

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

Meryem Benammar^{ib}, *Member, IEEE*, Pablo Piantanida^{ib}, *Senior Member, IEEE*,
and Shlomo Shamai^{ib}, *Life Member, IEEE*

Abstract—This work investigates the general two-user compound Broadcast Channel (BC) in which an encoder wishes to transmit two private messages W_1 and W_2 to two receivers while being oblivious to the actual channel realizations controlling the communication. The focus is on the characterization of the largest achievable rate region by resorting to more involved encoding and decoding techniques than the usual coding schemes of the standard BC. Involved decoding schemes are first explored, and an achievable rate region is derived based on the principle of *Interference Decoding* (ID), in which each receiver decodes its intended message and chooses to (non-uniquely) decode, or not, the interfering non-intended message. This decoding scheme 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 *No Interference Decoding* (NID)—fails to achieve the capacity region. Involved encoding schemes are later investigated, and an achievable rate region is derived based on *Multiple Description* (MD) coding wherein the encoder transmits a common description as well as multiple dedicated private descriptions to the many possible channel realizations of the users. It turns out that MD coding yields larger inner bounds than the single description scheme—*Common Description* (CD) coding—for a class of compound Multiple Input Single Output Broadcast Channels (MISO BC).

Index Terms—Channel capacity, channel state information, broadcast channels (BCs), compound channels, multiple description coding, multiple input single output broadcast channel (MISO BC).

I. INTRODUCTION

THE two-user Broadcast Channel (BC) –as first introduced by Cover in [1]– consists of an encoder which wishes to

Manuscript received October 20, 2014; revised November 25, 2015; accepted June 30, 2018. Date of publication September 23, 2019; date of current version December 23, 2019. This work was supported in part by the FP7 Network of Excellence in Wireless Communications (NEWCOM#). The work of S. Shamai was supported in part by the European Union’s Horizon 2020 Research and Innovation Programme, under Agreement 694630. This article was presented in part at the IEEE Information Theory Workshop, September 2013, and in part at the IEEE Information Theory Workshop, November 2014.

M. Benammar was with the Department of Telecommunications, CentraleSupélec, 91190 Gif-sur-Yvette, France. He is now with the Department of Electronics Optronics and Signal Processing, Institut Supérieur de l’Aéronautique et de l’Espace (ISAE-Supaero), 31055 Toulouse, France (e-mail: meryem.benammar@isae-supaeo.fr).

P. Piantanida is with CentraleSupélec, French National Center for Scientific Research (CNRS), Université Paris-Sud, 91192 Gif-sur-Yvette, France, and also with the Montreal Institute for Learning Algorithms (Mila), Université de Montréal, Montréal, QC H3T 1N8, Canada (e-mail: pablo.piantanida@centralesupelec.fr).

S. Shamai is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 3200003, Israel, (e-mail: sshlomo@ee.technion.ac.il).

Communicated by Daniela Tuninetti, Associate Editor for Communications. Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2019.2942615

transmit both a common message W_0 and two private messages W_1 (resp.) W_2 , each dedicated to a user observing the channel output Y (resp.) Z . Following this seminal work, intensive research was undertaken to characterize the capacity region of this setting which implies the design of efficient interference mitigation techniques.

In this work, we study the two-user compound BC wherein an encoder wishes to communicate two private messages W_1 (resp. W_2) to two users which can each observe the output of one of many possible channel statistics (Y_1, \dots, Y_J) (resp. (Z_1, \dots, Z_K)). The actual channel statistic controlling the communication is unknown to the transmitter, however, it is assumed to remain constant during the communication, and the aim is to ensure reliable communication whatever the channel realization. The compound channel model is relevant whenever the transmitter fails to acquire a perfect estimate of the channel but knows only a subset, or an interval, to which it belongs. Finite rate feedback from the receiver to the transmitter, which relies on a quantization step, might be the most realistic scenarios in which compound channels are encountered. It is also well understood that, when interested in the maximum error probability, the compound BC is equivalent to a BC with multiple users and common messages. Thus, the channel uncertainty in the compound BC is equivalent to multicasting in a multi-user scenario.

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 remains an open problem to date, Marton established in [3] an achievable rate region for the general two-user BC based on the notion of *random binning* and *superposition coding*, with common and private messages, which is commonly referred to as *Marton’s inner bound*. 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, e.g., 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 optimal coding scheme for such *ordered* BCs relies on *superposition coding* scheme at the encoder and allowing the user with the best channel observation to *decode* the interfering message, i.e., that of the opposite user, in addition to its intended message. 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], as well as the product and sum of two unmatched channels in [9]. For the latter channel models, it is rather *random binning* that proves to be

crucial for interference management, i.e., *precoding* against interference.

In the works listed herebefore, the channel statistic is perfectly known to the transmitter and thus, the encoder can exploit this knowledge to allow for an efficient interference mitigation scheme. Indeed, in all settings where Marton's inner bound is tight, the construction of the optimal auxiliary code depends on the *prior* knowledge of either the channel output statistic (e.g. deterministic and semi-deterministic BCs) or a function of these statistics (e.g. users' ordering in ordered single antennas BCs). When the transmitter has no Channel State Information (no-CSIT), very few capacity results are known for the compound BC among which are the two results in [10] and [11]. The coupling between interference and channel uncertainty calls for more involved coding schemes that Marton's coding scheme, which will be the focus of our contribution.

A. Related Works

We investigate in this work more involved encoding and decoding techniques than the usual coding schemes that proved to be capacity-achieving for some classes of BCs, namely, Interference Decoding (ID) and Multiple Description (MD) coding.

The idea behind ID was first introduced in [12] for Gaussian settings and formalized later in [13], and it consists in a combination of *non-unique decoding* with the possibility at each receiver to *decode, or not*, the interfering messages intended to the others users. The benefits of this decoding scheme results not only from non-unique decoding [14] but it follows essentially from allowing to *decode, or not*, interference. Note here that the straight-forward extension of the results of [13] to the BC is not general enough for it encompasses only *superposition coding* but fails at including *random binning*. Nevertheless, it provides an interesting insight on how to recover a *superposition coding* like inner bound while keeping a symmetric encoding, which will be crucial in the construction of our Interference Decoding scheme.

While ID allows to choose between *decoding or not* the interference, authors in [15] suggested a coding scheme which relies on *decoding part* of interference, which includes decoding interference or not decoding it. Though the setting investigated therein is fundamentally different from the compound BC we are investigating here, the results they suggest are strongly related to ours and will be thoroughly compared to them later on. The idea behind the coding scheme in [15] is to introduce an additional auxiliary codeword in the encoding step which is meant to carry part of the interfering message and to be decoded at only one user. This renders the extension of the inner bound to multiple users very hard since it involves extra superposition and binning operations. Thus, we will not pursue this direction in our work, but will restrict to Marton's encoding scheme whilst improving on Marton's inner bound through ID.

Authors in [16] derived an inner bound based on *coset codes* for the three users BC. The idea behind using coset codes is to allow the users to decode a *compressive* function of

the interfering messages of other users, for instance, $w_2 \oplus w_3$, and thus a complete cancellation of interference with less impediment to the information rates than *fully* decoding the interfering messages, i.e., decoding w_2 and w_3 . 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 is strictly suboptimal as compared to the coset coding. Such a coding technique based on coset codes, proves to be useful when there are many interfering messages, 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 messages, is strictly sub-optimal.

When the channels are *not ordered*, e.g. MISO BC, designing optimal coding schemes with no-CSIT calls for more involved *encoding* strategies. The intuition follows from the analysis of the effect of channel uncertainty on the *Degrees of Freedom* (DoF), which are very insightful to understand how interference should be managed in multiple antennas settings. 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 [17] with some cases of optimality. The outer bound derived therein was conjectured to be loose, but later Gou *et al.* [18] and Maddah-Ali [19] proved the optimal DoF region of the generic compound MISO BC, both in the complex and in the real settings, to match this outer bound. The achievability of the optimal DoF relies on either a linear or a non-linear coding scheme combined with *block expansion* (coding over many time slots) in [17] while the proof made in [19] resorts to number theory tools and consists in *Interference Alignment* (IA) over rational dimensions of the real numbers (see also [20]). When the states span an infinite set, i.e., in the ergodic setting, DoF become limited. Indeed, in [21], 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 various models of the amount and accuracy of CSI available at the transmitter, e.g. [22]. It turns out that richer *encoding strategies*, like *interference alignment* along with *block expansion* are crucial in dealing with interference, and thus, any optimal scheme for the finite power limited MISO BC should encompass such DoF-optimal schemes.

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 Marton's *random binning*, and *superposition coding* with the choice of *decoding or not* interference which we denote by ID. We show that for the compound BC –unlike the standard two-user BC– ID can strictly outperform Marton's standard coding scheme based only on *random binning*, and *superposition coding*, denoted

later by *No-Interference Decoding* (NID). This results from that ID allows a symmetric encoding and deals with the source's uncertainty by relegating the *asymmetric decoding* to the receive terminals. To illustrate clearly the gain of ID over NID, we investigate a class of discrete ordered compound Binary Erasure/ Binary Symmetric (BEC/BSC) BCs for which we derive the capacity region resorting to ID and show that NID yields a strictly sub-optimal rate region.

As for the class of non-ordered compounds BCs, more involved encoding schemes need to be investigated since we need to *precode* against interference rather than to *decode* it. Since each channel statistic involves a distinct interference signal, precoding against interference with only one common auxiliary code, i.e., Common Description (CD) coding, might be inefficient. For this reason, we look at the role that *Multiple Description* (MD) coding can play in the non-ordered compound BC, where the encoder precodes against interference *differently* for the many channel statistics of each user through private descriptions each tailored to one channel statistic. We follow a similar approach to that in [23] where MD coding has already proved to be useful over compound state-dependent channels. We prove that MD coding is beneficial as compared with CD coding [3] and we illustrate this for a class of compound Gaussian MISO BC under a specific *Dirty-Paper Coding* (DPC) scheme [24].

Finally, we discuss the relative behaviour 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 presents 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 is strictly sub-optimal. Next, in Section IV, we introduce MD coding and specialize it to 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 p.m.f will refer to probability mass function. Random variables (resp. their realizations) are denoted by upper (resp. lower) case letters. Vectors are denoted by bold font characters and RV stands for random variable, ARV for Auxiliary 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_X is specific to the of a RV X . $|\mathcal{X}|$ 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 joint probability distribution P_{XYZ} . If $P_{XYZ}(x|yz) = P_{X|Y}(x|y)$ for all x, y, z , then they form a Markov chain, which is denoted by $X \dashv\vdash Y \dashv\vdash Z$. 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 , $Y_2 \preceq Y_1$ means Y_1 is less noisy than Y_2 . \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 2×1 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

Consider the compound BC problem which consists in one source terminal and two distinct receivers each observing one of many possible channel outputs and where the source wishes to communicate two private messages W_1 and W_2 , one to each receiver. This setup is equivalent to a setting where each user is represented by multiple users that are interested in the same message W_1 or W_2 .

A. Definition of the Compound BC

- Consider a collection of n -th extensions of discrete memoryless BCs (defined by a pmf and the input and output alphabets)

$$\{\mathcal{W}_{j,k}^n\}_{j \in \mathcal{J}, k \in \mathcal{K}} = \{P_{Y_j^n Z_k^n | X^n}, \mathcal{X}^n, \mathcal{Y}^n, \mathcal{Z}^n\},$$

defined by the conditional p.m.fs:

$$P_{Y_j^n Z_k^n | X^n} = \prod_{i=1}^n P_{Y_{j,i} Z_{k,i} | X_i}.$$

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

The probability of error is given by:

$$P_e^{(n)}(j, k) \triangleq \mathbb{P} \left(\hat{W}_1^j \neq W_1 \cup \hat{W}_2^k \neq W_2 \right).$$

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

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log_2 M_{1n} \geq R_1 \text{ for all } l \in \{1, 2\}$$

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

The capacity region is the convex hull of the set of all achievable rate pairs (R_1, R_2) and is denoted as $\mathcal{C}_{\mathcal{J} \times \mathcal{K}}$.

B. Outer Bound on the Capacity Region of the Compound BC

We derive in this section an 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 region of the BC. It will be useful in the examples we investigate later.

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

$$\mathcal{R}_{\text{NEG}}^{(j,k)}(P_{QUVX}) \triangleq \left\{ (R_1, R_2) : \begin{aligned} R_1 &\leq I(QU; Y_j), \\ R_2 &\leq I(QV; Z_k), \\ R_1 + R_2 &\leq I(U; Y_j | QV) + I(QV; Z_k), \\ R_1 + R_2 &\leq I(QU; Y_j) + I(V; Z_k | QU) \end{aligned} \right\},$$

for a specific joint p.m.f 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 $\mathcal{C}_{\mathcal{J} \times \mathcal{K}}$ verifies:*

$$\mathcal{C}_{\mathcal{J} \times \mathcal{K}} \subseteq \mathcal{C}_o \triangleq \bigcup_{P_U P_V} \bigcap_{j \in \mathcal{J}, k \in \mathcal{K}} \bigcup_{P_{Q|UV}} \mathcal{R}_{\text{NEG}}^{(j,k)}(P_{QUVX}), \quad (1)$$

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

Remark 2. *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 ordered compound setups, the fact of optimizing the common auxiliary RV Q for each channel with index (j, k) , prevents 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: We need to recall that the proof in [4] of the outer bound for users' pair (j, k) , uses the specific choice of auxiliary RV:

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

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 (j, k) with the specific choice: $U_i = W_1, V_i = W_2$,

$$\begin{aligned} nR_1 &\leq \sum_{i=1}^n I(Q_i^{(j,k)} U_i; Y_{j,i}), \\ nR_2 &\leq \sum_{i=1}^n I(Q_i^{(j,k)} V_i; Z_{k,i}), \\ n(R_1 + R_2) &\leq \sum_{i=1}^n [I(U_i; Y_{j,i} | Q_i^{(j,k)} V_i) \\ &\quad + I(Q_i^{(j,k)} V_i; Z_{k,i})], \end{aligned}$$

$$\begin{aligned} &+ I(Q_i^{(j,k)} U_i; Y_{j,i}) \\ &+ I(Q_i^{(j,k)} V_i; Z_{k,i}), \end{aligned}$$

where $Q_i^{(j,k)} = (Y_{j,1}^{i-1}, Z_{k,i+1}^n)$. Thus, we could possibly factor the resulting joint p.m.f 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. ■

III. ID FOR THE COMPOUND BROADCAST CHANNEL

In this section, we derive an inner bound on the capacity region of the compound which relies on Marton's encoding scheme with common and private codewords, generated and mapped via superposition coding and random binning, but resorts to ID at the decoders.

A. ID Inner Bound

The inner bound we derive here shares common ideas with following works [25]. First, the idea used in [12] where – roughly speaking – each receiver is required to decode its intended message and is as well allowed to 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 [15].

Theorem 3 (ID inner bound). *An inner bound on the capacity region of the compound BC consists in the set of all rates (R_1, R_2) included in:*

$$\begin{aligned} \mathcal{R}_{\text{ID}} \triangleq & \bigcup_{P_{QUVX} \in \mathcal{P}} \bigcup_{\substack{(R_{0,1}, R_{0,2}) \\ \leq (R_1, R_2)}} \bigcup_{\substack{(T_1, T_2) \in \\ \mathbb{T}(P_{QUVX})}} \bigcap_{\substack{j \in \mathcal{J} \\ k \in \mathcal{K}}} \\ & \underbrace{\bigcup_{m=1}^4 \mathcal{T}_m^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2)}_{(4 \text{ Methods})}, \end{aligned}$$

where \mathcal{P} is the set of all input p.m.f's P_{QUVX} such that $(Q, U, V) \circlearrowleft X \circlearrowleft (Y_1, \dots, Y_J, Z_1, \dots, Z_K)$.

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

$$\begin{aligned} \mathbb{T}(P_{QUVX}) \triangleq & \left\{ (T_1, T_2) : \right. \\ & T_1 \geq R_1 - R_{0,1}, \\ & T_2 \geq R_2 - R_{0,2}, \\ & \left. T_1 + T_2 > R_1 - R_{0,1} + R_2 - R_{0,2} + I(U; V | Q) \right\}, \end{aligned}$$

Proof. The proof is relegated to Appendix B. Hereafter we summarize its main steps. The messages (W_1, W_2) are first

$$\begin{aligned}
\mathcal{T}_1^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2) &\triangleq \begin{cases} T_1 \leq I(U; Y_j|Q), \\ R_{0,1} + R_{0,2} + T_1 \leq I(QU; Y_j), \\ T_2 \leq I(V; Z_k|Q), \\ R_{0,1} + R_{0,2} + T_2 \leq I(QV; Z_k), \end{cases} \\
\mathcal{T}_2^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2) &\triangleq \begin{cases} T_1 \leq I(U; Y_j|Q), \\ R_{0,1} + R_{0,2} + T_1 \leq I(QU; Y_j), \\ T_2 \leq I(V; Z_kU|Q), \\ T_1 + T_2 \leq I(UV; Z_k|Q) + I(U; V|Q), \\ R_{0,1} + R_{0,2} + T_1 + T_2 \leq I(QUV; Z_k) + I(U; V|Q), \end{cases} \\
\mathcal{T}_3^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2) &\triangleq \begin{cases} T_1 \leq I(U; Y_jV|Q), \\ T_1 + T_2 \leq I(UV; Y_j|Q) + I(U; V|Q), \\ R_{0,1} + R_{0,2} + T_1 + T_2 \leq I(QUV; Y_j) + I(U; V|Q), \\ T_2 \leq I(V; Z_k|Q), \\ R_{0,1} + R_{0,2} + T_2 \leq I(QV; Z_k), \end{cases} \\
\mathcal{T}_4^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2) &\triangleq \begin{cases} T_1 \leq I(U; Y_jV|Q), \\ T_1 + T_2 \leq I(UV; Y_j|Q) + I(U; V|Q), \\ R_{0,1} + R_{0,2} + T_1 + T_2 \leq I(QUV; Y_j) + I(U; V|Q), \\ T_2 \leq I(V; Z_kU|Q), \\ T_1 + T_2 \leq I(UV; Z_k|Q) + I(U; V|Q), \\ R_{0,1} + R_{0,2} + T_1 + T_2 \leq I(QUV; Z_k) + I(U; V|Q). \end{cases}
\end{aligned}$$

split into two parts: common messages $(W_{0,1}, W_{0,2})$ transmitted through the common codeword Q , i.e., decoded by both users, and private messages $(W_{p,1}, W_{p,2})$ transmitted only through the private codewords U and V and intended to their respective decoders. Encoding is performed in the same fashion as Marton's encoding scheme through binning and superposition coding.

Each user introduces the union of two sets of constraints at the decoding, each set corresponds to decoding or not the interference. This results –in terms of achievable rates– in the union of four rate regions:

- The region $\mathcal{T}_1^{(j,k)}$ is the same rate region as obtained with Marton's inner bound, which does not involve decoding interfering messages,
- The region $\mathcal{T}_4^{(j,k)}$ in which both decoders decode their intended and the interfering messages,
- The regions $\mathcal{T}_2^{(j,k)}$ and $\mathcal{T}_3^{(j,k)}$ correspond to each destination decoding the interfering message in turns.

A similar rate region was also derived in [13], but it does not take advantage of Marton's encoding technique, i.e., random binning with common and private codewords, and thus fails at achieving even Marton's inner bound in the compound setting. \square

Remark 4. Consider the standard two-user BC, i.e., where $J = K = 1$. Observe that the rate region \mathcal{R}_{ID} contains Marton's rate region [3], which we will denote in the following as \mathcal{R}_{NID} . These regions are given by

$$\mathcal{R}_{ID} \triangleq \bigcup_{\substack{P_{QUVX} \in \mathcal{P} \\ (T_1, T_2) \in \mathbb{T}(P_{QUVX}) \\ (R_{0,1}, R_{0,2})}} \bigcup_{i=1}^4 \mathcal{T}_i(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2),$$

$$\mathcal{R}_{NID} \triangleq \bigcup_{\substack{P_{QUVX} \in \mathcal{P} \\ (T_1, T_2) \in \mathbb{T}(P_{QUVX}) \\ (R_{0,1}, R_{0,2})}} \mathcal{T}_1(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2).$$

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

$$\mathcal{R}_{ID} = \bigcup_{i=1}^4 \mathcal{R}_i = \mathcal{R}_{NID} \cup \left(\bigcup_{i=2}^4 \mathcal{R}_i \right),$$

where the regions $\mathcal{R}_k, k \in \{2, 3\}$ are respectively defined by the following sets of inequalities

$$\begin{aligned}
\mathcal{R}_2 = \{ (R_1, R_2) : \\
R_1 \leq I(QU; Y), \\
R_1 + R_2 \leq I(V; Z|UQ) + I(QU; Y), \\
R_1 + R_2 \leq I(QUV; Z) \},
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}_3 = \{ (R_1, R_2) : \\
R_2 \leq I(QV; Z), \\
R_1 + R_2 \leq I(U; Y|VQ) + I(QV; Z), \\
R_1 + R_2 \leq I(QUV; Y) \},
\end{aligned}$$

$$\begin{aligned}
\mathcal{R}_4 = \{ (R_1, R_2) : \\
R_1 + R_2 \leq I(QUV; Y), \\
R_1 + R_2 \leq I(QUV; Z) \},
\end{aligned}$$

¹For the interested reader a similar simplification through FME and bit recombination is presented in Appendix D.

and while \mathcal{R}_{NID} is equal to \mathcal{R}_1 and is defined by

$$\mathcal{R}_{\text{NID}} = \left\{ (R_0, R_1, R_2) : \right. \\ R_1 \leq I(QU; Y), \\ R_2 \leq I(QV; Z), \\ R_1 + R_2 \leq I(U; Y|Q) + I(QV; Z) \\ \quad - I(U; V|Q), \\ R_1 + R_2 \leq I(QU; Y) + I(V; Z|Q) \\ \quad \left. - I(U; V|Q) \right\}.$$

From the above rate regions, we observe that by taking $U = Q$, the region \mathcal{R}_{NID} contains \mathcal{R}_2 . Similarly, setting $V = Q$ makes \mathcal{R}_3 included in \mathcal{R}_{NID} , while $U = Q = V$ allows \mathcal{R}_4 to be included in \mathcal{R}_{NID} . Hence, using the ID strategy for the standard, i.e., non-compound BC, yields the same rate region as Marton's inner bound. This is explained by the fact that the apparent gain provided by choosing to decode or not the interference is recovered by an optimization of the auxiliary codewords distribution P_{QUV} .

For the compound BC setting, we observe similarly that the rate region \mathcal{R}_{ID} contains Marton's rate region [3], which we will denote in the following as \mathcal{R}_{NID} and which is given by:

$$\mathcal{R}_{\text{NID}} \triangleq \bigcup_{\substack{P_{QUVX} \\ (T_1, T_2) \in \mathbb{T}(P_{QUVX})}} \bigcap_{\substack{j \in \mathcal{J} \\ k \in \mathcal{K}}} \mathcal{T}_1^{(j,k)}(P_{QUVX}, R_{0,1}, R_{0,2}, T_1, T_2). \quad (2)$$

However, no evidence on the strict inclusion of \mathcal{R}_{NID} in \mathcal{R}_{ID} can be readily stated. To this end, we investigate in the following a compound BC for which we show that the rate region \mathcal{R}_{ID} from Theorem 3 is tight, i.e, achieves capacity, while the region \mathcal{R}_{NID} is strictly suboptimal.

B. ID is Optimal for a Class of Compound BCs

In this section, we investigate a compound BC model for which Marton's inner bound obtained through NID, is strictly sub-optimal compared to ID inner bound, in which users are allowed to decode or not the interference. For simplicity, we restrict our analysis to the case in which $J = 2$ and $K = 1$. This setting is complex enough to highlight the challenges of coding for compound settings. We first discuss a criterion for the construction of such a compound BC model and later, prove the strict optimality of ID.

1) Irrelevant Compound BC Models:

Characterizing the optimal coding for a compound BC might be rather challenging, as it can be trivial, depending on the class of BCs over which the compound setting is defined. We refer to as *irrelevant* models those of ordered BCs for which ID cannot strictly outperform NID.

Consider a compound BC with two possible BCs $X \rightarrow (Y_1, Z)$ and $X \rightarrow (Y_2, Z)$ where $Y_1 \preceq Y_2$, i.e., Y_1 is less noisy than Y_2 . Then, it follows that, whatever the auxiliary RVs $(Q, U, V) \sim P_{QUV}$:

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

TABLE I

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

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

Thus, NID yields the region

$$\mathcal{R}_{\text{NID}} = \left\{ (R_1, R_2) : \right. \\ R_1 \leq \min_{j=1,2} I(QU; Y_j), \\ R_2 \leq I(QV; Z), \\ R_1 + R_2 \leq \min_{j=1,2} I(U; Y_j|Q) + I(QV; Z) \\ \quad - I(U; V|Q), \\ R_1 + R_2 \leq \min_{j=1,2} I(QU; Y_j) + I(V; Z|Q) \\ \quad \left. - I(U; V|Q) \right\}, \\ = \left\{ (R_1, R_2) : \right. \\ R_1 \leq I(QU; Y_2), \\ R_2 \leq I(QV; Z), \\ R_1 + R_2 \leq I(U; Y_2|Q) + I(QV; Z) - I(U; V|Q), \\ R_1 + R_2 \leq I(QU; Y_2) + I(V; Z|Q) - I(U; V|Q) \left. \right\}.$$

This rate region corresponds to one obtained in the presence of only one BC channel: $X \rightarrow (Y_2, Z)$, i.e., non-compound setting. Since ID does not outperform NID for the two-user non-compound BC setting, this class of BC channels becomes irrelevant. Thus, if the possible channel outputs of user 1, Y_1 and Y_2 , are ordered, at least in the known sense of *less noisiness*, the resulting compound model does benefit from ID.

2) Compound Binary Erasure and Binary Symmetric BC:

In this section, we construct the simplest while relevant compound BC setting for which ID can be beneficial, i.e., for which the inclusion $\mathcal{R}_{\text{NID}} \subset \mathcal{R}_{\text{ID}}$ is *strict*.

In section III-B1, we showed that Y_1 and Y_2 need not be ordered in the strong sense of less-noisiness. In addition, we need to provide for some *inverse orderings*, among the BC channels of the compound setting, so as to impose a tradeoff between two orders of supersposition coding schemes. To this end, we consider a compound BC setting in which, for the BC: $X \rightarrow (Y_2, Z)$, Z is a *better* channel observation than Y_1 , while for the BC: $X \rightarrow (Y_2, Z)$ Y_2 is a *better* channel observation than Z .

To illustrate such a model, let us consider the *Binary Erasure Channel* (BEC) with erasure probability e and the *Binary Symmetric Channel* (BSC) with crossover probability p . These two channels allow for a variety of orderings between their outputs [6], depending on the pair (e, p) , as summarized in Table I. Define the compound BC with components:

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

In order to build a relevant compound BC setting, we choose Y_2 to be more capable than Y_1 , which requires: $4 p_1(1-p_1) < e_2 \leq H_2(p_1)$. Then we let Y_1 be a *physically degraded* version of Z , i.e., $p < p_1 < 0.5$, and Y_2 more capable than Z , i.e.,

$$4p(1-p) < 4p_1(1-p_1) < e_2 \leq H_2(p) \leq H_2(p_1). \quad (3)$$

Lemma 1. *The outer bound \mathcal{C}_o in (1) writes as:*

$$\mathcal{C}_o = \mathcal{C}_1 \cap \mathcal{C}_2 = \mathcal{C}_1,$$

where:

$$\mathcal{C}_1 = \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) : \begin{aligned} R_1 &\leq 1 - H_2(p_1 \star \alpha), \\ R_2 &\leq H_2(p \star \alpha) - H_2(p) \end{aligned} \right\}, \quad (4)$$

$$\mathcal{C}_2 = \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) : \begin{aligned} R_1 &\leq (1 - e_2) H_2(\alpha), \\ R_2 &\leq 1 - H_2(p \star \alpha), \\ R_1 + R_2 &\leq (1 - e_2) \end{aligned} \right\}. \quad (5)$$

Proof. Let us denote by $\bar{\mathcal{C}}_1$ the closure of \mathcal{C}_1 . We have that:

$$\bar{\mathcal{C}}_1 = \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) : \begin{aligned} R_1 &= 1 - H_2(p_1 \star \alpha), \\ R_2 &= H_2(p \star \alpha) - H_2(p) \end{aligned} \right\}.$$

In the following, we show that $\bar{\mathcal{C}}_1 \subseteq \mathcal{C}_2$. To this end, let $(R_1, R_2) \in \bar{\mathcal{C}}_1$, and let $\alpha \in [0 : 1]$ such that

$$R_1 = 1 - H_2(p_1 \star \alpha) \text{ and } R_2 = H_2(p \star \alpha) - H_2(p).$$

Let us show that $(R_1, R_2) \in \mathcal{C}_2$. We have that

$$\begin{aligned} &R_1 = 1 - H_2(p_1 \star \alpha) \\ &= 1 - H_2(p_1 \star \alpha) + H_2(p \star \alpha) - H_2(p) - R_2 \\ &= 1 - H_2(p) + H_2(p \star \alpha) - H_2(p_1 \star \alpha) - R_2 \\ &\stackrel{(a)}{\leq} 1 - H_2(p) - R_2 \\ &= (1 - H_2(p)) \left(1 - \frac{R_2}{1 - H_2(p)} \right) \\ &\stackrel{(b)}{\leq} (1 - e_2) \left(1 - \frac{R_2}{1 - H_2(p)} \right), \end{aligned} \quad (6)$$

where (a) results from that $p < p_1 < 0.5$ and (b) is stems from $e_2 \leq H_2(p)$, see (3). Next, since,

$$0 \leq 1 - \frac{R_2}{1 - H_2(p)} \leq 1,$$

then there exists $\beta \in [0 : 0.5]$ such that,

$$1 - \frac{R_2}{1 - H_2(p)} = H_2(\beta),$$

which implies that

$$R_2 = (1 - H_2(p)) (1 - H_2(\beta)).$$

With this definition of β , we have that

$$\begin{aligned} R_1 &= 1 - H_2(p_1 \star \alpha) \stackrel{(a)}{\leq} (1 - e_2) H_2(\beta) \\ R_2 &= (1 - H_2(p)) (1 - H_2(\beta)) \stackrel{(b)}{\leq} 1 - H_2(p \star \beta) \\ R_1 + R_2 &\leq 1 - H_2(p) \stackrel{(c)}{\leq} 1 - e_2, \end{aligned}$$

where (a) is a result of (7) and (b) stems from Mrs Gerber's lemma which yields

$$H_2(p \star \beta) \leq H_2(\beta)(1 - H_2(p)) + H_2(p),$$

while (c) is a consequence of (6) and $e_2 \leq H_2(p)$.

Thus, $(R_1, R_2) \in \mathcal{C}_2$ which ends the proof of the inclusion $\mathcal{C}_1 \subseteq \mathcal{C}_2$. \square

3) Evaluation of ID Inner Bound in Theorem 3:

In this section, we evaluate the inner bound \mathcal{R}_{ID} of Theorem 3, and show that for the class of compound BEC/BSC BC we construct following (3).

Theorem 5. *The rate region \mathcal{R}_{ID} of Theorem 3 is tight for the compound BEC/BSC under investigation, i.e.,*

$$\mathcal{R}_{ID} = \mathcal{C}_o = \mathcal{C}_1,$$

where \mathcal{C}_1 is as defined in Lemma 1.

Proof. The rate region \mathcal{R}_{ID} of Theorem 3 contains the rate region given by

$$\bigcup_{\substack{P_{QUVX} \in \mathcal{P} \\ (T_1, T_2) \in \mathbb{T}(P_{QUVX})}} \mathcal{T}_3^{(1)}(P_{QUVX}, T_1, T_2) \cap \mathcal{T}_4^{(2)}(P_{QUVX}, T_1, T_2),$$

where the intersection $\mathcal{T}_3^{(1)} \cap \mathcal{T}_4^{(2)}$ corresponds to decoding method (3) for the BC (Y_1, Z) and decoding method (4) for the BC (Y_2, Z) , i.e.,

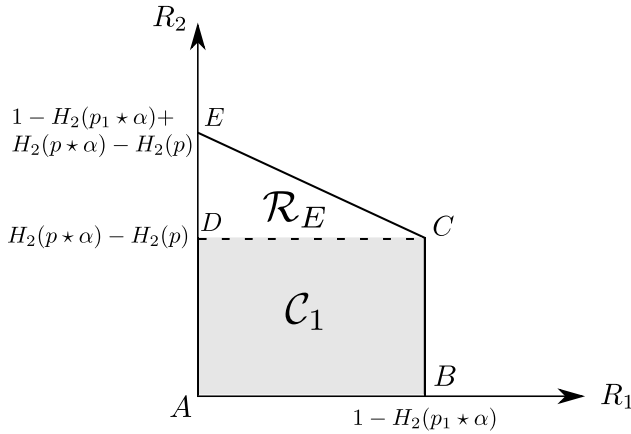
- For the BC channel (Y_1, Z) , user 1 which observes Y_1 , decodes only its intended messages, while user 2 which observes Z decodes its intended messages and interference as well
- For the BC channel (Y_2, Z) , user 1 (resp. user 2), observing Y_2 (resp. Z), decode both their intended messages as well as the interfering ones.

In Appendix D, it is shown that, after FME on the binning rates (T_1, T_2) and the rate splitting, the rate region $\mathcal{R}_{3,4}$ reduces to the set of rates satisfying:

$$\begin{aligned} R_1 &\leq I(QU; Y_1), \\ R_1 + R_2 &\leq I(QU; Y_1) + I(V; Z|QU), \\ R_1 + R_2 &\leq I(QU; Y_1) + I(UV; Y_2|Q), \\ R_1 + R_2 &\leq I(QUV; Y_2). \end{aligned}$$

Then, by letting: $V = X$, $\bar{Q} = (Q, U)$, and using the fact that Y_2 is more capable than Z , $\mathcal{R}_{3,4}$ becomes

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


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

Let $\bar{Q} \mapsto X \equiv \text{BSC}(\alpha)$, and $X \sim \text{Bern}(1/2)$. The rate region $\mathcal{R}_{3,4}$ writes then as

$$\begin{aligned} \mathcal{R}_{3,4} &= \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) : \right. \\ &R_1 \leq 1 - H_2(p_1 * \alpha), \\ &\left. R_1 + R_2 \leq 1 - H_2(p_1 * \alpha) + H_2(p * \alpha) - H_2(p) \right\}. \end{aligned} \quad (8)$$

In the following, we show that

$$\mathcal{R}_{3,4} = \mathcal{C}_1.$$

First, let us note that $\mathcal{R}_{3,4}$ is the union of two rate regions, \mathcal{C}_1 and \mathcal{R}_E , where \mathcal{R}_E is defined by

$$\begin{aligned} \mathcal{R}_E &\triangleq \bigcup_{\alpha \in [0:1]} \left\{ (R_1, R_2) : \right. \\ &R_2 \geq H_2(p * \alpha) - H_2(p), \\ &\left. R_1 + R_2 \leq 1 - H_2(p_1 * \alpha) + H_2(p * \alpha) - H_2(p) \right\}. \end{aligned}$$

Fig. 1 illustrates the two rate regions for a given $\alpha \in [0 : 1]$. The corner points C and D of \mathcal{R}_E belong to \mathcal{C}_1 . To show that the point E lies in the region \mathcal{C}_1 , note first that

$$E \triangleq (0, 1 - H_2(p_1 * \alpha) + H_2(p * \alpha) - H_2(p)).$$

Since Y_1 is physically degraded with respect to Z , i.e., $p \leq p_1$, and since α , $p * \alpha$ and $p_1 * \alpha$ are all included in the interval $[0 : 0.5]$, one can write that $-H_2(p_1 * \alpha) + H_2(p * \alpha) \leq 0$. 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 is then achievable by time sharing and convexity of the rate region \mathcal{C}_1 . Thus, since

$$\begin{aligned} \mathcal{R}_{3,4} &\subseteq \mathcal{R}_{\text{ID}} \subseteq \mathcal{C}_1, \\ \text{and } \mathcal{R}_{3,4} &= \mathcal{C}_1, \end{aligned}$$

the proof of the equality $\mathcal{R}_{\text{ID}} = \mathcal{C}_1$ is complete. \square

Remark 6. Authors in [15, Proposition 7] derived the capacity region of a specific class of three receiver broadcast channel where a common message is to be delivered to three users, observing each a channel output, Y_1, Y_2 and Z , and a private message is to be delivered only to the user with

observation Z . When the channel Z is less-noisy than Y_1 , Nair & El Gamal show that the capacity region is given by

$$\begin{aligned} R_1 &\leq \min\{I(U; Y_1), I(V; Y_2)\}, \\ R_2 &\leq I(X; Z|U), \\ R_1 + R_2 &\leq I(V; Y_2) + I(X; Z|V) \end{aligned}$$

with $U \text{---} V \text{---} X$. Though this setting is fundamentally different from the compound setting we are investigating, due to the degraded message set assumption, it turns out that when we assume that Z is less-noisy than Y_1 , the optimal scheme of [15, Proposition 7] yields an achievable rate region in our case, since, whatever Y_1 can decode, can be decoded as well by Z . Plus, we can prove that the resulting rate region is tight through a converse argument, outlined in the following,

$$\begin{aligned} nR_1 &\leq \sum_{i=1}^n I(W_1 Y_1^{i-1}; Y_{1,i}) \\ &\leq \sum_{i=1}^n I(W_1 Z^{i-1}; Y_{1,i}), \\ nR_1 &\leq \sum_{i=1}^n I(W_1 Y_{2,i+1}^n; Y_{2,i}) \\ &\leq \sum_{i=1}^n I(W_1 Y_{2,i+1}^n Z_1^{i-1}; Y_{2,i}), \\ nR_2 &\leq \sum_{i=1}^n I(X_i; Z_{1,i} | W_1 Z^{i-1}), \\ nR_1 + nR_2 &\leq \sum_{i=1}^n I(W_1 Y_{2,i+1}^n Z^{i-1}; Y_{2,i}) \\ &\quad + I(X_i; Y_{1,i} | W_1 Z^{i-1} Y_{2,i+1}^n), \end{aligned}$$

where the first inequality results from Nair's inequality for less-noisy channels [6], and where we define $V_i \triangleq (W_1, Y_{2,i+1}^n, Z^{i-1})$ and $U_i \triangleq (W_1, Z^{i-1})$. Thus, by letting $V = X$, we recover the capacity result in Theorem 5, through an alternative proof.

4) NID Inner Bound is Strictly Sub-Optimal:

In this section, we show that NID inner bound defined in (2) is strictly suboptimal for the class of compound BEC/BSC BC investigated. To this end, we first derive an outer bound on the rate region \mathcal{R}_{NID} and then, show that this outer bound is strictly included in \mathcal{R}_{ID} .

Lemma 2. The rate region \mathcal{R}_{NID} defined in (2) satisfies

$$\mathcal{R}_{\text{NID}} \subseteq \mathcal{R}_{\text{OuterNID}},$$

where $\mathcal{R}_{\text{OuterNID}}$ is the closure of the set of rate pairs satisfying

$$\mathcal{R}_{\text{OuterNID}} \triangleq \left\{ (R_1, R_2) : \right. \quad (9)$$

$$R_1 \leq \min_{j=1,2} I(\bar{Q}; Y_j), \quad (10)$$

$$\left. R_1 + R_2 \leq I(X; Z|\bar{Q}) + \min_{j=1,2} I(\bar{Q}; Y_j) \right\}, \quad (11)$$

for some joint p.m.f $P_{\bar{Q}X}$ where $\|\bar{Q}\| \leq 4$ and $X \sim \text{Bern}(1/2)$.

Proof. Let us note first that the NID rate region \mathcal{R}_{NID} defined in (2) is included in the following rate region

$$\begin{aligned} R_2 &\leq I(QV; Z), \\ R_1 &\leq \min_{j=1,2} I(QU; Y_j), \\ R_1 + R_2 &\leq I(V; Z|Q) + \min_{j=1,2} I(QU; Y_j) - I(U; V|Q), \\ R_1 + R_2 &\leq I(QV; Z) + I(U; Y_2|Q) - I(U; V|Q), \end{aligned}$$

where we have used the fact that $I(Q; Y_1) \leq I(Q; Z)$ since Y_1 is physically degraded with respect to Z . Next, this rate region is contained in the set of rate pairs satisfying

$$\begin{aligned} R_1 &\leq \min_{j=1,2} I(QU; Y_j), \\ R_1 + R_2 &\leq I(X; Z|QU) + \min_{j=1,2} I(QU; Y_j), \end{aligned}$$

by dropping some constraints and using the fact that, for all P_{QUVX} such that $(Q, U, V) \text{---} X \text{---} Z$,

$$I(V; Z|Q) - I(U; V|Q) \leq I(X; Z|QU).$$

Defining thus $\bar{Q} \triangleq (Q, U)$, yields the rate region $\mathcal{R}_{\text{OuterNID}}$. In Appendix E, we show that it suffices to evaluate $\mathcal{R}_{\text{OuterNID}}$ for all auxiliary RVs \bar{Q} that verify $\|\bar{Q}\| \leq 4$ and for $X \sim \text{Bern}(1/2)$. \square

In the following, we show that $\mathcal{R}_{\text{OuterNID}}$ defined in (11) is *strictly* included in \mathcal{C}_1 by comparing the closures of both regions. To this end, let us define the function $t(\cdot)$ on the interval $[0 : 1 - H_2(p)]$ as

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

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

$$\mathcal{C}(x) = \left\{ P_{XQ}, \quad Q \text{---} X \text{---} (Z, Y_1, Y_2) \right. \\ \left. X \sim \text{Bern}(1/2), I(X; Z|Q) \geq x \right\}. \quad (12)$$

The function $t(\cdot)$ characterizes the convex closure of the region $\bar{\mathcal{R}}_{\text{OuterNID}}$, i.e.,

$$(R_1, R_2) \in \bar{\mathcal{R}}_{\text{OuterNID}} \quad \text{iff} \quad R_1 = t(R_2).$$

Similarly, let us characterize the convex closure of $\bar{\mathcal{R}}_{\text{ID}}$, by the function $t_1(\cdot)$ defined on $[0 : 1 - H_2(p)]$ as

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

where $\mathcal{C}(x)$ is given in (12).

Lemma 3. *The function $t(\cdot)$ can be upper bounded as follows*

$$t(x) \leq \inf_{a \in [0:1]} t_a(x) \quad (13)$$

where

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

Further, $t_a(\cdot)$ can be written as follows

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

where

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

Proof. The full proof is given in Appendix F. \square

We can now compare the two regions NID and ID, through their closures, for all values of the rate R_2 .

Theorem 7. *The functions $t(\cdot)$, $t_1(\cdot)$ and $t_a(\cdot)$ satisfy the following*

1) For all $x \in [0 : H_2(p \star \alpha_0) - H_2(p)]$,

$$t(x) = \inf_{a \in [0:1]} t_a(x) = t_1(x),$$

2) For all $x \in [H_2(p \star \alpha_0) - H_2(p) : 1 - H_2(p)]$,

$$t(x) \leq \inf_{a \in [0:1]} t_a(x) < t_1(x),$$

where in 1) and 2) α_0 is the unique solution to

$$1 - H_2(p_1 \star \alpha_0) = (1 - e_2)(1 - H_2(\alpha_0)).$$

Thus, $\mathcal{R}_{\text{OuterNID}}$ and \mathcal{R}_{ID} are equal for $R_2 \in [0 : 1 - H_2(p_1 \star \alpha_0)]$ while, $\mathcal{R}_{\text{OuterNID}}$ is strictly contained in \mathcal{R}_{ID} for $R_2 \in [1 - H_2(p_1 \star \alpha_0) : 1 - H_2(p)]$.

Proof. 1) To proof part 1 of the Theorem 7, let us note first that, from Theorem 5, $t_1(\cdot)$ writes as:

$$t_1(x) = 1 - H_2(p_1 \star \alpha) |_{H_2(p \star \alpha) - H_2(p) = x}.$$

Next, let $x \in [0 : 1 - H_2(p)]$ and consider the specific choice of (Q_x, X)

$$X = Q_x \oplus U_x$$

where

$$Q_x \sim \text{Bern}(1/2) \text{ and } U_x \sim \text{Bern}(\alpha_x)$$

and $\alpha_x \in [0 : 1]$ is such that,

$$H_2(p \star \alpha_x) - H_2(p) = x.$$

It is clear that $P_{XQ_x} \in \mathcal{C}(x)$ and that

$$\begin{aligned} &\min\{I(Q_x; Y_1), I(Q_x; Y_2)\} \\ &= \min\{1 - H_2(p_1 \star \alpha), (1 - e_2)(1 - H_2(\alpha))\}. \end{aligned}$$

Thus,

$$t(x) \triangleq \sup_{P_{XQ} \in \mathcal{C}(x)} \min\{I(Q; Y_1), I(Q; Y_2)\} \quad (14)$$

$$\geq \min\{I(Q_x; Y_1), I(Q_x; Y_2)\} \quad (15)$$

$$= \min\{1 - H_2(p_1 \star \alpha_x), (1 - e_2)(1 - H_2(\alpha_x))\} |_{H_2(p \star \alpha_x) - H_2(p) = x}. \quad (16)$$

Let then α_0 be the unique solution to the equality

$$1 - H_2(p_1 \star \alpha_0) = (1 - e_2)(1 - H_2(\alpha_0)),$$

we have that, for all $x \leq H_2(p \star \alpha_0) - H_2(p)$,

$$\begin{aligned} &\min\{1 - H_2(p_1 \star \alpha_x), (1 - e_2)(1 - H_2(\alpha_x))\} \\ &= 1 - H_2(p_1 \star \alpha_x). \end{aligned} \quad (17)$$

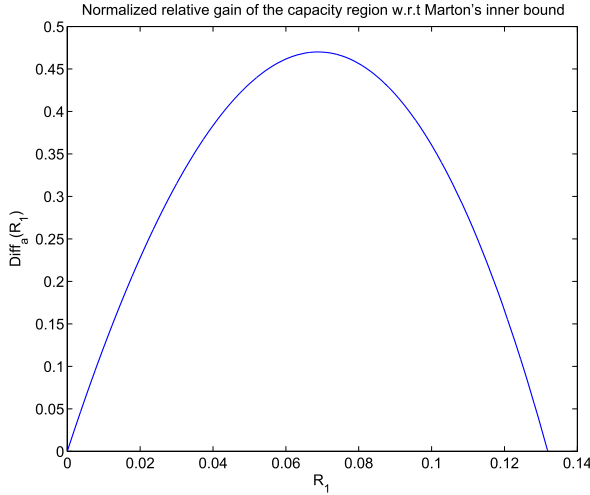


Fig. 2. $d_a(R_1)$ the normalized relative gain of ID (capacity region) with respect to NID for $a = 0.92$, $e_2 = 0.46$, $p = 0.1$ and $p_1 = 0.13$.

Combining (16) and (17) allows to write that, for all $x \leq H_2(p \star \alpha_0) - H_2(p)$,

$$t(x) \geq 1 - H_2(p_1 \star \alpha_x) = t_1(x).$$

This, along with (13), concludes the proof of $t(x) = t_1(x)$ on $[0 : H_2(p \star \alpha_0) - H_2(p)]$.

2) A closed form expression of the function $t_a(\cdot)$ for an arbitrary a brings about significant computational complexity, we thus only chose to plot it using stochastic optimization methods, i.e., Monte Carlo simulations. To this end, let $e_2 = 0.46$, $p = 0.1$ and $p_1 = 0.13$. It can be readily shown that these parameters verify (3).

In Fig. 2, we choose $a = 0.92$ and plot the normalized difference function:

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

over the interval of interest: $[0 : 1 - H_2(p_1 \star \alpha_0)]$. The function d_a being strictly positive, the claim of strict inclusion is thus shown. \square

We have investigated so far the role that evolved decoding techniques, namely *Interference Decoding*, play in the compound BC where the actual controlling the communication are ordered, however, the order is unknown to the transmitter. The decoding technique takes advantage of many possible decoding methods to alleviate the constraint of a fixed superposition order at the source, which allows the latter to apply a *symmetric* encoding rule regardless of which channels control the communication. In the sequel, we analyse a class of non-ordered compound BC to infer novel strategies when there is no specific order between users channels. In this case, we will not seek to optimize the decoding strategies but rather the encoding strategy.

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. This coding scheme is particularly beneficial when the many possible channels of the two users cannot be ordered. The main idea behind MD coding is to convey the message intended to the one user through a common description as well as a set of dedicated private descriptions. The common description is decoded whatever the actual channel realization, while the private descriptions are each decoded when a specific channel statistic is encountered. The common description –to be decoded in all cases– will be rate limited since its rate needs to be low enough to meet the decoding constraints of all possible channels, while the private descriptions encounter no such constraint.

Alike the first part of the work, we consider an elementary yet complex compound setting in which 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 inner bound for the compound BC, and the MD inner bound. We then show, for a class of compound MISO BC, and that MD coding outperforms the standard CD coding strategy. Finally, we analyse the behaviour of the obtained rate regions compared to the outer bound of Theorem 1.

A. Multiple Description (MD) Inner Bound

In the following, we derive an inner bound based on MD description coding that combines both a common description U_0 and two private descriptions U_1 and U_2 .

Theorem 8 (MD inner bound). *An inner bound on the capacity region of 2×1 compound BC is given by the set of rate pairs (R_1, R_2) satisfying:*

$$\begin{aligned} \mathcal{R}_{MD} = \bigcup_{P_{QU_0U_1U_2VX}} \{ & (R_1, R_2) : \\ & R_1 \leq I(U_0 U_1; Y_1|Q), \\ & R_1 \leq I(U_0 U_2; Y_2|Q), \\ & 2 R_1 \leq I(U_0 U_1; Y_1|Q) + I(U_0 U_2; Y_2|Q) \\ & \quad - I(U_1; U_2|QU_0), \\ & R_2 \leq I(V; Z|Q), \\ & R_1 + R_2 \leq I(U_0 U_1; Y_1|Q) + I(V; Z|Q) \\ & \quad - I(U_0 U_1; V|Q), \\ & R_1 + R_2 \leq I(U_0 U_2; Y_2|Q) + I(V; Z|Q) \\ & \quad - I(U_0 U_2; V|Q), \\ & 2R_1 + R_2 \leq I(U_0 U_1; Y_1|Q) + I(U_0 U_2; Y_2|Q) \\ & \quad + I(V; Z|Q) - I(U_0 U_1 U_2; V|Q) \\ & \quad - I(U_1; U_2|QU_0), \\ & 2 R_1 + 2 R_2 \leq I(U_0 U_1; Y_1|Q) + I(U_0 U_2; Y_2|Q) \\ & \quad + 2 I(V; Z|Q) - I(U_0 U_1; V|Q) - \\ & \quad \left. I(U_0 U_2; V|Q) - I(U_1; U_2|QU_0 V) \right\}, \end{aligned} \quad (18)$$

for some set of arbitrarily correlated RVs with joint p.m.f $P_{QU_0U_1U_2VX}$ such that the Markov chain $(Q, U_0, U_1, U_2, V) \rightarrow X \rightarrow (Y_1, Y_2, Z)$ holds.

Proof. The full proof is given in Appendix H. Yet, a brief outline of proof can be described as follows. The private descriptions U_1 and U_2 are superimposed over the common description U_0 , the three of which are binned against the interfering codeword V . The common variable Q is introduced mainly to allow for time sharing. \square

Remark 9. In this part of the work, we do not resort to rate splitting, i.e., the common codeword is used mainly for time sharing, since the resulting rate region would become rapidly intractable and thus, out of the scope of this part where we need to obtain closed form expressions of the MD coding inner bound.

Also, authors in [15, Proposition 5] resort to a codebook construction that appears to be similar to the one used in MD inner bound, and which yields the rate region

$$\begin{aligned}
R_0 &\leq \min\{I(U_0 U_1; Y_1), I(U_0 U_2; Y_2)\}, \\
2 R_0 &\leq I(U_0 U_1; Y_1) + I(U_0 U_2; Y_2) \\
&\quad - I(U_1; U_2|U_0), \\
R_0 + R_1 &\leq I(X; Z), \tag{19a} \\
R_0 + R_1 &\leq I(U_0 U_1; Y_1) + I(X; Z|U_0 U_1), \tag{19b} \\
R_0 + R_1 &\leq I(U_0 U_2; Y_2) + I(X; Z|U_0 U_2), \tag{19c} \\
2 R_0 + R_1 &\leq I(U_0 U_1; Y_1) + I(U_0 U_2; Y_2) + \\
&\quad I(X; Z|U_0 U_1 U_2) - I(U_1; U_2|U_0), \\
2 R_0 + 2 R_1 &\leq I(U_0 U_1; Y_1) + I(X; Z|U_0 U_1) \\
&\quad + I(U_0 U_2; Y_2) + I(X; Z|U_0 U_2) \\
&\quad - I(U_1; U_2|U_0), \\
2 R_0 + 2 R_2 &\leq I(U_0 U_1; Y_1) + I(U_0 U_2; Y_2) \\
&\quad + I(X; Z|U_0 U_1 U_2) + I(X; Z|U_0) \\
&\quad - I(U_1; U_2|U_0).
\end{aligned}$$

Though the two inner bounds seem similar at first sight, they are different in many respects. The two private descriptions U_1 and U_2 play different roles in both regions. In carry part of interference and allow therefore the users Y_1 and Y_2 to decode each a distinct part of interference, as opposed to our MD inner bound in which the private descriptions precodes each for interference but each is optimized for its dedicated channel, Y_1 or Y_2 . As such, it is hard to state that either region contains the other.

B. Common Description (CD) Inner Bound

Inspired by Marton's inner bound, we can derive what we call the "common description" CD coding inner bound, i.e., worst-case of Marton's inner bound for the compound BC, which consists of all rate pairs (R_1, R_2) verifying:

$$\begin{aligned}
R_1 &\leq \min_{j \in \{1,2\}} I(U; Y_j|Q), \\
R_2 &\leq I(V; Z|Q), \\
R_1 + R_2 &\leq \min_{j \in \{1,2\}} I(U; Y_j|Q) + I(V; Z|Q) \\
&\quad - I(U; V|Q),
\end{aligned} \tag{20}$$

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

Remark 10. Without time-sharing, the variable $Q \equiv \emptyset$, this inner bound imposes that, for both channels Y_1 and Y_2 in the compound setting, the decoded messages are the same. This does not allow to treat the two possible channels outputs differently, though they might observe two different interfering signals. When time-sharing is allowed, the performances of this region can be enhanced since it allows for different interference mitigation strategies across different time slots. The combination of the two techniques is denoted in literature as symbol or block expansion [17] 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 in (18) recovers the CD inner bound in (20) by setting both private descriptions to $U_1 \equiv \emptyset$ and $U_2 \equiv \emptyset$. In the following, we investigate whether MD coding can strictly outperform CD coding.

C. MD Coding Over the BC and the Compound Point-to-Point Channel

In this section, we show that for both the *single user* (non-broadcast) compound channel and the *standard two-user* (non-compound) BC, MD coding does not outperform CD coding.

As for the single user compound channel, let us assume that we have a compound model with two possible channel outputs denoted by Y_1 and Y_2 . In the following, we show that, for all joint p.m.f's $P_{Q U_0 U_1 U_2 X}$, the rate achieved by MD coding R_{MD} , is no greater than the rate achieved by CD coding R_{CD} , where R_{MD} and R_{CD} write as

$$\begin{aligned}
R_{MD} &\triangleq \max_{P_{Q U_0 U_1 U_2 X}} \min \{ I(U_0 U_1; Y_1|Q), I(U_0 U_2; Y_2|Q), \\
&\quad \frac{1}{2} [I(U_0 U_1; Y_1|Q) + I(U_0 U_2; Y_2|Q) \\
&\quad \quad - I(U_1; U_2|Q U_0)] \}, \\
R_{CD} &\triangleq \max_{P_{Q U_0 X}} \min \{ I(U_0; Y_1|Q), I(U_0; Y_2|Q) \}.
\end{aligned}$$

It is clear that $R_{CD} \leq R_{MD}$. TO prove the inverse inequality, we have that:

$$\begin{aligned}
I(U_0 U_1; Y_1|Q) &\leq I(U_0 U_1 U_2; Y_1|Q), \\
I(U_0 U_2; Y_2|Q) &\leq I(U_0 U_1 U_2; Y_2|Q),
\end{aligned}$$

and thus,

$$R_{MD} \leq \max_{P_{Q U_0 U_1 U_2 X}} \min_{j \in \{1,2\}} I(U_0 U_1 U_2; Y_j|Q).$$

By letting $U_0^* = (U_0 U_1 U_2)$, the desired inequality holds

$$R_{MD} \leq \max_{P_{Q U_0 X}} \min \{ I(U_0; Y_1|Q), I(U_0; Y_2|Q) \} = R_{CD}.$$

As for the standard two-user (non-compound) BC, i.e., $Y_1 \equiv Y_2$, let us denote \mathcal{R}_{MD} the MD inner bound in (18) and \mathcal{R}_{CD} the CD inner bound in (20). The inclusion $\mathcal{R}_{CD} \subseteq \mathcal{R}_{MD}$ is trivial. To show the converse, let us fix a joint p.m.f $P_{Q U_0 U_1 U_2 V X}$ and let us assume, for instance, that

$$\begin{aligned}
&I(U_0 U_1; Y_1|Q) - I(U_0 U_1; V|Q) \\
&\leq I(U_0 U_2; Y_1|Q) - I(U_0 U_2; V|Q).
\end{aligned}$$

It is easy to see that the choice $U^* = (U_0, U_2)$ allows us to obtain:

$$\begin{aligned} \mathcal{R}_{\text{MD}} \subseteq \bigcup_{P_{QUVX}} \{ & (R_1, R_2) : \\ & R_1 \leq I(U, Y|Q) \\ & R_2 \leq I(V; Z|Q) \\ & R_1 + R_2 \leq I(U, Y|Q) + I(V; Z|Q) - I(U; Z|Q) \}, \end{aligned}$$

which is equal to \mathcal{R}_{CD} .

Thus, MD coding does not outperform CD for setting in which channel uncertainty and interference are not coupled.

In the following, we show that for a class of compound Gaussian MISO BC, MD coding can be strictly beneficial as compared to CD coding.

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

Consider the compound Gaussian MISO BC which consists of a source equipped with 2 antennas and 2 single antenna receivers. Receiver 1 can observe one of two possible channel outputs, namely, Y_1 and Y_2 , and let Z be the channel output of the receiver 2, where, for $i = [1 : n]$ the difference outputs are given by

$$\begin{cases} y_{j,i} &= \mathbf{h}_j^T \mathbf{x}_i + n_{j,i}, \\ z_i &= \mathbf{g}^T \mathbf{x}_i + w_i, \end{cases} \quad (21)$$

for $j \in \{1, 2\}$, where: \mathbf{h}_j and \mathbf{g} are 2×1 generic real-valued 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 2×1 power limited channel input vector so that $\mathbb{E}[\mathbf{x}^T \mathbf{x}] \leq P$ and last, the noise sequences $\{n_{j,i}\}_{i \in [1:n]}$ and $\{w_i\}_{i \in [1:n]}$ are assumed to be i.i.d. draws according to a Gaussian distribution $\mathcal{N}(0, N)$.

The optimal transmit strategy for the non-ordered Gaussian MISO BC is to apply Dirty-Paper Coding (DPC) [24], [26], which is a non-linear coding technique that allows the encoder to precode against interference and suppress it at the decoder without having the decoder explicitly decode it.

In the following, in order to compare the two strategies of MD coding and CD coding, we will combine them with a DPC construction yielding thus two schemes that we will denote MD-DPC and CD-DPC. Besides, we will compare the CD to two variations of MD-DPC schemes, depending on the correlation between the private descriptions RVs we assume. In the first MD-DPC scheme, the private descriptions are *time-orthogonal*, in the sense that the encoder communicates *part of the time* a private description U_1 to cancel interference at user Y_1 , and a private description U_2 during the remaining part of the time to help user Y_2 , which annihilates the correlation cost, i.e., the mutual information $I(U_1; U_2|QU_0)$ becomes zero. In the second scheme, we consider an MD-DPC strategy where both private descriptions are transmitted across the whole time slot, however their correlation cost might not be null, i.e., the mutual information term $I(U_1; U_2|QU_0)$ is strictly positive.

A. Preliminaries and Definitions

In the sequel, we resort to DPC [24] 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:

$$\mathbf{X} = X_u \mathbf{B}_u + X_v \mathbf{B}_v, \quad (22a)$$

$$U_0 = X_u + \alpha X_v, \quad (22b)$$

$$V = X_v, \quad (22c)$$

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:

$$I(U_0; Y_j) - I(U_0; V) = \log_2 \left(\frac{h_{j,u}^2 P_u + N}{I_j(\alpha - \beta_j)^2 + N} \right), \quad (23)$$

where:

$$\beta_j = \frac{P_u h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N}, \quad (24)$$

$$I_j = \left(\frac{P_v}{P_u} \right) \frac{\left(h_{j,u}^2 P_u + N \right)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (25)$$

In the sequel, besides the common description U_0 which precodes against the interfering signal X_v , we are interested in introducing a private description U_j that will be required to precode as well against interference. If 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:

$$\mathbf{X} = (X_u + X_p) \mathbf{B}_u + X_v \mathbf{B}_v, \quad (26a)$$

$$U_0 = X_u + \alpha X_v, \quad (26b)$$

$$U_j = X_p + \alpha_j X_v, \quad (26c)$$

$$V = X_v, \quad (26d)$$

we can optimize the value of the private DPC parameter α_j to state the following result.

Lemma 4 (Optimizing the private descriptions).

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

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

$$\beta_j^x = \frac{(P_u - x) h_{j,u} h_{j,v}}{h_{j,u}^2 P_u + N}, \quad (27)$$

$$I_j^x = \left(\frac{P_v}{P_u - x} \right) \frac{\left(h_{j,u}^2 P_u + N \right)^2}{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}. \quad (28)$$

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

$$\begin{aligned} & \max_{\alpha_j \in \mathbb{R}} [I_{\alpha_j}(U_j; Y_j|U_0) - I_{\alpha_j}(U_j; V|U_0)] \\ & = I(X_p; Y_j|X_u X_v) = \frac{1}{2} \log_2 \left(\frac{h_{j,u}^2 x + N}{N} \right). \end{aligned}$$

The rest of the proof is relegated to Appendix I. \square

B. Common Description DPC (CD-DPC)

In the following, we evaluate the CD inner bound under a CD based DPC scheme for the channel model defined by (21). To this end, let us define the two following rate regions resulting from two DPC schemes:

$$\begin{aligned} \mathcal{R}_1(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v) & \triangleq \left\{ (R_1, R_2) : \right. \\ R_1 & \leq \max_{\alpha} \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left(\frac{h_{j,u}^2 P_u + N}{I_j(\alpha - \beta_j)^2 + N} \right), \\ R_2 & \leq \frac{1}{2} \log_2 \left(\frac{g_u^2 P_u + g_v^2 P_v + N}{g_u^2 P_u + N} \right) \left. \right\}, \end{aligned}$$

where β_j and I_j are given by (24) and (25).

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

$$\begin{aligned} \mathcal{R}_2(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v) & \triangleq \left\{ (R_1, R_2) : \right. \\ R_2 & \leq \frac{1}{2} \log_2 \left(\frac{g_v^2 P_v + N}{N} \right), \\ R_1 & \leq \min_{j=1,2} \frac{1}{2} \log_2 \left(\frac{h_{j,u}^2 P_u + h_{j,v}^2 P_v + N}{h_{j,v}^2 P_v + N} \right) \left. \right\}. \end{aligned}$$

Proposition 1 (CD inner bound). *An inner bound on the capacity region of the compound MISO BC defined in (21) is given by the set of rates satisfying:*

$$\begin{aligned} \mathcal{R}_{CD-MISO BC} & = \bigcup_{\substack{(P_u, P_v, \mathbf{B}_u, \mathbf{B}_v) \\ P_u + P_v \leq P \\ \|\mathbf{B}_u\| = \|\mathbf{B}_v\| = 1}} \mathcal{R}_1(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v) \cup \\ & \quad \mathcal{R}_2(\mathbf{B}_u, \mathbf{B}_v, P_u, P_v). \end{aligned} \quad (29)$$

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 (20). The rate region \mathcal{R}_1 is obtained by evaluating the corner point

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

using the following coding scheme:

$$\begin{aligned} \mathbf{X} & = X_u \mathbf{B}_u + X_v \mathbf{B}_v, \\ U & = X_u + \alpha X_v, \\ V & = X_v, \end{aligned}$$

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

As for the second rate region \mathcal{R}_2 , it results from the evaluation of the second corner point of CD with the following DPC scheme

$$\begin{aligned} \mathbf{X} & = X_u \mathbf{B}_u + X_v \mathbf{B}_v, \\ U & = X_u, \\ V & = X_v + \alpha X_u, \end{aligned}$$

in which the codeword V dirty-paper codes against the interference U ; the analysis follow in a similar manner. \square

C. MD-DPC With Orthogonal Private Descriptions

In the following, we evaluate the MD inner bound given in Theorem 8. To this end, we explore two different constructions of an MD-DPC scheme, depending on the existing correlation between the private descriptions. We will restrict our analysis to the specific corner point given by

$$\begin{aligned} R_1 & \leq \min_{j \in \{1,2\}} [I(U_0 U_j; Y_j|Q) - I(U_0 U_j; V|Q)], \\ 2 R_1 & \leq \sum_{j \in \{1,2\}} [I(U_0 U_j; Y_j|Q) - I(U_0 U_j; V|Q)] \\ & \quad - I(U_1; U_2|U_0 VQ), \\ R_2 & = I(V; Z|Q). \end{aligned} \quad (30)$$

The MD inner bound we derive in this section is based on the evaluation of (30) via a time-sharing argument [17], where each private description, U_1 (resp. U_2), is transmitted only part of the time. Let then Q be a binary valued time-sharing RV such that:

$$P_Q(1) \triangleq t \text{ and } P_Q(2) = 1 - t = \bar{t}.$$

For $Q = 1$, we let $U_2 = \emptyset$ and for $Q = 2$, we let $U_1 = \emptyset$, which annihilates the correlation cost, i.e.,

$$I(U_1; U_2|U_0 VQ) = 0.$$

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

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

where β_j^x and I_j^x are given by (27) and (28).

Proposition 2 (MD-DPC inner bound with orthogonal private descriptions). *An inner bound on the capacity region of the compound MISO BC defined in (21) is given by:*

$$\begin{aligned} \mathcal{R}_{MD-orth-MISO BC} & = \bigcup_{\substack{(x, P_u, P_v, \mathbf{B}_u, \mathbf{B}_v) \\ P_u + P_v \leq P \\ 0 \leq x \leq P_u \\ \|\mathbf{B}_u\| = \|\mathbf{B}_v\| = 1 \\ t \in [0:1]}} \mathcal{R}_u(\mathbf{B}_u, \mathbf{B}_v, x, t, P_u, P_v). \end{aligned}$$

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

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

Conversely, for $Q = 2$, let:

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

In this case, $I(U_1; U_2 | Q U_0) = 0$ since U_1 and U_2 are never transmitted in the same time slot. Hence, (30) reduces to

$$\begin{aligned}R_1 &\leq I(U_0; Y_1 | Q) - I(U_0; V | Q) + \\ &\quad t \left[I(U_1; Y_1 | U_0, Q=1) - I(U_1; V | U_0, Q=1) \right], \\ R_1 &\leq I(U_0; Y_2 | Q) - I(U_0; V | Q) + \\ &\quad \bar{t} \left[I(U_2; Y_2 | U_0, Q=2) - I(U_2; V | U_0, Q=2) \right], \\ R_2 &\leq I(V; Z | Q).\end{aligned}$$

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

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

where (a) follows similarly to the proof of Lemma 4 to maximize the private DPC parameters α_1 and α_2 . \square

D. MD-DPC With Correlated Private Descriptions

In this section, we allow the private descriptions U_1 and U_2 in (30) to be arbitrarily correlated and no longer time-orthogonal. Let the rate region \mathcal{R}_c defined by:

$$\begin{aligned}R_1 &\leq \min\{f_1(\alpha, x), f_2(\alpha, x)\}, \\ R_1 &\leq \frac{1}{2} \left[f_1(\alpha, x) + f_2(\alpha, x) + \frac{1}{2} \log_2(1 - \rho^2) \right], \\ R_2 &\leq \frac{1}{2} \log_2 \left(\frac{g_u^2 P_u + g_v^2 P_v + N}{g_u^2 P_u + N} \right),\end{aligned}$$

where:

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

and β_j^x and I_j^x are given by (27) and (28).

Proposition 3 (MD inner bound with correlated private descriptions). *An inner bound on the capacity region of the compound MISO BC is given by:*

$$\begin{aligned}\mathcal{R}_{MDcorr-MISOBC} &= \bigcup_{\substack{(x, P_u, P_v, \mathbf{B}_u, \mathbf{B}_v) \\ P_u + P_v \leq P \\ 0 \leq x \leq P_u \\ \|\mathbf{B}_u\| = \|\mathbf{B}_v\| = 1 \\ \rho \in [-1; 1] \\ \alpha \in \mathbb{R}}}\mathcal{R}_c(\mathbf{B}_u, \mathbf{B}_v, \alpha, x, \rho, P_u, P_v).\end{aligned}$$

Proof. To prove our claim, we resort to the MD coding inner bound letting, for the discrete memoryless case, the two ARVs U_1 and U_2 be arbitrarily correlated given Q , U_0 , and V .

Let us use the following coding scheme:

$$\begin{aligned}\mathbf{X} &= (X_u + X_{p_1} + X_{p_2}) \mathbf{B}_u + X_v \mathbf{B}_v, \\ U_0 &= X_u + \alpha X_v, \\ U_1 &= X_{p_1} + \alpha_1 X_v, \\ U_2 &= X_{p_2} + \alpha_2 X_v, \\ V &= X_v,\end{aligned}$$

where X_{p_1} and X_{p_2} are jointly Gaussian Random Variables, independent from all other variables, with covariance matrix:

$$K_{1,2} = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

such that $x = 2\sigma^2(1 + \rho)$. Then, the correlation cost of the two private descriptions is given by

$$I(U_1; U_2 | U_0 V Q) = I(X_{p_1}; X_{p_2}) = -\frac{1}{2} \log_2(1 - \rho^2).$$

Following similar calculations as in the proof of Lemma 4, the proof follows. \square

E. MD-DPC Strictly Outperforms CD-DPC

Let us now consider the compound Gaussian MISO BC model where the possible channels of user 1, \mathbf{h}_1 and \mathbf{h}_2 , are unit-norm orthogonal channels. Assume also that the second user's channel is quite accommodating such that \mathbf{g} is orthogonal to the *mean channel* of user 1,

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

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

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

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

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

Proof. The proof is relegated to Appendix J. \square

Remark 11. In order to derive the optimal value of η 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 η depends on the value of μ . 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 colinear to its channel. For intermediate values of μ , the weighted sum-rate is not necessarily maximized with either choices of η .

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 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:

$$R_1 \leq \frac{1}{2} \log_2 \left(\frac{P_u + 2N}{\sqrt{2N}(1 + g(x))} \right),$$

where the function $g(\cdot)$ is defined by

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

and where we have considered a time-sharing $t = \bar{t} = 0.5$. The function $g(\cdot)$ is not necessarily strictly decreasing in x for all values of η . However, it is clear that:

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

and since $0 \leq x \leq P_u$, we have

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

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, for instance, e.g. $P \geq 4N$, then for values of η close to -1 , i.e., R_2 close to second user's capacity, the gain is strictly positive and more significant.

3) *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 $\rho = 0$, then

the bound on R_1 can be written as follows:

$$\begin{aligned} R_{1,MD} &= \max_{\alpha \in \mathfrak{R}} \min \{ f_1(\alpha, x), f_2(\alpha, x) \} \\ &= \frac{1}{2} \log_2 \left(\frac{x}{4N} + 1 \right) \\ &= \frac{1}{2} \log_2 \left(\frac{P_u + 2N}{\frac{(P_u - x) 2N}{x + 2N} P(\eta) + 2N} \right) \\ &\quad - \frac{1}{2} \log_2 \left(\frac{x}{4N} + 1 \right) \\ &= \frac{1}{2} \log_2 \left(\frac{P_u + 2N}{\frac{(P_u - x) (x + 4N)}{x + 2N} P(\eta) + \frac{(x + 4N)}{2}} \right). \end{aligned}$$

Let us define,

$$h(x) \triangleq \frac{(P_u - x) (x + 4N)}{x + 2N} P(\eta) + (x + 4N),$$

we show in the following that, for some $x \in [0 : P_u]$ and $\eta \in [-1 : 1]$,

$$h(x) \leq h(0).$$

We have that

$$\begin{aligned} &h(x) - h(0) \\ &= \frac{(P_u - x) (x + 4N)}{x + 2N} P(\eta) + x - 2P(\eta) \\ &= P(\eta) \left[\frac{(P_u - x) (x + 4N)}{P_u} - 2 \right] + x, \end{aligned}$$

thus, it suffices that the two functions $\eta \mapsto P(\eta)$ and $x \mapsto \left[2 - \frac{(P_u - x) (x + 4N)}{P_u} \right]$, have values in intersecting intervals, which will be plot later for the example investigated.

F. Block Expansion and Plots

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 investigating 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 beam directions, power allocations and DPC parameters), then one could expect that the gain of multiple descriptions over one common description is still captured by the obtained bounds.

In Fig. 3, we plot the corresponding rate regions for SNR = 10 dB, and the assumptions made on the channels' structure. Fig. 3 shows four inner bounds on the capacity region of the compound Gaussian MISO BC. It can be seen that CD-DPC is strictly included in all MD-DPC inner bounds, i.e., MD-DPC with time-orthogonal descriptions and MD-DPC with arbitrary correlated private under both assumptions $\rho = 0$ (uncorrelated) and ρ variable (optimized). It can be seen that already at $\rho = 0$, MD-DPC is still strictly better than CD-DPC. It can be noticed as well that in this case, both MD-DPC bounds are equal, though this is not necessarily the case for more general settings.

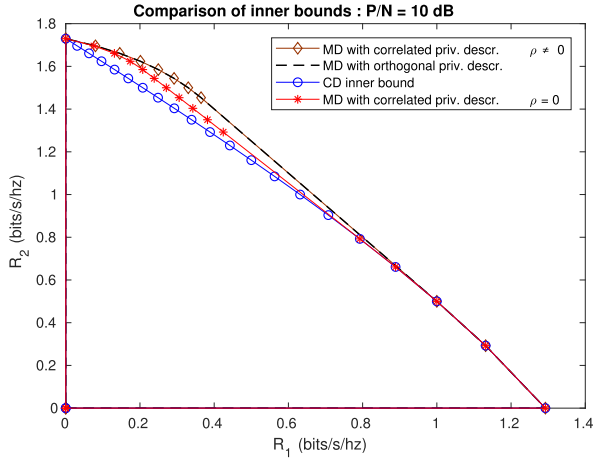


Fig. 3. Comparison of the CD-DPC and MD-DPC with uncorrelated and correlated private descriptions inner bounds with block expansion: SNR = 10 dB.

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 four rate regions.

Let us introduce the following channel matrices:

$$\begin{aligned} \mathbf{g}_{1,2} &\triangleq [\mathbf{g} \ \mathbf{h}_1 \ \mathbf{h}_2], \\ \mathbf{h}_{1,z} &\triangleq [\mathbf{h}_1 \ \mathbf{g}], \\ \mathbf{h}_{2,z} &\triangleq [\mathbf{h}_2 \ \mathbf{g}]. \end{aligned}$$

We then define the corresponding channel outputs to the channel $\mathbf{g}_{1,2}$, that has the same marginal p.d.f 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 12 (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:*

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

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

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

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

$$\begin{aligned} \mathcal{C}_{1,2} &= \bigcup_{\substack{(\mathbf{K}_u, \mathbf{K}_v) \\ \text{tr}(\mathbf{K}_u + \mathbf{K}_v) \leq P}} \left\{ (R_1, R_2) \in \mathbb{R}_+^2, \right. \\ R_1 &\leq \min_{j \in \{1,2\}} \frac{1}{2} \log_2 \left(\frac{\mathbf{h}_j^t (\mathbf{K}_u + \mathbf{K}_v) \mathbf{h}_j + N}{\mathbf{h}_j^t \mathbf{K}_v \mathbf{h}_j + N} \right), \\ R_2 &\leq \frac{1}{2} \log_2 \left(\frac{|\mathbf{g}_{1,2}^t \mathbf{K}_v \mathbf{g}_{1,2} + N \mathbf{I}_3|}{N^3} \right) \left. \right\} \end{aligned}$$

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

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

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 [27], [10]. \square

Remark 13. *The outer bound stated in Theorem 12 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 $(2 d_1 + d_2 \leq 2)$, thus the intersection of these two regions leads to the optimal DoF.*

In Fig. 4, we plot the inner and outer bound for intermediate SNR values. Although the gap to the outer bound suggests that the inner and outer regions do not meet, it is our belief that the inner bound might be tighter than the outer bound, and that better outer bounds could be derived for the compound Gaussian MISO BC.

VI. 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 rate regions. The question 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.

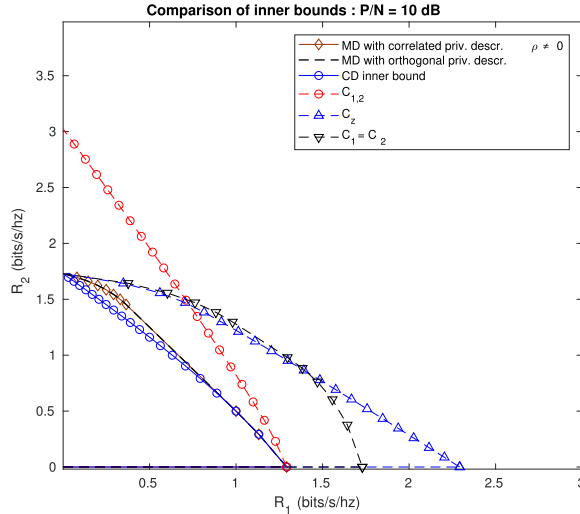


Fig. 4. Comparison of the inner bounds and the intersection of the outer bounds: SNR = 10 dB.

A. Can Multiple Descriptions or Interference Decoding Techniques Recover Each Other?

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 writes as

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

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 $\mathcal{R}_{3\text{-ARV}}$ defined by

$$\mathcal{R}_{3\text{-ARV}} \triangleq \bigcup_{\substack{P_{QU_1U_2VX} \\ (T_{1,1}, T_{1,2}, T_2) \in \\ \mathbb{T}(P_{QU_1U_2VX})}} \mathcal{M}(P_{QU_1U_2VX}, T_{1,1}, T_{1,2}, T_2), \quad (31)$$

where the rate region \mathcal{M} is defined by

$$\begin{aligned} T_2 &\leq I(V; Z|Q), \\ R_0 + T_2 &\leq I(QV; Z), \\ T_{1,1} &\leq I(U_1; Y_1|Q), \\ R_0 + T_{1,1} &\leq I(QU_1; Y_1), \\ T_{1,2} &\leq I(U_2; Y_2|Q), \\ R_0 + T_{1,2} &\leq I(QU_2; Y_2), \end{aligned}$$

and the set \mathbb{T} is given by

$$\mathbb{T}(P_{QU_1U_2VX}) \triangleq \left\{ (T_{1,1}, T_{1,2}, T_2) : \begin{aligned} T_2 &\geq R_2 \\ \min\{T_{1,1}, T_{1,2}\} &\geq R_1, \\ T_{1,1} - R_1 + T_2 - R_2 &> I(U_1; V|Q), \\ T_{1,2} - R_1 + T_2 - R_2 &> I(U_2; V|Q), \\ T_{1,1} - R_1 + T_{1,2} - R_1 &> I(U_1; U_2|Q), \\ T_{1,1} + T_{1,2} - 2R_1 + T_2 - R_2 &> I(U_1; U_2|Q) + \\ &I(U_1 U_2; V|Q) \end{aligned} \right\}.$$

Proof. The proof is relegated to Appendix K. \square

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:

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

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:

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

that is achieved by using ID, which yields:

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

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

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

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:

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

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. Conclusion

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 ID for which we derived an achievable rate region is derived by using *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 the standard two-user BC, the compound BC benefits from ID since 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, i.e, Nair & El Gamal outer bound. This outer bound is limited by the difficulty to write Csiszár & Körner's sum-identity for more than 2 users.

Surprisingly enough, ID not only outperforms 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 is 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 alleviate the uncertainty at the source.

Later, we studied an encoding technique with a more involved coding strategy, namely MD coding. The source transmits for each possible channel output, of the same user, 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 include 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 former 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 [28], 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 p.m.f of (X, Y) (resp. conditional distribution of Y given X , and marginal distribution of X).

Definition 14. For any sequence $x^n \in \mathcal{X}^n$ and any symbol $a \in \mathcal{X}$, $N(a|x^n)$ denotes the number of occurrences of a in x^n .

Definition 15. A sequence $x^n \in \mathcal{X}^n$ is called (strongly) δ -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 16. Let $x^n \in \mathcal{X}^n$. A sequence $y^n \in \mathcal{Y}^n$ is called (strongly) δ -typical (w.r.t. Y) given x^n if for all $a \in \mathcal{X}$, $b \in \mathcal{Y}$

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

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 [28]: For any sets \mathcal{X} , \mathcal{Y} , there exists a sequence $\{\delta_n\}_{n \in \mathbb{N}^*}$ such that the lemmas below hold.² From now on, typical sequences are understood with $\delta = \delta_n$. Typical sets are still denoted by $T_\delta^n(\cdot)$.

Lemma 5 ([28, Lemma 1.2.12]). *There exists a sequence $\eta_n \xrightarrow{n \rightarrow \infty} 0$ such that*

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

Lemma 6 ([28, Lemma 1.2.13]). *There exists a sequence $\eta_n \xrightarrow{n \rightarrow \infty} 0$ such that, for each $x^n \in T_\delta^n(X)$,*

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

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

²As a matter of fact, $\delta_n \rightarrow 0$ and $\sqrt{n} \delta_n \rightarrow \infty$ as $n \rightarrow \infty$.

Lemma 7 (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 8 (Joint typicality lemma [29]). *There exists a sequence $\eta_n \xrightarrow{n \rightarrow \infty} 0$ such that for all $x^n \in T_\delta^n(X)$*

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

APPENDIX B

SKETCH OF THE PROOF OF THEOREM 3

Let $(j, k) \in \mathcal{J} \times \mathcal{K}$ 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, k) .

A. Outline of Proof

The coding scheme we use is as follows:

- We consider three messages: a common message w_0 , and two private messages w_1 and w_2 ,
- We use three auxiliary RVs Q (resp. U and V) to code for the common message (resp. private messages),
- We perform random binning on the two auxiliary RV U and V which are superposed on Q ,
- The decoders resort to list decoding, which allows us to combine many decoding techniques, by intersecting different lists,
- The error probability is directly related to the list size, and thus, bounding the list size yields a bound on the average probability of error.

B. Detailed Proof

Codebook generation:

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

Fix the p.m.f's $P_Q, P_{U|Q}, P_{V|Q}$ and $P_{X|QUV}$, and let $T_1 \geq R_1$ and $T_2 \geq R_2$ be four positive real numbers.

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

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

and denote the set of all such codewords as \mathcal{C}_0 .

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

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

and set all these sequences randomly 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} sequences $v^n(l_2, w_0), l_2 \in [1 : 2^{nT_2}]$ following $P_{V|Q}^n(v^n(l_2, w_0))$ and set them randomly 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 $\mathcal{C}(W_1, W_0) \times \mathcal{C}(W_2, W_0)$ such that:

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

and then transmits: $x^n(q^n(W_0), u^n(L_1, W_0), v^n(L_2, W_0))$ which is generated randomly with p.m.f $P_{X^n|Q^n U^n V^n} = \prod_{i=1}^n P_{X|QUV}(x_i|q_i(W_0), u_i(L_1, W_0), v_i(L_2, W_0))$.

Decoding:

First, assume that no encoding error has occurred and let us note ϵ_0 the event of no error. Let then (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:

$$\mathcal{L}_1(y^n) \triangleq \left\{ (w_0, w_1) \mid \text{for } u^n(l_1, w_0) \in \mathcal{C}(w_1, w_0) \right. \\ \left. (q^n(w_0), u^n(l_1, w_0), y^n) \in T_\delta^n(QUY) \right\},$$

$$\mathcal{L}_2(y^n) \triangleq \left\{ (w_0, w_1) \mid \text{for some } w_2 \text{ s.t.} \right. \\ \left. \text{for some } v^n(l_2, w_0) \in \mathcal{C}(w_2, w_0) \right. \\ \left. \text{and some } u^n(l_1, w_0) \in \mathcal{C}(w_1, w_0), \right. \\ \left. (q^n(w_0), u^n(l_1, w_0), v^n(l_2, w_0), y^n) \right. \\ \left. \in T_\delta^n(QUVY) \right\}.$$

These two lists correspond to two different decoding strategies: non-unique decoding of the other user's message (interference), and no-interference decoding. Denote the intersection of these two lists by

$$\mathcal{L}^{(n)} \triangleq \mathcal{L}_1(y^n) \cap \mathcal{L}_2(y^n). \quad (32)$$

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 states this result.

Lemma 9. *For every $\epsilon_1 > 0$, there exists an integer N_1 , such that, for all $n \geq N_1$, the average probability of error is linked to the list size as follows:*

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

Proof. Proof is given in Appendix C. \square

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

Bounding the list size:

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

$$\mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} \leq \mathbb{E}[\|\mathcal{L}^{(n)}\|] - \mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 1\}. \quad (33)$$

On the other hand:

$$\mathbb{E}\|\mathcal{L}^{(n)}\| = \mathbb{P}\{(W_0, W_1) \in \mathcal{L}^{(n)}\} \\ + \sum_{(\hat{W}_0, \hat{W}_1) \neq (W_0, W_1)} \mathbb{P}\{(\hat{W}_0, \hat{W}_1) \in \mathcal{L}^{(n)}\}. \quad (34)$$

The next lemma provides a bound on the expected list size from the RHS of (34).

Lemma 10 (Bounding the probability of undetected errors). *For all $\epsilon_2 > 0$ and $\epsilon_3 > 0$, there exists an integer $N_2 > 0$ such that, for all $n \geq N_2$, the probability of decoding $(\hat{W}_0, \hat{W}_1) \neq (W_0, W_1)$, can be upper-bounded as follows:*

$$\sum_{\substack{(\hat{W}_0, \hat{W}_1) \neq \\ (W_0, W_1)}} \mathbb{P}\{(\hat{W}_0, \hat{W}_1) \in \mathcal{L}^{(n)}\} \leq \min\{I_1^{(n)}, I_2^{(n)}\}, \quad (35)$$

where:

$$\begin{aligned} I_1^{(n)} &\triangleq \exp_2(n [T_1 - I(U; Y|Q) + \epsilon_2]) \\ &\quad + \exp_2(n [R_0 + T_1 - I(QU; Y) + \epsilon_2]), \\ I_2^{(n)} &\triangleq \exp_2(n [T_1 - I(U; YV|Q) + \epsilon_3]) \\ &\quad + \exp_2(n [T_1 + T_2 - I(UV; Y|Q) \\ &\quad \quad - I(U; V|Q) + \epsilon_3]) \\ &\quad + \exp_2(n [R_0 + T_1 + T_2 - I(QUV; Y) \\ &\quad \quad - I(U; V|Q) + \epsilon_3]). \end{aligned}$$

Proof. Proof is given in Appendix C. \square

Hence, from (33), (34) and (35) we can write that:

$$\mathbb{P}\{\|\mathcal{L}^{(n)}\| \geq 2\} \leq \min\{I_1^{(n)}, I_2^{(n)}\}. \quad (36)$$

Then Lemma 1 and (36), imply that for n large enough:

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

Thus, provided that:

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

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 [29], the probability of encoding error can be upper bounded as n grows large enough as follows:

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

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. \square

APPENDIX C

THE PROBABILITY OF ERROR IS LINKED TO LIST SIZE

1) *Proof of Lemma 9:* Let us start by recalling:

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

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

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

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:

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

2) *Proof of Lemma 10:* Let $(\hat{W}_0, \hat{W}_1) \neq (W_0, W_1)$ be the supposedly decoded pair of messages. We have, recalling (32), that:

$$\mathbb{P}\{(\hat{W}_0, \hat{W}_1) \in \mathcal{L}^{(n)}\} \leq \min_{j=1,2} \mathbb{P}\{(\hat{W}_0, \hat{W}_1) \in \mathcal{L}_j(Y^n)\}.$$

For the first list, we have, following similar arguments of Lemma 8, that:

$$\begin{aligned} &\mathbb{P}\{(W_0, \hat{W}_1) \in \mathcal{L}_1(Y^n)\} \\ &= \mathbb{P}\{(q^n(W_0), u^n(\hat{L}_1, W_0), y^n) \in T_\delta^n(QUY) \\ &\quad \text{for some } \hat{L}_1 \in [1 : 2^{n(T_1 - R_1)}]\} \\ &\leq \sum_{\substack{\hat{L}_1 \in \\ [1 : 2^{n(T_1 - R_1)}]}} \mathbb{P}\{(q^n(W_0), u^n(\hat{L}_1, W_0), y^n) \in T_\delta^n(QUY)\} \\ &\leq \exp_2(n [T_1 - R_1 - I(U; Y|Q) + \epsilon_2]), \end{aligned}$$

and similarly, if moreover $\hat{W}_0 \neq W_0$,

$$\begin{aligned} &\mathbb{P}\{(\hat{W}_0, \hat{W}_1) \in \mathcal{L}_1(Y^n)\} \\ &\leq \exp_2(n [T_1 - R_1 - I(QU; Y) + \epsilon_2]). \end{aligned}$$

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

1) If $\hat{W}_0 = W_0$, $\hat{W}_1 \neq W_1$ and $\hat{L}_2 = L_2$ which implies $\hat{W}_2 = W_2$:

$$\begin{aligned} &\mathbb{P}\{(Q^n(W_0), U^n(\hat{L}_1, W_0), V^n(L_2, W_0), Y^n) \in \\ &\quad T_\delta^n(QUVY) \text{ for } \hat{L}_1 \in [1 : 2^{n(T_1 - R_1)}]\} \\ &\leq \exp_2(n [T_1 - R_1 + H(QUVY) - H(Q) \\ &\quad - H(U|Q) - H(YV|Q) + \epsilon_3]) \\ &= \exp_2(n [T_1 - R_1 - I(U; YV|Q) + \epsilon_3]), \end{aligned}$$

where we used the fact that, since $\hat{W}_1 \neq W_1$, then $U^n(\hat{L}_1, W_0)$ and $V^n(L_2, W_0)$ are independent conditionally on $Q^n(W_0)$.

2) If $\hat{W}_0 = W_0$, $\hat{W}_1 \neq W_1$, and $\hat{L}_2 \neq L_2$ then:

$$\begin{aligned} &\mathbb{P}\{(Q^n(W_0), U^n(\hat{L}_1, W_0), V^n(\hat{L}_2, W_0), Y^n) \in \\ &\quad T_\delta^n(QUVY) \text{ for } \hat{L}_1 \in [1 : 2^{n(T_1 - R_1)}]\} \\ &\leq \exp_2(n [T_1 - R_1 + H(QUVY) - H(Q) \\ &\quad - H(U|Q) - H(V|Q) - H(Y|Q) + \epsilon_3]) \\ &= \exp_2(n [T_1 - R_1 - I(UV; Y|Q) \\ &\quad - I(U; V|Q) + \epsilon_3]). \end{aligned}$$

3) Finally, if $\hat{W}_0 \neq W_0$, then whatever \hat{L}_1 and \hat{L}_2 :

$$\begin{aligned} &\mathbb{P}\{(Q^n(\hat{W}_0), U^n(\hat{L}_1, \hat{W}_0), V^n(\hat{L}_2, \hat{W}_0), Y^n) \in \\ &\quad T_\delta^n(QUVY)\} \\ &\leq \exp_2(n [H(QUVY) - H(Q) - H(U|Q) \\ &\quad - H(V|Q) - H(Y) + \epsilon_3]) \\ &= \exp_2(n [-I(QUV; Y) - I(U; V|Q) + \epsilon_3]). \end{aligned}$$

This ends the proof of Lemma 2. \square

APPENDIX D

PROOF OF ACHIEVABILITY OF THE CAPACITY

From Theorem 3, we can see that the region $\mathcal{R}_{\text{SNID}}$ verifies:

$$\bigcup_{\substack{P_{QUVX} \in \mathcal{P} \\ (T_1, T_2) \in \\ \mathbb{T}(P_{QUVX})}} \mathcal{T}_3^{(1)}(P_{QUVX}, T_1, T_2) \cap \mathcal{T}_4^{(2)}(P_{QUVX}, T_1, T_2) \subseteq \mathcal{R}_{\text{ID}}.$$

In this section, we evaluate the region given by

$$\mathcal{R}_{\text{ID}}^* \triangleq \bigcup_{\substack{P_{QUVX} \in \mathcal{P} \\ (T_1, T_2) \in \\ \mathbb{T}(P_{QUVX})}} \mathcal{T}_3^{(1)}(P_{QUVX}, T_1, T_2) \cap \mathcal{T}_4^{(2)}(P_{QUVX}, T_1, T_2)$$

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

$$\begin{aligned} T_2 &\leq I(V; ZU|Q), \\ T_1 + T_2 &\leq I(UV; Z|Q) + I(U; V|Q), \\ R_0 + T_1 + T_2 &\leq I(QUV; Z) + I(U; V|Q), \\ T_1 &\leq I(U; Y_2 V|Q), \\ T_1 + T_2 &\leq I(UV; Y_2|Q) + I(U; V|Q), \\ R_0 + T_1 + T_2 &\leq I(QUV; Y_2) + I(U; V|Q), \\ T_1 &\leq I(U; Y_1|Q), \\ R_0 + T_1 &\leq I(QU; Y_1), \\ T_1 &\geq R_1, T_2 \geq R_2, \\ T_1 + T_2 &> R_1 + R_2 + I(U; V|Q). \end{aligned}$$

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

$$\begin{aligned} T_2 &\leq I(V; ZU|Q), \\ T_1 &\leq \min\{I(U; Y_1|Q), I(U; Y_2 V|Q)\}, \\ R_0 + T_1 &\leq I(QU; Y_1), \\ T_1 + T_2 &\leq I(UV; Y_2|Q) + I(U; V|Q), \\ R_0 + T_1 + T_2 &\leq I(QUV; Y_2) + I(U; V|Q). \end{aligned}$$

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

$$\begin{aligned} R_2 &\leq I(V; ZU|Q), \\ R_1 &\leq \min\{I(U; Y_1|Q), I(U; Y_2 V|Q)\}, \\ R_0 + R_1 &\leq I(QU; Y_1), \\ R_1 + R_2 &\leq I(UV; Y_2|Q), \\ R_1 + R_2 &\leq I(V; Z|UQ) \\ &\quad + \min\{I(U; Y_1|Q), I(U; Y_2 V|Q)\}, \\ R_0 + R_1 + R_2 &\leq I(QUV; Y_2), \\ R_0 + R_1 + R_2 &\leq I(V; Z|UQ) + I(QU; Y_1). \end{aligned}$$

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

$$\begin{aligned} R_0 &= R_0^* + R_{01}^* + R_{02}^*, \\ R_1 &= R_1^* - R_{01}^* \geq 0, \\ R_2 &= R_2^* - R_{02}^* \geq 0, \\ R_{01}^* &\geq 0, R_{02}^* \geq 0. \end{aligned}$$

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:

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

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

$$\begin{aligned} R_0^* + R_1^* &\leq I(QU; Y_1), \\ R_0^* + R_1^* + R_2^* &\leq I(QU; Y_1) + I(UV; Y_2|Q), \\ R_0^* + R_1^* + R_2^* &\leq I(QU; Y_1) + I(V; ZU|Q), \quad (37) \\ R_0^* + R_1^* + R_2^* &\leq I(QU; Y_1) + I(V; Z|UQ) \\ &\quad + \min\{I(U; Y_1|Q), I(U; Y_2 V|Q)\}, \quad (38) \\ R_0^* + R_1^* + R_2^* &\leq I(QUV; Y_2), \\ R_0^* + R_1^* + R_2^* &\leq I(V; Z|UQ) + I(QU; Y_1). \quad (39) \end{aligned}$$

We clearly notice that the constraints: (37) and (38) are implied by (39), thus, the resulting region $\mathcal{R}_{\text{ID}}^*$ is defined by the following constraints:

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

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

APPENDIX E
CARDINALITY BOUNDS

Consider a pair of RVs (Q, X) following the joint p.m.f P_{QX} . Since the input is binary, let the four continuous functions on $P_{X|Q}$:

$$\begin{aligned} f_1(P_{X|Q}(0|q)) &= P_{X|Q}(0|q), \\ f_2(P_{X|Q}(0|q)) &= H(Z|Q=q) = H_2(p * P_{X|Q}(0|q)), \\ f_3(P_{X|Q}(0|q)) &= H(Y_1|Q=q) = H_2(p_1 * P_{X|Q}(0|q)), \\ f_4(P_{X|Q}(0|q)) &= H(X|Q=q) = H_2(P_{X|Q}(0|q)). \end{aligned}$$

By the usual consequence of Fenchel-Eggleston-Caratheodory theorem [28], we can construct an auxiliary RV Q' such that:

$$\begin{aligned} \sum_q P_Q(q) P_{X|Q}(0|q) &= \sum_{q'} P_{Q'}(q') P_{X|Q}(0|q') = P_X(0), \\ H(Z|Q) &= H(Z|Q'), \\ H(Y_1|Q) &= H(Y_1|Q'), \\ H(X|Q) &= H(X|Q'), \\ \|Q'\| &\leq 4. \end{aligned}$$

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

$$\begin{aligned} I(X; Z|Q) &= H(Z|Q) - H(Z|X) \\ &= H(Z|Q') - H_2(p) \\ &= I(X; Z|Q'), \\ I(Q; Y_1) &= H(Y_1) - H(Y_1|Q) \\ &= H_2(p_1 \star P_X(0)) - H(Y_1|Q') \\ &= I(Q'; Y_1), \\ I(Q; Y_2) &= (1 - e) (H(X) - H(X|Q)) \\ &= (1 - e) (H_2(P_X(0)) - H(X|Q')) \\ &= I(Q'; Y_2). \end{aligned}$$

Optimality of uniform input

In [6] the c -symmetric BC is defined as 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 channels, thus $X \sim \text{Bern}(1/2)$ is optimal for the compound BC as well. \square

APPENDIX F

PROOF OF PROPOSITION 3

We follow the method in [30] to write:

$$\begin{aligned} t(x) &\triangleq \sup_{p_{XQ} \in \mathcal{C}(x)} \min \{I(Q; Y_1), I(Q; Y_2)\} \\ &= \sup_{p_{XQ} \in \mathcal{C}(x)} \min_{a \in [0:1]} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)] \\ &\leq \min_{a \in [0:1]} \sup_{p_{XQ} \in \mathcal{C}(x)} [a I(Q; Y_1) + \bar{a} I(Q; Y_2)], \end{aligned}$$

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

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

Notice that:

- The case $a = 1$ was already studied in [30] and it was shown that:

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

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 [30] by finding out that:

$$\begin{aligned} t_0(x) &= \inf_{\lambda \in \mathcal{R}^+} [F_0(\lambda) - \lambda x] \\ &= (1 - e_2) \left(1 - \frac{x}{1 - H_2(p)} \right), \end{aligned}$$

where:

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

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

$$t_a(x) \leq a \sup_{\mathcal{C}(x)} I(Q; Y_1) + \bar{a} \sup_{\mathcal{C}(x)} I(Q; Y_2) \quad (40)$$

$$= a t_1(x) + \bar{a} t_0(x) \quad (41)$$

$$\geq t_1(x), \quad (42)$$

where (42) 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 (40) on t_a since it is rather loose, and we will hence bound more tightly the function t_a .

Proposition 5. *The function t_a satisfies the following properties:*

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

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

- (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. \square

The next result allows to transform the optimization of a rate region into optimizing one function denoted as $F_a(\lambda)$.

The following conclusions can be drawn:

- (a) The constraint in (12) can be transformed into:

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

- (b) We have that:

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

where

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

where (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 [31]. \square

APPENDIX G
PROOF OF PROPOSITION 5

Recall that:

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

We want to show that:

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

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

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. 1) We have that:

$$\mathcal{C}(x) = \left\{ P_{XQ} \in \mathcal{P} : Q \circlearrowleft X \circlearrowleft (Y, Z_1, Z_2), \right. \\ \left. X \sim \text{Bern}(1/2), I(X; Z|Q) = x \right\}.$$

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.

2) 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 $i \in \{1, 2\}$,

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

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

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

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

- $X \sim \text{Bern}(1/2)$.
- $Q \circlearrowleft X \circlearrowleft (Y, Z_1, Z_2)$ is a valid Markov chain.
- And the following equalities hold:

$$\begin{aligned} I(X; Z|Q) &= \alpha I(X_1; Z|Q_1) + (1 - \alpha) I(X_2; Z|Q_2) \\ &= \alpha x_1 + (1 - \alpha)x_2 = x. \end{aligned}$$

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

$$\begin{aligned} &\alpha t_a(x_1) + (1 - \alpha)t_a(x_2) \\ &= \alpha (a I(Q_1; Y_1) + \bar{a} I(Q_1; Y_2)) \\ &\quad + (1 - \alpha) (a I(Q_2; Y_1) + \bar{a} I(Q_2; Y_2)) \\ &= a I(Q_T; Y_1|T) + (1 - a) I(Q_T; Y_2|T) \\ &\leq a I(T Q_T; Y_1) + (1 - a) I(T Q_T; Y_2) \\ &= a I(Q; Y_1) + (1 - a) I(Q; Y_2) \\ &\leq \max_{P_{QX} \in \mathcal{C}(x)} [a I(Q; Y_1) + (1 - a) I(Q; Y_2)] \\ &= t_a(x), \end{aligned}$$

which concludes the proof of concavity.

3) This property follows from the concavity of t_a .

4) Monotony: Since t_a is concave, we have that:

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

Since,

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

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

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

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 in (8). Let W_1 be the message decoded by user 1, and let W_2 be the message decoded by user 2, and 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} sequences $u_0^n(l_1)$ where $l_1 \in [1 : 2^{nT_1}]$ 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 randomly to 2^{nR_1} bins $\mathcal{C}_0(w_1)$, each indexed with $w_1 \in [1 : 2^{nR_1}]$.

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

For each $u_0^n(l_1)$ where $l_1 \in [1 : 2^{nT_1}]$, generate $2^{n\hat{R}_j}$ sequences $u_j^n(s_j, l_1)$ with $s_j \in [1 : 2^{n\hat{R}_j}]$ 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 pair 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 pair of indices (s_1, s_2) such that

$$\begin{aligned} &(u_0^n(l_1), u_1^n(s_1, l_1), u_2^n(s_2, l_1), v^n(l_2)) \\ &\in T_\delta^n(U_0 U_1 U_2 V). \end{aligned}$$

It then transmits an sequence $x^n(u_0^n(l_1), u_1^n(s_1, l_1), u_2^n(s_2, l_1), v^n(l_2))$ which is generated via a stochastic mapping.

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

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

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:

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

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

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

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

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

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

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

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

$$\begin{cases} T_1 + \hat{R}_1 & \leq I(U_0 U_1; Y_j), \\ T_1 + \hat{R}_2 & \leq I(U_0 U_2; Y_j), \\ T_2 & \leq I(V; Z). \end{cases}$$

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

$$\begin{aligned} \hat{R}_1 &\geq 0, \\ \hat{R}_2 &\geq 0, \\ T_1 &\geq R_1, \\ T_2 &\geq R_2, \end{aligned}$$

the region given in (8) follows immediately. \square

APPENDIX I PROOF OF LEMMA 4

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

$$\begin{aligned} \mathbf{X} &= (X_u + X_p)\mathbf{B}_u + X_v\mathbf{B}_v \\ U_0 &= X_u + \alpha X_v, \\ U_1 &= X_p + \alpha_1 X_v, \\ V &= X_v, \end{aligned}$$

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 RVs 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: $R_{0,1} = I(U_0 U_1; Y) - I(U_0 U_1; V)$. Some algebraic manipulations yield

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

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

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

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 $\alpha_1 = \beta_1^o$ and $\alpha = \alpha^o$. 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 $\alpha_1 - \alpha_1^* = \alpha - \alpha^* = 0$.

Upon optimizing the polynomial $P(\alpha, \alpha_1)$ over α_1 , the resulting rate is given by the rather simple expression:

$$R_{0,1} = \frac{1}{2} \log_2 \left(\frac{h_u^2 P_u + N}{I_j^x \frac{N}{h_u^2 x + N} (\alpha - \beta^x)^2 + N} \right), \quad (43)$$

where I_j^x is given by (28). It can be readily checked that this expression corresponds to the following

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

where

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

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 $I(U_1; Y|U_0) - I(U_1; V|U_0)$. \square

APPENDIX J OPTIMIZATION OF COMMON DESCRIPTION INNER BOUND:

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

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

Since \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 both monotonic, one operand is increasing in $h_{1,u}^2$, while the other is decreasing in $h_{1,u}^2$. Thus, the maxmin point corresponds to the equality point. Which by simple algebraic calculations leads to the condition:

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

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

$$R_1 \leq \frac{1}{2} \log_2 \left(\frac{P_u + P_v + 2N}{P_v + 2N} \right). \quad (44)$$

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 (44), that $|h_{1,u}| = 1/\sqrt{2}$. Thus transmitting the first user's signal in the mean channel direction is an admissible optimal solution. Later in the proof, we show that this second corner point \mathcal{R}_2 is dominated by the first corner point of the CD inner bound \mathcal{R}_1 which is investigated below. 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 very involved and is not of central importance.

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

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

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*:

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

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:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \max_{\alpha} \min_{j \in \{1,2\}} \\ &\log_2 \left(\frac{P_u + 2N}{\frac{P_v}{P_u} \frac{(P_u + 2N)^2}{(P + 2N)} (\alpha - \alpha_j)^2 + 2N} \right) \\ &= \frac{1}{2} \max_{\alpha} \log_2 \left(\frac{P_u + 2N}{\frac{P_v}{P_u} \frac{(P_u + 2N)^2}{(P + 2N)} (|\alpha| + |\alpha_j|)^2 + 2N} \right) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2} \log_2 \left(\frac{P_u + 2N}{2N + P_v \frac{P_u}{P + 2N}} \right) \\ &= \frac{1}{2} \log_2 \left(\frac{P_u + P_v + 2N}{P_v + 2N} \right), \end{aligned}$$

where $\alpha_1 = -\alpha_2 = \frac{P_u}{P_u + 2N}$. It turns out then, that the optimization over the DPC parameter α yields $\alpha = 0$, meaning that the two interference signals to be precoded are orthogonal, and thus, precoding against both of them is impossible. 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:

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

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 α_1^* and α_2^* such that:

$$\begin{aligned} \alpha_1^* - \alpha_1 &= \frac{P_u}{P_u + 2N} \frac{(-\cos(\theta_v) + \sin(\theta_v)) c(\theta_v)}{s(\theta_v) + c(\theta_v)}, \\ \alpha_2^* - \alpha_1 &= \frac{P_u}{P_u + 2N} \frac{(\cos(\theta_v) - \sin(\theta_v)) c(\theta_v)}{s(\theta_v) - c(\theta_v)}, \end{aligned}$$

where $s(\theta_v)$ and $c(\theta_v)$ are given by

$$\begin{aligned} s(\theta_v) &= \sqrt{P_u + 2N + 2\sin^2(\theta_v)P_v}, \\ c(\theta_v) &= \sqrt{P_u + 2N + 2\cos^2(\theta_v)P_v}. \end{aligned}$$

The root that yields the greater rate is α_1^* . Then, we can rewrite with the following transformation $y = \sin(2\theta_v)$ that:

$$\begin{aligned} &(\alpha_1^* - \alpha_1)^2 \\ &= \frac{P_u^2}{(P_u + 2N)^2} \frac{(1-y) c^2(\theta_v)}{(s(\theta_v) + c(\theta_v))^2} \\ &= \frac{P_u^2}{(P_u + 2N)^2} \frac{(1-y) c^2(\theta_v)}{P + 2N + \sqrt{(P + 2N)^2 + (y^2 - 1)P_v^2}}. \end{aligned}$$

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:

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

where $P(y)$ is defined by

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

□

APPENDIX K

PROOF OF ACHIEVABILITY OF \mathcal{R}_{3-ARV}

Codebook generation:

We fix a p.m.f $P_{QU_1U_2VX}$. 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 p.m.f: $\prod_{i=1}^n P_Q(q_i(w_0))$. For each w_0 , generate 2^{nT_2} sequences $v^n(l_2, w_0)$ following the p.m.f $\prod_{i=1}^n P_{V|Q}(v_i(l_2, w_0)|q_i(w_0))$ 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:

$$\left(q^n(w_0), u_1^n(l_{1,1}, w_0), u_2^n(l_{1,2}, w_0), v^n(l_2, w_0) \right) \\ \in T_\delta^n(QU_1U_2V). \quad (45)$$

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 of \mathcal{M} . □

ACKNOWLEDGMENT

The authors are very grateful to the Associate Editor, Prof. Daniela Tuninetti, for her careful and meticulous reading of the paper which has deeply contributed to improve the presentation and quality of the present work.

REFERENCES

- [1] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
- [2] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.
- [3] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [4] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 2205–2209.
- [5] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 166–169, Mar. 1979.
- [6] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4207–4214, Sep. 2010.
- [7] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Problems Inf. Transmiss.*, vol. 16, no. 1, pp. 24–34, 1980.
- [8] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [9] A. A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Problems Inf. Transmiss.*, vol. 16, no. 1, pp. 3–23, 1980.
- [10] H.-F. Chong and Y.-C. Liang, "The capacity region of a class of two-user degraded compound broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 932–936.
- [11] H.-F. Chong and Y.-C. Liang, "The capacity region of a new class of K-receiver degraded compound broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1957–1961.
- [12] F. Baccelli, A. El Gamal, and D. N. C. Tse, "Interference networks with point-to-point codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2582–2596, May 2011.
- [13] B. Bandemer, A. El Gamal, and Y.-H. Kim, "Simultaneous nonunique decoding is rate-optimal," in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2012, pp. 9–16.
- [14] S. S. Bidokhti and V. M. Prabhakaran, "Is non-unique decoding necessary?" *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2594–2610, May 2014.
- [15] C. Nair and A. El Gamal, "The capacity region of a class of three-receiver broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.
- [16] A. Padakandla and S. S. Pradhan, "A new coding theorem for three user discrete memoryless broadcast channel," *CoRR*, vol. abs/1207.3146, pp. 1–23, Jul. 2012.
- [17] H. Weingarten, S. Shamai (Shitz), and G. Kramer, "On the compound MIMO broadcast channel," in *Proc. Annu. Inf. Theory Appl. Workshop UCSD*, Palo Alto, CA, USA, Jan. 2007, pp. 1–7.
- [18] T. Gou, S. A. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3286–3308, Jun. 2011.
- [19] M. A. Maddah-Ali, "On the degrees of freedom of the compound MISO broadcast channels with finite states," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2273–2277.
- [20] S. A. Jafar, *Interference Alignment: A New Look at Signal Dimensions in a Communication Network*. Boston, MA, USA: Now, 2011. [Online]. Available: <http://books.google.fr/books?id=GfwB7ItK4esC>
- [21] C. Huang, S. A. Jafar, S. Shamai (Shitz), and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849–857, Feb. 2012.
- [22] R. Tandon, S. A. Jafar, S. Shamai (Shitz), and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, Jul. 2013.
- [23] P. Piantanida and S. Shamai (Shitz), "On the capacity of compound state-dependent channels with states known at the transmitter," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 624–628.
- [24] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
- [25] M. Benammar and P. Piantanida, "On the role of interference decoding in compound broadcast channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2013, pp. 1–5.
- [26] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-49, no. 7, pp. 1691–1706, Jul. 2003.
- [27] H. Weingarten, T. Liu, S. Shamai (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, Nov. 2009.
- [28] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Akadémiai Kiado, 1982.
- [29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.
- [30] H. S. Witsenhausen and A. D. Wyner, "A conditional entropy bound for a pair of discrete random variables," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 493–501, Sep. 1975.
- [31] H. G. Eggleston, *Convexity*. Cambridge, U.K.: Cambridge Univ. Press, 1958.

Meryem Benammar (S'10–M'16) received the B.S. and M.Sc. in electrical engineering from CentraleSuplec in 2011, and the Ph.D. in Electrical engineering from CentraleSuplec in 2014. From 2014 to 2017, she was a researcher in Huawei's Mathematical and Algorithmic Sciences Lab in Paris, France. Since 2017, she is an associate professor in Information Theory with the Department of Optronics, Electronics and Signal Processing in ISAE-Suapéro, Toulouse, France. Her research interests include applications of information theory to many fields, including both source and channel coding for multi-terminal networks, state dependent transmissions, physical layer security and, since recently, statistical learning. She was awarded Eiffels excellence scholarship for engineering studies in France, Eleuthre Mascarts medal for outstanding records in engineering and masters studies, as well as a grant from the French ministry of higher education and research for PhD studies.

Pablo Piantanida (S'04–M'08–SM'16) received both B.Sc. in Electrical Engineering and the M.Sc. (with honors) from the University of Buenos Aires (Argentina) in 2003, and the Ph.D. from Université Paris-Sud (Orsay, France) in 2007. Since October 2007 he has joined the Laboratoire des Signaux et Systèmes (L2S), at CentraleSupélec together with CNRS (UMR 8506) and Université Paris-Sud, as an Associate Professor of Network Information Theory. He is currently associated with the Montreal Institute for Learning Algorithms (Mila) at Université de Montréal. He is an IEEE Senior Member and General Co-Chair of the 2019 IEEE International Symposium on Information Theory (ISIT). His research interests lie broadly in information theory and its interactions with other fields, including multi-terminal information theory, Shannon theory, machine learning and representation learning, statistical inference, cooperative communications, communication mechanisms for security and privacy.

Shlomo Shamai (Shitz) (S'80–M'82–SM'88–F'94–LF'18) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israeli Institute of Technology, in 1975, 1981, and 1986, respectively.

During 1975–1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, an URSI Fellow, a member of the Israeli Academy of Sciences and Humanities and a foreign member of the US National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award, the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering and the 2017 IEEE Richard W. Hamming Medal. He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 and 2015 European Commission FP7, Network of Excellence in Wireless COMMunications (NEWCOM++, NEWCOM#) Best Paper Awards, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, the 2014 EURASIP Best Paper Award (for the EURASIP Journal on Wireless Communications and Networking), and the 2015 IEEE Communications Society Best Tutorial Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors of the Information Theory Society. He has also served on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY and on the IEEE Information Theory Society Nominations and Appointments Committee.