



**HAL**  
open science

# On the compound broadcast channel: Multiple description coding and interference decoding

Meryem Benammar, Pablo Piantanida, Shlomo Shamai

► **To cite this version:**

Meryem Benammar, Pablo Piantanida, Shlomo Shamai. On the compound broadcast channel: Multiple description coding and interference decoding. *IEEE Transactions on Information Theory*, 2020, 66 (1), pp.38-64. 10.1109/tit.2019.2942615 . hal-03040015v1

**HAL Id: hal-03040015**

**<https://hal.science/hal-03040015v1>**

Submitted on 21 Sep 2020 (v1), last revised 4 Dec 2020 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Compound Broadcast Channel: Multiple Description Coding and Interference Decoding

Meryem Benammar and Pablo Piantanida and Shlomo Shamai (Shitz)

## Abstract

This work investigates the general two-user Compound Broadcast Channel (BC) where an encoder wishes to transmit common and private messages to two receivers while being oblivious to two possible channel realizations controlling the communication. The focus is on the characterization of the largest achievable rate region by resorting to more evolved *encoding* and *decoding* techniques than the conventional coding for the standard BC. The role of the decoder is first explored, and an achievable rate region is derived based on the principle of “Interference Decoding” (ID) where each receiver decodes its intended message and chooses to (non-uniquely) decode or not the interfering message. This inner bound is shown to be capacity achieving for a class of non-trivial compound BEC/BSC broadcast channels while the worst-case of Marton’s inner bound –based on “Non Interference Decoding” (NID)– fails to achieve the capacity region. The role of the encoder is then studied, and an achievable rate region is derived based on “Multiple Description” (MD) coding where the encoder transmits a common as well as multiple dedicated private descriptions to the many instances of the users channels. It turns out that MD coding outperforms the single description scheme –Common Description (CD) coding– for a class of compound Multiple Input Single Output Broadcast Channels (MISO BC).

The material in this paper was partially published in the IEEE Information Theory Workshop, Seville, September 9-13, 2013 and in the IEEE Information Theory Workshop, Tasmania, Australia, 2-5 November 2014. This research was partially supported by the FP7 Network of Excellence in Wireless COMMunications NEWCOM#.

Meryem Benammar and Pablo Piantanida are with the Department of Telecommunications, SUPELEC, 91192 Gif-sur-Yvette, France (e-mail: {meryem.benammar, pablo.piantanida}@supelec.fr).

Shlomo Shamai (Shitz) is with the Department of Electrical Engineering, Technion, Haifa, Israel, (e-mail: sshlomo@ee.technion.ac.il).

## I. INTRODUCTION

The two-user Broadcast Channel (BC) –as first introduced by Cover in [1]– consists of an encoder transmitting both a common message and two private messages to two users. Following this seminal work, intensive research was undertaken to characterize the capacity region of this setting for which the key feature in designing optimal codes is to allow for an efficient interference mitigation. In this work, we study the general two-user Compound BC where an encoder wishes to communicate one common message and two private messages to two users who can each observe one of many output channel statistics. The actual channel controlling the communication is unknown at the transmit side but assumed to remain constant during the communication and belongs to a known set of possible channels. Coding successfully for such a setting requires that the encoder must guarantee –whatever the channel realizations– reliable communication. Thus, it is well understood that the compound BC is equivalent to a BC with multiple users and common information. Our aim is to improve the understanding of how interference should be dealt within the current setting where both channel uncertainty and interference are coupled. To this end, we study alternative encoding and decoding techniques to the usual coding schemes that were proved to be capacity achieving for some broadcast channels.

Let us first briefly discuss the optimal coding schemes for the two-user BC, reported also partly in [2]. Although the capacity region of the BC still remains an open problem to this day, Marton established in [3] an inner bound on the general two-user BC based on the notion of *random binning* and *superposition coding* with common and private messages, commonly referred to as “Marton’s coding”. This inner bound remains the best hitherto known in literature while the best outer bound on the capacity region of the BC is due to Nair & El Gamal [4]. These two bounds were shown to coincide for several classes of “ordered” channels, citing here: degraded, less noisy, and more capable BCs (see [5] and references therein) and more recently [6], for essentially less noisy and essentially more capable BCs, the key feature being the use of *superposition coding* as an encoding strategy. Marton’s inner bound also proved to be capacity-achieving for some non-ordered channels: the deterministic and semi-deterministic BC in [3] and [7], the MIMO BC in [8] while the capacity region of a BC consisting of the product and sum of two unmatched channels is also reported in [9]. In these cases, it is *random binning* that proves to be crucial for interference management.

In the above mentioned works, the channel statistics are perfectly known to the transmitter and thus the encoder can exploit this knowledge to allow for an efficient interference mitigation scheme. In all the cases where Marton’s inner bound is tight, the construction of the optimizing auxiliary code depends on the prior knowledge of either the channel output statistics (e.g. deterministic and semi-deterministic BCs) or a function of these statistics (e.g. users’ ordering in ordered single antennas BCs). When the encoder is oblivious to any such information about the channel state –no channel state information (CSIT)–, the effect of interference coupled with channel uncertainty on Marton’s coding technique can be more stringent. This arises the necessity to explore encoding and decoding schemes that are powerful enough to deal with the effects of channel uncertainty.

#### A. Related Work

It is worth mentioning here that few works dealt lately with alternate decoding techniques. We cite here first [10], where the authors characterized the maximum rate region for general interference networks under a given code constraint. This work generalizes the technique of “Interference Decoding” (ID), which was already used in [11], and consists in an alternate strategy for treating interference at receive terminals. More precisely, ID combines *non-unique decoding* with the possibility at each receiver to decode or not the interfering messages intended to the others users. As a matter of fact, the gain of ID does not result from non-unique decoding [12] as much as it follows from decoding interference. Yet, the straight-forward extension of the results of this work [10] to the BC is not strong enough for it encompasses only *superposition coding* but not *random binning*. Nevertheless, it provides an interesting insight on how to recover a *superposition coding* like inner bound with alternative decoding strategies, while keeping a symmetric encoding which will be useful for ordered channels.

Later, authors in [13] derived an inner bound based on “Coset Codes” for the three users BC possibly enlarging the best-known known inner bound. Coset codes are *structured codes* that allow the destinations to decode a “compressive” function of the interfering messages and thus a complete cancellation of interference with less impediment to the information rates than fully decoding the interfering messages. A class of 3 users BCs is proposed where two links are interference free and for which the straightforward extension of Marton’s coding scheme, stays strictly suboptimal compared to the suggested rate region. Such a coding technique based

on Coset Codes, proves to be useful for three user BC, however, it does not enlarge Marton’s inner bound in the two user’s case. Yet this work presents the first class of 3 users BC for which Marton’s inner bound, with many common layers, is strictly sub-optimal.

When the channels are not ordered, e.g, MISO BC, the effect of channel uncertainty on the “Degrees of Freedom” (DoF) –insightful to understand how interference should be managed with no CSIT– is rather well understood. For finite state compound settings, Weingarten *et al.* had first derived both inner and outer bounds on the DoF region and on the sum-DoF of the compound MISO BC [14] with some cases of optimality. The outer bound derived therein was conjectured to be loose, but later Gou *et.al* [15] and Maddah-Ali [16] proved the optimal DoF region of the generic compound MISO BC, both in the complex and in the real settings, to perfectly match this outer bound. The achievability of the optimal DoF relies on either a Linear or a non-Linear coding scheme combined with “symbol extensions” in [14] while the proof made in [16] resorts to number theory tools and consists in interference alignment over rational dimensions of the real numbers (see also [17]). When the states span an infinite set, i.e., in the ergodic setting, DoF can experience severe loss. In [18], it is shown that with Rayleigh fading channels, the sum-DoF collapses to the number of transmit antennas: time-sharing is optimal. A few more works deal with alternate settings where various models of the amount and accuracy of CSI available at the transmitter are considered, e.g. [19]. It turns out that richer encoding strategies, like *Interference Alignment* (IA) along with block expansion (coding over many time slots) are crucial in dealing with interference, and thus, any optimal scheme for the finite power limited MISO BC should encompass such coding strategies.

## B. Our Contribution

In this work, we explore the role that two main interference mitigation techniques can play in the compound BC setup, and show that, by operating clever optimization either on the encoding or on the decoding side, we can alleviate the effect of uncertainty when coupled with interference in two different ways. We first start by deriving a rate region that takes advantage of the combination of each of ID, Marton’s *random binning*, and *superposition coding*. We prove that for the compound BC –unlike the standard two-user BC– ID can strictly outperform its antagonist “simpler” strategy, i.e., “Non Interference Decoding” (NID). The gain is due to the fact that ID allows for a symmetric encoding, and thus deals better with the source’s uncertainty

while relegating the “clever decoding” to the receive terminals. To illustrate clearly the role of this decoding, we investigate a class of discrete ordered compound BCs for which this improvement is strict and where ID is crucial to recreate superposition-coding like rate regions without specifying prior coding hierarchy and decoding orders.

However, if the channels are not ordered, then the ID gain is less explicit, and thus, more evolved encoding schemes need to be investigated. For this reason, we look at the role that “Multiple Description” (MD) coding can play in the non-ordered compound BC, where we allow each possible instance of the same user to decode a “private description” unintended for the other channels instances. We follow a similar approach to that in [20] where MD coding had been already proven to be useful over compound state-dependent channels. Such a scheme allows the encoder to treat differently the many channels instances of each user, and the resulting decoding constraints are therefore less stringent than the “Common Description” (CD) coding scheme [3]. Indeed, the introduction of several private descriptions results in a cost tantamount to their overall correlation. Therefore, the primary question that we aim to address here is whether this correlation is more harmful than the channel uncertainty. Our answer is mostly negative and this is stated by a class of compound MISO BC where we show that, under a specific “Dirty-Paper Coding” (DPC) scheme [21], MD coding can strictly outperform CD coding. By using a fraction of the power intended to superimpose private descriptions, each aligned for an instance of each user, can be strictly useful. Finally, we discuss the relative behavior of ID and MD coding techniques and present a brief example to support their exclusive inclusion.

The remainder of this paper is organized as follows. Section II plots the system model and provides basic definitions as well as a simple outer bound on the capacity region of a general compound BC. In Section III, we study the utility of ID for the Compound BC. We start by deriving the ID inner bound in Section III-A and show in Section III-B that ID is capacity achieving for a class of discrete Compound BCs while NID stays strictly sub-optimal. Next, in Section IV, we introduce MD coding for a Compound BC setup which is studied for the Compound Gaussian MISO BC in Section V. The performances of these two inner bounds are then compared to the outer bound presented in Section V-G. Last, we compare the relative behavior of both the ID and MD inner bounds in Section VI-A and end with summary and discussion in Section VI-B.

*Notations:* The term pmf will refer to probability mass function. Random variables and their realizations are denoted by upper resp. to lower case letters. Vectors are denoted by bold font characters and RV stands for random variable while FME stands for Fourier Motzkin Elimination.

For any sequence  $(x_i)_{i \in \mathbb{N}_+}$ , notation  $x_k^n$  stands for the collection  $(x_k, x_{k+1}, \dots, x_n)$ .  $x_1^n$  is simply denoted by  $x^n$ . Entropy is denoted by  $H(\cdot)$ , and mutual information by  $I(\cdot; \cdot)$  while differential entropy is denoted by  $h(\cdot)$ .  $\mathbb{E}$  resp.  $\mathbb{P}$  denote the expectation resp. the generic probability measure while the notation  $P$  is specific to the pmf of a RV.  $\|\mathcal{X}\|$  stands for the cardinality of the set  $\mathcal{X}$ . We denote typical and conditional typical sets by  $T_\delta^n(X)$  and  $T_\delta^n(Y|x^n)$ , respectively (see Appendix A for details). Let  $X, Y$  and  $Z$  be three RVs on some alphabets with probability distribution  $p$ . If  $p(x|yz) = p(x|y)$  for each  $x, y, z$ , then they form a Markov chain, which is denoted by  $X \text{---} Y \text{---} Z$ . The *binary entropy function*  $H_2$  is defined  $\forall x \in [0 : 1]$  by  $H_2(x) \triangleq -x \log_2(x) - (1-x) \log_2(1-x)$ , and the binary convolution operator  $(\star)$  as:  $x \star y \triangleq x(1-y) + (1-x)y$  for all  $(x, y) \in [0 : 1]^2$ . For two channels with outputs  $Y_1$  and  $Y_2$ ,  $\preceq$  stands for  $Y_1$  is less noisy than  $Y_2$ . On the other hand,  $\mathbf{h}^t$  is to be understood as the transpose of the real valued vector  $\mathbf{h}$ . Let  $\mathbf{B}_u$  be a unit norm  $2 \times 1$  column vector. We denote the scalar product between vectors  $\mathbf{B}_u$  and  $\mathbf{h}_j$  by  $h_{j,u} = \mathbf{h}_j^t \mathbf{B}_u$ .

## II. PROBLEM DEFINITION

The Compound Broadcast Channel model consists in one source terminal and two distinct receivers each observing one of many possible channel outputs. The source wishes to communicate two private messages, one to each receiver, while a common message is intended to both of them. This setup is equivalent to a setting where each user is represented by multiple users that are interested in the same message. As a matter of fact, this model is also equivalent to pairing up users from distinct groups, leading to a compound setup whose class of channels consists of all possible BCs created with possible pairs of users, and where the source is oblivious to the channel controlling the communication.

### A. Definition of the Compound Broadcast Channel (BC)

- Consider a collection of  $n$ -th extensions of discrete memoryless BCs:

$$\{\mathcal{W}_j^n\}_{j \in \mathcal{J}} = \{P_{Y_j^n Z_j^n | X^n} : \mathcal{X}^n \mapsto \mathcal{Y}^n \times \mathcal{Z}^n\}, \quad (1)$$

defined by the conditional pmfs:

$$P_{Y_j^n Z_j^n | X^n} = \prod_{i=1}^n P_{Y_{j,i} Z_{j,i} | X_i} . \quad (2)$$

- Users' pair of index  $j$  takes values in the finite set of indices  $\mathcal{J} = [1 : N]$ .
- An  $(M_{0n}, M_{1n}, M_{2n}, n)$ -code for this channel consists of: three sets of messages  $\mathcal{M}_0, \mathcal{M}_1$  and  $\mathcal{M}_2$ , an encoding function that assigns an  $n$ -sequence  $x^n(w_0, w_1, w_2)$  to each triple of messages  $(w_0, w_1, w_2) \in \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$  and decoding functions, one at each receiver, that assign to the received signal an estimate message pair  $(\hat{w}_{0k}, \hat{w}_k)$  in  $\mathcal{M}_0 \times \mathcal{M}_k$ , for  $k \in \{1, 2\}$  or an error.

The probability of error is given by:

$$P_e^{(n)}(j) \triangleq \mathbb{P} \left( \bigcup_{k \in \{1, 2\}} \{(\hat{W}_{0k}(j), \hat{W}_k(j)) \neq (W_0, W_k)\} \right) . \quad (3)$$

- A rate tuple  $(R_0, R_1, R_2)$  is said to be achievable if there exists an  $(M_{0n}, M_{1n}, M_{2n}, n)$ -code satisfying:

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log_2 M_{kn} \geq R_k \quad \forall k = \{0, 1, 2\} , \quad (4)$$

$$\limsup_{n \rightarrow \infty} \max_{j \in \mathcal{J}} P_e^{(n)}(j) = 0 . \quad (5)$$

The capacity region is the set of all achievable rate tuples.

### B. Outer Bound of the Capacity of the Compound BC

We derive in this section a simple and intuitive outer bound on the capacity region of the compound BC. This outer bound results from a straightforward extension to the compound setting of the best-known outer bound on the capacity of the BC. It will be useful in the examples we shall study later.

Let the rate region  $\mathcal{R}_{\text{NEG}}^{(j)}$  denote the outer bound derived in [4] applied to each pair of users with index “ $j$ ”. For the private message setup, the rate region is given by

$$\mathcal{R}_{\text{NEG}}^{(j)}(p_{QUVX}) : \begin{cases} R_1 \leq I(QU; Y_j) , \\ R_2 \leq I(QV; Z_j) , \\ R_1 + R_2 \leq I(U; Y_j | QV) + I(QV; Z_j) , \\ R_1 + R_2 \leq I(QU; Y_j) + I(V; Z_j | QU) . \end{cases} \quad (6)$$



for a specific pmf on  $p_{QUVX}$ . A simple outer bound on the capacity region of the compound BC is stated in the following theorem.

**Theorem 1** (Outer bound). *The capacity region of the two-user Compound BC  $\mathcal{C}_{\mathcal{J}}$  verifies:*

$$\mathcal{C}_{\mathcal{J}} \subseteq \bigcup_{P_U P_V} \bigcap_{j=1}^N \left( \bigcup_{P_{QX|UV}} \mathcal{R}_{NEG}^{(j)}(p_{QUVX}) \right), \quad (7)$$

where the channel input  $X$  is a deterministic mapping of  $\mathcal{Q} \times \mathcal{U} \times \mathcal{V}$ .

It is worth mentioning that when the Compound BC consists in only one BC, the outer bound [4] was not proven to be tight in general. For non-ordered compound setups, the fact of optimizing the common auxiliary RV  $Q$  for each channel with index  $j$ , prevents even further this outer bound from being tight since the encoder is oblivious to the actual channel realization. For instance, it cannot optimize the code for each of the possible channels instances. However, this bound can still be tight in some cases of interest as will be clarified later on.

*Proof:* A sketch of the proof is relegated to Appendix D. ■

### III. INTERFERENCE DECODING (ID) IN THE COMPOUND BROADCAST CHANNEL

We now derive an inner bound on the capacity region of the Compound BC resorting to a class of codes consisting of three RVs, each one encoding a message, generated and mapped via superposition coding and random binning.

#### A. Interference Decoding (ID) Inner Bound

The inner bound we derive here shares common ideas with following works [22]. First, the notion of ID used in [11] where –roughly speaking– each receiver is allowed to decode its intended message as well as (non-uniquely) decode or not the interfering message. Second, the fact that decoding “non-uniquely” the interfering message alleviates an extra constraint on the information rates yielding the same result as if the decoder would have to successively decode the interfering and the intended messages which is related to [23].

**Theorem 2** (ID inner bound). *An inner bound on the capacity region of the Compound BC consists in the set of all rates  $(R_0, R_1, R_2)$  included in:*

$$\mathcal{R}_{ID} \triangleq \bigcup_{p_{QUVX} \in \mathcal{P}} \underbrace{\bigcup_{(T_1, T_2) \in \mathbb{T}(p)} \bigcap_{j=1}^N \bigcup_{i_j=1}^4 \mathcal{T}_{i_j}^{(j)}(p, T_1, T_2)}_{\substack{\text{FME} \\ \text{(compound) (4 methods)}}}, \quad (8)$$

where  $\mathcal{P}$  is the set of all input pmfs  $p_{QUVX}$  such that  $(Q, U, V) \ominus X \ominus (Y_1, \dots, Y_N, Z_1, \dots, Z_N)$ .

The rate regions  $\mathcal{T}_{[1:4]}^{(j)}$  and the set  $\mathbb{T}$  are, respectively, defined as follows:

$$\mathcal{T}_1^{(j)}(p, T_1, T_2) : \begin{cases} T_1 \leq I(U; Y_j | Q), \\ R_0 + T_1 \leq I(QU; Y_j), \\ T_2 \leq I(V; Z_j | Q), \\ R_0 + T_2 \leq I(QV; Z_j), \end{cases} \quad (9)$$

$$\mathcal{T}_2^{(j)}(p, T_1, T_2) : \begin{cases} T_1 \leq I(U; Y_j | Q), \\ R_0 + T_1 \leq I(QU; Y_j), \\ T_2 \leq I(V; Z_j U | Q), \\ T_1 + T_2 \leq I(UV; Z_j | Q) + I(U; V | Q), \\ R_0 + T_1 + T_2 \leq I(QUV; Z_j) + I(U; V | Q), \end{cases} \quad (10)$$

$$\mathcal{T}_3^{(j)}(p, T_1, T_2) : \begin{cases} T_1 \leq I(U; Y_j V | Q), \\ T_1 + T_2 \leq I(UV; Y_j | Q) + I(U; V | Q), \\ R_0 + T_1 + T_2 \leq I(QUV; Y_j) + I(U; V | Q), \\ T_2 \leq I(V; Z_j | Q), \\ R_0 + T_2 \leq I(QV; Z_j), \end{cases} \quad (11)$$

$$\mathcal{T}_4^{(j)}(p, T_1, T_2) : \begin{cases} T_1 \leq I(U; Y_j V | Q), \\ T_1 + T_2 \leq I(UV; Y_j | Q) + I(U; V | Q), \\ R_0 + T_1 + T_2 \leq I(QUV; Y_j) + I(U; V | Q), \\ T_2 \leq I(V; Z_j U | Q), \\ T_1 + T_2 \leq I(UV; Z_j | Q) + I(U; V | Q), \\ R_0 + T_1 + T_2 \leq I(QUV; Z_j) + I(U; V | Q), \end{cases} \quad (12)$$

$$\mathbb{T}(p) = \left\{ (T_1, T_2) : \begin{aligned} T_1 &\geq R_1, \\ T_2 &\geq R_2, \\ T_1 + T_2 &> R_1 + R_2 + I(U; V|Q) \end{aligned} \right\}. \quad (13)$$

$$T_2 \geq R_2, \quad (14)$$

$$T_1 + T_2 > R_1 + R_2 + I(U; V|Q) \}. \quad (15)$$

*Proof:* The proof is relegated to Appendix B. ■

**Remark 3** (Main comments about the proof). *Each user introduces the union of two sets of constraints, corresponding to decoding or not the interference. This results –in terms of achievable rates– in the union of four rate regions:*

- 1) *The region  $\mathcal{T}_1^{(j)}$  is the same rate region as obtained with Marton's inner bound,*
- 2) *The region  $\mathcal{T}_4^{(j)}$  is obtained by letting the destinations to decode both the intended and the interfering message,*
- 3) *The regions  $\mathcal{T}_2^{(j)}$  and  $\mathcal{T}_3^{(j)}$  correspond to each destination decoding the interfering message at once.*

*A slightly similar rate region was also derived in [10] in a different context, but it does not take advantage of the encoding technique, and thus in our setting it fails at achieving even Marton's inner bound.*

**Remark 4** (Connection to the standard two-user BC). *Consider the standard two-user BC where  $\mathcal{J} = 1$ . Observe that by allowing both destinations to decode or not the message of the other user –ID scheme– we found a seemingly larger rate region  $\mathcal{R}_{s,ID}$  than that of Marton [3] which does not use the ID technique. Indeed, these regions are given by*

$$\mathcal{R}_{s,ID} \triangleq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbb{T}(p)} \left( \bigcup_{i=1}^4 \mathcal{T}_i(p, T_1, T_2) \right), \quad (16)$$

$$\mathcal{R}_{s,NID} \triangleq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbb{T}(p)} \mathcal{T}_1(p, T_1, T_2). \quad (17)$$

*It is clear that  $\mathcal{R}_{s,NID} \subseteq \mathcal{R}_{s,ID}$ , but the question is whether or not this inclusion is strict. To check this issue, we need to evaluate both regions and thus we resort to FME for  $(T_1, T_2)$ , and bit recombination between the private rates  $(R_1, R_2)$  and the common one  $R_0$ <sup>1</sup>. Since the unions commute, we can write that:*

<sup>1</sup>For the interested reader a similar calculation is done in Appendix E.

$$\begin{aligned}\mathcal{R}_{s,ID} &= \bigcup_{i=1}^4 \mathcal{R}_{s,i} \\ &= \mathcal{R}_{s,NID} \cup \left( \bigcup_{i=2}^4 \mathcal{R}_{s,i} \right),\end{aligned}\quad (18)$$

where  $\mathcal{R}_{[2:3]}$  are respectively defined by the following sets of inequalities:

$$\mathcal{R}_{s,2} : \begin{cases} R_0 + R_1 \leq I(QU; Y), \\ R_0 + R_1 + R_2 \leq I(V; Z|UQ) + I(QU; Y), \\ R_0 + R_1 + R_2 \leq I(QUV; Z), \end{cases} \quad (19)$$

$$\mathcal{R}_{s,3} : \begin{cases} R_0 + R_2 \leq I(QV; Z), \\ R_0 + R_1 + R_2 \leq I(U; Y|VQ) + I(QV; Z), \\ R_0 + R_1 + R_2 \leq I(QUV; Y), \end{cases} \quad (20)$$

$$\mathcal{R}_{s,4} : \begin{cases} R_0 + R_1 + R_2 \leq I(QUV; Y), \\ R_0 + R_1 + R_2 \leq I(QUV; Z), \end{cases} \quad (21)$$

while  $\mathcal{R}_{s,NID}$  is defined by

$$\mathcal{R}_{s,NID} = \mathcal{R}_{s,1} : \begin{cases} R_0 + R_1 \leq I(QU; Y), \\ R_0 + R_2 \leq I(QV; Z), \\ R_0 + R_1 + R_2 \leq I(U; Y|Q) + I(QV; Z) - I(U; V|Q), \\ R_0 + R_1 + R_2 \leq I(QU; Y) + I(V; Z|Q) - I(U; V|Q), \\ 2R_0 + R_1 + R_2 \leq I(QU; Y) + I(QV; Z) - I(U; V|Q). \end{cases} \quad (22)$$

From the above rate regions, we observe that by taking  $U = Q$ , the region  $\mathcal{R}_{s,NID}$  contains  $\mathcal{R}_{s,2}$ , and similarly, setting  $V = Q$  allows  $\mathcal{R}_{s,NID}$  to contain  $\mathcal{R}_{s,3}$  while  $U = Q = V$  allows it to contain  $\mathcal{R}_{s,4}$ . Hence, using the ID strategy in presence of a single channel per user yields the same rate region as Marton's inner bound. Indeed, the apparently gain provided by choosing to decode or not the interference is recovered by an optimization of the input distribution.

We can observe that by using ID in the compound setting, we get a seemingly larger region than Marton's worst-case inner bound, which is given by:

$$\mathcal{R}_{\text{NID}} \triangleq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbb{T}(p)} \left( \bigcap_{j=1}^N \mathcal{T}_1^{(j)}(p, T_1, T_2) \right). \quad (23)$$

It is clear that  $\mathcal{R}_{\text{NID}} \subseteq \mathcal{R}_{\text{ID}}$  but yet, no evidence on the strict inclusion has been stated here. In the sequel, we investigate a Compound BC for which the region based on the usual decoding in Marton's inner bound  $\mathcal{R}_{\text{NID}}$  fails at achieving the capacity while  $\mathcal{R}_{\text{ID}}$  from Theorem 2 is tight. The key point in this "strict inclusion", is that, if the optimizing input pmf varies from one channel to the other (e.g. in terms of superposition ordering of auxiliary RVs), then the joint optimization in the compound setup imposes a stringent limitation on the input pmf. This prevents  $\mathcal{R}_{\text{NID}}$  from reaching capacity for some compound models while the ID technique, allowing the choice between two decoding strategies, does not suffer such a loss.

### B. Interference Decoding is Optimal for a Class of Compound Broadcast Channels

In this section, we will construct a Compound BC model for which Marton's worst-case inner bound, obtained through NID, is strictly sub-optimal compared to ID inner bound where users are allowed to decode or not the interference. We first discuss a criterion for the construction of such a compound model and later, prove the optimality of ID. For simplicity, we restrict our analysis to the case  $\|\mathcal{J}\| = 2$  and private rates only, i.e.,  $R_0 = 0$ .

1) *Irrelevant compound models:* The difficult to characterize optimal coding for the Compound BC is inherent to the class of BCs in the set, i.e., the set of channel users over which we define the compound model. We shall refer to as "irrelevant" models those of ordered BCs for which Marton's worst-case inner bound is tight. As a matter of fact, Marton's inner bound achieves the capacity of every BC for which capacity is known.

Consider the class of broadcast channels:

$$\mathcal{W} = \{\mathcal{W}_1, \mathcal{W}_2\} = \{\mathcal{X} \mapsto (\mathcal{Y}_j, \mathcal{Z}_j)\}_{j \in \{1,2\}} , \quad (24)$$

where  $Y_2 \preceq Y_1$  and  $Z_1 \preceq Z_2$ . Then, it follows that, whatever the auxiliary RVs  $(Q, U) \sim p_{QU}$ :

$$I(QU; Y_2) \leq I(QU; Y_1) \quad , \quad I(U; Y_2|Q) \leq I(U; Y_1|Q) . \quad (25)$$

Thus, Marton's inner bound based on superposition coding and random binning yields the region

$$\left\{ \begin{array}{l} R_1 \leq \min_{j=1,2} I(QU; Y_j) , \\ R_2 \leq \min_{j=1,2} I(QV; Z_j) , \\ R_1 + R_2 \leq \min_{j=1,2} I(U; Y_j|Q) + \min_{j=1,2} I(QV; Z_j) - I(U; V|Q) , \\ R_1 + R_2 \leq \min_{j=1,2} I(QU; Y_j) + \min_{j=1,2} I(V; Z_j|Q) - I(U; V|Q) , \end{array} \right. \quad (26)$$

which reduces to:

$$\left\{ \begin{array}{l} R_1 \leq I(QU; Y_2) , \\ R_2 \leq I(QV; Z_1) , \\ R_1 + R_2 \leq I(U; Y_2|Q) + I(QV; Z_1) - I(U; V|Q) , \\ R_1 + R_2 \leq I(QU; Y_2) + I(V; Z_1|Q) - I(U; V|Q) . \end{array} \right. \quad (27)$$

This is the the rate region obtained by coding for only the pair of users corresponding to the channel  $(Y_2, Z_1)$ . Furthermore, it is straightforward to check that if the capacity of this channel is known (e.g. when  $Y_2$  and  $Z_1$  are ordered in the sense of “degradedness” or “less-noisiness”), then Marton’s inner bound achieves the capacity region of the Compound BC. Thus, if the marginals seen in set of users 1, i.e.,  $(Y_1, Y_2)$  are ordered at least in the known senses of “less noisiness”, and so are those in the set of users 2, i.e.,  $(Z_1, Z_2)$ , Marton’s inner bound for this setup leads to the capacity region of the “worst” BC formed by the worst pair of users in the set. Hence this class of compound models is irrelevant for our purpose.

2) *Compound Binary Erasure and Binary Symmetric BC*: In this section, we construct the simplest while relevant Compound BC setting, where:

- Set of user 2 contains only one channel instance, i.e.,  $Z_1 = Z_2 = Z$ .
- Set of user 1 is compound of two possible channel instances denoted by  $\{Y_j\}_{j \in \{1,2\}}$ .

Our aim is to show the desired “strict” inclusion  $\mathcal{R}_{\text{NID}} \subset \mathcal{R}_{\text{ID}}$ . To this end, we need to find a “relevant” compound BC where  $(Y_1, Y_2)$  are not strongly ordered (e.g. neither degraded nor less-noisy). Otherwise the resulting Compound BC would be formed by  $Z$  and the worst channel between  $(Y_1, Y_2)$ , for which it is straightforward to see that  $\mathcal{R}_{\text{NID}}$  achieves the capacity region.

Besides this argument, if we are to show the strict inclusion of Marton’s rate region with respect to the rate region obtained by ID, we need to provide for some inverse orderings in the compound channels formed by all possible pairs of users, so as to impose a tradeoff between two antagonist coding schemes for Marton’s coding scheme, i.e., two antagonist choices of auxiliary RVs at the encoder. One can then think of a setting where for instance the BC  $(Y_1, Z)$  has  $Z$  “better” than  $Y_1$  while the BC  $(Y_2, Z)$  is ordered in the opposite way, i.e  $Y_2$  is better than  $Z$ .

Consider the *Binary Erasure Channel* (BEC) with erasure probability  $e$  and the *Binary Symmetric Channel* (BSC) with crossover probability  $p$ . These have the particularity of allowing for a variety of orderings between the outputs [6], depending on  $(e, p)$ , as summarized in Table I.

TABLE I  
DIFFERENT ORDERINGS ALLOWED BY THE BEC(E)/BSC(P) BC

$0 \leq e \leq 2p$	$2p < e \leq 4p(1-p)$	$4p(1-p) < e \leq H_2(p)$	$H_2(p) < e \leq 1$
BSC degraded of BEC	BEC Less Noisy BSC	BEC More Capable BSC	BSC Ess. Less Noisy BEC

Define the Compound BC with components:

$$\mathcal{W} : \begin{cases} \mathcal{X} \mapsto \mathcal{Z} \equiv \text{BSC}(p), \\ \mathcal{X} \mapsto \mathcal{Y}_1 \equiv \text{BSC}(p_1), \\ \mathcal{X} \mapsto \mathcal{Y}_2 \equiv \text{BEC}(e_2). \end{cases} \quad (28)$$

We first start by imposing to  $Y_2$  to be more capable than  $Y_1$ , which requires:  $4p_1(1-p_1) < e_2 \leq H_2(p_1)$ . One possible choice is then to take  $Y_1$  as a *physically degraded* version of  $Z$ , i.e.,  $p < p_1 < 0.5$ , and  $Y_2$  more capable than  $Z$ , i.e.,

$$4p(1-p) < 4p_1(1-p_1) < e_2 \leq H_2(p) \leq H_2(p_1). \quad (29)$$

This choice fulfills the criteria stated for the construction of a relevant example. For this case, the simple outer bound enunciated in Section II-B writes as:

$$\mathcal{C}_{\mathcal{J}} \subset \mathcal{C}_1 \cap \mathcal{C}_2, \quad (30)$$

where:

$$\mathcal{C}_1 : \begin{cases} R_1 \leq 1 - H_2(p_1 \star \alpha), \\ R_2 \leq H_2(p \star \alpha) - H_2(p), \end{cases} \quad (31)$$

$$\mathcal{C}_2 : \begin{cases} R_1 \leq (1 - e_2) H_2(\alpha), \\ R_2 \leq 1 - H_2(p \star \alpha), \\ R_1 + R_2 \leq (1 - e_2). \end{cases} \quad (32)$$

We claim that the capacity region  $\mathcal{C}_1$  is strictly included in  $\mathcal{C}_2$ , for which we can compare:

$$\max_{(R_1, R_2) \in \mathcal{C}_1} \frac{R_1}{(1 - H_2(p)) - R_2} = \lim_{\alpha \rightarrow 1/2} \frac{1 - H_2(p_1 \star \alpha)}{1 - H_2(p \star \alpha)} \approx \frac{(1 - 2p_1)^2}{(1 - 2p)^2} \quad (33)$$

and

$$\max_{(R_1, R_2) \in \mathcal{C}_2} \frac{R_1}{(1 - H_2(p)) - R_2} \geq \frac{1 - e_2}{1 - H_2(p)}. \quad (34)$$

Our claim simply follows by noticing that from the assumptions on the parameters  $e_2$ ,  $p$  and  $p_1$ , we have that:

$$\frac{(1 - 2p_1)^2}{(1 - 2p)^2} \leq 1 \leq \frac{1 - e_2}{1 - H_2(p)} , \quad (35)$$

which shows the outer bound reduces to  $\mathcal{C}_1$ .

3) *Evaluation of the ID inner bound of Theorem 2:* We evaluate the proposed rate region  $\mathcal{R}_{\text{ID}}$  of Theorem 2, which satisfies:

$$\mathcal{R}_{\text{ID}} \supseteq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbb{T}(p)} \left( \mathcal{T}_3^{(1)}(p, T_1, T_2) \cap \mathcal{T}_4^{(2)}(p, T_1, T_2) \right) , \quad (36)$$

where  $\mathcal{T}_3^{(1)} \cap \mathcal{T}_4^{(2)}$  is defined by the set of inequalities:

$$\left\{ \begin{array}{l} T_2 \leq I(V; ZU|Q) , \\ T_1 + T_2 \leq I(UV; Z|Q) + I(U; V|Q) , \\ R_0 + T_1 + T_2 \leq I(QUV; Z) + I(U; V|Q) , \\ T_1 \leq I(U; Y_2V|Q) , \\ T_1 + T_2 \leq I(UV; Y_2|Q) + I(U; V|Q) , \\ R_0 + T_1 + T_2 \leq I(QUV; Y_2) + I(U; V|Q) , \\ T_1 \leq I(U; Y_1|Q) , \\ R_0 + T_1 \leq I(QY; Y_1) \\ T_1 \geq R_1 \quad , \quad T_2 \geq R_2 , \\ T_1 + T_2 > R_1 + R_2 + I(U; V|Q) . \end{array} \right. \quad (37)$$

This comes to choosing:  $i_1 = 3$ , i.e., using decoding method (3) for the BC (1), while the other channel gets the fourth decoding method:  $i_2 = 4$ . These constraints allow  $Z$  and  $Y_2$  to decode all messages, while forcing  $Y_1$  to decode only its own message. In Appendix E, it is shown after FME on  $(T_1, T_2)$ , bit recombination, and then setting  $R_0 = 0$ , that the previous rate region reduces to the set of rates satisfying:

$$\left\{ \begin{array}{l} R_1 \leq I(QU; Y_1) , \\ R_1 + R_2 \leq I(QU; Y_1) + I(V; Z|QU) , \\ R_1 + R_2 \leq I(QU; Y_1) + I(UV; Y_2|Q) , \\ R_1 + R_2 \leq I(QUV; Y_2) . \end{array} \right. \quad (38)$$



Then, letting:  $V = X$ ,  $\bar{Q} = (Q, U)$ , and using the fact that  $Y_2$  is *more capable* than  $Z$ , yields:

$$\begin{cases} R_1 \leq I(\bar{Q}; Y_1) , \\ R_1 + R_2 \leq I(\bar{Q}; Y_1) + I(X; Z|\bar{Q}) . \end{cases} \quad (39)$$

This achievable rate region coincides with the outer bound and thus provides the capacity region of the BC  $(Y_1, Z)$  for the considered setup. Letting then  $\bar{Q} \mapsto X \equiv \text{BSC}(\alpha)$ , and  $X \sim \text{Bern}(1/2)$  we get the following union over all  $\alpha \in [0 : 1]$  of:

$$\mathcal{R}_{\text{ID}} : \begin{cases} R_1 \leq 1 - H_2(p_1 \star \alpha) , \\ R_1 + R_2 \leq 1 - H_2(p_1 \star \alpha) + H_2(p \star \alpha) - H_2(p) . \end{cases} \quad (40)$$

In order to check that  $\mathcal{R}_{\text{ID}}$  is equal to the outer bound  $\mathcal{C}_1$ , we should first start by noticing that it is the exclusive union of two rate regions:  $\mathcal{C}_1$  and  $\mathcal{R}_E$  which are defined by

$$\mathcal{R}_E : \begin{cases} R_2 \geq H_2(p \star \alpha) - H_2(p) , \\ R_1 + R_2 \leq 1 - H_2(p_1 \star \alpha) + H_2(p \star \alpha) - H_2(p) . \end{cases} \quad (41)$$

As plot in Fig. 1, this region has four corner points among which 3 of them are clearly included in  $\mathcal{C}_1$ , i.e.,  $A$ ,  $B$ , and  $C$ .

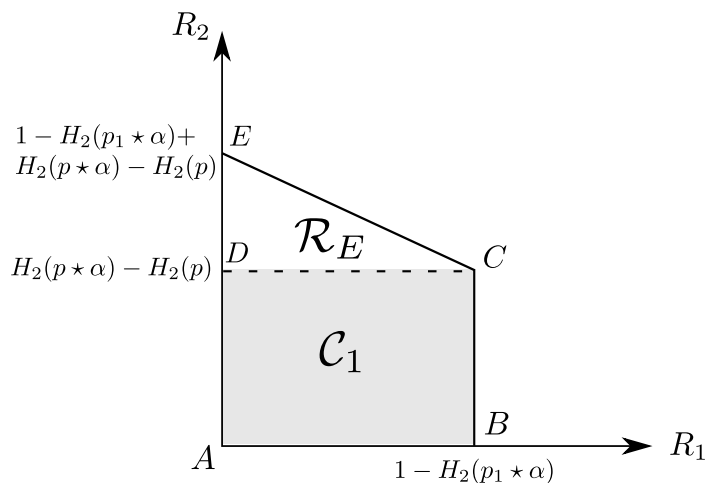


Fig. 1. Comparison between  $\mathcal{C}_1$  and  $\mathcal{R}_{\text{ID}}$ .

To show that the point  $E$  lies in the region  $\mathcal{C}_1$ , we first write that:

$$E = (0, 1 - H_2(p_1 \star \alpha) + H_2(p \star \alpha) - H_2(p)) . \quad (42)$$

Since  $Y_1$  is physically degraded with respect to  $Z$ , i.e.,  $p \leq p_1$ , and since  $\alpha$ ,  $p \star \alpha$  and  $p_1 \star \alpha$  are all included in the interval  $[0 : 0.5]$ , one can clearly write that:  $-H_2(p_1 \star \alpha) + H_2(p \star \alpha) \leq 0$ .

Hence, the point  $E$  is dominated by the point  $C_2 = (0; 1 - H_2(p))$ , which is already achievable in  $\mathcal{C}_1$ . The line between  $C$  and  $E$  can be achieved by the convexity of the rate region  $\mathcal{C}_1$ .

4) *Outer bound on Marton's inner bound:* When restricted to Marton's inner bound, the rate region in expression (23) is included in the union of next constraints:

$$\left\{ \begin{array}{l} T_2 \leq I(V; Z|Q) , \\ R_0 + T_2 \leq I(QV; Z) , \\ T_1 \leq \min_{j=1,2} I(U; Y_j|Q) , \\ R_0 + T_1 \leq \min_{j=1,2} I(QU; Y_j) , \\ T_1 \geq R_1 \quad , \quad T_2 \geq R_2 , \\ T_1 + T_2 > R_1 + R_2 + I(U; V|Q) . \end{array} \right. \quad (43)$$

Then, we perform FME on the rates  $T_1$  and  $T_2$ , bit recombination, and we set  $R_0 = 0$ , which yields the following rate region:

$$\left\{ \begin{array}{l} R_2 \leq I(QV; Z) \\ R_1 \leq \min_{j=1,2} I(QU; Y_j) , \\ R_1 + R_2 \leq I(V; Z|Q) + \min_{j=1,2} I(QU; Y_j) - I(U; V|Q) , \\ R_1 + R_2 \leq I(QV; Z) + I(U; Y_2|Q) - I(U; V|Q) , \end{array} \right. \quad (44)$$

where we have used the fact that:  $I(Q; Y_1) \leq I(Q; Z)$ , i.e., *physical degradedness*. As a matter of fact, the previous rate region is contained in the set of rates verifying:

$$\left\{ \begin{array}{l} R_1 \leq \min_{j=1,2} I(QU; Y_j) , \\ R_1 + R_2 \leq I(X; Z|QU) + \min_{j=1,2} I(QU; Y_j) , \end{array} \right. \quad (45)$$

because for each  $P_{QUVX} \in \mathcal{P}$  the next inequalities hold:

$$I(QV; Z) \leq I(X; Z) , \quad (46)$$

$$I(V; Z|Q) + \min_{j=1,2} I(QU; Y_j) - I(U; V|Q) \leq I(X; Z|QU) + \min_{j=1,2} I(QU; Y_j) \quad (47)$$

$$\leq I(X; Z) . \quad (48)$$

By letting  $\bar{Q} = (Q, U)$ , we obtain the following constraints:

$$\mathcal{R}_{\text{OuterNID}} : \left\{ \begin{array}{l} R_1 \leq \min_{j=1,2} I(\bar{Q}; Y_j) , \\ R_1 + R_2 \leq I(X; Z|\bar{Q}) + \min_{j=1,2} I(\bar{Q}; Y_j) . \end{array} \right. \quad (49)$$

In Appendix F, we show that it suffices to evaluate this bound for all auxiliary RVs  $\bar{Q}$  that verify  $\|\bar{Q}\| \leq 4$  and  $X \sim \text{Bern}(1/2)$ .

Though we might state such characteristics about the maximizing distribution, the optimization of this region turns out to be tricky since the usual bounding tools such as ‘‘Mrs. Gerber’s Lemma’’ leads only to the next lower bound:

$$\mathcal{R}_{\text{Lower,NID}} \subseteq \begin{cases} R_2 \leq H_2(p \star \alpha) - H_2(p) , \\ R_1 \leq \min\{1 - H_2(p_1 \star \alpha), \bar{e}_2(1 - H_2(\alpha))\} . \end{cases} \quad (50)$$

Fig. 2 plots a comparison between these two regions. This lower bound coincides with the capacity region  $\mathcal{R}_{\text{ID}}$  over the interval  $R_2 \in [0 : H_2(p \star \alpha_0) - H_2(p)]$  or equivalently  $R_1 \in [0 : 1 - H_2(p_1 \star \alpha_0)]$  where  $\alpha_0$  is given by:  $1 - H_2(p_1 \star \alpha_0) = (1 - e_2)(1 - H_2(\alpha_0))$  .

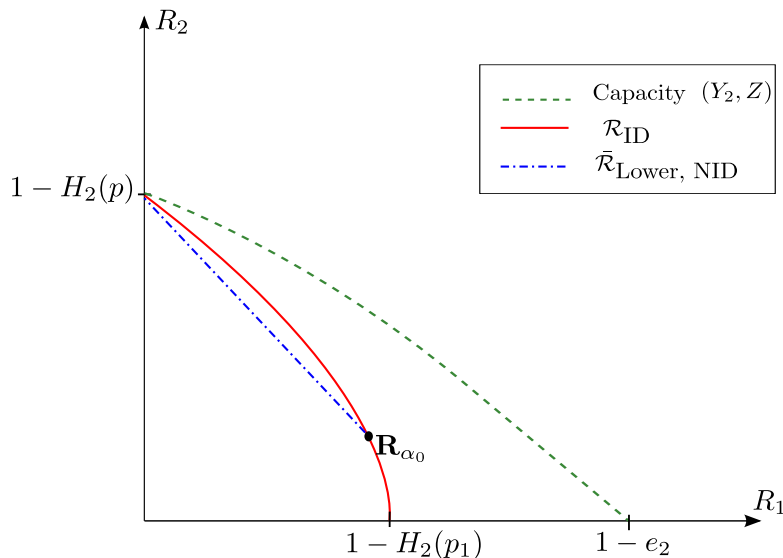


Fig. 2. Comparison between the rate region  $\mathcal{R}_{\text{ID}}$  and the convex closure of  $\mathcal{R}_{\text{Lower,NID}}$ .

In order to derive an upper bound, we study a looser outer bound to  $\mathcal{R}_{\text{Outer,NID}}$ , provided that the gap stays strict between the capacity region and this outer bound. Let us define the function  $t : [0 : 1 - H_2(p)] \mapsto \mathfrak{R}_+$  as:

$$t(x) \triangleq \sup_{p_{XQ} \in \mathcal{C}(x)} \min\{I(Q; Y_1), I(Q; Y_2)\} , \quad (51)$$

where the class  $\mathcal{C}(x)$  is given by

$$\mathcal{C}(x) = \{p_{XQ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Q}) : Q \text{ --- } X \text{ --- } (Z, Y_1, Y_2) \\ X \sim \text{Bern}(1/2) , I(X; Z|Q) \geq x\} . \quad (52)$$

The function  $t : x \mapsto t(x)$  characterizes the convex closure of the region  $\bar{\mathcal{R}}_{\text{Outer,NID}}$ , i.e.,  $(R_1, R_2) \in \bar{\mathcal{R}}_{\text{Outer,NID}}$  thus  $R_1 = t(R_2)$ . In the same way, define  $t_1$  over  $[0 : 1 - H_2(p)]$  by

$$t_1(x) \triangleq \sup_{p_{XQ} \in \mathcal{C}(x)} I(Q; Y_1) , \quad (53)$$

where  $t_1$  characterizes the convex closure of the region  $\bar{\mathcal{R}}_{\text{ID}}$ .

In the sequel, we work towards a closed form evaluation of an upper bound of  $t$  that would still be dominated by  $t_1$ .

5) *An upper bound on the function  $t(x)$* : We follow the method in [24] where

$$t(x) \triangleq \sup_{p_{XQ} \in \mathcal{C}(x)} \min \{I(Q; Y_1), I(Q; Y_2)\} \quad (54)$$

$$= \sup_{p_{XQ} \in \mathcal{C}(x)} \min_{a \in [0:1]} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] \quad (55)$$

$$\leq \min_{a \in [0:1]} \sup_{p_{XQ} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] , \quad (56)$$

for all  $x \in [0 : 1 - H_2(p)]$ . Let define for each  $a \in [0 : 1]$  and  $t_a \in [0 : 1 - H_2(p)]$ ,

$$t_a(x) \triangleq \sup_{p_{XQ} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] . \quad (57)$$

Notice that:

- The case  $a = 1$  was already studied in [24] and it was shown that:

$$t_1(x) = 1 - H_2(p_1 \star p_x) , \quad (58)$$

where  $H_2(p \star p_x) - H_2(p) = x$ .

- The case  $a = 0$  can be studied in a very similar fashion as in [24] by finding out that:

$$t_0(x) = \inf_{\lambda \in \mathcal{R}^+} [F_0(\lambda) - \lambda x] \quad (59)$$

$$= (1 - e_2) \left( 1 - \frac{x}{1 - H_2(p)} \right) , \quad (60)$$

where:

$$F_0(\lambda) = \max \{(1 - H_2(p)) \lambda, (1 - e_2)\} . \quad (61)$$

Now, to upper bound  $t_a$ , we could have written that:

$$t_a(x) \leq a \sup_{\mathcal{C}(x)} I(Q; Y_1) + \bar{a} \sup_{\mathcal{C}(x)} I(Q; Y_2) \quad (62)$$

$$= a t_1(x) + \bar{a} t_0(x) \quad (63)$$

$$\geq t_1(x) , \quad (64)$$

where (64) follows from what we have proved in Section III-B2, i.e.,  $t_0$  dominates  $t_1$  over the interval  $[0 : 1 - H_2(p)]$ . Thus, we cannot restrict ourselves to the upper bound in (62) on  $t_a$  since it is rather loose, and we will hence bound more tightly the function  $t_a$ .

**Proposition 1.** *The function  $t_a$  satisfies the following properties:*

(i) For all  $x \in [0 : 1 - H_2(p)]$ ,

$$t_a(x) = \max_{p_{XQ} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] , \quad (65)$$

(ii)  $t_a$  is concave in  $x$ ,

(iii)  $t_a$  can be described identically by its supporting lines,

(iv)  $t_a$  is decreasing in  $x$ .

*Proof:* The proof is relegated to Appendix G. ■

The next result is rather useful since it allows us to transform the optimization of a rate region into optimizing one quantity captured in  $F_a(\lambda)$ .

**Corollary 1.** *The following conclusions can be drawn:*

(a) The constraint in (52) can be transformed into:

$$I(X; Z|Q) = x . \quad (66)$$

(b) We have that:

$$t_a(x) = \inf_{\lambda \in \mathcal{R}^+} \left[ \max_{\mathcal{P}(\mathcal{X} \times \mathcal{Q})} [a I(Q; Y_1) + \bar{a} I(Q; Y_2) + \lambda I(X; Z|Q)] - \lambda x \right] \quad (67)$$

$$= \inf_{\lambda \in \mathcal{R}^+} [F_a(\lambda) - \lambda x] , \quad (68)$$

where

$$F_a(\lambda) \triangleq \max_{p_{XQ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Q})} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] . \quad (69)$$

*Proof:* (a) follows from the non-increasing property of  $t_a$  and (b) follows from the concavity of the function  $t_a$  since a concave function can be described by its supporting lines [25]. ■

The analysis of the function  $t_a$  for an arbitrary  $a$  brings about significant computational complexity, we thus only chose to plot it using stochastic optimization methods. We chose  $e_2 = 0.46$ ,  $p = 0.1$  and  $p_1 = 0.13$ . It can be readily shown that these parameters verify (29).

In Fig. 3, we chose  $a = 0.92$  and plot the normalized difference function:

$$d_a(R_1) = \frac{t_1^{-1}(R_1) - t_a^{-1}(R_1)}{\max(|t_1^{-1}(R_1) - t_a^{-1}(R_1)|)}, \quad (70)$$

over the interval of interest:  $[0 : 1 - H_2(p_1 \star \alpha_0)]$  where:  $1 - H_2(p_1 \star \alpha_0) = (1 - e_2)(1 - H_2(\alpha_0))$ .

The function  $d_a$  being strictly positive, the claim of strict inclusion is thus shown.

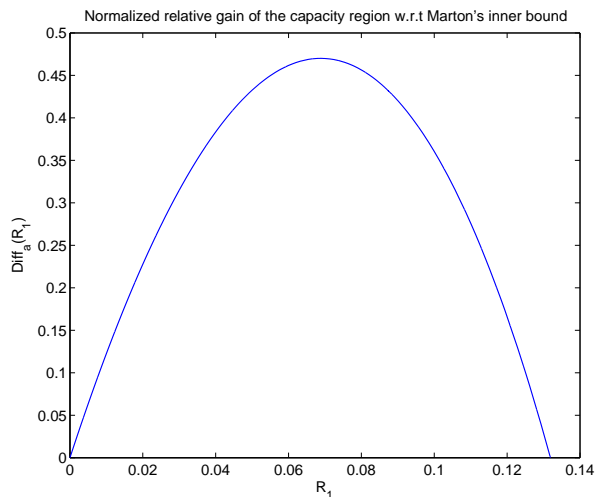


Fig. 3.  $d_a(R_1)$  the normalized relative gain of the capacity region with respect to Marton's inner bound for  $a = 0.92$ ,  $e_2 = 0.46$ ,  $p = 0.1$  and  $p_1 = 0.13$ .

We have investigated so far the role that alternative decoding techniques, namely “Interference Decoding”, play in the Compound BC where the users present a given hierarchy unknown at the encoder. The decoding technique takes advantage of the many possible decoding ways to alleviate the constraint of superposition coding at the source which allows the latter to apply a “symmetric” encoding rule regardless of which channel controls the communication. In the sequel, we analyze a class of non-ordered Compound BC to infer novel strategies when there is no specific order between channels users. In this case, we will not seek to optimize the decoder but rather the encoding technique.

#### IV. MULTIPLE DESCRIPTION CODING IN THE COMPOUND BROADCAST CHANNEL

In this section, we investigate a coding technique, referred to as “Multiple Description (MD) coding”, that can enhance the achievable rates in the Compound BC. The utility of this coding arises especially when no sort of order between the many possible instances of the users channels exists. The main idea behind MD coding is to convey the message intended to the

many instances of the same group of users, through a common description as well as a set of dedicated private descriptions which can be easily decoded each at their respective instances. The common description –to be decoded by all users– will suffer from the compound setup in that the rate has to be enough small to be decodable by all users in the same group whereas the private descriptions suffer no such loss. It is worth mentioning here that the introduction of private descriptions will also result in a loss tantamount to their “correlation cost”. We aim at exploring the utility of MD coding in the Compound BC setting.

In the sequel, for a matter of conciseness, we choose to address the Compound BC setting when only one user has two possible channels, namely  $Y_1$  or  $Y_2$ , whilst the other user suffers from no such uncertainty  $Z$ . We first derive two inner bounds on the capacity region to be compared: the Common Description (CD) inner bound that is equivalent to Marton’s worst-case inner bound, and the MD inner bound. We then specialize the bounds to the Compound MISO BC and show how MD coding outperforms the standard CD coding. Finally, we analyze the behavior of the obtained rate regions compared to our outer bound.

#### A. Multiple Description (MD) Inner Bound

**Theorem 5** (MD inner bound). *An inner bound on the capacity region of  $2 \times 1$  Compound BC is given by the set of rate pairs  $(R_1, R_2)$  satisfying:*

$$R_1 \leq I(U_0U_1; Y_1|Q) , \quad (71a)$$

$$R_1 \leq I(U_0U_2; Y_2|Q) , \quad (71b)$$

$$2R_1 \leq I(U_0U_1; Y_1|Q) + I(U_0U_2; Y_2|Q) - I(U_1; U_2|QU_0) , \quad (71c)$$

$$R_2 \leq I(V; Z|Q) , \quad (71d)$$

$$R_1 + R_2 \leq I(U_0U_1; Y_1|Q) + I(V; Z|Q) - I(U_0U_1; V|Q) , \quad (71e)$$

$$R_1 + R_2 \leq I(U_0U_1; Y_2|Q) + I(V; Z|Q) - I(U_0U_2; V|Q) , \quad (71f)$$

$$\begin{aligned} 2R_1 + R_2 \leq I(U_0U_1; Y_1|Q) + I(U_0U_2; Y_2|Q) + I(V; Z|Q) \\ - I(U_0U_1U_2; V|Q) - I(U_1; U_2|QU_0) , \end{aligned} \quad (71g)$$

$$\begin{aligned} 2R_1 + 2R_2 \leq I(U_0U_1; Y_1|Q) + I(U_0U_2; Y_2|Q) + 2I(V; Z|Q) \\ - I(U_0U_1; V|Q) - I(U_0U_2; V|Q) - I(U_1; U_2|QU_0V) , \end{aligned} \quad (71h)$$

for some set of arbitrarily correlated RVs of joint pmf:  $P_{QU_0U_1U_2VX}$  such that the Markov chain  $(Q, U_0, U_1, U_2, V) \text{---} X \text{---} (Y_1, Y_2, Z)$  holds.

*Proof:* The proof is given in Appendix H. ■

### B. Common Description (CD) Inner Bound

Inspired by Marton's inner bound, we can derive what we call the "common description" (CD) coding –worst-case of Marton's inner bound– that consists of all rate pairs  $(R_1, R_2)$  verifying:

$$R_1 \leq \min_{j \in \{1,2\}} I(U; Y_j | Q) , \quad (72)$$

$$R_2 \leq I(V; Z | Q) , \quad (73)$$

$$R_1 + R_2 \leq \min_{j \in \{1,2\}} I(U; Y_j | Q) + I(V; Z | Q) - I(U; V | Q) , \quad (74)$$

where  $U$ ,  $V$  and  $Q$  are arbitrarily correlated auxiliary RVs.

Without time-sharing, this inner bound imposes that both users in the compound setting decode the same set of variables and does not allow to treat the two possible outputs differently. However, time-sharing helps enhance the performance of this region since it allows for different signaling strategies across the time slots. The combination of the two techniques is denoted in literature as "symbol or block expansion" [14] and allows CD coding to achieve the optimal DoF for some classes of the compound MISO BC. It is easy to check that MD inner bound (71) recovers the CD inner bound (72) by setting both private descriptions equal to:  $U_1 \equiv \emptyset$  and  $U_2 \equiv \emptyset$ . Thus, implying that Marton's inner bound can achieve the optimal DoF for the compound  $2 \times 1$  Gaussian MISO BC, the question of whether MD inner bound can strictly improve on CD inner bound arises, and will be investigated in this section.

### C. MD Coding over the BC and the Compound Channel

In this section, we elaborate on the fact that CD coding performs at least as good as MD coding in both the BC and the Compound Channel.

As for the Compound Channel, let us assume that we have a compound model with two possible channel outputs denoted by  $Y_1$  and  $Y_2$ . We want to show that, for all joint pmfs  $P_{U_0U_1U_2X}$  there exists a common auxiliary RV  $U_0^*$  that yields a rate greater than the one achieved by using



MD coding. Let

$$R(p_{U_0U_1U_2X}) \triangleq \min \{I(U_0U_1; Y_1), I(U_0U_2; Y_2), \quad (75)$$

$$\frac{1}{2}[I(U_0U_1; Y_1) + I(U_0U_2; Y_2) - I(U_1; U_2|U_0)]\} , \quad (76)$$

where we have that:

$$I(U_0U_1; Y_1) \leq I(U_0U_1U_2; Y_1) , \quad (77)$$

$$I(U_0U_2; Y_2) \leq I(U_0U_1U_2; Y_2) , \quad (78)$$

$$I(U_0U_1; Y_1) + I(U_0U_2; Y_2) - I(U_1; U_2|U_0) \leq I(U_0U_1U_2; Y_1) + I(U_0U_1U_2; Y_2) , \quad (79)$$

and thus,

$$R(p_{U_0U_1U_2X}) \leq \min \{I(U_0U_1U_2; Y_1), I(U_0U_1U_2; Y_2)\} . \quad (80)$$

By letting  $U_0^* = (U_0U_1U_2)$ , the desired equality holds<sup>2</sup>:

$$\max_{p_{U_0U_1U_2X}} R(p_{U_0U_1U_2X}) = \max_{p_{U_0X}} \min \{I(U_0; Y_1), I(U_0; Y_2)\} . \quad (81)$$

As a matter of fact, for the case of the standard BC it turns out that MD coding do not help much neither. To check this, for  $Y_1 \equiv Y_2$ , fix a joint pmf  $p_{U_0U_1U_2|X}$  and let us assume that

$$I(U_0U_1; Y_1) - I(U_0U_1; V) \leq I(U_0U_2; Y_1) - I(U_0U_2; V) . \quad (82)$$

Then, it is easy to see that the choice  $U^* = (U_0U_2)$  and  $U_1^* = U_2^* = \emptyset$  allows us to get:

$$R(p_{U_0U_1U_2X}) \leq \max_{p_{U_0X}} \{I(U_0; Y_1) - I(U_0; V)\} . \quad (83)$$

Hence,

$$\max_{p_{U_0U_1U_2X}} R(p_{U_0U_1U_2X}) = \max_{p_{U_0X}} \min \{I(U_0; Y_1), I(U_0; Y_2)\} . \quad (84)$$

Needless to say that in the compound BC, the previous assertion is not true any longer since it is not known whether the inequalities:

$$I(U_0U_1; Y_1) - I(U_0U_1; V) \leq I(U^*; Y_1) - I(U^*; V) , \quad (85)$$

$$I(U_0U_2; Y_2) - I(U_0U_2; V) \leq I(U^*; Y_2) - I(U^*; V) , \quad (86)$$

$$\sum_{j=1}^2 [I(U_0U_j; Y_j) - I(U_0U_j; V)] - I(U_1; U_2|U_0V) \leq \sum_{j=1}^2 [I(U^*; Y_j) - I(U^*; V)] , \quad (87)$$

<sup>2</sup>The inequality in the inverse order is trivial by setting  $U_1 = U_2 = \emptyset$ .

still hold for some  $U^*$ , and this is the key reason for which MD is useful. However, MD proves to be useless in the cases of the BC and the Compound Channel while no evidence on its role in the Compound BC was stated. This motivates the following comparison between the CD and MD coding techniques for the  $2 \times 1$  Compound MISO BC.

## V. THE REAL COMPOUND MISO BC AND MD BASED DPC

The optimal transmit strategy for the non-ordered Gaussian MISO BC is to apply Dirty-Paper Coding (DPC) [26], [21], which is a non-linear coding technique that allows the decoder to suppress the interference. In the sequel, we derive the inner bounds resulting from an adequate use of the DPC scheme with the MD coding technique, referred to as MD-DPC, and later study a specific class of Compound MISO BC for which MDs are of consequent utility compared to the basic CD coding, referred to as CD-DPC.

Consider the Compound MISO BC which consists of a source equipped with 2 antennas and 2 single antenna receivers. Receiver 1 has two possible outputs, namely,  $Y_1$  and  $Y_2$ , and let  $Z$  be the channel output of the receiver 2, where these outputs at time  $t = [1, \dots, n]$  are given by

$$\begin{cases} y_{j,t} &= \mathbf{h}_j^t \mathbf{x}_t + n_{j,t} , \\ z_t &= \mathbf{g}^t \mathbf{x}_t + w_t , \end{cases} \quad (88)$$

for  $j \in \{1, 2\}$ , where:  $\mathbf{h}_j$  and  $\mathbf{g}$  are  $2 \times 1$  generic real channel vectors that are assumed to be constant throughout the transmission. Moreover, it is assumed that any subset of 2 channels among them are linearly independent;  $\mathbf{x}$  is the  $2 \times 1$  power limited channel input vector so that  $\mathbb{E}[\mathbf{x}^t \mathbf{x}] \leq P$  and last, the noise sequences  $\{n_{j,t}\}$  and  $\{w_t\}$  are assumed to be i.i.d. draws according to a standard Gaussian distribution  $\mathcal{N}(0, N)$ .

In this section, we will compare the CD to the MD inner bound under two different coding techniques depending on the correlation between the private auxiliary RVs. We first start with the case where the private descriptions are *uncorrelated* in the way that the encoder communicates part of the time a private description  $U_1$  to help user  $Y_1$  to decode the intended message, and a private description  $U_2$  during the remaining part of the time to help user  $Y_2$ . Later, we consider *arbitrary correlation* between the private descriptions in that both are transmitted all along time, resulting in a non-zero correlation cost.

### A. Preliminaries and Useful Definitions

In the sequel, we resort to DPC [27] in its vector formulation, thus some basic definitions and analytic formulas will be introduced herein to lighten the notation afterwards.

Let us consider the following coding scheme:

$$\begin{cases} U_0 &= X_u + \alpha X_v, \\ V &= X_v, \\ \mathbf{X} &= X_u \mathbf{B}_u + X_v \mathbf{B}_v, \end{cases} \quad (89)$$

where  $X_u \sim \mathcal{N}(0, P_u)$  and  $X_v \sim \mathcal{N}(0, P_v)$  are independent RVs such that  $P_u + P_v \leq P$ . It is then easy to check that:

$$I(U_0; Y_j) - I(U_0; V) = \log_2 \left( \frac{h_{j,u}^2 P_u + N}{I_j(\alpha - \beta_j)^2 + N} \right), \quad (90)$$

where:

$$\beta_j = \frac{P_u h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N} \quad \text{and} \quad I_j = \left( \frac{P_v}{P_u} \right) \frac{(h_{j,u}^2 P_u + N)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (91)$$

We now choose to transmit an additive private description  $X_p \sim \mathcal{N}(0, x)$  while keeping the total useful power equal to  $P_u$ , i.e.,  $0 \leq x \leq P_u$ . Then, with the following coding scheme:

$$\begin{cases} U_0 &= X_u + \alpha X_v, \\ U_j &= X_p + \alpha_j X_v, \\ \mathbf{X} &= (X_u + X_p) \mathbf{B}_u + X_v \mathbf{B}_v, \end{cases} \quad (92)$$

we can optimize the value of the private DPC parameter  $\alpha_j$  to state the following result.

**Lemma 1** (Optimizing the private descriptions).

$$\max_{\alpha_j \in \mathbb{R}} [I_{\alpha_j}(U_0 U_j; Y_j) - I_{\alpha_j}(U_0 U_j; V)] = \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + N}{\frac{I_j^x N (\alpha - \beta_j^x)^2}{h_{j,u}^2 x + N} + N} \right), \quad (93)$$

and where, for  $j \in \{1, 2\}$ , we have:

$$\beta_j^x = \frac{(P_u - x) h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N} \quad \text{and} \quad I_j^x = \left( \frac{P_v}{P_u - x} \right) \frac{(h_{j,u}^2 P_u + N)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (94)$$

*Proof:* The key point of the proof is that the private description, when optimized, yields an interference free link:

$$\max_{\alpha_j \in \mathbb{R}} [I_{\alpha_j}(U_j; Y_j | U_0) - I_{\alpha_j}(U_j; V | U_0)] = I(X_p; Y_j | X_u X_v) \quad (95)$$

$$= \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 x + N}{N} \right). \quad (96)$$

The rest of the proof is relegated to Appendix I. ■

### B. Common Description DPC (CD-DPC)

Consider the channel model defined by (88) and let us define the two following rate regions resulting from two antagonist DPC schemes:

$$\mathcal{R}_1 : \begin{cases} R_1 \leq \max_{\alpha} \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + N}{I_j(\alpha - \beta_j)^2 + N} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{g_u^2 P_u + g_v^2 P_v + N}{g_u^2 P_u + N} \right), \end{cases} \quad (97)$$

where  $\beta_j$  and  $I_j$  are given similarly by:

$$\beta_j = \frac{P_u h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N} \quad \text{and} \quad I_j = \left( \frac{P_v}{P_u} \right) \frac{(h_{j,u}^2 P_u + N)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (98)$$

The second rate region is given by the set of rate pairs satisfying:

$$\mathcal{R}_2 : \begin{cases} R_2 \leq \frac{1}{2} \log_2 \left( \frac{g_v^2 P_v + N}{N} \right), \\ R_1 \leq \min_{j=1,2} \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}{h_{j,v}^2 P_v + N} \right). \end{cases} \quad (99)$$

**Proposition 2** (CD inner bound). *An inner bound on the capacity region of the Compound MISO BC defined in (88) is given by the set of rates satisfying:*

$$\mathcal{R}_{CD-MISO BC} = \bigcup_{\substack{(P_u, P_v) \\ P_u + P_v \leq P}} \bigcup_{\substack{\mathbf{B}_u, \mathbf{B}_v \\ \|\mathbf{B}_u\|=1 \\ \|\mathbf{B}_v\|=1}} [\mathcal{R}_1(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v) \cup \mathcal{R}_2(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v)]. \quad (100)$$

*Proof:* First, note that the rate regions  $\mathcal{R}_1$  and  $\mathcal{R}_2$  are nothing but the two corner points of the CD rate region given in (72). The rate region  $\mathcal{R}_1$  is obtained by evaluating the corner point:

$$\begin{cases} R_1 \leq \min_{j \in \{1,2\}} I(U; Y_j | Q) - I(U; V | Q) \\ R_2 = I(V; Z | Q) \end{cases}, \quad (101)$$

using the following coding scheme:

$$\begin{cases} \mathbf{X} &= X_u \mathbf{B}_u + X_v \mathbf{B}_v , \\ U &= X_u + \alpha X_v = X_u + \alpha V , \end{cases} \quad (102)$$

where  $X_u \sim \mathcal{N}(0, P_u)$  and  $X_v \sim \mathcal{N}(0, P_u)$  are independent RVs such that  $P_u + P_v \leq P$ .

As for the second rate region  $\mathcal{R}_2$ , it results from the evaluation of the second corner point of CD under the antagonist coding scheme, where  $V$  dirty-paper codes the codewords  $U$ ; the calculations follow in a similar manner. ■

### C. MD-DPC with Uncorrelated Private Descriptions

In the sequel, we will evaluate the MD inner bound given in Theorem 5. To this end, we explore two different approaches for MD-DPC depending on the existing correlation between the private descriptions, for which it will be enough to study the specific corner points:

$$\begin{cases} R_1 &\leq \min_{j \in \{1,2\}} [I(U_0 U_j; Y_j | Q) - I(U_0 U_j; V | Q)] \\ 2R_1 &\leq \sum_{j \in \{1,2\}} [I(U_0 U_j; Y_j | Q) - I(U_0 U_j; V | Q)] - I(U_1; U_2 | U_0 V Q) \\ R_2 &= I(V; Z | Q) . \end{cases} \quad (103)$$

The MD inner bound we derive here is based on the evaluation of (103) via a time-sharing argument [14], where, unlike the common description, each private description is transmitted only part of the time. Both common and private descriptions apply a DPC scheme, but with difference parameters and signalings as will be clarified later. Let  $Q$  be a binary valued time-sharing RV such that:

$$\mathbb{P}(Q = 1) = 1 - \mathbb{P}(Q = 2) \triangleq t . \quad (104)$$

Let us define the following rate region  $\mathcal{R}_u$  as:

$$\mathcal{R}_u : \begin{cases} R_1 &\leq \max_{\alpha} \min_{j \in \{1,2\}} \left\{ \frac{1}{2} p_Q(j) \log_2 \left( \frac{h_{j,u}^2 x + N}{N} \right) \right. \\ &\quad \left. + \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + N}{I_j^x (\alpha - \beta_j^x)^2 + N + h_{j,u}^2 x} \right) \right\} , \\ R_2 &\leq \frac{1}{2} \log_2 \left( \frac{g_u^2 P_u + g_v^2 P_v + N}{g_u^2 P_u + N} \right) , \end{cases}$$

where  $\beta_j^x$  and  $I_j^x$  are chosen as follows:

$$\beta_j^x = \frac{(P_u - x) h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N} \quad \text{and} \quad I_j^x = \left( \frac{P_v}{P_u - x} \right) \frac{(h_{j,u}^2 P_u + N)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N} . \quad (105)$$

**Proposition 3** (MD-DPC inner bound with uncorrelated private descriptions). *An inner bound on the capacity region of the Compound MISO BC defined in (88) is given by:*

$$\mathcal{R}_{MDindep-MISO BC} = \bigcup_{t \in [0:1]} \bigcup_{\substack{(P_u, P_v) \\ P_u + P_v \leq P \\ 0 \leq x \leq P_u}} \bigcup_{\substack{\mathbf{B}_u, \mathbf{B}_v \\ \|\mathbf{B}_u\|=1 \\ \|\mathbf{B}_v\|=1}} \mathcal{R}_u(\mathbf{B}_u, \mathbf{B}_v, x, t, P_u, P_v). \quad (106)$$

*Proof:* For  $Q = 1$ , we let:

$$\begin{cases} \mathbf{X} &= (X_u + X_p) \mathbf{B}_u + X_v \mathbf{B}_v, \\ U_0 &= X_u + \alpha X_v, \\ U_2 &= \emptyset, \\ U_1 &= X_p + \alpha_1 X_v. \end{cases} \quad (107)$$

And alternately for  $Q = 2$ , let:

$$\begin{cases} \mathbf{X} &= (X_u + X_p) \mathbf{B}_u + X_v \mathbf{B}_v, \\ U_0 &= X_u + \alpha X_v, \\ U_1 &= \emptyset, \\ U_2 &= X_p + \alpha_2 X_v. \end{cases} \quad (108)$$

In this case, the correlation term becomes null since  $U_1$  and  $U_2$  are never activated in the same time slot. Hence, (103) becomes equal to:

$$R_1 \leq I(U_0; Y_1|Q) - I(U_0; V|Q) + t \left[ I(U_1; Y_1|U_0, Q=1) - I(U_1; V|U_0, Q=1) \right], \quad (109)$$

$$R_1 \leq I(U_0; Y_2|Q) - I(U_0; V|Q) + \bar{t} \left[ I(U_2; Y_2|U_0, Q=2) - I(U_2; V|U_0, Q=2) \right], \quad (110)$$

$$R_2 \leq I(V; Z|Q). \quad (111)$$

The key point is then to note that, for  $j \in \{1, 2\}$ :

$$I(U_0; Y_j|Q) - I(U_0; V|Q) \stackrel{(a)}{=} \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + N}{I_j^x (\alpha - \beta_j^x)^2 + N + h_{j,u}^2 x} \right), \quad (112)$$

where (a) is a result of that the CD suffers from the interference of the private description power  $h_{j,u}^2 x$  over both time slots in the exact same manner, be it from the private description  $U_1$  or from  $U_2$ . Finally, the result follows by using Lemma 1 to maximize the private DPC parameters  $\alpha_1$  and  $\alpha_2$ . ■

#### D. MD-DPC with Correlated Private Descriptions

In this section, we allow the private descriptions  $U_1$  and  $U_2$  in (103) to be arbitrarily correlated. Let the set of rate pairs  $\mathcal{R}_c$  defined by:

$$\mathcal{R}_c : \begin{cases} R_1 \leq \min\{f_1(\alpha, x), f_2(\alpha, x)\}, \\ R_1 \leq \frac{1}{2} \left[ f_1(\alpha, x) + f_2(\alpha, x) - \frac{1}{2} \log_2(2\pi e x) \right], \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{g_u^2 P_u + g_v^2 P_v + N}{g_u^2 P_u + N} \right), \end{cases}$$

where:

$$f_j(\alpha, x) \triangleq \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + N}{\frac{I_j^x N (\alpha - \beta_j^x)^2}{h_{j,u}^2 x + N} + N} \right), \quad (113)$$

and  $\beta_j^x$  and  $I_j^x$  are given similarly to (94) by:

$$\beta_j^x = \frac{(P_u - x) h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N} \quad \text{and} \quad I_j^x = \left( \frac{P_v}{P_u - x} \right) \frac{(h_{j,u}^2 P_u + N)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (114)$$

**Proposition 4** (MD inner bound with correlated private descriptions). *An inner bound on the capacity region of the Compound MISO BC is given by:*

$$\mathcal{R}_{MDcorr-MISO BC} = \bigcup_{\substack{\mathbf{B}_u, \mathbf{B}_v \\ \|\mathbf{B}_u\|=1 \\ \|\mathbf{B}_v\|=1}} \bigcup_{\substack{(P_u, P_v) \\ P_u + P_v \leq P}} \bigcup_{\alpha \in \mathbb{R}} \mathcal{R}_c(\mathbf{B}_u, \mathbf{B}_v, \alpha, x, P_u, P_v). \quad (115)$$

*Proof:* To prove our claim, we resort to the MD coding inner bound letting, in the single letter, the two ARV be  $U_1$  and  $U_2$  equal given  $Q$ ,  $U_0$ , and  $V$ . The correlation term becomes thus:

$$I(U_1; U_2 | Q U_0 V) = H(U_1 | Q U_0 V) = H(U_2 | Q U_0 V). \quad (116)$$

Let us use the following coding scheme:

$$\begin{cases} \mathbf{X} = (X_u + X_p) \mathbf{B}_u + X_v \mathbf{B}_v, \\ U_0 = X_u + \alpha X_v, \\ U_1 = X_p + \alpha_1 X_v, \\ U_2 = X_p + \alpha_2 X_v, \\ V = X_v. \end{cases} \quad (117)$$

It is then straightforward with the result of Lemma 1, that the achievable rates are those given in the proposition. ■

### E. MD-DPC strictly outperforms CD-DPC

Let us now be in the presence of the most stringent compound model where  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are unit-norm orthogonal channels. Assume also that the other user's channel is quite accommodating such that  $\mathbf{g}$  is orthogonal to the "mean channel" of user 1,

$$\mathbf{g} \perp \frac{1}{\sqrt{2}}(\mathbf{h}_1 + \mathbf{h}_2) = \mathbf{h}_{1,2}. \quad (118)$$

In order to show that MD-DPC strictly outperforms CD-DPC for this setting, we need to evaluate CD-DPC inner bound based on the corresponding channel models. Then, we show that the MD-DPC inner bound strictly outperforms it.

1) *CD-DPC inner bound:* We start by characterizing CD-DPC inner bound in a closed form.

**Proposition 5** (CD-DPC inner bound). *The CD-DPC inner bound writes as the set of rate pairs satisfying:*

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{P(\eta) + 2N} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{(1-\eta)P_v + 2N}{2N} \right), \end{cases} \quad (119)$$

for some  $\eta \in [-1 : 1]$ , where

$$P(\eta) \triangleq \frac{(1-\eta)P_u P_v}{P + 2N + \sqrt{(P + 2N)^2 + (\eta^2 - 1)P_v^2}}. \quad (120)$$

*Proof:* The proof is relegated to Appendix J. ■

**Remark 6.** *In order to derive the optimal value of  $\eta$  for the overall rate region, we look at the resulting weighted sum-rate. If we let  $\mu \in \mathfrak{R}_+$ , then the optimization of  $R_1 + \mu R_2$  over  $\eta$  depends on the value of  $\mu$ . For  $\mu = 0$ , the optimal choice is  $\eta = 1$  that is we have to transmit in a direction that is collinear with the mean channel  $\mathbf{h}_{1,2}$ , as for the case  $\mu \rightarrow \infty$ , the optimal choice is to let  $\eta = -1$ , which means to transmit the information for the second user in a direction that is collinear to its channel. For intermediate values of  $\mu$ , the weighted sum-rate is not necessarily maximized with either choices of  $\eta$ .*

We evaluate the two MD-DPC inner bounds as a function of  $x$ , the power dedicated to private descriptions, and compare them to the case  $x = 0$ , i.e., the CD-DPC inner bound. We let  $\mathbf{B}_u = \mathbf{h}_{1,2}$  and thus, by transmitting information to user 1 orthogonal to the channel of user 2.



2) *MD-DPC with correlated private descriptions outperforms CD-DPC:*

To evaluate the gain of MD-DPC inner bound with *arbitrarily correlated* private descriptions, note that if at least  $0 \leq x \leq (2\pi e)^{-1}$ , then the bound on  $R_1$  can be written as follows:

$$R_1 \leq \frac{1}{2} \max_{\alpha \in \mathbb{R}} \min \{ f_1(\alpha, x), f_2(\alpha, x) \} \quad (121)$$

$$= \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{(P_u - x) 2N}{x + 2N} \frac{P(\eta)}{P_u} + 2N} \right) \quad (122)$$

$$\stackrel{(a)}{\geq} \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{P(\eta) + 2N} \right), \quad (123)$$

where (a) follows from the fact that the function  $f : x \mapsto \frac{(P_u - x)}{x + 2N}$  is strictly decreasing in  $x$ . Indeed, the inequality in (a) is strict for non-degenerate power parameters  $P_v \neq 0$  and  $\eta \neq 1$ , which corresponds to  $R_2 \neq 0$  and yields the proof of the claim.

3) *MD-DPC with uncorrelated private descriptions outperforms CD-DPC:* As for MD-DPC inner bound with *uncorrelated private descriptions*, the constraint on the rate  $R_1$  writes as:

$$R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{(P_u - x) \sqrt{2N}}{\sqrt{x + 2N}} \frac{P(\eta)}{P_u} + \sqrt{2N} \sqrt{x + 2N}} \right), \quad (124)$$

for which we have considered a time-sharing  $t = \bar{t} = 0.5$ . Now, the function given by

$$g(x) \triangleq \frac{(P_u - x) P(\eta)}{\sqrt{x + 2N} P_u} + \sqrt{x + 2N}, \quad (125)$$

is not compulsorily strictly decreasing in  $x$  for all values of  $\eta$ . However, it is clear that:

$$g'(x) = \frac{(x + 2N)P(\eta) + (P_u + 2N)(P_u - P(\eta))}{2P_u(x + 2N)^{3/2}}, \quad (126)$$

and since  $0 \leq x \leq P_u$ , then:

$$g'(x) \leq \frac{1}{4P_u N \sqrt{2N}} (P_u + 2N)(P_u - 2P(\eta)). \quad (127)$$

Thus,  $P(\eta) > \frac{P_u}{2}$  suffices to have the function  $g$  strictly decreasing in  $x$ , and thus, the claim of strict optimality would be proved. Note that if e.g.  $P \geq 4N$ , then for values of  $\eta$  close to  $-1$ , i.e.,  $R_2$  close to second user's capacity, the gain is strictly positive and more significant.

4) *Comparison of the MD-DPC inner bounds:* An interesting question to investigate is whether the MD inner bound with correlated descriptions outperforms or not the same with uncorrelated descriptions. These two bounds compare differently following the values of the channel gains. The MD with uncorrelated private descriptions makes each user loose:

$$\frac{1}{4} \log_2 \left( \frac{\|\mathbf{h}\|^2 x + 2N}{2N} \right) \quad (128)$$

compared to the single rates of the MD-DPC with correlated descriptions. Whereas the latter, through the correlation coefficient, engenders a loss of

$$\frac{1}{4} \log_2 (2\pi e x) . \quad (129)$$

Thus, the relative behavior of these two bounds depends on the specific values of  $N$ ,  $P_u$  and  $\|\mathbf{h}\|^2$ . In Fig. 4, we plot the corresponding rate regions for  $\text{SNR} = 10$  dB,  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 2$  and the assumptions made on the channels' structure.

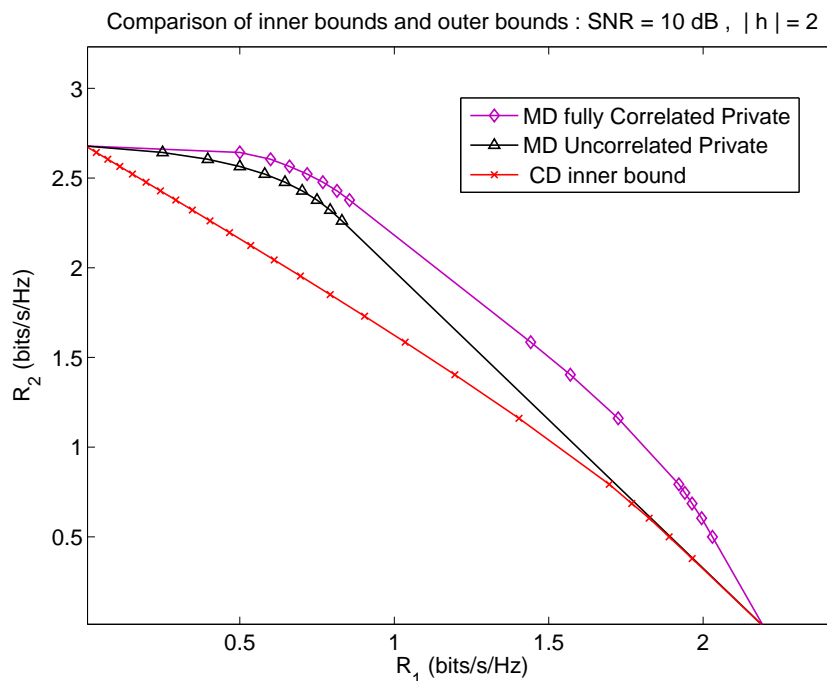


Fig. 4. Comparison of the CD-DPC and the MD-DPC inner bounds with uncorrelated and correlated private descriptions:  $\text{SNR} = 10$  dB,  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 2$ .

### F. Block Expansion

Last, the bounds we have studied so far did not allow for different encoding parameters across time slots. The reason is that the question we were exploring is one of the utility of private descriptions in the Compound MISO BC. Now, if we combine CD inner bound and MD inner bound with correlated private descriptions both with a time-sharing argument where in each time slot, a new coding scheme is used (in terms of beams, power allocations and DPC parameters), then one could expect that the behavior is still captured by the obtained bounds. Fig. 5 corroborates the previous statement.

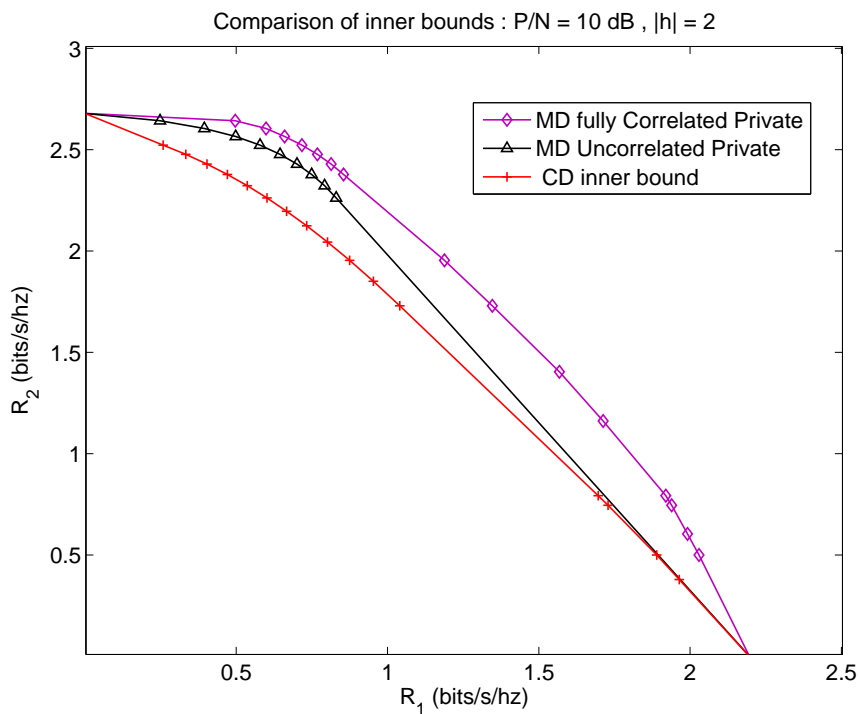


Fig. 5. Comparison of the CD-DPC and MD-DPC with uncorrelated and correlated private descriptions inner bounds with a time-sharing argument: SNR = 10 dB,  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 2$ .

Yet, Block Expansion does not enhance much the performance of MD-DPC coding scheme, the reason being these schemes allow already for good coding schemes, however, CD-DPC is much more enhanced by Block Expansion. Indeed, in the DoF analysis, Time Sharing is crucial for CD-DPC to be DoF optimal [14].

### G. Outer Bound on the Capacity of the Compound MISO BC

In this section, we present an outer bound on the capacity region of the Compound MISO BC which consists in the intersection of some rate regions.

Let us introduce the following channel matrices:

$$\mathbf{g}_{1,2} \triangleq [\mathbf{g} \ \mathbf{h}_1 \ \mathbf{h}_2] \ , \quad (130)$$

$$\mathbf{h}_{1,z} \triangleq [\mathbf{h}_1 \ \mathbf{g}] \ , \quad (131)$$

$$\mathbf{h}_{2,z} \triangleq [\mathbf{h}_2 \ \mathbf{g}] \ . \quad (132)$$

We then define the corresponding channel outputs to the channel  $\mathbf{g}_{1,2}$ , that has the same marginal as the output formed by the concatenation of  $[Z \ Y_1 \ Y_2]$ , as  $Z_{1,2}$ , and we define similarly the two outputs  $Y_{1,z}$  and  $Y_{2,z}$ . The following theorem gives the resulting outer bound.

**Theorem 7** (Outer bound on the capacity of the Compound MISO BC). *An outer bound on the capacity region of the Compound MISO BC is given by the set of rate pairs:*

$$\mathcal{O} = \mathcal{C}_1 \cap \mathcal{C}_2 \cap \mathcal{C}_{1,2} \cap \mathcal{C}_z \ , \quad (133)$$

where  $\mathcal{C}_j$  is the capacity region of the BC with outputs  $(Y_j, Z)$ , for  $j \in \{1, 2\}$ ,

$$\mathcal{C}_j = \bigcup_{\substack{(\mathbf{K}_u, \mathbf{K}_v) \\ \text{tr}(\mathbf{K}_u + \mathbf{K}_v) \leq P}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \right. \\ \left. R_1 \leq \frac{1}{2} \log_2 \left( \frac{\mathbf{h}_j^t \mathbf{K}_u \mathbf{h}_j + N}{N} \right) \right. \quad (134)$$

$$\left. R_2 \leq \frac{1}{2} \log_2 \left( \frac{\mathbf{g}^t (\mathbf{K}_u + \mathbf{K}_v) \mathbf{g} + N}{\mathbf{g}^t \mathbf{K}_u \mathbf{g} + N} \right) \right\} \quad (135)$$

or

$$R_1 \leq \frac{1}{2} \log_2 \left( \frac{\mathbf{h}_j^t (\mathbf{K}_u + \mathbf{K}_v) \mathbf{h}_j + N}{\mathbf{h}_j^t \mathbf{K}_v \mathbf{h}_j + N} \right) \quad (136)$$

$$\left. R_2 \leq \frac{1}{2} \log_2 \left( \frac{\mathbf{g}^t \mathbf{K}_v \mathbf{g} + N}{N} \right) \right\} \ , \quad (137)$$

$\mathcal{C}_{1,2}$  is the capacity region of the Compound MISO BC with outputs  $(Y_1, Z_{1,2})$  and  $(Y_2, Z_{1,2})$ ,

$$\mathcal{C}_{1,2} = \bigcup_{\substack{(\mathbf{K}_u, \mathbf{K}_v) \\ \text{tr}(\mathbf{K}_u + \mathbf{K}_v) \leq P}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \right. \\ \left. R_1 \leq \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left( \frac{\mathbf{h}_j^t (\mathbf{K}_u + \mathbf{K}_v) \mathbf{h}_j + N}{\mathbf{h}_j^t \mathbf{K}_v \mathbf{h}_j + N} \right), \right. \quad (138)$$

$$\left. R_2 \leq \frac{1}{2} \log_2 \left( \frac{|\mathbf{g}_{1,2}^t \mathbf{K}_v \mathbf{g}_{1,2} + N \mathbf{I}_3|}{N^3} \right) \right\} \quad (139)$$

and finally,  $\mathcal{C}_z$  is the capacity region of the Compound BC with outputs  $(Y_{1,z}, Z)$  and  $(Y_{2,z}, Z)$ ,

$$\mathcal{C}_z = \bigcup_{\substack{(\mathbf{K}_u, \mathbf{K}_v) \\ \text{tr}(\mathbf{K}_u + \mathbf{K}_v) \leq P}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \right. \\ \left. R_1 \leq \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left( \frac{|\mathbf{h}_{j,z}^t \mathbf{K}_u \mathbf{h}_{j,z} + N \mathbf{I}_2|}{N^2} \right) \right. \quad (140)$$

$$\left. R_2 \leq \frac{1}{2} \log_2 \left( \frac{\mathbf{g}^t (\mathbf{K}_u + \mathbf{K}_v) \mathbf{g} + N}{\mathbf{g}^t \mathbf{K}_u \mathbf{g} + N} \right) \right\}. \quad (141)$$

*Proof:* The proof is straightforward from the following observations. The fact that the capacity of the considered compound model is always included in the intersection of the capacities of the BCs  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , and that this setting is a degraded version of the setups where there is a least one user with an extra receive antenna, whose capacities are given in references [28], [29]. ■

**Remark 8.** *The outer bound stated in Theorem 7 is tight in the high SNR regime and thus is DoF optimal. To check this, notice that the bounds  $\mathcal{C}_1$ ,  $\mathcal{C}_2$  and  $\mathcal{C}_z$  attain each the points  $(d_1 \leq 1, d_2 \leq 1)$  by letting  $\mathbf{K}_u = \mathbf{g}^\perp \times (\mathbf{g}^\perp)^t$ . As for the bound  $\mathcal{C}_{1,2}$ , it achieves all the points  $(2d_1 + d_2 \leq 2)$ , thus the intersection of these two regions leads to the optimal DoF.*

In Fig. 6, we plot the inner and outer bound for intermediate SNR values. Although the gap with the outer bound suggests that the inner and outer regions do not meet, it is our belief that the inner bound is tight while our outer remains rather loose.

## VI. SUMMARY AND DISCUSSION

We start our conclusions with the analysis of the relative behavior of the MD and the ID inner bounds, to understand if there is any mutual inclusion between the two bounds. The question

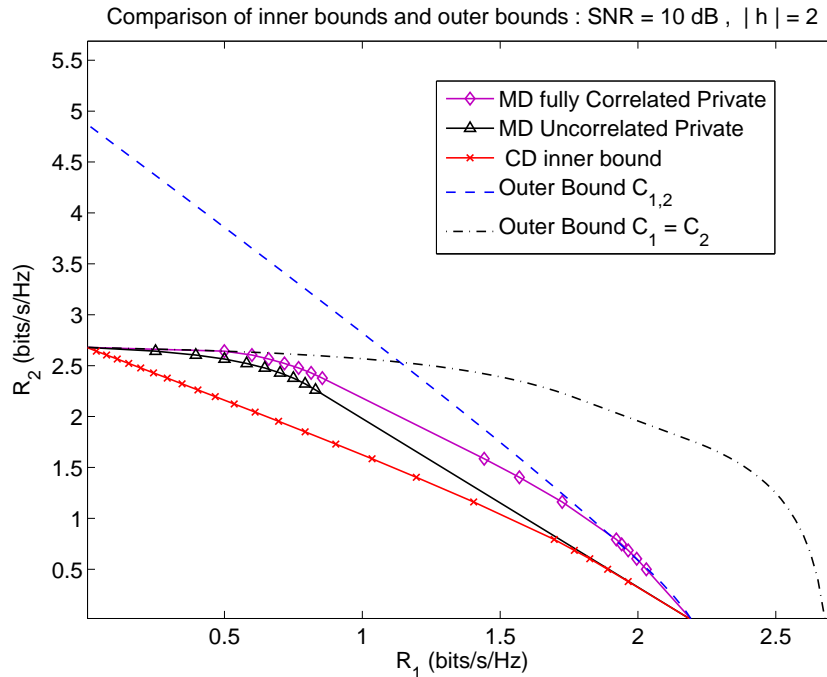


Fig. 6. Comparison of the inner bounds and the intersection of the outer bounds: SNR = 10 dB,  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 2$ .

we want to answer is whether introducing multiple descriptions, one for each instance in the compound setting, allows to recover the ID inner bound. We also would like to understand to what extent decoding interference is crucial for Marton's worst case inner bound.

#### A. Can Multiple Descriptions or Interference Decoding techniques recover each other?

For this sake, we evaluate the MD inner bound in the case of the discrete example studied in Section III-B and try to identify a set of auxiliary RVs yielding the capacity region. For the discrete Compound BC we studied earlier, we assumed that user 1 could observe one of two possible channel instances, namely,  $Y_1$  and  $Y_2$ , such that  $Y_2$  is more capable than both  $Y_1$  and  $Z$ , and  $Y_1$  be a degraded version of  $Z$ . The maximizing choice of auxiliary RVs led to  $Z$  and  $Y_2$  decoding all the signal and  $Y_1$  decoding only its intended information.

The capacity region is of the form:

$$\begin{cases} R_1 \leq I(Q; Y_1) , \\ R_1 + R_2 \leq I(Q; Y_1) + I(X; Z|Q) . \end{cases} \quad (142)$$

We next discuss a formulation of the MD inner bound that captures the intuition of the capacity achieving choice of auxiliary RV for ID inner bound. Indeed, the encoder does not transmit a common description to the two users interested in the same message, but communicate only private descriptions to them. However, in the present case the common auxiliary RV  $Q$  is no longer a time-sharing variable as it was the case in Section IV, it can carry common information to all receivers as well. With this, we can achieve the set of rate pairs satisfying:

$$\mathcal{R}_{3\text{-ARV}} = \bigcup_{pQU_1U_2VX} \left( \bigcup_{(T_{1,1}, T_{1,2}, T_2) \in \mathbb{T}(p)} \mathcal{M}(p, T_{1,1}, T_{1,2}, T_2) \right), \quad (143)$$

where  $\mathcal{M}$  and  $\mathbb{T}$  are respectively defined by the following:

$$\mathcal{M} : \begin{cases} T_2 \leq I(V; Z|Q), \\ R_0 + T_2 \leq I(QV; Z), \\ T_{1,1} \leq I(U_1; Y_1|Q), \\ R_0 + T_{1,1} \leq I(QU_1; Y_1), \\ T_{1,2} \leq I(U_2; Y_2|Q), \\ R_0 + T_{1,2} \leq I(QU_2; Y_2), \end{cases} \quad (144)$$

$$\mathbb{T} = \left\{ (T_{1,1}, T_{1,2}, T_2) : T_2 \geq R_2, \min\{T_{1,1}, T_{1,2}\} \geq R_1, \right. \quad (145)$$

$$T_{1,1} - R_1 + T_2 - R_2 > I(U_1; V|Q), \quad (146)$$

$$T_{1,2} - R_1 + T_2 - R_2 > I(U_2; V|Q), \quad (147)$$

$$T_{1,1} - R_1 + T_{1,2} - R_1 > I(U_1; U_2|Q), \quad (148)$$

$$T_{1,1} + T_{1,2} - 2R_1 + T_2 - R_2 > I(U_1; U_2|Q) + I(U_1U_2; V|Q) \left. \right\}. \quad (149)$$

*Proof:* The proof is relegated to Appendix L. ■

We know that an optimal transmission scheme to achieve the capacity region of the considered BEC/BSC requires both users  $Z$  and  $Y_2$  to decode all messages while restricting the weaker user  $Y_1$  to decode only the common message. Hence, we rely on this argument to build the straightforward extension of Marton's coding scheme, i.e.,  $V = U_2 = X$  and  $U_1 = Q$ , which along with rate splitting leads to the following achievable rate region:

$$\begin{cases} R_1 \leq I(Q; Y_1), \\ R_1 + R_2 \leq I(X; Z|Q) + I(X; Y_2|Q) + \min\{I(Q; Y_1), I(Q; Y_2)\} - H(X|Q). \end{cases} \quad (150)$$

In the general case, there is strong evidence that the above rate region induced by MD is strictly included in the capacity region given by:

$$\begin{cases} R_1 \leq I(Q; Y_1), \\ R_1 + R_2 \leq I(X; Z|Q) + I(Q; Y_1), \end{cases} \quad (151)$$

that is achieved by using ID, which yields:

$$\begin{cases} R_1 \leq I(Q; Y_1), \\ R_1 + R_2 \leq \min\{I(X; Z|Q), I(X; Y_2|Q)\} + I(Q; Y_1), \\ R_1 + R_2 \leq \min\{I(X; Z), I(X; Y_2)\}, \end{cases} \quad (152)$$

where  $Y_1$  is degraded with respect to  $Z$  and  $Y_2$  is more capable than  $Z$ . The inclusion results from the fact that there exist  $P_{X|Q}$  for which

$$I(X; Y_2|Q) - H(X|Q) < 0. \quad (153)$$

Thus, MD does not seem to be enough to achieve the capacity region of the compound model investigated in Section III-B. This is due to the fact that the cost engendered by precoding against interference prevents from decoding it which results in a loss proportional to its entropy. Therefore, it appears that ID outperforms MD in some cases.

On the other hand, in the MISO case, imposing users to decode interference is sub-optimal, at least from a DoF perspective, since ID introduces sum-rates constraints of the form

$$R_1 + R_2 \leq I(X; Y_1), \quad (154)$$

and thus, prevents the sum-DoF from reaching values greater than 1 which we already know is sub-optimal. Therefore, it is crucial to precode against interference.

Summarizing, since neither MD coding or ID seem to generalize all the results obtained herein one can benefit from the combination of both techniques and thus, from the optimization of both encoding and decoding schemes.

## B. Summary

In this work, we explored a decoding and a encoding technique for the two-user memoryless Compound Broadcast Channel (BC). We first studied the role of Interference Decoding (ID) where an achievable rate region is derived by using “single per-message description” codes constructed via *superposition coding* and *random binning*. At the decoders, the constraint of



decoding only the intended message is alleviated to allow each of the users to decode or not the other user's (interference) message. Unlike for the standard two-user BC, this strategy proves to be useful in compound setups, where channel uncertainty prevents the encoder from coding optimally for each possible BC formed by all pairs of channels in the set. A simple outer bound is also derived based on the best outer bound hitherto known on the capacity region of the two-user BC. This outer bound captures one of the most stringent effects of simultaneity of users over the random codes constructed: antagonist coding strategies.

Surprisingly enough, ID not only outperforms Non-Interference Decoding (NID) technique, i.e., Marton's worst-case rate region, but also allows to achieve the capacity of a class of non-trivial BC while NID stays strictly suboptimal. Thus, though the coding scheme is simple (in terms of the number of auxiliary variables involved and of the complexity of the encoding operation) the decoders' optimization allows to palliate the uncertainty at the source.

Later, we studied an encoding technique with a more evolved coding strategy, namely Multiple Description (MD) coding. The source transmits to the group of users, interested in the same message, common and private descriptions. For the specific case of the Compound MISO BC, resorting to MD is essential since a common description, i.e., applying DPC with a single description cannot accommodate the interference seen by each instance of the users channels in the set, unless combining it with a time-sharing argument. The key point in the MISO BC setting is that using a fraction of power to transmit the private descriptions is useful for all SNR ranges while turns out to be DoF optimal. Indeed, each private description creates an interference free link and thus each user can recover a part of its rate interference free.

Finally, we addressed the question of whether MD or ID may generalize each other. It appears that none of these schemes can perform well for ordered and non-ordered class of Compound BCs at once, mainly because the two strategies strongly rely on two different interference mitigation techniques: precoding against interference and decoding interference. The first results in a rate loss tantamount to a correlation cost while the latter results in an extra sum-rate constraint.

As a conclusion, it would be worth mentioning the benefits of combining these two schemes to yield a larger inner bound, and thus, full advantage would be taken from the joint optimization of the encoding technique (MD coding) and the decoding technique (ID).

APPENDIX A  
USEFUL NOTIONS AND AUXILIARY RESULTS

In this appendix we provide basic notions on some concepts used in this paper.

Following [30], we use in this paper *strongly typical sets* and the so-called *Delta-Convention*. Some useful facts are recalled here. Let  $X$  and  $Y$  be RVs on some finite sets  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively. We denote by  $p_{XY}$  (resp.  $p_{Y|X}$ , and  $p_X$ ) the joint pmf of  $(X, Y)$  (resp. conditional distribution of  $Y$  given  $X$ , and marginal distribution of  $X$ ).

**Definition 9.** For any sequence  $x^n \in \mathcal{X}^n$  and any symbol  $a \in \mathcal{X}$ , notation  $N(a|x^n)$  stands for the number of occurrences of  $a$  in  $x^n$ .

**Definition 10.** A sequence  $x^n \in \mathcal{X}^n$  is called (strongly)  $\delta$ -typical w.r.t.  $X$  (or simply typical if the context is clear) if

$$\left| \frac{1}{n} N(a|x^n) - P_X(a) \right| \leq \delta \text{ for each } a \in \mathcal{X} ,$$

and  $N(a|x^n) = 0$  for each  $a \in \mathcal{X}$  such that  $P_X(a) = 0$ . The set of all such sequences is denoted by  $T_\delta^n(X)$ .

**Definition 11.** Let  $x^n \in \mathcal{X}^n$ . A sequence  $y^n \in \mathcal{Y}^n$  is called (strongly)  $\delta$ -typical (w.r.t.  $Y$ ) given  $x^n$  if

$$\left| \frac{1}{n} N(a, b|x^n, y^n) - \frac{1}{n} N(a|x^n) P_{Y|X}(b|a) \right| \leq \delta \text{ for each } a \in \mathcal{X}, b \in \mathcal{Y} ,$$

and,  $N(a, b|x^n, y^n) = 0$  for each  $a \in \mathcal{X}$ ,  $b \in \mathcal{Y}$  such that  $P_{Y|X}(b|a) = 0$ . The set of all such sequences is denoted by  $T_\delta^n(Y|x^n)$ .

*Delta-Convention [30]:* For any sets  $\mathcal{X}, \mathcal{Y}$ , there exists a sequence  $\{\delta_n\}_{n \in \mathbb{N}^*}$  such that lemmas below hold.<sup>3</sup> From now on, typical sequences are understood with  $\delta = \delta_n$ . Typical sets are still denoted by  $T_\delta^n(\cdot)$ .

**Lemma 2** ([30, Lemma 1.2.12]). *There exists a sequence  $\eta_n \xrightarrow[n \rightarrow \infty]{} 0$  such that*

$$P_X^n(T_\delta^n(X)) \geq 1 - \eta_n .$$

<sup>3</sup>As a matter of fact,  $\delta_n \rightarrow 0$  and  $\sqrt{n} \delta_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

**Lemma 3** ([30, Lemma 1.2.13]). *There exists a sequence  $\eta_n \xrightarrow[n \rightarrow \infty]{} 0$  such that, for each  $x^n \in T_\delta^n(X)$ ,*

$$\left| \frac{1}{n} \log \|T_\delta^n(X)\| - H(X) \right| \leq \eta_n ,$$

$$\left| \frac{1}{n} \log \|T_\delta^n(Y|x^n)\| - H(Y|X) \right| \leq \eta_n .$$

**Lemma 4** (Asymptotic equipartition property). *There exists a sequence  $\eta_n \xrightarrow[n \rightarrow \infty]{} 0$  such that, for each  $x^n \in T_\delta^n(X)$  and each  $y^n \in T_\delta^n(Y|x^n)$ ,*

$$\left| -\frac{1}{n} \log P_X^n(x^n) - H(X) \right| \leq \eta_n ,$$

$$\left| -\frac{1}{n} \log P_{Y|X}^n(y^n|x^n) - H(Y|X) \right| \leq \eta_n .$$

**Lemma 5** (Joint typicality lemma [31]). *There exists a sequence  $\eta_n \xrightarrow[n \rightarrow \infty]{} 0$  such that*

$$\left| -\frac{1}{n} \log P_Y^n(T_\delta^n(Y|x^n)) - I(X; Y) \right| \leq \eta_n \text{ for each } x^n \in T_\delta^n(X) .$$

## APPENDIX B

### SKETCH OF THE PROOF OF THEOREM 2

Let  $j \in \mathcal{J}$  be the index of an arbitrary pair of users in the compound set. We first show the achievability of the union of the four regions for this channel  $\bigcup_{i \in [1:4]} \mathcal{T}_i$ . For convenience of notations we drop the index  $j$ .

#### A. Outline of Proof

The coding scheme we use is as follows:

- We use three auxiliary RVs, one for each message,
- We perform binning on the two auxiliary RV that code for the private messages, superposing them over the auxiliary RV coding for the common message,
- The decoding will introduce the principle of list decoding, which will allow us to combine two decoding techniques,
- The error probability will be shown to be directly related to the list size, and thus, bounding the list size will allow us to have a tight bound on the average probability of error,
- The intersection of the union of the regions comes from the fact that we use two different decoding functions at the two users.

## B. Detailed Proof

Codebook generation: The encoding is similar to that of Marton's coding with a common message.

Fix,  $P_Q, P_{U|Q}, P_{V|Q}$  and let  $T_1 \geq R_1$  and  $T_2 \geq R_2$  be four positive rates.

Generate  $2^{nR_0}$   $n$ -sequences  $q^n(w_0), w_0 \in \mathcal{M}_0$  following each the probability distribution:

$$P_Q^n(w_0) = \prod_{i=1}^n P_Q(q_i(w_0)) , \quad (155)$$

and set them all in  $\mathcal{C}_0$ .

For each  $q^n(w_0)$ , generate  $2^{nT_1}$   $n$ -sequences  $u^n(l_1, w_0), l_1 \in [1 : 2^{nT_1}]$  each following

$$P_{U|Q}^n(u^n(l_1, w_0)) = \prod_{i=1}^n P_{U|Q}(u_i(l_1, w_0)|q_i(w_0)) , \quad (156)$$

and map all these sequences in  $2^{nR_1}$  bins, each indexed with  $w_1 \in [1 : 2^{nR_1}]$ :  $\mathcal{C}(w_1, w_0)$ .

Generate similarly  $2^{nT_2}$   $n$ -sequences  $v^n(l_2, w_0), l_2 \in [1 : 2^{nT_2}]$  each following  $P_{V|Q}^n(v^n(l_2, w_0))$  and map them into  $2^{nR_2}$  bins:  $\mathcal{C}(w_2, w_0)$ .

### Encoding:

To send a message vector:  $(W_0, W_1, W_2)$ , the encoder first finds a pair of sequences  $(u^n(L_1, W_0), v^n(L_2, W_0))$  in the product bins  $\mathcal{C}(W_j, W_0)$  such that:

$$(q^n(W_0), u^n(L_1, W_0), v^n(L_2, W_0)) \in T_\delta^n(QUV) , \quad (157)$$

and then transmits:  $x^n(q^n(W_0), u^n(L_1, W_0), v^n(L_2, W_0))$  which is generated via a random mapping.

### Decoding:

First, assume that no “ encoding error:  $\epsilon_0$ ” has occurred, and note:  $(L_1, L_2)$  the chosen indices. For a matter of conciseness, we consider only Decoder 1.

Given a received sequence  $y^n$ , define the two lists:

$$\begin{aligned} \mathcal{L}_1(y^n) &\triangleq \left\{ (w_0, w_1) \mid (q^n(w_0), u^n(l_1, w_0), y^n) \in T_\delta^n(QUY) \text{ for } u^n(l_1, w_0) \in \mathcal{C}(w_1, w_0) \right\} \\ \mathcal{L}_2(y^n) &\triangleq \left\{ (w_0, w_1) \mid (q^n(w_0), u^n(l_1, w_0), v^n(l_2, w_0), y^n) \in T_\delta^n(QUVY) \right. \\ &\quad \left. \text{for some } w_2, v^n(l_2, w_0) \in \mathcal{C}(w_2, w_0), \text{ and } u^n(l_1, w_0) \in \mathcal{C}(w_1, w_0) \right\}. \end{aligned} \quad (159)$$

These lists correspond to two different decoding functions: “non-unique” decoding of the other user’s message, and “not” decoding it. Denote the intersection of these two lists by

$$\mathcal{L}^{(n)} \triangleq \mathcal{L}_1(y^n) \cap \mathcal{L}_2(y^n). \quad (160)$$

Analysis of the probability of error: To analyze the probability of error at user 1, we need to control the expected cardinality of the intersection of the above lists. The next lemma (shown in Appendix C ) states this result.

**Lemma 6.** *For every  $\epsilon_1 > 0$ , the average probability of error is linked to the list size as follows:*

$$P_e^{(n)} \leq \mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} + \epsilon_1 \quad (161)$$

for  $n > \exists N_1$  large enough.

Now, bounding the probability of error will mainly consist in bounding the decoding list size.

Bounding the list size:

On one hand, the list size being an integer valued RV, we can write:

$$\mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} \leq \mathbb{E}[\|\mathcal{L}^{(n)}\|] - \mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 1\}. \quad (162)$$

On the other hand:

$$\mathbb{E}\|\mathcal{L}^{(n)}\| = \mathbb{P}\{(W_0, W_1) \in \mathcal{L}^{(n)}\} + \sum_{(w_0, w_1) \neq (W_0, W_1)} \mathbb{P}\{(w_0, w_1) \in \mathcal{L}^{(n)}\}. \quad (163)$$

The next lemma provides a bound on the expected list size from the RHS of (163). The proof is relegated to Appendix C .

**Lemma 7** (Bounding the probability of undetected errors). *The probability of decoding  $(w_0, w_1) \neq (W_0, W_1)$ , can be upper-bounded as follows:*

$$\sum_{(w_0, w_1) \neq (W_0, W_1)} \mathbb{P}\{(w_0, w_1) \in \mathcal{L}^{(n)}\} \leq \min\{I_1^{(n)}, I_2^{(n)}\}, \quad (164)$$

for  $n$  large enough, i.e.  $n > N_2$ , for some  $N_2$ , where:

$$I_1^{(n)} \triangleq \exp_2(n [T_1 - I(U; Y|Q) + \epsilon_2]) + \exp_2(n [R_0 + T_1 - I(QU; Y) + \epsilon_2]), \quad (165)$$

$$I_2^{(n)} \triangleq \exp_2(n [T_1 - I(U; YV|Q) + \epsilon_3]) + \exp_2(n [T_1 + T_2 - I(UV; Y|Q) - I(U; V|Q) + \epsilon_3]) \\ + \exp_2(n [R_0 + T_1 + T_2 - I(QUV; Y) - I(U; V|Q) + \epsilon_3]). \quad (166)$$

Hence, from (162), (163) and (164) we can write that:

$$\mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} \leq \min\{I_1^{(n)}, I_2^{(n)}\}. \quad (167)$$

Then Lemma 1 and (167), imply that for  $n$  large enough:

$$P_e^{(n)} \leq \mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} + \epsilon_1 \leq \min\{I_1^{(n)}; I_2^{(n)}\} + \epsilon_1. \quad (168)$$

Thus, provided that:

$$\limsup_{n \rightarrow \infty} \min\{I_1^{(n)}, I_2^{(n)}\} = 0, \quad (169)$$

the probability of error at user 1, knowing that no encoding error occurred, will tend to 0 as  $n \rightarrow \infty$ .

Following the proof of the Covering lemma [31], the probability of encoding error can be upper bounded as  $n$  grows large enough as follows:

$$\mathbb{P}(\epsilon_0) \leq \exp_2(n [I(U; V|Q) - (T_1 - R_1 + T_2 - R_2) + \epsilon']). \quad (170)$$

The condition for no such error does not depend on the users pair index, and thus, it intersects the union of all regions, which concludes the proof.

## APPENDIX C

### THE PROBABILITY OF ERROR IS LINKED TO LIST SIZE

1) *Proof of Lemma 6:* Let us start by recalling:

$$\mathcal{L}_1(Y^n) \cap \mathcal{L}_2(Y^n) = \mathcal{L}^{(n)}. \quad (171)$$

Let  $(\hat{W}_0, \hat{W}_1)$  be the estimated messages at decoder 1, where

$$\begin{aligned} & \mathbb{P}\{(\hat{W}_0, \hat{W}_1) \neq (W_0, W_1)\} \\ &= \delta \mathbb{P}\{\exists(\hat{w}_0, \hat{w}_1) \neq (W_0, W_1) : (\hat{w}_0, \hat{w}_1) \in \mathcal{L}^{(n)} | (W_0, W_1) \in \mathcal{L}^{(n)}\} + \\ & (1 - \delta) \mathbb{P}\{\exists(\hat{w}_0, \hat{w}_1) \neq (W_0, W_1) : (\hat{w}_0, \hat{w}_1) \in \mathcal{L}^{(n)} | (W_0, W_1) \notin \mathcal{L}^{(n)}\} \end{aligned} \quad (172)$$

$$\leq \mathbb{P}\{\|\mathcal{L}^{(n)}\| > 1\} + (1 - \delta), \quad (173)$$

with  $(1 - \delta) \triangleq \mathbb{P}\{(W_0, W_1) \notin \mathcal{L}^{(n)}\}$ .

Then, following standard arguments, by the LLN and independence of codebooks, we can easily show that, for all  $\epsilon_1 > 0$ ,  $\exists N_1$  such that for  $n \geq N_1$ , we have  $(1 - \delta) \leq \epsilon_1$ .

This ends the proof of the statement:

$$P_e^{(n)} \leq \mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} + \epsilon_1. \quad (174)$$

2) *Proof of Lemma 7:* Let  $(w_0, w_1) \neq (W_0, W_1)$  be the supposedly decoded pair of messages.

We have, recalling (160), that:

$$\mathbb{P}\{(w_0, w_1) \in \mathcal{L}^{(n)}\} \leq \min_{j=1,2} \mathbb{P}\{(w_0, w_1) \in \mathcal{L}_j(Y^n)\} . \quad (175)$$

For the first list, we have, following similar arguments of Lemma 5, that:

$$\mathbb{P}\{(W_0, w_1) \in \mathcal{L}_1(Y^n)\} = \mathbb{P}\{(q^n(w_0), u^n(l_1, w_0), y^n) \in T_\delta^n(QUY) \text{ for } l_1 \in [1 : 2^{n(T_1-R_1)}]\} \quad (176)$$

$$\leq \sum_{l_1 \in [1 : 2^{n(T_1-R_1)}]} \mathbb{P}\{(q^n(w_0), u^n(l_1, w_0), y^n) \in T_\delta^n(QUY)\} \quad (177)$$

$$\leq \exp_2(n [T_1 - R_1 - I(U; Y|Q) + \epsilon_2]) , \quad (178)$$

and similarly, if moreover  $w_0 \neq W_0$ ,

$$\mathbb{P}\{(w_0, w_1) \in \mathcal{L}_1(Y^n)\} \leq \exp_2(n [T_1 - R_1 - I(QU; Y) + \epsilon_2]) . \quad (179)$$

Now, for the second list, i.e, decoding method, we know that:

1) If  $w_0 = W_0$ ,  $w_1 \neq W_1$  and  $l_2 = L_2$  which implies  $w_2 = W_2$ :

$$\begin{aligned} & \mathbb{P}\{(Q^n(W_0), U^n(l_1, W_0), V^n(L_2, W_0), Y^n) \in T_\delta^n(QUVY) \text{ for } l_1 \in [1 : 2^{n(T_1-R_1)}]\} \\ & \leq \exp_2(n [T_1 - R_1 + H(QUVY) - H(Q) - H(U|Q) - H(YV|Q) + \epsilon_3]) \end{aligned} \quad (180)$$

$$= \exp_2(n [T_1 - R_1 - I(U; YV|Q) + \epsilon_3]) , \quad (181)$$

where we used the fact that, since  $w_1 \neq W_1$ , then  $U^n(l_1, W_0)$  and  $V^n(L_2, W_0)$  are independent conditionally on  $Q^n(W_0)$ .

2) If  $w_0 = W_0$ ,  $w_1 \neq W_1$ , and  $l_2 \neq L_2$  then:

$$\begin{aligned} & \mathbb{P}\{(Q^n(W_0), U^n(l_1, W_0), V^n(l_2, W_0), Y^n) \in T_\delta^n(QUVY) \text{ for } l_1 \in [1 : 2^{n(T_1-R_1)}]\} \\ & \leq \exp_2(n [T_1 - R_1 + H(QUVY) - H(Q) - H(U|Q) - H(V|Q) - H(Y|Q) + \epsilon_3]) \end{aligned} \quad (182)$$

$$= \exp_2(n [T_1 - R_1 - I(UV; Y|Q) - I(U; V|Q) + \epsilon_3]) . \quad (183)$$

3) Finally, if  $w_0 \neq W_0$ , then whatever  $l_1$  and  $l_2$ :

$$\begin{aligned} & \mathbb{P}\{(Q^n(w_0), U^n(l_1, w_0), V^n(l_2, w_0), Y^n) \in T_\delta^n(QUVY)\} \\ & \leq \exp_2(n [H(QUVY) - H(Q) - H(U|Q) - H(V|Q) - H(Y) + \epsilon_3]) \end{aligned} \quad (184)$$

$$= \exp_2(n [-I(QUV; Y) - I(U; V|Q) + \epsilon_3]) . \quad (185)$$

This ends the proof of Lemma 2.

## APPENDIX D

## OUTER BOUND DERIVATION FOR THE COMPOUND BC

We need to recall that the proof in [4] of the outer bound for users' pair  $(k)$ , uses the specific choice of auxiliary RV:

$$\begin{cases} U_i = W_1, \\ V_i = W_2, \\ Q_i^{(k)} = (Y^{i-1,(k)}, Z_{i+1}^{n,(k)}) . \end{cases} \quad (186)$$

Here, we notice that the auxiliary RV  $(U_i, V_i)$  do not depend on the users' pair index. Thus, we can show that for all channel indices  $(k)$  with the specific choice:  $U_i = W_1, V_i = W_2$ ,

$$\mathcal{R}_{\text{NEG}}^{(k)}(p_{QUVX}) : \begin{cases} nR_1 \leq \sum_{i=1}^n I(Q_{k,i}U_i; Y_{k,i}) , \\ nR_2 \leq \sum_{i=1}^n I(Q_{k,i}V_i; Z_{k,i}) , \\ n(R_1 + R_2) \leq \sum_{i=1}^n [I(U_i; Y_{k,i}|Q_{k,i}V_i) + I(Q_{k,i}V_i; Z_{k,i})] , \\ n(R_1 + R_2) \leq \sum_{i=1}^n [I(Q_{k,i}U_i; Y_{k,i}) + I(V_i; Z_{k,i}|Q_{k,i}U_i)] , \end{cases} \quad (187)$$

where  $Q_{k,i} = (Y_{k,1}^{i-1}, Z_{k,i+1}^n)$ . Thus, we could possibly factor the resulting joint pmf on  $(U_i, V_i)$  over all compound channel indices, and let only the common variable choice vary from one channel to another. Moreover, we can show in the same fashion as in [4, Lemma 3.2], that the maximizing distribution of the input  $p_{X|QUV}$  is a deterministic mapping.

## APPENDIX E

## PROOF OF ACHIEVABILITY OF THE CAPACITY

From Theorem 2, we can see that the region  $\mathcal{R}_{\text{SNID}}$  verifies:

$$\mathcal{R}_{\text{SNID}} \supseteq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbf{T}(p)} \left( \mathcal{T}_3^{(1)}(p, T_1, T_2) \cap \mathcal{T}_4^{(2)}(p, T_1, T_2) \right) , \quad (188)$$

In this section, we evaluate the region thus obtained by:

$$\mathcal{R}_{\text{SNID}}^* \triangleq \bigcup_{p_{QUVX} \in \mathcal{P}} \bigcup_{(T_1, T_2) \in \mathbf{T}(p)} \left( \mathcal{T}_3^{(1)}(p, T_1, T_2) \cap \mathcal{T}_4^{(2)}(p, T_1, T_2) \right) , \quad (189)$$



where  $\mathcal{T}_3^{(1)} \cap \mathcal{T}_4^{(2)}$  is the subset of  $\mathfrak{R}_+^2$  defined by the inequalities:

$$\left\{ \begin{array}{l} T_2 \leq I(V; ZU|Q) \\ T_1 + T_2 \leq I(UV; Z|Q) + I(U; V|Q) \\ R_0 + T_1 + T_2 \leq I(QUV; Z) + I(U; V|Q) \\ T_1 \leq I(U; Y_2V|Q) \\ T_1 + T_2 \leq I(UV; Y_2|Q) + I(U; V|Q) \\ R_0 + T_1 + T_2 \leq I(QUV; Y_2) + I(U; V|Q) \\ T_1 \leq I(U; Y_1|Q) \\ R_0 + T_1 \leq I(QU; Y_1) \\ T_1 \geq R_1 \quad , \quad T_2 \geq R_2 \quad , \\ T_1 + T_2 > R_1 + R_2 + I(U; V|Q) \end{array} \right. \quad (190)$$

Recalling here that  $Y_1$  is physically degraded towards  $Z$ , we can first rewrite the decoding constraints as the following:

$$\left\{ \begin{array}{l} T_2 \leq I(V; ZU|Q) \\ T_1 \leq \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \\ R_0 + T_1 \leq I(QU; Y_1) \\ T_1 + T_2 \leq I(UV; Y_2|Q) + I(U; V|Q) \\ R_0 + T_1 + T_2 \leq I(QUV; Y_2) + I(U; V|Q) . \end{array} \right. \quad (191)$$

The, we can run FME over the binning rate pair  $(T_1, T_2)$  to get the following region:

$$\left\{ \begin{array}{l} R_2 \leq I(V; ZU|Q) \\ R_1 \leq \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \\ R_0 + R_1 \leq I(QU; Y_1) \\ R_1 + R_2 \leq I(UV; Y_2|Q) \\ R_1 + R_2 \leq I(V; Z|UQ) + \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \\ R_0 + R_1 + R_2 \leq I(QUV; Y_2) \\ R_0 + R_1 + R_2 \leq I(V; Z|UQ) + I(QU; Y_1) . \end{array} \right. \quad (192)$$

Later, we chose to apply *bit recombination* on the admissible rates  $(R_0, R_1, R_2)$  as follows:

$$\begin{cases} R_0 &= R_0^* + R_{01}^* + R_{02}^* , \\ R_1 &= R_1^* - R_{01}^* \geq 0 , \\ R_2 &= R_2^* - R_{02}^* \geq 0 , \\ R_{01}^* &\geq 0 \quad , \quad R_{02}^* \geq 0 . \end{cases} \quad (193)$$

It is straightforward that this bit recombination fits the decoding logic of the terminals, i.e., part of the private messages is mapped into the common message, enabling each terminal to still recover the totality of its intended message. The region writes thus as:

$$\begin{cases} R_2^* - R_{02}^* \leq I(V; ZU|Q) \\ R_1^* - R_{01}^* \leq \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \\ R_0^* + R_1^* + R_{02}^* \leq I(QU; Y_1) \\ R_1^* - R_{01}^* + R_2^* - R_{02}^* \leq I(UV; Y_2|Q) \\ R_1^* - R_{01}^* + R_2^* - R_{02}^* \leq I(V; Z|UQ) + \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \\ R_0^* + R_1^* + R_2^* \leq I(QUV; Y_2) \\ R_0^* + R_1^* + R_2^* \leq I(V; Z|UQ) + I(QU; Y_1) \\ R_1^* \geq R_{01}^* , \quad R_2^* \geq R_{02}^* \quad , \quad R_{01}^* \geq 0 , \quad R_{02}^* \geq 0 \end{cases} \quad (194)$$

Performing again FME over the splitting rate pair  $(R_{01}^*, R_{02}^*)$ , we get the following region:

$$R_0^* + R_1^* \leq I(QU; Y_1)$$

$$R_0^* + R_1^* + R_2^* \leq I(QU; Y_1) + I(UV; Y_2|Q)$$

$$R_0^* + R_1^* + R_2^* \leq I(QU; Y_1) + I(V; ZU|Q) \quad (195)$$

$$R_0^* + R_1^* + R_2^* \leq I(QU; Y_1) + I(V; Z|UQ) + \min\{I(U; Y_1|Q), I(U; Y_2V|Q)\} \quad (196)$$

$$R_0^* + R_1^* + R_2^* \leq I(QUV; Y_2)$$

$$R_0^* + R_1^* + R_2^* \leq I(U; Y|VQ) + I(QU; Y_1) . \quad (197)$$

We clearly notice that the constraints: (195) and (196) are implied by (197), thus, the resulting region  $\mathcal{R}_{\text{SNID}}^*$  is defined by the following constraints:

$$\begin{cases} R_0^* + R_1^* \leq I(QU; Y_1) \\ R_0^* + R_1^* + R_2^* \leq I(QU; Y_1) + I(UV; Y_2|Q) \\ R_0^* + R_1^* + R_2^* \leq I(QUV; Y_2) \\ R_0^* + R_1^* + R_2^* \leq I(V; Z|UQ) + I(QU; Y_1) . \end{cases} \quad (198)$$

Thus, letting  $R_0^* = 0$ , and noting the rate pairs as  $(R_1, R_2)$ , one gets the desired rate region.

## APPENDIX F

### CARDINALITY BOUNDS

Consider a pair of RVs  $(Q, X)$  following the joint probability distribution  $p_Q p_{X|Q}$ . Since the input is binary, let the four continuous functions on  $P_{X|Q}$ :

$$f_1(P_{X|Q}(0|q)) = P_{X|Q}(0|q) , \quad (199)$$

$$f_2(P_{X|Q}(0|q)) = H(Z|Q = q) = H_2(p \star P_{X|Q}(0|q)) , \quad (200)$$

$$f_3(P_{X|Q}(0|q)) = H(Y_1|Q = q) = H_2(p_1 \star P_{X|Q}(0|q)) , \quad (201)$$

$$f_4(P_{X|Q}(0|q)) = H(X|Q = q) = H_2(P_{X|Q}(0|q)) . \quad (202)$$

By the usual consequence of Fenchel-Eggleston-Caratheodory theorem [30], we can construct an auxiliary RV  $Q'$  such that:

$$\sum_q P_Q(q) P_{X|Q}(0|q) = \sum_{q'} P_{Q'}(q') P_{X|Q}(0|q') = P_X(0) , \quad (203)$$

$$H(Z|Q) = H(Z|Q') , \quad (204)$$

$$H(Y_1|Q) = H(Y_1|Q') , \quad (205)$$

$$H(X|Q) = H(X|Q') , \quad (206)$$

$$\|Q'\| \leq 4 . \quad (207)$$

Thus, we conclude that with this new auxiliary RV  $Q'$ , the region is unchanged:

$$I(X; Z|Q) = H(Z|Q) - H(Z|X) = H(Z|Q') - H_2(p) = I(X; Z|Q') , \quad (208)$$

$$I(Q; Y_1) = H(Y_1) - H(Y_1|Q) = H_2(p_1 \star P_X(0)) - H(Y_1|Q') = I(Q'; Y_1) , \quad (209)$$

$$I(Q; Y_2) = (1 - e) (H(X) - H(X|Q)) = \bar{e} (H_2(P_X(0)) - H(X|Q')) = I(Q'; Y_2) . \quad (210)$$

3) *Input uniformity*: In [6] lies a definition of the ‘‘c-symmetric broadcast channel’’ as being the BC formed by 2 c-symmetric channels. Following this same idea, and considering equivalently the Compound BC or the Compound Channel, we can say that the BC resulting from the simultaneity of two c-symmetric BC is c-symmetric.

As it is shown in [6, Lemma 2] that uniform input distribution is optimal for such a channel, we conclude that  $X \sim \text{Bern}(1/2)$  is optimal for the Compound BC.

APPENDIX G  
PROOF OF PROPOSITION 3

Recall that:

$$t_a(x) \triangleq \sup_{p_{QX} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] . \quad (211)$$

We want to show that: i)

1) For all  $x \in [0 : 1 - H_2(p)]$ ,

$$t_a(x) = \max_{p_{QX} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] \quad (212)$$

2)  $t_a$  is concave in  $x$ .

3)  $t_a$  can be described identically by its supporting lines.

4)  $t_a$  is decreasing in  $x$ .

*Proof:* i) We have that:

$$\mathcal{C}(x) = \left\{ p_{XQ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Q}) : Q \text{ --- } X \text{ --- } (Y, Z_1, Z_2), \quad (213)$$

$$X \sim \text{Bern}(1/2), I(X; Z|Q) = x \right\} . \quad (214)$$

Since, we have proved that the optimizing probabilities have a finite cardinality, the conditional mutual information being continuous,  $\mathcal{C}(x)$  is thus compact. As the probability space  $\mathcal{P}(\mathcal{X} \times \mathcal{Q})$  has a finite dimension, the set  $\mathcal{C}(x)$  is thus closed. Thus, the supremum is achieved.

ii) Concavity:

Let  $x_1, x_2 \in [0 : 1 - H_2(p)]$  and let  $\alpha \in [0 : 1]$ . Denote  $x = \alpha x_1 + (1 - \alpha) x_2$ . We need to show that:  $t_a(x) \geq \alpha t_a(x_1) + (1 - \alpha) t_a(x_2)$ .

Let for  $i \in \{1, 2\}$ ,

$$P_{X_i, Q_i} = \operatorname{argmax}_{p_{QX} \in \mathcal{C}(x_i)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] . \quad (215)$$

Define moreover:  $T \sim \text{Bern}(t)$  independent of all other RVs. Define

$$(X, Q_T) = \begin{cases} (X_1, Q_1) & \text{if } T = 0 , \\ (X_2, Q_2) & \text{if } T = 1 , \end{cases} \quad (216)$$

and by letting  $Q = (Q_T, T)$ , we have:

- $X \sim \text{Bern}(1/2)$ .
- $Q \text{ --- } X \text{ --- } (Y, Z_1, Z_2)$  is a valid Markov chain.

- And the following equalities hold:

$$I(X; Z|Q) = \alpha I(X_1; Z|Q_1) + (1 - \alpha) I(X_2; Z|Q_2) \quad (217)$$

$$= \alpha x_1 + (1 - \alpha) x_2 = x . \quad (218)$$

We thus have that:  $p_{XQ} \in \mathcal{C}(x)$ . Thus,

$$\begin{aligned} \alpha t_a(x_1) + (1 - \alpha) t_a(x_2) &= \alpha (a I(Q_1; Y_1) + \bar{a} I(Q_1; Y_2)) \\ &\quad + (1 - \alpha) (a I(Q_2; Y_1) + \bar{a} I(Q_2; Y_2)) \end{aligned} \quad (219)$$

$$= a I(Q_T; Y_1|T) + (1 - a) I(Q_T; Y_2|T) \quad (220)$$

$$\leq a I(TQ_T; Y_1) + (1 - a) I(TQ_T; Y_2) \quad (221)$$

$$= a I(Q; Y_1) + (1 - a) I(Q; Y_2) \quad (222)$$

$$\leq \max_{p_{QX} \in \mathcal{C}(x)} [a I(Q; Y_1) + (1 - a) I(Q; Y_2)] \quad (223)$$

$$= t_a(x) , \quad (224)$$

which concludes the proof of concavity.

iii) This property follows from the concavity of  $t_a$ .

iv) Monotony:

Since  $t_a$  is concave, we have that:

$$t'_a(x) \leq t'_a(0) = \lim_{x \rightarrow 0^+} \frac{t_a(x) - t_a(0)}{x} . \quad (225)$$

Since,

$$t_a(0) = a(1 - H_2(p_1)) + (1 - a)(1 - e_2) > t_a(x) , \quad (226)$$

for all  $x \in [0 : 1 - H_2(p)]$ , we have that:

$$t'_a(x) \leq t'_a(0) \leq 0 , \quad (227)$$

$t_a$  is thus decreasing in  $x$ . ■

## APPENDIX H

### PROOF OF ACHIEVABILITY OF MULTIPLE DESCRIPTION INNER BOUND

In this section, we establish the achievability of the MD inner bound (5). Let  $W_1$  be the message decoded by user 1, and let  $W_2$  be the message decoded by user 2, plus let  $R_1$  and  $R_2$

denote their respective rates. Let  $T_1$  and  $T_2$  denote the corresponding binning rates. We construct the following code.

Codebook generation:

Generate  $2^{nT_1}$   $n$ -sequences  $u_0^n(l_1)$ ,  $l_1 \in [1 : 2^{nT_1}]$  each following:

$$P_{U_0}^n(u_0^n(l_1)) = \prod_{i=1}^n P_{U_0}(u_{0,i}(l_1)) ,$$

and map all these sequences in  $2^{nR_1}$  bins, each indexed with  $w_1 \in [1 : 2^{nR_1}]$ :  $\mathcal{C}_0(w_1)$ .

Generate similarly  $2^{nT_2}$   $n$ -sequences  $v^n(l_2)$ ,  $l_2 \in [1 : 2^{nT_2}]$  each following  $P_V^n(v^n(l_2))$  and map them into  $2^{nR_2}$  bins:  $\mathcal{C}_v(w_2)$ .

For each  $u_0^n(l_1)$ ,  $l_1 \in [1 : 2^{nT_1}]$ , generate  $2^{n\hat{R}_j}$   $n$ -sequences  $u_j^n(s_j, l_1)$ ,  $s_j \in [1 : 2^{n\hat{R}_j}]$  following each:

$$P_{U_j|U_0}^n(u_j^n(s_j, l_1)) = \prod_{i=1}^n P_{U_j|U_0}(u_{j,i}(s_j, l_1)|u_{0,i}(l_1)) .$$

Encoding:

To send a message pair  $(W_1, W_2)$ , the encoder finds a couple of sequences  $u_0^n(l_1)$  and  $v^n(l_2)$  in the product bin  $\mathcal{C}_0(W_1) \times \mathcal{C}_v(W_2)$  and a couple of indices  $(s_1, s_2)$  such that

$$(u_0^n(l_1), u_1^n(s_1, l_1), u_1^n(s_2, l_1), v^n(l_2)) \in T_\delta^n(U_0 U_1 U_2 V) . \quad (228)$$

It then transmits an  $n$ -sequence  $x^n(u_0^n(l_1), u_1^n(s_1, l_1), u_1^n(s_2, l_1), v^n(l_2))$  which is generated via a random mapping.

Using the well known second order moment method, one can make the probability of the encoding error event arbitrarily close to 0 if:

$$\left\{ \begin{array}{l} T_1 - R_1 + \hat{R}_1 + \hat{R}_2 \geq I(U_1; U_2|U_0) , \\ T_1 - R_1 + T_2 - R_2 \geq I(U_0; V) , \\ T_1 - R_1 + \hat{R}_1 + T_2 - R_2 \geq I(U_0 U_1; V) , \\ T_1 - R_1 + \hat{R}_2 + T_2 - R_2 \geq I(U_0 U_2; V) , \\ T_1 - R_1 + \hat{R}_1 + \hat{R}_2 + T_2 - R_2 \geq I(U_0 U_1 U_2; V) + I(U_1; U_2|U_0) . \end{array} \right. \quad (229)$$

Decoding:

The second user, upon receiving the sequence  $z^n$ , looks for the unique index  $w_2$  such that for some  $v^n(l_2) \in \mathcal{C}_v(w_2)$ , the following holds:

$$(v^n(l_2), z^n) \in T_\delta^n(VZ) . \quad (230)$$

The probability of error in such a decoding rule is arbitrarily small provided that:

$$T_2 \leq I(V; Z) . \quad (231)$$

Concerning the two instances of the first user  $Y_1$  and  $Y_2$  let us assimilate each of them to a decoder. Decoder  $j$  finds the unique index  $l_1$  such that for some  $s_j$  where, the following joint typicality holds:

$$(u_0^n(l_1), u_j^n(s_j, l_1), y_j^n) \in T_\delta^n(U_0 U_j Y_j) . \quad (232)$$

The probability that the decoded  $l_1$  does not fall into the bin specified by  $w_1$  is made arbitrarily provided that:

$$T_1 + \hat{R}_j \leq I(U_0 U_j; Y_j) . \quad (233)$$

Then the overall decoding error events occur with arbitrary small probability provided that:

$$\left\{ \begin{array}{l} T_1 + \hat{R}_1 \leq I(U_0 U_1; Y_1) , \\ T_1 + \hat{R}_2 \leq I(U_0 U_2; Y_2) , \\ T_2 \leq I(V; Z) . \end{array} \right. \quad (234)$$

After running FME on the system of inequalities bearing in mind the natural encoding constraints:

$$\left\{ \begin{array}{l} \hat{R}_1 \geq 0 , \\ \hat{R}_2 \geq 0 , \\ T_1 \geq R_1 , \\ T_2 \geq R_2 , \end{array} \right. \quad (235)$$

the region given in (5) follows immediately.

## APPENDIX I

### PROOF OF LEMMA 1

We derive the optimal rate obtained when the following coding scheme is used:

$$\mathbf{X} = (X_u + X_p)\mathbf{B}_u + X_v\mathbf{B}_v , \quad U_0 = X_u + \alpha X_v , \quad (236)$$

$$V = X_v , \quad U_1 = X_p + \alpha_1 X_v , \quad (237)$$

where  $X_p \sim \mathcal{N}(0, x)$ ,  $X_u \sim \mathcal{N}(0, P_u - x)$  and  $X_v \sim \mathcal{N}(0, P_v)$  are pairwise independent RV and such that:  $P_u \leq P - P_v$ .

This means that we transmit two descriptions intended for user 1 making these two descriptions

compensate “jointly” the interference, hence, we are interested in computing the rate:  $R_{0,1} = I(U_0U_1; Y) - I(U_0U_1; V)$ . Some algebraic manipulations lead us to the following result:

$$R_{0,1} = \frac{1}{2} \log_2 \left( \frac{h_u^2 P_u + N}{\frac{P_v (h_u^2 P_u + N)}{h_u^2 P_u + h_v^2 P_v + N} P(\alpha, \alpha_1) + N} \right), \quad (238)$$

where the quadratic polynomial  $P(\alpha, \alpha_1)$  is given by:

$$P(\alpha, \alpha_1) = h_u^2 (\alpha_1 - \beta_1^x + \alpha - \beta^x)^2 + \frac{N}{x} (\alpha_1 - \beta_1^x)^2 + \frac{N}{P_u - x} (\alpha - \beta^x)^2, \quad (239)$$

and,  $\beta^x = \frac{(P_u - x) h_u h_v}{h_u^2 P_u + N}$  and  $\beta_1^x = \frac{x h_u h_v}{h_u^2 P_u + N}$ .

An interesting insight brought by this expression is that to achieve the optimal DoF, we need only have  $\alpha_1 + \alpha = \beta_1^o + \alpha^o$  rather than pairwise equality as might be suggested by the previous section. This translates perfectly the “joint” interference management of both decoded descriptions  $U_0$  and  $U_1$ , recovering trivially the optimal interference free rate as both descriptions cancel the interference fully each on their own  $\alpha_1 - \alpha_1^* = \alpha_0 - \alpha_0^* = 0$ .

Upon optimizing the polynomial  $P(\alpha, \alpha_1)$  over  $\alpha_1$ , the resulting rate is given by the rather simple expression:

$$R_{0,1} = \frac{1}{2} \log_2 \left( \frac{h_u^2 P_u + N}{\frac{P_v}{(P_u - x)} \frac{(h_u^2 P_u + N)^2}{(h_u^2 P_u + h_v^2 P_v + N)} \frac{N}{h_u^2 x + N} (\alpha - \beta^x)^2 + N} \right), \quad (240)$$

It can be readily checked that this expression corresponds to the following formulation of the rate:

$$R_{0,1} = I(U_0; Y) - I(U_0; V) + I(X_p; Y | X_u X_v), \quad (241)$$

where

$$I(X_p; Y | X_u X_v) = \frac{1}{2} \log_2 \left( \frac{h_u^2 x + N}{N} \right), \quad (242)$$

and where  $I(U_0; Y) - I(U_0; V)$  corresponds to the case where  $X_u$  dirty-paper codes  $X_v$  under the noise component variance:  $h_u^2 x + N$ .

This means that the optimal choice of the variable  $U_1$  is the one that maximizes the DPC term  $I(U_1; Y | U_0) - I(U_1; V | U_0)$ .



## APPENDIX J

## OPTIMIZATION OF COMMON DESCRIPTION INNER BOUND:

Let us first optimize the second corner point of the CD inner bound. We have that

$$\mathcal{R}_2 = \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \quad R_2 \leq \frac{1}{2} \log_2 \left( \frac{g_v^2 P_v + N}{N} \right), \right. \\ \left. R_1 \leq \min_{j=1,2} \frac{1}{2} \log_2 \left( \frac{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}{h_{j,v}^2 P_v + N} \right) \right\}. \quad (243)$$

We have that what maximizes  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are orthogonal and of unit norm, thus, we can write that:  $h_{1,u}^2 = 1 - h_{2,u}^2$  and  $h_{1,v}^2 = 1 - h_{2,v}^2$ . The rate  $R_2$  does not depend on the beam  $\mathbf{B}_u$ , thus, we start by optimizing the rate  $R_1$  over it. The two min operands are monotonic in inverse directions and have the same minimum value 0, thus, the maxmin point corresponds to the equality point.

Which by simple algebraic calculations leads to the condition:

$$h_{1,u}^2 = \frac{h_{1,v}^2 P_v + N}{P_v + 2N}, \quad (244)$$

and yields then a rate (independent of the beam  $\mathbf{B}_v$ ) equal to:

$$R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + P_v + 2N}{P_v + 2N} \right). \quad (245)$$

Note then that the maximizing beam direction  $\mathbf{B}_v = \mathbf{g}$ , thus one can easily check that this verifies:  $h_{1,v} = -1/\sqrt{2}$  and thus, from (245), that  $|h_{1,u}| = 1/\sqrt{2}$ . Thus transmitting the first user's signal in the mean channel direction is an admissible optimizing solution. Later in the proof, we show that this corner point is dominated by the first corner point of the CD inner bound. In the sequel, we will perform the optimization under the choice of  $h_{1,u} = 1/\sqrt{2}$  and  $g_u = 0$ , i.e., we transmit the signal intended to user 1 in the mean channel direction, which makes it orthogonal to the second user's channel; the optimality of which is given in Appendix K.

We can rewrite the first corner point of the CD inner bound as follows:

$$\mathcal{R}_1 = \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \right. \\ R_1 \leq \max_{\alpha \in \mathbb{R}} \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{P_v}{P_u} \frac{(P_u + 2N)^2}{P_u P_u + 2N + 2h_{j,v}^2 P_v} (\alpha - \alpha_j)^2 + 2N} \right) \\ \left. R_2 \leq \frac{1}{2} \log_2 \left( \frac{g_v^2 P_v + N}{N} \right) \right\} \quad (246)$$

where  $\alpha_j = \frac{\sqrt{2}P_u}{P_u + 2N}h_{j,v}$ . Since  $\|\mathbf{h}_j\| = \|\mathbf{B}_v\| = 1$  and,  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are orthogonal, we can let  $h_{1,v} = \cos(\theta_v)$  and  $h_{2,v} = \sin(\theta_v)$ .

The key point in the optimization is to solve the equation:

$$\frac{(\alpha - \alpha_1)^2}{P_u + 2N + 2 \cos(\theta_v)^2 P_v} = \frac{(\alpha - \alpha_2)^2}{P_u + 2N + 2 \sin(\theta_v)^2 P_v}. \quad (247)$$

The optimization of the rate of the first user  $R_1$  yields the following:

(i) If  $\cos^2(\theta_v) = \frac{1}{2}$  and  $\cos(\theta_v) = -\sin(\theta_v)$ , then the optimal rate is given by:

$$R_1 \leq \max_{\alpha} \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{P_v (P_u + 2N)^2}{P_u (P + 2N)} (\alpha - \alpha_j)^2 + 2N} \right) \quad (248)$$

$$= \max_{\alpha} \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{P_v (P_u + 2N)^2}{P_u (P + 2N)} (|\alpha| + |\alpha_j|)^2 + 2N} \right) \quad (249)$$

$$= \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N + P_v \frac{P_u}{P + 2N}} \right) \quad (250)$$

$$= \frac{1}{2} \log_2 \left( \frac{P_u + P_v + 2N}{P_v + 2N} \right) \quad (251)$$

where  $\alpha_1 = -\alpha_2 = \frac{P_u}{P_u + 2N}$ . It turns out then, that the optimization over the DPC parameter  $\alpha$  yields  $\alpha = 0$ , i.e. the dilemma at the transmitter is so strong that the optimal choice is not to cancel interference and send in a direction that does not yield privilege to any of the channel instances  $\mathbf{h}$ . A very important remark, is that this yields exactly the first corner point of the region.

(ii) If  $\cos^2(\theta_v) = \frac{1}{2}$  and  $\cos(\theta_v) = \sin(\theta_v)$ , then the optimal rate is given by:

$$R_1 \leq \frac{1}{2} \max_{\alpha} \min_{j \in \{1,2\}} \log_2 \left( \frac{P_u + 2N}{\frac{P_v (P_u + 2N)^2}{P_u (P + 2N)} (\alpha - \alpha_j)^2 + 2N} \right) \quad (252)$$

$$= \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right), \quad (253)$$

which corresponds to the point where  $h_{1,v} = h_{2,v}$  i.e.  $\alpha_1 = \alpha_2$ . Thus, we would have  $\mathbf{h}_1 - \mathbf{h}_2$  orthogonal to  $\mathbf{B}_v$ , but since  $\mathbf{h}_1 - \mathbf{h}_2$  is collinear to the second user's channel, then it means that no information is transmitted to it with the beam  $\mathbf{B}_v$ . The power optimization of this point corresponds to the corner point  $(C_1, 0)$ .

(iii) If  $\cos^2(\theta_v) \neq \frac{1}{2}$ , then there are two optimizing solutions  $\alpha_1^*$  and  $\alpha_2^*$  such that:

$$\alpha_1^* - \alpha_1 = \frac{P_u}{P_u + 2N} \frac{(-\cos(\theta_v) + \sin(\theta_v)) \sqrt{P_u/2 + N + \cos(\theta_v)^2 P_v}}{\sqrt{P_u + 2N + 2 \sin^2(\theta_v) P_v} + \sqrt{P_u + 2N + 2 \cos^2(\theta_v) P_v}}, \quad (254)$$

$$\alpha_2^* - \alpha_1 = \frac{P_u}{P_u + 2N} \frac{(\cos(\theta_v) - \sin(\theta_v)) \sqrt{P_u/2 + N + \cos(\theta_v)^2 P_v}}{\sqrt{P_u + 2N + 2 \sin^2(\theta_v) P_v} - \sqrt{P_u + 2N + 2 \cos^2(\theta_v) P_v}}. \quad (255)$$

The root that yields the greater rate is  $\alpha_1^*$ . Then, we can rewrite with the following transformation  $y = \sin(2\theta_v)$  that:

$$(\alpha_1^* - \alpha_1)^2 = \frac{2P_u^2}{(P_u + 2N)^2} \frac{\cos^2(\theta_v + \pi/4) (P_u/2 + N + \cos(\theta_v)^2 P_v)}{\left(\sqrt{P_u + 2N + 2 \sin^2(\theta_v) P_v} + \sqrt{P_u + 2N + 2 \cos^2(\theta_v) P_v}\right)^2} \quad (256)$$

$$= \frac{P_u^2}{2(P_u + 2N)^2} \frac{(1-y) (P_u/2 + N + \cos(\theta_v)^2 P_v)}{P + 2N + \sqrt{(P_u + 2N)(P + 2N + P_v) + y^2 P_v^2}} \quad (257)$$

$$= \frac{P_u^2}{2(P_u + 2N)^2} \frac{(1-y) (P_u/2 + N + \cos(\theta_v)^2 P_v)}{P + 2N + \sqrt{(P + 2N)^2 + (y^2 - 1)P_v^2}}. \quad (258)$$

Note that the value of  $y = -1$ , i.e.,  $\theta_v = -\pi/4$ , is included in this expression. Thus we drop the case distinctions  $\cos^2(\theta_v) = 1/2$  and  $\cos^2(\theta_v) \neq 1/2$ .

As a conclusion, CD inner bound writes as:

$$\mathcal{R}_{CD} = \bigcup_{y \in [-1:1]} \left\{ (R_1, R_2) \in \mathfrak{R}^2, \right. \\ \left. \begin{aligned} R_1 &\leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{\frac{P_v P_u (1-y)}{P + 2N + \sqrt{(P + 2N)^2 + (y^2 - 1)P_v^2}} + 2N} \right) \\ R_2 &\leq \frac{1}{2} \log_2 \left( \frac{(1-y)P_v + 2N}{2N} \right) \end{aligned} \right\}. \quad (259)$$

## APPENDIX K

### BEAMFORMING OPTIMIZATION FOR THE CD-DPC INNER BOUND

In this section, we show, that with a strong uncertainty over the channel instances of user 1, i.e.,  $\mathbf{h}_1$  and  $\mathbf{h}_2$  being orthogonal, when resorting to a CD-DPC, the source has no choice but to

send over the mean channel  $\mathbf{h}_{1,2}$ . The proof of this claim is quite evolved and requires the use of many analytical manipulations when solving optimization problems.

Let us use the following notations. We previously introduced  $\theta_v$  such that  $h_{1,v} = \cos(\theta_v)$  and  $h_{2,v} = \sin(\theta_v)$ . Since  $\|\mathbf{h}_j\| = \|\mathbf{B}_u\| = 1$  and,  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are orthogonal, we can similarly define  $\theta_u$  such that  $h_{1,u} = \cos(\theta_u)$  and  $h_{2,u} = \sin(\theta_u)$ . Let us define:

$$s_u \triangleq \frac{\sin(2\theta_u)}{|\sin(2\theta_u)|} \text{ and } s_v \triangleq \frac{\sin(2\theta_v)}{|\sin(2\theta_v)|}, \quad (260)$$

when  $\sin(2\theta_u) \neq 0$  and  $\sin(2\theta_v) \neq 0$ .

In this section, we prove that it is optimal to transmit the signal in the directions given by:  $\mathcal{B}_u = \mathbf{h}_{1,2}$ , which comes to having  $h_{1,u} = h_{1,v} = \frac{1}{\sqrt{2}}$ .

Thus, we need to solve the optimization problem given by:

$$\bigcup_{\mathcal{B}_u, \mathcal{B}_v} \begin{cases} R_1 \leq \max_{\alpha} \min_{j=1,2} \log_2 \left( \frac{1}{A_j(\alpha - \alpha_j)^2 + c_j} \right), \\ R_2 \leq \log_2 \left( 1 + \frac{g_v^2}{g_u^2 P_u + N} \right), \end{cases} \quad (261)$$

where:

$$A_j \triangleq \frac{P_v}{P_u} \frac{h_{j,u}^2 P_u + N}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}, \quad (262)$$

$$c_j \triangleq \frac{N}{h_{j,u}^2 P_u + N}, \quad (263)$$

$$\text{and } \alpha_j \triangleq P_u \frac{h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N}. \quad (264)$$

This, in part, requires solving the following optimization problem:

$$\max_{\alpha} \min_{j=1,2} \log_2 \left( \frac{1}{A_j \alpha^2 - 2B_j \alpha + D_j} \right), \quad (265)$$

where:

$$B_j \triangleq P_v \frac{h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}, \quad (266)$$

$$\text{and } D_j \triangleq \frac{h_{j,v}^2 P_v + N}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (267)$$

Finding the optimal DPC parameter  $\alpha$  to use requires solving the equation:

$$(A_1 - A_2)\alpha^2 - 2(B_1 - B_2)\alpha + (D_1 - D_2) = 0, \quad (268)$$

which yields:

- a) If  $A_1 = A_2$  and  $B_1 = B_2$  while  $D_1 \neq D_2$ , no solution exists,
- b) If  $A_1 = A_2$  and  $B_1 = B_2$  and  $D_1 = D_2$ , every  $\alpha$  is a solution,
- c) If  $A_1 = A_2$  and  $B_1 \neq B_2$ , then there exists only one solution:  $\alpha_{opt} = \frac{D_1 - D_2}{2(B_1 - B_2)}$ ,
- d) If  $A_1 \neq A_2$  and  $(B_1 - B_2)^2 = (A_1 - A_2)(D_1 - D_2)$ , then there exists only one solution:  
 $\alpha_{opt} = \frac{B_1 - B_2}{A_1 - A_2}$ ,
- e) If  $A_1 \neq A_2$  and  $(B_1 - B_2)^2 < (A_1 - A_2)(D_1 - D_2)$  no solution exists,
- f)  $A_1 \neq A_2$  and  $(B_1 - B_2)^2 > (A_1 - A_2)(D_1 - D_2)$  then there exist two solutions.

Next, we can deduce the optimal values of (265) to be used by the source as given in table II. Our aim is to show that, over all these cases, the optimal beamforming strategy is to let  $\mathbf{B}_u = \mathbf{h}_{1,2}$ .

#### A. Case of $A_1 = A_2$ and $B_1 = B_2$

We show in the following that if  $A_1 = A_2$  and  $B_1 = B_2$ , then it follows that  $D_1 = D_2$  and that  $h_{j,u} = -h_{j,v} = \frac{1}{\sqrt{2}}$ . Thus letting  $\mathbf{B}_u = \mathbf{h}_{1,2}$  would yield the optimal solution given by:

$$\min_{\alpha} \max_{j=1,2} A_j (\alpha - \alpha_j)^2 + c_j = c_1 = c_2 = \frac{2N}{P_u + 2N} . \quad (269)$$

Hereafter, the details of the proof.

First, note that since  $A_1 = A_2$ , then we can write that:

$$h_{j,v}^2 = \frac{h_{j,v}^2 P_u + N}{P_u + 2N} . \quad (270)$$

As such, we can say that:  $h_{1,v}h_{2,v} \neq 0$ . More over, all quantities write then as:

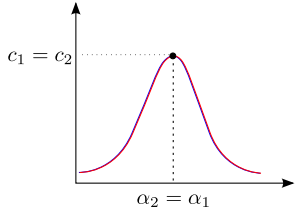
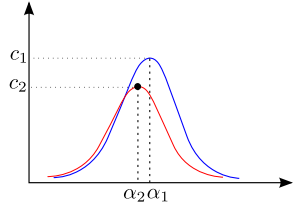
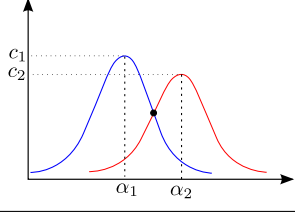
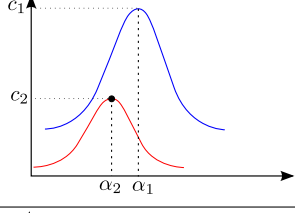
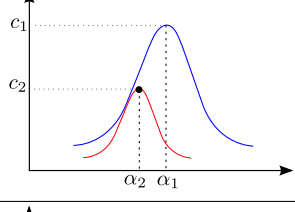
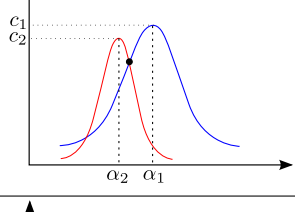
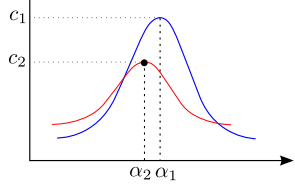
$$A_1 = A_2 = A \triangleq \frac{P_v P_u + 2N}{P_u P + 2N} , \quad (271)$$

$$\alpha_j = \frac{P_u}{P_u + 2N} \frac{h_{j,u}}{h_{j,v}} , \quad (272)$$

$$B_j = P_v \frac{h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N} = \frac{P_v}{P + 2N} \frac{h_{j,u}}{h_{j,v}} , \quad (273)$$

$$D_j = \frac{h_{j,v}^2 P_v + N}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N} = \frac{h_{j,v}^2 P_v + N}{h_{j,v}^2 (P + 2N)} . \quad (274)$$

TABLE II  
OPTIMAL DPC PARAMETER FOR THE CD-DPC

$A_1 = A_2$	$B_1 = B_2$		$\alpha_1$
	$B_1 \neq B_2$ $ c_1 - c_2  > A(\alpha_1 - \alpha_2)^2$		$\alpha_2$
	$B_1 \neq B_2$ $ c_1 - c_2  \leq A(\alpha_1 - \alpha_2)^2$		$\frac{D_1 - D_2}{2(B_1 - B_2)}$
$A_1 \neq A_2$	$\Delta < 0$		$\alpha_2$
	$\Delta = 0$		$\alpha_2$
	$\Delta > 0$ $ c_1 - c_2  \leq \min(A_1, A_2)(\alpha_1 - \alpha_2)^2$		$\frac{\alpha_1 + A_2 \alpha_1 - \alpha_2  - \sqrt{\Delta}}{A_1 - A_2}$
	$\Delta > 0$ $ c_1 - c_2  > \min(A_1, A_2)(\alpha_1 - \alpha_2)^2$		$\alpha_2$

Now, we have that:

$$B_1 = B_2 \Leftrightarrow P_v \left( \frac{h_{1,u}}{h_{1,v}} - \frac{h_{2,u}}{h_{2,v}} \right) = 0 , \quad (275)$$

$$\Leftrightarrow P_v = 0 \quad \text{or} \quad \cos(\theta_u) \sin(\theta_v) - \cos(\theta_v) \sin(\theta_u) = 0 , \quad (276)$$

$$\Leftrightarrow P_v = 0 \quad \text{or} \quad \sin(\theta_u - \theta_v) = 0 , \quad (277)$$

$$\Leftrightarrow P_v = 0 \quad \text{or} \quad \theta_u = \theta_v[\pi] , \quad (278)$$

but since  $\cos^2(\theta_v) = \frac{\cos^2(\theta_u)P_u + N}{P_u + 2N}$ , then, one can write that:

$$B_1 = B_2 \text{ and } A_1 = A_2 \quad \Rightarrow \quad P_v = 0 \text{ or } \cos^2(\theta_u) = \frac{1}{2} . \quad (279)$$

This implies then that:

$$c_1 = c_2 = \frac{P_v + 2N}{P + 2N} , \quad (280)$$

$$\alpha_1 = \alpha_2 = \pm \frac{P_u}{P_u + 2N} . \quad (281)$$

Thus in both cases of  $P_v = 0$  and  $P_v \neq 0$ , the optimal solution is given by:

$$\min_{\alpha} \max_{j=1,2} A_j (\alpha - \alpha_j)^2 + c_j = c_1 = c_2 = \frac{2N}{P_u + 2N} . \quad (282)$$

Note that, since  $\theta_u = \theta_v[\pi]$ , then  $2\theta_u = 2\theta_v[2\pi]$ , thus  $s_u = s_v$ .

Thus, as for the rate of user 2, two cases unfold:

- Case of  $s_u = s_v = 1$ , which corresponds to  $B_u = \mathbf{h}_{1,2}$ , and in this case  $B_v$  is co-linear to  $\mathbf{h}_{1,2}$  and thus orthogonal to user 2's channel  $\mathbf{g}$  leading to a zero achievable rate:

$$R_2 = 0 . \quad (283)$$

The power optimization of this point will yields the single capacity point  $(C_1, 0)$ .

- Case of  $s_u = s_v = -1$ , which corresponds to  $B_u \perp \mathbf{h}_{1,2}$ , and in this case  $B_v$  is co-linear to user 2's channel  $\mathbf{g}$  leading to the achievability of all rate pairs satisfying:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right) , \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{P + N}{P_u + N} \right) . \end{cases} \quad (284)$$

The set of rate pairs obtained can be shown to perform worse than time sharing as is explained hereafter. To show this, let  $\alpha \in [0 : 1]$  such that:

$$R_1 = \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right) = \frac{\alpha}{2} \log_2 \left( \frac{P + 2N}{2N} \right) . \quad (285)$$

We need to show that:

$$R_2 = \frac{1}{2} \log_2 \left( \frac{P+N}{P_u+N} \right) \leq \frac{(1-\alpha)}{2} \log_2 \left( \frac{P+N}{N} \right). \quad (286)$$

To see this, note that:

$$\frac{P_u+2N}{2N} = \frac{(P+2N)^\alpha}{(2N)^\alpha} \Rightarrow \frac{P_u}{N} + 1 = 2 \left( \frac{P}{2N} + 1 \right)^\alpha - 1 \quad (287)$$

$$\Rightarrow R_2 = \frac{1}{2} \log_2 \left( \frac{\frac{P}{N} + 1}{2 \left( \frac{P}{2N} + 1 \right)^\alpha - 1} \right) \quad (288)$$

$$\stackrel{(a)}{\Rightarrow} R_2 \leq \frac{1}{2} (1-\alpha) \log_2 \left( \frac{P+N}{N} \right), \quad (289)$$

where (a) is a result of that the function:

$$\begin{aligned} [1 : \infty[ &\mapsto \mathbb{R} \\ x &\mapsto 2(x+1)^\alpha - 1 - (2x+1)^\alpha \end{aligned} \quad (290)$$

by a quick function study, is positive. And thus:

$$2 \left( \frac{P}{2N} + 1 \right)^\alpha - 1 \geq \left( \frac{P}{N} + 1 \right)^\alpha, \quad (291)$$

which proves our claim.

To end the discussion of this case, it turns out that the optimal points obtained are the two single capacity points  $(C_1, 0)$  and  $(0, C_2)$ .

*B. Case of  $A_1 = A_2 = A$  and  $B_1 \neq B_2$  and  $|c_1 - c_2| \leq A(\alpha_1 - \alpha_2)^2$*

In this case, the optimal solution of the problem (268) is given by:

$$\alpha_{opt} = \frac{D_2 - D_1}{2(B_2 - B_1)} = \frac{c_2 - c_1}{2(\alpha_2 - \alpha_1)} + \frac{1}{2}(\alpha_2 + \alpha_1). \quad (292)$$

Thus, the minimum value of the function to optimize in (261) is given by:

$$F_{opt} \triangleq \frac{(c_2 - c_1)^2}{4A(\alpha_2 - \alpha_1)^2} + \frac{1}{2}(c_2 + c_1) + \frac{A}{4}(\alpha_2 - \alpha_1)^2, \quad (293)$$

where as for previously:

$$A = \frac{P_v P_u + 2N}{P_u P + 2N}, \quad (294)$$

$$c_j = \frac{N}{P_u + 2N} \frac{1}{h_{j,v}^2}, \quad (295)$$

$$\alpha_j = \frac{P_u}{P_u + 2N} \frac{h_{j,u}}{h_{j,v}}. \quad (296)$$



After some analytic manipulations we end up with the following expression of the optimal solution:

$$F_{opt} = \frac{1}{(P_u + 2N) \sin^2(2\theta_v)} \left[ \frac{N^2(P + 2N)}{P_u P_v} \frac{\cos^2(2\theta_v)}{\sin^2(\theta_u - \theta_v)} + \frac{P_u P_v}{P + 2N} \sin^2(\theta_u - \theta_v) + 2N \right], \quad (297)$$

Now, using the fact that:

$$\cos^2(\theta_v) = \frac{\cos^2(\theta_u) P_u + N}{P_u + 2N}, \quad (298)$$

we can write that:

$$\cos(2\theta_v) = \frac{P_u}{P_u + 2N} \cos(2\theta_u), \quad (299)$$

which implies

$$\sin(2\theta_u) = s_u \sqrt{1 - \cos^2(2\theta_u)} = s_u \sqrt{1 - \frac{(P_u + 2N)^2}{P_u^2} \cos^2(2\theta_v)}, \quad (300)$$

where we recall that:

$$s_u = \frac{\sin(2\theta_u)}{|\sin(2\theta_u)|} \quad \text{and} \quad s_v = \frac{\sin(2\theta_v)}{|\sin(2\theta_v)|}. \quad (301)$$

In the sequel, we define the two variables:

$$x \triangleq \cos^2(2\theta_v), \quad (302)$$

$$a \triangleq \frac{(P_u + 2N)^2}{P_u^2}. \quad (303)$$

As defined, and recalling (299), we can conclude that  $x$  lies in the set  $\left[0 : \frac{1}{a}\right]$ . To further simplify (297), we need to express the following quantity:

$$\sin^2(\theta_u - \theta_v) = \frac{1}{2P_u} \left[ P_u \left( 1 - x - s_u s_v \sqrt{(1-x) \left( 1 - \frac{(P_u + 2N)^2}{P_u^2} x \right)} \right) - 2Nx \right]. \quad (304)$$

Letting then:

$$\begin{aligned} g(x, s_u, s_v) &\triangleq 2P_u \sin^2(\theta_u - \theta_v) \\ &= P_u \left( 1 - x - s_u s_v \sqrt{(1-x) \left( 1 - \frac{(P_u + 2N)^2}{P_u^2} x \right)} \right) - 2Nx, \end{aligned} \quad (305)$$

one ends up with the following optimal function expressed in  $x$ ,  $s_u$  and  $s_v$ :

$$F_{opt}(x) = \frac{1}{(1-x)} \left[ \frac{2N^2(P + 2N)}{P_v} \frac{x}{g(x, s_u, s_v)} + \frac{P_v}{2(P + 2N)} g(x, s_u, s_v) + 2N \right]. \quad (306)$$

Now, the rate of the second user is given by:

$$R_2 = \log_2 \frac{1}{2} \left( 1 + \frac{g_v^2 P_v}{g_u^2 P_u + 2N} \right). \quad (307)$$

Then, we express:

$$g_v^2 = \cos^2 \left( \theta_v + \frac{\pi}{4} \right) = \frac{1 - \sin(2\theta_v)}{2} \quad (308)$$

$$= \frac{1 - s_v \sqrt{1-x}}{2}. \quad (309)$$

Similarly, we can show that:

$$g_u^2 = \frac{1}{2} - \frac{s_u}{2} \sqrt{1 - \frac{(P_u + 2N)^2}{P_u^2} x}, \quad (310)$$

Thus, the overall optimization problem is given as:

$$\bigcup_{(x, s_u, s_v) \in \mathcal{S}} \left\{ \begin{array}{l} R_1 \leq \frac{1}{2} \log_2 \left( \frac{1}{F_{opt}(x, s_u, s_v)} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_v(1 - s_v \sqrt{1-x})}{P_u(1 - s_u \sqrt{1-ax}) + N} \right) \end{array} \right\}, \quad (311)$$

where we define the optimization  $\mathcal{S}$  as:

$$\mathcal{S} \triangleq \left\{ x \in \left[ 0 : \frac{1}{a} \right], (s_u, s_v) \in \{-1, 1\}^2, \text{ s.t. } s_u s_v = 1 \Rightarrow x \neq 0 \right\}, \quad (312)$$

Hereafter, we study two distinct cases:  $s_u s_v = -1$  and  $s_u s_v = 1$ .

1)  $s_u s_v = +1$  :

we show in the following section that this case is impossible because  $s_u s_v = +1$  contradicts the existence of  $x$  such that

$$|c_1 - c_2| \leq A(\alpha_1 - \alpha_2)^2,$$

2)  $s_u s_v = -1$  :

When  $s_u s_v = -1$ , we express the first derivative of the function  $F_{opt}$  and show that it is always positive, leading us to the claim that  $F_{opt}$  is strictly increasing. Thus, the rate of user 1,  $R_1$  is decreasing in  $x$ .

If  $s_u = 1$  and  $s_v = -1$ , then  $R_2$  is easily shown to be decreasing in  $x$ , and thus, the optimal rate pair that is achievable is given by  $x = 0$  leading thus to:

$$\left\{ \begin{array}{l} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P + 2N}{P_v + 2N} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{P_v + N}{N} \right). \end{array} \right. \quad (313)$$

If  $s_u = 1$  and  $s_v = -1$ , then we can show that  $R_2$  can not be greater than:  $\log_2 \left( \frac{P_v + N}{N} \right)$ , and thus, the achievable rate region is dominated by (313).

To see this, note that:  $R_2$  is strictly increasing in  $x$  and thus, its maximum value is attained for  $x = \frac{1}{a}$ . Then, one can easily check that:

$$R_2 = \frac{1}{2} \log_2 \left( 1 + P_v \frac{1 - \sqrt{\frac{4N(P_u+N)}{P_u^2}}}{P_u + 2N} \right) \quad (314)$$

$$\leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{1}{P_u + 2N} \right) \quad (315)$$

$$\leq \frac{1}{2} \log_2 \left( 1 + \frac{P_v}{N} \right) , \quad (316)$$

which proves our claim.

Thus, the overall obtained rate region for this case, does not outperform the second corner point of CD-DPC, which is already included in the first corner point.

*C. Case of  $A_1 = A_2 = A$  and  $B_1 \neq B_2$  and  $|c_1 - c_2| > A(\alpha_1 - \alpha_2)^2$*

In this case, we show that the obtained rate region does not outperform time sharing.

To this end, we start by expressing the following:

$$|c_1 - c_2| > A(\alpha_1 - \alpha_2)^2 \quad \Leftrightarrow \quad N|\cos(2\theta_v)| > \frac{P_v P_u}{P + 2N} \sin^2(\theta_u - \theta_v) . \quad (317)$$

With the previous notations of the function  $g$  and  $x = \cos^2(2\theta_v)$ , we can rewrite this condition as:

$$N\sqrt{x} > \frac{P_v}{2(P + 2N)} g(x) . \quad (318)$$

If  $s_u s_v = 1$ , then we can show easily that the above condition is always verified even with  $P_v \neq 0$  and  $x \neq 0$ . To see this note that for  $x \in [0 : 1/a]$ :

$$g(x) = P_u(1 - x - \sqrt{(1-x)(1-ax)}) - 2Nx \leq P_u(1 - x - 1 + ax) - 2Nx \triangleq h(x) . \quad (319)$$

Then, it is easy to show that for  $x \in [0 : 1/a]$ :

$$N\sqrt{x} \geq h(x) , \quad (320)$$

since  $h$  is linear and  $h(0) = 0$  and  $h(1/a) = g(1/a) < N\sqrt{1/a}$ . Fig. 7 illustrates clearly our claim.

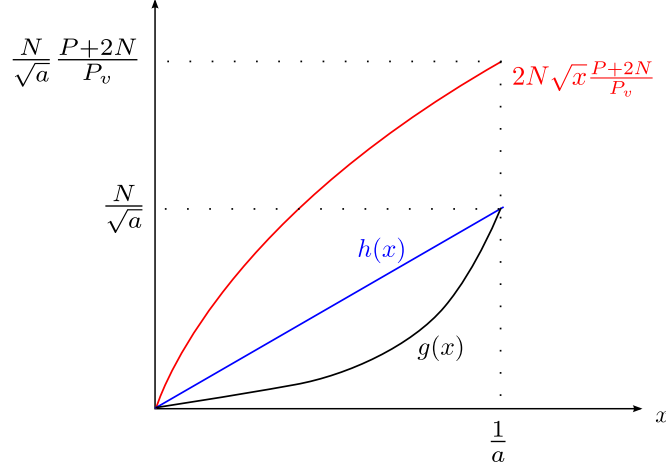


Fig. 7. Comparison of the functions  $h$ ,  $g$  and target upper bound.

Thus, since the condition is always verified, the optimal solution is given by the rate pairs  $(R_1, R_2)$  satisfying:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} (1 - \sqrt{x}) \right), \\ R_2 &\leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{1 - s_v \sqrt{1-x}}{P_u (1 - s_u \sqrt{1-ax}) + 2N} \right). \end{aligned} \quad (321)$$

If  $s_u = s_v = -1$ , then we show that the obtained rate region is included in the time sharing rate region. To this end, we choose to show this claim on a larger rate region given by:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} (1 - x) \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{1 - s_v \sqrt{1-x}}{P_u (1 - s_u \sqrt{1-ax}) + 2N} \right). \end{cases} \quad (322)$$

We proceed as follows to show that the obtained rate region for fixed  $P_u, P_v$  and  $N$ , is concave.

Let  $\alpha \in [0 : 1]$ , such that:

$$\frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} (1 - x) \right) \triangleq \alpha \log_2 \left( \frac{P_u + 2N}{2N} \right) + (1-\alpha) \log_2 \left( \frac{P_u + 2N}{2N} \left( 1 - \frac{1}{a} \right) \right). \quad (323)$$

Thus, we can show that:

$$1 - x = \left( 1 - \frac{1}{a} \right)^{1-\alpha}. \quad (324)$$

Thus, letting  $y \triangleq \left( 1 - \frac{1}{a} \right)$ , the previous rate of user 2 writes as:

$$R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \sqrt{1-y} \frac{1 + \sqrt{y^{1-\alpha}}}{P_u \left( 1 + \sqrt{y^{1-\alpha} - y} \right)} \right) \triangleq \frac{1}{2} \log_2 (f(\alpha)). \quad (325)$$

Our aim is to show that  $R_2$  is convex in  $\alpha$ , thus, we need to show that:

$$f''(\alpha)f(\alpha) - (f'(\alpha))^2 \geq 0 . \quad (326)$$

We have that:

$$f'(\alpha) = \frac{P_v \log(y) \sqrt{1-y}}{2P_u \left(1 + \sqrt{y^{1-\alpha} - y}\right)^2} \left( -\sqrt{y^{1-\alpha}} + \frac{\sqrt{y^{1-\alpha} + y}}{\sqrt{1-y^\alpha}} \right) . \quad (327)$$

It is easy to see that thus  $R_2$  is decreasing in  $\alpha$  since  $\log(y) \leq 0$ .

As for the second derivative, one can show that it writes as:

$$f''(\alpha) = \frac{P_v \log^2(y) \sqrt{1-y}}{4P_u (1-y^\alpha) \sqrt{1-y^\alpha}} \frac{1}{\left(1 + \sqrt{y^{1-\alpha} - y}\right)^3} G(\alpha) , \quad (328)$$

where  $G(\alpha)$  is given by:

$$\begin{aligned} G(\alpha) = & 2\sqrt{y^{1-\alpha} - y} \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right) \\ & + (1 + \sqrt{y^{1-\alpha} - y}) \left( (1 - y^\alpha) \sqrt{y^{1-\alpha}} (\sqrt{1-y^\alpha} - 1) + y^\alpha (y + \sqrt{y^{1-\alpha}}) \right) . \end{aligned} \quad (329)$$

Showing that  $R_2$  is convex in  $\alpha$ , i.e showing that (326) holds, amounts then to showing that:

$$\frac{P_v (1 + \sqrt{y^{1-\alpha}}) \sqrt{1-y} + P_u (1 + \sqrt{y^{1-\alpha} - y})}{\sqrt{1-y^\alpha} \sqrt{1-y}} G(\alpha) \geq P_v \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right)^2 . \quad (330)$$

We show the stronger result that consists in:

$$\frac{G(\alpha)}{\sqrt{1-y^\alpha}} \geq \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right)^2 , \quad (331)$$

which would yield the desired inequality.

Note here that since:

$$\begin{aligned} (1 - y^\alpha) \sqrt{y^{1-\alpha}} (\sqrt{1-y^\alpha} - 1) + y^\alpha (y + \sqrt{y^{1-\alpha}}) \\ \geq \sqrt{y^{1-\alpha}} (\sqrt{1-y^\alpha} - 1) + y^\alpha (y + \sqrt{y^{1-\alpha}}) \end{aligned} \quad (332)$$

$$\geq \sqrt{y^{1-\alpha}} (1 - y^\alpha - 1) + y^\alpha (y + \sqrt{y^{1-\alpha}}) \quad (333)$$

$$\geq 0 , \quad (334)$$

then,

$$G(\alpha) \geq 2\sqrt{y^{1-\alpha} - y} \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right) . \quad (335)$$

Hence ,we can write:

$$\begin{aligned} G(\alpha) - \sqrt{1-y^\alpha} \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right)^2 \\ \geq \sqrt{1-y^\alpha} \left( \sqrt{y^{1-\alpha}}(1 - \sqrt{1-y^\alpha}) + y \right) \left( \sqrt{y^{1-\alpha}}(1 + \sqrt{1-y^\alpha}) + y \right) \end{aligned} \quad (336)$$

$$\geq 0 , \quad (337)$$

this ends thus the proof. Thus,  $R_2$  being convex in  $\alpha$  and  $R_1$  linear in  $\alpha$  the trajectory of  $R_2(R_1)$  describes then a concave rate region.

If  $s_u = s_v = 1$ , we show that the obtained rate region is included in a suboptimal rate region compared to time sharing, studied earlier and given by:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right) , \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{P + N}{P_u + N} \right) . \end{cases} \quad (338)$$

To show this, note that the bound on  $R_1$  is trivial. However, concerning the second user's rate, note that as it writes:

$$R_2 = \frac{1}{2} \log_2 \left( 1 + P_v \frac{1 - \sqrt{1-x}}{P_u(1 - \sqrt{1-ax}) + 2N} \right) \triangleq \log_2 (1 + g(x)) , \quad (339)$$

$R_2$  is not always strictly monotonic. The sign of its first derivative in  $x$  is given by the sign of  $g'$ :

$$(P_u + 2N)\sqrt{1-ax} + P_u(a - 1 - a\sqrt{1-x}) , \quad (340)$$

that depends on the respective values of  $P_u$  and  $N$ . If there exists any point for which the first derivative is null  $x_{opt}$ , then it will imply that:

$$P_u(1 - \sqrt{1-ax}) + 2N = (P_u + 2N) \left( 1 - \sqrt{1-x} + 1 - \frac{1}{a} \right) \quad (341)$$

$$\geq (P_u + 2N) (1 - \sqrt{1-x}) \geq 0 . \quad (342)$$

Then, we can conclude that:

$$R_2 \leq \frac{1}{2} \log_2 \left( 1 + \frac{P_v}{P_u + 2N} \right) . \quad (343)$$

If no such point exists such that  $g'(x) = 0$ , then  $R_2$  is increasing in  $x$  and thus, the maximum value is obtained for  $x = 1/a$ , which clearly yields to the desired bound on  $R_2$ .

Now, if  $s_u s_v = -1$ , then two cases unfold following the signs of  $s_u$  and  $s_v$ . In both cases, the obtained rate region is shown not to outperform the optimal rate region we claim. If  $s_u = -1$  and  $s_v = 1$ , then, we can show that the rate region:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} (1 - \sqrt{x}) \right) \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{1 - \sqrt{1-x}}{P_u(1 + \sqrt{1-ax}) + 2N} \right) \leq \frac{1}{2} \log_2 \left( \frac{P + 2N}{P_u + 2N} \right), \end{cases} \quad (344)$$

is included in the sub-optimal rate region given by (284) which in turn does not outperform time sharing.

On the other side, if  $s_u = 1$  and  $s_v = -1$ , then we will show that the second corner point of CD-DPC inner bound contains the obtained rate region. In this case, it can be easily shown that both  $R_1$  and  $R_2$  are decreasing in  $x$ . However, not all values of  $x$  are admissible due to the constraint.

Let us start by characterizing the set of values such that:

$$N\sqrt{x} > \frac{P_v}{2(P + 2N)} \left( P_u(1 - x + \sqrt{(1-x)(1-ax)}) - 2Nx \right). \quad (345)$$

As done previously, we will solve only the simpler problem that yields a larger solution set and that is illustrated in Fig. 8:

$$N\sqrt{x} > \frac{P_v}{2(P + 2N)} (P_u(1 - x + 1 - ax) - 2Nx). \quad (346)$$

Solving this problem yields the following value of the infimum of all admissible beam

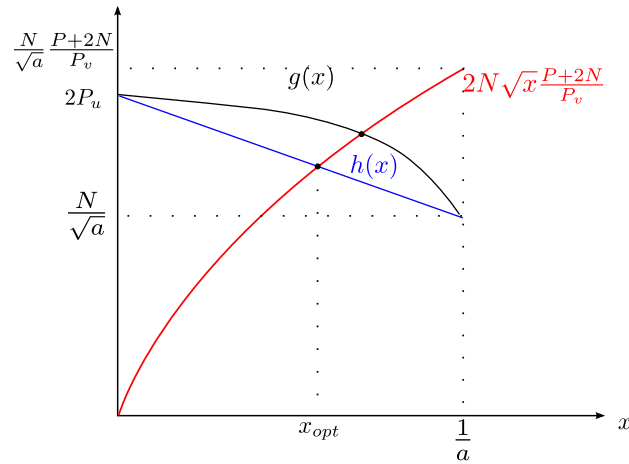


Fig. 8. Comparison of the functions  $h$ ,  $g$  and target upper bound.

directions:

$$x_{opt} = \frac{P_u^2}{2(P_u + N)^2(P_u + 2N)^2} \left[ 2(P_u + N)(P_u + 2N) + \frac{N^2(P + 2N)^2}{P_v^2} - \frac{N(P + 2N)}{P_v} \sqrt{\frac{N^2(P + 2N)^2}{P_v^2} + 4(P_u + N)(P_u + 2N)} \right]. \quad (347)$$

Since the solution of problem (346) yields a smaller value of the inf of admissible solutions, the resulting rate region is wider. However, we show that it still remains included in the second corner point of MD-DPC inner bound given by

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P + 2N}{P_u + 2N} \right), \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{P_v + N}{N} \right). \end{cases} \quad (348)$$

The bound on the rate  $R_2$  is quite trivial and requires no further proof. However, the bound on rate  $R_1$  requires showing that:

$$x_{opt} \geq \frac{P_u^2 P_v^2}{(P_u + 2N)^2 (P_v + 2N)^2}, \quad (349)$$

which can be shown through evolved bounding techniques. As a conclusion for these cases, no rate region outperforms the second corner point of CD-DPC inner bound.

*D. case of  $A_1 \neq A_2$  and  $(B_1 - B_2)^2 = (A_1 - A_2)(D_1 - D_2)$*

In this case, we start by showing that the above condition imposes:

$$\theta_u = \theta_v[\pi]. \quad (350)$$

Let us first quickly denote:

$$K_{Y_1} \triangleq \cos^2(\theta_u)P_u + \cos^2(\theta_v)P_v + N, \quad (351)$$

$$K_{Y_2} \triangleq \sin^2(\theta_u)P_u + \sin^2(\theta_v)P_v + N. \quad (352)$$



Next, recall that:

$$B_1 = \frac{P_v \cos(\theta_u) \cos(\theta_v)}{\cos^2(\theta_u)P_u + \cos^2(\theta_v)P_v + N} = \frac{P_v \cos(\theta_u) \cos(\theta_v)}{K_{Y_1}}, \quad (353)$$

$$B_2 = \frac{P_v \sin(\theta_u) \sin(\theta_v)}{\sin^2(\theta_u)P_u + \sin^2(\theta_v)P_v + N} = \frac{P_v \sin(\theta_u) \sin(\theta_v)}{K_{Y_2}}, \quad (354)$$

$$A_1 = \frac{P_v \cos^2(\theta_u)P_u + N}{P_u K_{Y_1}} = \frac{P_v}{P_u} \left( 1 - \frac{\cos^2(\theta_v)P_v}{K_{Y_1}} \right), \quad (355)$$

$$A_2 = \frac{P_v \sin^2(\theta_u)P_u + N}{P_u K_{Y_2}} = \frac{P_v}{P_u} \left( 1 - \frac{\sin^2(\theta_v)P_v}{K_{Y_2}} \right), \quad (356)$$

$$D_1 = \frac{\cos^2(\theta_v)P_v + N}{K_{Y_1}} = \left( 1 - \frac{\cos^2(\theta_u)P_u}{K_{Y_1}} \right), \quad (357)$$

$$D_2 = \frac{\sin^2(\theta_v)P_v + N}{K_{Y_2}} = \left( 1 - \frac{\sin^2(\theta_u)P_u}{K_{Y_2}} \right). \quad (358)$$

And that  $A_1 \neq A_2 \Rightarrow P_v \neq 0$ . Thus,

$$\begin{aligned} (B_1 - B_2)^2 &= (A_1 - A_2)(D_1 - D_2) \\ \Leftrightarrow \left( \frac{\cos(\theta_u) \cos(\theta_v)}{K_{Y_1}} - \frac{\sin(\theta_u) \sin(\theta_v)}{K_{Y_2}} \right)^2 \\ &= \left( \frac{\sin^2(\theta_v)}{K_{Y_2}} - \frac{\cos^2(\theta_v)}{K_{Y_1}} \right) \left( \frac{\sin^2(\theta_u)}{K_{Y_2}} - \frac{\cos^2(\theta_u)}{K_{Y_1}} \right) \end{aligned} \quad (359)$$

$$\Leftrightarrow K_{Y_1} K_{Y_2} \left( \sin(\theta_v) \cos(\theta_u) - \sin(\theta_u) \cos(\theta_v) \right)^2 = 0 \quad (360)$$

$$\Leftrightarrow K_{Y_1} K_{Y_2} \sin^2(\theta_u - \theta_v) = 0 \quad (361)$$

$$\Leftrightarrow \theta_u = \theta_v[\pi]. \quad (362)$$

The optimal solution of the system (265), is then given by:

$$R_1 = \log_2 \left( \frac{\min(\sin^2(\theta_v), \sin^2(\theta_u))P_u + N}{N} \right) \quad (363)$$

$$= \log_2 \left( \frac{(1 - |\cos(2\theta_v)|)P_u + 2N}{2N} \right) \quad (364)$$

$$\leq \log_2 \left( \frac{(1 - \sqrt{\cos^2(2\theta_v)})P_u + 2N}{2N} \right), \quad (365)$$

define then:

$$x \triangleq \cos^2(2\theta_v). \quad (366)$$

On the other side, note that:

$$R_2 = \log_2 \left( \frac{\cos^2(\theta_v + \pi/4)P + N}{\cos^2(\theta_v + \pi/4)P_u + N} \right) \quad (367)$$

$$= \log_2 \left( \frac{\sin(2\theta_v)P + 2N}{\sin(2\theta_v)P_u + 2N} \right) \quad (368)$$

$$\leq \log_2 \left( \frac{(1 - s_v\sqrt{1-x})P + 2N}{(1 - s_v\sqrt{1-x})P_u + 2N} \right), \quad (369)$$

if  $s_v = -1$ , then  $R_2$  is decreasing in  $x$ , and thus the optimal value is given by:

$$R_2 = \log_2 \left( \frac{P + 2N}{P_u + 2N} \right). \quad (370)$$

And since:

$$R_1 \leq \log_2 \left( \frac{P_u + 2N}{2N} \right), \quad (371)$$

then the obtained region is included in the set of rate pairs such that:

$$\begin{cases} R_1 \leq \log_2 \left( \frac{P_u + 2N}{2N} \right), \\ R_2 \leq \log_2 \left( \frac{P + 2N}{P_u + 2N} \right). \end{cases} \quad (372)$$

which was shown to perform less than time sharing. Now, if  $s_v = 1$ , then  $R_2$  is increasing in  $x$  where  $x \in [0 : 1]$ , and hence, the maximum is obtained for  $x = 1$ , which yields the same rate of user 2. For similar reasons, the obtained rate pair does not outperform time sharing.

Thus, the overall rate region obtained in this case, is included in mere time sharing.

*E. case of  $A_1 \neq A_2$  and  $(B_1 - B_2)^2 < (A_1 - A_2)(D_1 - D_2)$*

Since we have that:

$$(B_1 - B_2)^2 - (A_1 - A_2)(D_1 - D_2) = \frac{Pv^2}{K_{Y_1}K_{Y_2}} \left( \sin(\theta_v - \theta_u) \right)^2, \quad (373)$$

having  $(B_1 - B_2)^2 < (A_1 - A_2)(D_1 - D_2)$  is impossible.

*F. case of  $A_1 \neq A_2$ ,  $(B_1 - B_2)^2 > (A_1 - A_2)(D_1 - D_2)$  and  $|c_1 - c_2| \leq \min(A_1, A_2)(\alpha_1 - \alpha_2)^2$*

In this case, we can show that, the two possible optimum solutions are obtained for  $\theta_u = \pi/4$  or  $\theta_u = -\pi/4$ .

The case where  $\theta_u = \pi/4$  is the claimed optimal rate region. As for the case where  $\theta_u = -\pi/4$ , the resulting rate region consists of all rate pairs satisfying:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{P_u P_v \frac{1+y}{P+2N} + \sqrt{P+2N+(y^2-1)P_v}} + 2N \right) \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{1-y}{2(P_u+N)} \right) \end{cases}, \quad (374)$$

Note, that in this case, the maximum is achieved at both rates letting  $y = -1$ , thus, the resulting rate region can not outperform the rate region given by:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right) \\ R_2 \leq \frac{1}{2} \log_2 \left( \frac{P + N}{P_u + N} \right) \end{cases}, \quad (375)$$

which was clearly shown not to outperform time-sharing.

*G. case of  $A_1 \neq A_2$ ,  $(B_1 - B_2)^2 > (A_1 - A_2)(D_1 - D_2)$  and  $|c_1 - c_2| > \min(A_1, A_2)(\alpha_1 - \alpha_2)^2$*

In this peculiar last case, we show that the obtained rate region can not exceed time sharing neither. In this case, the resulting rate region writes as:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{(1 - |\cos(2\theta_u)|)P_u + 2N}{2N} \right) \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{(1 - s_v \sqrt{1 - \cos^2(2\theta_v)})}{P_u(1 - s_u \sqrt{1 - \cos^2(2\theta_u)}) + 2N} \right) \end{cases} \quad (376)$$

The case where  $s_u = -1$ , then we can show resorting to the same tools used on the analysis of the concavity in the previous sections that the obtained rate region when  $\cos^2(2\theta_u)$  spans the interval  $[0 : 1]$ , is concave for every value of  $\cos^2(2\theta_v)$ , thus, the resulting union can be at most concave. When  $s_u = 1$ , two cases unfold and  $R_1$  and  $R_2$  are both decreasing in  $\cos^2(2\theta_u)$ , and for a fixed beam direction  $\mathbf{B}_v$ , the optimal rate pair is given by:

$$\begin{cases} R_1 \leq \frac{1}{2} \log_2 \left( \frac{P_u + 2N}{2N} \right) \\ R_2 \leq \frac{1}{2} \log_2 \left( 1 + P_v \frac{(1 - s_v \sqrt{1 - \cos^2(2\theta_v)})}{2N} \right) \end{cases}. \quad (377)$$

## APPENDIX L

PROOF OF ACHIEVABILITY OF  $\mathcal{R}_{3-ARV}$ 

We fix a pmf  $p_{QU_1U_2VX}$ . Let  $R_0, R_1, R_2$  denote the message rates and  $T_{1,2}, T_{1,1}$  and  $T_2$  denote the binning rates. Generate  $2^{nR_0}$  sequences  $q^n(w_0)$ ,  $w_0 \in [1 : 2^{nR_0}]$  each following the pmf:  $\prod_{i=1}^n P_Q(q_i(w_0))$ . For each  $w_0$ , generate  $2^{nT_2}$  sequences  $v^n(l_2, w_0)$  following the pmf:  $\prod_{i=1}^n P_{V|Q}(v_i(l_2, w_0)|q_i(w_0))$  and map them randomly in  $2^{nR_2}$  bins  $B^n(w_2, w_0)$ . Generate similarly  $2^{nT_{1,1}}$  sequences  $u_1^n(l_{1,1}, w_0)$  and map them randomly in  $2^{nR_1}$  bins  $B_1^n(w_1, w_0)$  and  $2^{nT_{1,2}}$  sequences  $u_2^n(l_{1,2}, w_0)$  and map them in a distinct set of  $2^{nR_1}$  bins  $B_2^n(w_1, w_0)$ .

*Encoding:* for each message triple  $(w_0, w_1, w_2)$  to be transmitted, find in the product of all bins  $B^n(w_i, w_0)$ , a triple of sequences  $u_1^n(l_{1,1}, w_0), u_2^n(l_{1,2}, w_0), v^n(l_2, w_0)$  such that:

$$\left( q^n(w_0), u_1^n(l_{1,1}, w_0), u_2^n(l_{1,2}, w_0), v^n(l_2, w_0) \right) \in T_\delta^n(QU_1U_2V).$$

Send then a random mapping sequence:  $x^n(w_0, l_{1,1}, l_{1,2}, l_2)$ . The encoding is error free if all inequalities in  $\mathbb{T}$  are verified.

*Decoding:* Each receiver decodes its intended messages  $(w_0, w_j)$  by decoding the index  $l_j$  and non-uniquely the common message, yielding the constraints stated in  $\mathcal{M}$ .

## REFERENCES

- [1] T. Cover, "Broadcast channels," *Information Theory, IEEE Transactions on*, vol. 18, no. 1, pp. 2–14, 1972.
- [2] —, "Comments on broadcast channels," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2524–2530, 1998.
- [3] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *Information Theory, IEEE Transactions on*, vol. 25, no. 3, pp. 306–311, 1979.
- [4] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," in *Information Theory, 2006 IEEE International Symposium on*, 2006, pp. 2205–2209.
- [5] A. El Gamal, "The capacity of a class of broadcast channels," *Information Theory, IEEE Transactions on*, vol. 25, no. 2, pp. 166–169, 1979.
- [6] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4207–4214, 2010.
- [7] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Peredachi Inf.*, vol. 16, pp. 24–34, 1980.
- [8] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 3936–3964, 2006.
- [9] A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Peredachi Inf.*, vol. 16, pp. 3–23, 1980.

- [10] B. Bandemer, A. Gamal, and Y.-H. Kim, “Simultaneous nonunique decoding is rate-optimal,” in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, 2012, pp. 9–16.
- [11] F. Baccelli, A. El Gamal, and D. Tse, “Interference networks with point-to-point codes,” *Information Theory, IEEE Transactions on*, vol. 57, no. 5, pp. 2582–2596, 2011.
- [12] S. Bidokhti and V. Prabhakaran, “Is non-unique decoding necessary?” *Information Theory, IEEE Transactions on*, vol. 60, no. 5, pp. 2594–2610, 2014.
- [13] A. Padakandla and S. S. Pradhan, “A new coding theorem for three user discrete memoryless broadcast channel,” *CoRR*, vol. abs/1207.3146, 2012.
- [14] H. Weingarten, S. Shamai and G. Kramer, “On the compound MIMO broadcast channel,” *Information Theory and Applications (ITA 2007)*, UCSD, Palo-Alto, USA., Ed., Jan. 29-Feb. 2 2007.
- [15] T. Gou, S. Jafar, and C. Wang, “On the degrees of freedom of finite state compound wireless networks,” *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3286–3308, 2011.
- [16] M. Maddah-Ali, “On the degrees of freedom of the compound MISO broadcast channels with finite states,” in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, 2010, pp. 2273–2277.
- [17] S. Jafar, *Interference Alignment: A New Look at Signal Dimensions in a Communication Network*. Now Publishers, 2011. [Online]. Available: <http://books.google.fr/books?id=GfwB7ItK4esC>
- [18] C. Huang, S. Jafar, S. Shamai, and S. Vishwanath, “On degrees of freedom region of MIMO networks without channel state information at transmitters,” *Information Theory, IEEE Trans. on*, vol. 58, no. 2, pp. 849–857, feb. 2012.
- [19] R. Tandon, S. Jafar, S. Shamai Shitz, and H. Poor, “On the synergistic benefits of alternating CSIT for the MISO broadcast channel,” *Information Theory, IEEE Transactions on*, vol. 59, no. 7, pp. 4106–4128, 2013.
- [20] P. Piantanida and S. Shamai, “On the capacity of compound state-dependent channels with states known at the transmitter,” in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, 2010, pp. 624–628.
- [21] M. H. M. Costa, “Writing on dirty paper (corresp.),” *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, May 1983.
- [22] M. Benammar and P. Piantanida, “On the role of interference decoding in compound broadcast channels,” in *Information Theory Workshop (ITW), 2013 IEEE*, Sept 2013, pp. 1–5.
- [23] C. Nair and A. El Gamal, “The capacity region of a class of three-receiver broadcast channels with degraded message sets,” *Information Theory, IEEE Transactions on*, vol. 55, no. 10, pp. 4479–4493, 2009.
- [24] H. Witsenhausen and A. Wyner, “A conditional entropy bound for a pair of discrete random variables,” *Information Theory, IEEE Transactions on*, vol. 21, no. 5, pp. 493–501, 1975.
- [25] H. EGGLESTON, *Convexity*. Published by Cambridge University Press, 1958.
- [26] G. Caire and Shamai, “On the achievable throughput of a multi-antenna gaussian broadcast channel,” vol. IT-49, pp. 1691–1706, july 2003.
- [27] M. H. M. Costa, “Writing on dirty paper (corresp.),” *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, 1983.
- [28] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, “The capacity region of the degraded multiple-input multiple-output compound broadcast channel,” *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5011–5023, 2009.
- [29] H.-F. Chong and Y.-C. Liang, “The capacity region of a class of two-user degraded compound broadcast channels,” in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, 2013, pp. 932–936.

- [30] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Akadémiai Kiado, Budapest, 1982.
- [31] T. Cover and J. Thomas, *Elements of information theory (2nd Ed)*. Wiley-Interscience, 2006.