

# Approximate Capacity Region of the MISO Broadcast Channels with Delayed CSIT

Alireza Vahid, Mohammad Ali Maddah-Ali, and A. Salman Avestimehr

**Abstract**—We consider the problem of multiple-input single-output Broadcast Channels with Rayleigh fading where the transmitter has access to delayed knowledge of the channel state information. We first characterize the capacity region of this channel with two users to within constant number of bits for all values of the transmit power. The proposed signaling strategy utilizes the delayed knowledge of the channel state information and the previously transmitted signals, in order to create a signal of common interest for both receivers. This signal would be the quantized version of the summation of the previously transmitted signals. A challenge that arises in deriving the result for finite signal-to-noise ratio regimes is the correlation that exists between the quantization noise and the signal. To guarantee the independence of quantization noise and signal, we extend the framework of lattice quantizers with dither together with an interleaving step. For converse, we use the fact that the capacity region of this problem is upper-bounded by the capacity region of a physically degraded broadcast channel with no channel state information where one receiver has two antennas. Then, we derive an outer-bound on the capacity region of this degraded broadcast channel. Finally, we show how to extend our results to obtain the approximate capacity of the  $K$ -user multiple-input single-output Broadcast Channel with delayed knowledge of the channel state information at the transmitter to within  $2 \log_2(K+2)$  bits/s/Hz.

**Index Terms**—Broadcast channel, capacity, multiple-input single-output, delayed CSIT.

## I. INTRODUCTION

In wireless networks, receivers estimate the channel state information (CSI) and pass this information to the transmitters through feedback mechanisms. The extent to which channel state information is available at the transmitters has a direct impact on the capacity of wireless networks and the optimal strategies. In fast-fading scenarios where the coherence time of the channel is smaller than the delay of the feedback channel, providing the transmitters with up-to-date CSI is practically infeasible. Consequently, we are left with no choice but to understand the behavior of wireless networks under such constraint.

Alireza Vahid is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. Email: av292@cornell.edu.

A. Salman Avestimehr is with the Electrical Engineering Department, University of Southern California, Los Angeles, CA, USA. Email: avestimehr@ee.usc.edu.

Mohammad Ali Maddah-Ali is with Bell Labs, Alcatel-Lucent, Holmdel, NJ, USA. Email: mohammadali.maddah-ali@alcatel-lucent.com.

The work of A. S. Avestimehr and A. Vahid is in part supported by NSF Grants CAREER-0953117, CCF-1161720, NETS-1161904, AFOSR Young Investigator Program Award, ONR award N000141310094, and 2013 Qualcomm Innovation Fellowship.

The preliminary results of this work were presented at the Allerton Conference [1].

Our objective is to understand the effect of lack of up-to-date CSI on the *capacity* region of wireless networks by considering a fundamental building block, namely the multiple-input single-output (MISO) broadcast channel (BC). In the context of MISO BC, it has been shown that even completely stale CSIT (also known as delayed CSIT) can still be very useful and can change the scale of the capacity, measured by the degrees of freedom (DoF) [2]. The degrees of freedom by definition provides a first order approximation of the capacity, thus it is mainly useful in understanding the behavior of the capacity in high power regimes. However, it is not a suitable measure for practical settings with finite signal-to-noise ratio (SNR).

We first consider the two-user MISO BC and we focus on the effect of delayed CSIT at finite SNR regimes as opposed to the asymptotic DoF analysis. While there is a strong body of work on broadcast channels with perfect channel state information (see [3]–[5]), no capacity result has been reported for the delayed CSIT scenario. There are some prior results in the literature (for example [6]) that have proposed and analyzed several achievability strategies at finite SNR regimes. Nonetheless, characterizing the capacity region of the two-user MISO BC with delayed CSIT has remained open.

In this paper, we provide the first constant-gap approximation of the capacity region of the two-user MISO BC with delayed CSIT. We obtain an achievable scheme and an outer-bound on the capacity region, and we analytically show that they are within 1.81 bits/sec/Hz per user, for all values of the transmit power. Our numerical analysis shows that the gap is in fact smaller and in the worst case, it is at most 1.1 bits/sec/Hz per user.

The proposed achievability scheme for the two-user MISO BC with delayed CSIT has two phases. Phase 1 has two segments and in the first segment and the second segment, transmitter respectively sends messages intended for receivers one and two. In each one of these segments, the unintended receiver overhears and saves some signal (interference) which is only useful for the other receiver. At this point, transmitter can evaluate the overheard signals using the delayed CSI. In Phase 2, the transmitter will swap the overheard signals between the receivers by sending a signal of common interest. This signal is the quantized version of the summation of the previously transmitted signals. The swapping is performed by exploiting the overheard signals as available side-information at receivers side. The overall information that each receiver collects in the two phases is enough to decode the intended message. Although the above two phases follows the scheme of [2], as we will show, some important additional ingredients are needed to make an approximately optimal scheme.

To derive the outer-bound, we create a physically degraded BC by providing the received signal of one user one to the other user. Then, since we know that feedback does not enlarge the capacity region of a physically degraded BC [7], we ignore the delayed knowledge of the channel state information at the transmitter (*i.e.* no CSIT assumption). We then derive an outer-bound on the capacity region of this degraded broadcast channel which in turn serves as an outer-bound on the capacity region of the two-user MISO BC with delayed CSIT. We show that the achievable rate region and the outer-bound are within 1.81 bits/sec/Hz per user. Using numerical analysis, we can show that the gap is in fact smaller than 1.1 bits/sec/Hz per user.

We then consider the  $K$ -user MISO BC with delayed CSIT and we focus on the symmetric capacity. We show how to extend our ideas for the achievability and converse to this setting, and we show that for the symmetric capacity the gap between the achievable rate and the outer-bound is less than  $2 \log_2(K+2)$  bits/sec/Hz independent of the transmit power. We use this result to provide the approximate capacity of the  $K$ -user MISO BC with delayed CSIT.

In the literature, there have been several results on the impact of delayed CSIT in wireless networks. However, these results are either focused on the DoF region (*e.g.*, [8]–[10]) or capacity results for noiseless channels (*e.g.*, [11]–[14]). The present work provides constant-gap approximation of the capacity region of the two-user Complex Gaussian MISO BC with delayed CSIT which is of great importance in practical wireless settings. Existing results on constant-gap approximation of the capacity region of wireless networks mainly consider the scenario in which transmitters have perfect instantaneous knowledge of the channel state information (*e.g.*, [15]–[17]). Thus the current result opens the door for constant-gap approximation of the capacity region of wireless networks with delayed CSIT.

The rest of the paper is organized as follows. In Section II, we formulate our problem. In Section III, we present our main results. We describe our achievability strategy in Section IV. Section V is dedicated to deriving the outer-bound. In Section VI, we show that our inner-bound and outer-bound are within constant number of bits. We extend our results to the  $K$ -user MISO BC with delayed CSIT in Section VII. Finally, Section VIII concludes the paper.

## II. PROBLEM SETTING

We start by considering the two-user multiple-input single-output (MISO) complex Gaussian broadcast channel (BC) with Rayleigh fading as depicted in Fig. 1(a). The channel gains from the transmitter to receivers one and two are denoted by  $\mathbf{h}[t], \mathbf{g}[t] \in \mathbb{C}^{2 \times 1}$ , respectively, where the entries of  $\mathbf{h}[t]$  and  $\mathbf{g}[t]$  are distributed as i.i.d.  $\mathcal{CN}(0, 1)$  (independent across time, antenna, and users). At each receiver, the received signal can be expressed as follows.

$$y_1[t] = \mathbf{h}^\top[t] \mathbf{x}[t] + z_1[t], \quad y_2[t] = \mathbf{g}^\top[t] \mathbf{x}[t] + z_2[t], \quad (1)$$

where  $\mathbf{x}[t] \in \mathbb{C}^{2 \times 1}$  is the transmit signal subject to average power constraint  $P$ , *i.e.*  $\mathbb{E}[\mathbf{x}^\dagger[t] \mathbf{x}[t]] \leq P$  for  $P > 0$ . The

noise processes are independent from the transmit signal and are distributed i.i.d. as  $z_k[t] \sim \mathcal{CN}(0, 1)$ . Furthermore, we define

$$s_1[t] = \mathbf{h}^\top[t] \mathbf{x}[t], \quad s_2[t] = \mathbf{g}^\top[t] \mathbf{x}[t], \quad (2)$$

to be the noiseless versions of the received signals.

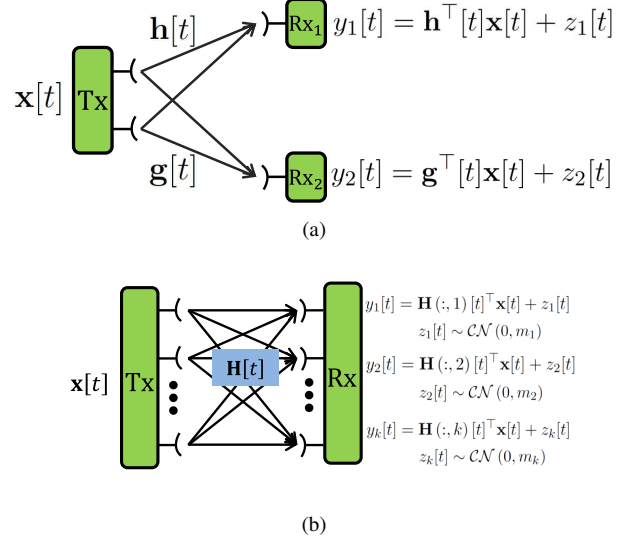


Fig. 1. (a) Two-user Multiple-Input Single-Output (MISO) Complex Gaussian Broadcast Channel; (b) Point-to-point complex Gaussian channel with  $k$  transmit and receive antennas and no CSIT, where  $z_j[t] \sim \mathcal{CN}(0, m_j)$ . Here,  $\mathbf{H}[t]$  denotes the channel transfer matrix at time  $t$ ; and  $\mathbf{H}(:, j)[t]$  denotes the channel gains from all transmit antennas to receive antenna  $j$  at time  $t$ .

Transmitter wishes to reliably communicate independent and uniformly distributed messages  $w_1 \in \{1, 2, \dots, 2^{nR_1}\}$  and  $w_2 \in \{1, 2, \dots, 2^{nR_2}\}$  to receivers 1 and 2, respectively, during  $n$  uses of the channel. We denote the channel state information at time  $t$  by  $(\mathbf{h}[t], \mathbf{g}[t])$  for  $t = 1, 2, \dots, n$ . The transmitter has access to the delayed (outdated) channel state information, meaning that at time instant  $t$ , the transmitter has access to  $(\mathbf{h}[\ell], \mathbf{g}[\ell])_{\ell=1}^{t-1}$  for  $t = 1, 2, \dots, n$ .

Due to the delayed knowledge of the channel state information, the encoded signal  $\mathbf{x}[t]$  is a function of both the messages and the previous channel realizations.

Each receiver  $k$ ,  $k = 1, 2$ , uses a decoding function  $\varphi_{k,n}$  to get the estimate  $\hat{w}_k$  from the channel outputs  $\{y_k[t] : t = 1, \dots, n\}$ . An error occurs whenever  $\hat{w}_k \neq w_k$ . The average probability of error is given by  $\lambda_{k,n} = \mathbb{E}[P(\hat{w}_k \neq w_k)]$ ,  $k = 1, 2$ , where the expectation is taken with respect to the random choice of the transmitted messages  $w_1$  and  $w_2$ .

We say that a rate pair  $(R_1, R_2)$  is achievable, if there exists a block encoder at the transmitter, and a block decoder at each receiver, such that  $\lambda_{k,n}$  goes to zero as the block length  $n$  goes to infinity,  $k = 1, 2$ . The capacity region  $\mathcal{C}$  is the closure of the set of the achievable rate pairs.

## III. STATEMENT OF THE MAIN RESULT

Our main contributions are: (1) characterization of the capacity region of the two-user MISO complex Gaussian BC with delayed CSIT to within 1.81 bits/sec/Hz; and (2)

characterizing the symmetric capacity of the  $K$ -user MISO complex Gaussian BC to within  $2\log_2(K+2)$  bits/sec/Hz.

For the two-user setting, the achievability scheme has two phases. Phase 1 has two segments and in the first segment and the second segment, the transmitter respectively sends messages intended for receiver one and receiver two. In each of these segments, the unintended receiver overhears and saves some signal (interference), which is only useful for the other receiver. Later, in the second phase, the transmitter evaluates what each receiver overheard about the other receiver's message using the delayed knowledge of the channel state information and provides these overheard signals efficiently to both receivers exploiting available side information at each receiver. Transmitter provides the overheard signals to the receivers by sending a signal of common interest. This way transmitter reduces the overall communication time and in turn, increases the achievable rate.

The outer-bound is derived based on creating a physically degraded broadcast channel where one receiver is enhanced by having two antennas. In this channel, feedback and in particular delayed knowledge of the channel state information, does not increase the capacity region. Thus, we can ignore the delayed knowledge of the channel state information and consider a degraded BC with no CSIT. This would provide us with the outer-bound.

We then show how to extend our arguments for achievability and converse to the  $K$ -user setting to derive approximate symmetric capacity under delayed CSIT assumption. Before stating our results, we need to define some notations.

**Definition 1.** For a region  $\mathcal{R} \subseteq \mathbb{R}^2$ , we define

$$\mathcal{R} \ominus (\tau, \tau) \triangleq \{(R_1, R_2) \mid R_1, R_2 \geq 0, (R_1 + \tau, R_2 + \tau) \in \mathcal{R}\}. \quad (3)$$

**Definition 2.** Consider an ergodic point-to-point complex Gaussian channel with  $k$  transmit antennas and  $\ell$  receive antennas where the channel gains are distributed as i.i.d.  $\mathcal{CN}(0, 1)$  (independent across time, antenna, and users). At receive antenna  $j$ , we assume an additive zero-mean Gaussian noise process of variance  $\sigma_j^2$ . Then under no CSIT assumption and for a total average transmit power of  $P$ , the ergodic capacity of this channel is denoted by

$$C_{k \times \ell}(P; \sigma_1^2, \sigma_2^2, \dots, \sigma_\ell^2). \quad (4)$$

For simplicity of notations, we drop  $P$ , and whenever noise variances are all equal to 1, we do not mention them. Fig. 1(B) depicts a point-to-point complex Gaussian channel with  $k$  transmit antennas and  $k$  receive antennas.

Consider the ergodic capacity of a point-to-point complex Gaussian channel with 2 transmit antennas and a single receive antenna denoted by  $C_{2 \times 1}(P; 1)$  as described in Definition 2. Then from [18], we have

$$C_{2 \times 1}(P; 1) = C_{2 \times 1} = \mathbb{E} \log_2 \left[ 1 + \frac{P}{2} \mathbf{g}^\dagger \mathbf{g} \right], \quad (5)$$

where  $\mathbf{g}$  is a 2 by 1 vector where entries are i.i.d.  $\mathcal{CN}(0, 1)$ .

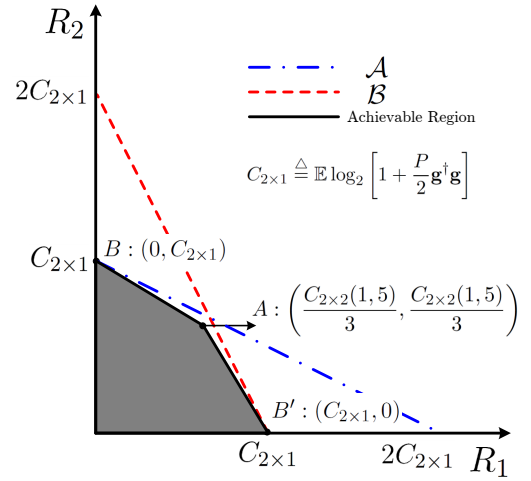


Fig. 2. The outer-bound on the capacity region of the two-user MISO BC with delayed CSIT is the intersection of  $\mathcal{A}$  and  $\mathcal{B}$ . We prove that the capacity region is within 1.81 bit/sec/Hz per user of this outer-bound. The achievable rate region is shown by the shaded area.  $C_{2 \times 1}$  and  $C_{2 \times 2}(1, 5)$  are given in Definition 2.

**Definition 3.** Based on  $C_{2 \times 1}$  described above, we define region  $\mathcal{A}$  and region  $\mathcal{B}$  as

$$\begin{aligned} \mathcal{A} &\triangleq \{(R_1, R_2) \geq 0 \mid R_1, R_2 \geq 0, R_1 + 2R_2 \leq 2C_{2 \times 1}\}, \\ \mathcal{B} &\triangleq \{(R_1, R_2) \geq 0 \mid R_1, R_2 \geq 0, 2R_1 + R_2 \leq 2C_{2 \times 1}\}. \end{aligned} \quad (6)$$

**Remark 1.** As we will show in Section V,  $\mathcal{A}$  ( $\mathcal{B}$ ) is an outer-bound on the capacity region of a two-user complex Gaussian MIMO BC with no CSIT where  $\mathbf{R}_{\mathbf{x}_1}$  ( $\mathbf{R}_{\mathbf{x}_2}$ ) has two antennas and  $\mathbf{R}_{\mathbf{x}_2}$  ( $\mathbf{R}_{\mathbf{x}_1}$ ) has only one antenna (additive noise processes all having zero-mean and variance 1). The corner points of  $\mathcal{A}$  are given by  $(0, C_{2 \times 1})$  and  $(2C_{2 \times 1}, 0)$ .

The following theorem states our contribution for the two-user MISO BC with delayed CSIT.

**Theorem 1.** The capacity region of the two-user MISO BC with delayed CSIT,  $\mathcal{C}$ , is within 1.81 bits/sec/Hz per user of  $(\mathcal{A} \cap \mathcal{B})$ , i.e.

$$(\mathcal{A} \cap \mathcal{B}) \ominus (1.81, 1.81) \subseteq \mathcal{C} \subseteq (\mathcal{A} \cap \mathcal{B}), \quad (7)$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are given in Definition 3.

**Remark 2.** Fig. 2 pictorially depicts our result for the two-user MISO BC with delayed CSIT. We have defined  $C_{2 \times 1}$  and  $C_{2 \times 2}(1, 5)$  in Definition 2. We analytically show that the achievable region is within 1.81 bits/sec/Hz per user of the outer-bound. Our numerical analysis shows that the gap is in fact smaller and in the worst case, it is at most 1.1 bits/sec/Hz per user.

We then consider the  $K$ -user MISO BC with delayed CSIT. We only focus on the symmetric capacity defined as follows.

**Definition 4.** The symmetric capacity of the  $K$ -user MISO BC with delayed CSIT,  $C_{\text{SYM},K\text{-user}}$ , is given by

$$C_{\text{SYM},K\text{-user}} \triangleq \sup \{R : R \leq R_j, \forall j = 1, 2, \dots, K, (R_1, \dots, R_K) \in \text{capacity region of } K\text{-user MISO BC with delayed CSIT}\}. \quad (8)$$

The following theorem states our contribution for the  $K$ -user setting.

**Theorem 2.** For the  $K$ -user MISO BC with delayed CSIT,  $C_{\text{SYM},K\text{-user}}$  is lower-bounded by

$$\frac{C_{K \times K}(1, 5, \dots, 1 + (j-1)(j+2), \dots, 1 + (K-1)(K+2))}{K \sum_{j=1}^K j^{-1}}, \quad (9)$$

and upper-bounded by  $(\sum_{j=1}^K j^{-1})^{-1} C_{K \times 1}$ .

We use Theorem 2 to provide the approximate capacity of the  $K$ -user MISO BC with delayed CSIT in Section VII.

**Remark 3.** For Theorem 2, our numerical analysis shows that the gap is at most 2.3 bits/sec/Hz per user for  $K \leq 20$  and  $P \leq 60$ dB. In general, we show that the gap is at most  $2 \log_2(K+2)$  bits/sec/Hz.

#### IV. ACHIEVABILITY PROOF OF THEOREM 1

In this section, we describe the achievability strategy of Theorem 1. To characterize the capacity region of the two-user MISO complex Gaussian BC to within 1.81 bits/sec/Hz per user, we need to show that the  $(\mathcal{A} \cap \mathcal{B}) \ominus (1.81, 1.81)$  is achievable.

We define rate region  $\mathcal{R}(D)$  for parameter  $D \geq 4$  as follows:

$$\mathcal{R}(D) \triangleq \left\{ (R_1, R_2) \left| \begin{array}{l} 0 \leq R_1 + \gamma R_2 \leq C_{2 \times 1} \\ 0 \leq \gamma R_1 + R_2 \leq C_{2 \times 1} \end{array} \right. \right\} \quad (10)$$

where  $\gamma = \left( \frac{3C_{2 \times 1}}{C_{2 \times 2}(1, 1+D)} - 1 \right)^1$ .

Later in Section VI, we show that  $(\mathcal{A} \cap \mathcal{B}) \ominus (1.81, 1.81) \subseteq \mathcal{R}(4)$ , and thus, to characterize the capacity region to within 1.81 bit/sec/Hz per user, it suffices to show that  $\mathcal{R}(4)$  is achievable.

The shadowed region in Fig. 2 corresponds to  $\mathcal{R}(4)$ , which is a polygon with corner points  $A$ ,  $B$ , and  $B'$ . Corner points  $B$  and  $B'$  of  $\mathcal{R}(4)$ , are achievable using the result on point-to-point MISO Gaussian channels with no CSIT [18]. Therefore, we only need to describe the achievability strategy for corner point  $A$ .

<sup>1</sup>As we will see later in this section and Section VI, we have  $\left( \frac{3C_{2 \times 1}}{C_{2 \times 2}(1, 1+D)} - 1 \right) > 0$  for  $D \geq 4$  and  $P > 0$ dB.

#### A. Transmission Strategy for Corner Point $A$

Here, we present the achievability strategy for corner point  $A$ . We show that for any  $\epsilon > 0$ , we can achieve

$$(R_1, R_2) = \left( \frac{C_{2 \times 2}(1, 5) - \epsilon}{3}, \frac{C_{2 \times 2}(1, 5) - \epsilon}{3} \right), \quad (11)$$

with vanishing decoding error probability as the communication length goes to infinity.

The achievability strategy is carried on over  $n$  blocks, each block consisting of 2 phases. Phase 1 has two segments each of length  $n$  channel uses, and Phase 2 is of length  $n$  channel uses. Denote by  $w_k^b$ , the message of user  $k$  in block  $b$ ,  $k \in \{1, 2\}$ ,  $b = 1, 2, \dots, n$ . Fix  $\epsilon > 0$  and set  $R = C_{2 \times 2}(1, 1+D) - \epsilon$ . We assume that  $w_k^b \in \{1, 2, \dots, 2^{nR}\}$  and that the messages are distributed uniformly and independently. The encoding is carried on as described below.

- **Encoding:** At the transmitter, the message of user  $k$  during block  $b$ , i.e.  $w_k^b$ , is mapped to a codeword of length  $n$  denoted by  $\mathbf{x}_k^{b,n}$  where any element of this codeword is drawn i.i.d. from  $\mathcal{CN}(0, P/2\mathbf{I}_2)^2$ .

- **Communication during segment  $j$  of Phase 1 of block  $b$ :** During this segment, the transmitter communicates  $\mathbf{x}_j^{b,n}$  from its two transmit antennas,  $j = 1, 2$ , and  $b = 1, 2, \dots, n$ . Receiver one obtains  $y_{1,1;j}^{b,n}$  and receiver two obtains  $y_{2,1;j}^{b,n}$ .

- **Communication during Phase 2 of block  $b$ :** Using the delayed CSIT, the transmitter creates

$$s^{b,n} = s_{2,1:1}^{b,n} + s_{1,1:2}^{b,n}, \quad (12)$$

where  $s_{2,1:1}^{b,n}$  is the received signal at  $R_{x_2}$  during the first segment of Phase 1 of block  $b$ ,  $y_{2,1:1}^{b,n}$ , minus the noise term as defined in (2), and  $s_{1,1:2}^{b,n}$  is the received signal at  $R_{x_1}$  during the second segment of Phase 1 of block  $b$ ,  $y_{1,1:2}^{b,n}$ , minus the noise term.

Note that  $s_{2,1:1}^{b,n} + s_{1,1:2}^{b,n}$  is useful for both receivers since each receiver can subtract its previously received signal to obtain what the other receiver has (up to the noise term). Therefore, the goal in this phase, is to provide  $s_{2,1:1}^{b,n} + s_{1,1:2}^{b,n}$  to both receivers with distortion  $D = 4$ .

**Remark 4.** We note that the idea of creating quantized version of previously received signals has been utilized previously for asymptotic degrees of freedom analysis [19]–[21]. However for finite SNR regime, we need to take into account the fact that the quantization noise is neither independently distributed over time nor is it independent from the signal; and we need to overcome these challenges.

The input signal to a lattice quantizer needs to be independently distributed over time (see [22], [23]). Thus, in order to quantize this signal using lattice quantizer, we need

$$(s_{2,1:1}^b[\ell] + s_{1,1:2}^b[\ell])_{\ell=1}^n$$

to be an independently distributed sequence. However, given message  $w_k^b$ , the transmit signal  $\mathbf{x}_k^{b,n}$  is correlated across time and as a result, the aforementioned signals are not independent anymore. In order to overcome this issue, we incorporate an

<sup>2</sup> $\mathbf{I}_2$  is the  $2 \times 2$  identity matrix.

interleaving step according to the following mapping which is depicted in Fig. 3.

$$\tilde{s}^b[t] = s_{2,1:1}^t[b] + s_{1,1:2}^t[b], \quad (13)$$

where  $b = 1, 2, \dots, n$  and  $t = 1, 2, \dots, n$ . It is important to notice that with this interleaving, the resulting signal at different time instants of a given phase at a given block are independent from each other. This is due to the fact that these signals are created from independent messages.

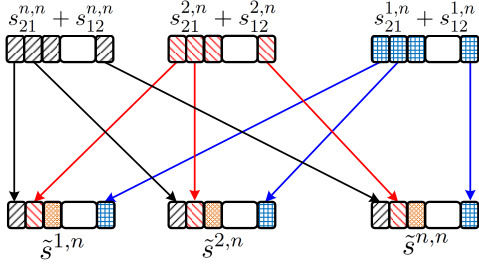


Fig. 3. The interleaving step: the resulting signal at different time instants of a given phase at a given block are independent from each other.

Note that given the previous channel realizations, the signal in (13) at each time has a Gaussian distribution but its variance varies from each time instant to the other. Thus, in order to be able to quantize it, we need a generalization of the rate-distortion function to include non-identically distributed sources. Below, we discuss this issue.

**Lemma 1.** [Rate distortion for non-identically distributed Gaussian Source] Consider a independently non-identically distributed Gaussian source  $\mathbf{u}$ , where at time instant  $t$ , it has zero mean and variance  $\sigma^2[t]$ . Assume that  $\sigma^2[t]$  is drawn from some i.i.d. continuous distribution with  $\mathbb{E}[\sigma^2] < \infty$ . The sequence of  $\sigma^2[t]$  is non-causally known by both encoder and decoder. Then, with squared-error distortion, we can quantize the signal at any rate greater than or equal to

$$\min_{D_\sigma: \mathbb{E}[D_\sigma] \leq D} \mathbb{E} \left[ \log_2 \frac{\sigma^2[t]}{D_\sigma} \right]^+, \quad (14)$$

and achieve distortion  $D$  (per sample), where the expectation is with respect to the distribution of  $\sigma^2$ .

*Proof sketch:* Suppose  $\sigma^2$  could only take  $m$  finite values  $\sigma_i^2$  with probability  $p_i$ ,  $i = 1, 2, \dots, m$ . The problem would then be similar to that of  $m$  parallel Gaussian channels where waterfilling gives the optimal solution ((Theorem 10.3.3 [24])). For the time instants where source has variance  $\sigma_i^2$ , we choose a distortion  $D_i$  such that  $\sum_{i=1}^m p_i D_i \leq D$ . Note that in order to derive the optimum answer, we need to optimize over the choice of  $D_i$ 's. The case where  $\sigma^2$  take values in a continuous set can be viewed as the limit of the discrete scenario using standard arguments. ■

It is easy to see that any rate greater than or equal to

$$\mathbb{E} \left[ \log_2 \left( 1 + \frac{\sigma^2[t]}{D} \right) \right], \quad (15)$$

is also achievable at distortion  $D$  (per sample). Basically, we have ignored the optimization over  $D$  and added a 1 to remove  $\max\{., 0\}$  (or  $.^+$ ).

Moreover, we would like the distortion to be independent of the signals. In order to have a distortion that is independent of the signal and is uncorrelated across time, we can incorporate lattice quantization with “dither” as described in [22]. Basically, dither is a random vector distributed uniformly over the basic Voronoi that is added to the signal before feeding it to the quantizer. At the decoder, this random vector will be subtracted.

From (15), we conclude that we can quantize  $\tilde{s}^{b,n}$  with squared-error distortion  $D$  at rate  $R_Q(D)$  (per sample), defined as

$$R_Q(D) \triangleq \mathbb{E} \left[ \log_2 \left( 1 + \frac{P}{2D} (\|\mathbf{g}\|_2^2 + \|\mathbf{h}\|_2^2) \right) \right], \quad (16)$$

where  $\mathbf{g}, \mathbf{h} \in \mathbb{C}^{2 \times 1}$  with i.i.d.  $\mathcal{CN}(0, 1)$  entries.

We can reliably communicate the quantization index over

$$\left\lceil \frac{nR_Q(4)}{C_{2 \times 1}} \right\rceil \quad (17)$$

time instants. Next, we need to show that given the appropriate choice of parameters, receivers can recover the corresponding messages with vanishing error probability as  $n \rightarrow \infty$ .

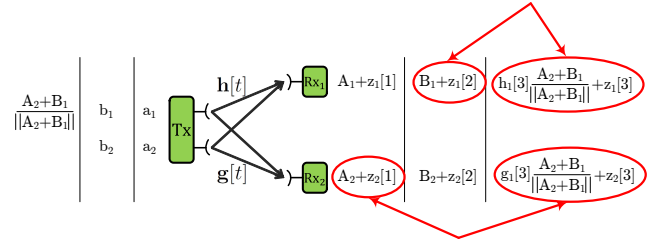


Fig. 4. The original three-slot Maddah-Ali-Tse protocol.

**Remark 5.** An attempt to directly implement Maddah-Ali-Tse protocol in the finite SNR regime faces some challenges. Maddah-Ali-Tse protocol is depicted in Fig. 4. At time 1, symbols  $a_1$  and  $a_2$  intended for receiver 1 are transmitted and receiver  $j$  obtains  $A_j + z_j[1]$  where  $z_j[1]$  is the additive noise and  $A_j$  is result of multiplying the channel vector to receiver  $j$  with the transmit signal. Similar story holds for time 2 where symbols  $b_1$  and  $b_2$  intended for receiver 2 are transmitted. During time 3,  $\frac{A_2+B_1}{\|A_2+B_1\|}$  is transmitted from the first antenna. For DoF purposes where we can ignore the noise terms, this scheme works by providing each receiver with sufficient information to decode its symbols.

Suppose due to the channel realization,  $A_2$  has a much higher power compared to  $B_1$ , i.e.  $\|A_2\| \gg \|B_1\|$ . The signal in time 3 is normalized to meet the power constraint and then transmitted. When receiver 2 removes  $A_2$ , it is left with  $B_1 + \tilde{z}_2[3]$  and the noise term may very well have a much higher power than  $B_1$ . Therefore,  $B_1$  is effectively eliminated and this results in a significant rate loss and thus a bigger gap.

To overcome this challenge, we incorporate quantization and allow the entire useful signal for each receiver to be quantized individually and then we create the XOR of the quantized bits as our transmit signal. This way we look at



many samples similar to  $A_2$  and  $B_1$ , and we effectively take into account the average power of such signals. In this manner, no signal will be eliminated due to have a lower power.

### B. Decoding

Upon completion of Phase 2 of block  $b$ , each receiver decodes the quantized signal. We know that as  $n \rightarrow \infty$ , this could be done with arbitrary small decoding error probability. Therefore, each receiver has access to  $\tilde{s}^{b,n} + z_Q^{b,n}$  where  $z_Q^{b,n}$  is the quantization noise with variance  $D$  which is independent of the transmit signals. Note that  $z_Q^b[t_1]$  and  $z_Q^b[t_2]$  are uncorrelated but not necessarily independent,  $t_1, t_2 = 1, 2, \dots, n$ ,  $t_1 \neq t_2$ .

Receiver 1 at the end of the  $n^{\text{th}}$  communication block, reconstructs signals by reversing the interleaving procedure described above, and removes  $y_{1,2}^{b,n}$  to obtain

$$\tilde{y}_{2,1:1}^{b,n} = y_{2,1:1}^{b,n} + \tilde{z}_Q^{b,n}, \quad (18)$$

here  $\tilde{z}_Q^{b,n}$  is the quantization noise with variance  $D$  which is independent of the transmit signals. Moreover,  $\tilde{z}_Q^b[\ell]_{\ell=1}^n$  is an independent sequence.

Note that since the messages are encoded at rate  $C_{2 \times 2}(1, 1 + D) - \epsilon$  for  $\epsilon > 0$ , if receiver one has access to  $y_{2,1:1}^{b,n}$  up to distortion  $D$ , it can recover  $w_1^b$  with arbitrary small decoding error probability as  $\epsilon \rightarrow 0$  and communication length goes to infinity. Thus, from  $y_{1,1:1}^{b,n}$  and  $\tilde{y}_{2,1:1}^{b,n}$ , receiver one can decode  $w_1^b$ ,  $b = 1, 2, \dots, n$ . Similar argument holds for receiver two.

An error may occur in either of the following steps: (1) if an error occurs in decoding message  $w_k^b$  provided required signals to the receiver,  $k = 1, 2$ ; (2) if an error occurs in quantizing  $\tilde{s}^{b,n}$ ; and (3) if an error occurs in decoding  $\tilde{s}^{b,n} + z_Q^{b,n}$  at either of the receivers,  $b = 1, 2, \dots, n$ . The probability of each one of such errors decreases exponentially in  $n$  (see [18], [25] and references therein). Using union bound and given that we have  $O(n^2)$  possible errors and the fact that each error probability decreases exponentially to zero, the total error probability goes to zero as  $n \rightarrow \infty$ .

### C. Achievable Rate

Using the achievable strategy described above, as  $n \rightarrow \infty$ , we can achieve a (symmetric) sum-rate point of

$$(R_1, R_2) = \left( \frac{C_{2 \times 2}(1, 5)}{2 + R_Q(4)/C_{2 \times 1}}, \frac{C_{2 \times 2}(1, 5)}{2 + R_Q(4)/C_{2 \times 1}} \right). \quad (19)$$

In Appendix A, we show that  $R_Q(4)/C_{2 \times 1} \leq 1$  for all values of  $P$ . Therefore, Phase 2 of each block has at most  $n$  time instants. Thus, we conclude that a (symmetric) sum-rate point of

$$(R_1, R_2) = \left( \frac{C_{2 \times 2}(1, 5)}{3}, \frac{C_{2 \times 2}(1, 5)}{3} \right), \quad (20)$$

is achievable.

## V. CONVERSE PROOF OF THEOREM 1

In this section, we provide the converse proof of Theorem 1. The converse consists of two main parts. In part 1, we show that the capacity region of the problem is included in the capacity region of a (stochastically) degraded BC, and in part 2, we derive an outer-bound on the capacity region of the degraded BC.

**Part 1:** We create the stochastically degraded BC as follows. We first provide the received signal of  $Rx_2$ , i.e.  $y_2[t]$ , to  $Rx_1$  as depicted in Fig. 5. Note that the resulting channel is *physically degraded*, and we know that for a physically degraded broadcast channel, feedback does not enlarge the capacity region [7]. Therefore, we can ignore the delayed knowledge of the channel state information at the transmitter (i.e. no CSIT assumption) and the correlation between the channel gains. The resulting channel is stochastically degraded.

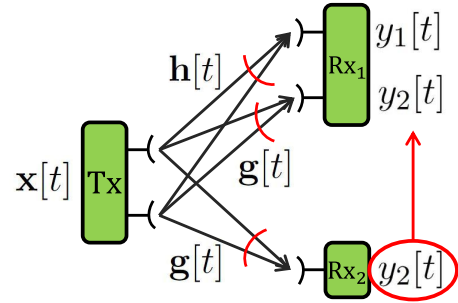


Fig. 5. By providing  $y_2^n$  to  $Rx_1$ , we create a physically degraded BC. We then ignore the delayed knowledge of the channel state information at the transmitter because for a physically degraded broadcast channel, feedback does not enlarge the capacity region. The resulting channel is stochastically degraded.

Thus, we form a *stochastically degraded* channel as shown in Fig. 5, and we formally define it in Definition 5 below. Let us denote an outer-bound on the capacity region of this channel by  $\mathcal{A}$ . The argument above shows that  $\mathcal{C} \subseteq \mathcal{A}$ .

**Definition 5.** The input-output relationship of a two-user stochastically degraded MIMO BC as depicted in Fig. 5, is given by

$$\tilde{y}_1[t] = \mathbf{H}[t]\mathbf{x}[t] + \mathbf{z}_1[t], \quad \tilde{y}_2[t] = \mathbf{g}^\top[t]\mathbf{x}[t] + z_2[t], \quad (21)$$

for

$$\mathbf{H}[t] = \begin{bmatrix} h_1[t] & h_2[t] \\ g_1[t] & g_2[t] \end{bmatrix}, \quad \mathbf{g}[t] = \begin{bmatrix} g_1[t] \\ g_2[t] \end{bmatrix}, \quad (22)$$

where  $h_j[t]$  and  $g_j[t]$  are distributed as i.i.d. (independent over time and from each other)  $\mathcal{CN}(0, 1)$  as described in Section II for  $j = 1, 2$ . We have

$$\tilde{\mathbf{y}}_1[t] = [\tilde{y}_{11}[t] \quad \tilde{y}_{12}[t]]^\top. \quad (23)$$

We assume  $\mathbf{z}_1[t] \in \mathbb{C}^{2 \times 1}$  and  $z_2[t] \sim \mathcal{CN}(0, 1)$ , and the transmit signal is subject to average power constraint  $P$ . We further assume that the transmitter has no knowledge of the channel gains besides their distributions (no CSIT assumption).

**Part 2:** In this part, we derive an outer-bound,  $\mathcal{A}$ , on the capacity region of the stochastically degraded broadcast channel.

For the stochastically degraded BC defined in Definition 5, suppose there exists encoders and decoders at the transmitter and receivers such that each message can be decoded at its corresponding receiver with arbitrary small decoding error probability.

$$\begin{aligned}
& n(R_1 + 2R_2 - 3\epsilon_n) \\
& \stackrel{\text{Fano}}{\leq} I(w_1; \tilde{y}_{11}^n, \tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) + 2I(w_2; \tilde{y}_2^n | \mathbf{H}^n, \mathbf{g}^n) \\
& \stackrel{(a)}{=} I(w_1; \tilde{y}_{11}^n, \tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) + I(w_2; \tilde{y}_{11}^n | \mathbf{H}^n, \mathbf{g}^n) \\
& \quad + I(w_2; \tilde{y}_{12}^n | \mathbf{H}^n, \mathbf{g}^n) \\
& = h(\tilde{y}_{11}^n, \tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) - h(\tilde{y}_{11}^n, \tilde{y}_{12}^n | w_1, w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \quad + h(\tilde{y}_{11}^n | \mathbf{H}^n, \mathbf{g}^n) - h(\tilde{y}_{11}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \quad + h(\tilde{y}_{12}^n | \mathbf{H}^n, \mathbf{g}^n) - h(\tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \stackrel{(b)}{=} h(\tilde{y}_{11}^n | \mathbf{H}^n, \mathbf{g}^n) - h(z_{11}^n | \mathbf{H}^n, \mathbf{g}^n) \\
& \quad + h(\tilde{y}_{12}^n | \mathbf{H}^n, \mathbf{g}^n) - h(z_{12}^n | \mathbf{H}^n, \mathbf{g}^n) \\
& \quad + h(\tilde{y}_{11}^n, \tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) - h(\tilde{y}_{11}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \quad - h(\tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \stackrel{(c)}{\leq} 2\mathbb{E} \log_2 \left[ 1 + \frac{P}{2} \mathbf{g}^\dagger \mathbf{g} \right] - I(\tilde{y}_{11}^n; \tilde{y}_{12}^n | w_2, \mathbf{H}^n, \mathbf{g}^n) \\
& \stackrel{(d)}{\leq} 2\mathbb{E} \log_2 \left[ 1 + \frac{P}{2} \mathbf{g}^\dagger \mathbf{g} \right] \stackrel{(5)}{=} 2C_{2 \times 1}, \tag{24}
\end{aligned}$$

where (a) follows from the fact that due to no CSIT assumption, we have

$$h(w_2 | \tilde{y}_{11}^n, \mathbf{H}^n, \mathbf{g}^n) \leq n\epsilon_n, \quad h(w_2 | \tilde{y}_{12}^n, \mathbf{H}^n, \mathbf{g}^n) \leq n\epsilon_n; \tag{25}$$

(b) holds since

$$\begin{aligned}
h(\tilde{y}_1^n | w_1, w_2, \mathbf{H}^n, \mathbf{g}^n) &= h(\tilde{y}_1^n | w_1, w_2, \mathbf{x}^n, \mathbf{H}^n, \mathbf{g}^n) \\
&= h(z_{11}^n, z_{12}^n | w_1, w_2, \mathbf{x}^n, \mathbf{H}^n, \mathbf{g}^n) \\
&= h(z_{11}^n | \mathbf{H}^n, \mathbf{g}^n) + h(z_{12}^n | \mathbf{H}^n, \mathbf{g}^n); \tag{26}
\end{aligned}$$

(c) follows from the results in [18]; and (d) follows from fact that mutual information is always positive. Dividing both sides by  $n$  and letting  $n \rightarrow \infty$ , we obtain the desired result. This completes the derivation of  $\mathcal{A}$ . Similarly, we can derive  $\mathcal{B}$ , and we have  $\mathcal{C} \subseteq \mathcal{A} \cap \mathcal{B}$  which completes the converse proof for Theorem 1.

## VI. GAP ANALYSIS

In this section, we evaluate the gap between our achievable rate-region and the outer-bound. We analytically prove that the gap is at most 1.81 bits/sec/Hz per user. We show that

$$(\mathcal{A} \cap \mathcal{B}) \ominus (1.81, 1.81) \subseteq \mathcal{R}(4), \tag{27}$$

where  $\mathcal{A}$  and  $\mathcal{B}$  are given in Definition 3, and  $\mathcal{R}(4)$  is given in (10).

Since the achievable rate region and the outer-bound (see Fig. 2) are formed by time sharing among the corresponding corner points (and thus, characterized by straight lines), we only need to consider the symmetric capacity,  $\mathcal{C}_{\text{SYM}}^2$ , defined in Definition 4.

We evaluate the gap between the inner-bound in (20), *i.e.*

$$(R_1, R_2) = \left( \frac{C_{2 \times 2}(1, 5)}{3}, \frac{C_{2 \times 2}(1, 5)}{3} \right). \tag{28}$$

and the symmetric point  $(\mathcal{C}_{\text{SYM}, 2\text{-user}}, \mathcal{C}_{\text{SYM}, 2\text{-user}})$ , obtained from Theorem 1.

A numerical evaluation of the gap between the sum-rate inner-bound and outer-bound is plotted in Fig. 6(a).

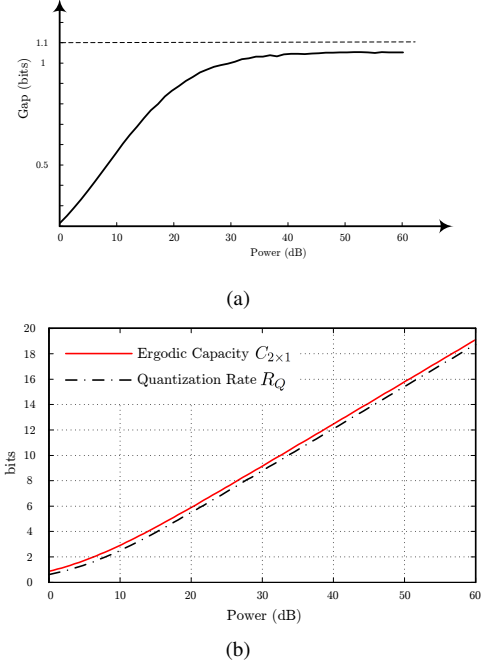


Fig. 6. (a) Numerical evaluation of the per-user gap between the sum-rate inner-bound and outer-bound; (b) Using numerical analysis, we can show that  $R_Q(3)/C_{2 \times 1} \leq 1$ .

To analyze the gap between the two bounds, we first study the gap between  $C_{2 \times 2}(1, 5)$  and  $2C_{2 \times 1}$ .

**Corollary 1.** Consider a MIMO point-to-point channel with 2 transmit antennas and 2 receive antennas as described in [18]. The only difference is that the additive noise at one antenna has variance 1 while the additive noise at the other antenna has variance  $(1 + D)$ . The ergodic capacity of this channel,  $C_{2 \times 2}(1, 1 + D)$ , satisfies

$$\begin{aligned}
& C_{2 \times 2}(1, 1 + D) \\
& \geq \left( \mathbb{E} \log_2 \det \left[ \mathbf{I}_2 + \frac{P}{2} \mathbf{H} \mathbf{H}^\dagger \right] - \log_2(1 + D) \right)^+. \tag{29}
\end{aligned}$$

The proof is provided at the top of the page in (30) where where (a) holds since the right hand side is obtained by evaluating the mutual information between the input and output, for a complex Gaussian input with covariance matrix  $\mathbb{E}[\mathbf{x}^\dagger \mathbf{x}] = P/2 \mathbf{I}_2$ .

Therefore, the gap between the sum-rate inner-bound and outer-bound can be upper-bounded by

$$\begin{aligned}
& \frac{4C_{2 \times 1}}{3} - \frac{2C_{2 \times 2}(1, 1 + D)}{3} \\
& \leq \frac{2(2C_{2 \times 1} - C_{2 \times 2} + \log_2(1 + D))}{3}, \tag{31}
\end{aligned}$$

$$\begin{aligned}
& C_{2 \times 2}(1, 1 + D) \\
& \stackrel{(a)}{\geq} \mathbb{E} \log_2 \det \left[ \mathbf{I}_2 + \frac{P}{2} \begin{bmatrix} h_{11} & h_{12} \\ h_{21}/\sqrt{1+D} & h_{22}/\sqrt{1+D} \end{bmatrix} \begin{bmatrix} h_{11}^\dagger & h_{21}^\dagger/\sqrt{1+D} \\ h_{12}^\dagger & h_{22}^\dagger/\sqrt{1+D} \end{bmatrix} \right] \\
& = \mathbb{E} \log_2 \det \left[ \mathbf{I}_2 + \frac{P}{2} \begin{bmatrix} |h_{11}|^2 + |h_{12}|^2 & (h_{11}h_{21}^\dagger + h_{12}h_{22}^\dagger)/\sqrt{1+D} \\ (h_{11}^\dagger h_{21} + h_{12}^\dagger h_{22})/\sqrt{1+D} & (|h_{21}|^2 + |h_{22}|^2)/(1+D) \end{bmatrix} \right] \\
& = \mathbb{E} \log_2 \det \left[ \begin{array}{cc} 1 + \frac{P}{2}|h_{11}|^2 + |h_{12}|^2 & \frac{P}{2}(h_{11}h_{21}^\dagger + h_{12}h_{22}^\dagger)/\sqrt{1+D} \\ \frac{P}{2}(h_{11}^\dagger h_{21} + h_{12}^\dagger h_{22})/\sqrt{1+D} & 1 + \frac{P}{2}(|h_{21}|^2 + |h_{22}|^2)/(1+D) \end{array} \right] \\
& \geq \mathbb{E} \log_2 \det \left[ \begin{array}{cc} 1 + \frac{P}{2}|h_{11}|^2 + |h_{12}|^2 & \frac{P}{2}(h_{11}h_{21}^\dagger + h_{12}h_{22}^\dagger)/\sqrt{1+D} \\ \frac{P}{2}(h_{11}^\dagger h_{21} + h_{12}^\dagger h_{22})/\sqrt{1+D} & (1 + \frac{P}{2}(|h_{21}|^2 + |h_{22}|^2))/(1+D) \end{array} \right] \\
& = \mathbb{E} \log_2 \det \underbrace{\left[ \mathbf{I}_2 + \frac{P}{2} \mathbf{H} \mathbf{H}^\dagger \right]}_{\triangleq C_{2 \times 2}} - \log_2(1 + D), \tag{30}
\end{aligned}$$

where

$$C_{2 \times 2} \triangleq \mathbb{E} \log_2 \det \left[ \mathbf{I}_2 + \frac{P}{2} \mathbf{H} \mathbf{H}^\dagger \right]. \tag{32}$$

**Remark 6.** While in this work we evaluate the gap for Rayleigh fading channels, our expressions for the inner-bounds and the outer-bounds hold for general i.i.d. channel realizations. The challenge to evaluate the gap for distributions other than Rayleigh fading, arises from determining the optimal covariance matrix for the transmit signal and evaluating the capacity result as discussed in [18] Section 4.1.

For  $P \leq 2$ , the sum-rate outer-bound is smaller than 2 bits (smaller than the gap itself). So, we assume  $P > 2$ . We have

$$\begin{aligned}
& 2C_{2 \times 1} - C_{2 \times 2} \\
& = 2\mathbb{E} \log_2 \left[ 1 + \frac{P}{2} \mathbf{g}^\dagger \mathbf{g} \right] - \mathbb{E} \log_2 \det \left[ \mathbf{I}_2 + \frac{P}{2} \mathbf{H} \mathbf{H}^\dagger \right] \\
& = 2\mathbb{E} \log_2 \left[ \frac{2}{P} + \mathbf{g}^\dagger \mathbf{g} \right] + 2\log_2 \left( \frac{P}{2} \right) \\
& \quad - \mathbb{E} \log_2 \det \left[ \frac{2}{P} \mathbf{I}_2 + \mathbf{H} \mathbf{H}^\dagger \right] - \log_2 \left( \frac{P^2}{4} \right) \\
& = 2\mathbb{E} \log_2 \left[ \frac{2}{P} + \mathbf{g}^\dagger \mathbf{g} \right] - \mathbb{E} \log_2 \det \left[ \frac{2}{P} \mathbf{I}_2 + \mathbf{H} \mathbf{H}^\dagger \right] \\
& \leq 2\mathbb{E} \log_2 [1 + \mathbf{g}^\dagger \mathbf{g}] - \mathbb{E} \log_2 \det [\mathbf{H} \mathbf{H}^\dagger]. \tag{33}
\end{aligned}$$

Thus, we have

$$\begin{aligned}
& \frac{4C_{2 \times 1}}{3} - \frac{2C_{2 \times 2}(1, 1 + D)}{3} \\
& \leq \frac{2}{3} (2\mathbb{E} \log_2 [1 + \mathbf{g}^\dagger \mathbf{g}] - \mathbb{E} \log_2 \det [\mathbf{H} \mathbf{H}^\dagger] + \log_2(1 + D)) \\
& \leq 3.62, \tag{34}
\end{aligned}$$

which implies that the gap is less than 1.81 bits per user independent of power  $P$ . We could also use numerical analysis to evaluate the gap. In particular, using numerical analysis,

we can show that  $R_Q(3)/C_{2 \times 1} \leq 1$  (see Fig. 6(b)). We have plotted

$$\frac{4C_{2 \times 1}}{3} - \frac{2C_{2 \times 2}(1, 1 + D)}{3} \tag{35}$$

in Fig. 6(a) for  $D = 4$ , and for  $P$  between 0 dB and 60 dB. As we can see, the sum-rate inner-bound and outer-bound are at most 2.2 bits (or 1.1 per user) away from each other for  $P$  between 0 dB and 60 dB.

## VII. SYMMETRIC CAPACITY OF THE $K$ -USER MISO BC WITH DELAYED CSIT

Now that we have presented our results for the two-user multiple-input single-output complex Gaussian broadcast channel with delayed CSIT, we consider the  $K$ -user setting as depicted in Fig. 7 (left). The channel matrix from the transmitter to the receivers is denoted by  $\mathbf{H} \in \mathbb{C}^{K \times K}$ , where the entries of  $\mathbf{H}[t]$  are distributed as i.i.d.  $\mathcal{CN}(0, 1)$  (independent across time, antenna, and users). The transmit signal  $\mathbf{x}[t] \in \mathbb{C}^{K \times 1}$  is subject to average power constraint  $P$ , i.e.  $\mathbb{E} [\mathbf{x}^\dagger[t] \mathbf{x}[t]] \leq P$  for  $P > 0$ . The noise processes are independent from the transmit signal and are distributed i.i.d. as  $z_k[t] \sim \mathcal{CN}(0, 1)$ . The input-output relationship in this channel is given by

$$[y_1[t] \quad \dots \quad y_K[t]]^\top = \mathbf{H}[t] \mathbf{x}[t] + [z_1[t] \quad \dots \quad z_K[t]]. \tag{36}$$

### A. Outer-bound

The derivation of the outer-bound in Theorem 2 is based on creating a degraded MIMO BC where  $R_{x_\ell}$  has access to the received signal of  $R_{x_j}$  for  $j \geq \ell$  similar to the converse proof provided in [2],  $\ell = 1, 2, \dots, K$ . This channel is physically degraded, and thus feedback does not enlarge the capacity region [7]. Therefore, we can ignore the delayed knowledge of the channel state information at the transmitter (i.e. no CSIT assumption) and the correlation between the channel gains. See



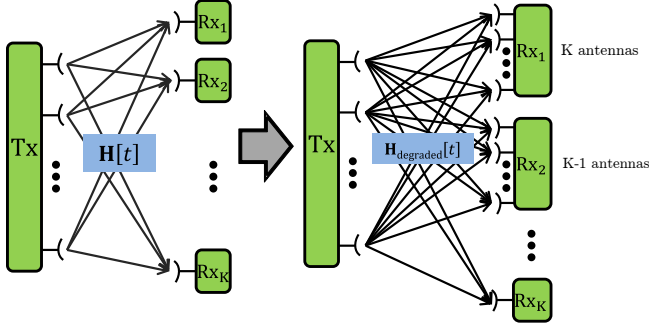


Fig. 7. Left:  $K$ -user Multiple-Input Single-Output (MISO) Complex Gaussian Broadcast Channel; and right: degraded MIMO BC.

Fig. 7 for a depiction. In this degraded channel, we denote the output signal of user  $j$  at time  $t$  by  $\tilde{y}_j^3$  where

$$\tilde{\mathbf{y}}_j[t] = [y_j[t] \ y_{j+1}[t] \ \dots \ y_K[t]]^\top. \quad (37)$$

For the resulting multiple-input multiple-output (MIMO) broadcast channel denote the channel matrix by  $\mathbf{H}_{\text{degraded}}^n$ . The capacity region of the  $K$ -user MISO BC with delayed CSIT (Fig. 7 left) is included in the capacity region of the degraded MIMO BC with no CSIT (Fig. 7 right). Under no CSIT assumption, the following result is known for the stochastically degraded MIMO BC described above.

**Lemma 2** ([26]). *Consider the  $K$ -user stochastically degraded MIMO BC described above and depicted in Fig. 7 (right). Then under no CSIT assumption, for  $j = 1, \dots, K-1$ , we have*

$$\begin{aligned} & h(\tilde{\mathbf{y}}_j^n | w_{j+1}, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \geq \frac{K-j+1}{K-j} h(\tilde{\mathbf{y}}_{j+1}^n | w_{j+1}, \dots, w_K, \mathbf{H}_{\text{degraded}}^n). \end{aligned} \quad (38)$$

Using Lemma 2, similar to the argument presented in (24), we get

$$\begin{aligned} & n \left( R_1 + \frac{K}{K-1} R_2 + \frac{K}{K-2} R_2 + \dots + K R_K - \epsilon_n \right) \\ & = I(w_1; \tilde{\mathbf{y}}_1^n | w_2, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad + \frac{K}{K-1} I(w_2; \tilde{\mathbf{y}}_2^n | w_3, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad + \dots + K I(w_K; \tilde{\mathbf{y}}_K^n | \mathbf{H}_{\text{degraded}}^n) \\ & \leq h(\tilde{\mathbf{y}}_1^n | w_2, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad - \frac{K}{K-1} h(\tilde{\mathbf{y}}_2^n | w_2, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad + \frac{K}{K-1} h(\tilde{\mathbf{y}}_2^n | w_3, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad - \frac{K}{K-2} h(\tilde{\mathbf{y}}_3^n | w_3, \dots, w_K, \mathbf{H}_{\text{degraded}}^n) \\ & \quad + \dots + K h(\tilde{\mathbf{y}}_K^n | \mathbf{H}_{\text{degraded}}^n) \\ & \stackrel{\text{Lemma 2}}{\leq} K h(\tilde{\mathbf{y}}_K^n | \mathbf{H}_{\text{degraded}}^n) \\ & \leq K \text{E} \log_2 \left[ 1 + \frac{P}{K} \mathbf{g}^\dagger \mathbf{g} \right] \stackrel{(2)}{=} K C_{K \times 1}, \end{aligned} \quad (39)$$

<sup>3</sup>for user  $K$ , we have  $\tilde{\mathbf{y}}_K[t] = \tilde{y}_K[t] = y_K[t]$

where  $\mathbf{g}$  is a  $K$  by 1 vector where entries are i.i.d.  $\mathcal{CN}(0, 1)$ . Then, for the symmetric sum-rate point we have  $R_1 = R_2 = \dots = R_K = R$ , which gives us

$$\left( \sum_{j=1}^K j^{-1} \right) R \leq K C_{K \times 1}, \quad (40)$$

which implies that

$$C_{\text{SYM}, K\text{-user}} \leq \frac{C_{K \times 1}}{\sum_{j=1}^K j^{-1}}. \quad (41)$$

This completes the converse proof of Theorem 2.

## B. Achievability

In this subsection, we focus on achievability and discuss the extension of our achievability results to the  $K$ -user MISO BC with delayed CSIT. The transmission strategy follows the steps of [2], but similar to the two-user case, some additional ingredients are needed to make an approximately optimal scheme. We first demonstrate the techniques and the required ingredients for the three-user MISO BC with delayed CSIT. Then, we discuss how the result is extended to the general case.

**Transmission strategy for the three-user MISO BC:** Each communication block includes three phases where Phase 1 and Phase 2 each have 3 segments, and Phase 3 has two segments. We focus on a specific block  $b$ . Fix  $\epsilon > 0$ .

• During segments 1, 2, and 3 of Phase 1 of block  $b$ , the message of user  $k$  ( $w_k^b$ ) is encoded as  $\mathbf{x}_k^{b,n}$  at rate

$$C_{3 \times 3}(1, 5, 11) - \epsilon, \quad (42)$$

where any element of this codeword is drawn i.i.d. from  $\mathcal{CN}(0, P/3\mathbf{I}_3)^4$ ; and  $C_{3 \times 3}(1, 5, 11)$  is given in Definition 2.

**Remark 7.** *The reason for the encoding rate of (42) becomes apparent as we describe the achievability strategy. Basically the quantization noise accumulates on top of the previous noises throughout some of the phases and thus, we need to adjust the encoding rate accordingly.*

Transmitter communicates these codewords and receiver one obtains  $y_{1,1:j}^{b,n}$ , receiver two obtains  $y_{2,1:j}^{b,n}$ , and receiver three obtains  $y_{3,1:j}^{b,n}$ ,  $j = 1, 2, 3$ .

**Definition 6.** *We define  $s_{k,p;j}^{b,n}$  to be the noiseless version of  $y_{k,p;j}^{b,n}$  for appropriate choice of indices. This is similar to Definition 2 for the two-user case.*

Note that at the end of the third segment of the first phase, transmitter has access to  $s_{1,1;j}^{b,n}$ ,  $s_{2,1;j}^{b,n}$ , and  $s_{3,1;j}^{b,n}$ . Due to the rate given in (42), if  $s_{2,1:1}^{b,n}$  and  $s_{3,1:1}^{b,n}$  are provided to receiver one with distortions 4 and 8 respectively, then receiver one will be able to decode its message with arbitrary small decoding error probability. Similarly receiver two is interested in  $s_{1,1:2}^{b,n}$  and  $s_{3,1:2}^{b,n}$ ; and receiver three is interested in  $s_{1,1:3}^{b,n}$  and  $s_{2,1:3}^{b,n}$ .

Based on the discussion above and the available signal at each receiver, we observe that  $s_{2,1:1}^{b,n} + s_{1,1:2}^{b,n}$  is of common

<sup>4</sup> $\mathbf{I}_3$  is the  $3 \times 3$  identity matrix.

interest to receivers one and two. Similarly,  $s_{3,1:1}^{b,n} + s_{1,1:3}^{b,n}$  is of common interest to receivers one and three; and  $s_{2,1:3}^{b,n} + s_{3,1:2}^{b,n}$  is of common interest to receivers two and three. Therefore, the goal would be to deliver these signals to their interested receivers during the following phases.

- Communication during Phase 2 Segment 1: Consider the communication during another block  $b'$ , and the corresponding signal  $s_{2,1:1}^{b',n} + s_{1,1:2}^{b',n}$ . Using Lemma 1, we quantize  $s_{2,1:1}^{b',n} + s_{1,1:2}^{b',n}$  and  $s_{2,1:1}^{b',n} + s_{1,1:2}^{b',n}$  at distortion 4, and we create the XOR of the resulting bits<sup>5</sup>. Then these bits will be encoded as  $\mathbf{v}_{2:1}^{b,n}$  at rate  $C_{3 \times 2}(1, 5) - \epsilon$ , and will be transmitted during the first segment of Phase 2 where  $C_{3 \times 2}(1, 5)$  is given in Definition 2. Receiver three obtains  $y_{3,2:1}^{b,n}$  that is of interest of users one and two.

- Communication during Phase 2 Segment 2: Consider the communication during another block  $b'$ , and the corresponding signal  $s_{3,1:1}^{b',n} + s_{1,1:3}^{b',n}$ . Using Lemma 1, we quantize  $s_{3,1:1}^{b',n} + s_{1,1:3}^{b',n}$  and  $s_{3,1:1}^{b',n} + s_{1,1:3}^{b',n}$  at distortion 4 and we create the XOR of the resulting bits. Then these bits will be encoded at rate  $C_{3 \times 2}(1, 5) - \epsilon$  denoted by  $\mathbf{v}_{2,2:2}^{b,n}$  and will be transmitted during the second segment of Phase 2. Receiver two obtains  $y_{2,2:2}^{b,n}$  that is of interest of users one and three.

- Communication during Phase 2 Segment 3: Consider the communication during another block  $b'$ , and the corresponding signal  $s_{2,1:3}^{b',n} + s_{3,1:2}^{b',n}$ . Using Lemma 1, we quantize  $s_{2,1:3}^{b',n} + s_{3,1:2}^{b',n}$  and  $s_{2,1:3}^{b',n} + s_{3,1:2}^{b',n}$  at distortion 4 and we create the XOR of the resulting bits. Then these bits will be encoded at rate  $C_{3 \times 2}(1, 5) - \epsilon$  denoted by  $\mathbf{v}_{2,3:3}^{b,n}$  and will be transmitted during the third segment of Phase 3. Receiver one obtains  $y_{1,2:3}^{b,n}$  that is of interest of users two and three.

We now create two signals that are of interest of all three receivers:

$$\begin{aligned} & \frac{1}{\sqrt{2}} s_{3,2:1}^{b,n} + \frac{1}{\sqrt{3}} s_{2,2:2}^{b,n} + \frac{1}{\sqrt{6}} s_{1,2:3}^{b,n}, \\ & \frac{1}{\sqrt{6}} s_{3,2:1}^{b,n} + \frac{-1}{\sqrt{3}} s_{2,2:2}^{b,n} + \frac{1}{\sqrt{2}} s_{1,2:3}^{b,n}. \end{aligned} \quad (43)$$

**Remark 8.** The choice of coefficients in (43) is such that all users are assigned equal powers and the linear combinations at the receivers remain independent. We also note that using such coefficients results in a  $1/3$  power loss for each user.

Note that any receiver has access to both signals in (43), then it can recursively recover the signals it is interested in, and decode the intended message.

- Communication during Phase 3 Segment 1: Using Lemma 1, we quantize

$$\left( \frac{1}{\sqrt{2}} s_{3,2:1}^{b,n} + \frac{1}{\sqrt{3}} s_{2,2:2}^{b,n} + \frac{1}{\sqrt{6}} s_{1,2:3}^{b,n} \right)$$

at distortion 5. Then these quantized bits will be encoded at rate  $C_{3 \times 1} - \epsilon$  and transmitted during the first segment of Phase 3.

<sup>5</sup>We note that we need these signals to be distributed independently, to handle this issue, we can incorporate an interleaving step similar to Section IV.

- Communication during Phase 3 Segment 2: Using Lemma 1, we quantize

$$\left( \frac{1}{\sqrt{6}} s_{3,2:1}^{b,n} + \frac{-1}{\sqrt{3}} s_{2,2:2}^{b,n} + \frac{1}{\sqrt{2}} s_{1,2:3}^{b,n} \right)$$

at distortion 5. Then these quantized bits will be encoded at rate  $C_{3 \times 1} - \epsilon$  and transmitted during the second segment of Phase 3.

Using the achievability strategy described above and for  $n \rightarrow \infty$  and  $\epsilon \rightarrow 0$ , it can be shown that a per-user rate of  $\frac{2}{11} C_{3 \times 3}(1, 5, 11)$  is achievable.

**Transmission strategy for the  $K$ -user MISO BC:** Now that we have described the transmission strategy for the 3-user MISO BC with delayed CSIT, we explain the transmission strategy for the general case  $K > 3$ . As mentioned before, the transmission strategy follows the steps of [2]. However, we highlight the key differences that are needed to derive near optimal results.

The scheme includes  $K$  phases. For simplicity, we do not go into details of the interleaving process. At the beginning of Phase  $j$ , transmitter has access to signals that are of interest of  $j$  receivers,  $j = 1, 2, \dots, K$ . There are a total of  $(K - j + 1) \binom{K}{j}$  such signals. Phase  $j$  has  $\binom{K}{j}$  segments. Consider a subset  $\mathcal{S}$  of the receivers where  $|\mathcal{S}| = j$ . Transmitter accumulates a total of  $(K - j + 1)$  signals that are of interest of all receivers in  $\mathcal{S}$  using  $(K - j + 1)$  different blocks (this is similar to Phase 2 for the 3-user MISO BC with delayed CSIT). Similar to Lemma 1, transmitter quantizes these signals at distortion  $(j + 2)$ , and creates the XOR of the resulting bits. The resulting bits are encoded at rate  $C_{K \times j}(1, 5, \dots, 1 + (j - 1)(j + 2)) - \epsilon$  and communicated during segment  $t_{\mathcal{S}}$  of Phase  $j$ .

**Remark 9.** The noise variance  $1 + (j - 1)(j + 2)$  results from the fact that each receiver has to solve  $(j - 1)$  equations of the signals that he is interested in. This step results in boosting up the noise variance.

Consider any subset  $\mathcal{S}'$  of the receivers where  $|\mathcal{S}'| = j + 1$ . Upon completion of Phase  $j$ , we observe that any receiver in  $\mathcal{S}'$  has a signal that is simultaneously of common interest of all other receivers in  $\mathcal{S}'$ . Transmitter has access to this signal (up to noise) using delayed knowledge of the channel state information. Transmitter forms  $j$  random linear combinations of such signals for each subset  $\mathcal{S}'$  where  $|\mathcal{S}'| = j + 1$ . Then transmitter creates  $j \binom{K}{j+1}$  signals that are simultaneously of interest of  $j + 1$  receivers. These signals (after being quantized) will be delivered in Phases  $j + 1, j + 2, \dots, K$ . The rest of the scheme is identical to that of [2] and is omitted due to space limitations.

Recursively solving the achievable rate over  $K$  phases, we can show that a per-user rate of

$$\frac{C_{K \times K}(1, 5, \dots, 1 + (j - 1)(j + 2), \dots, 1 + (K - 1)(K + 2))}{K \sum_{j=1}^K j^{-1}} \quad (44)$$

is achievable.

We are now ready to present the approximate capacity of the  $K$ -user MISO BC with delayed CSIT.

**Corollary 2.** *The capacity region of the  $K$ -user MIMO BC with delayed CSIT is within at most  $2\log_2(K+2)$  bits per user of the region described by:*

$$0 \leq R_{\pi_1} + \frac{|\mathcal{J}|}{|\mathcal{J}|-1} R_{\pi_2} + \frac{|\mathcal{J}|}{|\mathcal{J}|-2} R_{\pi_3} + \dots + |\mathcal{J}| R_{\pi_{|\mathcal{J}|}} \leq |\mathcal{J}| C_{K \times 1}, \quad \forall \mathcal{J} \subseteq \{1, 2, \dots, K\}, \quad (45)$$

for all permutations  $\pi$  of  $\mathcal{J}$ .

The proof is based on the results we provided in this section and thus, we only provide a proof sketch here. The derivation of the bounds is similar to that of Section VII-A for a degraded MIMO BC with  $K$  transmit antennas and  $|\mathcal{J}|$  receivers.

To obtain the achievability, we use the transmission strategy that we described in this section with a small modification. To explain why a modification is needed, consider  $|\mathcal{J}| = 2$ . Then we obtain  $\binom{K}{2}$  pair of bounds similar to:

$$\begin{aligned} R_1 + 2R_2 &\leq 2C_{K \times 1}, \\ 2R_1 + R_2 &\leq 2C_{K \times 1}. \end{aligned} \quad (46)$$

These bounds differ from the bounds in Sections IV and V since they have  $C_{K \times 1}$  rather than  $C_{2 \times 1}$  on the right hand side. Thus our transmission strategy needs to be modified to capture this difference.

Fix  $\mathcal{J}$ , then from (45) we obtain  $|\mathcal{J}|!$  bounds that will define a symmetric sum-rate point (symmetric for the users in  $\mathcal{J}$ ) of

$$R_i = \begin{cases} 0, & \text{if } i \notin \mathcal{J}, \\ \left( \sum_{j=1}^{|\mathcal{J}|} j^{-1} \right)^{-1} C_{K \times 1}, & \text{if } i \in \mathcal{J}. \end{cases} \quad (47)$$

It is sufficient to show we can achieve to within a constant gap of all such corner points.

In this section, we presented the transmission strategy of the symmetric sum-rate point of  $K$ -user MISO BC with delayed CSIT. It is easy to modify this result to achieve within a constant gap of the corner point in (47). Start with a  $|\mathcal{J}|$ -user MISO BC and re-write the transmission strategy of this section for that problem. Then, we need to modify the scheme by replacing  $C_{|\mathcal{J}| \times \ell}$  whenever it appears with  $C_{K \times \ell}$ ,  $1 \leq \ell \leq |\mathcal{J}| \leq K$  and the result would follow. Finally, we note that using an argument similar to the one presented in Section VI, the gap takes its maximum value at the symmetric sum-rate point of the  $K$ -user setting. Thus, the gap is at most  $2\log_2(K+2)$  bits per user.

### C. Gap Analysis

In Fig. 8, we have plotted the numerical analysis of the per user gap between the inner-bound and the outer-bound of Theorem 2 for  $P$  between 0 dB and 60 dB and  $K = 2, 3, 5, 10, 20$ .

For the per user gap, we analytically show that the outer-bound and lower-bound given in Theorem 2 are at most  $2\log_2(K+2)$  away from each other. However as depicted in Fig. 8, we can see that the per-user gap is in fact smaller than  $2\log_2(K+2)$ . The argument is similar to the one we presented in (30). To derive the gap, we first increase the noise term at

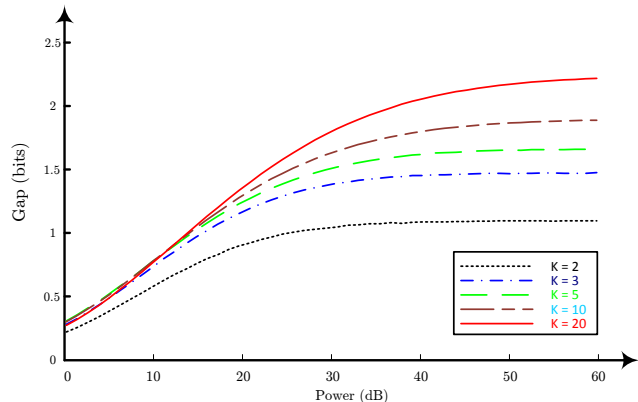


Fig. 8. Numerical evaluation of the per user gap between the inner-bound and the outer-bound in Theorem 2 for  $P$  between 0 dB and 60 dB and  $K = 2, 3, 5, 10, 20$ .

all receive antennas to  $1 + (K-1)(K+2)$ . Then similar to (30), we obtain

$$\begin{aligned} C_{K \times K} (1, 5, \dots, 1 + (j-1)(j+2), \dots, 1 + (K-1)(K+2)) \\ \geq C_{K \times K} - K \log_2 (1 + (K-1)(K+2)). \end{aligned} \quad (48)$$

Using (48) and considering the gap between  $C_{K \times K}$  and  $K C_{K \times 1}$ , we can show that the gap between the achievable symmetric rate and the symmetric capacity is bounded by  $2\log_2(K+2)$ .

## VIII. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we studied the capacity region of the multiple-input single-output complex Gaussian Broadcast Channels with delayed CSIT. We showed that a modification of the scheme introduced in [2], can be applied in the finite SNR regime to obtain an inner-bound that is within  $2\log_2(K+2)$  bits of the outer-bound. Therefore the gap scales as the logarithm of the number of users. This happens due to noise accumulation during the transmission strategy. Thus, an interesting future direction would be to figure out whether there exists a transmission strategy that results in constant gap (independent of channel parameters, transmission power and number of users) approximation of the capacity region.

Another direction is to consider a two-user MISO BC with delayed CSIT where the noise processes and the channel gains are not distributed as i.i.d. random variables. For example, consider the scenario where the noise processes have different variances. This model captures the scenario where users are located at different distances to the base station. For this setting, even the (generalized) DoF region is not known.

## APPENDIX A

DETERMINING  $D$  SUCH THAT  $R_Q(D)/C_{2 \times 1} \leq 1$ 

As mentioned in Section VI, we are interested in  $P > 2$ . Using Jensen's inequality, we have

$$\begin{aligned} R_Q(D) &= \mathbb{E} \left[ \log_2 \left( 1 + \frac{P}{2D} (\|\mathbf{g}\|_2^2 + \|\mathbf{h}\|_2^2) \right) \right] \\ &\leq \log_2 \left( 1 + \frac{P}{2D} \mathbb{E} [\|\mathbf{g}\|_2^2 + \|\mathbf{h}\|_2^2] \right) = \log_2 \left( 1 + \frac{2P}{D} \right). \end{aligned} \quad (49)$$

Moreover, from [18], we have

$$\begin{aligned} C_{2 \times 1} &= \int_0^\infty \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &= \int_0^1 \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &+ \int_1^\infty \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &= \sum_{m=1}^\infty \int_{2^{-m}}^{2^{1-m}} \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &+ \int_1^2 \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &+ \sum_{j=1}^{45} \int_{2^{+1(j-1)}}^{2^{+1j}} \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &+ \int_{6.5}^\infty \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &\stackrel{(a)}{\geq} \log_2(1 + P/2) \int_0^1 \lambda e^{-\lambda} d\lambda - \underbrace{\sum_{m=1}^\infty m \int_{2^{-m}}^{2^{1-m}} \lambda e^{-\lambda} d\lambda}_{<0.4} \\ &+ \log_2(1 + P/2) \int_1^2 \lambda e^{-\lambda} d\lambda \\ &+ \log_2(1 + P/2) \int_2^{6.5} \lambda e^{-\lambda} d\lambda \\ &+ \underbrace{\sum_{j=1}^{45} \log_2 \left[ \frac{1 + (2 + .1(j-1))P/2}{1 + P/2} \right] \int_{2^{+1(j-1)}}^{2^{+1j}} \lambda e^{-\lambda} d\lambda}_{>0.4} \\ &+ \int_{6.5}^\infty \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &> \log_2(1 + P/2) \int_0^\infty \lambda e^{-\lambda} d\lambda = \log_2(1 + P/2). \end{aligned} \quad (50)$$

where (a) holds since

$$\begin{aligned} &\sum_{m=1}^\infty \int_{2^{-m}}^{2^{1-m}} \log_2(1 + P\lambda/2) \lambda e^{-\lambda} d\lambda \\ &\geq \sum_{m=1}^\infty \int_{2^{-m}}^{2^{1-m}} \log_2(1 + 2^{-m}P/2) \lambda e^{-\lambda} d\lambda \\ &\geq \sum_{m=1}^\infty \int_{2^{-m}}^{2^{1-m}} [\log_2(1 + P/2) - \log_2(2^m)] \lambda e^{-\lambda} d\lambda \\ &= \log_2(1 + P/2) \int_0^\infty \lambda e^{-\lambda} d\lambda - \sum_{m=1}^\infty m \int_{2^{-m}}^{2^{1-m}} \lambda e^{-\lambda} d\lambda, \end{aligned} \quad (51)$$

and  $\sum_{m=1}^\infty m \int_{2^{-m}}^{2^{1-m}} \lambda e^{-\lambda} d\lambda$  converges since

$$\left\{ m \int_{2^{-m}}^{2^{1-m}} \lambda e^{-\lambda} d\lambda \right\}_{m=1}^\infty, \quad (52)$$

is a Cauchy sequence. Thus, we have  $R_Q(4)/C_{2 \times 1} < 1$ .

## REFERENCES

- [1] A. Vahid, M. Maddah-Ali, and A. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," in *51st Annual Allerton Conference on Communication, Control, and Computing*, 2013.
- [2] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4418–4431, 2012.
- [3] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in *International Symposium on Information Theory (ISIT)*, p. 174, 2004.
- [4] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.
- [5] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (corresp.)," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, 1984.
- [6] X. Yi and D. Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1–11, 2013.
- [7] A. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 379–381, 1978.
- [8] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSIT," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 199–203, IEEE, 2011.
- [9] R. Tandon, S. Mohajer, H. V. Poor, and S. Shamai, "Degrees of freedom region of the MIMO interference channel with output feedback and delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1444–1457, 2013.
- [10] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8336–8356, 2013.
- [11] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6093–6130, 2014.
- [12] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Communication through collisions: Opportunistic utilization of past receptions," in *Proceedings IEEE INFOCOM*, pp. 2553–2561, IEEE, 2014.
- [13] A. Vahid and R. Calderbank, "Impact of local delayed CSIT on the capacity region of the two-user interference channel," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 2421–2425, IEEE, 2015.
- [14] A. Vahid and R. Calderbank, "Two-user erasure interference channels with local delayed CSIT," *arXiv preprint arXiv:1503.03449*, 2012.
- [15] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, 2011.
- [16] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2788–2812, 2012.
- [17] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4566–4592, 2010.
- [18] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [19] M. Kobayashi, S. Yang, D. Gesbert, and X. Yi, "On the degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 2501–2505, IEEE, 2012.
- [20] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," *arXiv preprint arXiv:1205.3474*, 2012.

- [21] J. Chen, P. Elia, and S. A. Jafar, "On the vector broadcast channel with alternating CSIT: A topological perspective," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2579–2583, IEEE, 2014.
- [22] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Transactions on Information Theory*, vol. 42, no. 4, pp. 1152–1159, 1996.
- [23] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, 2002.
- [24] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [25] A. Gersho, "Asymptotically optimal block quantization," *IEEE Transactions on Information Theory*, vol. 25, no. 4, pp. 373–380, 1979.
- [26] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5354–5374, 2012.