

Duality, Achievable Rates and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels

Sriram Vishwanath, Nihar Jindal, and Andrea Goldsmith

Department of Electrical Engineering, Stanford University, Stanford, CA 94305

E-mail : *sriram, njindal, andrea@systems.stanford.edu*

June 24, 2003

Keywords—Broadcast channel, Channel Capacity, Dirty Paper Coding, Duality, MIMO systems

Abstract— We consider a multi-user multiple input multiple output (MIMO) Gaussian broadcast channel (BC), where the transmitter and receivers have multiple antennas. Since the MIMO broadcast channel is in general a non-degraded broadcast channel, its capacity region remains an unsolved problem. In this paper, we establish a duality between what is termed the “dirty paper” achievable region (the Caire-Shamai achievable region) for the MIMO broadcast channel and the capacity region of the MIMO multiple-access channel (MAC), which is easy to compute. Using this duality, we greatly reduce the computational complexity required for obtaining the dirty paper achievable region for the MIMO BC. We also show that the dirty paper achievable region achieves the sum-rate capacity of the MIMO BC by establishing that the maximum sum rate of this region equals an upper-bound on the sum rate of the MIMO BC.

I. INTRODUCTION

Multiple input multiple output (MIMO) systems have received a great deal of attention as a method to achieve very high data rates over wireless links. The capacity of single-user MIMO Gaussian channels was first studied by Foschini [1] and Telatar [2]. This work has also been extended to the MIMO multiple-access channel (MAC) [2–4]. The capacity of MIMO broadcast channels (BC), however, is an open problem due to the lack of a general theory on non-degraded broadcast channels. In pioneering work by Caire and Shamai [5], a set of achievable rates (the achievable region) for the MIMO broadcast channel was obtained by applying the “dirty paper” result [6] at the transmitter (or alternatively coding for non-causally known interference). It was also shown in [5, 7] that the sum rate MIMO BC capacity equals the maximum sum rate of this achievable region for the two user broadcast channel with an arbitrary number of transmit antennas ($t \geq 1$) and one receive antenna at each receiver ($r_1 = r_2 = 1$). However, computing this region is extremely complex and the approach used in [5, 7] to prove the optimality of dirty-paper coding for sum rate does not appear to work for the more general class of channels (i.e. an arbitrary number of users and receive antennas) which we consider.

In this paper, we consider a K -user MIMO Gaussian BC in which receiver j has $r_j \geq 1$ receive antennas and the transmitter has $t \geq 1$ transmit antennas. The achievable region for a general MIMO BC requires an extension of the Caire-Shamai region to multiple users and multiple receive antennas, which was done by Yu and Cioffi in [8]. We refer to this extension as the dirty paper region. We establish a duality between the dirty paper region of the MIMO BC and the capacity region of the MIMO MAC. In other words, we show that the dirty paper region is

exactly equal to the capacity region of the dual MIMO MAC, with the K transmitters having the same *sum* power constraint as the MIMO BC. We establish this duality by showing that all rates achievable in the dual MIMO MAC with power constraints whose sum equals the BC power constraint are also achievable in the MIMO BC, and vice versa. This duality is the multiple-antenna extension of the previously established duality between the scalar Gaussian BC and MAC [9]. Though we consider only the constant channel case, this duality can easily be shown to hold for fading multiple-antenna Gaussian BC’s and MAC’s, as it does in the scalar channel case. This duality has two important applications: it permits us to compute with ease the dirty paper region of the MIMO BC, which is very difficult to compute directly, and it allows us to show that the dirty paper achievable region achieves the sum rate capacity of the MIMO BC.

Finding the full capacity region of the MIMO BC is very difficult due to its non-degraded nature, but we are able to show that the dirty paper region achieves the sum rate capacity of the MIMO BC through the use of the Sato upper bound on the sum-rate capacity of broadcast channels [10]. The Sato upper bound was first applied to the MIMO BC by Caire and Shamai [5] to find the sum rate capacity of the ($t \geq 1, r_1 = r_2 = 1$) channel. Using Sato’s technique, we bound the sum rate capacity of the MIMO BC by considering the capacity when the K receivers perform joint signal detection (i.e. we consider a single-user $t \times (\sum_{j=1}^K r_j)$ antenna channel) when the noise at every antenna is correlated with the noise at every other antenna except for those at the same receiver, and we analytically show, employing a proof which generally constructs the worst-case noise covariance matrix, that this upper bound coincides exactly with the maximum achievable sum rate in the dirty paper region.

There has been parallel work on the optimality of the dirty paper region for sum capacity of the BC. In [11], the authors handle those channels for which the worst-case correlation of the noise in the Sato upper bound is non-singular, and in [12], the authors handle channels with multiple transmit antennas but only one antenna at each receiver. In this paper we deal with the most general case, with no restrictions on the number of antennas or on the worst-case noise.

Although the optimality of the dirty paper region has only been shown for sum rate (and trivially for the corner points of the region), the fact that the dirty paper region is equal to the dual MIMO MAC capacity region together with the fact that the scalar Gaussian BC capacity region is equal to the dual MAC capacity region leads us to believe that the dirty paper region may actually be the capacity region of the MIMO BC. Signifi-

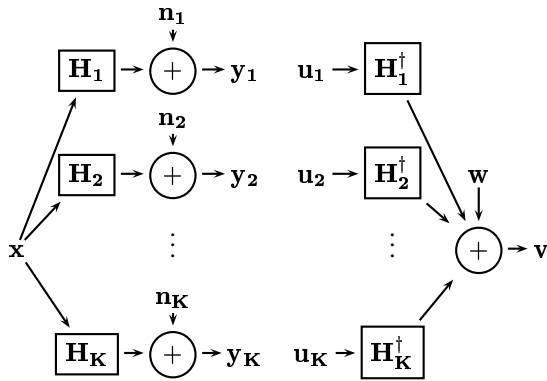


Fig. 1. System models of the MIMO BC(left) and the MIMO MAC (right) channels

cant progress towards proving this conjecture has recently been made [13] [14], but this hypothesis remains unproven.

The remainder of this paper is organized as follows. In Section II we describe the MIMO BC and the dual MIMO MAC. In Section III we summarize some background information, including the achievable “dirty paper” BC region, the MIMO MAC capacity region, and the duality of the scalar MAC and BC. We describe the MIMO MAC-BC duality result in Section IV. In Section V, we show that the dirty paper region achieves sum rate capacity of the MIMO BC and we provide a few illustrative examples in Section VI. We conclude with Section VII.

II. SYSTEM MODEL

We use boldface to denote matrices and vectors. $|\mathbf{S}|$ denotes the determinant and \mathbf{S}^{-1} the inverse of a square matrix \mathbf{S} . For any general matrix \mathbf{M} , \mathbf{M}^\dagger denotes the conjugate transpose and $\text{Tr}(\mathbf{M})$ denotes the trace. \mathbf{I} denotes the identity matrix and $\text{diag}(\lambda_i)$ denotes a diagonal matrix with the (i, i) entry equal to λ_i .

We consider a MIMO broadcast channel with a t -antenna transmitter and K receivers with r_1, \dots, r_K receive antennas, respectively. The transmitter sends independent information to each receiver. The broadcast channel is the system on the left in Fig. 1.

Let $\mathbf{x} \in \mathbb{C}^{t \times 1}$ be the transmitted vector signal and let $\mathbf{H}_k \in \mathbb{C}^{r_k \times t}$ be the channel matrix of receiver k where $\mathbf{H}_k(i, j)$ represents the channel gain from transmit antenna j to antenna i of receiver k . The white Gaussian noise at receiver k is represented by $\mathbf{n}_k \in \mathbb{C}^{r_k \times 1}$ where $\mathbf{n}_k \sim N(0, \mathbf{I})$. Let $\mathbf{y}_k \in \mathbb{C}^{r_k \times 1}$ be the received signal at receiver k . The received signal is mathematically represented as

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} = \mathbf{H}\mathbf{x} + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_K \end{bmatrix} \quad \text{where } \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_K \end{bmatrix}. \quad (1)$$

The matrix \mathbf{H} represents the channel gains of all receivers. The covariance matrix of the input signal is $\Sigma_x \triangleq \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. The transmitter is subject to an average power constraint P , which implies $\text{Tr}(\Sigma_x) \leq P$. We assume the channel matrix \mathbf{H} is constant and is known perfectly at the transmitter and at all receivers.

Now consider the dual multiple-access channel shown in the right half of Fig. 1. The dual channel is arrived at by converting the receivers in the BC into transmitters in the MAC and converting the t -antenna transmitter into a t -antenna receiver. Notice that the channel gains of the dual MAC are the same as that of the broadcast channel, i.e. $\mathbf{H}_k(i, j)$ corresponds to the gain from transmit antenna j to antenna i of receiver k in the BC and to the gain from antenna i of transmitter k to receive antenna j in the MAC.

Let $\mathbf{u}_k \in \mathbb{C}^{r_k \times 1}$ be the transmitted signal of transmitter k . Let $\mathbf{v} \in \mathbb{C}^{t \times 1}$ be the received signal and $\mathbf{w} \in \mathbb{C}^{t \times 1}$ the noise vector where $\mathbf{w} \sim N(0, \mathbf{I})$. The received signal is mathematically represented as

$$\begin{aligned} \mathbf{v} &= \mathbf{H}_1^\dagger \mathbf{u}_1 + \dots + \mathbf{H}_K^\dagger \mathbf{u}_K + \mathbf{w} \\ &= \mathbf{H}^\dagger \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_K \end{bmatrix} + \mathbf{w} \quad \text{where } \mathbf{H}^\dagger = [\mathbf{H}_1^\dagger \dots \mathbf{H}_K^\dagger]. \end{aligned}$$

In the dual MAC, each transmitter is subject to an individual power constraint of P_1, \dots, P_K , with $\sum_{i=1}^K P_i = P$ (i.e. the sum of the MAC power constraints equals the BC power constraint). We also assume perfect knowledge of the channel at the transmitters and the receiver in the dual MAC.

Lastly, we define the *cooperative system* to be the same as the broadcast channel, but with all receivers coordinating to perform joint detection. If the receivers are allowed to cooperate, the broadcast channel reduces to a single-user $t \times (\sum_{j=1}^K r_j)$ multiple-antenna system described by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \quad (2)$$

where $\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_K \end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_K \end{bmatrix}$. We call the capacity of this system the *cooperative capacity*.

We use $\mathcal{C}_{\text{BC}}(P, \mathbf{H})$, $\mathcal{C}_{\text{MAC}}(P_1, \dots, P_K, \mathbf{H}^\dagger)$ and $\mathcal{C}_{\text{coop}}(P, \mathbf{H})$ to denote the capacity regions of the MIMO BC, MIMO MAC, and cooperative system, respectively.

III. BACKGROUND

To obtain our results, we use the achievable region of the MIMO BC channel obtained in [5, 8] and results on the MIMO MAC capacity region [2–4] extensively in this paper. Hence, we first summarize these results and then state results on the duality of the scalar Gaussian BC and MAC [9].

A. Achievable BC Region - The Dirty Paper Region

An achievable region for the MIMO BC was first obtained in [5]. In [8], the region was extended to the more general multiple-user, multiple-antenna case using the following extension of dirty paper coding [6] to the vector case:

Lemma 1: [Yu, Cioffi] Consider a channel with $\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{s}_k + \mathbf{n}_k$, where \mathbf{y}_k is the received vector, \mathbf{x}_k the transmitted vector, \mathbf{s}_k the vector Gaussian interference, and \mathbf{n}_k the vector white Gaussian noise. If \mathbf{s}_k and \mathbf{n}_k are independent and non-causal knowledge of \mathbf{s}_k is available at the transmitter but

not at the receiver, then the capacity of the channel is the same as if \mathbf{s}_k is not present.

In the MIMO BC, this result can be applied at the transmitter when choosing codewords for different receivers. The transmitter first picks a codeword for receiver 1. The transmitter then chooses a codeword for receiver 2 with full (non-causal) knowledge of the codeword intended for receiver 1. Therefore receiver 2 does not see the codeword intended for receiver 1 as interference. Similarly, the codeword for receiver 3 is chosen such that receiver 3 does not see the signals intended for receivers 1 and 2 as interference. This process continues for all K receivers. Receiver 1 subsequently sees the signals intended for all other users as interference, Receiver 2 sees the signals intended for Users 3 through K as interference, etc. Since the ordering of the users clearly matters in such a procedure, the following is an achievable rate vector:

$$R_{\pi(i)} = \log \frac{|\mathbf{I} + \mathbf{H}_{\pi(i)}(\sum_{j \geq i} \boldsymbol{\Sigma}_{\pi(j)})\mathbf{H}_{\pi(i)}^\dagger|}{|\mathbf{I} + \mathbf{H}_{\pi(i)}(\sum_{j > i} \boldsymbol{\Sigma}_{\pi(j)})\mathbf{H}_{\pi(i)}^\dagger|} \quad i = 1, \dots, K. \quad (3)$$

The dirty-paper region $\mathcal{C}_{\text{DPC}}(P, \mathbf{H})$ is defined as the convex hull of the union of all such rates vectors over all positive semi-definite covariance matrices $\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K$ such that $\text{Tr}(\boldsymbol{\Sigma}_1 + \dots + \boldsymbol{\Sigma}_K) = \text{Tr}(\boldsymbol{\Sigma}_x) \leq P$ and over all permutations $(\pi(1), \dots, \pi(K))$:

$$\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \triangleq \text{Co} \left(\bigcup_{\pi, \boldsymbol{\Sigma}_i} \mathbf{R}(\pi, \boldsymbol{\Sigma}_i) \right) \quad (4)$$

where $\mathbf{R}(\pi, \boldsymbol{\Sigma}_i)$ is given by (3). The transmitted signal is $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_K$ and the input covariance matrices are of the form $\boldsymbol{\Sigma}_i = \mathbb{E}[\mathbf{x}_i \mathbf{x}_i^\dagger]$. The dirty paper-coding procedure yields statistically independent signals $\mathbf{x}_1, \dots, \mathbf{x}_K$, from which it follows $\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_1 + \dots + \boldsymbol{\Sigma}_K$.

One important feature to notice about the dirty paper rate equations in (3) is that the rate equations are neither a concave nor convex function of the covariance matrices. This makes finding the dirty paper region very difficult, because generally the entire space of covariance matrices which meet the power constraint must be searched over. In this paper we consider the dirty paper region subject to a transmit power constraint. Recent work [15] has characterized the dirty paper region subject to individual rate constraints (i.e. minimizing the transmit power required to achieve a certain set of rates).

B. MIMO MAC Capacity Region

The capacity region of a general MIMO MAC was obtained in [2–4] We now describe this capacity region for the dual MIMO MAC as defined in Section II. For any set of powers (P_1, \dots, P_K) , the capacity of the MIMO MAC is

$$\mathcal{C}_{\text{MAC}}(P_1, \dots, P_K; \mathbf{H}^\dagger) \triangleq \bigcup_{\{\text{Tr}(\mathbf{P}_i) \leq P_i \quad \forall i\}} \left\{ (R_1, \dots, R_K) : \right. \quad (5) \\ \left. \sum_{i \in S} R_i \leq \log |\mathbf{I} + \sum_{i \in S} \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i| \quad \forall S \subseteq \{1, \dots, M\} \right\}.$$

For $P > 0$, we denote by $\mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$ the following set

$$\mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger) \triangleq \bigcup_{\sum_{i=0}^K P_i \leq P} \mathcal{C}_{\text{MAC}}(P_1, \dots, P_K; \mathbf{H}^\dagger) \\ = \bigcup_{\{\sum_{i=1}^K \text{Tr}(\mathbf{P}_i) \leq P\}} \left\{ (R_1, \dots, R_K) : \right. \\ \left. \sum_{i \in S} R_i \leq \log |\mathbf{I} + \sum_{i \in S} \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i| \quad \forall S \subseteq \{1, \dots, M\} \right\} \quad (6)$$

By the argument provided in [4, Theorem 1], this region is convex. It can easily be shown that this region is the capacity region of a MAC when the transmitters have a sum power constraint instead of individual power constraints but are not allowed to cooperate. Additionally, the MIMO MAC rates can be shown to be a concave function of the covariance matrices. This implies that the boundary points (and the corresponding covariance matrices) of the sum power MIMO MAC capacity region can be found by a standard convex program (see [4] for a discussion of this with regards to the individual power constraint MIMO MAC, which is nearly identical in structure). This fact turns out to be quite important because later we show that the MIMO MAC sum power constraint region is equal to the dirty paper achievable region of the dual MIMO BC.

C. Duality of the Scalar Gaussian MAC and BC

Lastly, we state the duality result for scalar Gaussian MAC and BC channels [9].

Theorem 1 (Jindal, Vishwanath, Goldsmith) The capacity region of a scalar Gaussian BC with power P and channels $\bar{\mathbf{h}} = (h_1, \dots, h_K)$ is equal to the union of capacity regions of the dual MAC with powers (P_1, \dots, P_K) such that $\sum_{i=1}^K P_i = P$:

$$\mathcal{C}_{\text{BC}}(P; \bar{\mathbf{h}}) = \bigcup_{\sum_{i=1}^K P_i = P} \mathcal{C}_{\text{MAC}}(P_1, \dots, P_K; \bar{\mathbf{h}}). \quad (7)$$

The proof of this is obtained by showing that any set of rates achievable in the BC is also achievable in the MAC, and vice versa. One key point is that to achieve the same rate vector in the BC and MAC, the decoding order must in general be reversed, i.e if User 1 is decoded last in the BC then User 1 is decoded first in the MAC. In the next section, we will derive a similar result that equates the dirty paper BC achievable region with the union of MAC capacity regions for the MIMO channel we are considering.

IV. DUALITY OF THE MAC AND DIRTY PAPER BC REGION

In this section we show that the capacity region of the MIMO MAC with a sum power constraint of P for the K transmitters is the same as the dirty paper region of the dual MIMO BC with power constraint P . In other words, any rate vector that is achievable in the dual MAC with power constraints (P_1, \dots, P_K) is in the dirty paper region of the BC with power constraint $\sum_{i=1}^K P_i$. Conversely, any rate vector that is in the dirty paper region of the BC is also in the dual MIMO MAC region with the same total power constraint.

Theorem 2: The dirty paper region of a MIMO BC channel with power constraint P is equal to the capacity region of the dual MIMO MAC with sum power constraint P .

$$\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) = \mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger).$$

Proof: We first prove $\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \supseteq \mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$ by showing that every rate vector achieved by successive decoding in the MAC is also in the dirty paper region of the dual MIMO BC. More specifically, we show by the MAC to BC transformations below that for every set of MAC covariance matrices $\mathbf{P}_1, \dots, \mathbf{P}_K$ and any decoding order in the MAC, there exist BC covariance matrices $\Sigma_1, \dots, \Sigma_K$ using the same sum power as the MAC (i.e. $\sum_{i=1}^K \text{Tr}(\mathbf{P}_i) = \sum_{i=1}^K \text{Tr}(\Sigma_i)$) such that the MAC rates are achievable in the BC using the dirty paper coding method described in Section III-A. Each set of MAC covariance matrices corresponds to a K -dimensional polyhedron, as described in (5), with the $K!$ corner points of the polyhedron corresponding to performing successive decoding at the receiver in one of the $K!$ possible decoding orders. By the convexity of the dirty paper region (due to the convex hull operation), it is sufficient to show that the corner points of all polyhedrons (i.e. the successive decoding points) corresponding to all MAC covariance matrices are in the dirty paper region of the dual MIMO BC. Thus, with the MAC to BC transformations described below, this implies $\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \supseteq \mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$.

We complete the proof by showing $\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \subseteq \mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$. We prove this by showing (via the BC to MAC transformations below) that for every set of BC covariance matrices and any encoding order there exist MAC covariance matrices that achieve the same set of rates using the same sum power. The convexity of the MIMO MAC sum power constraint region thus implies that $\mathcal{C}_{\text{DPC}}(P, \mathbf{H}) \subseteq \mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$. This completes the proof, provided we have the transformations given below that map the MAC covariances to the BC covariances and vice versa. ■

Next, we explain some terminology used in the transformations, followed by the actual transformations. It is important to point out that the transformations require a reverse decoding/encoding order of the users in the dual MAC/BC channel. In other words, if User 1 is decoded first in the MAC (i.e. User 1 suffers interference of all other users' signals), then we must encode User 1's signal *last* (i.e. no interference from other users) in the BC to achieve the same rates using these transformations. Also, notice that the proof of duality only requires *existence* of BC covariance matrices which satisfy the rates achieved by a set of MAC covariance matrices, and vice versa. However, the below transformations actually provide equations for the transformed BC covariances as a function of the MAC covariances, and vice versa. This can be quite useful because it is generally much easier to find the optimal BC covariance matrices by finding the optimal MIMO MAC covariance matrices and then transforming the matrices to BC covariance matrices (due to the convex structure of the MIMO MAC rate equations) than it is to directly search for the optimum BC covariance matrices.

A. Terminology

First, we explain the terms *effective channel* and *flipped channel*. A single user MIMO system Θ with channel matrix \mathbf{H} , additive Gaussian noise with covariance \mathbf{X} , and additive independent Gaussian interference with covariance \mathbf{Z} is said to have an effective channel of $(\mathbf{X} + \mathbf{Z})^{-1/2}\mathbf{H}$. The set of rates achievable by Θ and a different system with channel matrix equal to the effective channel and with additive white noise of unit variance and no interference are the same. Also, the capacity of a system Θ_1 with effective channel matrix \mathbf{Y} and the capacity of system Θ_2 with effective channel matrix \mathbf{Y}^\dagger , termed the flipped channel, are the same [2]. In other words, for every transmit covariance Σ in Θ_1 , there exists a $\bar{\Sigma}$ in Θ_2 with $\text{Tr}(\bar{\Sigma}) \leq \text{Tr}(\Sigma)$ such that the rate achieved by $\bar{\Sigma}$ in Θ_2 is equal to the rate achieved by Σ in Θ_1 . In Appendix A we show that $\bar{\Sigma} = \mathbf{F}\mathbf{G}^\dagger \Sigma \mathbf{G}\mathbf{F}^\dagger$ meets this criterion where the Singular Value Decomposition (SVD) of \mathbf{Y} is $\mathbf{Y} = \mathbf{F}\mathbf{\Lambda}\mathbf{G}^\dagger$ where $\mathbf{\Lambda}$ is square and diagonal¹. Next, we describe the covariance transformations.

B. MAC to BC Transformation

In this section we derive a transformation that takes as inputs a set of MAC covariance matrices and a decoding order and outputs a set of BC covariances with the same sum power as the MAC covariances that achieve rates equal to the rates achieved in the MAC using the MAC covariance matrices and successive decoding with the specified decoding order.

Since the numbering of the users is arbitrary, we assume that User 1 is decoded first, User 2 second, and so on at the MAC receiver.

Let $\mathbf{A}_j \triangleq (\mathbf{I} + \mathbf{H}_j(\sum_{i=1}^{j-1} \Sigma_i)\mathbf{H}_j^\dagger)$ and $\mathbf{B}_j \triangleq (\mathbf{I} + \sum_{i=j+1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)$. The rate achieved by User j in the MAC for some arbitrary set of positive semi-definite covariance matrices $(\mathbf{P}_1, \dots, \mathbf{P}_K)$ is given by

$$\begin{aligned} R_j^M &= \log \frac{|\mathbf{I} + \sum_{i=j}^K (\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)|}{|\mathbf{I} + \sum_{i=j+1}^K (\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)|} \\ &= \log \left| \mathbf{I} + \left(\mathbf{I} + \sum_{i=j+1}^K (\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i) \right)^{-1} \mathbf{H}_j^\dagger \mathbf{P}_j \mathbf{H}_j \right| \\ &= \log \left| \mathbf{I} + \mathbf{B}_j^{-1} \mathbf{H}_j^\dagger \mathbf{P}_j \mathbf{H}_j \right|. \end{aligned}$$

Notice that \mathbf{B}_j represents the interference experienced by User j in the MAC. To simplify, we take the square root of \mathbf{B}_j^{-1} and use the property $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$. We also introduce $\mathbf{A}_j^{-1/2} \mathbf{A}_j^{1/2} = \mathbf{I}$ into the expression to get

$$R_j^M = \log \left| \mathbf{I} + \mathbf{B}_j^{-1/2} \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2} \mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2} \mathbf{A}_j^{-1/2} \mathbf{H}_j \mathbf{B}_j^{-1/2} \right|.$$

Treating $\mathbf{B}_j^{-1/2} \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2}$ as the effective channel of the system,

¹Note that the standard SVD command in MATLAB does not always return a square and diagonal matrix of singular values, so modification may be necessary to generate the flipped matrix correctly.

we flip the channel and find $\overline{\mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2}}$ such that

$$\begin{aligned} \overline{\text{Tr}(\mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2})} &\leq \text{Tr}(\mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2}) \\ R_j^M &= \log \left| \mathbf{I} + \mathbf{A}_j^{-1/2} \mathbf{H}_j \mathbf{B}_j^{-1/2} \overline{\mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2}} \mathbf{B}_j^{-1/2} \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2} \right|. \end{aligned}$$

Now consider the rate of User j in the BC assuming that the opposite encoding order is used (i.e. User 1 is encoded last, User 2 second to last, etc)

$$\begin{aligned} R_j^B &= \log \frac{\left| \mathbf{I} + \sum_{i=1}^j (\mathbf{H}_i \boldsymbol{\Sigma}_i \mathbf{H}_i^\dagger) \right|}{\left| \mathbf{I} + \sum_{i=1}^{j-1} (\mathbf{H}_i \boldsymbol{\Sigma}_i \mathbf{H}_i^\dagger) \right|} \\ &= \log \left| \mathbf{I} + \mathbf{A}_j^{-1} \mathbf{H}_j \boldsymbol{\Sigma}_j \mathbf{H}_j^\dagger \right| \\ &= \log \left| \mathbf{I} + \mathbf{A}_j^{-1/2} \mathbf{H}_j \boldsymbol{\Sigma}_j \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2} \right|. \end{aligned}$$

Here \mathbf{A}_j represents the interference experienced by User j in the BC. If we choose the BC covariances as

$$\boldsymbol{\Sigma}_1 = \mathbf{B}_1^{-1/2} \overline{\mathbf{P}_1} \mathbf{B}_1^{-1/2} \quad (8)$$

\vdots

$$\boldsymbol{\Sigma}_j = \mathbf{B}_j^{-1/2} \overline{\mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2}} \mathbf{B}_j^{-1/2} \quad (9)$$

\vdots

$$\boldsymbol{\Sigma}_K = \overline{\mathbf{A}_K^{1/2} \mathbf{P}_K \mathbf{A}_K^{1/2}} \quad (10)$$

clearly we see $R_j^M = R_j^B$. Additionally, it is easy to show that the resulting covariance matrices are all symmetric and positive semi-definite. In Appendix B, we show that the transformations given by (8)-(10) satisfy the sum trace constraint, or that $\sum_{i=1}^K \text{Tr}(\boldsymbol{\Sigma}_i) \leq \sum_{i=1}^K \text{Tr}(\mathbf{P}_i)$. Note that $\boldsymbol{\Sigma}_j$ depends only on $\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_{j-1}$, and hence the $\boldsymbol{\Sigma}_j$ can be computed sequentially in increasing order. By doing this for all K users, we find covariance matrices for the BC that achieve the same rate as in the MAC. If we substitute in the expression generating the flipped channel, the expression for the BC covariance matrix of the j -th user in (9) can be expanded as:

$$\boldsymbol{\Sigma}_j = \mathbf{B}_j^{-1/2} \mathbf{F}_j \mathbf{G}_j^\dagger \mathbf{A}_j^{1/2} \mathbf{P}_j \mathbf{A}_j^{1/2} \mathbf{G}_j \mathbf{F}_j^\dagger \mathbf{B}_j^{-1/2} \quad (11)$$

where the effective channel $\mathbf{B}_j^{-1/2} \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2}$ is decomposed using the SVD as $\mathbf{B}_j^{-1/2} \mathbf{H}_j^\dagger \mathbf{A}_j^{-1/2} = \mathbf{F}_j \boldsymbol{\Lambda}_j \mathbf{G}_j^\dagger$, where $\boldsymbol{\Lambda}_j$ is a square and diagonal matrix.

C. BC to MAC transformation

In this section we derive a transformation which, given a set of BC covariance matrices and an encoding order, outputs a set of MAC covariances with the same sum power as the BC covariances that achieve MAC rates (using successive decoding) equal to the rates achieved in the BC using the BC covariance matrices. These transformations are almost identical to the MAC-to-BC transformations. For the dirty paper encoding at the BC, we assume that User K is encoded first, User $K-1$ second, and so on in decreasing order. Along the same lines as the MAC-BC

transformation, we treat $\mathbf{A}_j^{-1/2} \mathbf{H}_j \mathbf{B}_j^{-1/2}$ as the effective channel and $\mathbf{B}_j^{1/2} \boldsymbol{\Sigma}_j \mathbf{B}_j^{1/2}$ as the covariance matrix. By flipping the effective channel, we obtain $\overline{\mathbf{B}_j^{1/2} \boldsymbol{\Sigma}_j \mathbf{B}_j^{1/2}}$ and obtain the transformation

$$\mathbf{P}_K = \mathbf{A}_K^{-1/2} \overline{\boldsymbol{\Sigma}_K} \mathbf{A}_K^{-1/2} \quad (12)$$

\vdots

$$\mathbf{P}_j = \mathbf{A}_j^{-1/2} \overline{\mathbf{B}_j^{1/2} \boldsymbol{\Sigma}_j \mathbf{B}_j^{1/2}} \mathbf{A}_j^{-1/2} \quad (13)$$

\vdots

$$\mathbf{P}_1 = \overline{\mathbf{B}_1^{1/2} \boldsymbol{\Sigma}_1 \mathbf{B}_1^{1/2}}. \quad (14)$$

As before, if we use the opposite decoding order in the MAC (i.e. User 1 decoded first, etc.), this transformation ensures that the rates of all users in the BC and MAC are equivalent along with the total power used in the BC and MAC. Also note that we can sequentially compute the \mathbf{P}_j 's in decreasing numerical order. If we substitute in the expression generating the flipped channel, the expression for the MAC covariance matrix of the j -th user in (13) can be expanded as:

$$\mathbf{P}_j = \mathbf{A}_j^{-1/2} \mathbf{F}_j \mathbf{G}_j^\dagger \mathbf{B}_j^{1/2} \boldsymbol{\Sigma}_j \mathbf{B}_j^{1/2} \mathbf{G}_j \mathbf{F}_j^\dagger \mathbf{A}_j^{-1/2} \quad (15)$$

where the effective channel $\mathbf{A}_j^{-1/2} \mathbf{H}_j \mathbf{B}_j^{-1/2}$ is decomposed using the SVD as $\mathbf{A}_j^{-1/2} \mathbf{H}_j \mathbf{B}_j^{-1/2} = \mathbf{F}_j \boldsymbol{\Lambda}_j \mathbf{G}_j^\dagger$, where $\boldsymbol{\Lambda}_j$ is a square and diagonal matrix.

D. MIMO MAC with Individual Power Constraints

We can also obtain the capacity region of a MIMO MAC with individual power constraints on each user from the dirty paper region of a dual MIMO BC. By Theorem 3 of [9], we can characterize the individual power constraint MIMO MAC capacity region as an intersection of sum power constraint MIMO MAC capacity regions. By duality, we know that the sum power constraint MIMO MAC capacity region is equal to the dirty paper achievable region of the dual MIMO BC.

Corollary 1: The capacity region of a MIMO MAC is the intersection of the scaled dirty paper regions of the MIMO BC. Mathematically, this is stated as:

$$\mathcal{C}_{\text{MAC}}(P_1, \dots, P_K, \mathbf{H}^\dagger) = \bigcap_{\alpha > 0} \mathcal{C}_{\text{DFC}} \left(\sum_{i=1}^K \frac{P_i}{\alpha_i}; [\sqrt{\alpha_1} \mathbf{H}_1^T \cdots \sqrt{\alpha_K} \mathbf{H}_K^T]^T \right). \quad (16)$$

Proof: This result can be obtained by a straightforward application of Theorem 3 from [9] to the MIMO MAC capacity region and the duality developed in Theorem 2. The scaled MIMO BC here refers to the channel where the matrix of each receiver \mathbf{H}_i is scaled by $\sqrt{\alpha_i}$. ■

V. SUM RATE CAPACITY OF BC CHANNELS

In the previous sections we showed that the dirty paper region of the BC and the union of the dual MAC capacity regions are equivalent. Now we show that the dirty paper broadcasting strategy is the capacity achieving strategy for the sum rate capacity of the MIMO BC. To do this, we make use of the duality of the

dirty paper region and the dual MAC to show that the dirty paper region achieves an upper bound on the sum rate capacity of the MIMO BC.

In [10], Sato presents an upper bound on the capacity region of general BCs. This bound utilizes the capacity of the *cooperative system* as defined in Section II. Since the cooperative system is the same as the BC, but with receiver coordination, the capacity of the cooperative system ($\mathcal{C}_{\text{coop}}(P, \mathbf{H})$) is an upper bound on the BC sum rate capacity ($\mathcal{C}_{\text{BC}}^{\text{sumrate}}(P, \mathbf{H})$). This bound is not tight in general, but by introducing noise correlation at the different receivers, we can get a much stronger bound.

Since the capacity region of a general BC depends only on the marginal transition probabilities of the channel (i.e. $p(y_i|x)$) and not on the joint distribution $p(y_1, \dots, y_K|x)$ [16, Theorem 14.6.1], correlation between the noise vectors at *different receivers* of the BC does not affect the BC capacity region. It does, however, affect the capacity of the cooperative system, which is still an upper bound on the sum rate of the BC. Therefore we retain $\mathbb{E}(\mathbf{n}_i \mathbf{n}_i^\dagger) = \mathbf{I}$, $1 \leq i \leq K$ as before (i.e. noise components at the multiple receive antennas within a single receiver are uncorrelated) and let $\mathbb{E}(\mathbf{n}_i \mathbf{n}_j^\dagger) \triangleq \mathbf{X}_{i,j} \prec \mathbf{I}$, since introducing noise correlation within a receiver affects the broadcast capacity region. Let \mathbf{Z} denote the noise covariance matrix in the cooperative system (i.e. $\mathbf{Z} = E[\mathbf{z}\mathbf{z}^T]$ where $\mathbf{z} = [\mathbf{n}_1 \dots \mathbf{n}_K]^T$) to define the set S to be all *non-singular* (or strictly positive definite) noise covariance matrices satisfying the Sato upper bound conditions

$$S = \left\{ \mathbf{Z} : \mathbf{Z} > 0, \mathbf{Z} = \begin{bmatrix} \mathbf{I} & \dots & \mathbf{X}_{K,1}^\dagger \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{K,1} & \dots & \mathbf{I} \end{bmatrix} \right\}. \quad (17)$$

Then for any $\mathbf{Z} \in S$, the cooperative capacity $\mathcal{C}_{\text{coop}}(P, \mathbf{Z}^{-1/2}\mathbf{H})$ is an upper bound to $\mathcal{C}_{\text{BC}}^{\text{sumrate}}(P, \mathbf{H})$, since receiver coordination can only increase capacity. Hence an upper bound on $\mathcal{C}_{\text{BC}}^{\text{sumrate}}(P, \mathbf{H})$ can be obtained as:

$$\mathcal{C}_{\text{BC}}^{\text{sumrate}}(P, \mathbf{H}) \leq \inf_{\mathbf{Z} \in S} \mathcal{C}_{\text{coop}}(P, \mathbf{Z}^{-1/2}\mathbf{H}).$$

From [2], the cooperative capacity is defined as:

$$\mathcal{C}_{\text{coop}}(P, \mathbf{Z}^{-1/2}\mathbf{H}) = \max_{\Sigma \geq 0, \text{Tr}(\Sigma) \leq P} \log |\mathbf{I} + \mathbf{Z}^{-1/2}\mathbf{H}\Sigma\mathbf{H}^\dagger\mathbf{Z}^{-1/2}|.$$

Using this definition, we write the Sato upper bound as:

$$\mathcal{C}_{\text{Sato}}(P, \mathbf{H}) \triangleq \inf_{\mathbf{Z} \in S} \max_{\Sigma \geq 0, \text{Tr}(\Sigma) \leq P} \log |\mathbf{I} + \mathbf{Z}^{-1/2}\mathbf{H}\Sigma\mathbf{H}^\dagger\mathbf{Z}^{-1/2}|. \quad (18)$$

Thus the region given by

$$\left\{ (R_1, \dots, R_K) : \sum_{i=1}^K R_i \leq \mathcal{C}_{\text{Sato}}(P, \mathbf{H}) \right\}$$

is an upper bound on $\mathcal{C}_{\text{BC}}(P, \mathbf{H})$. Next, we show that the sum rate capacity of the MIMO BC actually equals the Sato upper bound.

Theorem 3: The sum rate capacity of the MIMO BC equals the Sato upper bound. Furthermore, the dirty paper coding strategy achieves the sum rate capacity of the MIMO BC

$$\mathcal{C}_{\text{BC}}^{\text{sumrate}}(P, \mathbf{H}) = \mathcal{C}_{\text{DPC}}^{\text{sumrate}}(P, \mathbf{H}) = \mathcal{C}_{\text{Sato}}(P, \mathbf{H}). \quad (19)$$

Proof: Since the sum rate capacity of the MIMO BC can be no larger than the Sato upper bound, it is sufficient to show that the Sato upper bound is actually achievable in the MIMO BC using dirty paper coding. Note that, by duality, we know that the maximum sum rate of the MIMO MAC equals the maximum sum rate of the dirty paper achievable region. We therefore must show the inequality:

$$\begin{aligned} \mathcal{C}_{\text{DPC}}^{\text{sumrate}}(P, \mathbf{H}) &= \mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) \\ &= \max_{\{\mathbf{P}_j \geq 0, \sum_{i=1}^K \text{Tr}(\mathbf{P}_i) \leq P\}} \log |\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i| \\ &\geq \mathcal{C}_{\text{Sato}}(P, \mathbf{H}). \end{aligned} \quad (21)$$

We prove (21) by using Lagrangian duality to express both the Sato upper bound and the MIMO MAC sum rate capacity in different forms. Specifically, as shown in Appendix C, we can alternatively write the Sato upper bound defined in (18) as (see (58))

$$\begin{aligned} \mathcal{C}_{\text{Sato}}(P, \mathbf{H}) &= \inf_{\mathbf{Z} \in S} \min_{\mathbf{A}, \lambda} -\log |\mathbf{A}| + \text{Tr}(\mathbf{A}) + \lambda P - t \\ &\text{such that} \\ &\lambda \geq 0, \mathbf{A} \geq 0 \\ &\lambda \mathbf{Z} \geq \mathbf{H}\mathbf{A}\mathbf{H}^\dagger, \end{aligned} \quad (22)$$

and the MIMO MAC sum rate capacity (20) as

$$\begin{aligned} \mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) &= \min_{\mathbf{A}, \lambda} -\log |\mathbf{A}| + \text{Tr}(\mathbf{A}) + \lambda P - t \\ &\text{such that} \\ &\lambda \geq 0, \mathbf{A} \geq 0, \\ &\lambda \mathbf{I} \geq \mathbf{H}_i \mathbf{A} \mathbf{H}_i^\dagger \quad \forall i. \end{aligned} \quad (23)$$

Notice that the objective functions of (22) and (23) are the same, but the variables and constraints are different. We will show that $\mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) \geq \mathcal{C}_{\text{Sato}}(P, \mathbf{H})$ by constructing a feasible solution to the Sato upper bound dual problem from an optimal solution to the MIMO MAC sum rate dual problem.

Let $\lambda = \lambda_0$, $\mathbf{A} = \mathbf{A}_0$ be an optimizing solution to (23), i.e. $\mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) = -\log |\mathbf{A}_0| + \text{Tr}(\mathbf{A}_0) + \lambda_0 P - t$. Since (23) is a minimization of a convex function over a closed set, we know that the minimum is achieved, so a minimizing pair $(\lambda_0, \mathbf{A}_0)$ exists. We prove (21) by explicitly constructing a feasible set of variables $(\lambda, \mathbf{A}, \mathbf{Z})$ for the Sato upper bound (22) such that the objective functions in (23) and (22) are equal.

Let us first consider the choice of values of $(\lambda, \mathbf{A}, \mathbf{Z})$ as

$$\lambda = \lambda_0 \quad (24)$$

$$\mathbf{A} = \mathbf{A}_0 \quad (25)$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{I} & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0} & \dots & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_K^\dagger}{\lambda_0} \\ \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0} & \mathbf{I} & \dots & \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_K^\dagger}{\lambda_0} \\ \dots & \dots & \dots & \dots \\ \frac{\mathbf{H}_K \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0} & \frac{\mathbf{H}_K \mathbf{A}_0 \mathbf{H}_2^\dagger}{\lambda_0} & \dots & \mathbf{I} \end{bmatrix}. \quad (26)$$

As long as this choice of \mathbf{Z} is positive definite, it can be verified by the method used below that this set is feasible for (22) and that the objective functions of both minimizations are the same. Thus we have constructed the worst case noise \mathbf{Z} for the Sato upper bound and have shown (21), but only for the case when $\mathbf{Z} > 0$. However, in many practical cases this choice of \mathbf{Z} is singular, and hence, not a feasible choice of \mathbf{Z} for (22).

To circumvent this singularity, we construct instead a family of feasible points (i.e $\mathbf{Z} > 0$) by introducing an arbitrary parameter $\delta > 0$. This family of values of $(\lambda, \mathbf{A}, \mathbf{Z})$ is given by

$$\lambda = \lambda_0 + \delta \quad (27)$$

$$\mathbf{A} = \mathbf{A}_0 \quad (28)$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{I} & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_2^\dagger}{\lambda_0 + \delta} & \dots & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_K^\dagger}{\lambda_0 + \delta} \\ \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0 + \delta} & \mathbf{I} & \dots & \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_K^\dagger}{\lambda_0 + \delta} \\ \dots & \dots & \dots & \dots \\ \frac{\mathbf{H}_K \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0 + \delta} & \frac{\mathbf{H}_K \mathbf{A}_0 \mathbf{H}_2^\dagger}{\lambda_0 + \delta} & \dots & \mathbf{I} \end{bmatrix}. \quad (29)$$

We need to ensure that this set is feasible for (22). Since $(\lambda_0, \mathbf{A}_0)$ are an optimizing solution of (23), $(\lambda_0, \mathbf{A}_0)$ must satisfy the constraints in (23). Therefore we have that $\lambda = \lambda_0 + \delta > 0$ and $\mathbf{A} = \mathbf{A}_0 \geq 0$. Since the matrix \mathbf{Z} is block diagonal and symmetric by construction, we see that if $\mathbf{Z} > 0$ then $\mathbf{Z} \in \mathcal{S}$. We thus need to verify that $\mathbf{Z} > 0$ and that $\lambda \mathbf{Z} \geq \mathbf{H} \mathbf{A} \mathbf{H}^\dagger$.

Note that $\lambda_0 + \delta > 0$ and

$$\begin{aligned} (\lambda_0 + \delta) \mathbf{Z} - \mathbf{H} \mathbf{A}_0 \mathbf{H}^\dagger &= \text{diag}[(\lambda_0 + \delta) \mathbf{I} - \mathbf{H}_i \mathbf{A}_0 \mathbf{H}_i^\dagger] \\ &= \text{diag}[\lambda_0 \mathbf{I} - \mathbf{H}_i \mathbf{A}_0 \mathbf{H}_i^\dagger] + \delta \mathbf{I} \end{aligned} \quad (30)$$

Since $(\lambda_0, \mathbf{A}_0)$ are an optimizing solution of (23), we have $\lambda_0 \mathbf{I} - \mathbf{H}_i \mathbf{A}_0 \mathbf{H}_i^\dagger \geq 0$ for all i . This implies that $\text{diag}[\lambda_0 \mathbf{I} - \mathbf{H}_i \mathbf{A}_0 \mathbf{H}_i^\dagger] \geq 0$. Since $\delta > 0$, we have

$$\lambda \mathbf{Z} - \mathbf{H} \mathbf{A}_0 \mathbf{H}^\dagger = \text{diag}[\lambda_0 \mathbf{I} - \mathbf{H}_i \mathbf{A}_0 \mathbf{H}_i^\dagger] + \delta \mathbf{I} > 0. \quad (31)$$

This implies that $\lambda \mathbf{Z} > \mathbf{H} \mathbf{A}_0 \mathbf{H}^\dagger = \mathbf{H} \mathbf{A} \mathbf{H}^\dagger$. It thus remains to show that $\mathbf{Z} > 0$. Since $\mathbf{A} \geq 0$, we also have that $\mathbf{H} \mathbf{A} \mathbf{H}^\dagger \geq 0$. This implies that $\lambda \mathbf{Z} > 0$. Since $\lambda > 0$, we then get $\mathbf{Z} > 0$. Hence $(\lambda, \mathbf{A}, \mathbf{Z})$ form a feasible set of values for (22). Since the Sato upper bound is equal to the infimum (over the feasible set) of the objective function, we have

$$\begin{aligned} \mathcal{C}_{\text{Sato}}(P, \mathbf{H}) &\leq -\log |\mathbf{A}| + \text{Tr}(\mathbf{A}) + \lambda P - t \\ &= -\log |\mathbf{A}_0| + \text{Tr}(\mathbf{A}_0) + \lambda_0 P - t + \delta P \\ &= \mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) + \delta P. \end{aligned} \quad (32)$$

Since this holds for any $\delta > 0$, we get $\mathcal{C}_{\text{Sato}}(P, \mathbf{H}) \leq \mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H})$, which completes the proof of the theorem. ■

A nice property of the proof of the sum rate capacity proof is that it is *constructive* in the sense that the proof generates worst-case noise covariances for the Sato upper bound, assuming that the optimizing solution to the MIMO MAC sum rate problem is known. Specifically, in Equation (29), we explicitly construct a noise covariance for which the cooperative capacity is larger than the MIMO MAC sum rate capacity by an arbitrarily small amount δP . Though we show that the constructed matrix $\mathbf{Z} > 0$ when $\delta > 0$, as noted earlier when $\delta \geq 0$ we still are guaranteed

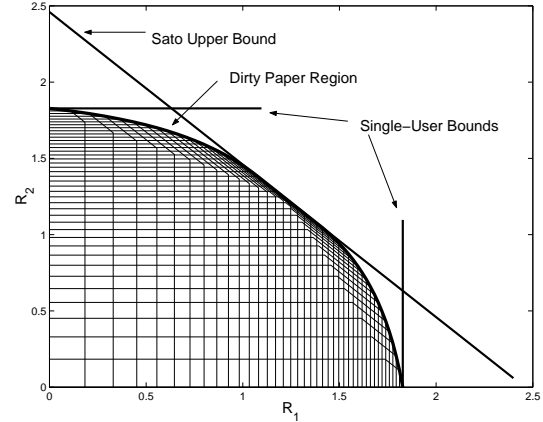


Fig. 2. Dirty paper broadcast region: $K = 2, t = 2, r_1 = r_2 = 1, H_1 = [1 \ .4], H_2 = [.4 \ 1], P = 10$

$\mathbf{Z} \geq 0$. If the constructed matrix \mathbf{Z} with $\delta = 0$ is strictly non-singular (i.e. $\mathbf{Z} > 0$), then the cooperative capacity with noise covariance \mathbf{Z} is *equal* to the MIMO MAC sum rate capacity. Therefore, in these cases \mathbf{Z} is in fact a worst case noise covariance for the Sato upper bound. Numerically, we also find \mathbf{Z} to be a worst case noise covariance for cases when \mathbf{Z} is singular when $\delta = 0$, as shown in Example 2 below.

The dirty paper BC region and the capacity regions of the dual MAC, along with the Sato upper bound and single-user bounds, are illustrated for a symmetric two user channel in Fig. 2. The dirty paper region is the union of the pentagons in the figure because the dirty paper region is formed as a union of the individual power constraint MAC regions. Since each receiver has only a single antenna, the dual MIMO MAC region with individual power constraints is a simple pentagon. The capacity upper bound is obtained by taking the intersection of the regions formed by the two single user optimum corner points (which are parallel to the axes) and the Sato upper bound, which is tight at the sum rate capacity. Note that the region formed by all three upper bounds is in fact quite close to the dirty paper achievable region. Also note that the boundary of the dirty paper achievable region has a straight line segment at the sum rate point. This characteristic of possessing a straight line segment (i.e. a time-division portion) at sum rate is also true for the MIMO MAC capacity region when the transmitters have more than one antenna [4].

VI. NUMERICAL EXAMPLES

In this section we provide two numerical examples to better illustrate the concepts discussed in the paper.

Example 1: Consider a two user broadcast channel with $P = 1$ and channel matrices

$$\mathbf{H}_1 = \begin{bmatrix} 1 & .8 \\ .5 & 2 \end{bmatrix}, \mathbf{H}_2 = \begin{bmatrix} .2 & 1 \\ 2 & .5 \end{bmatrix}. \quad (33)$$

The dirty paper achievable region is very difficult to compute without employing duality, as discussed in Section III-A. Thus, we find the dual MAC region using convex optimization techniques to obtain the achievable region in Figure 3.

To compute the Sato upper bound for the problem, we solve the dual problem to the MIMO MAC. Note that this problem

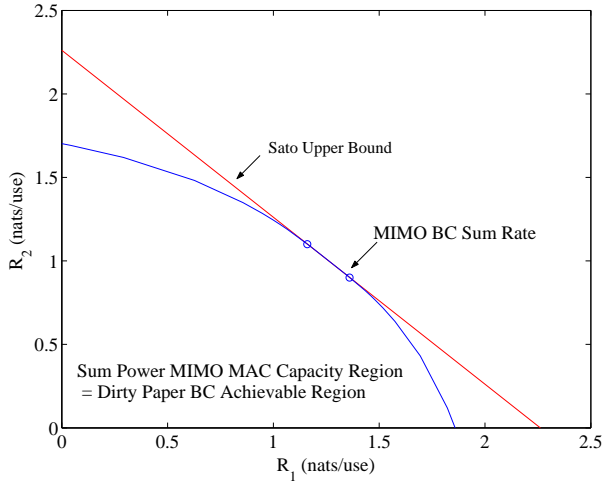


Fig. 3. Achievable region and Sato upper bound for Example 1

is a convex problem with linear matrix inequality constraints. There are many techniques in convex optimization literature to solve such problems. We use an easily available software called SDPSOL, developed by Boyd and Wu [17], to get that $\mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H}) = 2.2615$ nats/sec and obtain \mathbf{A}_0, λ_0 to be

$$\mathbf{A}_0 = \begin{bmatrix} 3.231 & 1.43 \\ 1.43 & 3.603 \end{bmatrix}$$

$$\lambda_0 = 1.2879.$$

From these, we obtain the worst case noise (using Equation (29) with $\delta = 0$)

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & .1332 & .4446 \\ 0 & 1 & .4478 & .0613 \\ .1332 & .4478 & 1 & 0 \\ .4446 & .0613 & 0 & 1 \end{bmatrix}$$

which is non-singular. Therefore, we find that $\mathcal{C}_{\text{coop}}(P, \mathbf{Z}^{-1/2}\mathbf{H}) = \mathcal{C}_{\text{Sato}}(P, \mathbf{H}) = \mathcal{C}_{\text{union}}^{\text{sumrate}}(P, \mathbf{H})$. The corresponding upper bound to the capacity region is shown in Figure 3.

The sum rate maximizing covariance matrices in the MAC are

$$\mathbf{P}_1 = \begin{bmatrix} .0720 & .1827 \\ .1827 & .4634 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 0 & .0026 \\ .0026 & .4646 \end{bmatrix}.$$

Notice that the sum rate is not maximized at a single point on the boundary of the capacity region, but it is actually maximized along a line segment. The corner points of this line segment are circled in Figure 3. In the MAC, the lower corner point of this line segment can be achieved using the above covariance matrices $\mathbf{P}_1, \mathbf{P}_2$ and by decoding User 1 last. The upper corner point of the line segment can be achieved using the *same* covariance matrices, but the *opposite* decoding order (i.e. decode User 2 last). Any other point on the line segment can be achieved by time-sharing between these two decoding orders.

We can use the MAC-BC transformations in (8) to find the corresponding sum rate capacity-achieving BC covariance matrices. Note, however, that the transformations depend on the decoding order in the MAC. If we assume that User 1 is decoded

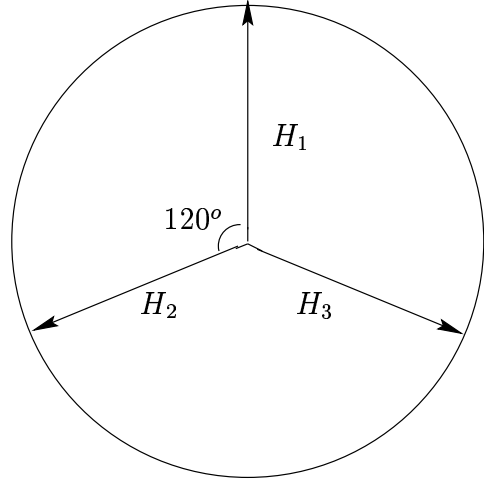


Fig. 4. Channel Parameters for Example 2

last in the MAC, the transformed BC covariances are:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} .0746 & .1932 \\ .1932 & .5004 \end{bmatrix}, \quad \mathbf{\Sigma}_2 = \begin{bmatrix} .4104 & -.0776 \\ -.0776 & .0147 \end{bmatrix}.$$

The lower corner point of the sum rate line segment is then achievable in the BC using these covariance matrices and by decoding User 2 last (i.e using dirty paper coding for User 2 to cancel out the signal of User 1). To achieve the upper corner point of the sum rate line segment, we must perform the MAC-BC transformations using the opposite order in the MAC. We then get

$$\mathbf{\Sigma}_1 = \begin{bmatrix} .0001 & -.0069 \\ -.0069 & .4841 \end{bmatrix}, \quad \mathbf{\Sigma}_2 = \begin{bmatrix} .4849 & .1225 \\ .1225 & .0309 \end{bmatrix}.$$

Clearly, these BC covariance matrices are different than those used to achieve the other corner point of the sum rate line segment. Therefore, we see that in the MAC the corner points of the sum rate boundary can be achieved by using the same set of covariance matrices and different decoding orders. In the BC, however, a different decoding order and different covariance matrices are needed to achieve the corner points of the sum rate boundary.

Example 2: Consider a three user broadcast channel, with two antennas at the transmitter ($t = 2$) and one antenna each per receiver ($r_1 = r_2 = r_3 = 1$). The channel matrices are given by $\mathbf{H}_1 = [0 \ 1]$, $\mathbf{H}_2 = [-\sqrt{3}/2 \ -1/2]$, $\mathbf{H}_3 = [\sqrt{3}/2 \ -1/2]$ and the total power constraint is $P = 1$. Note that the channels are unit vectors in Euclidean space, and are spaced 120 degrees apart, as shown in Figure 4. Also note that the channel matrix \mathbf{H}

$$\mathbf{H} = \begin{bmatrix} 0 & 1 \\ -\sqrt{3}/2 & -1/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} \quad (34)$$

has rank two, and that $\mathbf{H}_1 = -(\mathbf{H}_2 + \mathbf{H}_3)$.

First, let us consider the dual MAC problem. By the symmetric structure of these channels, it is clear that allocating equal power to each user maximizes the sum rate of this system. Thus

sum rate capacity is achieved with $P_1 = P_2 = P_3 = 1/3$ and any MAC decoding order.

Using the MAC-to-BC transformations in (8), we find covariances in the broadcast (corresponding to encoding User 1 last, User 2 second, and User 3 first) that achieve the same sum rate point on the capacity region to be

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 0 & 0 \\ 0 & .2857 \end{bmatrix}, \mathbf{\Sigma}_2 = \begin{bmatrix} .2187 & .1623 \\ .1623 & .1205 \end{bmatrix} \quad (35)$$

$$\mathbf{\Sigma}_3 = \begin{bmatrix} .2812 & -.1624 \\ -.1624 & .0937 \end{bmatrix} \quad (36)$$

and the sum rate capacity equals

$$\log |\mathbf{I} + \frac{1}{3}(\mathbf{H}_1^\dagger \mathbf{H}_1 + \mathbf{H}_2^\dagger \mathbf{H}_2 + \mathbf{H}_3^\dagger \mathbf{H}_3)| = .8109 \text{ nats/use.}$$

Now, let us find the worst case noise for the Sato upper bound. Note that, if $n_1 = -(n_2 + n_3)$, then $y_1 = -(y_2 + y_3)$. This implies that the received signal at one of the antennas is a linear combination of the signal at the other two antennas. Therefore, one receive antenna can be eliminated from the system without any loss in the cooperative capacity. Since $\mathbb{E}(n_1^2) = \mathbb{E}(n_2^2) = \mathbb{E}(n_3^2) = 1$, we require that $\mathbb{E}(n_i n_j) = -.5 \forall i \neq j$. Thus, a noise covariance matrix given by

$$\mathbf{Q}_{\text{cand}} = \begin{bmatrix} 1 & -.5 & -.5 \\ -.5 & 1 & -.5 \\ -.5 & -.5 & 1 \end{bmatrix} \quad (37)$$

corresponds to $n_1 = -(n_2 + n_3)$, which allows us to eliminate one receive antenna. This noise covariance is a candidate for the worst-case noise covariance for the Sato upper bound. Note that this noise covariance is singular. If we eliminate the third antenna output, we are left with the following two input, two output channel:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 \\ -\sqrt{3}/2 & -1/2 \end{bmatrix}. \quad (38)$$

The cooperative capacity with this noise covariance then is

$$\max_{\mathbf{\Sigma}} \log \left| \mathbf{I} + \begin{bmatrix} 1 & -.5 \\ -.5 & 1 \end{bmatrix}^{-1} \mathbf{G} \mathbf{\Sigma} \mathbf{G}^\dagger \right|. \quad (39)$$

To evaluate the expression above, we use the standard water-filling technique [2] and find the optimizing $\mathbf{\Sigma}$ to be $\begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix}$. The corresponding cooperative capacity is thus equal to .8109 nats/channel use. Since the cooperative capacity with noise covariance \mathbf{Q}_{cand} equals the MIMO MAC sum rate capacity, \mathbf{Q}_{cand} is a worst case noise for this problem.

Now, let us use the method for obtaining the worst case noise introduced in the proof of Theorem 3. Using SDPSOL we get $\mathbf{A}_0 = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}$ and $\lambda_0 = 2/3$. From (26) we can construct

the worst case noise covariance to be

$$\mathbf{Q}_0 = \begin{bmatrix} \mathbf{I} & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_2^\dagger}{\lambda_0} & \frac{\mathbf{H}_1 \mathbf{A}_0 \mathbf{H}_3^\dagger}{\lambda_0} \\ \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0} & \mathbf{I} & \frac{\mathbf{H}_2 \mathbf{A}_0 \mathbf{H}_3^\dagger}{\lambda_0} \\ \frac{\mathbf{H}_3 \mathbf{A}_0 \mathbf{H}_1^\dagger}{\lambda_0} & \frac{\mathbf{H}_3 \mathbf{A}_0 \mathbf{H}_2^\dagger}{\lambda_0} & \mathbf{I} \end{bmatrix} \quad (40)$$

$$= \begin{bmatrix} 1 & -.5 & -.5 \\ -.5 & 1 & -.5 \\ -.5 & -.5 & 1 \end{bmatrix}. \quad (41)$$

We immediately notice that $\mathbf{Q}_{\text{cand}} = \mathbf{Q}_0$. Thus, for this case, the singular noise covariance constructed from (26) actually is a worst case noise covariance, even though we are not guaranteed this when the noise is singular. It may in fact be that the noise covariance constructed using (26) is *always* a worst case noise covariance (and not just when it is non-singular), but we have not been able to prove this.

VII. CONCLUSION

In this paper, we established a duality relationship between two seemingly unrelated regions: the achievable region of the MIMO BC obtained using the dirty paper coding and the capacity region of the MIMO MAC. This duality allows us to easily find the dirty paper achievable region and the dirty paper covariance matrices which achieve the boundary of this region. Though the capacity region of the MIMO BC is unknown due to its non-degraded nature, we were able to show that the dirty paper achievable region achieves the sum rate capacity of the MIMO BC through use of the MAC-BC duality and the Sato upper bound. These results open up the possibility that the dirty paper region is the actual capacity region of the MIMO BC and also the possibility that other instances of duality exist in Gaussian multi-terminal networks.

APPENDIX

A. Covariance Matrix for Flipped Channel

Given a covariance matrix $\mathbf{\Sigma}$ for some channel \mathbf{H} , we wish to find a covariance matrix $\bar{\mathbf{\Sigma}}$ such that $\text{Tr}(\bar{\mathbf{\Sigma}}) \leq \text{Tr}(\mathbf{\Sigma})$ and

$$\log |\mathbf{I} + \mathbf{H} \mathbf{\Sigma} \mathbf{H}^\dagger| = \log |\mathbf{I} + \mathbf{H}^\dagger \bar{\mathbf{\Sigma}} \mathbf{H}|. \quad (42)$$

If the SVD of \mathbf{H} is $\mathbf{H} = \mathbf{F} \mathbf{\Lambda} \mathbf{G}^\dagger$ where $\mathbf{\Lambda}$ is square and diagonal, then we propose $\bar{\mathbf{\Sigma}} = \mathbf{F} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{F}^\dagger$. Using the identity $|\mathbf{I} + \mathbf{A} \mathbf{B}| = |\mathbf{I} + \mathbf{B} \mathbf{A}|$ and the fact that $\mathbf{F}^\dagger \mathbf{F} = \mathbf{I}$ and $\mathbf{G}^\dagger \mathbf{G} = \mathbf{I}$ we can write the capacity of the unflipped channel as:

$$\begin{aligned} \log |\mathbf{I} + \mathbf{H} \mathbf{\Sigma} \mathbf{H}^\dagger| &= \log |\mathbf{I} + \mathbf{F} \mathbf{\Lambda} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{\Lambda} \mathbf{F}^\dagger| \\ &= \log |\mathbf{I} + \mathbf{\Lambda} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{\Lambda}|. \end{aligned} \quad (43)$$

We can similarly write the capacity of the flipped channel with our candidate covariance matrix $\bar{\mathbf{\Sigma}}$ as:

$$\begin{aligned} \log |\mathbf{I} + \mathbf{H}^\dagger \bar{\mathbf{\Sigma}} \mathbf{H}| &= \log |\mathbf{I} + \mathbf{G} \mathbf{\Lambda} \mathbf{F}^\dagger \mathbf{F} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{F}^\dagger \mathbf{F} \mathbf{\Lambda} \mathbf{G}^\dagger| \\ &= \log |\mathbf{I} + \mathbf{G} \mathbf{\Lambda} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{\Lambda} \mathbf{G}^\dagger| \\ &= \log |\mathbf{I} + \mathbf{\Lambda} \mathbf{G}^\dagger \mathbf{\Sigma} \mathbf{G} \mathbf{\Lambda}|. \end{aligned} \quad (44)$$

Therefore the rate achieved by $\bar{\mathbf{\Sigma}}$ in the flipped channel is the same as the rate achieved by the original covariance matrix $\mathbf{\Sigma}$ in the unflipped channel. It thus only remains to show

that $\text{Tr}(\bar{\Sigma}) \leq \text{Tr}(\Sigma)$. To do so, we make use of the identity $\text{Tr}(\mathbf{A}\mathbf{B}) = \text{Tr}(\mathbf{B}\mathbf{A})$. Clearly, we can write

$$\text{Tr}(\bar{\Sigma}) = \text{Tr}(\mathbf{F}\mathbf{G}^\dagger \Sigma \mathbf{G}\mathbf{F}^\dagger) = \text{Tr}(\mathbf{G}^\dagger \Sigma \mathbf{G}) = \text{Tr}(\Sigma \mathbf{G}\mathbf{G}^\dagger). \quad (45)$$

We then use Gram-Schmidt to expand \mathbf{G} into a full unitary matrix:

$$\tilde{\mathbf{G}} = [\mathbf{G} \ \bar{\mathbf{G}}] \quad (46)$$

such that $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^\dagger = \tilde{\mathbf{G}}^\dagger\tilde{\mathbf{G}} = \mathbf{I}$. Using this unitary matrix, we can write

$$\begin{aligned} \text{Tr}(\Sigma) &= \text{Tr}(\Sigma \tilde{\mathbf{G}}\tilde{\mathbf{G}}^\dagger) \\ &= \text{Tr}(\Sigma \mathbf{G}\mathbf{G}^\dagger) + \text{Tr}(\Sigma \bar{\mathbf{G}}\bar{\mathbf{G}}^\dagger) \\ &\geq \text{Tr}(\Sigma \mathbf{G}\mathbf{G}^\dagger). \end{aligned} \quad (47)$$

To get (47), we used the fact that the matrix $\bar{\mathbf{G}}^\dagger \Sigma \bar{\mathbf{G}}$ is positive semi-definite, which implies $\text{Tr}(\Sigma \bar{\mathbf{G}}\bar{\mathbf{G}}^\dagger) = \text{Tr}(\bar{\mathbf{G}}^\dagger \Sigma \bar{\mathbf{G}}) \geq 0$.

B. Proof of trace constraint for transformations

In this section, we show that the MAC-BC transformations obtained in (8)-(10) satisfy the sum trace requirement. First, we compute

$$\begin{aligned} \text{Tr}(\Sigma_K) &\leq \text{Tr}(\mathbf{A}_K \mathbf{P}_K) \\ &= \text{Tr}(\mathbf{P}_K) + \sum_{i=1}^{K-1} \text{Tr}(\Sigma_i \mathbf{H}_i^\dagger \mathbf{P}_K \mathbf{H}_i). \end{aligned}$$

By adding $\sum_{i=1}^{K-1} \text{Tr}(\Sigma_i)$ to both sides we get

$$\sum_{i=1}^K \text{Tr}(\Sigma_i) \leq \text{Tr}(\mathbf{P}_K) + \sum_{i=1}^{K-1} \text{Tr}(\Sigma_i (\mathbf{I} + \mathbf{H}_i^\dagger \mathbf{P}_K \mathbf{H}_i)).$$

By the definition of Σ_j , we get

$$\begin{aligned} \text{Tr}(\Sigma_j (\mathbf{I} + \sum_{i=j+1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)) &= \text{Tr}(\Sigma_j \mathbf{B}_j) \\ &\leq \text{Tr}(\mathbf{A}_j \mathbf{P}_j) \\ &= \text{Tr}(\mathbf{P}_j) + \sum_{i=1}^{j-1} \text{Tr}(\Sigma_i \mathbf{H}_i^\dagger \mathbf{P}_j \mathbf{H}_i). \end{aligned}$$

Using this expression for $j = K - 1$ we get

$$\begin{aligned} \sum_{i=1}^K \text{Tr}(\Sigma_i) &\leq \sum_{l=K-1}^K \text{Tr}(\mathbf{P}_l) + \\ &\quad \sum_{i=1}^{K-2} \text{Tr}(\Sigma_i (\mathbf{I} + \sum_{l=K-1}^K \mathbf{H}_l^\dagger \mathbf{P}_l \mathbf{H}_l)). \end{aligned}$$

By induction, we can further show that

$$\sum_{i=1}^K \text{Tr}(\Sigma_i) \leq \sum_{l=j}^K \text{Tr}(\mathbf{P}_l) + \sum_{i=1}^{j-1} \text{Tr}(\Sigma_i (\mathbf{I} + \sum_{l=j}^K \mathbf{H}_l^\dagger \mathbf{P}_l \mathbf{H}_l)),$$

for any j . For $j = 1$, we get

$$\sum_{i=1}^K \text{Tr}(\Sigma_i) \leq \sum_{l=1}^K \text{Tr}(\mathbf{P}_l).$$

The same proof method can be used to show that the BC-MAC transformations in (12)-(14) also satisfy the trace constraints.

C. Finding the Lagrangian dual problem

We first find the dual problem of the maximum sumrate of the MIMO MAC.²

$$\max_{\mathbf{P}_i \in \mathcal{S}} \log |\mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i| \quad (48)$$

over the convex set $\mathcal{S} = \{\mathbf{P}_i : \mathbf{P}_i \geq 0 \ \forall i, \sum_{i=1}^K \{\text{Tr}(\mathbf{P}_i) \leq P\}$.

The problem given by (48) is a convex optimization problem, i.e., it has a concave objective function and a convex constraint set. Hence a convex Lagrangian dual minimization problem can be obtained that achieves the same optimum value at (48). For this, we rewrite (48) as

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{P}_i} & -\log |\mathbf{X}| \\ \text{such that} & \quad \mathbf{X} = \mathbf{I} + \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i \\ & \quad \sum_{i=1}^K \text{Tr}(\mathbf{P}_i) \leq P, \ \mathbf{P}_i \geq 0. \end{aligned}$$

Note that matrix inequalities are associated with dual variables that are matrices, while scalar inequalities are associated with scalar dual variables. The Lagrangian for this problem is:

$$\begin{aligned} \mathcal{L}(\mathbf{X}, \mathbf{P}_i, \mathbf{A}, \mathbf{S}_i, \lambda) &= -\log |\mathbf{X}| + \text{Tr}[\mathbf{A}(\mathbf{X} - \mathbf{I} - \sum_{i=1}^K \mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)] \\ &\quad + \lambda (\sum_{i=1}^K \text{Tr}(\mathbf{P}_i) - P) + \sum_{i=1}^K \text{Tr}(\mathbf{S}_i \mathbf{P}_i). \end{aligned} \quad (49)$$

The dual function is found by minimizing with respect to the primal variables \mathbf{X}, \mathbf{P}_i :

$$g(\mathbf{A}, \mathbf{S}_i, \lambda) = \inf_{\mathbf{X}, \mathbf{P}_i} \mathcal{L}(\mathbf{X}, \mathbf{P}_i, \mathbf{A}, \mathbf{S}_i, \lambda). \quad (50)$$

We obtain the optimality conditions by differentiating the Lagrangian (49) with respect to the primal variables to get

$$\begin{aligned} \lambda \mathbf{I} &= \mathbf{H}_i \mathbf{A} \mathbf{H}_i^\dagger + \mathbf{S}_i \quad \forall i \\ \mathbf{X}^{-1} &= \mathbf{A}. \end{aligned}$$

For any Lagrangians not satisfying these conditions, we get $g(\mathbf{A}, \mathbf{S}_i, \lambda) = -\infty$. For Lagrangians that do satisfy these conditions, we get

$$g(\mathbf{A}, \mathbf{S}_i, \lambda) = \log |\mathbf{A}| - \text{Tr}(\mathbf{A}) - \lambda P + t. \quad (51)$$

²This derivation is based on the dual problem found in [4]

The dual problem is then obtained by maximizing the dual function $g(\mathbf{A}, \mathbf{S}_i, \lambda)$ with respect to the dual variables:

$$\begin{aligned} & \max_{\mathbf{A}, \lambda} \log |\mathbf{A}| - \text{Tr}(\mathbf{A}) - \lambda P + t \quad (52) \\ \text{such that} \quad & \mathbf{A} \geq 0, \lambda \geq 0 \\ & \lambda \mathbf{I} \geq \mathbf{H}_i \mathbf{A} \mathbf{H}_i^\dagger \quad \forall i. \end{aligned}$$

Due to the convexity of the original optimization, the sum rate capacity of the MAC is equal to the solution of the dual problem above (52).

We now use Lagrangian duality to find an alternative form for the Sato upper bound. The Sato upper bound is originally defined in (18) as:

$$\mathcal{C}_{\text{Sato}}(P, \mathbf{H}) \triangleq \inf_{\mathbf{Z} \in \mathcal{S}} \max_{\Sigma \geq 0, \text{Tr}(\Sigma) \leq P} \log |\mathbf{I} + \mathbf{Z}^{-1} \mathbf{H} \Sigma \mathbf{H}^\dagger|. \quad (53)$$

We will find an alternate form for this upper bound by considering the Lagrangian dual of only the inner maximization (and not of the entire expression). For some fixed $\mathbf{Z} > 0$, consider the inner maximization:

$$\max_{\Sigma \geq 0, \text{Tr}(\Sigma) \leq P} \log |\mathbf{I} + \mathbf{Z}^{-1} \mathbf{H} \Sigma \mathbf{H}^\dagger|. \quad (54)$$

Note that the maximization in (54) is the capacity of a multi-antenna system with channel given by $\mathbf{Z}^{-1/2} \mathbf{H}$ and additive white Gaussian noise [2]. The capacity of a system with the conjugate transpose channel $\mathbf{H}^\dagger \mathbf{Z}^{-1/2}$ can be shown to be exactly equal to that in (48) (See [2] for proof). Therefore, (54) is equivalent to:

$$\max_{\mathbf{Q} \geq 0, \text{Tr}(\mathbf{Q}) \leq P} \log |\mathbf{I} + \mathbf{H}^\dagger \mathbf{Z}^{-1/2} \mathbf{Q} \mathbf{Z}^{-1/2} \mathbf{H}|. \quad (55)$$

Notice that this maximization is equivalent to the form of the MAC sumrate maximization in (48) with $K = 1$, $\mathbf{H}_1^\dagger = \mathbf{H}^\dagger \mathbf{Z}^{-1/2}$, and $\mathbf{P}_1 = \mathbf{Q}$. Therefore, the above maximization (55) is equivalent to:

$$\begin{aligned} & \max_{\mathbf{A}, \lambda} \log |\mathbf{A}| - \text{Tr}(\mathbf{A}) - \lambda P + t \quad (56) \\ \text{such that} \quad & \mathbf{A} \geq 0, \lambda \geq 0 \\ & \lambda \mathbf{I} \geq \mathbf{Z}^{-1/2} \mathbf{H} \mathbf{A} \mathbf{H}^\dagger \mathbf{Z}^{-1/2}. \end{aligned}$$

Clearly, we can multiply the last inequality by the term $\mathbf{Z}^{-1/2}$ on both sides and negate the objective function to get:

$$\begin{aligned} & \min_{\mathbf{A}, \lambda} -\log |\mathbf{A}| + \text{Tr}(\mathbf{A}) + \lambda P - t \quad (57) \\ \text{such that} \quad & \mathbf{A} \geq 0, \lambda \geq 0 \\ & \lambda \mathbf{Z} \geq \mathbf{H} \mathbf{A} \mathbf{H}^\dagger. \end{aligned}$$

Since this is equivalent to the inner maximization of the Sato upper bound for every $\mathbf{Z} \in \mathcal{S}$, we can rewrite the Sato upper bound as:

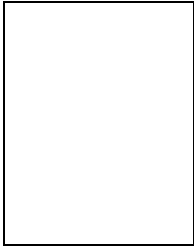
$$\begin{aligned} \mathcal{C}_{\text{Sato}}(P, \mathbf{H}) &= \inf_{\mathbf{Z} \in \mathcal{S}} \min_{\mathbf{A}, \lambda} -\log |\mathbf{A}| + \text{Tr}(\mathbf{A}) + \lambda P - t \\ & \quad \text{such that} \quad (58) \\ & \quad \lambda \geq 0, \mathbf{A} \geq 0 \\ & \quad \lambda \mathbf{Z} \geq \mathbf{H} \mathbf{A} \mathbf{H}^\dagger. \end{aligned}$$

I. ACKNOWLEDGMENT

The authors wish to thank Syed A. Jafar for helpful discussions regarding singular noise covariances.

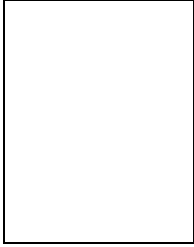
REFERENCES

- [1] G.J. Foschini and M.J. Gans, "On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas", *Wireless Personal Commun.*, vol 6.,311-335, 1998.
- [2] E. Telatar, "Capacity of Multi-antenna Gaussian Channels,"*European Trans. on Telecomm. (ETT)*, 10(6):585-596, November 1999.
- [3] S. Verdú, "Multiple-access Channels with Memory With and Without Frame Synchronization,"*IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 605-19, May 1989.
- [4] W. Yu, W. Rhee, S. Boyd, J. Cioffi, "Iterative Water-filling for Vector Multiple Access Channels", pp. 322, *Proc. IEEE Int. Symp. Inf. Theory, (ISIT)*, Washington DC, June 24-29, 2001.
- [5] G. Caire and S. Shamai, "On Achievable Rates in a Multi-antenna Broadcast Downlink", *38th Annual Allerton Conference on Commun., Control and Computing*, Monticello, IL, Oct. 4 - 6, 2000.
- [6] M. Costa, "Writing on Dirty Paper", *IEEE. Trans. Inf. Theory*, v. 29, pp. 439-441, May 1983.
- [7] G. Caire and S. Shamai, "On the Achievable Throughput of a Multi-antenna Gaussian Broadcast Channel", submitted to *IEEE Trans. on Inform. Theory*, July 2001.
- [8] W. Yu and J. Cioffi, "Trellis Precoding for the Broadcast Channel", *Proc. IEEE Global Tele. Conf. (Globecom)*, pp. 1344-1338, Nov., 2001.
- [9] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the Duality of Multiple-Access and Broadcast Channels", *Allerton Conference on Commun., Control, and Computing*, Oct. 3-5, 2001. Full version submitted to *IEEE Transactions on Information Theory* in August, 2002.
- [10] H. Sato, "An Outer Bound on the Capacity Region of Broadcast Channel", *IEEE. Trans. Inf. Theory*, v. 24, pp. 374-377, May 1978.
- [11] W. Yu and J. Cioffi, "Sum Capacity of a Gaussian Vector Broadcast Channel", *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 498, Lausanne, Switzerland, Jun. 30- Jul. 5, 2002.
- [12] P. Viswanath and D. Tse, "Sum Capacity of the Multi-Antenna Broadcast Channel", *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 497, Lausanne, Switzerland, Jun. 30- Jul. 5, 2002. Also submitted to *IEEE Transactions on Information Theory*.
- [13] S. Vishwanath, G. Kramer, S. Shamai, S. Jafar and A. Goldsmith, "Outer Bounds for Multi-Antenna Broadcast Channels", *Proc. DIMACS Workshop on Signal Processing for Wireless Transmission*, Rutgers, NJ, Oct. 2002.
- [14] D. Tse and P. Viswanath, "On the Capacity of the Multiple Antenna Broadcast Channel" *Proc. DIMACS Workshop on Signal Processing for Wireless Transmission*, Rutgers, NJ, Oct. 2002.
- [15] E. Jorswieck and H. Boche, "Rate balancing for the multi-antenna Gaussian broadcast channel," *Proc. of IEEE Intern. Symposium on Spread Spectrum Techniques and Applic.*, 2002, pg. 545-549.
- [16] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
- [17] L. Vandenberghe, S. Boyd and S.P. Wu "Determinant maximization with linear matrix inequality constraints", *SIAM Journal on Matrix Anal. and Appl.*, 19(2):499-533, 1998.
- [18] S. Vishwanath, N. Jindal and A. Goldsmith. "On the Capacity of Multiple Input Multiple Output Broadcast Channels", *Proc. IEEE Int. Conf. Commun.*, New York, NY April 2002.
- [19] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the Duality of Multiple-Access and Broadcast Channels", *Proc. IEEE Int. Symp. Inf. Theory*, pp. 500, Lausanne, Switzerland, Jun. 30- Jul. 5, 2002.

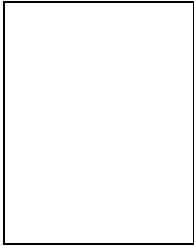


Sriram Vishwanath received his B.Tech. in Electrical Engineering from the Indian Institute of Technology (IIT) Madras in 1998, and an M.S. in Electrical Engineering from the California Institute of Technology, Pasadena in 1999. He is currently a Stanford graduate fellow working towards a Ph.D. in Electrical Engineering. His research interests include information and coding theory, with a focus on multiple antenna systems. His industry experience includes work at National Semiconductor Corporation, Santa Clara, CA, in the summer of 2000 and at the Lucent Bell

Labs, Murray Hill, NJ, during the summer of 2002.



Nihar Jindal received the B.S. degree in Electrical Engineering and Computer Science from U.C. Berkeley in 1999 and the M.S. degree in Electrical Engineering from Stanford University in 2001. He is currently a PhD student at Stanford University. His industry experience includes summer internships at Intel Corporation, Santa Clara, CA in 2000 and at Lucent Bell Labs, Holmdel, NJ in 2002. His research interests include multiple-antenna channels and multi-user information theory and their applications to wireless communication.



Andrea Goldsmith received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from U.C. Berkeley in 1986, 1991, and 1994, respectively. She was an Assistant Professor in the department of Electrical Engineering at Caltech from 1994-1999. In 1999 she joined the Electrical Engineering department at Stanford University, where she is currently an Associate Professor. Her industry experience includes affiliation with Maxim Technologies from 1986-1990, where she worked on packet radio and satellite communication systems, and with AT&T Bell Laboratories from 1991-1992, where she worked on microcell modeling and channel estimation. Her research includes work in capacity of wireless channels and networks, wireless information and communication theory, multiantenna systems, joint source and channel coding, cross-layer wireless network design, communications for distributed control, and adaptive resource allocation for cellular systems and ad-hoc wireless networks.

Dr. Goldsmith is a Terman Faculty Fellow at Stanford and a recipient of the Alfred P. Sloan Fellowship, the National Academy of Engineering Gilbreth Lectureship, a National Science Foundation CAREER Development Award, the Office of Naval Research Young Investigator Award, a National Semiconductor Faculty Development Award, an Okawa Foundation Award, and the David Griep Memorial Prize from U.C. Berkeley. She was an editor for the IEEE Transactions on Communications from 1995-2002, and has been an editor for the IEEE Wireless Communications Magazine since 1995. She is also an elected member of Stanford's faculty senate and the board of governors for the IEEE Information Theory Society.