Erhan Yilmaz (Industry, Turkey)
- Saad Kiani (Telenor Research Labs, Oslo)
- Marios Kountouris (Ass. Prof., SUPELEC, France)
- Agisilaos Papadogiannis (postdoc Chalmers U. Sweden)
- Hilde Skjevling (industry Norway)
- Alberto Suarez (NEC, UK) -co-supervised with Merouane Debbah
- Randa Zakhour (postdoc in NUS Singapore, UT Austin, Australia)
- Jabran Akhtar (Defense Labs, Norway)

# Chapter 5

# Project management

Our research has been and is funded primarily through collaborative EU and French ANR competitive funding programs, as well direct contracts with industrial research labs.

## 5.1 Current projects

- ”HARP” FP7 European funding
- ”E-CROPS” CHIST-ERA European funding
- ”LICORNE” ANR French funding
- ”SAPHYRE” FP7 European funding
- ”Novel Feedback design in distributed cooperative wireless network” funded by Orange Labs
- ”Virtual MIMO technologies for robust wireless mesh deployments” funded by Mitsubishi Electric Research Europe
- ”Interference Coordination for Cloud-based Radio Access Networks” funded by Intel Labs, Sophia Antipolis

## 5.2 Past projects

- ”ARTIST4G” FP7 European funding
- ”COOPCOM” FP6 European funding
- ”ORMAC” ANR French funding
- ”MIMO for 4G Networks” funded by mobile service provider SFR
- ”Theory and design of two way relaying networks” funded by Mitsubishi Electric Research Europe
- Several CIFRE projects funded by Orange Labs, Paris.
- ”MIMO systems” funded by Telenor, Oslo, Norway.

# Chapter 6

# A selection of three publications

In this last part of the document, we reproduce in full format three publications selected for their tutorial nature. Another interesting feature of these papers is to mark well the important evolution that research in the context of multiple antenna processing has undergone in the last ten years or so. Initially it has focussed on pure signal processing and information theoretic styudies on the point to point link, then widening up to encompass the multi-user channels, finally evolving to capture the impact of such system at the network level, with an explicit taking into account of inter-link interference issues and how they can be addressed by MIMO techniques. The first two papers have each won am IEEE Best Paper Award.

- D. Gesbert, M. Shafi, D. Shiu, P. Smith, " From theory to practice: An overview of space-time coded MIMO wireless systems ", IEEE Journal on Selected Areas on Communications (JSAC). April 2003, special issue on MIMO systems. (Recipient of the 2004 IEEE Best Tutorial Paper Award by IEEE Comm. Society).

- D. Gesbert, M. Kountouris, R. Heath, C-B. Chae, T. Salzer, "From Single User to Multiuser Communications: Shifting the MIMO paradigm", IEEE Signal Processing Magazine, Sept. 2007. Recipient of the 2012 Signal Processing Society Signal Magazine Best Paper Award.

- D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, W. Yu, ”Multi-cell MIMO cooperative networks: A new look at interference”, in IEEE Journal on Selected Areas in Communications, Dec. 2010.

# From Theory to Practice: An Overview of MIMO Space–Time Coded Wireless Systems

David Gesbert, *Member, IEEE*, Mansoor Shafi, *Fellow, IEEE*, Da-shan Shiu, *Member, IEEE*, Peter J. Smith, *Member, IEEE*, and Ayman Naguib, *Senior Member, IEEE*

*Tutorial Paper*

***Abstract*—This paper presents an overview of recent progress in the area of multiple-input–multiple-output (MIMO) space–time coded wireless systems. After some background on the research leading to the discovery of the enormous potential of MIMO wireless links, we highlight the different classes of techniques and algorithms proposed which attempt to realize the various benefits of MIMO including spatial multiplexing and space–time coding schemes. These algorithms are often derived and analyzed under ideal independent fading conditions. We present the state of the art in channel modeling and measurements, leading to a better understanding of actual MIMO gains. Finally, the paper addresses current questions regarding the integration of MIMO links in practical wireless systems and standards.**

***Index Terms*—Beamforming, channel models, diversity, multiple-input–multiple-output (MIMO), Shannon capacity, smart antennas, space–time coding, spatial multiplexing, spectrum efficiency, third-generation (3G), wireless systems.**

## I. Introduction

Digital communication using multiple-input–multiple-output (MIMO), sometimes called a “volume-to-volume” wireless link, has recently emerged as one of the most significant technical breakthroughs in modern communications. The technology figures prominently on the list of recent technical advances with a chance of resolving the bottleneck of traffic capacity in future Internet-intensive wireless networks. Perhaps even more surprising is that just a few years after its invention the technology seems poised to penetrate large-scale standards-driven commercial wireless products and networks such as broadband wireless access systems, wireless local area networks (WLAN), third-generation (3G)[1] networks and beyond.

MIMO systems can be defined simply. Given an arbitrary wireless communication system, we consider a link for which the transmitting end as well as the receiving end is equipped with multiple antenna elements. Such a setup is illustrated in Fig. 1. The idea behind MIMO is that the signals on the transmit (TX) antennas at one end and the receive (RX) antennas at the other end are “combined” in such a way that the quality (bit-error rate or BER) or the data rate (bits/sec) of the communication for each MIMO user will be improved. Such an advantage can be used to increase both the network’s quality of service and the operator’s revenues significantly.

A core idea in MIMO systems is *space–time* signal processing in which time (the natural dimension of digital communication data) is complemented with the spatial dimension inherent in the use of multiple spatially distributed antennas. As such MIMO systems can be viewed as an extension of the so-called *smart antennas*, a popular technology using antenna arrays for improving wireless transmission dating back several decades.

A key feature of MIMO systems is the ability to turn multipath propagation, traditionally a pitfall of wireless transmission, into a benefit for the user. MIMO effectively takes advantage of random fading [1]–[3] and when available, multipath delay spread [4], [5], for multiplying transfer rates. The prospect of many orders of magnitude improvement in wireless communication performance at no cost of extra spectrum (only hardware and complexity are added) is largely responsible for the success of MIMO as a topic for new research. This has prompted progress in areas as diverse as channel modeling, information theory and coding, signal processing, antenna design and multiantenna-aware cellular design, fixed or mobile.

This paper discusses the recent advances, adopting successively several complementing views from theory to real-world network applications. Because of the rapidly intensifying efforts in MIMO research at the time of writing, as exemplified by the numerous papers submitted to this special issue of JSAC, a complete and accurate survey is not possible. Instead this paper forms a synthesis of the more fundamental ideas presented over the last few years in this area, although some very recent progress is also mentioned.

Manuscript received June 1, 2002; revised December 5, 2002. The work of D. Gesbert was supported in part by Telenor AS, Norway.

D. Gesbert is with the Department of Informatics, University of Oslo, Blindern, 0316 Oslo, Norway (e-mail: gesbert@ifi.uio.no).

M. Shafi is with Telecom New Zealand, Wellington, New Zealand (e-mail: Mansoor.Shafi@telecom.co.nz).

D. Shiu is with Qualcomm, Inc., Campbell, CA 95008 USA (e-mail: dashiu@qualcomm.com).

P. J. Smith is with the Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand (e-mail: p.smith@elec.canterbury.ac.nz).

A. Naguib was with Morphics Technology, Inc., Campbell, CA 95008 USA. He is now with Qualcomm, Inc., Campbell, CA 95008 USA.

Digital Object Identifier 10.1109/JSAC.2003.809458

[1] Third-generation wireless UMTS-WCDMA.

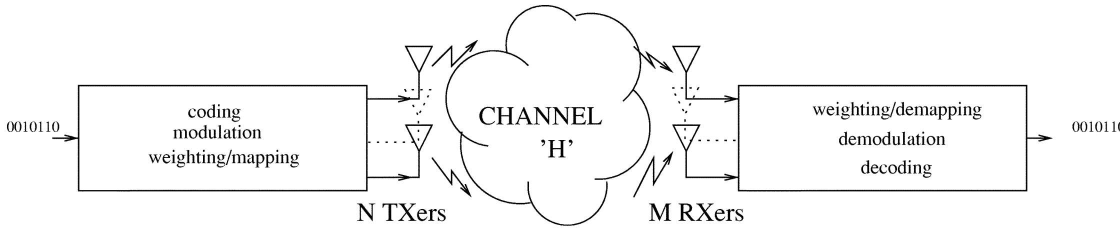



Fig. 1. Diagram of a MIMO wireless transmission system. The transmitter and receiver are equipped with multiple antenna elements. Coding, modulation, and mapping of the signals onto the antennas may be realized jointly or separately.

The article is organized as follows. In Section II, we attempt to develop some intuition in this domain of wireless research, we highlight the common points and key differences between MIMO and traditional smart antenna systems, assuming the reader is somewhat familiar with the latter. We comment on a simple example MIMO transmission technique revealing the unique nature of MIMO benefits. Next, we take an information theoretical stand point in Section III to justify the gains and explore fundamental limits of transmission with MIMO links in various scenarios. Practical design of MIMO-enabled systems involves the development of finite-complexity transmission/reception signal processing algorithms such as space–time coding and spatial multiplexing schemes. Furthermore, channel modeling is particularly critical in the case of MIMO to properly assess algorithm performance because of sensitivity with respect to correlation and rank properties. Algorithms and channel modeling are addressed in Sections IV and V, respectively. Standardization issues and radio network level considerations which affect the overall benefits of MIMO implementations are finally discussed in Section VI. Section VII concludes this paper.

## II. Principles of Space-Time (MIMO) Systems

Consider the multiantenna system diagram in Fig. 1. A compressed digital source in the form of a binary data stream is fed to a simplified transmitting block encompassing the functions of error control coding and (possibly joined with) mapping to complex modulation symbols (quaternary phase-shift keying (QPSK), M-QAM, etc.). The latter produces several separate symbol streams which range from independent to partially redundant to fully redundant. Each is then mapped onto one of the multiple TX antennas. Mapping may include linear spatial weighting of the antenna elements or linear antenna space–time *precoding*. After upward frequency conversion, filtering and amplification, the signals are launched into the wireless channel. At the receiver, the signals are captured by possibly multiple antennas and demodulation and demapping operations are performed to recover the message. The level of intelligence, complexity, and *a priori* channel knowledge used in selecting the coding and antenna mapping algorithms can vary a great deal depending on the application. This determines the class and performance of the multiantenna solution that is implemented.

In the conventional smart antenna terminology, only the transmitter or the receiver is actually equipped with more than one element, being typically the base station (BTS), where the extra cost and space have so far been perceived as more easily affordable than on a small phone handset. Traditionally, the intelligence of the multiantenna system is located in the weight selection algorithm rather than in the coding side although the development of *space–time codes (STCs)* is transforming this view.

Simple linear antenna array combining can offer a more reliable communications link in the presence of adverse propagation conditions such as multipath fading and interference. A key concept in smart antennas is that of beamforming by which one increases the average signal-to-noise ratio (SNR) through focusing energy into desired directions, in either transmit or receiver. Indeed, if one estimates the response of each antenna element to a given desired signal, and possibly to interference signal(s), one can optimally combine the elements with weights selected as a function of each element response. One can then maximize the average desired signal level or minimize the level of other components whether noise or co-channel interference.

Another powerful effect of smart antennas lies in the concept of *spatial diversity*. In the presence of random fading caused by multipath propagation, the probability of losing the signal vanishes exponentially with the number of decorrelated antenna elements being used. A key concept here is that of *diversity order* which is defined by the number of decorrelated spatial branches available at the transmitter or receiver. When combined together, leverages of smart antennas are shown to improve the coverage range versus quality tradeoff offered to the wireless user [6].

As subscriber units (SU) are gradually evolving to become sophisticated wireless Internet access devices rather than just pocket telephones, the stringent size and complexity constraints are becoming somewhat more relaxed. This makes multiple antenna elements transceivers a possibility at both sides of the link, even though pushing much of the processing and cost to the network's side (i.e., BTS) still makes engineering sense. Clearly, in a MIMO link, the benefits of conventional smart antennas are retained since the optimization of the multiantenna signals is carried out in a larger space, thus providing additional degrees of freedom. In particular, MIMO systems can provide a joint transmit-receive diversity gain, as well as an array gain upon coherent combining of the antenna elements (assuming prior channel estimation).

In fact, the advantages of MIMO are far more fundamental. The underlying mathematical nature of MIMO, where data is transmitted over a *matrix* rather than a vector channel, creates new and enormous opportunities beyond just the added diversity or array gain benefits. This was shown in [2], where the

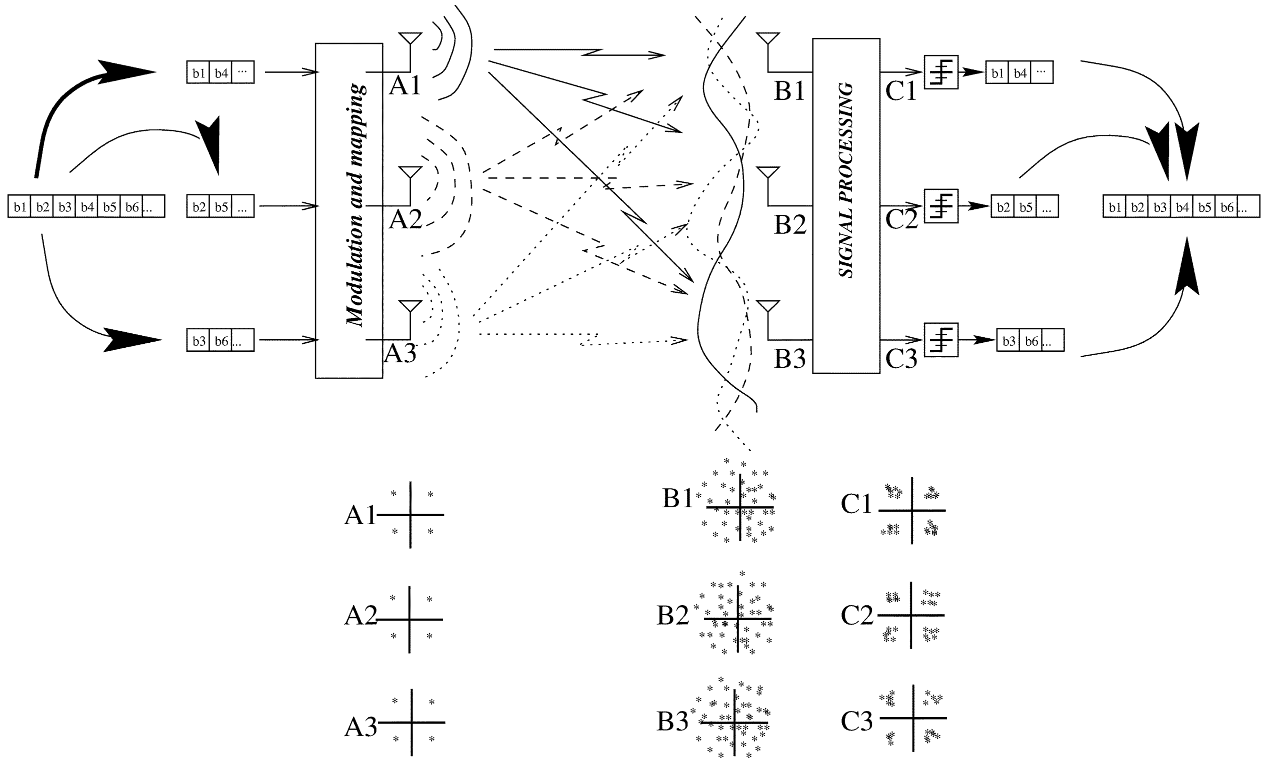



Fig. 2. Basic spatial multiplexing (SM) scheme with three TX and three RX antennas yielding three-fold improvement in spectral efficiency. Ai, Bi, and Ci represent symbol constellations for the three inputs at the various stages of transmission and reception.

author shows how one may under certain conditions transmit $\min(M, N)$ independent data streams simultaneously over the *eigenmodes* of a matrix channel created by $N$ TX and $M$ RX antennas. A little known yet earlier version of this ground breaking result was also released in [7] for application to broadcast digital TV. However, to our knowledge, the first results hinting at the capacity gains of MIMO were published by Winters in [8].

Information theory can be used to demonstrate these gains rigorously (see Section III). However, intuition is perhaps best given by a simple example of such a transmission algorithm over MIMO often referred to in the literature as V-BLAST[2] [9], [10] or more generically called here *spatial multiplexing*.

In Fig. 2, a high-rate bit stream (left) is decomposed into three independent $1/3$-rate bit sequences which are then transmitted simultaneously using multiple antennas, thus consuming one third of the nominal spectrum. The signals are launched and naturally mix together in the wireless channel as they use the same frequency spectrum. At the receiver, after having identified the mixing channel matrix through training symbols, the individual bit streams are separated and estimated. This occurs in the same way as three unknowns are resolved from a linear system of three equations. This assumes that each pair of transmit receive antennas yields a single scalar channel coefficient, hence flat fading conditions. However, extensions to frequency selective cases are indeed possible using either a straightforward multiple-carrier approach (e.g., in orthogonal frequency division multiplexing (OFDM), the detection is performed over each flat subcarrier) or in the time domain by combining the MIMO space–time detector with an equalizer (see for instance [11]–[13] among others). The separation is possible only if the equations are independent which can be interpreted by each antenna "seeing" a sufficiently different channel in which case the bit streams can be detected and merged together to yield the original high rate signal. Iterative versions of this detection algorithm can be used to enhance performance, as was proposed in [9] (see later in this paper for more details or in [14] of this special issue for a comprehensive study).

A strong analogy can be made with code-division multiple-access (CDMA) transmission in which multiple users sharing the same time/frequency channel are mixed upon transmission and recovered through their unique codes. Here, however, the advantage of MIMO is that the unique signatures of input streams ("virtual users") are provided by nature in a close-to-orthogonal manner (depending however on the fading correlation) without frequency spreading, hence at no cost of spectrum efficiency. Another advantage of MIMO is the ability to jointly code and decode the multiple streams since those are intended to the same user. However, the isomorphism between MIMO and CDMA can extend quite far into the domain of receiver algorithm design (see Section IV).

Note that, unlike in CDMA where user's signatures are quasi-orthogonal by design, the separability of the MIMO channel relies on the presence of rich multipath which is needed to make the channel spatially selective. Therefore, MIMO can be said to effectively *exploit* multipath. In contrast, some smart antenna systems (beamforming, interference rejection-based) will perform better in line-of-sight (LOS) or close to LOS conditions. This is especially true when the optimization criterion depends explicitly on angle of arrival/departure

[2]Vertical-Bell Labs Layered Space–Time Architecture.

parameters. Alternatively, diversity-oriented smart antenna techniques perform well in nonline-of-sight (NLOS), but they really try to mitigate multipath rather than exploiting it.

In general, one will define the *rank* of the MIMO channel as the number of independent equations offered by the above mentioned linear system. It is also equal to the algebraic rank of the $M \times N$ channel matrix. Clearly, the rank is always both less than the number of TX antennas and less than the number of RX antennas. In turn, following the linear algebra analogy, one expects that the number of independent signals that one may safely transmit through the MIMO system is at most equal to the rank. In the example above, the rank is assumed full (equal to three) and the system shows a *nominal* spectrum efficiency gain of three, with no coding. In an engineering sense, however, both the number of transmitted streams and the level of BER on each stream determine the link's efficiency (goodput[3] per TX antenna times number of antennas) rather than just the number of independent input streams. Since the use of coding on the multiantenna signals (a.k.a. space–time coding) has a critical effect on the BER behavior, it becomes an important component of MIMO design. How coding and multiplexing can be traded off for each other is a key issue and is discussed in more detail in Section IV.

## III. MIMO Information Theory

In Sections I and II, we stated that MIMO systems can offer substantial improvements over conventional smart antenna systems in either quality-of-service (QoS) or transfer rate in particular through the principles of spatial multiplexing and diversity. In this section, we explore the absolute gains offered by MIMO in terms of capacity bounds. We summarize these results in selected key system scenarios. We begin with fundamental results which compare single-input–single-output (SISO), single-input–multiple-output (SIMO), and MIMO capacities, then we move on to more general cases that take possible a priori channel knowledge into account. Finally, we investigate useful limiting results in terms of the number of antennas or SNR. We bring the reader's attention on the fact that we focus here on single user forms of capacity. A more general multiuser case is considered in [15]. Cellular MIMO capacity performance has been looked at elsewhere, taking into account the effects of interference from either an information theory point of view [16], [17] or a signal processing and system efficiency point of view [18], [19] to cite just a few example of contributions, and is not treated here.

### A. Fundamental Results

For a memoryless $1 \times 1$ (SISO) system the capacity is given by

$$C = \log_2(1 + \rho|h|^2) \quad \text{b/s/Hz} \tag{1}$$

where $h$ is the normalized complex gain of a fixed wireless channel or that of a particular realization of a random channel. In (1) and subsequently, $\rho$ is the SNR at any RX antenna. As we deploy more RX antennas the statistics of capacity improve and with $M$ RX antennas, we have a SIMO system with capacity given by

$$C = \log_2\left(1 + \rho \sum_{i=1}^{M} |h_i|^2\right) \quad \text{b/s/Hz} \tag{2}$$

where $h_i$ is the gain for RX antenna $i$. Note the crucial feature of (2) in that increasing the value of $M$ only results in a logarithmic increase in average capacity. Similarly, if we opt for transmit diversity, in the common case, where the transmitter does not have channel knowledge, we have a multiple-input–single-output (MISO) system with $N$ TX antennas and the capacity is given by [1]

$$C = \log_2\left(1 + \frac{\rho}{N} \sum_{i=1}^{N} |h_i|^2\right) \quad \text{b/s/Hz} \tag{3}$$

where the normalization by $N$ ensures a fixed total transmitter power and shows the absence of array gain in that case (compared to the case in (2), where the channel energy can be combined coherently). Again, note that capacity has a logarithmic relationship with $N$. Now, we consider the use of diversity at both transmitter and receiver giving rise to a MIMO system. For $N$ TX and $M$ RX antennas, we have the now famous capacity equation [1], [3], [21]

$$C_{\text{EP}} = \log_2\left[\det\left(\mathbf{I}_M + \frac{\rho}{N}\mathbf{H}\mathbf{H}^*\right)\right] \quad \text{b/s/Hz} \tag{4}$$

where $(*)$ means transpose-conjugate and $\mathbf{H}$ is the $M \times N$ channel matrix. Note that both (3) and (4) are based on $N$ equal power (EP) uncorrelated sources, hence, the subscript in (4). Foschini [1] and Telatar [3] both demonstrated that the capacity in (4) grows linearly with $m = \min(M, N)$ rather than logarithmically [as in (3)[. This result can be intuited as follows: the determinant operator yields a product of $\min(M, N)$ nonzero eigenvalues of its (channel-dependent) matrix argument, each eigenvalue characterizing the SNR over a so-called channel eigenmode. An eigenmode corresponds to the transmission using a pair of right and left singular vectors of the channel matrix as transmit antenna and receive antenna weights, respectivelly. Thanks to the properties of the $\log$, the overall capacity is the sum of capacities of each of these modes, hence the effect of capacity multiplication. Note that the linear growth predicted by the theory coincides with the transmission example of Section II. Clearly, this growth is dependent on properties of the eigenvalues. If they decayed away rapidly then linear growth would not occur. However (for simple channels), the eigenvalues have a known limiting distribution [22] and tend to be spaced out along the range of this distribution. Hence, it is unlikely that most eigenvalues are very small and the linear growth is indeed achieved.

With the capacity defined by (4) as a random variable, the issue arises as to how best to characterize it. Two simple summaries are commonly used: the mean (or ergodic) capacity [3], [21], [23] and capacity outage [1], [24]–[26]. Capacity outage measures (usually based on simulation) are often denoted $C_{0.1}$ or $C_{0.01}$, i.e., those capacity values supported 90% or 99% of the time, and indicate the system reliability. A full description of the capacity would require the probability density function

[3]The goodput can be defined as the error-free fraction of the conventional physical layer throughput.

or equivalent. Some results are available here [27] but they are limited.

Some caution is necessary in interpreting the above equations. Capacity, as discussed here and in most MIMO work [1], [3], is based on a "quasi-static" analysis where the channel varies randomly from burst to burst. Within a burst the channel is assumed fixed and it is also assumed that sufficient bits are transmitted for the standard infinite time horizon of information theory to be meaningful. A second note is that our discussion will concentrate on single user MIMO systems but many results also apply to multiuser systems with receive diversity. Finally, the linear capacity growth is only valid under certain channel conditions. It was originally derived for the independent and identically distributed (i.i.d.) flat Rayleigh fading channel and does not hold true for all cases. For example, if large numbers of antennas are packed into small volumes, then the gains in $H$ may become highly correlated and the linear relationship will plateau out due to the effects of antenna correlation [28]–[30]. In contrast, other propagation effects not captured in (4) may serve to reinforce the capacity gains of MIMO such as multipath delay spread. This was shown in particular in the case when the transmit channel is known [4] but also in the case when it is unknown [5].

More generally, the effect of the channel model is critical. Environments can easily be chosen which give channels where the MIMO capacities do not increase linearly with the numbers of antennas. However, most measurements and models available to date do give rise to channel capacities which are of the same order of magnitude as the promised theory (see Section V). Also the linear growth is usually a reasonable model for moderate numbers of antennas which are not extremely close-packed.

### *B. Information Theoretic MIMO Capacity*

*1) Background:* Since feedback is an important component of wireless design (although not a necessary one), it is useful to generalize the capacity discussion to cases that can encompass transmitters having some a priori knowledge of channel. To this end, we now define some central concepts, beginning with the MIMO signal model

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n}. \tag{5}$$

In (5), $\mathbf{r}$ is the $M \times 1$ received signal vector, $\mathbf{s}$ is the $N \times 1$ transmitted signal vector and $\mathbf{n}$ is an $M \times 1$ vector of additive noise terms, assumed i.i.d. complex Gaussian with each element having a variance equal to $\sigma^2$. For convenience we normalize the noise power so that $\sigma^2 = 1$ in the remainder of this section. Note that the system equation represents a single MIMO user communicating over a fading channel with additive white Gaussian noise (AWGN). The only interference present is *self-interference* between the input streams to the MIMO system. Some authors have considered more general systems but most information theoretic results can be discussed in this simple context, so we use (5) as the basic system equation.

Let $\mathbf{Q}$ denote the covariance matrix of $\mathbf{s}$, then the capacity of the system described by (5) is given by [3], [21]

$$C = \log_2 \left[\det\left(\mathbf{I}_M + \mathbf{H}\,\mathbf{Q}\,\mathbf{H}^*\right)\right] \quad \text{b/s/Hz} \tag{6}$$

where $tr(\mathbf{Q}) \leq \rho$ holds to provide a global power constraint. Note that for equal power uncorrelated sources $\mathbf{Q} = (\rho/N)\mathbf{I}_N$ and (6) collapses to (4). This is optimal when $\mathbf{H}$ is unknown at the transmitter and the input distribution maximizing the mutual information is the Gaussian distribution [3], [21]. With channel feedback $\mathbf{H}$ may be known at the transmitter and the optimal $\mathbf{Q}$ is not proportional to the identity matrix but is constructed from a waterfilling argument as discussed later.

The form of equation (6) gives rise to two practical questions of key importance. First, what is the effect of $\mathbf{Q}$? If we compare the capacity achieved by $\mathbf{Q} = (\rho/N)\mathbf{I}_N$ (equal power transmission or no feedback) and the optimal $\mathbf{Q}$ based on perfect channel estimation and feedback, then we can evaluate a maximum capacity gain due to feedback. The second question concerns the effect of the $\mathbf{H}$ matrix. For the i.i.d. Rayleigh fading case we have the impressive linear capacity growth discussed above. For a wider range of channel models including, for example, correlated fading and specular components, we must ask whether this behavior still holds. Below we report a variety of work on the effects of feedback and different channel models.

It is important to note that (4) can be rewritten as [3]

$$C_{\mathrm{EP}} = \sum_{i=1}^{m} \log_2\left(1 + \frac{\rho}{N}\lambda_i\right) \quad \text{b/s/Hz} \tag{7}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the nonzero eigenvalues of $\mathbf{W}$, $m = \min(M, N)$, and

$$\mathbf{W} = \begin{cases} \mathbf{H}\,\mathbf{H}^*, & M \leq N \\ \mathbf{H}^*\mathbf{H}, & N < M. \end{cases} \tag{8}$$

This formulation can be easily obtained from the direct use of eigenvalue properties. Alternatively, we can decompose the MIMO channel into m equivalent parallel SISO channels by performing a singular value decomposition (SVD) of $\mathbf{H}$ [3], [21]. Let the SVD be given by $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^*$, then $\mathbf{U}$ and $\mathbf{V}$ are unitary and $\mathbf{D}$ is diagonal with entries specified by $\mathbf{D} = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_m}, 0, \ldots, 0)$. Hence (5) can be rewritten as

$$\tilde{\mathbf{r}} = \mathbf{D}\tilde{\mathbf{s}} + \tilde{\mathbf{n}} \tag{9}$$

where $\tilde{\mathbf{r}} = \mathbf{U}^*\mathbf{r}$, $\tilde{\mathbf{s}} = \mathbf{V}^*\mathbf{s}$ and $\tilde{\mathbf{n}} = \mathbf{U}^*\mathbf{n}$. Equation (9) represents the system as m equivalent parallel SISO eigen-channels with signal powers given by the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$.

Hence, the capacity can be rewritten in terms of the eigenvalues of the sample covariance matrix $\mathbf{W}$. In the i.i.d. Rayleigh fading case, $\mathbf{W}$ is also called a Wishart matrix. Wishart matrices have been studied since the 1920s and a considerable amount is known about them. For general $\mathbf{W}$ matrices a wide range of limiting results are known [22], [31]–[34] as $M$ or $N$ or both tend to infinity. In the particular case of Wishart matrices, many exact results are also available [31], [35]. There is not a great deal of information about intermediate results (neither limiting nor Wishart), but we are helped by the remarkable accuracy of some asymptotic results even for small values of $M, N$ [36].

We now give a brief overview of exact capacity results, broken down into the two main scenarios, where the channel is either known or unknown at the transmitter. We focus on the two key questions posed above; what is the effect of feedback and what is the impact of the channel?

*2) Channel Known at the Transmitter (Waterfilling):* When the channel is known at the transmitter (and at the receiver), then $\mathbf{H}$ is known in (6) and we optimize the capacity over $\mathbf{Q}$ subject to the power constraint $tr(\mathbf{Q}) \leq \rho$. Fortunately, the optimal $\mathbf{Q}$ in this case is well known [3], [4], [21], [26], [37]–[39] and is called a waterfilling (WF) solution. There is a simple algorithm to find the solution [3], [21], [26], [37], [39] and the resulting capacity is given by

$$C_{\mathrm{WF}} = \sum_{i=1}^{m} \log_2(\mu\lambda_i)^+ \quad \text{b/s/Hz} \tag{10}$$

where $\mu$ is chosen to satisfy

$$\rho = \sum_{i=1}^{m} (\mu - \lambda_i^{-1})^+ \tag{11}$$

and "$+$" denotes taking only those terms which are positive. Since $\mu$ is a complicated nonlinear function of $\lambda_1, \lambda_2, \ldots, \lambda_m$, the distribution of $C_{\mathrm{WF}}$ appears intractable, even in the Wishart case when the joint distribution of $\lambda_1, \lambda_2, \ldots, \lambda_m$ is known. Nevertheless, $C_{\mathrm{WF}}$ can be simulated using (10) and (11) for any given $\mathbf{W}$ so that the optimal capacity can be computed numerically for any channel.

The effect on $C_{\mathrm{WF}}$ of various channel conditions has been studied to a certain extent. For example in Ricean channels increasing the LOS strength at fixed SNR reduces capacity [23], [40]. This can be explained in terms of the channel matrix rank [25] or via various eigenvalue properties. The issue of correlated fading is of considerable importance for implementations where the antennas are required to be closely spaced (see Section VI). Here, certain correlation patterns are being standardized as suitable test cases [41]. A wide range of results in this area is given in [26].

In terms of the impact of feedback (channel information being supplied to the transmitter), it is interesting to note that the WF gains over EP are significant at low SNR but converge to zero as the SNR increases [39], [40], [42]. The gains provided by WF appear to be due to the correlations in $\mathbf{Q}$ rather than any unequal power allocation along the diagonal in $\mathbf{Q}$. This was shown in [40], where the gains due to unequal power uncorrelated sources were shown to be small compared to waterfilling. Over a wide range of antenna numbers and channel models the gains due to feedback are usually less than 30% for SNR above 10 dB. From zero to 10 dB the gains are usually less than 60%. For SNR values below 0 dB, large gains are possible, with values around 200% being reported at $-10$ dB. These results are available in the literature, see for example [39], but some simulations are also given in Fig. 3 for completeness. The fact that feedback gain reduces at higher SNR levels can be intuitively explained by the following fact. Knowledge of the transmit channel mainly provides transmit array gain. In contrast, gains such as diversity gain and multiplexing gain do not require this knowledge as these gains can be captured by "blind" transmit schemes such as STCs and V-BLAST (see later). Since the relative importance of transmit array gain in boosting average SNR decreases in the high SNR region, the benefit of feedback also reduces.

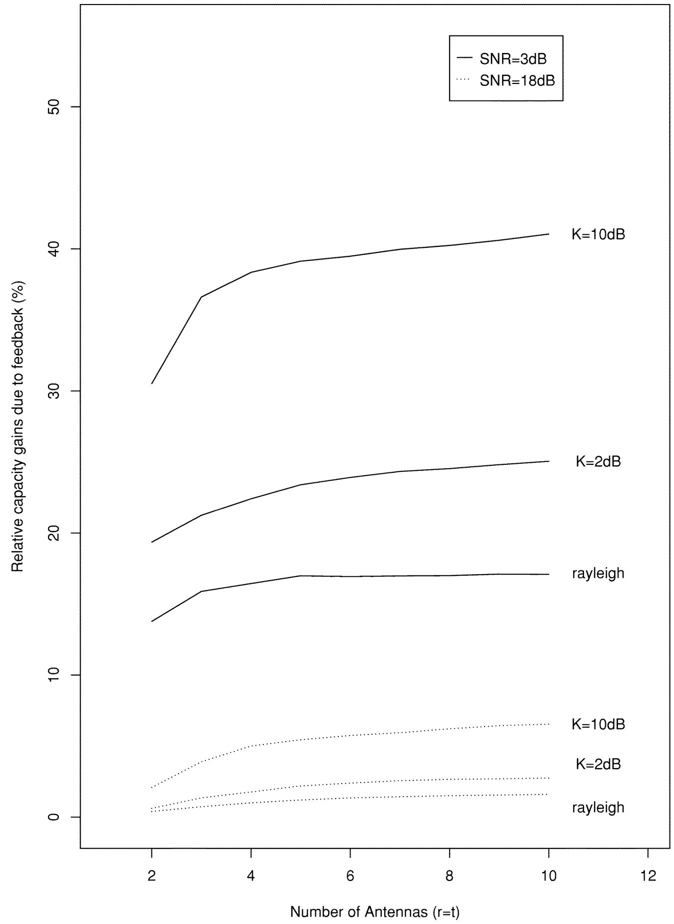



Fig. 3. Shows the percentage relative gains in capacity due to feedback at various SNR values, channel models ($K$ is the Ricean factor), and array sizes.

*3) Channel Unknown at the Transmitter:* Here, the capacity is given by $C_{\mathrm{EP}}$ in (4). This was derived by Foschini [1] and Telatar [3], [21] from two viewpoints. Telatar [3], [21] started from (6) and showed that $\mathbf{Q} = (\rho/N)I_N$ is optimal for i.i.d. Rayleigh fading. Foschini derived (4) starting from an equal power assumption. The variable, $C_{\mathrm{EP}}$, is considerably more amenable to analysis than $C_{\mathrm{WF}}$. For example, the mean capacity is derived in [3], [21] and the variance in [36] for i.i.d. Rayleigh fading, as well as [43]. In addition, the full moment generating function (MGF) for $C_{\mathrm{EP}}$ is given in [27] although this is rather complicated being in determinant form. Similar results include [44].

For more complex channels, results are rapidly becoming available. Again, capacity is reduced in Ricean channels as the relative LOS strength increases [25], [37] . The impact of correlation is important and various physical models and measurements of correlations have been used to assess its impact [26], [45]–[47]. For example, $C_{\mathrm{EP}}$ is shown to plateau out as the number of antennas increases in either sparse scattering environments [48] or dense/compact MIMO arrays [29], [30].

## C. Limiting Capacity Results

The exact results of Section III-B above are virtually all dependent on the i.i.d. Rayleigh fading (Wishart) case. For other scenarios exact results are few and far between. Hence, it is useful to pursue limiting results not only to cover a broader range of cases but also to give simpler and more intuitive results and to study the potential of very large scale systems. The surprising thing about limiting capacity results is their accuracy. Many authors have considered the limiting case where $M, N \to \infty$ and $M/N \to c$ for some constant $c$. This represents the useful case where the number of antennas grow proportionally at both TX and RX. Limiting results in this sense we denote as holding for "large systems." In particular, it covers the most interesting special case where $M = N$ and both become large. It turns out that results based on this limiting approach are useful approximations even down to $M = 2$! [36], [40], [49], [50]. We outline this work below, as well as results which are asymptotic in SNR rather than system size.

*1) Channel Known at the Transmitter:* Analytical results are scarce here but a nice analysis in [39] and [42] shows that $C_{\mathrm{WF}}/M$ converges to a constant, $\mu_{\mathrm{WF}}$, for "large systems" in both i.i.d. and correlated fading conditions. The value of $\mu_{\mathrm{WF}}$ is given by an integral equation. The rest of our "large system" knowledge is mainly based on simulations. For example, linear growth of $C_{\mathrm{WF}}$ is shown for Ricean fading in [40] as is the accuracy of Gaussian approximations to $C_{\mathrm{WF}}$ in both Rayleigh and Ricean cases.

In terms of SNR asymptotics for "large systems," [39] gives both low and high SNR results.

*2) Channel Unknown at the Transmitter:* In this situation, we have the capacity given in (4) as $C_{\mathrm{EP}}$. For "large systems" (assuming the Wishart case) the limiting mean capacity was shown to be of the form $M\mu_{\mathrm{EP}}$ [3] where $\mu_{\mathrm{EP}}$ depends on $M$, $N$ only through the ratio $c = M/N$. A closed form expression for $C_{\mathrm{EP}}$ was given in [23] and the accuracy of this result was demonstrated in [36] and [40]. The limiting variance is a constant [27], again dependent on $c$ rather than $M$ and $N$ individually. Convergence rates to this constant are indicated in [36] [40]. In fact, for a more general class of fading channels similar results hold and a central limit theorem can be stated [33], [34] as below

$$\lim_{M,\, N \to \infty} \left( \frac{C_{\mathrm{EP}} - E(C_{\mathrm{EP}})}{\sqrt{Var(C_{\mathrm{EP}})}} \right) = Z \qquad (12)$$

where $M/N \to c$ as $M, N \to \infty$ and $Z \sim N(0, 1)$ is a standard Gaussian random variable. See [33] and [34] for exact details of the conditions required for (12) to hold. Hence, for the Wishart case Gaussian approximations might be considered to $C_{\mathrm{EP}}$ using the exact mean and variance [3], [21], [36] or limiting values [23], [27]. These have been shown to be surprisingly accurate, even down to $M = 2$ [36], [40], not only for Rayleigh channels, but for Ricean channels as well. More general results which also cater for correlated fading can be found in [27], [39], and [42]. In [39] and [42], it is shown that $C_{\mathrm{EP}}/M$ converges to a constant, $\mu_{\mathrm{EP}}$, for "large systems" in both i.i.d. and correlated fading. The value of $\mu_{\mathrm{EP}}$ is obtained and it is shown that correlation always reduces $\mu_{\mathrm{EP}}$. In [27], a powerful technique is used to derive limiting results for the mean and variance in both i.i.d. and correlated fading.

Moving onto results which are asymptotic in SNR, [39] gives both low and high SNR capacity results for "large systems." It is shown that at high SNR, $C_{\mathrm{EP}}$, and $C_{\mathrm{WF}}$ are equivalent. For arbitrary values of $M$, $N$ high SNR approximations are given in [27] for the mean, variance, and MGF of $C_{\mathrm{EP}}$.

## IV. Transmission Over MIMO Systems

Although the information theoretic analysis can be bootstrapped to motivate receiver architectures (as was done, e.g., in [1], [2]), it usually carries a pitfall in that it does not reflect the performance achieved by actual transmission systems, since it only provides an upper bound realized by algorithms/codes with boundless complexity or latency. The development of algorithms with a reasonable BER performance/complexity compromise is required to realize the MIMO gains in practice. Here, we summarize different MIMO transmission schemes, give the intuition behind them, and compare their performance.

### *A. General Principles*

Current transmission schemes over MIMO channels typically fall into two categories: data rate maximization or diversity maximization schemes, although there has been some effort toward unification recently. The first kind focuses on improving the average capacity behavior. For example, in the example shown in Fig. 2, the objective is just to perform spatial multiplexing as we send as many independent signals as we have antennas for a specific error rate (or a specific outage capacity [2]).

More generally, however, the individual streams should be encoded jointly in order to protect transmission against errors caused by channel fading and noise plus interference. This leads to a second kind of approach in which one tries also to minimize the outage probability, or equivalently maximize the outage capacity.

Note that if the level of redundancy is increased between the TX antennas through joint coding, the amount of independence between the signals decreases. Ultimately, it is possible to code the signals so that the effective data rate is back to that of a single antenna system. Effectively, each TX antenna then sees a differently encoded, fully redundant version of the same signal. In this case, the multiple antennas are only used as a source of spatial diversity and not to increase data rate, or at least not in a *direct* manner.

The set of schemes aimed at realizing joint encoding of multiple TX antennas are called STCs. In these schemes, a number of code symbols equal to the number of TX antennas are generated and transmitted simultaneously, one symbol from each antenna. These symbols are generated by the *space–time encoder* such that by using the appropriate signal processing and decoding procedure at the receiver, the diversity gain and/or the coding gain is maximized. Fig. 4 shows a simple block diagram for STC.

The first attempt to develop STC was presented in [51] and was inspired by the delay diversity scheme of Wittneben [52]. However, the key development of the STC concept was originally revealed in [53] in the form of trellis codes, which required a multidimensional (vector) Viterbi algorithm at the receiver for decoding. These codes were shown to provide a diversity benefit equal to the number of TX antennas in addition to a coding gain that depends on the complexity of the code (i.e., number of states in the trellis) without any loss in bandwidth efficiency.

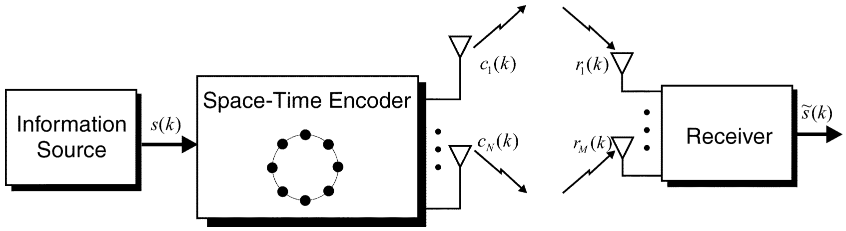



Fig. 4. Space–time coding.

Then, the popularity of STC really took off with the discovery of the so-called space–time block codes (STBCs). This is due to the fact that because of their construction, STBC can be decoded using simple linear processing at the receiver [in contrast to the vector Viterbi required for ST trellis codes (STTC)]. Although STBC codes give the same diversity gain as the STTC for the same number of TX antennas, they provide zero or minimal coding gain. Below, we will briefly summarize the basic concepts of STC and then extensions to the case of multiple RX antennas (MIMO case). As the reader will note, emphasis within space–time coding is placed on block approaches, which seem to currently dominate the literature rather than on trellis-based approaches. A more detailed summary of Sections IV-B and IV-C can be found in [54].

### *B. Maximizing Diversity With STTC*

For every input symbol $s_l$, a space–time encoder generates $N$ code symbols $c_{l1}, c_{l2}, \ldots, c_{lN}$. These $N$ code symbols are transmitted simultaneously from the $N$ transmit antennas. We define the code vector as $\mathbf{c}_l = [c_{l1}\ c_{l2} \cdots c_{lN}]^T$. Suppose that the *code vector* sequence

$$\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_L\}$$

was transmitted. We consider the probability that the decoder decides erroneously in favor of the legitimate code vector sequence

$$\tilde{\mathbf{C}} = \{\tilde{\mathbf{c}}_1, \tilde{\mathbf{c}}_2, \ldots, \tilde{\mathbf{c}}_L\}.$$

Consider a frame or block of data of length $L$ and define the $N \times N$ error matrix $\mathbf{A}$ as

$$\mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}}) = \sum_{l=1}^{L} (\mathbf{c}_l - \tilde{\mathbf{c}}_l)(\mathbf{c}_l - \tilde{\mathbf{c}}_l)^*. \tag{13}$$

If ideal channel state information (CSI) $\mathbf{H}(l)$, $l = 1, \ldots, L$ is available at the receiver, then it is possible to show that the probability of transmitting $\mathbf{C}$ and deciding in favor of $\tilde{\mathbf{C}}$ is upper bounded for a Rayleigh fading channel by [20]

$$P(\mathbf{C} \to \tilde{\mathbf{C}}) \le \left(\prod_{i=1}^{r} \beta_i\right)^{-M} \cdot (E_s/4N_o)^{-rM} \tag{14}$$

where $E_s$ is the symbol energy and $N_o$ is the noise spectral density, $r$ is the rank of the error matrix $\mathbf{A}$ and $\beta_i$, $i = 1, \ldots, r$ are the nonzero eigenvalues of the error matrix $\mathbf{A}$. We can easily see that the probability of error bound in (14) is similar to the probability of error bound for trellis coded modulation for fading channels. The term $g_r = \prod_{i=1}^{r} \beta_i$ represents the coding gain achieved by the STC and the term $(E_s/4N_o)^{-rM}$ represents a diversity gain of $rM$. Since $r \le N$, the overall diversity order is always less or equal to $MN$. It is clear that in designing a STTC, the rank of the error matrix $r$ should be maximized (thereby maximizing the diversity gain) and at the same time $g_r$ should also be maximized, thereby maximizing the coding gain.

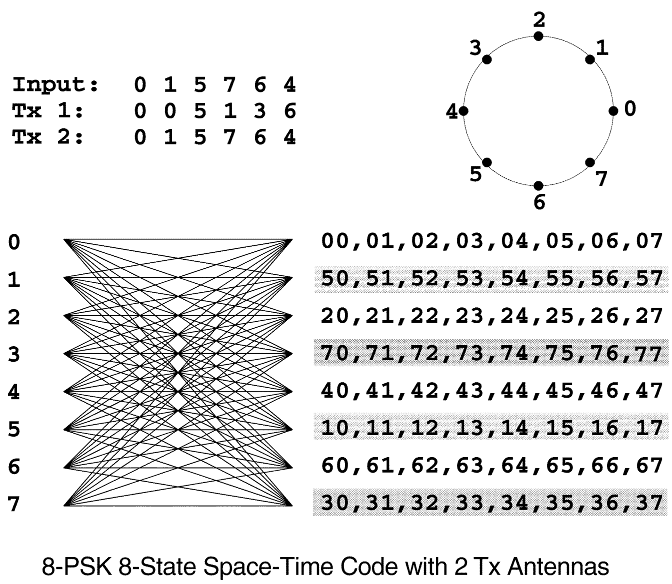



Fig. 5. The 8-PSK 8-state STC with two TX antennas.

As an example for STTCs, we provide an 8-PSK eight-state STC designed for two TX antennas. Fig. 5 provides a labeling of the 8-PSK constellation and the trellis description for this code. Each row in the matrix shown in Fig. 5 represents the edge labels for transitions from the corresponding state. The edge label $s_1 s_2$ indicates that symbol $s_1$ is transmitted over the first antenna and that symbol $s_2$ is transmitted over the second antenna. The input bit stream to the ST encoder is divided into groups of 3 bits and each group is mapped into one of eight constellation points. This code has a bandwidth efficiency of 3 bits per channel use.

Fig. 6 shows the performance of 4-PSK STTCs for two TX and one RX antennas with different number of states.

Since the original STTC were introduced by Tarokh *et al.* in [53], there has been extensive research aiming at improving the performance of the original STTC designs. These original STTC designs were hand crafted (according to the proposed design criteria) and, therefore, are not optimum designs. In recent years, a large number of research proposals have been published which propose new code constructions or perform systematic searches for different convolutional STTC or some variant of the original design criteria proposed by Tarokh *et al.* Examples of such work can be found in [55]–[60] (these are mentioned only as an example, there are many other published results that address

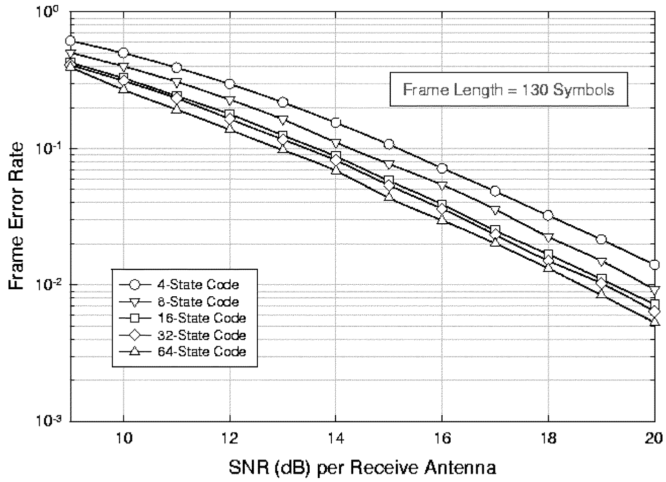



Fig. 6. Performance of 4-PSK STTCs with two TX and one RX antennas.

the same issue, too numerous to list here). These new code constructions provide an improved coding advantage over the original scheme by Tarokh *et al.*, however, only marginal gains were obtained in most cases.

### C. Maximizing Diversity With STBCs

When the number of antennas is fixed, the decoding complexity of space–time trellis coding (measured by the number of trellis states at the decoder) increases exponentially as a function of the diversity level and transmission rate [53]. In addressing the issue of decoding complexity, Alamouti [61] discovered a remarkable space–time block coding scheme for transmission with two antennas. This scheme supports maximum-likelihood (ML) detection based only on linear processing at the receiver. The very simple structure and linear processing of the Alamouti construction makes it a very attractive scheme that is currently part of both the W-CDMA and CDMA-2000 standards. This scheme was later generalized in [62] to an arbitrary number of antennas. Here, we will briefly review the basics of STBCs. Fig. 7 shows the baseband representation for Alamouti STBC with two antennas at the transmitter. The input symbols to the space–time block encoder are divided into groups of two symbols each. At a given symbol period, the two symbols in each group $\{c_1, c_2\}$ are transmitted simultaneously from the two antennas. The signal transmitted from antenna 1 is $c_1$ and the signal transmitted from antenna 2 is $c_2$. In the next symbol period, the signal $-c_2^*$ is transmitted from antenna 1 and the signal $c_1^*$ is transmitted from antenna 2. Let $h_1$ and $h_2$ be the channels from the first and second TX antennas to the RX antenna, respectively. The major assumption here is that $h_1$ and $h_2$ are scalar and constant over two consecutive symbol periods, that is

$$h_i(2nT) \approx h_i((2n+1)T), \qquad i = 1, 2.$$

We assume a receiver with a single RX antenna. we also denote the received signal over two consecutive symbol periods as $r_1$ and $r_2$. The received signals can be expressed as

$$r_1 = h_1 c_1 + h_2 c_2 + n_1 \tag{15}$$

$$r_2 = -h_1 c_2^* + h_2 c_1^* + n_2 \tag{16}$$

where $n_1$ and $n_2$ represent the AWGN and are modeled as i.i.d. complex Gaussian random variables with zero mean and power spectral density $N_o/2$ per dimension. We define the received signal vector $\mathbf{r} = [r_1 \ r_2^*]^T$, the code symbol vector $\mathbf{c} = [c_1 \ c_2]^T$, and the noise vector $\mathbf{n} = [n_1 \ n_2^*]^T$. Equations (15) and (16) can be rewritten in a matrix form as

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{c} + \mathbf{n} \tag{17}$$

where the channel matrix $\mathbf{H}$ is defined as

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}. \tag{18}$$

$\mathbf{H}$ is now only a virtual MIMO matrix with space (columns) and time (rows) dimensions, not to be confused with the purely spatial MIMO channel matrix defined in previous sections. The vector $\mathbf{n}$ is a complex Gaussian random vector with zero mean and covariance $N_o \cdot \mathbf{I}_2$. Let us define $\mathbf{C}$ as the set of all possible symbol pairs $\mathbf{c} = \{c_1, c_2\}$. Assuming that all symbol pairs are equiprobable, and since the noise vector $\mathbf{n}$ is assumed to be a multivariate AWGN, we can easily see that the optimum ML decoder is

$$\hat{\mathbf{c}} = \arg\min_{\hat{\mathbf{c}} \in \mathbf{C}} \|\mathbf{r} - \mathbf{H} \cdot \hat{\mathbf{c}}\|^2. \tag{19}$$

The ML decoding rule in (19) can be further simplified by realizing that the channel matrix $\mathbf{H}$ is always orthogonal regardless of the channel coefficients. Hence, $\mathbf{H}^*\mathbf{H} = \alpha \cdot \mathbf{I}_2$ where $\alpha = |h_1|^2 + |h_2|^2$. Consider the modified signal vector $\tilde{\mathbf{r}}$ given by

$$\tilde{\mathbf{r}} = \mathbf{H}^* \cdot \mathbf{r} = \alpha \cdot \mathbf{c} + \tilde{\mathbf{n}} \tag{20}$$

where $\tilde{\mathbf{n}} = \mathbf{H}^* \cdot \mathbf{n}$. In this case, the decoding rule becomes

$$\hat{\mathbf{c}} = \arg\min_{\hat{\mathbf{c}} \in \mathbf{C}} \|\tilde{\mathbf{r}} - \alpha \cdot \hat{\mathbf{c}}\|^2. \tag{21}$$

Since $\mathbf{H}$ is orthogonal, we can easily verify that the noise vector $\tilde{\mathbf{n}}$ will have a zero mean and covariance $\alpha N_o \cdot \mathbf{I}_2$, i.e., the elements of $\tilde{\mathbf{n}}$ are i.i.d. Hence, it follows immediately that by using this simple linear combining, the decoding rule in (21) reduces to two separate, and much simpler decoding rules for $c_1$ and $c_2$, as established in [61]. In fact, for the above $2 \times 1$ STBC, only two complex multiplications and one complex addition per symbol are required for decoding. Also, assuming that we are using a signaling constellation with $2^b$ constellation points, this linear combining reduces the number of decoding metrics that has to be computed for ML decoding from $2^{2b}$ to $2 \times 2^b$. It is also straightforward to verify that the SNR for $c_1$ and $c_2$ will be

$$\mathrm{SNR} = \frac{\alpha \cdot E_s}{N_o} \tag{22}$$

and, hence, a two branch diversity performance (i.e., a diversity gain of order two) is obtained at the receiver.

*MIMO Extensions:* Initially developed to provide transmit diversity in the MISO case, STCs are readily extended to the MIMO case. When the receiver uses $M$ RX antennas, the received signal vector $\mathbf{r_m}$ at RX antenna $m$ is

$$\mathbf{r}_m = \mathbf{H}_m \cdot \mathbf{c} + \mathbf{n}_m \tag{23}$$

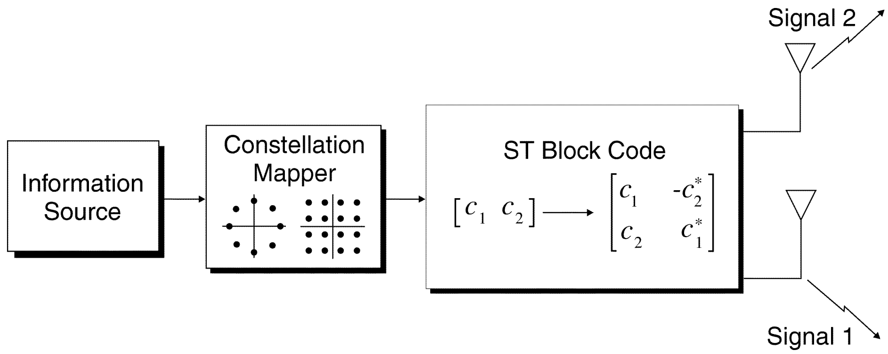



Fig. 7. Transmitter diversity with space–time block coding.

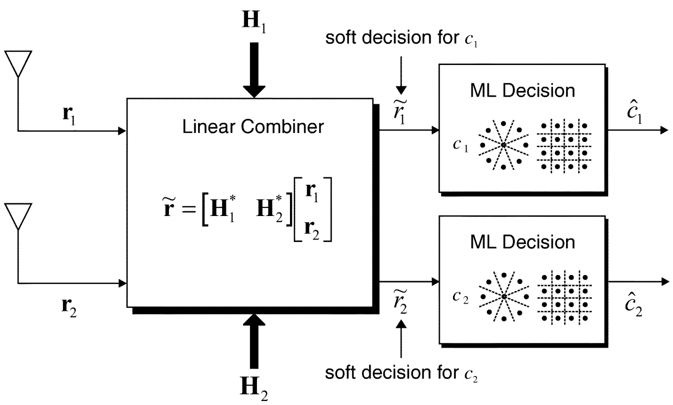



Fig. 8. Receiver for space–time block coding.

where $\mathbf{n}_m$ is the noise vector at the two time instants and $\mathbf{H}_m$ is the channel matrix from the two TX antennas to the $m$th receive antenna. In this case, the optimum ML decoding rule is

$$\hat{\mathbf{c}} = \arg\min_{\hat{\mathbf{c}} \in \mathbf{C}} \sum_{m=1}^{M} \|\mathbf{r}_m - \mathbf{H}_m \cdot \hat{\mathbf{c}}\|^2. \tag{24}$$

As before, in the case of $M$ RX antennas, the decoding rule can be further simplified by premultiplying the received signal vector $\mathbf{r}_m$ by $\mathbf{H}_m^*$. In this case, the diversity order provided by this scheme is $2M$. Fig. 8 shows a simplified block diagram for the receiver with two RX antennas. Note that the decision rule in (21) and (24) amounts to performing a hard decision on $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{r}}_M = \sum_{m=1}^{M} \mathbf{H}_m^* \mathbf{r}_m$, respectively. Therefore, as shown in Fig. 8, the received vector after linear combining, $\tilde{\mathbf{r}}_M$, can be considered as a soft decision for $c_1$ and $c_2$. Hence, in the case the STBC is concatenated with an outer conventional channel code, like a convolutional code, these soft decisions can be fed to the outer channel decoder to yield a better performance. Note also that for the above 2 $\times$ 2 STBC, the transmission rate is one while achieving the maximum diversity gain possible with two TX and two RX antennas (fourth order). However, concatenating a STBC with an outer conventional channel code (e.g., a convolutional or trellis coded modulation (TCM) code) will incur a rate loss. A very clever method to concatenate STBC based on the Alamouti scheme with an outer TCM or convolutional code was originally presented in [63]–[65]. In this approach, the cardinality of the inner STBC is enlarged to form an expanded orthogonal space–time signal set or constellation. This set is obtained by applying a unitary transformation to the original Alamouti scheme. Once this *expanded* space–time signal constellation is formed, the design of a good space–time TCM code based on this signal set is pretty much analogous to classic TCM code design. In other words, classic set partitioning techniques are used to partition signals within each block code subset. Thus, a combined STBC-TCM construct is generated and guaranteed to achieve full diversity by using a simple design rule that restricts the transition branches leaving from or arriving to each state to be labeled by codewords from the same block code subset. This rule is the same as the original design rule of STTC proposed by Tarokh *et al.* in [53]. A similar scheme was later presented in [66]. The extension of the above STBC to more than two TX antennas was studied in [62] and [67]–[69]. There, a general technique for constructing STBCs for $N > 2$ that provide the maximum diversity promised by the number of TX and RX antennas was developed. These codes retain the simple ML decoding algorithm based on only linear processing at the receiver [61]. It was also shown that for real signal constellations, i.e., PAM constellation, STBCs with transmission rate 1 can be constructed [62]. However, for general complex constellations like M-QAM or M-PSK, it *is not known* whether a STBC with transmission rate 1 and simple linear processing that will give the maximum diversity gain with $N > 2$ TX antennas *does exist or not*. Moreover, it was also shown that such a code where the number of TX antennas $N$ equals the number of both the number of information symbols transmitted and the number of time slots needed to transmit the code block *does not exist*. However, for rates <1, such codes can be found. For example, assuming that the transmitter unit uses four TX antennas, a rate 4/8 (i.e., it is a rate 1/2) STBC is given by

$$C = \begin{bmatrix} c_1 & -c_2 & -c_3 & -c_4 & c_1^* & -c_2^* & -c_3^* & -c_4^* \\ c_2 & c_1 & c_4 & -c_3 & c_2^* & c_1^* & c_4^* & -c_3^* \\ c_3 & -c_4 & c_1 & c_2 & c_3^* & -c_4^* & c_1^* & c_2^* \\ c_4 & c_3 & -c_2 & c_1 & c_4^* & c_3^* & -c_2^* & c_1^* \end{bmatrix}. \tag{25}$$

In this case, at time $t = 1$, $c_1$, $c_2$, $c_3$, $c_4$ are transmitted from antenna 1 through 4, respectively. At time $t = 2$, $-c_2$, $c_1$, $-c_4$, $c_3$ are transmitted from antenna 1 through 4, respectively, and so on. For this example, rewriting the received signal in a way analogous to (17) (where $\mathbf{c} = [c_1, \ldots, c_4]$) will yield a 8 $\times$ 4 virtual MIMO matrix $\mathbf{H}$ that is orthogonal i.e., the decoding is linear, and $\mathbf{H}^*\mathbf{H} = \alpha_4 \cdot \mathbf{I}$, where $\alpha_4 = 2 \cdot \sum_{i=1}^{4} |h_i|^2$ (fourth-order diversity). This scheme provides a 3-dB power gain that comes from the intuitive fact that eight time slots are used to transmit four information symbols. The power gain compensates for the rate loss.

As an alternative to the schemes above sacrificing code rate for orthogonality, it is possible to sacrifice orthogonality in an effort to maintain full rate one codes for $N > 2$. Quasi-orthogonal STBC were investigated for instance in [70] in which we can preserve the full diversity and full rate at the cost of a small loss in BER performance and some extra decoding complexity relative to truly orthogonal schemes.

*RX Channel Knowledge (or Lack of):* The decoding of ST block codes above requires knowledge of the channel at the receiver. The channel state information can be obtained at the receiver by sending training or pilot symbols or sequences to estimate the channel from each of the TX antennas to the receive antenna [71]–[78]. For one TX antenna, there exist differential detection schemes, such as differential phase-shift keying (DPSK), that neither require the knowledge of the channel nor employ pilot or training symbol transmission. These differential decoding schemes are used, for example, in the IS-54 cellular standard ($\pi/4$-DPSK). This motivates the generalization of differential detection schemes for the case of multiple TX antennas. A partial solution to this problem was proposed in [79] for the $2 \times 2$ code, where it was assumed that the channel is not known at the receiver. In this scheme, the detected pair of symbols at time $t-1$ are used to estimate the channel at the receiver and these channel estimates are used for detecting the pair of symbols at time $t$. However, the scheme in requires the transmission of known pilot symbols at the beginning and, hence, are not fully differential. The scheme in [79] can be thought as a joint data channel estimation approach which can lead to error propagation. In [80], a true differential detection scheme for the $2 \times 2$ code was constructed. This scheme shares many of the desirable properties of DPSK: it can be demodulated with or without CSI at the receiver, achieve full diversity gain in both cases, and there exists a simple noncoherent receiver that performs within 3 dB of the coherent receiver. However, this scheme has some limitations. First, the encoding scheme expands the signal constellation for nonbinary signals. Second, it is limited only to the $N = 2$ STBC for a complex constellation and to the case $N \leq 8$ for a real constellation. This is based on the results in [62] that the $2 \times 2$ STBC is an orthogonal design and complex orthogonal designs do not exist for $N > 2$. In [81], another approach for differential modulation with transmit diversity based on group codes was proposed. This approach can be applied to any number of antennas and any constellation. The group structure of theses codes greatly simplifies the analysis of these schemes, and may also yield simpler and more transparent modulation and demodulation procedures. A different nondifferential approach to transmit diversity when the channel is not known at the receiver is reported in [82] and [83], but this approach requires exponential encoding and decoding complexities. Additional generalizations on differential STC schemes are given in [84].

### D. STC in Frequency Selective Channels

Both STTC and STBC codes were first designed assuming a narrowband wireless system, i.e., a flat fading channel. However, when used over frequency selective channels a channel equalizer has to be used at the receiver along with the space–time decoder. Using classical equalization methods with space–time coded signals is a difficult problem. For example, for STTC designed for two TX antennas and a receiver with one RX antenna, we need to design an equalizer that will equalize two independent channels (one for each TX antenna) from *one* receive signal. For the case of the STBC, the nonlinear and noncausal nature of the code makes the use of classical equalization methods [such as the minimum mean square error (MMSE) linear equalizer, decision feedback equalizer (DFE), and maximum-likelihood sequence estimation (MLSE)] a challenging problem.

Initial attempts to address the problem for STTC made use of whatever structure was available in the space–time coded signal [85]–[87], where the structure of the code was used to convert the problem into one that can be solved using known equalization schemes. For the STBC, the equalization problem was addressed by modifying the original Alamouti scheme in such a way that the use over frequency selective channels, and hence the equalization, is a much easier task. For example, in [88], STBC was used in conjunction with OFDM. OFDM is used to convert the frequency selective channel into a set of independent parallel frequency-flat subchannels. The Alamouti scheme is then applied to two consecutive subcarriers (or two consecutive OFDM block). Note that more general code designs can be used [89].

In [90], the Alamouti scheme is imposed on a block basis (not on symbol basis as in the original scheme) and cyclic prefixes are added to each block. Using fast Fourier transform (FFT), a frequency-domain single carrier is used to equalize the channel. This is similar to OFDM except that it is a single carrier transmission system and the decisions are done in the time domain. A similar approach was proposed in [91], where the Alamouti scheme is imposed on block basis in the time domain and guard bands are added. The equalization is achieved by a clever combination of time domain filtering, conjugation, time reversal, and a SISO MLSE equalizer. This scheme is similar to that in [90] except that the equalization is now done in the time domain.

### E. Maximizing Data Rate Using Spatial Multiplexing

Spatial multiplexing, of which V-BLAST [2], [9] is a particular implementation approach, can be regarded as a special class of STBCs where streams of independent data are transmitted over different antennas, thus maximizing the average data rate over the MIMO system. One may generalize the example given in Section II in the following way: Assuming a block of independent data $\mathbf{C}$ of size $N \times L$ is transmitted over the $N \times M$ MIMO system, the receiver will obtain $\mathbf{Y} = \mathbf{HC} + \mathbf{N}$ where $\mathbf{Y}$ is of size $M \times L$. In order to perform symbol detection, the receiver must unmix the channel, in one of several various possible ways. Zero-forcing (ZF) techniques use a straight matrix inversion, a simple approach which can also result in poor results when the matrix $\mathbf{H}$ becomes very ill conditioned as in certain random fading events or in the presence of LOS (see Section V). The use of a MMSE linear receiver may help in this case, but improvements are found to be limited (1.5 to 2 dB in the $2 \times 2$ case) if knowledge of nontrivial noise/interference statistics (e.g., covariance matrix) are not exploited in the MMSE.

The optimum decoding method on the other hand is ML where the receiver compares all possible combinations of symbols which could have been transmitted with what is observed

$$\hat{\mathbf{C}} = \arg\min_{\hat{\mathbf{C}}} \|\mathbf{Y} - \mathbf{H}\hat{\mathbf{C}}\|. \tag{26}$$

The complexity of ML decoding is high, and even prohibitive when many antennas or high-order modulations are used. Enhanced variants of this like sphere decoding [92] have recently been proposed. Another popular decoding strategy proposed along side V-BLAST is known as nulling and canceling which gives a reasonable tradeoff between complexity and performance. The matrix inversion process in nulling and canceling is performed in layers where one estimates a row from $\mathbf{C}$, subtracts the symbol estimates from $\mathbf{Y}$, and continues the decoding successively [9]. Full details and analysis on this approach are provided in [14]. Note that the iterative nulling and canceling approach is reminiscent of the successive interference canceling (SIC) proposed for multiuser detection (MUD) in CDMA receivers [93]. In fact, any proposed MUD algorithm can be recast in the MIMO context if the input of the MIMO system are seen as virtual users. A difference here is that the separation is carried out in the spatial channel domain rather than the code domain, making its success dependent on channel realizations. On the other hand, the complexity of CDMA-SIC is much higher than in the MIMO case since the number of CDMA users may go well beyond the number of virtual users/antennas in a single MIMO link.

*Blind Detection:* When the channel is not known at the receiver (as well as at the transmitter) the joint detection of MIMO signals must resort to so-called "blind" approaches. Surprisingly, one may note that progress in this area has been initiated long before the results of [1]–[3], in the more general context of blind source separation (see for instance [94]). In these *blind array processing* techniques, the input sources are mixed linearly by a mixing matrix (here corresponding to the MIMO channel) and separated by exploiting higher order statistics of the receive array signals [95], [96], or covariance subspace estimation [97] and/or some alphabet (modulation format related) information [98] to cite just a few of the many contributions there. The price paid for avoiding channel training in blind approaches is in some limited loss of BER performance and more often in the increased computational complexity.

*1) Multiplexing Versus Diversity:* Pure spatial multiplexing allows for full independent usage of the antennas, however, it gives limited diversity benefit and is rarely the best transmission scheme for a given BER target. Coding the symbols within a block can result in additional coding and diversity gain which can help improve the performance, even though the data rate is kept at the same level. It is also possible to sacrifice some data rate for more diversity. In turn, the improved BER performance will buy more data rate indirectly through allowing higher level modulations, such as 16 QAM instead of 4 PSK, etc. The various tradeoffs between multiplexing and diversity have begun to be looked at, for instance in [99] and [100].

Methods to design such codes start from a general structure where one often assumes that a weighted linear combination of symbols may be transmitted from any given antenna at any given time. The weights themselves are selected in different fashions by using analytical tools or optimizing various cost functions [67], [101]–[103].

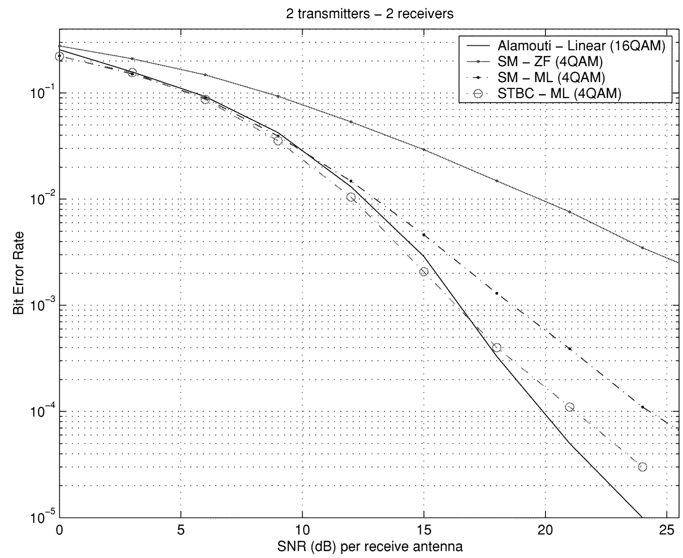



Fig. 9. BER comparisons for various transmission techniques over 2 $\times$ 2 MIMO. At high SNR, from top to bottom: Spatial multiplexing (SM)-ZF, SM-ML, STBC-ML, Alamouti STBC.

Spatial multiplexing and space–time block coding can be combined to give a transmission scheme that will maximize the average data rate over the MIMO channel and guarantee a minimum order of diversity benefit for each substream. In fact, the structure of the STBC can be exploited in a way such that the process of detecting and decoding successive steams or layers is a completely linear process. See [54] for more details.

*Numerical Comparisons:* In what follows, we compare four transmission strategies over a $2 \times 2$ MIMO system with ideally uncorrelated elements. All schemes result in the same nominal rate but offer different BER performance.

Fig. 9 plots the performance of the Alamouti code presented in Fig. 7, spatial multiplexing (SM) with ZF and with ML detection, and a spatial multiplexing scheme with ML decoding using precoding [103]. A 4-QAM constellation is used for the symbols except for the Alamouti code which is simulated under 16 QAM to keep the data rate at the same level as in the other schemes. It can be seen from the figure that spatial multiplexing with ZF returns rather poor results, while the curves for other coding-based methods are quite similar to each other. This is because using two independent streams and a ZF receiver in the $2 \times 2$ case leaves each substream starving for diversity. The Alamouti curve has the best slope at high SNR because it focuses entirely on diversity (order four). At lower SNR, the scheme combining spatial multiplexing with some block coding is the best one because ML decoding allows extraction of some diversity gain in addition to the rate (multiplexing) gain. Note that this benefit comes at the price of receiver complexity compared with Alamouti. In Section VI, we give more comparisons with system-based constraints.

It is important to note that as the number of antennas increases, the diversity effect will give diminishing returns. In contrast, the data rate gain of spatial multiplexing remains linear with the number of antennas. Therefore, for a larger number of antennas it is expected that more weight has to be put on spatial multiplexing and less on space–time coding. Interestingly, having a larger number of antennas does not need to result in a larger number of radio frequency (RF) chains. By using antenna selection techniques (see, e.g., [104]–[106]) it is possible to retain the benefits of a large MIMO array with just a subset of antennas being active at the same time.

### F. MIMO Systems With Feedback

One common aspect amongst the algorithms presented above is that they do not require any *a priori* channel information at the transmitter to extract either transmit diversity of multiplexing gains. Yet, the information theoretic analysis in Section III suggests that additional performance can be extracted from multiple antennas in the presence of channel state information at the transmitter (CSIT) through, e.g., waterfilling. It should be noted that although waterfilling may be optimal from an information theoretic point of view, it is not necessarily the best scheme using CSIT in practice. This is because the performance of real-world MIMO links are sensitive to BER performance rather than mutual information performance. Schemes that exploit CSIT to directly minimize BER-related metrics are therefore of interest, examples of which are found in [107] and [108].

One general drawback of approaches relying on complete and instantaneous CSIT at the transmitter rather than partial or statistical CSIT is feasibility and bandwidth overhead. This makes waterfilling or the equivalent difficult to realize in systems in which the acquisition of CSIT is dependent on a (typically low-rate) feedback channel from RX to TX, such as in frequency-division duplex (FDD) systems.[4] For a time division duplex (TDD) system feedback is not necessary, but only if the period for switching between a transmitter and a receiver ("ping-pong" time) is shorter than the channel coherence time, which may or may not be realized depending on the mobile's velocity (see Section V). In an effort to bring more performance and robustness to MIMO coding schemes at a reasonable cost of feedback bandwidth, a few promising solutions have been recently proposed to incorporate CSIT in the space–time transmit encoder. Solutions to reduce the feedback cost include using instantaneous yet partial (few bits) CSIT [109] or stastistics of CSIT, such as long term channel correlation information [110], [111], to name a few of the recent papers here.

## V. MIMO Channel Modeling

Because of the sensitivity of MIMO algorithms with respect to the channel matrix properties, channel modeling is particularly critical to assess the relative performance of the various MIMO architectures shown earlier in various terrains. Key modeling parameters, for which results from measurements of MIMO, as well as SISO can be exploited include path loss, shadowing, Doppler spread and delay spread profiles, and the Ricean $K$ factor distribution. Much more specific to MIMO and, hence, of interest here, are

- the joint antenna correlations at transmit and receive ends;
- the channel matrix singular value distribution.

[4]FDD is the main duplexing approach for 3G wireless (WCDMA, CDMA-2000).

In practice, the latter is more accurately represented by the *distribution* of eigenvalues of $\mathbf{HH}^*$, denoted $\{\lambda_1, \lambda_2, \ldots\}$. In what follows, we describe the impact of environmental parameters (LOS component, density of scattering) and antenna parameters (spacing, polarization) on the correlation/eigenvalue distribution.

### A. Pseudostatic Narrowband MIMO Channel

*1) LOS Component Model:* It is common to model a wireless channel as a sum of two components, a LOS component and a NLOS component. That is, $\mathbf{H} = \mathbf{H}_{\mathrm{LOS}} + \mathbf{H}_{\mathrm{NLOS}}$. The Ricean $K$ factor is the ratio between the power of the LOS component and the mean power of the NLOS component.

In conventional SISO wireless deployments, it is desirable that antennas be located where the channel between the transmitter and the receiver has as high a Ricean $K$ factor as possible. The higher the $K$ factor, the smaller the fade margin that needs to be allocated. For example, to guarantee service at 99% reliability, the fade margin for $K = 10$ is more than 10-dB lower than that for $K = 0$ (pure Rayleigh fading). Furthermore, as we mentioned earlier, certain beamforming techniques, especially those relying on angle-of-arrival (AOA) estimation are effective only if the LOS component dominates.

For MIMO systems, however, the higher the Ricean $K$ factor, the more dominant $\mathbf{H}_{\mathrm{LOS}}$ becomes. Since $\mathbf{H}_{\mathrm{LOS}}$ is a time-invariant, often low rank matrix [112], its effect is to drive up antenna correlation and drive the overall effective rank down (more precisely the singular value spread is up). High-$K$ channels show low useable spatial degrees of freedom and, hence, a lower MIMO capacity for the same SNR. For example, at $\rho = 6$ dB, the channel capacity for a (4, 4) MIMO channel with $K = 0$ is almost always higher than that with $K = 10$. Note, however, that this does not mean that one would intentionally place the antennas such that the LOS component diminishes. Near-LOS links typically enjoy both a more favorable path loss and less fading. In such cases, the resulting improvement in link budget may more than compensate the loss of MIMO capacity.

Recently, experimental measurements have been carried out to try to characterize the distribution of the $K$ factor in a coverage area [113]–[115]. In [113], an empirical model was derived for typical macrocell fixed-wireless deployment. The $K$ factor distribution was modeled as lognormal, with the median as a function of season, antenna heights, antenna beamwidth, and distance: $K \propto$ (antenna height)$^{0.46}$ (distance)$^{-0.5}$. Using this model, one can observe that the $K$ factor decreases as the distance increases. The implication, from a network deployment perspective, is that even though the use of MIMO does not materially improve the link throughput near the base station, where the signal strength is usually high enough to support the desired applications, it does substantially improve the quality of service in areas that are far away from the base station, or are physically limited to using low antennas.

In metropolitan areas, microcell deployment is popular. In a microcell, the base station antenna is typically at about the same height as street lamp posts, and the coverage radius is no more than a few hundred meters. Microcell channels frequently involve the presence of a LOS component and, thus, may be expected to be Ricean [116]. Similar to macrocells, in a microcell

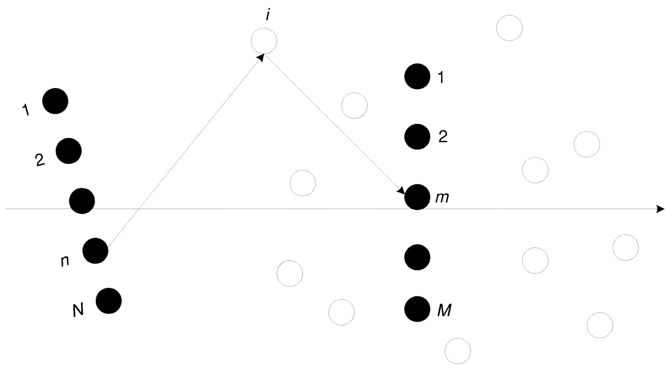



Fig. 10. Diagram to derive antenna correlation $\Psi$. The $i$th ($1 \leq i \leq n_s$) path from TX antenna $n$ to RX antenna $m$ goes through the $i$th single-bounce scatterer.

the $K$ factor declines when distance increases. Overall the $K$ factor observed in a microcell tends to be smaller than that in a macrocell.

In an indoor environment, many simulations [42] and measurements [117] have shown that typically the multipath scattering is rich enough that the LOS component rarely dominates. This plays in favor of in-building MIMO deployments (e.g., WLAN).

*2) Correlation Model for NLOS Component:* In the absence of a LOS component, the channel matrix reduces to $\mathbf{H}_{\mathrm{NLOS}}$ and is usually modeled with circularly-symmetric complex Gaussian random variables (i.e., Rayleigh fading). The elements of $\mathbf{H}_{\mathrm{NLOS}}$ can be correlated though, often due to insufficient antenna spacing, existence of few dominant scatterers and small AOA spreading. Antenna correlation is considered the leading cause of rank deficiency in the channel matrix, although as we see later, it may not always be so.

*Modeling of Correlation:* A full characterization of the second-order statistics of $\mathbf{H}_{\mathrm{NLOS}}$ is $\mathrm{cov}(\mathrm{vec}(\mathbf{H}_{\mathrm{NLOS}})) \equiv \Psi$, where $\mathrm{cov}$ and $\mathrm{vec}$ are the covariance and matrix vectorization operator (stacking the columns on top of each other), respectivelly. In the following, we will introduce commonly accepted models for $\mathrm{cov}(\mathrm{vec}(\mathbf{H}))$. Before that, let us first review a simple model shown in Fig. 10.

Consider a transmitter TX with $N$ antennas and a receiver RX with $M$ antennas. For simplicity, the antenna pattern is assumed to be omni-directional. Ignoring the rays that involve more than one scatterer, the channel gain between antenna $T_n$ and antenna $R_m$ is the summation of the contributions from each of the scatterers

$$h(R_m, T_n) = \sum_{i=1}^{n_s} r_i(R_m, T_n) \tag{27}$$

where $n_s$ is the number of scatterers and $r_i(R_m, T_n)$ is the complex amplitude associated with a ray emanated from antenna $T_n$, reflected by scatterer $i$, and then received at antenna $R_m$. The correlation between $h_{R_m, T_n}$ and $h_{R_{m'}, T_{n'}}$ can then be given by

$$\Psi(R_m T_n, R_{m'} T_{n'}) = \frac{\mathrm{E}\left[\sum_{i=1}^{n_s} r_i(R_m, T_n) \sum_{i=1}^{n_s} r_i(R_{m'}, T_{n'})^*\right]}{\sqrt{\mathrm{E}(|h(R_m, T_n)|^2)\mathrm{E}(|h(R_{m'}, T_{n'})|^2)}}. \tag{28}$$

An appropriate model for a macrocell deployment in a suburban environment is as follows [118]. The base station TX is elevated above urban clutter and far away from the scatterers, while on the other hand, the mobile terminal RX is surrounded by scatterers. Consider that infinitely many scatterers exist uniformly in azimuth angle around the mobile. Furthermore, consider that the amplitudes of the scattered rays are identical, whereas the phases of them are completely independent. Under these assumptions, one can easily show that $\Psi(R_m T_n, R_{m'} T_n) = J_0((2\pi/\lambda) D(R_m, R_{m'}))$, where $D(R_m, R_{m'})$ is the distance between antennas $R_m$ and $R_{m'}$. Hence, the decorrelation distance can be as low as half a wavelength.

It can be more involved to compute the correlation due to antenna separation at the base station, $\Psi(R_m T_n, R_m T_{n'})$. If the base station is higher than its surroundings, it is often the case that only waves transmitted within azimuth angle $\theta \in [\Theta - \Delta, \Theta + \Delta]$ can reach the mobile. Here, $\Theta$ and $\Delta$ correspond to the AOA and angle spread, respectively. Let us denote the distribution of scatterers in azimuth angle, as seen by the base station, by $p(\theta)$. This function $p(\theta)$ is referred to as a power azimuth distribution (PAD). Given $p(\theta)$, the spatial correlation function can be given by

$$\Psi(R_m T_n, R_m T_{n'}) = \int_{\Theta-\Delta}^{\Theta+\Delta} p(\theta) \exp\left(j \frac{2\pi \sin(\theta)}{\lambda} D(T_n, T_{n'})\right) d\theta \tag{29}$$

where $D(T_n, T_{n'})$ is the distance between base station antennas $T_n$ and $T_{n'}$.

Let us consider a particular choice of $p(\theta)$ which corresponds to the case where scatterers are uniformly distributed on a circle. The mobile is at the center of the circle. If the mobile is right at the broadside direction, i.e., $\Theta = 0$, then $\Psi(R_m T_n, R_m T_{n'}) \approx J_0((2\pi\Delta/\lambda) D(T_n, T_{n'}))$. On the other hand, if the mobile is at the inline direction, i.e., $\Theta = \pi/2$, then $|\Psi(R_m T_n, R_m T_{n'})| \approx J_0((2\pi/\lambda)\Delta^2 D(T_n, T_{n'}))$ [45]. It is apparent that at deployment, to obtain the highest diversity, one must ensure that the orientation of the base station antenna array is such that the mobiles are mostly distributed in the broadside direction. This is already common practice whenever possible. Note that in order for the antenna correlation to be low, one desires a large antenna spacing at the base station; on the other hand, phase-array beamforming will only perform well if the antennas are closely spaced in order to prevent spatial aliasing. Thus, at deployment one must make a choice between optimizing for beamforming or MIMO.

In addition to the PAD chosen above, there are a few other plausible PADs studied in the literature, e.g., uniform, truncated normal and Laplacian [119]. Different PADs naturally leads to different relations between antenna correlation and AOA or angle spread. Nevertheless, all point to the general trend that in order to reduce antenna correlation, one must increase the antenna separation, and ensure that $\Theta$ is as close to zero as possible.

Compared to macrocells, for microcell deployment, the uplink waves arriving at the base station may come predominantly from a few directions. In other words, $p(\theta)$ is nonzero in $[\Theta_0 - \Delta_0, \Theta_0 + \Delta_0] \cup [\Theta_1 - \Delta_1, \Theta_1 + \Delta_1] \cup \ldots$. Interestingly, as

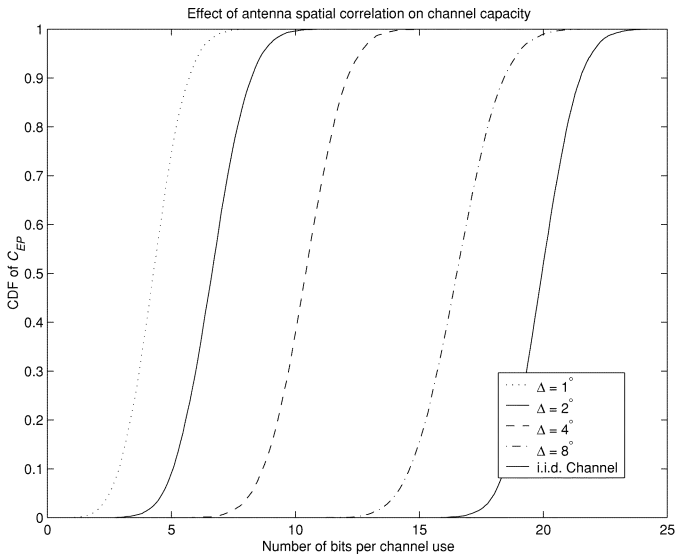



Fig. 11. Distribution of capacity as a function of angle spread for an (8, 8) system with $\rho = 8$. In producing this figure, transmit power is evenly divided, $\Theta = 0°$, $D(T_n, T_{n'}) = 3\lambda|n - n'|$, and $D(R_m, R_{m'}) = \lambda|m - m'|$.

long as the distribution of $\Theta_i$ is diverse enough, the antennas will become fairly uncorrelated, even with angle spreads $\Delta_i$ approaching zero [120].

### B. Impact of Spatial Correlation

The statistics of $\mathrm{vec}(\mathbf{H}_{\mathrm{NLOS}})$ given $\Psi$ is equal to that of $\Psi^{1/2}\mathrm{vec}(\mathbf{H}_w)$, where $\mathbf{H}_w$ is an $N$-by-$M$ matrix with i.i.d. circularly symmetric complex Gaussian entries. For convenience, it is common to approximate the correlation matrix $\Psi$ to be a Kronecker product of the two local correlation matrices. That is, let $\Psi^R$ and $\Psi^T$ denote the antenna correlation matrices at RX (mobile) and TX (base station), respectively; the approximation is $\mathrm{cov}(\mathrm{vec}(\mathbf{H}_{\mathrm{NLOS}})) \equiv \Psi \approx \Psi^R \otimes \Psi^T$. Under the assumption that the components of $\mathbf{H}_{\mathrm{NLOS}}$ are jointly Gaussian, the statistics of $\mathbf{H}_{\mathrm{NLOS}}$ is identical to those of $(\Psi^R)^{1/2}\mathbf{H}_w(\Psi^T)^{1/2}$. This is a useful form for mathematical manipulation. Fig. 11 shows the distribution of channel capacity of an (8, 8) system as a function of angle spread, assuming that the channel statistics can indeed be described by $(\Psi^R)^{1/2}\mathbf{H}_w(\Psi^T)^{1/2}$. In general, as the angle spread becomes narrower, the spatial correlation increases. As a result, the channel capacity decreases.

If the channel $\mathbf{H}$ can be described by $(\Psi^R)^{1/2}\mathbf{H}_w(\Psi^T)^{1/2}$, then an upper bound of channel capacity can be derived. The channel capacity given $\mathbf{H}_w$ and an SNR budget $\rho$ can be upper bounded by

$$C(\rho, \mathbf{H}_w) \le \max_{\rho_k} \sum_{k=1}^{\mathrm{rank}(\mathbf{H}_w)} \log_2(1 + \rho_k \upsilon_k^R \upsilon_k^T \lambda_k) \qquad (30)$$

where $\upsilon_k^R$, $\upsilon_k^T$ and $\lambda_k$ are the $k$th largest eigenvalues for $\Psi^R$, $\Psi^T$ and $\mathbf{H}_w\mathbf{H}_w^*$, respectively, and $\sum \rho_k = \rho$ [45].

Even though (29) is not a very tight bound, it does offer useful insights into the impact of spatial correlation on channel capacity. The higher the channel correlation, the more rapidly the sequence $\upsilon_k^R \upsilon_k^T$ diminishes toward zero. One can easily obtain an upper bound on the effective channel rank from the products of $\upsilon_k^R \upsilon_k^T$.

*1) Decoupling Between Rank and Correlation:* Though convenient, one must be careful in using the $\Psi \approx \Psi^R \otimes \Psi^T$ approximation. For instance a situation can arise where there is significant local scattering around both the BTS and the subscriber unit, causing uncorrelated fading at each end of the MIMO link and yet only a low rank is realized by the channel matrix. That may happen because the energy travels through a narrow "pipe." Mathematically, this is the case if the product of the scattering radius around the transmitter and that around the receiver divided by the TX–RX distance is small compared with the wavelength, as was modeled in [112]. Such a scenario is depicted in Fig. 12. Channels exhibiting at the same time antenna decorrelation (at both ends) *and* a low matrix rank are referred to as *pinhole* or *keyhole* channels in the literature [112], [121]. Pinhole channels can also result from certain rooftop diffraction effects [121]. However, most MIMO measurements carried out so far suggest that rank loss due to the pinhole effect is not common. In fact the results reported largely confirm the high level of dormant capacity of MIMO arrays, at least in urban or suburban environments. Indoor scenarios lead to even better results. Samples of analysis for UMTS type scenarios can be found in [122]–[126]. Measurements conducted at 2.5 GHz for broadband wireless access applications can be found in [115].

*2) Correlation Model Between Two Polarized Components:* Both reflection and diffraction processes are polarization sensitive, and can produce a rotation of the polarization of the scattered wave compared to the incident wave. This leads to the possibility of constructing a MIMO system using a pair of polarized antennas at both ends, with the two antennas potentially colocated and avoiding some of the issues above related to lack of richness in multipath.

Consider a MIMO channel using a pair of vertical and horizontal polarized antennas at both ends. A $2 \times 2$ matrix with equal-variance complex Gaussian entries clearly is not an appropriate narrowband channel model. First, the propagation environment may dictate that the pass losses for the two polarizations are different. Secondly, the cross-polar component is typically considerably weaker than the co-polar component. In general, the more sparse the scatterers, the lower the effect of cross polarization. Also, as distance between the two terminals increases, the cross polarization decreases. The cross-polarization ratio was found to be around 7.4 dB in macrocells in the 900 MHz band [127].

In typical outdoor environments with reasonable scattering, it has been found experimentally [127]–[129] that the co-polar and the cross-polar received components are almost uncorrelated. The mean correlation coefficients are around 0.1 or below, and were found to increase somewhat with range in microcells. Nevertheless, as the range increases, the power difference between the co-polar and the cross-polarization components increases. If the difference is high, regardless of the correlation between the co-polar and the cross-polar components, the effective rank of the $2 \times 2$ matrix will always be two.

Overall, the use of multipolarized antenna setups for MIMO opens the door to fairly compact MIMO designs while achieving enhanced robustness with respect to the multipath characteristics [130].

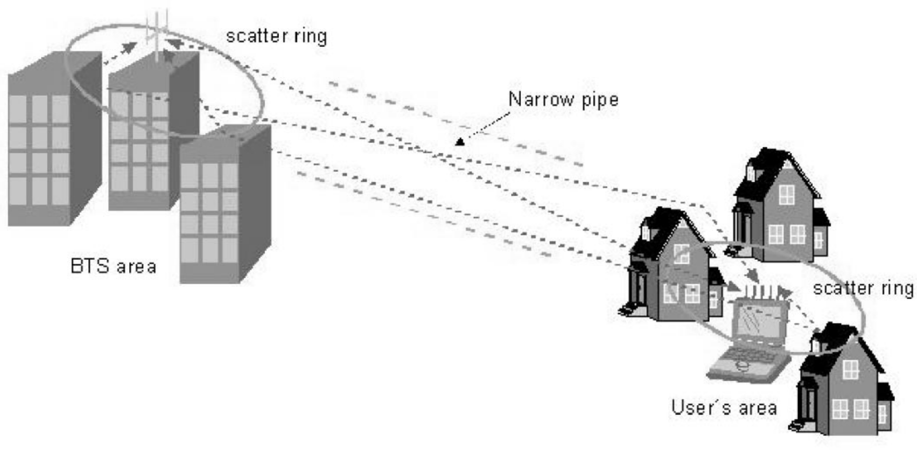



Fig. 12. An example of pin-hole realization. Reflections around the BTS and subscribers cause locally uncorrelated fading. However, because the scatter rings are too small compared to the separation between the two rings, the channel rank is low.

*3) Toward Using Orthogonal Antenna Patterns:* Antenna pattern diversity at either end of the MIMO link is particularly useful at a site where the waves are coming from diverse angles. Like polarization diversity, it allows for the collocation of antennas. Unlike polarization, where only two orthogonal modes are available, it is theoretically possible to utilize antennas with sharp patterns to obtain many more orthogonal modes. If the incoming waves do indeed distribute uniformly in AOA, a multimode antenna is expected to provide a large number of diversity branches in a very small physical footprint [131], although limited by the number of independent paths.

Since each antenna receives waves coming from different angles, in general one expects the average power, Doppler spectrum, and delay spread profile for each antenna pattern to be different. Thus, to model a MIMO system using antenna pattern diversity correctly, one must be careful in specifying the correlation matrix for $\mathrm{vec}(\mathbf{H})$; a matrix of correlated, equal-variance complex Gaussian entries may not be an appropriate model for such a MIMO channel.

*4) Effective Degrees of Freedom:* In Section III, we have shown that an $(M, N)$ channel can be decomposed into an equivalent system consisting of $\min(M, N)$ parallel SISO subchannels whose channel power gains are the eigenvalues $\lambda_k$ of $\mathbf{W}$. With an SNR so high that $\rho_k\lambda_k > 1 \ \forall k$, every additional 3 dB increase in signal power leads to an increase of $\min(M, N)$ b/s/Hz in channel capacity. However, the higher the correlation among the components of $\mathbf{H}$, in general the more widely spaced the primary support regions for the distributions of these eigenvalues. Effective degrees of freedom (EDOF) is a quantity defined to empirically observe the number of these SISO subchannels that effectively contribute to the channel capacity

$$\mathrm{EDOF} \equiv \frac{d}{d\delta} C_{\mathrm{EP}} \left(2^{\delta}\rho\right)\Big|_{\delta=0}. \tag{31}$$

Although the channel matrix $\mathbf{H}$ has rank $\min(M, N)$ with probability one in general, only the power allocated to EDOF out of these dimensions contributes to channel capacity. EDOF is considered a slowly time-varying property of the channel.

### *C. Time-Varying Wideband MIMO Channel*

Similar to the extension of a narrowband SISO channel model to a wideband SISO model, it is generally accepted that one can model a time-varying wideband MIMO channel as a sum of a LOS component and several delayed random fading components

$$\mathbf{H}(\tau) = \sum_{i=1}^{L} \mathbf{H}_i \delta(\tau - \tau_i)$$

where only $\mathbf{H}_1 = \mathbf{H}_{\mathrm{LOS}} + \mathbf{H}_{\mathrm{random}}$ contains a LOS component and a random fading component. Note that $\mathbf{H}(\tau)$ is a complex $M \times N$ matrix and $\mathbf{H}_i$ describes the linear transformation between the two antenna arrays at delay $\tau_i$, possibly using one of the previously mentioned flat fading models. This is simply a tapped delay line model where the channel coefficients at the $L$ delays are represented by matrices. Because the dimension of the antenna array is in general much smaller than the distance light travels between the taps, the short term statistics of these different taps are considered uncorrelated.

As mentioned before, the performance of MIMO techniques depends heavily on the spatial correlation of the antenna elements. For a terminal with limited space resources, MIMO works best when such a terminal is in a location where the decorrelation distance is short. Unfortunately, in such a low decorrelation distance environment, even if the terminal is moving at a reasonable speed, the channel matrix $\mathbf{H}$ can evolve at a very fast rate. This rate is also called *Doppler spread* and varies from a few Hertz in stationary applications to 200 Hz or so in fast mobile scenarios.

Clearly, the value of the Doppler spread multiplied by the number of simultaneous users will determine the traffic overhead incurred by channel feedback for cases where a MIMO or STC scheme is implemented that relies on some instantaneous form of CSIT. The Doppler spread also determines the timing requirement from the moment of channel measurement to the moment the transmitter adapts to the channel feedback. A full feedback of CSIT may quickly become prohibitive in practice and simpler rules for transmit adaptation of the MIMO signaling algorithm may be an attractive solution [132].

TABLE I
EXAMPLE OF STANDARDIZED MIMO CHANNELS FOR IEEE BODY 802.16

| SUI - 1 Channel | | | |
|---|---|---|---|
| | tap 1 | tap 2 | tap 3 |
| Delay ($\mu$s) | 0 | 0.4 | 0.8 |
| Power (dB) | 0 | -15 | -20 |
| K factor | 16 | 0 | 0 |
| Antenna correlation: $\Psi = 0.7$ | | | |

| SUI - 6 Channel | | | |
|---|---|---|---|
| | tap 1 | tap 2 | tap 3 |
| Delay ($\mu$s) | 0 | 14 | 20 |
| Power (dB) | 0 | -14 | -20 |
| K factor | 0 | 0 | 0 |
| Antenna correlation: $\Psi = 0.3$ | | | |

In a location where a LOS component dominates, even if the terminal is moving at a very high speed, the effective change in channel is actually small. Thus, the rate for full channel information feedback can be reasonable.

### D. Standardized Models

Recently, MIMO models have been standardized in IEEE 802.16 for fixed broadband wireless access and third-generation partnership project (3GPP) for mobile applications. The MIMO channel model adopted in IEEE 802.16 is described in [133]. In [133], a total of six typical models for (2, 2) macrocell fixed-wireless channel are proposed. The assumption made in the model includes vertical polarization only, the correlation matrix being the Kronecker product of the local correlation matrices, and every tap sharing the same antenna correlation. Table I shows two of the channel models proposed in [133]. Note that the SUI-1 channel is the most correlated channel and SUI-6 is the least correlated channel.

The discussions in 3GPP [41] are concerned with standardizing MIMO channel models, with the emphasis on definitions and ranges for the following:

- power azimuth spectrum and AOA for macrocells and microcells at zero mobility, pedestrian, and vehicular mobility;
- power delay profiles for the above cases;
- Ricean $K$-factor values for the above cases.

## VI. MIMO Applications in 3G Wireless Systems and Beyond

### A. Background

With MIMO-related research entering a maturing stage and with recent measurement campaign results further demonstrating the benefits of MIMO channels, the standardization of MIMO solutions in third generation wireless systems (and beyond) has recently begun, mainly in fora such as the International Telecommunications Union and the 3GPPs. Several techniques, seen as complementary to MIMO in improving throughput, performance and spectrum efficiency are drawing interest, especially as enhancements to present 3G mobile systems, e.g., high-speed digital packet access (HSDPA) [134]–[136]. These include adaptive modulation and coding, hybrid ARQ, fast cell selection, transmit diversity.

### B. MIMO in 3G Wireless Systems And Beyond

There is little commercial implementation of MIMO in cellular systems as yet and none is currently being deployed for 3G outside pure transmit diversity solutions for MISO. Current MIMO examples include the Lucent's BLAST chip and proprietary systems intended for specific markets such as Iospan Wireless' AirBurst system for fixed wireless access [137]. The earliest lab trials of MIMO have been demonstrated by Lucent Technologies several years ago.

In the case of 3GPP, some MIMO results are presented here, based on link level simulations of a combination of V-Blast and spreading code reuse [136]. Table II gives the peak data rates achieved by the down link shared channel using MIMO techniques in the 2-GHz band with a 5-MHz carrier spacing under conditions of flat fading. The gains in throughput that MIMO offer are for ideal conditions and are known to be sensitive to channel conditions. In particular, the conditions in urban channels that give rise to uncorrelated fading amongst antenna elements are known to be suitable for MIMO. The gains of MIMO come at the expense of increased receiver complexity both in the base station and in the handsets. Also various factors such as incorrect channel estimation, presence of correlation amongst antenna elements, higher Doppler frequencies, etc., will tend to degrade the ideal system performance. A brief discussion on some of the open issues and remaining hurdles on the way to a full scale commercialization of MIMO systems is contained below.

*1) Antenna Issues:* Antenna element numbers and interelement spacing are key parameters, especially the latter if the high spectral efficiencies of MIMO are to be realized. Base stations with large numbers of antennas pose environmental concerns. Hence, the antenna element numbers are limited to a modest number, say four, with an interelement spacing of around $10\lambda$. The large spacing is because base stations are usually mounted on elevated positions where the presence of local scatterers to decorrelate the fading cannot be always guaranteed. Using dual polarized antennas, four antennas can fit into a linear space of 1.5 m at $10\lambda$ spacing at 2 GHz. For the terminal, $1/2\ \lambda$ spacing is sufficient to ensure a fair amount of uncorrelated fading because the terminal is present amongst local scatterers and quite often there is no direct path. The maximum number of antennas on the terminal is envisaged to be four, though a lower number, say two, is an implementation option. Four dual polarized patch antennas can fit in a linear space of 7.5 cm. These antennas can easily be embedded in casings of lap tops. However, for handsets, even the fitting of two elements may be problematic. This is because, the present trend in handset design is to imbed the antennas inside the case to improve look and appeal. This makes spacing requirements even more critical.

*2) Receiver Complexity:* MIMO channel estimation results in increased complexity because a full matrix needs to be tracked per path delay (or per tone in OFDM) instead of a single coefficient. Since practical systems typically limit the number of antenna elements to a few, this added complexity is

TABLE II
PEAK DATA RATES OF VARIOUS MIMO ARCHITECTURES

| (M,N) | Tx technique | Code rate | Modulation | Rate/sub-stream | # sub- streams | Data rate |
|---|---|---|---|---|---|---|
| (1,1) | Conven-tional | 3/4 | 64QAM | 540 kbps | 20 | 10.8 Mbs |
| (2,2) | MIMO | 3/4 | 16QAM | 360 kbps | 40 | 14.4 Mbs |
| (2,2) | MIMO | 3/4 | QPSK | 180 kbps | 80 | 14.4 Mbs |
| (4,4) | MIMO | 1/2 | 8PSK | 540 kbps | 80 | 21.6 Mbs |

not seen as a bottle neck. Extra complexity comes from extra RF, hardware, and sophisticated receiver separation algorithms. A MIMO receiver should be dual mode to support non-MIMO mode. In the MIMO mode, it will have multiple RF chains (equal to the number of RX antennas), and additional baseband operations i.e., the space–time combiners and detector to eliminate spatial interference. The additional requirements increase the complexity of a (4,4) MIMO system to about twice that of a single antenna receiver [136], [138], [139]. There may also be additional processing (equalization or interference cancellation) needed due to dispersive channel conditions resulting from delay spread of the environment surrounding the MIMO receiver. The complexity impact of these is not yet fully accounted for.

Homodyne detection may provide direct conversion to baseband and, thus, avoid the need for SAW filters in the IF circuitry. This could reduce the RF complexity aspects of MIMO. Whilst the overall cost impact of MIMO complexity is not clear, one thing is clear: MIMO receivers are likely to cost more than conventional receivers and in the terminal the battery life may also be an issue.

*3) System Integration and Signaling:* The MIMO system needs to be integrated and be backward compatible with an existing non MIMO network. MIMO signaling imposes the support of special radio resource control (RRC) messages. The terminals need to know via broadcast down link signaling if a base station is MIMO capable. The base station also needs to know the mobile's capability, i.e., MIMO or non-MIMO. This capability could be declared during call set up. Handsets are also required to provide feedback to the base station on the channel quality so that MIMO transmission can be scheduled if the channel conditions are favorable. These downlink and uplink RRC messages are then mapped on to the layer 2 signaling messages [139].

*4) MIMO Channel Model:* The performance of a MIMO system is very much influenced by the underlying channel model especially the degree of correlation amongst the elements of the channel matrix, delay spread issues, etc. While the propagation models for conventional radio systems have been standardized in [140], there is no agreed MIMO channel model by the ITU as yet.

*5) CSI at Transmitter:* As shown earlier, the channel capacity is a function of the eigenmodes of the channel. The MIMO capacity will benefit from the transmitter having a knowledge of the channel state and may use water filling instead of equal power allocation [21], [39] or some partial form of feedback. Furthermore, knowing the channel correlation matrix, the transmitter could optimize channel coding, bit allocation per substream in addition to amplifier power management [141]. Various power allocation algorithms are discussed in [36] which are optimum during different channel conditions. The feedback of accurate and timely CSI to the transmitter is another open issue.

## VII. CONCLUSIONS AND FUTURE TRENDS

This paper reviews the major features of MIMO links for use in future wireless networks. Information theory reveals the great capacity gains which can be realized from MIMO. Whether we achieve this fully or at least partially in practice depends on a sensible design of transmit and receive signal processing algorithms. It is clear that the success of MIMO algorithm integration into commercial standards such as 3G, WLAN, and beyond will rely on a fine compromise between rate maximization (BLAST type) and diversity (space–time coding) solutions, also including the ability to adapt to the time changing nature of the wireless channel using some form of (at least partial) feedback. To this end more progress in modeling, not only the MIMO channel but its specific dynamics, will be required. As new and more specific channel models are being proposed it will useful to see how those can affect the performance tradeoffs between existing transmission algorithms and whether new algorithms, tailored to specific models, can be developed. Finally, upcoming trials and performance measurements in specific deployment conditions will be key to evaluate precisely the overall benefits of MIMO systems in real-world wireless scenarios such as UMTS.

## ACKNOWLEDGMENT

The authors would like to thank J. Akhtar for discussions and his help in the simulations. They also would like to thank Prof. D. Falconer, Prof. H. Bolcskei, Prof. A. Paulraj, and Dr. R. Kalbasi and Dr. D. Gore for reviewing the manuscript and for their constructive remarks.

## REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.

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

[3] E. Telatar, "Capacity of multiantenna Gaussian channels," AT&T Bell Laboratories, Tech. Memo., June 1995.

[4] G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, 1998.

[5] H. Bölcskei, 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] A. Paulraj and C. B. Papadias, "Space–time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[7] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception," Patent 5 345 599, 1994.

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

[9] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space–time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–15, 1999.

[10] G. J. Foschini, G. D. Golden, P. W. Wolniansky, and R. A. Valenzuela, "Simplified processing for wireless communication at high spectral efficiency," *IEEE J. Select. Areas Commun.—Wireless Commun. Series*, vol. 17, pp. 1841–1852, 1999.

[11] W. Choi, K. Cheong, and J. Cioffi, "Iterative soft interference cancellation for multiple antenna systems," in *Proc. Wireless Communications and Networking Conf.*, Chicago, IL, 2000, pp. 304–309.

[12] D. So and R. Cheng, "Detection techniques for V-BLAST in frequency selective channels," in *Proc. Wireless Communications and Networking Conf.*, 2002, pp. 487–491.

[13] A. Lozano and C. Papadias, "Layered space time receivers for frequency selective wireless channels," *IEEE Trans. Commun.*, vol. 50, pp. 65–73, Jan. 2002.

[14] G. J. Foschini, "Some basic layered space time architectures and their performance," *IEEE J. Select. Areas Commun.*—Special Issue on MIMO Systems, to be published.

[15] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Fundamental capacity of MIMO channels," *IEEE J. Select. Areas Commun.*—Special Issue on MIMO Systems, to be published.

[16] R. Blum, J. Winters, and N. Sollenberger, "On the capacity of cellular systems with MIMO," in *Proc. IEEE Vehicular Technology Conf.*, Atlantic City, NJ, Oct. 2001.

[17] R. Blum, "MIMO capacity with interference," IEEE J. Select. Areas Commun.—Special Issue on MIMO Systems, 2003, to be published.

[18] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Attainable throughput of an interference-limited multiple-input multiple-output cellular system," *IEEE Trans. Commun.*, vol. 48, pp. 1307–1311, Aug. 2001.

[19] H. Dai, A. Molisch, and H. V. Poor, "Downlink capacity of interference limited MIMO systems with joint detection," *IEEE Trans. Wireless Commun.*, submitted for publication.

[20] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.

[21] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Eur. Trans. Commun.*, vol. 10, no. 6, pp. 585–595, 1999.

[22] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *J. Multivariate Anal.*, vol. 55, no. 2, pp. 331–339, 1995.

[23] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Commun.*, vol. 48, pp. 1245–1248, Aug. 2000.

[24] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela, "Lifting the limits on high-speed wireless data access using antenna arrays," *IEEE Commun. Mag.*, vol. 39, pp. 156–162, Sept. 2001.

[25] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple-input multiple-output wireless channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, pp. 173–176, Feb. 1999.

[26] D. Shiu, *Wireless Communication Using Dual Antenna Arrays*, ser. International Series in Engineering and Computer Science. Norwell, MA: Kluwer, 1999.

[27] A. M. Sengupta and P. P. Mitra, "Capacity of multivariate channels with multiplicative noise: I. Random matrix techniques and large-$n$ expansions for full transfer matrices," *Phy. Arch.*, no. 0 010 081, 2000.

[28] N. Chiurtu, B. Rimoldi, and E. Telatar, "Dense multiple antenna systems," in *Proc. IEEE Information Theory Workshop*, 2001, pp. 108–109.

[29] D. Gesbert, N. Christophersen, and T. Ekman, "Capacity limits of dense palm-sized MIMO arrays," in *Proc. Globecom Conf.*, 2002.

[30] S. Wei, D. Goeckel, and R. Janaswami, "On the capacity of fixed length linear arrays under bandlimited correlated fading," in *Proc. CISS*, Princeton, NJ, Apr. 2002.

[31] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1989.

[32] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *J. Multivariate Analysis*, vol. 12, pp. 1–38, 1982.

[33] V. L. Girko, *Theory of Random Determinants*. Norwell, MA: Kluwer, 1990.

[34] ——, *Theory of Linear Algebraic Equations With Random Coefficients*. New York: Allerton, 1996.

[35] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.

[36] P. J. Smith and M. Shafi, "Waterfilling methods for MIMO systems," in *Proc. 3rd Australian Communication Theory Workshop*, Canberra, Australia, 2002, AusCTW 2002.

[37] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space–time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, pp. 85–87, 2001.

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

[39] C. N. Chuah, D. Tse, J. M. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.

[40] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *Proc. Int. Conf. Communications, ICC 2002*, 2002, pp. 406–410.

[41] "A standardized set of MIMO radio propagation channels," Lucent, Nokia, Siemens, Ericsson, Jeju, Korea, 3GPP TSG-RAN WG1 23, Nov. 19–23, 2001.

[42] C.-N. Chuah, J. M. Kahn, and D. Tse, "Capacity of indoor multiantenna array systems in indoor wireless environment," in *Proc. GLOBECOM, 98*, Sydney, Australia, 1998, pp. 1894–1899.

[43] O. Oyman, R. Nabar, H. Bolcskei, and A. Paulraj, "Lower bounds on the ergodic capacity of Rayleigh fading MIMO channels," in *Proc. IEEE GLOBECOM*, Taiwan, 2002.

[44] A. Grant, "Rayleigh fading multiple antenna channels," *EURASIP J. Appl. Signal Processing*, pp. 316–329, Mar. 2002.

[45] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[46] M. Stoytchev, H. F. Safar, A. L. Moustakas, and S. H. Simon, "Compact antenna arrays for MIMO applications," in *Proc. IEEE Int. Symp. Antennas and Propagation*, 2001, pp. 708–711.

[47] D. Chizhik, F. Rashid-Farrokhi, F. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Commun. Lett.*, vol. 4, pp. 337–339, Nov. 2000.

[48] L. Hanlen and M. Fu, "Multiple antenna wireless communication systems: Capacity limits for sparse scattering," in *Proc. 3rd Australian Communication Theory Workshop, AusCTW 2002*, Canberra, Australia, 2002.

[49] H. Ge, K. D. Wong, and J. C. Liberti, "Characterization of multiple-input multiple-output (MIMO) channel capacity," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Orlando, FL, 2002.

[50] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2172–2178, Nov. 2000.

[51] N. Seshadri and J. H. Winters, "Two schemes for improving the performance of frequency-division duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Information Networks*, vol. 1, pp. 49–60, Jan. 1994.

[52] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE ICC'93*, vol. 3, Geneva, Switzerland, 1993, pp. 1630–1634.

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

[54] A. Naguib, N. Seshadri, and R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Mag.*, vol. 17, pp. 76–92, May 2000.

[55] M. P. Fitz and J. V. Krogmeier, "Further results on space–time codes for Rayleigh fading," in *Proc. Allerton*, Sept. 1998, pp. 391–400.

[56] Q. Yan and R. S. Blum, "Optimum space–time convolutional codes for quasi-static slow fading channels," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, Sept. 2000, pp. 1351–1355.

[57] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space–time trellis-coded modulation," *IEEE Commun. Lett.*, vol. 1, pp. 20–22, Jan. 2000.

[58] A. R. Hammons and H. E. Gamal, "On the theory of space–time codes for psk modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 524–542, Mar. 2000.

[59] Z. Chen, J. Yuan, and B. Vucetic, "Improved space–time trellis coded modulation scheme on slow fading channels," in *Proc. ISIT*, 2001.

[60] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "A rank criterion for QAM space–time codes," in *Proc. IEEE Int. Symp. Information Theory*, Dec. 2000, pp. 3062–3079.

[61] S. Alamouti, "Space block coding: A simple transmitter diversity technique for wireless communications," *IEEE J. Select. Areas. Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

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

[63] S. Siwamogsatam and M. Fitz, "Improved high rate space–time TCM via orthogonality and set partitioning," in *Proc. Int. Symp. Wireless Personal Multimedia Communications*, Alborg, Denmark, Sept. 2001.

[64] S. Siwamogsatam and M. Fitz, "Improved high rate space–time TCM via concatenation of expanded orthogonal block codes and MTCM," in *Proc. IEEE Int. Conf. Communication*, New York, Apr. 2002.

[65] S. Siwamogsatam and M. Fitz, "Improved high rate space–time TCM from an expanded STB-MTCM construction," *IEEE Trans. Information Theory*, to be published.

[66] N. Seshadri and H. Jafarkhani, "Super-orthogonal space–time trellis codes," in *Proc. IEEE Int. Conf. Communication*, New York, Apr. 2002.

[67] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–460, Mar. 1999.

[68] G. Ganesan and P. Stoica, "Space–time diversity using orthogonal and amicable orthogonal designs," *Wireless Personal Commun.*, vol. 18, pp. 165–178, Aug. 2001.

[69] H. Jafarkhani, "A quasi orthogonal space time block code," *IEEE Trans. Commun.*, vol. 49, pp. 1–4, Jan. 2001.

[70] O. Tirkkonen, A. Boariu, and A. Hottinen, "Minimal nonorthogonality rate 1 space–time block code for 3+ tx antennas," in *Proc. IEEE Int. Symp. Spread Spectrum Technology*, 2000, pp. 429–432.

[71] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space–time coding based modem for high data rate wireless communication," *IEEE J. Select. Areas. Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.

[72] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh faded channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 683–693, Nov. 1991.

[73] S. Sampei and T. Sunaga, "Rayleigh fading compensation method for 16 QAM in digital land mobile radio channels," in *Proc. IEEE VTC'89*, vol. I, San Francisco, CA, May 1989, pp. 640–646.

[74] M. L. Moher and J. H. Lodge, "TCMP-A modulation and coding strategy for Ricean fading channels," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 1347–1355, Dec. 1989.

[75] R. J. Young, J. H. Lodge, and L. C. Pacola, "An implementation of a reference symbol approach to generic modulation in fading channels," in *Proc. Int. Mobile Satellite Conf.*, Ottawa, ON, Canada, June 1990, pp. 182–187.

[76] J. Yang and K. Feher, "A digital Rayleigh fade compensation technology for coherent OQPSK system," in *Proc. IEEE Vehicular Technology Conf.*, Orlando, FL, May 1990, pp. 732–737.

[77] C. L. Liu and K. Feher, "A new generation of Rayleigh fade compensated $\pi/4$-QPSK coherent modem," in *Proc. IEEE Vehicular Technology Conf.*, Orlando, FL, May 1990, pp. 482–486.

[78] A. Aghamohammadi, H. Meyr, and G. Asheid, "A new method for phase synchronization and automatic gain control of linearly modulated signals on frequency-flat fading channel," *IEEE Trans. Commun.*, vol. 39, pp. 25–29, Jan. 1991.

[79] V. Tarokh, S. M. Alamouti, and P. Poon, "New detection scheme for transmit diversity with no channel estimation," in *Proc. Int. Conf. Universal Personal Communications, ICUPC '98*, Oct. 1998, pp. 917–920.

[80] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 3, pp. 1043–1047, July 2000.

[81] B. L. Hughes, "Differential space–time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 145–149, Nov. 2000.

[82] B. M. Hochwald and T. L. Marzetta, "Unitary space–time modulation for multiple antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[83] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldons, and R. Urbanke, "Systematic design of unitary spacetime constellation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1962–1973, Sept. 2000.

[84] J. Yuan and X. Shao, "New differential space time block coding schemes with two three and four transmit antennas," *IEEE J. Select. Areas Commun.*, submitted for publication.

[85] C. Fragouli, N. Al-Dhahir, and S. Diggavi, "Pre-filtered space–time $m$-bcjr equalizer for frequency selective channels," *IEEE Trans. Commun.*, vol. 50, pp. 742–753, May 2002.

[86] A. Naguib, "Equalization of transmit diversity space–time coded signals," in *Proc. IEEE Global Telecommunications Conf., 2000. GLOBECOM '00*, vol. 2, 2000, pp. 1077–1082.

[87] G. Bauch and A. F. Naguib, "Map equalization of space–time coded signals over frequency selective channels," in *Proc. IEEE Wireless Communication and Networking Conf. WCNC'99*, vol. 1, New Orleans, LA, Sept. 1999, pp. 261–265.

[88] A. Liu, G. B. Giannakis, A. Scaglione, and S. Barbarossa, "Decoding and equalization of unknown multipath channels based on block precoding and transmit diversity," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, 1999, pp. 1557–1561.

[89] H. Bolcskei and A. Paulraj, "Space-frequency codes for broadband fading channels," presented at the IEEE Int. Symp. Information Theory, Washington, DC, June 2001.

[90] N. Al-Dhahir, "Single-carrier frequency domain equalization for space–time block coded transmission over frequency selective fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 304–306, July 2001.

[91] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proc. IEEE ICC'2000*, New Orleans, LA, 2000, pp. 307–311.

[92] M. O. Damen, A. Chkeif, and J. C. Belfiore, "Lattice codes decoder for space–time codes," *IEEE Commun. Lett.*, vol. 4, pp. 161–163, May 2000.

[93] S. Verdu, "Multiuser detection," in *Advances in Statistical Signal Processing*, V. Poor, Ed. Greenwich, CT: JAI, 1993, pp. 369–409.

[94] P. Comon and P. Chevalier, "Source separation: Models, concepts, algorithms and performance," in *Unsupervised Adaptive Filtering*. ser. Adaptive and learning systems for communications signal processing and control, S. Haykin, Ed. New York: Wiley, 2000, vol. I, Blind Source Separation, pp. 191–236.

[95] J. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEEE Proceedings*, pt. F, vol. 140, pp. 362–370, 1993.

[96] C. Papadias, "A multiuser Kurtosis algorithm for blind source separation," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2000, pp. 3144–3147.

[97] P. Loubaton, E. Moulines, and P. Regalia, "Subspace methods for blind identification and deconvolution," in *Signal Processing Advances in Wireless and Mobile Communications*, G. Giannakis, J. Hua, P. Stoica, and L. Tong, Eds. Englewood Cliffs, NJ: Prentice-Hall, 2001, ch. 3.

[98] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.

[99] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, submitted for publication.

[100] R. Heath and A. Paulraj, "Diversity versus multiplexing in narrowband MIMO channels: A tradeoff based on Euclidian distance," IEEE Trans. Commun., 2001, submitted for publication.

[101] B. Hassibi and B. Hochwald, "High rates codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[102] S. Sandhu and A. Paulraj, "Unified design of linear space–time block-codes," in *IEEE GLOBECOM Conf.*, vol. 2, 2001, pp. 1073–1077.

[103] M. O. Damen, A. Tewfik, and J. C. Belfiore, "A construction of a space–time code based on number theory," *IEEE Trans. Inform. Theory*, vol. 48, pp. 753–760, Mar. 2002.

[104] A. Molisch, M. Z. W. J. Winters, and A. Paulraj, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int. Conf. Communications*, 2001, pp. 570–574.

[105] D. Gore and A. Paulraj, "MIMO antenna subset selection with space–time coding," *IEEE Trans. Signal Processing*, vol. 50, pp. 2580–2588, Oct. 2002.

[106] R. W. Heath, S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, pp. 142–143, Apr. 2001.

[107] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[108] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted muse criterion," *IEEE Trans. Commun.*, pp. 2198–2206, Dec. 2001.

[109] J. Akhtar and D. Gesbert, "Partial feedback based space–time block coding," *IEEE Trans. Commun*., submitted for publication.

[110] S. Vishwanath, S. Jafar, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. ICC*, Helsinki, Finland, 2001, pp. 2266–2270.

[111] S. Simon and A. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Select. Areas Commun.*—Special Issue on MIMO Systems, to be published.

[112] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, Dec. 2002.

[113] L. J. Greenstein, S. Ghassemzadeh, V. Erceg, and D. G. Michelson, "Theory, experiments, and statistical models," in *Proc. WPMC 99 Conf.*, Amsterdam, Sept. 1999.

[114] D. S. Baum, D. A. Gore, R. U. Nabar, S. Panchanathan, K. V. S. Hari, V. Erceg, and A. J. Paulraj, "Measurements and characterization of broadband MIMO fixed wireless channels at 2.5 GHz," in *Proc. ICPWC 2000*, Hyderabad, Dec. 2000, pp. 203–206.

[115] S. Pitchaiah, V. Erceg, D. Baum, R. Krishnamoorthy, and A. Paulraj, "Modeling of multiple-input multiple-output (MIMO) radio channel based on outdoor measurements conducted at 2.5GHz for fixed BWA applications," in *Proc. Int. Conf. Communications*, 2002.

[116] E. Green, "Radio link design for microcellular systems," *BT Tech. J.*, vol. 8, no. 1, pp. 85–96, 1990.

[117] R. Stridh, B. Ottersten, and P. Karlsson, "MIMO channel capacity on a measured indoor radio channel at 5.8 GHz," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Oct. 2000, pp. 733–737.

[118] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[119] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," in *Proc. Inst. Elect. Eng., Radar, Sonar Navigation*, vol. 145, 1998, pp. 32–41.

[120] P. Driessen and G. J. Foschini, "On the capacity formula for multiple input-multiple output wireless channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, pp. 173–76, Feb. 1999.

[121] D. Chizhik, G. Foschini, and R. A. Valenzuela, "Capacities of multielement transmit and receive antennas: Correlations and keyholes," *Electron. Lett.*, pp. 1099–1100, 2000.

[122] C. C. Martin, J. Winters, and N. Sollenberger, "Multiple input multiple output (MIMO) radio channel measurements," in *Proc. IEEE Vehicular Technology Conf.*, Boston, MA, 2000, pp. 418–421.

[123] J. Ling, D. Chizhik, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple transmitter multiple receiver capacity survey in Manhattan," *Electron. Lett.*, vol. 37, pp. 1041–1042, Aug. 2001.

[124] R. Buehrer, S. Arunachalam, K. Wu, and A. Tonello, "Spatial channel models and measurements for IMT-2000 systems," in *Proc. IEEE Vehicular Technology Conf.*, May 2001, pp. 342–346.

[125] J. P. Kermoal, P. E. Mogensen, S. H. Jensen, J. B. Andersen, F. Frederiksen, T. Sorensen, and K. Pedersen, "Experimental investigation of multipath richness for multielement transmit and receive antenna arrays," in *Proc. IEEE Vehicular Technology Conf.-Spring*, 2000, pp. 2004–2008.

[126] J. R.-M. L. Schumacher and L. Berger, "Recent advances in propagation characterization and multiple antenna processing in the 3GPP framework," presented at the URSI Radio Conf., Aug. 2002.

[127] S. Kozono, T. Tsuruhara, and M. Sakamoto, "Base station polarization diversity reception for mobile radio," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 301–306, 1984.

[128] A. M. D. Turkmani, A. A. Arowojolu, P. A. Jefford, and C. J. Kellett, "An experimental evaluation of the performance of two-branch space and polarization diversity schemes at 1800 MHz," *IEEE Trans. Veh. Technol.*, vol. VT-44, pp. 318–326, 1995.

[129] P. C. F. Eggers, J. Toftgård, and A. M. Oprea, "Antenna systems for base station diversity in urban small and micro cells," *IEEE J. Select. Areas Commun.*, pp. 1046–1057, 1993.

[130] R. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, and A. Paulraj, "Performance of multiantenna signaling techniques in the presence of polarization diversity," *IEEE Trans. Commun.*, pp. 2553–2562, Oct. 2002.

[131] M. Wennström and T. Svantesson, "An antenna solution for MIMO channels: The switched parasitic antenna," in *Proc. IEEE Symp. Personal Indoor and Mobile Radio Communication (PIMRC) 2001*, San Diego, CA, 2001, pp. 159–163.

[132] S. Catreux, D. Gesbert, V. Erceg, and R. Heath, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commun. Mag.*, vol. 40, pp. 108–115, June 2002.

[133] V. Erceg *et al.*, "Channel models for fixed wireless applications," IEEE, Tech. Rep., IEEE 802.16 Work Group, 2001.

[134] 3GPP, "Physical layer aspects of ultra high speed downlink packet access, release 4,", Tech. Rep. 3G TR25.848 v.4.0, 2001–2003.

[135] 3GPP, "Tx diversity solutions for multiple antennas release 5,", Tech. Rep. 3G TR 25.869 v 1.0.0, 2001–2006.

[136] 3GPP, "Multiple-input multiple output antenna processing for HSDPA,", Tech. Rep. 3GPP TR 25.876 v0.0.1, 2001–2011.

[137] D. Gesbert, L. Haumonte, H. Bolcskei, and A. Paulraj, "Technologies and performance for nonline-of-sight broadband wireless access networks," *IEEE Commun. Mag.*, vol. 40, pp. 86–95, Apr. 2002.

[138] 3GPP, "Practical aspects of multiple architectures for HSDPA,", Tech. Rep. TSGR1#16 (00)1219.

[139] 3GPP, "MIMO system integration and signalling in HSDPA,", Tech. Rep. TSGR1#19(01)0305, Feb. 27–Mar 2, 2001.

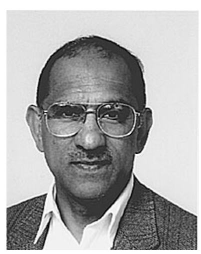
[140] "Guidelines for the evaluation of imt 2000, ITU recommendation," ITU, Tech. Rep. M 1225.

[141] 3GPP, "Alternatives in MIMO link design,", Tech. Rep. TSGR1#19 (01)0333, Feb. 27–Mar. 2, 2001.

**David Gesbert** (S'96–M'99) received the Ph.D. degree from Ecole Nationale Superieure des Telecommunications, Paris, France, in 1997.

From 1993 to 1997, he was with France Telecom Research, Paris, where he was involved in the development and study of signal processing algorithms for digital radio communications systems, with emphasis on blind array processing. In 1997 and 1998, he was a Postdoctoral Fellow at the Informations Systems Lab, Stanford University, Stanford, CA. From April 1997 to April 1998, he was the recipient of a French Defense DGA/DRET Postdoctoral Fellowship. In October 1998, he was a cofounder, then a Research Manager at Iospan Wireless, formerly Gigabit Wireless Inc., San Jose, CA, a startup company designing high-speed wireless internet access networks using smart antennas (MIMO), OFDM, and other state-of-the-art applied wireless signal processing research (now Intel). In 2001, he joined the Signal and Image Processing Group, Department of Informatics at the University of Oslo, Oslo, Norway, as an Associate Professor in parallel to his other activities. He has published about 50 conference and journal papers and several patents all in the area of signal processing and communications. His research interests are in the area of signal processing for high-speed digital communications, signal detection, smart antennas and MIMO, multiuser communications, cellular optimization, and resource allocation algorithms.

Dr. Gesbert has served on the Technical Program Committee of various IEEE conferences. He is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



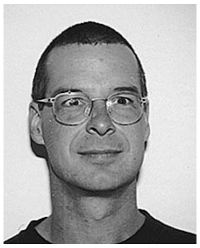
**Mansoor Shafi** (S'69–A'70–M'82–SM'87–F'93) received the B.Sc. degree in electrical engineering from Engineering University, Lahore, Pakistan, and the Ph.D. degree from the University of Auckland, Auckland, New Zealand, in 1970 and 1979, respectively.

From 1975 to 1979, he was a Junior Lecturer at the University of Auckland. Since 1979, he has been with Telecom New Zealand, where he holds the position of Principal Advisor Wireless Systems. In 1980, he held a Postdoctoral Fellowship at MacMaster University, Hamilton, ON, Canada. His research interests are in wireless communication systems. He has published widely in many aspects of wireless communications covering radio propagation, signal processing, cellular systems optimization, MIMO systems, etc. He serves as a New Zealand delegate to the meetings of ITU-R that are concerned with the standardization of IMT-2000 systems and beyond.

**Da-Shan Shiu** (S'95–M'99) received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1993 and the Ph.D. degree in electrical engineering and computer sciences from the University of California at Berkeley in 1999.

His current research interests include wireless optical communications, space–time signal processing, smart antenna, and spread-spectrum communications. He is currently with Qualcomm Inc., Campbell, CA. His current research interests include wireless optical communications, space–time signal processing, smart antenna, and spread-spectrum communications.

Dr. Shiu is a member of Phi Tau Phi.



**Peter J. Smith** (M'93) received the B.Sc. degree in mathematics and the Ph.D. degree in statistics from the University of London, London, U.K., in 1983 and 1988, respectively.

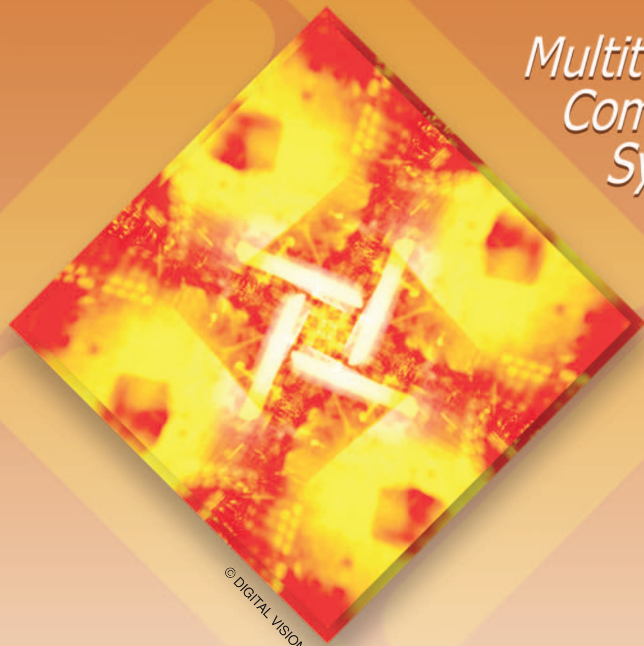
From 1983 to 1986, he was with the Telecommunications Laboratories, GEC Hirst Research Centre, Wembley, U.K. From 1988 to 2001, he was a Lecturer in statistics and Consulting Statistician at Victoria University of Wellington, Wellington, New Zealand. He is currently a Senior Lecturer in the Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand. His research interests include the statistical aspects of design and analysis for communication systems, especially MIMO systems, antenna arrays, and mobile radio.

**Ayman Naguib** (S'91–M'96–SM'00) received the B.Sc. degree (with honors) and the M.S. degree in electrical engineering from Cairo University, Cairo, Egypt, in 1987 and 1990, respectively, and the M.S. degree in statistics and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1993 and 1996, respectively.

From 1987 to 1989, he spent his military service at the Signal Processing Laboratory, The Military Technical College, Cairo. From 1989 to 1990, he was a Research and Teaching Assistant in the Communication Theory Group, Cairo University. From 1990 to 1995, he was a Research and Teaching Assistant in the Information Systems Laboratory, Stanford University. In 1996, he joined AT&T Labs, Florham Park, NJ, as a principal member of technical staff, where he was a Member of the research team in AT&T Labs that pioneered the field space–time coding. In September 2000, he joined Morphics Technology, Inc., as a Technical Leader and Manager of Core Technology. In October 2002, he joined Qualcomm, Inc., Campbell, CA, where he is a Senior Staff Engineer working future wireless systems. His current research interests include space–time coding, signal processing, and coding for wireless communications. He holds five U.S. patents and seven other pending patent applications in the area of space–time coding and signal processing.

Dr. Naguib is currently serving as an Associate Editor for CDMA and space–time systems, IEEE TRANSACTIONS ON COMMUNICATIONS.

Multiterminal Communication Systems





[David Gesbert, Marios Kountouris, Robert W. Heath Jr., Chan-Byoung Chae, and Thomas Sälzer]

# Shifting the MIMO Paradigm

[From single-user to multiuser communications]

The last ten years have witnessed the transition of multiple-input multiple-output (MIMO) communication from a theoretical concept to a practical technique for enhancing performance of wireless networks [1]. Point-to-point (single-user) MIMO communication promises large gains for both channel capacity and reliability, essentially via the use of space-time codes (diversity gain oriented) combined with stream multiplexed transmission (rate maximization oriented). In such a traditional single-user view of MIMO systems, the extra spatial degrees of freedom (DoF) brought by the use of multiple antennas are exploited to expand the dimensions available for signal processing and detection, thus acting mainly as a physical (PHY) layer performance booster. In this approach, the link layer protocols for multiple access (uplink and downlink) indirectly reap the performance benefits of MIMO antennas in the form of greater per-user rates or more reliable channel quality despite not requiring full awareness of the MIMO capability.

The situation with multiuser MIMO (MU-MIMO) techniques is radically different as these techniques imply the use of spatial sharing of the channel by the users, thus deeply affecting the design of the multiple access protocol. In spatial multiple access, the resulting multiuser interference is handled by the multiple antennas, which, in addition to providing per-link diversity, also give the DoF necessary for spatial separation of the users (see e.g. [1] Part IV). In practice, MU-MIMO schemes with good complexity/performance tradeoffs can be implemented to realize these ideas. On the uplink or multiple access channel (MAC), the development of MU-MIMO techniques appears as a generalization of known single-user

*Digital Object Identifier 10.1109/MSP.2007.904815*



1053-5888/07/$25.00©2007IEEE

MIMO (SU-MIMO) concepts to the multiuser case. As usual in information theory, the downlink or broadcast channel (BC) case is by far the most challenging one. Information theory reveals that the optimum transmit strategy for the MU-MIMO BC involves a theoretical preinterference cancellation technique known as dirty paper coding (DPC) combined with an implicit user scheduling and power loading algorithm. In that respect, the role played by seminal papers such as [2] was fundamental. In turn, several practical strategies have recently been proposed to approach the rates promised in the MU-MIMO channel involving concepts such as linear and nonlinear channel-aware precoding, channel state feedback, and multiuser receivers. A number of corresponding scheduling and user selection algorithms have also been proposed, leveraging features of different MU-MIMO strategies.

MU-MIMO techniques and performance have begun to be intensely investigated because of several key advantages over SU-MIMO communications.

- MU-MIMO schemes allow for a direct gain in multiple access capacity [proportional to the number of base station (BS) antennas] thanks to so-called multiuser multiplexing schemes.
- MU-MIMO appears more immune to most of propagation limitations plaguing SU-MIMO communications such as channel rank loss or antenna correlation. Although increased correlation still affects per-user diversity, this may not be a major issue if multiuser diversity [3] can be extracted by the scheduler instead. Additionally, line-of-sight propagation, which causes severe degradation in single-user spatial multiplexing schemes, is no longer a problem in multiuser setting.
- MU-MIMO allows the spatial multiplexing gain at the BS to be obtained without the need for multiple antenna terminals, thereby allowing the development of small and cheap terminals while intelligence and cost is kept on the infrastructure side.

The advantages above unfortunately come at a price. Perhaps the most substantial cost is due to the fact that MU-MIMO requires (although benefits from) channel state information at transmitter (CSIT) to properly serve the spatially multiplexed users. CSIT, while not essential in SU-MIMO communication channels, is of critical importance to most downlink multiuser precoding techniques. The need for CSIT feedback places a significant burden on uplink capacity in most systems, exacerbated in systems with wideband [e.g. orthogonal frequency division multiplexing (OFDM)] communication or high mobility (such as 3GPP-LTE [4], WiMax [5], etc.). Finally, another challenge related to MU-MIMO cross-layer design lies in the complexity of the scheduling procedure associated with the selection of a group of users that will be served simultaneously. Optimal scheduling involves exhaustive search whose complexity is exponential in the group size and depends on the choice of precoding, decoding, and channel state feedback technique.

Inspection of recent literature reveals several different schools of thought on the MU-MIMO downlink, each advocating a different combination of precoding, feedback, and scheduling strategies. Precoding strategies include linear minimum mean square error (MMSE) or zero-forcing (ZF) techniques and nonlinear approaches. Examples of the latter are vector perturbation, DPC techniques, and Tomlinson-Harashima precoding (THP) (a number of references are listed below). Many different feedback strategies have been suggested, including vector quantization, dimension reduction, adaptive feedback, statistical feedback, and opportunistic spatial division multiple access (SDMA). Finally, a number of scheduling disciplines have been suggested, including max-rate techniques, greedy user selection, and random user selection.

## PROMISES AND CHALLENGES OF MU-MIMO NETWORKS

### *LESSONS LEARNED FROM MULTIUSER INFORMATION THEORY*

#### SYSTEM AND SIGNAL MODEL

Progress in the field of multiuser information theory has been instrumental in understanding the fundamental nature and limits of the gains associated with exploiting multiple antennas in wireless networks, often also suggesting ideas for actual algorithms. We now review some aspects of MU-MIMO information theory with an eye for the key lessons learned from this field towards practical system design. A complete study of MU-MIMO information theoretic progress is beyond the scope of this article. Good references on the topic include [6] and [1, Ch. 18 and 19].

We focus on the communication between a BS or an access point equipped with $N$ antennas, and $U$ active terminals, where each active user $k$ is equipped with $M_k$ antennas. Among all terminals, the set of active users is roughly defined by the set of users simultaneously downloading or uploading packets during one given scheduling window. The length of the window is arbitrary but should not exceed the maximum latency expected by the application (likely as small as a few tens of milliseconds to several hundred milliseconds). By all means the active users over one given window will be a small subset of the connected users, themselves forming a small subset of the subscribers. We consider both the uplink and downlink but will emphasize on the challenges associated with the downlink for several reasons explained later.

In the uplink, the received signal at the BS can be written as

$$\mathbf{y} = \sum_{k=1}^{U} \mathbf{H}_k^T \mathbf{x}_k + \mathbf{n}, \tag{1}$$

where $\mathbf{x}_k$ is the $M_k \times 1$ user signal vector, possibly encompassing power-controlled, linearly combined, constellation symbols. $\mathbf{H}_k \in \mathbb{C}^{M_k \times N}$ represents the flat-fading channel matrix and $\mathbf{n}$ is the independent and identically distributed (i.i.d.), unit-variance, additive Gaussian noise vector at the BS. We assume that the receiver $k$ has perfect and instantaneous knowledge of the channel $\mathbf{H}_k$. We focus on the flat-fading model here for the sake of exposition. Wideband models, using OFDM for example, can be accommodated by using a dependency on a frequency index. The transpose operator is simply used by convention for consistence with the downlink notation and does not presume a reciprocal link.

In the downlink illustrated in Figure 1, the received signal at the $k$th receiver can be written as

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x} + \mathbf{n}_k \qquad \text{for} \qquad k = 1, \ldots, U, \tag{2}$$

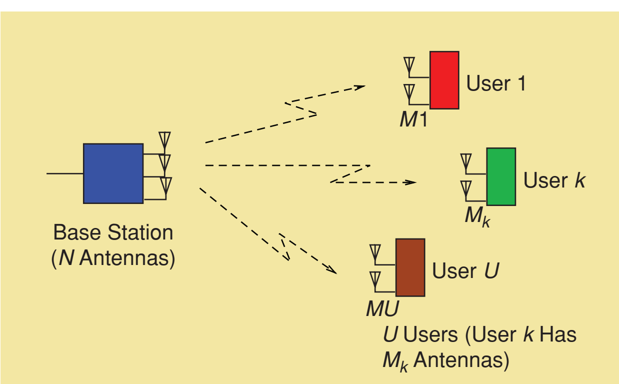
where $\mathbf{H}_k \in \mathbb{C}^{M_k \times N}$ represents the downlink channel and $\mathbf{n}_k \in \mathbb{C}^{M_k \times 1}$ is the additive Gaussian noise at receiver $k$. We assume that each receiver also has perfect and instantaneous knowledge of its own channel $\mathbf{H}_k$. The transmitted signal $\mathbf{x}$ is a function of the multiple users' information data, an example of which takes the superposition form $\mathbf{x} = \sum_k \mathbf{x}_k$ where $\mathbf{x}_k$ is the signal carrying, possibly nonlinearly encoded, user $k$'s message, with covariance $\mathbf{Q}_k = \mathbb{E}(\mathbf{x}_k\mathbf{x}_k^H)$, with $\mathbb{E}(\cdot)$ the expectation operator. The power allocated to user $k$ is therefore given by $P_k = \mathrm{Tr}(\mathbf{Q}_k)$, where Tr is the trace operator. Under a sum power constraint at the BS, the power allocation needs to maintain $\sum_k P_k \leq P$.

Assuming a unit variance for the noise, it is now known that the capacity region for a given matrix channel realization can be written as [7]:

$$\mathcal{C}_{BC} = \bigcup_{P_1,..P_U \; s.t. \; \sum_k P_k = P} \left\{ (R_1, ..R_U) \in \Re^{+U}, R_i \leq \log_2 \frac{\det\left[\mathbf{I} + \mathbf{H}_i(\sum_{j\geq i}\mathbf{Q}_j)\mathbf{H}_i^H\right]}{\det\left[\mathbf{I} + \mathbf{H}_i(\sum_{j>i}\mathbf{Q}_j)\mathbf{H}_i^H\right]} \right\}, \tag{3}$$

where the expression should in turn be optimized over each possible user ordering. Although difficult to realize in practice, the computation of the region above is facilitated by exploiting the so-called duality results between the BC and the much simple to obtain MAC capacity region, which stipulate that the BC region can be calculated through the union of regions of the dual MAC with all uplink power allocation vectors meeting the sum power constraint $P$ [8], [9].

The fundamental role played by the multiple antennas at either the BS or the users in expanding the channel capacity is best apprehended by examining how the sum rate (the point yielded by the maximum $\sum_k R_k$ in the capacity region) scales with the number of active users.




**[FIG1] Downlink of a multiuser MIMO network. A BS communicates simultaneously with several multiple antenna terminals.**

Assuming a block fading channel model and an homogeneous network where all users have the same signal-to-noise ratio (SNR), the scaling law of the sum rate capacity of MIMO Gaussian BC, denoted as $\mathcal{R}^{DPC}$ for $M_k = M$, fixed $N$ and $P$, and large $U$ is given by [10]

$$\lim_{U\to\infty} \frac{\mathbb{E}(\mathcal{R}^{DPC})}{N \log\log(UM)} = 1. \tag{4}$$

The result in (4) indicates that, with full CSIT, the system can enjoy a multiplexing gain of $N$, obtained by the BS sending data to $N$ carefully selected users out of $U$. Since each user exhibits $M$ independent fading coefficients, the total number of DoF for multiuser diversity is $UM$, thus giving the extra gain $\log\log(UM)$.

In contrast with (4), the capacity obtained in a situation where the BS is deprived from the users' channel information is reduced to (in the high SNR regime)

$$\mathbb{E}(\mathcal{R}^{NoCSIT}) \approx \min(M, N) \log SNR. \tag{5}$$

## DESIGN LESSONS

Information theory highlights several fundamental aspects of MU-MIMO systems, which are in contrast much with the conventional SU-MIMO setting. First, the results above advocate for serving multiple users simultaneously in a SDMA fashion, with a suitably chosen precoding scheme at the transmitter. Although the multiplexing gain is limited by the number of transmit antennas, the number of simultaneously served users is, in principle, arbitrary. How many and which users should effectively be served with nonzero power at any given instant is the problem addressed by the resource allocation algorithm. Unlike in the single-user setting, the spatial multiplexing of different data streams can be done while users are equipped with single antenna receivers, thus enabling the capacity gains of MIMO while maintaining a low cost for user terminals. Having multiple antennas at the terminal can thus be viewed as optional equipment allowing extra diversity gain for certain users or giving the flexibility toward interference canceling and multiplexing of several data streams to such users (but reducing the number of other users served simultaneously). In addition to yielding MIMO multiplexing gains without the need for MIMO user terminals, the multiuser setup presents the advantage of being immune with respect to the possible ill-behavior of the propagation channel, which often plagues SU-MIMO communications, i.e., rank loss due to small spacing and/or the presence of strong line-of-sight component thanks to the wide physical separation between the users.

Finally, also in contrast with the conventional SU-MIMO setting, the multiplexing factor $N$ in the downlink comes at the condition of channel knowledge at the transmitter. In the uplink this multiplexing gain is more easily extracted because the BS can be safely assumed to have uplink channel knowledge and simply implements a classical multiuser receiver to separate the contributions of the selected users in (1).

In the downlink, in the absence of CSIT, user multiplexing is generally not possible, as the BS just does not know in which direction to form spatial beams. Thus, the complete lack of channel state information (CSI) knowledge reduces the multiplexing gain to unity [11]. The exception lies in scenarios with terminal devices having enough antennas to remove costream interference at the receiver ($M_k \geq N$). In the latter case, the base may decide to either multiplex several streams to a single user or spread the streams over multiple users, achieving an equivalent multiplexing gain in both cases. This is conditioned however on the individual user channels to be full rank. Hence, the advantage of having CSIT in MU-MIMO lies in the possibility of not only serving single antenna users but also relaxing the dependence on single-user channel full rank.

**THE ADVANTAGES OF MU-MIMO TECHNIQUES AND PERFORMANCE OVER SU-MIMO COMMUNICATIONS UNFORTUNATELY COME AT A PRICE.**

### MU-MIMO AND RESOURCE ALLOCATION

One of the fundamental lessons learned from information theoretic studies is that resource allocation techniques help to exploit the gains of MU-MIMO systems. From a multiuser information theoretic perspective, the capacity region boundary is achieved by serving all $U$ active users simultaneously, where $U$ is possibly a large number. The resource that should be allocated to each one, in the form of, e.g., $P_k$, is surely dependent on the instantaneous channel conditions and may vary greatly from user to user. The fact that the multiplexing gain is limited to $N$ also suggests that the number of users effectively served with nonzero $P_k$ at any given instant of time is directly related to the number of antennas at the BS, which is considerably less than the number of active cell users. Studies show in fact that the optimal number of users with nonzero allocated power for any given realization of the channel is upper bounded by $N^2$ [12]. In the remainder of the article we shall refer to this subset of users as the selected users. When restricting to linear precoding techniques such as ZF, the number of served users is directly limited by the number of DoF at the BS, $N$. This motivates the need to pick a good set of users, which is the aim of the resource allocation algorithm. In particular, the scheduler selects among all possible active users, for each channel realization, an optimal subgroup of terminals and respective power levels within the subgroup, so as to maximize a given performance metric. Such a metric can be the sum rate or the realization of per-user rate targets while minimizing transmit power. To maximize the sum rate, the scheduler algorithm looks for users that exhibit a compromise between a high level of instantaneous SNR (to maximize multiuser diversity [3]) and a good separability of their spatial signatures to facilitate user multiplexing. Practical and low complexity algorithms to solve the user scheduling problem are presented later in this article.

## MU-MIMO SCHEMES WITH PERFECT CHANNEL KNOWLEDGE AT THE TRANSMITTER

### *LINEAR PRECODING*

Linear precoding is a generalization of traditional SDMA, where users are assigned different precoding matrices at the transmitter. The precoders are designed jointly based on CSI of all the users based on any number of designs, including ZF and MMSE.

From a practical point of view, the relevant criteria are error probability, sum rate, signal-to-interference-plus-noise ratio (SINR), etc. The difficulty of designing capacity-optimal downlink precoding, mainly due to the coupling between power and beamforming and the user ordering, has lead to several different approaches ranging from transmit power minimization while maintaining individual SINR constraints to worst case SINR maximization under a power constraint. Duality and iterative algorithms are often used to provide solutions [13].

Consider the transmitted signal for user $k$ given by $\mathbf{W}_k\mathbf{s}_k$, where $\mathbf{W}_k$ denotes the precoding matrix for the $k$th user and $\mathbf{s}_k$ is the symbol vector. We assume that service will be provided to a set of $K$ selected users (among all active ones). Scheduling algorithms as discussed in the sequel can be applied to perform this selection across possible subsets. The received signal vector at the $k$th user is

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{W}_k\mathbf{s}_k + \mathbf{H}_k \sum_{l=1, l\neq k}^{K} \mathbf{W}_l\mathbf{s}_l + \mathbf{n}_k. \tag{6}$$

We assume that each user has $M_k$ antennas and will decode the $S_k \leq M_k$ streams that constitute its data. The goal of linear precoding is to design $\{\mathbf{W}_k\}_{k=1}^{K}$ based on the channel matrix knowledge, so a given performance metric is maximized for each stream.

One of the simplest approaches for finding the precoder is to premultiply the transmitted signal by a suitably normalized ZF or MMSE inverse of the multiuser matrix channel [14], [15]. In this case, it can be assumed for simplification that $M_k = S_k = 1$. Thus, $\mathbf{H}_k = \mathbf{h}_k$ is a row vector and $\mathbf{W}_k$ (the precoding vector for the $k$th user) is chosen as the $k$th column of the right pseudoinverse (or MMSE inverse) of the composite channel $\left[\mathbf{h}_1^T, \mathbf{h}_2^T, \ldots, \mathbf{h}_K^T\right]^T$. In the case when the selected users are not sufficiently separable, this approach may result in inefficient use of transmit power, causing a large rate loss with respect to the optimum sum capacity solution. This problem, however, is shown to be fixed by the scheduler when the number of active users to choose from is large enough so near-orthogonal users with good SNR conditions can be found. An additional disadvantage is that this approach does not readily extend to multiple receive antennas or streams without further degradation.

A generalization of the ZF or MMSE beamforming is to combine linear beamforming with a suitable power control policy to maximize the sum rate or realize individual SINR requirements for each user. Several approaches have been proposed, including maximizing the jointly achievable SINR margin under a total power constraint and minimizing the total transmission power while satisfying a set of SINR constraints [13]. Another generalization of ZF beamforming (ZFBF) is provided by block diagonalization (BD), which assumes $M_k = S_k \geq 1$ and $\sum_{k=1}^{K} M_k = N$. The idea is to choose $\mathbf{W}_k$ such that $\mathbf{H}_l\mathbf{W}_k = 0$, $\forall l \neq k$,thus precanceling the interference in (6) so that $\mathbf{y}_k = \mathbf{H}_k\mathbf{W}_k\mathbf{s}_k + \mathbf{n}_k$. If we define $\tilde{\mathbf{H}}_k$ as

$$\tilde{\mathbf{H}}_k = \left[\mathbf{H}_1^T \cdots \mathbf{H}_{k-1}^T \mathbf{H}_{k+1}^T \cdots \mathbf{H}_K^T\right]^T, \tag{7}$$

then any suitable $\mathbf{W}_k$ lies in the null space of $\tilde{\mathbf{H}}_k$. Let the singular value decomposition (SVD) of $\tilde{\mathbf{H}}_k$ be $\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k\tilde{\mathbf{D}}_k \left[\tilde{\mathbf{V}}_k^{(1)}\tilde{\mathbf{V}}_k^{(0)}\right]^H$, where $\tilde{\mathbf{U}}_k$ and $\tilde{\mathbf{D}}_k$ are the left singular vector matrix and the matrix of singular values of $\tilde{\mathbf{H}}_k$, respectively, and $\tilde{\mathbf{V}}_k^{(1)}$ and $\tilde{\mathbf{V}}_k^{(0)}$ denote the right singular matrices, each corresponding to nonzero singular values and zero singular values, respectively. Any precoder $\mathbf{W_k}$ that is a linear combination of the columns of $\tilde{\mathbf{V}}_k^{(0)}$ will satisfy the null constraint. Assuming that $\tilde{\mathbf{H}}_k$ is full rank, the transmitter requires that the number of transmit antennas is at least the sum of all users' receive antennas to satisfy the dimensionality constraint required to cancel interference for each user [16]. Under the BD constraint, $\mathbf{W}_k$ can be further optimized based on waterfilling. If excess antennas are available, eigenmode selection or antenna subset selection can be used to further improve performance [17].

A disadvantage of BD is that it requires $M_k = S_k$. This can be solved by including the receive processing in the problem formulation. For example, with a linear receive combining matrix $\mathbf{V}_k$ for user $k$, the received signal can be expressed as

$$\mathbf{y}_k = \mathbf{V}_k^H\mathbf{H}_k\mathbf{W}_k\mathbf{s}_k + \mathbf{V}_k^H\mathbf{H}_k \sum_{l=1,l\neq k}^{K} \mathbf{W}_l\mathbf{s}_l + \mathbf{V}_k^H\mathbf{n}_k. \tag{8}$$

The design problem then becomes selecting $\{\mathbf{W}_k, \mathbf{V}_k\}_{k=1}^K$ jointly such that $\mathbf{V}_k^H\mathbf{H}_k \sum_{l=1,l\neq k}^{K} \mathbf{W}_l = \mathbf{0},\ \forall k$. This is difficult to solve in closed form, thus several iterative solutions have been proposed, including, e.g., [18], [19]. In such approaches, the transmitter generally computes a new effective channel for each user $k$ using the initial receive combining vector. Using this new effective channel, the transmitter recomputes the transmit filter $\mathbf{W}_k$ to enforce a zero interference condition, and the receive filter $\mathbf{V}_k$ for each user. The algorithm repeats this process until satisfying a convergence criterion. To extend this algorithm to multiple data streams for each user, the matrix of right singular vectors is used based on the number of data streams and is used to calculate the effective channel matrix [18]–[20]. To avoid the use of extra feedback between the users and the BS, the computation of all filters (transmit and receive) normally takes place at the BS. After this computation, either the users must acquire the effective combined channel or information about the transmit filters must be sent [19].

### *NONLINEAR PRECODING*

Linear precoding provides reasonable performance but may remain far from DPC-like precoding strategies when the available set of active users to choose from is small. Nonlinear precoding involves additional transmit signal processing to improve error rate performance. In this section, we discuss two representative methods, one based on perturbation [21], the other based on a spatial extension of THP [22].

Vector perturbation uses a modulo operation at the transmitter to perturb the transmitted signal vector to avoid the transmit power enhancement incurred by ZF methods [21]. Finding the optimal perturbation involves solving a minimum distance type problem and thus can be implemented using sphere encoding or full search-based algorithms.

Let $\mathbf{H}$ denote a $K \times N$ multiuser composite channel, assuming each user has a single receive antenna. The idea of perturbation is to find a perturbing vector $\mathbf{p}$ from an extended constellation to minimize the transmit power. The perturbation $\mathbf{p}$ is found by solving

$$\mathbf{p} = \arg \min_{\mathbf{p}' \in A\mathcal{C}\mathbb{Z}^K} \|\mathbf{G}(\mathbf{s} + \mathbf{p}')\|^2, \tag{9}$$

where $\mathbf{G}$ is a some transmit matrix such that $\mathrm{Tr}(\mathbf{G}^H\mathbf{G}) \leq P$, $\mathbf{s}$ is a modulated transmitted signal vector, and the scalar $A$ is chosen depending on the original constellation size (e.g., $A = 2$ for QPSK), and $\mathcal{C}\mathbb{Z}^K$ is the $K$-dimensional complex lattice. ZF or MMSE precoder can be used for the transmit matrix $\mathbf{G}$. A set of points is used to represent symbols that are congruent to the symbol in the fundamental region. After predistortion using ZF or MMSE precoder, the resulting constellation region also becomes distorted and thus it takes more power to transmit the original point than before distortion. Among the equivalent points, if the transmitter sends the point that is the one closest to the origin to minimize transmit power, the receiver finds its equivalent image inside the fundamental constellation region using a modulo operation. This problem can be regarded as $K$-dimensional integer-lattice least squares problem and thus search based algorithms can be implemented. There are other methods to simplify the search based methods [23].

Several algorithms have also been proposed based on variations of THP [22], [24]. THP was originally proposed for use with $Z$ point one-dimension pulse amplitude modulation (PAM) signal as a temporal equalization. For this constellation, THP is the same as the inverse channel filter except that an offset-free modulo $2Z$ adder is used. If the result of the summation is greater than $Z$, $2Z$ is subtracted until the final result is smaller than $Z$. Similarly, if the result of the summation is less than $-Z$, $2Z$ is added until satisfying the peak constraint. While in the original THP, a single channel is equalized with respect to time, spatial equalization is required for MIMO channels.

So far, we reviewed linear and nonlinear MU-MIMO solutions to approximate the sum capacity. In Figure 2, we compare sum capacity and achievable sum rates for DPC, coordinated beamforming [19], time sharing single-user closed loop MIMO (choosing only one user having the best channel quality and applying the SVD), and ZFBF with the dimensionality constraint [25]. In this case, no scheduling algorithm is required for DPC, coordinated beamforming, and ZFBF. We investigate scheduling issues below. Note that for the $(T, 1, T)$ scenario (i.e., the user has only one receive antenna while the BS has $T$ transmit antennas and there are $T$ active users in the network), there is a big gap between DPC and ZFBF, but this gap is decreased when the receivers have multiple antennas. For additional tradeoff analysis between linear and nonlinear precoding strategies, see also [26].

RESOURCE ALLOCATION TECHNIQUES HELP TO EXPLOIT THE GAINS OF MU-MIMO SYSTEMS.

In the following section, we consider the problem of choosing a subset of users for transmission in the MIMO BC. A brute-force complete search over all possible combinations of users guarantees maximizing the throughput, but the computational complexity is prohibitive when the number of users is large. Due to the complexity of the search process, both optimal and suboptimal approaches are considered. A key idea for low-complexity multiuser scheduling is that of greedy search.

### *OPTIMAL SCHEDULING FOR THE MU-MIMO DOWNLINK*

The previous theoretical capacity results illustrate that, in general, the MIMO BC results in transmission to more than one user at a time. The problem of selecting a subset of users for transmission is a user scheduling problem, and the gain is achieved in a form of multiuser diversity. In this section we summarize some scheduling algorithms for different MU-MIMO solutions.

Linear beamforming can achieve the sum capacity when the number of active users in the system is large [10], [25], [27]. In [25], the users are equipped with only one receive antenna, and ZFBF is performed at the transmitter. Analogous to BD, this full search-based user selection algorithm can be extended to the multiple stream scenario. For simplicity, in this section, we assume that the number of receive antennas is equal to the number of data streams, where the postcoder V is not needed, and thus BD can be implemented.

Suppose $\mathcal{U} = \{1, 2, \cdots, U\}$ is the set of all users, and $\mathcal{A}_k$ one possible subset of selected users in $\mathcal{U}$. Let $\mathcal{A}$ be the set including all possible $\mathcal{A}_i$, i.e., $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2, \cdots\}$. Then total achievable rate with BD is given by

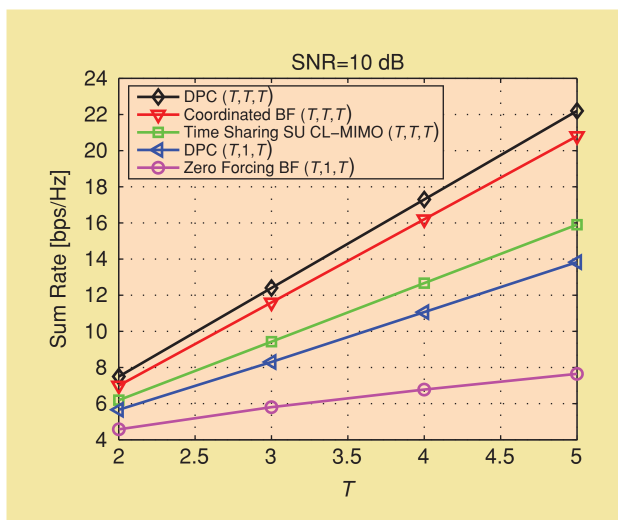
$$R_{BD|\mathcal{A}_k}(\mathrm{H}_{\mathcal{A}_k}, P, \sigma^2) = \max_{\sum_{j\in\mathcal{A}_k}\mathrm{Tr}(\mathrm{Q}_j)\le P} \sum_{j\in\mathcal{A}_k} \log\left|\mathrm{I} + \frac{\mathrm{H}_j \mathrm{W}_j \mathrm{Q}_j \mathrm{W}_j^H \mathrm{H}_j^H}{\sigma^2}\right|, \quad (10)$$

where $\mathrm{Q}_j = \mathbb{E}(\mathrm{x}_j \mathrm{x}_j)^H$ is the input covariance matrix for the user $j$, $\mathrm{W}_j$ is the precoding matrix earlier defined, and the same noise variance $\sigma^2$ is assumed at all users. Therefore, the maximum total sum rate with BD is given by $R_{BD}(\mathrm{H}_{1,\ldots,U}, P, \sigma^2) = \max_{\mathcal{A}_k\in\mathcal{A}} R_{BD|\mathcal{A}_k}(\mathrm{H}_{\mathcal{A}_k}, P, \sigma^2)$. Denote $\mathcal{S}$ as the maximum number of users to be supported. For the case of BD, $\mathcal{S} \le N$. Thus, the cardinality of $\mathcal{A}$ is $\sum_{i=1}^{\mathcal{S}} C_U^i$, where $C_a^b$ is the combination of $a$ choosing $b$. Hence, it is clear that the exhaustive search over all possible combinations is computationally prohibitive when the number of users in the system is increased, and thus low-complexity user selection algorithm is desired.

### *GREEDY AND ITERATIVE METHODS FOR USER GROUPING*

The complexity of the optimal scheduling is high, thus there has been several suboptimal algorithms that were proposed to reduce the computational complexity for user group selection [25], [27]–[29].

In the capacity-based greedy user selection algorithm, the transmitter chooses the first user with the highest channel capacity. Then, it finds the next user that provides the maximum sum rate from the remaining unselected users. The algorithm is repeated until $K$ users are selected. Clearly, the complexity of the capacity-based greedy user selection is no more than $U \times K$ user sets, which greatly reduces the complexity compared to the exhaustive search method explained in the previous section. Note that the full search method needs to consider roughly $\mathcal{O}(U^K)$ possible user sets. The sum rate can be obtained under a number of transmit schemes, including optimal nonlinear precoders. Scheduling for the nonlinear precoders mentioned previously is an ongoing topic of research, though few results have appeared, including a greedy user selection for ZF DPC (ZFDPC), which has been proposed in [27].




**[FIG2] Ergodic sum capacity and achievable sum rate as a function of the number of users, the number of transmit/receive antennas. ($T_1$, $T_2$, $T_3$) denotes the number of transmit antennas at the BS, the number of receive antennas at the user, and the number of active users in the network, respectively. Coordinated BF refers to the method presented in [19].**

## LIVING WITH PARTIAL CHANNEL KNOWLEDGE AT THE TRANSMITTER

### *QUANTIZATION-BASED TECHNIQUES*

Quantization is the first idea that comes to mind when dealing with source compression, in this case, the random channel matrix or the corresponding precoders being the possible sources. The amount of feedback depends on the frequency of feedback (generally a fraction of the coherence time), the number of parameters being quantized, and the resolution of the quantizer. Most research focuses on reducing the number of parameters and the required resolution. The feedback problem has been solved in SU-MIMO communication systems using a concept known as limited feedback precoding [30]. The key idea of this line of research has been to quantize the precoder for a MIMO channel and not simply the channel coefficients. The challenge of extending this work to the multiuser channel is that the transmit precoder depends on the channels of the other users in the system.

**CHANNEL QUALITY METRIC DESIGN IS ONE OF THE LARGELY OPEN CHALLENGES IN MU-MIMO.**

Other methods for reducing feedback in MU-MIMO channels assume a single receive antenna at the mobile—extensions to multiple receive antennas is an ongoing research topic. Some of the main results on this subject are due to [31], [32], where the random codebook and Grassmannian quantization ideas are used to quantize the direction of each user's channel $\mathbf{h}_k$. The main observation in [31] is that the feedback requirements scale linearly both as the number of transmit antennas grows and as a function of the SNR (in dB), unlike the single-user case. The reason is that quantization error introduces an SINR floor since it prohibits perfect inter-user cancellation. Thus, this error must diminish for higher SNRs to allow for a balancing between the noise and the residual interference due to channel quantizing. An improvement can be obtained by quantizing the channel vector and a certain received SINR upper bound that is a function of the error between the true and quantized channel [33]. This increases the performance of the system and helps in user selection. Thresholds based on sum rate constraints on the feedback channel can also be used to reduce required feedback yet maintain capacity scaling [34].

### *DIMENSION REDUCTION AND PROJECTION TECHNIQUES*

In addition to quantization-based approaches where the channel metric is discretized, dimension reduction techniques can be used that involve projecting the matrix channel onto one or more basis vectors known to the transmitter and receiver. In that way, the CSI matrix of size $M \times N$ is mapped into a $p$-dimensional vector with $1 \leq p \leq M \times N$, thus reducing the dimensionality of the CSI to $p$ complex scalars (which in turn may be quantized). Once the projection is carried out, the receiver feeds back a metric $\varphi_k = f(\mathbf{H}_k)$, which is typically related to the square magnitude of the projected signal. Antenna selection methods fall into this category. In this case, the projection is carried out by the terminal itself. Alternatively, the projection can be the result of using a particular precoder at the BS. A good example of this approach is given by a class of algorithms using unitary precoders. We now review this approach when $M_k = 1$ and the BS serves $N$ users. In this case, the $k$th user channel is a $1 \times N$ row vector denoted by $\mathbf{h}_k$. The BS designs an arbitrary unitary precoder $\mathbf{Q}$ of size $N \times N$, further scaled for power constraint. Each terminal identifies the projection of its vector channel onto the precoder by $\mathbf{h}_k\mathbf{Q}$, and reports an index and a scalar metric expressing the SINR measured under an optimal beamforming vector selection.

$$\varphi_k = \max_{1 \leq i \leq N} \frac{|\mathbf{h}_k \mathbf{q}_i|^2}{\sigma^2 + \sum_{j \neq i} |\mathbf{h}_k \mathbf{q}_j|^2}, \tag{11}$$

where $\mathbf{q}_i$ denotes the $i$th column of $\mathbf{Q}$. The scheduling algorithm then consists in opportunistically assigning to each beamformer $\mathbf{q}_i$ the user that has selected it and has reported the highest SINR.

When the unitary precoder must be designed without any form of CSIT a priori, a scaled identity matrix can be used. In this case, the algorithm falls back to assigning a different selected user to each base antenna. In the small number of user case, the performance of such scheme is plagued by inter-user interference. Fortunately, interference tends to decrease as the number of users to choose in the cell becomes high.

When the dynamics of the system are limited (low mobility), the use of a fixed set of precoders may result in severe unfairness between the users. This problem can be alleviated by the randomization of the beamforming vectors. The so-called opportunistic random beamforming (ORBF) was initially proposed for single-user setting [35] and later generalized in [36]. The performance of these methods is illustrated below. The idea of [36] can be recast in the context above, assuming this time that $\mathbf{Q}$ is randomly generated at each scheduling period according to an isotropic distribution while preserving the unitary constraint. The intuition behind that scheme is that the columns $\mathbf{q}_i$, $i = 1, \ldots, N$, are like orthogonal beams, and if there are enough users in the cell, each beam will be aligned with a given user's channel while simultaneously being nearly orthogonal to the other selected users' channels. With this scheme, it is possible to spatially multiplex $N$ users with a level of feedback given by one scalar and one index. In the case of a large number of active users, opportunistic multibeam schemes are shown to yield an optimal capacity growth of $N \log \log U$ for fixed $N$, which is precisely the scaling obtained with full CSIT, as shown in (4).

## DEALING WITH SPARSE NETWORKS

A limitation of fixed or random opportunistic beamforming approaches is that the optimal capacity scaling emerges for a large, sometimes impractical, number of simultaneously active users in the cell. The performance degrades with a decreasing number of users (sparse networks), and this degradation is amplified when the number of transmit antennas increases, as intuition also reveals. The lack of robustness of these approaches in cases with small to moderate number of users is a serious problem that can be resolved by modifying the random beams for a better matching with the actual users' channels. This can be done at little or no extra feedback cost by one of several means. In one approach, the unitary constraint is relaxed by introducing a power control across the beams. The SINR feedback is used to adjust the power allocated to each beam [37] or simply to turn off certain beams [38], thus reducing interuser interference when the random beams are not well aligned with users' channels. In Figure 3, we compare the robustness of the single-beam ORBF [35] and multibeam ORBF [36], both with SINR feedback with respect to the number of active users in the cell. With four antennas at the BS, at 10 dB SNR, simulations suggest that at least 12 simultaneously active users are required for the multibeam gains to kick in. Whether this condition is met in practice or not is an interesting open research problem whose solution is likely to depend on the considered traffic, operational scenario, and delay constraint. With less users, the lack of CSIT destroys the benefits of user multiplexing. Interestingly, a strategy allowing for beam power control in multibeam ORBF [37] allows for a smooth transition between TDMA and SDMA regions, as shown in the figure.

Yet another approach is to exploit the second-order statistics of the channel, either in the temporal or in the spatial domain. The time domain approach consists in exploiting the natural temporal correlation of the channel to help refine the beams over time [39], [40]. In the spatial domain, statistics give information about spatial separability, which is instrumental to a proper beamforming design. Such aspect is described below.

**INFORMATION THEORY HIGHLIGHTS SEVERAL FUNDAMENTAL ASPECTS OF MU-MIMO SYSTEMS.**

### *USE OF SPATIAL STATISTICAL FEEDBACK*

In practical, especially outdoor, networks, the i.i.d. channel model used so far does not hold, and each user tends to exhibit different channel statistics. The advantage of statistical CSI is its long coherence time compared with that of the fading channel. Several forms of statistical CSI are even reciprocal (i.e., holds for both uplink and downlink frequency) such as second-order correlation matrix, power of Ricean component, etc., and do not necessitate any feedback. Overall, spatial channel statistics reveal a great deal of information on the macroscopic nature of the underlying channel, including the multipath's mean angle of arrival/departure and its angular spread. More generally, a substantial amount of channel distribution information (CDI) is revealed by channel statistics, which can be used to infer knowledge on mean user separability. Clearly however, in fading channels, the CDI ought to be complemented with some form of instantaneous channel quality information (CQI) to extract multiuser diversity gain. Combining CDI and CQI can yield partial CSIT, which is very well suited to solving the scheduling stage of the MU-MIMO problem. It is an open topic for research, but some leads are presented below.

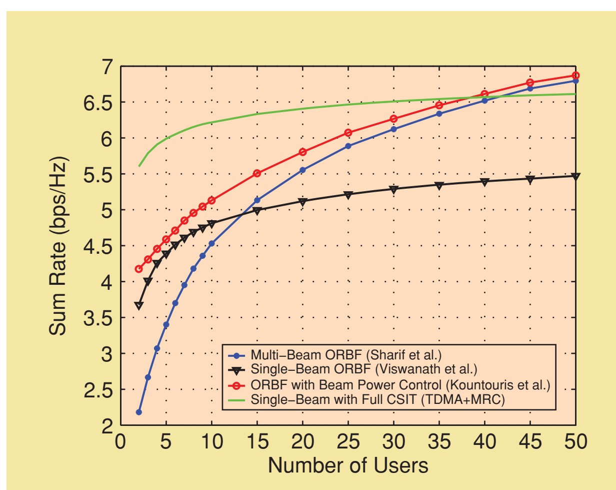
Consider the downlink of a network with single antenna mobiles, where the BS exhibits correlated transmit antennas. The channel is modeled as correlated Ricean fading, i.e., the channel vector of $k$th user satisfies $\mathbf{h}_k \sim \mathcal{CN}(\bar{\mathbf{h}}_k, \mathbf{R}_k)$, where $\bar{\mathbf{h}}_k \in \mathbb{C}^{1\times N}$ and $\mathbf{R}_k \in \mathbb{C}^{N\times N}$ are the mean value and transmit covariance matrix, respectively, known to the BS. A general form of CQI is

$$\gamma_k = \left\|\mathbf{h}_k \mathbf{Q}_k\right\|^2, \tag{12}$$

where $\mathbf{Q}_k \in \mathbb{C}^{N\times L}$ is a training matrix containing $L$ orthonormal vectors $\{\mathbf{q}_{ki}\}_{i=1}^{L}$. Conditioned on the CQI feedback, a coarse estimate of the instantaneous channel realization and channel correlation at the transmitter can be calculated as the conditional expectations

$$\widehat{\mathbf{h}}_k = \mathbb{E}(\mathbf{h}_k|\gamma_k) \qquad \widehat{\mathbf{R}}_k = \mathbb{E}\left(\mathbf{h}_k^H\mathbf{h}_k|\gamma_k\right), \tag{13}$$

which can be used to provide an MMSE estimate of the instantaneous SINR [41]. Note that with $\mathbf{Q}_k = \mathbf{I}$, (12) falls back to a channel norm feedback.




**[FIG3] The sum rate is compared for random beamforming schemes with SINR feedback. Multibeam (SDMA) random beamforming outperforms the single-beam (TDMA) when the number of active users is sufficient. Power control over the random beams allows for a smooth transition between TDMA and SDMA. TDMA with full CSIT outperforms partial feedback schemes for a small number of users but fails to provide multiplexing gain when this number increases.**

Similarly, a maximum-likelihood (ML) estimation framework maximizing the log-likelihood function of the probability density function (pdf) of $\mathbf{h}_k$ under the scalar constraint (12) can be formulated [42]. Let $L = 1$, $\mathbf{h}_k \sim \mathcal{CN}(0, \mathbf{R}_k)$ and CQI feedback $\gamma_k = |\mathbf{h}_k \mathbf{q}_k|^2$. The solution to the ML problem

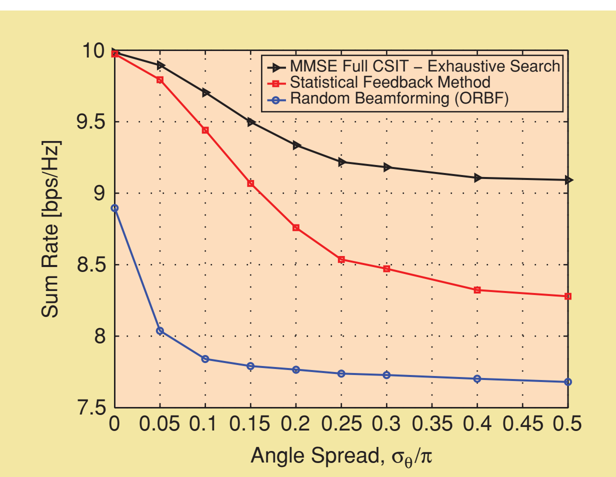
$$\max_{\mathbf{h}_k} \mathbf{h}_k \mathbf{R}_k \mathbf{h}_k^H \ s.t. \ |\mathbf{h}_k \mathbf{q}_k|^2 = \gamma_k \tag{14}$$

is given by the (dominant) generalized eigenvector associated with the largest positive generalized eigenvalue of the Hermitian matrix pair $(\mathbf{R}_k, \mathbf{q}_k \mathbf{q}_k^H)$. Once the coarse channel estimation is performed by the BS, it can be used to select up to $N$ users according to any number of previously described performance metric based on CSIT. As a second stage, more complete CSIT may be requested by the BS only to the small set of selected users for a more accurate precoding design. The performance exceeds that of random beamforming but depends on the level of antenna correlation, i.e., angle spread $\sigma_\theta$, as is shown in Figure 4. Certain techniques above are suited to specific deployments scenarios. For instance, opportunistic schemes are suited to densely populated networks. Schemes using temporal statistics are better suited to low mobility (indoor) setting, while the exploitation of spatial statistics would be more effective in outdoor cases where the elevation of the BS above the clutter decreases the angle spread of multipath and gives rise to Ricean models.

**THE IMPACT AND DESIGN OF AN OPTIMAL FORM OF CSIT UNDER FINITE RATE FEEDBACK IS STILL AN OPEN AND EXCITING PROBLEM.**

## SYSTEM ISSUES

Although it is now widely recognized that MIMO techniques, in their generality, will be a key element in the evolution of broadband wireless access systems, applications of MU-MIMO solutions have yet to emerge. While spatial diversity and basic SU-MIMO techniques are available in several products and standards, adaptive antenna solutions, including MU-MIMO, are mostly considered for time division duplex (TDD) systems in low and moderate mobility where CSI can be obtained from estimation in the uplink. We believe, however, that the promise is such that these techniques will be eventually available in most systems.




**[FIG4] Sum rate as a function of the angle spread $\sigma_\theta$ at the BS, where the number of transmit antennas is 2, the average SNR = 10 dB, and the number of active users in the cell is 50.**

Note that codebook based precoding schemes for SU- and MU-MIMO are emerging in existing and future standards [4]. MU-MIMO systems may have the potential to achieve the spectrum efficiency requirements set by operators for the next generation of mobile communication systems [43]. Practical MU-MIMO applications are still challenging however, and further studies seem needed to get a deeper understanding of the related tradeoffs and system gains (number of antennas, choice of algorithm, etc.).

When it comes to the crucial CSIT issue, one problem with designing feedback metrics is that the SINR measurement depends, among other things, on the number of other terminals being simultaneously scheduled along with the user making the measurement. Certain metrics (such as those in, e.g., [33], [36]) assume a fixed number of scheduled SDMA users. However, in practice, methods allowing fast transitions between TDMA and SDMA modes will be required. In such cases, the number of simultaneous users and the available power for each of them will generally be unknown at the terminal. Channel quality metric design in this scenario is one of the largely open challenges in MU-MIMO.

Also, opportunistic scheduling in MU-MIMO not only requires feedback for CSIT but also signaling of scheduling decisions to the terminal. The feedback and control loop in MU-MIMO introduces a nonnegligible overhead and latency in the system, which must carefully be weighed against the capacity gains expected from such techniques. Certain scenarios look promising (e.g., broadband best-effort internet access); others are more questionable, such as Voice over Internet Protocol (VoIP), where small packets are to be delivered with tight delay constraints. In addition, a poorly designed feedback channel can suffer from delays and cause the reported channel quality metrics to the transmitter to be outdated, bringing further degradations [44].

Another fundamental aspect is the impact of realistic traffic models and system loads, especially on schemes relying on high user loads (e.g., random beamforming). In recent wireless systems based on MIMO-OFDMA [5], opportunistic scheduling can be performed in up to three dimensions, namely, time, frequency, and space. Different types of traffic are likely to have different constraints with respect to the available DoF for the scheduler. For example, real-time services typically have tight delay constraints and limit the flexibility of the scheduler in the time domain. One may then wonder how many effective users are available for selection by the scheduler in each of these dimensions and how to take advantage of the different DoF to satisfy the QoS constraints for different types of traffic?

## DISCUSSION

MU-MIMO networks reveal the unique opportunities arising from a joint optimization of antenna combining techniques with resource allocation protocols. Furthermore, it brings robustness with respect to multipath richness, allowing for compact antenna spacing at the BS and, crucially, yielding the diversity and multiplexing gains without the need for multiple antenna user terminals. To realize these gains, however, the BS should be informed with the user's channel coefficients, which may limit practical application to TDD or low-mobility settings. To circumvent this problem and reduce feedback load, combining MU-MIMO with opportunistic scheduling seems a promising direction. The success for this type of scheduler is strongly traffic and QoS-dependent, however. A number of complementary approaches geared toward feedback reduction were proposed, which may restore the robustness of MU-MIMO techniques with respect to a wider range of application and environments. These results and other performance studies with low feedback schemes suggest that MU-MIMO transmitters can cope with very coarse channel information. From a theoretical point of view, the impact and design of an optimal form of CSIT under finite rate feedback is still an open and exciting problem.

**MU-MIMO NETWORKS REVEAL THE OPPORTUNITIES ARISING FROM A JOINT OPTIMIZATION OF ANTENNA COMBINING TECHNIQUES WITH RESOURCE ALLOCATION PROTOCOLS.**

## AUTHORS

***David Gesbert*** (gesbert@eurecom.fr) is with Eurecom Institute, France. He obtained the Ph.D degree from ENST, France in 1997. Before joining Eurecom, he was with the Information Systems Laboratory, Stanford University, then a founding engineer of Iospan Wireless Inc, San Jose, CA, (now Intel), then with the Department of Informatics, University of Oslo. He has published 120 papers and several patents, all in the area of signal processing, communications, and wireless networks. He is a co-editor of five special issues on wireless networks and communications theory for *JSAC*, *EURASIP*, and *Wireless Communications Magazine*. He is a member of the IEEE Signal Processing for Communications Technical Committee. He co-authored three award-winning papers in 2003, 2004, and 2005. He coauthored the book *Space Time Wireless Communications: From Parameter Estimation to MIMO Systems* (Cambridge Press, 2006).

***Marios Kountouris*** (kountour@eurecom.fr) received the diploma in electrical and computer engineering from the National Technical University of Athens, Greece in 2002 and the M.S. degree in electrical engineering from ENST Paris, France in 2004. He is currently working toward the Ph.D. degree at ENST—Eurecom Institute, France, funded by France Telecom R&D. During the summer of 2004, he interned with Samsung Advanced Institute of Technology, Korea, developing MU-MIMO precoding and scheduling techniques for 3GPP, and IEEE 802.16e systems. His research interests include precoding and resource allocation for multiuser MIMO systems and cross-layer design for wireless networks. He is a Student Member of the IEEE.

***Robert W. Heath Jr.*** (rheath@ece.utexas.edu) received the B.S. and M.S. degrees from the University of Virginia, Charlottesville, in 1996 and 1997, respectively, and the Ph.D. from Stanford University in 2002, all in electrical engineering. From 1998 to 2001, he was with Iospan Wireless Inc, San Jose, California. In 2003, he founded MIMO Wireless Inc. Since January 2002, he has been with the Department of Electrical and Computer Engineering at The University of Texas at Austin, where he is an associate professor and a member of the Wireless Networking and Communications Group. His research interests cover a broad range of MIMO communication, including limited feedback techniques, antenna design, relaying, ad hoc networking, and scheduling algorithms as well as 60 GHz communication techniques. He is an editor for *IEEE Transactions on Communication*, a former associate editor for *IEEE Transactions on Vehicular Technology*, and a member of the Signal Processing for Communications Technical Committee of the IEEE Signal Processing Society. He is a Senior Member of the IEEE.

***Chan-Byoung Chae*** (cbchae@ece.utexas.edu) is a Ph.D. student at the University of Texas, Austin. Prior to joining UT, he was a research engineer at the Telecommunication R&D center, Samsung Electronics from 2001 to 2005. He was a visiting research engineer at WING Lab, Aalborg University in Denmark, in 2004. He participated in the IEEE 802.16e standardization where he made several contributions and filed a number of related patents from 2004 to 2005. He was awarded the Bronze Prize in the 1996 Samsung Humantech International Paper Contest and the KSEA scholarship in 2007. His current research interests include capacity analysis and interference management in wireless mobile networks and all aspects of MIMO communications.

***Thomas Sälzer*** (thomas.salzer@orange-ftgroup.com) is with France Telecom R&D since 2004 and currently responsible for standard activities related to mobile radio systems. He obtained degrees from the University of Stuttgart, Germany, and ENST Paris, France in 2000 as well as a Ph.D. from INSA Rennes, France in 2004. He worked for Mitsubishi Electric ITE from 2000 to 2004, where he was involved in European research projects. He holds several patents and is author and co-author of papers in the field of MIMO, OFDM, and cross-layer design.

## REFERENCES


[1] H. Bölcskei, D. Gesbert, C. Papadias, and A.J. van der Veen, Eds., *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, Cambridge, U.K.: Cambridge Univ. Press, 2006.

[2] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[3] R. Knopp and P. Humblet, "Information capacity and power control in single cell multi-user communications," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, USA, June 1995, pp. 331–335.

[4] 3GPP, "Long Term Evolution, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; General description," *TS 36.201 v1.0.0*, Mar. 2007.

[5] IEEE, "Air interface for fixed and mobile broadband wireless access systems amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands," IEEE Standard 802.16e-2005, Feb. 2006.

[6] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[7] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[8] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 768–783, May 2004.

[9] P. Viswanath and D.N. Tse, "Sum capacity of the vector Gaussian channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[10] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 11–15, Jan. 2007.

[11] S.A. Jafar and A. Goldsmith, "Isotropic fading vector broadcast channels: The scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 848–857, Mar. 2005.

[12] W. Yu and W. Rhee, "Degrees of freedom in wireless multiuser spatial multiplex systems with multiple antennas," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1744–1753, Oct. 2006.

[13] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[14] C.B. Peel, B.M. Hochwald, and A.L. Swindlehurst, "A vector-perturbation technique for near capacity multiantenna multiuser communication—part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[15] M. Joham, W. Utschick, and J. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[16] Q. Spencer, A.L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 462–471, Feb. 2004.

[17] R. Chen, R.W. Heath Jr., and J.G. Andrews, "Transmit selection diversity for unitary precoded multiuser spatial multiplexing systems with linear receivers," *IEEE Trans. Signal Processing*, vol. 55, no. 3, pp. 1159–1571, Mar. 2007.

[18] Z. Pan, K.-K. Wong, and T.-S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969–1973, Nov. 2004.

[19] C.-B. Chae, D. Mazzarese, and R.W. Heath Jr., "Coordinated beamforming for multiuser MIMO systems with limited feedforward," in *Proc. Asilomar Conf. Sign., Syst. Computers*, Oct.–Nov. 2006, pp. 1511–1515.

[20] L. Choi and R.D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 773–786, July 2003.

[21] B.M. Hochwald, C.B. Peel, and A.L. Swindlehurst, "A vector-perturbation technique for near capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[22] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[23] C. Windpassinger, R.F.H. Fischer, and J.B. Huber, "Lattice-reduction-aided broadcast precoding," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2057–2060, Dec. 2004.

[24] R.F. Fischer, *Precoding and Signal Shaping for Digital Transmission*, New York: Wiley, 2002.

[25] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[26] F. Boccardi, F. Tosato, and G. Caire, "Precoding Schemes for the MIMO-GBC," in *Proc. Int. Zurich Sem. Commun.*, ETH Zurich, Switzerland, Feb. 2006.

[27] G. Dimić and N.D. Sidiropoulos, "On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm," *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[28] Z. Tu and R.S. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 370–372, Aug. 2003.

[29] Z. Shen, J.G. Andrews, R.W. Heath Jr., and B.L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3658–3663, Sept. 2006.

[30] D.J. Love, R.W. Heath Jr., W. Santipach, and M.L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 54–59, Oct. 2003.

[31] N. Jindal, "MIMO broadcast channels with finite rate feedback," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5059, Nov. 2006.

[32] P. Ding, D. Love, and M. Zoltowski, "Multiple antenna broadcast channels with shape feedback and limited feedback," *IEEE Trans. Signal Processing*, June 2006.

[33] T. Yoo, N. Jindal, and A. Goldsmith, "Finite-rate feedback MIMO broadcast channels with a large number of users," in *Proc. IEEE Int. Symp. Inform. Theory*, Seattle, WA, USA, July 2006, pp. 3879–3891.

[34] K. Huang, R.W. Heath Jr., and J. Andrews, "Space division multiple access with a sum feedback rate constraint," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3879–3891, July 2007.

[35] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[36] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[37] M. Kountouris and D. Gesbert, "Robust multi-user opportunistic beamforming for sparse networks," in *Proc. IEEE Workshop Signal Processing Adv. Wireless Commun.*, New York, USA, June 2005, pp. 975–979.

[38] J. Wagner, Y.-C. Liang, and R. Zhang, "On the balance of multiuser diversity and spatial multiplexing gain in random beamforming," *IEEE Trans. Wireless Commun.,* 2007, to be published.

[39] M. Kountouris and D. Gesbert, "Memory-based opportunistic multi-user beamforming," in *Proc. IEEE Int. Symp. Inform. Theory*, Adelaide, Australia, Sept. 2005, pp. 1426–1430.

[40] D. Avidor, J. Ling, and C. Papadias, "Jointly opportunistic beamforming and scheduling (JOBS) for downlink packet access," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, June 2004, pp. 2959–2964.

[41] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Signal Processing*, Mar. 2007.

[42] M. Kountouris, D. Gesbert, and L. Pittman, "Transmit correlation-aided opportunistic beamforming and scheduling," in *Proc. Europ. Signal Processing Conf.*, Florence, Italy, Sept. 2006.

[43] "NGMN: 'Next Generation Mobile Networks Beyond HSPA and EVDO—A white paper', V3.0," [Online]. Available: http://www.ngmn-cooperation.com, Dec. 2006.

[44] M. Kobayashi, G. Caire, and D. Gesbert, "Transmit diversity vs. opportunistic beamforming in data packet mobile downlink transmission," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 151–157, Jan. 2007.


[SP]

# Multi-Cell MIMO Cooperative Networks: A New Look at Interference

David Gesbert, Stephen Hanly, Howard Huang, Shlomo Shamai Shitz, Osvaldo Simeone, and Wei Yu

***Abstract*—This paper presents an overview of the theory and currently known techniques for multi-cell MIMO (multiple input multiple output) cooperation in wireless networks. In dense networks where interference emerges as the key capacity-limiting factor, multi-cell cooperation can dramatically improve the system performance. Remarkably, such techniques literally *exploit* inter-cell interference by allowing the user data to be jointly processed by several interfering base stations, thus mimicking the benefits of a large virtual MIMO array. Multi-cell MIMO cooperation concepts are examined from different perspectives, including an examination of the fundamental information-theoretic limits, a review of the coding and signal processing algorithmic developments, and, going beyond that, consideration of very practical issues related to scalability and system-level integration. A few promising and quite fundamental research avenues are also suggested.**

***Index Terms*—Cooperation, MIMO, cellular networks, relays, interference, beamforming, coordination, multi-cell, distributed.**

## I. INTRODUCTION

### A. Dealing with interference: conventional and novel approaches

FADING and interference are the two key challenges faced by designers of mobile communication systems. While fading puts limits on the coverage and reliability of any point-to-point wireless connection, e.g., between a base station and a mobile terminal, interference restricts the reusability of the spectral resource (time, frequency slots, codes, etc.) in space, thus limiting the overall spectral efficiency expressed in bits/sec/Hz/base station. At least, so has been the conventional view until recent findings in the area of cooperative transmission. Two basic scenarios are envisioned for cooperation in wireless networks. The first one assumes a (virtual) MIMO model for cooperative transmission over otherwise interfering links and will be the focus of this paper, while in the second relays are exploited. There exist interesting conceptual bridges between the two setups however, as will be made clearer in Section III and beyond.

Relay-based cooperative techniques try to mitigate detrimental propagation conditions from a transmitter to a receiver by allowing communication to take place via a third party device (mobile or base) acting as a relay. Initially developed relay-based cooperative transmission protocols have proved to be instrumental in mitigating fading effects (both path loss and multipath related) in point-to-point and point-to-multipoint communications. So-called amplify-forward, decode-forward, compress-forward cooperation schemes exploit available relay nodes to offer a powerful extra diversity dimension [1].

While conventional diversity and relaying schemes greatly improve the link-level performance and reliability, they do little to increase the quality of service to users placed in severe inter-cell interference-dominated areas, such as the cell boundary areas of current cellular networks. Instead, interference should be dealt with using specific tools such as "virtual" or "network" MIMO, so as to maximize the number of co-channel links that can coexist with acceptable quality of service. In the high SNR regime (achieved in, say, a small cell scenario), this figure of merit corresponds to the maximum number of concurrent interference-free transmissions and is referred to as the multiplexing gain of the network, or the number of *degrees of freedom* in the information-theoretic terminology.

The conventional non-cooperative approach to interference, via spatial reuse partitioning, prevents the reuse of any spectral resource within a certain cluster of cells. Typically, the frequency re-use factor is much less than unity, so that the level of co-channel interference is low. Thus, interference is controlled by fixing the frequency reuse pattern and the maximum power spectral density levels of each base station. Current designs do allow for full frequency re-use in each cell (typically for Code Division Multiple Access (CDMA) or frequency hopping spread spectrum systems) but this results in very severe interference conditions at the cell edge, causing a significant data rate drop at the terminals and a strong lack of fairness across cell users. Some interference mitigation is offered by limited inter-cell coordination, which is conventionally restricted to scheduling or user assignment mechanisms (e.g. cell breathing) or soft handover techniques. Inter-cell interference is treated as noise at the receiver side and is handled by resorting to improved point-to-point communications between the base station (BS) and the mobile station (MS), using efficient coding and/or single-link multiple-antenna techniques [2]. This approach to dealing with interference may be characterized as *passive*.

Manuscript received 10 January 2010; revised 1 July 2010. The review of this tutorial article was coordinated by Senior Editor Pamela Cosman.

D. Gesbert is with EURECOM, 06904 Sophia Antipolis, France (e-mail: gesbert@eurecom.fr).

S. Hanly is with the National University of Singapore Department of Electrical and Computer Engineering (e-mail: elehsv@nus.edu.sg).

H. Huang is with Bell Labs, Alcatel-Lucent, NJ., USA (e-mail: hchuang@alcatel-lucent.com).

S. Shamai Shitz is with Technion, Israel (e-mail: sshlomo@ee.technion.ac.il).

O. Simeone is with the Center for Wireless Communication and Signal Processing Research (CWCSPR) New Jersey Institute of Technology (NJIT) (e-mail: osvaldo.simeone@njit.edu).

Wei Yu is with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto (e-mail: weiyu@comm.utoronto.ca).

Digital Object Identifier 10.1109/JSAC.2010.1012xx.

In contrast, the emerging view on network design advocates a more *proactive* treatment of interference, which can be accomplished through some form of interference-aware multi-cell coordination, at the base station side. Although the complexity associated with the coordination protocols can vary greatly, the underlying principle is the same: Base stations no longer tune separately their physical and link/MAC layer parameters (power level, time slot, subcarrier usage, beamforming coefficients etc.) or decode independently of one another, but instead coordinate their coding or decoding operations on the basis of global channel state and user data information exchanged over backhaul links among several cells. Coordination protocols can exploit pre-existing finite capacity backhaul links (e.g., 802.16 WiMax, 4G LTE) or may require a design upgrade to accommodate the extra information sharing overhead. There are several possible degrees of cooperation, offering a trade-off between performance gains and the amount of overhead placed on the backhaul and over-the-air feedback channels. The different possible levels of cooperation are illustrated in detail in Section II.

### B. From multi-user to multi-cell MIMO

The history of base station cooperation can be traced back to previous decades, with the concept of *macroscopic diversity* whereby one or more mobiles communicate their messages through multiple surrounding base stations to provide diversity against long-term and short-term fading. In conventional CDMA networks, soft-handoff allows a mobile to communicate simultaneously with several base stations, and selection diversity is used to select the best of these connections at any given time. Such selection diversity increases both coverage and capacity [3], and combined with power control, it allows full frequency re-use in each cell. However, full frequency re-use comes at a price: CDMA capacity is then critically constrained by inter-cell interference, and the per-cell capacity in a network of interfering cells is much less than that of a single isolated cell. This reduction in capacity is measured by the so-called “f-factor” [4]. We will see that full base station cooperation essentially removes this interference penalty.

By “full base station cooperation” we mean that all base stations are effectively connected to a central processing site, as depicted in Figure 1 for the downlink scenario. On the downlink, the network is effectively a MIMO broadcast channel with distributed antennas. First steps toward full base station cooperation were taken in [5], [6], [7] for the uplink, which is effectively a MIMO multiple access channel. In these works, the base stations cooperate to decode each user. In [6], the model is a CDMA network, with single-user matched filter (SUMF) decoding, but the received signals from a mobile, at each base station, are maximal ratio combined before decoding. With such a global receiver there are no wasted signals causing pure interference: All received signals carry useful information for the global decoder, and hence interference is *exploited*. In [6], it is shown that with the optimal power control, such base station cooperation eliminates the inter-cell interference penalty. In other words, a network of interfering cells has the same per-cell capacity (in numbers of users) as a single, isolated cell. This result

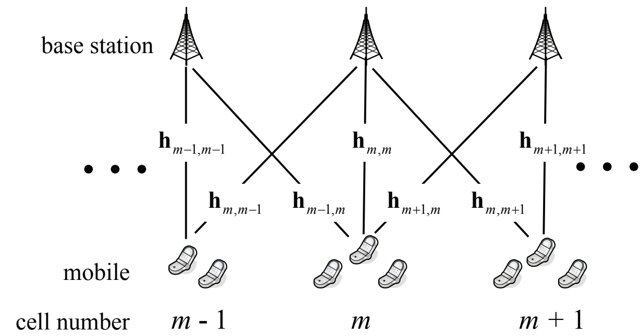



Fig. 1. A linear Wyner-type model with inter-cell interference spans $L_\ell = L_r = 1$ and $K = 3$ MSs per cell.

was extended to CDMA networks with more sophisticated multi-user receivers (decorrelator and MMSE receivers) in [7]. Again, the interference is fully eliminated and the achievable number of simultaneous users is same as if the cells were isolated from each other, although in this case the per-cell capacity benefits further from the more sophisticated multi-user detection.

The story is not so clear-cut if we consider more fundamental, information-theoretic models, where particular physical layers (*e.g.*, CDMA) or receiver structures, are not assumed *a priori*. Such information-theoretic results will be surveyed in Section III. However, similar conclusions do hold at high SNR, in terms of the degrees of freedom, as will be seen. Pioneering work on the information-theoretic capacity of the uplink of cellular networks with full base station cooperation was done in the early 90s [8], [9]. In these works, it was shown that with full base station cooperation, the traditional approach of frequency re-use partitioning is inherently suboptimal. Wyner [9] introduced a linear array model, and a hexagonal cell model, which have become known as Wyner models of cellular systems, and these are very tractable for information-theoretic analysis. In [8], it was shown that at high SNR, the capacity of a cellular system with fractional frequency re-use is less than a system with full frequency re-use, by exactly the re-use factor. This is equivalent to saying that full base station cooperation reduces the inter-cell interference penalty (or “f-factor”) to zero.

Although rich in content and ideas, [8], [9] stopped just short of spelling out the connections between the multi-user multi-cell channel and the MIMO channel. Communication over the spatial modes of the point-to-point MIMO channel (so-called spatial multiplexing) was formalized later in the mid 90s in [10], [11], then gradually extended to multi-user (MU-) MIMO channels. It was at that stage only that the downlink of the multi-cell cooperative channel, first investigated for the downlink in 2001 [12], was recognized to be almost identical to the so-called *broadcast MIMO channel*, if one ignores the power constraint at the individual base stations. On the uplink, there is no difference between ideal[1] multi-cell MIMO decoding and decoding over a multi-user MIMO channel.

[1]An ideal multi-cell MIMO channel is one where all base stations are connected via infinite capacity backhaul links.

Thus, a network of $M$ ideally connected $J$-antenna base stations can serve a total of $MJ$ terminals in an interference-free manner simultaneously, regardless of how strong the interference is. To achieve this remarkable result, multi-user spatial precoding and decoding is involved on the downlink and uplink respectively, much akin to techniques used over the MU-MIMO channel [13]. To date, progress in the area of multi-cell MIMO cooperation continues to parallel that realized in the more general MU-MIMO area. Nevertheless, this domain of communications provides specific and tough scientific challenges to communication theorists, although, remarkably, it is already being considered within industry and standardization fora.

### C. Challenges of multi-cell MIMO

Despite their promise, multi-cell MIMO systems still pose a number of challenges both theoretical and practical, several of which are described in this paper. First, a thorough understanding of the information-theoretic capacity of multi-cell MIMO system accounting for fading and path loss effects, even with an ideal backhaul, is yet to be obtained. As reviewed in this paper, capacity results exist for simplified interference models. Such results provide intuition for the general performance behavior but are difficult to extend to general channel models. Second, as multi-cell channels may involve a large number of antennas and users, algorithm development work is required to reduce the complexity of currently proposed precoding and decoding schemes. Optimal precoding over the broadcast (downlink) MIMO channel as well as optimal joint decoding over the uplink involve non-linear computationally intensive operations [14], [15] which scale poorly with the size of the network. Third, the equivalence between multi-cell systems and MIMO systems only holds in the case of ideal backhaul conditions. Practical cooperation schemes must operate within the constraints of finite capacity, finite latency communications links between base stations.

Deriving good theoretic performance bounds for MIMO cooperation over a channel with limited information exchange capability between the cooperating transceivers is a difficult task. As shown in this paper some results are available for a few simplified network models. From a practical point-of-view, a major research goal is to find good signal processing and coding techniques that approach ideal cooperative gains while relying on mostly local channel state information and local user data. This problem, referred to here as *distributed cooperation*, is as challenging as it is important. Efficient partial feedback representation methods building on classical MIMO research [16] are also desirable. From a system-level perspective, simulations indicate that substantial gains in capacity and increased fairness across cell locations will be accrued from the adoption of multi-cell MIMO techniques. Yet, a number of important practical issues must be addressed before a very realistic assessment of system gains can be made, such as the impact of imperfect synchronization between base stations, imperfect channel estimation at the receiver side, and network latency. Such aspects are addressed at the end of the paper along with a review of current field experiments.

### D. Scope and organization of paper

The theoretical treatments of interference-limited channels on the one hand, and of cooperation protocols on the other hand, are still maturing, mostly due to the inherent complexity of the problem. Nevertheless, the literature in these areas has grown to be extremely rich. For this reason, we do not attempt complete coverage of those domains of research. Instead, we focus on the adaptation of multi-antenna processing principles to the context of multi-cell cooperation. We refer to the obtained framework as *multi-cell processing* (MCP).

In Section II, we begin with the mathematical models for the network and signals, and the way that information is exchanged between the cells. Basic notations for multi-cell cooperation are given. Next, in Section III, key information-theoretic results are surveyed that establish closed-form expressions for sum rate bounds for several important interference and backhaul models. In Sec. IV, the focus is the design of practical MCP techniques, assuming an ideal backhaul. Sec. V deals with the problem of finite capacity backhaul and considers the feasibility of scalable and distributed MIMO cooperation, using such concepts as partial feedback, distributed optimization, Turbo base stations, and clustering. Sec. VI addresses system-level implementation issues due to expected imperfections at the physical layer. Initial tests with prototypes are reported. Finally, Sec. VII provides perspectives and suggestions for promising research avenues in this area.

Throughout this paper we adopt the following notations: $[\mathbf{x}]_k$ represents the $k$th element of vector $k$; $\mathbf{1}_N$, and $\mathbf{0}_N$ are $N \times 1$ vectors of all ones and all zeros, respectively; $[a, b]$, where $a \leq b$ are integer, is the interval $[a, ..., b]$; $(\cdot)^\dagger$ represent the conjugate transpose of its argument. $\mathbf{I}$ denotes the identity matrix.

## II. Modelling multi-cell cooperation

### A. System model

We consider a multi-cell network comprising $M$ cooperating base stations assigned with the same carrier frequency. Each cell serves $K$ users. The base stations are equipped with $J$ antennas each. Due to lack of space, we mostly focus on base station-side interference control: The users have single-antenna terminals, unless otherwise specified. Multiple antenna terminals can be considered to allow for the spatial multiplexing of mutiple data streams per user, or to give user-side multi-cell interference cancellation capability. The latter turns out to be useful especially in the context of *interference coordination*. This scenario is addressed in Section IV, but is otherwise excluded. The base stations can assume any geometry, however, strongly structured cell models can help the theoretical analysis of cooperation, as is discussed in Section III.

In the uplink, the received signal at the $m$th BS, $m \in [1, M]$ can be written as

$$\mathbf{y}_m = \sum_{l=1}^{M} \sum_{k=1}^{K} \mathbf{h}_{m,l,k} x_{l,k} + \mathbf{z}_m, \tag{1}$$

where $x_{l,k}$ is the symbol transmitted by the $k$-th MS in the $l$th cell, $\mathbf{h}_{m,l,k}$ denotes the $J$ element channel vector from the $k$-th user of cell $l$ towards the $m$th BS, $\mathbf{z}$ is the noise vector containing additive noise and any inter-cell interference not accounted for by the $M$ cooperating cells alone, for instance if the networks features more than $M$ BSs.

The model for the downlink can be easily obtained from the one above. We will reuse some symbols, but their meaning will be clear from the context. The signal received at the $k$-th user in the $m$-th cell is written as:

$$y_{m,k} = \sum_{l=1}^{M} \mathbf{h}_{l,m,k}^{\dagger} \mathbf{x}_l + z_m, \tag{2}$$

where $\mathbf{x}_l$ is the transmitter signal vector with $J$ elements transmitted from the $l$-th BS containing possibly precoded (beamformed) information symbols for several users. Note that, as a convention, the downlink channel vector from the $l$-th BS towards the $k$-th user in the $m$-th cell is denoted by the complex conjugated form of the corresponding uplink channel $\mathbf{h}_{l,m,k}$. This is done to allow the exploration of interesting duality results between uplink and downlink, as seen in Sections III and IV. Note that this represents a convenient writing convention rather than an actual assumption on physical reciprocity of the uplink and downlink channels. Where/if such an assumption of reciprocity (e.g., TDD system) is actually needed will be made clear in the paper.

### *B. The different levels of multi-cell cooperation*

In this paper we distinguish simpler forms of multi-cell coordination from those requiring a greater level of information sharing between cells.

*1) Interference coordination:* The performance of current cellular networks can already be improved if the BSs share the channel state information of both the direct and interfering links, obtained from the users via feedback channels (see Fig 2). The availability of channel state information (CSI) allows BSs to coordinate in their signaling strategies, such as power allocation and beamforming directions, in addition to user scheduling in time and frequency. This basic level of coordination requires a relatively modest amount of backhaul communication and can be quite powerful if enough users co-exist in the system (multi-user diversity) [17]. No sharing of transmission data, or signal-level synchronization between the base stations is necessary. We refer to such schemes as *interference-coordination*. In this case, the downlink signal at the $l$-th base, $\mathbf{x}_l$, is a combination of symbols intended for its $K$ users alone.

*2) MIMO cooperation:* On the other hand, when base stations are linked by high-capacity delay-free links, they can share not only channel state information, but also the full data signals of their respective users (see Fig. 3). A more powerful form of cooperation can be achieved. In this scenario, the concept of an individual serving base for one terminal disappears since the network as a whole, or at least a group of cells, is serving the user. The combined use of several BS antennas belonging to different cells to send or receive multiple user data streams mimicks transmission over a MIMO channel and is referred to here as *MIMO cooperation*. In principle, MIMO cooperation transforms the multi-cell network into a multi-user MIMO (MU-MIMO) channel for which all propagation links (including interfering ones) are *exploited* to carry useful

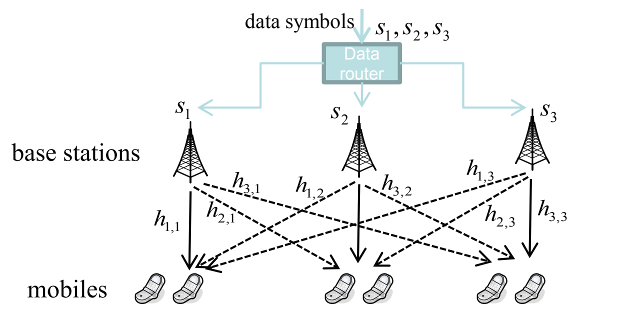



Fig. 2. Illustration of interference coordination for the downlink. The BSs acquire and exchange channel state information (but not the data symbols) pertaining to all relevant direct and interfering links, so as to optimize jointly their transmission parameters (time-frequency scheduling, power level, beamforming

data, upon appropriate precoding/decoding. In this case, the downlink signal, $\mathbf{x}_l$, is a combination of symbols intended for all $MK$ users. In contrast, interference-coordination schemes try to *mitigate* the generated interference, but they cannot really *exploit* it. For instance, beamforming may be used in each cell if the base stations are equipped with multiple antennas. In this case the beams typically try to strike a compromise between eliminating the inter-cell interference and maximizing the received energy to/from the user within the cell of interest. Ideally the choice of such beams across multiple cells is coordinated.

Although some interference-coordination ideas are promising, they are touched upon rather briefly in this paper, mainly in Section IV. MIMO cooperation schemes are the main focus of our attention.

*3) Rate-limited MIMO cooperation:* In the intermediate case, the base stations are linked by limited-capacity backhaul links. Typically, channel state information is shared first, then only a substream of user data or a quantized version of the antenna signals are shared among the base stations, which allows partial interference cancellation. Such hybrid scenarios are investigated from an information-theoretic point of view in Section III, then from an algorithmic perspective in Sections IV and V.

*4) Relay-assisted cooperation:* Instead of cooperating through backhaul links, it is also possible to consider channel models in which a separate relay node is available to assist the direct communication within each cell. Relay communication is relevant to the multi-cell MIMO network because it can be beneficial not only in strengthening the effective direct channel gain between the BS and the remote users, but also in helping with intercell interference mitigation. Relay enabled cooperation is studied in Sec. III-E.

## III. CAPACITY RESULTS FOR MULTI-CELL MIMO COOPERATION

In this section, we address the impact of cooperation on cellular systems from an information-theoretic standpoint. We

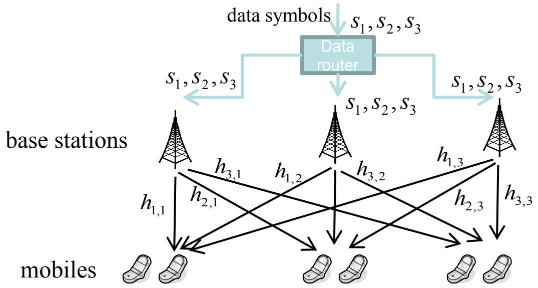



Fig. 3. Illustration of multi-cell MIMO for the downlink. The BSs, each equipped with $J$ antennas, acquire and share channel state information and user data, so as to mimic the behavior of a large MIMO array with $MJ$ antennas

mostly focus on the performance of MCP but we also consider the interplay of such techniques with cooperation in the form of relaying at the Mobile Stations (MSs) level.

### A. Introduction

The analysis of MCP was started in the early works [8][9] for the uplink and [12] for the downlink. The analysis in these works is based on the assumption that the BSs are connected via unrestricted backhaul links (*error-free and unlimited capacity*) to a Central Processor (CP) and focuses on models that, in information theoretic terms, can be seen as *symmetric Gaussian multiple-access or broadcast interference channels*. In these models, typically referred to as *Wyner-type models*, a number of users per cell is served by a single-antenna BS, as in a multiple access or broadcast channel, and interference takes place only between adjacent cells, as in partially connected interference networks [18]. Both models where cells are arranged on a line or in a more conventional bidimensional geometry can be considered, where the first class may model systems deployed along a highway or long corridor (see [19] for an implementation-based study), while the second applies to more general scenarios.

In this section, we consider the multi-cell MIMO scenario of Fig. 3 (i.e. with backhaul links allowing for some exchange of CSI and data symbols information), however we will focus on simplified cellular models that extend the Wyner-type models considered in the initial works [8], [9], [12]. Specifically, we consider the presence of limited-capacity and limited-connectivity backhaul links, fading channels and the interplay of MCP with relaying. We will focus on the *per-cell sum-rate* as the criterion of interest. It should be noted that, while the considered models capture some of the main features and practical constraints of real cellular systems, such as the locality of interference and constrained backhaul links, other features such as user-dependent path loss are not accounted for (see, e.g., [20]). Therefore, the models at hand can be seen as useful simplifications of real cellular settings, that enable insights and intuition to be obtained via analysis. It should be also noted that the use of sum rate as a system metric may mask other interesting features of multi-cell cooperation such as improving the balancing of user's quality of service from cell center to cell edge.

### B. The Linear Wyner Model

In this section, we review a basic system model for multiple cell networks introduced in [8], [9]. We focus the attention on *linear* Wyner-type models, as done in the original works. Extension of the given results to planar models is possible, though not always straightforward and we refer to [21] for further discussion on this point. An extension of the model to include relays is discussed in Section III-E of the present paper. A linear Wyner-type model is sketched in Fig. 1. We now present the corresponding signal models for uplink and downlink.

*1) Uplink:* A general linear Wyner-type model is characterized by $M$ cells arranged on a line (as for a highway or corridor), each equipped with a single-antenna ($J = 1$) base station (BS) and $K$ single-antenna MSs. In this class of models, inter-cell interference at a given BS is limited to $L_\ell$ BSs on the left and $L_r$ on the right. Considering the uplink, the received signal (1) at the $m$th BS, $m \in [1, M]$, at a given time instant $t \in [1, n]$ ($n$ is the size of the transmitted block) can then be specialized as

$$y_m(t) = \sum_{l=-L_r}^{L_\ell} \mathbf{h}_{m,m-l}^T(t)\mathbf{x}_{m-l}(t) + z_m(t), \qquad (3)$$

where $\mathbf{x}_m(t)$ is the $K \times 1$ (complex) vector of signals transmitted by the $K$ MSs in the $m$th cell, the $K \times 1$ vector $\mathbf{h}_{m,l}(t)$ contains the channel gains $\{h_{m,l,k}\}$ towards the $m$th BS from mobiles in the $l$th cell (see Fig. 1 for an illustration) and $z_m(t)$ is complex symmetric Gaussian noise with unit power and uncorrelated over $m$ and time. We assume equal per-user power constraints

$$\frac{1}{n}\sum_{t=1}^{n} |[\mathbf{x}_m(t)]_k|^2 \leq \frac{P}{K}, \qquad (4)$$

for all $m \in [1, M]$ and $k \in [1, K]$, so that the *per-cell power constraint* is given by $P$. Notice that model (3) assumes full frame and symbol-level synchronization among cells and users, even though extensions of the available results to the asynchronous case may be possible following, e.g., [22].

The model (3) discussed above reduces to the following special cases that will be referred to throughout this section:

- *Gaussian Wyner model*: This corresponds to a static scenario with symmetric inter-cell interference and cell-homogeneous channel gains, i.e., we have $L_\ell = L_r = L$ and $\mathbf{h}_{m,m-k}(t) = \alpha_k \mathbf{1}_K$ with $\alpha_k = \alpha_{-k}$ and $\alpha_0 = 1$. By cell-homogeneous, we mean that the channel gains do not depend on the cell index $m$, but only on the distance between interfering cells (see also discussion below on edge effects). Parameter $L$ can be referred to as the *inter-cell interference span*. Moreover, inter-cell gains $\alpha_k \geq 0$, $k \in [1, L]$, are deterministic (no fading) and generally known to all terminals. It is remarked that in this class of models, all users in the same cell share the same path loss. We also emphasize that the original model in [8], [9] had $L = 1$, so that the system referred to here as

Gaussian Wyner model is to be seen as an extension of [8], [9];

- *Gaussian soft-handoff model*: This corresponds to a static cell-homogeneous system like the Gaussian Wyner model, in which, however, there is no symmetry in the inter-cell channel gains. Specifically, we have inter-cell interference only from the left cells as $L_\ell = L$, $L_r = 0$ and $\mathbf{h}_{m,m-k}(t) = \alpha_k \mathbf{1}$, with $\alpha_0 = 1$, where, as above, $\alpha_k \geq 0$ are deterministic quantities. This model accounts for a scenario in which users are placed at the border of the cell so that inter-cell interference is relevant only on one side of the given cell. In a number of works, including [23], [24], the Gaussian soft-handoff model is studied with $L = 1$, which can be seen as describing a soft-handoff situation between two adjacent cells;
- *Fading Wyner model*: This model incorporates fading, accounted for by random channel gains $\mathbf{h}_{m,k}(t)$, in the Gaussian Wyner model. In particular, we have $L_\ell = L_r = L$ and $\mathbf{h}_{m,m-k}(t) = \alpha_k \tilde{\mathbf{h}}_{m,m-k}(t)$ where vectors $\tilde{\mathbf{h}}_{m,m-k}(t)$, $t \in [1, n]$, are independent over $m$ and $k$ and distributed according to a joint distribution $\pi_k$ with the power of each entry of $\tilde{\mathbf{h}}_{m,m-k}(t)$ normalized to one. For simplicity, similar to the Gaussian Wyner model, *statistical* symmetric inter-cell interference is assumed, i.e., $\alpha_k = \alpha_{-k}$ (and $\alpha_0 = 1$) and $\pi_k = \pi_{-k}$. As for temporal variations, two scenarios are typical: (*i*) *Quasi-static fading*: Channels $\tilde{\mathbf{h}}_{m,m-k}(t)$ are constant over the transmission of a given codeword (i.e., for $t \in [1, n]$); (*ii*) *Ergodic fading*: Channels $\tilde{\mathbf{h}}_{m,m-k}(t)$ vary in an ergodic fashion along the symbols of the codeword. The ergodic model was studied in [25] with $L = 1$;
- *Fading soft-handoff model*: This model is the fading counterpart of the Gaussian soft-handoff model, and has $L_1 = L$, $L_2 = 0$, and $\mathbf{h}_{m,m-k}(t) = \alpha_k \tilde{\mathbf{h}}_{m,m-k}(t)$ where $\tilde{\mathbf{h}}_{m,m-k}(t)$ are independent and modelled as for the fading Wyner model. This scenario was considered in [26], [27] (under more general conditions on the joint distribution of vectors $\tilde{\mathbf{h}}_{m,m-k}(t)$).

In order to remove edge effects, we will focus on the regime of a large number of cells, i.e., $M \to \infty$. This way, all cells see exactly the same inter-cell interference scenario, possibly in a statistical sense, as discussed above. An alternative approach, considered, e.g., in [8], [23], would be to consider a system in which cells are placed on a circle, which would exhibit homogeneous inter-cell interference for any finite $M$. It is noted that, however, the two models coincide in the limit of large $M$ and, in practice, results for the two models are very close for relatively small values of $M$ [21].

We now rewrite model (3) in a more compact matricial form. We drop dependence on time $t$ for simplicity. To proceed, construct a $M \times MK$ channel matrix $\mathbf{H}$ such that $m$th row collects all channel gains to $m$th BS, i.e., $[\mathbf{h}_{m,1}^T, \mathbf{h}_{m,2}^T, ..., \mathbf{h}_{m,m}^T, \mathbf{h}_{m,m+1}^T, ..., \mathbf{h}_{m,M}^T]$, where $\mathbf{h}_{m,m-k}^T$ with $k \notin [-L_r, L_\ell]$ are to be considered as zero. We can then write the $M \times 1$ vector of received signals $\mathbf{y} = [y_1, ..., y_M]^T$ as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{5}$$

where $\mathbf{x} = [\mathbf{x}_1^T \cdots \mathbf{x}_M^T]^T$ is the vector of transmitted signals and $\mathbf{z}$ the uncorrelated vector of unit-power Gaussian noises. From the definition above, it is clear that, in general, $\mathbf{H}$ is a finite-band matrix (in the sense that only a finite number of diagonals have non-zero entries). Moreover, it is not difficult to see that for Gaussian Wyner and Gaussian soft-handoff models, matrix $\mathbf{H}$ has a block-Toeplitz structure, which will be useful in the following.

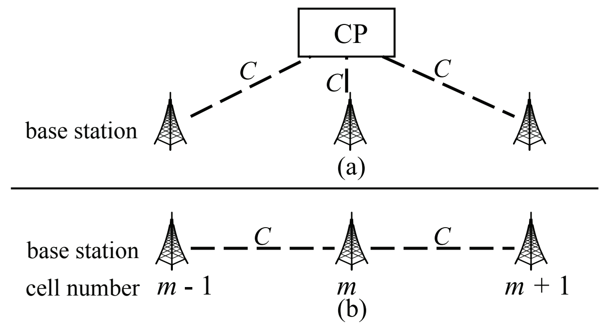



Fig. 4. Backhaul models for MCP: (a) Central processor (CP) with finite-capacity backhaul links (of capacity $C$); (b) Local finite-capacity backhaul between adjacent BSs (of capacity $C$, uni- or bi-directional). Dashed lines represent backhaul links.

*2) Downlink:* Define as $\mathbf{y}_m$ the $K \times 1$ vector of signals received by the $K$ MSs in the $m$th cell, $\mathbf{y} = [\mathbf{y}_1^T \cdots \mathbf{y}_M^T]^T$, and $\mathbf{x}$ as the $M \times 1$ transmitted signal by the BSs. We then have from (2)

$$\mathbf{y} = \mathbf{H}^\dagger \mathbf{x} + \mathbf{z}, \tag{6}$$

where $\mathbf{z}$ is the vector of unit-power uncorrelated complex Gaussian noise and channel matrix $\mathbf{H}$ is defined as above. We assume a *per-BS (and thus per-cell) power constraint* $\frac{1}{n}\sum_{t=1}^{n} |[\mathbf{x}(t)]_m|^2 \leq P$ for all $m \in [1, M]$.

*3) Multi-Cell Processing:* For both uplink and downlink, we will consider the two following models for the backhaul links that enable MCP, see Fig. 4.

- *Central processor (CP) with finite-capacity backhaul* (Fig. 4-(a)): In this case, all BSs are connected to a CP for joint decoding (for uplink) or encoding (for downlink) via finite-capacity backhaul links of capacity $C$ [bits/channel use]. Recall that the original works [8], [9], [12] assume unlimited backhaul capacity, i.e., $C \to \infty$;
- *Local finite-capacity backhaul between adjacent BSs* (Fig. 4-(b)): Here BSs are connected to their neighboring BSs only via finite-capacity links of capacity $C$ [bits/channel use], that may be uni- or bi-directional.

It is noted that two models coincide in the case of unlimited backhaul capacity $C \to \infty$. Also, we remark that another popular model assumes that only BSs within a certain cluster of cells are connected to a CP for decoding. This model will be considered as well, albeit briefly, below.

## C. Capacity Results for the Wyner Uplink Model

In the rest of this section, we elaborate on the per-cell sum-rate achievable for the uplink of the Wyner-type models without relays reviewed above. When not stated otherwise, we will focus on the *Gaussian Wyner model*. Fading models are discussed in Sec. III-C5. Throughout, we assume that channel state information (CSI) on gains $\{\alpha_k\}$ is available at all nodes.

*1) Single-Cell Processing (SCP) and Spatial Reuse:* Consider at first a baseline scheme, where Single-Cell Processing (SCP) is performed, so that each BS decodes individually its own $K$ users by treating users in other cells as Gaussian noise. A standard technique to cope with inter-cell interference is *spatial reuse*, that consists in activating at a given time (or equivalently in a given subband) only one cell every $F \geq 1$ cells. Parameter $F$ is referred to as the *spatial reuse factor*. SCP with special reuse is easily seen to achieve the per-cell sum-rate in Eq. 7 where $L$ is the inter-cell interference span. Rate (7) is obtained by either letting all users in a given cell transmit at the same time with power $FP/K$, which we refer to as Wide-Band (WB) transmission, or by intra-cell TDMA, whereby each user in a cell transmits with power $FP$ for a fraction of time $1/K$. Notice that such power allocations satisfy the per-block power constraint (4), due to the fact that each cell transmits for a fraction $1/F$ of the time.

A few remarks are in order. First, as seen in (7), if the reuse factor $F$ is larger than the inter-cell interference span $L$, SCP with spatial reuse completely eliminates inter-cell interference and provides a non-interference-limited behavior with per-cell multiplexing gain[2] equal to $1/F$, whereas otherwise the system operates in the interference-limited regime [8], [9]. From this, we conclude that the presence of inter-cell interference, if handled via SCP, leads to a rate degradation with respect to an interference-free system at high SNR given by a factor of $L$ [8]. In the low SNR regime, instead, where noise dominates inter-cell interference, using the formalism of [28][3], it can be seen that inter-cell interference does not cause any increase in the minimum (transmit) energy-per-bit necessary for reliable communications $E_b/N_{0\min}$, which equals $\ln 2 = -1.59$dB, as for interference-free channels. However, if one observes also the slope of the spectral efficiency $S_0$ [bits/s/Hz/(3dB)], which accounts for a higher-order expansion of the spectral efficiency as the SNR $P \to 0$, the loss due to inter-cell interference is seen also in the low-SNR regime. In fact, we have for rate (7):

$$\frac{E_b}{N_0}_{\min} = \ln 2 \text{ and } S_0 = \begin{cases} \frac{2}{F\left(1+4\sum_{k=1}^{\lfloor L/F \rfloor} \alpha_{kF}^2\right)} & \text{if } F \leq L \\ \frac{2}{F} & \text{if } F > L \end{cases}, \tag{8}$$

where we recall that interference-free channels have $S_0 = 2$. The conclusions here are related to the analysis in [29] on the suboptimality of TDMA for multiuser channels. As shown below, MCP allows to overcome the limitations of SCP and spatial reuse discussed here.

*2) Unlimited Backhaul:* Assume at first unlimited backhaul links to a CP, i.e., $C \to \infty$. The per-cell sum-capacity $R_{MCP}(P)$ with MCP in this scenario (for any $M$) is given by [9]:

$$R_{MCP}(P) = \frac{1}{M} \log_2 \det \left( \mathbf{I}_M + \frac{P}{K} \mathbf{H}\mathbf{H}^\dagger \right) \tag{9a}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \log_2 \left( 1 + \frac{P}{K} \lambda_i(\mathbf{H}\mathbf{H}^\dagger) \right) \tag{9b}$$

$$= \int_0^\infty \log_2 \left( 1 + \frac{P}{K} x \right) dF_{\mathbf{H}\mathbf{H}^\dagger}(x), \tag{9c}$$

where $\lambda_i(\mathbf{H}\mathbf{H}^\dagger)$ denotes eigenvalues of the argument matrix and $F_{\mathbf{H}\mathbf{H}^\dagger}(x)$ is the empirical distribution of such eigenvalues:

$$F_{\mathbf{H}\mathbf{H}^\dagger}(x) = \frac{1}{M} \sum_{m=1}^{M} 1(\lambda_i(\mathbf{H}\mathbf{H}^\dagger) \leq x). \tag{10}$$

The per-cell capacity (9) is achieved by performing ideal multi-user detection at the CP (which can in practice be realized by following approaches such as [30]). Moreover, it can be attained via both an intra-cell TDMA scheme where users transmit with power $P$ for a fraction of time $1/K$ and by the WB scheme (whereby all users transmit with full power $P/K$ at all times). It is noted that the optimality of TDMA is strictly dependent on the per-block power constraint (4), and would not hold under more restrictive conditions, such as peak or per-symbol power constraints. More general conditions under which TDMA is optimal, under per-block power constraints, can be found in [8]. For instance, from [8], it is found that TDMA would generally not be optimal in scenarios where users had different intra- and inter-cell channel gains, such as in fading scenarios (see Sec. III-C5). For the Gaussian Wyner model of interest here, using Szego's theorem, we get that for $M \to \infty$ rate (9) can be written [9] in a simple integral form as in (11). Expression (11) can be interpreted by considering the case $K = 1$ (without loss of generality, given the optimality of intra-cell TDMA) and identifying the signal received at the CP as the output (for each time instant) of a Linear Time Invariant (LTI) filter, whose input is given by the signals transmitted by the MSs and whose impulse response is $\delta_m + \sum_{k=1}^{L} \alpha_k \delta_{m-k} + \sum_{k=1}^{L} \alpha_k \delta_{m+k}$ ($\delta_m$ is the Kronecker delta). This integral cannot be evaluated in closed form in general. It should be noted that in other scenarios, such as the Gaussian soft-handoff model with $L = 1$, the corresponding integral can be instead calculated in closed form [23][24]. Notice that multiplexing gain of the MCP capacity (11) is one, as for an interference-free scenario. Moreover, the minimum energy-per-bit is given by

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{\left(1 + 2\sum_{k=1}^{L} \alpha_k^2\right)}, \tag{12}$$

showing an energy gain due to MCP with respect to SCP and to an interference-free system given by $(1 + 2\sum_{k=1}^{L} \alpha_k^2)$ (parameter $S_0$ is not reported here for lack of space but can be obtained from [21]).

*3) Limited Backhaul to the CP:* Consider now the scenario in Fig. 4-(a), where the BSs are connected to a CP via finite-capacity links. At first, we remark that the achievable per-cell

[2]The per-cell multiplexing gain is defined as the limit $\lim_{P\to\infty} R(P)/\log P$, where $R(P)$ is the given achievable per-cell sum-rate. A system is said to be interference-limited if the multiplexing gain is zero, and non-interference-limited otherwise.

[3]Reference [28] proposes to expand an achievable rate $R$ as a function of the energy-per-bit $E_b = P/R$ as $R \simeq \frac{S_0}{3dB}\left(\frac{E_b}{N_0}_{dB} - \frac{E_b}{N_0}_{\min,dB}\right)$, where $N_0$ is the noise spectral density (normalized to 1 here) and $\frac{E_b}{N_0}_{\min} = \frac{1}{\dot{R}(0)}$ and $S_0 = (2\ln 2)\frac{(\dot{R}(0))^2}{(-\ddot{R}(0))}$, where $R(P)$ is the considered rate (in bits/channel use) as a function of the power $P$.

$$R_{SCP}(P,F) = \begin{cases} \frac{1}{F}\log_2\left(1+\frac{FP}{1+2FP\sum_{k=1}^{\lfloor L/F \rfloor}\alpha_{kF}^2}\right) & \text{if } F \le L \\ \frac{1}{F}\log_2\left(1+FP\right) & \text{if } F > L \end{cases}, \tag{7}$$

$$R_{MCP}(P) = \int_0^1 \log_2\left(1+P\left(1+2\sum_{k=1}^{L}\alpha_k\cos(2\pi k\theta)\right)^2\right)d\theta \tag{11}$$

sum-rate is limited by the cut-set upper bound

$$R_{UB}(P,C) = \min\{C,\ R_{MCP}(P)\}. \tag{13}$$

Moreover, here, the performance depends on the knowledge of codebooks used by the MSs at the BSs. Assume at first that the BSs are unaware of the codebooks used by the MSs (*oblivious BSs*). In [31], this scenario is considered, and a per-cell achievable sum-rate is derived for a strategy whereby the BSs simply compress (to $C$ bits/channel use) and forward the received signals to the CP. Compression is followed by (random) binning, exploiting the fact that the other BSs have correlated information, according to standard techniques in distributed source coding (see, e.g., [32]). Decoding at the CP is done by *jointly*[4] decompressing the signals forwarded by the BSs and decoding the codewords transmitted by the users. A simple expression is found for this achievable rate in [31]:

$$R_{OBL}(P,C) = R_{MCP}(P(1-2^{-r})), \tag{14}$$

where $R_{MCP}$ is defined in (11) and $r$ is the solution of the fixed-point equation:

$$R_{OBL}(P(1-2^{-r})) = C - r. \tag{15}$$

In other words, the finite-capacity links entail a SNR loss of $(1-2^{-r})$ with respect to the unlimited backhaul capacity (11). It is noted that parameter $r$ has the interpretation of the amount of capacity $C$ that is wasted to forward channel noise to the CP [31], [32]. Also, we remark that rate (14) does not match the upper bound (13) in general. However, this is not always the case, and thus optimality of (14) is proved, in the regimes with $C \to \infty$ (in which compression noise becomes negligible), on the one hand, and $P \to \infty$ (in which the performance is limited by $C$), on the other. It is also interesting to point out that for low-SNR, the power loss of the oblivious scheme at hand with respect to (11) is quantified by calculating parameter $E_b/N_{0\min}$ as $E_b/N_{0\min} = \ln 2(1+2\sum_{k=1}^{L}\alpha_k^2)^{-1}(1-2^{-C})$. Comparing this with (12), one sees that in the low-SNR regime, the loss of (14) with respect to (11) is neatly quantified by $1-2^{-C}$. As a final remark, the optimal multiplexing gain of one is achieved if the backhaul capacity $C$ scales as $\log P$, which coincides with the optimal behavior predicted by the upper bound (13).

We now consider a different scenario where BSs are informed about the codebooks used by the MSs both in the same cell and in the interfering cells. In [31], a scheme is considered where partial decoding is carried out at the BSs. According to this approach, each MS splits its message and transmitted power into two parts: The first is intended to be decoded locally by the in-cell BS (with possible joint decoding also of the signals from the interfering BSs) and transmitted over the limited link to the CP, while the second part is processed according to the oblivious scheme and is decoded by the CP as discussed above. This scheme is shown to provide advantages over the oblivious rate (14) when the inter-cell interference is low (it is easy to see that it is optimal for $\alpha_k = 0,\ k > 0$) and for small $C$. Another strategy that exploits codebook knowledge at the BSs is the structured coding scheme proposed in [36] and reviewed below.

In [36], it is proposed that the BSs, rather than decoding the individual messages (or parts thereof) of the MSs as in [31], decode instead a function of such messages or, more precisely, of the corresponding transmitted codewords. The key idea that enables this operation is the use of structured, rather than randomly constructed, codes. Each MS employs the same nested lattice code and the signal received at any $m$th BS can be written from (11) as $y_m = \sum_{k=-L}^{L}\alpha_k x_{m-k} + z_m$. Recalling that a lattice code is a discrete group, the (modulo[5]) sum of the lattice codewords $x_{m-k}$, weighted by *integer* coefficients, is still a codeword in the same lattice code and can thus be decoded by the $m$th BS. The problem is that the channel coefficients $\alpha_k$ are generally not integers. The $m$th BS can however decode an arbitrary linear combination $\sum_{k=-L}^{L} b_k x_{m-k}$ with $b_k \in \mathbb{Z}$ (and by symmetry $b_k = b_{-k}$) and $b_0 \neq 0$ and treat the remaining part of the signal as Gaussian noise. The index of the decoded codeword can then be sent to the CP, that decodes based on all received linear combinations. This leads to the achievable rate [36] shown in (16) where $\mathcal{B} = \{(b_0,..,b_L) \in \mathbb{Z} : b_0 \neq 0$ and $b_0^2 + 2\sum_{k=1}^{L} b_k^2 \le 1 + P(1+2\sum_{k=1}^{L}\alpha_k^2)\}$. As shown in [36], this rate may outperform (14) for low and high inter-cell interference ([36] considers the case $L=1$). Moreover, [36] proves that rate (16) can be improved by superimposing additional messages to the lattice codewords.

*4) Local BS backhaul:* In this section, we turn to the model in Fig. 4-(b), where BSs are connected only to their neighboring BS via finite-capacity links. At first, for reference, we consider the related cluster-decoding setting of [37], where each BS, say the $m$th, can decode based not only on the locally received signal $y_m$ but also on the received signals from $i_\ell$ BSs on the left ($y_{m-k}$ with $k \in [1, i_\ell]$) and $i_r$ BSs on the right ($y_{m+k}$ with $k \in [1, i_r]$). Notice that this accounts for a situation where unlimited capacity backhaul links connect

[4]It is interesting to notice that while joint decompression/ decoding yields no performance benefits for regular interference-free systems [33], this is not the case in the presence of interference (see also [34], [35]).

[5]The modulo operation is taken with respect to the coarse lattice forming the nested lattice code.

$$R_{LAT} = \min\left\{C,\ \max_{(b_0,..,b_L)\in\mathcal{B}} -\log\left(b_0^2 + 2\sum_{k=1}^{L} b_k^2 - \frac{P(b_0 + 2\sum_{k=1}^{L}\alpha_k b_k)^2}{1 + P(1 + 2\sum_{k=1}^{L}\alpha_k^2)}\right)\right\}, \tag{16}$$

BSs, but only within a certain range of cells. Reference [37] obtains the maximum multiplexing gain of this setting for a Gaussian soft-handoff model with $L = 1$ with intra-cell TDMA (or equivalently $K = 1$). The model of [37] also assumes that MSs are aware, before choosing the transmitted codewords, of the messages of the MSs in $J_\ell$ cells on the left and $J_r$ cells on the right. This is a simple way to account for cooperation at the MS level, and will be further discussed in Sec. III-E. The maximum multiplexing gain is given by

$$\frac{J_\ell + J_r + i_\ell + i_r + 1}{J_\ell + J_r + i_\ell + i_r + 2}, \tag{17}$$

showing that with clustered decoding the multiplexing gain is generally less than one, but larger than $1/2$, as achievable with SCP and spatial reuse (see Sec. III-C1). Moreover, this shows that (for the soft-handoff model), left and right side informations have the same impact on the multiplexing gain, and the same applies to cooperation at the MSs or cluster decoding. Multiplexing gain (17) is achieved by *successive interference cancellation* at the BSs, where BSs exchange information about the decoded signals (see also below), and Dirty Paper Coding (DPC)-based cooperation at the users. It is noted that this scheme requires knowledge of the codebooks used in adjacent cells by both BSs and MSs. A model with cluster decoding at the BSs, but no cooperation amongst the mobiles, is considered in [38, Section IV], where similar general conclusions about the multiplexing gain are obtained.

In the presence of finite-capacity backhaul, the inter-BS links can be used to provide limited-rate information about the received signal or a processed version thereof to adjacent BSs. Such "relaying" has in general the double purpose of providing information about the useful signal of the recipient but also of the interference. This observation has also been made in the context of interference relay channels (see review in [39]). Along these lines, it is noted that the model and techniques at hand are very related to interference channels with "conferencing" decoders studied in [34], [40]. Consider, as in [41], a soft-handoff model with $L = 1$ and unidirectional backhaul links allowing information to be passed to the right. Assuming knowledge of only the local codebook, a successive decoding scheme can be devised in which each BS decodes the local message and sends the quantized decoded codeword to the neighboring (right) BS for interference mitigation. It is not difficult to see that such a scheme has zero multiplexing gain since it is not able to fully mitigate the interference. This is in contrast with the case where BSs have information about the codebooks used in adjacent cells. In this case, as in [37] (see discussion above), it becomes possible to perform joint decoding of the local message and of (possibly part of) the interfering message, and to use the backhaul link to convey directly hard information (messages) rather than soft information about the decoded codewords. This allows a non-interference limited behavior to be attained: Specifically, assuming that $C$ grows like $\beta \log P$, the multiplexing gain $\min(1,\, 0.5 + \beta)$ can be attained [41].

*5) Fading Channels:* In this section, we discuss available results on the sum-rate of fading Wyner and soft-handoff models. We consider both quasi-static and ergodic fading below.

*a) Quasi-static Fading:* With quasi-static fading, the *outage capacity* is typically used as a performance criterion [42]. This is, generally speaking, the maximum rate that guarantees reliable transmission for a given percentage of channel realizations (the measure of whose complement is referred to as outage probability). This setting implies either lack of CSI at the users (so that rate adaptation is not possible) or inelastic constant-rate applications. Using such a criterion in a large-scale cellular system with MCP proves to be challenging: In fact, on the one hand, defining outage as the event where *any* of the users' messages are not correctly decoded leads to uninteresting results as the number of cells $M$ grows large; On the other hand, defining individual outage events, as studied in [43] for a two-user MAC, appears to be analytically intractable for large systems (see [44] for related work).

A tractable performance criterion is instead obtained by considering the achievable per-cell sum-rate (9) for given channel realizations in the limit as the number of users per cell $K$ and/or the number of cells $M$ grow large, where the limit is defined in an almost sure sense. It is noted that such per-cell sum-rate is achievable by appropriate choice of distinct rates by the MSs, and such choice depends on the current realization of the channel matrices. The practical significance of this criterion is thus limited to instances in which, thanks to appropriate signaling, such rate adaptation is possible. Therefore, we review these results below as they are practically more relevant in the context of ergodic channels.

*b) Ergodic Fading:* With ergodic fading, the per-cell sum-capacity is given by the expectation of (9c) with respect to the distribution of $\mathbf{H}$, i.e., $R^{erg}_{MCP}(P) = \mathbb{E}[R_{MCP}(P)]$. It is noted that such rate can be attained, due to the (stochastic) symmetry of the considered model (neglecting edge effects by taking the limit $M \to \infty$) by equal rate allocation to all users. Moreover, it is achieved by a WB scheme (all users transmit at the same time), the rate of intra-cell TDMA being generally smaller. This is in contrast with Gaussian (unfaded) models, as discussed in Sec. III-C1, and is in line with standard results for multiple access channels [45], [46]. It is also noted that with SCP, when treating interference as noise as in (7), intra-cell TDMA may instead be advantageous over WB when intercell interference takes place and exceeds a given threshold [46]. Some performance comparison between intra-cell TDMA and spatial reuse in the presence of MCP for the soft-handoff model with $L = 1$ and Rayleigh fading can be found in [47].

To evaluate $R^{erg}_{MCP}(P)$, one can either use approximations based on bounding techniques as in [25] or the fact that, if $F_{\mathbf{HH}^\dagger}(x)$ converges almost surely in some asymptotic regime of interest to some limiting distribution (spectrum) $F(x)$, then $R^{erg}_{MCP}(P)$ converges to (9c) with $F_{\mathbf{HH}^\dagger}(x) = F(x)$ (see, e.g.,

[48]). We discuss below two such regimes.

Consider first the asymptotics with respect to $K$ and $M$ (with the inter-cell interference span $L$ kept fixed). Let us assume that the distribution $\pi_k$ of vectors $\bar{\mathbf{h}}_{m,k}$ (recall Sec. III-B) is such that each channel vector can be seen as a realization of a stationary and ergodic process with unit power and mean $0 \leq \mu \leq 1$ (and thus variance $1-\mu^2$). In this case, it can be verified that matrix $\mathbf{H}\mathbf{H}^{\dagger}$ converges almost surely to a deterministic Toeplitz matrix due to the strong law of large numbers [8]. Now, using Szego's theorem, similarly to (11), we have that for $K \to \infty$ and $M \to \infty$ (taken in this order) we obtain (18) (see [25]).

Notice that we recover (11) for $\mu = 1$, which corresponds to an unfaded scenario. It can be proved, similarly to [25], that (18) is decreasing in $\mu^2$, which implies that fading is beneficial in the limit of a large number of users. It is remarked that this may not be the case for a finite number of users $K$, as can easily be seen by noticing that for $K = 1$ and no inter-cell interference, one obtains a point-to-point link for which fading is known not to increase the rate [25], [26]. It is noted that the potential benefits of fading are related to the independence of the fading gains towards different BSs and thus cannot be mimicked by the MSs [25][23] (see also [8, Section 5.1.2] for a discussion on the effect of fading on the multiplexing gain, and on the power offset term). Moreover, from Jensen's inequality, it can be seen that (18) is an upper bound on the ergodic per-cell capacity for any number of users $K$ [25]. Finally, rate (18) does not depend on the actual (stationary and ergodic) distribution $\pi_k$ but only on its first two moments.

Consider now a regime where $K$ is fixed and $M$ grows to infinity. An approximation of the limiting spectrum $F(x)$, and thus of the per-cell rate $R_{MCP}^{erg}(P)$, has been obtained in [49] for the fading Wyner model with $L = 1$ and Rayleigh fading using free probability tools. Such approximation is seen to be accurate only for small values of the interference gain $\alpha_1$. In [26], [27], exact results on the convergence of per-cell rate (9a) are studied for the fading soft-handoff model. Almost sure convergence to a limit that depends on the Lyapunov exponent of a certain product of matrices is shown (see also [50] for related work). A central limit theorem is also proven in [27] along with a corresponding large deviation result, providing evidence to the fact that, given the limited randomness present in matrix $\mathbf{H}$ (due to the banded structure), convergence is slower than in classical random matrix theory (see, e.g., [51]). Finally, [52] characterizes the high-SNR behavior (in the sense of [53]) of (9a) as $M$ grows large and $K = 1$ user and $L = 1$. Performance bounds are also provided for $K > 1$. The result shows that such behavior depends on the specific distribution $\pi_k$, lending evidence to the conclusion that, in the case of finite-band matrices, the limit spectrum depends on the entries' distribution, unlike standard random matrix theory [25], [48]. We also remark that in the fading soft-handoff model with Rayleigh fading, $L = 1$, and intra-cell TDMA (or equivalently $K = 1$), the ergodic rate $R_{MCP}^{erg}(P)$ can be found in a compact integral form as shown in [54], [23]. Reference [47] obtains related bounds for $K > 1$.

Finally, we point to [46], where the effect of fading on a Wyner model with ideal cooperation only between adjacent cell sites is studied.

*c) MIMO Fading Models:* Another extension is to consider multiple antennas at the BS, with fading from each antenna to each user. Uplink models comparing SCP to MCP in this context are considered in [55], [56]. In [55] an asymptotic regime is considered in which the number of antennas at the base station, and the number of mobiles, grow large together, in a circular Wyner model. It is shown that the degrees of freedom depend on the system loading (number of users per base station antenna), but, if SCP and MCP are both optimally loaded (respectively), then MCP gains over SCP by a factor of three, but the gap can be reduced to a factor of two via the use of a re-use factor of two, with even and odd cells in separate bands.

*d) Channel uncertainty:* Channel uncertainty has not been adequately treated in the network MIMO literature to date. Its importance can be seen from the point to point MIMO channel where it is known that the number of transmit antennas should not exceed the number of symbols in the coherence block [57]. The reason is that part of the block of symbols must be used for training so that the MIMO channels can be measured at the receiver. If the coherence time is long relative to the symbol period then the number of antennas can be large, but if the coherence time is one symbol duration then one antenna is optimal. Note that the limiting asymptotics discussed in c) implicitly assume that the coherence time is growing with the number of users. Thus, channel uncertainty has implications for MIMO scalability, and we will discuss this issue further in Section V.

*6) Numerical Results:* We now focus on a numerical example for a Gaussian Wyner model with $L = 1$. Fig. 5 shows the per-cell sum-rate achievable by the techniques discussed above, namely SCP with spatial reuse $F = 1$ and $F = 2$ (7), ideal MCP (11), oblivious processing at the BSs (14) with backhaul capacity $C = 5$ and lattice coding (16) with $C = 5$. We have $P = 15dB$ and the inter-cell interference power gain $\alpha_1^2$ is varied. It can be seen that SCP with spatial reuse $F = 1$ provides interference-limited performance, while with $F = 2$ inter-cell interference is eliminated, but at the cost of possibly reducing the achievable rate. MCP provides remarkable performance gains, and can potentially benefit from larger inter-cell power gains $\alpha_1^2$. When the backhaul capacity is restricted to $C = 5$ (which is of the order of the per-cell achievable rates at hand), it is seen that by choosing the best between the oblivious BS scheme and the lattice-based scheme, one performs fairly close to the bound of ideal MCP. Moreover, lattice-based coding has performance advantages over oblivious processing for sufficiently large or small interference, large $P$ (not shown here) and moderate $C$. Increasing the capacity $C$ to say $C = 8$, leads to an almost ideal rate with the oblivious strategy (consistently with its asymptotic optimality for $P \to \infty$), while lattice coding does not improve its performance.

### *D. Capacity Results for the Wyner Downlink Model*

In this section, we review corresponding results for the downlink. Reference results using SCP and frequency reuse are similar to Sec. III-C1 and need not be discussed here. We focus, as for the uplink, on Gaussian models and briefly discuss the impact of fading in Sec. III-D3.

$$R^{erg}_{MCP}(P) = \int_0^1 \log_2 \left( \begin{array}{c} 1 + P(1-\mu^2)\left(1+2\sum_{k=1}^{L}\alpha_k^2\right) \\ +P\mu^2\left(1+2\sum_{k=1}^{L}\alpha_k\cos(2\pi k\theta)\right)^2 \end{array} \right) d\theta \quad (18)$$

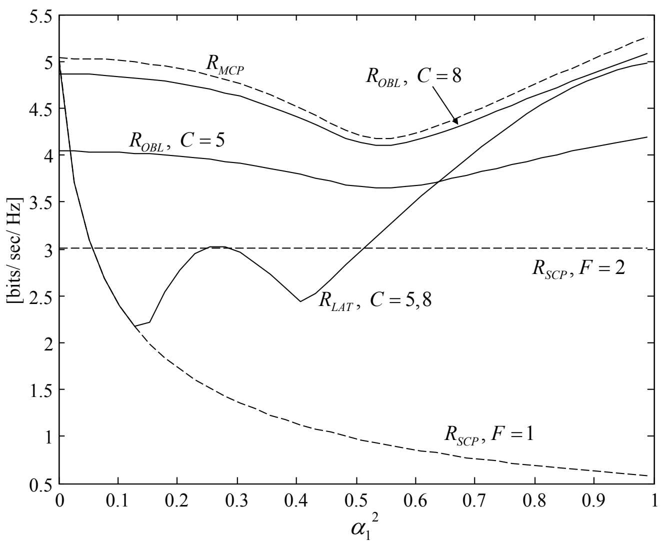



Fig. 5. Uplink of a Gaussian Wyner model with $L = 1$: Per-cell sum-rate achievable by SCP with spatial reuse $F = 1$ and $F = 2$ (7), ideal MCP (11), oblivious processing at the BSs (14) with backhaul capacity $C = 5$ and lattice coding (16) with $C = 5$ versus the inter-cell interference power gain $\alpha_1^2$ ($P = 15dB$).

*1) Unlimited Backhaul:* Consider first the case of unlimited backhaul. Reference [12] derives achievable rates based on the linear precoding dirty paper coding strategy of [58]. The per-cell sum-capacity is instead derived in [23] using the uplink-downlink duality results of [59] as

$$R_{MCP}(P) = \frac{1}{M}\min_{\mathbf{\Lambda}}\max_{\mathbf{D}}\log_2\frac{\det\left(\mathbf{\Lambda}+\frac{P}{K}\mathbf{HDH}^\dagger\right)}{\det(\mathbf{\Lambda})}, \quad (19)$$

with $\mathbf{\Lambda}$ and $\mathbf{D}$ being diagonal $MK \times MK$ matrices with the constraints $tr(\mathbf{\Lambda}) \le M$ and $tr(\mathbf{D}) \le M$. This rate is known to be achieved by dirty paper coding at the CP. For the Gaussian Wyner (circulant) model, it can be shown that the per-cell sum-capacity (19) is exactly equal to the corresponding capacity for the uplink (11) for $M \to \infty$. It follows that, as for the uplink, intra-cell TDMA, where only one user is served per-cell, is optimal with Gaussian (unfaded) channels.

*2) Limited Backhaul to a CP:* Similarly to the uplink, strategies to be used in the presence of limited backhaul to a CP depend on the level of codebook information available at the BSs. For oblivious BSs, reference [60] proposes to perform joint DPC under individual power constraint at the CP and then send the obtained codewords to the corresponding BSs via the backhaul links. The BSs simply transmit the compressed DPC-codewords. Since the transmitted quantization noise decreases the overall SNR seen by the MSs, joint DPC at the CP is designed to meet lower SNR values and tighter power constraints than those of the unlimited setup [23]. The resulting per-cell rate is shown to be equal to (11) but with a degraded SNR

$$\bar{P} = \frac{P}{1+\frac{1+P\left(1+2\sum_{k=1}^{L}\alpha_k^2\right)}{2^C-1}} \quad (20)$$

due to quantization noise. Similarly to the corresponding result (14) for the uplink, this rate is generally suboptimal but it achieves cut-set bound (13) (which is still a valid bound also for the downlink) for $C \to \infty$ (where the compression noise is dominant). However, unlike for the uplink, this rate is not optimal for $P \to \infty$: This fact can be understood by noticing that in the high-SNR regime, the compression noise dominates the performance, and, in the downlink, the compression noise is dealt with independently by each MS, unlike in the uplink, where decompression is performed jointly at the CP. Interestingly, for low-SNR the power loss in terms of $E_b/N_{0\,\min}$ turns out to be exactly the same as for the uplink, being given by $(1-2^{-C})$. Moreover, as for the uplink rate (14), optimal multiplexing gain of 1 per-cell is achieved if $C \sim \log P$.

Reference [60] also considers the case where the BSs possess codebook information about adjacent BSs belonging to a given cluster and proposes to perform DPC within the given cluster. The main conclusion of [60] is that the oblivious scheme is the preferred choice for small-to-moderate SNRs or when the backhaul capacity $C$ is allowed to increase with the SNR. On the other hand, for high SNR values and fixed capacity $C$, a system with oblivious BSs is limited by the quantization noise, and knowledge of the codebooks at the BSs becomes the factor dominating the performance.

*3) Fading Channels:* Following the discussion in Sec. III-C5, here we focus on the ergodic fading scenario. For this setting, the per-cell sum-capacity is given by $R^{erg}_{MCP}(P) = \mathbb{E}[R_{MCP}(P)]$ using (19). Evaluating this quantity is not an easy task due to the min-max operation involved. In [23], upper and lower bounds on $R^{erg}_{MCP}(P)$ are derived for the fading soft-handoff model with $L = 1$ and Rayleigh fading, along with asymptotic SNR characterizations. An important finding from such analysis is that, for large number $K$ of users per cell, the per-cell sum-rate capacity scales as $\log\log K$, which is the same type of scaling as for interference-free systems. A suboptimal scheme is then proposed in [61] based on zero-forcing (ZF) beamforming and a simple user selection (scheduling) rule whereby one user is served in each cell at any given time in an intra-cell TDMA fashion. It is found that, even this suboptimal scheme is able to achieve the same optimal scaling law of $\log\log K$ with Rayleigh fading.

An illustration of the achievable per-cell rates in a fading Wyner model with $L = 1$ and $\alpha_1^2 = 0.4$ is shown in Fig. 6. Specifically, the per-cell achievable rates with SCP and spectral reuse $F = 1$ and $F = 2$ (obtained similarly to Sec. III-C1), with ideal MCP (shown is the upper bound of [23]) and with the ZF beamforming and scheduling scheme of [61] are plotted versus the power $P$ and for $K = 50$ users per cell. The interference-limited behavior of SCP with $F = 1$

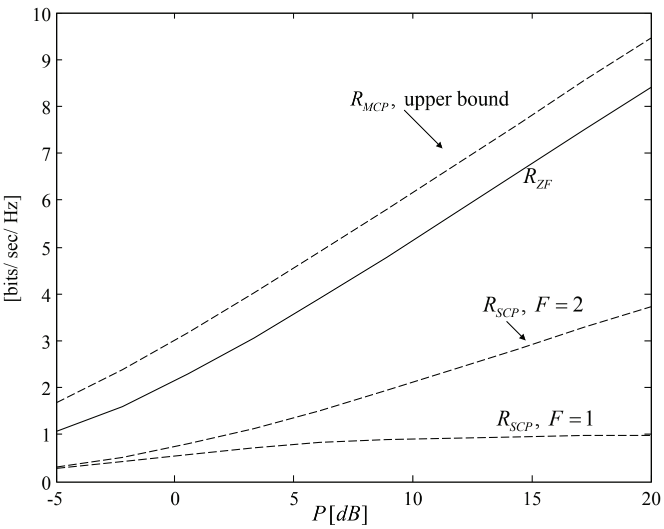



Fig. 6. Downlink of a fading Wyner model with $L = 1$: Per-cell sum-rates achievable with SCP and spectral reuse $F = 1$ and $F = 2$, ideal MCP, and with the ZF beamforming and scheduling scheme of [61] are plotted versus the power $P$ ($K = 50$, $\alpha_1^2 = 0.4$).

is apparent and so is the performance gain achievable via MCP. It is also interesting to notice that the suboptimal ZF beamforming scheme performs relatively close to the upper bound set by ideal MCP.

### *E. Relay-aided Models*

In modern cellular systems, the presence of dedicated relays is considered to be instrumental in extending coverage by enabling multi-hop communications or, more generally, cooperation at the MS level [62]. Here, we briefly review a model that accounts for the presence of dedicated relays, one per cell, in a Wyner-type setting (first considered in [63]). We focus on the uplink for simplicity and assume that the users are sufficiently far from the BSs so that the direct link from MS to BS can be neglected. We thus end up with two Wyner-type models, one from the users to the relays and one from the relays to the BSs (see Fig. 7). We will refer to these as first and second hop, respectively. We assume that the relays are full-duplex so that they can transmit and receive at the same time. The protocols we consider, except when stated otherwise, work by pipelining transmission on the two hops: The mobiles send a new message to the relays in every block, while the relays transmit to the BSs a signal obtained by processing the samples received in the previous block. Given the assumption of no direct link between mobiles and BSs, it is easy to see that results with half-duplex relays are immediately derived by halving the spectral efficiencies obtained for the full-duplex case (a new message can only be sent once every two blocks).

Denoting as $L_\ell$ and $L_r$ the maximum inter-cell interference spans on the left and right, respectively, for the two hops, we can write the signal model, similar to (5), as follows. The $M \times 1$ signal received at the relays can be written as

$$\mathbf{y}_R = \mathbf{H}\mathbf{x} + \mathbf{T}\mathbf{x}_R + \mathbf{z}_R, \tag{21}$$

where $\mathbf{H}$ is defined as in (5), and contains the channel gains for the first hop (MSs-relays), $\mathbf{x}$ is as in (5) and $\mathbf{z}_R$ is the Gaussian noise. The new element here is the $M \times 1$ signal $\mathbf{x}_R$ transmitted by the relays. The possible interference among relays in different cells is accounted for by matrix $\mathbf{T}$. Here, we assume that there is interference only between relays in adjacent cells, and that such interference is symmetric, so that $\mathbf{T}$ is a symmetric Toeplitz matrix with first row equal to $[0\ \mu\ \mathbf{0}_{M-2}^T]$, where $\mu$ represent the inter-relay gains. Finally, the signal received at the BSs is given by

$$\mathbf{y} = \mathbf{H}_R\mathbf{x}_R + \mathbf{z}, \tag{22}$$

where now $\mathbf{H}_R$ is the matrix containing the channel gains from relays to BSs and is defined similarly to $\mathbf{H}$ (see also Fig. 7). We assume *per-relay (and thus per-cell) power constraint* $\frac{1}{n}\sum_{t=1}^{n} |[\mathbf{x}_R(t)]_m|^2 \leq Q$, for $m \in [1, M]$.

Consider now the performance of cooperation in cellular networks in the presence of dedicated relay stations, following the uplink model discussed above. Depending on whether one assumes SCP or MCP, the system can be seen as an interference network with relays or as a multiple access channel with multiple relays and a multiple-antenna receiver, respectively. We remark that, in both cases, general conclusive results are unavailable even in the simple two-user cases considered in [39], [64]. Analysis, in terms of achievable per-cell sum-rate and corresponding upper bounds, has been pursued by assuming different transmission strategies and intra-cell TDMA (or equivalently $K = 1$). Specifically, reference [63] considers half-duplex amplify-and-forward (AF) processing at the relays, [65] studies half-duplex decode-and-forward (DF) relays, [66] full-duplex AF operation, [67] full-duplex DF and [68] full-duplex compress-and-forward, CF. In the following, we briefly review some results for the full duplex case.

In [66], the performance of AF with both SCP and MCP is studied. Relays simply delay the received symbol by at least one time unit, amplify and forward it, sample by sample. Closed-form analytical expressions are obtained for the per-cell sum-rate based on the observation that the received signal can be seen as the output of a two-dimensional LTI channel via Szego's theorem. The performance of both SCP and MCP is shown to be independent of the time-delay applied by the relays. It is observed that the rates of both schemes are decreasing with the intra-relay interference factor $\mu$. It is also shown that using the full power $Q$ of the relays is unconditionally optimal only for the MCP scheme, while this is not the case with SCP.

In [68], CF relaying with SCP and MCP is studied. Here, the relays operate in blocks, as explained in Sec. III-B, by collecting a number of received samples and compressing the received signal using a vector quantizer. Each BS for SCP, or the CP for MCP, decodes based on the quantized signals received from either the local relay (for SCP) or all the relays (for MCP). For MCP, due to the correlation among the received signals at the relays, distributed compression techniques are applied similarly to [31]. Moreover, the CF scheme with MCP exploits side information available at the CP regarding the signals transmitted by the relays (which are in fact decoded at the CP). It is proved that the scheme can completely remove the effect of the inter-relay interference. It

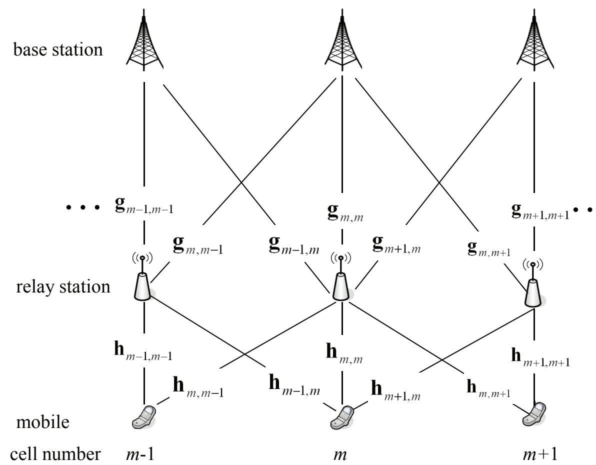



Fig. 7. Relay-aided linear Wyner models with inter-cell interference spans $L_\ell = L_r = 1$ and $K = 3$ MSs per cell.

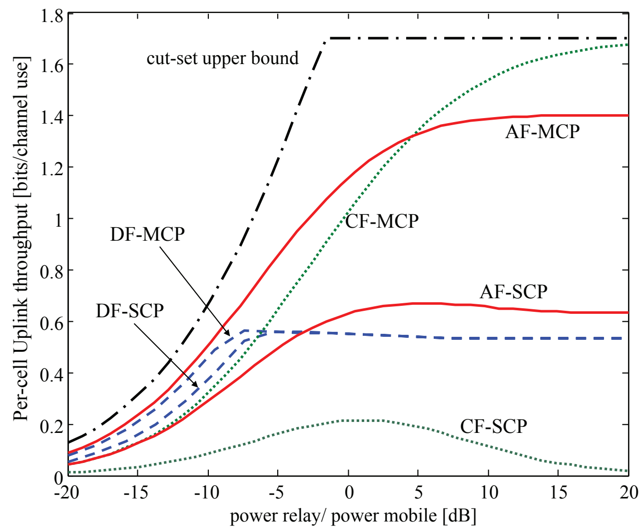



Fig. 8. Per-cell sum-rates achieved by different relaying schemes with SCP or MCP versus the ratio $Q/P$ between the power constraint at the relays ($Q$) and that at the MSs ($P$).

is noted that, in the nomenclature of the standard IEEE 802.16j [62], both CF and AF, which are non-regenerative relaying schemes, classify as transparent relaying strategies in that no knowledge of their presence is required at the mobiles. Also, the relays do not require information regarding the codebooks used by the terminals.

Finally, in [67], a regenerative relaying scheme based on DF is considered. Here, codebook information is required at the relays and generally the proposed schemes are non-transparent. The idea is to use rate splitting at the mobile in a similar manner to the standard technique for interference channels [69] so that each relay decodes not only the message of the local mobile (recall that we are assuming intra-cell TDMA), but also part of the message of the adjacent mobiles. This way, the relays can cooperate while transmitting towards the BSs by beamforming the common information.

A comparison among the performance of the schemes described above is shown in Fig. 8 versus the ratio $Q/P$ between the power constraint at the relays ($Q$) and that at the MSs ($P$). A first observation is the interplay between SCP or MCP (i.e., cooperation at the BSs) and cooperation via dedicated relays through different strategies. Specifically, it can be seen that if SCP is deployed, DF is advantageous with respect to CF, and also with respect to AF, if the power of the sources is sufficiently larger than that of the relays. It is noted that CF performs very poorly due to its inability to beamform the users' signals towards the BSs, unlike DF and AF. However, if MCP is in place, the situation is remarkably different in that DF is outperformed by both CF and AF unless the sources' power is sufficiently larger than that of the relays. This is because DF is limited by the performance bottleneck due to the need to decode at the relay stations, which prevents the system from benefiting from MCP. Finally, it is seen that the proposed CF scheme performs close to optimal if the relay power is sufficiently large.

We finally recall a different model for cooperation at the mobile level that does not involve dedicated relays but inter-mobile transmission. Namely, [70] models the inter-user links as orthogonal to the main uplink channel, whereas a generalized feedback model (in the sense of [71]) is considered in [67].

## F. Conclusions from the Information-Theoretic Models

This section has illustrated, via information-theoretic arguments, the advantages of cooperation in cellular systems. Cooperation among the BSs (or MCP) has been shown to be able to potentially increase the sum-capacity of the network by an amount proportional to the inter-cell interference span (i.e., number of BSs interfered by a local transmission) with respect to standard single-cell strategies with spatial reuse. While initial work demonstrated such benefits under idealistic conditions, in terms of, e.g., absence of fading and perfect backhaul, more recent research has confirmed the promises of MCP under more practical conditions.

In this section, we have reviewed more recent research that includes practical constraints such as limited backhaul bandwidth, localized base station clustering, and the effect of fading. We conclude that with oblivious BSs there is an $E_b/N_{0\min}$ penalty incurred by limited backhaul, but the capacity is unaffected provided that the backhaul bandwidth scales as $\log$ SNR. Not surprisingly, the impact of BS clustering is not significant provided that the clusters are sufficiently large, see equation (17). On the other hand, the effect of fading is somewhat surprising. As well known for general MIMO links, fading provides the degrees of freedom. But with MCP, it turns out that with a large number of users and BSs, the capacity is *increasing* in the variance of the fading.

The performance benefits of cooperation at the MS level have been reviewed as well, along with considerations regarding the strong interplay between the design of relaying strategies and MCP techniques. Here, there is much scope for further research, and most of the results reported are of a preliminary nature: Even the simple relay channel is an open problem in information theory. But achieveable rates can easily be calculated for particular schemes such as AF, CF, and DF. We conclude that for sufficiently powerful relays, CF is the best technique under MCP. These results show that the MCP model with relays is stimulating new efforts in network information theory.

The presentation has also briefly touched upon the potential gains achievable by exploiting novel transmission strategies such as structured codes. Other advanced techniques, not discussed here, such as interference alignment are also expected to have an important role to play in cooperative cellular systems (see, e.g., [72]). Other related issues of interest are the impact of imperfect channel state information and robust coding strategies [73]. This area remains an active and fertile field for research and is briefly addressed in the next sections.

## IV. TRANSMISSION AND CODING TECHNIQUES

This section provides an overview of transmission and signaling strategies for practical multi-cell MIMO networks, in which the base-stations cooperate. The nature of cooperation (interference coordination or MCP) determines the suitable strategies in various cases. Some of these strategies are straightforward extensions of traditional MIMO signaling techniques, while many others require novel and nontrivial ideas. This section also reviews the possible optimization methods. The optimization space involves scheduling, power allocation, transmit and receive beamforming, as well as choices of transmission strategies. One of the objectives of this section is to highlight the difference between single-cell MIMO techniques and multi-cell techniques. When appropriate, the possible distributed implementation of an algorithm is mentioned, since distributed processing is a primary challenge for the design of multi-cell MIMO networks. However this issue is visited in greater detail in Section V. Below we distinguish between the techniques involving CSI exchange only (interference coordination) and the MCP schemes which require both CSI and user data exchange, and provide an overview of coding, precoding and optimization strategies in each of these cases.

### *A. Interference coordination strategies*

Consider first a basic level of coordination where only the channel state information of the direct and interfering channels are shared among the BSs, a setup illustrated earlier in Fig. 2. The availability of channel state information allows the transmission strategies across the different cells to adapt to the channel state jointly. Transmission strategies can include scheduling, power control, beamforming, as well as advanced coding methods specifically designed for interference mitigation.

*1) Coordinated power control:* In an interference-limited cellular network, joint power control and scheduling across the multiple BSs that adapts to the channel condition of the entire network can bring improvement over traditional per-cell power control. This is especially evident when cellular topology is such that cells significantly overlap.

The resource allocation problem in the multi-cell setting has been studied extensively in the literature [17], [74], [75], [76], [77], [78], [79]. In the following, we consider a simple scenario where both the BS and the remote users are equipped with a single antenna to illustrate the main challenge in multi-cell power control. In this setting, there is a surprising result for the special case of an arbitrary two-cell set-up where the optimum sum-rate maximizing power allocation policy is in fact binary, i.e. the optimum strategy involves either both cells operating at maximum allowed power or one cell being completely shut down [74]. This result does not extend to more than two cells however.

In a more general setting, consider an orthogonal frequency-division multiple-access (OFDMA) system in which multiple users within each cell are separated in the frequency domain. Note that the multiple access across the multiple cells is not orthogonal since we allow for full reuse of the frequency tones from one cell to the next. The joint power control and scheduling problem is that of deciding which user should be served and how much power should be used on each frequency tone. Mathematically, in a multi-cell network with $M$ cells, $K$ users per cell, and $N$ OFDM tones, let $h^n_{l,m,k}$ denote the channel response between the $l$th BS and the $k$th user in the $m$th cell in tone $n$. Let $P^n_l$ denote the power allocation at the $l$th BS and $n$th tone. The multi-cell downlink weighted rate maximization problem is

$$\begin{aligned} \max \quad & \sum_{l=l}^{M}\sum_{k=1}^{K}\alpha_{lk}R_{lk} \\ \text{s.t.} \quad & R_{lk}=\sum_{n\in\mathcal{N}_{lk}}\log\left(1+\frac{P^n_l|h^n_{l,l,k}|^2}{\sum_{j\neq l}P^n_j|h^n_{j,l,k}|^2+1}\right) \end{aligned} \tag{23}$$

where $\mathcal{N}_{lk}$ denotes the set of frequency tones in which the $k$th user in the $l$th cell is scheduled. Here, $\alpha_{lk}$ signals the priority of each user, whose value is typically determined by higher-level protocols, and the background noise variance is assumed to be one without loss of generality. Further, either peak or total power constraints are typically imposed in addition.

Numerically finding the global optimal solution to the above optimization problem is known to be a difficult problem [80]. No convex reformulation of the above problem is known, even in the simpler case of fixed scheduling. In [81], [77], [78], [82], an approach which iterates between scheduling and power allocation has been proposed, but the core difficulty, namely the nonconvexity of the signal-to-interference-and-noise (SINR) expression remains.

One approach for solving the power allocation problem is to let each cell independently optimize its own transmission power in a game theoretical model, where the multiple cells eventually converge to a competitive optimum (e.g., [83], [84], [85]). However, further performance gain can be obtained if cells cooperate.

One idea is to encourage an interfering transmitter to lower its transmit power whenever it causes too much interference to neighboring transmitter-receiver pairs. Toward this end, a promising approach is to devise a mechanism to measure the impact of each transmitter's interference on its neighbors' transmissions, then to coordinate BSs based on the exchange of these measures. This idea is called *interference pricing*, which has been proposed for the power spectrum adaptation problem for the wireless ad-hoc network [86], [87], [88], [89], the digital subscriber line network [90], and is also applicable to wireless multi-cell networks [78], [82]. As shown in these studies, coordinating power control can already yield appreciable improvement in the overall sum rate as compared to a non-coordinated system.

*2) Coordinated beamforming:* When the BSs are equipped with multiple antennas, the availability of additional spatial dimensions allows the possibility of coordinating beamforming vectors across the BSs, further improving the overall performance. This idea has been explored in [91], [76], [92], [93], [94].

The optimization problem associated with multi-cell joint scheduling, beamforming and power allocation inherits the nonconvex structure of the multi-cell power control problem discussed above. However, there is a particular formulation that enjoys efficient and global optimal solution — this is when the problem is formulated as the minimization of the transmit power across the BSs subject to SINR constraints in a frequency flat channel for the case where the remote users are equipped with a single antenna only. This formulation is most applicable to constant bit-rate applications with fixed quality-of-service constraints.

Let $\mathbf{w}_{l,k}$ be the downlink transmit beamforming vector for the $k$th user in the $l$th cell, the downlink SINR for the $k$th user in the $l$th cell can be expressed as:

$$\Gamma_{l,k} = \frac{|\mathbf{h}_{l,l,k}^{\dagger}\mathbf{w}_{l,k}|^2}{\sum_{n\neq k}|\mathbf{h}_{l,l,k}^{\dagger}\mathbf{w}_{l,n}|^2 + \sum_{j\neq l,n}|\mathbf{h}_{j,l,k}^{\dagger}\mathbf{w}_{j,n}|^2 + 1} \tag{24}$$

where $\mathbf{h}_{j,l,k}$ is now the vector channel from the $j$th BS to the $k$th user in the $l$th cell. Let $\gamma_{l,k}$ be the SINR target for the $k$th user in the $l$th cell. We can formulate, for example, a total downlink transmit power minimization problem as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{l=1}^{M}\sum_{k=1}^{K} ||\mathbf{w}_{l,k}||^2 \qquad (25)\\ \text{subject to} \quad & \Gamma_{l,k} \geq \gamma_{l,k}, \quad \forall l = 1\cdots M,\ k = 1\cdots K \end{aligned}$$

where the minimization is over the $\mathbf{w}_{l,k}$'s, which implicitly include both transmit direction and transmit power optimization. For simplicity, we assume that the set of SINR targets are feasible.

Intuitively, coordinating beamforming vectors across the BSs are beneficial when the number of BS antennas exceeds the number of simultaneous users in each cell, in which case the BS has spare spatial dimensions for interference mitigation. In the case where the number of spare dimensions exceeds the number of dominant interferers in every cell, a complete nulling of interference within each cell is possible using a per-cell zero-forcing solution. Insight into the optimal cell loading under coordinated beamforming has been obtained in [95] using large systems analysis.

The key challenge to coordinated beamforming is to coordinate the BSs in such a way as to enable them to find an optimal solution jointly without excessive exchange of channel state information. This turns out to be possible using a tool known as uplink-downlink duality.

The transmit downlink beamforming problem for the multi-cell system is first considered in the classic work of [96], where an algorithm for iteratively optimizing the beamforming vectors and power allocations is proposed. The key idea is to consider a virtual dual uplink network with transmitters and receivers reversed (so that the uplink channels are the Hermitian transpose of the original downlink channels). The algorithm of [96] proposes to use the optimal uplink receiver beamformers (which are easy to find using the minimum mean-squared error (MMSE) criterion) as the downlink transmit beamformer, then to iterate between the beamformer update step and the power update step to satisfy the target SINRs. The optimality of this algorithm can be established for the single-cell network using several different techniques based on convex optimization methods [97], [98], [99], [100], [101]. In particular, the semidefinite relaxation approach of [98] and the second-order cone programming reformulation of [101] also lead to new and more efficient numerical algorithms for finding the optimal beamformers and powers. Further, it is possible to show that uplink-downlink duality is an example of Lagrangian duality in optimization [59].

The use of convex optimization ideas for establishing duality and for optimal beamforming can be extended to the multi-cell setting [102], [92]. One consequence of the duality result is that it suggests a way of implementing optimal multi-cell beamforming and power control in a distributed fashion for a time-division duplex (TDD) system, where channel reciprocity guarantees that the actual uplink channels are identical to the virtual dual uplink channels. In this case, the optimal transmit beamformers for the downlink can just simply be set as the MMSE receive beamformers for the uplink. Together with a distributed downlink power control step, this provides a distributed and optimal solution to the problem (25) [92].

The duality result can be further extended to account for the optimization objective of minimizing per-BS or per-antenna powers. The idea is to set up the optimization problem as that of minimizing the weighted sum power, where the weights can be adjusted to tradeoff powers among different BS antennas, and where the weights enter the dual channel as scaling factors for the dual virtual noise variances [59], [92]. In addition, duality also holds for the case where the remote users are equipped with multiple receive antennas as well [103]. However, the iterative updating of transmit beamformer, receive beamformer and the power is no longer guaranteed to converge to the global optimal solution; only a local optimal solution is guaranteed in this case.

Finally, duality holds not only for the power minimization problem formulated in (25), but also for the complementary problem of rate region maximization subject to power constraints (e.g., [104], [105]). This latter problem is of interest in variable rate-adaptive applications. However, as both uplink and downlink networks are MIMO multi-cell interference networks, finding the optimal solution to either problem is a challenging task. In this realm, [106] used an approach based on the first-order condition of the optimization problem, and [107] used a rate profile approach to reach the boundary points of the rate region. In addition, much work has also been done to identify solutions from a competitive (e.g., [108]) or egotistic vs. altruistic points of view [109], [93], [110].

Although competitive optimal solutions are not global optimal solutions for the entire network, they nevertheless can offer improvement over existing static networks.

*3) Coding for interference mitigation:* So far, we have focused on transmission strategies which treat intercell interference as noise. For interference-limited networks, it is possible to further improve these strategies by considering the possibility of detecting then subtracting the interference. In currently deployed cellular networks, interference signals are typically too weak to be detected by out-of-cell users. The key to make interference decoding work is to specifically design transmit signals to facilitate detection at neighboring cells — as suggested by information theoretical results on the interference channel.

The largest known achievable rate region for the two-user interference channel is the celebrated Han-Kobayashi region [69] derived based on the idea of splitting each user's transmit signal into a common message, which is decodable by all receivers, and a private message, which is decodable by the intended receiver only. In other words, by lowering the rate of part of the transmitted message to allow it to be decoded by out-of-cell users, the overall interference level would be reduced, enabling a higher overall rate. The recent work of [111] provides further insights into this scheme by showing that a particular common-private splitting can get within one bit to an outer bound of the two-user interference channel. The key insight is to set the private message power seen at the opposite receiver to be at the background noise level, whereas anything above that should be decoded. Although the outer bound of [111] applies only to the two-user single-antenna case, in a multi-cell MIMO network, adjacent cells can be paired and the optimal beamforming and power splitting problem can be solved together to produce significant performance gain for the overall network [112].

Finally, for an interference channel with more than two transmitters, it is also possible to specifically design transmit signals so that the interferences are always constrained at confined subspaces at each receiver. This allows the receiver to efficiently reject the interference. This idea, known as interference alignment, has been shown to achieve significantly improved multiplexing gain for the MIMO interference network, where both the transmitters and the receivers are equipped with multiple antennas [72]. Practical implementation of these ideas for wireless networks is an active area that is currently attracting much research.

### *B. Coding strategies for MCP networks*

The coding and optimization strategies considered in the previous sections require the sharing of the channel state information only. Significant further improvement in data rates is possible, if, in addition, the BSs are synchronized and the data streams for all the active users or the received signals at all antennas are shared between the BSs via high-capacity backhaul links [113], [114], [115], [116], [117], [118]. This setup is illustrated in Fig. 3. Many coding strategies have been proposed in the literature for this setting (e.g., [119], [120], [121], [122], [123].) The antennas from all the BSs are in this case effectively pooled together to form a giant antenna array. The uplink channel can then be modeled as a multiple access channel with multiple transmitters and a single multi-antenna receiver. The downlink channel can be modeled as a broadcast channel with a single multi-antenna transmitter and multiple receivers.

*1) Uplink:* The capacity region of the uplink multiple-access channel is achieved with superposition coding and successive decoding [15]. The idea is to decode each user's codeword based on the observation sequence of the entire antenna array (using a linear beamformer across the BSs), then to subtract the decoded codewords in a successive fashion. To achieve this multiple-access channel capacity region, the cooperating BSs theoretically need to share their observation sequence, which requires infinite backhaul capacity.

There is an important special case where the multiple-access channel capacity can be approximately achieved by just the linear detection of each user's individual message using a receive beamformer across all BSs, without the nonlinear successive decoding step. This happens when the interfering links are much weaker than the direct links (but the interference level is still much stronger than the background noise). Consider the following example where the channel matrix $\mathbf{H}$ between the $K$ single-antenna BSs and $K$ remote users is near diagonal:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}. \tag{26}$$

The capacity region of this multiple-access channel is almost a rectangle with each user achieving close to its interference-free capacity. This is because a joint receiver across the BSs can simply employ a zero-forcing receiver with rows of $\mathbf{H}^{-1}$ as the beamformers. As $\mathbf{H}^{-1}$ is nearly diagonal, it produces minimal noise enhancement. Thus, the single-user interference-free capacity can be nearly achieved for all users with just linear decoding, without the successive decoding step.

Note that for the diagonally dominant network, the above network-wide zero-forcing strategy is superior to an alternative strategy where each BS performs detection based on the received signal at its own antennas only, but BSs share the decoded bits for interference subtraction. In this case, the BSs must follow a particular decoding order with intercell interference subtracted successively. This alternative strategy is clearly suboptimal, because it achieves the single-user interference-free bound only for the last user in any particular decoding order, but not for earlier users. In contrast, the linear strategy mentioned earlier achieves the single-user bound simultaneously for every user in a diagonally dominant interference network.

*2) Downlink:* The capacity region of the downlink broadcast channel is achieved with a dirty-paper coding strategy at the encoder [14]. The idea is to fix an encoding order, then transmit each user's codeword using a transmit beamformer across all the antennas at all cooperating BSs, and successively encode each user's codeword while treating the messages already encoded as known interference. From an information theoretical point of view, the known interference can be completely pre-subtracted without using extra power at the transmitter. This is called dirty-paper coding [124]. Dirty-paper coding can be approximately implemented in practice using Tomlinson-Harashima precoding or lattice precoding strategies (see e.g., [125], [126]).

When the channel matrix associated with the interference network is diagonally dominant, the zero-forcing strategy is again near optimal. Consider again the single-antenna case:

$$\mathbf{y} = \mathbf{H}^{\dagger}\mathbf{x} + \mathbf{z}. \tag{27}$$

The zero-forcing strategy precodes $\mathbf{x} = (\mathbf{H}^{\dagger})^{-1}\mathbf{u}$, where $\mathbf{u}$ is the information symbol. When $\mathbf{H}^{\dagger}$ is near diagonal, it produces minimal power enhancement at the transmitter, resulting in a near rectangular achievable rate region. Note that this network-wide zero-forcing strategy requires joint transmit beamforming across the BSs, but no dirty-paper coding. This is again superior to the alternative strategy of dirty-paper coding without joint beamforming for the diagonally dominant interference network, analogous to the uplink case discussed earlier.

*3) Optimization:* For a cellular network with an arbitrary topology and a general channel matrix, the optimization of a network-wide beamforming vector together with the successive decoding or dirty-paper precoding orders becomes a relevant question. Consider first the uplink channel:

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_k + \mathbf{z} \tag{28}$$

where $\mathbf{y}$ is the network-wide receive signal, and $\mathbf{x}_k$ is the transmit signal for user $k$, who may be equipped with multiple antennas as well, and the noise vector $\mathbf{z}$ has a normalized unit variance. Let the optimization problem be formulated as that of maximizing the weighted sum rate $\sum_k \alpha_k R_k$. Because of the polymatroid structure of the multiple-access channel capacity region, the optimal decoding order is completely determined by the relative values of $\alpha_k$ [127]. The user with the smallest $\alpha_k$ should be decoded first; the user with the largest $\alpha_k$ last.

Without loss of generality, let $\alpha_1 \leq \alpha_2 \cdots \leq \alpha_K$. The resulting weighted sum rate can be expressed as

$$\sum_{k=1}^{K} \alpha_k R_k = \sum_{k=1}^{K} \alpha_k \log \frac{\det\left(\sum_{j=k}^{K} \mathbf{H}_j \mathbf{S}_j \mathbf{H}_j^{\dagger} + \mathbf{I}\right)}{\det\left(\sum_{j=k+1}^{K} \mathbf{H}_j \mathbf{S}_j \mathbf{H}_j^{\dagger} + \mathbf{I}\right)} \tag{29}$$

where $\mathbf{S}_k$ is the transmit covariance matrix of user $k$. The above rate expression is a convex function of $\mathbf{S}_k$. Thus, the weighted sum rate optimization problem for the uplink can be solved efficiently. The eigenvectors of the resulting optimal $\mathbf{S}_k$ give the optimal transmit beamformers. The network-wide receive beamformers for user $k$ are the MMSE beamformers with interference from the first $k-1$ users subtracted.

For the downlink channel

$$\mathbf{y}_k = \mathbf{H}_k^{\dagger}\mathbf{x} + \mathbf{z}, \tag{30}$$

(where again the noise variance is normalized to one), although a straightforward formulation of the achievable rate region does not result in a convex formulation, a key result known as uplink-downlink duality [128] enables the downlink transmit covariance optimization problem to be translated to the uplink. Uplink-downlink duality guarantees that the capacity region of the downlink channel is identical to the capacity region of the dual uplink, where the transmitters and the receivers are interchanged, and the channel matrices are Hermitian transpose of each other, and where the same sum-power constraint is applied to both. Thus, to find the optimal downlink beamformer, one only needs to solve the optimal uplink problem with a sum power constraint, then use the covariance transformation technique of [128] to translate the optimal uplink solution to the downlink.

The duality result established in [128] solves the optimal downlink beamforming problem with a sum power constraint across all the antennas. In a multi-cell network, the power usages across the BSs cannot easily be traded with each other. In addition, each antenna is typically constrained by the linearity of its power amplifier, and hence is peak power constrained. Thus, a more sensible approach is to apply a per-BS or per-antenna power constraint at each cell.

The uplink-downlink duality result can be generalized to accommodate the per-antenna power constraint [59] as mentioned in Section III. The additional ingredient is to recognize that transmit power constraints for the downlink are reflected in the dual uplink as the noise covariances. In particular, for the weighted per-antenna power minimization problem for the downlink, its dual uplink would have its noise variances scaled by the same weights. Further, to enforce per-antenna power constraints, one would need to search over all such weights in the downlink. This amounts to searching over all possible noise variances.

More precisely, for a downlink broadcast channel with per-antenna power constraint $P_i$ in each of its antennas, the dual uplink is a multiple-access channel with the same sum power constraint $\sum_i P_i$, but whose noise covariance matrix is a diagonal matrix with $q_{ii}$ on its diagonal and constrained by

$$\sum_i q_{ii} P_i \leq \sum_i P_i. \tag{31}$$

Numerically, the weighted rate sum maximization problem for the downlink becomes a minimax problem in the uplink with maximization over uplink transmit covariances and minimization over uplink receiver noise covariances. This minimax problem is concave in transmit covariance and convex in noise covariance, so it can be solved using convex optimization techniques.

The discussion so far focuses on capacity maximization. When practical coding and modulation schemes are used, an SNR gap needs to be included in the achievable rate computation. Unfortunately, accurate expressions of the SNR gap in the multiuser setting are not easy to obtain. Furthermore, although duality still holds with the inclusion of gap, the dual uplink problem is no longer tractable. The issue is that with an additional gap term, (29) is no longer a concave function of the transmit covariance matrices. Work on finding the approximate optimal ordering and beamformers for the single-receive-antenna case includes [129], but the optimization problem in its full generality remains open.

### C. Coding Strategies with Rate-Limited Cooperation

In this section, we focus on channel models where the BSs cooperate via rate-limited backhaul links as in Fig. 4-(b), or via independent relay nodes with rate-limited connections to the BSs. These channel models are practically relevant, but the information theoretical capacities of these channels are often unknown, except for certain simplified models as mentioned in Section III. This situation is not really surprising considering the fact that the capacity of even the simplest single-transmitter single-receiver and single-relay channel is still open. Thus, instead of capacity analysis, this section focuses on effective interference mitigation techniques in these settings.

*1) Receiver Cooperation:* In the uplink direction, receiver cooperation can be realized either with a dedicated relay node with fixed-capacity links to the BSs, or with rate-limited conferencing links between BSs which act as relays for each other. In these so-called relay-interference channels, the objective of the relay strategy is typically to mitigate interference, rather than to enhance direct transmission. Well-known strategies such as decode-and-forward and compress-and-forward can both be employed toward this goal.

Consider first a two-user interference channel employing Han-Kobayashi style common-private information splitting. Consider a practical regime of interest where the interfering links are "weak", but where interference is still stronger than background noise. In this case, the rates of the common messages are typically constrained by the interfering links. Thus, when the receivers are equipped with conferencing links, the common message rates can be effectively increased if each receiver decodes the common message from its own transmitter, then forwards a bin index of the common message to the other receiver. Such a decode-and-forward strategy allows each conferencing bit to increase the common information rate (and hence the overall achievable rate) by one bit, up to a limit. This strategy can be shown to be sum capacity achieving in the asymptotic high SNR regime for a simpler Z-interference channel [40]. A more sophisticated coding strategy, which consists of a two-round conferencing with quantization as the first step and binning as the second step, can in fact be proved to be within 2 bits to the capacity region of this channel model for all interference regimes [34].

The decode-and-forward strategy discussed above can be thought of as an *interference-forwarding* strategy, as the relay decodes and then forwards part of the signal that would have caused interference. The knowledge of the interference can either help the interfered transmit-receive pair subtract the interference, hereby increasing its direct transmission rate, or help the interfering transmitter-receiver pair increase its common message rate. This interference-forwarding strategy has been used in various studies, including interference channel models with a dedicated relay node [130], [131], [132], [133], [134].

In existing multi-cell networks where the Han-Kobayashi style common-private information splitting is not deployed, interference mitigation can be effectively carried out using compress-and-forward or amplify-and-forward strategies. An interesting result in this area is due to the works [135], [136], [137] that show that when a relay observes the precise interference sequence of a transmitter-receiver pair, every relaying bit to the receiver can increase the direct transmission rate by one bit in the noiseless limit. This can be achieved using a compress-and-forward strategy where the relay quantizes its observation of the interference with Wyner-Ziv coding [138], and the receiver first decodes the quantized version of the interference, then subtracts part of the interference before decoding the direct transmission. In fact, the asymptotic optimality of compress-and-forward in the noiseless limit continues to hold when the relay observes a linear combination of the transmitted signal and the interference. This idea can be further extended to show that a single relay can help both transmitter-receiver pairs of an interference channel using a universal strategy called generalized hash-and-forward [139]. Interestingly, although amplify-and-forward is typically not optimal in these settings, the amplify-and-forward strategy can be significantly improved with nonlinear amplification [140].

*2) Transmitter Cooperation:* In the downlink direction, when the BSs are equipped with rate-limited backhaul links at the transmitter, the BSs can still cooperate using a variety of techniques. One idea is to share part of the common information among the transmitters (assuming a Han-Kobayashi coding strategy is deployed), which allows the transmitters to cooperatively send shared common messages; this idea has been pursued in [141], [142]. Another idea is to share part of the private message, which allows the possibility of partial zero-forcing or dirty-paper coding at the transmitter; these possibilities have been explored in [143], [105], [144], [142]. In certain high SNR and interference-limited asymptotic regimes, it is possible to show that each cooperation bit can improve the direct transmission rate sum by one bit [144]. However, in general, the question of which transmission strategies should be adopted in specific cases remains very much open.

## V. SCALABLE COOPERATIVE SCHEMES

The potential benefits associated with exploiting or eliminating interference in cellular networks are huge. However, there are several practical hurdles which need to be overcome, over which we now draw the interested reader's attention.

In this section, we address the important issue of *scalability*. The first models of base station cooperation were centralized in nature, and a natural implementation would consist of a central processing unit, or controller, to which all the base stations are directly connected. The downside to this is that it has a single point of failure and would be an expensive infrastructure to build. Such an architecture would place enormous demands on the back-haul network, as all traffic would have to be routed to and from the central node, causing excessive delays. Besides the problem of user data sharing, there is also the issue of channel state information at the transmitter (CSIT) which also must be shared amongst base stations, and between mobiles and base stations. This is an additional signaling burden associated with MCP. Thus, when it comes to an assessment of the real advantages of MCP in realistic networks, a fundamental question arises: Might it be that the capacity increase due to MCP is outweighed by the signaling overhead it implies?

The information theoretic picture, examined in Section III, reveals that the capacity of the backhaul should grow in proportion to the capacity targeted on the over-the-air section of the network, to avoid being a bottleneck for traffic. Nevertheless, the complete answer to our question seems highly system and scenario dependent and is the focus of ongoing research. A simpler yet related problem would be: how to design practical

MCP schemes whose overhead scales favorably when the size of the network grows large? This section considers research that has attempted to reduce the overhead required to achieve most of the benefits of cooperation.

One can distinguish two lines of research devoted to overhead reduction. The first deals with deriving efficient representations of the channel state information, which is conveyed to precoding and decoding algorithms. In the second, (perfect or possibly partial) CSI is assumed and attention is focused instead on implementing scalable cooperation schemes via *distributed* precoding and decoding algorithms. There is not a great deal of difference between trying to obtain efficient channel representations in multi-cell MIMO or in MU-MIMO setups. Since a rich body of literature already exists for this problem, we simply refer the reader to past special journal issues on this topic such as [16], [145]. As a note of caution we point out that existing work on limited CSIT representation for MU-MIMO systems does not take into account the specifics of the multi-cell MIMO channel, such as the different channels from each base station to each user, whose path loss coefficient depends on the user's location in the network. In what follows, we assume that a CSI model already exists at the base stations through feedback channels. We present some concepts related to *distributed precoding and decoding* and *clustering*.

### A. Impact of channel uncertainty

*1) Network capacity:* As discussed in Section III-C5, channel uncertainty affects the scalability of point to point MIMO channels. The number of transmit antennas that can effectively be used is limited by the coherence duration in symbol times. What are the implications for network MIMO? Recently, this issue has been explored in [146], where random matrix theory is exploited to obtain tractable formulas for per-cell rates, involving parameters such as the number of base stations and the coherence duration in symbol times. It is claimed that the per-cell rate can in some cases decrease with the number of base stations, due to the cost of measuring the extra channel parameters. This conclusion may impact the optimal cluster size to use in network MIMO (see Section V-C for a discussion of clustering in the context of MCP). On the other hand, this analysis does not take into account the impact of intercluster interference, leaving open further research on this issue.

*2) Downlink: Distributed precoding with partial information sharing:* The general problem of distributed multi-cell precoding, whereby the $l$-th base station must design optimally its transmit beamforming vectors on the basis of partially shared CSIT and partially shared user data is largely open. Interestingly, in the case of fully shared user data (MIMO cooperation), this problem can be shown to fall within the framework of team decision theory, which reviews optimization problems in which different agents (here, the base stations) must act cooperatively despite not sharing the same view of the system state [147]. In our context, the problem can be formalized as follows: the users are assumed to feedback their channel state information to all base stations, in a broadcast fashion. As the distance between the user and surrounding bases differ, the quality of feedback for a given channel coefficient is unequal at different base stations. An optimization problem, by which the beamforming vectors are designed taking into account the locally received CSIT feedback as well as the expected quality level for the feedback received at other bases, is formulated [147]. The obtained beamformers can range smoothly from fully distributed to fully centralized, depending on the feedback model.

Rather than a partial sharing of CSIT along with fully shared user data, another particular framework for distributed precoding assumes a partial sharing of the user data, but under perfect CSIT sharing. A possible practical model for this is as follows: the finite backhaul links are used to convey two types of traffic. The first type is routed in the interference-coordination mode, i.e. a message to user $k$ in cell $m$ is routed to base station $m$ alone, while the second type of traffic is duplicated to all cooperating bases, in the MIMO fashion. The first and second types are referred to as private and common, respectively. An optimization problem can be formulated by which the total user rate is optimally split across private and common information, as a function of the finite backhaul capacity and of the channel state information [148]. By comparing private and common information rates, one can assess the value of MIMO cooperation depending on the interference strength model.

### B. Distributed processing using Turbo Base Stations

*1) Uplink: distributed decoding:* We now consider the problem of uplink decoding of multiple base station signals jointly. The fact that the complexity of the general multi-user detection problem grows exponentially with the number of users [149] raises a question of scalability: A priori, it looks as though the multiuser decoding of all users might be intractable as the size of the network grows large. On the other hand, the very localized structure of the interference offers hope of salvation from the apparent intractability, and it motivates the search for decentralized algorithms.

An interesting question to ask is whether the global uplink task of demodulating all the users' data symbols can be distributed across a network of interacting base stations. In this context, each base station is individually performing local computations, and then passing the results to immediate neighbors for further processing. It is very natural to try and apply well known message passing techniques from coding theory, such as the iterative method of Turbo decoding.

A first step in this direction is taken in [150], which considers an uplink multi-user detection (MUD) problem. Each base station first does an independent MUD to try and separate the desired same-cell user from the co-channel interferers in other cells. However, the desired user is also heard at the neighboring base stations, and to gain the benefits of macrodiversity, the base stations share their log-likelihood ratios. The base station controller computes an *a posteriori* log-likelihood ratio using the log-likelihoods from the local base stations, which is in essence the first step of the Turbo-decoder. Later works, such as [151], provided an explicit connection to Turbo-decoding, and propose iterative, message passing algorithms.

A related problem is to find the most likely sequence of bits transmitted in the network. This problem has a simplified

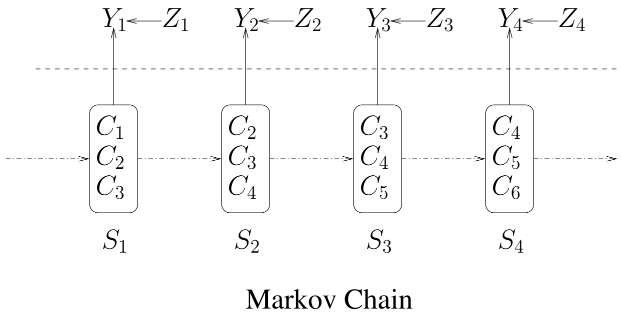



Fig. 9. Hidden Markov model of linear cellular array

structure due to the local interference coupling, which may make it amenable to a solution via dynamic programming [152].

Turbo decoding is an example of belief propagation on a graph. Communication on the uplink of a one dimensional, linear cellular array model, with one user per cell (as considered in [8], [9]), can be modeled by a Markov chain moving to the right along the linear array, see Figure 9. Each state of the Markov chain corresponds to the choice of code words in three consecutive cells. For example, state $S_1$ in Figure 9 consists of the codewords chosen by users 1,2, and 3, respectively. Each base station observes a noisy version of the superposition of the three signals, so it is a hidden Markov model, and a one-dimensional probabilistic graph can be associated with this model. The (forward-backward) BCJR algorithm [153] can be applied to compute the MAP estimates of the codewords [154]. This is a one dimensional graphical model, with clustering to provide the Markov structure. Although simple, this model allows an exploration of issues such as distributed computation, parallelism, complexity, and accuracy [154], [155], [156], [157], [158], [30], [159].

Since the complexity of the BCJR grows exponentially with the size of the state space, Gaussian models are considered in [155], [156], where linear estimation techniques are optimal. The analogous problem is Kalman smoothing, and a forward-back Kalman smoother is proposed. Note that the delay and complexity are linear in the network size, but the local nature of the interference can be exploited. In [157] a parallelized version of the forward-backward algorithm is proposed, which allows base stations to make estimates at any time. If the coupling between base stations is weak, or the noise is strong, then accurate estimates are obtained after a small number of message passings. Thus, the complexity and delay, per base station, need not grow with the array size in practice.

More realistic two dimensional cellular array models are more problematic. Forward-backward methods no longer apply, and the associated probabilistic graph models now have loops. The uplink decoding problem is considered in [30], and two graphical models are investigated. The belief propagation algorithm is applied, and it is found that in spite of the loops, error rates near the single user lower bound are obtained, for fading channels. The numerical complexity per base station is a constant, independent of the network size.

Since the complexity of the sum-product algorithm (*i.e.* belief propagation) grows exponentially with the number of variable nodes connected to a function node, it is of interest to look for suboptimal approaches that reduce the complexity, especially when there are many interfering users per cell. In this case, the computation of the log-likelihood messages sent from a variable node to a function node is in essence an MUD computation. In [160], symbols are grouped, and *a posteriori* probabilities are computed within a group, treating the interference from the other groups as Gaussian noise, with the mean and variance determined from the *a priori* probabilities. Thus, a reduced complexity group-wise MUD scheme is proposed. This paper also incorporates an LDPC (Low Density Parity Check) code, so that the graph is in time as well as space.

*2) Downlink: Distributed beamforming:* The downlink is a broadcast channel in which all base station antennas are pooled. If attention is restricted to linear techniques, then the problem to be solved is that of macroscopic beamforming. As was exposed earlier in Section IV, duality between the uplink and downlink allows some of the above methods to be used on the downlink, also. The downlink beamforming problem can be recast as an equivalent, virtual, uplink estimation problem, in which the downlink data symbols to be transmitted become observables in the uplink problem, and belief propagation on the virtual uplink graph finds the samples to be transmitted by each base station, *i.e.* the outputs of the global precoder. These samples are obtained by the sum-product message passing algorithm [159].

### C. Limited cooperation via clustering

Current cellular networks typically connect base transceiver stations (BTS's) to base station controllers (BSC's), and in some implementations the BSC handles the base-band signal processing and encoding/decoding [150]. It is therefore very natural to consider clustered models, in which the processing is done locally at the BSC, which is connected to the adjacent base stations. The collection of base stations served by a BSC forms a cluster, each cluster behaving as a network MIMO system, but now there is interference between adjacent clusters. In this case, there is not a single, centralized node, but many nodes, each independently encoding and decoding signals for the mobiles in the local cluster. The advantages of the clustered model are 1) relevance to currently implemented systems 2) reduced computational complexity 3) reduced demands on the back-haul network since only neighboring base stations (i.e. those which are mutually interfering the most) are engaged in cooperation, and 4) increased robustness to node failures (a base station can be served by more than one system controller). The disadvantages when compared with full network-wide cooperation are 1) increased levels of intercell interference in some areas (since adjacent clusters will interfere), 2) reduced diversity, and 3) lower capacity. Tradeoffs between these factors have been considered in a number of research papers.

It is well known that network MIMO has the capability to eliminate intercell interference. In models in which interference is treated as noise, a notion of effective bandwidths can be developed, which allows a definition of user capacity

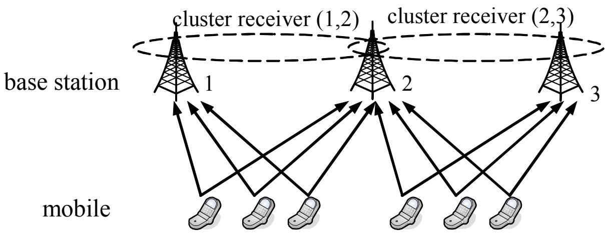



Fig. 10. Cluster decoding using pairs of base stations

region [6], [161], [7]. It can be shown that in these models, base station cooperation and optimal power control effectively eliminate inter-cell interference [162], [5], [6], [7]. In other words, the user capacity region of a network of $K$ cooperating base stations is the same as that of $K$ non-interfering (isolated) cells, as was pointed out in Section III. However, this relies on global cooperation. Consider instead a very simple model consisting of three base stations labelled $1, 2$, and 3, as depicted in Figure 10. Suppose base stations 1 and 2 cooperate to decode the users between them (cluster $1, 2$), and base stations 2 and 3 cooperate to decode the users between them (cluster $2, 3$). With full cooperation, there is a capacity limit on the sum of the effective bandwidths in the two cells. With the limited cooperation described here, the user capacity region is reduced, with additional constraints imposed by each cluster. It is shown in Theorem 9.6 in [5] that the user capacity region is the intersection of three regions: one corresponding to each of the two-receiver clusters, and one corresponding to the three-receiver system (the user capacity region under global processing).

Extensions of the above simple three antenna model to more complex networks, including linear and planar models, are considered in [163]. In these models, MCP "receivers" are associated with clusters of antennas, and adjacent receivers share common antennas, as in the simple three antenna model above. In [163] the focus is the information-theoretic capacity, and since the transmitters interfere at nearby receivers, the techniques used come from the theory of interference channels. Upper and lower bounds on achievable rates are derived.

The impact of clustering on the information-theoretic capacity of multi-cell processing (MCP) has been considered more recently in [38], [23], [164], and in other works surveyed in Section III-C4 (*i.e.* capacity results for the uplink, with local base station backhaul). In [23], the degrees of freedom are shown to be reduced by a factor of $N/(N+1)$ or $(N-1)/(N+1)$, when $N$ is even or odd, respectively, where $N$ is the number of base stations in each cluster. In [38] a similar result is obtained for a limiting regime in which the number of antennas at each base station grows large, in proportion to the number of users in each cell. In this case, the corresponding result is $(N-2)/N$, for $N \geq 3$: The limiting asymptotics wash out the effect of even or odd parity. Recently, [164] has undertaken a large system analysis of MCP, with clustered decoding, fair user scheduling, and a realistic path loss model.

A practical way to reduce inter-cluster interference is to use frequency planning. For example, one can employ two frequency bands, and by employing appropriate power control in each band, and alternating the roles of each band in adjacent clusters, the impact of inter-cluster interference can be mitigated, whilst maintaining full frequency re-use in each cell [165], [166]. Another approach to clustering is to limit the base station cooperation to the users that really need it *i.e.* those users near the cell boundaries. The problem then becomes that of grouping users into appropriate clusters for joint MUD. These ideas related to dynamic clustering have started to be investigated in [167], [168] among others.

Clustering reduces the information available to encoders and decoders alike. Information-theoretic approaches in which encoders and decoders are limited to knowledge of the codebooks of users in adjacent cells only are to be found in [154], [41], [37]. Clustering can also create unfairness for mobiles that happen to lie near the boundary of a cluster. A way to treat this problem is to introduce a family of clusters, so that every cell gets a turn at being on a cluster boundary. A round robin across the clusters provides fairness to all the users [38], [165], [166]. Dynamic clustering based on channel strength information also helps to mitigate the unfairness effects in the long run. Clustering has also recently been considered in the context of linear precoding. In [169] mobiles are classified as cluster interior or cluster edge users. Within a cluster, block diagonalization subject to per base station power constraints is performed, but, as in coordinated beamforming, nulls are also steered to neighboring edge users.

## VI. Real-world implementation and performance

The previous sections have addressed the theoretical performance of cooperative networks, including some non-ideal assumptions such as limited backhaul bandwidth and channel uncertainty. In this section, we discuss these and other topics related to the real-world implementation of cooperative techniques in cellular networks. We discuss the practical aspects of system implementation and present system-level simulations and prototypes which hint at the potential and problems of real-world cooperative cellular networks.

### A. System implementation

In the practical implementation of any coherent wireless communication system, issues including synchronization and channel estimation need to be addressed. In addition, downlink MU-MIMO transmission requires channel state information at the base station transmitters, and cooperative networks require an enhanced backhaul network connecting base stations with each other or with a central processor.

*1) Synchronization:* Downlink MIMO cooperation across multiple base stations requires tight synchronization so that there is ideally no carrier frequency offset (CFO) between the local oscillators at the base stations. Sufficient synchronization could be achieved using commercial GPS (global positioning system) satellite signals for outdoor base stations [170]. For indoor base stations, the timing signal could be sent from an outdoor GPS receiver using a precisely timed network protocol. In the absence of a GPS signal, each base could correct its offset based on CFO estimatation and feedback from the mobiles [171]. On the uplink, CFO results in interference between subcarriers of an OFDM transmission. Techniques for joint detection and CFO compensation in uplink coordinated multi-cell MIMO OFDM systems have been proposed in [172].

*2) Channel estimation:* Coherent combining at the receiver or coherent pre-combining at the transmitter provides SNR gain when the channel state is known. Sufficient resources must be allocated to pilot signals to ensure reliable estimation of the channel state, accounting for the fact that the estimation is performed on each transmit/receive antenna pair with no combining benefit. In the context of network coordination with spatially distributed bases, the extent of the coordination could depend on the range of reliable channel estimation, and there is a tradeoff between increasing the coordination network size at the expense of increased pilot overhead.

For estimation of channels at the transmitter in time-division duplex (TDD) networks, one can rely on the reciprocity of the uplink and downlink channels so that channel estimation on the uplink can be used for downlink transmission. In this sense, channel estimation at the receiver and at the TDD transmitter face similar challenges. However, the TDD system faces additional challenges if the number of users is much larger than the total spatial degrees of freedom, and if the users transmitting uplink data are not the same as those receiving downlink data. Pilot signals and protocols should be designed to address these issues without resulting in excessive training overhead. For example, these issues could be addressed by allowing only high priority users receiving downlink data to transmit uplink pilots, regardless of whether they have uplink data to send.

Estimation of channels at the base station transmitters in frequency-division duplex (FDD) networks face much greater challenges, as mentioned in Section V. In FDD networks, the channel estimates obtained at the mobile receiver must be conveyed to the transmitter, typically over a limited-bandwidth uplink feedback channel. While quantized channel estimates could be fed back, current cellular standards such as LTE [173] implement transmitter "codebooks" consisting of fixed precoding (i.e., beamforming) vectors. Under these standards, the mobile estimates the downlink channel and feeds back an index to its desired precoding vector. In cooperative networks, these codebooks would be designed to contain codeword vectors up to size $MJ$, where $M$ is the number of bases and $J$ is the number antennas per base. Because the codebooks for cooperative networks contain more codewords than for conventional networks, additional feedback bits will be required to index the codebooks, most likely leading to an increase in the uplink feedback rate.

Note that the problems for obtaining channel estimates at the transmitter are encountered in single-cell MU-MIMO downlink transmission, but they are more complicated in multi-cell coordinated networks due to the size of the networks and the potential latency introduced in distributing the estimates across the bases.

*3) Backhaul issues:* Strategies for rate-limited cooperation described in Sections III and V require a high-bandwidth, low-latency backhaul network for connecting the base stations with each other or with a central processor [174][143]. Compared to a conventional network with no coordination, interference coordination techniques shown in Figure 2 require the sharing of channel state information among cooperating bases. MIMO cooperation requires the sharing of both channel state information and user data. As shown in Figure 3, the data symbols of all users must be known at all cooperating bases. With coordination among a cluster of $L$ base stations, the data is sent to these base stations results in a factor of $L$ increase in the backhaul bandwidth. Compared to the exchange of data, the bandwidth required for exchanging channel state information are minimal for the case of moderate mobile speeds [175]. Of course the bandwidth requirements for exchanging channel information increase for higher mobile speeds and more frequency selective channels.

As an alternative to sending the data signals and beamforming weights separately to the bases, one could send a quantized baseband signal. A linearly quantized signal was shown to achieve a significant fraction of the ideal unquantized sum-rate performance in an uplink coordinated network [175].

In the context of the 3GPP LTE-Advanced standard [176], network coordination techniques are known as coordinated multi-point (CoMP) transmission or reception. Downlink CoMP transmission requires standardization of signaling and will be addressed as a study item starting in September of 2010 for possible consideration in Release 11 of LTE-Advanced. On the other hand, uplink CoMP reception can be implemented in a proprietary fashion, and could be introduced earlier as a vendor-specific feature.

### B. Simulated System Performance

Network coordination was studied for an indoor network with eight access points arranged in a line and using a TDD framing structure based on WiMAX [177]. Detailed simulations that model the physical layer of the network employed joint zero-forcing precoding and MMSE detection across all access points for the downlink and uplink, respectively. Results confirm that a multiple-fold increase in spectral efficiency is achievable for both the uplink and downlink with conventional channel estimation based on linear interpolation. Interpolation based on minimum mean squared error (MMSE) was also considered but was shown to have nearly identical performance. It is potentially more accurate, but because it requires the estimation of the channel's time-frequency covariance, it is also potentially less robust for higher-speed mobiles.

The downlink cooperative performance of a large multi-cell FDD network was evaluated in the context of 3GPP LTE parameters [178]. Using pilot signals with powers set according to LTE simulation assumptions, mobiles could not reliably acquire the pilots from multiple base stations, and as a result, cooperation could occur among only a limited set of bases. This was a major limiting factor, reducing the throughput gain by 50% compared to the ideal theoretical performance. Limited uplink feedback for conveying channel estimates to the base was another important limiting factor, reducing throughput by about 30%. Overall, the performance gains of network coordination in terms of mean throughput were about 20%. These relatively pessimistic results highlight the importance of designing efficient pilot signaling to enable effective channel acquisition and estimation for larger networks and higher mobility users.

### *C. Prototypes and testbeds*

The feasibility of cooperative techniques have been demonstrated in “over-the-air" networks of limited size. A downlink cooperative network with four distributed base antennas serving two users was implemented using zero-forcing precoding [179] as described in Section IV. The proposed system showed significant gains in mean sum-rate capacity (as a function of measured SINR) compared to a conventional time-multiplexed baseline.

Two outdoor testbeds for implementing network coordination have been developed under the EASY-C project (Enablers for Ambient Services and Systems Part C- Cellular networks), a collaboration between academia and industry for the research and development of LTE-Advanced technologies. One testbed in Berlin, Germany, consists of four base station sites (seven sectors) connected through a high-speed optical fiber network [180]. An even larger testbed consists of ten base station sites (28 sectors) distributed in downtown Dresden, Germany [181]. Network coordination has been recently demonstrated over limited portions of each testbed.

Using two distributed base antennas and two users, the Berlin testbed demonstrated downlink network coordination for an FDD LTE trial system [182]. It accounted for many practical implementation aspects including synchronization, CSI uplink feedback, limited modulation and coding schemes, and a finite-bandwidth backhaul connection between the bases. Zero-forcing precoding based on limited CSI feedback was implemented jointly across the two bases. The Dresden testbed demonstrated a similarly detailed field trial for an LTE uplink system, also consisting of two bases and two users [183]. MMSE detection was performed jointly across the bases. In both systems, the users had low mobility (or were stationary), and the systems were isolated so there was no intercell interference. In these relatively benign environments, network coordination was shown to provide significant performance gains over the conventional interference-limited strategy. In particular, it is claimed that MCP can provide median rate gains on the order of at least 50 percent, as well as increased fairness, and improved diversity, taking into account the practical constraints of their system.

Some testbeds are currently testing MCP principles together with the use of relays, in the scenario of so-called mesh networks [184] . Recently a large integrated research project called ARTIST4G funded by the EU and comprising over 20 academic and industrial partners throughout Europe was launched and is fully dedicated to the development of multi-cell cooperation techniques in future cellular networks. These testbeds and projects allow the exploration of system-level issues discussed in this section as well as broader issues that include hybrid ARQ, resource allocation, and user scheduling.

## VII. CONCLUSIONS AND FUTURE DIRECTIONS

Although the underlying MIMO theoretic concepts are well understood, cooperative systems are still in their infancy and much further research is required in order to fully understand these systems and to practically achieve the full benefits of multi-base cooperation. Unlike standard MIMO systems where the cost of multi-antenna processing lies in the extra hardware and software at individual devices, cooperative MIMO techniques do not necessarily require extra antennas. Rather, the cost lies in the additional exchange of information (user data and channel state) between the devices engaged in the cooperation, or between the devices and the central controller in a centralized architecture. Furthermore, the information exchange is subject to tight delay constraints which are difficult to meet over a large network. MIMO-cooperation offers additional benefits over simpler beamforming coordination schemes, but it requires user data sharing among several BSs and more complex precoding and decoding.

This tutorial began, in Section III, with an extensive review of capacity results for the classical Wyner model and its variants, including limited backhaul bandwidth, localized clustered MCP, and relay assisted MCP. The main conclusions are summarized in Section III-F. Although the Wyner model is mathematically tractable, attention must now steer to more realistic models of cellular communication.

Fading is included in the Wyner model, but the fading parameters are always assumed to be known perfectly at the mobiles and/or base stations. Future work must consider the impact of channel uncertainty, and the cost of measuring the channels in the network. Channel measurement issues may impact the optimal size of clusters in clustered MCP. Bounds on capacity under channel uncertainty are needed, and the coupling of channel uncertainty with limited backhaul bandwidth is an important area yet to be explored. Information-theoretic models provide tractable, elegant capacity formulas that are amenable to optimization, and performance bounds against which practical schemes can be compared. More importantly, however, they provide insights into the key performance bottlenecks, which can then be addressed in more practically oriented research.

Section IV reviewed the transmission and coding techniques required to approach the information-theoretic limits. This included a review of the celebrated uplink-downlink duality theory for the MIMO broadcast channel, which in a rough sense is the model for MCP on the downlink. However, network MIMO has additional constraints, such as limited backhaul bandwidth, the need for decentralized processing, and per base station power constraints. Recent research has included per base station power constraints, and introduced notions such as coordinated beamforming, along with the development of the associated Lagrangian optimization theory. Coordinated beamforming is intermediate between SCP, where only local information is used, and MCP, where global information is available to a central processor. In coordinated beamforming, the BS knows the data and channel state of the users in its own cell, but it also knows the channel state of users in adjacent cells, and this enables a joint optimization problem to be solved.

One challenge for the future is to move these ideas from theory to practice. Joint optimization across many cells may be problematic when channels are changing due to mobility. One approach may be to reformulate the problem in terms of channel statistics, rather than require instantaneous channel knowledge. Another approach may be to look for simplified, suboptimal beamforming structures which nevertheless come close to optimality in practical settings.

Other challenges addressed include coordinated power control, and the multi-cell joint problems of scheduling, power control, and rate allocation across the frequency spectrum. Many of these problems are computationally intractable, in general, and the way forward may be to look for structure in real-world networks that allows the problems to be solved in polynomial time. Recent work on fractional frequency reuse in OFDMA systems provides a new set of techniques that could be applied to network MIMO in a joint multi-cell optimization. Recent work at the cutting edge of network information theory, including interference alignment, network coding, and the recent progress on the interference channel, all provide new ways to approach the fundamental problem: how do we achieve maximum spectral efficiency in a multiple cell network?

Sections V and VI address a few of the fundamental and practical issues, such as scalability, synchronization, and channel estimation. It is highly unlikely that a future network MIMO system will be built according to a centralized architecture. Recent research has considered the problem of distributing the network-wide optimization problems, so that much of the processing can be done locally, with limited communication between nearby nodes. One option is clustered MCP, in which small clusters of BSs collaborate together on uplink decoding and downlink beamforming. Turbo base stations provide another approach, in which soft information is passed between adjacent BSs, allowing iterative, probabilistic graph-based methods to provide decentralized solutions to similar problems. Other interesting approaches lending themselves to distributed implementations are game and team decision theoretic approches.

Behind such problems, a recurrent and quite fundamental issue is associated with the aquisition of channel state information. An important open question is to determine just how much channel state information is needed at each particular node in the network, including information that has been measured at other nodes in the network. This question gives rise to a fundamental feedback resource allocation problem. Cooperation gains go at the expense of feedback resource, hence such a cost is justified when interference is strong enough. More generally, a fundamental trade-off between cooperation and information exchange exists which remains to be explored theoretically.

Another important problem in practice is that of synchronizing the BSs so that there is no carrier frequency offset. GPS offers one approach, but future work may consider methods of distributed clock synchronization. From a practical point of view, distributed precoding and decoding at multiple bases, which are designed to offer cooperation gains while exploiting primarily local channel state and user data information are of high interest and will attract significant research efforts in the years to come.

## VIII. Acknowledgements

The authors gracefully acknowledge the many constructive comments made by the anonymous reviewers on this paper. David Gesbert acknowledges the partial support of the European Commission seventh framework programme (FP7/2007-2013) under grant agreements no 247223 (ARTIST4G). Wei Yu acknowledges the support of Natural Science and Engineering Research Council (NSERC) of Canada through the Canada Research Chairs Program. Stephen Hanly acknowledges support from a National University of Singapore startup grant.

## References

[1] G. Kramer, I. Marić, and R. D. Yates, "Cooperative communications," *Found. Trends Netw.*, vol. 1, no. 3, pp. 271–425, 2006.

[2] D. Gesbert, M. Shafi, P. Smith, D. Shiu, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J Sel. Areas Commun.*, special Issue on MIMO systems, Vol. 21, No. 3, pp281-302, April 2003.

[3] A. J. Viterbi, A. M. Viterbi, K. S. Gilhousen, and E. Zehavi, "Soft handoff extends CDMA cell coverage and increases reverse link capacity," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 8, pp. 1281–1288, Oct. 1994.

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

[5] S. V. Hanly, "Information capacity of radio networks," Ph.D. dissertation, Cambridge University, Aug. 1993.

[6] ——, "Capacity and power control in spread spectrum macrodiversity radio networks," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 247–256, Feb. 1996.

[7] S. V. Hanly and D. N. Tse, "Resource pooling and effective bandwidths in cdma networks with multiuser receivers and spatial diversity," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1328–1351, May 2001.

[8] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommunications Systems*, vol. 1, no. 1, pp. 1–42, Mar. 1993.

[9] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov 1994.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," AT & T Bell Laboratories, Tech. Rep. #BL0112170-950615-07TM, 1995.

[11] G. J. Foschini, "Layered space-time architecture for wireless communication," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn. 1996.

[12] S. Shamai and B. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 3, 2001, pp. 1745–1749 vol.3.

[13] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Mag.*, Vol. 24, No. 5, pp36-46, Sept. 2007.

[14] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. Wiley, 1991.

[16] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun., Vol. 26*, Oct. 2008.

[17] D. Gesbert, S. G. Kiani, A. Gjendemsj, and G. E.Oien, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proc. of the IEEE*, vol. 95, no. 5, pp. 2393–2409, Dec. 2007.

[18] A. Carleial, "Interference channels," *Information Theory, IEEE Transactions on*, vol. 24, no. 1, pp. 60–70, Jan 1978.

[19] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network mimo: Overcoming intercell interference in indoor wireless systems," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, Nov. 2007, pp. 83–87.

[20] G. Caire, S. Ramprashad, H. Papadopoulos, C. Pepin, and C.-E. Sundberg, "Multiuser mimo downlink with limited inter-cell cooperation: Approximate interference alignment in time, frequency and space," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, Sept. 2008, pp. 730–737.

[21] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, U. Spagnolini, and S. Shamai, “An information theoretic view of distributed antenna processing in cellular systems,” in *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*. Boston, MA: Auerbach Publications, CRC Press, 2007.

[22] S. Verdu, “The capacity region of the symbol-asynchronous Gaussian multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. 35, no. 4, pp. 733–751, Jul 1989.

[23] O. Somekh, B. Zaidel, and S. Shamai, “Sum rate characterization of joint multiple cell-site processing,” *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4473–4497, Dec. 2007.

[24] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. E. Smee, and R. Padovani, “Downlink macro-diversity in cellular networks,” in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, June 2007, pp. 1–5.

[25] O. Somekh and S. Shamai, “Shannon-theoretic approach to a Gaussian cellular multiple-access channel with fading,” *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1401–1425, Jul 2000.

[26] N. Levy, O. Zeitouni, and S. Shamai, “On information rates of the fading Wyner cellular model via the thouless formula for the strip,” *submitted [arXiv:0806.2991v1]*.

[27] ——, “Central limit theorem and large deviations of the fading Wyner cellular model via product of random matrices theory,” *Problems of Information Transmission*, vol. 45, no. 1, pp. 5–22, 2009.

[28] S. Verdu, “Spectral efficiency in the wideband regime,” *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1319–1343, Jun 2002.

[29] G. Caire, D. Tuninetti, and S. Verdu, “Suboptimality of TDMA in the low-power regime,” *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 608–620, April 2004.

[30] E. Aktas, J. S. Evans, and S. V. Hanly, “Distributed decoding in a cellular multiple access channel,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 241–250, Jan. 2008.

[31] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, “Uplink macro diversity of limited backhaul cellular network,” *submitted [arXiv:0805.4620]*.

[32] V. Prabhakaran, D. Tse, and K. Ramachandran, “Rate region of the quadratic Gaussian CEO problem,” in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, June-2 July 2004, p. 119.

[33] A. E. Gamal, “Achievability for discrete memoryless systems,” lecture notes, School of Information Theory, Evanston, IL, Aug. 2009.

[34] I.-H. Wang and D. N. C. Tse, “Interference mitigation through limited receiver cooperation,” submitted to *IEEE Trans. Inf. Theory*, preprint at arXiv:0911.2053, 2009.

[35] O. Simeone, O. Somekh, E. Erkip, S. Shamai, and H. V. Poor, “Multirelay channel with non-ergodic link failures,” in *Proc. IEEE Workshop on Networking and Information Theory (ITW 2009)*, June 10-12 2009.

[36] B. Nazer, A. Sanderovich, M. Gastpar, and S. Shamai, “Structured superposition for backhaul constrained cellular uplink,” in *Proc. IEEE Symposium on Information Theory (ISIT 2009)*, pp. 1530-1534, June 28 - July 3 2009.

[37] A. Lapidoth, N. Levy, S. Shamai, and M. A. Wigger, “A cognitive network with clustered decoding,” in *Proc. IEEE Symposium on Information Theory (ISIT 2009)*, June 28 - July 3 2009, pp. 596–600.

[38] N. Bacha, J. S. Evans, and S. V. Hanly, “On the capacity of MIMO cellular networks with macrodiversity,” in *Proc. Australian Communication Theory Workshop*, Feb. 2006, pp. 105–109.

[39] O. Sahin, E. Erkip, and O. Simeone, “Interference channel with a relay: Models, relaying strategies, bounds,” in *Proc. Information Theory and Applications Workshop (ITA 2009)*, Feb. 8 - 13 2009.

[40] L. Zhou and W. Yu, “Gaussian Z-interference channel with a relay link: Achievability region and asymptotic sum capacity,” submitted to *IEEE Trans. Inform. Theory* preprint available at arXiv:1006.5087, June 2010.

[41] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, “Local base station cooperation via finite-capacity links for the uplink of linear cellular networks,” *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 190–204, Jan. 2009.

[42] L. Ozarow, S. Shamai, and A. Wyner, “Information theoretic considerations for cellular mobile radio,” *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[43] R. Narasimhan, “Individual outage rate regions for fading multiple access channels,” in *Proc. IEEE International Symposium on Information Theory. (ISIT 2007)*. pp. 1571–1575, June 2007.

[44] N. Levy, O. Zeitouni, and S. Shamai, “Information theoretic aspects of users’ activity in a Wyner-like cellular model,” IEEE Transactions on Information Theory, Vol. 56, No. 5, pp 2241-2248, May 2010.

[45] R. G. Gallager, “An inequality on the capacity region of multi-access multi-path channels,” in *Communications and Cryptography—Two Sides of One Tapestry*. Norwell, MA: Kluver Academic, 1994, pp. 129–139.

[46] S. Shamai and A. Wyner, “nformation-theoretic considerations for symmetric, cellular, multiple-access fading channels. I,” *Information Theory, IEEE Transactions on*, vol. 43, no. 6, pp. 1877–1894, Nov 1997.

[47] Y. Liang and A. Goldsmith, “Symmetric rate capacity of cellular systems with cooperative base stations,” in *Global Telecommunications Conference, 2006. GLOBECOM ’06. IEEE*, 27 2006-Dec. 1 2006, pp. 1–5.

[48] O. Somekh, O. Simeone, B. M. Zaidel, H. V. Poor, and S. Shamai, “On the spectrum of large random hermitian finite-band matrices,” in *Proc. Information Theory and Applications Workshop (ITA 2008)*, Jan. 27- Feb. 1 2008.

[49] N. A. Letzepis, “Gaussian cellular multiple access channels,” Ph.D. dissertation, Univ. South Australia, 2006.

[50] T. Holliday, A. Goldsmith, and P. Glynn, “Capacity of finite state channels based on Lyapunov exponents of random matrices,” *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3509–3532, Aug. 2006.

[51] A. M. Tulino and S. Verdú, “Random matrix theory and wireless communications,” Now Publishers, 2004.

[52] N. Levy, O. Somekh, S. Shamai, and O. Zeitouni, “On certain large random hermitian jacobi matrices with applications to wireless communications,” *submitted [arXiv:0806.2674]*.

[53] A. Lozano, A. Tulino, and S. Verdu, “High-snr power offset in multiantenna communication,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[54] A. Narula, “Information theoretic analysis of multiple-antenna transmission diversity,” Ph.D. dissertation, MIT, Cambridge, MA, 1997.

[55] M. Bacha, J. Evans, and S. Hanly, “On the capacity of cellular networks with mimo links,” in *ICC ’06. IEEE International Conference on Communications*, vol. 3, June 2006, pp. 1337–1342.

[56] D. Aktas, M. Bacha, J. Evans, and S. Hanly, “Scaling results on the sum capacity of cellular networks with mimo links,” *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3264–3274, July 2006.

[57] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links,” *IEEE Trans. on Information*, vol. 49, no. 4, pp. 951–963, Apr 2003.

[58] G. Caire and S. Shamai, “On the achievable throughput of a multi-antenna Gaussian broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[59] W. Yu and T. Lan, “Transmitter optimization for the multi-antenna downlink with per-antenna power constraints,” *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[60] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, “Downlink multicell processing with limited backhaul capacity,” *EURASIP Journal on Advances in Signal Processing*, article ID 840814, Jan. 27- Feb. 1 2009.

[61] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, and S. Shamai, “Cooperative multicell zero-forcing beamforming in cellular downlink channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3206–3219, July 2009.

[62] “http://wirelessman.org/relay/index.html.”

[63] O. Simeone, O. Somekh, Y. Bar-Ness, and U. Spagnolini, “Uplink throughput of TDMA cellular systems with multicell processing and amplify-and-forward cooperation between mobiles,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2942–2951, August 2007.

[64] G. Kramer and A. van Wijngaarden, “On the white Gaussian multiple-access relay channel,” in *Proc. IEEE International Symposium on Information Theory (ISIT 2000)*, p. 40, June, 2000.

[65] O. Simeone, O. Somekh, Y. Bar-Ness, and U. Spagnolini, “Throughput of low-power cellular systems with collaborative base stations and relaying,” *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 459–467, Jan. 2008.

[66] O. Somekh, O. Simeone, H. V. Poor, and S. Shamai, “Cellular systems with full-duplex amplify-and-forward relaying and cooperative base-stations,” in *Proc. IEEE International Symposium on Information Theory (ISIT 2007)*, June 2007, pp. 16–20.

[67] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai, “Uplink sum-rate analysis of a multi-cell system with feedback,” in *Proc. Forty-Fifth Annual Allerton Conference on Communication, Control, and Computing*, Sept. 23 - 26 2008.

[68] O. Somekh, O. Simeone, H. Poor, and S. Shamai, "Cellular systems with full-duplex compress-and-forward relaying and cooperative base stations," in *Proc. IEEE International Symposium on Information Theory, (ISIT 2008)*, July 2008, pp. 2086–2090.

[69] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[70] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai, "Throughput of cellular systems with conferencing mobiles and cooperative base-stations," *Eurasip Journal on Wireless Communications and Networking*, article ID 652325, Jan. 2008.

[71] A. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 841–850, Nov 1982.

[72] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[73] C. W. T. Gou, S. A. Jafar, "On the degrees of freedom of finite state compound wireless networks - settling a conjecture by weingarten et. al," *submitted*.

[74] A. Gjendemsjoe, D. Gesbert, G. Oien, and S. Kiani, "Binary power control for sum rate maximization over multiple interfering links," *IEEE Trans. Wireless Commun.*, Vol. 7, No. 8, pp 3164-3173, Aug. 2008.

[75] S. G. Kiani and D. Gesbert, "Optimal and distributed scheduling for multicell capacity maximization," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 288–297, Jan. 2008.

[76] M. Vemula, D. Avidor, J. Ling, and C. Papadias, "Inter-cell coordination, opportunistic beamforming and scheduling," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 12, Istanbul, Turkey, June 2006, pp. 5319–5324.

[77] A. L. Stolyar and H. Viswanathan, "Self-organizing dynamic fractional frequency reuse in OFDMA systems," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Phoenix, AZ, U.S.A., April 2008, pp. 691–699.

[78] ——, "Self-organizing dynamic fractional frequency reuse for best-effort traffic through distributed inter-cell coordination," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Rio de Janeiro, Brazil, April 2009, pp. 1287–1295.

[79] V. Tralli, R. Veronesi, and M. Zorzi, "Power-shaped advanced resource assignment (PSARA) for fixed broadband wireless access systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2207–2220, Nov. 2004.

[80] Z. Q. Luo and S. Zhang, "Dynamic spectrum management: complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[81] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2835–2848, July 2009.

[82] W. Yu, T. Kwon, and C. Shin, "Joint scheduling and dynamic power spectrum optimization for wireless multicell networks," in *Proc. 44th Conf. Info. Science Sys. (CISS)*, Princeton, NJ, March 2010.

[83] Z. Han, Z. Ji, and K. J. R. Liu, "Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 6, pp. 1079–1090, Aug. 2007.

[84] T. Alpcan, T. Basar, and S. Dey, "A power control game based on outage probabilities for multicell wireless data networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 890–899, April 2006.

[85] A. Leshem and E. Zehavi, "Cooperative game theory and the Gaussian interference channel," *IEEE J. Select. Areas Commun.*, vol. 26, no. 7, pp. 1078–1088, Sept. 2008.

[86] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 24, no. 5, pp. 1074-1084, May 2006.

[87] J. Yuan and W. Yu, "Distributed cross-layer optimization of wireless sensor networks: A game theoretic approach," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, San Francisco, U.S.A., 2006.

[88] C. Shi, R. A. Berry, and M. L. Honig, "Distributed interference pricing for OFDM wireless networks with non-separable utilities," in *Proc. Conf. Info. Science Sys. (CISS)*, March 2008, pp. 755–760.

[89] F. Wang, M. Krunz, and S. Cui, "Price-based spectrum management in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 2, pp. 74–87, Feb. 2008.

[90] W. Yu, "Multiuser water-filling in the presence of crosstalk," in *Information Theory and Applications Workshop*, San Diego, U.S.A., Jan. 2007.

[91] R. Veronesi, V. Tralli, J. Zander, and M. Zorzi, "Distributed dynamic resource allocation for multicell SDMA packet access net," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2772–2783, Oct. 2006.

[92] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[93] R. Zakhour, Z. K. M. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell MIMO channels," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Barcelona, Spain, April 2009.

[94] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MIMO transmitter optimization for inter-cell interference mitigation," 2009, preprint: arXiv:0909.1344v1.

[95] R. Zakhour and S. Hanly, "Base station cooperation on the downlink: Large systems analysis," 2010, preprint: arXiv:1006.3360v1.pdf.

[96] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[97] E. Visotsky and U. Madhow, "Optimal beamforming using transmit antenna arrays," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, vol. 1, July 1999, pp. 851–856.

[98] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, 2002.

[99] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 18–28, Jan. 2004.

[100] ——, "Iterative multiuser uplink and downlink beamforming under SINR contraints," *IEEE Trans. Signal Process.*, vol. 53, pp. 2324–2334, July 2005.

[101] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[102] R. Stridh, M. Bengtsson, and B. Ottersten, "System evaluation of optimal downlink beamforming with congestion control in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 743–751, April 2006.

[103] B. Song, R. Cruz, and B. Rao, "Network duality for multiuser MIMO beamforming networks and applications," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 618–630, March 2007.

[104] J. Yang and D. K. Kim, "Multi-cell uplink-downlink beamforming throughput duality based on lagrangian duality with per-base station power constraints," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 277–279, April 2008.

[105] P. Marsch and G. Fettweis, "On downlink network MIMO under a constrained backhaul and imperfect channel knowledge," in *Proc. IEEE Global Telecommun. Conf. (Globecom)*, Nov. 2009.

[106] L. Venturino, N. Prasad, and X. Wang, "Coordinated linear beamforming in downlink multi-cell wireless networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1451–1461, April 2010.

[107] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," to appear in *IEEE Trans. Signal Processing*, 2010, preprint available at arXiv:0910.2771.

[108] G. Scutari, D. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Select. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Sept. 2008.

[109] E. Larsson and E. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE J. Select. Areas Commun.*, vol. 26, no. 9, pp. 1059–1069, Sept. 2008.

[110] K. M. Ho and D. Gesbert, "Balancing egoism and altruism on the interference channel: The mimo case," in *Proc. IEEE International Conf. on Communications (ICC)*, May 2010.

[111] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 5534–5562, Dec. 2008.

[112] H. Dahrouj and W. Yu, "Interference mitigation with joint beamforming and common information decoding in multicell systems," in *Proc. IEEE Int. Sym. Inf. Theory (ISIT)*, Austin, TX, June 2010, pp. 2068–2072.

[113] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *EURASIP J. Wireless Commun. Networking*, vol. 2004, no. 2, pp. 222 – 235, Dec. 2004, article ID 202654.

[114] H. Zhang, H. Dai, and Q. Zhou, "Base station cooperation for multiuser MIMO: Joint transmission and BS selection," in *Proc. Conf. Info. Science Sys. (CISS)*, Princeton, NJ, March 2004.

[115] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *EURASIP J. Wireless Commun. Networking*, vol. 2008, no. 5, Jan. 2008, article ID 586878.

[116] O. Simeone, O. Somekh, G. Kramer, S. Shamai, and H. V. Poor, "Throughput of cellular systems with conferencing mobiles and cooperative base stations," *EURASIP J. Wireless Commun. Networking*, 2008, article ID 652325.

[117] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[118] G. Foschini, M. Karakayali, and R. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEE Proc. Commun.*, vol. 153, no. 4, pp. 548–555, Aug. 2006.

[119] F. Boccardi and H. Huang, "Limited downlink network coordination in cellular networks," in *IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Sept. 2007, pp. 1–5.

[120] S. Venkatesan, "Coordinating base stations for greater uplink spectral efficiency in a cellular network," in *IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Athens, Greece, Sept. 2007, pp. 1–5.

[121] J. Zhang, R. Chen, J. G. Andrews, and R. W. Heath, "Coordinated multi-cell MIMO systems with cellular block diagonalization," in *Conf. Record of the Forty-First Asilomar Conf. on Signals, Systems and Computers*, Nov. 2007, pp. 1669–1673.

[122] C. Botella, G. Pinero, A. Gonzalez, and M. de Diego, "Coordination in a multi-cell multi-antenna multi-user W-CDMA system: A beamforming approach," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4479–4485, Nov. 2008.

[123] H. Huang and M. Trivellato, "Performance of multiuser MIMO and network coordination in downlink cellular networks," in *Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, Berlin, April 2008, pp. 85–90.

[124] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[125] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 3820–3833, Nov. 2005.

[126] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1220–1230, July 2005.

[127] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels. Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[128] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 768–783, May 2004.

[129] C.-H. F. Fung, W. Yu, and T. J. Lim, "Precoding for the multiantenna downlink: Multiuser SNR gap and optimal user ordering," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 188–197, Jan. 2007.

[130] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel," in *Proc. Global Telecommun. Conf. (Globecom)*, Nov. 2007, pp. 1627–1631.

[131] ——, "On achievable rates for interference relay channel with interference cancelation," in *Conf. Record Forty-First Asilomar Conf. Signals, Systems and Computers*, Nov. 2007, pp. 805–809.

[132] O. Sahin, O. Simeone, and E. Erkip, "Interference channel with an out-of-band relay," submitted to *IEEE Trans. Inf. Theory*, 2009, preprint available at arXiv:1007.0267.

[133] I. Marić, R. Dabora, and A. Goldsmith, "On the capacity of the interference channel with a relay," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 554-558, July 2008.

[134] R. Dabora, I. Marić, and A. Goldsmith, "Relay strategies for interference-forwarding," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Porto, Portugal, May 2008, pp. 46–50.

[135] Y.-H. Kim and T. M. Cover, "Capacity of a class of deterministic relay channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2007, pp. 591–595.

[136] Y.-H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1328–1329, March 2008.

[137] ——, "Coding techniques for primitive relay channels," in *Allerton Conf. Commun. Control and Computing*, Sept. 2007.

[138] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[139] P. Razaghi and W. Yu, "Universal relaying for the interference channel," in *Information Theory and Applications Workshop (ITA)*, San Diego, U.S.A., Jan. 2010.

[140] M. N. Khormuji, A. Zaidi, and M. Skoglund, "Interference management using nonlinear relays," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 1924–1930, July 2010.

[141] I. Marić, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 3536–3548, Oct. 2007.

[142] H. Bagheri, A. S. Motahari, and A. K. Khandani, "Zero-forcing for the symmetric interference channel with conferencing encoders," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, June 2010, pp. 370–374.

[143] P. Marsch and G. Fettweis, "On base station cooperation schemes for downlink network MIMO under a constrained backhaul," *Proc. IEEE Global Telecommunications Conference (Globecom'08)*, Dec. 2008.

[144] S. Hari and W. Yu, "Partial zero-forcing precoding for the interference channel with partially cooperating transmitters," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, June 2010, pp. 2283–2287.

[145] R. Heath, D. Love, V. Lau, D. Gesbert, B. Rao, and M. Andrews (Editors), "Exploiting limited feedback in tomorrow's communication networks," *Special Issue of IEEE J. Sel. Areas Commun., Vol. 26*, No. 8, pp 1337-1340, Oct. 2008.

[146] J. Hoydis, M. Kobayashi, and M. Debbah, "On the optimal number of cooperative base stations in network MIMO systems," 2010, preprint: arXiv:1003.0332v1.

[147] R. Zakhour and D. Gesbert, "Team decision for the cooperative MIMO channel with imperfect CSIT sharing," in *The Information Theory and Applications (ITA) Workshop, San Diego CA.*, Feburary 2010.

[148] ——, "On the value of data sharing in constrained-backhaul network MIMO," in *Proc. International Zurich Seminar on Communications*, March 2010.

[149] S. Verdú, *Multiuser Detection.* Norwell, MA: Cambridge University Press, 1998.

[150] M. C. Valenti and B. D. Woerner, "Iterative multiuser detection, macrodiversity combining, and decoding for the TDMA cellular uplink," *IEEE J. Sel. Areas Commun.*, vol. 19, pp. 1570–1583, Aug. 2001.

[151] T. Mayer, H. Jenkac, and J. Hagenauer, "Turbo base-station cooperation for intercell interference cancellation," in *Proc. IEEE International Conference on Communications*, vol. 11, June 2006, pp. 4977 – 4982.

[152] L. R. Welburn, J. K. Cavers, and K. W. Sowerby, "A computational paradigm for space-time multiuser detection," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1595–1604, Sep. 2004.

[153] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[154] A. J. Grant, S. V. Hanly, J. S. Evans, and R. Müller, "Distributed decoding for Wyner cellular systems," in *Proc. 5th Australian Communications Theory Workshop, Newcastle, Australia*, Feb. 2004, pp. 77–81.

[155] B. L. Ng, J. S. Evans, S. V. Hanly, and A. J. Grant, "Information capacity of Wyner's cellular network with LMMSE receivers," in *Proc. IEEE International Conference on Communications, Paris, France*, vol. 1, June 2004, pp. 583 – 587.

[156] ——, "Distributed linear multiuser detection in cellular networks," in *Proc. 5th Australian Communications Theory Workshop, Newcastle, Australia*, Feb. 2004, pp. 127–132.

[157] B. L. Ng, J. S. Evans, and S. V. Hanly, "Distributed linear multiuser detection in cellular networks based on Kalman smoothing," in *Proc. IEEE Global Telecommunications Conference, Dallas, USA*, vol. 1, Dec. 2004, pp. 134 – 138.

[158] ——, "Distributed downlink beamforming in cellular networks," in *Proc. IEEE International Symposium on Information Theory*, Jun. 2007, pp. 6–10.

[159] B. L. Ng, J. S. Evans, S. V. Hanly, and D. Aktas, "Distributed downlink beamforming with cooperative base stations," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5491–5499, Dec. 2008.

[160] S. Bavarian and J. K. Cavers, "Reduced-complexity belief propagation for system-wide mud in the uplink of cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 541–549, Apr. 2008.

[161] D. N. Tse and S. V. Hanly, "Linear multi-user receivers: Effective interference, effective bandwidth, and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[162] S. V. Hanly, "Macro-diversity for spread spectrum mobile radio: Capacity and power control," in *Proc. IEEE Second International Symposium on Spread Spectrum Techniques and Applications*, Nov. 1992, pp. 349–352.

[163] N. Levy and S. S. (Shitz), "Clustered Local Decoding for Wyner-Type Cellular Models," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4967–4985, November 2009.

[164] H. Huh, S. H. Moon, Y. T. Kim, I. Lee, and G. Caire, "Multicell MIMO downlink with cell cooperation and fair scheduling: A large system analysis," 2010, preprint: arXiv:1006.2162v1.pdf.

[165] G. Caire, S. Ramprashad, H. Papadopoulos, C. Pepin, and C.-E. W. Sundberg, “Multiuser MIMO Downlink with Limited Inter-cell Cooperation: Approximate Interference Alignmnent in Time, Frequency and Space,” in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, September 2008, pp. 730–737.

[166] E. Katranaras, M. Imran, and R. Hoshyar, “Sum-rate of linear cellular systems with clustered joint processing,” *Vehicular Technology Conference-VTC 2009, Spring, 2009*, vol. 8, no. 4, pp. 1910–1921.

[167] P. Marsch and G. Fettweis, “A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul,” in *Proc. IEEE International Conference on Communications*, June 2007, pp. 975–979.

[168] A. Papadogiannis, D. Gesbert, and E. Hardouin, “A dynamic clustering approach in wireless networks with multi-cell cooperative processing,” in *Proc. IEEE Intern. Conf. on Comm. (ICC)*, pp. 4033-4037, May 2008.

[169] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, “Networked mimo with clustered linear precoding,” *IEEE Trans. Wireless Commun.*, April 2009.

[170] V. Jungnickel, T. Wirth, M. Schellmann, T. Haustein, and W. Zirwas, “Sychronization of cooperative base stations,” *Proc. International Symposium on Wireless Communication Systems (ISWCS’08)*, October 2008.

[171] B. Zarikoff and J. Cavers, “Multiple frequency offset estimation for the downlink of coordinated MIMO systems,” *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 901–912, 2008.

[172] V. Kotzsch, J. Holfeld, and G. Fettweis, “Joint detection and CFO compensation in asynchronous multi-user MIMO OFDM systems,” *Proc. 69th IEEE Vehicular Technology Conference (VTC’09)*, vol. 2009, April 2009.

[173] S. Sesia, I. Toufik, and M. Baker, Eds., *LTE: The UMTS Long Term Evolution*. John Wiley and Sons, 2009.

[174] P. Marsch and G. Fettweis, “On uplink network MIMO under a constrained backhaul and imperfect channel knowledge,” *Proc. IEEE International Conference on Communications (ICC’09)*, 2009.

[175] D. Samardzija and H. Huang, “Determining backhaul bandwidth requirements for network MIMO,” *Proc. 17th European Signal Processing Conference (EUSIPCO’09)*, 2009.

[176] B. Clerckx, A. Lozano, S. Sesia, C. van Rensburg, and C. B. Papadias, “3GPP LTE and LTE Advanced,” *EURASIP Journal on Wireless Communications and Networking*, 2009.

[177] S. Venkatesan, H. Huang, A. Lozano, and R. Valenzuela, “A WiMAX-based implementation of network mimo for indoor wireless systems,” *Eurasip Journal on Advances in Signal Processing, special issue on Multiuser MIMO transmission with limited feedback, cooperation and coordination*, no. Article ID 963547, 2009.

[178] S. Annapureddy, A. Barbieri, S. Geirhofer, S. Mallik, and A. Gorokhov, “Coordinated joint transmission in WWAN,” *IEEE Communication Theory Workshop*, 2010.

[179] D. Samardzija, H. Huang, T. Sizer, and R. Valenzuela, “Experimental downlink multiuser MIMO system with distributed and coherently-coordinated transmit antennas,” *Proc. International Conference on Communications (ICC’09)*, June 2007.

[180] V. Jungnickel, L. Thiele, M. Schellmann, T. Wirth, T. Haustein, O. Koch, W. Zirwas, and E. Schulz, “Interference aware scheduling in the multiuser MIMO-OFDM downlink,” *IEEE Commun. Mag.*, vol. 47, no. 6, pp. 56 – 66, June 2009.

[181] R. Irmer, H.-P. Mayer, A. Weber, V. Braun, M. Schmidt, M. Ohm, N. Ahr, A. Zoch, C. Jandura, P. Marsch, and G. Fettweis, “Multisite field trial for LTE and advanced concepts,” *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 92 – 98, February 2009.

[182] V. Jungnickel, A. Forck, S. Jaeckel, F. Bauermeister, S. Schiffermueller, S. Schubert, S. Wahls, L. Thiele, T. Haustein, W. Kreher, J. Mueller, H. Droste, and G. Kadel, “Field trials using coordinated multi-point transmission in the downlink,” *Proc. 3rd Int. Workshop on Wireless Distributed Networks (WDN), held in conjunction with IEEE PIMRC*, 2010.

[183] M. Grieger, P. Marsch, Z. Rong, and G. Fettweis, “Field trial results for a coordinated mutli-point (CoMP) uplink in cellular systems,” *Proc. ITG/IEEE Workshop on Smart Antennas (WSA’10)*, 2010.

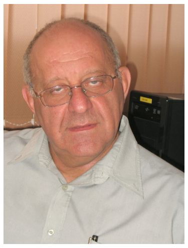
[184] “The openair wireless experimentation platform,” in *Available under http://www.openairinterface.org/*.

**David Gesbert** (IEEE SM) is Professor in the Mobile Communications Dept., EURECOM, France. He obtained the Ph.D degree from Ecole Nationale Superieure des Telecommunications, France, in 1997. From 1997 to 1999 he has been with the Information Systems Laboratory, Stanford University. In 1999, he was a founding engineer of Iospan Wireless Inc, San Jose, Ca.,a startup company pioneering MIMO-OFDM (now Intel). Between 2001 and 2003 he has been with the Department of Informatics, University of Oslo as an adjunct professor. D. Gesbert has published about 170 papers and several patents all in the area of signal processing, communications, and wireless networks.

D. Gesbert was a co-editor of several special issues on wireless networks and communications theory, for JSAC (2003, 2007, 2010), EURASIP Journal on Applied Signal Processing (2004, 2007), Wireless Communications Magazine (2006). He served on the IEEE Signal Processing for Communications Technical Committee, 2003-2008. He’s an associate editor for IEEE Transactions on Wireless Communications and the EURASIP Journal on Wireless Communications and Networking. He authored or co-authored papers winning the 2004 IEEE Best Tutorial Paper Award (Communications Society) for a 2003 JSAC paper on MIMO systems, 2005 Best Paper (Young Author) Award for Signal Proc. Society journals, and the Best Paper Award for the 2004 ACM MSWiM workshop. He co-authored the book "Space time wireless communications: From parameter estimation to MIMO systems", Cambridge Press, 2006.

**Stephen Hanly** (M’98) received a B.Sc. (Hons) and M. Sc. from the University of Western Australia, and the Ph.D. degree in mathematics in 1994 from Cambridge University, UK. From 1993 to 1995, he was a Post-doctoral member of technical staff at AT&T Bell Laboratories. From 1996 to 2009 he was on the research and teaching staff at the University of Melbourne. He is presently an Associate Professor in the Department of Electrical and Computer Engineering at the National University of Singapore. He was an Associate Editor for IEEE Transactions on Wireless Communications from 2005-2009, and is guest editor for the IEEE Journal on Selected Areas special issue on “Cooperative Communications in MIMO Cellular Networks”. In 2005, he was the technical co-chair for the IEEE International Symposium on Information Theory held in Adelaide, Australia. He was a co-recipient of the best paper award at the Infocom 1998 conference, and the 2001 Joint IEEE Communications Society and Information Theory Society best paper award, both for his work with David Tse (Berkeley). His research interests are in the areas of information theory, signal processing, and wireless networking.

**Howard Huang** received a BSEE degree from Rice University in 1991 and a Ph.D. in electrical engineering from Princeton University in 1995. Since then, he has been a researcher at Bell Labs, in Holmdel, New Jersey, currently as a Distinguished Member of Technical Staff in the wireless access domain of Alcatel-Lucent. His research interests include wireless communication theory and cellular system design. He has taught as an adjunct professor at Columbia University and is a Senior Member of the IEEE.

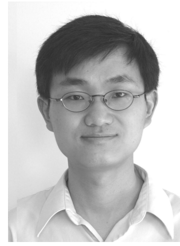


**Shlomo Shamai Shitz** received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai Shitz is an IEEE Fellow, and the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the SHANNON THEORY OF THE IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served on the Board of Governors of the Information Theory Society.

**Osvaldo Simeone** received the M.Sc. degree (with honors) and the Ph.D. degree in Information Engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005 respectively. He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), at the New Jersey Institute of Technology (NJIT), Newark, New Jersey, where he is an Assistant Professor. His current research interests concern the cross-layer analysis and design of wireless networks with emphasis on information-theoretic, signal processing and queuing aspects. Specific topics of interest are: cognitive radio, cooperative communications, ad hoc, sensor, mesh and hybrid networks, distributed estimation and synchronization. Dr. Simeone is the co-recipient of the best paper awards of IEEE SPAWC 2007 and IEEE WRECOM 2007. He currently serves as an Editor for IEEE Trans. Commun.



**Wei Yu** (S'97-M'02-SM'08) received the B.A.Sc. degree in Computer Engineering and Mathematics from the University of Waterloo, Waterloo, Ontario, Canada in 1997 and M.S. and Ph.D. degrees in Electrical Engineering from Stanford University, Stanford, CA, in 1998 and 2002, respectively. Since 2002, he has been with the Electrical and Computer Engineering Department at the University of Toronto, Toronto, Ontario, Canada, where he is now an Associate Professor and holds a Canada Research Chair in Information Theory and Digital Communications. His main research interests include multiuser information theory, optimization, wireless communications and broadband access networks.

Prof. Wei Yu is currently an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS. He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2004 to 2007, and a Guest Editor for a number of special issues for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING. He is member of the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society. He received the IEEE Signal Processing Society Best Paper Award in 2008, the McCharles Prize for Early Career Research Distinction in 2008, the Early Career Teaching Award from the Faculty of Applied Science and Engineering, University of Toronto in 2007, and the Early Researcher Award from Ontario in 2006.