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

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

Tutorial Paper

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

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

I. Introduction

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

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

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

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

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

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

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

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

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

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

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

Digital Object Identifier 10.1109/JSAC.2003.809458

[1]Third-generation wireless UMTS-WCDMA.

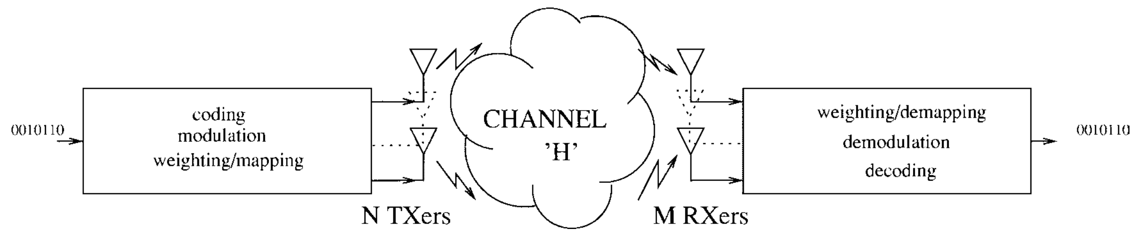

Fig. 1. Diagram of a MIMO wireless transmission system. The transmitter and receiver are equipped with multiple antenna elements. Coding, modulation, and mapping of the signals onto the antennas may be realized jointly or separately.

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

II. Principles of Space-Time (MIMO) Systems

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

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

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

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

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

In fact, the advantages of MIMO are far more fundamental. The underlying mathematical nature of MIMO, where data is transmitted over a *matrix* rather than a vector channel, creates new and enormous opportunities beyond just the added diver-

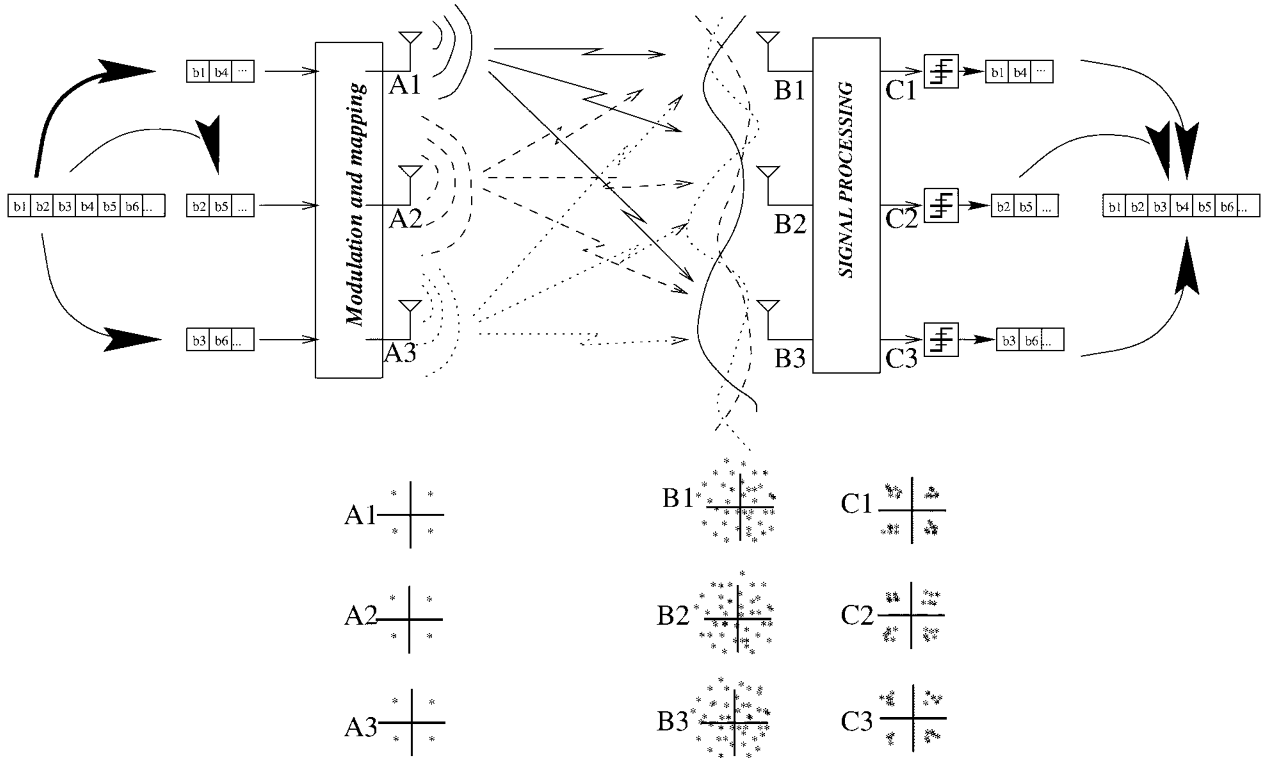

Fig. 2. Basic spatial multiplexing (SM) scheme with three TX and three RX antennas yielding three-fold improvement in spectral efficiency. Ai, Bi, and Ci represent symbol constellations for the three inputs at the various stages of transmission and reception.

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

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

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

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

Note that, unlike in CDMA where user's signatures are quasi-orthogonal by design, the separability of the MIMO channel relies on the presence of rich multipath which is needed to make the channel spatially selective. Therefore, MIMO can be said to effectively *exploit* multipath. In contrast, some smart antenna systems (beamforming, interference rejection-based) will perform better in line-of-sight (LOS) or close to LOS conditions. This is especially true when the optimizaparameters. Alternatively, diversity-oriented smart antenna techniques perform well in nonline-of-sight (NLOS), but they really try to mitigate multipath rather than exploiting it.

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

III. MIMO Information Theory

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

A. Fundamental Results

For a memoryless 1×1 (SISO) system the capacity is given by

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

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

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

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

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

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

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

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

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

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

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

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

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

B. *Information Theoretic MIMO Capacity*

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

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n}. \qquad (5)$$

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 additive white Gaussian noise (AWGN). The only interference present is *self-interference* between the input streams to the MIMO system. Some authors have considered more general systems but most information theoretic results can be discussed in this simple context, so we use (5) as the basic system equation.

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

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

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

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

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

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

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

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

$$\tilde{\mathbf{r}} = \mathbf{D}\tilde{\mathbf{s}} + \tilde{\mathbf{n}} \qquad (9)$$

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

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

We now give a brief overview of exact capacity results, broken down into the two main scenarios, where the channel is

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase® Smarter legal research.