

A Proof of the Converse for the Capacity of Gaussian MIMO Broadcast Channels

Mehdi Mohseni
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: mmohseni@stanford.edu

John M. Cioffi
Department of Electrical Engineering
Stanford University
Stanford, CA 94305, USA
Email: cioffi@stanford.edu

Abstract—The paper provides a proof of the converse for the capacity region of the Gaussian MIMO broadcast channel under total average transmit power constraint. The proof uses several ideas from earlier works on the problem including the recent converse proof by Weingarten, Steinberg and Shamai. First the duality between Gaussian multiple access and broadcast channels is employed to show that every point on the boundary of the dirty paper coding region can be represented as the optimal solution to a convex optimization problem. Using the optimality conditions for this convex problem, a degraded broadcast channel is constructed for each point. It is then shown that the capacity region for this degraded broadcast channel contains the capacity region of the original channel. Moreover, the same point lies on the boundary of the dirty paper coding region for this degraded channel. Finally, the standard entropy power inequality is used to show that this point lies on the boundary of the capacity region of the degraded channel as well and consequently it is on the boundary of the capacity region of the original channel.

I. INTRODUCTION

Consider a memoryless Gaussian multiple-input multiple-output (MIMO) broadcast channel (BC) with $K = 2$ receivers. Assume that the transmitter has t antennas and each receiver has r antennas. $K = 2$ receivers and equal number of receive antennas are chosen for simplicity purposes. The proof can be easily extended to the case of more than two receivers with different numbers of antennas. The received symbols of user $k = 1, 2$ at transmission i can be expressed in terms of the transmitted symbols and channel coefficients as

$$\mathbf{y}_k(i) = H_k^T \mathbf{x}(i) + \mathbf{z}_k(i), \quad (1)$$

where $\mathbf{x}(i) \in \mathbb{R}^t$ is the vector of transmitted symbols and $\mathbf{y}_k(i) \in \mathbb{R}^r$ is the vector of received symbols. The noise vectors $\mathbf{z}_k(i)$ for $k = 1, 2$ and $i = 1, 2, \dots$ are i.i.d. white Gaussian noise with covariance matrix I_r . The matrices $H_k^T \in \mathbb{R}^{r \times t}$, $k = 1, 2$, represent the channel gains, where the entry $H_k(i, j)$ represents the channel gain from transmit antenna i to receive antenna j of

user k . A total average power constraint is assumed on transmitted symbols, i.e., for every codeword of length n , $\mathbb{E}(\sum_{i=1}^n \mathbf{x}(i)^T \mathbf{x}(i)) \leq nP$. From this point on, we shall refer to this BC as OBC (for original BC).

In [1], Caire and Shamai used Costa's "writing on dirty paper" result to establish an achievable rate region for the OBC channel, commonly referred to as the "Dirty Paper Coding" (DPC) region. They conjectured that this achievable rate region is the capacity. In pioneering works presented in [2], [3] and [7], the optimality of the DPC scheme was established for the sum-rate. Progress toward establishing this conjecture in general was made in [3] and [4]. By introducing the Degraded Same Marginal (DSM) outer bound concept, the proof of the conjecture was reduced to that for the degraded Gaussian MIMO BC. The conjecture was finally proved in [6], where the DPC region was proved to be equal to the capacity region for a class of BCs referred to as Aligned Degraded MIMO BC (ADBC) channels. This equality was proved through the definition of the enhanced ADBC channel and its existence theorem, which made it possible to employ the entropy power inequality. The converse was then generalized to the larger class of Aligned MIMO BC (AMBC) channels. Using the result for the AMBC channels, the proof was extended to the OBC channel by showing that the capacity region of the OBC channel can be expressed as the limit of the capacity regions of AMBC channels as some of the eigenvalues of the noise covariance matrices go to infinity.

In this paper, we give a simpler proof of the aforementioned conjecture. While our proof employs several ideas from previous works including the recent converse proof in [6], there are several key differences. Our converse proof applies directly to the OBC channel and does not require the concept of AMBC, ADBC channels or the use of the limiting argument. Although our proof uses the concept of enhanced ADBC channel, it is obtained via

duality theory and the optimality conditions for a convex optimization problem, which is quite different from the proof of the enhanced channel existence theorem in [6].

In the next section, by means of duality, every boundary point of the DPC region of OBC is represented as the solution to a convex optimization problem. The proof of the converse is given in Section III.

A word on notations: We use upper case letters for matrices and boldface letters for vectors. The i th element of a vector \mathbf{a} is denoted by a_i . The (i, j) entry of a matrix A is denoted by $A(i, j)$. A^T is the transpose of A and $|A|$ is its determinant. I_r denotes a r by r identity matrix. $\mathbb{E}(\cdot)$ and $\text{tr}(\cdot)$ denote the expectation and trace operations, respectively. For a symmetric matrix A , $A \succeq 0$ and $A \succ 0$ mean that A is positive semi-definite and positive definite, respectively.

II. CHARACTERIZATION OF THE DPC REGION

We first review the characterization of the DPC region for the OBC channel derived in [1]. Let the OBC channel input covariance matrix be $\bar{S} = \mathbb{E}(\mathbf{x}\mathbf{x}^T)$ and π be a permutation on $\{1, 2\}$. Then for any π and any set of positive semi-definite matrices \bar{S}_k , $k = 1, 2$, such that $\bar{S} = \sum_k \bar{S}_k$ and $\text{tr}(\bar{S}) \leq P$, the rates \bar{R}_k as given below for $k = 1, 2$ are achievable using the DPC scheme.

$$\bar{R}_{\pi(k)} = \frac{1}{2} \log \frac{\left| H_{\pi(k)}^T \left(\sum_{i \geq k} \bar{S}_{\pi(i)} \right) H_{\pi(k)} + I_r \right|}{\left| H_{\pi(k)}^T \left(\sum_{i > k} \bar{S}_{\pi(i)} \right) H_{\pi(k)} + I_r \right|}. \quad (2)$$

Recall that the permutation π specifies the encoding order. The message of user $\pi(1)$ is encoded first while the message of user $\pi(2)$ is encoded second. In the DPC scheme, users' codewords are independent and are added up to form the transmitted codeword. Hence, $\bar{S} = \sum_k \bar{S}_k$, where \bar{S}_k is the covariance matrix of the transmitted codeword for user k . Let the set $\mathcal{F}(\pi, \{\bar{S}_k\})$ be a rectangle in \mathbb{R}_+^2 defined by

$$\mathcal{F}(\pi, \{\bar{S}_k\}) = \{\mathbf{R} \in \mathbb{R}_+^2 : R_k \leq \bar{R}_k, \quad k = 1, 2\},$$

then the DPC region for OBC denoted by \mathcal{R}_{DPC}^{OBC} , is the convex hull of the union of all such sets over all permutations and admissible covariance matrices, i.e.,

$$\mathcal{R}_{DPC}^{OBC} = \overline{\bigcup_{\pi, \{\bar{S}_k\} : \bar{S}_k \succeq 0 \quad \forall k, \sum_k \text{tr}(\bar{S}_k) \leq P} \mathcal{F}(\pi, \{\bar{S}_k\})}. \quad (3)$$

The rate terms \bar{R}_k in (2) are not convex functions of $\{\bar{S}_k\}$ and therefore it is very difficult to directly characterize the boundary points of the region \mathcal{R}_{DPC}^{OBC} . This difficulty can be overcome using the duality theory introduced in [7]. It was shown there that the DPC region of OBC is equal to the capacity region of a dual multiple

access channel (MAC) under sum power constraint. This dual MAC is obtained by reversing the roles of the transmitter and the receivers in OBC. Specifically, the output of the dual MAC is given by

$$\mathbf{y} = \sum_k H_k \mathbf{x}_k + \mathbf{z}, \quad (4)$$

where $\mathbf{y} \in \mathbb{R}^t$ is the received vector, $\mathbf{x}_k \in \mathbb{R}^r$ is the transmitted vector of user k and \mathbf{z} is the receiver's Gaussian noise with covariance matrix I_t . We shall refer to this dual MAC as OMAC.

The duality result states that the DPC region in (3) is equal to the capacity region of OMAC under the same sum power constraint P . Additionally, any rate tuple $\bar{\mathbf{R}}$ in \mathcal{R}_{DPC}^{OBC} as given in (2) is achievable in OMAC by successive decoding and the decoding order is the reverse of π . In order to describe this region, for a given set of transmit covariance matrices $\{S_k\}$ for users of OMAC, define the set

$$\mathcal{G}(\{S_k\}) = \left\{ \mathbf{R} \in \mathbb{R}_+^2 : \sum_{k \in J} R_k \leq \frac{1}{2} \log \left| \sum_{k \in J} H_k S_k H_k^T + I_t \right| \quad \forall J \subseteq \{1, 2\} \right\}. \quad (5)$$

This set contains all achievable rates for OMAC when the transmit covariance matrix of user k is S_k . The capacity region of the OMAC channel under sum power constraint can be expressed as

$$\mathcal{C}_{sum}^{OMAC} = \bigcup_{\{S_k\} : S_k \succeq 0 \quad \forall k, \sum_k \text{tr}(S_k) \leq P} \mathcal{G}(\{S_k\}). \quad (6)$$

It is not hard to show that this region is closed. Furthermore, based on concavity of the $\log|\cdot|$ function, this region as given in (6), is convex and no convexification is required. Therefore, each of its boundary points can be found by maximizing $\mu_1 R_1 + \mu_2 R_2$ over \mathcal{C}_{sum}^{OMAC} for some weights $\mu_1, \mu_2 \geq 0$ [9].

The optimality of the DPC region was established for the sum-rate point ($\mu_1 = \mu_2$) and the corner points ($\mu_1 = 0$ or $\mu_2 = 0$) (see [2], [3] and [7]). Thus, it only remains to show its optimality for $\mu_1 > \mu_2 > 0$ or $\mu_2 > \mu_1 > 0$. Without loss of generality, assume that $\mu_1 > \mu_2 > 0$. The following lemma characterizes the boundary point \mathbf{R}^* corresponding to given weight vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$.

Lemma 2.1: The boundary point \mathbf{R}^* maximizing $\mu_1 R_1 + \mu_2 R_2$ over \mathcal{C}_{sum}^{OMAC} for $\mu_1 > \mu_2 > 0$ is unique and is given by

$$\begin{aligned} R_1^* &= \frac{1}{2} \log |H_1 S_1^* H_1^T + I_t|, & (7) \\ R_2^* &= \frac{1}{2} \log \frac{|H_1 S_1^* H_1^T + H_2 S_2^* H_2^T + I_t|}{|H_1 S_1^* H_1^T + I_t|}, & (8) \end{aligned}$$

$$\begin{aligned} \mathcal{L}(S_1, S_2, \Phi_1, \Phi_2, \lambda) &= (\mu_1 - \mu_2) \frac{1}{2} \log |H_1 S_1 H_1^T + I_t| + \mu_2 \frac{1}{2} \log |H_1 S_1 H_1^T + H_2 S_2 H_2^T + I_t| \\ &+ \text{tr}(S_1 \Phi_1) + \text{tr}(S_2 \Phi_2) - \lambda(\text{tr}(S_1) + \text{tr}(S_2) - P). \end{aligned} \quad (16)$$

where S_1^* and S_2^* are solutions to the optimization problem:

$$\text{Maximize} \quad (\mu_1 - \mu_2) \frac{1}{2} \log |H_1 S_1 H_1^T + I_t| \quad (9)$$

$$+ \mu_2 \frac{1}{2} \log |H_1 S_1 H_1^T + H_2 S_2 H_2^T + I_t|$$

$$\text{Subject to} \quad \text{tr}(S_1) + \text{tr}(S_2) \leq P, \quad (10)$$

$$S_1, S_2 \succeq 0. \quad (11)$$

Furthermore, for any optimal S_1^*, S_2^* , there exists $\lambda^* > 0$ and positive semi-definite matrices Φ_1^*, Φ_2^* such that $\text{tr}(S_1^* \Phi_1^*) = \text{tr}(S_2^* \Phi_2^*) = 0$ and they jointly satisfy the following Karush-Kuhn-Tucker (KKT) optimality conditions,

$$H_1^T Q_1 H_1 + \frac{1}{\lambda^*} \Phi_1^* - I_r = \mathbf{0}, \quad (12)$$

$$H_2^T Q_2 H_2 + \frac{1}{\lambda^*} \Phi_2^* - I_r = \mathbf{0}, \quad (13)$$

where Q_1 and Q_2 are defined as

$$\begin{aligned} Q_1 &= \frac{\mu_1 - \mu_2}{2\lambda^*} (H_1 S_1^* H_1^T + I_t)^{-1} \\ &+ \frac{\mu_2}{2\lambda^*} (H_1 S_1^* H_1^T + H_2 S_2^* H_2^T + I_t)^{-1} \end{aligned} \quad (14)$$

$$Q_2 = \frac{\mu_2}{2\lambda^*} (H_1 S_1^* H_1^T + H_2 S_2^* H_2^T + I_t)^{-1}. \quad (15)$$

Proof: The proof of the first part is based on the *Polymatroid* structure of the set $\mathcal{G}(\{S_k\})$ and follows directly from [5]. Recall that this point is achievable by successive decoding with user 2's message being decoded before user 1's.

The feasible region of the optimization problem in (9) defined by the constraints (10) and (11) is convex, closed and compact for any norm on the space of symmetric matrices. Also the cost function in (9) is concave and continuous in S_1 and S_2 . Hence, the optimization problem is convex and there exists S_1^*, S_2^* that achieve the maximum. In addition, the Slater condition holds and the feasible region has an interior point for any $P > 0$. Thus, any optimal solution of (9) must satisfy the KKT optimality conditions and vice versa [9]. KKT conditions are obtained from the Lagrangian of (9) given in (16) above, where $\lambda \geq 0$ is the dual variable associated with the sum power constraint in (10) and the matrices $\Phi_1, \Phi_2 \succeq 0$ are the dual variables associated with the

positive semi-definite constraints on S_1 and S_2 given in (11).

KKT conditions state that, for any optimal solution S_1^*, S_2^* of (9), there exists dual feasible variables λ^*, Φ_1^* and Φ_2^* for which the derivatives of the Lagrangian with respect to S_1 and S_2 are zero. In addition, $\text{tr}(S_1^* \Phi_1^*) = \text{tr}(S_2^* \Phi_2^*) = 0$ and $\lambda^*(\text{tr}(S_1^*) + \text{tr}(S_2^*) - P) = 0$ (Complementary Slackness conditions). Furthermore, any set of feasible primal and dual variables that satisfy these conditions is optimal.

It is not hard to show that the optimal λ^* is positive for this optimization problem. Dividing the derivatives of the Lagrangian with respect to S_1 and S_2 by λ^* , yields the left-hand side terms in (12) and (13), respectively. Also since $\lambda^* > 0$, the complementary slackness conditions require that $\text{tr}(S_1^*) + \text{tr}(S_2^*) = P$. ■

Figure 1 sketches the DPC region of the OBC channel and shows a boundary point (R_1^*, R_2^*) for some μ_1, μ_2 .

III. PROOF OF THE CONVERSE

Theorem 3.1: \mathcal{R}_{DPC}^{OBC} is the capacity region.

The steps of the proof are as follows. First, based on the optimality conditions given in Lemma 2.1, a degraded MIMO BC is defined for each point \mathbf{R}^* on the boundary of \mathcal{R}_{DPC}^{OBC} that maximizes $\mu_1 R_1 + \mu_2 R_2$, for a given μ_1, μ_2 . Using duality again, it is shown that the boundary of the DPC region for this degraded BC is tangent to the boundary of the DPC region of OBC at \mathbf{R}^* (see Lemma 3.1). Using the entropy power inequality and the degradedness property of this BC, we then prove that \mathbf{R}^* also lies on the boundary of the capacity region of this degraded BC (Lemma 3.2). Furthermore, we show that the capacity region of this degraded BC contains the capacity region of the OBC channel. Since \mathbf{R}^* is on the boundary of the capacity region of the degraded BC channel, it must also lie on the boundary of the capacity region of the OBC channel. The same argument can be used for other boundary points of the DPC region of OBC to show the DPC region is actually the capacity region.

We now give the details of the proof. Consider the boundary point \mathbf{R}^* of \mathcal{R}_{DPC}^{OBC} corresponding to a given $\mu_1 > \mu_2 > 0$ as characterized in Lemma 2.1. In the following we define a degraded BC corresponding the boundary point of \mathcal{R}_{DPC}^{OBC} under consideration.

Definition 3.1: For a given weight vector $\boldsymbol{\mu}$ and its corresponding boundary point \mathbf{R}^* of \mathcal{R}_{DPC}^{OBC} , define the DBC($\boldsymbol{\mu}$) channel as

$$\mathbf{y}_k = \mathbf{x} + \mathbf{z}_k \quad k = 1, 2, \quad (17)$$

where \mathbf{x}, \mathbf{y}_1 and $\mathbf{y}_2 \in \mathbb{R}^t$ are the channel input and output vectors, respectively, and $\mathbf{z}_1, \mathbf{z}_2$ are Gaussian noise vectors with covariance matrices Q_1 and Q_2 as defined in (14) and (15). We further assume the same average total transmit power P for this channel.

It is not hard to see that $Q_1 \succ Q_2 \succ 0$. This choice of Q_1 and Q_2 ensures that DBC($\boldsymbol{\mu}$) is *statistically degraded*, a property we shall use to establish its capacity region later.

The following lemma shows that \mathbf{R}^* is on the boundary of the DPC region of DBC($\boldsymbol{\mu}$).

Lemma 3.1: The point \mathbf{R}^* maximizes $\mu_1 R_1 + \mu_2 R_2$ over the DPC region of DBC($\boldsymbol{\mu}$) denoted by $\mathcal{R}_{DPC}^{DBC(\boldsymbol{\mu})}$.

Proof: Similar to Lemma 2.1, the boundary point \mathbf{R}^o that maximizes $\mu_1 R_1 + \mu_2 R_2$ over the DPC region of DBC($\boldsymbol{\mu}$) can be characterized. Let DMAC($\boldsymbol{\mu}$) be the dual MAC of DBC($\boldsymbol{\mu}$). This MAC has channel matrices $Q_k^{-1/2}$ for user k and white Gaussian noise \mathbf{z} with covariance matrix I_t , and is given by

$$\mathbf{y} = \sum_k Q_k^{-1/2} \mathbf{x}_k + \mathbf{z}. \quad (18)$$

Let Γ_1^* and Γ_2^* be the optimal input covariance matrices of user 1 and 2 in DMAC($\boldsymbol{\mu}$) that achieve the boundary point \mathbf{R}^o . Then \mathbf{R}^o is given by the same equations as in (7) and (8) with Γ_k^* and $Q_k^{-1/2}$ replacing S_k^* and H_k , respectively for $k = 1, 2$. This point is obtained by successive decoding with user 2's message being decoded before user 1's. Additionally, Γ_1^* and Γ_2^* are optimal solutions to an optimization problem obtained from (9) by the same replacements. Let the scalar γ and the matrices Ψ_1 and Ψ_2 be the corresponding dual variables of this optimization problem. The KKT conditions are obtained exactly in the same way. Set $\gamma^* = \lambda^*$, $\Psi_1^* = \Psi_2^* = \mathbf{0}$ and

$$\begin{aligned} \Gamma_1^* &= Q_1^{1/2} H_1 S_1^* H_1^T Q_1^{1/2}, \\ \Gamma_2^* &= Q_2^{1/2} H_2 S_2^* H_2^T Q_2^{1/2}, \end{aligned}$$

where S_1^* , S_2^* and λ^* are primal and dual optimal solutions of problem (9). It is not hard to show that the given primal and dual variables are feasible and satisfy the KKT conditions of the optimization problem for \mathbf{R}^o , hence, they are optimal. Moreover, $\mathbf{R}^o = \mathbf{R}^*$ for this set of input covariance matrices. Thus, \mathbf{R}^* lies on the boundary of $\mathcal{R}_{DPC}^{DBC(\boldsymbol{\mu})}$. ■

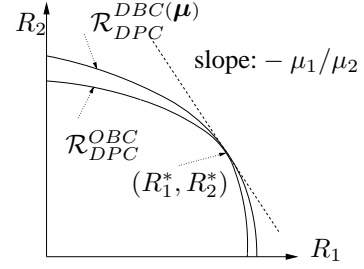


Fig. 1. \mathbf{R}^* lies on the boundary of DPC regions of OBC and DBC($\boldsymbol{\mu}$).

Figure 1 shows the DPC regions for both OBC and the degraded DBC($\boldsymbol{\mu}$) defined for the point \mathbf{R}^* .

Before proceeding to show that \mathbf{R}^* also lies on the boundary of the capacity region of DBC($\boldsymbol{\mu}$), consider the transmit covariance matrices for DBC($\boldsymbol{\mu}$) that achieve \mathbf{R}^* . Denote these matrices by $\bar{\Gamma}_1^*$ and $\bar{\Gamma}_2^*$. Using the transformation formulae given in [7] between each user's covariance matrix in the MAC and its dual BC, one can obtain $\bar{\Gamma}_1^*$ and $\bar{\Gamma}_2^*$ in terms of Γ_1^* and Γ_2^* and ultimately in terms of S_1^* , S_2^* and λ^* . After some algebra, the expressions for these matrices can be simplified to

$$\bar{\Gamma}_1^* = \frac{\mu_1}{2\lambda^*} I_t - Q_1 - \bar{\Gamma}_2^*, \quad (19)$$

$$\bar{\Gamma}_2^* = \frac{\mu_2}{2\lambda^*} (H_1 S_1^* H_1^T + I_t)^{-1} - Q_2. \quad (20)$$

By some matrix manipulation, it can be verified that $\bar{\Gamma}_1^*, \bar{\Gamma}_2^* \succeq 0$. Also $\bar{\Gamma}_1^* + \bar{\Gamma}_2^* = Q_1 H_1 S_1^* H_1^T + Q_2 H_2 S_2^* H_2^T$, hence, from equations (12) and (13), $\bar{\Gamma}_1^* + \bar{\Gamma}_2^*$ has the same trace, P , as $S_1^* + S_2^*$. Furthermore, they achieve \mathbf{R}^* in DBC($\boldsymbol{\mu}$) by DPC scheme. Recall that to achieve \mathbf{R}^* in DMAC($\boldsymbol{\mu}$), user 2's message is decoded before user 1's, thus, for DBC($\boldsymbol{\mu}$), user 1's message is encoded before user 2's.

Lemma 3.2: The point \mathbf{R}^* , which was shown to be on the boundary of the DPC region of DBC($\boldsymbol{\mu}$) is also on the boundary of its capacity region denoted by $\mathcal{C}^{DBC(\boldsymbol{\mu})}$.

Proof: Consider $\mathcal{C}(e^{nR_1}, e^{nR_2}, n)$, an arbitrary sequence of codes each with block length n and rates (R_1, R_2) for DBC($\boldsymbol{\mu}$) such that the average probability of decoding error, $P_e^{(n)}$, vanishes as $n \rightarrow \infty$. One way to verify that (R_1^*, R_2^*) lies on the boundary of $\mathcal{C}^{DBC(\boldsymbol{\mu})}$ is to show that for any such sequence of codes with $R_2 = R_2^*$, R_1 cannot take values greater than R_1^* .

Note that $Q_1 \succ Q_2$, therefore DBC($\boldsymbol{\mu}$) is a degraded BC and its capacity region is equal to the capacity region of a *physically degraded* BC given by

$$\begin{aligned} \mathbf{y}_2 &= \mathbf{x} + \mathbf{z}_2, \\ \mathbf{y}_1 &= \mathbf{y}_2 + \mathbf{z}'_1, \end{aligned}$$

where \mathbf{z}_2 and \mathbf{z}'_1 are independent Gaussian noises with covariance matrices equal to Q_2 and $Q_1 - Q_2$, respectively. The capacity region of the degraded BC is well known [8]. For the sequence of codes under consideration, assume $R_2 = R_2^*$. By Fano's inequality we have,

$$nR_1 \leq I(W_1; \mathbf{y}_1^n) + n\epsilon_{1n}, \quad (21)$$

$$nR_2^* \leq I(W_2; \mathbf{y}_2^n | W_1) + n\epsilon_{2n}, \quad (22)$$

where W_1, W_2 are intended messages for user 1 and user 2, respectively and $\epsilon_{1n}, \epsilon_{2n}$ go to zero as $n \rightarrow \infty$. After expanding the mutual information term in (22) as $I(W_2; \mathbf{y}_2^n | W_1) = h(\mathbf{y}_2^n | W_1) - h(\mathbf{y}_2^n | W_1, W_2)$ and from the equalities $h(\mathbf{y}_2^n | W_1, W_2) = nh(\mathbf{z}_2)$ and $R_2^* = \frac{1}{2} \log |\bar{\Gamma}_2^* + Q_2| - \frac{1}{2} \log |Q_2|$, the following lower bound on $h(\mathbf{y}_2^n | W_1)$ is obtained

$$h(\mathbf{y}_2^n | W_1) + n\epsilon_{2n} \geq \frac{n}{2} \log(2\pi e)^t |\bar{\Gamma}_2^* + Q_2|. \quad (23)$$

Now since \mathbf{z}'_1 is independent of $(W_1, W_2, \mathbf{z}_2^n)$, and conditioned on W_1 , $\mathbf{y}_1^n = \mathbf{y}_2^n + \mathbf{z}'_1$ and \mathbf{y}_2^n have densities, the entropy power inequality [8] can be applied to obtain

$$e^{\frac{2}{nt} h(\mathbf{y}_1^n | W_1)} \geq e^{\frac{2}{nt} h(\mathbf{y}_2^n | W_1)} + e^{\frac{2}{nt} h(\mathbf{z}'_1)}. \quad (24)$$

From the inequalities in (23), (24) and the expression for $h(\mathbf{z}'_1)$, one can obtain

$$e^{\frac{2}{nt} h(\mathbf{y}_1^n | W_1)} \geq 2\pi e \left(|\bar{\Gamma}_2^* + Q_2|^{\frac{1}{t}} + |Q_1 - Q_2|^{\frac{1}{t}} \right) - \delta_n,$$

where δ_n goes to zero as $n \rightarrow \infty$. However, the expression for $\bar{\Gamma}_2^*$ in (20) reveals that two matrix expressions on the right-hand side, $(\bar{\Gamma}_2^* + Q_2)$ and $(Q_1 - Q_2)$, are scaled versions of each other. Consequently,

$$|\bar{\Gamma}_2^* + Q_2|^{\frac{1}{t}} + |Q_1 - Q_2|^{\frac{1}{t}} = |\bar{\Gamma}_2^* + Q_1|^{\frac{1}{t}},$$

which yields a lower bound on $h(\mathbf{y}_1^n | W_1)$ as given below

$$h(\mathbf{y}_1^n | W_1) \geq \frac{n}{2} \log(2\pi e)^t |\bar{\Gamma}_2^* + Q_1| - n\delta'_n,$$

where δ'_n goes to zero as $n \rightarrow \infty$. This inequality together with (21) enables us to obtain an upper bound on R_1 as

$$\begin{aligned} R_1 &\leq \frac{1}{n} h(\mathbf{y}_1^n) - \frac{1}{n} h(\mathbf{y}_1^n | W_1) + \epsilon_{1n} \\ &\leq h(\mathbf{y}_1) - \frac{1}{2} \log(2\pi e)^t |\bar{\Gamma}_2^* + Q_1| + \delta'_n + \epsilon_{1n}. \end{aligned}$$

However, $\mathbf{y}_1 = \mathbf{x} + \mathbf{z}_2 + \mathbf{z}'_1$ is the transmitted vector corrupted by additive Gaussian noise with covariance matrix Q_1 and there is an average power constraint on \mathbf{x} , $\mathbb{E}(\mathbf{x}^T \mathbf{x}) \leq P$. Consequently, $h(\mathbf{y}_1)$ is maximized by a Gaussian \mathbf{x} with zero mean and covariance matrix that

has the same eigenvectors as the noise covariance matrix Q_1 and its eigenvalues water-fill the eigenvalues of Q_1 [8]. From (19), it can be seen that $\bar{\Gamma}_1^* + \bar{\Gamma}_2^* + Q_1 = \frac{\mu_1}{2\lambda^*} I_t$, which states that the transmit covariance matrix $\bar{\Gamma}_1^* + \bar{\Gamma}_2^*$ satisfies the water-filling conditions. Hence $h(\mathbf{y}_1) \leq \frac{1}{2} \log(2\pi e)^t |\bar{\Gamma}_1^* + \bar{\Gamma}_2^* + Q_1|$, that yields

$$R_1 \leq \frac{1}{2} \log \frac{|\bar{\Gamma}_1^* + \bar{\Gamma}_2^* + Q_1|}{|\bar{\Gamma}_2^* + Q_1|} + \delta'_n + \epsilon_{1n} = R_1^* + \delta'_n + \epsilon_{1n}.$$

as $n \rightarrow \infty$, δ'_n and ϵ_{1n} go to zero and R_1 could be at most R_1^* . ■

To complete the proof and show that \mathbf{R}^* is on the boundary of the capacity region of OBC, it remains to show that the capacity region of DBC(μ) contains the capacity region of OBC. Since \mathbf{R}^* is on the boundary of the capacity region of DBC(μ) that contains the capacity region of OBC, it must be on the boundary of the capacity region of OBC as well. Equalities given in (12) and (13) guarantee that $H_k^T Q_k H_k \leq I_r$ for $k = 1, 2$. Therefore, the capacity region of DBC(μ) cannot be smaller than the capacity region of OBC. This is based on the observation that receiver k of DBC(μ) can successfully decode any code originally designed for OBC after multiplying its output by H_k^T and adding an independent Gaussian noise with covariance matrix $I_r - H_k^T Q_k H_k$ to it. This completes the converse proof.

ACKNOWLEDGMENT

The authors would like to thank professor El Gamal for his careful comments on the earlier draft of this paper.

REFERENCES

- [1] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," *38th annual Allerton conf. on commun., control and computing*, Oct. 2000.
- [2] W. Yu and J.M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. on Inform. Theory*, Vol. 50, No. 9, pp. 1875-1892, Sept. 2004.
- [3] P. Viswanath and D. Tse "On the capacity of the multiple antenna broadcast channel," *DIMACS Workshop on Signal Processing for Wireless Commun.*, Oct. 2002.
- [4] S. Vishwanath, G. Kramer, S. Shamai, S. Jafar and A. Goldsmith, "Outer bounds for multi-antenna broadcast channels," *DIMACS Workshop on Signal Processing for Wireless Commun.*, Oct. 2002.
- [5] D. Tse and S. Hanly, "Multi-access fading channels-Part I: polymatroid structure, optimal resource allocation and throughput capacities" *IEEE Trans. Inform. Theory*, vol.44, no.7, pp.2796-2815, Nov. 1998.
- [6] H. Weingarten, Y. Steinberg and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," *Proc. CISS*, 2004.
- [7] S. Vishwanath, N. Jindal, A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inform. Theory*, vol.49, No.10, pp.2658-2668, Oct. 2003.
- [8] T. Cover and J. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
- [9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.