

A Power Scaling Analysis of Norm-Based Antenna Selection Techniques

Peter J. Smith, *Senior Member, IEEE*, Timothy W. King, *Student Member, IEEE*,
Lee M. Garth, *Senior Member, IEEE*, and Mischa Dohler, *Senior Member, IEEE*

Abstract—In this paper we consider transmit and receive selection methods designed to achieve high channel capacities in a single-user MIMO link. A variety of radio channels are considered, including i.i.d. Rayleigh, correlated Rayleigh and Ricean fading environments. Also considered is the presence of imperfect channel state information (CSI) and a simplified waterfilling scheme. In all cases, we evaluate the performance of optimal selection, simple norm-based selection and other benchmark selection techniques. The major contribution is a general approach to analyzing the capacity of the norm-based selection schemes via a simple power scaling factor. We are able to assess the impact of different channels, imperfect CSI and power allocation using this power scaling factor. Furthermore, the analysis is valid for all scenarios: transmit selection, receive selection and joint transmit-receive selection. Results are shown which compare the capacity performance over a wide range of cases. A notable conclusion is that optimal selection, which is computationally intensive, is outperformed at low signal-to-noise-ratios by the simple norm-based approach with power allocation.

Index Terms—Multiple-input multiple-output systems, antenna selection techniques, Rayleigh fading, Ricean fading.

I. INTRODUCTION

GIVEN the advantages of wireless multiple-input multiple-output (MIMO) systems using transmit (TX) and receive (RX) antenna arrays, much attention has been paid to these systems and their potential capacity gains. However, to achieve these gains, an increasing number antenna elements at either end as well as the associated baseband complexity to process MIMO capacity-approaching codes are required. In practical systems, size constraints and power consumption of the MIMO array also become important issues. For example, with the continuous miniaturization of cellular mobile terminals (MTs), the space for accommodating several antenna elements diminishes. Also, each radio frequency (RF) chain requires proper amplification prior to transmission. Whilst

Manuscript received February 20, 2007; revised November 20, 2007; accepted February 14, 2008. The associate editor coordinating the review of this letter and approving it for publication was H. Shin. This work was partially supported by the Foundation for Research, Science and Technology of New Zealand, Telecom New Zealand Limited and a University of Canterbury Doctoral Scholarship. The material in this paper was presented in part at the IEEE Int'l. Conf. on Communications, Istanbul, Turkey, June 2006.

P. J. Smith and T. W. King are with the Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand (e-mail: p.smith@elec.canterbury.ac.nz; twk20@student.canterbury.ac.nz).

L. M. Garth is with the Advanced Systems and Technology Division, BAE Systems, Merrimack, NH 03054 USA (e-mail: lee.garth@ieee.org).

M. Dohler is with the Telecommunications Technologies Centre of Catalonia, Parc Mediterrani de la Tecnologia, Av. Canal Olímpic S/N, 08860 – Castelldefels, Barcelona, Spain (e-mail: mischa.dohler@cttc.es).

Digital Object Identifier 10.1109/TWC.2008.070211.

highly linear amplifiers are very power efficient, they are also very expensive. Today's typical 3GGP MT transmitter, for instance, uses a cheap non-linear amplifier per antenna, which consumes between 20% and 30% of the MT's power budget. Maintaining many of these RF chains is hence very power consuming.

In this paper we therefore consider a simple yet powerful approach for reducing the system complexity and RF chain power consumption whilst retaining a considerable portion of the increased rates from using MIMO technology. Using the principle of antenna selection at the TX and/or RX side, RF chains can be optimally assigned to a subset of the available antennas. This provides greater spatial capacity gains without a large increase in the system's hardware needs. The only caveat is that a feedback channel is required from the RX to the TX, if TX antenna selection is to be performed; however, all modern communication systems have feedback channels available.

Antenna selection has its historical roots in selection combining which is usually performed at the RX side for single-input-multiple-output (SIMO) channels (see e.g. [1]). This technique has also been generalized to the TX side for multiple-input-single-output (MISO) channels [2]. In the meantime, many researchers have proposed algorithms for more general antenna selection [3]–[6]. Some of these approaches focus on selection to minimize outage probabilities or error rates [7], [8]. We consider selection methods which target high capacity values as in [4], [9]. Previous work in this area includes simple methods [3], [5], [10], [11], variants of which are discussed in this paper, as well as complex methods [12], which may give slightly better performance than the simple methods, but with much greater computational complexity.

In this paper we propose to build on the ideas in [3], [4], [5], analyzing and extending the algorithms that were discussed. We demonstrate in our statistical analysis of the resulting channel capacities that the norm-based selection process can be modelled as a simple scaling of the signal-to-noise-ratio (SNR), resulting in a performance gain. Thus, we propose the SNR scaling factor as a *very simple metric* which can be used to compare the performances of the norm-based algorithms. Note that this approach is extremely general and can be used for a variety of channels, selection scenarios, and also when channel estimation errors are considered. Furthermore, our analysis is valid for all scenarios: transmit selection, receive selection and joint transmit-receive selection. In this paper we use the scaling factor to assess the impact of i.i.d. Rayleigh,



Fig. 1. Antenna Selection for MIMO Systems.

correlated Rayleigh and Ricean fading environments as well as the presence of imperfect channel state information (CSI) and a simplified waterfilling scheme.

This paper is organized as follows. In Sec. II we review the MIMO channel model, the channel capacity for antenna selection systems and a model for imperfect channel state information (CSI). We introduce our various selection algorithms in Sec. III, followed by a statistical analysis of the algorithms in Sec. IV. Then we explore the effects of imperfect channel knowledge in Sec. V. We verify our analysis using Monte Carlo simulations in Sec. VI. We conclude in Sec. VII.

II. MIMO CHANNEL MODEL AND CAPACITY

We focus on a single user MIMO link with n transmitters and m receivers with channel observations of the form

$$\mathbf{r} = \mathbf{H}\mathbf{u} + \mathbf{n} \quad (1)$$

where vectors \mathbf{r} , \mathbf{u} , and \mathbf{n} contain the received observations, the transmitted symbols and the channel noise, respectively. Without loss of generality we assume that $\mathbb{E}\{|n_i|^2\} = 1$. The $m \times n$ channel matrix \mathbf{H} is used to model a flat-fading Gaussian channel with elements h_{ij} having complex Gaussian distributions. We focus on the i.i.d. Rayleigh case but do discuss semi-correlated (SC) Rayleigh and Ricean channels. We assume that this CSI, whether perfect or imperfect, is known at both the transmitter and receiver. The signal-to-noise ratio (SNR) has the form $\mathbb{E}\{\|\mathbf{u}\|^2\}$. If $q = \min(m, n)$, the channel capacity for this system is then

$$C_{\text{full}} = \log_2 \left| \mathbf{I}_q + \frac{\text{SNR}}{n} \mathbf{W} \right| \quad (2)$$

where $|\cdot|$ represents the matrix determinant, \mathbf{I}_q is the $q \times q$ identity matrix, and the $q \times q$ matrix \mathbf{W} is given by

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger, & \text{for } m \leq n \\ \mathbf{H}^\dagger\mathbf{H}, & \text{for } m > n \end{cases} \quad (3)$$

where $(\cdot)^\dagger$ denotes the conjugate transpose. To simplify the notation and without loss of generality, from now on we assume that the number of columns is not less than the number of rows in the channel or selection matrices we define.

If we choose to use only r of the receive antennas and t of the transmit antennas as shown in Fig. 1, we are left with a $r \times t$ submatrix \mathbf{S} of the main channel matrix \mathbf{H} . The channel capacity then has the form

$$C_{\text{sel}} = \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{S}\mathbf{S}^\dagger \right|. \quad (4)$$

We now define the unordered column norms of \mathbf{H} to be $P_c^{(u)} = \sum_{i=1}^m |h_{ic}|^2$, for $c = 1, \dots, n$, and the unordered row norms to be $Q_r^{(u)} = \sum_{j=1}^n |h_{rj}|^2$, for $r = 1, \dots, m$.

Ordering the column norms of \mathbf{H} yields the ordered set $\{P_{1:n} > P_{2:n} > \dots > P_{n:n}\}$, where $P_{j:n}$ denotes the j -th largest column norm. We also define $\mu_{j:n}$ to be the mean of $P_{j:n}$. For ease of notation, we use P_j in place of $P_{j:n}$ and μ_j in place of $\mu_{j:n}$ throughout.

In the case of imperfect CSI the estimated channel, $\widehat{\mathbf{H}}$, is commonly modelled as

$$\mathbf{H} = \rho \widehat{\mathbf{H}} + \sqrt{1 - \rho^2} \mathbf{E} \quad (5)$$

where correlation $\rho = \text{corr}(h_{ij}, \widehat{h}_{ij})$ and \mathbf{E} is an error matrix which is statistically identical to \mathbf{H} . The column and row norms for $\widehat{\mathbf{H}}$ are defined as before with $\widehat{P}_c^{(u)} = \sum_{i=1}^m |\widehat{h}_{ic}|^2$, for $c = 1, \dots, n$, and $\widehat{Q}_r^{(u)} = \sum_{j=1}^n |\widehat{h}_{rj}|^2$, for $r = 1, \dots, m$.

III. SELECTION ALGORITHMS

We now describe a variety of existing and novel transmit and/or receive selection algorithms. All of these algorithms involve selecting a $r \times t$ submatrix \mathbf{S} from the $m \times n$ channel matrix \mathbf{H} . Equations (2) and (4) define the capacity relationships for the full and selection channels, respectively. We focus on the scenarios where selection is performed either at the TX or at both the TX and RX. Analogous RX-only selection methods can easily be obtained from the respective TX selection methods.

A. Algorithms for Transmit Selection

The optimal selection method involves computing the channel capacity for all possible submatrices of \mathbf{H} and picking the \mathbf{S}_{OSA} which yields the highest channel capacity. We can write this Optimal Selection Algorithm (OSA) as

$$\mathbf{S}_{\text{OSA}} = \arg \max_{\{\mathbf{S} \in \mathbf{H}\}} \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{S}\mathbf{S}^\dagger \right| \quad (6)$$

where set $\{\mathbf{S} \in \mathbf{H}\}$ contains all $m \times t$ submatrices of \mathbf{H} . The second method is named the Arbitrary Selection Algorithm (ASA) in which an arbitrary \mathbf{S}_{ASA} is chosen.

Note that the OSA requires an exhaustive search over all possible submatrices of \mathbf{H} , yielding a large computational complexity. As a simple suboptimal alternative, consider forming \mathbf{S}_{NSA} using the columns of \mathbf{H} with the t biggest column norms, P_j , $j = 1 \dots t$. Other researchers [3], [5] have proposed this simple selection method, which we call the Norm-based Selection Algorithm (NSA). We will show, using analysis and simulations, that the capacity for the NSA comes close to that for the OSA.

B. Algorithms for Transmit-Receive Selection

All the TX selection algorithms (OSA, ASA and NSA) can be extended for RX antenna selection. The main difference is that the calculations are now over the rows of \mathbf{H} rather than the columns as in the TX selection case. When jointly implementing both TX and RX selection, the OSA method involves searches over both the rows and columns. The ASA method is equivalent to selecting an arbitrary $r \times t$ submatrix of \mathbf{H} . For the NSA approach, RX selection can be performed independently of TX selection (i.e., the selection of columns does not affect the selection of rows) or jointly with TX

selection. In this joint case the selection of the RX antennas is performed on \mathbf{S} before or after the TX selection. Intuitively, we should first select over the dimension with the greatest choice (i.e., select TX antennas first if $m \leq n$ and RX antennas first if $m > n$). In principle we could also select the rows and columns iteratively. However, the differences observed between such methods are very small. We will show in the analysis and simulation section that joint selection using the NSA slightly outperforms independent selection with the same algorithm.

C. “Poor Man’s Waterfilling”

Note that (2) and (4) assume equal transmit power levels among the TX antennas. For the case of TX selection, if we select t antennas with corresponding norms P_1, P_2, \dots, P_t , consider the novel power allocation method where the power is allocated to each antenna proportional to its column norm. Hence, if an antenna has column norm P_j , it is allocated a proportion $P_j / \sum_{k=1}^t P_k$ of the total power. Hence, the norm of a vector containing the column norms becomes $tP_j^2 / \sum_{k=1}^t P_k$. Due to its low complexity compared to standard waterfilling techniques, we have dubbed this “Poor Man’s Waterfilling” (PMWF). In Sec. IV-F we will analyze PMWF and show that in conjunction with NSA, it outperforms OSA at low SNRs.

IV. ALGORITHM ANALYSIS

In this paper, we do not analyze the OSA method because it is an extremely difficult nonlinear problem. The analytical properties of the ASA method are well known, as the method leads to the standard capacity results [13]. Here we concentrate our analysis primarily on the NSA method. The aim of this analysis is to develop effective methods to calculate the gains made by norm-based antenna selection. Our statistical analysis is based mainly on the fact that capacity, although exactly defined by the joint distributions of the elements of *all* possible submatrices of \mathbf{H} , is strongly affected by the moments of these elements. Finally, note that in Secs. IV-A–IV-D we only consider TX selection, whereas in Sec. IV-E we extend our analysis to joint TX-RX selection.

A. Norm-Based Selection Algorithm

Consider the simplest case of a (2,2) Rayleigh channel matrix with column norms $P_1^{(u)}, P_2^{(u)}$. Assuming the first column $[h_{11}, h_{21}]^T$ has the largest norm and we are selecting one TX antenna, then $\mathbf{S} = [h_{11}, h_{21}]^T$. Note that the elements of \mathbf{S} are no longer Gaussian since their distribution is conditioned on $P_1^{(u)} > P_2^{(u)}$. The exact distribution of \mathbf{S} can be obtained by computing the conditional joint density $f(h_{11}, h_{21} | P_1^{(u)} > P_2^{(u)})$. Clearly, the elements of \mathbf{S} will have a larger variance than those of \mathbf{H} , since they were selected to have the largest column norm.

The derivation of the probability density function (p.d.f.) of the column with the k -th largest norm in the general $m \times n$ case is given in Appendix A. In Appendix A, it is also shown that the columns retain their isotropic nature, so the zero mean complex Gaussian entries in \mathbf{H} are replaced by zero mean complex isotropic entries in \mathbf{S} . Overall, the elements of \mathbf{S}

have increased variances relative to \mathbf{H} but retain the isotropic structure.

The approach taken here is to approximate the entries of \mathbf{S} with zero mean complex Gaussians with the correct variances. There is considerable support for the accuracy of this approach in the literature [14], [15], [16], [17]. Analytical work in this area includes [14], [15], [17] and simulation evidence is given in [16]. In [14], [15] it was found that the asymptotic capacity of MIMO systems is governed solely by the second order statistics of the channel coefficients. Due to the accuracy of the asymptotic approximations it may be expected that the second order statistics will also accurately describe channel capacity [14] for non-Gaussian channels. In a similar vein, expansions derived in [17] employ only the first few moments of the channel coefficients. Lastly, in [16] simulations over a wide range of isotropic non-Gaussian channels showed that for fixed values of the second order statistics, the ergodic capacity is highly insensitive to the exact distribution unless the channel exhibits very severe fading.

Hence, we heuristically model the effect of TX selection as a simple scaling of the original elements of \mathbf{H} . Column j of \mathbf{S} has norm P_j (after rearranging columns). Following the above approach, we model \mathbf{S} as having elements with the same distribution as those of \mathbf{H} , but with variances $E\{P_j\}/m = \mu_j/m$ rather than 1 for column j . This gives us a power scaling due to the selection process. Hence, we approximate \mathbf{S} by

$$\mathbf{S} \approx \mathbf{V} \text{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}, \dots, \sqrt{\mu_t}) / \sqrt{m} \quad (7)$$

where the $m \times t$ matrix \mathbf{V} is statistically identical to an $m \times t$ submatrix of \mathbf{H} picked at random. Defining $\mathbf{M} = \text{diag}(\mu_1, \mu_2, \dots, \mu_t)/m$, equation (7) leads to the approximate capacity

$$C_{\text{sel,pa}} \approx \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{V} \mathbf{M} \mathbf{V}^\dagger \right| \quad (8)$$

where the subscript, “pa”, denotes the fact that the diagonal matrix, \mathbf{M} , can be interpreted as performing power allocation over the antennas. We refer to this technique as the power allocation (PA) approximation.

We can further approximate (8) by replacing the power allocation matrix \mathbf{M} by a single power scaling factor

$$P_{\text{av}} = \frac{1}{m t} \sum_{j=1}^t \mu_j. \quad (9)$$

This leads to the capacity

$$C_{\text{sel,ps}} \approx \log_2 \left| \mathbf{I}_r + P_{\text{av}} \frac{\text{SNR}}{t} \mathbf{V} \mathbf{V}^\dagger \right| \quad (10)$$

and a very simple interpretation of the NSA as providing a power scaling of P_{av} . This technique is referred to as a power scaling (PS) approximation. To implement both the PA and PS approximations, we require $\mu_1, \mu_2, \dots, \mu_t$ and the statistics of (8) and (10) for the various channels. Note that the mean capacity of (10) is known for all the channels considered here [13]–[19]. We now consider these statistics for a variety of channel distributions.

B. Rayleigh Flat-Fading Channels

The Rayleigh channel defined in Sec. II has Gaussian matrix elements h_{ij} , and thus its column norm powers $P_j^{(u)}$ are i.i.d. complex χ^2 distributed with m degrees of freedom¹ and a mean value of $E\{P_j^{(u)}\} = m$. The complex χ^2 p.d.f. and the cumulative distribution function (c.d.f.) are, respectively [20]

$$f_\chi(x; m) = \frac{x^{m-1}}{(m-1)!} \exp(-x) \quad (11)$$

$$F_\chi(x; m) = 1 - \exp(-x) \sum_{k=0}^{m-1} \frac{x^k}{k!}. \quad (12)$$

From (11) and (12) we can calculate the moments of the ordered column norms using simple order statistics. In general, we can write [21]

$$E\{P_j^\ell\} = \frac{n!}{(n-j)!(j-1)!} \int_{-\infty}^{\infty} x^\ell F(x)^{n-j} [1-F(x)]^{j-1} f(x) dx. \quad (13)$$

Combining (13) with (11) and (12), in [22, p. 38] it is shown that

$$E\{P_j^\ell\} = \frac{n!}{(n-j)!(j-1)!(m-1)!} \sum_{r=0}^{n-j} (-1)^r \binom{n-j}{r} \times \sum_{s=0}^{(m-1)(j+r-1)} \frac{c_s(j+r-1)(\ell+m+s-1)!}{(j+r)^{\ell+m+s}} \quad (14)$$

where $c_s(N)$ is the coefficient of x^s in $\left(\sum_{v=0}^{m-1} \frac{x^v}{v!}\right)^N$.

Although we provide exact solutions for computing μ_j and general $E\{P_j^\ell\}$, simpler approximate results might also be useful. The obvious approach is to use the well-known quantile approximations [22], [21] which are of the form $\mu_j \approx F^{-1}\left(\frac{n+1-j-\alpha}{n-\beta}\right)$, where $F(\cdot)$ is the c.d.f. of each column norm. Unfortunately, for these to be accurate, especially when $j=1$, the correction factors α and β need to be chosen differently for each value of n and j . Hence, a single accurate approximation is not available, and we prefer to use the exact results.

C. Semi-Correlated Rayleigh Flat Fading Channels

Consider now a SC Rayleigh channel matrix given by:

$$\mathbf{H}_{sc} = \mathbf{R}^{1/2} \mathbf{H} \quad (15)$$

where \mathbf{R} represents the $m \times m$ spatial correlation matrix at the receiver. The column norms of \mathbf{H}_{sc} are now given by $P_j^{(u)} = \mathbf{h}_j^\dagger \mathbf{R} \mathbf{h}_j$, $j = 1, \dots, n$, where \mathbf{h}_j is the j -th column of \mathbf{H} . This is a well known quadratic form [23] and can be rewritten as $P_j^{(u)} = \sum_{i=1}^m \lambda_i |h_{ij}|^2$ where $\lambda_1, \dots, \lambda_m$ are the eigenvalues of \mathbf{R} . The p.d.f. and c.d.f. of $P_j^{(u)}$ are, respectively [23]

$$f(x) = \sum_{j=1}^m \frac{b_j}{\lambda_j} \exp\left(-\frac{x}{\lambda_j}\right) \quad (16)$$

$$F(x) = \sum_{j=1}^m b_j \left[1 - \exp\left(-\frac{x}{\lambda_j}\right)\right] \quad (17)$$

¹Note that this is slightly different from a real χ^2 distribution with $2m$ degrees of freedom.

where

$$b_j = \lambda_j^{m-1} \prod_{k=1, k \neq j}^m (\lambda_j - \lambda_k)^{-1}.$$

As before, the capacity approximations (8) and (10) can be extended to this case. Note that \mathbf{V} remains semi-correlated. If we define $\mathbf{V} = \mathbf{R}^{1/2} \mathbf{U}$, where \mathbf{U} is an i.i.d. Rayleigh channel matrix, then the capacity approximations can be rewritten as

$$C_{\text{sel,pa}} \approx \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{R}^{1/2} \mathbf{U} \mathbf{M} \mathbf{U}^\dagger \mathbf{R}^{1/2} \right| \quad (18)$$

$$C_{\text{sel,ps}} \approx \log_2 \left| \mathbf{I}_r + P_{\text{av}} \frac{\text{SNR}}{t} \mathbf{R}^{1/2} \mathbf{U} \mathbf{U}^\dagger \mathbf{R}^{1/2} \right|. \quad (19)$$

Again, we can derive μ_1, \dots, μ_t and, hence, \mathbf{M} and P_{av} using order statistics, as (13) and (9) are perfectly applicable to this situation, with the χ^2 distribution replaced by the quadratic form distribution in (17). In Appendix B we obtain the following closed-form expression for μ_j in this semi-correlated case

$$\mu_j = \frac{n!}{(j-1)!(n-j)!} \sum_{t=0}^{n-j} (-1)^t \binom{n-j}{t} \sum_{K(t+j-1)} c_{K(t+j-1)} \times \sum_{s=1}^m \frac{b_s}{\lambda_s} \left(d_{K(t+j-1)} + \frac{1}{\lambda_s}\right)^{-2} \quad (20)$$

where

$$\left(\sum_{r=1}^m b_r e^{-x/\lambda_r}\right)^L = \sum_{K(L)} c_{K(L)} \exp\{-d_{K(L)} x\}$$

with

$$c_{K(L)} = L! \prod_{i=1}^m \frac{b_i^{k_i}}{k_i!} \quad d_{K(L)} = \sum_{i=1}^m \frac{k_i}{\lambda_i}$$

and the sum is over all partitions of L (i.e., $K(L) = (k_1, \dots, k_m)$ where $\sum_{i=1}^m k_i = m$).

D. Ricean Channels

An i.i.d. Ricean channel can be expressed as the sum of a line-of-sight (LOS) channel and a Rayleigh flat-fading channel as follows

$$\mathbf{H}_{\text{rice}} = a \mathbf{H}_{\text{LOS}} + b \mathbf{H} \quad (21)$$

where $a^2 + b^2 = 1$ and the entries of \mathbf{H}_{LOS} and \mathbf{H} have unit variance. If we assume the common model of a unit rank matrix \mathbf{H}_{LOS} , then we can assume without loss of generality that the elements of \mathbf{H}_{LOS} are all unity. Hence, the j -th column norm of \mathbf{H}_{rice} is given by $P_j^{(u)} = [a \mathbf{1} + b \mathbf{h}_j]^\dagger [a \mathbf{1} + b \mathbf{h}_j]$, where $\mathbf{1} = [1, \dots, 1]^T$ and \mathbf{h}_j is again the j -th column of the i.i.d. Rayleigh matrix \mathbf{H} . This is a non-central quadratic form, and its p.d.f. and c.d.f. are respectively given by [23]

$$f(x) = \frac{1}{b^2} \left(\frac{2x}{b^2 \delta}\right)^{(m-1)/2} I_{m-1} \left(\frac{2x \sqrt{\delta}}{b^2}\right) \times \exp\left[-\left(\frac{2x}{b^2} + \delta\right) / 2\right] \quad (22)$$

$$F(x) = \exp\left(-\frac{\delta}{2}\right) \sum_{j=0}^{\infty} \frac{(\delta/2)^j}{j! (m+j-1)! 2^{m+j}} \times \int_0^{2x/b^2} u^{m+j-1} \exp\left(-\frac{u}{2}\right) du \quad (23)$$

where $\delta = 2m a^2/b^2$ and $I_k(\cdot)$ is a modified Bessel function. Note that the integral in the c.d.f. can be given as a finite sum, but there is no closed-form solution for the c.d.f. which avoids either numerical integration or an infinite series. Again, (13) can be used to obtain $\mu_1, \mu_2, \dots, \mu_t$ and the corresponding capacity approximations (7) and (10) can be computed.

E. Transmit-Receive Antenna Selection

As mentioned in Sec. III-B, the TX selection algorithms can be extended for RX antenna selection. We denote the TX and RX power scaling factors by $P_{av,t}$ and $P_{av,r}$, respectively. If we perform joint TX and RX selection, we can leverage the results from Secs. IV-A and IV-D to approximate the capacity. Assuming that TX selection occurs first, we have the selected matrix $\mathbf{S}_t \approx \mathbf{V}\mathbf{M}^{1/2}$ as in (7). For the i.i.d. Rayleigh case, the row selection process is then exactly the same as TX selection for the SC Rayleigh channel, since the form of the rows of \mathbf{S}_t match the columns of \mathbf{H}_{sc} in (15). Hence, we can evaluate both (8) and (10) for the i.i.d. Rayleigh case.

For the SC Rayleigh and i.i.d. Rician channels similar methods also lead to capacity approximations of the form given in (8). For space reasons these derivations are omitted here. However, the SNR scaling approximation (10) is simple to derive for all of these channels. After TX selection we have $\mathbf{S}_t \approx P_{av,t}^{1/2} \mathbf{V}$. Similarly, after RX selection we have $\mathbf{S} \approx P_{av,r}^{1/2} \mathbf{S}_t = P_{av,r}^{1/2} P_{av,t}^{1/2} \mathbf{V}$. Hence, we can use (10) with the new power scaling $P_{av} = P_{av,r} P_{av,t}$.

F. Poor Man's Waterfilling (for TX Selection)

After the initial antenna selection, other techniques can be used to further enhance the capacity of the system. These include techniques like conventional waterfilling, where the transmit power is allocated along the eigenmodes of the channel. We propose a new simpler algorithm to achieve suboptimal gains, but at a highly reduced complexity level.

"Poor Man's Waterfilling" (PMWF) employs the values of the column norms P_j to allocate power to the system. We focus on a power scaling approximation to the effect of PMWF. After TX selection has been performed, column j in \mathbf{S} has norm P_j . Using PMWF this norm is scaled by $tP_j/\sum_{k=1}^t P_k$, and the resulting matrix is denoted \mathbf{S}_{pmwf} . The total norm of \mathbf{S}_{pmwf} is $\|\mathbf{S}_{pmwf}\|^2 = \sum_{j=1}^t tP_j^2/\sum_{k=1}^t P_k$. The average modulus squared value of any element in \mathbf{S}_{pmwf} is therefore $\|\mathbf{S}_{pmwf}\|^2/(mt)$, and the mean value of this gives the power scaling factor:

$$P_{av,pmwf} = \frac{1}{m} \mathbb{E} \left\{ \frac{\sum_{j=1}^t P_j^2}{\sum_{k=1}^t P_k} \right\}. \quad (24)$$

For TX-RX selection using PMWF we have $P_{av} = P_{av,r} P_{av,pmwf}$. Unfortunately, the exact evaluation of (24) is very difficult, as the p.d.f. of a sum of χ^2 order statistics is extremely complex (see e.g., [24]). But, we can use the so-called "delta method" to good effect [25]. Using a Taylor

series expansion, we let

$$\begin{aligned} \mathbb{E} \left(\frac{X}{Y} \right) &\triangleq \mathbb{E}[g(X, Y)] \\ &= \mathbb{E} \left[\sum_{m,n} \frac{g^{(m,n)}(\mu_X, \mu_Y)}{m!n!} (X - \mu_X)^m (Y - \mu_Y)^n \right] \end{aligned} \quad (25)$$

where

$$g^{(m,n)}(\mu_X, \mu_Y) = \left. \frac{\partial^{m+n}}{\partial x^m \partial y^n} g(x, y) \right|_{x=\mu_X, y=\mu_Y}$$

and we approximate $\mathbb{E}[g(X, Y)]$ by taking only the first few terms. If we take terms up to order two ($m+n \leq 2$), we get

$$\begin{aligned} \mathbb{E} \left(\frac{X}{Y} \right) &\approx \frac{\mu_X}{\mu_Y} + \frac{\mu_X}{\mu_Y^3} \text{var}(Y) - \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\mu_Y^2} \\ &= \frac{\mu_X \mathbb{E}(Y^2)}{\mu_Y^3} - \frac{\text{cov}(X, Y)}{\mu_Y^2}. \end{aligned} \quad (26)$$

Comparing (24) with (26), we see that we need to compute $\mu_X = \sum_{j=1}^t \mathbb{E}(P_j^2)$, $\mu_Y = \sum_{j=1}^t \mathbb{E}(P_j)$, $\mathbb{E}(Y^2) = \sum_{j=1}^t \sum_{k=1}^t \mathbb{E}(P_j P_k)$, and $\text{cov}(X, Y) = \sum_{j=1}^t \sum_{k=1}^t \mathbb{E}(P_j^2 P_k) - \mu_X \mu_Y$. But, from (14), we already know components $\mathbb{E}(P_j)$, $\mathbb{E}(P_j^2)$, and $\mathbb{E}(P_j^3)$ for the i.i.d. Rayleigh fading case. In Appendix C, for this same fading case we compute $\mathbb{E}\{P_j^s P_k^t\}$, for $j \neq k$, yielding the required values for components $\mathbb{E}\{P_j P_k\}$ and $\mathbb{E}\{P_j^2 P_k\}$. Due to the tedious analysis required for other channels, in this paper we limit our analysis to the i.i.d. Rayleigh channel for this particular power allocation method.

V. EFFECTS OF IMPERFECT CHANNEL STATE INFORMATION

In this section we assess the impact of imperfect CSI on TX selection via a simple power scaling approach. Taking column j of (5) gives

$$\mathbf{h}_j = \rho \hat{\mathbf{h}}_j + \sqrt{1 - \rho^2} \mathbf{e}_j \quad (27)$$

with \mathbf{h}_j , $\hat{\mathbf{h}}_j$, \mathbf{e}_j being the j -th columns of \mathbf{H} , $\hat{\mathbf{H}}$ and \mathbf{E} respectively. Taking the norms of the columns in (27) gives

$$P_j^{(u)} = \rho^2 \hat{P}_j^{(u)} + (1 - \rho^2) \|\mathbf{e}_j\|^2 + \Delta \quad (28)$$

where Δ is the cross product term. Since $\mathbb{E}\{\hat{P}_j^{(u)}\} = \mathbb{E}\{\|\mathbf{e}_j\|^2\} = m$, we can rewrite (28) as

$$P_j^{(u)} = \rho^2 (\hat{P}_j^{(u)} - m) + m + \epsilon_j \quad (29)$$

where $\epsilon_j = \Delta + (1 - \rho^2)(\|\mathbf{e}_j\|^2 - m)$. In the process of TX selection we are selecting columns of \mathbf{H} on the basis of measurements of $\hat{\mathbf{H}}$ which leads to the area of order statistics called concomitants [21]. Note that (29) is a simple linear model, and in the situation where ϵ_j is independent of $\hat{P}_j^{(u)}$ we have [21]

$$\mu_{[j]} = \rho^2 (\mu_j - m) + m \quad (30)$$

where μ_j is as before but $\mu_{[j]}$ represents the mean norm of the column of \mathbf{H} selected on the basis of $\hat{\mathbf{H}}$. In our scenario, ϵ_j is uncorrelated with $\hat{P}_j^{(u)}$ but not independent. Hence we apply (30) as an approximate result. Since the column norms

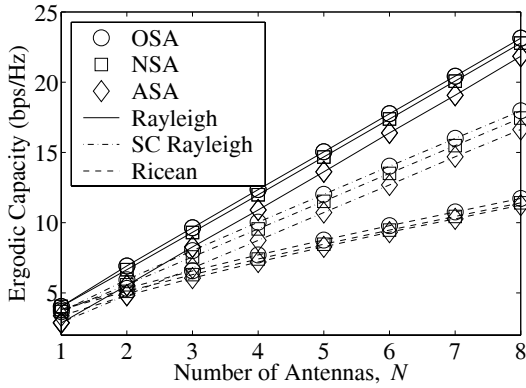


Fig. 2. Comparison of selection schemes for a $(N, N+2)$ choose (N, N) system (SNR = 10dB).

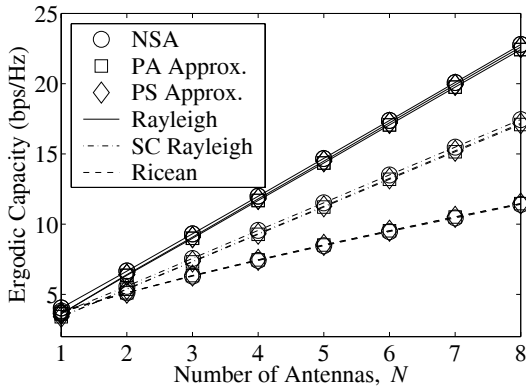


Fig. 3. Approximations for NSA for a $(N, N+2)$ choose (N, N) system (SNR = 10dB).

now have means $\mu_{[j]}$ we can adjust $P_{av,t}$ to give $P_{av,CSI} \triangleq \sum_{j=1}^t \mu_{[j]} / (mt)$ which can be rewritten as

$$P_{av,CSI} = 1 - \rho^2 + \rho^2 P_{av,t}. \quad (31)$$

Again the effect of another system factor, in this case CSI, is simply accommodated in the power scaling factor.

VI. COMPUTER SIMULATIONS

Figure 2 shows a comparison of the key selection algorithms OSA, NSA, ASA for three different channels: i.i.d. Rayleigh, SC Rayleigh and Ricean. All curves are for an SNR of 10dB and show capacity vs. N in a $(N, N+2)$ system where an (N, N) subsystem is to be selected. Parameter values are selected for ease of comparison rather than physical reasons. The SC Rayleigh channel has an exponential correlation structure [26] at the receiver with $r = 0.8$ giving the correlation between adjacent antennas. The Ricean channel has K-factor $K = 10$ dB and is defined in [19]. Figure 2 shows that NSA offers a reasonable proportion of the benefits of OSA at a much reduced complexity. This holds true for all of the channel models.

Figure 3 shows the PA and PS approximations to NSA for the same three channels and SNR value as in Fig. 2. Note that both approximations are very accurate and PS performs at least as well as PA for all cases. Hence, since PS is simpler we concentrate on this approximation in the following simulations.

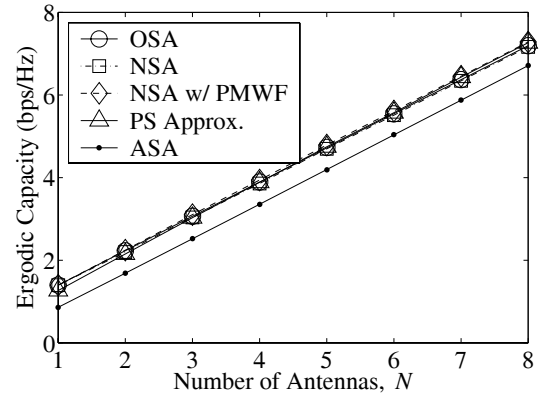


Fig. 4. PMWF comparison for a $(N, N+2)$ choose (N, N) system for an i.i.d. Rayleigh channel (SNR = 0dB).

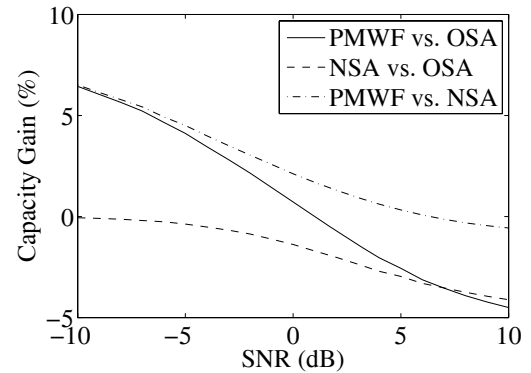


Fig. 5. Percentage capacity improvement of PMWF, NSA and OSA relative to each other in a $(4,2)$ choose $(2,2)$ system for an i.i.d. Rayleigh channel.

The results in Figs. 2 and 3 show the performances at an SNR of 10dB. In Fig. 4 we look at the low SNR case (SNR = 0dB) and show that NSA with PMWF is superior to OSA in this scenario. In addition, the PS approximation is excellent here. Note that NSA with PMWF outperforms OSA for low SNR, but for higher SNRs OSA becomes superior again, although the difference is small. This comparison is shown in more detail in Fig. 5 for an i.i.d. Rayleigh channel. We observe that as the SNR varies from -10dB to 10dB the percentage improvement of OSA over NSA is at most 4%. Also, NSA with PMWF offers improvements over OSA of up to about 6.5% at -10dB. These gains reduce with increasing SNR and at 10dB OSA has a slight superiority. Overall, the performance of the simple norm-based approaches compare very favorably with OSA.

The NSA and PMWF by design are low complexity algorithms involving the computation of column norms and ordering only. This gives a complexity order of $m \times n$. OSA, on the other hand, is highly intensive requiring $\binom{m}{r} \times \binom{n}{t}$ matrix determinants, in addition to ordering and other matrix operations. Since the determinant is order r^3 , OSA has a complexity order of $\binom{m}{r} \times \binom{n}{t} \times r^3$. Hence, the comparable performance of NSA with PMWF at low SNR, but with much reduced complexity, is quite notable.

In (31) we showed a quadratic relationship between the PS factor and ρ , which is a measure of the CSI quality. This result is verified in Fig. 6 which shows the PS approximation

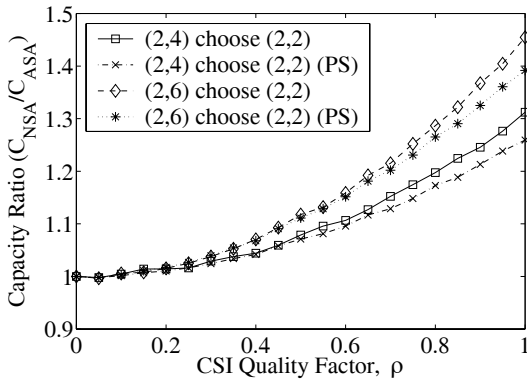


Fig. 6. Effect of imperfect channel state information on antenna selection for an i.i.d. Rayleigh channel (SNR = 0dB).

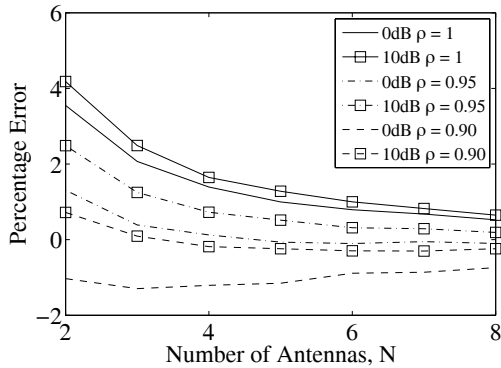


Fig. 7. Percentage error of the PS approach at various SNR values with imperfect CSI in a $(N, N + 2)$ choose (N, N) system for an i.i.d. Rayleigh channel

lower bounding the simulated values and a quadratic shape. Results here are for two scenarios, and the plot is of the relative capacity increase due to NSA to show both cases on the same scale. Note the increased benefits due to selection as the number of redundant antennas increases from 2 to 4. Again the low SNR region is considered.

In Fig. 7 we take a closer look at the errors of the PS approach for both perfect and imperfect CSI cases in an i.i.d. Rayleigh channel over a range of SNR values and system sizes. In all cases we observe that the percentage errors of the PS approach decay with system size and are usually less than 4%. The value of the SNR does not have a marked effect but the % errors at 10dB are greater than those at 0dB. It is interesting that increasing the channel estimation error or reducing ρ does not reduce the accuracy of the PS approach. This is because the small errors in the PS approach for handling NSA and imperfect CSI tend to cancel each other out.

In Fig. 8 we look at the behavior of joint TX-RX selection at an SNR of 10dB. As shown, the benefit of joint selection at the TX and RX rather than independent selection at both ends is very slight. Again, the PS approximation is excellent and the improvement over TX selection is only marginal.

In Sec. IV-C the semi-correlated case was analyzed where selection occurs across independent columns, but each column contains correlated entries. The opposite case is also of interest but is difficult to handle analytically since order statistics

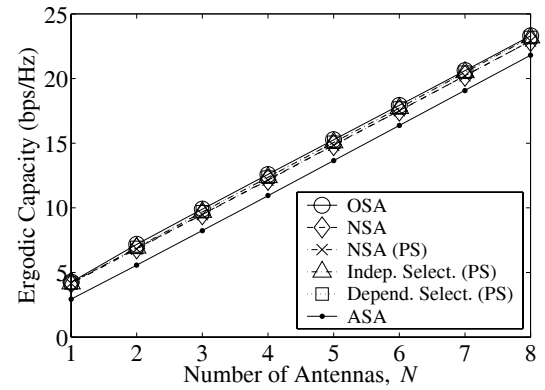


Fig. 8. TX-RX selection for a $(N + 1, N + 1)$ choose (N, N) system for an i.i.d. Rayleigh channel (SNR = 10dB).

of correlated samples have complicated distributions. Hence, we handle this case by simulation. Since the OSA approach selects columns based on both the size and the relationship between the columns, it might be expected that OSA would increase its superiority over NSA in this correlated case since there is an intrinsic dependence between the columns that OSA could exploit. On the other hand, in the limiting case of perfect correlation between the columns, each column is identical, and so it does not matter which columns are chosen and ASA, NSA and OSA all become equivalent. We investigate this issue in the simulations.

The correlated scenario is also important in another sense. Consider the gains in capacity offered by selecting a few antennas from a densely packed array of a large number of antennas occupying a fixed space. In this situation, the columns become correlated due to the small antenna spacing. As the array size increases, there are two opposing factors: the selection gain due to greater choice versus the increased correlation between the columns.

To explore this situation we consider an equi-spaced linear array of fixed length. The decorrelation distance is assumed to be half a wavelength so that the correlation between channel coefficients separated by half a wavelength is $\rho(0.5) = 0.5$. We assume the well-known exponential model for correlation, where $\rho(d) = \exp(-ad)$ and a is chosen to match the decorrelation distance. With this correlation model we are able to study both the relative performance of NSA and OSA and also the competition between reduced spacing and increased antenna numbers.

Figure 9 shows the mean capacities achieved by ASA, NSA and OSA in a $(2,4)$ choose $(2,2)$ system and a $(4,8)$ choose $(4,4)$ system, where the TX selects half of the available antennas. The x-axis is the total antenna array span in wavelengths, and as predicted, when the intra-antenna spacing is small all three methods converge. For systems of this size NSA works satisfactorily, giving more than 70% of the gains of OSA over ASA. Figure 10 shows a more complex picture. In Fig. 10 we consider $(2, N)$ select $(2, 2)$ and $(4, N)$ select $(4, 4)$ scenarios. The x-axis is N , where the N antennas are constrained to a total array length of one wavelength. Figure 10 shows the increasing superiority of OSA over NSA and the potential decrease in performance of NSA as N

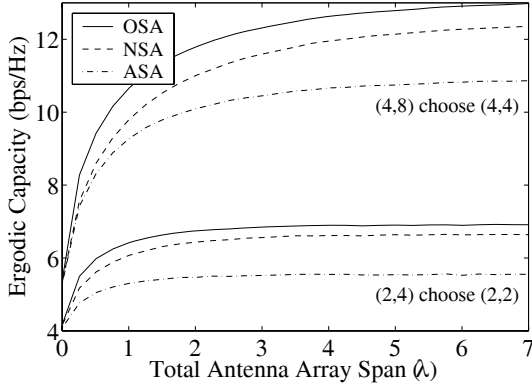


Fig. 9. Effect of intra-antenna spacing for semi-correlated Rayleigh channel (SNR = 10dB).

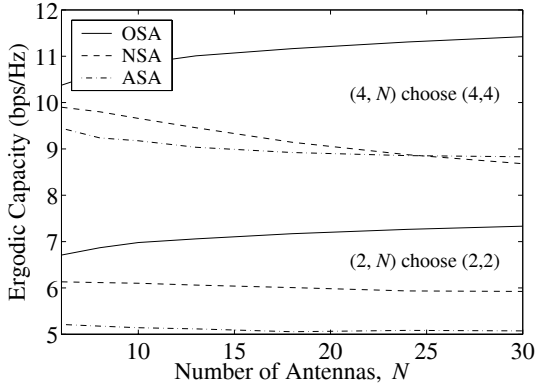


Fig. 10. Effect of number of antennas for semi-correlated Rayleigh channel (SNR = 10dB).

increases. The first effect is caused by the ability of OSA to use the correlation structure of the channel. The second effect is due to the fact that in heavily correlated scenarios NSA tends to pick consecutive antennas since the antenna with the highest norm is often surrounded by antennas with similar large norms. However, due to the high correlation, the channel gains are similar, leading to reduced capacity. Hence, in densely packed arrays or highly correlated scenarios NSA needs to be altered to account for correlation patterns in addition to the simple norms.

VII. CONCLUSION

NSA is seen to be an important selection method, since its complexity is far less than OSA, its performance is close to OSA and, with the addition of PMWF, it can outperform OSA at low SNR levels. Hence, an analysis method for NSA is important, and the power scaling method we have developed is able to approximate the capacity for NSA over a variety of channels and system scenarios. The beauty of the PS approach is its generality and its ability to gauge the effect of selection in a single number. In particular, analytical results clearly show the effects of PMWF and imperfect CSI in simple closed-form expressions.

APPENDIX A

ISOTROPIC ANALYSIS OF LARGEST COLUMN DENSITY

Consider the $m \times n$ matrix \mathbf{H} in the i.i.d. Rayleigh case. Define an arbitrary column of \mathbf{H} by $\mathbf{X} = [X_1, X_2, \dots, X_m]^T$ and its norm by $R = \sum_{i=1}^m |X_i|^2$. In NSA this column is chosen for inclusion in \mathbf{S} if it has one of the t largest column norms. Hence, the distribution of \mathbf{S} is affected by the selection of the “largest” columns and is no longer i.i.d. complex Gaussian. The distribution of the column with the k -th largest norm is derived as follows

$$f_{\mathbf{X}}(\mathbf{x}|k) = \int_0^\infty f_{\mathbf{X}}(\mathbf{x}|R=r) f_{P_k}(r) dr \quad (32)$$

where conditioning on k implies the k -th largest norm, $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$, and P_k is the k -th largest norm as before. Using Bayes rule, the conditional density in (32) can be simplified as

$$f_{\mathbf{X}}(\mathbf{x}|R=r) = \frac{f_R(r|\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{f_R(r)}. \quad (33)$$

Now, conditioned on \mathbf{x} , R has a degenerate distribution and $f_{\mathbf{X}}(\mathbf{x})$ is simply the joint density of m independent complex Gaussians. Hence, we have

$$f_{\mathbf{X}}(\mathbf{x}|R=r) = \frac{\delta(r - \sum_{i=1}^m |x_i|^2) \exp(-\sum_{i=1}^m |x_i|^2)}{\pi^m f_R(r)}. \quad (34)$$

Substituting (34) into (32) gives

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}|k) &= \int_0^\infty \frac{\delta(r - \sum_{i=1}^m |x_i|^2) \exp(-\sum_{i=1}^m |x_i|^2)}{\pi^m f_R(r)} \\ &\quad \times \frac{n!}{(n-k)!(k-1)!} F_R(r)^{k-1} [1 - F_R(r)]^{n-k} f_R(r) dr \\ &= \frac{n!}{\pi^m (n-k)!(k-1)!} \exp\left(-\sum_{i=1}^m |x_i|^2\right) \\ &\quad \times F_R\left(\sum_{i=1}^m |x_i|^2\right)^{k-1} \left[1 - F_R\left(\sum_{i=1}^m |x_i|^2\right)\right]^{n-k} \end{aligned} \quad (35)$$

and so the joint density of the k -th largest column remains isotropic as it is solely a function of $\sum_{i=1}^m |x_i|^2$.

APPENDIX B

NORM MEAN CALCULATION FOR SEMI-CORRELATED RAYLEIGH CHANNEL

To calculate $\mu_j = E\{P_j\}$ for the semi-correlated Rayleigh channel, note that the density of P_j is given by [22]

$$f_{P_j}(x) = \frac{n!}{(j-1)!(n-j)!} F(x)^{n-j} [1 - F(x)]^{j-1} f(x).$$

Substituting for $F(x)$ and $f(x)$ in the semi-correlated case gives

$$\begin{aligned} \mu_j &= \frac{n!}{(j-1)!(n-j)!} \int_0^\infty x \left(1 - \sum_{t=1}^m b_t e^{-x/\lambda_t}\right)^{n-j} \\ &\quad \times \left(\sum_{r=1}^m b_r e^{-x/\lambda_r}\right)^{j-1} \sum_{s=1}^m \frac{b_s}{\lambda_s} e^{-x/\lambda_s} dx. \end{aligned} \quad (36)$$

Expanding the binomial term produces

$$\begin{aligned}
\mu_j &= \frac{n!}{(j-1)!(n-j)!} \int_0^\infty x \sum_{t=0}^{n-j} (-1)^t \binom{n-j}{t} \\
&\quad \times \left(\sum_{r=1}^m b_r e^{-x/\lambda_r} \right)^{t+j-1} \sum_{s=1}^m \frac{b_s}{\lambda_s} e^{-x/\lambda_s} dx \\
&= \frac{n!}{(j-1)!(n-j)!} \sum_{t=0}^{n-j} (-1)^t \binom{n-j}{t} \sum_{K(t+j-1)} c_{K(t+j-1)} \\
&\quad \times \sum_{s=1}^m \frac{b_s}{\lambda_s} \int_0^\infty x e^{-(d_{K(t+j-1)}+1/\lambda_s)x} dx \\
&= \frac{n!}{(j-1)!(n-j)!} \sum_{t=0}^{n-j} (-1)^t \binom{n-j}{t} \sum_{K(t+j-1)} c_{K(t+j-1)} \\
&\quad \times \sum_{s=1}^m \frac{b_s}{\lambda_s} \left(d_{K(t+j-1)} + \frac{1}{\lambda_s} \right)^{-2} \quad (37)
\end{aligned}$$

where

$$\left(\sum_{r=1}^m b_r e^{-x/\lambda_r} \right)^L = \sum_{K(L)} c_{K(L)} \exp \{ -d_{K(L)} x \}$$

with

$$c_{K(L)} = L! \prod_{i=1}^m \frac{b_i^{k_i}}{k_i!} \quad d_{K(L)} = \sum_{i=1}^m \frac{k_i}{\lambda_i}$$

and the sum is over all partitions of L (i.e., $K(L) = (k_1, \dots, k_m)$ where $\sum_{i=1}^m k_i = m$).

APPENDIX C

HIGHER-ORDER MOMENT CALCULATION FOR I.I.D. RAYLEIGH CHANNEL

To analyze PMWF, we require $E\{P_i^s P_j^t\}$. Without loss of generality, we consider $i > j$, so $0 < X \triangleq P_i < Y \triangleq P_j < \infty$. The joint density of (P_i, P_j) or (X, Y) is [22]

$$\begin{aligned}
f_{X,Y}(x,y) &= \frac{n!}{(n-i)!(i-j-1)!(j-1)!} F(x)^{n-i} \\
&\quad \times [F(y) - F(x)]^{i-j-1} [1 - F(y)]^{j-1} f(x) f(y).
\end{aligned}$$

Substituting for $f(x)$ and $F(x)$ using (11) and (12) gives

$$\begin{aligned}
f_{X,Y}(x,y) &= \Delta \left(1 - e^{-x} \sum_{k=0}^{m-1} \frac{x^k}{k!} \right)^{n-i} \\
&\quad \times \left(e^{-x} \sum_{h=0}^{m-1} \frac{x^h}{h!} - e^{-y} \sum_{\ell=0}^{m-1} \frac{y^\ell}{\ell!} \right)^{i-j-1} \\
&\quad \times \left(e^{-y} \sum_{u=0}^{m-1} \frac{y^u}{u!} \right)^{j-1} \frac{x^{m-1} e^{-x}}{(m-1)!} \frac{y^{m-1} e^{-y}}{(m-1)!} \quad (38)
\end{aligned}$$

where $\Delta = n! / [(n-i)!(i-j-1)!(j-1)!]$. Expanding the binomial terms, the first two bracketed terms in (38), yields

$$\begin{aligned}
f_{X,Y}(x,y) &= \frac{\Delta}{[(m-1)!]^2} \sum_{k=0}^{n-i} \sum_{\ell=0}^{i-j-1} (-1)^{k+\ell} \binom{n-i}{k} \binom{i-j-1}{\ell} \\
&\quad \times x^{m-1} y^{m-1} \left(\sum_{h=0}^{m-1} \frac{x^h}{h!} \right)^{k+i-j-\ell-1} \left(\sum_{u=0}^{m-1} \frac{y^u}{u!} \right)^{\ell+j-1} \\
&\quad \times e^{-(k+i-j-\ell)x} e^{-(\ell+j)y}. \quad (39)
\end{aligned}$$

Expanding the two power series in (39), we then have

$$\begin{aligned}
f_{X,Y}(x,y) &= \frac{\Delta}{[(m-1)!]^2} \sum_{k=0}^{n-i} \sum_{\ell=0}^{i-j-1} (-1)^{k+\ell} \binom{n-i}{k} \binom{i-j-1}{\ell} \\
&\quad \times \sum_{h=0}^{N(i,j,k,\ell,m)} c_h(N(i,j,k,\ell,m)) \sum_{u=0}^{N(j,\ell,m)} c_u(N(j,\ell,m)) \\
&\quad \times x^{m+h-1} y^{m+u-1} e^{-(k+i-j-\ell)x} e^{-(\ell+j)y} \quad (40)
\end{aligned}$$

where $N(i,j,k,\ell,m) = (m-1)(k+i-j-\ell-1)$, $N(j,\ell,m) = (m-1)(\ell+j-1)$ and $c_s(N)$ is defined after (14). The calculation of $E\{X^s Y^t\}$ is now straightforward using the result

$$\begin{aligned}
&\int_0^\infty \int_x^\infty x^a y^b e^{-cx} e^{-gy} dy dx \\
&= g^{-(b+1)} \sum_{v=0}^b \frac{b!}{(b-v)!} \frac{(a+b-v)!}{(c+1)^{a+b-v+1}}
\end{aligned}$$

where $a, b \in \mathbb{Z}^+$. We can finally write the joint moment as

$$\begin{aligned}
E\{P_i^s P_j^t\} &= \frac{\Delta}{[(m-1)!]^2} \sum_{k=0}^{n-i} \sum_{\ell=0}^{i-j-1} (-1)^{k+\ell} \binom{n-i}{k} \binom{i-j-1}{\ell} \\
&\quad \times \sum_{h=0}^{N(i,j,k,\ell,m)} c_h(N(i,j,k,\ell,m)) \sum_{u=0}^{N(j,\ell,m)} c_u(N(j,\ell,m)) \\
&\quad \times \sum_{v=0}^{m+r+t-1} \frac{\Psi(m,u,t,h,s,v)}{\Upsilon(m,u,t,v,k,i,j,\ell,h,s)} \quad (41)
\end{aligned}$$

where

$$\begin{aligned}
\Psi(m,u,t,h,s,v) &= (m+u+t-1)! (2m+u+h+s+t-2-v)! \\
\Upsilon(m,u,t,v,k,i,j,\ell,h,s) &= (m+u+t-1-v)! (k+i-j-\ell+1)^{2m+u+h+s+t} \\
&\quad \times (\ell+j)^{m+u+t}.
\end{aligned}$$

REFERENCES

- [1] J. Cheng and T. Berger, "Capacity and performance analysis for hybrid selection/maximal-ratio combining in Nakagami fading with unequal fading parameters and branch powers," in *Proc. IEEE Int'l. Conf. on Communications*, Anchorage, USA, May 2003, pp. 3031-3035.
- [2] M. Z. Win and J. H. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1926-1934, Nov. 2001.
- [3] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68-73, Oct. 2004.
- [4] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 322-324, Aug. 2002.
- [5] Z. Zhou, Y. Dong, X. Zhang, W. Wang, and Y. Zhang, "A novel antenna selection scheme in MIMO systems," in *Proc. Int'l. Conf. on Communications, Circuits and Systems*, Chengdu, China, June 2004, pp. 190-194.
- [6] A. Ghrayeb, "A survey on antenna selection for MIMO communication systems," in *Proc. 2nd IEEE Int'l. Conf. on Information and Communication Technologies*, vol. 2, Damascus, Syria, Apr. 2006, pp. 2104-2109.
- [7] R. Heath and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int'l. Conf. on Communications*, Helsinki, Finland, June 2001, pp. 2276-2280.
- [8] Z. Chen, B. Vucetic, and J. Yuan, "Space-time trellis codes with transmit antenna selection," *IEE Electron. Lett.*, vol. 39, pp. 854-855, May 2003.
- [9] A. Molisch, M. Win, and J. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int'l. Conf. on Communications*, Helsinki, Finland, June 2001, pp. 570-574.

- [10] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 339–347, Feb. 2004.
- [11] L. Dai, S. Sfar, and K. Lettaief, "Optimal antenna selection based on capacity maximization for MIMO systems in correlated channels," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 563–573, Mar. 2006.
- [12] A. Molisch and X. Zhang, "FFT-based hybrid antenna selection schemes for spatially correlated MIMO channels," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 36–38, Jan. 2004.
- [13] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun. Related Technol.*, vol. 10, pp. 585–595, Nov.-Dec. 1999.
- [14] V. Veeravalli, Y. Liang, and A. Sayeed, "Correlated MIMO wireless channels: capacity, optimal signaling and asymptotics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2058–2072, June 2005.
- [15] J. Perez, J. Ibanez, L. Vielva, and I. Santamaria, "Approximate closed-form expression for the ergodic capacity of MISO and SIMO systems," in *Proc. IEEE Vehicular Technology Conf.*, Stockholm, Sweden, May 2005, pp. 910–913.
- [16] S. H. Choi, P. Smith, B. Allen, W. Q. Malik, and M. Shafi, "Severely fading MIMO channels: Models and mutual information," in *Proc. IEEE Int'l. Conf. on Communications*, Glasgow, Scotland, June 2007, pp. 4628–4633.
- [17] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, July 2004.
- [18] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.
- [19] M. Kang and M.-S. Alouini, "On the capacity of MIMO Rician channels," in *Proc. 40th Allerton Conf. on Communications, Control, and Computing*, Monticello, IL, USA, Oct. 2002, pp. 936–945.
- [20] N. L. Johnson and S. Kotz, Eds., *Continuous Univariate Distributions-1*. Hoboken, NJ, USA: John Wiley and Sons Inc., 1970.
- [21] H. David and H. Nagaraja, eds., *Order Statistics*. Hoboken, NJ, USA: John Wiley and Sons Inc., 2003.
- [22] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, Eds., *A First Course in Order Statistics*. Hoboken, NJ, USA: John Wiley and Sons Inc., 1992.
- [23] N. L. Johnson and S. Kotz, Eds., *Continuous Univariate Distributions-2*. Hoboken, NJ, USA: John Wiley and Sons Inc., 1970.
- [24] K. Alam and K. T. Wallenius, "Distribution of a sum of order statistics," *Scand. J. Stat., Theory Appl.*, vol. 6, pp. 123–126, 1979.
- [25] J. Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. Belmont, CA, USA: Duxbury Press, 1994.
- [26] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sept. 2001.



Peter J. Smith received the B.Sc. degree in Mathematics and the Ph.D. degree in Statistics from the University of London, London, U.K., in 1983 and 1988, respectively. From 1983 to 1986, he was with the Telecommunications Laboratories at GEC Hirst Research Centre. From 1988 to 2001, he was a lecturer in statistics and consulting statistician at Victoria University of Wellington, Wellington, New Zealand. He is currently an associate professor in the Department of Electrical and Computer Engineering at the University of Canterbury, Christchurch, New Zealand. His research interests include the statistical aspects of design and analysis for communication systems, especially antenna arrays, MIMO systems and cognitive radio.



Timothy W. King was born in Wellington, New Zealand, in 1981. He received the B.E. (with first class honors) in Electrical and Electronic Engineering from the University of Canterbury, Christchurch, New Zealand, in 2004. He is currently working towards his Ph.D. in Electrical and Computer Engineering at the same university. His research interests lie in the areas of wireless communications and information theory. He also has industrial experience in cryptography and DSL systems. Specific research topics he has worked on include the capacity of MIMO broadcast channels, cooperative systems and antenna selection techniques in MIMO systems.



Lee M. Garth received the B.S.E. degree (magna cum laude) from Princeton University, Princeton, NJ, in 1987, and the M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, in 1989 and 1996, respectively. He has had summer employment with Raytheon Company, GTE Corporation, and MITRE Corporation. From 1990 to 1996, he was a Senior Engineer at Techno-Sciences, Inc., Urbana, IL. From 1996 to 2000, he was a member of the Advanced Data Communications Group of Bell Laboratories within Lucent Technologies, Holmdel, NJ. From 2000 to February 2008, he was a faculty member in the Department of Electrical and Electronic Engineering at the University of Canterbury, Christchurch, New Zealand. In 2006 he held a visiting appointment at the Samsung Advanced Institute of Technology in South Korea. Since March 2008 he has been a Senior Principal Engineer in the Advanced Systems and Technologies Division of BAE Systems, Merrimack, New Hampshire. His research interests include signal detection, array processing, adaptive equalization, and statistical signal processing with applications to communications systems. Dr. Garth is a Senior Member of the IEEE and a Member of Tau Beta Pi.



Mischa Dohler is now a Senior Researcher with CTTC in Barcelona. Prior to this, he was a Senior Research Expert in the R&D department of France Telecom working on cooperative communication systems and wireless sensor networks from June 2005 to February 2008. From September 2003 to June 2005, he was a lecturer at King's College London, Centre for Telecommunications Research. He obtained his Ph.D. in Telecommunications from King's College London, UK, in 2003, his Diploma in Electrical Engineering from Dresden University of Technology, Germany, in 2000, and his M.Sc. degree in Telecommunications from King's College London, UK, in 1999. In the framework of the Mobile VCE, he has pioneered research on distributed cooperative space-time encoded communication systems, dating back to December 1999. Prior to Telecommunications, he studied Physics in Moscow. He has won various competitions in Mathematics and Physics, and participated in the 3rd round of the International Physics Olympics for Germany. He has been Student Representative of the IEEE UKRI Section, member of the Student Activity Committee of IEEE Region 8 and the London Technology Network Business Fellow for King's College London. He has published more than 100 technical journal and conference papers at a citation h-index of 11 and citation g-index of 22, holds several patents, co-edited and contributed to several books, has given numerous international short-courses, and participated in standardization activities. He has been TPC member and co-chair of various conferences, is technical chair of IEEE PIMRC 2008 to be held in Cannes, France, and is editor for *IEEE Communications Letters*, *IEEE Transactions on Vehicular Technology*, *IEEE Communications Magazine*, *IEEE Wireless Communications*, *IET Communications* (former *IEE Proceedings in Communications*), the *Elsevier PHYCOM - Physical Communication* - journal, the *EURASIP JWCN* journal, and the *Bentham Science Recent Patents on Computer Science* journal. He is a Senior Member of the IEEE.