Int. J. Communications, Network and System Sciences, 2010, 3, 1-18
doi:10.4236/ijcns.2010.31001 Published Online January 2010 (http://www.SciRP.org/journal/ijcns/).

Achievable Rate Regions for Orthogonally Multiplexed MIMO Broadcast Channels with Multi-Dimensional Modulation

Marthe KASSOUF, Harry LEIB
Department of Electrical and Computer Engineering, McGill University, Montreal, Canada
Email: *marthe.kassouf@mail.mcgill.ca, harry.leib@mcgill.ca*
Received July 23, 2009; *revised September* 12, 2009; *accepted November* 27, 2009

Abstract

In this work, we consider a multi-antenna channel with orthogonally multiplexed non-cooperative users, and present its achievable information rate regions with and without channel knowledge at the transmitter. With an informed transmitter, we maximize the rate for each user. With an uninformed transmitter, we consider the optimal power allocation that causes the fastest convergence to zero of the fraction of channels whose mutual information is less than any given rate as the transmitter channel knowledge converges to zero. We assume a deterministic space and time dispersive multipath channel with multiple transmit and receive antennas, generating an orthogonally multiplexed Multiple-Input Multiple-Output (MIMO) broadcast system. Under limited transmit power; we consider different user specific space-time modulation formats that represent assignments of signal dimensions to transmit antennas. For the two-user orthogonally multiplexed MIMO broadcast channels, the achievable rate regions, with and without transmitter channel knowledge, evolve from a triangular region at low SNR to a rectangular region at high SNR. We also investigate the maximum sum rate for these regions and derive the associated power allocations at low and high SNR. Furthermore, we present numerical results for a two-user system that illustrate the effects of channel knowledge at the transmitter, the multi-dimensional space-time modulation format and features of the multipath channel.

Keywords: MIMO Channels, Broadcast Systems, Capacity Region, Space-Time Coding

1. Introduction

Multiple-Input Multiple-Output (MIMO) systems, employing multiple antennas at the transmitter and receiver, have been shown to yield significant capacity gains for single-user channels [1]. A gain in the capacity of MIMO channels is also observed when increasing the number of multipath components [2–4]. Furthermore, channel knowledge at the transmitter has been shown to increase capacity more significantly at low SNR [5,6]. These favorable features trigerred a considerable interest in the application of MIMO technology to multi-user systems as well.

The capacity region of the two-user scalar orthogonal broadcast channel (BC) is shown in [7] to be a rectangle generated by the set of jointly achievable mutual information rate pairs. A larger capacity region may be obtained by allowing multi-user data superposition instead of simple time sharing [8]. Assuming perfect channel state information (CSI) at transmitter and receiver, the optimality of Code Division Multiple Access (CDMA) with successive decoding has been established in [9,10] for flat and frequency selective fading channels. MIMO broadcast channels (BCs) belong to the class of non-degraded broadcast channels, thus, making the evaluation of their capacity regions very difficult. Superposition coding does not apply to non-degraded broadcast channels because users may employ different rates making successive decoding quite difficult if not impossible [11]. However, this reference shows that a capacity region for broadcast channels can be achieved by using a coding technique, nicknamed dirty paper coding (DPC) [12], where the interference is non-causally known to the transmitter and unknown to the receiver. The optimality of DPC in terms of maximizing the sum rate was proved in [13] for a constant two-user BC with single-antenna receivers, and known channel at the transmitter as well as all receivers. Generalizations of results from [13] to

Copyright © 2010 SciRes.

systems with arbitrary number of users and multiple transmit and receive antennas has been carried out independently in [14] and [15]. The sum rate optimality of DPC for Gaussian MIMO BCs has been investigated in [16–18] using the duality [19] between the DPC rate region of the MIMO BC and the capacity region of a Gaussian MIMO MAC with similar power constraint. In [20] it was shown that the DPC rate region is in fact the MIMO BC capacity region. Scaling laws of the sum rate for block fading Rayleigh MIMO BCs with large number of users are considered in [21] using DPC, Time Division Multiple Access (TDMA) and beamforming. The rate balancing problem (i.e. the selection of the capacity region boundary point that satisfies given constraints on the ratios between the users' rates) is considered in [22], which also provides optimal and suboptimal algorithms for MIMO BCs employing Orthogonal Frequency Division Multiplexing (OFDM) transmission.

In this work, we consider a MIMO BC with orthogonally multiplexed non-cooperating users who employ space-time modulation. As in [23,24], we assume a non-fading space and time dispersive multipath environment. These schemes model the downlink of cellular communication systems with orthogonal user multiplexing. We consider a deterministic channel model since it provides an insight to the behaviour of the capacity region with respect to the number of antenna and multipath components, and often serves as a first step towards the study of fading channels. We investigate the achievable rate region of such orthogonally multiplexed broadcast schemes with multi-dimensional space-time modulation, where a transmitter attempts simultaneously to transfer information to several users without mutual interference. When the channel is known at the transmitter, we consider the optimal power allocation that maximizes the rate for each user. We also consider the power allocation for each user that causes the fastest convergence to zero of the fraction of channels whose mutual information is less than any given rate, as the transmitter channel knowledge goes to zero. For both cases, we investigate the maximum sum rate. Considering a two-user broadcast system, we investigate the asymptotic behaviour of the achievable rate regions at low and high SNR, and provide the optimum power allocations that correspond to the maximum sum rate. Illustrative numerical results are provided for users having different propagation channels, using different multi-dimensional space-time modulation schemes and employing different number of antennas. This paper is structured as follows. Section 2 presents the system model. The capacity region with known channel at the transmitter is investigated in Section 3. The case of unknown channel at the transmitter is considered in Section 4. Section 5 presents some illustrative numerical results. The conclusions follow in Section 6.

2. MIMO Broadcast Multipath Channels with Space-Time Modulation

In this paper, column vectors and matrices are represented by lower-case and upper-case bold letters. The dth component of a vector $\mathbf{a}$ is denoted by $[\mathbf{a}]_d$ Furthermore, $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$, $\otimes$ denotes the matrix Kronecker product, and $\cdot$ denotes the matrix product. We use the following superscripts: $*$ for complex conjugate, T for matrix transpose, and $\dagger$ for Hermitian conjugate. The $vec()$ operator denotes the stack in a single column vector of matrix columns or a set of column vectors. The direct sum of n matrices $\{\mathbf{A}_i\}_{i=1}^n$ is denoted by $diag_{i=1}^n[\mathbf{A}_i]$. The vertical stack of n matrices with equal number of columns $\{\mathbf{B}_i\}_{i=1}^n$ in a single matrix is denoted by $stack_{i=1}^n[\mathbf{B}_i]$. The n-square identity matrix is denoted by $\mathbf{I}_n$. The n-dimensional vectors $\mathbf{e}_{i;n}$ for $i=1,2,\cdots,n$ are defined as $\mathbf{e}_{i;n}=\left[\delta(i,1),\delta(i,2),\cdots,\delta(i,n)\right]^T$, with the Kronecker symbol defined by $\delta(i,j)=1$ if $i=j$, and $\delta(i,j)=0$ otherwise. For a scalar a, we have $\{a\}^+=\max(0,a)$. Unless otherwise specified, the function $\log()$ denotes the base-2 logarithm, and the superscript (k) refers to the kth user in the system.

We consider K orthogonally multiplexed users, each with power $P^{(k)}$ satisfying the constraint $\sum_{k=1}^K P^{(k)} = P_T$ and affected by independent interference. Let D_T denote the total number of signal dimensions, with user k occupying a sub-space of dimensionality $D^{(k)}$, where $\sum_{k=1}^K D^{(k)} = D_T$. Each user employs a different signal sub-space. This model corresponds to an orthogonally multiplexed MIMO broadcast channel (BC) without user cooperation. For user k, the propagation medium consists of $L_t^{(k)}$ time resolvable multipath clusters following the 3GPP space and time dispersive channel model [25]. The signal paths of same cluster have equal propagation delays and are resolved in space only. For user k we define the transmitted and received signal vectors $\mathbf{s}^{(k)}(t)=[s_1^{(k)}(t),s_2^{(k)}(t),\cdots,s_{N_T}^{(k)}(t)]^T$ and $\mathbf{z}^{(k)}(t)=[z_1^{(k)}(t),\ z_2^{(k)}(t),\cdots,z_{N_R^{(k)}}^{(k)}(t)]^T$, respectively, with N_T and $N_R^{(k)}$ denoting the number of transmit and receive antennas.

The continuous time channel model is specified by

$$\mathbf{z}^{(k)}(t)=\sum_{l=1}^{L_t^{(k)}}\mathbf{F}_l^{(k)}\cdot\mathbf{s}^{(k)}\left(t-\tau_l^{(k)}\right)+\mathbf{i}^{(k)}(t) \qquad (1)$$

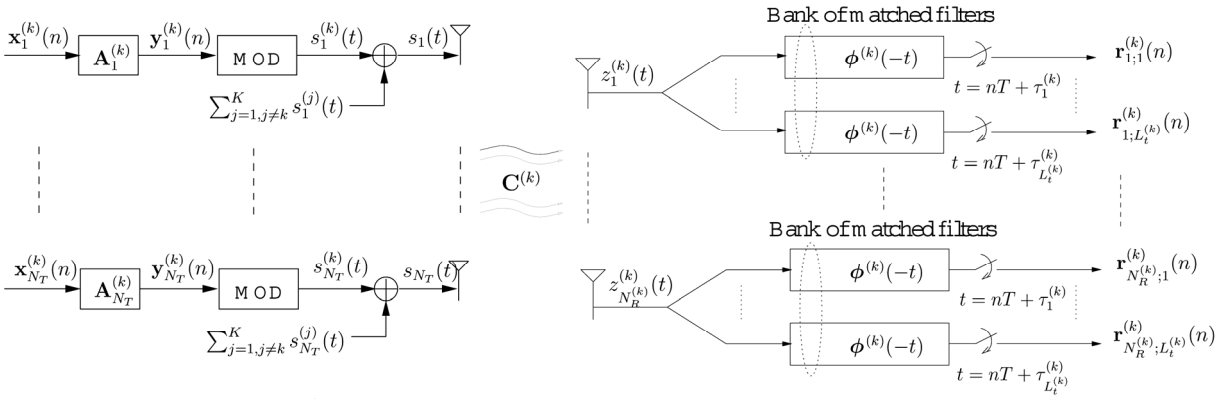

Figure 1. System model for user *k*.

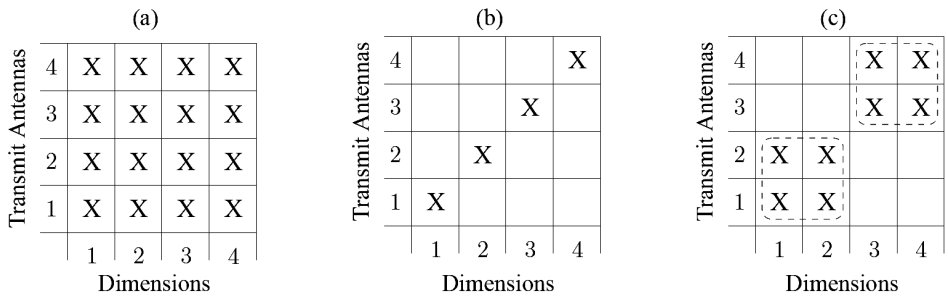

Figure 2. Examples of dimension allocation schemes for user k with $N_T = 4$ transmit antennas and $D^{(k)} = 4$ dimensions: a) ATA, b) OTA, c) POTA with $N_G^{(k)} = 2$ TOGs.

where $\tau_l^{(k)}$ and $\mathbf{F}_l^{(k)}$ denote the propagation delay and the $N_R^{(k)} \times N_T$ channel propagation matrix associated with cluster l ($l \in \{1,\cdots,L_t^{(k)}\}$). The interference vector $\mathbf{i}^{(k)}(t) = [i_1^{(k)}(t), i_2^{(k)}, \cdots, i_{N_R^{(k)}}^{(k)}(t)]^T$ is complex valued zero mean white Gaussian, with autocovariance matrix $E[\mathbf{i}^{(k)}(t)\cdot(\mathbf{i}^{(k)}(t-\tau))^{\dagger}] = N_o\delta(\tau)\mathbf{I}_{N_R^{(k)}}$ where $\delta(\tau)$ denotes the Dirac's delta function.

Consider the system model from [26] illustrated in Figure 1. Assume a modulation process for user k that partitions the transmit antennas into groups called Transmit Orthogonal Groups (TOGs), each sharing a given subset of its $D^{(k)}$ signal dimensions. Hence each TOG employs a different signal sub-space. Let $N_G^{(k)}$ denote the number of TOGs ($1 \le N_G^{(k)} \le N_T$), and $n_i^{(k)}$ denote the number of transmit antennas in the ith TOG, assumed to be adjacent. We assume an equal number of signal dimensions per TOG, $D_G^{(k)} = \dfrac{D^{(k)}}{N_G^{(k)}}$ and define $\tilde{D}^{(k)} = N_T D_G^{(k)}$. The $D_G^{(k)}$-dimensional complex input vector $\mathbf{x}_j^{(k)}(n)$ for transmit antenna j and the corresponding $D^{(k)}$-dimensional signal space vector $\mathbf{y}_j^{(k)}(n)$ are given by $\mathbf{y}_j^{(k)}(n) = (\mathbf{A}_j^{(k)})^T \cdot \mathbf{x}_j^{(k)}(n)$, with $\mathbf{A}_j^{(k)} = \mathbf{e}_{t_j^{(k)};N_G^{(k)}}^T \otimes \mathbf{I}_{D_G^{(k)}}$ and $t_j^{(k)}$ ($t_j^{(k)} \in \{1\cdots,N_G^{(k)}\}$) indicating the TOG to which antenna j belongs. The $D_G^{(k)} \times D^{(k)}$ matrix $\mathbf{A}_j^{(k)}$ determines the signal dimensions (indexed by $d \in \{1,\cdots,D^{(k)}\}$) that can be used on transmit antenna j by making $[\mathbf{y}_j^{(k)}(n)]_d = 0$ if the dth signal dimension is not used on antenna j, or $[\mathbf{y}_j^{(k)}(n)]_d = [\mathbf{x}_j^{(k)}(n)]_{d'}$ for $d' \in \{1,\cdots,D_G^{(k)}\}$ otherwise. The multi-dimensional space-time modulation format is determined by the matrices $\{\mathbf{A}_j^{(k)}\}_{j=1}^{N_T}$. With $N_G^{(k)} = 1$ and $N_G^{(k)} = N_T$ we have the space-time modulation formats Aggregate Transmit Antenna (ATA) and Orthogonal Transmit Antenna (OTA), respectively. The space-time coding system [27] and the Alamouti transmit diversity scheme with two transmit antennas [28] are examples of ATA. The orthogonal transmit diversity technique of IS 2000 [29,30] is an example of OTA. The

more general case $(1 < N_G^{(k)} < N_T)$ corresponds to Partially Orthogonal Transmit Antenna (POTA), that can be viewed as a combination of ATA and OTA. Examples of ATA, OTA and POTA are illustrated in Figures 2(a), 2(b) and 2(c), respectively.

The transmitted waveform through antenna j is given by $s_j^{(k)}(t) = \sum_{n=-\infty}^{\infty} (\mathbf{y}_j^{(k)}(n))^T \cdot \phi^{(k)}(t - nT)$ where $\phi^{(k)}(t) = \left[\phi_1^{(k)}(t), \phi_2^{(k)}, \cdots, \phi_{D^{(k)}}^{(k)}(t)\right]^T$ denotes user k's real orthogonal basis functions and $1/T$ is the symbol rate. We assume no inter-symbol interference and perfect synchronization at every receive antenna. Assuming perfect multipath time resolvability for each user as well as perfect orthogonality between users, we have

$$\int_{-\infty}^{\infty} \phi_d^{(k)}\left(t - \tau_l^{(k)} - nT\right) \phi_{d'}^{(k')}\left(t - \tau_{l'}^{(k')} - n'T\right) dt$$
$$= \delta(k,k')\delta(l,l')\delta(d,d')\delta(n,n')$$

$\forall k, k', \forall l, l' \in \left\{1, \cdots, L_t^{(k)}\right\}$, $\forall d, d' \in \left\{1, \cdots, D^{(k)}\right\}$ and $\forall n, n'$. Since different TOGs employ different signal subspaces, user k's system is equivalent to $N_G^{(k)}$ parallel MIMO systems each consisting of a TOG and the receiver. The overall Maximum Likelihood detection complexity for user k is the sum of his TOGs complexities, each being exponential in $n_i^{(k)}$. By reducing each $n_i^{(k)}$ the overall complexity decreases. It is minimum, and linear in N_T when $n_i^{(k)} = 1$ which corresponds to OTA. Thus, increasing $N_G^{(k)}$ lowers complexity and increases parallelism.

Let $\mathbf{r}_{m;l}^{(k)}(n)$ and $\mathbf{v}_{m;l}^{(k)}(n)$ denote the $D^{(k)}$-dimensional discrete time channel output and noise vectors of the lth time resolvable cluster received on the mth antenna.

Hence, the vectors $\mathbf{x}^{(k)}(n) = vec(\{\mathbf{x}_j^{(k)}(n)\}_{j=1}^{N_T})$ and $\mathbf{r}^{(k)}(n) = vec(\{\mathbf{r}_{m;l}^{(k)}(n)\}_{m=1}^{N_R^{(k)}}\}_{l=1}^{L_t^{(k)}})$ denote the input and output of the discrete time channel defined by

$$\mathbf{r}^{(k)}(n) = \mathbf{H}^{(k)} \cdot \mathbf{x}^{(k)}(n) + \mathbf{v}^{(k)}(n) \tag{2}$$

with noise vector $\mathbf{v}^{(k)}(n) = vec(\{\{\mathbf{v}_{m;l}^{(k)}(n)\}_{m=1}^{N_R^{(k)}}\}_{l=1}^{L_t^{(k)}})$ .Using $\mathbf{A}^{(k)} = diag_{j=1}^{N_T}[\mathbf{A}_j^{(k)}]$, we have $\mathbf{y}^{(k)}(n) = vec(\{\mathbf{y}_j^{(k)}(n)\}_{j=1}^{N_T})$ $= [\mathbf{A}^{(k)}]^T \cdot \mathbf{x}^{(k)}(n)$. Assuming perfect multipath resolvability, the $L_t^{(k)} N_R^{(k)} D^{(k)} \times \tilde{D}^{(k)}$ discrete time channel matrix $\mathbf{H}^{(k)}$ can be seen as the stack of $N_R^{(k)} D^{(k)} \times \tilde{D}^{(k)}$ submatrices each associated with a time resolvable cluster. Therefore, we have

$$\mathbf{H}^{(k)} = stack_{l=1}^{L_t^{(k)}}[(\mathbf{F}_l^{(k)} \otimes \mathbf{I}_{D^{(k)}}) \cdot [\mathbf{A}^{(k)}]^T]$$
$$= stack_{l=1}^{L_t^{(k)}}[(\mathbf{F}_l^{(k)} \otimes \mathbf{I}_{D^{(k)}})] \cdot [\mathbf{A}^{(k)}]^T$$

or, equivalently,

$$\mathbf{H}^{(k)} = \left[\mathbf{C}^{(k)} \otimes \mathbf{I}_{D^{(k)}}\right] \cdot \left(\mathbf{A}^{(k)}\right)^T \tag{3}$$

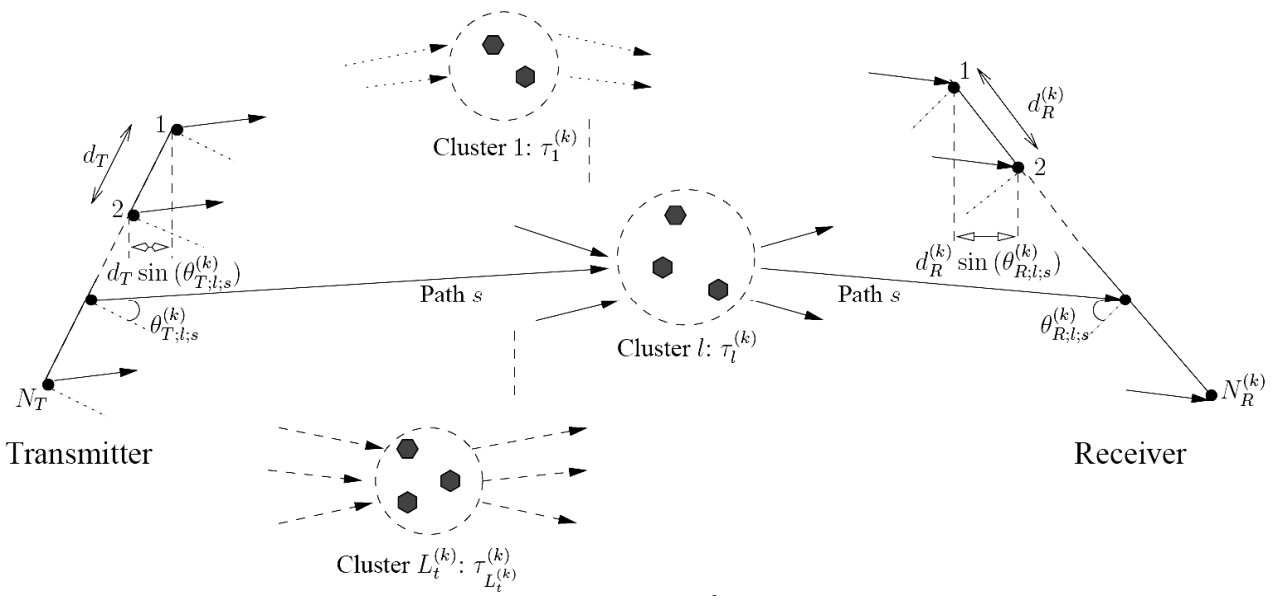

Figure 3. Space and time dispersive MIMO channel model for user k.

with $\mathbf{C}^{(k)} = stack_{l=1}^{L_t^{(k)}}[\mathbf{F}_l^{(k)}]$. Using the 3GPP spatial channel model [25] illustrated in Figure 3, let $S_l^{(k)}$ denote the number of propagation paths in cluster l, with path s $(s = 0,\cdots,S_l^{(k)}-1)$ characterized by the gain coefficient $G_{l;s}^{(k)}$, angle of departure (AOD) $\theta_{T;l;s}^{(k)}$ and angle of arrival (AOA) $\theta_{R;l;s}^{(k)}$.The total number of propagation paths for user k is $\sum_{l=1}^{L_t^{(k)}} S_l^{(k)}$. The spaces between adjacent transmit and receive antennas are denoted by d_T and $d_R^{(k)}$. We use the notations $\alpha_{l;s}^{(k)} = -2\pi \frac{d_R^{(k)}}{\lambda}$ $\sin(\theta_{R;l;s}^{(k)})$ and $\beta_{l;s}^{(k)} = -2\pi \frac{d_T}{\lambda}\sin(\theta_{T;l;s}^{(k)})$ with λ denoting the signal wavelength. The space signature vectors at the receiver and transmitter are given by $\mathbf{a}_{l;s}^{(k)} = [1, e^{j\alpha_{l;s}^{(k)}},\cdots,e^{j(N_R^{(k)}-1)\alpha_{l;s}^{(k)}}]^T$ and $\mathbf{b}_{l;s}^{(k)} = [1, e^{-j\beta_{l;s}^{(k)}},\cdots,$ $e^{-j(N_T-1)\beta_{l;s}^{(k)}}]^T$ and from [25] we can write $\mathbf{F}_l^{(k)} =$ $\sum_{s=0}^{S_l^{(k)}-1} G_{l;s}^{(k)}\mathbf{a}_{l;s}^{(k)}\cdot(\mathbf{b}_{l;s}^{(k)})^\dagger$. Subsequently, we define the $L_t^{(k)}N_R^{(k)}\times n_i^{(k)}$ propagation matrix $\mathbf{C}_i^{(k)}$ describing the propagation between the ith TOG and the receiver of user k, such that $\mathbf{C}^{(k)} = [\mathbf{C}_1^{(k)}\mathbf{C}_2^{(k)}\cdots\mathbf{C}_{N_G^{(k)}}^{(k)}]$. Let $r_i^{(k)}$ denote the rank of $\mathbf{C}_i^{(k)}$ and define $r_T^{(k)} = \sum_{i=1}^{N_G^{(k)}} r_i^{(k)}$. We also use $\mathbf{M}_i^{(k)}$ and $\mathbf{Q}_i^{(k)}$ to denote the $n_i^{(k)}$-square unitary and diagonal matrices associated with the eigenvalue decomposition $(\mathbf{C}_i^{(k)})^\dagger\cdot\mathbf{C}_i^{(k)} = (\mathbf{M}_i^{(k)})^\dagger\cdot\mathbf{Q}_i^{(k)}\cdot\mathbf{M}_i^{(k)}$.

Assuming a memoryless channel, we can drop the dependencies on the time index n for the remainder of this paper. Moreover, we consider orthogonally multiplexed broadcast MIMO channels with two users (K = 2) for simplicity. The results can be generalized to broadcast systems with an arbitrary number of users. In the next section, we investigate the capacity region assuming that the transmitter and both receivers have perfect knowledge of the channel propagation matrices $\{\mathbf{C}^{(k)}\}_{k=1}^2$ and the multi-dimensional space-time modulation formats $\{\mathbf{A}^{(k)}\}_{k=1}^2$.

3. Capacity Region with Known Channel at the Transmitter

Let $\sum^{(k)} = E[\mathbf{x}^{(k)}\cdot(\mathbf{x}^{(k)})^\dagger]$ denote the input covariance matrix of user k constrained by $Tr(\sum^{(k)}) \le P^{(k)}$. For a fixed $\sum^{(k)}$, the input/output average mutual information for (2) is maximized by a Gaussian input distribution and it is given by [1]

$$I(\mathbf{x}^{(k)},\mathbf{r}^{(k)}) = \log(|\mathbf{I}_{L_t^{(k)}N_R^{(k)}D^{(k)}} + \frac{1}{N_0}\mathbf{H}^{(k)}\cdot\sum^{(k)}\cdot(\mathbf{H}^{(k)})^\dagger|) \tag{4}$$

From [1,31], Shannon capacity is obtained by maximizing $I(\mathbf{x}^{(k)},\mathbf{r}^{(k)})$ over all positive semidefinite input covariance matrices $\sum^{(k)}$ satisfying $Tr(\sum^{(k)}) \le P^{(k)}$. From [32], we have that this capacity is obtained using a water-filling power allocation [31], and it is given by

$$C^{(k)} = \frac{1}{N_G^{(k)}}\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{n_i^{(k)}}\{\log(\mu^{(k)}\lambda_{i;l}^{(k)})\}^+ \quad bps/\text{Hz} \tag{5}$$

where $\lambda_{i;1}^{(k)} \ge \lambda_{i;2}^{(k)} \ge \cdots \ge \lambda_{i;n_i^{(k)}}^{(k)} \ge 0$ denote the (real and non-negative) ordered eigenvalues of $(\mathbf{C}_i^{(k)})^\dagger\cdot\mathbf{C}_i^{(k)}$, and the constant $\mu^{(k)}$ satisfies $\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{n_i^{(k)}}\{\mu^{(k)} - (\lambda_{i;l}^{(k)})^{-1}\}^+$ $= \frac{N_G^{(k)}P^{(k)}}{N_0 D^{(k)}}$. The corresponding input signal $\mathbf{x}^{(k)}$ is zero mean Gaussian with a block diagonal covariance matrix $\sum^{(k)} = diag_{i=1}^{N_G^{(k)}}[\sum_i^{(k)}]$ where the $n_i^{(k)}D_G^{(k)}$-square input covariance matrix of TOG i is given by

$$\sum_i^{(k)} = (\mathbf{M}_i^{(k)}\otimes\mathbf{I}_{D_G^{(k)}})^\dagger$$
$$\cdot(N_0 diag_{l=1}^{n_i^{(k)}}[\{\mu^{(k)} - \left(\lambda_{i;l}^{(k)}\right)^{-1}\}^+]\otimes\mathbf{I}_{D_G^{(k)}})\cdot(\mathbf{M}_i^{(k)}\otimes\mathbf{I}_{D_G^{(k)}})$$

that, using [33], reduces to

$$\sum_i^{(k)} = [(\mathbf{M}_i^{(k)})^\dagger$$
$$\cdot(N_0 diag_{l=1}^{n_i^{(k)}}[\{\mu^{(k)} - (\lambda_{i;l}^{(k)})^{-1}\}^+])\cdot\mathbf{M}_i^{(k)}]\otimes\mathbf{I}_{D_G^{(k)}} \tag{6}$$

All rate pairs $(R^{(1)},R^{(2)})$ such that $R^{(1)} \le I(\mathbf{x}^{(1)},\mathbf{r}^{(1)})$ and $R^{(2)} \le I(\mathbf{x}^{(2)},\mathbf{r}^{(2)})$ are achievable, and the capacity region is the closure of all such rate pairs $(R^{(1)},R^{(2)})$. We specify the power allocation between the two users by $\alpha_p = \frac{P^{(1)}}{P_T}$, which is the fraction of power allocated to user one. The fraction of power allocated to user two is $1-\alpha_p = \frac{P^{(2)}}{P_T}$. Using (5) and the notation $C^{(k)}(\alpha_p)$ for $C^{(k)}, k = 1,2,$ the boundary of the capacity region is a parametric curve in α_p defined by

$$\begin{cases} C^{(1)}\left(\alpha_p\right)= \\ \dfrac{1}{N_G^{(1)}}\sum_{i=1}^{N_G^{(1)}}\sum_{l=1}^{n_i^{(1)}}\left\{\log\left(\mu^{(1)}\lambda_{i;l}^{(1)}\right)\right\}^+ \quad bps/\mathrm{Hz} \\ C^{(2)}\left(\alpha_p\right)= \\ \dfrac{1}{N_G^{(2)}}\sum_{i=1}^{N_G^{(2)}}\sum_{l=1}^{n_i^{(2)}}\left\{\log\left(\mu^{(2)}\lambda_{i;l}^{(2)}\right)\right\}^+ \quad bps/\mathrm{Hz} \end{cases} \tag{7}$$

such that $\sum_{i=1}^{N_G^{(1)}}\sum_{l=1}^{n_i^{(1)}}\{\mu^{(1)}-\left(\lambda_{i;l}^{(1)}\right)^{-1}\}^+=\dfrac{N_G^{(1)}\alpha_p P_T}{N_0 D^{(1)}}$ and $\sum_{i=1}^{N_G^{(2)}}\sum_{l=1}^{n_i^{(2)}}\{\mu^{(2)}-(\lambda_{i;l}^{(2)})^{-1}\}^+=\dfrac{N_G^{(2)}(1-\alpha_p)P_T}{N_0 D^{(2)}}$ with $\alpha_p\in[0,1]$. Using a time-sharing argument as in [31], we have that the capacity region is convex-$\cap$, and also continuous since continuity is an underlying property of convexity [34]. Moreover, as α_p increases in the interval [0, 1], $P^{(1)}$ increases and $P^{(2)}$ decreases, leading to an increase in $C^{(1)}(\alpha_p)$ and decrease in $C^{(2)}(\alpha_p)$, respectively. Hence, $C^{(1)}(\alpha_p)$ is monotonically increasing with α_p while $C^{(2)}(\alpha_p)$ is monotonically decreasing with α_p. It follows that $C^{(2)}(\alpha_p)$ is a monotonically decreasing function of $C^{(1)}(\alpha_p)$.

In order to assess the transmission performance of a multi-user system using a comparison of capacity regions, we introduce the following definition:

Definition 1: A capacity or rate region is said to be larger (respectively, smaller) than another region if the former contains (respectively, is contained in) the latter for the same power P_T.

For a fixed α_p (or $P^{(k)}$), [32] shows that for a single user system $C^{(k)}(\alpha_p)$ is maximized by ATA, and for $N_G^{(k)}\geq 2$ merging TOGs cannot decrease $C^{(k)}(\alpha_p)$. It follows that, for k = 1, 2 and for a given α_p, maximum, intermediate and minimum values for $C^{(k)}(\alpha_p)$ are respectively obtained when user k employs ATA, POTA and OTA, provided that the POTA system can be obtained from OTA by merging TOGs. A straightforward application of this result to the boundary points of the capacity region (7) of the orthogonally multiplexed MIMO BC leads to the following theorem:

Theorem 1: With channel known at the transmitter, the capacity region is largest, intermediate and smallest when the users employ ATA, POTA and OTA, respectively.

Considering a transmitter with N_T transmit antennas, ATA represents a transmission strategy where there are no constraints on the assignment of a signal dimension to the transmit antennas. As such, with ATA a signal dimension can be used on all the antennas. For POTA, a signal dimension is constrained to be used only on a subset of antennas, while for OTA it is constrained to be used only on one antenna. Thus, as $N_G^{(k)}$ increases the assignment of a signal dimension to transmit antennas is more constrained, yielding a decrease or no change in the capacity.

We now investigate the sum capacity. Given P_T, we use $\alpha_{p;s}$ to denote the power allocation that maximizes the sum capacity. Define the maximum sum capacity $C_s=\max_{\alpha_p\in[0,1]}(C^{(1)}(\alpha_p)+C^{(2)}(\alpha_p))$, and let $(C_s^{(1)},C_s^{(2)})$ denote the point of the capacity region boundary that corresponds to $\alpha_{p;s}$. We have $C_s=C_s^{(1)}+C_s^{(2)}$. Using the convexity property and considering the two-dimensional plane defined by the axes $R^{(1)}$ and $R^{(2)}$, the point $(C_s^{(1)},C_s^{(2)})$ corresponds to the intersection of the boundary of the capacity region $(C^{(1)}(\alpha_p),C^{(2)}(\alpha_p))$ with the affine function $R^{(1)}+R^{(2)}=C_s$. Thus, $(C_s^{(1)},C_s^{(2)})$, is the point at which the support line to $(C^{(1)}(\alpha_p),C^{(2)}(\alpha_p))$ has a slope equal to -1, and is located at the most right of the capacity region.

Let $(E_b)_T^{(k)}$ denote the transmitted energy per bit when operating at capacity limits. We have $(E_b)_T^{(k)}=\dfrac{P^{(k)}}{C^{(k)}(\alpha_p)D^{(k)}}$. The SNR per bit $(E_b/N_0)_T^{(k)}$ referenced at the transmitter satisfies

$$\frac{P^{(k)}}{N_0 D^{(k)}}=C^{(k)}(\alpha_p)(E_b/N_0)_T^{(k)} \tag{8}$$

Next, we consider the capacity region and sum capacity at low and high SNR.

3.1. Capacity Region at Low SNR

For user k at low SNR only the eigenmodes corresponding to the maximum eigenvalue $\tilde{\lambda}^{(k)}=\max\{\{\lambda_{i;j}^{(k)}\}_{j=1}^{n_i^{(k)}}\}_{i=1}^{N_G^{(k)}}$ are active. Let $\tilde{r}^{(k)}$ denote the number of elements of $\{\{\lambda_{i;j}^{(k)}\}_{j=1}^{n_i^{(k)}}\}_{i=1}^{N_G^{(k)}}$ that are equal to $\tilde{\lambda}^{(k)}$. Thus, using (8) and (5) we have

$$C^{(k)}(\alpha_p)=\frac{\tilde{r}^{(k)}}{N_G^{(k)}}\log(1+\tilde{\lambda}^{(k)}\frac{N_G^{(k)}C^{(k)}(\alpha_p)}{\tilde{r}^{(k)}}(E_b/N_0)_T^{(k)}).$$

Equivalently, we can write

$$(E_b/N_0)_T^{(k)}=\frac{2^{\frac{N_G^{(k)}C^{(k)}(\alpha_p)}{\tilde{r}^{(k)}}}-1}{\frac{N_G^{(k)}C^{(k)}(\alpha_p)}{\tilde{r}^{(k)}}\tilde{\lambda}^{(k)}}\approx\frac{N_G^{(k)}(\ln 2)^2}{2\tilde{r}^{(k)}\tilde{\lambda}^{(k)}}C^{(k)}(\alpha_p)+\frac{\ln 2}{\tilde{\lambda}^{(k)}} \tag{9}$$

showing that $C^{(k)}(\alpha_p)$ is linear in $(E_b/N_0)_T^{(k)}$ with slope $\frac{2\tilde{r}^{(k)}\tilde{\lambda}^{(k)}}{N_G^{(k)}(\ln 2)^2}$. As $C^{(k)}(\alpha_p)\to 0$, we have $(E_b/N_0)_T^{(k)}\to\frac{\ln 2}{\tilde{\lambda}^{(k)}}$ (which corresponds to $-1.6-10\log_{10}(\tilde{\lambda}^{(k)})$ dB). Furthermore, (7) becomes

$$\begin{cases}C^{(1)}(\alpha_p)=\\ \frac{\tilde{r}^{(1)}}{N_G^{(1)}}\log(1+\tilde{\lambda}^{(1)}\frac{N_G^{(1)}\alpha_p P_T}{\tilde{r}^{(1)}N_0D^{(1)}})\approx\frac{\tilde{\lambda}^{(1)}}{\ln 2}\frac{\alpha_p P_T}{N_0D^{(1)}}\\ C^{(2)}(\alpha_p)=\\ \frac{\tilde{r}^{(2)}}{N_G^{(2)}}\log(1+\tilde{\lambda}^{(2)}\frac{N_G^{(2)}(1-\alpha_p)P_T}{\tilde{r}^{(2)}N_0D^{(2)}})\approx\frac{\tilde{\lambda}^{(2)}}{\ln 2}\frac{(1-\alpha_p)P_T}{N_0D^{(2)}}\end{cases} \tag{10}$$

Hence, we have $\alpha_p P_T=\frac{N_0D^{(1)}\ln 2}{\tilde{\lambda}^{(1)}}C^{(1)}(\alpha_p)$ leading to a triangular capacity region (in the positive quadrant of the two-dimensional plane defined by the axes $R^{(1)}$ and $R^{(2)}$) bounded by the line

$$R^{(2)}=-\frac{\tilde{\lambda}^{(2)}D^{(1)}}{\tilde{\lambda}^{(1)}D^{(2)}}R^{(1)}+\frac{\tilde{\lambda}^{(2)}}{\ln 2}\frac{P_T}{N_0D^{(2)}} \tag{11}$$

As $P_T/N_0\to 0$, the segment (11) converges to the point (0, 0).

We investigate the sum capacity at low SNR by considering three possible cases depending on the channel parameters:

1) If $\frac{\tilde{\lambda}^{(2)}D^{(1)}}{\tilde{\lambda}^{(1)}D^{(2)}}>1$, then the right most support line $R^{(1)}+R^{(2)}=C_s$ intersects the capacity region boundary at $(C_s^{(1)},C_s^{(2)})=(0,\frac{\tilde{\lambda}^{(2)}}{\ln 2}\frac{P_T}{N_0D^{(2)}})$, which corresponds to $\alpha_{p;s}=0$ with total transmit power allocated to user two and a sum capacity $C_s=\frac{\tilde{\lambda}^{(2)}}{\ln 2}\frac{P_T}{N_0D^{(2)}}$.

2) If $\frac{\tilde{\lambda}^{(2)}D^{(1)}}{\tilde{\lambda}^{(1)}D^{(2)}}<1$ then the right most support line $R^{(1)}+R^{(2)}=C_s$ intersects the capacity region boundary at $(C_s^{(1)},C_s^{(2)})=(\frac{\tilde{\lambda}^{(1)}}{\ln 2}\frac{P_T}{N_0D^{(1)}},0)$, which corresponds to $\alpha_{p;s}=1$ with total transmit power allocated to user one and a sum capacity $C_s=\frac{\tilde{\lambda}^{(1)}}{\ln 2}\frac{P_T}{N_0D^{(1)}}$.

3) If $\frac{\tilde{\lambda}^{(2)}D^{(1)}}{\tilde{\lambda}^{(1)}D^{(2)}}=1$ then the support line with slope -1 lies on the boundary of the capacity region (11), thus, maximizing the sum capacity for any power allocation α_p. Hence, $\alpha_{p;s}$ may take arbitrary value in [0, 1] and $(C_s^{(1)},C_s^{(2)})$ could be any point of the segment (11). The sum capacity is given by $C_s=\frac{\tilde{\lambda}^{(2)}}{\ln 2}\frac{P_T}{N_0D^{(2)}}=\frac{\tilde{\lambda}^{(1)}}{\ln 2}\frac{P_T}{N_0D^{(1)}}$.

3.2. Capacity Region at High SNR

At high SNR, all $r_T^{(k)}D_G^{(k)}$ eigenmodes are active and the channel capacity from (5) becomes $C^{(k)}(\alpha_p)=\frac{1}{N_G^{(k)}}\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{r_i^{(k)}}\log(\mu^{(k)}\lambda_{i;l}^{(k)})$ with $\mu^{(k)}=\frac{N_G^{(k)}P^{(k)}}{r_T^{(k)}N_0D^{(k)}}+\frac{1}{r_T^{(k)}}[\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{r_i^{(k)}}(\lambda_{i;l}^{(k)})^{-1}]$. We define $A^{(k)}=\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{r_i^{(k)}}\log(\lambda_{i;l}^{(k)})$ and $d^{(k)}=\frac{1}{r_T^{(k)}}\sum_{i=1}^{N_G^{(k)}}\sum_{l=1}^{r_i^{(k)}}\left(\lambda_{i;l}^{(k)}\right)^{-1}$ yielding,

$$C^{(k)}(\alpha_p)=\frac{A^{(k)}}{N_G^{(k)}}+\frac{r_T^{(k)}}{N_G^{(k)}}\log(\frac{N_G^{(k)}P^{(k)}}{r_T^{(k)}N_0D^{(k)}}+d^{(k)}).$$

Equivalently, we have

$$C^{(k)}(\alpha_p)=X^{(k)}+\frac{r_T^{(k)}}{N_G^{(k)}}\log(1+\frac{N_G^{(k)}}{r_T^{(k)}}\frac{P^{(k)}}{N_0D^{(k)}d^{(k)}}) \tag{12}$$

where

$$X^{(k)}=\frac{A^{(k)}}{N_G^{(k)}}+\frac{r_T^{(k)}}{N_G^{(k)}}\log(d^{(k)}). \tag{13}$$

Using (8), we can write

$$(E_b/N_0)_T^{(k)} = \frac{2^{\left[\frac{N_G^{(k)}C^{(k)}(\alpha_p)-A^{(k)}}{r_T^{(k)}}\right]} - d^{(k)}}{\frac{N_G^{(k)}C^{(k)}(\alpha_p)}{r_T^{(k)}}} \approx \frac{2^{\left[\frac{N_G^{(k)}C^{(k)}(\alpha_p)-A^{(k)}}{r_T^{(k)}}\right]}}{\frac{N_G^{(k)}C^{(k)}(\alpha_p)}{r_T^{(k)}}} \quad (14)$$

As $C^{(k)}(\alpha_p) \to \infty, (E_b/N_0)_T^{(k)}$ increases exponentially, making $(E_b/N_0)_T^{(k)}$ dB linear in $C^{(k)}(\alpha_p)$ with slope $10\log_{10}(2)\frac{N_G^{(k)}}{r_T^{(k)}} \approx 3\frac{N_G^{(k)}}{r_T^{(k)}}$. Furthermore, we prove the following theorem in the appendix

Theorem 2: As $P_T/N_0 \to \infty$, the asymptotic capacity region with known channel at the transmitter becomes rectangular, defined by the points (0, 0), $(0, \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$ $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), 0)$ and $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$.

Hence, regardless of the space-time modulation format the capacity region of the orthogonally multiplexed MIMO broadcast channel converges to a rectangle, similar to that of orthogonal broadcast channels [7]. Next we investigate the sum capacity at high SNR for the following cases:

1) $0 < \alpha_p < 1$: Using (12), we can write $\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)} = \frac{dC^{(2)}(\alpha_p)}{d\alpha_p}\frac{d\alpha_p}{dC^{(1)}(\alpha_p)}$, which yields

$$\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)} = -\frac{d^{(1)}N_0D^{(1)} + \frac{N_G^{(1)}}{r_T^{(1)}}\alpha_p P_T}{d^{(2)}N_0D^{(2)} + \frac{N_G^{(2)}}{r_T^{(2)}}(1-\alpha_p)P_T}. \quad (15)$$

The denominator of (15) becomes zero for $\alpha_p = 1 + \frac{r_T^{(2)}d^{(2)}N_0D^{(2)}}{N_G^{(2)}P_T}$ which is strictly larger than 1, thus, making $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ differentiable for all $\alpha_p \in (0,1)$. It follows that the support line at every point of the capacity region boundary with $\alpha_p \neq 0,1$ is unique and equal to the tangential line [34]. As $P_T/N_0 \to \infty$, we have from (15) that $\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)} \to -\frac{N_G^{(1)}}{r_T^{(1)}}\frac{r_T^{(2)}}{N_G^{(2)}}\frac{\alpha_p}{(1-\alpha_p)}$. The power allocation that maximizes the sum capacity can be obtained by solving $[\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)}]_{\alpha_p=\alpha_{p;s}} = -1$, yielding $\alpha_{p;s} = \{\frac{N_G^{(2)}}{r_T^{(2)}} + \frac{N_0}{P_T}(d^{(2)}D^{(2)} - d^{(1)}D^{(1)})\}\{\frac{N_G^{(1)}}{r_T^{(1)}} + \frac{N_G^{(2)}}{r_T^{(2)}}\}^{-1}$, and as $P_T/N_0 \to \infty$

$$\alpha_{p;s} = \frac{N_G^{(2)}}{r_T^{(2)}}\frac{r_T^{(1)}}{N_G^{(1)}}[1 + \frac{N_G^{(2)}}{r_T^{(2)}}\frac{r_T^{(1)}}{N_G^{(1)}}]^{-1}. \quad (16)$$

Furthermore, from Theorem 2 we have that the capacity region boundary points corresponding to $\alpha_p \in (0,1)$ converge toward the upper right corner of the limiting rectangle. Therefore, we have

$(C_s^{(1)}, C_s^{(2)}) \to (\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$ and $\frac{C_s}{\log(P_T/N_0)} \to (\frac{r_T^{(1)}}{N_G^{(1)}} + \frac{r_T^{(2)}}{N_G^{(2)}})$. It is seen that both C_s and $\alpha_{p;s}$ depend on the users space-time modulation formats as well as the ranks of the TOGs propagation matrices without being dependent on the channel eigenvalues $\lambda_{i;l}^{(k)}$.

2) $\alpha_p = 0$: From (15) the right derivative is such that

$$[\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)}]_{\alpha_p=0}^{+} = -d^{(1)}N_0D^{(1)}[d^{(2)}N_0D^{(2)} + \frac{N_G^{(2)}}{r_T^{(2)}}P_T]^{-1}$$

The support line at $(C^{(1)}(0), C^{(2)}(0))$ is not unique and has a slope that varies in the interval $[[\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)}]_{\alpha_p=0}^{+}, \infty]$. The lower bound of this interval is larger than -1 for P_T/N_0 sufficiently large and hence, no support line at the point $(C^{(1)}(0), C^{(2)}(0))$ can have a slope equal to -1. It follows that $(C^{(1)}(0), C^{(2)}(0))$ (for which $\alpha_p = 0$) cannot be an intersection of the capacity region boundary with the affine function $R^{(1)} + R^{(2)} = C_s$, yielding $\alpha_{p;s} \neq 0$.

3) $\alpha_p = 1$: From (15) the left derivative is such that

$$[\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)}]_{\alpha_p=1}^{-} = -(d^{(1)}N_0D^{(1)} + \frac{N_G^{(1)}}{r_T^{(1)}}P_T)[d^{(2)}N_0D^{(2)}]^{-1}$$

The support line at $(C^{(1)}(1), C^{(2)}(1))$ is not unique and it has a slope in $(-\infty, [\frac{dC^{(2)}(\alpha_p)}{dC^{(1)}(\alpha_p)}]_{\alpha_p=1}^{-}] \cup [0, \infty)$ that does not include -1 if P_T/N_0 is sufficiently large. Hence, no support line at the point $(C^{(1)}(1), C^{(2)}(1))$ can have a slope equal to -1. It follows that $(C^{(1)}(1), C^{(2)}(1))$ (for

which $\alpha_p = 1$) cannot be an intersection of the capacity region boundary with the affine function $R^{(1)} + R^{(2)} = C_s$, yielding $\alpha_{p;s} \neq 1$.

4. Rate Region with Unknown Channel at the Transmitter

In this section, we assume that the transmitter has no information about the channel matrices $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$.Without channel knowledge at the transmitter, [1, 35] advocate to uniformly distribute the transmit power among all antennas. In [32], we represented the lack of channel knowledge at the transmitter in a single user system by an uninformative a-prior probability distribution on the channel propagation matrix, and considered the following optimality criterion:

Definition 2 (Optimality Criterion 1) An input covariance matrix $\sum^{(k)}$ subject $Tr(\sum^{(k)}) \leq P^{(k)}$ is said to be optimal in sense 1 if, as the transmitter channel knowledge converges to zero, it causes the fastest convergence to zero of the fraction of channels for which the input/output mutual information is below any specific value R, $0 < R < \infty$.

In [32] we considered input covariance matrices of the form $\sum^{(k)} = diag_{i=1}^{N_G^{(k)}}[\sum_i^{(k)}]$ where $\sum_i^{(k)} = \mathbf{S}_i^{(k)} \otimes \mathrm{I}_{D_G^{(k)}}$ and $\mathbf{S}_i^{(k)}$ is $n_i^{(k)} \times n_i^{(k)}$ Hermitian positive semidefinite, similarly to the water-filling matrix (6), and have shown that Optimality Criterion 1 is satisfied using a zero mean Gaussian input vector $\mathrm{x}^{(k)}$ of independent components, with input covariance matrix of TOG i given by

$$\sum_i^{(k)} = \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \mathrm{I}_{n_i^{(k)}} \otimes \mathrm{I}_{D_G^{(k)}} \equiv \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \mathrm{I}_{n_i^{(k)} D_G^{(k)}} . \quad (17)$$

Using this uniform power allocation for each user in each TOG, the transmission rate (4) for user $k\left(k = 1, 2\right)$ was shown to be [32]

$$I_u^{(k)} = \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} \sum_{j=1}^{n_i^{(k)}} \log(1 + \frac{1}{N_0} \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \lambda_{i;j}^{(k)}) \text{ bps / Hz} \quad (18)$$

which depends on α_p through $P^{(k)}$ and can be subsequently denoted as $I_u^{(k)}(\alpha_p)$ with $\alpha_p \in [0,1]$.

From [32], the capacity $C^{(k)}(\alpha_p)$ and the transmission rate $I_u^{(k)}(\alpha_p)$ present several common properties, such as continuity and convexity as well as similar asymptotic behaviour. Using (18), the boundary of the rate region with unknown channel at the transmitter is given by

$$\begin{cases} I_u^{(1)}\left(\alpha_p\right) = \\ \frac{1}{N_G^{(1)}} \sum_{i=1}^{N_G^{(1)}} \sum_{j=1}^{n_i^{(1)}} \log(1 + \frac{\alpha_p N_G^{(1)} P_T}{N_0 D^{(1)}} \frac{\lambda_{i;j}^{(1)}}{N_T}) \\ I_u^{(2)}\left(\alpha_p\right) = \\ \frac{1}{N_G^{(2)}} \sum_{m=1}^{N_G^{(2)}} \sum_{l=1}^{n_m^{(2)}} \log(1 + \frac{\left(1-\alpha_p\right) N_G^{(2)} P_T}{N_0 D^{(2)}} \frac{\lambda_{m;l}^{(2)}}{N_T}) \end{cases} \quad (19)$$

Let $(E_b / N_0)_{T;u}^{(k)}$ denote the SNR per bit referenced at the transmitter. When operating at rate $I_u^{(k)}(\alpha_p)$, we have as in (8)

$$\frac{P^{(k)}}{N_0 D^{(k)}} = I_u^{(k)}(\alpha_p)(E_b / N_0)_{T;u}^{(k)} . \quad (20)$$

As for the case of known channel at the transmitter, $I_u^{(1)}(\alpha_p)$ is monotonically increasing with α_p while $I_u^{(2)}(\alpha_p)$ is monotonically decreasing with α_p. Thus, $I_u^{(2)}(\alpha_p)$ is monotonically decreasing with $I_u^{(1)}(\alpha_p)$. Under similar transmit power constraint and for a given α_p, we have $I_u^{(k)}(\alpha_p) \leq C^{(k)}(\alpha_p)$, with equality achieved when the following necessary and sufficient conditions are satisfied:

Theorem 3: The capacity region $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ with transmitter channel knowledge is equal to the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ without transmitter channel knowledge if and only if for each user the TOG propagation matrices are full column rank (i.e. $r_i^{(k)} = n_i^{(k)}$) with all eigenmodes being active and with equal eigenvalues $\lambda_{i;l}^{(k)} = \lambda_i^{(k)}$, such that

$$(\lambda_i^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_0 N_T D^{(k)}} = \text{Constant}$$

for all $l = 1, \cdots, n_i^{(k)}$ and all $i = 1, \cdots, N_G^{(k)}, k = 1, 2.$

The proof of this theorem can be found in the appendix.

In [32], we conjectured that the single-user information transmission rate $I_u^{(k)}(\alpha_p)$ is maximum, intermediate and minimum with ATA, POTA and OTA, respectively. Since this statement holds for any α_p, it can be easily extended to the orthogonally multiplexed MIMO BC as follows:

Conjecture 1: Using the uniform power allocation for each user with input covariance matrix $\sum^{(k)} = = diag_{i=1}^{N_G^{(k)}}[\sum_i^{(k)}]$ and $\sum_i^{(k)}$ given in (17), the rate

region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ is largest, intermediate and smallest when the users employ ATA, POTA and OTA, respectively.

Next, we investigate the asymptotic behaviour of the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and the maximum sum rate at low and high SNR. For the remainder of this section, we use $\alpha'_{p;s}$ to denote the value of α_p that corresponds to the maximum sum rate. The associated point on the boundary of the rate region is denoted by $(I_{u;s}^{(1)}, I_{u;s}^{(2)})$ and the maximum sum rate by $I_{u;s} = I_{u;s}^{(1)} + I_{u;s}^{(2)}$. We also prove Conjecture 1 in these extreme SNR regimes.

4.1. Rate Region at Low SNR

At low SNR, (18) reduces to

$$I_u^{(k)}(\alpha_p) = \frac{1}{N_G^{(k)}} \log\left(\prod_{i=1}^{N_G^{(k)}} \prod_{j=1}^{n_i^{(k)}} [1 + \frac{1}{N_0} \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \lambda_{i;j}^{(k)}]\right) \approx \frac{1}{N_G^{(k)}} \log\left(1 + \sum_{i=1}^{N_G^{(k)}} \sum_{j=1}^{n_i^{(k)}} [\frac{1}{N_0} \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \lambda_{i;j}^{(k)}]\right).$$

By defining, $B^{(k)} = N_G^{(k)} \sum_{i=1}^{N_G^{(k)}} \sum_{j=1}^{n_i^{(k)}} \frac{\lambda_{i;j}^{(k)}}{N_T} \equiv \frac{N_G^{(k)}}{N_T} \sum_{i=1}^{N_G^{(k)}} Tr[(\mathbf{C}_i^{(k)})^\dagger \cdot \mathbf{C}_i^{(k)}]$, we can write

$$I_u^{(k)}(\alpha_p) = \frac{1}{N_G^{(k)}} \log\left(1 + \frac{P^{(k)} B^{(k)}}{N_0 D^{(k)}}\right) \tag{21}$$

and, using (20), we have

$$(E_b / N_0)_{T;u}^{(k)} = \frac{2^{N_G^{(k)} I_u^{(k)}(\alpha_p)} - 1}{B^{(k)} I_u^{(k)}(\alpha_p)} \approx \frac{(N_G^{(k)})^2 (\ln 2)^2}{2B^{(k)}} I_u^{(k)}(\alpha_p) + \frac{N_G^{(k)} \ln 2}{B^{(k)}}. \tag{22}$$

Thus, $I_u^{(k)}(\alpha_p)$ is linear in $(E_b / N_0)_{T;u}^{(k)}$ with slope $\frac{2B^{(k)}}{(N_G^{(k)})^2 (\ln 2)^2}$. Furthermore, if $I_u^{(k)}(\alpha_p) \to 0$ then $(E_b / N_0)_{T;u}^{(k)} \to (E_b / N_0)_{u;\min}^{(k)} \triangleq \frac{N_G^{(k)} \ln 2}{B^{(k)}}$ (which corresponds to $-1.6 + 10\log_{10}(N_G^{(k)}) - 10\log_{10}(B^{(k)})\,\text{dB}$). Since $Tr[(\mathbf{C}^{(k)})^\dagger \cdot \mathbf{C}^{(k)}] = \sum_{i=1}^{N_G^{(k)}} Tr[(\mathbf{C}_i^{(k)})^\dagger \cdot \mathbf{C}_i^{(k)}]$, we have $B^{(k)} = \frac{N_G^{(k)}}{N_T} Tr[(\mathbf{C}^{(k)})^\dagger \cdot \mathbf{C}^{(k)}]$, leading to $(E_b / N_0)_{u;\min}^{(k)} = \frac{N_T \ln 2}{Tr[(\mathbf{C}^{(k)})^\dagger \cdot \mathbf{C}^{(k)}]}$ which is independent of the space-time modulation format. Hence, the slope of $I_u^{(k)}(\alpha_p)$ (which is equal to $\frac{2}{N_G^{(k)} (\ln 2)^2} \frac{Tr[(\mathbf{C}^{(k)})^\dagger \cdot \mathbf{C}^{(k)}]}{N_T}$ at $(E_b / N_0)_{u;\min}^{(k)}$ is decreasing with $N_G^{(k)}$, and we have that $I_u^{(k)}(\alpha_p)$ is maximum, intermediate and minimum with ATA, POTA and OTA, respectively, proving Conjecture 1 at low SNR.

One can also see that $\frac{B^{(k)}}{N_G^{(k)}} = \sum_{i=1}^{N_G^{(k)}} \sum_{l=1}^{n_i^{(k)}} \frac{\lambda_{i;l}^{(k)}}{N_T} \le \tilde{\lambda}^{(k)} \sum_{i=1}^{N_G^{(k)}} \frac{r_i^{(k)}}{N_T}$. Using $r_i^{(k)} \le n_i^{(k)}$ (and hence $\sum_{i=1}^{N_G^{(k)}} r_i^{(k)} \le N_T$),we have $\frac{B^{(k)}}{N_G^{(k)}} \le \tilde{\lambda}^{(k)}$ showing that $\frac{N_G^{(k)} \ln 2}{B^{(k)}} \ge \frac{\ln 2}{\tilde{\lambda}^{(k)}}$. Equality is achieved when all eigenvalues $\{\{\lambda_{i;l}^{(k)}\}_{l=1}^{n_i^{(k)}}\}_{i=1}^{N_G^{(k)}}$ are equal and $r_i^{(k)} = n_i^{(k)}$ for all $i \in \{1, \cdots, N_G^{(k)}\}$. Thus, at low SNR, $C^{(k)}(\alpha_p)$ grows with a steeper slope than $I_u^{(k)}(\alpha_p)$ with respect to $(E_b / N_0)_T^{(k)}$. Subsequently, we refer to the ratio $\frac{B^{(k)}}{N_G^{(k)}}$ as average eigenvalue for user k.

From (21) we have

$$\begin{cases} I_u^{(1)}(\alpha_p) \approx \frac{B^{(1)}}{N_G^{(1)} \ln 2} \frac{\alpha_p P_T}{N_0 D^{(1)}} \\ I_u^{(2)}(\alpha_p) \approx \frac{B^{(2)}}{N_G^{(2)} \ln 2} \frac{(1-\alpha_p) P_T}{N_0 D^{(2)}} \end{cases}$$

yielding a triangular rate region (in the positive quadrant of the two-dimensional plane defined by the axes $R^{(1)}$ and $R^{(2)}$) and bounded by the line

$$R^{(2)} = -\frac{B^{(2)}}{N_G^{(2)}} \frac{N_G^{(1)}}{B^{(1)}} \frac{D^{(1)}}{D^{(2)}} R^{(1)} + \frac{B^{(2)}}{N_G^{(2)} \ln 2} \frac{P_T}{N_0 D^{(2)}} \tag{23}$$

Since $\frac{B^{(k)}}{N_G^{(k)}} \le \tilde{\lambda}^{(k)}$, it can be seen from (11) and (23) that the capacity region bounded by (11) contains the one bounded by (23). Similar to (11) the segment (23) reduces to the point (0, 0) as $P_T / N_0 \to 0$. We distinguish the following cases:

1) If $\frac{B^{(2)} / N_G^{(2)}}{B^{(1)} / N_G^{(1)}} \frac{D^{(1)}}{D^{(2)}} > 1$, then (23) has a steeper slope

than −1. The right most support line with slope −1 intersects the rate region boundary at $(I_{u;s}^{(1)}, I_{u;s}^{(2)}) = (0, \frac{B^{(2)}}{N_G^{(2)}\ln 2}\frac{P_T}{N_0 D^{(2)}})$ yielding a maximum sum rate $I_{u;s} = \frac{B^{(2)}}{N_G^{(2)}\ln 2}\frac{P_T}{N_0 D^{(2)}}$ with power allocation $\alpha'_{p;s} = 0$.

2) If $\frac{B^{(2)}/N_G^{(2)}}{B^{(1)}/N_G^{(1)}}\frac{D^{(1)}}{D^{(2)}} < 1$, then the right most support line with slope −1 has a steeper slope than (23) and intersects the rate region boundary at $(I_{u;s}^{(1)}, I_{u;s}^{(2)}) = (\frac{B^{(1)}}{N_G^{(1)}\ln 2}\frac{P_T}{N_0 D^{(1)}}, 0)$, yielding $I_{u;s} = \frac{B^{(1)}}{N_G^{(1)}\ln 2}\frac{P_T}{N_0 D^{(1)}}$ with $\alpha'_{p;s} = 1$.

3) If $\frac{B^{(2)}/N_G^{(2)}}{D^{(2)}} = \frac{B^{(1)}/N_G^{(1)}}{D^{(1)}}$, the right most support line with slope −1 lies on the boundary of the triangular rate region (23), thus, maximizing the sum rate at every point. Hence, $\alpha'_{p;s}$ can take arbitrary value in [0, 1] and the maximum sum rate is $I_{u;s} = \frac{B^{(2)}}{N_G^{(2)}\ln 2}\frac{P_T}{N_0 D^{(2)}} = \frac{B^{(1)}}{N_G^{(1)}\ln 2}\frac{P_T}{N_0 D^{(1)}}$.

Comparison with Section 3 shows that at low SNR, the sum rate maximization is determined by the average eigenvalues $B^{(k)}/N_G^{(k)}$ when the channel is unknown at the transmitter, while being determined by the maximum eigenvalues $\tilde{\lambda}^{(k)}$ with known channel at the transmitter.

4.2. Rate Region at High SNR

For large P_T/N_0, (18) becomes $I_u^{(k)}(\alpha_p) \approx \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} \sum_{j=1}^{r_i^{(k)}} \log(\frac{1}{N_0}\frac{N_G^{(k)}P^{(k)}}{N_T D^{(k)}}\lambda_{i;j}^{(k)})$, and with $M^{(k)} = \sum_{i=1}^{N_G^{(k)}} \sum_{j=1}^{r_i^{(k)}} \log(\frac{N_G^{(k)}\lambda_{i;j}^{(k)}}{N_T})$ we have

$$I_u^{(k)}\left(\alpha_p\right) = \frac{M^{(k)}}{N_G^{(k)}} + \frac{r_T^{(k)}}{N_G^{(k)}}\log\left(\frac{P^{(k)}}{N_0 D^{(k)}}\right). \quad (24)$$

Using (20) and (24), we can write

$$\left(E_b/N_0\right)_{T;u}^{(k)} \approx \frac{2^{\left[\frac{N_G^{(k)} I_u^{(k)}(\alpha_p) - M^{(k)}}{r_T^{(k)}}\right]}}{I_u^{(k)}(\alpha_p)}. \quad (25)$$

As $I_u^{(k)}(\alpha_p) \to \infty, (E_b/N_0)_{T;u}^{(k)}$ increases exponentially, making $\left(E_b/N_0\right)_{T;u}^{(k)}$ dB linear in $I_u^{(k)}\left(\alpha_p\right)$ with slope $10\log_{10}\left(2\right)\frac{N_G^{(k)}}{r_T^{(k)}} \approx 3\frac{N_G^{(k)}}{r_T^{(k)}}$. Therefore, for large P_T/N_0 the rate of change of the SNR in dB with the transmission rate remains unchanged with and without channel knowledge at the transmitter. By using a proof similar to that of Theorem 2 with equation (24) instead of (12), we can easily prove the following theorem:

Theorem 4: As $P_T/N_0 \to \infty$, the rate region with uniform power allocation becomes rectangular, defined by the points $(0,0)$, $(0, \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$, $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), 0)$ and $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$.

Comparison with Theorem 2 shows that the limiting rectangle is the same with and without channel knowledge at the transmitter. From [32] and using the uniform power allocation $\sum^{(k)} = diag_{i=1}^{N_G^{(k)}}[\sum_i^{(k)}]$ with $\sum_i^{(k)}$ given in (17) and for a given α_p, the necessary and sufficient conditions for the equality of $C^{(k)}(\alpha_p)$ and $I_u^{(k)}(\alpha_p)$ at high SNR are that $r_i^{(k)} = n_i^{(k)}$ for all $i = 1, \cdots, N_G^{(k)}$, $k = 1, 2$. Since these conditions hold for every α_p, they are also necessary and sufficient for the equality of the capacity region $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ and the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ at high SNR, yielding the following theorem.

Theorem 5: Similar asymptotic capacity and rate regions are obtained with and without transmitter channel knowledge if and only if all users have full column rank TOG propagation matrices $r_i^{(k)} = n_i^{(k)}$, for all $i = 1, \cdots, N_G^{(k)}$ and $k = 1, 2$.

The conditions of Theorem 5 are satisfied whenever OTA is used, and if the channel propagation matrix has full column rank $(r_T^{(k)} = N_T)$ whenever ATA is used. Finally, one can show that the conditions of Theorem 3 reduce to those of Theorem 5 for high values of P_T/N_0.

Consider now the effect of the space-time modulation format. The term $\frac{r_T^{(k)}}{N_G^{(k)}}\log(P^{(k)}/N_0)$ is dominant in (24), and hence the impact of the space-time modulation

format is determined through the ratio $\frac{r_T^{(k)}}{N_G^{(k)}}$. Since $r_i^{(k)} = \min(n_i^{(k)}, L_t^{(k)} N_R^{(k)})\ \forall i$ for 3GPP channels [32], $\frac{r_T^{(k)}}{N_G^{(k)}}$ takes the values $r_{T;A}^{(k)} = \min(N_T, L_t^{(k)} N_R^{(k)})$ with ATA, $\frac{r_{T;P}^{(k)}}{N_G^{(k)}} = \sum_{i=1}^{N_G^{(k)}} \min(\frac{n_i^{(k)}}{N_G^{(k)}}, \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}})$ with POTA, and $\frac{r_{T;O}^{(k)}}{N_T} = \sum_{i=1}^{N_T} \min(\frac{1}{N_T}, \frac{L_t^{(k)} N_R^{(k)}}{N_T}) = 1$ with OTA. Since $\min(\frac{n_i^{(k)}}{N_G^{(k)}}, \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}}) \le \frac{n_i^{(k)}}{N_G^{(k)}}$, we have $\frac{r_{T;P}^{(k)}}{N_G^{(k)}} \le \sum_{i=1}^{N_G^{(k)}} \frac{n_i^{(k)}}{N_G^{(k)}} = \frac{N_T}{N_G^{(k)}}$. Similarly, using $\min(\frac{n_i^{(k)}}{N_G^{(k)}}, \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}}) \le \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}}$ we have $\frac{r_{T;P}^{(k)}}{N_G^{(k)}} \le \sum_{i=1}^{N_G^{(k)}} \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}} = L_t^{(k)} N_R^{(k)}$. It follows that $\frac{r_{T;P}^{(k)}}{N_G^{(k)}} \le \min(\frac{N_T}{N_G^{(k)}}, L_t^{(k)} N_R^{(k)}) \le r_{T;A}^{(k)}$. Also, using $\min(\frac{n_i^{(k)}}{N_G^{(k)}}, \frac{L_t^{(k)} N_R^{(k)}}{N_G^{(k)}}) \ge \frac{1}{N_G^{(k)}}$ one can see that $\frac{r_{T;P}^{(k)}}{N_G^{(k)}} \ge \sum_{i=1}^{N_G^{(k)}} \frac{1}{N_G^{(k)}} = 1 = \frac{r_{T;O}^{(k)}}{N_T}$. Thus, $\frac{r_T^{(k)}}{N_G^{(k)}}$ is decreasing with increasing $N_G^{(k)}$, and so is $I_u^{(k)}(\alpha_p)$. It follows easily that the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ is largest, intermediate and smallest when the users employ ATA, POTA and OTA, respectively. This proves Conjecture 1 at high SNR.

We now investigate the maximum sum rate with uniform power allocation at high SNR by considering the following three cases:

1) $0 < \alpha_p < 1$: Using (24) with $P^{(1)} = \alpha_p P_T$ and $P^{(2)} = (1-\alpha_p) P_T$, we can write

$$\begin{aligned} \frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)} &= \frac{dI_u^{(2)}(\alpha_p)}{d\alpha_p} \frac{d\alpha_p}{dI_u^{(1)}(\alpha_p)} \\ &= -\frac{r_T^{(2)}}{N_G^{(2)}} \frac{N_G^{(1)}}{r_T^{(1)}} \frac{\alpha_p}{(1-\alpha_p)} \end{aligned} \tag{26}$$

which exists, thus, making $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ differentiable for all $\alpha_p \in (0,1)$. The power allocation maximizing the sum rate can be obtained by solving $[\frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)}]_{\alpha_p = \alpha'_{p;s}} = -1$ yielding

$$\alpha'_{p;s} = \frac{N_G^{(2)}}{r_T^{(2)}} \frac{r_T^{(1)}}{N_G^{(1)}} [1 + \frac{N_G^{(2)}}{r_T^{(2)}} \frac{r_T^{(1)}}{N_G^{(1)}}]^{-1} \tag{27}$$

which is identical to (16). Furthermore, using Theorem 4 it can be easily shown that $(I_{u;s}^{(1)}, I_{u;s}^{(2)}) = (\frac{r_T^{(1)}}{N_G^{(1)}} \log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}} \log(\frac{P_T}{N_0}))$, thus, yielding $\frac{I_{u;s}}{\log(P_T / N_0)} \to (\frac{r_T^{(1)}}{N_G^{(1)}} + \frac{r_T^{(2)}}{N_G^{(2)}})$ as $P_T / N_0 \to \infty$. Similarly to the capacity region with known channel at the transmitter, both $I_{u;s}$ and $\alpha'_{p;s}$ depend on the users space-time modulation formats as well as the ranks of the TOGs propagation matrices without being dependent on the channel eigenvalues $\lambda_{i;l}^{(k)}$.

2) $\alpha_p = 0$: From (26) the right derivative is such that $[\frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)}]^+_{\alpha_p=0} \to 0$ when $P_T / N_0 \to \infty$. The support line at $(I_u^{(1)}(0), I_u^{(2)}(0))$ is not unique and has a slope in the range $[[\frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)}]^+_{\alpha_p=0}, \infty)$ which does not include -1. As in Subsection 3.2 we can show in this case that $\alpha'_{p;s} \ne 0$.

3) $\alpha_p = 1$: From (26) we have $[\frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)}]^-_{\alpha_p=1} \to -\infty$ when $P_T / N_0 \to \infty$. The support line at $(I_u^{(1)}(1), I_u^{(2)}(1))$ is not unique, with a slope in the set $(-\infty, [\frac{dI_u^{(2)}(\alpha_p)}{dI_u^{(1)}(\alpha_p)}]^-_{\alpha_p=1}] \cup [0, \infty)$ that does not include -1. Hence also in this case $\alpha'_{p;s} \ne 1$.

5. Numerical Results

For numerical calculations, we assume equal number of antenna elements per TOG for both users, $n_G^{(k)} = \frac{N_T}{N_G^{(k)}}$. We also assume $d_T = d_R^{(k)} = \frac{\lambda}{2}$, $N_T = 4$ and $G_{l;s}^{(k)} = 1$, for all $s \in \{0, \cdots, S_l^{(k)} - 1\}$, $l \in \{1, \cdots, L_t^{(k)}\}$ and $k = 1, 2$.

We consider the 3GPP spatial channel model of Figure 3 from the standardization document [25] with the following parameters for user $k = 1, 2$:

$\theta_{BS}^{(k)}$: The angle of the line-of-sight (LOS) direction between the base station and user k with respect to the antenna array normal at the transmitter.

$\theta_{MS}^{(k)}$: The angle of the LOS direction between user k and the base station with respect to the antenna array normal at the receiver.

$\overline{\theta}_{T;l}^{(k)}$: The mean AOD of cluster l.

$\overline{\theta}_{R;l}^{(k)}$: The mean AOA of cluster l

$\hat{\theta}_{T;l;s}^{(k)}$: The offset AOD from $\overline{\theta}_{T;l}^{(k)}$ of path s in cluster l.

$\hat{\theta}_{R;l;s}^{(k)}$: The offset AOA from $\overline{\theta}_{R;l}^{(k)}$ of path s in cluster l.

Hence, the AOD and AOA of path s in cluster l are given by $\theta_{T;l;s}^{(k)} = \theta_{BS}^{(k)} + \overline{\theta}_{T;l}^{(k)} + \hat{\theta}_{T;l;s}^{(k)}$ and, $\theta_{R;l;s}^{(k)} = \theta_{MS}^{(k)} + \overline{\theta}_{R;l}^{(k)} + \hat{\theta}_{R;l;s}^{(k)}$, respectively. For numerical results, we fix the mean AODs and mean AOAs for all clusters and $k = 1, 2$. The values of the offset AODs and AOAs are chosen from the simulation model presented in [25]. We consider a macrocell environment with a root mean square (RMS) angle spread of 2° at the base station and RMS angle spread of 35° at the receiver with the following characterization:

User one: $D^{(1)} = 4$, $\theta_{BS}^{(1)} = 20^o$ and $\theta_{MS}^{(1)} = -15^o$. Furthermore, we assume $L_t^{(1)} = 3$ with the following clustering structure:

$S_1^{(1)} = 3$, mean AOD $\overline{\theta}_{T;1}^{(1)} = 4^o$ and mean AOA $\overline{\theta}_{R;1}^{(1)} = 38^o$. The offset AODs and AOAs in degrees are given by $(\hat{\theta}_{T;1;1}^{(1)}, \hat{\theta}_{R;1;1}^{(1)}) = (0.2826, 4.9447)$, $(\hat{\theta}_{T;1;2}^{(1)}, \hat{\theta}_{R;1;2}^{(1)}) = (1.3594, 23.7899)$ and $\left(\hat{\theta}_{T;1;3}^{(1)}, \hat{\theta}_{R;1;3}^{(1)}\right) = \left(-3.0389, -53.1816\right)$.

$S_2^{(1)} = 2$, mean AOD $\overline{\theta}_{T;2}^{(1)} = 2^o$ and mean AOA $\overline{\theta}_{R;2}^{(1)} = 10^o$. The offset AODs and AOAs in degrees are given by $\left(\hat{\theta}_{T;2;1}^{(1)}, \hat{\theta}_{R;2;1}^{(1)}\right) = \left(-0.4984, -8.7224\right)$ and $\left(\hat{\theta}_{T;2;2}^{(1)}, \hat{\theta}_{R;2;2}^{(1)}\right) = \left(4.3101, 75.4274\right)$.

$S_3^{(1)} = 1$, mean AOD $\overline{\theta}_{T;3}^{(1)} = -3^o$ and mean AOA $\overline{\theta}_{R;3}^{(1)} = -20^o$. The offset AOD and AOA in degrees are given by $\left(\hat{\theta}_{T;3;1}^{(1)}, \hat{\theta}_{R;3;1}^{(1)}\right) = \left(1.0257, 17.9492\right)$.

User two: $D^{(2)} = 4, \theta_{BS}^{(2)} = -10^o$ and $\theta_{MS}^{(2)} = 5^o$. Furthermore, we assume $L_t^{(2)} = 2$ with the following clustering structure:

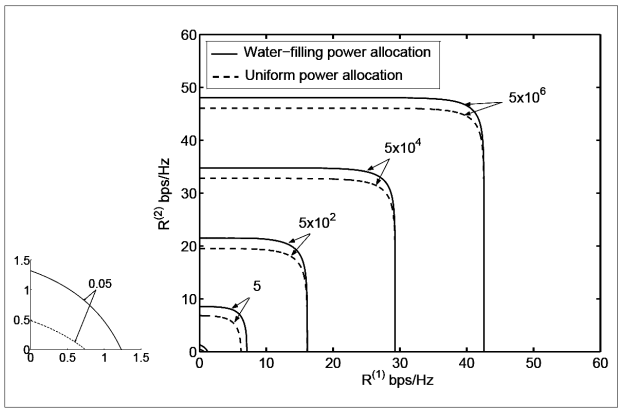

Figure 4. The capacity and uniform power allocation rate regions for two-user orthogonally multiplexed MIMO broadcast channel with user one employing POTA and user two employing ATA and $\frac{P_T}{N_0 D^{(1)}} = 0.05, 5, 5\times10^2, 5\times10^4, 5\times10^6$.

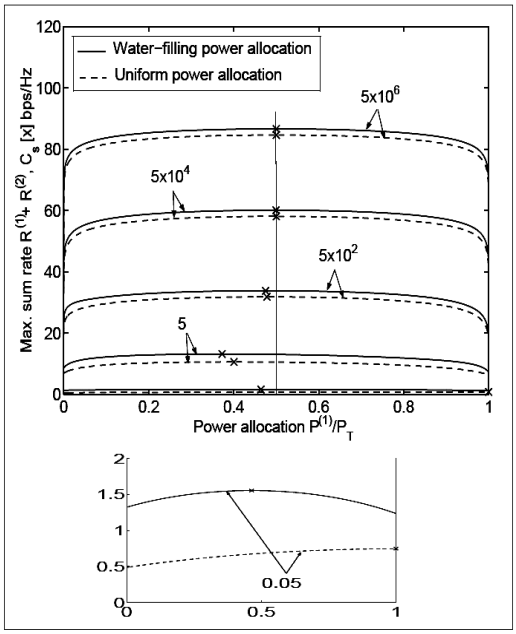

Figure 5. Maximum sum rate for the two-user orthogonally multiplexed MIMO broadcast channel with user one employing POTA and user two employing ATA and $\frac{P_T}{N_0 D^{(1)}} = 0.05, 5, 5\times10^2, 5\times10^4, 5\times10^6$. The line $\alpha_{p,s} = 0.5$ denotes the power allocation (16), (27) that maximizes the sum rate at high SNR and "x" denote the maximum sum rate points.

$S_1^{(2)} = 2$, mean AOD $\overline{\theta}_{T;1}^{(2)} = 3^o$ and mean AOA $\overline{\theta}_{R;1}^{(2)} = 17^o$. The offset AODs and AOAs in degrees are given by $(\hat{\theta}_{T;1;1}^{(2)}, \hat{\theta}_{R;1;1}^{(2)}) = (-0.0894, -1.5649)$ and $(\hat{\theta}_{T;1;2}^{(2)}, \hat{\theta}_{R;1;2}^{(2)}) = (1.7688, 30.9538)$.

$S_2^{(2)} = 2$, mean AOD $\bar{\theta}_{T;2}^{(2)} = -4.5^o$ and mean AOA $\bar{\theta}_{R;2}^{(2)} = -25^o$. The offset AODs and AOAs in degrees are given by $(\hat{\theta}_{T;2;1}^{(2)}, \hat{\theta}_{R;2;1}^{(2)}) = (-0.7431, -13.0045)$ and $(\hat{\theta}_{T;2;2}^{(2)}, \hat{\theta}_{R;2;2}^{(2)}) = (2.2961, 40.1824)$.

Effect of transmit power: The regions $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ are illustrated in Figure 4 for low, intermediate and high values of P_T / N_0 when user one employs POTA and user two employs ATA. We consider single antenna receivers for both users, $(N_R^{(1)}, N_R^{(2)}) = (1,1)$. Figure 4 illustrates the convergence of the capacity and rate regions with P_T / N_0 toward a rectangle, and that the equality at high SNR of $I_u^{(1)}(\alpha_p)$ and $C^{(1)}(\alpha_p)$ is satisfied since $r_i^{(1)} = n_i^{(1)} = 2, i = 1,2$, following Theorem 5. However, the rate regions with and without channel knowledge at the transmitter are not equal at high SNR because user two employs ATA, and $r_T^{(2)} = 2$ while $N_T = 4$. The corresponding maximum sum rate plots are presented in Figure 5 with respect to α_p, where "x" denotes the maximum sum capacity points. We see the convergence at high SNR of these points toward the line corresponding to (16), (27) given by $\alpha_{p;s} = 0.5$.

Effect of space-time modulation: The regions $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ are illustrated in Figure 6 for users employing space-time modulation formats similar to those of Figure 2. For this example, the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and capacity region $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ for OTA coincide. From Figure 6 we see that the largest capacity and rate regions are obtained when both users employ ATA, while POTA yields a smaller region, and the smallest regions are obtained with OTA. These results reinforce Theorem 1 and Conjecture 1.

Effect of the number of receive antennas: Figure 7 shows that increasing the number of receive antennas results in an expansion of the regions $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$.

Effect of multipath propagation: Figure 8 considers two users employing POTA with single antenna receivers $(N_R^{(1)}, N_R^{(2)}) = (1,1)$. When the total number of multipath components increases from $(S^{(1)}, S^{(2)}) = (3,1)$ to $(S^{(1)}, S^{(2)}) = (6,4)$ (where $(S^{(1)}, S^{(2)}) = (3,1)$ is obtained by considering the paths of the first time resolvable cluster for user one and the first path for user two), an expansion of the rate regions is observed, with and without channel knowledge at the transmitter.

In all figures, we see that the rate region with unknown channel at the transmitter is contained in the capacity region for channel knowledge at the transmitter.

6. Conclusions

This paper considered orthogonally multiplexed MIMO broadcast systems with multi-dimensional space-time modulation over a deterministic multipath additive Gaussian channel. We showed that the largest capacity region is achieved when each user employs all his signal dimensions on all transmit antennas (which corresponds to ATA space-time modulation format). The capacity

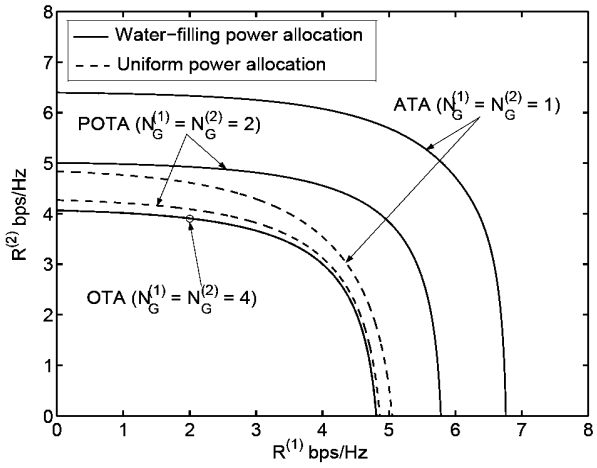

Figure 6. Two-user orthogonally multiplexed MIMO broadcast channel with users employing $(N_R^{(1)}, N_R^{(2)}) = (1,1)$ antennas and different space-time modulation formats with $\frac{P_T}{N_0 D^{(1)}} = 2$.

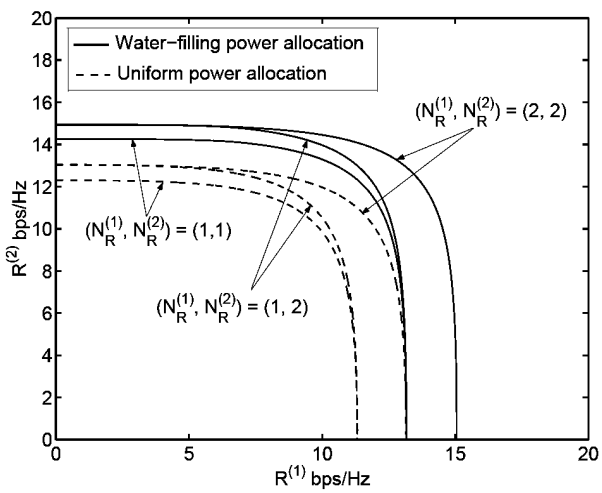

Figure 7. Two-user orthogonally multiplexed MIMO broadcast channel with users employing ATA and different numbers of antenna elements with $\frac{P_T}{N_0 D^{(1)}} = 40$.

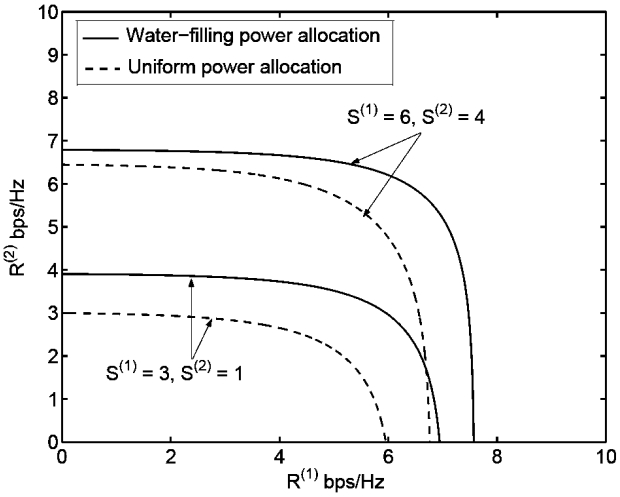

Figure 8. Two-user orthogonally multiplexed MIMO broadcast channel with users employing POTA with $(N_R^{(1)}, N_R^{(2)}) = (1,1)$ and different numbers of propagation paths with $\frac{P_T}{N_0 D^{(1)}} = 7$.

region with informed transmitter and the rate region with uninformed transmitter using a uniform power allocation are triangular at low SNR and become rectangular at high SNR. At high SNR these regions become the same if and only if all users have full column rank TOG propagation matrices.

We also investigated the power allocation among users that maximizes the sum rate, and provided explicit expressions for such power allocation and the corresponding maximum sum rate at low and high SNR. At high SNR the power allocation that maximizes the sum capacity is determined by the users' space-time modulation format and ranks of the TOGs propagation matrices. However, when the channel is known (respectively unknown) at the transmitter the sum rate is maximized at low SNR by an arbitrary power allocation between users if they have equal ratios of maximum (respectively average) eigenvalue to signal space dimensionality; otherwise it is maximized by allocating the total transmit power to one user only.

Numerical results for a two-user system using some examples from the 3GPP spatial channel model show that the capacity region with an informed transmitter and the rate region with an uninformed transmitter using a uniform power allocation expand when the number of transmit antennas per TOG or the number of receive antennas increases. Furthermore, these numerical results show that an increase in the number of multipath components leads to a rate region expansion with known and unknown channel at the transmitter.

In this paper we assumed that users do not share signal dimensions, resulting in an orthogonal multiplexed system without interference between users. Future work will explore systems where some dimensions are shared between users, and hence interference plays a major role.

Appendix

Proof of Theorem 2

Fix α_p and take $P_T / N_0 \to \infty$. We distinguish the following cases:

$\alpha_p = 0$: Since no power is allocated to user one $C^{(1)}(0) = 0$, and (12) yields $C^{(2)}(0) \to X^{(2)} + \frac{r_T^{(2)}}{N_G^{(2)}} \log(\frac{N_G^{(2)}}{r_T^{(2)}} \frac{P_T}{N_0 d^{(2)} D^{(2)}})$ with $X^{(2)}$ given in (13). Hence, $(\frac{C^{(1)}(0)}{\log(P_T / N_0)}, \frac{C^{(2)}(0)}{\log(P_T / N_0)}) \to (0, \frac{r_T^{(2)}}{N_G^{(2)}})$

$\alpha_p = 1$: Similarly to the previous case, we can show that $(\frac{C^{(1)}(1)}{\log(P_T / N_0)}, \frac{C^{(2)}(1)}{\log(P_T / N_0)}) \to (\frac{r_T^{(1)}}{N_G^{(1)}}, 0)$.

$0 < \alpha_p < 1$: Using (12) we have $C^{(k)}(\alpha_p) \to X^{(k)} + \frac{r_T^{(k)}}{N_G^{(k)}} \log(\frac{N_G^{(k)}}{r_T^{(k)}} \frac{P^{(k)}}{N_0 D^{(k)} d^{(k)}})$, leading to $[(\frac{C^{(1)}(\alpha_p)}{\log(P_T / N_0)}, \frac{C^{(2)}(\alpha_p)}{\log(P_T / N_0)})]_{0<\alpha_p<1} \to (\frac{r_T^{(1)}}{N_G^{(1)}}, \frac{r_T^{(2)}}{N_G^{(2)}})$. Thus, the rate of change of both $C^{(1)}(\alpha_p)$ and $C^{(2)}(\alpha_p)$ with P_T / N_0 in dB is independent of α_p. In the limit, all the points $[(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))]_{0<\alpha_p<1}$ converge toward the intersection point of the capacity region boundary with the function

$$R^{(2)} = \frac{r_T^{(2)}}{N_G^{(2)}} \frac{N_G^{(1)}}{r_T^{(1)}} R^{(1)} \tag{28}$$

in the two-dimensional plane defined by the axes $R^{(1)}$ and $R^{(2)}$. The capacity region boundary is characterized by three points $(0, \frac{r_T^{(2)}}{N_G^{(2)}} \log(\frac{P_T}{N_0})), (\frac{r_T^{(1)}}{N_G^{(1)}} \log(\frac{P_T}{N_0}), 0)$ and $(\frac{r_T^{(1)}}{N_G^{(1)}} \log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}} \log(\frac{P_T}{N_0}))$. Since, by definition, the capacity region is the closure of all the set of achievable rate pairs, it follows that its boundary contains all the points $(R^{(1)}, \frac{r_T^{(2)}}{N_G^{(2)}} \log(\frac{P_T}{N_0}))$ with $0 \le R^{(1)} \le \frac{r_T^{(1)}}{N_G^{(1)}} \log(\frac{P_T}{N_0})$,

and $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), R^{(2)})$ with $0 \le R^{(2)} \le \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0})$. Consequently, the limiting capacity region is a rectangle with lower left corner (0, 0) and upper right corner $(\frac{r_T^{(1)}}{N_G^{(1)}}\log(\frac{P_T}{N_0}), \frac{r_T^{(2)}}{N_G^{(2)}}\log(\frac{P_T}{N_0}))$.

Proof of Theorem 3

The proof of this theorem consists first in deriving the necessary and sufficient conditions for the equality of $I_u^{(k)}(\alpha_p)$ and $C^{(k)}(\alpha_p)$ for a given α_p. The necessary and sufficient conditions for equality of the regions $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ follow by varying the power allocation α_p in the interval [0, 1]. Assuming a fixed power allocation α_p, we have

<u>Lemma</u> 1: $C^{(k)}(\alpha_p) = I_u^{(k)}(\alpha_p)$ if and only if the corresponding TOGs propagation matrices have full column rank ($r_i^{(k)} = n_i^{(k)}$) and $\lambda_{i;l}^{(k)} = \lambda_i^{(k)}$ for all $l \in \{1, \cdots, n_i^{(k)}\}$ such that

$$(\lambda_i^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}} = \text{Constant}\ \forall i, \quad i = 1 \cdots, N_G^{(k)}. \tag{29}$$

Proof: 1) Assume that $C^{(k)}(\alpha_p) = I_u^{(k)}(\alpha_p)$. Due to the uniqueness of the input covariance matrix that maximizes the average mutual information [36], the covariance matrices associated with $C^{(k)}(\alpha_p)$ and $I_u^{(k)}(\alpha_p)$ must be equal. Hence, (6) and (7) yield

$$N_0 diag_{l=1}^{n_i^{(k)}}[\left\{\mu^{(k)} - \left(\lambda_{i;l}^{(k)}\right)^{-1}\right\}^+] = \frac{N_G^{(k)} P^{(k)}}{N_T D^{(k)}} \mathbf{I}_{n_i^{(k)}}. \tag{30}$$

From (30), $(\lambda_{i;l}^{(k)})^{-1} = \mu^{(k)} - \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}} \forall l$, leading to full rank TOG propagation matrices. Therefore, $r_i^{(k)} = n_i^{(k)}$ and all eigenmodes in TOG i are active. Furthermore, (30) also leads to

$$(\lambda_{i;l}^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}} = \mu^{(k)} = \text{Constant} \quad \forall l, \quad l = 1 \cdots, n_i^{(k)}.$$

It follows that $\lambda_{i;l}^{(k)} = \lambda_i^{(k)}$, and the water-filling constant becomes

$$\mu^{(k)} = (\lambda_i^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}} \qquad \forall i, \quad i = 1 \cdots, N_G^{(k)} \tag{31}$$

making $(\lambda_i^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}}$ equal for all TOGs.

2) Consider the channel of user k with $N_G^{(k)}$ TOGs, $r_i^{(k)} = n_i^{(k)}$ and $\lambda_{i;l}^{(k)} = \lambda_i^{(k)}$ for all $l \in \left\{1, \cdots, n_i^{(k)}\right\}$ such that (29) is satisfied. Equation (18) can be written as

$$I_u^{(k)}(\alpha_p) = \frac{N_T}{N_G^{(k)}} \log[\prod_{i=1}^{N_G^{(k)}} \prod_{l=1}^{n_i^{(k)}} ((\lambda_i^{(k)})^{-1} + \frac{1}{N_T} \frac{N_G^{(k)} P^{(k)}}{N_0 D^{(k)}})^{\frac{1}{N_T}}] + \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} n_i^{(k)} \log(\lambda_i^{(k)}).$$

Using the arithmetic-geometric inequality [37], we have

$$I_u^{(k)}(\alpha_p) \le \frac{N_T}{N_G^{(k)}} \log[\frac{1}{N_T} \sum_{i=1}^{N_G^{(k)}} \sum_{l=1}^{n_i^{(k)}} ((\lambda_i^{(k)})^{-1} + \frac{1}{N_T} \frac{N_G^{(k)} P^{(k)}}{N_0 D^{(k)}})] + \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} n_i^{(k)} \log(\lambda_i^{(k)})$$

with equality achieved since $(\lambda_i^{(k)})^{-1} + \frac{1}{N_T} \frac{N_G^{(k)} P^{(k)}}{N_0 D^{(k)}}$ is constant for all $i = 1, \cdots, N_G^{(k)}$. Hence,

$$I_u^{(k)}(\alpha_p) = \frac{N_T}{N_G^{(k)}} \log[\frac{1}{N_T} (\sum_{i=1}^{N_G^{(k)}} n_i^{(k)} (\lambda_i^{(k)})^{-1}) + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}}] + \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} n_i^{(k)} \log(\lambda_i^{(k)}). \tag{32}$$

If the channel is known at the transmitter, (29) shows that there exists a water-filling power allocation where all eigenmodes are active, and all TOGs contribute to the overall channel capacity with power $\frac{n_i^{(k)} P^{(k)}}{N_T}$ allocated to each TOG i, for $i = 1, \cdots, N_G^{(k)}$. Thus, the water-filling solution leading to the channel capacity results in a constant $\mu_p^{(k)} = (\lambda_i^{(k)})^{-1} + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}}$, leading to

$$C^{(k)}(\alpha_p) = \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} \sum_{l=1}^{n_i^{(k)}} \log(\mu_p^{(k)} \lambda_i^{(k)})$$

with $\sum_{i=1}^{N_G^{(k)}} \sum_{l=1}^{n_i^{(k)}} (\mu_p^{(k)} - (\lambda_i^{(k)})^{-1}) = \frac{N_G^{(k)} P^{(k)}}{N_0 D^{(k)}}$. Thus, we have $\mu_p^{(k)} = \frac{1}{N_T} (\sum_{i=1}^{N_G^{(k)}} n_i^{(k)} (\lambda_i^{(k)})^{-1}) + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}}$ and

$$C^{(k)}(\alpha_p) = \frac{N_T}{N_G^{(k)}} \log[\frac{1}{N_T}(\sum_{i=1}^{N_G^{(k)}} n_i^{(k)} (\lambda_i^{(k)})^{-1}) + \frac{N_G^{(k)} P^{(k)}}{N_T N_0 D^{(k)}}] + \frac{1}{N_G^{(k)}} \sum_{i=1}^{N_G^{(k)}} n_i^{(k)} \log(\lambda_i^{(k)}) \quad (33)$$

which is similar to (32) and, hence, yielding $C^{(k)}(\alpha_p) = I_u^{(k)}(\alpha_p)$.

This proves Lemma 1. Now we return to the proof of Theorem 3.

If $C^{(k)}(\alpha_p) = I_u^{(k)}(\alpha_p)$ for all $\alpha_p \in [0,1]$, then (7) and (19) are identical, and the equality of the rate region $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ and capacity region $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ is straightforward. If the regions are equal, then for any value of α_p the corresponding point $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p))$ cannot be outside the rectangle with lower left corner (0, 0) and upper right corner $(C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ since $I_u^{(k)}(\alpha_p) \le C^{(k)}(\alpha_p)$. In such case, since (7) and (19) coincide one can easily show that $(I_u^{(1)}(\alpha_p), I_u^{(2)}(\alpha_p)) = (C^{(1)}(\alpha_p), C^{(2)}(\alpha_p))$ using the fact that the capacity region is convex-$\cap$ with $C^{(2)}(\alpha_p)$ being monotonically decreasing with α_p and $C^{(1)}(\alpha_p)$. Consequently, the equality of the capacity regions is equivalent to the equality $C^{(k)}(\alpha_p) = I_u^{(k)}(\alpha_p)$ for all $\alpha_p \in [0,1]$.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," European Transactions on Telecommunications, Vol. 10, No. 6, pp. 585–595, November-December 1999.

[2] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," IEEE Transactions on Communications, Vol. 46, No. 3, pp. 357–366, March 1998.

[3] G. G. Raleigh and V. K. Jones, "Multivariate modulation and coding for wireless communication," IEEE Journal on Selected Areas in Communications, Vol. 17, No. 5, pp. 851–866, May 1999.

[4] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," IEEE Transactions on Communications, Vol. 50, No. 2, pp. 225–234, February 2002.

[5] W. Bliss, K. W. Forsythe, A. O. Hero III, and A. F. Yegulalp, "Environmental issues for MIMO capacity," IEEE Transactions on Signal Processing, Vol. 50, No. 9, pp. 2128–2142, September 2002.

[6] C-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," IEEE Transactions on Information Theory, Vol. 48, No. 3, pp. 637–650, March 2002.

[7] T. M. Cover, "Broadcast channels," IEEE Transactions on Information Theory, Vol. 18, No. 1, pp. 2–14, January 1972.

[8] T. M. Cover, "Comments on broadcast channels," IEEE Transactions on Information Theory, Vol. 44, No. 6, pp. 2524–2530, October 1998.

[9] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels - Part I: Ergodic capacity," IEEE Transactions on Information Theory, Vol. 47, No. 3, pp. 1083–1102, March 2001.

[10] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels - Part II: Outage capacity," IEEE Transactions on Information Theory, Vol. 47, No. 3, pp. 1103–1127, March 2001.

[11] G. Caire and S. Shamai, "On the multiple antenna broadcast channel," in Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, Vol. 2, pp. 1188–1193, November 2001.

[12] M. H. M. Costa, "Writing on dirty paper," IEEE Transactions on Information Theory, Vol. IT-29, No. 3, pp. 439–441, May 1983.

[13] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," IEEE Transactions on Information Theory, Vol. 49, No. 7, pp. 1691–1706, July 2003.

[14] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," IEEE Transactions on Information Theory, Vol. 50, No. 9, pp. 1875–1892, September 2004.

[15] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," IEEE Transactions on Information Theory, Vol. 49, No. 8, pp. 1912–1921, August 2003.

[16] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," IEEE Transactions on Information Theory, Vol. 49, No. 10, pp. 2658–2668, October 2003.

[17] S.Vishwanath, N. Jindal, and A. Goldsmith, "On the capacity of multiple input multiple output broadcast channels," in IEEE International Conference on Communications (ICC), Vol. 3, pp. 1444–1450, 2002.

[18] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," IEEE Journal on Selected Areas in Communications, Vol. 21, No. 5, pp. 684–702, June 2003.

[19] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," IEEE Transactions on Information Theory, Vol. 50, No. 5, pp. 768–783, May 2004.

[20] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," IEEE Transactions on Information Theory, Vol. 52, No. 9, pp. 3936–3964, Septem-

ber 2006.

[21] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," IEEE Transactions on Communications, Vol. 55, No. 1, pp. 11–15, January 2007.

[22] P. Tejera, W. Utschick, J. A. Nossek, and G. Bauch, "Rate balancing in multiuser MIMO OFDM systems," IEEE Transactions on Communications, Vol. 57, No. 5, pp. 1370–1380, May 2009.

[23] M. Kassouf and H. Leib, "Shannon capacity regions for orthogonally multiplexed MIMO broadcast channels with informed transmitters," in IEEE Wireless Communications and Networking Conference (WCNC), Vol. 1, pp. 357–362, 2004.

[24] M. Kassouf and H. Leib, "The capacity region of orthogonally multiplexed MIMO broadcast channels for open and closed loop systems," in Proceedings of Conference on Information Sciences and Systems (CISS), March 2004.

[25] 3GPP TR 25.996 V 7.0.0, "Spatial channel model for multiple input multiple output (MIMO) simulations (Release 7)," Technical specification group radio access network, 3rd Generation Partnership Project, June 2007.

[26] S. D. Blostein and H. Leib, "Multiple antenna systems: Their role and impact in future wireless access," IEEE Communications Magazine, Vol. 41, No. 7, pp. 94–101, July 2003.

[27] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," IEEE Transactions on Information Theory, Vol. 44, No. 2, pp. 744–765, March 1998.

[28] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE Journal on Selected Areas in Communications, Vol. 16, No. 8, pp. 1451–1458, October 1998.

[29] R. T. Derryberry, S. D. Gray, D. M. Ionescu, G. Mandyam, and B. Raghothaman, "Transmit diversity in 3G CDMA systems," IEEE Communications Magazine, Vol. 40, No. 4, pp. 68–75, April 2002.

[30] R. A. Soni and R. M. Buehrer, "On the performance of open-loop transmit diversity techniques for IS-2000 systems: A comparative study," IEEE Transactions on Wireless Communications, Vol. 3, No. 5, pp. 1602–1615, September 2004.

[31] T. M. Cover and J. A. Thomas, "Elements of information theory," John Wiley & Sons, Inc., 1991.

[32] M. Kassouf and H. Leib, "Information rates for multidimensional modulation over multiple antenna wireless channels," European Transactions on Telecommunications, Vol. 20, No. 5, pp. 463–481, August 2009.

[33] D. A. Harville, "Matrix algebra from a statistician's perspective," Springer-Verlag New York, Inc., 1997.

[34] R. Webster, Convexity, Oxford University Press INC., New York, 1994.

[35] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," IEEE Wireless Personal Communications, Vol. 6, No. 3, pp. 311–335, March 1998.

[36] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multiantenna Gaussian broadcast channels," IEEE Transactions on Information Theory, Vol. 51, No. 4, pp. 1570–1580, April 2005.

[37] I. S. Gradshteyn and I. M. Ryzhik, "Table of integrals, series, and products," Academic Press, corrected and enlarged edition, 1980.