On The Duality of Gaussian Multiple-Access and Broadcast Channels

Nihar Jindal, Sriram Vishwanath, and Andrea Goldsmith

Department of Electrical Engineering

Stanford University, Stanford, CA 94305

njindal, sriram, andrea@systems.stanford.edu

August 29, 2002

Abstract

We define a duality between Gaussian multiple-access channels and Gaussian broadcast channels. The dual channels we consider have the same channel gains and the same noise power at all receivers. We show that the capacity region of a broadcast channel (both constant and fading) can be written in terms of the capacity region of the dual multiple-access channel, and vice versa. We can use this result to find the capacity region of the MAC if the capacity region of only the broadcast channel is known, and vice versa. For fading channels we show duality under ergodic capacity, but duality also holds for different capacity definitions for fading channels such as outage capacity and minimum rate capacity. Using duality, many results known for only one of the two channels can be extended to the dual channel as well.

Keywords

Broadcast Channel, Multiple-Access Channel, Fading Channels, Capacity Region, Duality

I. Introduction

The Gaussian multiple-access channel (MAC) and the Gaussian broadcast channel (BC) have been studied extensively over the years, but no relationship between the two channels has previously been discovered. In this paper we show that the scalar[1] Gaussian MAC and BC are duals of each other and as a result the capacity regions of the BC and the MAC with the same channel gains (i.e. the channel gain of receiver j in the BC equals the channel gain of transmitter j in the MAC) and the same noise power at every receiver (i.e. the receiver in the MAC and each receiver in the BC have the same noise power) are very closely related.

The Gaussian MAC and the Gaussian BC have two fundamental differences. In the MAC, *each transmitter* has an individual power constraint, whereas in the BC there is only a single power constraint on the transmitter. In addition, signal and interference come from *different* transmitters in the MAC and are therefore multiplied by *different* channel gains (known as the near-far effect) before being received, whereas in the BC the entire received signal comes from the same source and therefore has the same channel gain.

Though the channels differ in some fundamental aspects, there is a striking similarity between the coding/decoding scheme used to achieve the capacity of the Gaussian MAC and BC. In the MAC, each user transmits Gaussian codewords which are scaled by the channel and then "added" in the air. Decoding is done using successive decoding with interference cancellation, in which one user's codeword is decoded

[1] A similar duality is established between the multiple-input, multiple-output (MIMO) MAC and BC in [1, 2]. This result is briefly discussed in Section VIII-A

and then subtracted from the received signal, then the next user is decoded and subtracted out, etc. For the Gaussian BC, superposition coding is optimal. Superposition coding, which involves superimposing information for stronger receivers on top of information intended for weaker users, is optimal in general for degraded broadcast channels, and is therefore optimal for the Gaussian BC which is a degraded broadcast channel. In the Gaussian BC, superposition coding simplifies to transmitting the sum of independent Gaussian codewords (one codeword per user). The receivers also perform successive decoding with interference cancellation, with the caveat that each user can only decode and subtract out the codewords of users with smaller channel gains than themselves. In both the MAC and BC, the received signal is a sum of Gaussian codewords and successive decoding with interference cancellation is performed. The similarity in the encoding/decoding process for the Gaussian MAC and BC hints at the relationship between the channels.

We first show that the capacity region of a K-user Gaussian MAC is a subset of the capacity region of the dual Gaussian BC with power constraint equal to the *sum* of the MAC power constraints. This result holds for a constant (AWGN) or fading channel and is the fundamental relationship between the capacity regions of the dual channels. Building on this relationship, we derive an expression for the capacity region of the BC in terms of the dual MAC which shows that the BC capacity region is equal to the *union* of dual MAC capacity regions over all power constraint vectors such that their sum is equal to the BC power constraint. If we let $\mathcal{C}_{BC}(P;\boldsymbol{h})$ represent the BC capacity region and $\mathcal{C}_{MAC}(P_1,\dots,P_K;\boldsymbol{h})$ represent the MAC capacity region, this can be stated as:

$$\mathcal{C}_{BC}(\overline{P};\boldsymbol{h}) = \bigcup_{\{P_i\}_1^K:\ \sum_{i=1}^K P_i=\overline{P}} \mathcal{C}_{MAC}(P_1,\dots,P_K;\boldsymbol{h}).$$

This leads to the conclusion that the uplink (MAC) and downlink (BC) channels differ only due to the fact that power constraints are placed on each transmitter in the MAC instead of on all transmitters.

We also find an expression for the capacity region of the MAC in terms of the capacity region of the dual BC which shows that the MAC capacity region is equal to an *intersection* of dual BC capacity regions. Using the same notation as above, this can be expressed as:

$$\mathcal{C}_{MAC}(\overline{P}_1,\dots,\overline{P}_K;\boldsymbol{h}) = \bigcap_{\{\alpha_i\}_{i=1}^K:\ \alpha_i>0} \mathcal{C}_{BC}\left(\sum_{i=1}^K \frac{\overline{P}_i}{\alpha_i};\sqrt{\alpha_1}h_1,\dots,\sqrt{\alpha_K}h_K\right).$$

In addition to the constant AWGN channel, we consider flat-fading channels with perfect channel state information (CSI) at all transmitters and receivers. We show that the duality which holds for constant channels also holds for ergodic capacity of fading channels. In addition we show that duality holds for a number of different capacity definitions, namely outage capacity and minimum rate capacity. Though the ergodic capacity regions [3] [4] and outage capacity regions [5] [6] of both the MAC and BC have previously been found, the duality ties these previously independent results together. Minimum rate capacity has only been found for the BC [7], but using duality we can find the minimum rate capacity of the MAC as well. It appears to be very challenging to characterize the minimum rate capacity of the MAC directly because it is a combination of ergodic and zero-outage capacity, but duality provides a nice way of avoiding this difficulty by using the solution to the dual channel instead. We also consider frequency-selective channels (channels with inter-symbol interference, or ISI) and we show that duality holds for the BC and MAC with ISI as well.

The remainder of this paper is organized as follows. In Section II we describe the dual Gaussian BC and MAC. In Section III we show that the constant Gaussian BC and MAC are duals. In Section IV we extend the results on constant channels to fading channels and show that the fading BC and MAC are also duals with respect to ergodic capacity. In Sections V and VI we show duality also holds for outage and minimum rate capacity, respectively. Frequency-selective channels are examined in Section VII. We consider some extensions of this duality in Section VIII, followed by our conclusions.

II. System Model

The notation used in this paper is as follows: Boldface is used to denote vectors. $\mathbb{E}_H$ is used to denote expectation over the random variable H and lower case h denotes a realization of H. For vectors $\boldsymbol{h}$ and $\boldsymbol{\alpha}$, we use $\boldsymbol{\alpha h}$ and $\frac{\boldsymbol{\alpha}}{\boldsymbol{h}}$ to refer to component-wise multiplication and division. Additionally, inequalities with respect to vectors are also component-wise.

We consider two different discrete time systems as shown in Fig. 1, where i denotes the time index. The system to the left is a broadcast channel: a one-to-many system, where the transmitter sends independent information to each receiver by broadcasting signal $X[i]$ to K different receivers simultaneously. Each receiver is assumed to suffer from flat-fading, i.e. the desired signal $X[i]$ is multiplied by a possibly time-varying channel gain[2] $h_j[i]$, and white Gaussian noise $n_j[i]$ is added to the received signal. We let $\boldsymbol{h}[i] = (h_1[i], \dots, h_K[i])$ denote the vector of channel gains at time i.

The system to the right is a multiple access channel: a many-to-one system, where K independent transmitters each sends a signal $X_j[i]$ to a single receiver. The received signal is the sum of the K transmitted signals (each scaled by the channel gain) and additive Gaussian noise $n[i]$.

[2] In general the channel gain may be complex, but assuming perfect phase information at the receivers, without loss of generality we consider real channel gains only.

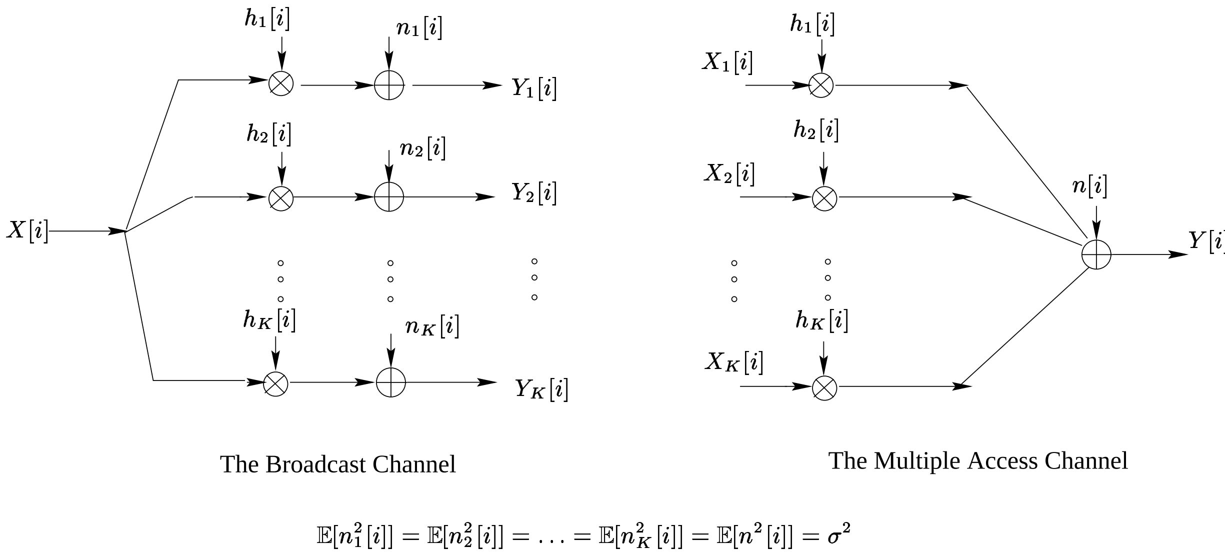


Fig. 1. System Models

Mathematically, the two systems can be described as:

$$\text{BC: } Y_j[i] = h_j[i]X[i] + n_j[i] \qquad\qquad \text{MAC: } Y[i] = \sum_{j=1}^{K} h_j[i]X_j[i] + n[i].$$

Notice that the noise power of each receiver in the BC and the single receiver in the MAC is equal to σ^2. Also note that the term $h_j[i]$ is the channel gain of receiver j in the BC (downlink) and of transmitter j in the MAC (uplink). We call this BC the *dual* of the MAC, and vice versa.

We consider two different models in this paper: constant and time-varying channel gains. In the constant channel, the channel gains $h_j[i]$ are constant for all i and these values are assumed to be known at all the transmitters and the receivers in the MAC and BC. In the fading channel, the channel gains $(H_1[i], \ldots, H_K[i])$ are a jointly stationary and ergodic random process. The dual fading channels actually need to only have the same fading distribution (as opposed to having the same instantaneous channel gains) because the realization of the fading process does not affect the ergodic capacity region. In this paper we assume perfect CSI at all transmitters and receivers, i.e. that all transmitters and receivers know $\boldsymbol{h}[i]$ perfectly at time i.

The dual channels we consider are not only interesting from a conceptual standpoint, but in fact they also resemble a time-division duplexed (TDD) cellular system quite well. In such a system, the channel gains on the uplink and downlink are identical, assuming that the channels are not changing too rapidly. The assumption of equal noise power at all mobiles (i.e. receivers in the MAC) is also quite realistic, but the noise power at the base station (i.e. the receiver in the BC) may or may not be equal to the noise power at the mobiles.

III. DUALITY OF THE CONSTANT MAC AND BC

Before establishing the duality of the constant MAC and BC, we first formally define the capacity regions of both channels.

A. Capacity Region of the Multiple-Access Channel

From [8], the capacity region of a Gaussian multiple-access channel with channel gains $\boldsymbol{h} = (h_1, \ldots, h_K)$ and power constraints $\overline{\boldsymbol{P}} = (\overline{P_1}, \ldots, \overline{P_K})$ is

$$\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{h}) = \left\{ \mathbf{R} : \sum_{j \in S} R_j \leq \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j \in S} h_j^2 \overline{P}_j \right) \quad \forall S \subseteq \{1, \ldots, K\} \right\}. \tag{1}$$

The capacity region of the constant MAC is a K-dimensional polyhedron, and successive decoding with interference cancellation can achieve all corner points of the capacity region [8]. Every decoding order corresponds to a different corner point of the capacity region, and consequently there are $K!$ corner points in the capacity region. Given a decoding order $(\pi(1), \pi(2), \ldots, \pi(K))$ in which User $\pi(1)$ is decoded first, User $\pi(2)$ is decoded second, etc., the rates of the corresponding corner point are:

$$R_{\pi(j)} = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 \overline{P}_{\pi(j)}}{\sigma^2 + \sum_{i=j+1}^{K} h_{\pi(i)}^2 \overline{P}_{\pi(i)}} \right) \quad j = 1, \ldots, K. \tag{2}$$

We will use this form of the rates throughout this paper. The capacity region of the MAC is in fact equal to the convex hull of these $K!$ corner points and other smaller rate points corresponding to the rate equations in (2) but with the powers of any subset of users set to zero.

B. Capacity Region of the Broadcast Channel

From [8], the capacity region of a Gaussian broadcast channel with channel gains $\boldsymbol{h} = (h_1, \ldots, h_K)$ and power constraint $\overline{P}$ is

$$\mathcal{C}_{BC}(\overline{P}; \boldsymbol{h}) = \left\{ \mathbf{R} : R_j \leq \frac{1}{2} \log \left(1 + \frac{h_j^2 P_j^B}{\sigma^2 + h_j^2 \sum_{k=1}^{K} P_k^B \mathbf{1}[h_k > h_j]} \right) \quad j = 1, \ldots, K \right\} \tag{3}$$

over all power allocations such that $\sum_{j=1}^{K} P_j^B = \overline{P}$. Additionally, any rate vector taking the form of (3) with equality lies on the boundary of the capacity region.

Any set of rates in the capacity region is achievable by successive decoding with interference cancellation, in which users decode and subtract out signals intended for other users before decoding their own signal. To achieve the boundary points of the BC capacity region, the signals are encoded such that the strongest user can decode all users' signals, the second strongest user can decode all users' signals except for the strongest user's signal , etc. The "strongest" user refers to the user with the largest channel gain magnitude, or the largest value of h_i^2.

Though successive decoding can achieve any point in the broadcast channel capacity region, we will find it convenient to consider rate vectors (some of which may lie in the interior of the capacity region) achieved by use of "dirty-paper coding" to simplify the MAC-BC duality. As seen in [9], the capacity region of the broadcast channel is also achievable via "dirty-paper coding", in which the transmitter essentially "pre-subtracts" (similar to pre-coding) certain users' codewords (instead of receivers decoding and subtracting out other users' signals). When using this technique, the codeword of the weakest user is first selected. Then the codeword of the next weakest user is selected, but since the weakest user's codeword is already known, it can be pre-subtracted (using dirty paper coding [10] or coding for known interference techniques [11]) such that the second receiver does not experience any interference from the signal intended for the weakest user. Using this procedure for all users, it is clear that the strongest user is the last user to be encoded and his encoding will be done with knowledge of all other users' codewords. This allows the codewords of all other users to be pre-subtracted at the transmitter so that the strongest receiver experiences no interference (similar to the user who is *decoded* last in successive decoding). This technique is optimal when the strongest user is encoded last. This optimal strongest-user-last *encoding* order for dirty paper coding is the same as the optimal *decoding* order used in successive decoding. which is the same order as is used in successive decoding.

The rates achieved by dirty-paper coding (abbreviated as DPC) and successive decoding are identical when the optimal strongest-user-last encoding/decoding order is used. The two schemes do differ, however, when sub-optimal decoding/encoding orders are used. Though using a sub-optimal decoding/encoding order is strictly sub-optimal in the sense that the rates achievable are strictly in the interior of the capacity region, we will find it convenient to consider such points when establishing the duality between the MAC and BC. Since the signals are "pre-subtracted" at the transmitter when using dirty-paper coding, any encoding order can be used without changing the structure of the rate equations. With successive decoding, however, when a sub-optimal decoding order is used it must be ensured that all users who are supposed to decode and subtract out a certain user's signal are actually able to do so. This in fact limits the rates achievable using successive decoding with a sub-optimal decoding order. For this reason, we will consider the rates achieved by using a sub-optimal encoding order with dirty paper coding (instead of

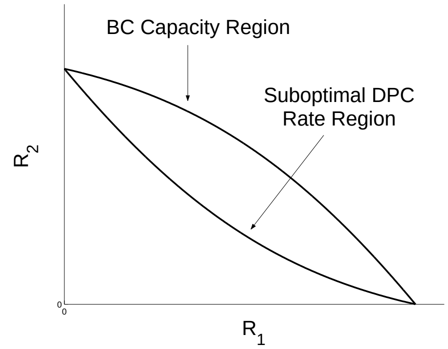


Fig. 2. Capacity region and suboptimal encoding order DPC rate region for BC with $h_1^2 > h_2^2$.

successive decoding) and we will soon see that these sub-optimal points correspond to boundary points of the dual MAC capacity region. In Fig. 2, the capacity region of a broadcast channel in which $h_1^2 > h_2^2$ is shown. The suboptimal region corresponding to the rates achieved under DPC when the weaker of the two users is encoded last is also shown in the figure. This suboptimal region is seen to be considerably smaller than the actual capacity region, but the reasons for considering such a region will become clear in Section III-C. The successive decoding region corresponding to the suboptimal weakest-user-last decoding order is even smaller than the suboptimal DPC region.

Assuming encoding order $(\pi(1), \pi(2), \ldots, \pi(K))$ in which the codeword of User $\pi(1)$ is encoded first, the rates achieved in the BC are:

$$R^B_{\pi(j)} = \frac{1}{2}\log\left(1 + \frac{h^2_{\pi(j)} P^B_{\pi(j)}}{\sigma^2 + h^2_{\pi(j)} \sum_{i=j+1}^K P^B_{\pi(i)}}\right). \tag{4}$$

Clearly these rates are achievable and thus are in the BC capacity region. In fact, any rates of the form of (4) for any encoding order $\pi(\cdot)$ and any power allocation such that $\sum_{i=1}^N P_i^B = \overline{P}$ lie in $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{h})$. Notice that the rate equations in (4) are almost identical to the rate equations along the boundary of the capacity region, the only difference being the order of encoding/decoding and the subsequent interference at each receiver. Additionally, notice the similarity between the above rate equations and the rate equations for the successive decoding corner points of the MAC in (2).

C. Subset Relationship between MAC and BC

We now establish the fundamental relationship between the capacity region of the dual MAC and BC. We later use the ideas of this result to characterize the capacity region of the BC in terms of the dual MAC capacity region.

Theorem 1: The capacity region of a constant Gaussian MAC with power constraints $\boldsymbol{P} = (P_1, \ldots, P_K)$

is a subset of the capacity region of the dual BC with power constraint $P = \mathbf{1} \cdot \boldsymbol{P}$:

$$\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h}) \subseteq \mathcal{C}_{BC}(\mathbf{1}\cdot\boldsymbol{P}; \boldsymbol{h}). \tag{5}$$

Furthermore, the boundaries of the two regions intersect at exactly one point if the channel gains of all K users are distinct ($h_i^2 \neq h_j^2$ for all $i \neq j$).

Proof: Due to the convexity of the BC capacity region and because the MAC capacity region is equal to the convex hull of successive decoding points in the form of (2), it suffices to show that every successive decoding point in the MAC is also in the dual BC capacity region. We will show that every corner point of the MAC capacity region is achievable in the dual BC using the same sum power and dirty paper coding (which implies it is also achievable using standard successive decoding) using an encoding order in the BC which is *opposite* the decoding order used in the MAC.

Let us consider the successive decoding point of the MAC with power constraints $(P_1^M, \ldots, P_K^M)$ corresponding to decoding order $(\pi(1), \ldots, \pi(K))$ for some permutation $\boldsymbol{\pi}(\cdot)$ of $(1, \ldots, K)$. The rate of User $\pi(j)$ in the MAC at this successive decoding point is

$$R_{\pi(j)}^M = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^M}{\sigma^2 + \sum_{i=j+1}^K h_{\pi(i)}^2 P_{\pi(i)}^M} \right).$$

Assuming that the opposite encoding order is used in the BC (i.e. User $\pi(1)$ encoded last, etc.), the rate of User $\pi(j)$ in the dual BC when powers $(P_1^B, \ldots, P_K^B)$ are used is

$$R_{\pi(j)}^B = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^B}{\sigma^2 + h_{\pi(j)}^2 \sum_{i=1}^{j-1} h_{\pi(i)}^2 P_{\pi(i)}^B} \right).$$

By defining A_j and B_j as

$$A_j = \sigma^2 + h_{\pi(j)}^2 \sum_{i=1}^{j-1} P_{\pi(i)}^B, \quad B_j = \sigma^2 + \sum_{i=j+1}^{K} h_{\pi(i)}^2 P_{\pi(i)}^M, \tag{6}$$

we can rewrite the rates in the MAC and BC as

$$R_{\pi(j)}^M = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^M}{\sigma^2 + \sum_{i=j+1}^K h_{\pi(i)}^2 P_{\pi(i)}^M} \right) = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^M}{B_j} \right) \tag{7}$$

$$R_{\pi(j)}^B = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^B}{\sigma^2 + h_{\pi(j)}^2 \sum_{i=1}^{j-1} P_{\pi(i)}^B} \right) = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)}^2 P_{\pi(j)}^B}{A_j} \right). \tag{8}$$

If we define $P_{\pi(j)}^B$ as:

$$P_{\pi(j)}^B = P_{\pi(j)}^M \frac{A_j}{B_j}, \quad j = 1, \ldots, K \tag{9}$$

clearly we will have $R^B_{\pi(j)} = R^M_{\pi(j)}$ for all j. From (9) it is clear that $P^B_j \geq 0$ for all users, and in Appendix A we show that $\sum_{j=1}^K P^M_j = \sum_{j=1}^K P^B_j$.

We will refer to the mapping described in (9) as the MAC-BC transformation. From the mapping in (9), it is not clear that the transformations can actually be performed because it seems as if calculating the broadcast powers depends on other broadcast powers. However, the transformation is indeed constructive and to compute the BC powers, the transformations must be performed in numerical order, starting with user $\pi(1)$:

$$\begin{aligned} P^B_{\pi(1)} &= P^M_{\pi(1)} \frac{\sigma^2}{\sigma^2 + \sum_{i=2}^K h^2_{\pi(i)} P^M_{\pi(i)}} \\ P^B_{\pi(2)} &= P^M_{\pi(2)} \frac{\sigma^2 + h^2_{\pi(2)} P^B_{\pi(1)}}{\sigma^2 + \sum_{i=3}^K h^2_{\pi(i)} P^M_{\pi(i)}} \\ &\cdots \\ P^B_{\pi(K)} &= P^M_{\pi(K)} \frac{\sigma^2 + h^2_{\pi(K)} \sum_{i=1}^{K-1} P^B_{\pi(i)}}{\sigma^2}. \end{aligned}$$

Notice that $P^B_{\pi(1)}$ depends only on the MAC powers, $P^B_{\pi(2)}$ depends on the MAC powers and $P^B_{\pi(1)}$, etc.

Since this transformation works for any decoding order in the MAC, all successive decoding points in the MAC are achievable in the dual BC using the same sum power. Thus we get $\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{h}) \subseteq \mathcal{C}_{BC}(\mathbf{1}\cdot\boldsymbol{P};\boldsymbol{h})$.

It only remains to show that the MAC and dual BC capacity region boundaries meet at exactly one point. Consider the MAC successive decoding point corresponding to the permutation $\pi(.)$ such that $h^2_{\pi(i)} \geq h^2_{\pi(i+1)}$ for $i = 1, \ldots, K-1$, or decoding in order of *decreasing* channel gains (i.e. weakest user last). If we apply the MAC-BC transformation (9), we will find BC powers $(P^B_1, \ldots, P^B_K)$ such that $\sum_{i=1}^K P^B_i = \sum_{i=1}^K P^M_i$ and the same rates are achieved using the opposite decoding order (i.e. strongest user last). From the definition of $\mathcal{C}_{BC}(\overline{P};\boldsymbol{h})$, the rate vector is therefore on the boundary of $\mathcal{C}_{BC}(\sum_{i=1}^K P^B_i;\boldsymbol{h})$. Therefore, the MAC and BC boundaries intersect at this point. In fact, if the channel gains of all K users are distinct, this is the *only* point where the boundaries of the two regions meet [3]. All other corner points of the MAC capacity region lie strictly in the interior of the dual BC capacity region. The proof of this is in Appendix B. ■

Fig. 3 shows the relationship between the capacity regions of the dual MAC and BC described in Theorem 1 for a 2-user channel. Notice that the MAC capacity region lies within the dual BC capacity region and the boundaries of the two capacity regions meet at the corner point of the MAC where the

[3] If all channel gains are not distinct, then the MAC and BC boundaries will meet along a hyperplane.

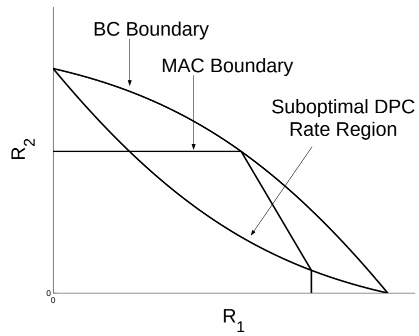


Fig. 3. Capacity regions for a constant MAC and its dual BC with $h_1^2 > h_2^2$.

weaker user (User 2 in the figure) is decoded last. The suboptimal BC region is also shown here to illustrate the purpose of considering suboptimal encoding orders in the BC. The corner point of the MAC capacity region achieved by decoding the stronger user last is exactly equal to the rate vector achieved in the broadcast channel when using dirty-paper coding with the sub-optimal (weaker user last) encoding order. This is apparent in the figure because the boundary of the suboptimal BC region coincides exactly with the second corner point of the MAC capacity region.

If the power constraints $(P_1, \ldots, P_K)$ were varied but their sum $\sum_{i=1}^{K} P_i$ was kept constant, the result would be a *different* MAC capacity region that intersects the dual BC boundary at a *different* point. This is the main idea behind the characterization of the BC capacity region in terms of the dual MAC capacity region.

D. Multiple-Access Channel To Broadcast Channel

In this section we show that the capacity region of a Gaussian BC can be characterized in terms of capacity regions of the dual MAC.

Theorem 2: The capacity region of a constant Gaussian BC with power constraint $\overline{P}$ is equal to the union of capacity regions of the dual MAC with power constraints $(P_1, \ldots, P_K)$ such that $\sum_{j=1}^{K} P_j = \overline{P}$:

$$\mathcal{C}_{BC}(\overline{P}; \boldsymbol{h}) = \bigcup_{\{\boldsymbol{P}:\ \mathbf{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h}). \tag{10}$$

Proof: From Theorem 1, we have $\mathcal{C}_{BC}(\mathbf{1}\cdot\boldsymbol{P}; \boldsymbol{h}) \supseteq \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h})$, which implies $\mathcal{C}_{BC}(P; \boldsymbol{h}) \supseteq \bigcup_{\{\boldsymbol{P}:\ \mathbf{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h})$. It thus remains to show that this inequality also holds in the opposite direction, or that $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{h}) \subseteq \bigcup_{\{\boldsymbol{P}:\ \mathbf{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h})$. We will do this by using the transformations in (9) in the opposite direction to get MAC powers from BC powers (termed the BC-MAC transformation).

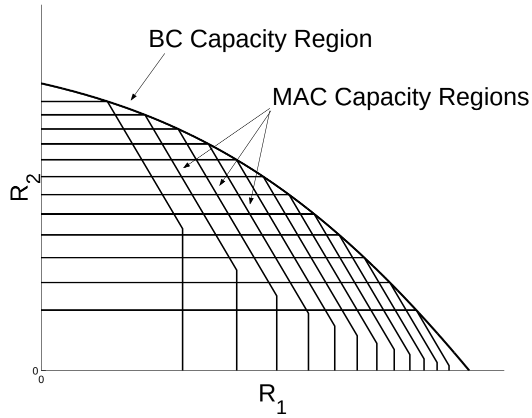


Fig. 4. Constant BC capacity in terms of the dual MAC

The boundary of the BC capacity region is achieved by successive decoding in order of increasing channel gains. If $\pi(.)$ is the permutation such that $h^2_{\pi(i)} \geq h^2_{\pi(i+1)}$ for $i = 1, \ldots, K-1$, then the rate equations in (8) correspond to the rates on the boundary of the BC capacity region. By using the opposite decoding order in the MAC (i.e. decoding in order of decreasing channel gains) and by using MAC powers according to the BC-MAC transformation

$$P^M_{\pi(j)} = P^B_{\pi(j)} \frac{B_j}{A_j}, \quad j = 1, \ldots, K \tag{11}$$

the same rates can be achieved in the BC and MAC using the same sum power (since $\sum_{j=1}^M P_j^M = \sum_{j=1}^M P_j^B$ by Appendix A). Therefore any point on the boundary of the BC capacity region is in the capacity region of the dual MAC with power constraints $(P_1, \ldots, P_K)$ as defined by the BC-MAC transformation. This implies that any point on the boundary of the BC capacity region lies in $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h})$ for some $\boldsymbol{P}$ such that $\mathbf{1} \cdot \boldsymbol{P} = \overline{P}$. ■

The power transformations introduced in (9) and (11) are the key to establishing the relationship between the MAC and BC. The transformations show that any rates achievable in the BC are also achievable in the MAC using the same sum power, and vice versa.

Theorem 2 is illustrated in Fig. 4, where $\mathcal{C}_{MAC}(P_1, P - P_1; h_1, h_2)$ is plotted for different values of P_1. The BC capacity region boundary is shown in bold in the figure. Notice that each MAC capacity region boundary touches the BC capacity region boundary at a *different* point.

If we carefully examine the union expression in the characterization of the BC in terms of the dual MAC in (10), it is easy to see that the union of MAC's is nothing more than the capacity region of the MAC with a *sum* power constraint $P = \sum_{i=1}^K$ instead of *individual* power constraints $(P_1, \ldots, P_K)$.

This is the channel where the transmitters are not allowed to transmit cooperatively (i.e each transmitter encodes its data independently) but the transmitters are allowed to draw from a common power source (i.e. the sum of the transmit power is bounded instead of individual bounds on transmit power). Therefore, Theorem 2 implies that *the capacity region of the MAC with sum power constraint P equals the capacity region of the dual BC with power constraint P*.

Though the capacity regions of the sum power constraint uplink (MAC) and downlink (BC) are equivalent, the optimal decoding orders on the downlink and uplink are the opposite of each other. From the BC-MAC transformations and from Theorem 1, we discover that boundary points of the BC capacity region are achievable in the MAC using successive decoding in order of *decreasing* channel gains. In the BC, it is optimal to give maximum priority to the *strongest* user, whereas in the sum power MAC, it is optimal to give priority to the *weakest* user.

This is easy to understand if we consider a simple two-user MAC where $h_1^2 > h_2^2$. The first user to be decoded will be at a disadvantage because he will experience interference from the other user. However, this allows the second user to be decoded without any interference from the first. Therefore, choosing the correct decoding order is a balance between harming the first user to be decoded and helping the second user. If the weaker user, User 2 in this example, is decoded first, then total interference seen by User 2 would be noise power plus interference from User 1, or $h_1^2 P_1$. Since $h_1^2 > h_2^2$, this interference term could be very large relative to the signal of User 2, which is only $h_2^2 P_2$. This could severely limit the rate of User 2. It would allow User 1 to be decoded without any interference, but since the interference due to the weaker user is not very strong, there is not much to be gained from decoding the weaker user first. If instead User 1 was decoded first in the MAC, the interference from User 2 would not be as large relative to User 1's signal because $h_1^2 > h_2^2$. Therefore decoding User 1 first would not significantly reduce the rate of User 1, but it would significantly increase the rate achievable by the weaker user since the strong interference term will have been removed from the received signal by the time User 2 is decoded. Therefore it intuitively makes sense to decode stronger users before weaker users when considering uplink capacity with a sum power constraint.

In the broadcast channel, on the other hand, the strongest user is optimally decoded last because the strongest user is able to decode all other users' signals due to the degraded nature of the channel. In other words, any signal that can be decoded by a user can also be decoded by any stronger user due to the degraded structure. Therefore, not decoding the strongest user last in the BC is akin to throwing away information that the strongest user has about signals of weaker users.

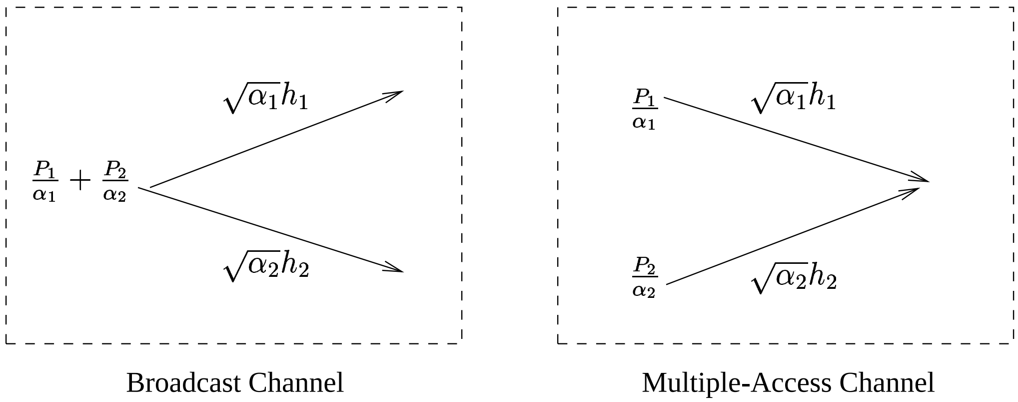


Fig. 5. Scaled dual channels

E. Broadcast Channel to Multiple Access Channel

In this section we show that the capacity region of the MAC can be characterized in terms of the capacity region of the dual BC. In order to derive this relationship, we make use of a concept called *channel scaling*. Since h_j^2 and P_j always appear as a product in the constant MAC capacity expressions (1), we can arbitrarily scale h_j^2 by a constant and scale P_j by the inverse of the constant and still get the same capacity region. Specifically,

$$\begin{aligned}
\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{h}) &= \left\{ \mathbf{R} : \sum_{j\in S} R_j \leq \frac{1}{2}\log\left(1 + \frac{1}{\sigma^2}\sum_{j\in S} h_j^2 P_j\right) \quad \forall S \subseteq \{1,\dots,K\} \right\} \\
&= \left\{ \mathbf{R} : \sum_{j\in S} R_j \leq \frac{1}{2}\log\left(1 + \frac{1}{\sigma^2}\sum_{j\in S} (\sqrt{\alpha_j}h_j)^2 \left(\frac{P_j}{\alpha_j}\right)\right) \quad \forall S \subseteq \{1,\dots,K\} \right\} \\
&= \mathcal{C}_{MAC}\left(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{h}\right) \quad \forall \boldsymbol{\alpha} > 0. \qquad (12)
\end{aligned}$$

The scaled dual channels are shown in Fig. 5. The scaling of the channel and the power constraints clearly negate each other in the multiple-access channel. However, the dual scaled BC is affected by channel scaling and the capacity region of the scaled BC is a function of $\boldsymbol{\alpha}$ since channel scaling affects the power constraint as well as the channel gains of each user relative to all other users.

By applying Theorem 1 to the scaled MAC and the scaled BC, we get $\mathcal{C}_{MAC}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{h}) \subseteq \mathcal{C}_{BC}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{h})$ for all $\boldsymbol{\alpha} > 0$. Since scaling does not affect the MAC capacity region (12), we find that the unscaled MAC capacity region is a subset of the scaled BC capacity region *for all scalings:*

$$\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{h}) \subseteq \mathcal{C}_{BC}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{h}\right) \quad \forall \boldsymbol{\alpha} > 0 \qquad (13)$$

and the boundaries of the MAC capacity region and each scaled BC capacity region intersect. In fact,

$\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{h})$ and $\mathcal{C}_{BC}(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{h})$ intersect at the corner point of the MAC corresponding to decoding in decreasing order of scaled gains $\alpha_i h_i^2$, opposite the optimal decoding order of the *scaled* BC.

In order to characterize the capacity region of the MAC in terms of the BC, we first establish a general theorem (Theorem 3 below) which characterizes *individual transmit power constraint* rate regions of a Gaussian multiple-access channel in terms of *sum transmit power constraint* rate regions. Since the capacity region of the K-user constant MAC is a relatively simple K-dimensional polyhedron, it is quite straightforward using the definitions of the BC and MAC capacity regions to prove that the constant MAC capacity region can be obtained from the capacity region of its dual BC. However, we present a more general theorem here which is also applicable to the more challenging fading channel setting. For the same reasons, our definition below considers a somewhat general notion of a rate region in order to establish a theorem that can be applied to the capacity region of the constant MAC, the ergodic capacity region of the flat-fading MAC, the outage and minimum rate capacity regions of the flat-fading MAC, and the capacity region of the frequency-selective MAC. Before our theorems, we first define the notion of a rate region and the conditions that the rate regions must satisfy in order for the theorems to hold.

Definition 1: Let a K-dimensional rate vector be written as $\boldsymbol{R} = (R_1, \ldots, R_K)$ where R_j is the rate of transmitter j over a given channel to a receiver. Let $\boldsymbol{P} = (P_1, \ldots, P_K)$ be the vector of transmit power constraints and let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ be a vector of scaling constants. We define a rate region $R(\boldsymbol{P})$ as a mapping from a power constraint vector $\boldsymbol{P}$ to a set in $\mathcal{R}_+^K$ which satisfies the conditions stated in Definition 2 below. The $\boldsymbol{\alpha}$-scaled version of the channel is the channel in which the channel gain from transmitter i to the receiver is scaled by $\sqrt{\alpha_i}$. We denote the rate region of the scaled channel as $R_{\boldsymbol{\alpha}}(\boldsymbol{P})$.

Definition 2: We consider K-dimensional rate regions $R(\boldsymbol{P}) \subseteq \mathcal{R}_+^K$ that satisfy the following conditions:

1. $R(\boldsymbol{P}) = R_{\boldsymbol{\alpha}}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}})\ \forall\, \boldsymbol{\alpha} > 0, \boldsymbol{P} > 0$.
2. $S = \{(\boldsymbol{R}, \boldsymbol{P}) | \boldsymbol{P} \in \mathcal{R}_+^K,\ \boldsymbol{R} \in R(\boldsymbol{P})\}$ is a convex set.
3. For all $\boldsymbol{P} \in \mathcal{R}_+^K$, $R(\boldsymbol{P})$ is a closed, convex region.
4. $R(\boldsymbol{P})$ is monotonically increasing, or $\boldsymbol{P_1} \geq \boldsymbol{P_2}$ implies $R(\boldsymbol{P_1}) \supseteq R(\boldsymbol{P_2})$.
5. If $(R_1, \ldots, R_K) \in R(P_1, P_2, \ldots, P_K)$, then $(0, R_2, \ldots, R_K) \in R(0, P_2, \ldots, P_K)$. This condition must hold for all K transmitters.
6. If $\boldsymbol{R} \in R(\boldsymbol{P})$, then $\boldsymbol{R'} \leq \boldsymbol{R}$ implies $\boldsymbol{R'} \in R(\boldsymbol{P})$.
7. $R(\boldsymbol{P})$ is unbounded in every direction as $\boldsymbol{P}$ increases, or $\forall j,\ \max_{R_j \in R(\boldsymbol{P})} R_j \to \infty$ as $P_i \to \infty$.
8. $R(\boldsymbol{P})$ is finite for all $\boldsymbol{P} > 0$.

These conditions on the rate region $R(\boldsymbol{P})$ are very general and are satisfied by nearly any capacity region or rate region. Finally, we define the notion of a sum power constraint rate region.

Definition 3: For any scaling $\boldsymbol{\alpha}$,we define the *sum* power constraint rate region $R^{sum}_{\boldsymbol{\alpha}}(P_{sum})$ as:

$$R^{sum}_{\boldsymbol{\alpha}}(P_{sum}) \triangleq \bigcup_{\{\boldsymbol{P} |\ \boldsymbol{P} \in \mathcal{R}^{K}_{+},\ \mathbf{1} \cdot \boldsymbol{P} \leq P_{sum}\}} R_{\boldsymbol{\alpha}}(\boldsymbol{P}). \tag{14}$$

Having established these definitions, we now state a theorem about rate regions and channel scaling.

Theorem 3: Any rate region $R(\boldsymbol{P})$ satisfying the conditions of Definition 2 is equal to the intersection over all strictly positive scalings of the sum power constraint rate regions for any strictly positive power constraint vector $\boldsymbol{P}$:

$$R(\boldsymbol{P}) = \bigcap_{\boldsymbol{\alpha} > 0} R^{sum}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}} \right). \tag{15}$$

Proof: Appendix D contains the proof of a multi-receiver generalization (Theorem 14) of this result, but the proof applies to the MAC if we set N, the number of receivers, to 1. ■

Applying Theorem 3 to the capacity region of the constant MAC, we can use the fact that the sum power constraint MAC is equal to the capacity region of the dual BC to characterize the individual power constraint MAC in terms of the scaled dual BC.

Theorem 4: The capacity region of a constant Gaussian MAC is equal to the intersection of the capacity regions of the scaled dual BC over all possible channel scalings:

$$\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h}) = \bigcap_{\boldsymbol{\alpha} > 0} \mathcal{C}_{BC}(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}). \tag{16}$$

Proof: To prove this theorem, we first invoke Theorem 3 to establish

$$\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h}) = \bigcap_{\boldsymbol{\alpha} > 0} \mathcal{C}^{sum}_{MAC}(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}). \tag{17}$$

In Appendix C we show that the region $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{h})$ satisfies the conditions of Definition 2, as required by Theorem 3. By Theorem 2, we have

$$\mathcal{C}_{BC}(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}) = \bigcup_{\{\hat{\boldsymbol{P}}:\ \mathbf{1} \cdot \hat{\boldsymbol{P}} = \mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\}} \mathcal{C}_{MAC}(\hat{\boldsymbol{P}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}) \tag{18}$$

for all $\boldsymbol{\alpha} > 0$. Due to the fact that $\mathcal{C}_{MAC}(\boldsymbol{P}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h})$ is an increasing function of power, we can equivalently write the union over all power vectors that do not exceed the sum power constraint (instead of meeting the constraint exactly). From Definition 3, we then see that

$$\mathcal{C}^{sum}_{MAC}(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}) = \mathcal{C}_{BC}(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}} \boldsymbol{h}) \tag{19}$$

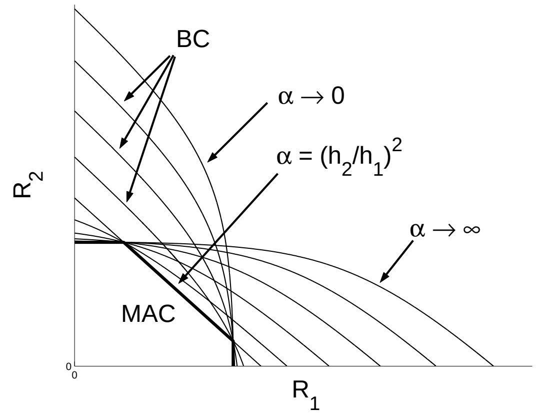


Fig. 6. Constant MAC capacity in terms of the dual BC

for all $\boldsymbol{\alpha} > 0$. By substituting $\mathcal{C}_{BC}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{h})$ for $\mathcal{C}_{MAC}^{sum}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{h})$ in (17), we get the desired result.

■

Theorem 4 is illustrated for a 2-user channel in Fig. 6. Although we consider channel scaling of all K users in Theorem 3, scaling $K-1$ users is sufficient because scaling by $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{K-1}, \alpha_K)$ is equivalent to scaling by $(\frac{\alpha_1}{\alpha_K}, \ldots, \frac{\alpha_{K-1}}{\alpha_K}, 1)$. We therefore let $\alpha_2 = 1$ and only let α_1 (denoted by α in the figure) vary. In the figure we plot $\mathcal{C}_{BC}(\frac{P_1}{\alpha} + P_2; \sqrt{\alpha}h_1, h_2)$ for a range of values of $\alpha > 0$. Note that since the constant MAC region is a pentagon, the BC channel characterized by $\alpha = (h_2/h_1)^2$ and the limit of the BC channels as $\alpha \to 0$ and $\alpha \to \infty$ are sufficient to form the pentagon. When $\alpha = (h_2/h_1)^2$, the channel gains of both users are the same and the BC capacity region is bounded by a straight line segment because the capacity region can be achieved by time-sharing between single-user transmission. This line segment corresponds exactly with the forty-five degree line bounding the MAC capacity region. As $\alpha \to 0$, the total transmit power $\frac{P_1}{\alpha} + P_2$ tends to infinity but the channel gain of User 1 goes to zero. These effects negate each other and cause $R_1 \to \log(1 + \frac{h_1^2 P_1}{\sigma^2})$ and $R_2 \to \infty$. As $\alpha \to \infty$, the total amount of power converges to P_2 and the channel gain of User 1 becomes infinite. This causes $R_1 \to \infty$ and $R_2 \to \log(1 + \frac{h_2^2 P_2}{\sigma^2})$. These two limiting capacity regions bound the vertical and horizontal line segments, respectively, of the MAC capacity region.

Additionally, by Theorem 1, all scaled BC capacity regions except the channel corresponding to $\alpha = (h_2/h_1)^2$ intersect the MAC at exactly one of the two corner points of the MAC region. In fact, all scaled BC capacity regions with $\alpha > (h_2/h_1)^2$ intersect the MAC at the point where user 2 is decoded last in the MAC (i.e. upper left corner), and all scaled BC capacity regions with $\alpha < (h_2/h_1)^2$ intersect the MAC at the corner point where user 1 is decoded last (i.e. lower right corner).

A general K-user constant MAC capacity region is the intersection of $2^K - 1$ hyperplanes (each corresponding to a different subset of $\{1, \dots, K\}$). Therefore, in general, only $2^K - 1$ different scaled BC capacity regions are needed to get the MAC capacity region[4]. One of these regions corresponds to $\boldsymbol{\alpha}$ such that $\alpha_i h_i^2 = \alpha_j h_j^2$ for all i, j. The other necessary scalings correspond to limiting capacity regions as one or more of the components of $\boldsymbol{\alpha}$ are taken to infinity.

IV. Duality of the Fading MAC and BC

We now move on to the flat-fading BC and MAC and show that duality holds for the ergodic capacity regions (subject to an average power constraint) of the dual flat-fading MAC and BC, assuming perfect channel state information (CSI) at all transmitters and receivers. We will soon see that establishing duality for the flat-fading MAC and BC is very similar to the process we used for the constant versions of the channel. We again use the idea of the MAC-BC and BC-MAC transformations, but we have to perform these transformations on a per fading state basis. Before discussing duality, we first formally define the ergodic capacity regions of the fading MAC and BC.

A. Ergodic Capacity Region of the Multiple-Access Channel

We define a power policy $\mathcal{P}_{MAC}$ over all possible fading states as a function that maps from a joint fading state $\boldsymbol{h} = (h_1, \dots, h_K)$ to the transmitted power $P_j^M(\boldsymbol{h})$ for each user. Let $\mathcal{F}_{\mathcal{MAC}}$ denote the set of all power policies satisfying the K individual average power constraints: $\mathcal{F}_{MAC} = \{\mathcal{P}_{MAC} : \mathbb{E}_{\boldsymbol{H}}[P_j^M(\boldsymbol{h})] \le \overline{P}_j \quad 1 \le j \le K\}$.

From Theorem 2.1 of [4], the ergodic capacity region of the multiple-access channel with perfect CSI and power constraints $\overline{\boldsymbol{P}} = (\overline{P}_1, \dots, \overline{P}_K)$ is:

$$\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H}) = \bigcup_{\mathcal{P}_{MAC} \in \mathcal{F}_{MAC}} \mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H}) \tag{20}$$

where

$$\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H}) = \tag{21}$$

$$\left\{ \boldsymbol{R} : \sum_{j \in S} R_j \le \mathbb{E}_{\boldsymbol{H}} \left[\frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j \in S} h_j^2 P_j^M(\boldsymbol{h}) \right) \right], \forall S \subset \{1, \dots, K\} \right\}. \tag{22}$$

Each power policy $\mathcal{P}_{MAC}$ corresponds to a pentagon shaped region $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$, similar to the form of the constant MAC capacity region. By taking the union over all power policies, the ergodic capacity region is seen to have a curved boundary [4]. The optimal MAC power policies are the policies that

[4] Theorem 4 can alternatively be derived based on this idea.

achieve points on the boundary of $\mathcal{C}_{MAC}(\overline{P}_1, \ldots, \overline{P}_K; \boldsymbol{H})$ and every point on the boundary is achieved using a *different* optimal power policy.

The pentagon $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ is the capacity region when powers are allocated according to power policy $\mathcal{P}_{MAC}$. The corner points of $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ correspond to successive decoding with a *fixed* decoding order in all fading states. Alternatively, the same power policy can be used but with a *different* decoding order in every fading state. Let $\pi_{\boldsymbol{h}}(.)$ be the decoding order used in fading state $\boldsymbol{h}$. In fading state $\boldsymbol{h}$, User $\pi_{\boldsymbol{h}}(1)$ is decoded first, etc. To denote the place in the decoding order of User k we use the notation $\pi_{\boldsymbol{h}}^{-1}(k)$. Then the rates

$$R_j \leq \mathbb{E}_{\boldsymbol{H}} \left[R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h})) \right] \qquad j = 1, \ldots, K. \tag{23}$$

are also in $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ for any decoding order function $\pi_{\boldsymbol{h}}$ where the instantaneous rates $R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h}))$ are defined as[5]

$$R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h})) = \frac{1}{2} \log \left(1 + \frac{h_j^2 P_j^M(\boldsymbol{h})}{\sigma^2 + \sum_{k=\pi_{\boldsymbol{h}}^{-1}(j)+1}^{M} h_{\pi_{\boldsymbol{h}}(k)}^2 P_{\pi_{\boldsymbol{h}}(k)}^M(\boldsymbol{h})} \right), \quad j = 1, \ldots, K. \tag{24}$$

The fact that these rate vectors are in $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ can be verified by noticing that the inequalities in the definition of $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ are satisfied in every fading state (because the "rates" in each fading state are defined by these inequalities). Equivalently,

$$\sum_{j \in S} R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h})) \leq \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j \in S} h_j^2 P_j^M(\boldsymbol{h}) \right) \quad \forall S \subset \{1, \ldots, K\}, \ \forall \boldsymbol{h} \tag{25}$$

and therefore the inequalities are satisfied with the expectation over $\boldsymbol{H}$ as in (21). If $\pi_{\boldsymbol{h}}$ is constant for all fading states, then the rates in (23) correspond to one of the $K!$ corner points of $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$. If $\pi_{\boldsymbol{h}}$ is not constant for all fading states, then the rates in (23) lie in the convex hull of the $K!$ corner points.

Due to the convexity of $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ and from the definition of $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H})$, we see that $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H})$ is equal to the convex hull of all rate vectors in the form of (23) over all power policies satisfying the power constraints and all *fixed* decoding order functions. The convex hull operation (i.e. time-sharing) is not needed between different power policies [4], but it may be needed between different fixed decoding orders and the same power policy to achieve points on the boundary of $\mathcal{C}_{MAC}(\mathcal{P}_{MAC}; \boldsymbol{H})$ for some specific power policy[6]. We have also verified that the rates corresponding to non-fixed decoding

[5] Though the idea of instantaneous rates are not needed to define ergodic capacity, they are introduced for simplicity of notation and will be explicitly used in later sections on outage and minimum rate capacity.

[6] Time-sharing between different fixed decoding orders is only necessary when there is a non-zero probability of the fading gains of two users being equal [4].

order functions also lie in the capacity region, which will be of use to us later when establishing duality.

B. Ergodic Capacity Region of the Broadcast Channel

We define a power policy $\mathcal{P}_{BC}$ over all possible fading states as a function over all joint fading states that maps from a joint fading state $\boldsymbol{h} = (h_1, \ldots, h_K)$ to the transmitted power $P_j^B(\boldsymbol{h})$ for each user. Let $\mathcal{F}_{BC}$ denote the set of all power policies satisfying the average power constraint: $\mathcal{F}_{BC} = \{\mathcal{P}_{BC} : \mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^B(\boldsymbol{h})] \leq \overline{P}\}$. Here the expectation is taken over the joint fading distribution $\boldsymbol{H} = (H_1, \ldots, H_K)$. From Theorem 1 of [3], the ergodic capacity region of the BC with perfect CSI and power constraint $\overline{P}$ is the union over all power policies in $\mathcal{F}_{BC}$:

$$\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H}) = \bigcup_{\mathcal{P}_{BC} \in \mathcal{F}_{BC}} \mathcal{C}_{BC}(\mathcal{P}_{BC}; \boldsymbol{H}) \tag{26}$$

where

$$\mathcal{C}_{BC}(\mathcal{P}_{BC}; \boldsymbol{H}) = \left\{ \boldsymbol{R} : R_j \leq \mathbb{E}_{\boldsymbol{H}} \left[\frac{1}{2} \log \left(1 + \frac{h_j^2 P_j^B(\boldsymbol{h})}{\sigma^2 + h_j^2 \sum_{k=1}^K P_k^B(\boldsymbol{h}) \mathbf{1}[h_k^2 > h_j^2]} \right) \right], \quad j = 1, \ldots, K \right\}. \tag{27}$$

The optimal power policies are the policies that achieve points on the boundary of $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$ and every boundary point of $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$ is achieved using a *different* optimal power policy.

Successive decoding can be used in each fading state to achieve any point in the ergodic capacity region. Though the optimal decoding order in the BC is strongest user last (i.e. the strongest user decodes and subtracts out all other signals before decoding his own) in every fading state, a different (and sub-optimal) encoding order can be used in combination with dirty paper coding in each fading state as described for the constant channel in (4). We consider sub-optimal encoding orders because any decoding order is possible in the MAC (since there is no single decoding order which is always optimal for the MAC, as there is for the BC) and allowing for any encoding order in the dual BC helps in establishing the duality between the MAC and BC. Let $\pi_{\boldsymbol{h}}(.)$ be the encoding order used in fading state $\boldsymbol{h}$. In fading state $\boldsymbol{h}$, User $\pi_{\boldsymbol{h}}(1)$ is encoded first, etc. We use the inverse notation $\pi_{\boldsymbol{h}}^{-1}(k)$ (i.e. $\pi_{\boldsymbol{h}}(\pi_{\boldsymbol{h}}^{-1}(k)) = k$) to denote the place in the encoding order of User k. Then the rates

$$R_j \leq \mathbb{E}_{\boldsymbol{H}} \left[R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h})) \right], \quad j = 1, \ldots, K \tag{28}$$

are achievable using power policy $P_j^B(\boldsymbol{h})$ and thus are in $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$ where the instantaneous rates

$R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h}))$ are defined as

$$R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h})) = \frac{1}{2}\log\left(1 + \frac{h_j^2 P_j^B(\boldsymbol{h})}{\sigma^2 + h_j^2 \sum_{k=\pi_{\boldsymbol{h}}^{-1}(j)+1}^{K} P_{\pi_{\boldsymbol{h}}(k)}^B(\boldsymbol{h})}\right), \quad j = 1, \dots, K. \tag{29}$$

This can easily be seen by the following: By keeping the sum power in each fading state constant but reallocating the power between the users and performing successive decoding with the optimum decoding order in each state, the same instantaneous rates can be achieved (and possibly strictly exceeded) for every user in each fading state. Therefore it follows that the same average rates can also be achieved by use of successive decoding and some different power policy in the form of (27).

If the encoding order function in each state goes from the weakest to strongest user (i.e. if the encoding order function satisfies the condition that $\pi_{\boldsymbol{h}}^{-1}(i) > \pi_{\boldsymbol{h}}^{-1}(k)$ implies $h_i^2 \geq h_k^2$ for all fading states and all i, k,) then the encoding order is optimal. Therefore the optimal rates defined in (27) can also be expressed in the form of the capacity definition in (28).

C. Subset Relationship of the Fading MAC and BC

We now show that the ergodic capacity region of the MAC lies within the ergodic capacity region of the dual BC.

Theorem 5: The ergodic capacity region of a flat-fading Gaussian MAC with power constraints $\overline{\boldsymbol{P}} = (\overline{P}_1, \dots, \overline{P}_K)$ is a subset of the ergodic capacity region of the dual BC with power constraint $P = \mathbf{1} \cdot \overline{\boldsymbol{P}}$:

$$\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H}) \subseteq \mathcal{C}_{BC}(\mathbf{1} \cdot \overline{\boldsymbol{P}}; \boldsymbol{H}). \tag{30}$$

Proof: The main idea of this proof is to show that for every MAC power policy there is an equivalent BC power policy derived by using the MAC-BC transformation in every fading state. Any boundary point of $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H})$ is achievable by time-sharing between rate points of the form of (23):

$$R_j = \mathbb{E}_{\boldsymbol{H}}\left[\frac{1}{2}\log\left(1 + \frac{h_j^2 P_j^M(\boldsymbol{h})}{\sigma^2 + \sum_{k=\pi_{\boldsymbol{h}}^{-1}(j)+1}^{K} h_{\pi_{\boldsymbol{h}}(k)}^2 P_{\pi_{\boldsymbol{h}}(k)}^M(\boldsymbol{h})}\right)\right] \quad j = 1, \dots, K. \tag{31}$$

Let us first consider points in the MAC that are achieved without time-sharing, or rates of the form of (31) for some power policy satisfying the power constraints and decoding order function. We wish to show that this same rate is achievable in the dual BC using the same sum power. By applying the MAC-BC transformation *in each fading state*, we can find a BC power policy with the same sum power and a decoding order function (which will be opposite the MAC decoding order function) such that the "instantaneous" rates in the form of (28) of every user are the same as in the MAC.

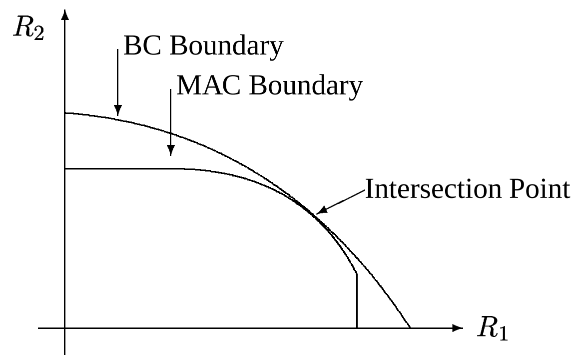


Fig. 7. Capacity regions for the dual fading MAC and BC

If power policy $P_j^M(\boldsymbol{h})$ and decoding order function $\pi_{\boldsymbol{h}}(.)$ are used in the MAC, then the instantaneous rates $R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h}))$ defined in (23) can be achieved. If we define a BC power policy by

$$P^B_{\pi_{\boldsymbol{h}}(j)}(\boldsymbol{h}) = P^M_{\pi_{\boldsymbol{h}}(j)}(\boldsymbol{h}) \frac{A_j(\boldsymbol{h}, \pi_{\boldsymbol{h}})}{B_j(\boldsymbol{h}, \pi_{\boldsymbol{h}})}, \quad j = 1, \ldots, K \tag{32}$$

for all fading states $\boldsymbol{h}$, then by our earlier results on the duality of the constant BC and MAC in Section III the same instantaneous rates can be achieved in the BC using the decoding order opposite of $\pi_{\boldsymbol{h}}(.)$ in every fading state. Of course, this implies that the same average rates are achievable in the BC. Additionally, the power policies will use the same sum power in each fading state, and therefore the same average sum power, or $\mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^M(\boldsymbol{h})] = \mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^B(\boldsymbol{h})]$. Thus we see that for any power policy in the MAC, there is a power policy in the BC with the same sum power that achieves the same average rates.

If we consider more general rate vectors in the MAC that are achieved by time-sharing between different power policies and decoding orders, we can find a dual BC power policy for each power policy used in the MAC. The dual BC power policy will achieve the same instantaneous and average rates as the MAC power policy, and time-sharing can also be performed in the BC to achieve the same overall average rates as in the MAC. This completes the proof. ∎

We conjecture that the boundaries of the ergodic capacity region of the MAC and of the dual BC meet at one point, as they do for the constant channel case (Theorem 1). We are able to show this for the $K = 2$ case, but not for arbitrary K. This point is further discussed in Section IV-F.

Fig. 7 illustrates the subset relationship established in Theorem 5 for the ergodic capacity regions of the dual flat-fading MAC and BC for a 2-user channel. Due to the fading, the ergodic capacity region of the MAC is bounded by straight line segments connected by a curved section as opposed to the pentagon-like capacity region of the constant MAC. The BC and MAC intersect in the curved portion of the MAC boundary.

D. Multiple-Access Channel To Broadcast Channel

We now characterize the ergodic capacity region of the BC in terms of the dual MAC.

Theorem 6: The ergodic capacity region of a fading Gaussian BC with power constraint $\overline{P}$ is equal to the union of ergodic capacity regions of the dual MAC with power constraints $(P_1, \ldots, P_K)$ such that $\mathbf{1}\cdot\boldsymbol{P} = \overline{P}$:

$$\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H}) = \bigcup_{\mathbf{1}\cdot\boldsymbol{P}=\overline{P}} \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H}). \tag{33}$$

Proof: We have already seen that $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H}) \subseteq \mathcal{C}_{BC}(\mathbf{1}\cdot\boldsymbol{P}; \boldsymbol{H})$ for all power constraints $\boldsymbol{P}$. This implies $\bigcup_{\mathbf{1}\cdot\boldsymbol{P}=\overline{P}} \mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H}) \subseteq \mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$. Therefore it only remains to show that $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$ is a subset of the union of dual MAC capacity regions.

Any point in $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$ is achievable by rates of the form:

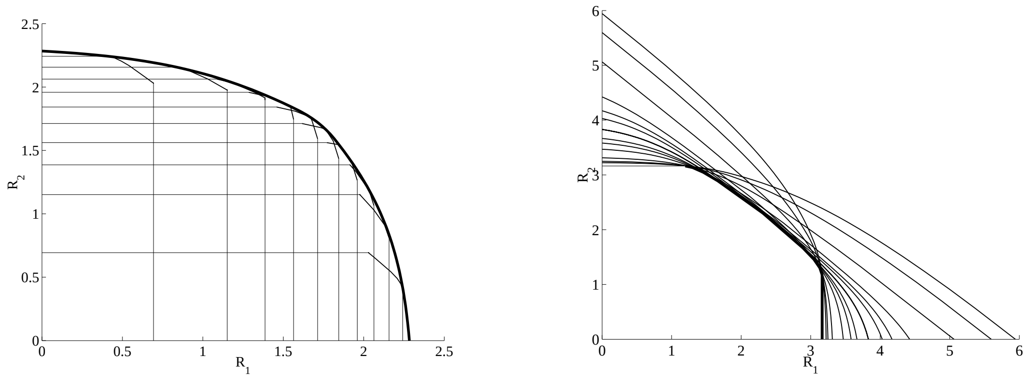
$$R_j \leq \mathbb{E}_{\boldsymbol{H}}\left[\frac{1}{2}\log\left(1 + \frac{h_j^2 P_j^B(\boldsymbol{h})}{\sigma^2 + h_j^2 \sum_{k=\pi_{\boldsymbol{h}}^{-1}(j)+1}^{K} P_{\pi_{\boldsymbol{h}}(k)}^B(\boldsymbol{h})}\right)\right], \quad j = 1, \ldots, K \tag{34}$$

for some power policy and the *optimal* decoding order function (i.e. strongest user last). For any BC power policy, we can find a MAC power policy using the BC-MAC transformation and the opposite decoding order (i.e. weakest user last) that achieves the same instantaneous rates (and average rates) as in the BC, and that uses the same sum power. In other words, we can find a MAC power policy $P_j^M(\boldsymbol{h})$ that achieves the same rates as the BC such that $\mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^M(\boldsymbol{h})] = \mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^B(\boldsymbol{h})]$. If we let $\overline{\boldsymbol{P}_j} = \mathbb{E}_{\boldsymbol{H}}[P_j^M(\boldsymbol{h})]$ for all j, then this rate vector lies in $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H})$. Since this is true for any point in $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H})$, we get $\mathcal{C}_{BC}(\overline{P}; \boldsymbol{H}) \subseteq \bigcup_{\mathbf{1}\cdot\overline{\boldsymbol{P}}=\overline{P}} \mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{H})$. ■

Fig. 8(a) illustrates Theorem 6. As we saw for constant channels, we find that the ergodic capacity region of the MAC with a *sum* power constraint P equals the ergodic capacity region of the dual BC with power constraint P.

E. Broadcast Channel to Multiple Access Channel

In order to characterize the ergodic capacity region of the MAC in terms of the dual BC, we use channel scaling again. Channel scaling by the factor $\sqrt{\boldsymbol{\alpha}}$ for fading channels refers to the ergodic capacity of a channel with power constraints $\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}$ and the fading distribution defined as $\tilde{\boldsymbol{H}} = \sqrt{\boldsymbol{\alpha}}\boldsymbol{H}$. It is easy to verify that the channel scaling relationship in (13) also holds for fading channels, or that $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H}) = \mathcal{C}_{MAC}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{H})$ for all $\boldsymbol{\alpha} > 0$. Combining this result with Theorem 5, we get $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H}) \subseteq \mathcal{C}_{BC}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{H})\ \forall \boldsymbol{\alpha} > 0$. Using Theorem 3, we can find an expression for the ergodic capacity region of the MAC in terms of the dual BC.



(a) Fading MAC to BC

(b) Fading BC to MAC

Fig. 8. Duality of the fading MAC and BC

Theorem 7: The ergodic capacity region of a fading MAC is equal to the intersection of the ergodic capacity regions of the dual BC over all scalings:

$$\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{H}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{BC}(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\sqrt{\boldsymbol{\alpha}}\boldsymbol{H}) \tag{35}$$

Proof: The proof of this is identical to the proof for the constant channel version of this in Theorem 4. The fact that the ergodic capacity region of the MAC satisfies the conditions of Theorem 3 can be verified using the arguments used for the constant MAC capacity region in Appendix C. See Section IV-F for a discussion of Theorem 3 as applied to ergodic capacity. ■

Theorem 7 is illustrated in Figure 8(b). The MAC ergodic capacity region cannot be characterized by a finite number of BC regions as it was for the constant MAC capacity region in Section III-E. The BC capacity regions where $\alpha \to 0$ and $\alpha \to \infty$ still limit the vertical and horizontal line segments of the MAC ergodic capacity region. The curved section of the MAC boundary, however, is intersected by many different scaled BC ergodic capacity regions.

F. Convex Optimization Interpretation

If we consider the boundary points of $\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{H})$ from a convex optimization viewpoint, we can gain some additional insight into the MAC-BC duality. In fact, the proof of Theorem 3 in Appendix D is based on convex optimization. Since the region $\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{H})$ is closed and convex, we can fully characterize the region by the following convex maximization:

$$\max_{\boldsymbol{R}\in\mathcal{C}_{MAC}(\boldsymbol{P};\boldsymbol{H})} \boldsymbol{\mu}\cdot\boldsymbol{R} \quad \text{such that:} \quad \boldsymbol{P} \leq \overline{\boldsymbol{P}} \tag{36}$$

over all non-negative priority vectors $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$ such that $\boldsymbol{\mu} \cdot \mathbf{1} = 1$. In [4] it is shown that a rate vector is on the boundary surface of $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H})$ if and only if it is the solution to the above maximization for some $\boldsymbol{\mu}$. The optimization in (36) is shown in Figure 9 for a 2-user MAC capacity region.

Since (36) is a convex problem, we know that the solution to the original optimization also maximizes the Lagrangian function $\boldsymbol{\mu} \cdot \boldsymbol{R} + \sum_{i=1}^{K} \lambda_i (P_i - \overline{P}_i)$ for the *optimal* Lagrangian multipliers $\boldsymbol{\lambda^*} = (\lambda_1^*, \ldots, \lambda_K^*)$. The optimal Lagrange multipliers $\boldsymbol{\lambda^*}$ can be interpreted as the power prices of the K users, or alternatively λ_i^* is the sensitivity of the function $\boldsymbol{\mu} \cdot \boldsymbol{R}$ to a change in the power constraint $\overline{P_i}$. For each non-negative priority vector $\boldsymbol{\mu}$, there exists an optimum Lagrange multiplier $\boldsymbol{\lambda^*}$. If for some $\boldsymbol{\mu}$ we have $\lambda_1^* = \lambda_2^* = \cdots = \lambda_K^*$, then each power constraint is equally "hard". In other words, each power constraint is equally restrictive, which implies that the solution is sum-power optimal in the sense that having individual power constraints $(\overline{P}_1, \ldots, \overline{P}_K)$ is no more restrictive than having a sum power constraint $\sum_{i=1}^{K} \overline{P}_i$. Therefore, the maximum value of $\boldsymbol{\mu} \cdot \boldsymbol{R}$ in the sum power constraint MAC capacity region and in the individual power constraint MAC capacity region are *equal* for any $\boldsymbol{\mu}$ such that the optimal Lagrangian multipliers are all equal. Since the capacity regions of the sum power constraint MAC and the dual BC are equivalent as established in Theorem 6, this implies that the boundaries of the MAC (with individual power constraints) and the dual BC touch at any point on the MAC boundary where $\lambda_1^* = \lambda_2^* = \cdots = \lambda_K^*$. Therefore, verifying that the MAC and BC capacity region boundaries touch at least one once (as conjectured in Section IV-D) is equivalent to showing that for some $\boldsymbol{\mu}$, the optimal Lagrangian multipliers are all equal.

The optimal Lagrange multipliers are also the key to establishing that the capacity region of the MAC is equal to the intersection of all dual scaled BC capacity regions. The important observation to make with regards to channel scaling is that if $\boldsymbol{\lambda^*}$ is the optimal Lagrange multiplier for some priority vector $\boldsymbol{\mu}$ for the unscaled MAC, then $\frac{\boldsymbol{\lambda^*}}{\boldsymbol{\alpha}}$ is the optimal Lagrange multiplier for the MAC scaled by $\boldsymbol{\alpha}$. Therefore we can scale the channel appropriately so that $\frac{\lambda_i^*}{\alpha_i}$ are equal for all i, which implies that the point is on the boundaries of both the individual and sum power constraint MAC's. Using this method, every point on the boundary of $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H})$ can be shown to be on the boundary of the sum power MAC (and therefore of the dual BC) for some scaling. The proof of Theorem 3 in Appendix D is based on this idea.

If we examine the points where the MAC and BC capacity region boundaries touch, we find that there is also a fundamental relationship between the power policies used to achieve these points. The optimal power policies (i.e. boundary achieving power policies) for the fading MAC and BC are established in

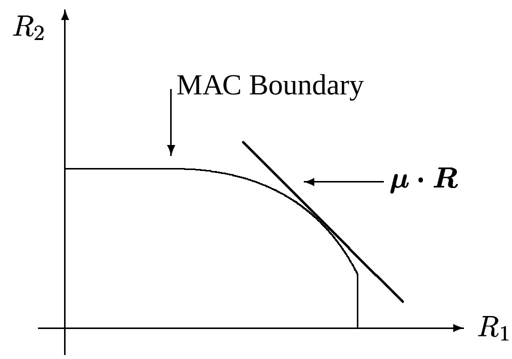


Fig. 9. MAC capacity region optimization

[4] and [3] respectively. Given a priority vector $\boldsymbol{\mu}$, it is possible to find the optimal power policy which maximizes $\boldsymbol{\mu} \cdot \boldsymbol{R}$ in both the MAC and the BC. Due to the duality of these channels, it comes as no surprise that the optimal power policies derived independently for the BC and MAC are related by the MAC-BC (9) and BC-MAC (11) transformations at the points where the BC and MAC capacity region boundaries touch. It is quite straightforward to verify that the optimal MAC power policies can be found from knowledge of the optimal BC power policies, and vice-versa.

G. Decoding Order

The duality of the flat-fading MAC and BC leads to some interesting observations about the optimum decoding order in the BC. By the power transformations, any point on the boundary of the BC ergodic capacity region is also on the boundary of the MAC ergodic capacity region for some set of power constraints whose sum equals the BC power constraint. Additionally, it is easy to show from the proof of Theorem 3 that the MAC and BC ergodic capacity regions are "tangential" at the point where the boundaries touch in the sense that the weighted rate sum $\boldsymbol{\mu} \cdot \boldsymbol{R}$ at the intersection point is equal to the maximum of $\boldsymbol{\mu} \cdot \boldsymbol{R}$ in $\mathcal{C}_{MAC}(\boldsymbol{P}; \boldsymbol{H})$ and in $\mathcal{C}_{BC}(\boldsymbol{P}; \boldsymbol{H})$ for the same $\boldsymbol{\mu}$.

From results on the ergodic capacity region of the MAC [4, 12], it is known to be optimal to decode users in order of increasing priority. In other words, in order to maximize $\mu_1 R_1 + \mu_2 R_2$ with $\mu_1 > \mu_2$, user 1 should be decoded last in the MAC in all fading states. Therefore, a fixed decoding order in all fading states is optimal for the MAC. Suppose we consider a point on the boundary of the BC capacity region which is also a boundary point of a dual MAC capacity region at the point where $\mu_1 R_1 + \mu_2 R_2$ is maximized in the MAC with $\mu_1 > \mu_2$. Then the optimum decoding order in the MAC to achieve this point will decode user 1 last in every fading state. By the BC-MAC transformation, we can find a BC power policy that achieves the same point. Curiously, the transformation tells us that the decoding order should be reversed in the BC, or that user 1 should always be decoded first in the BC and user 2 should always be decoded last. Therefore, by the transformations, we see that boundary points can be achieved in the BC by decoding in order of decreasing priority. From basis results on the BC, however, it is known that the strongest user should always be decoded last in the BC. This apparent inconsistency is resolved by the fact that in optimal BC power policies only one of the two users transmits (which means decoding order is irrelevant) in fading states in which this inconsistency arises. In other words, if user 1 has a stronger channel than user 2, then only user 1 receives data in the BC. Otherwise, we would have an inconsistency with regards to the decoding order.

In summary, we have established that in the MAC it is optimal to decode users in order of *increasing* priority, while in the BC it is optimal to decode users in order of *decreasing* priority.

H. Symmetric Channels

While duality holds for any joint fading distribution, we can make some particularly powerful statements with regards to symmetric fading distributions. If the joint fading distribution is symmetric and all K transmitters in the MAC have the same power constraint, then the optimal Lagrange multipliers corresponding to the sum rate capacity of the MAC (the maximum of $\boldsymbol{\mu} \cdot \boldsymbol{R}$ where $\mu_1 = \mu_2 = \cdots = \mu_K = \frac{1}{K}$) are all equal by a symmetry argument. As discussed in Section IV-F, this implies that the unscaled MAC and the unscaled BC (i.e. $\boldsymbol{\alpha} = \boldsymbol{1}$) ergodic capacity regions meet at the maximum sum rate point of their capacities. In this scenario the optimal power policies in the dual channels are identical, since only the user with the largest fading gain transmits in each fading state [3] [4]. The uplink is generally assumed to be a more challenging transmission medium than the downlink in cellular-type systems, but this result indicates that they are theoretically equivalent when the channel is symmetric.

For arbitrary non-symmetric fading distributions (even with symmetric power constraints), this fact is not generally true. For an arbitrary fading distribution, the boundaries of the dual channels may meet at a point where $\boldsymbol{\mu} \cdot \boldsymbol{R}$ is maximized for unequal values of $\boldsymbol{\mu}$ or possibly may not meet at all (for $K > 2$). By the convexity of the BC and MAC capacity regions, either situation would imply that the sum rate capacity of the uplink is smaller than the sum rate capacity of the downlink.

V. Duality of Outage Capacity

In this section we show that duality holds for outage capacity in addition to ergodic capacity. The outage capacity region (denoted $\mathcal{C}_{MAC}^{out}(\overline{\boldsymbol{P}}, \boldsymbol{P^{out}}; \boldsymbol{H})$ and $\mathcal{C}_{BC}^{out}(\overline{P}, \boldsymbol{P^{out}}; \boldsymbol{H})$) is defined as the set of rates that can be maintained for user j for a fraction P_j^{out} of the time, or in all but P_j^{out} of the fading states [5, 6]. Outage capacity is concerned with situations in which each user (in either the BC or MAC) desires a constant rate a certain percentage of the time. The zero-outage capacity [5, 13][7] is a special case of outage capacity where a constant rate must be maintained in *all* fading states, or where $\boldsymbol{P^{out}} = 0$.

There are two different notions of outage: common outage and individual outage. In common outage, an outage is declared for all users simultaneously, so all transmission is stopped at certain times. With individual outage, outages need not be declared simultaneously so users are allowed to declare outages individually. We will state theorems for the more general individual outage scenario, but it is easy to see that all theorems extend to common outage as well. In [5, 6] it is shown that superposition coding with successive decoding achieves the outage capacity of the BC and successive decoding achieves the outage capacity of the MAC. Due to this fact, we can use the BC-MAC and the MAC-BC power transformations with outage capacity.

Any rate vector in the outage capacity region of the MAC or BC is achievable via some power policy and decoding order function. With outage, we are concerned with the *instantaneous* rates achieved in every fading state as opposed to the ergodic capacity region where we were concerned with the *average* rates. For any power policy, the instantaneous rates in a fading state are a direct function of the power allocated to each user in that specific fading state. By the BC-to-MAC and MAC-to-BC transformations, we know that for any BC power policy, we can find a MAC power policy such that the same *instantaneous* rates are achieved for all users in *every* fading state. This implies that the outage states (i.e. states where the rates of certain users are zero) are also preserved by the transformation because no power is allocated to users when they are in outage. We thus get the following theorem.

Theorem 8: For a fading Gaussian BC and the dual MAC, the outage capacity region of the MAC is a subset of the outage capacity region of the BC for any outage vector $\boldsymbol{P^{out}}$:

$$\mathcal{C}^{out}_{MAC}(\overline{\boldsymbol{P}}, \boldsymbol{P^{out}}; \boldsymbol{H}) \subseteq \mathcal{C}^{out}_{BC}(\boldsymbol{1} \cdot \overline{\boldsymbol{P}}, \boldsymbol{P^{out}}; \boldsymbol{H}) \quad . \tag{37}$$

Proof: A rate vector $\boldsymbol{R}$ is in the MAC outage region $\mathcal{C}^{out}_{MAC}(\overline{\boldsymbol{P}}, \boldsymbol{P^{out}}; \boldsymbol{H})$ if and only if there exists a power policy satisfying power constraints $\overline{\boldsymbol{P}}$ and decoding order function $\pi(\boldsymbol{h})$ (or a generalized time-sharing strategy[8] between multiple power policies and decoding order functions) such that $Pr[R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h})) < R_j] \leq P_j^{out}$ for all j [6].

Similarly, a rate vector $\boldsymbol{R}$ is in the BC outage region $\mathcal{C}^{out}_{BC}(\overline{P}, \boldsymbol{P^{out}}; \boldsymbol{H})$ if and only if there exists a power policy satisfying power constraint $\overline{P}$ and decoding order function $\pi(\boldsymbol{h})$ (or a generalized time-share

[7] Zero-outage capacity is referred to as *delay-limited capacity* in [13].

[8] Here time-sharing represents a more generalized notion of time-sharing in *each* fading state as opposed to time-sharing between two different power policies that are defined for all fading states. See [5] [6] for more details.

strategy between multiple power policies and decoding order functions) such that $Pr[R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h})) < R_j] \leq P_j^{out}$ for all j [5].

Using the MAC-to-BC power transformation, for any MAC power policy $P^M(\boldsymbol{h})$ and decoding order function $\pi(\boldsymbol{h})$, there exists a BC power policy $P^B(\boldsymbol{h})$ such that

$$R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h})) = R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi^*(\boldsymbol{h})), \quad j = 1, \dots, K, \tag{38}$$

where $\pi^*(\boldsymbol{h})$ is opposite $\pi(\boldsymbol{h})$ in every fading state and $\mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^M(\boldsymbol{h})] = \mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^K P_j^B(\boldsymbol{h})]$. Since the instantaneous rates of every user are preserved by the transformations, the outage constraints will also be satisfied in the BC. Rate vectors in the MAC outage region achievable by time-sharing between multiple power policies and decoding order functions can be achieved in the dual BC by finding the transformed BC power policy/decoding order function pair for each pair in the MAC and then time-sharing between these policies. Therefore any rate vector in the MAC outage capacity region must also be in the dual BC outage capacity region with the same outage constraint. ■

As expected, we also find that the BC outage capacity region is equal to the union of the dual MAC outage capacity regions over all MAC's with the same sum power constraint.

Theorem 9: The outage capacity region of a constant Gaussian BC with power constraint $\overline{P}$ and outage constraint $\boldsymbol{P^{out}}$ is equal to the union of outage capacity regions of the dual MAC with outage constraint $\boldsymbol{P^{out}}$ and power constraints $(P_1, \dots, P_K)$ such that $\sum_{j=1}^K P_j = \overline{P}$:

$$\mathcal{C}_{BC}^{out}(\overline{P}, \boldsymbol{P^{out}}; \boldsymbol{H}) = \bigcup_{\{\boldsymbol{P}:\, \boldsymbol{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{MAC}^{out}(\boldsymbol{P}, \boldsymbol{P^{out}}; \boldsymbol{H}). \tag{39}$$

Proof: This theorem follows in a straightforward manner from the BC-MAC and MAC-BC transformations. For any MAC power policy, we can find a BC power policy using the same sum power which achieves the same instantaneous rates in each fading state. Similarly, for any BC power policy we can find an equivalent MAC power policy. ■

Theorem 9 characterizes the outage capacity of the BC in terms of the outage capacity of the dual MAC. We can also characterize the outage capacity of the MAC in terms of the outage capacity of the scaled BC.

Theorem 10: The outage capacity region of a fading MAC channel is equal to the intersection of the outage capacity regions of the scaled dual BC over all scalings:

$$\mathcal{C}_{MAC}^{out}(\boldsymbol{P}, \boldsymbol{P^{out}}; \boldsymbol{H}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{BC}^{out}(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}, \boldsymbol{P^{out}}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{H}) \tag{40}$$

Proof: Using the arguments of Appendix C, we see that Theorem 3 can be applied. Thus the proof follows that of Theorem 4. ■

Though the outage capacity region has been characterized for both the BC and MAC [5, 6], the MAC region can be quite difficult to find numerically. Duality, however, allows the region to easily be found numerically via the dual BC outage capacity region.

As mentioned earlier, outage capacity with *common* outage also adheres to Theorems 8, 9, and 10. The proofs of these theorems are identical to the proofs for independent outage.

VI. Duality of Minimum Rate Capacity

Minimum rate capacity was first introduced for the fading broadcast channel in [7]. In this section we will establish the duality of the minimum rate capacity regions of the MAC and BC and use this duality to find the previously unknown minimum rate capacity region of the MAC.

In minimum rate capacity, long-term average rates are maximized subject to an average power constraint and an additional constraint that requires the instantaneous rates of all users to meet or exceed the defined minimum rates in *all* fading states. Minimum rate capacity is essentially a combination of zero-outage capacity (minimum rates are maintained in all fading states) and ergodic capacity (long-term average rates in excess of the minimum rates are maximized).

The minimum rate capacity is thus defined as the maximum ergodic capacity that can be obtained while ensuring that a set of minimum rates $\boldsymbol{R^*} = (R_1^*, \ldots, R_K^*)$ is maintained for all users in *all* fading states. The minimum rate capacity of a fading broadcast channel with power constraint $\overline{P}$ and minimum rate vector $\boldsymbol{R^*} = (R_1, \ldots, R_K)$ is formally defined as [7, Theorem 1]:

$$\mathcal{C}_{BC}^{min}(\overline{P}, \boldsymbol{R^*}) = \left\{ \bigcup_{\mathcal{P} \in \mathcal{F}'_{BC}} \mathcal{C}_{BC}(\mathcal{P}) \right\} \bigcap \{\boldsymbol{R} \geq \boldsymbol{R^*}\} \tag{41}$$

where $\mathcal{F}'_{BC}$ is the set of power policies satisfying the power constraint and the minimum rate constraints:

$$\mathcal{F}'_{BC} = \{\mathcal{P}_{BC} : \mathbb{E}_{\boldsymbol{H}}[\sum_{j=1}^{K} P_j^B(\boldsymbol{h})] \leq \overline{P} \text{ and } R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h})) \geq R_j^* \; \forall j, \boldsymbol{h}\} \tag{42}$$

with $\pi(\boldsymbol{h})$ assumed to be the strongest-user-last decoding order in each fading state and $\mathcal{C}_{BC}(\mathcal{P})$ as defined in (27). Since $\mathcal{C}_{BC}(\mathcal{P})$ extends from the corresponding optimal rate vector to the origin (i.e. below the minimum rates), the intersection with the set $\{\boldsymbol{R} \geq \boldsymbol{R^*}\}$ is necessary to ensure that the average rates are larger than the minimum rates. From this definition of the minimum rate capacity region, it is easy to see that a rate vector $\boldsymbol{R}$ is in the minimum rate capacity region $\mathcal{C}_{BC}^{min}(\overline{P}, \mathbf{R}^*; \boldsymbol{H})$ if and only if there exists a power policy that satisfies the power constraint $\overline{P}$ and the minimum rate constraints in every fading state, such that $R_j \leq \mathbb{E}_{\boldsymbol{H}}\left[R_j^B(\boldsymbol{h}, P^B(\boldsymbol{h}), \pi(\boldsymbol{h}))\right]$ (where $\pi(\boldsymbol{h})$ is again assumed to be the optimal decoding order) for all K users.

Clearly the minimum rate vector $\boldsymbol{R^*}$ must be in the zero-outage capacity of the channel $\mathcal{C}_{BC}^{out}(\overline{P}, \mathbf{0}; \boldsymbol{H})$ for the minimum rates to be achievable in all fading states[9]. Additionally, every minimum rate vector $\boldsymbol{R^*}$ defines a different minimum rate capacity region.

We can similarly define the minimum rate capacity region of a fading multiple-access channel as the maximum ergodic capacity that can be obtained while ensuring that a set of minimum rates $\boldsymbol{R^*} = (R_1^*, \ldots, R_K^*)$ are maintained for all users in *all* fading states. Formally, this becomes

$$\mathcal{C}_{MAC}^{min}(\overline{\boldsymbol{P}}, \boldsymbol{R^*}) = \left\{ \bigcup_{\pi(\boldsymbol{h}),\, \mathcal{P} \in \mathcal{F}'_{\mathcal{MAC}}} \mathcal{C}_{MAC}(\mathcal{P}, \pi(\boldsymbol{h})) \right\} \bigcap \{\boldsymbol{R} \geq \boldsymbol{R^*}\} \tag{43}$$

where $\mathcal{F}'_{\mathcal{MAC}}$ is defined analogous to $\mathcal{F}'_{BC}$ but $\mathcal{C}_{MAC}(\mathcal{P}, \pi(\boldsymbol{h})$ is taken to be the rectangular region $\{R_j \leq \mathbb{E}_{\boldsymbol{H}}\left[R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h}))\right] \;\; j = 1, \ldots, K\}$ instead of the standard pentagon region defined in Section IV-A. Notice that for the MAC we must also take the union over all decoding order functions because it is not clear what the optimum decoding order should be in each fading state. As with the minimum rate capacity region of the BC, it is clear that a rate vector $\boldsymbol{R}$ is in the minimum rate capacity region $\mathcal{C}_{MAC}^{min}(\overline{\boldsymbol{P}}, \mathbf{R}^*)$ if and only if there exists a power policy and decoding order function that satisfy the power constraint $\overline{P}$ and the minimum rate constraints in every fading state, such that $R_j \leq \mathbb{E}_{\boldsymbol{H}}\left[R_j^M(\boldsymbol{h}, P^M(\boldsymbol{h}), \pi(\boldsymbol{h}))\right]$ for all K users.

With this definition of the BC and MAC minimum rate capacity regions, it is easy to see that the minimum rate capacity regions are duals as well because the MAC-to-BC and BC-to-MAC transformations preserve instantaneous rates. If a BC power policy satisfies the minimum rates in all fading states, then the transformed MAC power policy will correspond to the same instantaneous rates and thus will satisfy the minimum rates in all fading states. This statement clearly holds when going from a MAC power policy to a BC power policy. We thus get the following two theorems, proofs of which are identical to the proofs of Theorems 8 and 9.

Theorem 11: For a fading Gaussian BC and the dual MAC, the minimum rate capacity region of the MAC is a subset of the minimum rate capacity region of the BC:

$$\mathcal{C}_{MAC}^{min}(\overline{\boldsymbol{P}}, \boldsymbol{R^*}; \boldsymbol{H}) \subseteq \mathcal{C}_{BC}^{min}(\mathbf{1} \cdot \overline{\boldsymbol{P}}, \boldsymbol{R^*}; \boldsymbol{H}) \tag{44}$$

[9] This implies that the minimum rate capacity region is non-zero only for fading distributions with non-zero zero-outage capacity regions.

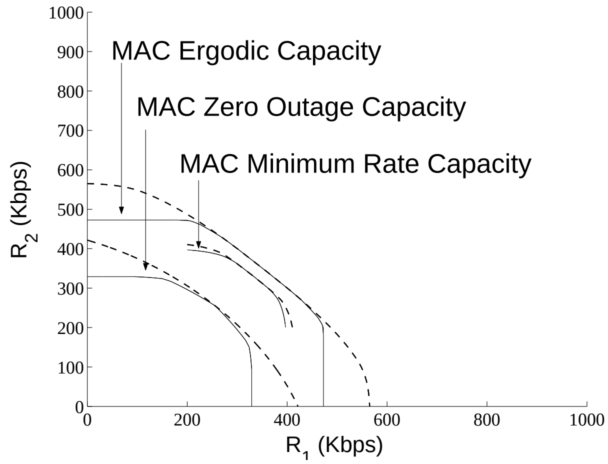


Fig. 10. MAC Minimum Rate (200 Kbps Min Rate), Ergodic, and Zero-Outage Capacity Regions

Theorem 12: The minimum rate capacity region of a fading Gaussian BC is given by:

$$\mathcal{C}_{BC}^{min}(\overline{P}, \boldsymbol{R^*}; \boldsymbol{H}) = \bigcup_{\{\boldsymbol{P}:\, \boldsymbol{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{MAC}^{min}(\boldsymbol{P}, \boldsymbol{R^*}; \boldsymbol{H}). \tag{45}$$

The most interesting application of duality to minimum rate capacity is characterizing the minimum rate capacity region of the MAC via the dual BC. Directly finding the minimum rate capacity of the MAC appears to be quite difficult, but duality allows us to do so directly from the dual BC.

Theorem 13: The minimum rate capacity region of a fading Gaussian MAC is given by:

$$\mathcal{C}_{MAC}^{min}(\boldsymbol{P}, \boldsymbol{R^*}; \boldsymbol{H}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{BC}^{min}(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}, \boldsymbol{R^*}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{H}). \tag{46}$$

Proof: If we consider the MAC and BC minimum rate capacity regions as defined in (43) and (41) *without* the intersection with the region $\{\boldsymbol{R} \geq \boldsymbol{R^*}\}$ so that the regions extend to the origin, then Theorem 3 can be applied (using the arguments of Appendix C). By using Theorem 12 and then performing the intersection operation we get the desired result. ■

The MAC minimum rate capacity region for a discrete fading distribution channel is plotted in Figure 10. In the figure, the MAC ergodic, zero-outage, and minimum rate capacity regions are all shown. The corresponding dual BC capacity regions (ergodic, zero-outage, and minimum rate) are shown by the dotted lines. The minimum rate capacity region is shown for symmetric minimum rates of 200 Kbps for each user. In the figure, all three MAC capacity regions were calculated using duality, i.e. by taking the intersection of scaled BC capacity regions. However, we show only the unscaled BC capacity regions in the plot for simplicity.

Much of the intuition which applies to the BC minimum rate capacity region [7] also appears to extend to the MAC minimum rate capacity region. The MAC minimum rate capacity region lies between the zero-outage and ergodic capacity regions because minimum rate capacity is a combination of zero-outage and ergodic capacity. This also occurs for the BC minimum rate capacity region.

Notice that Theorems 12 and 13 are consistent with our results on the zero outage capacity of the dual BC and MAC. By Theorem 9 we know that for any $\boldsymbol{R^*} \in \mathcal{C}_{BC}^{out}(\overline{P}; \boldsymbol{H})$ there is a set of power constraints $\overline{\boldsymbol{P}}$ such that $\boldsymbol{R^*} \in \mathcal{C}_{MAC}^{out}(\overline{\boldsymbol{P}}, \boldsymbol{0}; \boldsymbol{H})$. Without this Theorem 12 would not be valid for all $\boldsymbol{R^*} \in \mathcal{C}_{BC}^{out}(\overline{P}, \boldsymbol{0}; \boldsymbol{H})$. In addition, by Theorem 10 we know that any $\boldsymbol{R^*} \in \mathcal{C}_{MAC}^{out}(\overline{\boldsymbol{P}}, \boldsymbol{0}; \boldsymbol{H})$ is also in $\mathcal{C}_{BC}^{out}(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}, \boldsymbol{0}; \sqrt{\boldsymbol{\alpha}}\boldsymbol{H}) \forall \boldsymbol{\alpha} \geq 0$. This result is necessary for Theorem 13 to be valid.

VII. Frequency Selective Channels

Duality can also be shown to hold for frequency selective (ISI) channels. We consider broadcast and multiple-access channels with time-invariant, finite-length impulse responses and additive Gaussian noise [14] [15]. The model we consider is:

$$\begin{aligned} \text{BC:} \quad Y_j[i] = \sum_{k=0}^{L} h_j[k] X[i-k] + n_j[i] \quad &= \quad h_j[i] * x[i] + n_j[i] \qquad j = 1, \ldots, K \\ \text{MAC:} \quad Y[i] = \sum_{j=1}^{K} \sum_{k=0}^{L} h_j[k] X_j[i-k] + n[i] \quad &= \quad \sum_{j=1}^{K} (h_j[i] * x_j[i]) + n[i] \end{aligned}$$

where the finite-length impulse response of Receiver j in the BC and Transmitter j in the dual MAC is $\{h_j[i]\}_{i=0}^{L}$. Here we assume that L is the maximum finite impulse response duration over all K channels. Additionally, we assume the noise at each receiver to be stationary Gaussian noise with power σ^2 and no correlation in time (i.e. white noise)[10].

The capacity region of the two-user broadcast channel with ISI is stated in [15, Equations 32-24] and the capacity region of the two-user MAC with ISI is stated in [14, Equation 9] [11]. These definitions are analogous to those given in (26) and (27) for the flat-fading BC and (20) and (21) for the flat-fading MAC, with the exception that power policies for the frequency-selective channel are given with respect to *frequency* (ω) instead of the *fading state* ($\boldsymbol{h}$).

By using the MAC-to-BC and BC-to-MAC transformations for each frequency (instead of for each fading state), we see that the duality results developed in Section IV for the flat-fading MAC and BC

[10] Colored noise can be absorbed into the channel gain spectrum.

[11] Interestingly, the authors of [14] used the concept of channel scaling in order to find the optimal power allocation policy of the frequency-selective MAC. This turns out to be the same channel scaling which is used to characterize the MAC in terms of the dual BC.

also hold for the frequency-selective BC and MAC. In other words, the MAC capacity region with power constraint $\overline{\boldsymbol{P}}$ is a subset of the dual BC capacity region with power constraint $\mathbf{1} \cdot \overline{\boldsymbol{P}}$ (analogous to Theorem 5) and the BC capacity region with power constraint $\overline{P}$ is equal to the union of dual MAC capacity regions over all power constraints $\boldsymbol{P}$ such that $\mathbf{1} \cdot \boldsymbol{P} = \overline{P}$ (analogous to Theorem 6). Since the MAC capacity region is closed and convex [14] and because it satisfies the channel scaling relationship defined in (12), the MAC capacity region with ISI is equal to the intersection of the dual BC capacity region over all possible scalings (analogous to Theorem 7). For the ISI channel, channel scaling by the factor $\sqrt{\boldsymbol{\alpha}}$ refers to multiplying impulse response $h_j[i]$ by the constant $\sqrt{\alpha_j}$.

VIII. Extensions

The focus of this paper has been to characterize the duality of the scalar Gaussian multiple-access and broadcast channels. There are, however, many possible directions in which this notion of duality may be extended. In this section we discuss a few possibilities for such extensions. It appears there are two main directions in which duality may apply: a more general set of Gaussian channels, and non-Gaussian channels.

A. Multiple-Antenna MAC and BC

In this paper we have focused on the scalar Gaussian MAC and BC, but the multiple-antenna (multiple-input, multiple-output or MIMO) versions of the Gaussian MAC and BC are also of great interest. In a related paper, we show that the multiple-input, multiple-output (MIMO) Gaussian BC and MAC are duals [1]. Since the MIMO BC is a non-degraded broadcast channel, its capacity region is not known. However, an achievable region for the MIMO BC based on dirty-paper coding [10] [16] is known. Additionally, the capacity region of the MIMO MAC is known [17] [18] [19]. In [1] it is shown that the capacity region of the MIMO MAC and the dirty-paper achievable region of the BC are duals, or that the MIMO BC achievable region is equal to the union of the dual MIMO MAC capacity regions and the MIMO MAC capacity region is equal to the intersection of the scaled MIMO BC achievable regions. The duality between the dirty paper region and the MIMO MAC seems to indicate that the dirty paper region is actually the capacity region of the MIMO, but the dirty paper region is known to be optimal only at the sum rate point and the optimality of the full region is still an open question.

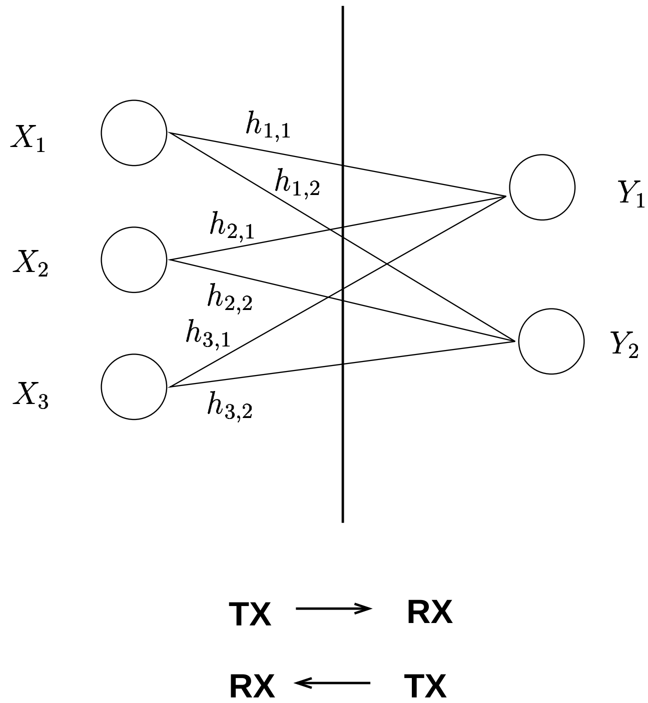


Fig. 11. Multi-Terminal Gaussian Network

B. Gaussian Multi-Terminal Networks

The Gaussian MAC and BC can be generalized to a model in which there are multiple transmitters (similar to the MAC) *and* multiple receivers (similar to the BC) subject to additive Gaussian noise. Such networks are referred to as Gaussian multi-terminal networks [8, Section 14.10]. In this section we discuss networks in which nodes can either be transmitters or receivers, and not both simultaneously. Though very few results are known about such general networks, it appears that duality may apply to these networks as well.

In Figure 11, a three-transmitter, two receiver channel is shown, if transmission is considered from left-to-right. We define the dual channel for this network as the channel associated with transmission from right-to-left (i.e. two transmitters, three receivers) with the *same channel gains* $h_{i,j}$ between all nodes. As before, we assume that every receiver suffers from Gaussian noise with the same power.

The dual broadcast and multiple-access channels can be seen as a specialization of multi-terminal networks in which there is only a single node on the left. If transmission occurs from left to right, then the channel is a two-user broadcast channel and if the nodes on the right transmit then the channel is a two-user multiple access channel. Theorem 1 states that the capacity regions of the BC and the dual sum power constraint MAC are the same. This is equivalent to stating that the capacity regions for left to right communication (BC) and right to left communication (MAC) are the same if the same sum transmit power constraint is applied to both channels.

In the general multi-terminal setting we consider, any transmitter is allowed to communicate with any receiver. In the channel shown in Figure 11, the capacity region is six-dimensional (because there are six possible receiver-transmitter pairs). It is then tempting to conjecture that Theorem 1 extends to general Gaussian multi-terminal networks, or that that the six-dimensional capacity region governing transmission from left-to-right when sum power constraint P is imposed on the three transmitters on the left is the same as the capacity region for transmission from right-to-left when the same sum power constraint P is imposed on the two transmitters on the right. Unfortunately, this conjecture cannot be confirmed since the capacity region of a general multi-terminal network is not known. One trivial case for which duality holds is for completely symmetric channels in which the channels from left-to-right and right-to-left are indistinguishable (An example of such a channel is the two-transmitter, two-receiver channel considered in [20] in which the channel gains from X_1 to Y_1 and from X_2 to Y_2 are 1 and the gains from X_1 to Y_2 and from X_2 to Y_1 are equal to a constant a).

If the nodes on the left and right were considered to be multiple-antennas of single users (i.e. a single transmitter with 3 antennas communicating to a receiver with 2 antennas, or vice versa), then duality holds due to the reciprocity of multiple-antenna Gaussian links [18]. By the reciprocity result we know that the capacity of a channel with gain matrix H is equal to the capacity of the channel with gain matrix equal to the transpose (or Hermitian transpose since conjugation of the channel matrix has no effect) of H. This hints that some sort of left-right vs. right-left duality may hold, but this has yet to be confirmed.

Though we are unable to verify an extension of Theorem 2 to multi-terminal networks, Theorem 3 (which characterizes a MAC rate region as an intersection of sum power constraint MAC rate regions) does apply to rate regions of a multi-terminal network. Before stating the theorem, we first define the notion of a rate region of a K-transmitter, N-receiver multi-terminal network.

Definition 4: Let a KN-dimensional rate vector be written as

$$\boldsymbol{R} = (R_{1,1}, \ldots, R_{1,N}, R_{2,1}, \ldots, R_{2,N}, \ldots, R_{K,1}, \ldots, R_{K,N})$$

where $R_{i,j}$ is the rate from transmitter i to receiver j. Let $\boldsymbol{P} = (P_1, \ldots, P_K)$ be the vector of transmit power constraints and let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$ be a vector of scaling constants. We define a rate region $R(\boldsymbol{P})$ as a mapping from a power constraint vector $\boldsymbol{P}$ to a set in $\mathcal{R}_+^{\mathcal{KN}}$. The $\boldsymbol{\alpha}$-scaled version of the multi-terminal channel is the multi-terminal channel in which the channel gain from transmitter i to *each* receiver is scaled by $\sqrt{\alpha_i}$. We denote the rate region of the scaled channel as $R_{\boldsymbol{\alpha}}(\boldsymbol{P})$.

As with the MAC, we only consider rate regions which satisfy the conditions of Definition 2. Though these conditions were developed for single-receiver systems, they apply to multi-receiver systems with only one minor modification: in condition 7, we require that the capacity of *every* transmitter-receiver link goes to infinity as the power constraint of the transmitter becomes large, or equivalently as $P_i \to \infty$, we have $R_{i,j} \to$ for $j = 1, \ldots, N$. Notice that these conditions are very minimal and would be expected to be satisfied by any standard rate region definition. The sum power constraint region for multi-terminal networks is defined by Definition 3. Having stated these definitions, we are able to state the theorem.

Theorem 14: Any rate region $R(\boldsymbol{P})$ of a K-transmitter, N-receiver multi-terminal network satisfying the conditions of Definition 2 is equal to the intersection over all strictly positive scalings of the sum power constraint rate regions for any strictly positive power constraint vector $\boldsymbol{P}$:

$$R(\boldsymbol{P}) = \bigcap_{\boldsymbol{\alpha}>0} R_{\boldsymbol{\alpha}}^{sum}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right). \tag{47}$$

Proof: See Appendix D. ■

This theorem allows us to characterize an individual power constraint rate region in terms of a sum power constraint region. In intuitive terms, Theorem 14 characterizes the individual power constraint rate region of left-to-right communication in terms of the sum power constraint region of left-to-right communication as well. With the MAC (i.e. $N = 1$), we were able to extend this further because the sum power constraint MAC was shown to be equivalent to the dual BC (i.e. left-to-right and right-to-left communication are equivalent). However, as discussed above, no such connection has been established for general multi-terminal networks.

C. *Duality of Non-Gaussian Channels*

Although we have treated only Gaussian channels in this paper, it would be very interesting to see if duality holds between general broadcast and multiple-access channels. Since there is no general theorem on the capacity region of the non-degraded broadcast channel, such duality could be helpful in this respect. However, formulating a dual channel for an arbitrary broadcast or multiple-access channel appears quite difficult. It is entirely possible that duality only applies to Gaussian channels, in which case it would be illuminating to determine what special properties of Gaussian channels allows duality to work. The Gaussian channel seems to have a number of properties which appear to lead to duality. Among these is the fact that successive interference cancellation is optimal for the BC and the MAC and the fact that a Gaussian distribution is the capacity-achieving distribution for both channels. Though it may appear unlikely, we still hold some hope hat duality exists for a broader class of broadcast/multiple-access channels than the Gaussian channels.

IX. Conclusion

We have defined a duality between the Gaussian MAC and BC by establishing fundamental relationships between the capacity regions of the MAC and BC with the same channel gains and the same noise power at all receivers. We first established that the MAC capacity region is a subset of the capacity region of the dual BC with power constraints equal to the sum of the MAC power constraints. For constant channels, we also showed that these dual capacity regions intersect at exactly one point, namely the successive decoding point on the boundary of the MAC where users are decoded in order of *decreasing* channel gains. Interestingly, this is exactly opposite of the optimal decoding order used to achieve BC capacity.

Based on this result relating the capacity regions of the dual channels, we also found explicit expressions characterizing the BC capacity region in terms of the capacity region of the dual MAC and vice versa. We showed that the capacity region of the BC is equal to the *union* of the capacity regions of the dual MAC, where the union is taken over all MAC power constraints which sum to equal the BC power constraint. This result indicates that the capacity region of a multiple-access channel with a sum power constraint instead of individual power constraints (but without any transmitter coordination) is equal to the capacity region of the dual BC with the same power constraint. In establishing this equivalence, we also found an explicit formula relating the power distribution in the BC and in the MAC to achieve the same rate point.

Using the concept of channel scaling, we found that the MAC capacity region is equal to the *intersection* of dual BC capacity regions over all positive channel scalings. In doing so, we proved a more general result that characterizes the individual transmit power constraint capacity of a multi-terminal Gaussian network in terms of the sum transmit power constraint capacity of the same network, assuming some basic conditions on the capacity region.

In addition to considering constant channels, we showed that duality extends to ergodic, outage, and minimum rate capacity for flat-fading channels and to frequency-selective channels as well. Using duality, results that are known for one of the two channels can often be extended to the dual channel, i.e. if the Shannon capacity with certain constraints is known for the BC, then the capacity with the same constraints can easily be found for the MAC as well. We in fact find the previously unknown minimum rate capacity region of the MAC by using duality and known results on the minimum rate capacity region of the BC.

We also briefly discussed a number of intriguing extensions of duality, including multi-terminal Gaussian networks and non-Gaussian channels. It remains to be seen if duality is a result of the special structure of the Gaussian MAC and BC or if it has any deeper information theoretic implications.

APPENDIX

A. Proof of Transformations

We show that if $P_j^M A_j = P_j^B B_j$ for all j, where A_j and B_j are defined as

$$A_j = \sigma^2 + h_j^2 \sum_{i=1}^{j-1} P_i^B, \quad B_j = \sigma^2 + \sum_{i=j+1}^{K} h_i^2 P_i^M,$$

then $\sum_{i=1}^K P_i^M = \sum_{i=1}^K P_i^B$. We do this by inductively showing that

$$\sum_{i=1}^{j} P_i^B = \frac{\sigma^2}{B_j} \sum_{i=1}^{j} P_i^M. \tag{48}$$

For $j = 1$, this clearly holds by definition: $P_1^B = \frac{A_1 P_1^M}{B_1} = \frac{\sigma^2}{B_1} P_1^M$. Assume (48) holds for j. We then must show it holds for $j+1$.

$$\begin{aligned}
\sum_{i=1}^{j+1} P_i^B &= \sum_{i=1}^{j} P_i^B + P_{j+1}^B \\
&= \sum_{i=1}^{j} P_i^B + \frac{A_{j+1} P_{j+1}^M}{B_{j+1}} \\
&= \sum_{i=1}^{j} P_i^B + \frac{P_{j+1}^M(\sigma^2 + h_{j+1}^2 \sum_{i=1}^{j} P_i^B)}{B_{j+1}} \\
&= \frac{\sigma^2 P_{j+1}^M + (P_{j+1}^M h_{j+1}^2 + B_{j+1}) \sum_{i=1}^{j} P_i^B}{B_{j+1}} \\
&= \frac{\sigma^2 P_{j+1}^M + B_j \sum_{i=1}^{j} P_i^B}{B_{j+1}} \\
&= \frac{\sigma^2 P_{j+1}^M + \sigma^2 \sum_{i=1}^{j} P_i^M}{B_{j+1}} \\
&= \frac{\sigma^2 \sum_{i=1}^{j+1} P_i^M}{B_{j+1}}.
\end{aligned}$$

(add (a) above the equations to explain here). By using (48) for $j = K$ and the fact that $B_K = \sigma^2$, we get $\sum_{i=1}^K P_i^B = \sum_{i=1}^K P_i^M$ as desired.

B. Proof of the MAC Decoding Order Lemma

We show that the constant MAC and BC capacity regions meet at exactly one point if all channel gains are distinct. Alternatively, we show that every corner point other than the one corresponding to decoding in order of decreasing channel gains lies strictly in the interior of the dual BC capacity region. In this proof we use the fact that the dual BC capacity region is equal to the sum power constraint MAC capacity region (Theorem 1). We explicitly show that the sum power needed to achieve any strictly positive rate vector $\boldsymbol{R}$ using the decoding order in which the weakest user is last is *strictly less* than the sum power needed to achieve the same rate vector using any other decoding order at the receiver, assuming successive decoding is used. This implies that points on the boundary of the sum power MAC cannot be achieved by successive decoding with a decoding order other than the order of decreasing channel gains. Therefore all corner points of the individual power constraint MAC other than the optimal decoding order point are in the interior of the dual BC capacity region.

Assume a decoding order such that $h_i^2 < h_j^2$ but User i is decoded directly *before* User j. We will show that the sum power needed to achieve any strictly positive rate vector is strictly less if the decoding order of Users i and j is reversed, or if User i is decoded directly *after* User j. Users i and j do not affect users decoded after them because their signals are subtracted out, but they do contribute interference $h_i^2 P_i + h_j^2 P_j$ to all users decoded before them. All users which are decoded after Users i and j are seen as interference to both Users i and j. We denote this interference by I. The rates of Users i and j then are

$$
\begin{aligned}
R_i &= \frac{1}{2}\log\left(1+\frac{h_i^2 P_i}{h_j^2 P_j + \sigma^2 + I}\right) \\
R_j &= \frac{1}{2}\log\left(1+\frac{h_j^2 P_j}{\sigma^2 + I}\right)
\end{aligned}
$$

if User i is decoded before User j. The power required by Users i and j to achieve their rates are

$$P_i = \frac{h_j^2 P_j + \sigma^2 + I}{h_i^2}(e^{2R_i}-1), \quad P_j = \frac{\sigma^2 + I}{h_j^2}(e^{2R_j}-1)$$

and the sum of their powers is

$$P_i + P_j = \frac{\sigma^2+I}{h_i^2}(e^{2R_i}-1) + \frac{\sigma^2+I}{h_j^2}(e^{2R_j}-1) + \frac{\sigma^2+I}{h_i^2}(e^{2R_i}-1)(e^{2R_j}-1).$$

If User i is decoded directly after User j instead of before him, then the powers needed to achieve R_i and R_j are

$$P_i' = \frac{\sigma^2 + I}{h_i^2}(e^{2R_i}-1), \quad P_j' = \frac{h_i^2 P_i + \sigma^2 + I}{h_j^2}(e^{2R_j}-1)$$

and the sum power is

$$P_i' + P_j' = \frac{\sigma^2+I}{h_i^2}(e^{2R_i}-1) + \frac{\sigma^2+I}{h_j^2}(e^{2R_j}-1) + \frac{\sigma^2+I}{h_j^2}(e^{2R_i}-1)(e^{2R_j}-1).$$

Clearly $P_i + P_j - P'_i - P'_j = (\frac{\sigma^2+I}{h_i^2} - \frac{\sigma^2+I}{h_j^2})(e^{2R_i} - 1)(e^{2R_j} - 1) > 0$ since $h_i^2 < h_j^2$ and $R_i, R_j > 0$ by assumption. Therefore we have $P_i + P_j > P'_i + P'_j$. This fact means that Users i and j can achieve the same rates using *less sum power* by switching the decoding order of Users i and j and switching their powers from P_i and P_j to P'_i and P'_j. The rates of users decoded after i and j are unaffected by such a switch. However, as noted above, Users i and j do contribute interference to all users decoded before them. If we expand the interference contribution of Users i and j, we find

$$h_i^2 P_i + h_j^2 P_j = (\sigma^2 + I)(e^{2(R_i+R_j)} - 1) = h_i^2 P'_i + h_j^2 P'_j.$$

so the rates of all users decoded earlier are unaffected. Therefore by switching the decoding order of Users i and j and changing the powers to P'_i and P'_j (but not altering the rest of the decoding order or power allocations), we can achieve the same set of rates for all K users using strictly less sum power.

If the decoding order is such that no user has a smaller gain than the next user being decoded (i.e. the first user decoded has a larger channel gain than the second, the second has a larger channel gain than the third, etc.), then the decoding order is optimum in the sum power sense.

C. Verification of Rate Region Conditions

It only remains to show that capacity region of the constant MAC $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ meets the conditions specified in Theorem 3. All conditions are satisfied by any reasonable definition of a capacity region, but we explicitly verify them for this case.

1. The scaling property of $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ is established in (12).

2. The set S is convex if for any $x, y \in S$ and $\theta \in [0, 1]$, $\theta x + (1 - \theta) y \in S$. Let $\boldsymbol{r} \in C_{MAC}(\overline{\boldsymbol{P}})$ and $\boldsymbol{t} \in C_{MAC}(\overline{\boldsymbol{Q}})$. We wish to show that $\theta \boldsymbol{r} + (1-\theta)\boldsymbol{t} \in C_{MAC}(\theta\overline{\boldsymbol{P}} + (1-\theta)\overline{\boldsymbol{Q}})$. By time-sharing between the schemes used to achieve $\boldsymbol{r}$ and $\boldsymbol{t}$, we use power $\theta\overline{\boldsymbol{P}} + (1-\theta)\overline{\boldsymbol{Q}}$ and achieve rate $\theta\boldsymbol{r} + (1-\theta)\boldsymbol{t}$, which verifies the convexity of the set.

3. The region $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ is closed by definition and is convex due to a time-sharing argument.

4. $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ is an increasing function of power because any rate achievable with a smaller power constraint is also achievable with a larger power constraint because all power need not be used.

5. If some set of rates are achievable by transmitters 2 through K while transmitter 1 is also sending information, then those same rates are achievable in the absence of transmitter 1's signal because any set of codebooks which work in the presence of interference will not perform worse in the absence of interference is present.

6. If transmission is halted for some fraction of time, then any smaller rate vector can be achieved.

7. Additional power allows for additional rate on any link by transmitting a codeword which can be decoded (and thus subtracted off) by all K receivers, even when treating the rest of the received signal as noise.

8. $\mathcal{C}_{MAC}(\overline{\boldsymbol{P}};\boldsymbol{h})$ is bounded by the individual capacities of each link (i.e. each transmitter-receiver pair), which are finite due to the basic properties of Gaussian channels.

D. Proof of Theorems 3 and 14

We wish to show that for any strictly positive[12] power constraint $\overline{\boldsymbol{P}} = (\overline{P}_1, \dots, \overline{P}_K) > 0$

$$R(\overline{\boldsymbol{P}}) = \bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\boldsymbol{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right), \tag{49}$$

where the sum power constraint capacity region is defined as

$$R^{sum}_{\boldsymbol{\alpha}}(P_{sum}) \triangleq \bigcup_{\{\boldsymbol{P}|\ \boldsymbol{P}\in\mathcal{R}^{\mathcal{K}}_{+},\ \boldsymbol{1}\cdot\boldsymbol{P}\le P_{sum}\}} R_{\boldsymbol{\alpha}}(\boldsymbol{P}). \tag{50}$$

Before beginning the proof, we first restate the conditions required of $R(\boldsymbol{P})$:

1. $R(\boldsymbol{P}) = R_{\boldsymbol{\alpha}}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}})\ \forall\ \boldsymbol{\alpha} > 0, \boldsymbol{P} > 0$.
2. $S = \{(\boldsymbol{R},\boldsymbol{P})|\boldsymbol{P}\in\mathcal{R}^{\mathcal{K}}_{+},\ \boldsymbol{R}\in R(\boldsymbol{P})\}$ is a convex set.
3. For all $\boldsymbol{P}\in\mathcal{R}^{\mathcal{K}}_{+}$, $R(\boldsymbol{P})$ is a closed, convex region.
4. $R(\boldsymbol{P})$ is monotonically increasing, or $\boldsymbol{P_1}\ge\boldsymbol{P_2}$ implies $R(\boldsymbol{P_1})\supseteq R(\boldsymbol{P_2})$.
5. If $(R_{1,1},\dots,R_{1,N},R_{2,1},\dots,R_{K,N})\in R(P_1,P_2,\dots,P_K)$, then $(0,\dots,0,R_{2,1},\dots,R_{K,N})\in R(0,P_2,\dots,P_K)$. This condition must hold for all K transmitters.
6. If $\boldsymbol{R}\in R(\boldsymbol{P})$, then $\boldsymbol{R'}\le\boldsymbol{R}$ implies $\boldsymbol{R'}\in R(\boldsymbol{P})$.
7. $R(\boldsymbol{P})$ is unbounded in every direction as $\boldsymbol{P}$ increases, or $\forall i,j,\ \max_{R_{i,j}\in R(\boldsymbol{P})} R_{i,j}\to\infty$ as $P_i\to\infty$.
8. $R(\boldsymbol{P})$ is finite for all $\boldsymbol{P}>0$.

By condition 1, we know that scaling does not affect the individual power constraint capacity region, or $R(\boldsymbol{P}) = R_{\boldsymbol{\alpha}}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}})\ \forall\boldsymbol{\alpha}>0$. However, *scaling does affect the sum power constraint capacity region,* and in general the sum power constraint capacity region is different for every scaling. Though the sum power constraint capacity region is different for every scaling, we show that their intersection is equal to the capacity region of the unscaled region $R(\boldsymbol{P})$ with individual power constraints.

From the definition of the sum power constraint capacity region (50), it is clear that the sum power constraint capacity region is larger than the individual power constraint capacity region, or $R(\boldsymbol{P}) \subseteq$

[12]If the power constraint of some transmitter is zero, then we can eliminate the user and consider the $K-1$ user problem.

$R^{sum}(\mathbf{1}\cdot\boldsymbol{P})$. This also holds for scaled channels: $R_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}) \subseteq R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}})\ \forall\boldsymbol{\alpha} > 0$. Since $R(\overline{\boldsymbol{P}}) = R_{\boldsymbol{\alpha}}(\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})$, we know that $R(\overline{\boldsymbol{P}}) \subseteq R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}})\ \forall\boldsymbol{\alpha} > 0$. This in turn implies that $R(\overline{\boldsymbol{P}}) \subseteq \bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)$. To complete the proof, we must show that this inequality also holds in the opposite direction.

Since $R(\overline{\boldsymbol{P}})$ is a closed and convex region, it is completely characterized by the following maximization [21, p. 135]

$$\max_{\boldsymbol{R}} \boldsymbol{\mu}\cdot\boldsymbol{R} \quad s.t. \quad \boldsymbol{R} \in R(\overline{\boldsymbol{P}}) \tag{51}$$

over non-negative priorities $\boldsymbol{\mu} = (\mu_{1,1}, \dots, \mu_{K,N})$ such that $\boldsymbol{\mu} \geq 0$ and $\mathbf{1}\cdot\boldsymbol{\mu} = 1$. We define the sum priority of transmitter i as $\mu_i = \sum_{j=1}^{N} \mu_{i,j}$ for $i = 1, \dots, K$.

Since at least one component of $\boldsymbol{\mu}$ must be strictly positive for $\mathbf{1}\cdot\boldsymbol{\mu} = 1$ to hold, we assume without loss of generality that $\mu_K > 0$. (We can simply renumber the users to ensure this condition holds).

For every $\boldsymbol{\mu} \geq 0$, we will show[13]

$$\max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R} \geq \sup_{\boldsymbol{R}\in\bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu}\cdot\boldsymbol{R}. \tag{52}$$

This implies $R(\overline{\boldsymbol{P}}) \supseteq \bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)$ because $R(\overline{\boldsymbol{P}})$ is completely characterized by $\max \boldsymbol{\mu}\cdot R$ (51).

We will essentially show that for every $\boldsymbol{\mu}$ (or roughly every point on the boundary of $R(\overline{\boldsymbol{P}})$) there exists an $\boldsymbol{\alpha}$ such that the boundaries of $R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)$ and $R(\overline{\boldsymbol{P}})$ meet at the point where $\boldsymbol{\mu}\cdot\boldsymbol{R}$ is maximized[14]

The optimization in (51) is equivalent to

$$\max_{(\boldsymbol{R},\boldsymbol{P})\in S} \boldsymbol{\mu}\cdot\boldsymbol{R} \quad s.t. \quad \boldsymbol{P} = \overline{\boldsymbol{P}}. \tag{53}$$

Since $R(\overline{\boldsymbol{P}})$ is monotonically increasing, we can equivalently perform the maximization in (53) with constraint $\boldsymbol{P} \leq \overline{\boldsymbol{P}}$:

$$\max_{(\boldsymbol{R},\boldsymbol{P})\in S} \boldsymbol{\mu}\cdot\boldsymbol{R} \quad s.t. \quad \boldsymbol{P} \leq \overline{\boldsymbol{P}}. \tag{54}$$

Consider the above maximization for some fixed $\boldsymbol{\mu}$. By the convexity of the set S and the convexity of the power constraint, this is a convex optimization problem. Furthermore, the maximization takes on some optimal value p^* by the feasibility of the constraint set. The optimal value is finite due to the

[13] We take a sup instead of a max over the sum power constraint capacity region because we have not verified that it is a closed region.

[14] For $\boldsymbol{\mu}$ such that $\mu_i = 0$ for some i there may be no $\boldsymbol{\alpha}$ for which the boundaries of the $R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)$ and $R(\overline{\boldsymbol{P}})$ meet, but there instead the boundaries become arbitrarily close as $\boldsymbol{\alpha}$ approaches some limit. This was in fact seen in the constant channel case where we took α to 0 and ∞ to bound the vertical segments of the MAC capacity region.

assumption that $R(\overline{\boldsymbol{P}})$ is finite. The Lagrangian is formed by adding the weighted sum of the constraints to the objective function:

$$L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}) = \boldsymbol{\mu}\cdot\boldsymbol{R} - \lambda_1(P_1 - \overline{P}_1) - \cdots - \lambda_K(P_K - \overline{P}_K) \tag{55}$$

where the weights $\boldsymbol{\lambda} = (\lambda_1,\ldots,\lambda_K)$ are the Lagrangian multipliers. The Lagrangian dual function is

$$g(\boldsymbol{\lambda}) = \sup_{(\boldsymbol{R},\boldsymbol{P})\in S} L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}) \tag{56}$$

and for any $(\boldsymbol{R},\boldsymbol{P}) \in S$ satisfying $\boldsymbol{P} \le \overline{\boldsymbol{P}}$, we have $\boldsymbol{\mu}\cdot\boldsymbol{R} \le g(\boldsymbol{\lambda})$. This implies $p^* \le g(\boldsymbol{\lambda})$ for any $\boldsymbol{\lambda} \ge 0$. Notice that the supremum is taken over the entire set S *without* taking the power constraints into effect. Additionally, the dual function $g(\boldsymbol{\lambda})$ is a convex function of $\boldsymbol{\lambda}$ since $g(\boldsymbol{\lambda})$ is the pointwise supremum of affine (and therefore convex) functions of $\boldsymbol{\lambda}$. This implies that the function $g(\boldsymbol{\lambda})$ is continuous because g is defined over all $\boldsymbol{\lambda}$ and any convex function is continuous on the relative interior of its domain [22].

By minimizing the dual function over all non-negative Lagrange multipliers, we get an upper bound d^* on the optimal value p^*. Due to the convexity and feasibility of the problem, this bound is tight [22] [21]:

$$d^* \triangleq \min_{\boldsymbol{\lambda}\ge 0} g(\boldsymbol{\lambda}) \triangleq g(\boldsymbol{\lambda^*}) = p^* \tag{57}$$

where $\boldsymbol{\lambda^*} = (\lambda_1^*,\ldots,\lambda_K^*)$ are the optimum Lagrange multipliers which lead to d^*. It can easily be shown that $g(\boldsymbol{\lambda}) = \infty$ if any $\lambda_i = \infty$, so we know $\boldsymbol{\lambda^*}$ is finite.

We now show that $\mu_i > 0$ implies $\lambda_i^* > 0$. Assume $\lambda_i = 0$. Since $\mu_i > 0$, we know $\mu_{i,j} > 0$ for some receiver j. Choose $P_j = \overline{P}_j$ for all $j \ne i$ and let P_i be arbitrarily large. Additionally, choose all rates to be zero except for $R_{i,j}$. Due to the unbounded condition on $R(\boldsymbol{P})$, letting P_i be arbitrarily large implies $R_{i,j}$ can be made arbitrarily large while still maintaining $(\boldsymbol{R},\boldsymbol{P}) \in S$. For this choice of $(\boldsymbol{R},\boldsymbol{P})$, we have $L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}) = \mu_{i,j}R_{i,j}$. This in turn implies $g(\boldsymbol{\lambda}) = \infty$. Since $p^* = \min_{\boldsymbol{\lambda}\ge 0} g(\boldsymbol{\lambda}) = g(\boldsymbol{\lambda^*})$ is finite, we have $\lambda_i^* \ne 0$.

We now show that $\mu_i = 0$ implies $\lambda_i^* = 0$. Without loss of generality, assume $\mu_1 = 0$. Consider the maximization of $L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda})$. For any $(\boldsymbol{R},\boldsymbol{P}) \in S$, define $\boldsymbol{Q}$ and $\boldsymbol{T}$ as

$$Q_{k,l} = \begin{cases} R_{k,l} & k \ne 1 \\ 0 & k = 1 \end{cases}, \quad T_k = \begin{cases} P_k & k \ne 1 \\ b & k = 1 \end{cases}$$

where b is any positive constant. Since $\mu_1 = 0$, we see that $L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}) = L(\boldsymbol{Q},\boldsymbol{T},\boldsymbol{\lambda})$ and condition 5 implies $(\boldsymbol{Q},\boldsymbol{T}) \in S$. Therefore we can equivalently perform the maximization of L assuming the rates from transmitter 1 are zero and over any $P_1 \geq 0$. Therefore we can decompose $g(\boldsymbol{\lambda})$ as

$$
\begin{aligned}
g(\boldsymbol{\lambda}) &= \sup_{P_1 \geq 0,\ ((0,\ldots,0,R_{2,1},\ldots,R_{K,N}),(0,P_2,\ldots,P_K)) \in S} \boldsymbol{\mu} \cdot \boldsymbol{R} - \lambda_1(P_1 - \overline{P}_1) - \cdots - \lambda_K(P_K - \overline{P}_K) \\
&= \sup_{P_1 \geq 0} \sup_{((0,\ldots,0,R_{2,1},\ldots,R_{K,N}),(0,P_2,\ldots,P_K)) \in S} \boldsymbol{\mu} \cdot \boldsymbol{R} - \lambda_1(P_1 - \overline{P}_1) - \cdots - \lambda_K(P_K - \overline{P}_K) \\
&= \sup_{P_1 \geq 0} -\lambda_1(P_1 - \overline{P}_1) + \\
&\qquad \sup_{((0,\ldots,0,R_{2,1},\ldots,R_{K,N}),(0,P_2,\ldots,P_K)) \in S} \boldsymbol{\mu} \cdot \boldsymbol{R} - \lambda_2(P_2 - \overline{P}_2) - \cdots - \lambda_K(P_K - \overline{P}_K) \\
&= \lambda_1 \overline{P}_1 + \sup_{((0,\ldots,0,R_{2,1},\ldots,R_{K,N}),(0,P_2,\ldots,P_K)) \in S} \boldsymbol{\mu} \cdot \boldsymbol{R} - \lambda_2(P_2 - \overline{P}_2) - \cdots - \lambda_K(P_K - \overline{P}_K).
\end{aligned}
$$

The Lagrangian variable λ_1 only affects the first term $\lambda_1 \overline{P}_1$, and therefore for any $\lambda_1 > 0$, we have $g(\lambda_1, \lambda_2, \ldots, \lambda_K) > g(0, \lambda_2, \ldots, \lambda_K)$. Since $d^* = \min_{\boldsymbol{\lambda} \geq 0} g(\boldsymbol{\lambda})$, this implies $\lambda_1^* = 0$.

Now consider the scaled MAC with α_i defined as

$$
\alpha_i = \begin{cases} \frac{\lambda_K^*}{\lambda_i^*} & \text{if } \lambda_i^* > 0 \\ c & \text{if } \lambda_i^* = 0 \end{cases} \tag{58}
$$

for $i = 1, \ldots, K$ and where $c > 0$ is some positive constant. Notice that $\boldsymbol{\alpha}$ is an implicit function of c by this definition. Since $\lambda_K^* > 0$ due to the fact that $\mu_K > 0$, we have $\alpha_i > 0$ for all i. We will now consider the sum power constraint capacity region of the scaled MAC with $\boldsymbol{\alpha}$ as defined above. Consider the following optimization on the scaled MAC:

$$
\sup_{\boldsymbol{R}} \boldsymbol{\mu} \cdot \boldsymbol{R} \quad s.t. \;\; \boldsymbol{R} \in R_{\boldsymbol{\alpha}}^{sum}(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}). \tag{59}
$$

By the definition of the sum power constraint capacity region, $\boldsymbol{R} \in R_{\boldsymbol{\alpha}}^{sum}(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})$ is equivalent to $\boldsymbol{R} \in R_{\boldsymbol{\alpha}}(\boldsymbol{P})$ for any $\boldsymbol{P}$ satisfying $\mathbf{1} \cdot \boldsymbol{P} \leq \mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}$. The above maximization (59) can thus be rewritten as

$$
\sup_{\boldsymbol{P} \in R_K^+,\ \boldsymbol{R} \in R_{\boldsymbol{\alpha}}(\boldsymbol{P})} \boldsymbol{\mu} \cdot \boldsymbol{R} \quad s.t. \;\; \mathbf{1} \cdot \boldsymbol{P} \leq \mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}. \tag{60}
$$

We denote the solution to this by $d_{\boldsymbol{\alpha}}^*$. Because there is a sum power constraint, there is only one Lagrange multiplier and the Lagrangian therefore is:

$$
\begin{aligned}
H_{\boldsymbol{\alpha}}(\boldsymbol{R}, \boldsymbol{P}, \nu) &= \boldsymbol{\mu} \cdot \boldsymbol{R} - \nu(\mathbf{1} \cdot \boldsymbol{P} - \mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}) && (61) \\
&= \boldsymbol{\mu} \cdot \boldsymbol{R} - \nu(P_1 - \frac{\overline{P}_1}{\alpha_1}) - \cdots - \nu(P_K - \frac{\overline{P}_K}{\alpha_K}) && (62) \\
&= \boldsymbol{\mu} \cdot \boldsymbol{R} - \frac{\nu}{\alpha_1}(\alpha_1 P_1 - \overline{P}_1) - \cdots - \frac{\nu}{\alpha_K}(\alpha_K P_K - \overline{P}_K) && (63)
\end{aligned}
$$

and the corresponding Lagrangian dual function is:

$$f_{\boldsymbol{\alpha}}(\nu) = \sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R_{\boldsymbol{\alpha}}(\boldsymbol{P})} H_{\boldsymbol{\alpha}}(\boldsymbol{R}, \boldsymbol{P}, \nu). \tag{64}$$

Again, the dual function satisfies $f_{\boldsymbol{\alpha}}(\nu) \geq d^*_{\boldsymbol{\alpha}}$ for all $\nu \geq 0$ (and in fact the minimum of the dual function is equal to $d^*_{\boldsymbol{\alpha}}$ by convexity). Due to the fact that $R_{\boldsymbol{\alpha}}(\boldsymbol{P}) = R(\boldsymbol{\alpha P})$ and $\boldsymbol{\alpha} > 0$, we can simplify the dual function as:

$$\begin{aligned}
f_{\boldsymbol{\alpha}}(\nu) &= \sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R_{\boldsymbol{\alpha}}(\boldsymbol{P})} H_{\boldsymbol{\alpha}}(\boldsymbol{R}, \boldsymbol{P}, \nu) && (65)\\
&= \sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R_{\boldsymbol{\alpha}}(\boldsymbol{P})} \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}(\alpha_1 P_1 - \overline{P}_1) - \cdots - \frac{\nu}{\alpha_K}(\alpha_K P_K - \overline{P}_K) && (66)\\
&= \sup_{\boldsymbol{\alpha P}\in R_K^+,\ \boldsymbol{R}\in R(\boldsymbol{\alpha P})} \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}(\alpha_1 P_1 - \overline{P}_1) - \cdots - \frac{\nu}{\alpha_K}(\alpha_K P_K - \overline{P}_K) && (67)\\
&= \sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R(\boldsymbol{P})} \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}(P_1 - \overline{P}_1) - \cdots - \frac{\nu}{\alpha_K}(P_K - \overline{P}_K) && (68)\\
&= g(\frac{\nu}{\boldsymbol{\alpha}}) && (69)
\end{aligned}$$

where g is the Lagrangian dual function of the individual power constraint unscaled MAC (56).

If we evaluate the dual function with $\nu = \lambda^*_K$, we get

$$f_{\boldsymbol{\alpha}}(\lambda^*_K) = g(\frac{\lambda^*_K}{\boldsymbol{\alpha}}) = g(\frac{\lambda^*_K}{\alpha_1}, \ldots, \frac{\lambda^*_K}{\alpha_K}).$$

Having established this, there are now two cases to consider: $\mu_i > 0$ for all i and $\mu_i = 0$ for some i.

If $\mu_i > 0$ for all i, then $\alpha_i = \frac{\lambda^*_K}{\lambda^*_i}$ for all i. Thus

$$f_{\boldsymbol{\alpha}}(\lambda^*_K) = g(\frac{\lambda^*_K}{\alpha_1}, \ldots, \frac{\lambda^*_K}{\alpha_K}) = g(\lambda^*_1, \ldots, \lambda^*_K) = p^*.$$

Since $f_{\boldsymbol{\alpha}}(\lambda^*_K)$ is an upper bound to $d^*_{\boldsymbol{\alpha}}$, we have $d^*_{\boldsymbol{\alpha}} \leq p^*$. This implies

$$\max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R} \geq \sup_{\boldsymbol{R}\in R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})} \boldsymbol{\mu}\cdot\boldsymbol{R} \geq \sup_{\boldsymbol{R}\in\bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu}\cdot\boldsymbol{R} \tag{70}$$

where the second inequality follows from $R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}) \supseteq \bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)$.

If $\mu_i = 0$ for some i, we must consider $\boldsymbol{\alpha}$-scalings for *different* values of c. Assume without loss of generality that $\mu_i = 0$ for $i = 1, \ldots, K$ and $\mu_i > 0$ for $i = K+1, \ldots, K$. We established earlier that $\mu_i = 0$ implies $\lambda^*_i = 0$. Therefore, we have $p^* = g\left(\lambda^*_1, \ldots, \lambda^*_K, \lambda^*_{K+1}, \ldots, \lambda^*_K\right) =$

$g\left(0,\ldots,0,\lambda^*_{K+1},\ldots,\lambda^*_K\right)$. Using the fact that $f_{\boldsymbol{\alpha}}(\nu) = g\left(\frac{\nu}{\boldsymbol{\alpha}}\right)\ \forall \boldsymbol{\alpha} > 0$, we have

$$\begin{aligned} f_{\boldsymbol{\alpha}}\left(\lambda^*_K\right) &= g\left(\frac{\lambda^*_K}{\boldsymbol{\alpha}}\right) \\ &= g\left(\frac{\lambda^*_K}{c},\ldots,\frac{\lambda^*_K}{c},\frac{\lambda^*_K}{\alpha_{K+1}},\ldots,\frac{\lambda^*_K}{\alpha_K}\right) \\ &= g\left(\frac{\lambda^*_K}{c},\ldots,\frac{\lambda^*_K}{c},\lambda^*_{K+1},\ldots,\lambda^*_K\right). \end{aligned}$$

Here $f_{\boldsymbol{\alpha}}(\lambda^*_K)$ is a function of the constant c because $\boldsymbol{\alpha}$ depends on c, as defined in (58). For a fixed $\boldsymbol{\alpha}$ (i.e. a fixed c), the optimum value of the sum power constraint region satisfies $d^*_{\boldsymbol{\alpha}} \leq f_{\boldsymbol{\alpha}}(\lambda^*_K)$.

We now show the desired result by contradiction. Assume

$$\sup_{\boldsymbol{R}\in\bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu}\cdot\boldsymbol{R} \ > \ \max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R}, \tag{71}$$

or equivalently

$$\left(\sup_{\boldsymbol{R}\in\bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu}\cdot\boldsymbol{R}\right) = \left(\max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R}\right) + \epsilon \tag{72}$$

for some $\epsilon > 0$. Since $R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}) \supseteq \bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)\ \forall\boldsymbol{\alpha} > 0$, (72) implies that for all $\boldsymbol{\alpha} > 0$

$$\left(\sup_{\boldsymbol{R}\in R^{sum}_{\boldsymbol{\alpha}}(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})} \boldsymbol{\mu}\cdot\boldsymbol{R}\right) \ \geq \ \left(\max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R}\right) + \epsilon. \tag{73}$$

This implies that for all $\boldsymbol{\alpha} > 0$, $d^*_{\boldsymbol{\alpha}} \geq p^* + \epsilon = g\left(0,\ldots,0,\lambda^*_{K+1},\ldots,\lambda^*_K\right) + \epsilon$. However, earlier we established that $d^*_{\boldsymbol{\alpha}} \leq g\left(\frac{\lambda^*_K}{c},\ldots,\frac{\lambda^*_K}{c},\lambda^*_{K+1},\ldots,\lambda^*_K\right)$. Thus we have that $g\left(\frac{\lambda^*_K}{c},\ldots,\frac{\lambda^*_K}{c},\lambda^*_{K+1},\ldots,\lambda^*_K\right) \geq g\left(0,\ldots,0,\lambda^*_{K+1},\ldots,\lambda^*_K\right) + \epsilon$ for all c. Since g is a convex function, the function g must lie *beneath* the line between $g\left(0,\ldots,0,\lambda^*_{K+1},\ldots,\lambda^*_K\right)$ and $g\left(1,\ldots,1,\lambda^*_{K+1},\ldots,\lambda^*_K\right)$. This contradicts

$$g\left(\frac{\lambda^*_K}{c},\ldots,\frac{\lambda^*_K}{c},\lambda^*_{K+1},\ldots,\lambda^*_K\right) \geq g\left(0,\ldots,0,\lambda^*_{K+1},\ldots,\lambda^*_K\right) + \epsilon \quad \forall c \geq \lambda_K \tag{74}$$

Thus (71) must be false and therefore

$$\max_{\boldsymbol{R}\in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu}\cdot\boldsymbol{R} \geq \sup_{\boldsymbol{R}\in\bigcap_{\boldsymbol{\alpha}>0} R^{sum}_{\boldsymbol{\alpha}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu}\cdot\boldsymbol{R} \tag{75}$$

for all $\boldsymbol{\mu}$ such that $\mu_i = 0$ for some i.

We have now shown that the above relationship (75) holds for *all* $\boldsymbol{\mu} \geq 0$ and the proof is complete.

REFERENCES

[1] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the capacity of multiple input multiple output broadcast channels," in *Proceedings of Int. Conf. Commun.*, pp. 1444–1450, April 2002.

[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," Submitted to IEEE Trans. Inform. Theory Aug. 2002.

[3] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels–Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1083–1102, March 2001.

[4] D. Tse and S. Hanly, "Multiaccess fading channels–Part I:Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796–2815, November 1998.

[5] L. Li and A. Goldsmith, "Outage capacities and optimal power allocation for fading multiple-access channel," *IEEE Trans. Inform. Theory*, 1999. In revision.

[6] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels–Part II: Outage capacity," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1103–1127, March 2001.

[7] N. Jindal and A. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," in *Proceedings of IEEE Globecom*, pp. 1296–1292, November 2001. Also submitted for publication to IEEE Trans. Inform. Theory.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[9] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *Proceedings of IEEE Globecom*, pp. 1344–1348, October 2001.

[10] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[11] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," in *International Symposium on Information Theory and its Applications*, pp. 681–684, November 2000.

[12] S. Vishwanath, S. Jafar, and A. Goldsmith, "Optimum power and rate allocation strategies for multiple access fading channels," in *Proceedings of IEEE Vehicular Technology Conference*, pp. 2888–2892, May 2001.

[13] S. Hanly and D. Tse, "Multiaccess fading channels–Part II:Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2816–2831, November 1998.

[14] R. Cheng and S. Verdu, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, May 1993.

[15] A. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 47, pp. 211–219, Jan. 2001.

[16] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *38th Annual Allerton Conference on Commun., Control and Computing*, pp. 1188–1193, 2000.

[17] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for vector multiple access channels," in *Proceedings of IEEE International Symposium on Information Theory*, p. 322, 2001.

[18] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecomm. ETT*, vol. 10, pp. 585–596, November 1999.

[19] S. Verdu, "Multiple-access channels with memory with and without frame synchronism," *IEEE Trans. Inform. Theory*, vol. 35, pp. 605–619, May 1989.

[20] H. Sato, “On degraded two-user gaussian channels,” *IEEE Trans. Inform. Theory*, vol. 24, pp. 637–640, Sept 1978.

[21] D. Luenberger, *Optimization by Vector Space Methods.* Wiley, 1969.

[22] S. Boyd and L. Vandenberghe, *Introduction to convex optimization with engineering applications.* Course Reader, 2001.