Bounds on the Outage-Constrained Capacity Region of Space-Division Multiple-Access Radio Systems

Haipeng Jin

Center for Wireless Communications, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92097, USA
Email: jin@cwc.ucsd.edu

Anthony Acampora

Center for Wireless Communications, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92097, USA
Email: acampora@ece.ucsd.edu

Space-division multiple-access (SDMA) systems that employ multiple antenna elements at the base station can provide much higher capacity than single-antenna-element systems. A fundamental question to be addressed concerns the ultimate capacity region of an SDMA system wherein a number of mobile users, each constrained in power, try to communicate with the base station in a multipath fading environment. In this paper, we express the capacity limit as an outage region over the space of transmission rates $R_1, R_2, \ldots, R_n$ from the n mobile users. Any particular set of rates contained within this region can be transmitted with an outage probability smaller than some specified value. We find outer and inner bounds on the outage capacity region for the two-user case and extend these to multiple-user cases when possible. These bounds provide yardsticks against which the performance of any system can be compared.

Keywords and phrases: outage capacity, space-division multiple-access, capacity region.

1. INTRODUCTION

Time-varying multipath fading is a fundamental phenomenon affecting the availability of terrestrial radio systems, and strategies to abate or exploit multipath are crucial. Recent information theoretic research [1, 2] has shown that in most scattering environments, a multiple-antenna-element (MAE) array is a practical and effective technique to exploit the effect of multipath fading and achieve enormous capacity advantages.

Next-generation wireless systems are intended to provide high voice quality and high-rate data services. At the same time, the mobile units must be small and lightweight. It appears that base station complexity is the preferred strategy for meeting the requirements of the next-generation systems. In particular, MAE arrays can be installed at the base stations to provide higher capacity.

A primary question to be addressed is the ultimate capacity limit of a single cell with constrained user power and multipath fading. Since there are multiple users in the cell, the capacity limit is expressed as a region of allowable transmission rates such that information can be reliably transmitted by user 1 at rate R_1, user 2 at rate R_2, and so on. For a static channel condition with fixed fade depth at each mobile, the capacity region of the multiple-access channel is then the set of all rate vectors $\mathbf{R} = \{R_1, R_2, \ldots, R_n\}$ that can be achieved with arbitrarily small error probability [3, 4]. However, when the channel is time varying due to the dynamic nature of the wireless communication environment, the capacity region is characterized differently depending on the delay requirements of the mobiles and the coherence time of the channel fading. Two important notions are used in the dynamic channel case [2, 5]: ergodic capacity and outage capacity. Ergodic capacity is defined for channels with long-term delay constraint, meaning that the transmission time is long enough to reveal the long-term ergodic properties of the fading channel. The ergodic capacity is given by the appropriately averaged mutual information. In practical communication systems operating on fading channels, the ergodic assumption is not necessarily satisfied. For example, in the cases with real-time applications over wireless channels, stringent delay constraints are demanded and the ergodicity requirement cannot be fulfilled. No significant channel variability occurs during a transmission. There may be nonzero outage probability associated with any value of actual transmission rate, no matter how small. Here, we have to consider the information rate that can be maintained under all channel conditions, at least, within a certain outage probability [6, 7, 8, 9]. The maximal rate that can be achieved with a given outage

percentile p is defined as the p percent outage capacity. Both the ergodic capacity and outage capacity notions originated from single-user case. They are easily extended to cases with multiple users.

The outage capacity issue for single-user multiple-input multiple-output (MIMO) case was studied extensively. In [1], the authors characterized the outage capacity for a point-to-point MIMO channel subject to flat Rayleigh fading. The cumulative distribution functions for the outage capacity were presented such that given a specific outage probability, we will know at what rate information can be transmitted over the MIMO channel. Biglieri et al. [10] considered the outage capacity of a MIMO system for different delay and transmit power constraints.

However, there are limited results on the outage capacity region of multiple-access system with multiple antennas at the base station. Most previous studies are constrained to either fixed channel condition [11, 12] or ergodic capacity region [13]. The space-division multiple-access (SDMA) capacity regions under fixed channel conditions were considered in [11] with both independent decoding and joint decoding schemes. An iterative algorithm was proposed in [12] to maximize the sum capacity of a time invariant Gaussian MIMO multiple-access channel. The ergodic capacity region for MIMO multiple-access channel with covariance feedback has been studied in [13].

In this paper, we consider the outage capacity region for single-cell flat fading SDMA systems with MAEs at the base station and multiple mobiles, each with a single antenna element. From the outage capacity region, we can determine with what outage probability a certain rate vector can be transmitted with arbitrarily small error probabilities. Specifically, we derive outer and inner bounds on the outage capacity region for the two-user case and explain how the same principles can be extended to multiple-user case.

To find an outer bound on the capacity region, cooperation among the geographically separated mobile stations is assumed to take place via a virtual central processor, and there is a total power constraint. The total capacity is found for all combinations of $n, n-1, \ldots, 2, 1$ mobile stations in the system. For example, if there are two mobiles, the capacity of each, in isolation, is found, along with the combination of both, treated as having a single transmitter with two geographically remote antenna elements.

By definition, any realizable approach for which the capacity region can be found forms an achievable inner bound to the capacity region. Time sharing among users provides a simple inner bound [14]. We also derive a tighter inner bound by allowing users to transmit at the same time while performing joint decoding at the base station. It is noted that construction of the inner bound also provides a method for achieving the inner bound.

Most of our derivations and discussions will be focused on the two-user case since the results in this case can be easily displayed graphically and provide significant physical insight. However, along our way, we will point out wherever the results are extensible to cases involving more than two mobiles.
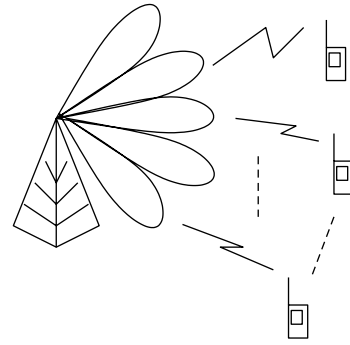

FIGURE 1: Space-division multiple-access systems.

This paper is organized as follows. Section 2 contains a description of the channel model which we use, and also introduces the concept of outage capacity region. Outer and inner bounds on the outage capacity region are derived in Section 3. Numerical results are presented in Section 4, along with discussions.

2. CHANNEL MODEL AND DEFINITION OF THE OUTAGE CAPACITY REGION

We consider a single-cell system where a number of geographically separated mobile users communicate with the base station. The system is shown in Figure 1. We consider the reverse link from the mobiles to the base station and model this as a multiple-access system. All mobile are equipped with single antenna element, and the base station is equipped with multiple receive antenna elements to exploit spatial diversity. We assume that the channel between each mobile station and each base station antenna element is subject to flat Rayleigh fading and that the fading between each mobile and each base station antenna element is independent of the fading between other mobile-base station pairs. Also present at each antenna element is additive white Gaussian noise (AWGN). Rayleigh fading between mobile j and base station antenna element i is represented by a zero mean, unit variance complex Gaussian random variable $H_{ij} = N(0, \sigma^2) + \sqrt{-1} N(0, \sigma^2)$, where $E|H_{ij}|^2 = 2\sigma^2 = 1$. The noise components observed at all the receiving antennas are identical, and independent white Gaussian distributed with power σ_n^2. Each mobile's transmitted power is limited such that the average received signal-to-noise ratio at each base station antenna element is ρ if only one mobile is transmitting.

Suppose there are n mobile stations and m antenna elements at the base station. Then the received signal can be represented as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is an m-element vector representing the received signals, $\mathbf{x}$ is a vector with n elements, each element representing the signal transmitted by one mobile station, $\mathbf{H}$ is the channel fading matrix with $m \times n$ complex Gaussian elements, and $\mathbf{n}$ is the received AWGN noise vector with covariance $\sigma_n^2 \mathbf{I}_{m \times m}$.

Consider first the simple case where the channel conditions are fixed, that is, $\mathbf{H}$ is constant. The capacity region of the multiple-access channel is then the set of all rate vectors $\mathbf{R} = \{R_1, R_2, \ldots, R_n\}$ satisfying [3]

$$\sum_{k \in T} R_k \leq I(\mathbf{y}; x_k, \ k \in T \mid x_l, \ l \in \bar{T}), \qquad (2)$$

where I stands for mutual information, T denotes any subset of $\{1, 2, \ldots, n\}$, $\bar{T}$ its complement, and ρ denotes the power limitation. In essence, the sum rate of any subset of the mobiles $\{1, 2, \ldots, n\}$ needs to be smaller than the mutual information between the transmitted and received signals if only the mobiles with the subset are transmitting. We can express the capacity region concisely as follows:

$$C(\mathbf{H}, \rho) = \left\{ \mathbf{R} \, \middle\| \, \sum_{k \in T} R_k \leq I(\mathbf{y}; x_k, \ k \in T \mid x_l, \ l \in \bar{T}) \right\}. \quad (3)$$

All the rate vectors $\mathbf{R}$ within the region $C(\mathbf{H}, \rho)$ can be achieved with arbitrarily small probability.

Now, since the channel matrix is a set of random variables, the capacity region is random and a key goal is to find the cumulative distribution functions of the regions, from which we can determine the probability that a specific rate vector can be transmitted. With this in mind, we define the outage capacity region as

$$\mathcal{C}(p, \rho) = \{\mathbf{R} \| \operatorname{Prob}(\mathcal{N}) \geq p\}, \qquad (4)$$

where $\mathcal{N}$ is defined as the set of the channel conditions under which the rate vector $\mathbf{R}$ is achievable with arbitrarily small error probability. The set $\mathcal{N}$ can be written as

$$\mathcal{N} = \{\mathbf{H} \| \mathbf{R} \in C(\mathbf{H}, \rho)\}. \qquad (5)$$

In (4), $1 - p$ is the outage probability and ρ represents the power limitation. The outage capacity region $\mathcal{C}(p, \rho)$ contains all the rate vectors that can be achieved with a probability greater than or equal to p. Alternatively speaking, the probability that a rate vector contained in $\mathcal{C}(p, \rho)$ cannot be achieved is less than $1 - p$.

To simplify notation, we write $\mathcal{C}(p, \rho)$ as $\mathcal{C}(p)$ and $C(\mathbf{H}, \rho)$ as $C(\mathbf{H})$, with the implication that the power limitation is always specified by ρ unless otherwise stated.

For the two-mobile case, we define $\mathbf{h}_1$ and $\mathbf{h}_2$ as the channel response vectors from mobiles 1 and 2, respectively, to the base station antenna elements, and they can be expressed as $\mathbf{h}_1 = \{H_{11}, H_{21}, \ldots, H_{m1}\}^T$ and $\mathbf{h}_2 = \{H_{12}, H_{22}, \ldots, H_{m2}\}^T$, where T denotes the transpose operation. As a result, the channel fading matrix $\mathbf{H} = [\mathbf{h}_1 \mathbf{h}_2]$. Then the outage capacity region is given as

$$\mathcal{C}(p) = \{(R_1, R_2) \| \operatorname{Prob}(\mathcal{N}) \geq p\}, \qquad (6)$$

where $\mathcal{N}$ is the set of channel response matrices which satisfy the following three conditions:

$$\begin{aligned} R_1 &\leq C_1(\mathbf{h}_1), \\ R_2 &\leq C_2(\mathbf{h}_2), \\ R_1 + R_2 &\leq C_{12}(\mathbf{H}), \end{aligned} \qquad (7)$$

where $C_1(\mathbf{h}_1)$, $C_2(\mathbf{h}_2)$, and $C_{12}(\mathbf{H})$ define the capacity region under the fading condition specified by $\mathbf{H}$. For convenience, we write this as follows:

$$\mathcal{N} = \left\{ \mathbf{H} \, \middle\| \, \begin{array}{l} R_1 \leq C_1(\mathbf{h}_1) \\ R_2 \leq C_2(\mathbf{h}_2) \\ R_1 + R_2 \leq C_{12}(\mathbf{H}) \end{array} \right\}. \qquad (8)$$

In (6), $\mathcal{C}(p)$ has the same interpretation as in (4); it consists of the rate pairs (R_1, R_2) simultaneously achievable with a probability greater than or equal to p, that is, the outage probability is smaller than or equal to $1 - p$.

3. OUTAGE CAPACITY BOUNDS

In this section, we derive bounds on the outage capacity region with a focus on the two-mobile case. The base station is equipped with m antenna elements. As we will see, most of our derivations are not constrained by the number of mobiles, and thus are applicable to cases with an arbitrary number of mobiles.

3.1. Outer bound

To obtain an outer bound on the outage capacity region, we start out by defining the following rate regions:

$$\begin{aligned} \mathcal{B}_1(p) &= \{(R_1, R_2) \| \operatorname{Prob}(\mathcal{M}_1) \geq p\}, \\ \mathcal{B}_2(p) &= \{(R_1, R_2) \| \operatorname{Prob}(\mathcal{M}_2) \geq p\}, \\ \mathcal{B}_a(p) &= \{(R_1, R_2) \| \operatorname{Prob}(\mathcal{M}_a) \geq p\}, \end{aligned} \qquad (9)$$

where $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_a$ are different sets of channel conditions. The three sets are defined by the following conditions, respectively:

$$\begin{aligned} \mathcal{M}_1 &= \{\mathbf{H} \| R_1 \leq C_1(\mathbf{h}_1)\}, \\ \mathcal{M}_2 &= \{\mathbf{H} \| R_2 \leq C_2(\mathbf{h}_2)\}, \\ \mathcal{M}_a &= \{\mathbf{H} \| R_1 + R_2 \leq C_{12}(\mathbf{H})\}. \end{aligned} \qquad (10)$$

The values $C_1(\mathbf{h}_1)$, $C_2(\mathbf{h}_2)$, and $C_{12}(\mathbf{H})$ are the same as those in (8); they jointly define the multiple-access capacity region under the fading condition $\mathbf{H}$. As a result of the definition, the rate pairs in $\mathcal{B}_1(p)$ and $\mathcal{B}_2(p)$ only satisfy the constraint on the individual rates R_1 and R_2, respectively; the rate pairs in $\mathcal{B}_a(p)$ only satisfy the constraint on the sum rate $R_1 + R_2$. We can now prove the following:

Claim 1. $\mathcal{C}(p) \subset \mathcal{B}_1(p)$, $\mathcal{C}(p) \subset \mathcal{B}_2(p)$, and $\mathcal{C}(p) \subset \mathcal{B}_a(p)$.
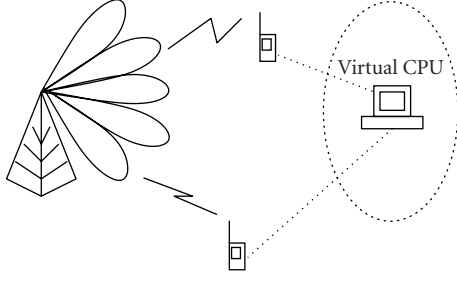
FIGURE 2: Space-division multiple-access systems with coordinated users.

Proof. For a given (R_1, R_2), if $\mathbf{H} \in \mathcal{N}$, where $\mathcal{N}$ is defined in (8), then $C_1(\mathbf{h}_1) \geq R_1$, $C_2(\mathbf{h}_2) \geq R_2$, and $C_{12}(\mathbf{H}) \geq R_1 + R_2$. By definition, $\mathbf{H} \in \mathcal{M}_1$, $\mathbf{H} \in \mathcal{M}_2$, and $\mathbf{H} \in \mathcal{M}_a$. As a result, $\mathcal{N} \subset \mathcal{M}_1$. This implies that $\text{Prob}(\mathcal{M}_1) \geq p$ if $\text{Prob}(\mathcal{N}) \geq p$. Therefore, $\mathcal{C}(p) \subset \mathcal{B}_1(p)$. Similarly, we can show that $\mathcal{C}(p) \subset \mathcal{B}_2(p)$ and $\mathcal{C}(p) \subset \mathcal{B}_a(p)$. $\square$

If a set is contained in each of several sets, then it is also contained in the intersection of those sets. Consequently, we can obtain the following outer bound for the outage capacity region.

Claim 2. $\mathcal{C}(p) \subset \mathcal{U}(p)$, *where* $\mathcal{U}(p) = \mathcal{B}_1(p) \cap \mathcal{B}_2(p) \cap \mathcal{B}_a(p)$.

In order to obtain the outer bound given in Claim 2, we need to evaluate $C_1(\mathbf{h}_1)$, $C_2(\mathbf{h}_2)$, and $C_{12}(\mathbf{H})$ under every specific fading condition $\mathbf{H}$. The sum capacity $C_{12}(\mathbf{H})$ is usually difficult to find. Fortunately, upper bounds on the sum capacity are easily obtained. We can use these upper bounds to find looser outer bounds on the outage capacity region that are easy to evaluate.

An upper bound $C'_{12}(\mathbf{H})$ on the sum capacity $C_{12}(\mathbf{H})$ can be obtained by assuming that both users are connected via some error free channel to a central coordinator as shown in Figure 2. We also assume that the virtual transmitter formed this way has perfect knowledge about the channel; thus singular value decomposition and water-filling techniques [2, 3] can be used to achieve the highest possible capacity. In water-filling, more power is allocated to better subchannels with higher signal-to-noise ratio so as to maximize the sum of data rates in all subchannels. If we define $\mathcal{N}'$ as

$$\mathcal{N}' = \left\{ \mathbf{H} \,\middle|\, \begin{array}{l} R_1 \leq C_1(\mathbf{h}_1) \\ R_2 \leq C_2(\mathbf{h}_2) \\ R_1 + R_2 \leq C_w(\mathbf{H}) \end{array} \right\}, \tag{11}$$

where $C_w(\mathbf{H})$ is the water-filling capacity under channel condition $\mathbf{H}$ and it is always greater than the actual sum capacity $C_{12}(\mathbf{H})$, it follows that the set $\mathcal{N}$ defined in (8) is always a subset of $\mathcal{N}'$. Consequently, we can use $\mathcal{N}'$ to define the following outer bound on the outage capacity region:

$$\mathcal{C}'(p) = \{(R_1, R_2) \,\|\, \text{Prob}(\mathcal{N}') \geq p\} \supset \mathcal{C}(p). \tag{12}$$

Now we define the following region:

$$\begin{aligned} \mathcal{B}'_a(p) &= \{(R_1, R_2) \,\|\, \text{Prob}(\mathcal{M}'_a) \geq p\}, \\ \mathcal{M}'_a &= \{\mathbf{H} \,\|\, R_1 + R_2 \leq C_w(\mathbf{H})\}. \end{aligned} \tag{13}$$

Region $\mathcal{B}'_a(p)$ contains all the rate pairs whose sum rates are constrained by the water-filling capacity. Following the steps used to prove Claim 1, the following readily shown.

Claim 3. $\mathcal{C}'(p) \subset \mathcal{B}_1(p)$, $\mathcal{C}'(p) \subset \mathcal{B}_2(p)$, *and* $\mathcal{C}'(p) \subset \mathcal{B}'_a(p)$.

Consequently, the following claim provides an outer bound on the outage capacity region.

Claim 4. $\mathcal{C}(p) \subset \mathcal{C}'(p) \subset \{\mathcal{B}_1(p) \cap \mathcal{B}_2(p) \cap \mathcal{B}'_a(p)\}$.

The outer bounds given in both Claims 2 and 4 are very easy to evaluate since the outer bounds consist of a set of regions defined by straight lines.

The above derivation can be used to an outer bound for multiple-user outage capacity region. The only difference is that the outer bound on the outage capacity region will be defined by a series of planes rather than straight lines, and is therefore in the shape of a polyhedra instead of a polygon as in the two-user case. Each plane will correspond to an outer bound on the capacity of one combination of users chosen from the entire set of users. For example, if there are three mobile users, then we can find the bounding rate regions for each of the following combinations of the users: {1}, {2}, {3}, {1, 2}, {1, 3}, {2, 3}, and {1, 2, 3}, and then take their intersection as the outer bound as we have done in Claims 2 and 4.

In deriving the outer bound in Claim 4, we assumed that the mobiles are coordinated by a central processing unit, and the channel condition is known at both the base station and the virtual coordinated transmitter. Thus, for our outer bound, the SDMA system is reduced to a point-to-point MIMO system. For such a system, it has been shown [2] that the forward and reverse channels are reciprocal and have the same capacity. Therefore, the outage capacity region outer bound given by Claim 4 for the multiple-access channel is also a bound for the broadcast channel.

3.2. Time-share bound

We now turn our attention to obtaining inner bounds on the outage capacity region for the two-mobile case. As we have said previously, any realizable approach for which the capacity region can be found forms an achievable inner bound to the capacity region. One such inner bound is the time-share bound, attained by time-sharing the base station between the two mobiles [14]. For every fading state $(\mathbf{h}_1, \mathbf{h}_2)$, if only mobile 1 is allowed to transmit, then it can achieve capacity $C(\mathbf{h}_1)$. Similarly, if only mobile 2 is allowed to transmit, it can achieve capacity $C(\mathbf{h}_2)$. An achievable time-share capacity region is given by $\{(R_1, R_2) \,\|\, R_1 \leq aC(\mathbf{h}_1),\ R_2 \leq (1-a)C(\mathbf{h}_2),\ 0 \leq a \leq 1\}$.

We define an outage capacity region

$$\mathcal{C}_{\mathcal{TS}}(p) = \{(R_1, R_2) \,||\, \text{Prob}(\mathcal{S}) \geq p\}, \tag{14}$$

where

$$\mathcal{S} = \{\mathbf{H} \,||\, R_1 \leq aC(\mathbf{h}_1), \; R_2 \leq (1-a)C(\mathbf{h}_2), \; 0 \leq a \leq 1\}. \tag{15}$$

Then, $\mathcal{C}_{\mathcal{TS}}(p)$ contains all the rate pairs that can be achieved by time sharing with an outage probability smaller than $1 - p$. Thus, $\mathcal{C}_{\mathcal{TS}}(p)$ is an inner bound on the outage capacity region defined by (6) since the time-sharing capacity region is an achievable region, and an achievable is, by definition, an inner bound on the actual capacity region. The following claim reiterates this observation.

Claim 5. $\mathcal{C}_{\mathcal{TS}}(p) \subset \mathcal{C}(p)$.

The boundary of the region $\mathcal{C}_{\mathcal{TS}}(p)$ is defined by all (R_1, R_2) pairs that can be achieved with an outage probability exactly equal to p, as expressed in the following condition:

$$\begin{aligned}
p &= \text{Prob}\left(\{\mathbf{H} \,||\, R_1 \leq aC(\mathbf{h}_1), \right. \\
&\qquad\qquad \left. R_2 \leq (1-a)C(\mathbf{h}_2), \; 0 \leq a \leq 1\}\right) \\
&= \text{Prob}\left(\{\mathbf{h}_1 \,||\, R_1 \leq aC(\mathbf{h}_1)\}\right) \\
&\qquad\qquad \times \text{Prob}\left(\{\mathbf{h}_2 \,||\, R_2 \leq (1-a)C(\mathbf{h}_2)\}\right),
\end{aligned} \tag{16}$$

since $\mathbf{h}_1$ and $\mathbf{h}_2$ are independent random vectors.

We now show how the boundary given in (16) can be derived in closed form. Each value $C(\mathbf{h}_1)$ and $C(\mathbf{h}_2)$ is the capacity of an AWGN channel with a single antenna element at the transmitter and m antenna elements at the receiver. Thus [1],

$$C(\mathbf{h}_i) = \log\left(1 + \rho |\mathbf{h}_i|^2\right), \tag{17}$$

where $|\mathbf{h}_i|^2 = \sum_{j=1}^{m} |H_{ij}|^2$ is a random variable following chi-square distribution with $2m$ degrees of freedom, and m is the number of receive antennas at the base station. The complementary cumulative distribution function $\bar{F}(\cdot)$ of $|\mathbf{h}_i|^2$ is given as [15] follows:

$$\bar{F}(t) = \text{Prob}\left(|\mathbf{h}_i|^2 > t\right) = \sum_{k=0}^{m-1} \frac{t^k e^{-t/2\sigma^2}}{k!(2\sigma^2)^k} = \sum_{k=0}^{m-1} \frac{t^k e^{-t}}{k!}. \tag{18}$$

We have $2\sigma^2 = 1$ in the above equation because the Rayleigh fading gain H_{ij} between mobile i and base station antenna element j is zero mean, unit variance complex Gaussian random variable, as specified in Section 2. As a result, the boundary of $\mathcal{L}_{\mathcal{TS}}(p)$ is given as follows:

$$\begin{aligned}
p &= \left\{\sum_{k=0}^{m-1} \frac{1}{k!}\left(\frac{e^{R_1/a} - 1}{\rho}\right)^k \exp\left(-\frac{e^{R_1/a} - 1}{\rho}\right)\right\} \\
&\quad \times \left\{\sum_{l=0}^{m-1} \frac{1}{l!}\left(\frac{e^{R_2/(1-a)} - 1}{\rho}\right)^l \exp\left(-\frac{e^{R_2/(1-a)} - 1}{\rho}\right)\right\} \\
&= \bar{F}\left(\frac{e^{R_1/a} - 1}{\rho}\right)\bar{F}\left(\frac{e^{R_2/(1-a)} - 1}{\rho}\right).
\end{aligned} \tag{19}$$

Given a certain probability p, the maximum R_1 can be found by setting R_2 to zero and a to 1 in (19). Then for every R_1 between the maximum and zero, we can always sweep out the possible R_2's by varying a between 0 and 1 and solving the equations numerically.

The same principle and derivations can be applied to multiple-user cases to obtain time-sharing bounds. The only difference is that the boundary will be defined by multiple rates and the condition in (19) will be given by the product of multiple complementary cumulative functions.

3.3. Joint decoding inner bound

The time-share inner bound may be quite pessimistic since one mobile may transmit at any time. A tighter inner bound may be obtained by allowing the two mobiles to transmit simultaneously, but without the coordinating virtual central processing unit that was used to obtain an outer bound. We now find such an inner bound by allowing the base station to jointly detect the information from both mobile stations. In this way, an achievable capacity region for SDMA under a particular channel condition $(\mathbf{h}_1, \mathbf{h}_2)$ is given by [3, 11]

$$C_1'(\mathbf{h}_1) = \log\left|1 + \rho\mathbf{h}_1^H\mathbf{h}_1\right| = \log\left|1 + \rho a_1\right|, \tag{20}$$

$$C_2'(\mathbf{h}_2) = \log\left|1 + \rho\mathbf{h}_2^H\mathbf{h}_2\right| = \log\left|1 + \rho a_2\right|, \tag{21}$$

$$\begin{aligned}
C_a'(\mathbf{H}) &= \log\left|\mathbf{I} + \rho[\mathbf{h}_1 \; \mathbf{h}_2]^H[\mathbf{h}_1 \; \mathbf{h}_2]\right| \\
&= \log\left(1 + \rho a_1 + \rho a_2 + \rho^2 a_1 a_2 - \rho^2 a_1 a_2 \cos^2\theta\right),
\end{aligned} \tag{22}$$

where

$$\begin{aligned}
a_1 &= |\mathbf{h}_1|^2, \\
a_2 &= |\mathbf{h}_2|^2, \\
\theta &= \cos^{-1}\left\{\frac{|\mathbf{h}_1^H \cdot \mathbf{h}_2|}{|\mathbf{h}_1| \cdot |\mathbf{h}_2|}\right\}.
\end{aligned} \tag{23}$$

The scalars a_1 and a_2 are the squared magnitude of the vectors $\mathbf{h}_1$ and $\mathbf{h}_2$, respectively, and θ is the angle between $\mathbf{h}_1$ and $\mathbf{h}_2$. The three scalars a_1, a_2, and θ are independently distributed random variables. a_1 and a_2 are chi-square distributed with $2m$ degrees of freedom and θ is uniformly distributed over $(0, 2\pi)$. Thus, their joint probability density function is given by the product of the individual probability density functions [16]:

$$\text{pdf}(a_1, a_2, \cos 2\theta) = g(a_1)g(a_2)h(\cos 2\theta), \tag{24}$$

where

$$g(x) = \frac{1}{\sigma^n 2^{n/2} \Gamma(n/2)} x^{n/2 - 1} e^{-x/2\sigma^2}, \quad 0 \geq x < \infty, \tag{25}$$

$$h(x) = \frac{1}{\pi\sqrt{1 - x^2}}, \quad |x| \leq 1. \tag{26}$$

In (25), the function $\Gamma(\cdot)$ is the Gamma function as defined in [15]. As with the time-sharing case, we can define a rate region based on this joint decoding achievable capacity region:

$$\mathcal{C}_{\mathcal{LB}}(p) = \{(R_1, R_2) \,||\, \text{Prob}(\mathcal{U}) \geq p\}, \tag{27}$$

where

$$\mathcal{U} = \left\{ \mathbf{H} \left\| \begin{array}{c} R_1 \le C_1'(\mathbf{h}_1) \\ R_2 \le C_2'(\mathbf{h}_2) \\ R_1 + R_2 \le C_a'(\mathbf{H}) \end{array} \right. \right\}. \quad (28)$$

The region $\mathcal{C}_{\mathcal{LB}}(p)$ contains all the rate pairs that fall into the achievable region with an outage probability smaller than $1 - p$. Again, since it is derived from an achievable region, $\mathcal{C}_{\mathcal{LB}}(p)$ forms an inner bound to the actual outage capacity region.

The boundary of the region $\mathcal{C}_{\mathcal{LB}}(p)$ is defined by all the rate pairs that can be achieved with a probability $\text{Prob}(\mathcal{U})$ exactly equal to p. Next, we show how this probability can be expressed in terms of R_1 and R_2. First, we define the following dummy variables:

$$t_1 = e^{C_1'(\mathbf{h}_1)} = 1 + \rho a_1,$$

$$t_2 = e^{C_2'(\mathbf{h}_2)} = 1 + \rho a_2,$$

$$t_a = e^{C_{12}'(\mathbf{H})} = 1 + \rho(a_1 + a_2) + \rho^2 a_1 a_2 - \rho^2 a_1 a_2 \cos^2 \theta \quad (29)$$

$$= 1 + \rho(a_1 + a_2) + \frac{\rho^2 a_1 a_2}{2} - \frac{\rho^2 a_1 a_2 \cos 2\theta}{2}.$$

The values of t_1, t_2, and t_a are functions of a_1, a_2, and θ, and their joint distribution is given by

$$f(t_1, t_2, t_a) = g(a_1)g(a_2)h(\cos 2\theta)|\mathbf{J}|$$

$$= g\left(\frac{t_1 - 1}{\rho}\right)g\left(\frac{t_2 - 1}{\rho}\right)$$

$$\times h\left(\frac{2(t_1 + t_2 + 1 - t_a)}{(t_1 - 1)(t_2 - 1)} + 1\right)\frac{2}{\rho^2(t_1 - 1)(t_2 - 1)}, \quad (30)$$

where the square matrix $\mathbf{J}$ is the Jacobian matrix of the transform from $(a_1, a_2, \cos 2\theta)$ to (t_1, t_2, t_a) given by

$$\mathbf{J} = \begin{bmatrix} \dfrac{1}{\rho} & 0 & 0 \\ 0 & \dfrac{1}{\rho} & 0 \\ \dfrac{2(t_a - t_2)}{(t_1 - 1)^2(t_2 - 1)} & \dfrac{2(t_a - t_1)}{(t_1 - 1)(t_2 - 1)^2} & \dfrac{-2}{(t_1 - 1)(t_2 - 1)} \end{bmatrix}. \quad (31)$$

As a result, the probability $\text{Prob}(\mathcal{U})$ defining the boundary of the capacity bound given by (27) can be evaluated as follows:

$$\text{Prob}(\mathcal{U}) = \text{Prob}(R_1 \le C_1', \ R_2 \le C_2', \ R_1 + R_2 \le C_a')$$

$$= \text{Prob}(e^{R_1} \le t_1, \ e^{R_2} \le t_2, \ e^{R_1 + R_2} \le t_a)$$

$$= \int_{e^{R_1}}^{\infty} dt_1 \int_{e^{R_2}}^{\infty} dt_2 \int_{e^{R_1 + R_2}}^{\infty} f(t_1, t_2, t_a) dt_a \quad (32)$$

$$= T_1 + T_2 + T_3,$$

where T_1, T_2, and T_3 in the above equation are defined as follows:

$$T_1 = \int_{e^{R_1}}^{a - e^{R_1}} dt_1 \int_{e^{R_2}}^{a - t_1} dt_2 \int_{e^{R_1 + R_2}}^{c} f(t_1, t_2, t_a) dt_a,$$

$$T_2 = \int_{e^{R_1}}^{a - e^{R_2}} dt_1 \int_{a - t_1}^{\infty} dt_2 \int_{b}^{c} f(t_1, t_2, t_a) dt_a, \quad (33)$$

$$T_3 = \int_{a - e^{R_2}}^{\infty} dt_1 \int_{e^{R_2}}^{\infty} dt_2 \int_{b}^{c} f(t_1, t_2, t_a) dt_a.$$

The variables a, b, and c in the above equations are defined as follows:

$$a = e^{R_1 + R_2} + 1,$$

$$b = t_1 + t_2 - 1, \quad (34)$$

$$c = t_1 + t_2 - 1 + (t_1 - 1)(t_2 - 1).$$

In the appendix, we will show that T_2 and T_3 are fairly easy to evaluate. Following the steps used to prove Claim 5, it can be shown the following.

Claim 6. $\mathcal{L}(p) = \{(R_1, R_2) \| T_2 + T_3 \ge p\} \subset \mathcal{C}_{\mathcal{LB}}(p) \subset \mathcal{C}(p).$

The rate region $\mathcal{L}(p)$ gives another inner bound on the outage capacity region. The boundary of the region $\mathcal{L}(p)$ is formed by the rate pairs that satisfy the following equation:

$$p = T_2 + T_3. \quad (35)$$

Given every possible R_1, we can trace out the corresponding R_2 by solving (35) numerically using the expressions obtained for T_2 and T_3 in the appendix.

Although the approach used in this section may, in principle, also be applied to multiple-mobile cases, the increased complexity needed to derive the joint probability distribution functions and evaluate the relevant probabilities prevents us from obtaining closed-form expressions for multiple-mobile cases similar to those shown in (A.2) and (35).

4. NUMERICAL RESULTS

The bounding techniques introduced in the previous section will now be applied to generate numerical results for inner and outer bounds on the outage capacity region for various values of number of antennas, outage percentage, and signal-to-noise ratio. Here, we focus our attention only on the two-mobile case since they are graphically friendly and offer significant physical insight.

Shown in Figure 3 is the outer bound on the outage capacity region with 2 and 16 antenna elements at the base station, plotted for outage probabilities of 1%, 10%, and 50%. In all cases, the signal-to-noise ratio ρ is 10 dB. Since these are outer bounds on the outage capacity region, no

(a)

(b)

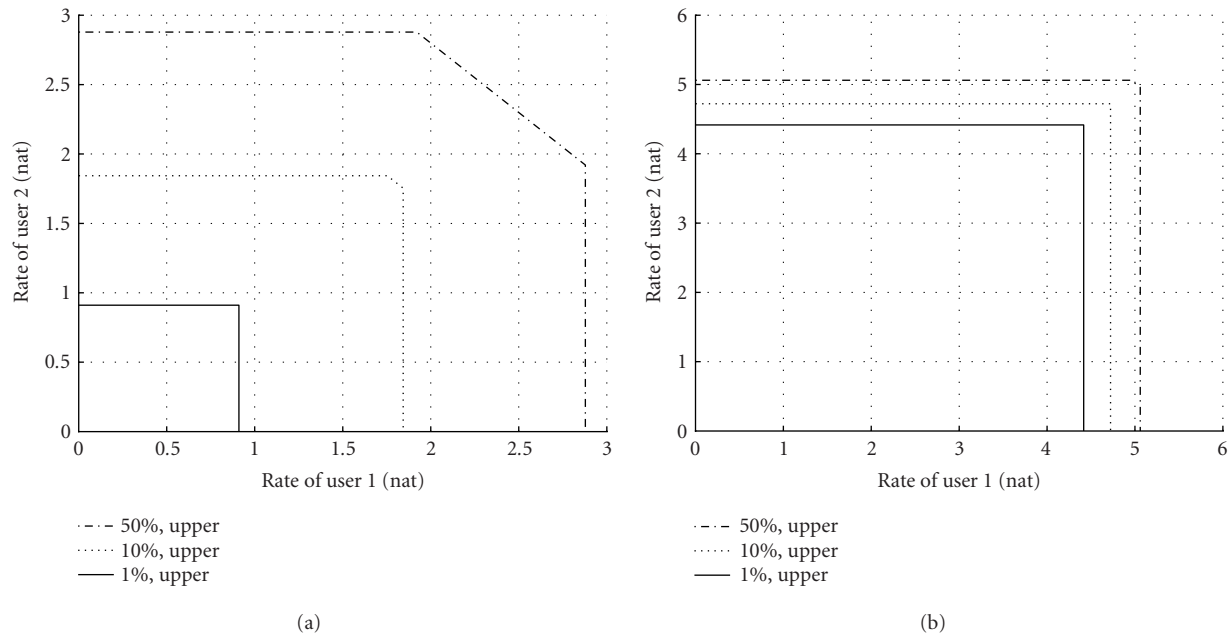FIGURE 3: Outer bound on outage capacity regions for (a) 2 and (b)16 antenna elements at the base station with SNR=10 dB for both cases.
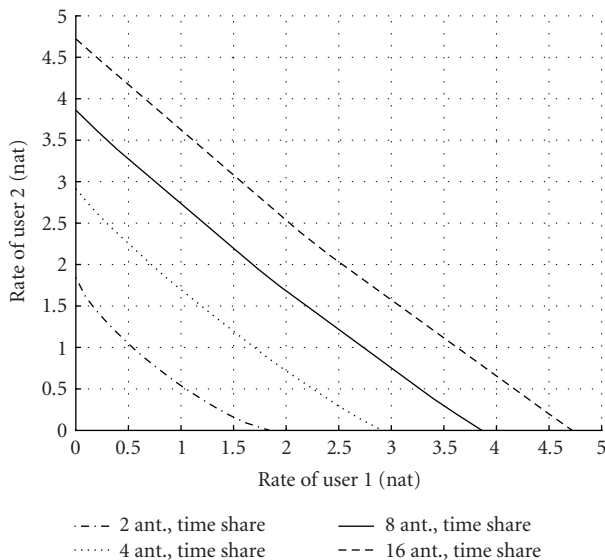

FIGURE 4: Time-share bound on 10% outage capacity region with SNR=10 dB.

rate pair outside the capacity region can ever be achieved with an outage probability smaller than the designated value p. Rate pairs inside the outer bound region may or may not be achievable with an outage probability smaller than p. For example, with 2 antenna elements at the base station, any rate pair with one rate higher than 0.9 nat per second (1 nat = 1.44 bites) has an outage probability greater than 1%. We note that the number of antenna elements has a very significant effect, not only on significantly enlarging

the outer bound on the capacity region but also in reducing the differences between low outage and high outage objectives.

Figure 4 shows the time-share bound for the 10% outage capacity region for base station with 2, 4, 8, and 16 antenna elements, respectively. Since these are inner bounds, all rate pairs within the bound can be achieved with an outage probability smaller than 10%. For example, with 2 antenna elements, both mobiles can transmit at 0.76 nat per second while achieve an outage probability smaller than 10%. With 16 antenna elements at the base station, both mobiles can transmit at 2.25 nat per second while achieving an outage probability of 10%. We notice that all the time-sharing bounds are concave.

When both users are allowed to transmit at the same time and joint decoding is performed at the base station, we get a tighter inner bound on the outage capacity region. Figure 5 shows the joint decoding inner bounds for the 10% and 1% outage capacity regions with 2, 4, 8, and 16 antenna elements at the base station. With 2 antenna elements at the base station, both mobiles can simultaneously transmit at 0.72 nat per second with an outage probability smaller than 1%; with 16 antenna elements at the base station, the mobiles can simultaneously transmit at 2.65 nat per second with an outage probability smaller than 1%. The inner bound given by Claim 6 is obviously much tighter than the time-sharing bound in Claim 5. Unfortunately, this joint decoding bound cannot be easily obtained for more than two mobiles. Once again, we notice that increasing the number of antenna elements greatly reduces the separation between the 1% outage and 10% outage results.

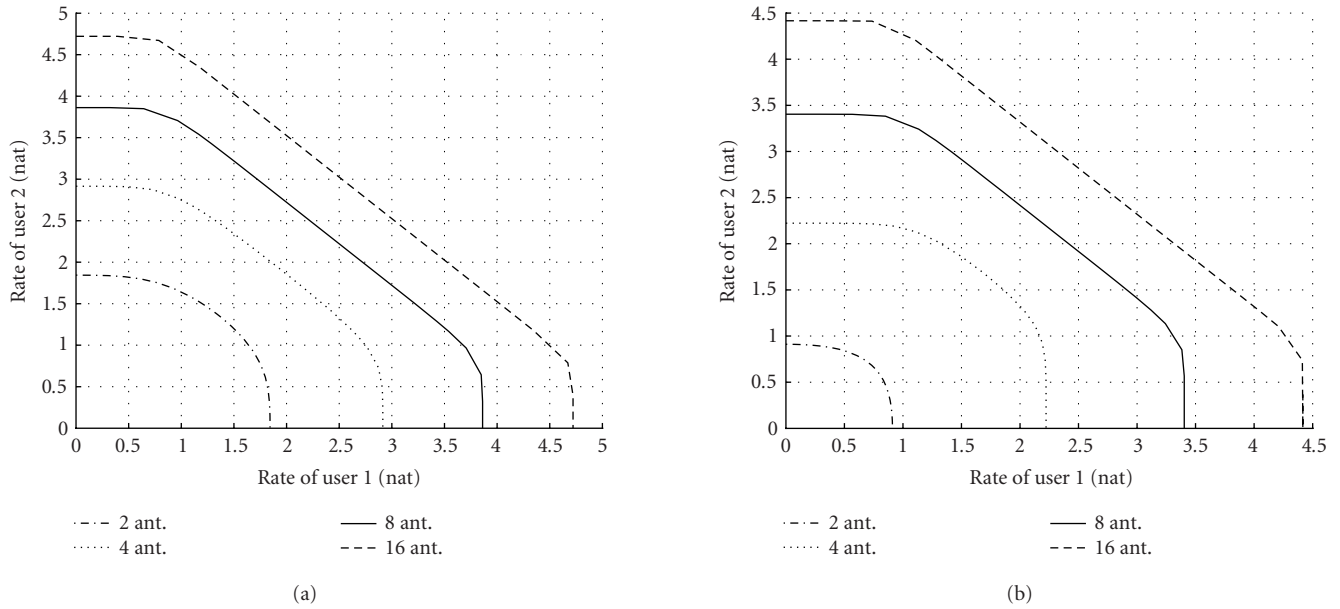Figure 6 shows both the outer and inner bounds on the

FIGURE 5: Inner bound on (a) 10% and (b) 1% outage capacity regions with different numbers of antenna elements at the base stations and where SNR=10 dB.
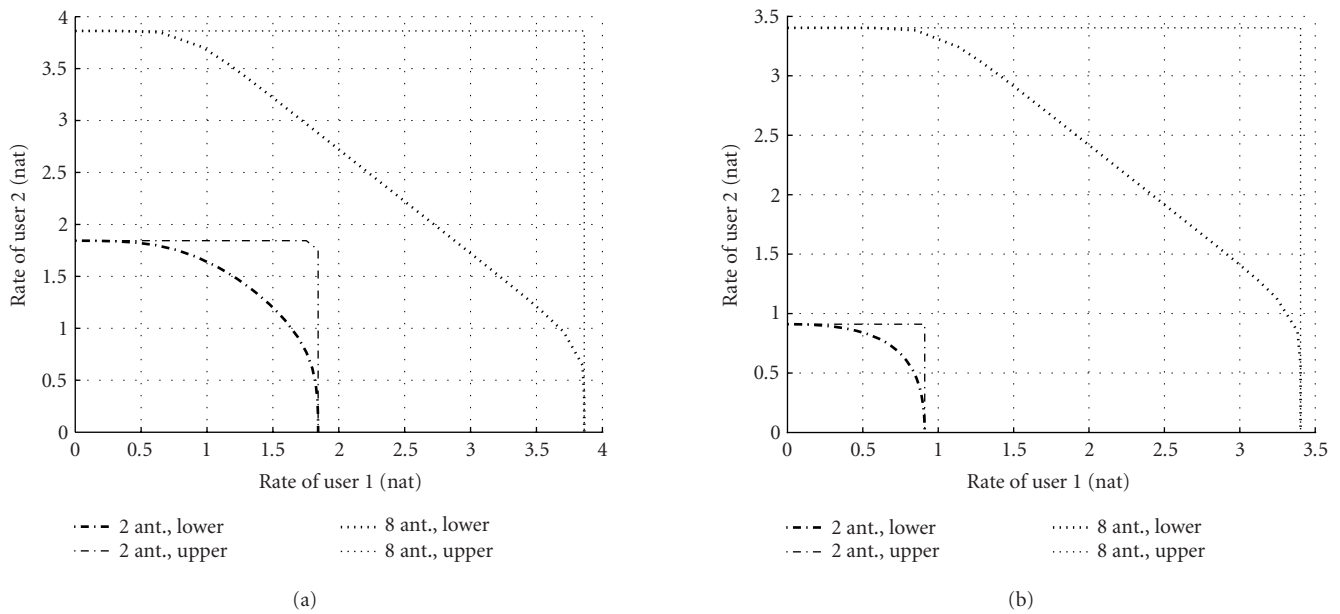

FIGURE 6: Both outer and inner bounds on (a) 10% and (b) 1% outage capacity regions for 2 and 8 antenna elements at the base stations with SNR=10 dB.

10% and 1% outage capacity regions with both 2 and 8 antenna elements. It can be seen from the plot that the inner bound is reasonably tight for the 2 antenna elements case. The bounds are very tight near the corners where one mobile is transmitting at its maximum allowable rate. We also notice that when the required outage probability is low, as in

the 1% case, there is more performance improvement by increasing the number of antenna elements than that when the required outage probability is high, as in the 10% case. This is no surprise since the cumulative distribution function of the allowable rates is much sharper when the number of antenna elements is large.
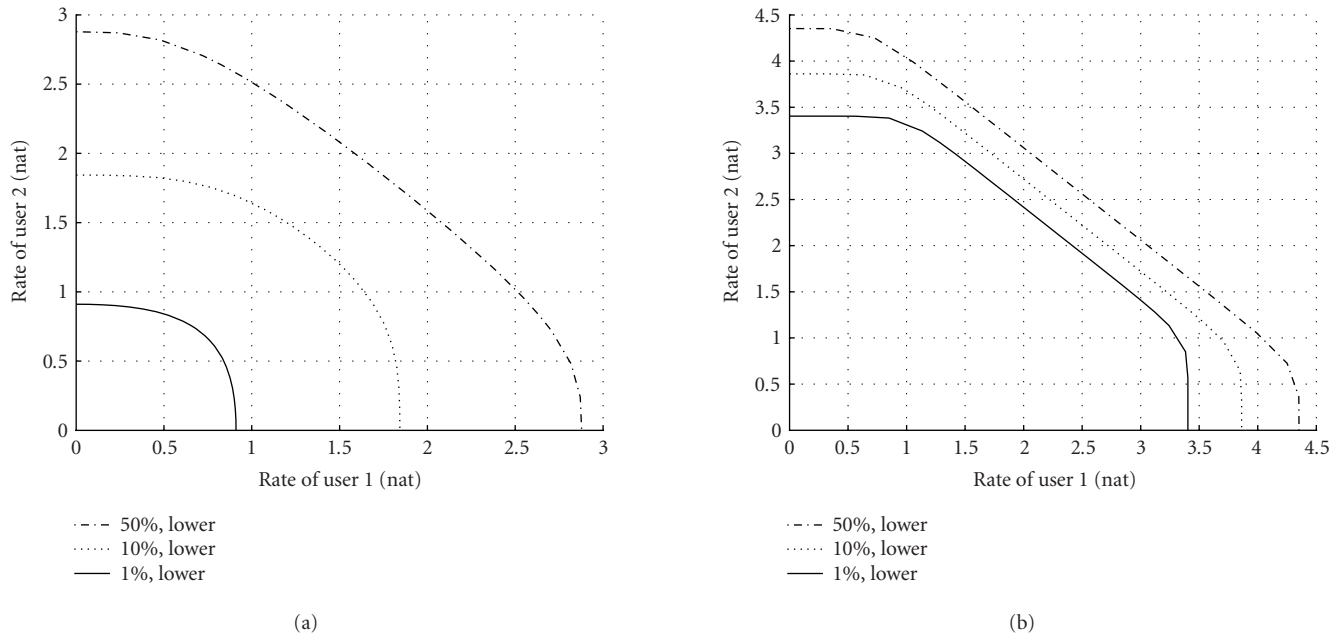
(a)

(b)

FIGURE 7: Inner bound on capacity region for (a) 2 and (b) 8 antenna elements under different outage probability with SNR=10 dB.
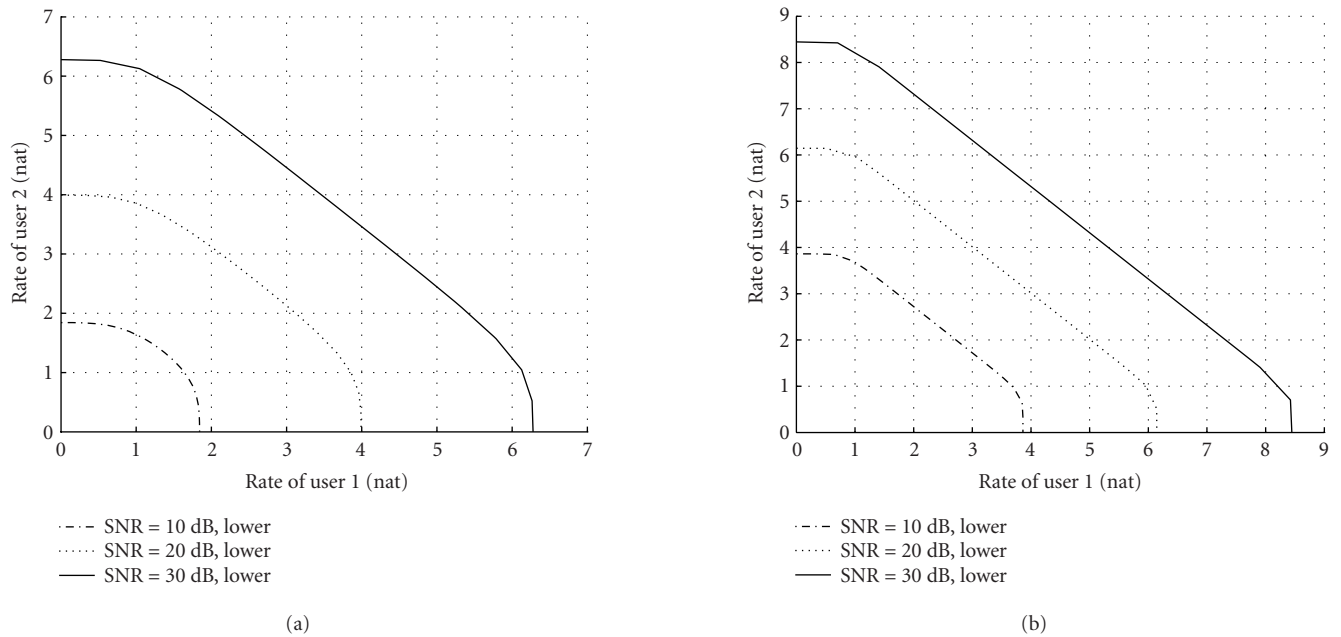

(a)

(b)

FIGURE 8: Inner bound on 10% outage capacity region under different SNRs for (a) 2 and (b) 8 antenna elements.

Figure 7 shows the tight inner bound for both 2 and 8 antenna elements at the base station with different outage probabilities. We can see from the plot that when the number of antenna elements is large, the outage capacity regions at different outage probabilities are not significantly different. This can be explained by noting that a large number of antenna elements at the base station is very efficient at combating severe fading, thereby keeping the allowable transmitting rates

relatively constant, and producing a sharper cumulative distribution function for the allowable rates.[1]

---

[1]Note that the capacity depends on chi-square distributed variables with mean M and variance $2M$ through a log operation in (17). As a result, the distance between different outage capacities is determined by the ratio of different percentage points along the chi-square CDF. Or, equivalently, the distance is determined by the normalized CDF curve.

Figure 8 shows the tight inner bound for the 10% outage capacity region with 2 or 8 antenna elements at the base station and different SNRs. From (22), we would expect, for large SNR and any fading condition, that the capacity region should increase linearly in each dimension as ρ increases linearly in dB. As a result, we would also expect the outage capacity region to increase linearly in each dimension as ρ increases. This trend is confirmed from the regions shown in Figure 8.

5. CONCLUSION

In this paper, we have studied the use of MAE array at the base station to increase system capacity. The fundamental question addressed is the ultimate capacity achievable with a MAE array equipped base station communicating with multiple mobile stations. Since there are multiple mobiles, the capacity is expressed as a region over the space of transmission rates from the mobiles. Any particular set of rates contained in the region can be transmitted with a certain outage probability.

We have obtained both outer and inner bounds for the outage capacity regions. The outer bounds indicates what is beyond the capability of the SDMA system while the inner bounds indicates what is achievable. As expected, the use of multiple antennas at the base station can greatly increase the allowable transmission rates from the mobiles. For example, with 2 and 16 antenna elements and joint decoding at the base station, two mobiles can simultaneously transmit at 0.72 nat per second and 2.65 nat per second, respectively, and achieve an outage probability smaller than 1%.

Our results show that although the capacity region can be expanded by allowing higher outage probability, the increase in allowable rates is much greater when the number of antenna elements is small. In cases with a large number of antenna elements at the base station, the strong protection against severe fading provided by the antenna array can keep the maximum allowable rates relatively constant. We also observe that the outage capacity regions can increase dramatically with an increase in the signal-to-noise ratio; the capacity regions expand almost linearly in every dimension as the signal-to-noise ratio increases.

The capacity region bounds derived in this paper provide a yardstick against which the performance of any space-division multiple-access technique can be compared.

APPENDIX

EVALUATION OF T_2 AND T_3 IN SECTION 3.3

Now we examine T_2 and T_3 more closely. It is easily verified that

$$
\begin{aligned}
\int_b^c h & \left(\frac{2(t_1 + t_2 + 1 - t_a)}{(t_1 - 1)(t_2 - 1)} + 1 \right) dt_a \\
&= \int_{-1}^1 h(x) \frac{(t_1 - 1)(t_2 - 1)}{2} dx \\
&= \frac{(t_1 - 1)(t_2 - 1)}{2}.
\end{aligned}
\tag{A.1}
$$

Using this result and the complementary cumulative distribution function for chi-square random variables given in (18), we can further simplify the expressions for T_2 and T_3:

$$
\begin{aligned}
T_3 &= \int_{a-e^{R_2}}^\infty dt_1 \int_{e^{R_2}}^\infty g\left(\frac{t_1 - 1}{\rho}\right) g\left(\frac{t_2 - 1}{\rho}\right) \frac{1}{\rho^2} dt_2 \\
&= \int_{(a-e^{R_2}-1)/\rho}^\infty g(x) dx \int_{(e^{R_2}-1)/\rho}^\infty g(y) dy \\
&= \bar{F}\left(\frac{a - e^{R_2} - 1}{\rho}\right) \bar{F}\left(\frac{e^{R_2} - 1}{\rho}\right),
\end{aligned}
$$

$$
\begin{aligned}
T_2 &= \int_{(e^{R_1}-1)/\rho}^{(a-e^{R_2}-1)/\rho} g(x) dx \int_{(a-\rho x-2)/\rho}^\infty g(y) dy \\
&= \int_{x_1}^{x_2} g(x) dx \bar{F}\left(\frac{a - 2}{\rho} - x\right) \\
&= \int_{x_1}^{x_2} g(x) \sum_{k=0}^{m-1} \frac{((a-2)/\rho - x)^k}{k!} e^{x-(a-2)/\rho} dx \\
&= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \int_{x_1}^{x_2} \frac{x^{m-1}((a-2)/\rho - x)^k}{(m-1)!k!} dx \\
&= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \int_{x_1}^{x_2} \frac{x^{m-1}}{(m-1)!k!} \sum_{l=0}^k \frac{k!}{l!(k-l)!} \left(\frac{a-2}{\rho}\right)^{k-l} (-x)^l dx \\
&= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \sum_{l=0}^k \left(\frac{a-2}{\rho}\right)^{k-l} \frac{(-1)^l}{l!(k-l)!(m-1)!} \int_{x_1}^{x_2} x^{m+l-1} dx \\
&= \sum_{k=0}^{m-1} e^{-(a-2)/\rho} \sum_{l=0}^k \left(\frac{a-2}{\rho}\right)^{k-l} \frac{(-1)^l}{l!(k-l)!(m-1)!} \frac{(x_2^{n+l} - x_1^{n+l})}{n+l} \\
&= \sum_{l=0}^{m-1} e^{-(a-2)/\rho} \frac{(-1)^l (x_2^{n+l} - x_1^{n+l})}{l!(m-1)!(n+l)} \sum_{k=0}^{m-1} \left(\frac{a-2}{\rho}\right)^{k-l} \frac{1}{(k-l)!} \\
&= \sum_{l=0}^{m-1} e^{-(a-2)/\rho} \frac{(-1)^l}{l!(m-1)!(n+l)} \left(\frac{e^{R_1} - 1}{\rho}\right)^{n+l} \left(e^{(n+l)R_2} - 1\right) \\
&\quad \times \sum_{k=0}^{m-1-l} \left(\frac{a-2}{\rho}\right)^k \frac{1}{k!}.
\end{aligned}
\tag{A.2}
$$

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[3] T. M. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[4] R. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, New York, NY, USA, 1968.

[5] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[6] D. N. C. Tse and S. V. Hanly, "Part I: Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.

[7] S. V. Hanly and D. N. C. Tse, "Part II: Multiaccess fading channels. II. Delay-limited capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2816–2831, 1998.

[8] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. I. Ergodic capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1083–1102, 2001.

[9] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels. II. Outage capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1103–1127, 2001.

[10] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1273–1289, 2001.

[11] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1468–1476, 1998.

[12] W. Yu, W. Rhee, and J. Cioffi, "Optimal power control in multiple access fading channels with multiple antennas," in *Proc. IEEE International Conference on Communications (ICC '01)*, pp. 575–579, Helsinki, Finland, June 2001.

[13] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Vector MAC capacity region with covariance feedback," in *Proc. IEEE International Symposium on Information Theory (ISIT '01)*, p. 54, Washington, DC, USA, June 2001.

[14] A. S. Acampora, "The ultimate capacity of frequency reuse communication satellites," *Bell System Technical Journal*, vol. 59, no. 7, pp. 1089–1122, 1980.

[15] J. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1995.

[16] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 1991.

Haipeng Jin received his B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1996. He received his M.S. and Ph.D. degrees in electrical engineering from the University of California, San Diego, in 2000 and 2003, respectively. He is currently working with the Standards Group, Qualcomm Inc., San Diego, Calif, where he participates in the research and development of wireless Internet standards.
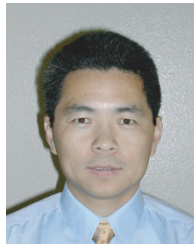
Anthony Acampora is a Professor of electrical and computer engineering at the University of California, San Diego, and is involved in numerous research projects addressing various issues at the leading edge of telecommunication networks. From 1995 through 1999, he was Director of UCSD's Center for Wireless Communications. Prior to joining the faculty at UCSD in 1995, he was a Professor of electrical engineering at Columbia University and Director of the Center for Telecommunications Research, a national engineering research center. He joined the faculty at Columbia in 1988 following a 20-year career at AT&T Bell Labs, most of which was spent in basic research as a contributing researcher and research manager. At Columbia, he was involved in research and education programs concerning broadband networks, wireless access networks, network management, optical networks, and multimedia applications. He received his Ph.D. degree in electrical engineering from the Polytechnic Institute of Brooklyn and is a Fellow of the IEEE. Professor Acampora has published over 160 papers, holds 30 patents, and has authored a textbook entitled *An Introduction to Broadband Networks*.