

## Capacity of MIMO Channels With Antenna Selection

Shahab Sanayei, *Member, IEEE*, and  
Aria Nosratinia, *Senior Member, IEEE*

**Abstract**—This correspondence studies the capacity of multiple-input–multiple-output (MIMO) channels in the presence of antenna selection. Antenna selection reduces the complexity of the radio devices and requires only a small amount of channel state feedback. For high signal-to-noise ratio (SNR), we define *excess rate* as the constant term in the expansion of the ergodic capacity in terms of SNR. It is shown that this value is representative of the channel state information (CSI) at the transmitter. The asymptotic behavior of the excess rate is then analyzed for three cases: complete CSI, no CSI, and partial CSI at transmitter (antenna selection). While water-filling provides a excess rate that increases logarithmically in  $M$  (the number of transmit antennas), the excess rate of transmit antenna selection behaves only like  $\log(\log M)$ . For the low SNR case, we use the concept of *channel gain*, a measure introduced by Verdú. We show that channel gain for antenna selection increases only logarithmically in  $M$  as opposed to water-filling channel gain which increases linearly in  $M$ . The same techniques are also applied to the receive selection, and corresponding results are noted in high- and low-SNR regimes. The methodology developed in this correspondence, although motivated by antenna selection, is fairly general and can be used for any system where partial CSI is available at the MIMO transmitter.

**Index Terms**—Antenna selection, capacity scaling, channel state information (CSI), multiple-input–multiple-output (MIMO).

### I. INTRODUCTION

The cost and complexity of multiple radio frequency (RF)-chains, power amplifiers, and low-noise amplifiers (LNA) is an important practical issue in multiple-input–multiple-output (MIMO) systems. One solution to this cost/complexity problem is antenna subset selection [2]–[5].

At the receive side, antenna subset selection reduces the complexity. At the transmit side, antenna subset selection reduces complexity and increases capacity compared with an open loop MIMO system, using only a minimal amount of feedback [6]–[9]. Fast and efficient algorithms have been devised to determine the selected antenna subsets [10]–[12]. However, despite the practical importance of antenna selection, the information theoretic properties of the resulting channels remains a mostly unexplored territory. A few notable exceptions exist [13], [14], however, to date closed form expressions for capacity have not been available.

In this correspondence, we analyze the capacity of the antenna selection channel in the high-SNR and low-SNR regimes. In the high-SNR regime, we define the notion of *excess rate* as the constant term in the high-SNR expansion of the capacity expression, and demonstrate that it is directly related to transmit-side channel state information (CSI). This concept was first introduced in [8], [9] and is closely related to a similar concept that was independently proposed by Lozano *et al.* [15]. We are able to draw conclusions about the behavior of the system in the asymptote of large number of transmit antennas, and draw comparisons between the antenna selection capacity and water-filling capacity. Our

results also have interesting implications in the design and analysis of other MIMO systems beyond antenna selection. For example, the water-filling capacity ( $C_{\text{wf}}$ ) has the same growth rate as the capacity of the uninformed transmitter ( $C$ ) [16], but nevertheless  $C_{\text{wf}} > C$  with non-vanishing difference at high SNR.<sup>1</sup> The difference is an *excess rate* that is due to channel state feedback. The excess rate is not limited to water-filling; when partial CSI is available at the transmitter the excess rate is still there, but is smaller. Examples of partial CSI include transmit antenna selection and channel covariance feedback. The growth of this excess rate with the number of transmit antennas is a measure of how effectively CSI is being used by a given method.

The above developments, the reader may recall, were in the high-SNR regime. In the low-SNR regime we employ a similar methodology for analysis. In particular, we look at the power series expansion of the capacity around  $SNR = 0$ , where the coefficient of the first-order term is called *channel gain*,<sup>2</sup> a quantity that is related to the channel state information. Using this notion, we show that at low SNR, the optimal selection strategy at the transmitter is to select exactly one transmit antenna, regardless of other parameters. This is a new result that is reminiscent of, but distinct from, the well-known water-filling result at low SNR.

Finally, motivated by the symmetry inherent in the problem, we analyze the receive side antenna selection in a manner similar to transmit selection. Even if no multiplexing is lost due to selection, receive antenna selection incurs a loss of receive power due to de-selected antennas, thus, unlike transmit selection which may increase capacity (with respect to the open loop system that uses all antennas), receive selection always incurs a capacity loss.

In each of the low- and high-SNR regimes, we focus on three important cases: full CSI at transmitter, no CSI at transmitter, and antenna-selection CSI at transmitter. In the latter case, capacity analysis naturally depends on the antenna selection algorithm. Optimal antenna selection is not only practically difficult, it also induces channel distributions that do not lead to a tractable formulation. Thankfully there exist antenna selection algorithms that deliver capacities almost indistinguishable from optimal selection [8], [11]. In this correspondence, we undertake the analysis of antenna selection using these algorithms.

We use the following notation.  $\mathbb{E}[\cdot]$  refers to expected value of a random variable,  $I_N$  denotes the  $N \times N$  identity matrix,  $(x)^+ = \max\{x, 0\}$ , and  $\gamma \approx 0.57721566$  is the Euler-Mascheroni constant. We use  $a_n \stackrel{\circ}{=} b_n$  to denote the asymptotic equivalence of  $a_n$  and  $b_n$  defined as  $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$ . We use the natural logarithm throughout this correspondence so the capacity unit is in Nats/Sec/Hz. The chi-square distribution with  $2p$  degrees of freedom is shown by  $\chi_{2p}^2$ , and the maximum of  $n$  independent  $\chi_{2p}^2$  distributions is denoted by  $\hat{\chi}_{2p,n}^2$ . With an abuse of notation, we show random variables following the latter distribution also with  $\hat{\chi}_{2p,n}^2$ .

### II. SYSTEM MODEL

We assume a frequency nonselective (flat) linear time invariant fading channel between  $M$  transmit and  $N$  receive antennas. The signal model is

$$y(t) = Hx(t) + n(t) \quad (1)$$

where  $y(t)$  represents the  $N \times 1$  received vector sampled at time  $t$ , and  $x(t)$  represents the  $M \times 1$  vector transmitted by the antennas with power constrain  $\mathbb{E}[x^H x] \leq \rho$ , where  $\rho$  is the average SNR (per channel use),  $n(t)$  is the  $N \times 1$  additive circularly symmetric complex

<sup>1</sup>We assume there are more transmit than receive antennas.

<sup>2</sup>Terminology due to Verdú [1]

Manuscript received February 21, 2005; revised January 19, 2007.  
S. Sanayei is with Huawei Technologies, Plano, TX 75075 USA (e-mail: sanayei@ieee.org).

A. Nosratinia is with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: aria@utdallas.edu).

Communicated by B. Hassibi, Associate Editor for Communications.

Color versions of Figures 2–4 in this correspondence are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2007.907476

- 
- 1) Let  $\mathcal{S}_1 \equiv \{\text{all columns of } H\}$  and  $\tilde{P}_1 = I_N$ .
  - 2) choose  $h_1 = \arg \max_{h \in \mathcal{S}_1} \|h\|_2$ ,  $\tilde{H}_1 = h_1$ .
  - 3) for  $i = 2 : L$ 
    - a)  $\mathcal{S}_i = \mathcal{S}_{i-1} - \{\tilde{h}_{i-1}\}$
    - b)  $\tilde{P}_i = I - \frac{\rho}{L} \tilde{H}_{i-1} (I + \frac{\rho}{L} \tilde{H}_{i-1}^H \tilde{H}_{i-1})^{-1} \tilde{H}_{i-1}^H$
    - c)  $\tilde{h}_i = \arg \max_{h \in \mathcal{S}_i} h^H \tilde{P}_i h$
    - d)  $\tilde{H}_i = [\tilde{H}_{i-1} \tilde{h}_i]$
  - 4)  $\tilde{H} = \tilde{H}_L$
- 

Fig. 1. Antenna selection algorithm.

Gaussian noise vector with zero mean and covariance matrix equal to  $I_N$  (the  $N \times N$  identity matrix) and  $H$  is the  $N \times M$  channel matrix, whose  $ij$ th element is the scalar channel between the  $i$ th receive and  $j$ th transmit antenna. We assume that the elements of  $H$  are independent and have complex Gaussian distribution with zero mean and unit variance. We also assume that  $H$  is perfectly known at the receiver but it is not necessarily known at the transmitter. For antenna selection, we assume there is a rate-limited feedback channel from receiver to transmitter so that a subset of transmit antennas can be selected by the receiver and furthermore we assume that the feedback channel is without error or delay.

### III. TRANSMIT ANTENNA SELECTION

We consider a transmit antenna selection scheme where a subset of transmit antennas are used for transmission with equal power. Optimal transmit antenna selection via exhaustive search among all  $\binom{M}{L}$  combinations ( $L$  is the number of antennas to be selected) has complexity  $O(M^L)$ , which is impractical for large number of transmit antennas. One may reduce this complexity by employing a successive selection scheme, i.e., a greedy algorithm that at each step maximizes the capacity of the selected subchannel. A very similar methodology was mentioned in [10] for receive antenna selection. In this work, starting from the original channel matrix, the algorithm removes antennas one after another in a way that the capacity loss is minimized. Gharavi-Alkhansari and Greshman [12] showed that an incremental successive selection leads to less computational complexity. Simulation results show that this successive selection captures almost all the capacity of optimal antenna selection in a wide range of signal-to noise ratios (SNRs). Therefore, we adopt the latter algorithm for our analysis.

The input of the algorithm consists of  $\rho$  (the given SNR),  $L$  (the desired number of transmit antennas to be selected) and  $H$  (the original channel matrix). The output of the algorithm is  $\tilde{H}$  (the channel matrix associated with selected transmit antennas). This algorithm, shown in Fig. 1, is the basis for the following analysis.

#### A. A Framework for High SNR Analysis

Using the Sherman-Morrisson formula for determinants [17], for the selected channel  $\tilde{H}$  we have

$$\det \left( I_N + \frac{\rho}{L} \tilde{H} \tilde{H}^H \right) = \prod_{i=1}^L \left( 1 + \frac{\rho}{L} \tilde{h}_i^H \tilde{P}_i \tilde{h}_i \right). \quad (2)$$

As  $\rho \rightarrow \infty$ ,  $\tilde{P}_i \rightarrow P_i$ , where,  $P_i = I_N - \tilde{H}_{i-1} \left( \tilde{H}_{i-1}^H \tilde{H}_{i-1} \right)^{-1} \tilde{H}_{i-1}^H$  is a projection matrix of rank  $N - i + 1$ . When  $M$  is also large, at each selection step, the distribution of the remaining channel vectors can still be well approximated by a circularly symmetric Gaussian distribution. Our simulations verify that for large  $M$  the Gaussianity assumption provides a good approximation for the actual distribution of the remaining columns.<sup>3</sup> Using this assumption we can approximate the statistics of the right side of (2). We know that for an uncorrelated

<sup>3</sup>This asymptotic result has eluded a proof, and remains a conjecture. Even for simple cases it is an open problem and there is no known analytical result.

complex Gaussian vector  $x$  and a projection matrix  $P$ ,  $x^H P x$  has  $\chi^2$  distribution with  $\text{rank}(P)$  degrees of freedom. Hence for large  $\rho$  and large  $M$ , we use the following approximation:

$$\det(\tilde{H} \tilde{H}^H) \sim \prod_{i=1}^L \chi_{2(N-i+1), M-i+1}^2 \quad (3)$$

where  $\tilde{\chi}_{2p, n}^2$  stands for a random variable which is the maximum of  $n$  independent  $\chi_{2p}^2$  random variables. The probability density function (pdf) of this random variable can be computed in closed form [18]

$$f_{\tilde{\chi}_{2p, n}^2}(x) = \frac{n x^{p-1}}{(p-1)!} e^{-x} (1 - e^{-x} e_p(x))^{n-1} \quad (4)$$

where  $e_p(x) = \sum_{k=0}^{p-1} \frac{x^k}{k!}$ .

#### B. Excess Rate of MIMO Systems

We introduce the concept of excess rate as a measure of effectiveness of channel state information at the transmitter. We start with the capacity expression for a general MIMO system. Under the flat fading assumption, given a general channel matrix, the ergodic capacity of the MIMO channel is calculated as follows [19]:

$$C = \mathbb{E} \left[ \max_{\text{tr}(Q) \leq \rho} \log(\det(I_N + H Q H^H)) \right] \quad (5)$$

where  $Q = \mathbb{E}[x x^H]$  is the covariance of the transmitted vector  $x$ . We consider three different cases: First, uninformed transmission in which CSI is perfectly known at receiver, but not at transmitter. Second, informed transmission in which CSI is perfectly known both at transmitter and receiver. Third, transmission using an optimal subset of transmit antennas selected by the receiver. The only information available at the transmitter is the indices of the selected transmit antennas. Fig. 2 shows the ergodic capacity and the excess rate for the above three cases.

**Uninformed transmitter:** As shown in [19], when CSI is available only at the receiver, the covariance matrix that maximizes the capacity is of the form  $Q = \frac{1}{M} I_N$ , hence, the ergodic capacity is

$$\begin{aligned} C &= \mathbb{E} \left[ \log \left( \det \left( I_N + \frac{\rho}{M} H H^H \right) \right) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^m \log \left( 1 + \frac{\rho}{M} \lambda_i \right) \right] \end{aligned}$$

where  $\lambda_1, \dots, \lambda_m$  are ordered nonzero eigenvalues of the Wishart matrix  $H H^H$  [19], and

$$m = \text{rank}(H) = \min\{M, N\}$$

is the degrees of freedom of the MIMO channel, assuming channel is full-rank. As shown in [20],  $C = m \log \rho + O(1)$ . So the ergodic capacity grows linearly with  $m$ . Now we notice that in the asymptotic expansion of  $C$  there is a constant term that does not vanish as  $\rho \rightarrow \infty$ . Thus we define the excess rate as follows:

$$G \triangleq \lim_{\rho \rightarrow \infty} (C - m \log \rho). \quad (6)$$

For uninformed transmission, we have

$$\begin{aligned} G &= \lim_{\rho \rightarrow \infty} \left( \mathbb{E} \left[ \sum_{i=1}^m \log \left( 1 + \frac{\rho}{M} \lambda_i \right) \right] - m \log \rho \right) \\ &= \lim_{\rho \rightarrow \infty} \mathbb{E} \left[ \sum_{i=1}^m \log \left( \frac{1}{\rho} + \frac{\lambda_i}{M} \right) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^m \log \lambda_i \right] - m \log M \end{aligned} \quad (7)$$

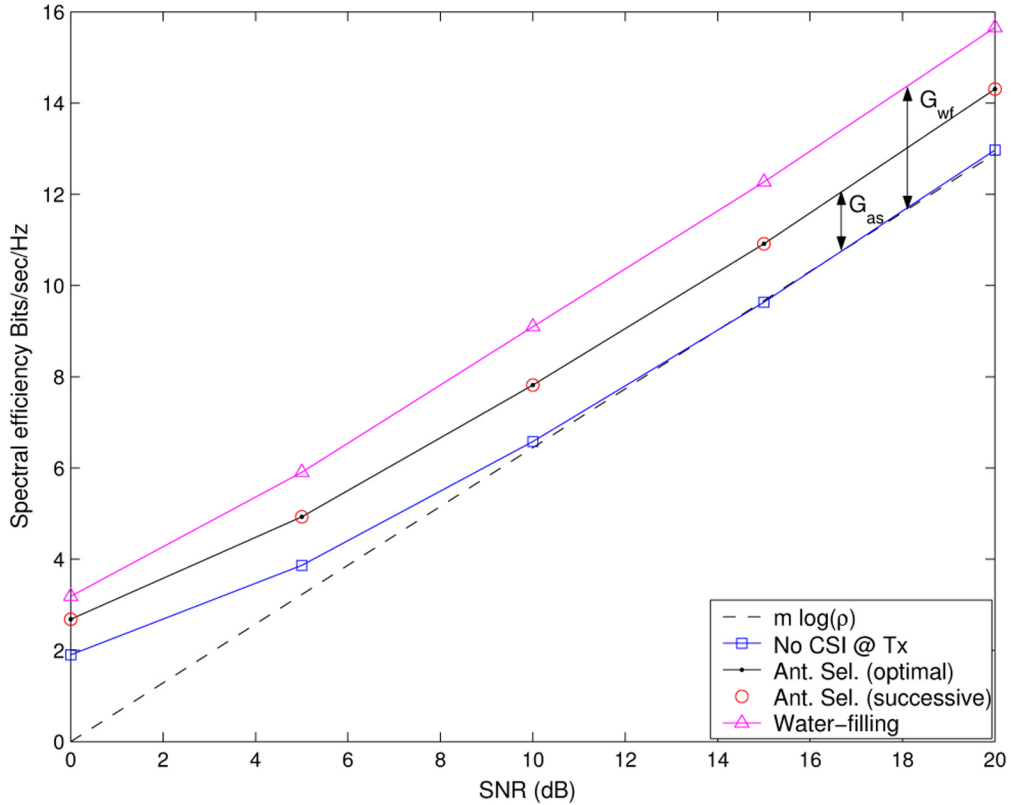


Fig. 2. Excess rate of antenna selection:  $M = 8$ ,  $N = L = 2$ .

in the last step, the exchange of expectation and limit is allowed by the monotone convergence theorem.<sup>4</sup>

**Informed transmitter** (water-filling capacity): In this case the channel state information is available at transmitter. As addressed in [19] the water-filling capacity of the MIMO channel is

$$C_{\text{wf}} = \mathbb{E} \left[ \sum_{i=1}^m (\log(\mu \lambda_i))^+ \right] \quad (8)$$

where  $\mu$  should satisfy  $\rho = \sum_{i=1}^m (\mu - \lambda_i^{-1})^+$ . In the large-SNR scenario, all the eigenmodes of the channel are used by the beamformer, hence,  $\mu = \frac{\rho + \sum_{i=1}^m \lambda_i^{-1}}{m}$  and the water-filling capacity is equal to

$$\begin{aligned} C_{\text{wf}} &= \mathbb{E} \left[ \sum_{i=1}^m (\log(\mu \lambda_i)) \right] \\ &= m \mathbb{E} \left[ \log \left( \rho + \sum_{i=1}^m \lambda_i^{-1} \right) \right] \\ &\quad + \mathbb{E} \left[ \sum_{i=1}^m \log \lambda_i \right] - m \log m. \end{aligned} \quad (9)$$

For large  $\rho$  we have  $C_{\text{wf}} \approx m \log \rho$ . In other words, availability of CSI at transmitter side has no impact on the logarithmic growth rate of the ergodic capacity, because the growth rate only depends

<sup>4</sup>Due to the fact that the  $\lambda_i$ 's have continuous cumulative density function (cdf) with an exponential tail, the logarithmic moment of eigenvalues exists. The details of the existence argument is omitted for the sake of brevity.

on the rank of the channel matrix. Now, we similarly calculate the excess rate for informed transmission

$$G_{\text{wf}} = \mathbb{E} \left[ \sum_{i=1}^m \log \lambda_i \right] - m \log m \quad (10)$$

$$\Delta G = G_{\text{wf}} - G = m \log(M/m). \quad (11)$$

In the asymptote of large SNR, this is the maximum amount of excess rate obtainable by providing CSI at the transmitter. We note that if  $M \leq N$  then  $\Delta G = 0$ , thus, channel state information at transmitter cannot provide any excess rate asymptotically. This result agrees with one's intuition that beamforming is not very effective with a small number of transmit antennas. In the sequel, we only consider the interesting case of  $M > N$ . In particular, we are interested to understand the behavior of the excess rate when  $M \gg N$ . In these cases, (11) suggests that at high SNR, the capacity gain can be used as an information-theoretic metric to evaluate any method that uses channel state information at the transmitter.

**Antenna selection:** In the high-SNR regime, one is interested in the case  $L \geq N$ , to maintain the degrees of freedom of the channel and prevent excessive rate loss. Suppose we have selected  $L$  ( $L \geq N$ ) out of  $M$  transmit antennas ( $M \gg N$ ) then if the selected channel is  $\tilde{H}$ , the excess rate is

$$\begin{aligned} \tilde{G} &= \lim_{\rho \rightarrow \infty} \mathbb{E} \left[ \log \left( \det \left( I_N + \frac{\rho}{L} \tilde{H} \tilde{H}^H \right) \right) \right] - N \log \rho \\ &= \lim_{\rho \rightarrow \infty} \mathbb{E} \left[ \log \left( \det \left( \frac{1}{\rho} I_N + \frac{1}{L} \tilde{H} \tilde{H}^H \right) \right) \right] \\ &= \mathbb{E} \left[ \log \left( \det(\tilde{H} \tilde{H}^H) \right) \right] - N \log L. \end{aligned} \quad (12)$$

## IV. ASYMPTOTIC BEHAVIOR OF EXCESS RATE

In this section we explore the behavior of excess rate, for large  $M$ , in the case of informed, uninformed, and antenna selection transmitter.

**Uninformed transmitter:** in the case  $M > N$ , (7) can be rewritten as

$$G = \mathbb{E} \left[ \log \det(HH^H) \right] - N \log M. \quad (13)$$

It is known [20] that  $\det(HH^H) \sim \prod_{i=1}^N \chi_{2(M-i+1)}^2$ , therefore, [21]

$$G = \sum_{i=1}^N (\psi(M-i+1) - \log M) \quad (14)$$

where  $\psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}$  is the di-gamma function. We have [20]

$$\lim_{M \rightarrow \infty} G = 0. \quad (15)$$

**Informed transmitter:** Using (11) and (15) for large  $M$  we have

$$G_{\text{wf}} \stackrel{\circ}{=} N \log \left( \frac{M}{N} \right) \stackrel{\circ}{=} N \log M. \quad (16)$$

**Antenna selection:** Using the results of Section III-A, we can evaluate the excess rate for antenna selection. Using (12)

$$\begin{aligned} \tilde{G} &= \mathbb{E} \left[ \log \det(\tilde{H}\tilde{H}^H) \right] - N \log L \\ &\stackrel{\circ}{=} \sum_{i=1}^L \mathbb{E} \left[ \log (\tilde{\chi}_{2(N-i+1), M-i+1}^2) \right] - N \log L. \end{aligned} \quad (17)$$

Equation (17) suggests that for large  $M$ , selecting more than  $N$  antennas does not provide any further gain. In the previous section we argued that  $L$  cannot be less than  $N$ , hence for large  $M$  the optimal value for  $L$  is  $N$ . Henceforth, we assume  $L = N$ . To evaluate the asymptotic behavior of  $\tilde{G}$ , we only need to evaluate  $\mathbb{E}[\log X]$ , where  $X \sim \tilde{\chi}_{p,n}^2$ . We use the following result from order statistics [18].

*Definition 1:* A cdf  $F$  is said to belong to the domain of maximal attraction of a nondegenerate cdf  $U$  if there exist sequences  $\{a_n\}$  and  $\{b_n > 0\}$  such that

$$\lim_{n \rightarrow \infty} F^n(a_n + b_n x) = U(x) \quad (18)$$

at all continuity points of  $U(x)$ .

*Theorem 1:* Let  $X_{(n)}$  be the maximum of  $n$  i.i.d. random variables  $\{X_i\}_{i=1}^n$ , each with cdf  $F$ . If  $F$  belongs to the domain of maximal attraction of  $U$ , then

$$\frac{X_{(n)} - a_n}{b_n} \xrightarrow{d} W \quad (19)$$

where  $W$  is a random variable whose cdf is  $U$ . Moreover,  $U$  can only be one of the following three distributions:

$$\begin{aligned} U_1(x) &= \exp(-e^{-x}) \quad -\infty < x < \infty \\ U_2(x) &= \begin{cases} \exp(-x^\alpha), & x > 0, \alpha > 0 \\ 0, & x \leq 0 \end{cases} \\ U_3(x) &= \begin{cases} 1, & x > 0 \\ \exp(-(-x)^\alpha) & x \leq 0, \alpha > 0. \end{cases} \end{aligned}$$

These distributions are also known as Gumbel, Fréchet, and Weibull distributions, respectively.

Theorem 1 is analogous to the central limit theorem which was for the normalized sum of independent and identically distributed (i.i.d.)

random variables. Although unlike the central limit theorem, the normalization constants  $a_n$  and  $b_n$  are not necessarily unique. One way to find the constant  $a_n$  is to solve the following equation [18]:

$$\begin{aligned} a_n &= F^{-1} \left( 1 - \frac{1}{n} \right) \\ b_n &= F^{-1} \left( 1 - \frac{1}{ne} \right) - F^{-1} \left( 1 - \frac{1}{n} \right) \end{aligned}$$

where  $F^{-1}$  is the inverse function. It is known that the cdf of a  $\chi_{2p}^2$  random variable is equal to  $F(x) = 1 - e^{-x} e_p(x)$ , where  $e_p(x) = \sum_{k=0}^{p-1} \frac{x^k}{k!}$ . Choosing  $a_n = \log n + \log \left( \log \left( \frac{n^{p-1}}{(p-1)!} \right) \right)$  and  $b_n = 1$  we have

$$\lim_{n \rightarrow \infty} F^n(a_n + b_n x) = \exp(-e^{-x}). \quad (20)$$

We use (20) to evaluate the logarithmic moment of  $\tilde{\chi}_{p,n}^2$  which is key to our analysis. In the above formulation, we choose  $F$  to be the cdf of a chi-square random variable with  $p$  degrees of freedom. Jensen's inequality provides an upper bound on the logarithmic moment and furthermore the following theorem states that this bound is asymptotically tight.

*Theorem 2:* Let  $\{X_n\}_{i=1}^n$  be positive random variables with finite mean and variance,  $X_{(n)}$ ,  $W$ ,  $\{a_n\}$  and  $\{b_n > 0\}$  are defined as in Theorem 1 and furthermore,  $\frac{|a_n|}{b_n} \rightarrow \infty$ , then as  $n \rightarrow \infty$

$$\log(\mathbb{E}[X_{(n)}]) - \mathbb{E}[\log X_{(n)}] \rightarrow 0.$$

*Proof:* See Appendix I.

So to calculate the logarithmic moment it is sufficient to evaluate the mean of the above random variable. It is known that when the limiting distribution is of the first kind (Gumbel distribution) then Theorem 1 can also be used to evaluate the moments of  $X_{(n)}$  (in fact for this case, convergence in distribution implies convergence in moments [22]). In particular, we have

$$\mathbb{E} \left[ \frac{X_{(n)} - a_n}{b_n} \right] \rightarrow \mathbb{E}[W] = \gamma \quad (21)$$

$$\mathbb{E} \left[ \left( \frac{X_{(n)} - a_n}{b_n} \right)^2 \right] \rightarrow \mathbb{E}[W^2] = \frac{\pi^2}{6}. \quad (22)$$

Hence, the asymptotic growth of the logarithmic moment of the extreme order statistics of chi-square random variables is given by

$$\begin{aligned} \mathbb{E}[\log(X_{(n)})] &\stackrel{\circ}{=} \log(\mathbb{E}[X_{(n)}]) \\ &\stackrel{\circ}{=} \log \left( \log n + \log \left( \log \left( \frac{n^{p-1}}{(p-1)!} \right) \right) + \gamma \right) \\ &\stackrel{\circ}{=} \log(\log n). \end{aligned} \quad (23)$$

Now we can evaluate the behavior of  $\tilde{G}$  in (17):

$$\begin{aligned} \tilde{G} &\stackrel{\circ}{=} \sum_{i=1}^N \mathbb{E} \left[ \log (\tilde{\chi}_{2(N-i+1), M-i+1}^2) \right] - N \log N \\ &\stackrel{\circ}{=} \sum_{i=1}^N \log \left( \frac{\log(M-i+1) + \log \left( \log \left( \frac{(M-i+1)^{N-i}}{(N-i)!} \right) \right) + \gamma}{N} \right). \end{aligned} \quad (24)$$

Hence

$$\tilde{G} \stackrel{\circ}{=} N \log(\log M).$$

Thus, the excess rate for transmit antenna selection behaves like  $O(\log(\log M))$ . Also, for large number of transmit antennas (24) can

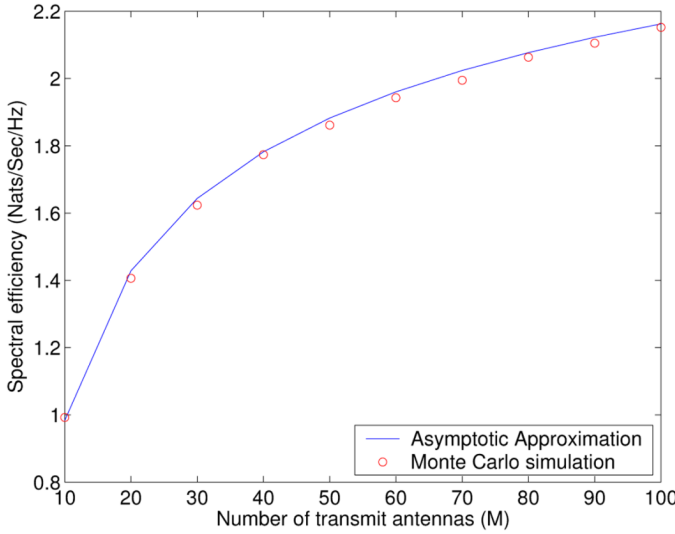


Fig. 3. Excess rate of antenna selection ( $N = 2$  and SNR = 30 dB).

be used as an approximate formula for ergodic capacity of transmit antenna selection at high SNR, in fact

$$\tilde{C} \approx N \log \rho + \tilde{G}. \quad (25)$$

Fig. 3 compares our result with computer simulations. We run the simulation for SNR = 30 dB and  $N = 2$ . For each point on the plot, the excess rate is calculated by averaging over 5000 different channel realizations and compared to the results obtained from (24). Simulations match our analysis very well, thus, the asymptotic formula proves to be a useful tool for the evaluating of the capacity of transmit antenna selection.

## V. TRANSMIT ANTENNA SELECTION: LOW SNR

In the low-SNR regime, we use the concept of *channel gain*, introduced by Verdú [1], which is essentially the slope of the linear term in the Taylor expansion of the ergodic capacity, namely

$$\Gamma \triangleq \left. \frac{\partial C(\rho)}{\partial \rho} \right|_{\rho=0} \quad (26)$$

since  $C = \Gamma \rho + O(\rho^2)$ ,  $\Gamma$  can be considered as an information-theoretic metric for evaluating the spectral efficiency at low SNR [1].

**Uninformed transmitter:** In this case the channel gain [1] is:

$$\Gamma = \mathbb{E} \left[ \frac{\|H\|_F^2}{M} \right] = N \quad (27)$$

we notice that in this case the channel gain does not depend on  $M$ , so increasing the number of transmit antenna will not affect the capacity.

**Informed transmitter:** When CSI is fully provided at the transmitter, at low SNR, the beamformer only uses the eigenmode of the channel associated with the largest eigenvalue of  $H^H H$  hence, the channel gain is [1]

$$\Gamma_{\text{wf}} = \mathbb{E} \left[ \lambda_{\max} \left( H^H H \right) \right]. \quad (28)$$

It was first shown in [23] that  $\lambda_{\max} \triangleq \left( \sqrt{M} + \sqrt{N} \right)^2$  when  $M$  or  $N$  are large. Thus for the case  $M \gg N$ , we have  $\Gamma_{\text{wf}} \triangleq M$ .

**Antenna selection:** Suppose we are selecting  $L$  transmit antennas with equal power-splitting among them. In the low-SNR scenario, the channel gain is

$$\tilde{\Gamma} = \mathbb{E} \left[ \frac{\|\tilde{H}\|_F^2}{L} \right] = \frac{\mathbb{E} \left[ \sum_{i=1}^L \|\tilde{h}_i\|_2^2 \right]}{L}$$

where  $\tilde{H}$  is the selected channel, and  $\tilde{h}_1, \dots, \tilde{h}_L$  are the  $L$  columns of  $\tilde{H}$ . Thus antenna selection in low SNR leads to selecting  $L$  antennas with the highest norms. This is also consistent with the successive antenna selection algorithm presented in Fig. 1 for  $\rho \approx 0$ . Moreover, the channel gain is maximized when  $L = 1$ , because  $\frac{\mathbb{E} \left[ \sum_{i=1}^L \|\tilde{h}_i\|_2^2 \right]}{L} \leq \max_i \left\{ \|\tilde{h}_i\|_2^2 \right\}$ . This suggests that the optimal transmit antenna selection strategy in low SNR is to select only one transmit antenna whose channel vector is of the highest norm. In other words,  $\tilde{H} = h_j$  where  $\|h_j\| = \max \{ \|h_1\|, \dots, \|h_M\| \}$ . Now in order to evaluate the channel gain for low SNR transmit antenna selection we need to evaluate  $\mathbb{E} \left[ \|\tilde{H}\|_F^2 \right]$ . The random variable  $\|\tilde{H}\|_F^2$  is distributed according to  $\tilde{X}_{2N,M}$  defined in Section III-A. For  $M \gg N$ , using Theorem 1 we have

$$\tilde{\Gamma}_{\text{opt}} \triangleq \log M + \log \left( \log \left( \frac{M^{N-1}}{(N-1)!} \right) \right) + \gamma \triangleq \log M \quad (29)$$

## VI. RECEIVE ANTENNA SELECTION

Up to this point, the emphasis of the correspondence has been on the transmitter side. However, it is not difficult to see that due to the reciprocity of electromagnetic propagation, the problem is highly symmetric, therefore we can use the framework developed thus far to also address receive antenna selection.

In particular, consider the following setup: An MIMO system with  $M$  transmit and  $N$  receive antennas, such that  $N \gg M$ . We wish to choose  $L$  out of  $N$  receive antennas in a way to maximize the retained capacity. The algorithm depicted in Fig. 1 will perform the antenna selection, with the notable difference that we now must select *rows* and not columns of  $H$ . As before, we call the selected channel  $\tilde{H}$ . Assuming receive antenna selection with  $L = M$  and no CSI at transmitter, the capacity of the system is

$$C = \log \det \left( I_L + \frac{\rho}{M} \tilde{H} \tilde{H}^H \right).$$

Similarly to the previous case, we can write

$$\begin{aligned} \log \det \left( I + \frac{\rho}{M} \tilde{H} \tilde{H}^H \right) &\approx \log \prod_{i=1}^M \left( 1 + \frac{\rho}{M} \tilde{\chi}_{2M, N-i+1}^2 \right) \\ &\approx M \log \rho - M \log M \\ &\quad + \sum_{i=1}^M \log \tilde{\chi}_{2M, N-i+1}^2 \\ &= M \log \rho + \hat{G} \end{aligned}$$

where  $\tilde{\chi}_{2M, N-i+1}^2$  is as earlier defined, the maximum of  $N - i + 1$  chi-square random variables. Recall that in transmit selection the SNR scales by the factor  $\rho/L$  in (2), i.e., the fewer the selected antennas, the more power can be sent through each antenna. The result was that

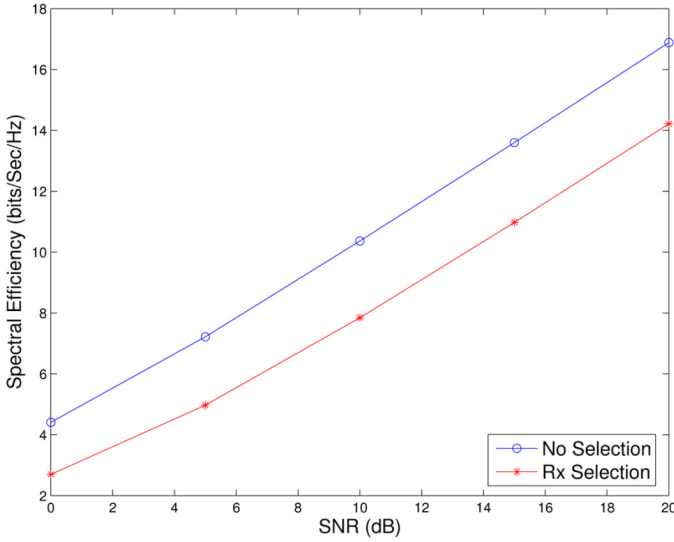


Fig. 4. Receive antenna selection in a  $2 \times 8$  system.

the capacity actually increases through transmit antenna selection,<sup>5</sup> an increase that was characterized by  $\tilde{G}$ .

In receive selection, selecting fewer antennas will result in smaller receive power, but antennas that are selected enjoy better channel distributions than the original MIMO channel. However, the loss of power cannot be made up by the improvement in the channel distributions. Therefore the capacity of the receive antenna selection is less than the capacity of the full-scale system. Unlike transmit selection, no additional capacity is obtained by selecting down to the best antennas, a result that is not surprising because information is being lost by receive selection. The above results are demonstrated on a  $2 \times 8$  system in Fig. 4.

In the low-SNR regime, we once again use the concept of channel gain, which for the receive antenna selection leads to

$$\tilde{\Gamma} = \mathbb{E} \left[ \frac{\|\tilde{H}\|_F^2}{M} \right] = \mathbb{E} \left[ \frac{\sum_{i=1}^L \|\tilde{h}_i\|_2^2}{M} \right].$$

It is evident that, unlike transmit-antenna selection, there is no penalty for selecting more antennas, in fact the more antennas selected, the higher the capacity. Calculating this value using Theorem 1, we have

$$\tilde{\Gamma} \stackrel{\circ}{=} \frac{L}{M} \log N. \quad (30)$$

## VII. CONCLUSION

In this correspondence, we investigate the behavior of the capacity of antenna selection in the asymptote of large number of transmit antennas. For high SNR, we introduce the concept of *excess rate*, an information theoretic metric for schemes that use channel state information at the transmitter. It is shown that this quantity describes the advantage gained by having channel state information at the transmitter. We present new results in order statistics that are useful for asymptotic analysis of antenna selection problems. By exploring the behavior of the excess rate, we show that the optimal number of selected antennas,

<sup>5</sup>Assuming the multiplexing gain of the system is unaffected by selection, which we ensured via our transmit-selection assumption of  $N \geq L \geq M$

when number of transmit antennas  $M$  is large, is exactly equal to the number of receive antennas  $N$ . Simulations verify the accuracy of the analysis. For low SNR, we show that the optimal selection strategy is to choose only one transmit antenna with the highest channel norm. Also, we evaluate the channel gain and show its logarithmic behavior in the asymptote of large number of antennas.

## APPENDIX

*Lemma 1:* If  $\frac{|a_n|}{b_n} \rightarrow \infty$ , then  $\log \left( \frac{X(n)}{a_n} \right) \xrightarrow{d} 0$ .

*Proof:* For every  $\epsilon > 0$  and  $\delta > 0$

$$\begin{aligned} Pr \left[ \left| \frac{X(n)}{a_n} - 1 \right| > \epsilon \right] &= Pr \left[ \left| \frac{X(n) - a_n}{b_n} \right| > \epsilon \frac{|a_n|}{b_n} \right] \\ &\leq \frac{\mathbb{E} \left[ \left( \frac{X(n) - a_n}{b_n} \right)^2 \right]}{\epsilon^2 \left( \frac{a_n}{b_n} \right)^2} \\ &< \frac{\mathbb{E}[W^2] + \delta}{\epsilon^2 \left( \frac{a_n}{b_n} \right)^2} \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ . Thus,  $\frac{X(n)}{a_n} \xrightarrow{i.p.} 1$ , hence,  $\frac{X(n)}{a_n} \xrightarrow{d} 1$ . Now using the continuous mapping theorem [24], we conclude that  $\log \left( \frac{X(n)}{a_n} \right) \xrightarrow{d} 0$ .

*Lemma 2:* Let  $\mu_n = \mathbb{E}[X(n)]$ , if  $\frac{|a_n|}{b_n} \rightarrow \infty$ , then  $\log \left( \frac{\mu_n}{a_n} \right) \rightarrow 0$ .

*Proof:* From (21), we have  $\frac{\mu_n - a_n}{b_n} \rightarrow 0$ , also we have  $\frac{b_n}{a_n} \rightarrow 0$ . By multiplying two sides, we get  $\frac{\mu_n}{a_n} - 1 \rightarrow 0$ , hence,  $\log \left( \frac{\mu_n}{a_n} \right) \rightarrow 0$ .

*Definition 2:* The sequence of random variables  $\{X_n\}$  is called *uniformly integrable* if [24]

$$\lim_{c \rightarrow \infty} \limsup_n \int_{|X_n| > c} |x| dF_{X_n}(x) < \infty. \quad (31)$$

*Lemma 3:* If the random variables  $\{X_n\}$  have finite mean, then uniform integrability is equivalent to the following condition:

$$\lim_{c \rightarrow \infty} \limsup_n \int_{|X_n| > c} Pr[|X_n| > x] dx < \infty. \quad (32)$$

*Proof:* Using integration by part, we have

$$\int_{|X_n| > c} |x| dF_{X_n}(x) = -x(1 - F_{X_n}(x)) \Big|_c^\infty + \int_c^\infty Pr[|X_n| > x] dx. \quad (33)$$

Since  $\mathbb{E}[X_n] < \infty$ , we have  $\lim_{c \rightarrow \infty} c(1 - F_{X_n}(c)) = 0$  and this proves the lemma.

*Theorem 3:* If the sequence of random variables  $\{X_n\}$  is uniformly integrable, then  $X_n \xrightarrow{d} X$  implies  $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$ .

*Proof:* See [24].

*Proof of Theorem 2:* Let  $Y_n = \log \left( \frac{X_n}{a_n} \right)$ , then by Lemma 1 we have  $Y_n \xrightarrow{d} 0$ , we show that  $Y_n$  is uniformly integrable, and to do so we use the alternative form of uniform integrability provided by Lemma 2

$$\begin{aligned} \int_c^\infty Pr[|Y_n| > y] dy &= \int_c^\infty Pr[Y_n > y] dy \\ &\quad + \int_c^\infty Pr[Y_n < -y] dy \\ &= I_1 + I_2. \end{aligned} \quad (34)$$

We evaluate each integral

$$\begin{aligned}
 I_1 &= \int_c^\infty \Pr[Y_n > y] dy \\
 &= \int_c^\infty \Pr[X_n > a_n e^y] dy \\
 &= \int_c^\infty \Pr\left[\frac{X_n - a_n}{b_n} > \frac{a_n}{b_n}(e^y - 1)\right] dy \\
 &= \int_\alpha^\infty \Pr\left[\frac{X_n - a_n}{b_n} > \frac{a_n}{b_n}t\right] \frac{dt}{t+1} \quad (35)
 \end{aligned}$$

where in the last equation  $t := e^y - 1$  and  $\alpha = e^c - 1$ . Let  $\eta_n \triangleq \frac{a_n}{b_n}$ , therefore,  $\eta_n \rightarrow \infty$  as  $n \rightarrow \infty$ . From (22) we have  $\mathbb{E}\left[\left(\frac{X_n - a_n}{b_n}\right)^2\right] \rightarrow \frac{\pi^2}{6}$ . Thus using Chebyshev's inequality, for every  $\delta > 0$  we have

$$\begin{aligned}
 I_1 &\leq \frac{1}{\eta_n^2} \int_\alpha^\infty \frac{\mathbb{E}\left[\left(\frac{X_n - a_n}{b_n}\right)^2\right]}{t^2(t+1)} dt \\
 &\leq \frac{1}{\eta_n^2} \int_\alpha^\infty \frac{\pi^2/6 + \delta}{t^2(t+1)} dt \\
 &\leq \frac{1}{\eta_n^2(\alpha+1)} \int_\alpha^\infty \frac{\pi^2/6 + \delta}{t^2} dt \\
 &= \frac{\pi^2/6 + \delta}{\eta_n^2 \alpha(\alpha+1)} \rightarrow 0. \quad (36)
 \end{aligned}$$

Also in a similar way for all  $\delta > 0$  we have

$$\begin{aligned}
 I_2 &= \int_c^\infty \Pr[Y_n < -y] dy \\
 &= \int_c^\infty \Pr[X_n < a_n e^{-y}] dy \\
 &= \int_c^\infty \Pr\left[\frac{a_n - X_n}{b_n} > \frac{a_n}{b_n}(1 - e^{-y})\right] dy \\
 &\leq \left(\frac{b_n}{a_n}\right)^2 \int_c^\infty \frac{\mathbb{E}\left[\left(\frac{a_n - X_n}{b_n}\right)^2\right]}{(1 - e^{-y})^2} dy \\
 &\leq \left(\frac{b_n}{a_n}\right)^2 \int_c^\infty \frac{\pi^2/6 + \delta}{(1 - e^{-y})^2} dy \\
 &= \left(\frac{b_n}{a_n}\right)^2 \left(\frac{\pi^2}{6} + \delta\right) \left(\frac{e^{-c}}{1 - e^{-c}} + \log(1 - e^{-c})\right) \\
 &\rightarrow 0 \quad (37)
 \end{aligned}$$

from (36) and (37) we conclude that  $I_1 + I_2 \rightarrow 0$  thus,  $Y_n$  is uniformly integrable hence  $Y_n$  converges in mean, namely,  $\mathbb{E}[\log X_{(n)}] - \log a_n \rightarrow 0$ . On the other hand, by Lemma 2 we have  $\log(\mathbb{E}[X_{(n)}]) - \log a_n \rightarrow 0$ , thus, we have  $\mathbb{E}[\log X_{(n)}] - \log(\mathbb{E}[X_{(n)}]) \rightarrow 0$ .

## REFERENCES

- [1] S. Verdú, "Spectral efficiency in wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1319–1343, Jun. 2002.
- [2] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
- [3] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, pp. 1773–1776, Dec. 1999.
- [4] R. Heath and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. Int. Conf. Commun.*, Helsinki, Finland, Jun. 2003, pp. 2276–2280.

- [5] D. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE ICASSP*, Istanbul, Turkey, May 2000, pp. 2785–2788.
- [6] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.
- [7] D. Gore, R. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, pp. 491–493, Nov. 2002.
- [8] S. Sanayei and A. Nosratinia, "Asymptotic capacity gain of transmit antenna selection," in *Proc. WNCG Symp.*, Austin, TX, Oct. 2003.
- [9] S. Sanayei and A. Nosratinia, "Asymptotic capacity analysis of transmit antenna selection," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, pp. 241–241.
- [10] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems," in *Proc. IEEE ICASSP*, Orlando, FL, May 2002, pp. 2875–2860.
- [11] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms," *IEEE Trans. Signal Process.*, pp. 2796–2807, Nov. 2003.
- [12] M. Gharavi-Alkhanisari and A. Greshman, "Fast antenna selection in MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, pp. 339–347, Feb. 2003.
- [13] A. Gorokhov, D. Gore, and A. Paulraj, "Receive antenna selection for MIMO flat fading channels: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2687–2696, Oct. 2003.
- [14] M. Win, A. Molisch, and Winters, "Capacity of MIMO systems with antenna selection," in *Proc. Int. Conf. Commun.*, 2001, vol. 2, pp. 570–574.
- [15] A. Lozano, A. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun. 2004, pp. 288–288.
- [16] D. Bliss, K. Forsythe, A. Hero, and A. Yegulalp, "Environmental issues for MIMO capacity," *IEEE Trans. Signal Process.*, vol. 50, pp. 2128–2142, Sep. 2002.
- [17] G. Kéri, "The Sherman-Morrisson formula for the determinant and its application for optimizing the quadratic functions on conditions sets given by extreme generators," in *Optimization Theory: Recent Developments from Mátraháza*, F. Giannessi, P. Pardalos, and T. Rapcsák, Eds. Boston, MA: Kluwer, 2001.
- [18] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*. New York: Wiley, 1992.
- [19] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.
- [20] G. J. Foschini and M. J. Gans, "On limits of wireless communication in fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.
- [21] O. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels: Insights into diversity-multiplexing tradeoff," *IEEE Trans. Signal Process.*, vol. 51, pp. 2784–2795, Nov. 2003.
- [22] J. Galambos, *Asymptotic Theory of Order Statistics*. New York: Krieger, 1987.
- [23] S. Geman, "A limit theorem for the norm of the random matrices," *Ann. Probabil.*, vol. 8, pp. 252–261, 1980.
- [24] P. Billingsley, *Probability and Measure*. New York: Wiley, 1995.