On the Duality of Gaussian Multiple-Access and Broadcast Channels

Nihar Jindal, *Student Member, IEEE*, Sriram Vishwanath, and Andrea Goldsmith, *Senior Member, IEEE*

***Abstract*—We define a duality between Gaussian multiple-access channels (MACs) and Gaussian broadcast channels (BCs). 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 the BC (both constant and fading) can be written in terms of the capacity region of the dual MAC, and *vice versa*. We can use this result to find the capacity region of the MAC if the capacity region of only the BC 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.**

***Index Terms*—Broadcast channel (BC), channel capacity, duality, fading channels, multiple-input multiple-output (MIMO) systems, multiple-access channel (MAC).**

I. Introduction

In this paper, we show that the scalar Gaussian multiple-access channel (MAC) and broadcast channel (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 that 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 and then subtracted from the received signal, then the next user is decoded and subtracted out, and so on. Since the Gaussian BC is a degraded BC, superposition coding is optimal [1]. 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 the Gaussian BC is equal to the capacity region of the dual Gaussian MAC subject to the same *sum* power constraint instead of the standard individual power constraints. Alternatively, the capacity region of the Gaussian BC is equal to the *union* of capacity regions of the dual MAC, where the union is taken over all individual power constraints that sum up to the BC power constraint. If we let $\mathcal{C}_{\rm BC}(P;\boldsymbol{h})$ represent the BC capacity region and $\mathcal{C}_{\rm MAC}(P_1,\ldots,P_K;\boldsymbol{h})$ represent the MAC capacity region, this can be stated as

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

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 jointly.

We then use the equivalence of the BC and sum power constraint MAC to find an expression for the capacity region of the individual power constraint MAC in terms of the capacity region of the dual BC. We show that the MAC capacity region is equal to the *intersection* of dual BC capacity regions. Using the same notation as above, this can be expressed as

$$\mathcal{C}_{\rm MAC}(\overline{P}_1,\ldots,\overline{P}_K;\boldsymbol{h}) = \bigcap_{\{\alpha_i\}_{i=1}^K:\alpha_i>0} \mathcal{C}_{\rm BC}\left(\sum_{i=1}^K \frac{\overline{P}_i}{\alpha_i};\alpha_1 h_1,\ldots,\alpha_K h_K\right). \quad (2)$$

This result follows from a general theorem characterizing the individual power constraint MAC in terms of the sum power constraint MAC capacity region, which is equal to the dual BC capacity region by (1).

Manuscript received September 7, 2002; revised September 25, 2003. This work was supported in part by the Office of Naval Research under Grants N00014-99-1-0578 and N00014-01-1-0003. The work of S. Vishwanath was supported by a Stanford Graduate Fellowship.

N. Jindal and A. Goldsmith are with the Department of Electrical Engineering, Stanford University, Stanford, CA, 94305 USA (e-mail: njindal@systems.stanford.edu; andrea@systems.stanford.edu).

S. Vishwanath was with the Department of Electrical Engineering, Stanford University, Stanford, CA, 94305 USA. He is now with the Department of Electrical and Computer Engineering, University of Texas, Austin, TX, 78712 USA (e-mail: sriram@ece.utexas.edu).

Communicated by D. N. C. Tse, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2004.826646

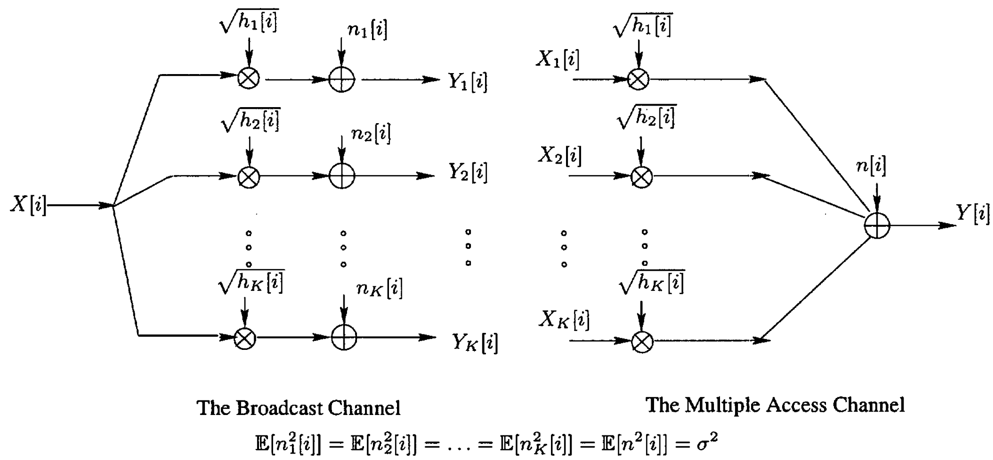


Fig. 1. System models.

In addition to the constant additive white Gaussian noise (AWGN) channel, we consider flat-fading channels with perfect channel state information (CSI) at all transmitters and receivers. Fading channels can be decomposed into a set of parallel constant channels, one for each fading state. Using the duality of constant channels established in (1), we show that duality holds for the ergodic capacity region of fading channels as well. We also show that the relationship in (2) holds for fading channels. Duality also holds for outage capacity and minimum-rate capacity. Though the ergodic capacity regions [2], [3] and outage capacity regions [4], [5] of both the MAC and BC have previously been found, duality ties these results together. Minimum-rate capacity has only been found for the BC [6], but using duality we can find the minimum-rate capacity of the MAC as well.

Duality is an exciting new concept that gives great insight into the similarities between the Gaussian MAC and BC, as well as their capacities and optimal transmission strategies. It also opens up the possibility that a more general information-theoretic duality exists between the MAC and BC [7]. Duality has also been very useful in proving new results, most prominently for the multiple-antenna BC. In [8], duality is extended to multiple-antenna Gaussian channels and in [8], [9] this duality has been used to find the sum capacity of the multiple-antenna BC. Duality also greatly simplifies numerical computation of the multiple-antenna BC sum rate capacity and achievable region [10], [11].

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. We consider some extensions of this duality in Section VII, 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 BC: 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[1] $h_j[i]$, and white Gaussian noise $n_j[i]$ is added to the received signal. We let $\boldsymbol{h}[i] = (h_1[i], \ldots, h_K[i])$ denote the vector of channel gains at time i.

The system to the right is a MAC: a many-to-one system, where K independent transmitters each send 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 $n[i]$.

Mathematically, the two systems can be described as

$$\text{BC:} \quad Y_j[i] = \sqrt{h_j[i]}X[i] + n_j[i]$$

$$\text{MAC:} \quad Y[i] = \sum_{j=1}^{K} \sqrt{h_j[i]}X_j[i] + n[i].$$

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 need 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

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

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 do not change too rapidly.

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 MAC

From [1], the capacity region of a Gaussian MAC with channel gains $\boldsymbol{h} = (h_1, \ldots, h_K)$ and power constraints $\overline{\boldsymbol{P}} = (\overline{P_1}, \ldots, \overline{P_K})$ is

$$\mathcal{C}_{\text{MAC}}(\overline{\boldsymbol{P}}; \boldsymbol{h}) = \left\{ \mathbf{R} : \sum_{j \in S} R_j \leq \frac{1}{2} \log \left(1 + \frac{1}{\sigma^2} \sum_{j \in S} h_j \overline{P}_j \right) \right.$$
$$\left. \forall S \subseteq \{1, \ldots, K\} \right\}. \quad (3)$$

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 [1]. 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

$$R_{\pi(j)} = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)} \overline{P}_{\pi(j)}}{\sigma^2 + \sum_{i=j+1}^{K} h_{\pi(i)} \overline{P}_{\pi(i)}} \right), \quad j = 1, \ldots, K. \quad (4)$$

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 all other rate vectors that lie below this convex hull (i.e., are componentwise less than or equal to a rate vector in the convex hull).

B. Capacity Region of the BC

From [1], the capacity region of a Gaussian BC with channel gains $\boldsymbol{h} = (h_1, \ldots, h_K)$ and power constraint $\overline{P}$ is

$$\mathcal{C}_{\text{BC}}(\overline{P}; \boldsymbol{h}) = \left\{ \mathbf{R} : R_j \leq \frac{1}{2} \log \left(1 + \frac{h_j P_j^B}{\sigma^2 + h_j \sum_{k=1}^{K} P_k^B \mathbf{1}[h_k > h_j]} \right) \quad j = 1, \ldots, K \right\} \quad (5)$$

over all power allocations such that $\sum_{j=1}^{K} P_j^B = \overline{P}$. Additionally, any rate vector taking the form of (5) 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 h_i.

As seen in [12], the capacity region of the BC is also achievable via "dirty-paper coding," in which the transmitter "presubtracts" (similar to precoding for equalization) certain users' codewords instead of receivers decoding and subtracting out other users' signals. When users are encoded in order of increasing channel gains, this technique achieves capacity and is equivalent to successive decoding. Though suboptimal, dirty-paper coding can be performed using any other encoding order as well.[2] 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

$$R_{\pi(j)}^B = \frac{1}{2} \log \left(1 + \frac{h_{\pi(j)} P_{\pi(j)}^B}{\sigma^2 + h_{\pi(j)} \sum_{i=j+1}^{K} P_{\pi(i)}^B} \right). \quad (6)$$

Clearly, these rates are achievable and thus are in the BC capacity region. In fact, any rates of the form of (6) for any encoding order $\pi(\cdot)$ and any power allocation such that $\sum_{i=1}^{N} P_i^B = \overline{P}$ lie in $\mathcal{C}_{\text{BC}}(\overline{P}; \boldsymbol{h})$. Note that if $\pi(\cdot)$ is in order of increasing channel gains, then the rate vector lies on the boundary of the capacity region.

C. MAC to BC

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

Theorem 1: 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 $\sum_{j=1}^{K} P_j = \overline{P}$

$$\boxed{\mathcal{C}_{\text{BC}}(\overline{P}; \boldsymbol{h}) = \bigcup_{\{\boldsymbol{P} : \mathbf{1} \cdot \boldsymbol{P} = \overline{P}\}} \mathcal{C}_{\text{MAC}}(\boldsymbol{P}; \boldsymbol{h}).} \quad (7)$$

Proof: We will show that every point on the boundary of the BC capacity region is a corner point of the dual MAC for some set of powers with the same sum power and that every corner point of the MAC for every set of powers is in the dual BC capacity region with the same sum power. Since the MAC and BC capacity regions are convex, this suffices to prove the result.

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

[2] Note that with successive decoding, when a suboptimal decoding order is used it must be ensured that all users who are supposed to decode and subtract out a certain user's signal have a large enough channel gain to do so. This limits the rates achievable using successive decoding with a sub-optimal decoding order.

$(1, \ldots, K)$. The rate of User $\pi(j)$ in the MAC at this successive decoding point is

$$R^M_{\pi(j)} = \frac{1}{2}\log\left(1 + \frac{h_{\pi(j)} P^M_{\pi(j)}}{\sigma^2 + \sum_{i=j+1}^{K} h_{\pi(i)} P^M_{\pi(i)}}\right).$$

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 $(P^B_1, \ldots, P^B_K)$ are used is

$$R^B_{\pi(j)} = \frac{1}{2}\log\left(1 + \frac{h_{\pi(j)} P^B_{\pi(j)}}{\sigma^2 + h_{\pi(j)} \sum_{i=1}^{j-1} P^B_{\pi(i)}}\right).$$

By defining A_j and B_j as

$$A_j = \sigma^2 + h_{\pi(j)} \sum_{i=1}^{j-1} P^B_{\pi(i)}, \quad B_j = \sigma^2 + \sum_{i=j+1}^{K} h_{\pi(i)} P^M_{\pi(i)} \tag{8}$$

we can rewrite the rates in the MAC and BC as

$$R^M_{\pi(j)} = \frac{1}{2}\log\left(1 + \frac{h_{\pi(j)} P^M_{\pi(j)}}{B_j}\right) \tag{9}$$

$$R^B_{\pi(j)} = \frac{1}{2}\log\left(1 + \frac{h_{\pi(j)} P^B_{\pi(j)}}{A_j}\right). \tag{10}$$

Thus, if the powers satisfy

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

then the rates in the MAC using powers $(P^M_1, \ldots, P^M_K)$ and decoding order $(\pi(1), \ldots, \pi(K))$ are the same as the rates in the BC using powers $(P^B_1, \ldots, P^B_K)$ and encoding order $(\pi(K), \ldots, \pi(1))$. In Appendix A, we show that if the powers satisfy (11), then

$$\sum_{j=1}^{K} P^M_j = \sum_{j=1}^{K} P^B_j.$$

We now need only show that given a set of MAC powers and a MAC decoding order, there exist a set of BC powers satisfying (11), and *vice versa*. To compute BC powers from MAC powers, the relationship in (11) must be evaluated in numerical order, starting with User $\pi(1)$

$$\begin{aligned}
P^B_{\pi(1)} &= P^M_{\pi(1)} \frac{\sigma^2}{\sigma^2 + \sum_{i=2}^{K} h_{\pi(i)} P^M_{\pi(i)}} \\
P^B_{\pi(2)} &= P^M_{\pi(2)} \frac{\sigma^2 + h_{\pi(2)} P^B_{\pi(1)}}{\sigma^2 + \sum_{i=3}^{K} h_{\pi(i)} P^M_{\pi(i)}} \\
&\cdots \\
P^B_{\pi(K)} &= P^M_{\pi(K)} \frac{\sigma^2 + h_{\pi(K)} \sum_{i=1}^{K-1} P^B_{\pi(i)}}{\sigma^2}.
\end{aligned} \tag{12}$$

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. Therefore, any successive decoding point of the MAC region for any set of powers $(P^M_1, \ldots, P^M_K)$ with $\sum_{i=1}^{K} P^M_i = P$ is in the dual BC capacity region.

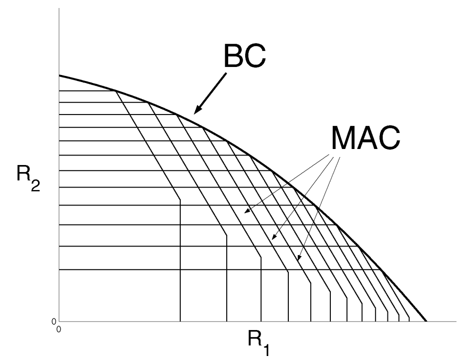


Fig. 2. Constant BC capacity in terms of the dual MAC.

Similarly, MAC powers can be derived from BC powers starting with User $\pi(K)$ downwards

$$\begin{aligned}
P^M_{\pi(K)} &= P^B_{\pi(K)} \frac{\sigma^2}{\sigma^2 + h_{\pi(K)} \sum_{i=1}^{K-1} P^B_{\pi(i)}} \\
P^M_{\pi(K-1)} &= P^B_{\pi(K-1)} \frac{\sigma^2 + h_{\pi(K)} P^M_{\pi(K)}}{\sigma^2 + h_{\pi(K-1)} \sum_{i=1}^{K-2} P^B_{\pi(i)}} \\
&\cdots \\
P^M_{\pi(1)} &= P^B_{\pi(K)} \frac{\sigma^2 + \sum_{i=2}^{K} h_{\pi(i)} P^M_{\pi(i)}}{\sigma}.
\end{aligned} \tag{13}$$

If we consider only permutations corresponding to encoding in order of increasing channel gain, we see that any point on the boundary of the BC capacity region is in the dual MAC region for some set of MAC powers with the same sum power. □

Note that we refer to (12) and (13) as the MAC-BC transformations and BC-MAC transformations, respectively.

Corollary 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 $P = \mathbf{1}\cdot\boldsymbol{P}$

$$\mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P}; \boldsymbol{h}) \subseteq \mathcal{C}_{\mathrm{BC}}(\mathbf{1}\cdot\boldsymbol{P}; \boldsymbol{h}). \tag{14}$$

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

Proof: See Appendix B. □

Theorem 1 is illustrated in Fig. 2, where $\mathcal{C}_{\mathrm{MAC}}(P_1, P - P_1; h_1, h_2)$ 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, as specified by Corollary 1.

If we carefully examine the union expression in the characterization of the BC in terms of the dual MAC in (7), it is easy to see that the union of MACs is equal to the capacity region of the MAC with a *sum* power constraint $P = \sum_{i=1}^{K} P_i$ 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 transmits an independent message) but the transmitters are allowed to draw from a common power source. Therefore, Theorem 1 implies that the capacity region of the

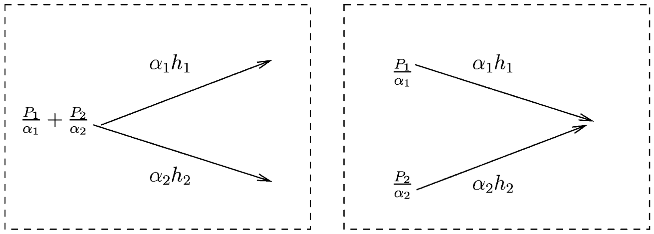


Fig. 3. Scaled dual channels.

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 (i.e., encode last) to the strongest user, whereas in the sum power MAC, it is optimal to give priority (i.e., decode last) to the weakest user.

D. BC to MAC

In this subsection, 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 and P_j always appear as a product in the constant MAC capacity expression (3), we can scale h_j by any positive constant α_j and scale P_j by the inverse of α_j without affecting the capacity region. Therefore, $\mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{h})=\mathcal{C}_{\rm MAC}\left(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h}\right)$ for any vector of constants $\boldsymbol{\alpha}>0$. The scaled dual channels are shown in Fig. 3. The scaling of the channel and the power constraints clearly negate each other in the MAC. However, the dual 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 Corollary 1 to the scaled MAC and the scaled BC, we find that

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

and the boundaries of the MAC capacity region and each scaled BC capacity region intersect. In fact, $\mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{h})$ and $\mathcal{C}_{\rm BC}(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h})$ intersect at the corner point of the MAC corresponding to decoding in decreasing order of scaled gains $\alpha_i h_i$, 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 2) that characterizes individual transmit power constraint rate regions of a Gaussian MAC in terms of sum transmit power constraint rate regions. We could directly establish a relationship between the MAC and the scaled BC for the constant channel. However, we present a more general theorem here that is applicable to fading channels as well. Before stating the theorem, 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 $\boldsymbol{R}=(R_1,\dots,R_K)$ where R_j is the rate of transmitter j. Let $\boldsymbol{P}=(P_1,\dots,P_K)$ be the vector of transmit power constraints and let $\boldsymbol{\alpha}=(\alpha_1,\dots,\alpha_K)$ 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 $\mathcal{R}_+^K$ that satisfies the conditions stated in Definition 2 that follows. 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 α_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) If $\boldsymbol{P_1}\geq\boldsymbol{P_2}$ then $R(\boldsymbol{P_1})\supseteq R(\boldsymbol{P_2})$.
5) If $(R_1,\dots R_K)\in R(P_1,P_2,\dots,P_K)$, then for any i
$$(R_1,\dots,R_{i-1},0,R_{i+1},\dots,R_K)\in R(P_1,\dots,P_{i-1},0,P_{i+1},\dots,P_K).$$
6) If $\boldsymbol{R}\in R(\boldsymbol{P})$ and $\boldsymbol{R}'\leq\boldsymbol{R}$, then $\boldsymbol{R}'\in R(\boldsymbol{P})$.
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 $R_{\boldsymbol{\alpha}}^{\rm sum}(P_{\rm sum})$ as

$$R_{\boldsymbol{\alpha}}^{\rm sum}(P_{\rm sum})\triangleq\bigcup_{\{\boldsymbol{P}|\ \boldsymbol{P}\in\mathcal{R}_+^K,\ \boldsymbol{1}\cdot\boldsymbol{P}\leq P_{\rm sum}\}}R_{\boldsymbol{\alpha}}(\boldsymbol{P}).\tag{15}$$

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

Theorem 2: 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 $\boldsymbol{P}$

$$R(\boldsymbol{P})=\bigcap_{\boldsymbol{\alpha}>0}R_{\boldsymbol{\alpha}}^{\rm sum}\left(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right).\tag{16}$$

Proof: See Appendix D. □

We now apply Theorem 2 to the capacity region of the constant MAC.

Theorem 3: 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

$$\boxed{\mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{h})=\bigcap_{\boldsymbol{\alpha}>0}\mathcal{C}_{\rm BC}\left(\boldsymbol{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h}\right).}\tag{17}$$

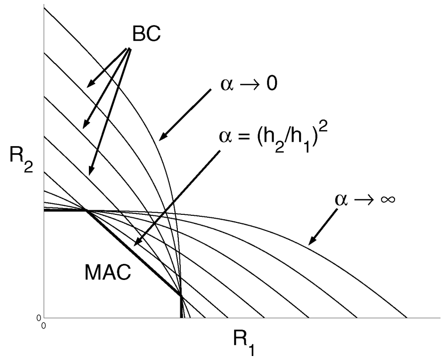


Fig. 4. Constant MAC capacity in terms of the dual BC.

Proof: In Appendix C, we show that the region $\mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P};\boldsymbol{h})$ satisfies the conditions of Definition 2. Therefore, by Theorem 2 we get

$$\mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P};\boldsymbol{h}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{\mathrm{MAC}}^{\mathrm{sum}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h}\right). \tag{18}$$

By Theorem 1

$$\mathcal{C}_{\mathrm{MAC}}^{\mathrm{sum}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h}\right) = \mathcal{C}_{\mathrm{BC}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha h}\right)$$

for any $\boldsymbol{\alpha} > 0$. Thus, the result follows. □

Theorem 3 is illustrated for a two-user channel in Fig. 4. Although we consider channel scaling of all K users in Theorem 2, 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 $\mathcal{C}_{\mathrm{BC}}(\frac{P_1}{\alpha} + P_2; \alpha h_1, h_2)$ for a range of values of $\alpha > 0$. Since the constant MAC region is a pentagon, the BC characterized by $\alpha = (h_2/h_1)$ and the limit of the BCs as $\alpha \to 0$ and $\alpha \to \infty$ are sufficient to form the pentagon. When $\alpha = (h_2/h_1)$, 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 45° 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 $R_1 \to \log(1 + \frac{h_1 P_1}{\sigma^2})$ 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 $R_2 \to \log(1 + \frac{h_2 P_2}{\sigma^2})$. These two limiting capacity regions bound the vertical and horizontal line segments, respectively, of the MAC capacity region boundary.

Additionally, by Corollary 1, all scaled BC capacity regions except the channel corresponding to $\alpha = (h_2/h_1)$ intersect the MAC at exactly one of the two corner points of the MAC region. Scaled BC capacity regions with $\alpha > (h_2/h_1)$ 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 $\alpha < (h_2/h_1)$ 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 $\{1, \ldots, K\}$). Therefore, in general, only $2^K - 1$ different scaled BC capacity regions are needed to get the MAC capacity region. One of these regions corresponds to $\boldsymbol{\alpha}$ such that $\alpha_i h_i = \alpha_j h_j$ 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 CSI at all transmitters and receivers. Flat-fading channels can be decomposed into an infinite set of parallel, independent channels, one for each joint fading state. The ergodic capacity for both the MAC and the BC is shown to be the "average" of the capacities of each of these independent channels. We can then use the duality of the MAC and BC for each fading state to show that duality holds for ergodic capacity as well. We first define the ergodic capacity regions of the fading MAC and BC, and then explicitly state the duality between the two regions.

A. Ergodic Capacity Region of the MAC

We define a power policy $\mathcal{P}_{\mathrm{MAC}}$ over all possible fading states as a function that maps from a joint fading state $\boldsymbol{h} = (h_1, \ldots, h_K)$ to the transmitted power $P_j^M(\boldsymbol{h})$ for each user. Let $\mathcal{F}_{\mathrm{MAC}}$ denote the set of all power policies satisfying the K individual average power constraints

$$\mathcal{F}_{\mathrm{MAC}} = \{\mathcal{P}_{\mathrm{MAC}} : \mathbb{E}_{\boldsymbol{H}}[P_j^M(\boldsymbol{h})] \le \overline{P}_j \quad 1 \le j \le K\}.$$

From [3, Theorem 2.1], the ergodic capacity region of the MAC with perfect CSI and power constraints $\overline{\boldsymbol{P}} = (\overline{P}_1, \ldots, \overline{P}_K)$ is

$$\mathcal{C}_{\mathrm{MAC}}(\overline{\boldsymbol{P}};\boldsymbol{H}) = \bigcup_{\mathcal{P}_{\mathrm{MAC}} \in \mathcal{F}_{\mathrm{MAC}}} \mathcal{C}_{\mathrm{MAC}}(\mathcal{P}_{\mathrm{MAC}};\boldsymbol{H}). \tag{19}$$

By [3, Lemma 3.8]

$$\mathcal{C}_{\mathrm{MAC}}(\mathcal{P}_{\mathrm{MAC}};\boldsymbol{H}) = \{\mathbb{E}_{\boldsymbol{H}}[\boldsymbol{R}(\boldsymbol{h})] : \boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P}^M(\boldsymbol{h});\boldsymbol{h})\, \forall \boldsymbol{h}\} \tag{20}$$

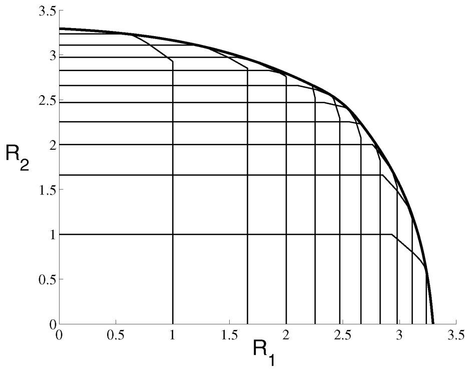
where $\boldsymbol{R}(\boldsymbol{h})$ is the rate vector of all K users as a function of the joint fading state and $\mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P}^M(\boldsymbol{h});\boldsymbol{h})$ is the constant MAC capacity region. Therefore, the ergodic capacity is clearly the average of the instantaneous rate[3] vector $\boldsymbol{R}(\boldsymbol{h})$.

B. Ergodic Capacity Region of the BC

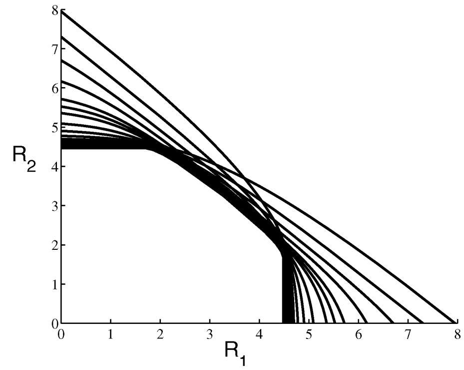
We define a power policy $\mathcal{P}_{\mathrm{BC}}$ over all possible fading states as a function over all joint fading states that maps from a joint fading state $\boldsymbol{h} = (h_1, \ldots, h_K)$ to the transmitted power $P_j^B(\boldsymbol{h})$ for each user. Let $\mathcal{F}_{\mathrm{BC}}$ denote the set of all power policies satisfying the average power constraint

$$\mathcal{F}_{\mathrm{BC}} = \left\{\mathcal{P}_{\mathrm{BC}} : \mathbb{E}_{\boldsymbol{H}}\left[\sum_{j=1}^{K} P_j^B(\boldsymbol{h})\right] \le \overline{P}\right\}.$$

[3]Note that the notion of an instantaneous rate is used only for mathematical convenience and the transmitted rate can, in fact, be constant over all states



(a) Fading MAC to BC



(b) Fading BC to MAC

Fig. 5. Duality of the fading MAC and BC.

From [2, Theorem 1], the ergodic capacity region of the BC with perfect CSI and power constraint $\overline{P}$ is the union over all power policies in $\mathcal{F}_{\rm BC}$

$$\mathcal{C}_{\rm BC}(\overline{P};\boldsymbol{H}) = \bigcup_{\mathcal{P}_{\rm BC}\in\mathcal{F}_{\rm BC}} \mathcal{C}_{\rm BC}(\mathcal{P}_{\rm BC};\boldsymbol{H}) \tag{21}$$

where

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

From this definition it follows that a rate vector $\boldsymbol{R}$ is in $\mathcal{C}_{\rm BC}(\overline{P};\boldsymbol{H})$ if and only if there exists $\boldsymbol{R}(\boldsymbol{h})$ and a $\boldsymbol{P}^B(\boldsymbol{h})$ in $\mathcal{F}_{\rm BC}$ such that $\boldsymbol{R} \le \mathbb{E}_{\boldsymbol{H}}[\boldsymbol{R}(\boldsymbol{h})]$ with $\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm BC}(\sum_{i=1}^K \boldsymbol{P}_i^B(\boldsymbol{h});\boldsymbol{h})$ for all $\boldsymbol{h}$.

C. MAC to BC

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

Theorem 4: 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 $(P_1,\dots,P_K)$ such that $\mathbf{1}\cdot\boldsymbol{P} = \overline{P}$

$$\mathcal{C}_{\rm BC}(\overline{P};\boldsymbol{H}) = \bigcup_{\mathbf{1}\cdot\boldsymbol{P}=\overline{P}} \mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{H}). \tag{22}$$

Proof: We show that any rate vector in the union of MAC regions is in the dual BC capacity region, and *vice versa*. By the definition of $\mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{H})$, a rate vector $\boldsymbol{R}$ is in the ergodic capacity region of the MAC if and only if $\boldsymbol{R} = \mathbb{E}_{\boldsymbol{H}}[\boldsymbol{R}(\boldsymbol{h})]$ where $\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm MAC}(\boldsymbol{P}^M(\boldsymbol{h});\boldsymbol{h})$ for all $\boldsymbol{h}$. By Theorem 1

$$\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm BC}\left(\sum_{i=1}^K \boldsymbol{P}_i^M(\boldsymbol{h});\boldsymbol{h}\right), \qquad \forall \boldsymbol{h}.$$

Therefore, $\boldsymbol{R} \in \mathcal{C}_{\rm BC}(\overline{P};\boldsymbol{H})$.

Similarly, a rate vector $\boldsymbol{R}$ is in the ergodic capacity region of the BC if and only if there exists $\boldsymbol{R}(\boldsymbol{h})$ such that $\boldsymbol{R} \le \mathbb{E}_{\boldsymbol{H}}[\boldsymbol{R}(\boldsymbol{h})]$ with

$$\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm BC}\left(\sum_{i=1}^K \boldsymbol{P}_i^B(\boldsymbol{h});\boldsymbol{h}\right), \qquad \forall \boldsymbol{h}$$

for some $\boldsymbol{P}^B(\boldsymbol{h})$ in $\mathcal{F}_{\rm BC}$. Applying Theorem 1 to each fading state, we get that $\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm MAC}(\boldsymbol{P}^M(\boldsymbol{h});\boldsymbol{h})$ for some $\boldsymbol{P}^M(\boldsymbol{h})$ such that

$$\sum_{i=1}^K \boldsymbol{P}_i^M(\boldsymbol{h}) = \sum_{i=1}^K \boldsymbol{P}_i^B(\boldsymbol{h})$$

for each $\boldsymbol{h}$. Therefore,

$$\mathbb{E}_{\boldsymbol{H}}\left[\sum_{i=1}^K \boldsymbol{P}_i^M(\boldsymbol{h})\right] = \mathbb{E}_{\boldsymbol{H}}\left[\sum_{i=1}^K \boldsymbol{P}_i^B(\boldsymbol{h})\right] = \overline{P}.$$

If we let $Q_i = \mathbb{E}_{\boldsymbol{H}}[\boldsymbol{P}_i^M(\boldsymbol{h})]$, then

$$\boldsymbol{R} \in \mathcal{C}_{\rm MAC}(Q_1,\dots,Q_K;\boldsymbol{H}) \in \bigcup_{\mathbf{1}\cdot\boldsymbol{P}=\overline{P}} \mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{H})$$

since $\sum_{i=1}^K Q_i = \overline{P}$. □

Intuitively, for any MAC power policy, we can use the MAC-BC transformations in each fading state to find a BC power policy that achieves the same rates in each fading state, and therefore the same average rates, using the same sum power in each state. Alternatively, for any BC power policy, we can find a dual MAC power policy that achieves the same rates while using the same sum power.

Fig. 5(a) illustrates Theorem 4. The pentagon-like regions are the dual MAC ergodic capacity regions, while the region denoted with a bold line is the BC ergodic capacity region. 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.

Corollary 2: 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 $P = \mathbf{1}\cdot\overline{\boldsymbol{P}}$

$$\mathcal{C}_{\rm MAC}(\overline{\boldsymbol{P}};\boldsymbol{H}) \subseteq \mathcal{C}_{\rm BC}(\mathbf{1}\cdot\overline{\boldsymbol{P}};\boldsymbol{H}). \tag{23}$$

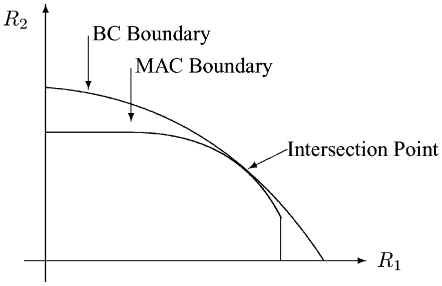


Fig. 6. Capacity regions for the dual fading MAC and BC.

Proof: This result is a direct consequence of Theorem 4. 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 (Corollary 1). We are able to show this for the $K = 2$ case, but not for arbitrary K. □

Fig. 6 illustrates the subset relationship established in Corollary 2 for the ergodic capacity regions of the dual flat-fading MAC and BC for a two-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. BC to MAC

In order to characterize the ergodic capacity region of the MAC in terms of the dual BC, we again use channel scaling. Channel scaling by the factor $\boldsymbol{\alpha}$ 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 $\tilde{\boldsymbol{H}} = \boldsymbol{\alpha H}$. It is easy to see that channel scaling does not affect the ergodic capacity region of a fading MAC, or that $\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H}) = \mathcal{C}_{\text{MAC}}(\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha H})$ for all $\boldsymbol{\alpha} > 0$. Using Theorem 2, we can find an expression for the ergodic capacity region of the MAC in terms of the dual BC.

Theorem 5: 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

$$\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{\text{BC}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}};\boldsymbol{\alpha H}\right). \tag{24}$$

Proof: The proof of this is identical to the proof for the constant channel version of this in Theorem 3. The fact that the ergodic capacity region of the MAC satisfies the conditions of Theorem 2 can be verified using the arguments used for the constant MAC capacity region in Appendix C. See Section IV-E for a discussion of Theorem 2 as applied to ergodic capacity. □

Theorem 5 is illustrated in Fig. 5(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-D. 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.

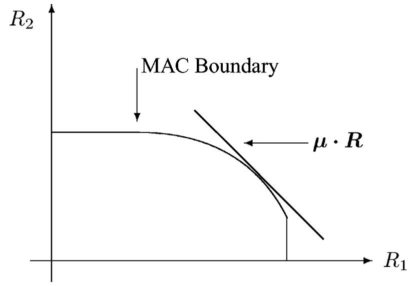


Fig. 7. MAC capacity region optimization.

E. Convex Optimization Interpretation

If we consider the boundary points of $\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H})$ from a convex optimization viewpoint, we can gain some additional insight into the MAC-BC duality and Theorem 5. Since the region $\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H})$ is closed and convex, we can fully characterize the region by the following convex maximization:

$$\max_{\boldsymbol{R}\in\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H})} \boldsymbol{\mu}\cdot\boldsymbol{R}, \qquad \text{such that} \quad \boldsymbol{P} \leq \overline{\boldsymbol{P}} \tag{25}$$

over all nonnegative priority vectors $\boldsymbol{\mu} = (\mu_1,\ldots,\mu_K)$ such that $\boldsymbol{\mu}\cdot\mathbf{1} = 1$. This maximization is shown pictorially in Fig. 7. Since (25) 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 maximum of $\boldsymbol{\mu}\cdot\boldsymbol{R}$ to a change in the power constraint $\overline{P_i}$.

For each nonnegative priority vector $\boldsymbol{\mu}$, there exists an optimum Lagrange multiplier $\boldsymbol{\lambda^*}$. If for some $\boldsymbol{\mu}$ we have $\lambda_1^* > \lambda_2^*$, then constraint $\overline{P}_1$ is more restrictive than constraint $\overline{P}_2$. In this scenario, increasing $\overline{P}_1$ while decreasing $\overline{P}_2$ by the same amount would lead to an increase in the maximum-weighted sum rate $\boldsymbol{\mu}\cdot\boldsymbol{R}$. On the other hand, if $\lambda_1^* = \lambda_2^* = \cdots = \lambda_K^*$, then each power constraint is equally "hard" and no tradeoff of power between different users would increase the maximum. Thus, 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 4, 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 $\lambda_1^* = \lambda_2^* = \cdots = \lambda_K^*$.

By scaling the channel gains, we can force the Lagrangians to be equal. If $\boldsymbol{\lambda^*}$ is the optimal Lagrange multiplier for some priority vector $\boldsymbol{\mu}$ for the unscaled MAC, then $\boldsymbol{\alpha\lambda^*}$ is the optimal Lagrange multiplier for the MAC scaled by $\boldsymbol{\alpha}$. Therefore, we can scale the channel appropriately so that $\alpha_i\lambda_i^*$ are equal for all i. Using this method, every point on the boundary of $\mathcal{C}_{\text{MAC}}(\boldsymbol{P};\boldsymbol{H})$ can be shown to be on the boundary of the sum power MAC (and, therefore, of the dual BC) for some scaling vector. The proof of Theorem 2 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 [3] and [2], respectively. Given a priority vector $\boldsymbol{\mu}$, it is possible to find the optimal power policy that maximizes $\boldsymbol{\mu} \cdot \boldsymbol{R}$ in both the MAC and the BC. Due to the duality of these channels, the optimal power policies derived independently for the BC and MAC are related by the MAC-BC (12) and BC-MAC (13) transformations at the points where the BC and MAC capacity region boundaries touch.

F. Optimal MAC/BC Decoding Order

The duality of the flat-fading MAC and BC leads to some interesting observations about the optimum decoding order in the BC and MAC. By duality, 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 2 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 $\mathcal{C}_{\rm MAC}(\boldsymbol{P};\boldsymbol{H})$ and in $\mathcal{C}_{\rm BC}(\boldsymbol{P};\boldsymbol{H})$ for the same $\boldsymbol{\mu}$.

From results on the ergodic capacity region of the MAC [3], [13], it is optimal to decode users in order of increasing priority μ_i 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 that is also a boundary point of a dual MAC capacity region. The optimal MAC and BC power policies will be related by the power transformations given earlier. Additionally, the decoding order in the BC and the MAC are opposite in each fading state. Therefore, by duality, we see that boundary points can be achieved in the BC by decoding in order of *decreasing* priority. From basic results on the BC, however, users should be decoded in order of increasing channel gain in every fading state. This apparent inconsistency is resolved by the fact that a user is allocated power in the BC only if all users with larger priority have smaller channel gains. Thus, decoding in order of decreasing priority is equivalent to decoding in order of increasing channel gain for the optimal BC power allocation policy.

Using duality in the form of Theorem 3, every boundary point of the MAC is also a boundary point of a scaled BC. In the $\boldsymbol{\alpha}$-scaled BC, users are decoded in increasing order of $\alpha_i h_i$. Since $\alpha_i = \frac{\lambda_K^*}{\lambda_i^*}$ for the correct scaling (see proof of Theorem 2 for justification), users are decoded in order of *increasing* $\frac{h_i}{\lambda_i^*}$ in the BC. By duality, the opposite decoding order should be used in the MAC. Thus, users should be decoded in order of *decreasing* $\frac{h_i}{\lambda_i^*}$. Again, the apparent inconsistency with decoding users in the MAC in order of increasing priority is resolved by the optimal power allocation policy.

G. Symmetric Channels

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 symmetry. As discussed in Section IV-E, this implies that the unscaled MAC and the unscaled BC (i.e., $\boldsymbol{\alpha} = \mathbf{1}$) 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 [2], [3]. For asymmetric fading distributions and/or power constraints, the uplink sum rate capacity is generally strictly less than the downlink sum rate capacity.

H. Frequency-Selective Channels

Duality easily extends to frequency-selective (intersymbol interference (ISI)) channels as well. BCs and MACs with time-invariant, finite-length impulse responses and additive Gaussian noise were considered in [14], [15]. The dual channels have the same impulse response on the uplink and downlink, and the same noise power at each receiver. Similar to flat-fading channels, frequency-selective channels can be decomposed into a set of parallel independent channels, one for each frequency. Using the duality of each of these independent channels, it is easy to establish that the capacity region of the BC is equal to the capacity region of the dual MAC with a sum power constraint. Furthermore, it is also straightforward to verify that the conditions of Theorem 2 hold, and thus the capacity region of the MAC is equal to an intersection of scaled BC capacity regions.[4]

V. Duality of Outage Capacity

In this section, we show that duality holds for the outage capacity of fading channels. The outage capacity region (denoted $\mathcal{C}_{\rm MAC}^{\rm out}(\overline{\boldsymbol{P}}, \boldsymbol{P}^{\rm out};\boldsymbol{H})$ and $\mathcal{C}_{\rm BC}^{\rm out}(\overline{P}, \boldsymbol{P}^{\rm out};\boldsymbol{H})$) is defined as the set of rates that can be maintained for user j for a fraction $1 - P_j^{\rm out}$ of the time, or in all but $P_j^{\rm out}$ of the fading states [4], [5]. 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] [4], [16] is a special case of outage capacity where a constant rate must be maintained in *all* fading states, or where $\boldsymbol{P}^{\rm out} = 0$.

By definition, a rate vector $\boldsymbol{R}$ is in $\mathcal{C}_{\rm MAC}^{\rm out}(\overline{\boldsymbol{P}}, \boldsymbol{P}^{\rm out};\boldsymbol{H})$ if and only if there exists a power policy $\boldsymbol{P}(\boldsymbol{h})$ satisfying the power constraints $\overline{\boldsymbol{P}}$ and a rate function $\boldsymbol{R}(\boldsymbol{h})$ such that

$$\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\rm MAC}(\boldsymbol{P}(\boldsymbol{h}), \boldsymbol{h})$$

for all $\boldsymbol{h}$ and

$$\Pr[\boldsymbol{R}_j(\boldsymbol{h}) \geq \boldsymbol{R}_j] \geq 1 - P_j^{\rm out}.$$

The BC outage capacity region is defined similarly, except that the power policy must only satisfy a sum power constraint and $\boldsymbol{R}(\boldsymbol{h})$ must be in $\mathcal{C}_{\rm BC}(\sum_{i=1}^K \boldsymbol{P}_i(\boldsymbol{h}), \boldsymbol{h})$ for all $\boldsymbol{h}$. By applying duality to each fading state, it is clear that every rate vector in the MAC outage capacity region is achievable in the dual BC, and *vice versa*. Thus, the outage capacity region of the BC is

[4]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 that is used to characterize the MAC in terms of the dual BC.

[5]Zero-outage capacity is referred to as *delay-limited capacity* in [16].

equal to the sum power constraint outage capacity region of the MAC, with the same outage vector $\boldsymbol{P}^{\text{out}}$

$$\mathcal{C}_{\text{BC}}^{\text{out}}(\overline{P}, \boldsymbol{P}^{\text{out}}; \boldsymbol{H}) = \bigcup_{\{\boldsymbol{P}:\mathbf{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{\text{MAC}}^{\text{out}}(\boldsymbol{P}, \boldsymbol{P}^{\text{out}}; \boldsymbol{H}). \quad (26)$$

Using Theorem 2, it also follows that the MAC outage capacity region is equal to the intersection of the scaled BC outage capacity regions

$$\mathcal{C}_{\text{MAC}}^{\text{out}}(\boldsymbol{P}, \boldsymbol{P}^{\text{out}}; \boldsymbol{H}) = \bigcap_{\boldsymbol{\alpha}>0} \mathcal{C}_{\text{BC}}^{\text{out}}\left(\mathbf{1}\cdot\frac{\boldsymbol{P}}{\boldsymbol{\alpha}}, \boldsymbol{P}^{\text{out}}; \boldsymbol{\alpha}\boldsymbol{H}\right). \quad (27)$$

Though the outage capacity region has been characterized for both the BC and MAC [4], [5], 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.

There is also a more stringent notion of outage capacity in which outages must be declared simultaneously for all users (referred to as common outage). In this situation, there is only one outage probability (for all users). It is also straightforward to show that duality extends to common outage as well.

VI. Duality of Minimum-Rate Capacity

Minimum-rate capacity was first introduced for the fading BC in [6]. In this section, we establish the duality of the minimum-rate capacity regions of the MAC and BC and use this duality to find the 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 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 $\boldsymbol{R}^* = (R_1^*, \dots, R_K^*)$ is maintained for all users in *all* fading states. From this definition, it follows that a rate vector $\boldsymbol{R}$ is in $\mathcal{C}_{\text{BC}}^{\min}(\overline{P}, \boldsymbol{R}^*; \boldsymbol{H})$ if and only if there exists $\boldsymbol{R}(\boldsymbol{h})$ such that $\boldsymbol{R} \leq \mathbb{E}_{\boldsymbol{H}}[\boldsymbol{R}(\boldsymbol{h})]$ with

$$\boldsymbol{R}(\boldsymbol{h}) \in \mathcal{C}_{\text{BC}}\left(\sum_{i=1}^{K} \boldsymbol{P}_i^B(\boldsymbol{h}); \boldsymbol{h}\right), \qquad \forall \boldsymbol{h}$$

and $\boldsymbol{R}(\boldsymbol{h}) \geq \boldsymbol{R}^*$ for all $\boldsymbol{h}$, for some $\boldsymbol{P}^B(\boldsymbol{h})$ satisfying the average sum power constraint. The MAC minimum-rate capacity region is defined analogously. With this definition, it is easy to see that the BC minimum-rate capacity region equals the sum power MAC minimum-rate capacity region

$$\mathcal{C}_{\text{BC}}^{\min}(\overline{P}, \boldsymbol{R}^*; \boldsymbol{H}) = \bigcup_{\{\boldsymbol{P}:\mathbf{1}\cdot\boldsymbol{P}=\overline{P}\}} \mathcal{C}_{\text{MAC}}^{\min}(\boldsymbol{P}, \boldsymbol{R}^*; \boldsymbol{H}). \quad (28)$$

Since every point on the boundary of the MAC minimum-rate capacity region is also on the boundary of the scaled BC minimum-rate capacity region, we can also find the optimal power policy and decoding order for the MAC. As always in the BC, users are decoded in order of increasing channel gain. In the case of a scaled channel, this corresponds to decoding in order of *increasing* $\frac{h_i}{\lambda_i^*}$. By duality, users in the MAC should be decoded in order of *decreasing* $\frac{h_i}{\lambda_i^*}$. The optimal MAC power allocation policy can be derived from the optimal power policy of the scaled BC.

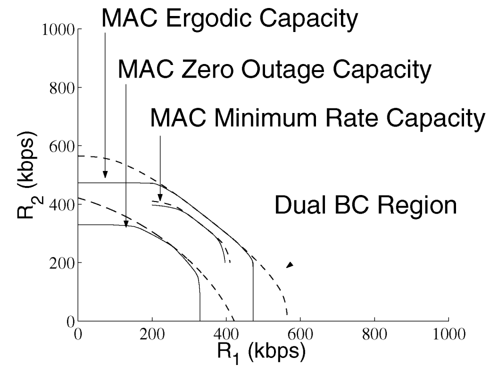


Fig. 8. MAC minimum-rate (200 kb/s minimum rate), ergodic, and zero-outage capacity regions.

The MAC minimum-rate capacity region for a discrete fading distribution channel is plotted in Fig. 8. In the figure, the MAC ergodic, zero-outage, and minimum-rate capacity region boundaries are all shown. The corresponding dual BC capacity region (ergodic, zero-outage, and minimum rate) boundaries are indicated with dotted lines. The minimum-rate capacity region is shown for symmetric minimum rates of 200 kb/s 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. Note that the MAC minimum-rate capacity region lies between the zero-outage and ergodic capacity regions as it does for the BC minimum-rate capacity region [6].

VII. Extensions

The focus of this paper has been to characterize the duality of the scalar Gaussian MACs and BCs. There are, however, many possible directions in which duality can 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 dealt exclusively with 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 MIMO Gaussian BC and MAC are duals [8]. The MIMO BC is a nondegraded BC, for which a general expression for the capacity region remains unknown. However, an achievable region for the MIMO BC based on dirty-paper coding [17], [18] is known. In [8], 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

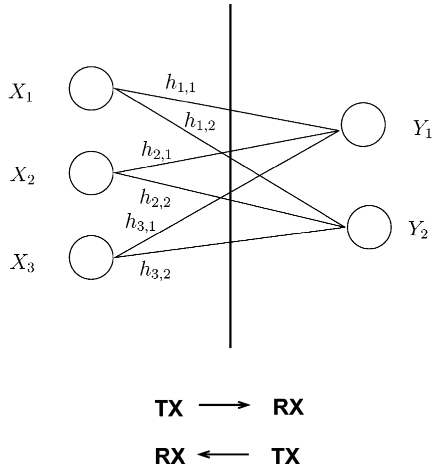


Fig. 9. Multiterminal Gaussian network.

intersection of the scaled MIMO BC achievable regions. This duality is established with matrix versions of the MAC-BC and BC-MAC power transformations given in Section III-C. Furthermore, this duality result is used in [8], [9] to show that dirty-paper coding achieves the sum-rate capacity of the multiple-antenna BC. 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.

B. Gaussian Multiterminal Networks

The Gaussian MAC and BC can be generalized to a model in which there are multiple transmitters *and* multiple receivers subject to additive Gaussian noise. Such networks are referred to as Gaussian multiterminal networks [1, Sec. 14.10]. In this subsection, we discuss networks in which nodes can either be transmitters or receivers, but not both simultaneously.

In Fig. 9, 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 BCs and MACs can be seen as a specialization of multiterminal 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 BC and if the nodes on the right transmit then the channel is a two-user MAC. 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, multiterminal setting we consider, any transmitter is allowed to communicate with any receiver. In the channel shown in Fig. 9, 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 multiterminal networks, or 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 multiterminal network is not known. Interestingly, Theorem 2 can easily be extended to multiple-receiver channels. This allows characterization of a multiple-transmitter/multiple-receiver rate region (with individual transmitter power constraints) in terms of the sum transmit power constraint capacity regions.

If the nodes on the left and right were considered to be multiple antennas of single users (i.e., a single transmitter with three antennas communicating to a receiver with two antennas, or *vice versa*), then duality holds due to the reciprocity of multiple-antenna Gaussian links [19]. 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 a broader duality may hold for general Gaussian networks, but this has yet to be confirmed.

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 BCs and MACs. Since the capacity region of the nondegraded BC is unknown, such a duality could perhaps be helpful in this respect. In [7], a setup for dual discrete memoryless broadcast and MACs is proposed. Furthermore, a duality is established between a limited set of deterministic (i.e., noiseless) BCs and MACs. This appears to be a promising avenue of research, but there is also indication that this setup does not allow for the possibility of duality for all BCs.

VIII. 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. This duality allows us to express the capacity region of the Gaussian BC in terms of the Gaussian MAC, and *vice versa*. We also showed that this duality extends to fading channels. Though this paper deals with scalar channels, duality has also been extended to MIMO channels and has been used to find the sum rate capacity of the MIMO BC. Furthermore, we conjecture that duality applies to general multiterminal Gaussian networks as well.

Duality provides an insightful connection between the Gaussian MAC and BC. This relationship is not only conceptually powerful, but has also been of great use in establishing new results for Gaussian channels. A number of other information-theoretic dualities (e.g., source/channel coding [20], channel coding/rate distortion [21], MAC/Slepian–Wolf [1, Sec. 14.5]) have been established over the decades. It remains to be seen if the multiple-access/broadcast duality is a result of the special structure of Gaussian channels or if it has 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

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

then $\sum_{i=1}^K P_i^M = \sum_{i=1}^K P_i^B$. For notational simplicity we assume $\pi(i) = i$ in this section. We do this by inductively showing that

$$\sum_{i=1}^{j} P_i^B = \frac{\sigma^2}{B_j} \sum_{i=1}^{j} P_i^M. \tag{30}$$

The base case $(j = 1)$ holds by definition

$$P_1^B = \frac{A_1 P_1^M}{B_1} = \frac{\sigma^2}{B_1} P_1^M.$$

Assume (30) holds for j. For $j+1$ we get

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

where (a) follows from the inductive hypothesis. By using (30) 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.

APPENDIX B
PROOF OF COROLLARY 14

The fact that $\mathcal{C}_{\mathrm{MAC}}(\boldsymbol{P}; \boldsymbol{h}) \subseteq \mathcal{C}_{\mathrm{BC}}(\boldsymbol{1} \cdot \boldsymbol{P}; \boldsymbol{h})$ follows trivially from Theorem 1. Also, the fact that the boundaries of the regions meet at the point where users are decoded in order of decreasing channel gains in the MAC follows from the MAC-BC transformations and the fact that the opposite decoding order is used in the BC. It only remains to show that the MAC and dual BC capacity region boundaries meet at only this point if the channel gains of all K users are distinct[6] and that all other corner points of the MAC capacity region lie strictly in the interior of the dual BC capacity region.

We show this by proving that every successive decoding point other than the one corresponding to decoding in order of decreasing channel gains lies strictly in the interior of the sum power constraint MAC capacity region (i.e., the dual BC capacity region). We 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. This implies that points on the boundary of the sum power MAC can only be achieved by successive decoding in order of decreasing channel gain. 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 there exist i and j such that $h_i < h_j$ but User i is decoded directly *before* User j. This is easily seen to be true if and only if decoding is not done in order of decreasing channel gains. We will show that the sum power needed to achieve any strictly positive rate vector is strictly less 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 $h_i P_i + h_j P_j$ to all users decoded before them. All users that 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

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

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

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

and the sum of their powers is

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

If User i is decoded directly after User j instead of before him, then the required sum power is

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

Clearly,

$$P_i + P_j - P_i' - P_j' = \left(\frac{\sigma^2 + I}{h_i} - \frac{\sigma^2 + I}{h_j}\right)(e^{2R_i} - 1)(e^{2R_j} - 1) > 0$$

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

since $h_i < h_j$ 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

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

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. Thus, the point lies in the interior of the sum power constraint MAC, so it is not on the boundary of the dual BC capacity region.

Appendix C
Verification of Rate Region Conditions

In this appendix, we show that the capacity region of the constant MAC $\mathcal{C}_{\mathrm{MAC}}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ meets the conditions specified in Theorem 2. 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 $\mathcal{C}_{\mathrm{MAC}}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ follows from the definition of the capacity region in (3).

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 $\boldsymbol{r} \in C_{\mathrm{MAC}}(\overline{\boldsymbol{P}})$ and $\boldsymbol{t} \in C_{\mathrm{MAC}}(\overline{\boldsymbol{Q}})$. We wish to show that

$$\theta \boldsymbol{r} + (1-\theta)\boldsymbol{t} \in C_{\mathrm{MAC}}(\theta \overline{\boldsymbol{P}} + (1-\theta)\overline{\boldsymbol{Q}}).$$

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

4) $\mathcal{C}_{\mathrm{MAC}}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ 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 each user transmits an independent message.

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 that can be decoded (and thus subtracted off) by all K receivers, even when treating the rest of the received signal as noise.

8) $\mathcal{C}_{\mathrm{MAC}}(\overline{\boldsymbol{P}}; \boldsymbol{h})$ 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.

Appendix D
Proof of Theorem 2

We wish to show that for any strictly positive[7] power constraint $\overline{\boldsymbol{P}} = (\overline{P}_1, \ldots, \overline{P}_K) > 0$

$$R(\overline{\boldsymbol{P}}) = \bigcap_{\boldsymbol{\alpha} > 0} R^{\mathrm{sum}}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}} \right) \tag{31}$$

where the sum power constraint capacity region is defined as

$$R^{\mathrm{sum}}_{\boldsymbol{\alpha}}(P_{\mathrm{sum}}) \triangleq \bigcup_{\{\boldsymbol{P} |\ \boldsymbol{P} \in \mathcal{R}^K_+,\ \mathbf{1}\cdot\boldsymbol{P} \le P_{\mathrm{sum}}\}} R_{\boldsymbol{\alpha}}(\boldsymbol{P}). \tag{32}$$

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) $S = \{(\boldsymbol{R}, \boldsymbol{P}) | \boldsymbol{P} \in \mathcal{R}^K_+,\ \boldsymbol{R} \in R(\boldsymbol{P})\}$ is a convex set.
3) For all $\boldsymbol{P} \in \mathcal{R}^K_+$, $R(\boldsymbol{P})$ is a closed, convex region.
4) If $\boldsymbol{P_1} \ge \boldsymbol{P_2}$, then $R(\boldsymbol{P_1}) \supseteq R(\boldsymbol{P_2})$.
5) If $(R_1, \ldots R_K) \in R(P_1, P_2, \ldots, P_K)$, then for any i

$$(R_1, \ldots, R_{i-1}, 0, R_{i+1}, \ldots, R_K) \in R(P_1, \ldots, P_{i-1}, 0, P_{i+1}, \ldots, P_K).$$

6) If $\boldsymbol{R} \in R(\boldsymbol{P})$ and $\boldsymbol{R}' \le \boldsymbol{R}$, then $\boldsymbol{R}' \in R(\boldsymbol{P})$.
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 $\boldsymbol{P} > 0$.

From Condition 1 and the definition of the sum power constraint capacity region (32), it is clear that

$$R(\overline{\boldsymbol{P}}) = R_{\boldsymbol{\alpha}} \left(\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}} \right) \subseteq R^{\mathrm{sum}}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}} \right), \qquad \forall \boldsymbol{\alpha} > 0.$$

This implies that $R(\overline{\boldsymbol{P}}) \subseteq \bigcap_{\boldsymbol{\alpha}>0} R^{\mathrm{sum}}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}} \right)$. 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 [22, p. 135]:

$$\max_{\boldsymbol{R}} \boldsymbol{\mu} \cdot \boldsymbol{R} \quad s.t. \ \ \boldsymbol{R} \in R(\overline{\boldsymbol{P}}) \tag{33}$$

over priorities $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$ such that $\boldsymbol{\mu} \ge 0$ and $\mathbf{1}\cdot\boldsymbol{\mu} = 1$. Since at least one component of $\boldsymbol{\mu}$ must be strictly positive for $\mathbf{1}\cdot\boldsymbol{\mu} = 1$ to hold, without loss of generality we assume $\mu_K > 0$. For every $\boldsymbol{\mu} \ge 0$, we show[8]

$$\max_{\boldsymbol{R} \in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu} \cdot \boldsymbol{R} \ge \sup_{\boldsymbol{R} \in \bigcap_{\boldsymbol{\alpha}>0} R^{\mathrm{sum}}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}} \right)} \boldsymbol{\mu} \cdot \boldsymbol{R}. \tag{34}$$

This implies $R(\overline{\boldsymbol{P}}) \supseteq \bigcap_{\boldsymbol{\alpha}>0} R^{\mathrm{sum}}_{\boldsymbol{\alpha}} \left(\mathbf{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}} \right)$ because $R(\overline{\boldsymbol{P}})$ is completely characterized by $\max \boldsymbol{\mu} \cdot R$(33). We essentially show that for every $\boldsymbol{\mu}$ (or roughly every point on the boundary of

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

[8] 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.

$R(\overline{\boldsymbol{P}})$) there exists an $\boldsymbol{\alpha}$ such that the boundaries of $R_{\boldsymbol{\alpha}}^{\text{sum}}(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})$ and $R(\overline{\boldsymbol{P}})$ meet at the point where $\boldsymbol{\mu}\cdot\boldsymbol{R}$ is maximized.

The optimization in (33) is equivalent to

$$\max_{(\boldsymbol{R},\boldsymbol{P})\in S} \boldsymbol{\mu}\cdot\boldsymbol{R} \text{ s.t. } \boldsymbol{P}\leq\overline{\boldsymbol{P}} \tag{35}$$

where the set S is defined in Condition 2. Consider the above maximization for some fixed $\boldsymbol{\mu}$. Since the objective function is linear and the set S is convex, this is a convex optimization problem (see [23] for a general reference on convex optimization and Lagrangian duality). Furthermore, the maximization takes on some optimal value p^* by the feasibility of the constraint set. The optimal value is finite due to the assumption that $R(\overline{\boldsymbol{P}})$ is finite. The Lagrangian is formed by adding the weighted sum of the constraints to the objective function

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

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

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

By the above definition, for any $(\boldsymbol{R},\boldsymbol{P})\in S$ satisfying $\boldsymbol{P}\leq\overline{\boldsymbol{P}}$ and $\boldsymbol{\lambda}\geq 0$, we have $\boldsymbol{\mu}\cdot\boldsymbol{R}\leq L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda})\leq g(\boldsymbol{\lambda})$. This implies $p^*\leq g(\boldsymbol{\lambda})$ for any $\boldsymbol{\lambda}\geq 0$. 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 $\boldsymbol{\lambda}$.

By minimizing the dual function over all nonnegative 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 [23], [22]

$$d^* \triangleq \min_{\boldsymbol{\lambda}\geq 0} g(\boldsymbol{\lambda}) \triangleq g(\boldsymbol{\lambda^*}) = p^* \tag{36}$$

where $\boldsymbol{\lambda^*}=(\lambda_1^*,\ldots,\lambda_K^*)$ are the optimum Lagrange multipliers that lead to d^*. In what follows, we show that λ_i^* is finite and strictly positive if $\mu_i>0$ and λ_i^* is zero if $\mu_i=0$.

First consider i such that $\mu_i>0$. If $\lambda_i=\infty$, then with $\boldsymbol{R}=\boldsymbol{P}=0$ we get $L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda})\geq\lambda_i\overline{P}_i$. Thus, $g(\boldsymbol{\lambda})=\infty$, which implies that λ_i^* must be finite. Now assume $\lambda_i=0$. Choose $P_j=\overline{P}_j$ for all $j\neq i$ and let P_i be arbitrarily large. Additionally, choose all rates to be zero except for R_i. For this choice of $(\boldsymbol{R},\boldsymbol{P})$, we have $L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda})=\mu_iR_i$. Due to the unbounded condition on $R(\boldsymbol{P})$, letting P_i be arbitrarily large implies R_i can be made arbitrarily large while still maintaining $(\boldsymbol{R},\boldsymbol{P})\in S$. This, in turn, implies $g(\boldsymbol{\lambda})=\infty$. Since $p^*=g(\boldsymbol{\lambda^*})$ is finite, we must have $\lambda_i^*\neq 0$.

We now show that $\mu_i=0$ implies $\lambda_i^*=0$. Since $g(\boldsymbol{\lambda})$ is the supremum of the Lagrangian and because $\mu_i=0$, it follows from Condition 5 that $R_i=P_i=0$ to achieve $g(\boldsymbol{\lambda})$ if $\lambda_i>0$. Thus, for any $\boldsymbol{\lambda}$ with $\lambda_i>0$ and any $(\boldsymbol{R},\boldsymbol{P})$ with $R_i=P_i=0$, we have

$$\begin{aligned}L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}) &= \boldsymbol{\mu}\cdot\boldsymbol{R} - \sum_{j\neq i}\lambda_jP_j + \sum_j\lambda_j\overline{P}_j\\ &> \boldsymbol{\mu}\cdot\boldsymbol{R} - \sum_{j\neq i}\lambda_jP_j + \sum_{j\neq i}\lambda_j\overline{P}_j\\ &= L(\boldsymbol{R},\boldsymbol{P},\boldsymbol{\lambda}')\end{aligned}$$

where $\boldsymbol{\lambda}'=\boldsymbol{\lambda}$ except that $\lambda_i'=0$. Thus, $g(\boldsymbol{\lambda})>g(\boldsymbol{\lambda}')$, which implies that $\lambda_i^*=0$.

Now consider the *scaled* MAC with α_i defined as

$$\alpha_i = \begin{cases}\frac{\lambda_K^*}{\lambda_i^*}, & \text{if } \lambda_i^*>0\\ c, & \text{if } \lambda_i^*=0\end{cases} \tag{37}$$

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:

$$\sup_{\boldsymbol{R}} \boldsymbol{\mu}\cdot\boldsymbol{R} \text{ s.t. } \boldsymbol{R}\in R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right). \tag{38}$$

By the definition of the sum power constraint capacity region, $\boldsymbol{R}\in R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)$ 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 maximization (38) can thus be rewritten as

$$\sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R_{\boldsymbol{\alpha}}(\boldsymbol{P})} \boldsymbol{\mu}\cdot\boldsymbol{R} \text{ s.t. } \mathbf{1}\cdot\boldsymbol{P}\leq\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}. \tag{39}$$

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

$$\begin{aligned}H_{\boldsymbol{\alpha}}(\boldsymbol{R},\boldsymbol{P},\nu) &= \boldsymbol{\mu}\cdot\boldsymbol{R} - \nu\left(\mathbf{1}\cdot\boldsymbol{P}-\mathbf{1}\cdot\frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)\\ &= \boldsymbol{\mu}\cdot\boldsymbol{R} - \nu\left(P_1-\frac{\overline{P}_1}{\alpha_1}\right) - \cdots\\ &\quad - \nu\left(P_K-\frac{\overline{P}_K}{\alpha_K}\right)\\ &= \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}\left(\alpha_1P_1-\overline{P}_1\right) - \cdots\\ &\quad - \frac{\nu}{\alpha_K}\left(\alpha_KP_K-\overline{P}_K\right)\end{aligned}$$

and the corresponding Lagrangian dual function is

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

Again, the dual function satisfies $f_{\boldsymbol{\alpha}}(\nu)\geq d_{\boldsymbol{\alpha}}^*$ for all $\nu\geq 0$. Due to the fact that $R_{\boldsymbol{\alpha}}(\boldsymbol{P})=R(\boldsymbol{\alpha}\boldsymbol{P})$ and $\boldsymbol{\alpha}>0$, we can simplify the dual function as

$$\begin{aligned}f_{\boldsymbol{\alpha}}(\nu) &= \sup_{\boldsymbol{\alpha}\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R(\boldsymbol{\alpha}\boldsymbol{P})} \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}(\alpha_1P_1-\overline{P}_1) - \cdots\\ &\quad - \frac{\nu}{\alpha_K}(\alpha_KP_K-\overline{P}_K)\\ &= \sup_{\boldsymbol{P}\in R_K^+,\ \boldsymbol{R}\in R(\boldsymbol{P})} \boldsymbol{\mu}\cdot\boldsymbol{R} - \frac{\nu}{\alpha_1}(P_1-\overline{P}_1) - \cdots\\ &\quad - \frac{\nu}{\alpha_K}(P_K-\overline{P}_K)\\ &= g\left(\frac{\nu}{\boldsymbol{\alpha}}\right)\end{aligned}$$

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

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

$$f_{\boldsymbol{\alpha}}(\lambda_K^*) = g\left(\frac{\lambda_K^*}{\boldsymbol{\alpha}}\right) = g\left(\frac{\lambda_K^*}{\alpha_1}, \ldots, \frac{\lambda_K^*}{\alpha_K}\right).$$

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

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

$$f_{\boldsymbol{\alpha}}(\lambda_K^*) = g\left(\frac{\lambda_K^*}{\alpha_1}, \ldots, \frac{\lambda_K^*}{\alpha_K}\right) = g(\lambda_1^*, \ldots, \lambda_K^*) = p^*.$$

Since $f_{\boldsymbol{\alpha}}(\lambda_K^*)$ is an upper bound to $d_{\boldsymbol{\alpha}}^*$, we have $d_{\boldsymbol{\alpha}}^* \le p^*$. This implies

$$\max_{\boldsymbol{R} \in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu} \cdot \boldsymbol{R} \ge \sup_{\boldsymbol{R} \in R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu} \cdot \boldsymbol{R} \ge \sup_{\boldsymbol{R} \in \bigcap_{\boldsymbol{\alpha} > 0} R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu} \cdot \boldsymbol{R}$$

where the second inequality follows from

$$R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}\right) \supseteq \bigcap_{\boldsymbol{\alpha} > 0} R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right).$$

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 $i = 1, \ldots, L$ and $\mu_i > 0$ for $i = L+1, \ldots, K$. We established earlier that $\mu_i = 0$ implies $\lambda_i^* = 0$. Therefore, we have

$$\begin{aligned} p^* &= g\left(\lambda_1^*, \ldots, \lambda_L^*, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) \\ &= g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right). \end{aligned}$$

Using the fact that $f_{\boldsymbol{\alpha}}(\nu) = g\left(\frac{\nu}{\boldsymbol{\alpha}}\right) \ \forall \boldsymbol{\alpha} > 0$, we have

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

Here $f_{\boldsymbol{\alpha}}(\lambda_K^*)$ is a function of the constant c because $\boldsymbol{\alpha}$ depends on c, as defined in (37). For a fixed $\boldsymbol{\alpha}$ (i.e., a fixed c), the optimum value of the sum power constraint region satisfies $d_{\boldsymbol{\alpha}}^* \le f_{\boldsymbol{\alpha}}(\lambda_K^*)$.

We now show the desired result by contradiction. Assume

$$\sup_{\boldsymbol{R} \in \bigcap_{\boldsymbol{\alpha} > 0} R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu} \cdot \boldsymbol{R} > \max_{\boldsymbol{R} \in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu} \cdot \boldsymbol{R}. \tag{40}$$

Since $R_{\boldsymbol{\alpha}}^{\text{sum}}(\boldsymbol{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}}) \supseteq \bigcap_{\boldsymbol{\alpha} > 0} R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)$, $\forall \boldsymbol{\alpha} > 0$, this implies that for some $\epsilon > 0$

$$\left(\sup_{\boldsymbol{R} \in R_{\boldsymbol{\alpha}}^{\text{sum}}(\boldsymbol{1} \cdot \frac{\overline{\boldsymbol{P}}}{\boldsymbol{\alpha}})} \boldsymbol{\mu} \cdot \boldsymbol{R}\right) \ge \left(\max_{\boldsymbol{R} \in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu} \cdot \boldsymbol{R}\right) + \epsilon$$

for all $\boldsymbol{\alpha} > 0$. This implies that for all $\boldsymbol{\alpha} > 0$

$$d_{\boldsymbol{\alpha}}^* \ge p^* + \epsilon = g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) + \epsilon.$$

However, earlier we established that

$$d_{\boldsymbol{\alpha}}^* \le f_{\boldsymbol{\alpha}}(\lambda_K^*) = g\left(\frac{\lambda_K^*}{c}, \ldots, \frac{\lambda_K^*}{c}, \lambda_{L+1}^*, \ldots, \lambda_K^*\right)$$

for all $c > 0$. Thus, we have that

$$g\left(\frac{\lambda_K^*}{c}, \ldots, \frac{\lambda_K^*}{c}, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) \ge g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) + \epsilon$$

for all c. Since g is a convex function, $g\left(\beta, \ldots, \beta, \lambda_{L+1}^*, \ldots, \lambda_K^*\right)$ must lie *beneath* the line between

$$g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) \text{ and } g\left(1, \ldots, 1, \lambda_{L+1}^*, \ldots, \lambda_K^*\right)$$

(which is finite) for $0 \le \beta \le 1$. This contradicts

$$g\left(\frac{\lambda_K^*}{c}, \ldots, \frac{\lambda_K^*}{c}, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) \ge g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right) + \epsilon, \quad \forall c \ge \lambda_K^*.$$

In other words, the convexity of g implies that as c becomes large, the value of $g\left(\frac{\lambda_K^*}{c}, \ldots, \frac{\lambda_K^*}{c}, \lambda_{L+1}^*, \ldots, \lambda_K^*\right)$ must become arbitrarily close to $g\left(0, \ldots, 0, \lambda_{L+1}^*, \ldots, \lambda_K^*\right)$. Thus, (40) must be false and therefore,

$$\max_{\boldsymbol{R} \in R(\overline{\boldsymbol{P}})} \boldsymbol{\mu} \cdot \boldsymbol{R} \ge \sup_{\boldsymbol{R} \in \bigcap_{\boldsymbol{\alpha} > 0} R_{\boldsymbol{\alpha}}^{\text{sum}}\left(\boldsymbol{1} \cdot \frac{\boldsymbol{P}}{\boldsymbol{\alpha}}\right)} \boldsymbol{\mu} \cdot \boldsymbol{R} \tag{41}$$

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

We have now shown that the relationship (41) holds for *all* $\boldsymbol{\mu} \ge 0$ and the proof is complete.

REFERENCES

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

[2] 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, Mar. 2001.

[3] D. N. C. 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, Nov. 1998.

[4] 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, Mar. 2001.

[5] L. Li, N. Jindal, and A. Goldsmith, "Outage capacities and optimal power allocation for fading multiple-access channels," *IEEE Trans. Inform. Theory*, submitted for publication.

[6] N. Jindal and A. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2895–2909, Nov. 2003.

[7] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality between general multiple-access/broadcast channels," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, July 2003, p. 313.

[8] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2658–2668, Oct. 2003.

[9] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink–downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[10] H. Viswanathan, S. Venkatesan, and H. C. Huang, "Downlink capacity evaluation of cellular networks with known interference cancellation," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 802–811, June 2003.

[11] N. Jindal, S. Jafar, S. Vishwanath, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," in *Proc. Asilomar Conf. Signals, Systems, and Computing*, Pacific Grove, CA, Nov. 2002, pp. 1518–1522.

[12] W. Yu and J. Cioffi, "Trellis preceding for the broadcast channel," in *Proc. Global Communications Conf.*, Oct. 2001, pp. 1344–1348.

[13] S. Vishwanath, S. Jafar, and A. Goldsmith, "Optimum power and rate allocation strategies for multiple access fading channels," in *Proc. Vehicular Technology Conf.*, May 2000, pp. 2888–2892.

[14] R. Cheng and S. Verdú, "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] S. Hanly and D. N. C. Tse, "Multiaccess fading channels–Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2816–2831, Nov. 1998.
[17] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 439–441, May 1983.
[18] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Proc. 38th Annu. Allerton Conf. Communication, Control and Computing*, Oct. 2000, pp. 1188–1193.
[19] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
[20] S. Pradhan and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1181–1203, May 2003.
[21] T. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1629–1638, June 2002.
[22] D. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
[23] S. Boyd and L. Vandenberghe, *Introduction to Convex Optimization with Engineering Applications*, 2001, Course Reader.