From Theory to Practice: An overview of MIMO space-time coded wireless systems

D. Gesbert (1), M. Shafi (2), D. S. Shiu (3), P. Smith (4), A. Naguib (3)

(1) Corresponding Author: Department of Informatics, University of Oslo,

Gaustadalleen 23, P.O.Box 1080 Blindern, 0316 Oslo, Norway

gesbert@ifi.uio.no, Tel: +47 97537812, Fax: +47 22 85 24 01

(2) Telecom New Zealand

49-55 Tory Street Wellington, New Zealand

Mansoor.Shafi@telecom.co.nz, Ph: +64-4-382-5249, Fax: +64-4-385-8110

(3) Qualcomm Inc., USA

675 Campbell Technology Parkway Suite 200 Campbell, CA 95008, USA

dashiu@qualcomm.com Ph: +1-408-557-1006, Fax: +1-408-557-1001

(4) Dept of Electrical and Elec. Engineering University of Canterbury

Christchurch, New Zealand

p.smith@elec.canterbury.ac.nz, Ph: +64-3-364-2987 ext. 7157 Fax: +64-3-364-2761

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.

Keywords

Spectrum efficiency, wireless systems, MIMO, smart antennas, diversity, Shannon capacity, space-time coding, channel models, 3G, beamforming, spatial multiplexing.

D. Gesbert acknowledges the support in part of Telenor AS, Norway

I. Introduction

Digital communication using MIMO (multiple-input multiple-output), 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), 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], [2], [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 multi-antenna-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

[1]Third generation wireless UMTS-WCDMA.

fundamental ideas presented over the last few years in this area, although some recent progress is also mentioned.

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 VI. Section VII concludes this paper.

II. Principles of Space-time (MIMO) systems

Consider the multi-antenna 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 (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 *pre-coding*. 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 multi-antenna 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 multi-antenna system is located in the weight selection algorithm rather than in the coding side although the development of *space-time codes* 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 vs. quality trade-off 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 multi-antenna 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 author shows how one may under certain conditions transmit $\min(M, N)$ independent data streams simultaneously over the *eigen-modes* 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 J. 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 (eg. in 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], [12], [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 parameters. Alternatively, diversity-oriented smart antenna techniques perform well in 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

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

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 multi-antenna 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), 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 multi-user 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 1×1 (SISO) system the capacity is given by [20]:

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

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, ρ is the SNR at any RX antenna. As we deploy

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

more RX antennas the statistics of capacity improve and with M RX antennas we have a single-input multiple-output (SIMO) system with capacity given by [20]:

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

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 MISO system with N TX antennas and the capacity is given by [1]:

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

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]

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

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 $min(M, N)$ non-zero eigenvalues of its (channel-dependent) matrix argument, each eigenvalue characterizing the SNR over a so-called channel eigen-mode. 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 respectively. 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 [23], [3], [21] and capacity outage [1], [24], [25], [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 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 multi-user 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], [29], [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

B.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 σ^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 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]

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

where $tr(\mathbf{Q}) \leq \rho$ holds to provide a global power constraint. Note that for equal power uncorrelated sources $\mathbf{Q} = \frac{\rho}{N}\mathbf{I}_N$ 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. Firstly, what is the effect of $\mathbf{Q}$? If we compare the capacity achieved by $\mathbf{Q} = \frac{\rho}{N}\mathbf{I}_N$ (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]

$$C_{EP} = \sum_{i=1}^{m} \log_2(1 + \frac{\rho}{N}\lambda_i) \quad \text{bits/sec/Hz} \tag{7}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the non zero 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 $\mathbf{D} = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_m}, 0, \ldots, 0)$. 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 $\lambda_1, \lambda_2, \cdots, \lambda_m$.

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 1920's and a considerable amount is known about them. For general $\mathbf{W}$ matrices a wide range of limiting results are known [31], [22], [32], [33], [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 ?

B.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], [21], [37], [26], [38], [39], [4] and is called a waterfilling (WF) solution. There is a simple algorithm to find the solution [3], [21], [37], [26], [39] and the resulting capacity is given by

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

where μ 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 μ is a complicated non-linear function of $\lambda_1, \lambda_2, \cdots, \lambda_m$, the distribution of C_{WF} appears intractable, even in the Wishart case when the joint distribution of $\lambda_1, \lambda_2, \cdots, \lambda_m$ is known. Nevertheless C_{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_{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 [40], [23]. 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 [40], [39], [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 $10dB$. From zero to $10dB$ the gains are usually less than 60%. For SNR values below $0dB$, large gains are possible, with values around 200% being reported at $-10dB$. 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 space-time codes 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.

B.3 Channel unknown at the transmitter

Here the capacity is given by C_{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 $\mathbf{Q} = \frac{\rho}{N} I_N$ is optimal for i.i.d. Rayleigh fading. Foschini derived (4) starting from an equal power assumption. The variable, C_{EP}, is considerably more amenable to analysis than C_{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_{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 [45], [46], [47], [26]. For example C_{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 Subsection 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 $M = 2!$ [40], [36], [49], [50]. We outline this work below as well as results which are asymptotic in SNR rather than system size.

C.1 Channel known at the transmitter

Analytical results are scarce here but a nice analysis in [39], [42] shows that C_{WF}/M converges to a constant, μ_{WF}, for "large systems" in both i.i.d. and correlated fading conditions. The value of μ_{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_{WF} is shown for Ricean fading in [40] as is the accuracy of Gaussian approximations to C_{WF} in both Rayleigh and Ricean cases.

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

C.2 Channel unknown at the transmitter

In this situation we have the capacity given in (4) as C_{EP}. For "large systems" (assuming the Wishart case) the limiting mean capacity was shown to be of the form $M\mu_{EP}$ [3] where μ_{EP} depends on M, N only through the ratio $c = M/N$. A closed form expression for C_{EP} was given in [23] and the

accuracy of this result was demonstrated in [40], [36]. 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 [40], [36]. 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

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

where $M/N \to c$ as $M, N \to \infty$ and $Z \sim N(0,1)$ is a standard Gaussian random variable. See [33], [34] for exact details of the conditions required for (12) to hold. Hence for the Wishart case Gaussian approximations might be considered to C_{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 $M = 2$ [40], [36], not only for Rayleigh channels but for Ricean channels as well. More general results which also cater for correlated fading can be found in [42], [39], [27]. In [42], [39] it is shown that C_{EP}/M converges to a constant, μ_{EP}, for "large systems" in both i.i.d. and correlated fading. The value of μ_{EP} is obtained and it is shown that correlation always reduces μ_{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_{EP} and C_{WF} are equivalent. For arbitrary values of M, N high SNR approximations are given in [27] for the mean, variance and MGF of C_{EP} .

IV. Transmission over MIMO systems

Although the information theoretic analysis can be bootstrapped to motivate receiver architectures (as was done eg. 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 *space-time codes* (STC). 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. Figure 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. Then, the popularity of STC really took off with the discovery of the so-called *space time block codes* (STBC). 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, IV-C can be found in [54].

B. Maximizing diversity with space-time trellis codes

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

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

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

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

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) $\mathbf{H}(l),\ l = 1, \cdots, L$ 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]

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

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 $\beta_i, i = 1, \cdots, r$ 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 space-time code and the term $(E_s/4N_o)^{-rM}$ represents a diversity gain of rM. Since $r \leq N$, the overall diversity order is always less or equal to MN. It is clear that in designing a space-time trellis code, 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.

As an example for space-time trellis codes, we provide an 8-PSK 8-state ST code designed for 2 TX antennas. Figure 5 provides a labeling of the 8-PSK constellation and the trellis description for this code. Each row in the matrix shown in Figure 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 8 constellation points. This code has a bandwidth efficiency of 3 bits per channel use.

Figure 6 shows the performance of 4-PSK space-time trellis codes for 2 TX and 1 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], [56], [57], [58], [59], [60] (these are mentioned only as an example, there are many other published results that address 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 space-time block codes

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 space-time block codes. Figure 7 shows the baseband representation for Alamouti space-time block coding (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 maximum likelihood (ML) decoder is

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

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 independent and identically distributed. 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 2×1 space-time block code, 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 2×2^b. It is also straightforward to verify that the SNR for c_1 and c_2 will be

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

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, space-time codes 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

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

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 mth 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 pre-multiplying the received signal vector $\mathbf{r}_m$ by $\mathbf{H}_m^*$. In this case, the diversity order provided by this scheme is $2M$. Figure 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 Figure 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 space-time block code (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 2×2 STBC, the transmission rate is 1 while achieving the maximum diversity gain possible with two TX and two RX antennas (4th order). However, concatenating a STBC with an outer conventional channel code (eg. a convolutional or 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], [64], [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 2 TX antennas was studied in [67], [68], [69], [62]. There, a general technique for constructing space-time block codes 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, space-time block codes 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 space-time block code 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 4 TX antennas, a rate 4/8 (i.e. it is a rate 1/2) space-time block code 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 $\mathbf{c} = [c_1, .., c_4]$) will yield a 8×4 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 $\alpha_4 = 2 \cdot \sum_{i=1}^{4} |h_i|^2$ (4th order diversity). This scheme provides a 3 dB power gain that comes from the intuitive fact that 8 time slots are used to transmit 4 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], [72], [73], [74], [75], [76], [77], [78]. For one TX antenna, there exist differential detection schemes, such as 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 ($\frac{\pi}{4}$-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 2×2 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 [79] 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 2×2 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 non-binary signals. Second, it is limited only to the $N = 2$ space-time block code 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 2×2 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 non-differential approach to transmit diversity when the channel is not known at the receiver is reported in [82], [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 narrow band 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 2 TX antennas and a receiver with 1 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 non-linear and non-causal 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), 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], [86], [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 consectutive 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 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 space-time block codes 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 un-mix 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 (1.5dB to 2dB in the 2×2 case) if knowledge of non-trivial noise/interference statistics (eg. 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 SIC (successive interference canceling) proposed for multi-user 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], [2], [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 traning in blind approaches is in some limited loss of BER performance and more often in the increased computational complexity.

E.1 Multiplexing vs. 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 16QAM instead of 4PSK etc. The various trade-offs between multiplexing and diversity have begun to be looked at, for instance in [99], [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], [102], [103].

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 sub-stream. 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 2×2 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 2×2 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 RF chains. By using antenna selection techniques (see for example [104], [105], [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 eg. 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], [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

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

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

In practice the latter is more accurately represented by the *distribution* of eigenvalues of $\mathbf{HH}^*$, denoted $\{\lambda_1, \lambda_2, ..\}$. 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. Pseudo-static narrowband MIMO channel

A.1 Line-of-sight component model

It is common to model a wireless channel as a sum of two components, a LOS component and a non-line-of-sight (NLOS) component. That is, $\mathbf{H} = \mathbf{H}_{\text{LOS}} + \mathbf{H}_{\text{NLOS}}$. 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 $\mathbf{H}_{\text{LOS}}$ becomes. Since $\mathbf{H}_{\text{LOS}}$ 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], [114], [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 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 (eg. WLAN).

A.2 Correlation model for non-line-of-sight component

In the absence of a LOS component, the channel matrix reduces to $\mathbf{H}_{\text{NLOS}}$ and is usually modeled with circularly-symmetric complex Gaussian random variables (i.e. Rayleigh fading). The elements of $\mathbf{H}_{\text{NLOS}}$ 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 $\mathbf{H}_{\text{NLOS}}$ is $\text{cov}(\text{vec}(\mathbf{H}_{\text{NLOS}})) \equiv \Psi$, where cov and vec are the covariance and matrix vectorization operator (stacking the columns on top of each other) respectively. In the following, we will introduce commonly accepted models for $\text{cov}(\text{vec}(\mathbf{H}))$. Before that, let us first review a simple model shown in Figure 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

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

An appropriate model for a macro-cell 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 $\Psi(R_m T_n, R_{m'} T_n) = J_0(\frac{2\pi}{\lambda} D(R_m, R_{m'}))$, 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, Θ and Δ 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 $\Psi(R_m T_n, R_m T_{n'}) \approx J_0(\frac{2\pi\Delta}{\lambda} D(T_n, T_{n'}))$. On the other hand, if the mobile is at the inline direction, i.e. $\Theta = \pi/2$, then $|\Psi(R_m T_n, R_m T_{n'})| \approx J_0(\frac{2\pi}{\lambda}\Delta^2 D(T_n, T_{n'}))$ [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, eg. 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 Θ 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 long as the distribution of Θ_i is diverse enough, the antennas will become fairly uncorrelated, even with angle spreads Δ_i approaching zero [120].

B. Impact of spatial correlation

The statistics of $\mathrm{vec}(\mathbf{H}_{\mathrm{NLOS}})$ given Ψ 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 Ψ to be a Kronecker product of the two local correlation matrices. That is, let Ψ^R and Ψ^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 $(\Psi^R)^{\frac{1}{2}}\mathbf{H}_w(\Psi^T)^{\frac{1}{2}}$. This is a useful form for mathematical manipulation. Figure 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 $(\Psi^R)^{\frac{1}{2}}\mathbf{H}_w(\Psi^T)^{\frac{1}{2}}$. 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 $(\Psi^R)^{\frac{1}{2}}\mathbf{H}_w(\Psi^T)^{\frac{1}{2}}$, then an upper bound of channel capacity can be derived. The channel capacity given $\mathbf{H}_w$ and an SNR budget ρ can be upper bounded by:

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

where υ_k^R, υ_k^T and λ_k are the kth largest eigenvalues for Ψ^R, Ψ^T and $\mathbf{H}_w\mathbf{H}_w^*$, respectively, and $\sum \rho_k = \rho$ [45].

Even though (30) 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 $\upsilon_k^R \upsilon_k^T$.

B.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], [123], [124], [125], [126]. Measurements conducted at 2.5GHz for broadband wireless access applications can be found in [115].

B.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-by-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][128][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-by-2 matrix will always be 2.

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

B.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 vec($\mathbf{H}$); a matrix of correlated, equal-variance complex Gaussian entries may not be an appropriate model for such a MIMO channel.

B.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 λ_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 $\min(M, N)$ bits/sec/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:

$$\text{EDOF} \equiv \frac{d}{d\delta} C_{EP}(2^{\delta}\rho)|_{\delta=0} \tag{31}$$

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 $\mathbf{H}_1 = \mathbf{H}_{LOS} + \mathbf{H}_{random}$ 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 τ_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 Hz in stationary applications to 200Hz 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].

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 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 Third Generation 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 third generation partnership projects (3GPP's). 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 eg. high speed digital packet access (HSDPA) [134], [135], [136]. These include adaptive modulation and coding, hybrid ARQ, fast cell selection, transmit diversity.

B. MIMO in 3G wireless systems & 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 2GHz 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.

B.1 Antenna issues

Antenna element numbers and inter element 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 4, with an inter element 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, 4 antennas can fit into a linear space of 1.5m at 10 lambda spacing at 2 GHz. For the terminal, half lambda spacing is sufficient to ensure a fair amount of un-correlated 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 4, though a lower number, say 2, 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.

B.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 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 base band 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.

B.3 System integration and signalling

The MIMO system needs to be integrated and be backwards compatible with an existing non MIMO network. MIMO signalling imposes the support of special radio resource control (RRC) messages. The terminals need to know via broadcast down link signalling 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 down link and up link RRC messages are then mapped on to the layer 2 signalling messages [139].

B.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. Whilst the propagation models for conventional radio systems have been standardised in [140], there is no agreed MIMO channel model by the ITU as yet.

B.5 Channel state information at transmitter

As shown earlier, the channel capacity is a function of the eigen-modes 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 [39], [21] 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 channel state information 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 trade-offs between existing tranmission 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.

Acknowledgments

The authors would like to thank J. Akhtar for discussions and his help in the simulations. We also thank Profs. D. Falconer, H. Bolcskei, A. Paulraj and Drs. R. Kalbasi, 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 Personal Communications*, vol. 6, pp. 311–335, March 1998.

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

[3] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," technical memorandum, AT&T Bell Laboratories, June 1995.

[4] G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communications," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, 1998.

[5] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Comm.*, Feb. 2002.

[6] A. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Proc. 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," *U. S. Patent*, no. 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. JSAC-5(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," *Electronics Letters*, vol. 35, no. 1, pp. 14–15, 1999.

[10] R. A. V. G. J. Foschini, G. D. Golden and P. W. Wolniansky, "Simplified processing for wireless communication at high spectral efficiency," *IEEE Journal on Selected Areas in Communications - Wireless Communications Series*, vol. 17, no. 11, 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 Conference*, 2000.

[12] D. So and R. Cheng, "Detection techniques for v-blast in frequency selective channels," in *Proc. Wireless Communications and Networking Conference*, 2002.

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

[14] G. J. Foschini, "Some basic layered space time architectures and their performance," *IEEE Journal on Selected Areas in Communications, Special Issue on MIMO systems*, 2003.

[15] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Fundamental Capacity of MIMO Channels," *IEEE Journal on Selected Areas in Communications, Special Issue on MIMO systems*, 2003.

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

[17] R. Blum, "MIMO capacity with interference," *IEEE Journal on Selec. Areas in Comm., Special Issue on MIMO systems*, 2003.

[18] P. F. D. S. Catreux and L. J. Greenstein, "Attainable throughput of an interference-limited multiple-input multiple-output cellular system," *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1307–1311, 2001.

[19] H. Dai, A. Molisch, and H. V. Poor, "Downlink capacity of interference limited MIMO systems with joint detection," *IEEE Transactions on Wireless Communications, submitted*, 2002.

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

[21] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Communications*, 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," *Journal of Multivariate Analysis*, 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 Transactions on Communications*, vol. 48, no. 8, pp. 1245–1248, 2000.

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

[25] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple-input multiple-output wireless channels: a geometric interpretation," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 173–176, 1999.

[26] D. Shiu, *Wireless Communication using Dual Antenna Arrays*. Kluwer International Series in Engineering and Computer Science, 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," *Physics Archive*, no. 0010081, 2000.

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

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

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

[31] A. Edelman, *Eigenvalues and condition numbers of random matrices*. PhD thesis, Massachusetts Institute of Technology, 1989.

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

[33] V. L. Girko, *Theory of Random Determinants*. Kluwer Academic Publishers, 1990.

[34] V. L. Girko, *Theory of Linear Algebraic Equations with Random Coefficients*. New York, Allerton Press, 1996.

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

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

[37] A. L. F. R. Farrokhi, G. J. Foschini and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Communications Letters*, vol. 5, no. 3, pp. 85–87, 2001.

[38] D. N. C. T. P. Viswanath and V. Anantharam, "Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Transactions on Information Theory*, vol. 47, no. 1, 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. Inf. Theory*, vol. 48, pp. 637–650, March 2002.

[40] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *International Conference on Communications*, ICC 2002, IEEE, 2002.

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

[42] J. M. K. C-N. Chuah and D. Tse, "Capacity of indoor multi-antenna array systems in indoor wireless environment," in *Globecom*, 98, Sydney, pp. 1894–1899, 1998.

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

[44] A. Grant, "Rayleigh fading multiple antenna channels," *EURASIP Journal of Applied Signal Processing*, pp. 316–329, March 2002.

[45] M. J. G. D. Shiu, G. J. Foschini and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, 2000.

[46] A. L. M. M. Stoytchev, H. F. Safar and S. H. Simon, "Compact antenna arrays for MIMO applications," in *IEEE Intern. Symposium on Antennas and Propagation*, 2001.

[47] J. L. D. Chizhik, F. Rashid-Farrokhi and A. Lozano, "Effect of antenna separation on the capacity of blast in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, 2000.

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

[49] M. B. H. Ge, K. D. Wong and J. C. Liberti, "Characterization of Multiple-Input Multiple-Output (MIMO) Channel Capacity," in *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC), Orlando, FL.*, 2002.

[50] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE Journal on Selected Areas on Communications*, vol. 18, pp. 2172–2178, November 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," *International Journal of 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), pp. 1630–1634, 1993.

[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. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[54] A. Naguib, N. Seshadri, and R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Magazine*, may 2000.

[55] J. G. M. P. Fitz and J. V. Krogmeier, "Further results on space-time codes for rayleigh fading," in *Proc. Allerton*, p. 391400, September 1998.

[56] Q. Yan and R. S. Blum, "Optimum space-time convolutional codes for for quasistatic slow fading channels," in *WCNC*, September 2000.

[57] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis-coded modulation," *IEEE Commun. Letters*, vol. 1, p. 2022, January 2000.

[58] A. R. Hammons and H. E. Gamal, "On the theory of space-timec odes for psk modulation," *IEEE Trans. Information Theory*, vol. 46, pp. 524–542, March 2000.

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

[60] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "A rank criterion for QAM space-time codes," *IEEE Intern. Symposium on Info. Theory*, 2000.

[61] S. Alamouti, "Space Block Coding: A Simple Transmitter Diversity Technique for Wireless Communications," *IEEE Journal on Selec. Areas. Commun.*, vol. 16, pp. 1451–1458, October 1998.

[62] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. 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 *International Symposium on Wireless Personal Multimedia Communications*, (Alborg, Denmark), September 2001.

[64] S. Siwamogsatam and M. Fitz, "Improved high rate space–time TCM via concatenation of expanded orthogonal block codes and MTCM," in *IEEE International Conference on Communication*, (New York), April 2002.

[65] S. Siwamogsatam and M. Fitz, "Improved high rate space–time TCM from an expanded STB-MTCM Construction," *IEEE Trans. Information Theory*, Accepted for publication.

[66] N. Seshadri and H. Jafarkhani, "Super-Orthogonal Space-Time Trellis Codes," in *IEEE International Conference on Communication*, (New York), April 2002.

[67] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes for wireless communications: Performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, March 1999.

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

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

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

[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 Journal on Selec. Areas. Commun.*, vol. 16, pp. 1459–1478, October 1998.

[72] J. K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Faded Channels," *IEEE Trans. Veh. Technology*, vol. VT-40, pp. 683–693, November 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), pp. 640–646, May 1989.

[74] M. L. Moher and J. H. Lodge, "TCMP-A Modulation and Coding Strategy for Rician Fading Channels," *IEEE Journal on Selected Areas in Commun.*, vol. JSAC-7, pp. 1347–1355, December 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. of International Mobile Satellite Conf.*, (Ottawa), pp. 182–187, June 1990.

[76] J. Yang and K. Feher, "A digital Rayleigh Fade Compensation Technology for Coherent OQPSK System," in *Proc. IEEE VTC'90*, (Orlando,Fl), pp. 732–737, May 1990.

[77] C. L. Liu and K. Feher, "A New Generation of Rayleigh Fade Compensated $\frac{\pi}{4}$-QPSK Coherent Modem," in *Proc. IEEE VTC'90*, (Orlando,Fl), pp. 482–486, May 1990.

[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. COM-39, pp. 25–29, January 991.

[79] V. Tarokh, S. M. Alamouti, and P. Poon, "New Detection Scheme for Transmit Diversity with no Channel Estimation." Intern. Conf. Universal Personal Communications, ICUPC '98, 1998.

[80] V. Tarokh and H. Jafarkhani, "A Differential Detection Scheme for Transmit Diversity." IEEE Journal on Selec. Areas. Commun., July, 2000.

[81] B. L. Hughes, "Differential Space-Time Modulation." IEEE Trans. Information Theory, Nov., 2000.

[82] B. M. Hochwald and T. L. Marzetta, "Unitary Space-Time Modulation for Multiple Antenna Communications in Rayleigh Flat Fading." to appear in IEEE Trans. Information Theory, March 2000.

[83] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldons, and R. Urbanke, "Systematic Design of Unitary Space-Time Constellation," *IEEE Trans. Information Theory, 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 Journal on Selec. Areas in Comm.*, Submitted, May 2002.

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

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

[87] G. Bauch and A. F. Naguib, " Map Equalization of Space-time Coded Signals over Frequency Selective Channels ," in *Proc. IEEE Wireless Commun. and Networking Conf. WCNC'99*, vol. 1, (New Orleans, Louisiana), pp. 261–265, September 1999.

[88] A. Liu, G. B. Giannakis, A. Scaglione, and S. Barbarossa, "Decoding and Equalization of Unkown Multipath Channels Based on Block Precoding and Transmit Diversity," in *Asilomar Conference on Signals, Systems, and Computers*, pp. 1557–1561, 1999.

[89] H. Bolcskei and A. Paulraj, "Space-frequency codes for broadband fading channels," in *IEEE International Symposium on Information Theory, Washington D. C.*, June 2001.

[90] N. Al-Dhahir, "Single-Carrier Frequency Domain Equalization for Space-Time Block Coded Transmission Over Frequency Selective Fading Channels," *IEEE Communication Letters*, 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.

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

[93] S. Verdu, "Multi-user detection," in *Advances in statistical signal processing* (V. Poor, ed.), pp. 369–409, JAI Press, 1993.

[94] P. Comon and P. Chevalier, "Source separation: Models, concepts, algorithms and performance," in *Unsupervised Adaptive Filtering, Vol. I, Blind Source Separation* (S. Haykin, ed.), Series on Adaptive and learning systems for communications signal processing and control, pp. 191–236, Wiley, 2000.

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

[96] C. Papadias, "A multi-user kurtosis algorithm for blind source separation," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2000.

[97] P. Loubaton, E. Moulines, and P. Regalia, *Subspace methods for blind identification and deconvolution (Chap. 3), in Signal Processing Advances in Wireless and Mobile Communications. Giannakis, Hua, Stoica, Tong, Eds.* Prentice Hall, 2001.

[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. Inf. Theory*, Submitted 2002.

[100] R. Heath and A. Paulraj, "Multiplexing versus diversity in MIMO antenna channels," *submitted*, 2001.

[101] B. Hassibi and B. Hochwald, "High rates codes that are linear in space and time," *Submitted to IEEE Trans. on Information Theory*, 2000.

[102] S. Sandhu and A. Paulraj, "Unified design of linear space-time block-codes," *IEEE Globecom Conference*, 2001.

[103] M. O. Damen, A. Tewfik, and J. C. Belfiore, "A construction of a space-time code based on number theory," *IEEE Trans. on Information Theory*, March 2002.

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

[105] D. Gore and A. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Tr. Signal Processing*, October 2002.

[106] R. W. Heath, S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Communications Letters*, April 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*, May 2002.

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

[109] J. Akhtar and D. Gesbert, "Partial feedback based space-time block coding," *IEEE Trans. Comm.*, submitted 2002.

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

[111] S. Simon and A. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE JSAC, Special Issue on MIMO systems*, 2003.

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

[113] L. J. Greenstein, S. Ghassemzadeh, V. Erceg, and D. G. Michelson, "Theory, experiments, and statistical models," in *WPMC 99 Conference Proceedings*, Amsterdam, September 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 *Proceedings of ICPWC 2000*, Hyderabad, December 2000.

[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. International Conference on Communications*, 2002.

[116] E. Green, "Radio link design for microcellular systems," *BT Tech. J*, vol. 8(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 *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, October 2004.

[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," *IEE Proc. Radar, Sonar Navigation*, vol. 145, pp. 32–41, 1998.

[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, February 1999.

[121] D. Chizhik, G. Foschini, and R. A. Valenzuela, "Capacities of multi-element transmit and receive antennas: Correlations and keyholes," *Electronic Letters*, pp. 1099–1100, 2000.

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

[123] J. Ling, D. Chizhik, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple transmitter multiple receiver capacity survey in manhattan," *Electronic Letters*, vol. 37, 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 Conference*, May 2001.

[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 multi-element transmit and receive antenna arrays," in *Proc. IEEE Spring VTC*, 2000.

[126] J. R.-M. L. Schumacher, L. Berger, "Recent advances in propagation characterisation and multiple antenna processing in the 3gpp framework," in *Proc. of the URSI radio conference*, August 2002.

[127] S. Kozono, T. Tsuruhara, and M. Sakamoto, "Base station polarisation diversity reception for mobile radio," *IEEE Trans. Veh. Tech.*, pp. 301–6, 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 polarisationi diversity schemes at 1800 mhz," *IEEE Trans. Veh. Tech.*, pp. 318–26, 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. Selected Areas in Commun.*, pp. 1046–57, 1993.

[130] R. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, and A. Paulraj, "Performance of multi-antenna signaling techniques in the presence of polarization diversity," *IEEE Trans. Comm.*, October 2002.

[131] M. Wennström and T. Svantesson, "An antenna solution for mimo channels: The switched parasitic antenna," in *IEEE Symposium on Personal Indoor and Mobile Radio Communication (PIMRC) 2001*, San Diego, 2001.

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

[133] V. Erceg *et al.*, "Channel models for fixed wireless applications," tech. rep., IEEE 802.16 work group technical report, 2001.

[134] 3GPP, "Physical layer aspects of utra high speed downlink packet access, release 4," Tech. Rep. 3G TR25.848 v.4.0, 2001-03.

[135] 3GPP, "Tx diversity solutions for multiple antennas release 5," Tech. Rep. 3G TR 25.869 v 1.0.0, 2001-06.

[136] 3GPP, "Multiple-input multiple output antenna processing for hsdpa," Tech. Rep. 3GPP TR 25.876 v0.0.1, 2001-11.

[137] D. Gesbert, L. Haumonte, H. Bolcskei, and A. Paulraj, "Technologies and performance for non line-of-sight broadband wireless access networks," *IEEE Communications Magazine*, April 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 r recommendation," Tech. Rep. M 1225.

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

LIST OF FIGURE CAPTIONS:

- 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.
- Fig. 2. Basic spatial multiplexing (SM) scheme with 3 TX and 3 RX antennas yielding three-fold improvement in spectral efficiency. Ai, Bi, Ci represent symbol constellations for the three inputs at the various stages of transmission and reception.
- Fig. 3. Figure showing the percentage relative gains in capacity due to feedback at various SNR values, channel models (K is the Ricean factor) and array sizes.
- Fig. 4. Space-Time Coding.
- Fig. 5. 8-PSK 8-state space-time code with 2 TX antennas.
- Fig. 6. Performance of 4-PSK Space-Time Trellis Codes with 2 TX and 1 RX Antennas
- Fig. 7. Transmitter Diversity with Space-Time Block Coding
- Fig. 8. Receiver for Space-Time Block Coding
- Fig. 9. Bit Error Rate (BER) comparisons for various transmission techniques over 2×2 MIMO. At high SNR, from top to bottom: Spatial multiplexing (SM)-ZF, SM-ML, STBC-ML, Alamouti STBC.
- Fig. 10. Diagram to derive antenna correlation Ψ. 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.
- 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 $D(R_m, R_{m'}) = \lambda|m - m'|$.
- 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.

LIST OF TABLE CAPTIONS:

- Table 1. Example of standardized MIMO channels for IEEE body 802.16.
- Table 2. Peak Data rates of Various MIMO Architectures.

David Gesbert See the guest editorial for this special issue.

Mansoor Shafi See the guest editorial for this special issue.

Da-shan Shiu See the guest editorial for this special issue.

Peter Smith See the guest editorial for this special issue.

Ayman Naguib A. Naguib received the B.Sc Degree (with honors) and the M. S. degree EE from Cairo University, Cairo, Egypt, in 1987 and 1990 respectively, and the M. S. degree in statistics and the Ph.D. degree in EE 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, Egypt. 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 Laboratories, 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, a new startup company developing next generation signal processors for wireless communications as a Technical Leader and Manager of Core Technology. His current research interests include, space-time signal processing and coding for wireless communications. Dr. Naguib holds 5 US patents and 7 other pending patent applications in the area of space-time coding and signal processing. Dr. Naguib is a senior member of IEEE and is currently serving as an Associate Editor for CDMA and Space-Time Systems, IEEE Transactions on Communications.

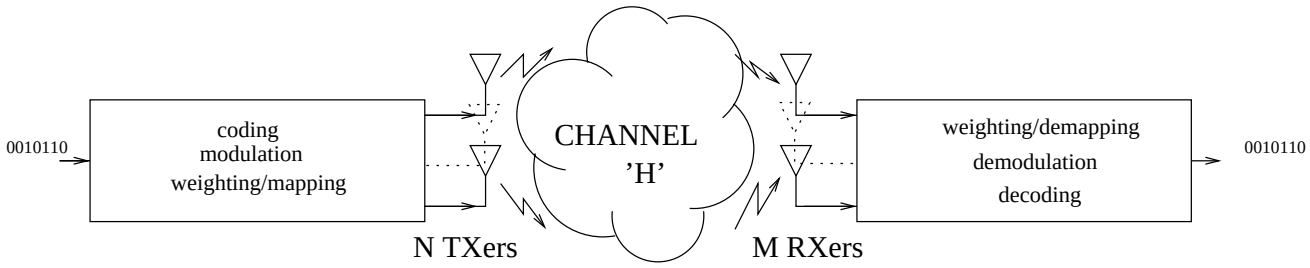

Fig. 1. Diagram of a MIMO wireless transmission system. The transmitter and receiver are equipped with multiple antenna elements. Coding, modulation and mapping of the signals onto the antennas may be realized jointly or separately.

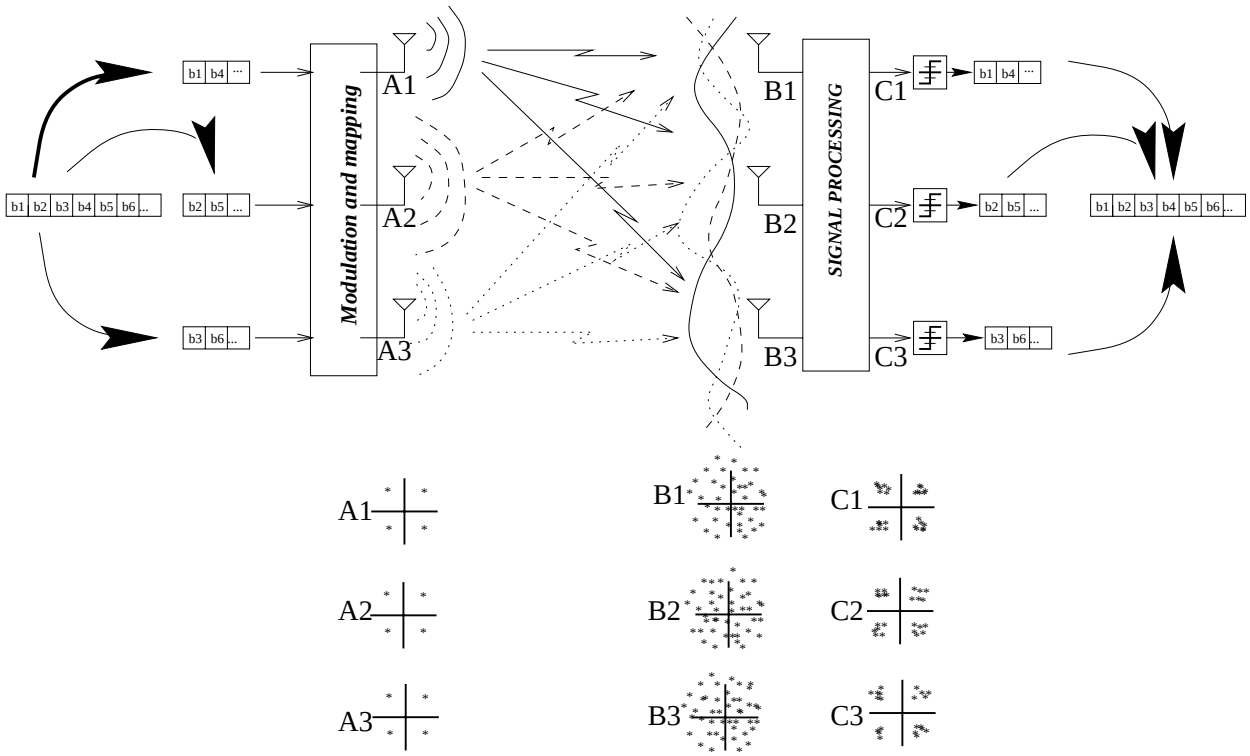

Fig. 2. Basic spatial multiplexing (SM) scheme with 3 TX and 3 RX antennas yielding three-fold improvement in spectral efficiency. Ai, Bi, Ci represent symbol constellations for the three inputs at the various stages of transmission and reception.

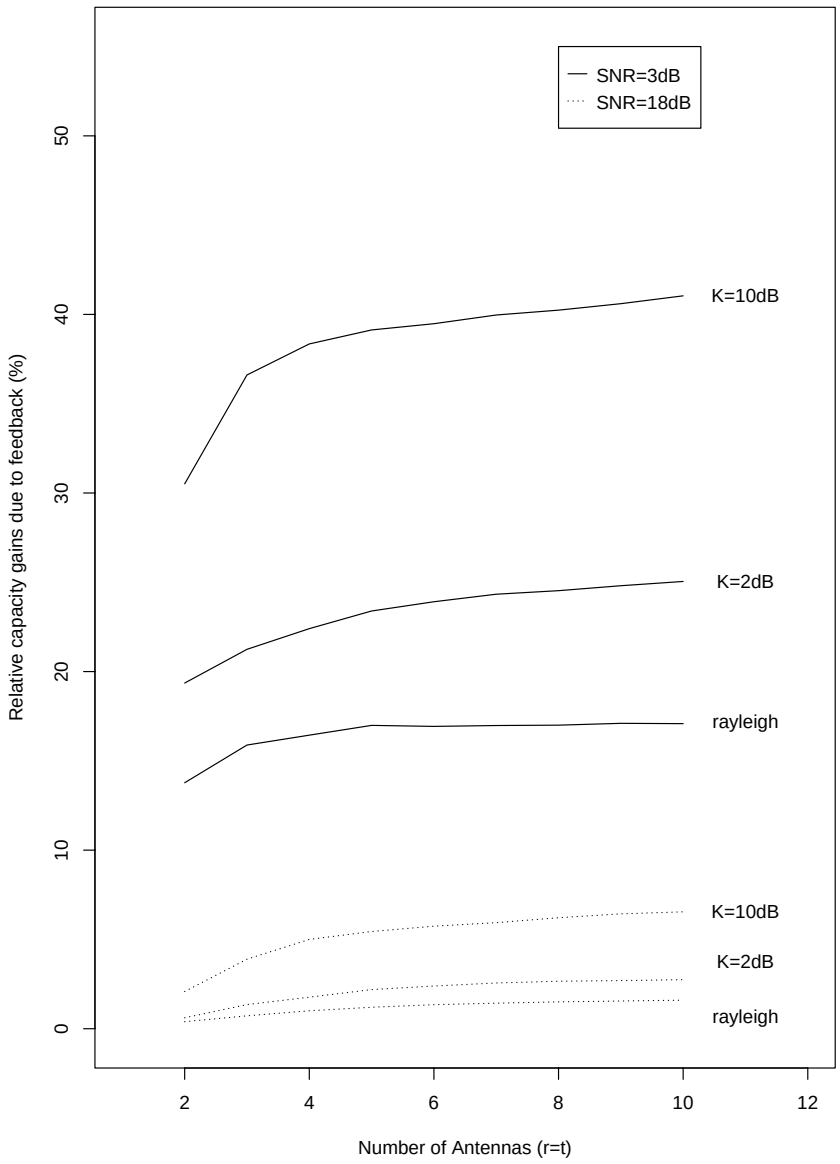

Fig. 3. Figure showing the percentage relative gains in capacity due to feedback at various SNR values, channel models (K is the Ricean factor) and array sizes.

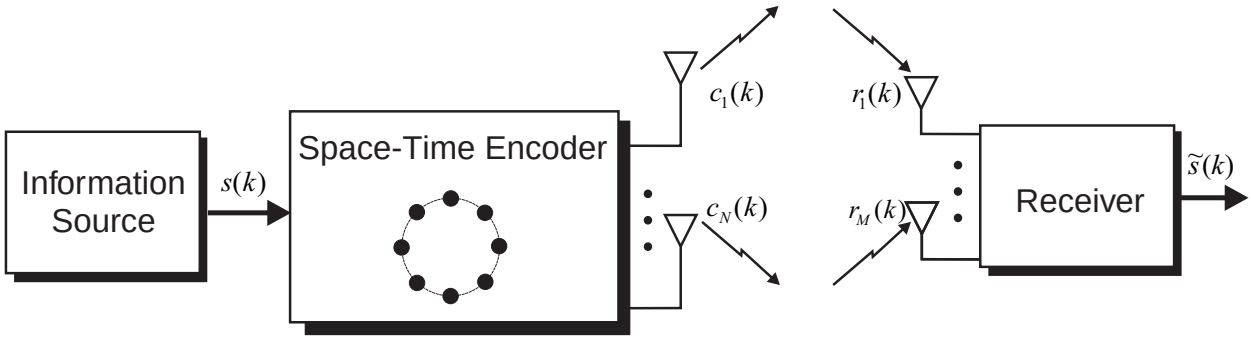

Fig. 4. Space-Time Coding

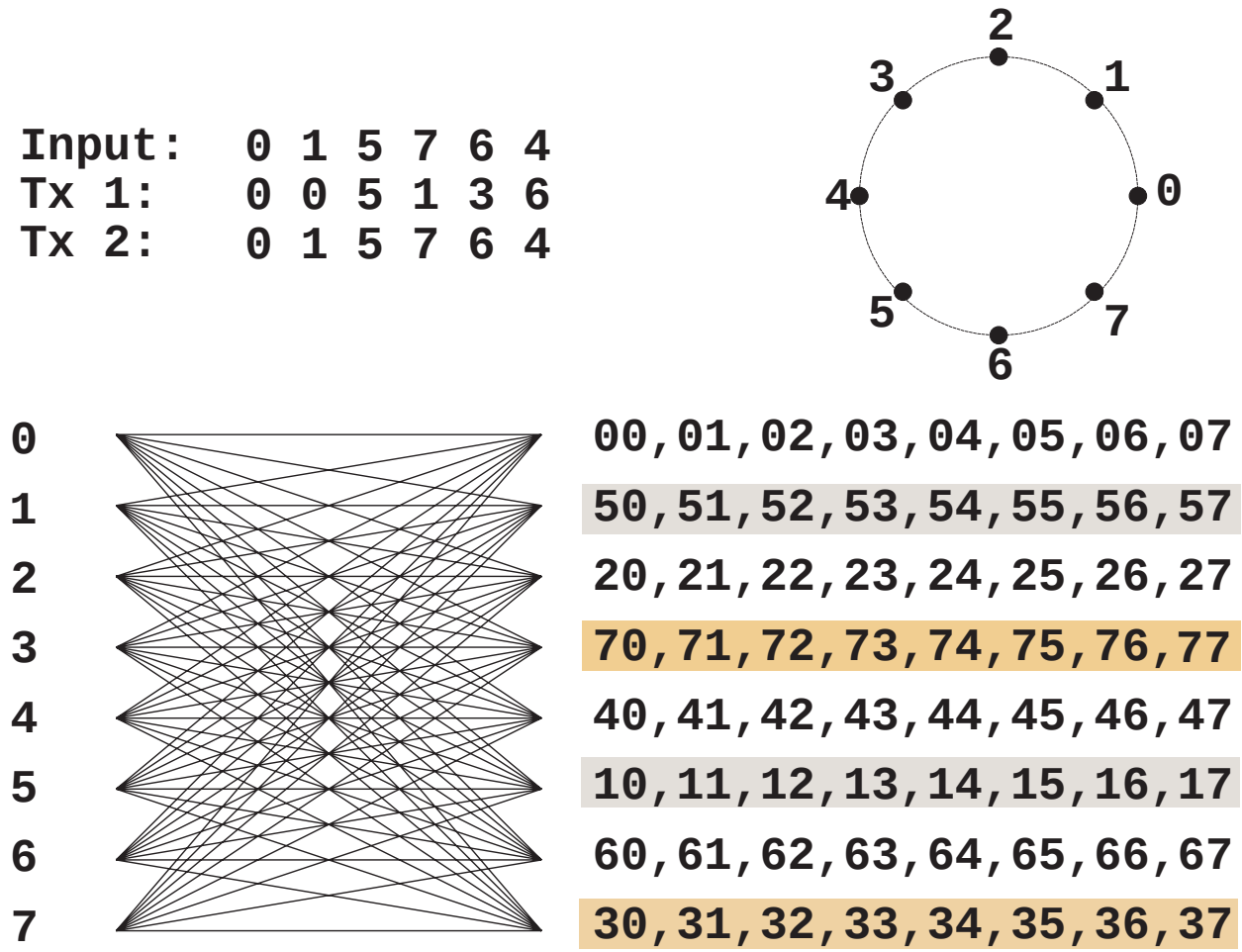

Fig. 5. 8-PSK 8-state space-time code with 2 TX antennas

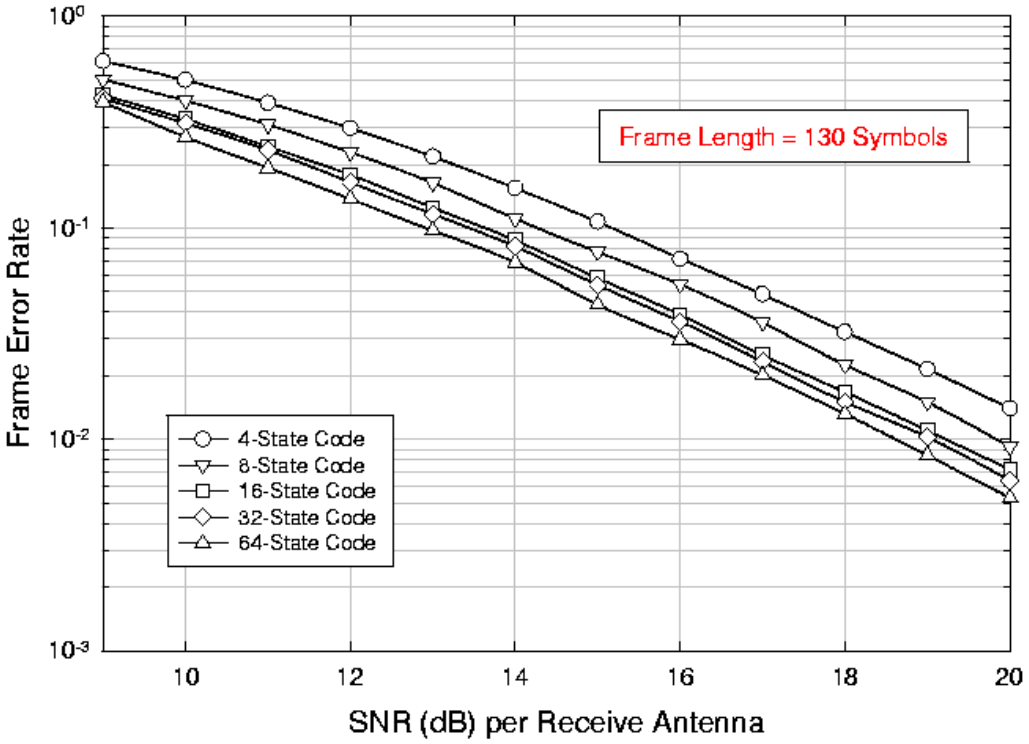

Fig. 6. Performance of 4-PSK Space-Time Trellis Codes with 2 TX and 1 RX Antennas

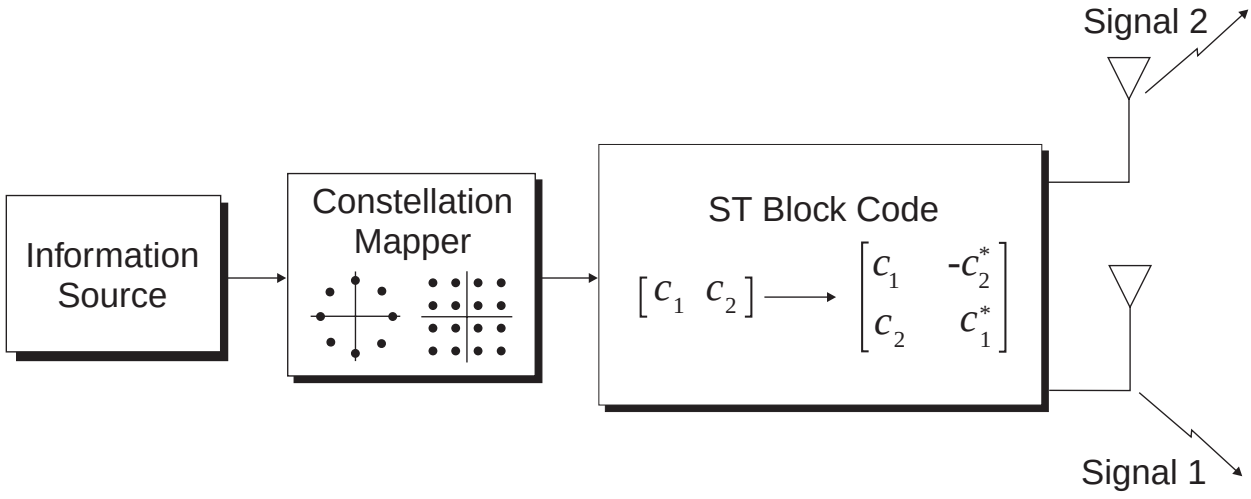

Fig. 7. Transmitter Diversity with Space-Time Block Coding

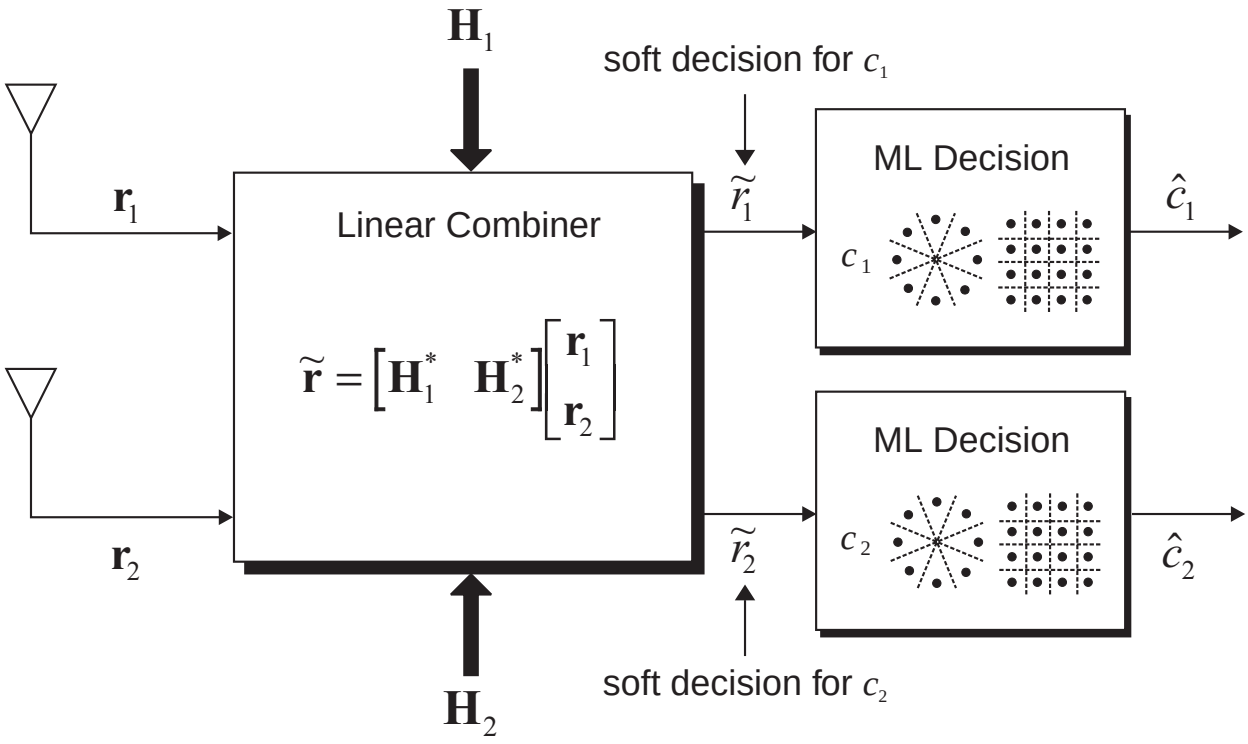

Fig. 8. Receiver for Space-Time Block Coding

TABLE I

Example of standardized MIMO channels for IEEE body 802.16

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

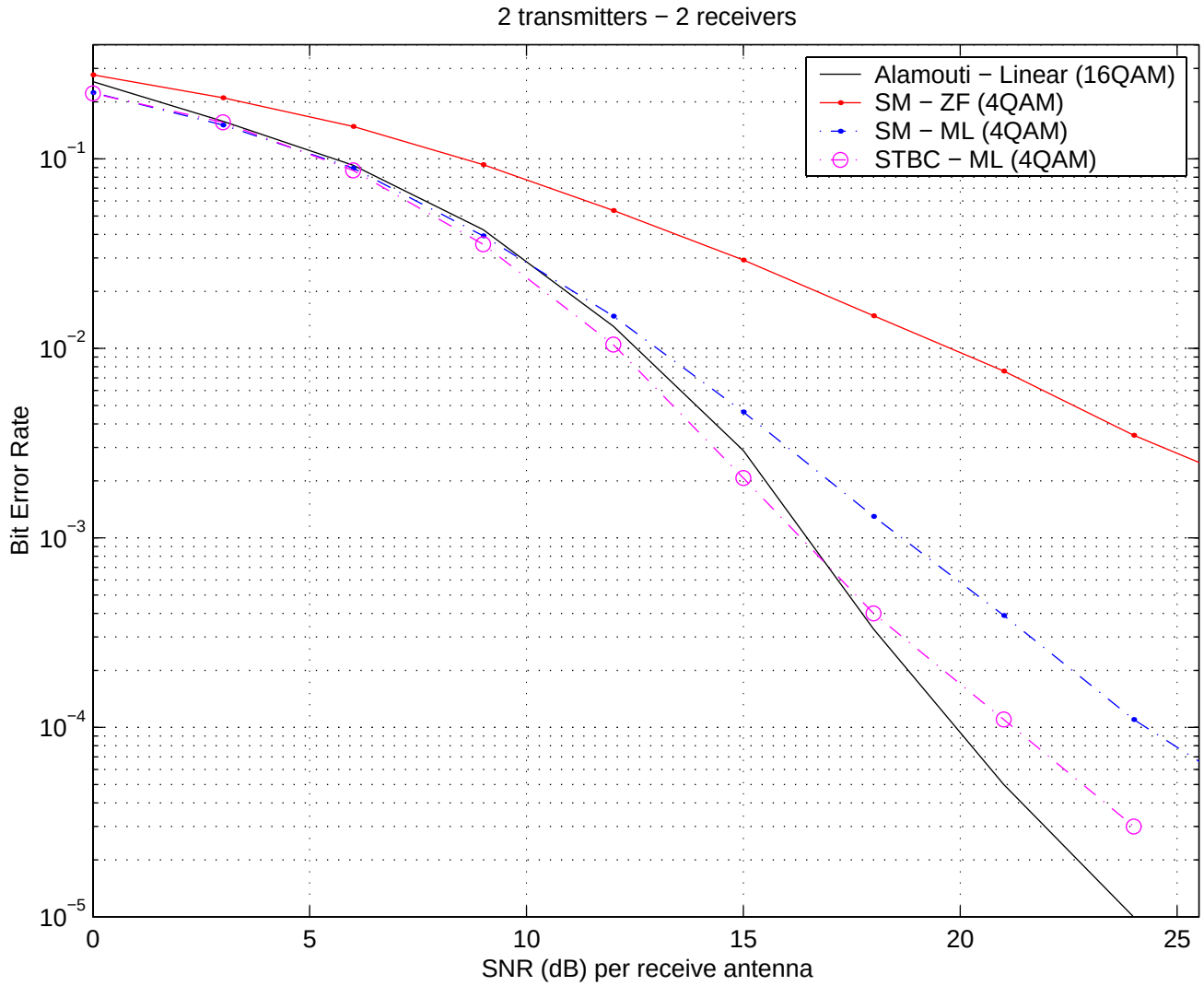

Fig. 9. Bit Error Rate (BER) comparisons for various transmission techniques over 2×2 MIMO. At high SNR, from top to bottom: Spatial multiplexing (SM)-ZF, SM-ML, STBC-ML, Alamouti STBC.

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

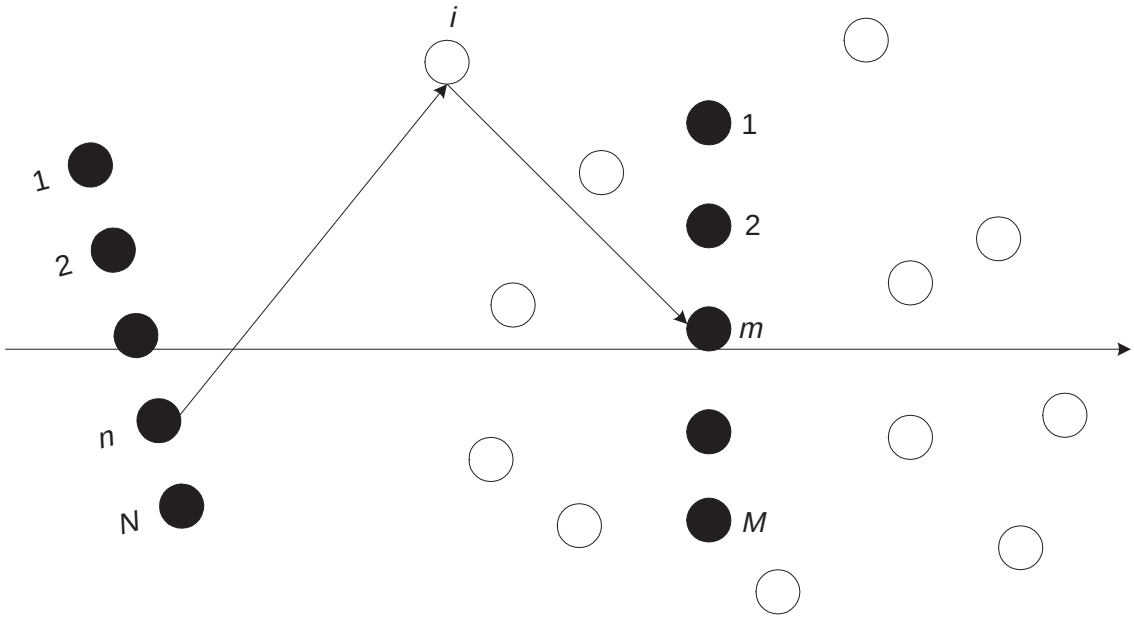

Fig. 10. Diagram to derive antenna correlation Ψ. The i-th ($1 \leq i \leq n_s$) path from TX antenna n to RX antenna m goes through the i-th single-bounce scatterer.

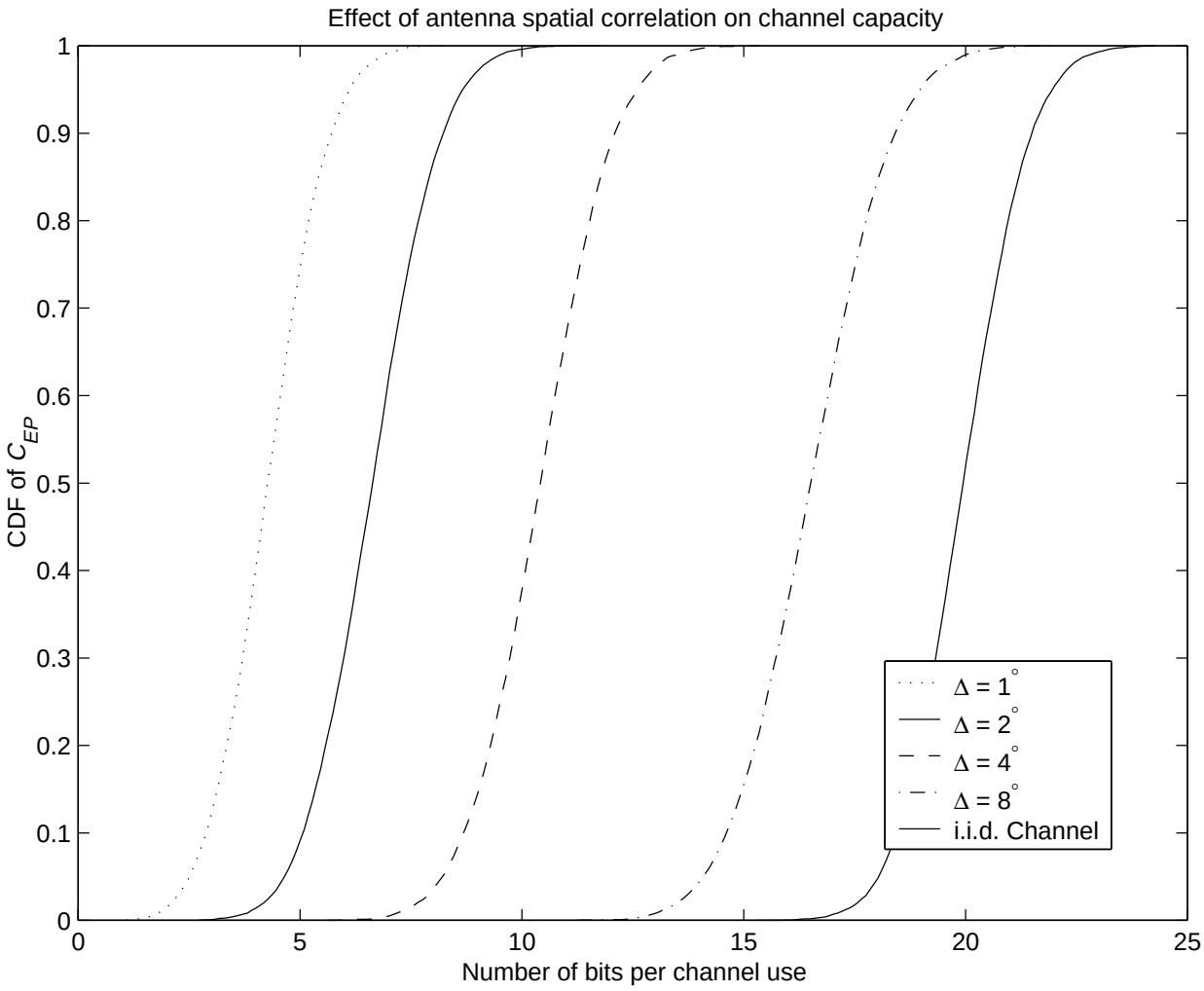

Fig. 11. Distribution of capacity as a function of angle spread for an (8, 8) system with $\rho = 8$. In producing this figure, transmit power is evenly divided, $\Theta = 0^\circ$, $D(T_n, T_{n'}) = 3\lambda|n - n'|$, and $D(R_m, R_{m'}) = \lambda|m - m'|$.

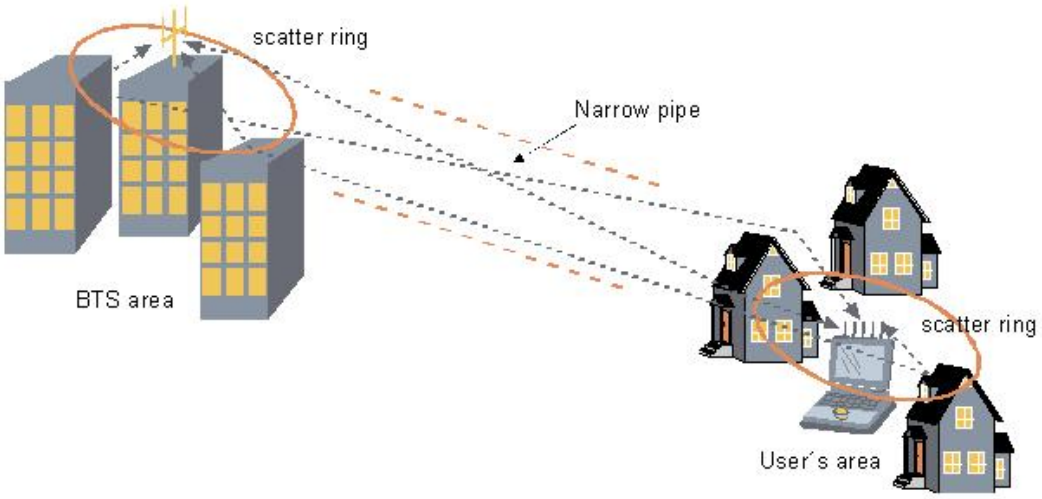

Fig. 12. An example of pin-hole realization. Reflections around the BTS and subscribers cause locally uncorrelated fading. However, because the scatter rings are too small compared to the separation between the two rings, the channel rank is low.