Capacity Region and Optimum Power Control Strategies for Fading Gaussian Multiple Access Channels With Common Data

Nan Liu and Sennur Ulukus

***Abstract*—A Gaussian multiple access channel (MAC) with common data is considered. Capacity region when there is no fading is known in an implicit form. We provide an explicit characterization of the capacity region and provide a simpler encoding/decoding scheme than that mentioned in work by Slepian and Wolf. Next, we give a characterization of the ergodic capacity region when there is fading, and both the transmitters and the receiver know the channel perfectly. Then, we characterize the optimum power allocation schemes that achieve arbitrary rate tuples on the boundary of the capacity region. Finally, we provide an iterative method for the numerical computation of the ergodic capacity region and the optimum power control strategies.**

***Index Terms*—Capacity region, common data, correlated data, fading channels, multiple access channel (MAC), power control.**

I. Introduction

CORRELATED data arises naturally in many applications of wireless communications. It arises mainly for three reasons: the observed data may be correlated (as in sensor networks) [4]–[7]; the correlated data may be created by communication between the transmitters (as in user cooperation diversity) [8], [9]; and the correlated data may result from decoding the data coming from previous stages of a larger network (as in relaying and multihopping in ad-hoc wireless networks) [10]–[13]. In this paper, we consider the transmission of correlated data in a multiple access channel (MAC). However, even in the simple MAC, finding capacity results for the transmission of arbitrarily correlated data is known to be extremely difficult [5], [14]–[17]. Therefore, in this paper, we constrain ourselves to a special kind of correlated data, correlated data in the sense of Slepian and Wolf [2], which we will call *common data*. In this MAC, the two transmitters each have their individual messages, which will be denoted by W_1 and W_2, respectively. Also, there is a common message W_0, which is known to both transmitters. All three messages are independent. The goal is to determine the rates R_0, R_1, and R_2, at which all three messages can be decoded with negligible error. The capacity will be a volume in the three-dimensional space. This model includes the traditional MAC as a special case, when $R_0 = 0$. It also includes the two-transmitter one-receiver point-to-point system as a special case, when $R_1 = R_2 = 0$, except that we have individual power constraints for the two transmit antennas here, instead of a single sum power constraint, as one would have in a point-to-point system [18].

Slepian and Wolf established the capacity region of the MAC with common data for discrete memoryless channels in [2]. Prelov and van der Meulen gave the capacity expression for a Gaussian MAC with common data in [3]. The characterization of the capacity region in [3] is implicit, in that the capacity region is expressed as a union of regions, and the boundary points on the capacity region are not determined explicitly. We first provide an explicit characterization of the capacity region and provide a simpler encoding/decoding scheme, compared with that mentioned in [2]; our encoding/decoding scheme is specially tailored for the Gaussian channel. We then concentrate on the case where there is fading in the channel and obtain a characterization of the ergodic capacity region. We also characterize the optimum power allocation schemes that achieve the rate tuples on the boundary of the capacity region. Finally, we provide an iterative method for the numerical computation of the ergodic capacity region, and the optimum power control strategies.

II. System Model

The Gaussian MAC we consider in this paper has two transmitters and one receiver. Without fading, the inputs and the output are related as

$$Y = X_1 + X_2 + Z \tag{1}$$

where Z is a Gaussian random variable with zero mean and unit variance. Transmitters 1 and 2 are subject to power constraints $\bar{P}_1$ and $\bar{P}_2$, respectively. We have three independent messages W_0, W_1, and W_2. Transmitter 1 knows W_0 and W_1, and transmitter 2 knows W_0 and W_2. Therefore, X_1 is a function of W_0, W_1, and X_2 is a function of W_0, W_2.

A rate triplet (R_1, R_2, R_0) is *achievable* if there exists a sequence of $((2^{nR_0} \times 2^{nR_1}, 2^{nR_0} \times 2^{nR_2}), n)$ codes with average probability of error approaching zero as n goes to infinity. Here, the probability of error is the probability that any of the three messages is decoded incorrectly. The *capacity region* is the closure of the set of achievable (R_1, R_2, R_0).

Paper approved by A. Host-Madsen, the Editor for Multiuser Communications of the IEEE Communications Society. Manuscript received June 3, 2005; revised March 10, 2006. This work was supported in part by the National Science Foundation under Grants ANI 02-05330, CCR 03-11311, CCF 04-47613, and CCF 05-14846, and in part by the Army Research Laboratory/Collaborative Technology Alliance (ARL/CTA) under Grant DAAD 19-01-2-0011. This paper was presented in part at the 42nd Annual Allerton Conference on Communications, Control, and Computing, Monticello, IL, September 2004.

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: nkancy@umd.edu; ulukus@umd.edu).

Digital Object Identifier 10.1109/TCOMM.2006.881370

With fading, the inputs and the output are related as

$$Y(k) = \sqrt{H_1(k)}X_1(k) + \sqrt{H_2(k)}X_2(k) + Z(k) \quad (2)$$

where $X_i(k)$ and $H_i(k)$ are the transmitted symbol and the fading process of user i, and $Z(k)$ is the zero-mean unit-variance Gaussian noise sample, at time k. $H_1(k)$ and $H_2(k)$ are jointly stationary and ergodic, and the stationary distribution has continuous density. The user signals are subject to average power constraints of $\bar{P}_1$ and $\bar{P}_2$. We assume that both the transmitters and the receiver know $H_1(k)$ and $H_2(k)$ for all k. The *ergodic capacity region* is the closure of the set of achievable rates in this scenario. For notational convenience, let $C(x) = (1/2)\log(1+x)$. All logarithms are defined with respect to base e.

III. Capacity Region Without Fading

The capacity region of the Gaussian MAC with common data is all triplets (R_1, R_2, R_0) [3]

$$R_1 \le C(\alpha\bar{P}_1) \quad (3)$$
$$R_2 \le C(\beta\bar{P}_2) \quad (4)$$
$$R_1 + R_2 \le C(\alpha\bar{P}_1 + \beta\bar{P}_2) \quad (5)$$
$$R_0 + R_1 + R_2 \le C\left(\bar{P}_1 + \bar{P}_2 + 2\sqrt{(1-\alpha)(1-\beta)\bar{P}_1\bar{P}_2}\right) \quad (6)$$

for some α and β such that $0 \le \alpha \le 1$ and $0 \le \beta \le 1$.

An alternative representation of the capacity region is obtained by defining $P_1 = \alpha\bar{P}_1$, $P_2 = \beta\bar{P}_2$. With these definitions, the capacity region is all triplets $[R_1, R_2, R_0]$ such that

$$R_1 \le C(P_1) \quad (7)$$
$$R_2 \le C(P_2) \quad (8)$$
$$R_1 + R_2 \le C(P_1 + P_2) \quad (9)$$
$$R_0 + R_1 + R_2 \le C(P_1 + P_2 + P_0) \quad (10)$$

for some $0 \le P_1 \le \bar{P}_1$, $0 \le P_2 \le \bar{P}_2$ and $P_0 = (\sqrt{\bar{P}_1 - P_1} + \sqrt{\bar{P}_2 - P_2})^2$. For fixed P_1, P_2, let $\mathcal{B}(P_1, P_2)$ denote the set of all rate triplets that satisfy (7)–(10). In the set $\mathcal{B}(P_1, P_2)$, certain points are of interest, which we define here: $Q = (0, 0, C(P_1 + P_2 + P_0))$, $S = (C(P_1), 0, C(P_1 + P_2 + P_0) - C(P_1))$, $T = (C(P_1), C(P_1 + P_2) - C(P_1), C(P_1 + P_2 + P_0) - C(P_1 + P_2))$ and the expressions for points V and U are the same as those for points S and T when the roles of users 1 and 2 are swapped. An example of $\mathcal{B}(P_1, P_2)$ and the corresponding points Q, S, T, U, V are shown in Fig. 1. The capacity region is the union of $\mathcal{B}(P_1, P_2)$ over all P_1, P_2 satisfying $0 \le P_1 \le \bar{P}_1$ and $0 \le P_2 \le \bar{P}_2$.

We can interpret the capacity region in (7)–(10) in the following way. Transmitter 1 spends power P_1 for transmitting its individual message W_1, and the remaining power $\bar{P}_1 - P_1$ for transmitting the common message W_0. Similarly, transmitter 2 spends power P_2 for transmitting its individual message W_2, and the remaining power $\bar{P}_2 - P_2$ for transmitting the common message. Since the common message is known to both transmitters, the effective received power for the common message is P_0, which may also be interpreted as the beamforming gain as in a two-transmitter one-receiver point-to-point system.

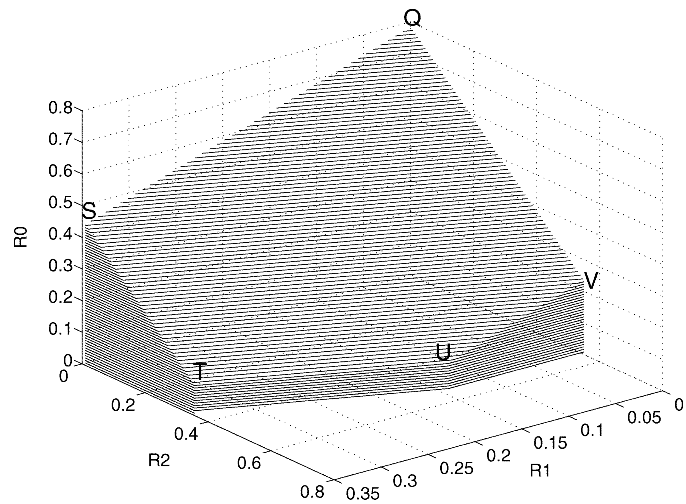


Fig. 1. $\mathcal{B}(P_1, P_2)$.

Both capacity region representations above are implicit, in the sense that one has to vary some variables in their valid intervals and take the union of regions corresponding to each valid allocation of these variables in order to obtain the capacity region. Next, we seek an explicit characterization of the capacity region. Let the rate pair (R_1, R_2) be such that it satisfies the conditions

$$R_1 \le C(\bar{P}_1), \quad R_2 \le C(\bar{P}_2), \quad R_1 + R_2 \le C(\bar{P}_1 + \bar{P}_2). \quad (11)$$

Let us define $c_1 = e^{2R_1} - 1$, $c_2 = e^{2R_2} - 1$, and $c = e^{2(R_1+R_2)} - 1$. Then, the powers P_1 and P_2 in representation (7)–(10) have to satisfy

$$P_1 \ge c_1, \quad P_2 \ge c_2, \quad P_1 + P_2 \ge c. \quad (12)$$

For a fixed pair (R_1, R_2), the largest possible R_0^* achievable is

$$R_0^* = \max_{P_1, P_2} C\left(\bar{P}_1 + \bar{P}_2 + 2\sqrt{(\bar{P}_1 - P_1)(\bar{P}_2 - P_2)}\right) - R_1 - R_2 \quad (13)$$

where the maximization in (13) is over all P_1, P_2 that satisfy (12). Note that (R_1, R_2, R_0^*) is on the boundary of the capacity region.

To solve the maximization problem in (13), it suffices to maximize $f(P_1, P_2) \triangleq (\bar{P}_1 - P_1)(\bar{P}_2 - P_2)$ subject to (12). Let P_1^* and P_2^* be the solution to this maximization problem. Then, (P_1^*, P_2^*) lies on the line $P_1 + P_2 = c$, since $f(P_1, P_2)$ is monotonically decreasing in both P_1 and P_2. Hence, it suffices to maximize $f(P_1, P_2)$ subject to the constraints that $P_1 + P_2 = c$ and $c_1 \le P_1 \le c - c_2$. Given that $P_1 + P_2 = c$, $f(P_1, P_2)$ becomes a quadratic form, and the validity of the following can be checked easily.

1) When $c_2 > (\bar{P}_2 - \bar{P}_1 + c)/2$

$$P_1^* = c - c_2, \quad P_2^* = c_2. \quad (14)$$

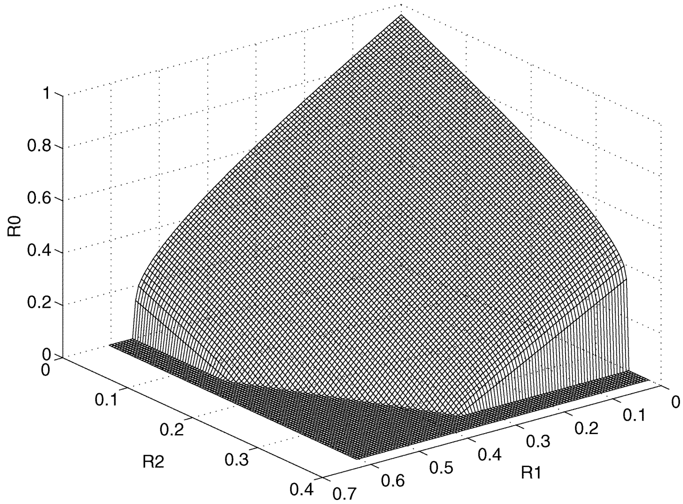


Fig. 2. Capacity region of the Gaussian MAC with common data.

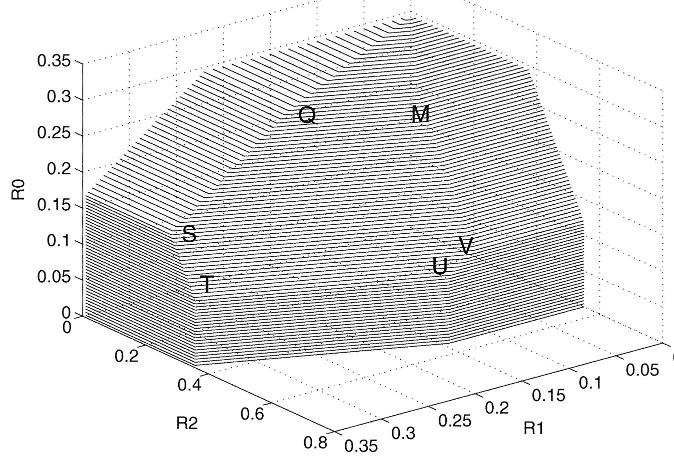


Fig. 3. $\mathcal{D}(P_1, P_2)$.

Moreover, point U on $\mathcal{B}(P_1^*, P_2^*)$ is the (R_1, R_2, R_0^*) point.

2) When $c_1 > (\bar{P}_1 - \bar{P}_2 + c)/2$

$$P_1^* = c_1, \quad P_2^* = c - c_1. \tag{15}$$

Moreover, point T on $\mathcal{B}(P_1^*, P_2^*)$ is the (R_1, R_2, R_0^*) point.

3) In all other cases

$$P_1^* = \frac{\bar{P}_1 - \bar{P}_2 + c}{2}, \quad P_2^* = \frac{\bar{P}_2 - \bar{P}_1 + c}{2}. \tag{16}$$

Moreover, some point on the line segment TU of $\mathcal{B}(P_1^*, P_2^*)$ is the (R_1, R_2, R_0^*) point.

This characterization is explicit, because for a fixed-rate pair (R_1, R_2), we can calculate R_0^* such that (R_1, R_2, R_0^*) is on the boundary of the capacity region. With this characterization, we can easily plot the capacity region of the Gaussian MAC with common data. An example is shown in Fig. 2 with $\bar{P}_1 = 2$ and $\bar{P}_2 = 1$.

It is interesting to note that all points on the capacity region are achieved by some point on the line segment TU of $\mathcal{B}(P_1, P_2)$ for some $0 \le P_1 \le \bar{P}_1$, $0 \le P_2 \le \bar{P}_2$. All other points of $\mathcal{B}(P_1, P_2)$, for example, points Q, S, and V are never on the boundary of the capacity region unless they coincide with point T or U.

Let us define $\mathcal{D}(P_1, P_2)$ to be the set of (R_1, R_2, R_0) such that

$$R_0 \le C(P_0) \tag{17}$$
$$R_1 \le C(P_1) \tag{18}$$
$$R_2 \le C(P_2) \tag{19}$$
$$R_0 + R_1 \le C(P_0 + P_1) \tag{20}$$
$$R_0 + R_2 \le C(P_0 + P_2) \tag{21}$$
$$R_1 + R_2 \le C(P_1 + P_2) \tag{22}$$
$$R_0 + R_1 + R_2 \le C(P_0 + P_1 + P_2) \tag{23}$$

for a fixed P_1, P_2, and $P_0 = (\sqrt{\bar{P}_1 - P_1} + \sqrt{\bar{P}_2 - P_2})^2$. In the set $\mathcal{D}(P_1, P_2)$, certain points are of interest, which we define here: Q, M, S, T, U, and V are the points (R_1, R_2, R_0) where [(17), (20), (23)], [(17), (21), (23)], [(18), (20), (23)], [(18), (22), (23)], [(19), (22), (23)], [(19), (21), (23)] are all satisfied with equality, respectively. An example of $\mathcal{D}(P_1, P_2)$ and the corresponding points Q, M, S, T, U, V are shown in Fig. 3.

Note that for any given P_1 and P_2, $\mathcal{D}(P_1, P_2)$ is a strict subset of $\mathcal{B}(P_1, P_2)$, since there are extra constraints involved in the definition of $\mathcal{D}(P_1, P_2)$. However, the capacity region of the Gaussian MAC with common data can also be written as the union of $\mathcal{D}(P_1, P_2)$ over all $0 \le P_1 \le \bar{P}_1$ and $0 \le P_2 \le \bar{P}_2$. This is because the coordinates of the points on line segment TU of $\mathcal{B}(P_1, P_2)$ are exactly the same as those on line segment TU of $\mathcal{D}(P_1, P_2)$. Since only the line segment TU appears on the final capacity region, the union of $\mathcal{D}(P_1, P_2)$ over all $0 \le P_1 \le \bar{P}_1$ and $0 \le P_2 \le \bar{P}_2$ gives the same capacity region.

$\mathcal{D}(P_1, P_2)$ is very similar to the capacity region of the three-user Gaussian MAC with independent messages. This suggests that encoding and decoding schemes similar to those of the three-user Gaussian MAC with independent messages can be used to achieve the points on the boundary of the capacity region of the Gaussian MAC with common data. To achieve a rate triplet (R_1, R_2, R_0) on the boundary of the capacity region, we first calculate P_1^*, P_2^* according to (14), (15), or (16). Depending on the values of (R_1, R_2, R_0), we want to achieve either point T or U, or some point on the line segment TU of region $\mathcal{D}(P_1^*, P_2^*)$. Points T and U can be achieved by successive decoding, and the remaining points on the line segment TU can be achieved by time sharing, just as in a three-user Gaussian MAC with independent messages.

More specifically, to achieve point T [similarly, point U], we generate three independent random codebooks C_0, C_1, and C_2, of sizes $(2^{nR_0'}, n)$, $(2^{nR_1'}, n)$, and $(2^{nR_2'}, n)$, respectively, where (R_1', R_2', R_0') is the coordinates of point T [similarly, point U]. Each entry of these codebooks is generated according to a zero-mean, unit-variance Gaussian random variable. When the messages to be transmitted are $W_0 = w_0$, $W_1 = w_1$, and $W_2 = w_2$, transmitter 1 transmits the sum of the w_0th row of

C_0 scaled by $\sqrt{\bar{P}_1 - P_1^*}$ and the w_1th row of C_1 scaled by $\sqrt{P_1^*}$, and transmitter 2 transmits the sum of the w_0th row of C_0 scaled by $\sqrt{\bar{P}_2 - P_2^*}$ and the w_2th row of C_2 scaled by $\sqrt{P_2^*}$. The effective received power for W_0, W_1, and W_2 are $P_0^* = (\sqrt{\bar{P}_1 - P_1^*} + \sqrt{\bar{P}_2 - P_2^*})^2$, P_1^*, and P_2^*, respectively. The receiver treats the received signal as if it comes from a three-user Gaussian MAC with independent messages, and successively decodes in the order of W_0 first, then W_2, and finally W_1 (similarly, W_0 first, then W_1, and finally W_2). The encoding scheme proposed in [2] generates two large correlated codebooks, instead of three small independent codebooks as we do here. The decoding scheme proposed in [2] uses joint maximum-likelihood (ML) detection of two codewords coming from the two large codebooks, while in our case, we can reduce the complexity by successive decoding, i.e., by applying ML detection to one codeword from a small codebook at a time, while treating other undecoded codewords as noise. If the aim is to achieve some interior point on the line segment TU, then time sharing is used between points T and U. This simpler encoding/decoding scheme is possible because we have a Gaussian channel.

Yet another way to write the capacity region, which will be useful in the development of the fading case in the next section, is the following. The capacity region is all triplets (R_1, R_2, R_0) such that inequalities (7)–(10) hold true for some P_1, P_2, $P_0 \geq 0$, $0 \leq \rho \leq 1$ such that $P_1 + \rho^2 P_0 = \bar{P}_1$ and $P_2 + (1-\rho)^2 P_0 = \bar{P}_2$. This representation of the capacity region can be interpreted as follows: P_1, P_2, and P_0 are the received powers for messages W_1, W_2, and W_0, respectively. In order for the received power for the common message to be P_0, transmitter 1 spends $\rho^2 P_0$ power, and transmitter 2 spends $(1-\rho)^2 P_0$ power. Note that the two powers add up to less than P_0, which is to be expected, because there is a beamforming gain for the common message. Transmitter 1 spends a total of $P_1 + \rho^2 P_0$ power, and this must equal the power constraint $\bar{P}_1$, and transmitter 2 spends a total of $P_2 + (1-\rho)^2 P_0$ power, and this must equal $\bar{P}_2$. Here, ρ can be interpreted as the "portion" of the received power of the common message that comes from transmitter 1.

IV. Capacity Region in Fading

Consider the system model in (2), in the simple case when $H_1(k) = h_1$ and $H_2(k) = h_2$ for all k. Using the representation of the capacity region with P_0, P_1, P_2, and ρ, the capacity region is the set of all triplets (R_1, R_2, R_0) such that inequalities (7)–(10) hold true for some P_1, P_2, $P_0 \geq 0$, $0 \leq \rho \leq 1$ such that $(1/h_1)P_1 + (\rho^2/h_1)P_0 = \bar{P}_1$ and $(1/h_2)P_2 + ((1-\rho)^2/h_2)P_0 = \bar{P}_2$. Here, again, P_1, P_2, and P_0 are all received powers.

Now, we consider the case where the channel is time-varying and both the transmitters and the receiver track the channel perfectly. Let us denote the channel state as a vector $\mathbf{h} = [h_1, h_2]^T$. Let $\mathbf{p} = [p_1, p_2, p_0]^T$ be a mapping from the channel state space $\mathcal{H}$ to the received power vector in $\mathbb{R}_+^3$. Also, let us define ρ to be a mapping from $\mathcal{H}$ to [0,1]. Then, heuristically, when the channel state is $\mathbf{h}$, $p_1(\mathbf{h})/h_1$ is the power that transmitter 1 uses for W_1, and $\rho(\mathbf{h})^2 p_0(\mathbf{h})/h_1$ is the power that transmitter 1 uses for W_0. Similarly, $p_2(\mathbf{h})/h_2$ is the power that transmitter 2 uses for W_2, and $(1-\rho(\mathbf{h}))^2 p_0(\mathbf{h})/h_2$ is the power that transmitter 2 uses for W_0. Let $\mathcal{C}_f(\mathbf{p}, \rho)$ be the set of (R_1, R_2, R_0) such that

$$R_1 \leq E\left[C\left(p_1(\mathbf{h})\right)\right] \triangleq f_1(\mathbf{p}, \rho) \tag{24}$$

$$R_2 \leq E\left[C\left(p_2(\mathbf{h})\right)\right] \triangleq f_2(\mathbf{p}, \rho) \tag{25}$$

$$R_1 + R_2 \leq E\left[C\left(p_1(\mathbf{h}) + p_2(\mathbf{h})\right)\right] \triangleq f_3(\mathbf{p}, \rho) \tag{26}$$

$$R_0 + R_1 + R_2 \leq E\left[C\left(p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})\right)\right] \triangleq f_4(\mathbf{p}, \rho) \tag{27}$$

where the expectation is taken over the joint stationary distribution of the fading states h_1 and h_2.

Theorem 1: The ergodic capacity region of the fading Gaussian MAC with common data when perfect channel state information is available at the transmitters and the receiver is

$$\mathcal{C}(\bar{P}_1, \bar{P}_2) = \bigcup_{(\mathbf{p},\rho) \in \mathcal{F}} \mathcal{C}_f(\mathbf{p}, \rho) \tag{28}$$

where

$$\mathcal{F} = \left\{(\mathbf{p}, \rho) : p_0(\mathbf{h}), p_1(\mathbf{h}), p_2(\mathbf{h}) \geq 0, 0 \leq \rho(\mathbf{h}) \leq 1 \quad \forall \mathbf{h} \right.$$
$$E\left[\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho(\mathbf{h})^2 p_0(\mathbf{h})}{h_1}\right] \leq \bar{P}_1$$
$$\left. E\left[\frac{p_2(\mathbf{h})}{h_2} + \frac{(1-\rho(\mathbf{h}))^2 p_0(\mathbf{h})}{h_2}\right] \leq \bar{P}_2 \right\}. \tag{29}$$

A proof of *Theorem 1* is given in Appendix A.

To explicitly characterize the capacity region, we solve for the boundary surface of the capacity region. As in [19], the boundary surface of the capacity region $\mathcal{C}(\bar{P}_1, \bar{P}_2)$ is the closure of all points $\mathbf{R}^* = (R_1^*, R_2^*, R_0^*)$ such that $\mathbf{R}^*$ is a solution to the problem

$$\max_{\mathbf{R}} \quad \mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0 \quad \text{subject to } \mathbf{R} \in \mathcal{C}(\bar{P}_1, \bar{P}_2) \tag{30}$$

for some $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_0]^T \in \mathbb{R}_+^3$. This optimization problem is equivalent to

$$\max_{(\mathbf{R}, \tilde{P}_1, \tilde{P}_2)} \mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0$$
$$\text{subject to } (\mathbf{R}, \tilde{P}_1, \tilde{P}_2) \in \mathcal{L}, \tilde{P}_1 \leq \bar{P}_1, \tilde{P}_2 \leq \bar{P}_2 \tag{31}$$

where

$$\mathcal{L} = \left\{(\mathbf{R}, \tilde{P}_1, \tilde{P}_2) : \tilde{P}_1, \tilde{P}_2 \in \mathbb{R}_+, \mathbf{R} \in \mathcal{C}(\tilde{P}_1, \tilde{P}_2)\right\}. \tag{32}$$

Lemma 1: $\mathcal{L}$ is a convex set.

A proof of *Lemma 1* is given in Appendix B.

Due to the convexity of $\mathcal{L}$, there exist Lagrange multipliers $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]^T \in \mathbb{R}_+^2$ such that $\mathbf{R}^*$ is a solution to the optimization problem

$$\max_{(\mathbf{R}, \tilde{P}_1, \tilde{P}_2) \in \mathcal{L}} \mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0 - \lambda_1 \tilde{P}_1 - \lambda_2 \tilde{P}_2. \tag{33}$$

Since $\mathcal{C}(\tilde{P}_1, \tilde{P}_2)$ is a union over $\mathcal{C}_f(\mathbf{p}, \rho)$, we first express $(\mathbf{R}, \tilde{P}_1, \tilde{P}_2)$ in terms of $(\mathbf{p}, \rho)$ and then optimize over $(\mathbf{p}, \rho)$. It can be seen that the capacity region is unchanged if we replace the two power constraint inequalities with equalities in (29). Hence

$$\tilde{P}_1 = E\left[\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho(\mathbf{h})^2 p_0(\mathbf{h})}{h_1}\right] \quad (34)$$

$$\tilde{P}_2 = E\left[\frac{p_2(\mathbf{h})}{h_2} + \frac{(1-\rho(\mathbf{h}))^2 p_0(\mathbf{h})}{h_2}\right]. \quad (35)$$

Instead of considering all $\mathbf{R} \in \mathcal{C}(\tilde{P}_1, \tilde{P}_2)$, it suffices to consider $\mathbf{R} \in \mathcal{C}_f(\mathbf{p}, \rho)$ that maximizes $\mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0$ for each $(\mathbf{p}, \rho)$. Thus, we first focus on the following problem:

$$\max_{\mathbf{R}} \quad \mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0 \quad \text{subject to } \mathbf{R} \in \mathcal{C}_f(\mathbf{p}, \rho) \quad (36)$$

where $\mathcal{C}_f(\mathbf{p}, \rho)$ is a region with a shape as in Fig. 1. Due to the nature of $\mathcal{C}_f(\mathbf{p}, \rho)$, when $\mu_0 \geq \max(\mu_1, \mu_2)$, point $Q = [0, 0, f_4(\mathbf{p}, \rho)]$ achieves the maximum. When $\mu_1 \geq \mu_0 \geq \mu_2$, point $S = [f_1(\mathbf{p}, \rho), 0, f_4(\mathbf{p}, \rho) - f_1(\mathbf{p}, \rho)]$ achieves the maximum. When $\mu_1 \geq \mu_2 \geq \mu_0$, point $T = [f_1(\mathbf{p}, \rho), f_3(\mathbf{p}, \rho) - f_1(\mathbf{p}, \rho), f_4(\mathbf{p}, \rho) - f_3(\mathbf{p}, \rho)]$ achieves the maximum. When $\mu_2 \geq \mu_1 \geq \mu_0$, point $U = [f_3(\mathbf{p}, \rho) - f_2(\mathbf{p}, \rho), f_2(\mathbf{p}, \rho), f_4(\mathbf{p}, \rho) - f_3(\mathbf{p}, \rho)]$ achieves the maximum. When $\mu_2 \geq \mu_0 \geq \mu_1$, point $V = [0, f_2(\mathbf{p}, \rho), f_4(\mathbf{p}, \rho) - f_2(\mathbf{p}, \rho)]$ achieves the maximum. Hence, the optimization problem as defined in (36) is solved, and the solution is expressed in terms of $(\mathbf{p}, \rho)$.

We are ready to solve the optimization problem in (33) now. According to the solution to the optimization problem in (36), we have five cases: 1) $\mu_0 \geq \max(\mu_1, \mu_2)$; 2) $\mu_1 \geq \mu_0 \geq \mu_2$; 3) $\mu_1 \geq \mu_2 \geq \mu_0$; 4) $\mu_2 \geq \mu_1 \geq \mu_0$; and 5) $\mu_2 \geq \mu_0 \geq \mu_1$. We will concentrate on the first three cases, since case 4) is the same as case 3), and case 5) is the same as case 2) by swapping indices 1 and 2.

1) When $\mu_0 \geq \max(\mu_1, \mu_2)$, the optimization problem in (33) is equivalent to

$$\min_{\mathbf{p}\geq\mathbf{0}, 0\leq\rho\leq 1} E\Bigg[-\mu_0 \log(1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})) + \lambda_1\left(\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho(\mathbf{h})^2}{h_1} p_0(\mathbf{h})\right) + \lambda_2\left(\frac{p_2(\mathbf{h})}{h_2} + \frac{(1-\rho(\mathbf{h}))^2}{h_2} p_0(\mathbf{h})\right)\Bigg]. \quad (37)$$

Since the cost function is an expectation and the probability distributions are nonnegative, it suffices to consider the minimization for a fixed channel state $\mathbf{h} = (h_1, h_2)$, i.e.,

$$\min_{\mathbf{p}(\mathbf{h})\geq\mathbf{0}, 0\leq\rho(\mathbf{h})\leq 1} -\mu_0 \log(1+p_1(\mathbf{h})+p_2(\mathbf{h})+p_0(\mathbf{h})) +\lambda_1\left(\frac{p_1(\mathbf{h})}{h_1}+\frac{\rho(\mathbf{h})^2}{h_1}p_0(\mathbf{h})\right)+\lambda_2\left(\frac{p_2(\mathbf{h})}{h_2}+\frac{(1-\rho(\mathbf{h}))^2}{h_2}p_0(\mathbf{h})\right). \quad (38)$$

Though the cost function is not convex in $(\mathbf{p}(\mathbf{h}), \rho(\mathbf{h}))$, it is a quadratic function of $\rho(\mathbf{h})$ when $\mathbf{p}(\mathbf{h})$ is fixed. The optimal $\rho^*(\mathbf{h})$ is

$$\rho^*(\mathbf{h}) = \frac{\frac{h_1}{\lambda_1}}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}. \quad (39)$$

Since the dependencies of the cost functions on $\rho(\mathbf{h})$ in all three cases are the same, $\rho^*(\mathbf{h})$ is, in fact, the optimal solution for all three cases. Thus, we proceed with $\rho^*(\mathbf{h})$ in place of $\rho(\mathbf{h})$ and the problem becomes convex. We write the Karush–Kuhn–Tucker (KKT) necessary conditions as follows:

$$-\frac{\mu_0}{1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})} + \frac{1}{\frac{h_1}{\lambda_1}} - \omega_1(\mathbf{h}) = 0 \quad (40)$$

$$-\frac{\mu_0}{1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})} + \frac{1}{\frac{h_2}{\lambda_2}} - \omega_2(\mathbf{h}) = 0 \quad (41)$$

$$-\frac{\mu_0}{1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})} + \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}} - \omega_0(\mathbf{h}) = 0 \quad (42)$$

$$p_1(\mathbf{h}), p_2(\mathbf{h}), p_0(\mathbf{h}), \omega_0(\mathbf{h}), \omega_1(\mathbf{h}), \omega_2(\mathbf{h}) \geq 0 \quad (43)$$

$$\omega_0(\mathbf{h}) p_0(\mathbf{h}) = \omega_1(\mathbf{h}) p_1(\mathbf{h}) = \omega_2(\mathbf{h}) p_2(\mathbf{h}) = 0 \quad (44)$$

where $\omega_0(\mathbf{h})$, $\omega_1(\mathbf{h})$, and $\omega_2(\mathbf{h})$ are the complementary slackness variables. The KKTs have a unique solution, and thus the solution is the global optimum. Let us define two regions in $\mathbb{R}_+^2$

$$\mathcal{R}_1 = \left\{(x, y) : x + y \geq \frac{1}{\mu_0}\right\} \quad (45)$$

$$\mathcal{R}_2 = \left\{(x, y) : x + y < \frac{1}{\mu_0}\right\}. \quad (46)$$

Then, the optimum solution is

$$p_0(\mathbf{h}) = \begin{cases} \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - 1, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{R}_1 \\ 0, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{R}_2 \end{cases} \quad (47)$$

$$p_1(\mathbf{h}) = 0 \quad (48)$$

$$p_2(\mathbf{h}) = 0. \quad (49)$$

The transmit powers can be found by dividing these received powers with corresponding channel gains. As seen from (48) and (49), in the case of $\mu_0 \geq \max(\mu_1, \mu_2)$, the transmitters use their entire power to transmit the common message; they do not allocate any power to transmit their individual messages. When $(h_1/\lambda_1) + (h_2/\lambda_2) \geq (1/\mu_0)$, i.e., the combined channel is good enough, the transmitters transmit the common message using beamforming as if we have a two-transmitter one-receiver point-to-point system. When the channel is poor, i.e., $(h_1/\lambda_1) + (h_2/\lambda_2) < (1/\mu_0)$, the transmitters both keep silent and save their powers for better channel states. This is shown in Fig. 4.

2) When $\mu_1 \geq \mu_0 \geq \mu_2$, the optimization problem in (33) is equivalent to

$$\min_{\mathbf{p}\geq\mathbf{0}, 0\leq\rho\leq 1} E\Bigg[-\mu_0 \log(1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})) - (\mu_1 - \mu_0)\log(1 + p_1(\mathbf{h})) + \lambda_1\left(\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho(\mathbf{h})^2}{h_1} p_0(\mathbf{h})\right) + \lambda_2\left(\frac{p_2(\mathbf{h})}{h_2} + \frac{(1-\rho(\mathbf{h}))^2}{h_2} p_0(\mathbf{h})\right)\Bigg]. \quad (50)$$

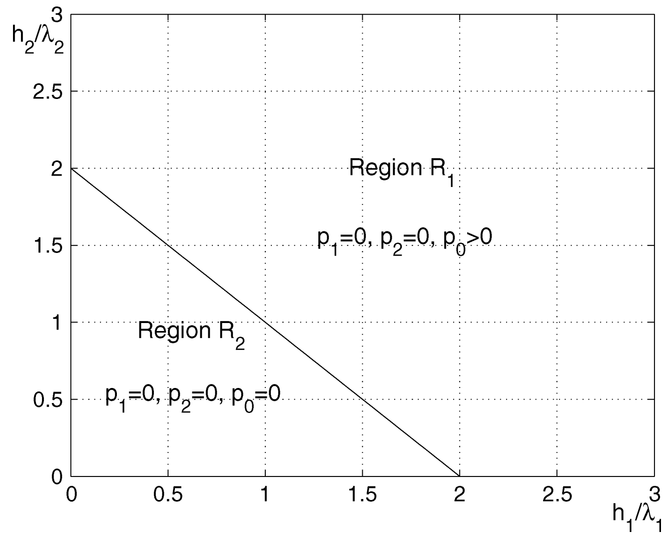


Fig. 4. power control policy in the case of $\mu_0 \geq \max(\mu_1, \mu_2)$.

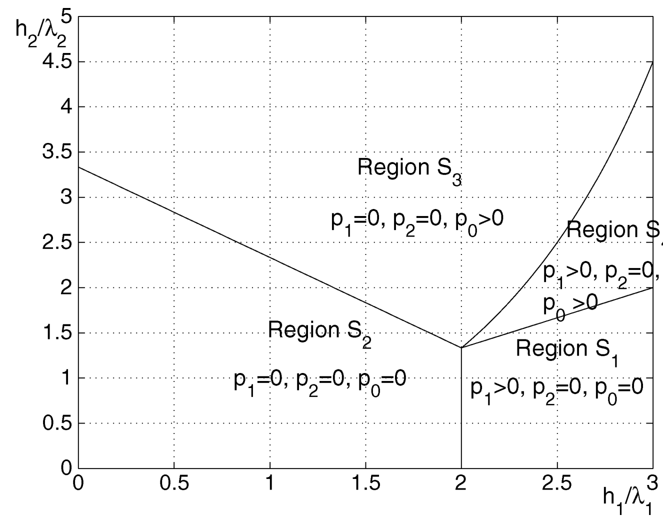


Fig. 5. power control policy in the case of $\mu_1 \geq \mu_0 \geq \mu_2$.

Following the same argument as in case 1), let us define four regions in $\mathbb{R}_+^2$

$$\mathcal{S}_1 = \left\{(x,y) : x \geq \frac{1}{\mu_1}, \frac{y}{x} < \frac{\mu_1}{\mu_0} - 1\right\} \tag{51}$$

$$\mathcal{S}_2 = \left\{(x,y) : x < \frac{1}{\mu_1}, x+y < \frac{1}{\mu_0}\right\} \tag{52}$$

$$\mathcal{S}_3 = \left\{(x,y) : \frac{1}{x} - \frac{1}{x+y} \geq \mu_1 - \mu_0, x+y \geq \frac{1}{\mu_0}\right\} \tag{53}$$

$$\mathcal{S}_4 = \left\{(x,y) : \frac{1}{x} - \frac{1}{x+y} < \mu_1 - \mu_0, \frac{y}{x} \geq \frac{\mu_1}{\mu_0} - 1, x+y \geq \frac{1}{\mu_0}\right\}. \tag{54}$$

Then, the optimal solution is

$$p_0(\mathbf{h}) = \begin{cases} \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - 1, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{S}_3 \\ \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - (\mu_1 - \mu_0) & \\ \quad \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1}, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{S}_4 \\ 0, & \text{otherwise} \end{cases} \tag{55}$$

$$p_1(\mathbf{h}) = \begin{cases} \mu_1 \frac{h_1}{\lambda_1} - 1, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{S}_1 \\ (\mu_1 - \mu_0) & \\ \quad \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1} - 1, & \text{if } \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{S}_4 \\ 0, & \text{otherwise} \end{cases} \tag{56}$$

$$p_2(\mathbf{h}) = 0. \tag{57}$$

Again, the transmit powers are found by dividing these with appropriate channel gains. As seen from (57), in the case of $\mu_1 \geq \mu_0 \geq \mu_2$, transmitter 2 never uses its power to transmit its individual message. When both channels are poor, no one transmits. When the channel of the first transmitter is much better than that of the second transmitter, transmitter 1 transmits only its individual message and transmitter 2 keeps silent. When the channel of the second transmitter is much better than that of the first transmitter, both transmitters cooperate using beamforming to transmit the common message. When both channels are more or less equally good, both common message and individual message from transmitter 1 are transmitted. These regions are shown in Fig. 5.

3) When $\mu_1 \geq \mu_2 \geq \mu_0$, the optimization problem in (33) is equivalent to

$$\begin{aligned} \min_{\mathbf{p} \geq \mathbf{0}, 0 \leq \rho \leq 1} E\Big[& -\mu_0 \log(1 + p_1(\mathbf{h}) + p_2(\mathbf{h}) + p_0(\mathbf{h})) \\ & - (\mu_2 - \mu_0) \log(1 + p_1(\mathbf{h}) + p_2(\mathbf{h})) \\ & - (\mu_1 - \mu_2) \log(1 + p_1(\mathbf{h})) \\ & + \lambda_1 \left(\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho(\mathbf{h})^2}{h_1} p_0(\mathbf{h})\right) \\ & + \lambda_2 \left(\frac{p_2(\mathbf{h})}{h_2} + \frac{(1 - \rho(\mathbf{h}))^2}{h_2} p_0(\mathbf{h})\right)\Big]. \end{aligned} \tag{58}$$

Let us define eight regions in $\mathbb{R}_+^2$

$$\mathcal{T}_1 = \left\{(x,y) : x < \frac{1}{\mu_1}, y < \frac{1}{\mu_2}, x+y < \frac{1}{\mu_0}\right\} \tag{59}$$

$$\mathcal{T}_2 = \left\{(x,y) : x+y \geq \frac{1}{\mu_0}, \frac{1}{y} - \frac{1}{x+y} \geq \mu_2 - \mu_0, \frac{1}{x} - \frac{1}{x+y} \geq \mu_1 - \mu_0\right\} \tag{60}$$

$$\mathcal{T}_3 = \left\{(x,y) : x \geq \frac{1}{\mu_1}, \frac{y}{x} < \min\left(\frac{\mu_1}{\mu_2}, \frac{\mu_1}{\mu_0} - 1\right)\right\} \tag{61}$$

$$\mathcal{T}_4 = \left\{(x,y) : y \geq \frac{1}{\mu_2}, \frac{x}{y} < \frac{\mu_2}{\mu_0} - 1, \frac{1}{x} - \frac{1}{y} \geq \mu_1 - \mu_2\right\} \tag{62}$$

$$\mathcal{T}_5 = \left\{(x,y) : x+y \geq \frac{1}{\mu_0}, \frac{1}{x} - \frac{1}{x+y} < \mu_1 - \mu_0, \frac{\mu_1}{\mu_0} - 1 \leq \frac{y}{x} < \sqrt{\frac{\mu_1 - \mu_0}{\mu_2 - \mu_0}}\right\} \tag{63}$$

$$\mathcal{T}_6 = \left\{(x,y) : x+y \geq \frac{1}{\mu_0}, \frac{x}{y} \geq \frac{\mu_2}{\mu_0} - 1, \frac{1}{y} - \frac{1}{x+y} < \mu_2 - \mu_0, \frac{1}{x} - \frac{1}{y} \geq \mu_1 - \mu_2\right\} \tag{64}$$

$$\mathcal{T}_7 = \Big\{(x,y) : y \geq \frac{1}{\mu_2}, \frac{1}{x} - \frac{1}{y} < \mu_1 - \mu_2, \frac{x}{y} < \min\left(\frac{\mu_2}{\mu_1}, \frac{\mu_2}{\mu_0} - 1\right)\Big\} \tag{65}$$

$$\mathcal{T}_8 = \Big\{(x,y) : x + y \geq \frac{1}{\mu_0}, \frac{1}{y} - \frac{1}{x+y} < \mu_2 - \mu_0, \frac{1}{x} - \frac{1}{y} < \mu_1 - \mu_2, \frac{\mu_2}{\mu_0} - 1 \leq \frac{x}{y} < \min\left(\sqrt{\frac{\mu_2-\mu_0}{\mu_1-\mu_0}}, \left(\frac{c+\sqrt{c^2+4}}{2}\right)^{-1}\right) \text{ where } c = \frac{\mu_1 - \mu_2}{\mu_0}\Big\}. \tag{66}$$

Then, the optimal solution is

$$p_0(\mathbf{h}) = \begin{cases} \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_2 \\ \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - (\mu_1 - \mu_0) \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1}, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_5 \\ \mu_0\left(\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}\right) - (\mu_2 - \mu_0) \times \left(\frac{1}{\frac{h_2}{\lambda_2}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1}, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_6 \bigcup \mathcal{T}_8 \\ 0, & \text{otherwise} \end{cases} \tag{67}$$

$$p_1(\mathbf{h}) = \begin{cases} \mu_1 \frac{h_1}{\lambda_1} - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_3 \\ (\mu_1 - \mu_0) \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1} - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_5 \\ (\mu_1 - \mu_2) \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_2}{\lambda_2}}\right)^{-1} - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_7 \bigcup \mathcal{T}_8 \\ 0, & \text{otherwise} \end{cases} \tag{68}$$

$$p_2(\mathbf{h}) = \begin{cases} \mu_2 \frac{h_2}{\lambda_2} - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_4 \\ (\mu_2 - \mu_0) \times \left(\frac{1}{\frac{h_2}{\lambda_2}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1} - 1, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_6 \\ \mu_2 \frac{h_2}{\lambda_2} - (\mu_1 - \mu_2) \times \left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_2}{\lambda_2}}\right)^{-1}, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_7 \\ (\mu_2 - \mu_0)\left(\frac{1}{\frac{h_2}{\lambda_2}} - \frac{1}{\frac{h_1}{\lambda_1} + \frac{h_2}{\lambda_2}}\right)^{-1} - (\mu_1 - \mu_2)\left(\frac{1}{\frac{h_1}{\lambda_1}} - \frac{1}{\frac{h_2}{\lambda_2}}\right)^{-1}, & \text{if} \left(\frac{h_1}{\lambda_1}, \frac{h_2}{\lambda_2}\right) \in \mathcal{T}_8 \\ 0, & \text{otherwise.} \end{cases} \tag{69}$$

As in the previous two cases, the transmit powers are found by dividing these with the corresponding channel gains. There are two subcases in the case of $\mu_1 \geq \mu_2 \geq \mu_0$. When $(1/\mu_1)+(1/\mu_2) \leq (1/\mu_0)$, i.e., μ_0 is very small, the common message never gets transmitted due to its small weight. When both channels are poor, no one transmits.

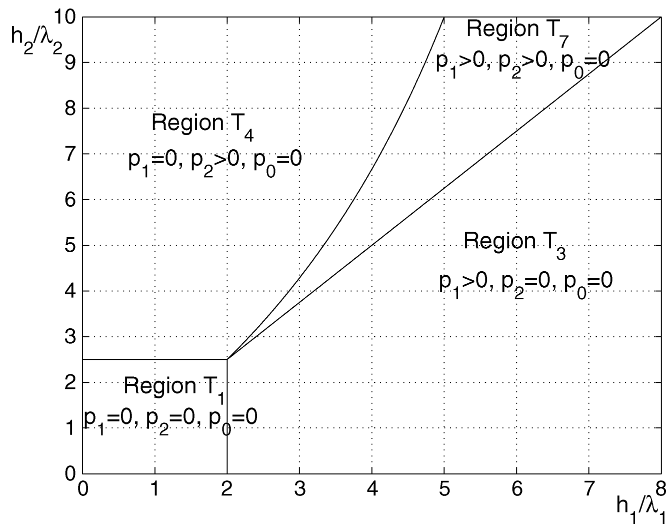


Fig. 6. power control policy in the case of $\mu_1 \geq \mu_2 \geq \mu_0$ and $(1/\mu_1) + (1/\mu_2) \leq (1/\mu_0)$.

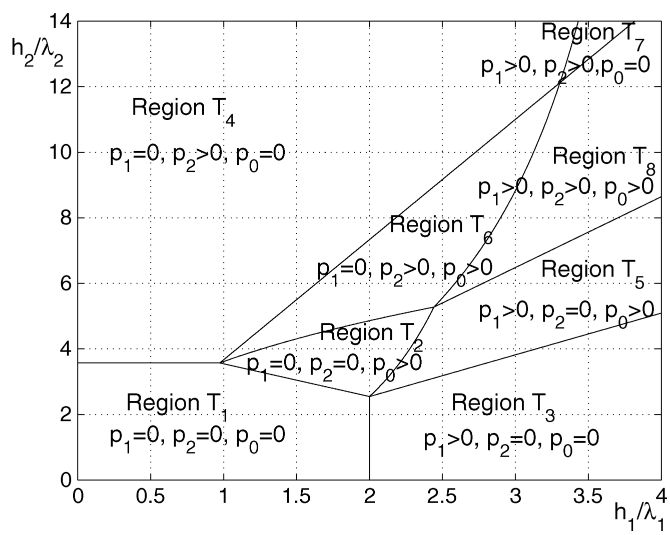


Fig. 7. power control policy in the case of $\mu_1 \geq \mu_2 \geq \mu_0$ and $(1/\mu_1) + (1/\mu_2) > (1/\mu_0)$.

When channel of the first transmitter is much better than that of the second transmitter, individual message W_1 is transmitted only. When channel of the second transmitter is much better than that of the first transmitter, individual message W_2 is transmitted only. When both channels are more or less equally good, both individual messages are transmitted. These regions are shown in Fig. 6.

In the other subcase of $(1/\mu_1) + (1/\mu_2) > (1/\mu_0)$, all three messages get a chance to be transmitted. These regions are shown in Fig. 7.

Thus far, we have solved the optimization problem in (33) in terms of the Lagrange multipliers $\boldsymbol{\lambda}$. Next, we need to solve for $\boldsymbol{\lambda}$. Since there is no duality gap, we will solve for $\boldsymbol{\lambda}$ by solving the dual problem, i.e., we will find $\boldsymbol{\lambda}$ that maximizes the dual function, $g(\boldsymbol{\lambda})$. The maximizer of the dual function enables the power policies to satisfy the power constraints with equalities due to the uniqueness of the optimal p_0, p_1, p_2, ρ for each given

$\boldsymbol{\lambda}$. We will solve the dual problem by using the subgradient method [20]. For our problem

$$u(\boldsymbol{\lambda}) \triangleq \begin{bmatrix} E\left[\frac{p_1(\mathbf{h})}{h_1} + \frac{\rho^*(\mathbf{h})^2 p_0(\mathbf{h})}{h_1}\right] - \bar{P}_1 \\ E\left[\frac{p_2(\mathbf{h})}{h_2} + \frac{(1-\rho^*(\mathbf{h}))^2 p_0(\mathbf{h})}{h_2}\right] - \bar{P}_2 \end{bmatrix} \tag{70}$$

is a subgradient of the dual function and the set $\{\boldsymbol{\lambda} : \boldsymbol{\lambda} \geq 0, g(\boldsymbol{\lambda}) > -\infty\} = \{\boldsymbol{\lambda} : \boldsymbol{\lambda} > 0\}$. We start from an arbitrary point $\boldsymbol{\lambda}(0) \in \{\boldsymbol{\lambda} : \boldsymbol{\lambda} > 0\}$. At iteration k, we have available $\boldsymbol{\lambda}(k-1)$ from the previous iteration, and we compute (p_0, p_1, p_2, ρ) by setting $\boldsymbol{\lambda} = \boldsymbol{\lambda}(k-1)$. Then, using the (p_0, p_1, p_2, ρ) we obtained, we compute the subgradient vector $u(\boldsymbol{\lambda}(k-1))$ by (70) and update $\boldsymbol{\lambda}$ using

$$\boldsymbol{\lambda}(k) = \max\left[\boldsymbol{\lambda}(k-1) + s(k)u\left(\boldsymbol{\lambda}(k-1)\right), \boldsymbol{\epsilon}\right] \tag{71}$$

where $s(k)$ is a positive scalar stepsize at step k, and $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2]^T$ is a positive vector very close to zero so that $\boldsymbol{\lambda}(k)$ stays in $\{\boldsymbol{\lambda} : \boldsymbol{\lambda} > 0\}$. We stop when both components of vector $u(\boldsymbol{\lambda}(k))$ are small enough. In [20], it is proved that for small enough step sizes, this algorithm converges.

Due to the strict concavity of the log function, the Lagrange multipliers are unique. The uniqueness of the Lagrange multipliers ensures that the boundary rate triplet that solves (30) is unique for all $\boldsymbol{\mu}$ vectors except for the following three singular cases: $\mu_0 = \mu_1 = \mu_2 = 0$; $\mu_1 > \mu_0 = \mu_2 = 0$; and $\mu_2 > \mu_0 = \mu_1 = 0$. Thus, by varying the $\boldsymbol{\mu}$ vector over all possible values, and expressing the rates in limiting expressions for the singular cases, we obtain the entire boundary surface of the capacity region. In the process, we also obtain the power control policies that achieve the rate tuples on the boundary.

V. Simulations

In this section, we present simulation results for a two-user Gaussian MAC with common data in the presence of fading. The channel gains are assumed to be independent, identically distributed (i.i.d.) exponential with mean 1, independent across the two users. In our simulations, we use the subgradient method, and we picked the stepsize $s(k)$ by method (a) in [20, p. 508].

In Fig. 8, we show the ergodic capacity region of this two-user Gaussian MAC with common data in fading. The power constraints are $\bar{P}_1 = 2$ and $\bar{P}_2 = 1$. We calculated the rate triplets on the boundary of the capacity region by varing $\boldsymbol{\mu}$ over all possible values. It is straightforward to see that point R is the solution to case 1), which is independent of μ_0. Points between R and S are the solutions to case 2). Points between T and U are solutions to subcase 1 of case 3) and case 4). Points between V and R are solutions to case 5). All points on the surface of $RSTUV$ are solutions to subcase 2 of case 3) and case 4). Surface YST is the singular case of $\mu_1 > \mu_0 = \mu_2 = 0$, and surface UVZ is the singular case of $\mu_2 > \mu_0 = \mu_1 = 0$.

We next compare the achievable rate $\mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0$ under different power allocation schemes. We choose $\mu_1 = 0.45$, $\mu_2 = 0.35$, and $\mu_0 = 0.2$ which corresponds to an interesting case where all three rates, R_0, R_1, and R_2, are nonzero, i.e., subcase 2 of case 3). In Fig. 9, we plot the achievable rate as a function of the sum of the power constraints, i.e., $\bar{P}_1 + \bar{P}_2$. In this experiment, we assume that the power constraints are the same for both users, i.e., $\bar{P}_1 = \bar{P}_2$. The top-most curve in Fig. 9 corresponds to the rate achieved by the optimum power allocation scheme we developed in this paper. It is numerically solved by using the subgradient method. The "optimal channel-independent power control" curve corresponds to the solution of the following problem:

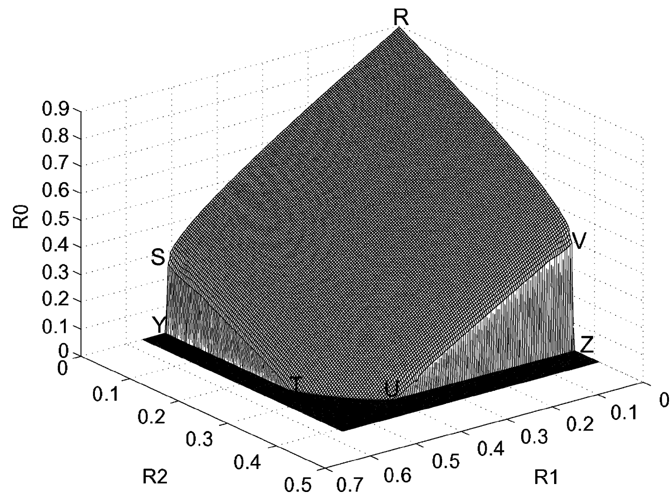


Fig. 8. Ergodic capacity region of the Gaussian MAC with common data in fading.

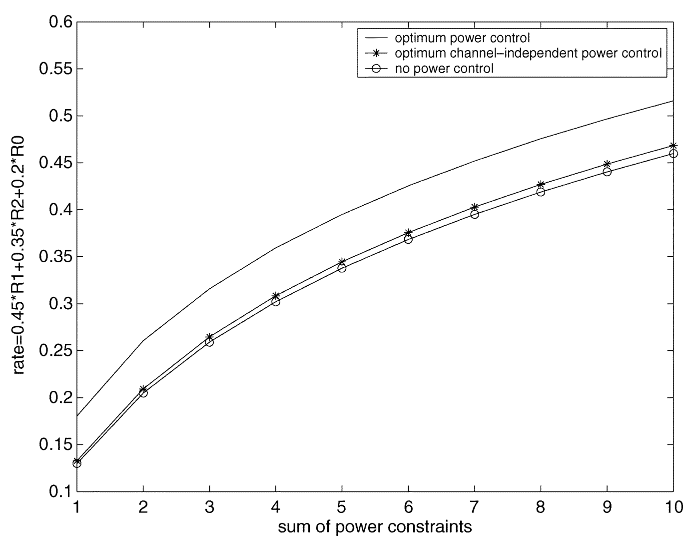


Fig. 9. A weighted sum of rates with and without power control.

$$\max_{0 \leq P_1 \leq \bar{P}_1, 0 \leq P_2 \leq \bar{P}_2} E\Big[\mu_0 \log\Big(1 + h_1\bar{P}_1 + h_2\bar{P}_2 + 2\sqrt{h_1 h_2(\bar{P}_1 - P_1)(\bar{P}_2 - P_2)}\Big) + (\mu_2 - \mu_0)\log(1 + h_1 P_1 + h_2 P_2) + (\mu_1 - \mu_2)\log(1 + h_1 P_1)\Big] \tag{72}$$

where we choose P_1 and P_2 to maximize the expectation in (72). Note that P_1 and P_2 are constants, and not functions of the channel realizations. This corresponds to the largest achievable rate $\mu_1 R_1 + \mu_2 R_2 + \mu_0 R_0$ when there is no channel state information at the transmitters, i.e., the transmitters only know the statistics of the channel gains. This maximization is solved numerically by searching over all admissible P_1 and P_2. The lowest curve in Fig. 9 corresponds to the case where we choose $P_1 = P_2 = P_0$, with $P_0 = (\sqrt{\bar{P}_1 - P_1} + \sqrt{\bar{P}_2 - P_2})^2$. This corresponds to a case where the transmitters do not know the channel realizations or the channel statistics. Consequently, the transmitters use "equal" powers for all three messages. For this instance, we see from Fig. 9 that there is a relatively large performance gain due to adjusting the transmit powers according to the channel realizations. For this particular fading distribution, using optimum channel-independent power control provides only a small gain over choosing "equal" powers for all three messages.

VI. Conclusion

In this paper, we study the Gaussian MAC with common data. In the case of no fading, we provide an explicit characterization of the capacity region, and a simpler encoding/decoding scheme. In the case of fading, we characterize the ergodic capacity region, as well as the power control policies that achieve the rate tuples on the boundary of the capacity region. As expected, the common message enjoys a beamforming gain. Hence, if all three rates are weighted equally, i.e., we are interested in the sum capacity, then we would always transmit only the common message using beamforming.

Appendix

A. Proof of Theorem 1

The achievability part follows from an argument similar to [19] and thus is omitted.

For the converse, we develop a series of bounds on the achievable rates.

$$\begin{aligned} nR_1 &= H(W_1|\mathbf{H}^n) && (73)\\ &\leq H(W_1|Y^n, \mathbf{H}^n) + I(W_1; Y^n|\mathbf{H}^n) && (74)\\ &\stackrel{(a)}{\leq} n\epsilon_n + I(W_1; Y^n|\mathbf{H}^n) && (75)\\ &\stackrel{(b)}{\leq} n\epsilon_n + I\left(W_1; Y^n|W_0, X_2^n, \mathbf{H}^n\right) && (76) \end{aligned}$$

where (a) follows from Fano's inequality [21], and (b) follows from the fact that W_1 and (W_0, X_2^n) are independent, conditioned on $\mathbf{H}^n$.

$$\begin{aligned} &I\left(W_1; Y^n|W_0, X_2^n, \mathbf{H}^n\right)\\ &\stackrel{(c)}{\leq} I\left(X_1^n; Y^n|W_0, X_2^n, \mathbf{H}^n\right) && (77)\\ &\stackrel{(d)}{\leq} \sum_{i=1}^{n} I(X_{1i}; Y_i|X_{2i}, W_0, \mathbf{H}_i) && (78)\\ &= \sum_{i=1}^{n} \int_{\mathcal{H}} p_{\mathbf{H}_i}(\mathbf{h}) I(X_{1i}; Y_i|X_{2i}, W_0, \mathbf{h}) d\mathbf{h} && (79)\\ &= \sum_{i=1}^{n} \int_{\mathcal{H}} p_{\mathbf{H}_i}(\mathbf{h}) \left(h(\sqrt{h_1}X_{1i} + Z_i|W_0, \mathbf{h}) - \frac{1}{2}\log(2\pi e) \right) d\mathbf{h} && (80) \end{aligned}$$

where (c) follows from the data processing inequality [21] and (d) follows from the usual converse argument that upper bounds the mutual information of n-sequences by the sum of the mutual informations of the single letters, based on the fact that the channel is memoryless conditioned on the channel fading coefficients. In (80), $h(\cdot)$ denotes the differential entropy.

$$\begin{aligned} &h(\sqrt{h_1}X_{1i} + Z_i|W_0, \mathbf{h})\\ &\quad = E_s\left[h(\sqrt{h_1}X_{1i} + Z_i|W_0 = s, \mathbf{h}) \right] && (81)\\ &\quad \stackrel{(e)}{\leq} \frac{1}{2}\log(2\pi e)\left(h_1 E_s\left[V(X_{1i}|W_0 = s, \mathbf{h}) \right] + 1 \right) && (82) \end{aligned}$$

where $V(\cdot)$ is the variance of a random variable and (e) follows from the fact that given the variance, Gaussian distribution maximizes the entropy, and applying Jensen's inequality [21] afterwards. Then

$$\begin{aligned} R_1 &\leq \frac{1}{n}\sum_{i=1}^{n} \int_{\mathcal{H}} p_{\mathbf{H}_i}(\mathbf{h}) \frac{1}{2}\\ &\quad \times \log\left(1 + h_1 E_s\left[V(X_{1i}|W_0 = s, \mathbf{h}) \right]\right) d\mathbf{h} + \epsilon_n && (83)\\ &\stackrel{(f)}{=} \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h}) \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2}\\ &\quad \times \log\left(1 + h_1 E_s\left[V(X_{1i}|W_0 = s, \mathbf{h}) \right]\right) d\mathbf{h} + \epsilon_n && (84)\\ &\stackrel{(g)}{\leq} \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h}) \frac{1}{2}\\ &\quad \times \log\left(1 + h_1 \frac{1}{n}\sum_{i=1}^{n} E_s\left[V(X_{1i}|W_0 = s, \mathbf{h}) \right]\right) d\mathbf{h} + \epsilon_n && (85) \end{aligned}$$

where in writing (f), we define $\mathbf{H}$ to be a random variable whose distribution is the same as the stationary distribution of $\mathbf{H}_i$, and (g) follows from the concavity of the function $\log(1 + x)$.

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} E_s\left[V(X_{1i}|W_0 = s, \mathbf{h}) \right]\\ &\quad = \frac{1}{n}\sum_{i=1}^{n} \left(E_s\left[E\left[X_{1i}^2|W_0 = s, \mathbf{h} \right] \right.\right.\\ &\quad\quad \left.\left. - E^2\left[X_{1i}|W_0 = s, \mathbf{h} \right] \right] \right) && (86)\\ &\quad = \frac{1}{n}\sum_{i=1}^{n} E\left[X_{1i}^2|\mathbf{h} \right]\\ &\quad\quad - \frac{1}{n}\sum_{i=1}^{n} E_s\left[E^2\left[X_{1i}|W_0 = s, \mathbf{h} \right] \right] && (87) \end{aligned}$$

Let us define $P_1(\mathbf{h}) = (1/n)\sum_{i=1}^{n} E[X_{1i}^2|\mathbf{h}]$ and $(1 - \alpha(\mathbf{h}))P_1(\mathbf{h}) = (1/n)\sum_{i=1}^{n} E_s[E^2[X_{1i}|W_0 = s, \mathbf{h}]]$ and by definition, $0 \leq \alpha(\mathbf{h}) \leq 1$. Hence

$$R_1 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h}) \frac{1}{2}\log\left(1 + h_1\alpha(\mathbf{h})P_1(\mathbf{h})\right) d\mathbf{h} + \epsilon_n. \quad (88)$$

Let us define $P_2(\mathbf{h}) = (1/n)\sum_{i=1}^{n} E[X_{2i}^2|\mathbf{h}]$ and $(1 - \beta(\mathbf{h}))P_2(\mathbf{h}) = (1/n)\sum_{i=1}^{n} E_s[E^2[X_{2i}|W_0 = s, \mathbf{h}]]$. Then, a symmetric argument gives

$$R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h}) \frac{1}{2}\log\left(1 + h_2\beta(\mathbf{h})P_2(\mathbf{h})\right) d\mathbf{h} + \epsilon_n. \quad (89)$$

Following arguments similar to (73)–(84), we get an inequality akin to (85) as shown in (90)–(92) at the bottom of the page, where (h) follows from the fact that, without loss of generality, we may consider encoders that depend only on the current channel state. Then, it follows that, conditioned on the common message W_0 and the current channel state $\mathbf{H}_i = \mathbf{h}$, X_{1i} and X_{2i} are independent.

For the case of $R_0 + R_1 + R_2$, again, by following similar arguments, we get an inequality akin to (85) as

$$R_0+R_1+R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2} \times \log\left(1+\frac{1}{n}\sum_{i=1}^{n} V\left(\sqrt{h_1}X_{1i}+\sqrt{h_2}X_{2i}|\mathbf{h}\right)\right)d\mathbf{h} + \epsilon_n. \quad (93)$$

Now, we have (94)–(100), shown at the bottom of the page. Hence

$$R_0 + R_1 + R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log\Big(1 + h_1P_1(\mathbf{h}) + h_2P_2(\mathbf{h}) +2\sqrt{h_1h_2\left(1-\alpha(\mathbf{h})\right)\left(1-\beta(\mathbf{h})\right)P_1(\mathbf{h})P_2(\mathbf{h})}\Big)\, d\mathbf{h} + \epsilon_n. \quad (101)$$

The power constraints of the system are

$$\frac{1}{n}\sum_{i=1}^{n} X_{1i}^2 \leq \bar{P}_1, \quad \frac{1}{n}\sum_{i=1}^{n} X_{2i}^2 \leq \bar{P}_2 \quad \text{with probability 1.} \quad (102)$$

Hence

$$\int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})P_1(\mathbf{h})d\mathbf{h} = \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{n}\sum_{i=1}^{n} E\left[X_{1i}^2|\mathbf{h}\right] d\mathbf{h} = \frac{1}{n}\sum_{i=1}^{n} E\left[X_{1i}^2\right] \leq \bar{P}_1 \quad (103)$$

$$\int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})P_2(\mathbf{h})d\mathbf{h} = \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{n}\sum_{i=1}^{n} E\left[X_{2i}^2|\mathbf{h}\right] d\mathbf{h} = \frac{1}{n}\sum_{i=1}^{n} E\left[X_{2i}^2\right] \leq \bar{P}_2. \quad (104)$$

The rates triplets (R_1, R_2, R_0) have to satisfy

$$R_1 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log(1+h_1\alpha(\mathbf{h})P_1(\mathbf{h}))\, d\mathbf{h}+\epsilon_n \quad (105)$$

$$R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log(1+h_2\beta(\mathbf{h})P_2(\mathbf{h}))\, d\mathbf{h}+\epsilon_n \quad (106)$$

$$R_1 + R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log\left(1 + \frac{1}{n}\sum_{i=1}^{n} E_s\Big[V\Big(\sqrt{h_1}X_{1i} + \sqrt{h_2}X_{2i}|W_0 = s, \mathbf{h}\Big)\Big]\right)d\mathbf{h} + \epsilon_n \quad (90)$$

$$\stackrel{(h)}{=} \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log\left(1 + h_1\frac{1}{n}\sum_{i=1}^{n} E_s\Big[V\Big(\sqrt{h_1}X_{1i}|W_0 = s, \mathbf{h}\Big)\Big] + h_2\frac{1}{n}\sum_{i=1}^{n} E_s\Big[V\Big(\sqrt{h_2}X_{2i}|W_0 = s, \mathbf{h}\Big)\Big]\right)d\mathbf{h} + \epsilon_n \quad (91)$$

$$= \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log(1 + h_1\alpha(\mathbf{h})P_1(\mathbf{h}) + h_2\beta(\mathbf{h})P_2(\mathbf{h}))\, d\mathbf{h} + \epsilon_n \quad (92)$$

$$\frac{1}{n}\sum_{i=1}^{n} V(\sqrt{h_1}X_{1i} + \sqrt{h_2}X_{2i}|\mathbf{h}) \quad (94)$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} E\left[(\sqrt{h_1}X_{1i} + \sqrt{h_2}X_{2i})^2|\mathbf{h}\right] \quad (95)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E_s\left[E\left[(\sqrt{h_1}X_{1i} + \sqrt{h_2}X_{2i})^2|W_0 = s, \mathbf{h}\right]\right] \quad (96)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(h_1E\left[X_{1i}^2|\mathbf{h}\right] + h_2E\left[X_{2i}^2|\mathbf{h}\right] + 2\sqrt{h_1h_2}E_s\left[E[X_{1i}X_{2i}|W_0 = s, \mathbf{h}]\right]\right) \quad (97)$$

$$\leq h_1P_1(\mathbf{h}) + h_2P_2(\mathbf{h}) + 2\sqrt{h_1h_2}\frac{1}{n}\sum_{i=1}^{n}\left(E_s\left[E^2[X_{1i}|W_0 = s, \mathbf{h}]\right]\right)^{\frac{1}{2}}\left(E_s\left[E^2[X_{2i}|W_0 = s, \mathbf{h}]\right]\right)^{\frac{1}{2}} \quad (98)$$

$$\leq h_1P_1(\mathbf{h}) + h_2P_2(\mathbf{h}) + 2\sqrt{h_1h_2\left(\frac{1}{n}\sum_{i=1}^{n} E_s\left[E^2[X_{1i}|W_0 = s, \mathbf{h}]\right]\right)\left(\frac{1}{n}\sum_{i=1}^{n} E_s\left[E^2[X_{2i}|W_0 = s, \mathbf{h}]\right]\right)} \quad (99)$$

$$= h_1P_1(\mathbf{h}) + h_2P_2(\mathbf{h}) + 2\sqrt{h_1h_2\left(1-\alpha(\mathbf{h})\right)\left(1-\beta(\mathbf{h})\right)P_1(\mathbf{h})P_2(\mathbf{h})} \quad (100)$$

$$R_1+R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log(1+h_1\alpha(\mathbf{h})P_1(\mathbf{h}) + h_2\beta(\mathbf{h})P_2(\mathbf{h}))\, d\mathbf{h}+\epsilon_n \quad (107)$$

$$R_0+R_1+R_2 \leq \int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})\frac{1}{2}\log\Big(1+h_1P_1(\mathbf{h})+h_2P_2(\mathbf{h}) +2\sqrt{h_1h_2\left(1-\alpha(\mathbf{h})\right)\left(1-\beta(\mathbf{h})\right)P_1(\mathbf{h})P_2(\mathbf{h})}\Big)\, d\mathbf{h}+\epsilon_n \quad (108)$$

for some $\alpha(\mathbf{h})$ and $\beta(\mathbf{h})$ that map state space to [0,1] and $P_1(\mathbf{h})$ and $P_2(\mathbf{h})$ that satisfy

$$\int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})P_1(\mathbf{h})d\mathbf{h} \leq \bar{P}_1 \quad (109)$$

$$\int_{\mathcal{H}} p_{\mathbf{H}}(\mathbf{h})P_2(\mathbf{h})d\mathbf{h} \leq \bar{P}_2. \quad (110)$$

We make the following variable changes:

$$p_1(\mathbf{h}) = h_1\alpha(\mathbf{h})P_1(\mathbf{h}) \quad (111)$$

$$p_2(\mathbf{h}) = h_2\beta(\mathbf{h})P_2(\mathbf{h}) \quad (112)$$

$$p_0(\mathbf{h}) = \left(\sqrt{h_1(1-\alpha(\mathbf{h}))P_1(\mathbf{h})}+\sqrt{h_2(1-\beta(\mathbf{h}))P_2(\mathbf{h})}\right)^2 \quad (113)$$

$$\rho(\mathbf{h}) = \frac{\sqrt{h_1\left(1-\alpha(\mathbf{h})\right)P_1(\mathbf{h})}}{\sqrt{p_0(\mathbf{h})}}. \quad (114)$$

Thus

$$R_1 \leq E\left[C\left(p_1(\mathbf{h})\right)\right] \quad (115)$$

$$R_2 \leq E\left[C\left(p_2(\mathbf{h})\right)\right] \quad (116)$$

$$R_1+R_2 \leq E\left[C\left(p_1(\mathbf{h})+p_2(\mathbf{h})\right)\right] \quad (117)$$

$$R_0+R_1+R_2 \leq E\left[C\left(p_1(\mathbf{h})+p_2(\mathbf{h})+p_0(\mathbf{h})\right)\right] \quad (118)$$

for some $\rho(\mathbf{h})$ that maps the state space to [0,1], and some $p_1(\mathbf{h})$, $p_2(\mathbf{h})$ and $p_0(\mathbf{h})$ that satisfy

$$E\left[\frac{p_1(\mathbf{h})}{h_1}+\frac{\rho(\mathbf{h})^2p_0(\mathbf{h})}{h_1}\right] \leq \bar{P}_1 \quad (119)$$

$$E\left[\frac{p_2(\mathbf{h})}{h_2}+\frac{\left(1-\rho(\mathbf{h})\right)^2p_0(\mathbf{h})}{h_2}\right] \leq \bar{P}_2 \quad (120)$$

$$p_0(\mathbf{h}), p_1(\mathbf{h}), p_2(\mathbf{h}) \geq 0. \quad (121)$$

B. *Proof of Lemma 1*

Let $(\mathbf{R}^a, \tilde{P}_1^a, \tilde{P}_2^a)$ and $(\mathbf{R}^b, \tilde{P}_1^b, \tilde{P}_2^b)$ be two elements in set $\mathcal{L}$. To prove that set $\mathcal{L}$ is convex, we need to show that for any $0 \leq \theta \leq 1$, $(\theta\mathbf{R}^a+(1-\theta)\mathbf{R}^b, \theta\tilde{P}_1^a+(1-\theta)\tilde{P}_1^b, \theta\tilde{P}_2^a+(1-\theta)\tilde{P}_2^b)$ is in set $\mathcal{L}$.

For $i=a$ or b, $(\mathbf{R}^i, \tilde{P}_1^i, \tilde{P}_2^i) \in \mathcal{L}$ means that $\mathbf{R}^i \in \mathcal{C}_f(\mathbf{p}^i, \rho^i)$ for some $(\mathbf{p}^i, \rho^i)$ such that

$$E\left[\frac{p_1^i(\mathbf{h})}{h_1}+\frac{\rho^i(\mathbf{h})^2p_0^i(\mathbf{h})}{h_1}\right] \leq \tilde{P}_1^i,$$

$$E\left[\frac{p_2^i(\mathbf{h})}{h_2}+\frac{\left(1-\rho^i(\mathbf{h})\right)^2p_0^i(\mathbf{h})}{h_2}\right] \leq \tilde{P}_2^i \quad (122)$$

$$p_0^i(\mathbf{h}), p_1^i(\mathbf{h}), p_2^i(\mathbf{h}) \geq 0, \quad 0 \leq \rho^i(\mathbf{h}) \leq 1 \quad (123)$$

i.e., there exist $(\mathbf{p}^i(\mathbf{h}), \rho^i(\mathbf{h}))$ that satisfy (122) and (123) and

$$R_1^i \leq E\left[C\left(p_1^i(\mathbf{h})\right)\right] \quad (124)$$

$$R_2^i \leq E\left[C\left(p_2^i(\mathbf{h})\right)\right] \quad (125)$$

$$R_1^i+R_2^i \leq E\left[C\left(p_1^i(\mathbf{h})+p_2^i(\mathbf{h})\right)\right] \quad (126)$$

$$R_0^i+R_1^i+R_2^i \leq E\left[C\left(p_1^i(\mathbf{h})+p_2^i(\mathbf{h})+p_0^i(\mathbf{h})\right)\right]. \quad (127)$$

Let

$$p_1(\mathbf{h}) = \theta p_1^a(\mathbf{h})+(1-\theta)p_1^b(\mathbf{h}) \quad (128)$$

$$p_2(\mathbf{h}) = \theta p_2^a(\mathbf{h})+(1-\theta)p_2^b(\mathbf{h}) \quad (129)$$

$$p_0(\mathbf{h}) = \theta p_0^a(\mathbf{h})+(1-\theta)p_0^b(\mathbf{h}) \quad (130)$$

$$\rho_1(\mathbf{h}) = \sqrt{\frac{\theta p_0^a(\mathbf{h})\rho^a(\mathbf{h})^2+(1-\theta)p_0^b(\mathbf{h})\rho^b(\mathbf{h})^2}{\theta p_0^a(\mathbf{h})+(1-\theta)p_0^b(\mathbf{h})}} \quad (131)$$

$$1-\rho_2(\mathbf{h}) = \sqrt{\frac{\theta p_0^a(\mathbf{h})(1-\rho^a(\mathbf{h}))^2+(1-\theta)p_0^b(\mathbf{h})(1-\rho^b(\mathbf{h}))^2}{\theta p_0^a(\mathbf{h})+(1-\theta)p_0^b(\mathbf{h})}}. \quad (132)$$

It is straightforward to verify that $\rho_1(\mathbf{h}) \geq \rho_2(\mathbf{h})$ for all possible $\mathbf{h}$, θ, $p_0^a(\mathbf{h})$, $p_0^b(\mathbf{h})$, $\rho^a(\mathbf{h})$, $\rho^b(\mathbf{h})$. Due to the concavity of the log function

$$\theta R_1^a+(1-\theta)R_1^b \leq E\left[C\left(p_1(\mathbf{h})\right)\right] \quad (133)$$

$$\theta R_2^a+(1-\theta)R_2^b \leq E\left[C\left(p_2(\mathbf{h})\right)\right] \quad (134)$$

$$\left(\theta R_1^a+(1-\theta)R_1^b\right)+\left(\theta R_2^a+(1-\theta)R_2^b\right) \leq E\left[C\left(p_1(\mathbf{h})+p_2(\mathbf{h})\right)\right] \quad (135)$$

$$\left(\theta R_0^a+(1-\theta)R_0^b\right)+\left(\theta R_1^a+(1-\theta)R_1^b\right)+\left(\theta R_2^a+(1-\theta)R_2^b\right) \leq E\left[C\left(p_1(\mathbf{h})+p_2(\mathbf{h})+p_0(\mathbf{h})\right)\right]. \quad (136)$$

Also, it is easy to check that

$$E\left[\frac{p_1(\mathbf{h})}{h_1}+\frac{\rho_1(\mathbf{h})^2p_0(\mathbf{h})}{h_1}\right] \leq \theta\tilde{P}_1^a+(1-\theta)\tilde{P}_1^b \quad (137)$$

$$E\left[\frac{p_2(\mathbf{h})}{h_2}+\frac{\left(1-\rho_1(\mathbf{h})\right)^2p_0(\mathbf{h})}{h_2}\right] \leq E\left[\frac{p_2(\mathbf{h})}{h_2}+\frac{\left(1-\rho_2(\mathbf{h})\right)^2p_0(\mathbf{h})}{h_2}\right] \leq \theta\tilde{P}_2^a+(1-\theta)\tilde{P}_2^b. \quad (138)$$

From (133)–(138), we see that $\theta\mathbf{R}^a+(1-\theta)\mathbf{R}^b \in \mathcal{C}_f([p_1,p_2,p_0]^T, \rho_1)$. Also, $p_0, p_1, p_2 \geq 0$, $0 \leq \rho_1 \leq 1$ and satisfy the power constraints of $\theta\tilde{P}_1^a+(1-\theta)\tilde{P}_1^b$ and $\theta\tilde{P}_2^a+(1-\theta)\tilde{P}_2^b$.

Hence

$$\theta\mathbf{R}^a+(1-\theta)\mathbf{R}^b \in \mathcal{C}\left(\theta\tilde{P}_1^a+(1-\theta)\tilde{P}_1^b, \theta\tilde{P}_2^a+(1-\theta)\tilde{P}_2^b\right) \quad (139)$$

and $(\theta\mathbf{R}^a+(1-\theta)\mathbf{R}^b, \theta\tilde{P}_1^a+(1-\theta)\tilde{P}_1^b, \theta\tilde{P}_2^a+(1-\theta)\tilde{P}_2^b) \in \mathcal{L}$ as desired. Thus, $\mathcal{L}$ is convex.

REFERENCES

[1] N. Liu and S. Ulukus, "Ergodic capacity region of fading Gaussian multiple access channels with common data," in *Proc. 42nd Annu. Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sep. 2004, pp. 773–782.

[2] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, Sep. 1973.

[3] V. V. Prelov and E. C. van der Meulen, "Asymptotic expansion for the capacity region of the multiple-access channel with common information and almost Gaussian noise," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 1991, p. 300.

[4] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[5] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[6] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 155–170, Jan. 2006.

[7] A. Murugan, P. Gopala, and H. El Gamal, "Correlated sources and wireless channels: Cooperative source-channel coding," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 988–998, Aug. 2004.

[8] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[9] I. Maric, R. D. Yates, and G. Kramer, "The discrete memoryless compound multiple access channel with conferencing encoders," in *Proc. IEEE Int. Symp. Inf. Theory*, Sep. 2005, pp. 407–410.

[10] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2000, p. 22.

[11] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," in *Proc. IEEE INFOCOM*, New York, NY, Jun. 2002, pp. 1577–1586.

[12] A. F. Dana, M. Sharif, R. Gowaikar, B. Hassibi, and M. Effros, "Is broadcast channel plus multiaccess optimal for Gaussian wireless networks?," in *Proc. 37th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2003, pp. 1748–1752.

[13] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[14] G. Dueck, "A note on the multiple access channels with correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 2, pp. 232–235, Mar. 1981.

[15] W. Kang and S. Ulukus, "A single-letter upper bound for the sum rate of multiple access channels with correlated sources," in *Proc. 39th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct. 2005, pp. 1732–1736.

[16] ——, "An outer bound for multiple access channels with correlated sources," in *Proc. 40th Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 2006, pp. 240–244.

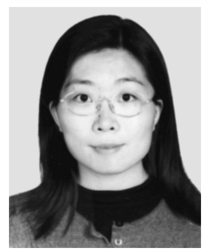
[17] S. S. Pradhan, S. Choi, and K. Ramchandran, "An achievable rate region for multiple access channels with correlated messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2004, p. 108.

[18] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

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

[20] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.



Nan Liu was born in Dalian, China, on December 17, 1978. She received the B.E. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2001. She is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Maryland, College Park.

Her research interests include network information theory and wireless communication theory.

Sennur Ulukus received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1991 and 1993, respectively, and the Ph.D. degree in electrical and computer engineering from Rutgers University, New Brunswick, NJ, in 1998.

During her Ph.D. studies, she was with the Wireless Information Network Laboratory (WINLAB), Rutgers University. From 1998 to 2001, she was a Senior Technical Staff Member at AT&T Labs-Research in NJ. In 2001, she joined the University of Maryland at College Park, where she is currently an Associate Professor in the Department of Electrical and Computer Engineering, with a joint appointment to the Institute for Systems Research (ISR). Her research interests are in wireless communication theory and networking, network information theory for wireless networks, and signal processing for wireless communications.

Dr. Ulukus is a recepient of the 2005 NSF CAREER Award, and a co-recepient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications. She serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.