On the Capacity of Multiple Input Multiple Output Broadcast Channels

Sriram Vishwanath, Nihar Jindal, and Andrea Goldsmith
Department of Electrical Engineering, Stanford University, Stanford, CA 94305

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. The duality also enables us to translate previously known results for the MIMO MAC to 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 Telatar [3] and Foschini [4]. This work has also been extended to the MIMO multiple-access channel (MAC) [2, 3, 15]. 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 [8] at the transmitter (also known as coding for non-causally known interference). It was also shown in [5, 6] that the sum rate MIMO BC capacity equals the maximum sum rate of this achievable region for the two user broadcast channel with two transmit antennas ($t = 2$) and one receive antenna at each receiver ($r_1 = r_2 = 1$). However, computing this region is extremely complex and this approach does not appear to work for the more general class of channels which we consider.

In this paper, we consider a K-user MIMO Gaussian BC in which receiver j has r_j receive antennas and the transmitter has t 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 [7]. 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 [1]. 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.

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 same sum rate as the actual MIMO BC capacity region through the use of the Sato upper bound on the sum-rate capacity of broadcast channels [9]. The same upper bound is used in [5] to find the sum rate capacity of the ($t = 2, r_1 = r_2 = 1$) channel. We upper 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 that this upper bound coincides exactly with the maximum sum rate in the dirty paper region.

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. However, this hypothesis has remains to be proved.

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 and show that the dirty paper BC region achieves sum rate MIMO BC capacity in Section V. We conclude with Section VI.

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 $\mathbf{M}$, let $\mathbf{M}^{\dagger}$ denote the conjugate transpose, and Tr($\mathbf{M}$) the trace. $\mathbf{I}$ denotes the identity matrix and diag(λ_i) denotes a diagonal matrix with the (i, i) entry equaling λ_i.

We consider a MIMO broadcast channel with a single t-antenna transmitter and K receivers with $r_1, \ldots, r_K$ 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

$$\begin{bmatrix} \mathbf{y_1} \\ \vdots \\ \mathbf{y_K} \end{bmatrix} = \mathbf{Hx} + \begin{bmatrix} \mathbf{n_1} \\ \vdots \\ \mathbf{n_K} \end{bmatrix} \quad \text{where} \quad \mathbf{H} = \begin{bmatrix} \mathbf{H_1} \\ \vdots \\ \mathbf{H_K} \end{bmatrix} \tag{1}$$

The matrix $\mathbf{H}$ represents the channel gains of all receivers. The

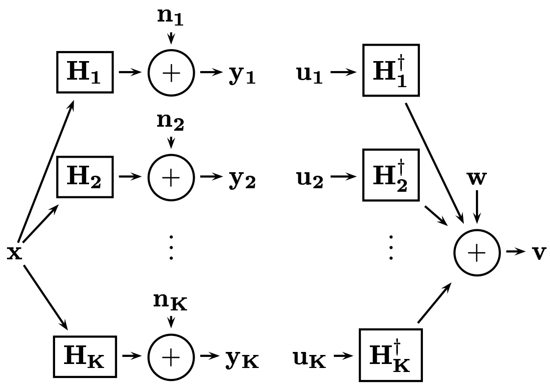

Fig. 1. System models of the BC MIMO(left) and the MAC MIMO (right) channels

covariance matrix of the input signal is $\mathbf{\Sigma}_x \triangleq \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. The transmitter is subject to an average power constraint P, which implies $\text{Tr}(\mathbf{\Sigma}_x) \leq P$. 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

$$\begin{aligned} \mathbf{v} &= \mathbf{H_1}^\dagger \mathbf{u_1} + \ldots + \mathbf{H}_K^\dagger \mathbf{u}_K + \mathbf{w} \\ &= \mathbf{H}^\dagger \begin{bmatrix} \mathbf{u_1} \\ \vdots \\ \mathbf{u}_K \end{bmatrix} + \mathbf{w} \quad \text{where} \quad \mathbf{H}^\dagger = \begin{bmatrix} \mathbf{H_1}^\dagger & \ldots & \mathbf{H}_K^\dagger \end{bmatrix}. \end{aligned}$$

In the dual MAC, each transmitter is subject to an individual power constraint of $P_1, \ldots, P_K$, 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

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{2}$$

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 $\mathcal{C}_{\text{BC}}(P, \mathbf{H})$, $\mathcal{C}_{\text{MAC}}(P_1, \ldots, P_K, \mathbf{H}^\dagger)$ 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, 7] and results on the MIMO MAC capacity region [2,3,15] extensively in this paper. Hence, we summarize these results first, and then state results on the duality of the scalar Gaussian BC and MAC [1].

A. Achievable BC Region - The Dirty Paper Region

An achievable region for the MIMO BC was first obtained in [5]. In [7], the region was extended to the more general multiple-user, multiple-antenna case using the following extension of the "dirty paper result" [8] 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. Since the ordering of the users clearly matters in such a procedure, the following is an achievable set of rates

$$R_{\pi(i)} = \frac{1}{2}\log\frac{|\mathbf{I} + \mathbf{H}_{\pi(i)}(\sum_{j \geq i}\mathbf{\Sigma}_{\pi(j)})\mathbf{H}_{\pi(i)}^\dagger|}{|\mathbf{I} + \mathbf{H}_{\pi(i)}(\sum_{j > i}\mathbf{\Sigma}_{\pi(j)})\mathbf{H}_{\pi(i)}^\dagger|} \quad i = 1, \ldots, K.$$

The dirty-paper region $\mathcal{C}_{\text{dirtypaper}}(P, \mathbf{H})$ is defined as the union of all such rates vectors over all covariance matrices $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_K$ such that $\text{Tr}(\mathbf{\Sigma}_1 + \ldots \mathbf{\Sigma}_K) = \text{Tr}(\mathbf{\Sigma}_x) \leq P$ and over all permutations $(\pi(1), \ldots, \pi(K))$. The transmitted signal is $\mathbf{x} = \mathbf{x_1} + \ldots + \mathbf{x_K}$ and the input covariance matrices are of the form $\mathbf{\Sigma}_i = \mathbb{E}[\mathbf{x_i}\mathbf{x_i}^\dagger]$.

B. MIMO MAC Capacity Region

The capacity region of a general MIMO MAC was obtained in [2, 3, 15] We now describe this capacity region for the dual MIMO MAC as defined in Section II. For any set of powers $(P_1, \ldots, P_K)$, the capacity of the MIMO MAC is

$$\mathcal{C}_{\text{MAC}}(P_1, \ldots, P_K; \mathbf{H}^\dagger) \triangleq \bigcup_{\{\text{Tr}(\mathbf{P_1}) \leq P_i \;\; \forall i\}} \Big\{ (R_1, \ldots, R_K) : \sum_{i \in S} R_i \leq \frac{1}{2}\log|\mathbf{I} + \sum_{i \in S}\mathbf{H}_i^\dagger P_i \mathbf{H}_i| \;\; \forall S \subseteq \{1, \ldots, M\} \Big\}$$

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

$$\mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger) = \bigcup_{\sum_{i=0}^K P_i \leq P} \mathcal{C}_{\text{MAC}}(P_1, \ldots, P_K; \mathbf{H}^\dagger). \tag{3}$$

It can be easily 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.

C. Duality of the Scalar Gaussian MAC and BC

Lastly, we state the duality result for scalar Gaussian MAC and BC channels [1].

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

$$\mathcal{C}_{\text{BC}}(P; \bar{h}) = \bigcup_{\sum_{i=1}^{K} P_i = P} \mathcal{C}_{\text{MAC}}(P_1, \ldots, P_K; \bar{h}). \quad (4)$$

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 total 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 $(P_1, \ldots, P_K)$ 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 the capacity region of the dual MIMO MAC with sum power constraint P.

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

Proof: We prove this by showing that for every set of MAC covariance matrices $\mathbf{P}_1, \ldots, \mathbf{P}_K$ and the corresponding set of MAC rates, there exist BC covariance matrices $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_K$ using the same sum power as the MAC (i.e. $\sum_{i=1}^{K}$ Tr $(\mathbf{P}_i)$ = $\sum_{i=1}^{K}$ Tr $(\mathbf{\Sigma}_i)$) such that the MAC rates are achievable in the BC using the dirty paper coding method described in Section III-A. We also show that the inverse of this statement holds true, or that for every set of BC covariance matrices there exist MAC covariance matrices that achieve the same set of rates using the same sum power. It is important to point out that we reverse the decoding/encoding order of the users in the dual MAC/BC channel. In other words, if User 1 is decoded first in the MAC, then we must encode User 1's signal last in the BC to achieve the same rates using these transformations. This completes the proof, provided we have the appropriate transformations 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.

A. Terminology

First, we explain the terms *effective channel* and *flipped channel*. A single user MIMO system $\mathfrak{S}$ 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 $\mathfrak{S}$ and a different system with channel matrix equal to the effective channel, additive white noise of unit variance, and no interference are the same. Also, the capacity of a system $\mathfrak{S}_1$ with effective channel matrix $\mathbf{Y}$ and the capacity of system $\mathfrak{S}_2$ with effective channel matrix $\mathbf{Y}^\dagger$, termed the flipped channel, are the same [3]. In other words, for every transmit covariance $\mathbf{\Sigma}$ in $\mathfrak{S}_1$, there exists a $\overline{\mathbf{\Sigma}}$ in $\mathfrak{S}_2$ with $\text{Tr}(\overline{\mathbf{\Sigma}}) = \text{Tr}(\mathbf{\Sigma})$ such that the rate achieved by $\overline{\mathbf{\Sigma}}$ in $\mathfrak{S}_2$ is equal to the rate achieved by $\mathbf{\Sigma}$ in $\mathfrak{S}_1$. It can easily be shown that $\overline{\mathbf{\Sigma}} = \mathbf{F}\mathbf{G}^\dagger\mathbf{\Sigma}\mathbf{G}\mathbf{F}^\dagger$ meets this criterion where the SVD of $\mathbf{Y}$ is $\mathbf{Y} = \mathbf{F}\mathbf{\Lambda}\mathbf{G}^\dagger$. Next, we describe the transformations.

B. MAC to BC Transformation

In this section we prove the existence of BC covariances that achieve the same rates as a set of MAC covariances and that use the same sum power.

Since the numbering of the users is arbitrary and because successive decoding at the receiver is known to be optimal for the MAC, we assume that User 1 is decoded first, User 2 second, and so on at the receiver.

Let $\mathbf{A}_j \triangleq (\mathbf{I} + \mathbf{H}_j(\sum_{l=1}^{j-1} \mathbf{\Sigma}_l)\mathbf{H}_j^\dagger)$ and $\mathbf{B}_j \triangleq (\mathbf{I} + \sum_{l=j+1}^{K} \mathbf{H}_l^\dagger \mathbf{P}_l \mathbf{H}_l)$. The rate achieved by User j in the MAC is given by

$$\begin{aligned} R_j^{\text{M}} &= \frac{1}{2} \log \frac{|\mathbf{I} + \sum_{i=j}^{K}(\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)|}{|\mathbf{I} + \sum_{i=j+1}^{K}(\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i)|} \\ &= \frac{1}{2} \log |I + (\mathbf{I} + \sum_{i=j+1}^{K} (\mathbf{H}_i^\dagger \mathbf{P}_i \mathbf{H}_i))^{-1} \mathbf{H}_j^\dagger \mathbf{P}_j \mathbf{H}_j| \\ &= \frac{1}{2} \log |I + \mathbf{B}_j^{-1} \mathbf{H}_j^\dagger \mathbf{P}_j \mathbf{H}_j|. \end{aligned}$$

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

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

Treating $\mathbf{B}_j^{-1/2}\mathbf{H}_j^\dagger\mathbf{A}_j^{-1/2}$ as the effective channel of the system, we flip the channel and find $\overline{\mathbf{A}_j^{1/2}\mathbf{P}_j\mathbf{A}_j^{1/2}}$ such that

$$\text{Tr}(\overline{\mathbf{A}_j^{1/2}\mathbf{P}_j\mathbf{A}_j^{1/2}}) = \text{Tr}(\mathbf{A}_j^{1/2}\mathbf{P}_j\mathbf{A}_j^{1/2})$$

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

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)

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

If we choose $\mathbf{\Sigma}_j$ as

$$\mathbf{\Sigma}_j = \mathbf{B}_j^{-1/2}\overline{\mathbf{A}_j^{1/2}\mathbf{P}_j\mathbf{A}_j^{1/2}}\mathbf{B}_j^{-1/2} \tag{5}$$

clearly we see $R_j^{\mathrm{M}} = R_j^{\mathrm{B}}$. By doing this for all K users, we find covariance matrices for the BC that achieve the same rate as in the MAC. In Appendix A, we show that the transformations given in (5) satisfy the sum trace constraint, or that $\sum_{i=1}^{K}$ Tr $(\mathbf{P}_i)$ = $\sum_{i=1}^{K}$ Tr $(\mathbf{\Sigma}_i)$.

C. BC to MAC transformation

In this section we prove the existence of MAC covariances that achieve the same rates as a set of BC covariances and that use the same sum power. For the dirty paper encoding at the BC, we assume that User K is encoded first, $K-1$ 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 $\mathbf{B}_j^{1/2}\mathbf{\Sigma}_j\mathbf{B}_j^{1/2}$ as the covariance matrix. By flipping the effective channel, we obtain $\overline{\mathbf{B}_j^{1/2}\mathbf{\Sigma}_j\mathbf{B}_j^{1/2}}$ and obtain the transformation

$$\mathbf{P}_j = \mathbf{A}_j^{-1/2}\overline{\mathbf{B}_j^{1/2}\mathbf{\Sigma}_j\mathbf{B}_j^{1/2}}\mathbf{A}_j^{-1/2} \tag{6}$$

As before, if we use the opposite decoding order in the MAC, this transformation ensures that the rates of all users in the BC and MAC are equivalent as is the total power used in the BC and MAC. ■

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. We use the superscript "sumrate" to denote the maximum sum rate of the rate region under consideration.

In [9], 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 ($\mathcal{C}_{\mathrm{coop}}(P,\mathbf{H})$) is an upper bound on the BC sum rate capacity ($\mathcal{C}_{\mathrm{BC}}^{\mathrm{sumrate}}(P,\mathbf{H})$). We now show that the bound can be tightened by introducing noise correlation.

Since the capacity region of a general BC depends only on the marginal transition probabilities (i.e. $p(y_i|x)$) and not on the entire joint distribution $p(y_1,\ldots,y_K|x)$, we can introduce correlation between the noise vectors at different receivers of the BC without affecting the BC capacity region. This correlation 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 $\mathbb{E}(\mathbf{n}_i\mathbf{n}_i^\dagger) = \mathbf{I}$, $i=1,\ldots,K$ to maintain the same marginal transition probabilities, but we let $\mathbb{E}(\mathbf{n_i}\mathbf{n_j}^\dagger) \triangleq \mathbf{X}_{i,j} \prec \mathbf{I}$ for $i \neq j$. By searching over all feasible positive definite noise covariance matrices, we get the following bound

$$\mathcal{C}_{\mathrm{BC}}^{\mathrm{sumrate}}(P,\mathbf{H}) \leq \min_{S}\mathcal{C}_{\mathrm{coop}}(P,\mathbf{Q}_K^{-1/2}\mathbf{H})$$

where

$$S = \left\{\mathbf{Q}_K : \mathbf{Q}_K > 0, \mathbf{Q}_K = \begin{bmatrix} \mathbf{I} & \ldots & \mathbf{X}_{K,1}^\dagger \\ \vdots & \vdots & \vdots \\ \mathbf{X}_{K,1} & \ldots & \mathbf{I} \end{bmatrix}\right\}.$$

By using the fact that the capacity region of the dual MIMO MAC equals the capacity region of the dirty paper region of the MIMO BC (and therefore the maximum sum rate of the MAC and the dirty paper region are equivalent), we are able to show that this bound is tight for the BC MIMO.

Theorem 3: The maximum sum rate in the dirty paper region of the MIMO BC equals the Sato upper bound on the sum rate capacity of broadcast channels and is therefore the actual sum rate capacity of the MIMO BC.

$$\begin{aligned} \min_{\{\mathbf{Q}_K\in S\}}\mathcal{C}_{\mathrm{coop}}(P,\mathbf{Q}_K^{-1/2}\mathbf{H}) &= \mathcal{C}_{\mathrm{dirtypaper}}^{\mathrm{sumrate}}(P,\mathbf{H}) \\ &= \mathcal{C}_{\mathrm{BC}}^{\mathrm{sumrate}}(P,\mathbf{H}) \end{aligned}$$

Proof: We will prove that the maximum sum rate in the dual MIMO MAC capacity region with a sum power constraint P is equal to the Sato upper bound. Due to the duality of the MAC and BC, the maximum sum rate in the dirty paper region also equals the Sato upper bound, which implies that the dirty paper region achieves the sum rate capacity of the MIMO BC.

We now show that the capacity of the cooperative system with worst case correlated noise is equal to the maximum sum rate capacity of the MIMO MAC, or that

$$\begin{aligned} &\min_{\mathbf{Q}_K\in S}\max_{\{\mathbf{\Sigma}_{\mathrm{coop}}:\mathrm{Tr}(\mathbf{\Sigma}_{\mathrm{coop}})\leq P\}}\frac{1}{2}\log\frac{|\mathbf{Q}_K+\mathbf{H}\mathbf{\Sigma}_{\mathrm{coop}}\mathbf{H}^\dagger|}{|\mathbf{Q}_K|} \\ &= \max_{\{\mathbf{P}_j:0\leq\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{P}_i)\leq P\}}\frac{1}{2}\log|\mathbf{I}+\sum_{i=1}^{K}\mathbf{H}_i^\dagger\mathbf{P}_i\mathbf{H}_i|. \end{aligned}$$

Note that the capacity of a cooperative system with channel $\mathbf{H}$ and noise covariance $\mathbf{Q}_K$ can be rewritten as

$$\begin{aligned} &\max_{\{\mathbf{\Sigma}_{\mathrm{coop}}:\mathrm{Tr}(\mathbf{\Sigma}_{\mathrm{coop}})\leq P\}}\frac{1}{2}\log\frac{|\mathbf{Q}_K+\mathbf{H}\mathbf{\Sigma}_{\mathrm{coop}}\mathbf{H}^\dagger|}{|\mathbf{Q}_K|} = \\ &\max_{\{\mathbf{\Sigma}_{\mathrm{coop}}:\mathrm{Tr}(\mathbf{\Sigma}_{\mathrm{coop}})\leq P\}}\frac{1}{2}\log|\mathbf{I}+\mathbf{Q}_K^{-1/2}\mathbf{H}\mathbf{\Sigma}_{\mathrm{coop}}\mathbf{H}^\dagger(\mathbf{Q}_K^{-1/2})^\dagger| \end{aligned}$$

where $\mathbf{Q}_K^{-1/2}$ is a matrix such that $(\mathbf{Q}_K^{-1/2})^\dagger\mathbf{Q}_K^{-1/2} = \mathbf{Q}_K^{-1}$. We obtain the flipped channel $\mathbf{H}^\dagger(\mathbf{Q}_K^{-1/2})^\dagger$ from the effective channel $\mathbf{Q}_K^{-1/2}\mathbf{H}$ and a covariance matrix $\overline{\mathbf{\Sigma}_{\text{coop}}} = \mathbf{\Delta}_K$ such that

$$\begin{aligned} & \mathcal{C}_{\text{coop}}(P, \mathbf{Q}_K^{-1/2}\mathbf{H}) \qquad (7) \\ = \; & \max_{\mathbf{\Delta}_K \in T} \frac{1}{2} \log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{-1/2}\mathbf{H}| \end{aligned}$$

where T is the set $\{\mathbf{\Delta}_K : \text{Tr}(\mathbf{\Delta}_K) \le P\}$. Note that

$$\begin{aligned} C_{\text{sato}} \triangleq \; & \min_{\mathbf{Q}_K \in S} \max_{\mathbf{\Delta}_K \in T} \log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{-1/2}\mathbf{H}| \\ = \; & \max_{\mathbf{\Delta}_K \in T} \min_{\mathbf{Q}_K \in S} \log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{-1/2}\mathbf{H}| \end{aligned} \tag{8}$$

i.e., the order in which maximization and minimization is done can be interchanged. We use the max-min version of the bound in the proof. Equation (8) can be proved using Ky Fan's Theorem [11], which is an extension of von Neumann's Minimax Theorem [12]. This theorem can also be found in [13, Page 11]. The conditions required for interchangeability are
1. The sets S and T are compact and convex.
2. The function $\log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{-1/2}\mathbf{H}|$ is continuous and convex in $\mathbf{Q}_K$ and continuous and concave in $\mathbf{\Delta}_K$.
These conditions can be easily verified. For a proof of the second condition and a proof of interchangeability of maximum and minimum for a case similar to this, see [14].

Note that the sum rate capacity of the MAC channel $\mathcal{C}_{\text{union}}^{\text{sumrate}}$ for given covariance matrices $\mathbf{P}_1, \ldots, \mathbf{P}_K$ can be rewritten as

$$\frac{1}{2} \log \left| \mathbf{I} + \mathbf{H}^\dagger \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{P}_K \end{bmatrix} \mathbf{H} \right| \tag{9}$$

In Appendix B, we show that a $\mathbf{Q}_K^{0\;1/2} \in S$ can be found such that

$$\begin{bmatrix} \mathbf{P}'_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{P}'_K \end{bmatrix} = (\mathbf{Q}_K^{0\;-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{0\;-1/2} \tag{10}$$

for every $\mathbf{\Delta}_K \in T$. Here, $\mathbf{P}'_1, \ldots, \mathbf{P}'_K$ are symmetric positive definite matrices that satisfy

$$\sum_{i=1}^{K} \text{Tr}(\mathbf{P}'_i) = P \tag{11}$$

Substituting (10) in the expression for $\mathcal{C}_{\text{sato}}$ (8) we get

$$\begin{aligned} \mathcal{C}_{\text{sato}} &= \max_{\mathbf{\Delta}_K \in T} \min_{\mathbf{Q}_K \in S} \log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{-1/2}\mathbf{H}| \\ &\le \max_{\mathbf{\Delta}_K \in T} \log |\mathbf{I} + \mathbf{H}^\dagger (\mathbf{Q}_K^{0\;-1/2})^\dagger \mathbf{\Delta}_K \mathbf{Q}_K^{0\;-1/2}\mathbf{H}| \\ &= \max_{T} \frac{1}{2} \log \left| \mathbf{I} + \mathbf{H}^\dagger \begin{bmatrix} \mathbf{P}'_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{P}'_K \end{bmatrix} \mathbf{H} \right| \\ &\le \text{Sum Rate Capacity of MAC} \end{aligned}$$

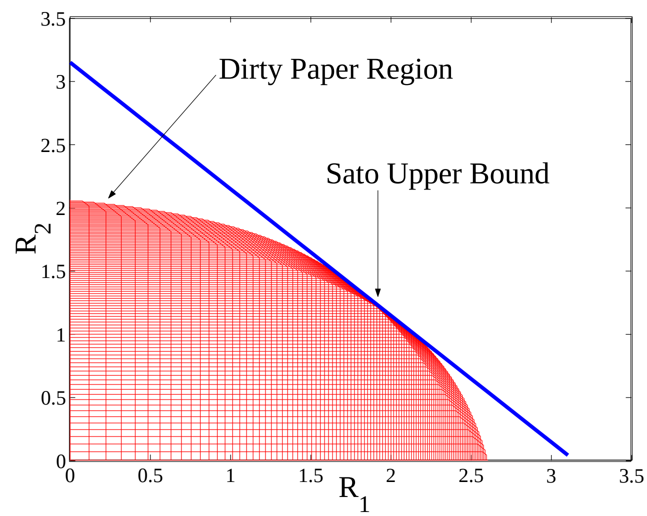

Fig. 2. Dirty paper broadcast region - achievability and converse

Since, from duality, we have that the sum rate capacity for the MAC is less than or equal to the cooperative capacity bound, we have the result. We prove Equations (10) and (11) in Appendix B. ■

The dirty paper BC region and the capacity regions of the dual MAC, along with the Sato upper bound are illustrated for a two receiver BC in which the transmitter has two antennas and each receiver has a single antenna ($K = 2, t = 2, r_1 = r_2 = 1$) in Fig. 2. The channel matrix is

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0.5} \\ \mathbf{0.2} & \mathbf{0.8} \end{bmatrix}$$

and the power constraint is 10. Notice that the dirty paper region is the union of the dual MAC capacity regions. The boundary of the dirty paper region is a straight line segment at the sum rate point, where the region touches the Sato upper bound.

This region is significantly different than the capacity region of a scalar BC. In a scalar BC, the capacity region boundary does not contain line segments. Additionally, the sum rate point in a scalar BC is achieved by allocating all power to one user (the user with the largest channel gain). In the example MIMO channel, however, notice that the sum rate point is actually achieved by allocating power to both users. However, the shape of the dirty paper region is very similar to the shape of the MIMO MAC capacity region when the transmitters have more than one antenna [2].

There are two important implications of Theorems 2 and 3. First, it is easy to verify that the sum power constraint MAC capacity region $\mathcal{C}_{\text{union}}(P, \mathbf{H}^\dagger)$ is convex due to a time-sharing argument, and hence the dirty paper region is also convex. Additionally, since the dual MAC capacity region can be posed as a convex optimization problem, the equivalence of the dirty paper region and the dual MAC allows us to use convex optimization techniques (such as the interior point method in [16]) to calculate the entire dirty paper region and techniques such as iterative water-filling [7] to calculate the sum rate capacity of the MIMO BC.

VI. CONCLUSION

In this paper, we established a duality relationship between two seemingly different regions : the achievable region of the MIMO BC obtained using the "dirty paper" coding strategy and the capacity region of the MIMO MAC. This duality makes the previously intractable problem of finding this "dirty paper" achievable region much easier to solve. Though the capacity region of the MIMO BC is unknown due to its non-degraded nature, we were able to show that, for sum rate, the boundary of the"dirty paper" achievable region and the MIMO BC capacity region are the same. These results open up the possibility that the dirty paper region is the actual capacity region of the BC MIMO and also the possibility that other instances of "duality" exist in multi-terminal networks.

APPENDIX

A. Proof of Trace Constraint for Covariance Transformations

In this section, we show that the transformations obtained in (5) satisfy the sum trace requirement. Throughout this proof, we utilize the linearity of the trace and the property $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ when both $\mathbf{AB}$ and $\mathbf{BA}$ are defined. First, we compute

$$\begin{aligned}\mathrm{Tr}(\mathbf{\Sigma}_K) &= \mathrm{Tr}(\mathbf{A}_K\mathbf{P}_K)\\ &= \mathrm{Tr}(\mathbf{P}_K) + \sum_{i=1}^{K-1}\mathrm{Tr}(\mathbf{\Sigma}_i\mathbf{H}_K^\dagger\mathbf{P}_K\mathbf{H}_K)\end{aligned}$$

By adding $\sum_{i=1}^{K-1}\mathrm{Tr}(\mathbf{\Sigma}_i)$ to both sides, we get

$$\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{\Sigma}_i) = \mathrm{Tr}(\mathbf{P}_K) + \sum_{i=1}^{K-1}\mathrm{Tr}(\mathbf{\Sigma}_i(\mathbf{I}+\mathbf{H}_K^\dagger\mathbf{P}_K\mathbf{H}_K))$$

Note that

$$\begin{aligned}&\mathrm{Tr}(\mathbf{\Sigma}_j(\mathbf{I}+\sum_{i=j+1}^{K}\mathbf{H}_i^\dagger\mathbf{P}_i\mathbf{H}_i))\\ &= \mathrm{Tr}(\mathbf{\Sigma}_j\mathbf{B}_j)\\ &= \mathrm{Tr}(\mathbf{B}_j^{-1/2}\overline{\mathbf{A}_j^{1/2}\mathbf{P}_j\mathbf{A}_j^{1/2}}\mathbf{B}_j^{1/2})\\ &= \mathrm{Tr}(\mathbf{A}_j\mathbf{P}_j)\\ &= \mathrm{Tr}(\mathbf{P}_j) + \sum_{i=1}^{j-1}\mathrm{Tr}(\mathbf{\Sigma}_i\mathbf{H}_j^\dagger\mathbf{P}_j\mathbf{H}_j) \qquad (12)\end{aligned}$$

Hence

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

Assume the inductive hypothesis

$$\begin{aligned}&\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{\Sigma}_i)\\ &= \sum_{l=j+1}^{K}\mathrm{Tr}(\mathbf{P}_l) + \sum_{i=1}^{j}\mathrm{Tr}(\mathbf{\Sigma}_i(\mathbf{I}+\sum_{l=j+1}^{K}\mathbf{H}_l^\dagger\mathbf{P}_l\mathbf{H}_l)).\end{aligned}$$

By (12), we get

$$\begin{aligned}&\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{\Sigma}_i)\\ &= \sum_{l=j}^{K}\mathrm{Tr}(\mathbf{P}_l) + \sum_{i=1}^{j-1}\mathrm{Tr}(\mathbf{\Sigma}_i(\mathbf{I}+\sum_{l=j}^{K}\mathbf{H}_l^\dagger\mathbf{P}_l\mathbf{H}_l)).\end{aligned}$$

For $j=1$, we have

$$\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{\Sigma}_i) = \sum_{l=1}^{K}\mathrm{Tr}(\mathbf{P}_l).$$

A similar method can be used to show that (6) also satisfy the trace constraints.

B. Proof of Equation (10)

We prove (10) by recursive reduction. We assume that the result holds for a $K-1$ user system and prove it for the K user one. Note that we can rewrite $\mathbf{Q}_K^0 \in S$ as

$$\mathbf{Q}_K^0 = \mathbf{Q}_K^{0\;1/2}(\mathbf{Q}_K^{0\;1/2})^\dagger = \begin{bmatrix}\mathbf{I} & \mathbf{X}^\dagger\\ \mathbf{X} & \mathbf{Q}_{K-1}^0\end{bmatrix} \qquad (13)$$

where, using the notation of Section III-*A*,

$$\mathbf{X}^\dagger = [\mathbf{X}_{2,1}^\dagger, \ldots, \mathbf{X}_{K,1}^\dagger]$$

and

$$\mathbf{Q}_{K-1}^0 = \begin{bmatrix}\mathbf{I} & \mathbf{X}_{3,2}^\dagger & \ldots & \mathbf{X}_{K,2}^\dagger\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{X}_{K,2} & \ldots & \mathbf{X}_{K,K-1} & \mathbf{I}\end{bmatrix}.$$

$\mathbf{Q}_K^0$ also must satisfy

$$\mathbf{\Delta}_K = (\mathbf{Q}_K^{0\;1/2})^\dagger\begin{bmatrix}\mathbf{P}_1' & \mathbf{0}\\ \mathbf{0} & \mathbf{D}_{K-1}\end{bmatrix}\mathbf{Q}_K^{0\;1/2} \qquad (14)$$

where $\mathbf{P}_1'$ and $\mathbf{D}_{K-1}$ are symmetric positive semidefinite matrices that satisfy

$$\mathrm{Tr}(\mathbf{P}_1') + \mathrm{Tr}(\mathbf{D}_{K-1}) = P \qquad (15)$$

To prove this, we invoke the block LDU factorization [10] of $\mathbf{\Delta}_K$, which is given by

$$\mathbf{\Delta}_K = \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{B} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_{K-1}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{B}^\dagger\\ \mathbf{0} & \mathbf{I}\end{bmatrix} \qquad (16)$$

where $\mathbf{A}$ and $\mathbf{\Delta}_{K-1}$ are symmetric positive semidefinite matrices. We pick $\mathbf{Q}_K^{0\;1/2}$ to have the structure

$$\mathbf{Q}_K^{0\;1/2} = \begin{bmatrix}(\mathbf{I}-\mathbf{X}^\dagger\mathbf{Q}_{K-1}^{0\;-1}\mathbf{X})^{1/2} & \mathbf{X}^\dagger((\mathbf{Q}_{K-1}^0)^{-1/2})^\dagger\\ \mathbf{0} & (\mathbf{Q}_{K-1}^0)^{1/2}\end{bmatrix} \qquad (17)$$

where $(\mathbf{I}-\mathbf{X}^\dagger \mathbf{Q}^0_{K-1}{}^{-1}\mathbf{X})^{1/2}$ is the symmetric square root of $(\mathbf{I}-\mathbf{X}^\dagger \mathbf{Q}^0_{K-1}{}^{-1}\mathbf{X})$ and $\mathbf{Q}^0_{K-1}{}^{-1/2}$ is a block upper triangular square root of $\mathbf{Q}^0_{K-1}$. Note that this choice of $\mathbf{Q}^{1/2}_K$ satisfies the requirement in (13) above. Also note that $\mathbf{Q}^0_{K-1}$ is positive definite and hence invertible. This is due to the recursive construction methodology employed in this proof and the fact that $\mathbf{Q}^0_1 = \mathbf{I}$.

Therefore, a sufficient condition for $\mathbf{Q}^0_K{}^{1/2}$ to be full rank, and hence for $\mathbf{Q}^0_K{}^{1/2}(\mathbf{Q}^0_K{}^{1/2})^\dagger$ to be positive definite is that the spectral radius of $\mathbf{Q}^0{}^{-1/2}_{K-1}\mathbf{X}$ is less than unity. Let us suppose the optimum $\mathbf{Q}^0{}^{-1/2}_{K-1}$ is known. Now, we construct $(\mathbf{Q}^0{}_{K-1})^{-1/2}\mathbf{X}$ as follows.

Let the singular value decomposition (SVD) of $\mathbf{B}$ equal $\mathbf{U\Phi V}$, where $\mathbf{\Phi} = \text{diag}(\phi_i)$ where $\phi_i \geq 0$. We choose $(\mathbf{Q}^0{}_{K-1})^{-1/2}\mathbf{X}$ to equal $\mathbf{U\Psi V}$, where $\mathbf{\Psi} = \text{diag}(\phi_i/\sqrt{(1+\phi_i^2)})$. By this construction, note that all singular values of $(\mathbf{Q}^0{}_{K-1})^{-1/2}\mathbf{X}$ are less than unity.

Defining variables $\mathbf{Y}$, $\mathbf{Z}$, and $\mathbf{R}$ as

$$\mathbf{Y} = (\mathbf{I}-\mathbf{X}^\dagger \mathbf{Q}^0_{K-1}{}^{-1}\mathbf{X})^{-1/2}$$
$$\mathbf{Z} = \begin{bmatrix} \mathbf{Y} & \mathbf{0} \\ \mathbf{BY} & (\mathbf{Q}^0_{K-1}{}^{1/2})^\dagger \end{bmatrix}$$
$$\mathbf{R} = \begin{bmatrix} \mathbf{Y} & \mathbf{0} \\ \mathbf{0} & (\mathbf{Q}^0_{K-1}{}^{1/2})^\dagger \end{bmatrix}$$

we can rewrite Equation (16) as follows

$$\mathbf{\Delta}_K = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{I} \end{bmatrix} \mathbf{R}\mathbf{R}^{-1} \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}_{K-1} \end{bmatrix} \mathbf{R}^{-1\dagger}\mathbf{R}^\dagger \begin{bmatrix} \mathbf{I} & \mathbf{B}^\dagger \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

which equals

$$\mathbf{\Delta}_K = \mathbf{Z} \begin{bmatrix} \mathbf{Y}^{-1}\mathbf{A}\mathbf{Y}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{Q}^0_{K-1}{}^{-1/2})^\dagger \mathbf{\Delta}_{K-1} \mathbf{Q}^0_{K-1}{}^{-1/2} \end{bmatrix} \mathbf{Z}^\dagger$$

For the particular choice of $\mathbf{X}$ given above, we find that $\mathbf{BY} = (\mathbf{Q}^0{}_{K-1})^{-1/2}\mathbf{X}$, which implies $\mathbf{Z}^\dagger = (\mathbf{Q}^0_K)^{1/2}$. Defining $\mathbf{Y}^{-1}\mathbf{A}\mathbf{Y}^{-1} \triangleq \mathbf{P}'_1$ and $(\mathbf{Q}^0{}^{-1/2}_{K-1})^\dagger \mathbf{\Delta}_{K-1}\mathbf{Q}^0{}^{-1/2}_{K-1} \triangleq \mathbf{D}_{K-1}$ makes the expression for $\mathbf{\Delta}_K$ above and that given by Equation (14) identical.

Substituting (14) in (7) and calling $[\mathbf{H}_2 \ldots \mathbf{H}_K] = \mathbf{H}_2^{K-1}$, we obtain

$$\mathcal{C}_{\text{coop}}(P, \mathbf{Q}^0_K{}^{-1/2}\mathbf{H}) = \frac{1}{2}\log|\mathbf{I} + \mathbf{H}_1^\dagger \mathbf{P}_1 \mathbf{H}_1 + (\mathbf{H}_2^{K-1})^\dagger (\mathbf{Q}^0{}^{-1/2}_{K-1})^\dagger \mathbf{\Delta}_{K-1} \mathbf{Q}^0{}^{-1/2}_{K-1}\mathbf{H}_2^{K-1}|$$

Thus, our objective in (10) now simplifies to finding the optimal $\mathbf{Q}^0_{K-1}$, i.e., finding $\mathbf{Q}^0_{K-1}$ such that

$$(\mathbf{Q}^0{}^{-1/2}_{K-1})^\dagger \mathbf{\Delta}_{K-1} \mathbf{Q}^0{}^{-1/2}_{K-1} = \mathbf{D}_{K-1} = \begin{bmatrix} \mathbf{P}'_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{P}'_K \end{bmatrix}$$

for some positive semidefinite matrices $\mathbf{P}'_2, \ldots, \mathbf{P}'_K$ with

$$\text{Tr}(\mathbf{\Delta}_{K-1}) = \text{Tr}(\mathbf{P}'_2) + \ldots + \text{Tr}(\mathbf{P}'_K)$$

Note that this is exactly the same objective as in (10) with $\mathbf{Q}^0{}_{K-1}$ instead of $\mathbf{Q}^0_\mathbf{K}$ and $\mathbf{\Delta}_{K-1}$ instead of $\mathbf{\Delta}_K$, and with $K-1$ users (Users 2,3… K) instead of K users. Thus, the same procedure as in (13)-(18) can be applied to reduce the problem to a $K-2$ user one. When reduced to a 2 user (User $K-1$ and User K) system, the optimum $\mathbf{Q}^0_1$ equals $\mathbf{I}$. Thus, one can run the algorithm backwards to generate the optimal Δ^0_K and thus, this recursive reduction establishes the theorem.

Now, we show that the trace constraint is also preserved in each step of this recursive reduction, i.e. that

$$\begin{aligned} \text{Tr}(\mathbf{\Delta}_K) &= \text{Tr}\left((\mathbf{Q}^0_K{}^{1/2})^\dagger \begin{bmatrix} \mathbf{P}'_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{K-1} \end{bmatrix} \mathbf{Q}^0_K{}^{1/2}\right) \\ &= \text{Tr}\left(\mathbf{Q}^0_K \begin{bmatrix} \mathbf{P}'_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{K-1} \end{bmatrix}\right) \\ &= \text{Tr}\left(\begin{bmatrix} \mathbf{P}'_1 & \mathbf{X}^\dagger \mathbf{D}_{K-1} \\ \mathbf{X}\mathbf{P}_1 & \mathbf{Q}^0_K \mathbf{D}_{K-1} \end{bmatrix}\right) \\ &= \sum_{i=1}^{K} \text{Tr}\left(\mathbf{P}'_i\right) \end{aligned}$$

REFERENCES

[1] 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. Also in preparation for journal submission.

[2] 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.

[3] E. Telatar, “Capacity of Multi-antenna Gaussian Channels,” European Trans. on Telecomm. ETT, 10(6):585-596, November 1999.

[4] G.J. Foschini and M.J. Gans, “ On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas”, Wireless Personal Communications, vol 6.,311-335, 1998.

[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] 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.

[7] W. Yu and J. Cioffi, “Trellis Precoding for the Broadcast Channel”, Proc. IEEE Globecom, pp. 1344-1338, Nov., 2001.

[8] M. Costa, “Writing on Dirty Paper”, IEEE. Trans. Inf. Theory, v. 29, pp. 439-441, May 1983.

[9] H. Sato, “An Outer Bound on the Capacity Region of Broadcast Channel”, IEEE. Trans. Inf. Theory, v. 24, pp. 374-377, May 1978.

[10] R.A. Horn and C.R. Johnson, “Matrix Analysis”, Cambridge University Press. See Also “Topics in Matrix Analysis”, Cambridge University Press.

[11] K. Fan, “Minimax Theorems”, Proc. Natl. Acad. Sci. 39, pp. 42-47, 1953.

[12] A. von Neumann, “Zur Theorie der Gesellshaftspiele”, Math. Ann. 100, pp. 295-320, 1928

[13] H.W. Kuhn and G.P. Szego, “Differential Games and Related Topics”, North Holland Publishing Company, 1971.

[14] B. Hassibi and B.M. Hochwald, “How Much Training is Needed in Multiple-Antenna Wireless Links?”, Technical Memorandum, Bell Laboratories, April 2000.

[15] S. Verdu, “Multiple-access Channels with Memory With and Without Frame Synchronism,” IEEE Trans. Inf. Theory, vol. 35, no. 3, pp. 605-19, May 1989.

[16] L. Vandenberghe, S. Boyd and S.P. Wu “Determinant maximization with linear matrix inequality constraints”, SIAM Journal on Matrix Analysis and Applications, 19(2):499-533, 1998.