

# Capacity Bounds for the Gaussian Interference Channel

Abolfazl S. Motahari, *Student Member, IEEE*, and Amir K. Khandani, *Member, IEEE*

Coding & Signal Transmission Laboratory ([www.cst.uwaterloo.ca](http://www.cst.uwaterloo.ca))

{abolfazl,khandani}@cst.uwaterloo.ca

## Abstract

The capacity region of the two-user Gaussian Interference Channel (IC) is studied. Three classes of channels are considered: weak, one-sided, and mixed Gaussian IC. For the weak Gaussian IC, a new outer bound on the capacity region is obtained that outperforms previously known outer bounds. The sum capacity for a certain range of channel parameters is derived. For this range, it is proved that using Gaussian codebooks and treating interference as noise is optimal. It is shown that when Gaussian codebooks are used, the full Han-Kobayashi achievable rate region can be obtained by using the naive Han-Kobayashi achievable scheme over three frequency bands (equivalently, three subspaces). For the one-sided Gaussian IC, an alternative proof for the Sato's outer bound is presented. We derive the full Han-Kobayashi achievable rate region when Gaussian codebooks are utilized. For the mixed Gaussian IC, a new outer bound is obtained that outperforms previously known outer bounds. For this case, the sum capacity for the entire range of channel parameters is derived. It is proved that the full Han-Kobayashi achievable rate region using Gaussian codebooks is equivalent to that of the one-sided Gaussian IC for a particular range of channel parameters.

## Index Terms

Gaussian interference channels, capacity region, sum capacity, convex regions.

## I. INTRODUCTION

ONE of the fundamental problems in Information Theory, originating from [1], is the full characterization of the capacity region of the interference channel (IC). The simplest form of IC is the two-user case in which two transmitters aim to convey independent messages to their corresponding receivers through a common channel. Despite some special cases, such as very strong and strong interference, where the exact capacity region has been derived [2], [3], the characterization of the capacity region for the general case is still an open problem.

A limiting expression for the capacity region is obtained in [4] (see also [5]). Unfortunately, due to excessive computational complexity, this type of expression does not result in a tractable approach to fully characterize the capacity region. To show the weakness of the limiting expression, Cheng and Verdú have shown that for the Gaussian Multiple Access Channel (MAC), which can be considered as a special case of the Gaussian IC, the limiting expression fails to fully characterize the capacity region by relying only on Gaussian distributions [6]. However, there is a point on the boundary of the capacity region of

<sup>1</sup>An earlier version of this work containing all the results is reported in *Library and Archives Canada Technical Report UW-ECE 2007-26, Aug. 2007* (see [http://www.cst.uwaterloo.ca/pub\\_tech\\_rep.html](http://www.cst.uwaterloo.ca/pub_tech_rep.html) for details).

the MAC that can be obtained directly from the limiting expression. This point is achievable by using simple scheme of Frequency/Time Division (FD/TD).

The computational complexity inherent to the limiting expression is due to the fact that the corresponding encoding and decoding strategies are of the simplest possible form. The encoding strategy is based on mapping data to a codebook constructed from a unique probability density and the decoding strategy is to treat the interference as noise. In contrast, using more sophisticated encoders and decoders may result in collapsing the limiting expression into a single letter formula for the capacity region. As an evidence, it is known that the joint typical decoder for the MAC achieves the capacity region [7]. Moreover, there are some special cases, such as strong IC, where the exact characterization of the capacity region has been derived [2], [3] where decoding the interference is the key idea behind this result.

In their pioneering work, Han and Kobayashi (HK) proposed a coding strategy in which the receivers are allowed to decode part of the interference as well as their own data [8]. The HK achievable region is still the best inner bound for the capacity region. Specifically, in their scheme, the message of each user is split into two independent parts: the common part and the private part. The common part is encoded such that both users can decode it. The private part, on the other hand, can be decoded only by the intended receiver and the other receiver treats it as noise. In summary, the HK achievable region is the intersection of the capacity regions of two three-user MACs, projected on a two-dimensional subspace.

The HK scheme can be directly applied to the Gaussian IC. Nonetheless, there are two sources of difficulties in characterizing the full HK achievable rate region. First, the optimal distributions are unknown. Second, even if we confine the distributions to be Gaussian, computation of the full HK region under Gaussian distribution is still difficult due to numerous degrees of freedom involved in the problem. The main reason behind this complexity is the computation of the cardinality of the time-sharing parameter.

Recently, reference [9], Chong *et al.* has presented a simpler expression with less inequalities for the HK achievable region. Since the cardinality of the time-sharing parameter is directly related to the number of inequalities appearing in the achievable rate region, the computational complexity is decreased. However, finding the full HK achievable region is still prohibitively complex.

Regarding outer bounds on the capacity region, there are three main results known. The first one obtained by Sato [10] is originally derived for the degraded Gaussian IC. Sato has shown that the capacity region of the degraded Gaussian IC is outer bounded by a certain degraded broadcast channel whose capacity region is fully characterized. In [11], Costa has proved that the capacity region of the degraded Gaussian broadcast channel is equivalent to that of the one-sided weak Gaussian IC. Hence, Sato outer bound can be used for the one-sided Gaussian IC as well.

The second outer bound obtained for the weak Gaussian IC is due to Kramer [12]. Kramer outer bound is based on the fact that removing one of the interfering links enlarges the capacity region. Therefore, the capacity region of the two-user Gaussian IC is inside the intersection of the capacity regions of the underlying one-sided Gaussian ICs. For the case of weak Gaussian IC, the underlying one-sided IC is weak, for which the capacity region is unknown. However, Kramer has used the outer bound obtained by Sato to derive an outer bound for the weak Gaussian IC.

The third outer bound due to Etkin, Tse, and Wang (ETW) is based on the Genie aided technique [13]. A genie that provides some extra information to the receivers can only enlarge the capacity region. At first glance, it seems a clever genie must

provide some information about the interference to the receiver to help in decoding the signal by removing the interference. In contrast, the genie in the ETW scheme provides information about the intended signal to the receiver. Remarkably, reference [13] shows that their proposed outer bound outperforms Kramer bound for certain range of parameters. Moreover, using a similar method, [13] presents an outer bound for the mixed Gaussian IC.

In this paper, by introducing the notion of admissible ICs, we propose a new outer bounding technique for the two-user Gaussian IC. The proposed technique relies on an extremal inequality recently proved by Liu and Viswanath [14]. We show that by using this scheme, one can obtain tighter outer bounds for both weak and mixed Gaussian ICs. More importantly, the sum capacity of the Gaussian weak IC for a certain range of the channel parameters is derived.

The rest of this paper is organized as follows. In Section II, we present some basic definitions and review the HK achievable region when Gaussian codebooks are used. We study the time-sharing and the convexification methods as means to enlarge the basic HK achievable region. We investigate conditions for which the two regions obtained from time-sharing and concavification coincide. Finally, we consider an optimization problem based on extremal inequality and compute its optimal solution.

In Section III, the notion of an admissible IC is introduced. Some classes of admissible ICs for the two-user Gaussian case is studied and outer bounds on the capacity regions of these classes are computed. We also obtain the sum capacity of a specific class of admissible IC where it is shown that using Gaussian codebooks and treating interference as noise is optimal.

In Section IV, we study the capacity region of the weak Gaussian IC. We first derive the sum capacity of this channel for a certain range of parameters where it is proved that users should treat the interference as noise and transmit at their highest possible rates. We then derive an outer bound on the capacity region which outperforms the known results. We finally prove that the basic HK achievable region results in the same enlarged region by using either time-sharing or concavification. This reduces the complexity of the characterization of the full HK achievable region when Gaussian codebooks are used.

In Section V, we study the capacity region of the one-sided Gaussian IC. We present a new proof for the Sato outer bound using the extremal inequality. Then, we present methods to simplify the HK achievable region such that the full region can be characterized.

In Section VI, we study the capacity region of the mixed Gaussian IC. We first obtain the sum capacity of this channel and then derive an outer bound which outperforms other known results. Finally, by investigating the HK achievable region for different cases, we prove that for a certain range of channel parameters, the full HK achievable rate region using Gaussian codebooks is equivalent to that of the one-sided IC. Finally, in Section VII, we conclude the paper.

#### A. Notations

Throughout this paper, we use the following notations. Vectors are represented by bold faced letters. Random variables, matrices, and sets are denoted by capital letters where the difference is clear from the context.  $|A|$ ,  $tr\{A\}$ , and  $A^t$  represent the determinant, trace, and transpose of the square matrix  $A$ , respectively.  $I$  denotes the identity matrix.  $\mathbb{N}$  and  $\mathbb{R}$  are the sets of nonnegative integers and real numbers, respectively. The union, intersection, and Minkowski sum of two sets  $U$  and  $V$  are represented by  $U \cup V$ ,  $U \cap V$ , and  $U + V$ , respectively. We use  $\gamma(x)$  as an abbreviation for the function  $0.5 \log_2(1 + x)$ .

## II. PRELIMINARIES

### A. The Two-user Interference Channel

*Definition 1 (two-user IC):* A two-user discrete memoryless IC consists of two finite sets  $\mathcal{X}_1$  and  $\mathcal{X}_2$  as input alphabets and two finite sets  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$  as the corresponding output alphabets. The channel is governed by conditional probability distributions  $\omega(y_1, y_2 | x_1, x_2)$  where  $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$  and  $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ .

*Definition 2 (capacity region of the two-user IC):* A code  $(2^{nR_1}, 2^{nR_2}, n, \lambda_1^n, \lambda_2^n)$  for the two-user IC consists of the following components for User  $i \in \{1, 2\}$ :

- 1) A uniform distributed message set  $\mathcal{M}_i \in [1, 2, \dots, 2^{nR_i}]$ .
- 2) A codebook  $\mathcal{X}_i = \{\mathbf{x}_i(1), \mathbf{x}_i(2), \dots, \mathbf{x}_i(2^{nR_i})\}$  where  $\mathbf{x}_i(\cdot) \in \mathcal{X}_i^n$ .
- 3) An encoding function  $F_i : [1, 2, \dots, 2^{nR_i}] \rightarrow \mathcal{X}_i$ .
- 4) A decoding function  $G_i : \mathcal{Y}_i \rightarrow [1, 2, \dots, 2^{nR_i}]$ .
- 5) The average probability of error  $\lambda_i^n = \mathbb{P}(G_i(\mathbf{y}_i) \neq \mathcal{M}_i)$ .

A rate pair  $(R_1, R_2)$  is achievable if there is a sequence of codes  $(2^{nR_1}, 2^{nR_2}, n, \lambda_1^n, \lambda_2^n)$  with vanishing average error probabilities. The capacity region of the IC is defined to be the supremum of the set of achievable rates.

Let  $\mathcal{C}_{IC}$  denote the capacity region of the two-user IC. The limiting expression for  $\mathcal{C}_{IC}$  can be stated as [5]

$$\mathcal{C}_{IC} = \lim_{n \rightarrow \infty} \text{closure} \left( \bigcup_{\mathbb{P}(\mathbf{X}_1^n) \mathbb{P}(\mathbf{X}_2^n)} \left\{ (R_1, R_2) \mid \begin{array}{l} R_1 \leq \frac{1}{n} \mathbf{I}(\mathbf{X}_1^n, \mathbf{Y}_1^n) \\ R_2 \leq \frac{1}{n} \mathbf{I}(\mathbf{X}_2^n, \mathbf{Y}_2^n) \end{array} \right\} \right). \quad (1)$$

In this paper, we focus on the two-user Gaussian IC which can be represented in standard form as [15], [16]

$$\begin{aligned} y_1 &= x_1 + \sqrt{a}x_2 + z_1, \\ y_2 &= \sqrt{b}x_1 + x_2 + z_2, \end{aligned} \quad (2)$$

where  $x_i$  and  $y_i$  denote the input and output alphabets of User  $i \in \{1, 2\}$ , respectively, and  $z_1 \sim \mathcal{N}(0, 1)$ ,  $z_2 \sim \mathcal{N}(0, 1)$  are standard Gaussian random variables. Constants  $a \geq 0$  and  $b \geq 0$  represent the gains of the interference links. Furthermore, Transmitter  $i$ ,  $i \in \{1, 2\}$ , is subject to the power constraint  $P_i$ . Achievable rates and the capacity region of the Gaussian IC can be defined in a similar fashion as that of the general IC with the condition that the codewords must satisfy their corresponding power constraints. The capacity region of the two-user Gaussian IC is denoted by  $\mathcal{C}$ . Clearly,  $\mathcal{C}$  is a function of the parameters  $P_1$ ,  $P_2$ ,  $a$ , and  $b$ . To emphasize this relationship, we may write  $\mathcal{C}$  as  $\mathcal{C}(P_1, P_2, a, b)$  as needed.

*Remark 1:* Since the capacity region of the general IC depends only on the marginal distributions [16], the ICs can be classified into equivalent classes in which channels within a class have the same capacity region. In particular, for the Gaussian IC given in (2), any choice of joint distributions for the pair  $(z_1, z_2)$  does not affect the capacity region as long as the marginal distributions remain Gaussian with zero mean and unit variance.

Depending on the values of  $a$  and  $b$ , the two-user Gaussian IC is classified into weak, strong, mixed, one-sided, and degraded Gaussian IC. In Figure 1, regions in  $ab$ -plane together with their associated names are shown. Briefly, if  $0 < a < 1$  and  $0 < b < 1$ , then the channel is called *weak Gaussian IC*. If  $1 \leq a$  and  $1 \leq b$ , then the channel is called *strong Gaussian IC*. If either  $a = 0$  or  $b = 0$ , the channel is called *one-sided Gaussian IC*. If  $ab = 1$ , then the channel is called *degraded Gaussian IC*. If either  $0 < a < 1$  and  $1 \leq b$ , or  $0 < b < 1$  and  $1 \leq a$ , then the channel is called *mixed Gaussian IC*. Finally, the *symmetric Gaussian IC* (used throughout the paper for illustration purposes) corresponds to  $a = b$  and  $P_1 = P_2$ .

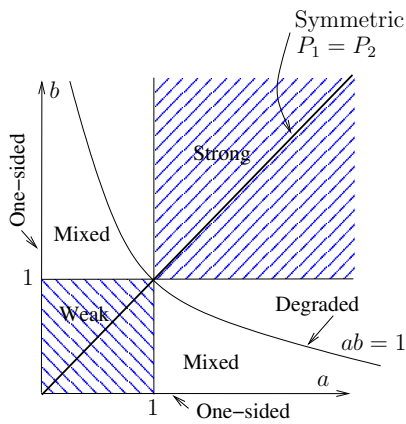


Fig. 1. Classes of the two-user ICs.

Among all classes shown in Figure 1, the capacity region of the strong Gaussian IC is fully characterized [3], [2]. In this case, the capacity region can be stated as the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$\begin{aligned} R_1 &\leq \gamma(P_1), \\ R_2 &\leq \gamma(P_2), \\ R_1 + R_2 &\leq \min\{\gamma(P_1 + aP_2), \gamma(bP_1 + P_2)\}. \end{aligned}$$

### B. Support Functions

Throughout this paper, we use the following facts from convex analysis. There is a one to one correspondence between any closed convex set and its support function [17]. The support function of any set  $D \in \Re^m$  is a function  $\sigma_D : \Re^m \rightarrow \Re$  defined as

$$\sigma_D(\mathbf{c}) = \sup\{\mathbf{c}^t \mathbf{R} | \mathbf{R} \in D\}. \quad (3)$$

Clearly, if the set  $D$  is compact, then the sup is attained and can be replaced by max. In this case, the solutions of (3) correspond to the boundary points of  $D$  [17]. The following relation is the dual of (3) and holds when  $D$  is closed and convex

$$D = \{\mathbf{R} | \mathbf{c}^t \mathbf{R} \leq \sigma_D(\mathbf{c}), \forall \mathbf{c}\}. \quad (4)$$

For any two closed convex sets  $D$  and  $D'$ ,  $D \subseteq D'$ , if and only if  $\sigma_D \leq \sigma_{D'}$ .

### C. Han-Kobayashi Achievable Region

The best inner bound for the two-user Gaussian IC is the full HK achievable region denoted by  $\mathcal{C}_{HK}$  [8]. Despite having a single letter formula,  $\mathcal{C}_{HK}$  is not fully characterized yet. In fact, finding the optimum distributions achieving boundary points of  $\mathcal{C}_{HK}$  is still an open problem. We define  $\mathcal{G}$  as a subset of  $\mathcal{C}_{HK}$  where Gaussian distributions are used for codebook generation. Using a shorter description of  $\mathcal{C}_{HK}$  obtained in [9],  $\mathcal{G}$  can be described as follows.

Let us first define  $\mathcal{G}_0$  as the collection of all rate pairs  $(R_1, R_2) \in \mathbb{R}_+^2$  satisfying

$$R_1 \leq \psi_1 = \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (5)$$

$$R_2 \leq \psi_2 = \gamma \left( \frac{P_2}{1 + b\alpha P_1} \right), \quad (6)$$

$$R_1 + R_2 \leq \psi_3 = \min \{ \psi_{31}, \psi_{32}, \psi_{33} \}, \quad (7)$$

$$2R_1 + R_2 \leq \psi_4 = \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left( \frac{\alpha P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{\beta P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right), \quad (8)$$

$$R_1 + 2R_2 \leq \psi_5 = \gamma \left( \frac{\beta P_2}{1 + b\alpha P_1} \right) + \gamma \left( \frac{P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right) + \gamma \left( \frac{\alpha P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right), \quad (9)$$

for fixed  $\alpha \in [0, 1]$  and  $\beta \in [0, 1]$ .<sup>1</sup>  $\psi_3$  is the minimum of  $\psi_{31}$ ,  $\psi_{32}$ , and  $\psi_{33}$  defined as

$$\psi_{31} = \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left( \frac{\beta P_2}{1 + b\alpha P_1} \right), \quad (10)$$

$$\psi_{32} = \gamma \left( \frac{\alpha P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right), \quad (11)$$

$$\psi_{33} = \gamma \left( \frac{\alpha P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left( \frac{\beta P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right). \quad (12)$$

$\mathcal{G}_0$  is a polytope and a function of four variables  $P_1$ ,  $P_2$ ,  $\alpha$ , and  $\beta$ . To emphasize this relation, we may write  $\mathcal{G}_0(P_1, P_2, \alpha, \beta)$  as needed. It is convenient to represent  $\mathcal{G}_0$  in a matrix form as  $\mathcal{G}_0 = \{ \mathbf{R} | \mathbf{A}\mathbf{R} \leq \Psi(P_1, P_2, \alpha, \beta) \}$  where  $\mathbf{R} = (R_1, R_2)^t$ ,  $\Psi = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5)^t$ , and

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix}^t.$$

Equivalently,  $\mathcal{G}_0$  can be represented as the convex hull of its extreme points, i.e.,  $\mathcal{G}_0(P_1, P_2, \alpha, \beta) = \text{conv} \{r_1, r_2, \dots, r_K\}$ , where it is assumed that  $\mathcal{G}_0$  has  $K$  extreme points. It is easy to show that  $K \leq 7$ .

Now,  $\mathcal{G}$  can be defined as a region obtained from enlarging  $\mathcal{G}_0$  by making use of the time-sharing parameter, i.e.,  $\mathcal{G}$  is the collection of all rate pairs  $\mathbf{R} = (R_1, R_2)^t$  satisfying

$$\mathbf{A}\mathbf{R} \leq \sum_{i=1}^q \lambda_i \Psi(P_{1i}, P_{2i}, \alpha_i, \beta_i), \quad (13)$$

where  $q \in \mathbb{N}$  and

$$\sum_{i=1}^q \lambda_i P_{1i} \leq P_1, \quad (14)$$

$$\sum_{i=1}^q \lambda_i P_{2i} \leq P_2, \quad (15)$$

$$\sum_{i=1}^q \lambda_i = 1, \quad (16)$$

$$\lambda_i \geq 0, (\alpha_i, \beta_i) \in [0, 1]^2; \forall i \in \{1, \dots, q\}. \quad (17)$$

It is easy to show that  $\mathcal{G}$  is a closed, bounded and convex region. In fact, the capacity region  $\mathcal{C}$  which contains  $\mathcal{G}$  is inside the rectangle defined by inequalities  $R_1 \leq \gamma(P_1)$  and  $R_2 \leq \gamma(P_2)$ . Moreover,  $(0, 0)$ ,  $(\gamma(P_1), 0)$ , and  $(0, \gamma(P_2))$  are extreme

<sup>1</sup>In the HK scheme, two independent messages are encoded at each transmitter, namely the *common message* and the *private message*.  $\alpha$  and  $\beta$  are the parameters that determine the amount of power allocated to the common and private messages for the two users, i.e.,  $\alpha P_1$ ,  $\beta P_2$  and  $(1 - \alpha)P_1$ ,  $(1 - \beta)P_2$  of the total power is used for the transmission of the private/common messages to the first/second users, respectively.

points of both  $\mathcal{C}$  and  $\mathcal{G}$ . Hence, to characterize  $\mathcal{G}$ , we need to obtain all extreme points of  $\mathcal{G}$  that are in the interior of the first quadrant (the same argument holds for  $\mathcal{C}$ ). In other words, we need to obtain  $\sigma_{\mathcal{G}}(c_1, c_2)$ , the support function of  $\mathcal{G}$ , either when  $1 \leq c_1$  and  $c_2 = 1$  or when  $c_1 = 1$  and  $1 \leq c_2$ .

We also define  $\mathcal{G}_1$  and  $\mathcal{G}_2$  obtained by enlarging  $\mathcal{G}_0$  in two different manners.  $\mathcal{G}_1$  is defined as

$$\mathcal{G}_1(P_1, P_2) = \bigcup_{(\alpha, \beta) \in [0, 1]^2} \mathcal{G}_0(P_1, P_2, \alpha, \beta). \quad (18)$$

$\mathcal{G}_1$  is not necessarily a convex region. Hence, it can be further enlarged by the convex hull operation.  $\mathcal{G}_2$  is defined as the collection of all rate pairs  $\mathbf{R} = (R_1, R_2)^t$  satisfying

$$\mathbf{R} = \sum_{i=1}^{q'} \lambda_i \mathbf{R}_i \quad (19)$$

where  $q' \in \mathbb{N}$  and

$$A\mathbf{R}_i \leq \Psi(P_{1i}, P_{2i}, \alpha_i, \beta_i), \quad (20)$$

$$\sum_{i=1}^{q'} \lambda_i P_{1i} \leq P_1, \quad (21)$$

$$\sum_{i=1}^{q'} \lambda_i P_{2i} \leq P_2, \quad (22)$$

$$\sum_{i=1}^{q'} \lambda_i = 1, \quad (23)$$

$$\lambda_i \geq 0, (\alpha_i, \beta_i) \in [0, 1]^2; \forall i \in \{1, \dots, q'\}. \quad (24)$$

It is easy to show that  $\mathcal{G}_2$  is a closed, bounded and convex region. In fact,  $\mathcal{G}_2$  is obtained by using the simple method of TD/FD. To see this, let us divide the available frequency band into  $q'$  sub-bands where  $\lambda_i$  represents the length of the  $i$ 'th band and  $\sum_{i=1}^{q'} \lambda_i = 1$ . User 1 and 2 allocate  $P_{1i}$  and  $P_{2i}$  in the  $i$ 'th sub-band, respectively. Therefore, all rate pairs in  $\mathcal{G}_0(P_{1i}, P_{2i}, \alpha_i, \beta_i)$  are achievable in the  $i$ 'th sub-band for fixed  $(\alpha_i, \beta_i) \in [0, 1]^2$ . Hence, all rate pairs in  $\sum_{i=1}^{q'} \lambda_i \mathcal{G}_0(P_{1i}, P_{2i}, \alpha_i, \beta_i)$  are achievable provided that  $\sum_{i=1}^{q'} \lambda_i P_{1i} \leq P_1$  and  $\sum_{i=1}^{q'} \lambda_i P_{2i} \leq P_2$ .

Clearly, the chain of inclusions  $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{G} \subseteq \mathcal{C}_{HK} \subseteq \mathcal{C}$  always holds.

#### D. Concavification Versus Time-Sharing

In this subsection, we follow two objectives. First, we aim at providing some necessary conditions such that  $\mathcal{G}_2 = \mathcal{G}$ . Second, we bound  $q$  and  $q'$  which are parameters involved in the descriptions of  $\mathcal{G}$  and  $\mathcal{G}_2$ , respectively. However, we derive the required conditions for the more general case where there are  $M$  users in the system. To this end, assume an achievable scheme for an  $M$ -user channel with the power constraint  $\mathbf{P} = [P_1, P_2, \dots, P_M]$  is given. The corresponding achievable region can be represented as

$$D_0(\mathbf{P}, \Theta) = \{\mathbf{R} | A\mathbf{R} \leq \Psi(\mathbf{P}, \Theta)\}, \quad (25)$$

where  $A$  is a  $K \times M$  matrix and  $\Theta \in [0, 1]^M$ .  $D_0$  is a polyhedron in general, but for the purpose of this paper, it suffices to assume that it is a polytope. Since  $D_0$  is a convex region, the convex hull operation does not lead to a new enlarged region. However, if the extreme points of the region are not a concave function of  $\mathbf{P}$ , it is possible to enlarge  $D_0$  by using two

different methods which are explained next. The first method is based on using the time sharing parameter. Let us denote the corresponding region as  $D$  which can be written as

$$D = \left\{ \mathbf{R} \mid A\mathbf{R} \leq \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i), \sum_{i=1}^q \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^q \lambda_i = 1, \lambda_i \geq 0, \Theta_i \in [0, 1]^M \forall i \right\}, \quad (26)$$

where  $q \in \mathbb{N}$ .

In the second method, we use TD/FD to enlarge the achievable rate region. This results in an achievable region  $D_2$  represented as

$$D_2 = \left\{ \mathbf{R} = \sum_{i=1}^{q'} \lambda_i \mathbf{R}_i \mid A\mathbf{R}_i \leq \Psi(\mathbf{P}_i, \Theta_i), \sum_{i=1}^{q'} \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^{q'} \lambda_i = 1, \lambda_i \geq 0, \Theta_i \in [0, 1]^M \forall i \right\}, \quad (27)$$

where  $q' \in \mathbb{N}$ . We refer to this method as concavification. It can be readily shown that  $D$  and  $D_2$  are closed and convex, and  $D_2 \subseteq D$ . We are interested in situations where the inverse inclusion holds.

The support function of  $D_0$  is a function of  $\mathbf{P}$ ,  $\Theta$ , and  $\mathbf{c}$ . Hence, we have

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \max\{\mathbf{c}^t \mathbf{R} \mid A\mathbf{R} \leq \Psi(\mathbf{P}, \Theta)\}. \quad (28)$$

For fixed  $\mathbf{P}$  and  $\Theta$ , (28) is a linear program. Using strong duality of linear programming, we obtain

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \min\{\mathbf{y}^t \Psi(\mathbf{P}, \Theta) \mid A^t \mathbf{y} = \mathbf{c}, \mathbf{y} \geq 0\}. \quad (29)$$

In general,  $\hat{\mathbf{y}}$ , the minimizer of (29), is a function of  $\mathbf{P}$ ,  $\Theta$ , and  $\mathbf{c}$ . We say  $D_0$  possesses *the unique minimizer property* if  $\hat{\mathbf{y}}$  merely depends on  $\mathbf{c}$ , for all  $\mathbf{c}$ . In this case, we have

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \hat{\mathbf{y}}^t(\mathbf{c}) \Psi(\mathbf{P}, \Theta), \quad (30)$$

where  $A^t \hat{\mathbf{y}} = \mathbf{c}$ . This condition means that for any  $\mathbf{c}$  the extreme point of  $D_0$  maximizing the objective  $\mathbf{c}^t \mathbf{R}$  is an extreme point obtained by intersecting a set of specific hyperplanes. A necessary condition for  $D_0$  to possess the unique minimizer property is that each inequality in describing  $D_0$  is either redundant or active for all  $\mathbf{P}$  and  $\Theta$ .

*Theorem 1:* If  $D_0$  possesses the unique minimizer property, then  $D = D_2$ .

*Proof:* Since  $D_2 \subseteq D$  always holds, we need to show  $D \subseteq D_2$  which can be equivalently verified by showing  $\sigma_D \leq \sigma_{D_2}$ .

The support function of  $D$  can be written as

$$\sigma_D(\mathbf{c}, \mathbf{P}) = \max\{\mathbf{c}^t \mathbf{R} \mid \mathbf{R} \in D\}. \quad (31)$$

By fixing  $\mathbf{P}$ ,  $\mathbf{P}_i$ 's,  $\Theta_i$ 's, and  $\lambda_i$ 's, the above maximization becomes a linear program. Hence, relying on weak duality of linear programming, we obtain

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \min_{A^t \mathbf{y} = \mathbf{c}, \mathbf{y} \geq 0} \mathbf{y}^t \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i). \quad (32)$$

Clearly,  $\hat{\mathbf{y}}(\mathbf{c})$ , the solution of (29), is a feasible point for (32) and we have

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \hat{\mathbf{y}}^t(\mathbf{c}) \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i). \quad (33)$$

Using (30), we obtain

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^q \lambda_i \sigma_{D_0}(\mathbf{c}, \mathbf{P}_i, \Theta_i). \quad (34)$$

Let us assume  $\hat{\mathbf{R}}_i$  is the maximizer of (28). In this case, we have

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^q \lambda_i \mathbf{c}^t \hat{\mathbf{R}}_i. \quad (35)$$

Hence, we have

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \mathbf{c}^t \sum_{i=1}^q \lambda_i \hat{\mathbf{R}}_i. \quad (36)$$

By definition,  $\sum_{i=1}^q \lambda_i \hat{\mathbf{R}}_i$  is a point in  $D_2$ . Therefore, we conclude

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \sigma_{D_2}(\mathbf{c}, \mathbf{P}). \quad (37)$$

This completes the proof.  $\square$

*Corollary 1 (Han [18]):* If  $D_0$  is a polymatroid, then  $D=D_2$ .

*Proof:* It is easy to show that  $D_0$  possesses the unique minimizer property. In fact, for given  $\mathbf{c}$ ,  $\hat{\mathbf{y}}$  can be obtained in a greedy fashion independent of  $\mathbf{P}$  and  $\Theta$ .  $\square$

In what follows, we upper bound  $q$  and  $q'$ .

*Theorem 2:* The cardinality of the time sharing parameter  $q$  in (26) is less than  $M + K + 1$ , where  $M$  and  $K$  are the dimensions of  $\mathbf{P}$  and  $\Psi(\mathbf{P})$ , respectively. Moreover, if  $\Psi(\mathbf{P})$  is a continuous function of  $\mathbf{P}$ , then  $q \leq M + K$ .

*Proof:* Let us define  $E$  as

$$E = \left\{ \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i) \mid \sum_{i=1}^q \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^q \lambda_i = 1, \lambda_i \geq 0, \Theta_i \in [0, 1]^M \forall i \right\}. \quad (38)$$

In fact,  $E$  is the collection of all possible bounds for  $D$ . To prove  $q \leq M + K + 1$ , we define another region  $E_1$  as

$$E_1 = \{(\mathbf{P}', \mathbf{S}') \mid 0 \leq \mathbf{P}', \mathbf{S}' = \Psi(\mathbf{P}', \Theta'), \Theta' \in [0, 1]^M\}. \quad (39)$$

From the direct consequence of the Caratheodory's theorem [19], the convex hull of  $E_1$  denoted by  $\text{conv } E_1$  can be obtained by convex combinations of no more than  $M + K + 1$  points in  $E_1$ . Moreover, if  $\Psi(\mathbf{P}', \Theta')$  is continuous, then  $M + K$  points are sufficient due to the extension of the Caratheodory's theorem [19]. Now, we define the region  $\hat{E}$  as

$$\hat{E} = \{\mathbf{S}' \mid (\mathbf{P}', \mathbf{S}') \in \text{conv } E_1, \mathbf{P}' \leq \mathbf{P}\}. \quad (40)$$

Clearly,  $\hat{E} \subseteq E$ . To show the other inclusion, let us consider a point in  $E$ , say  $S = \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i)$ . Since  $(\mathbf{P}_i, \Psi(\mathbf{P}_i, \Theta_i))$  is a point in  $E_1$ ,  $\sum_{i=1}^q \lambda_i (\mathbf{P}_i, \Psi(\mathbf{P}_i, \Theta_i))$  belongs to  $\text{conv } E_1$ . Having  $\sum_{i=1}^q \lambda_i \mathbf{P}_i \leq \mathbf{P}$ , we conclude  $\sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i) \in \hat{E}$ . Hence,  $E \subseteq \hat{E}$ . This completes the proof.  $\square$

*Corollary 2 (Etkin, Parakh, and Tse [20]):* For the  $M$ -user Gaussian IC where users use Gaussian codebooks for data transmission and treat the interference as noise, the cardinality of the time sharing parameter is less than  $2M$ .

*Proof:* In this case,  $D_0 = \{\mathbf{R} \mid \mathbf{R} \leq \Psi(\mathbf{P})\}$  where both  $\mathbf{P}$  and  $\Psi(\mathbf{P})$  have dimension  $M$  and  $\Psi(\mathbf{P})$  is a continuous function of  $\mathbf{P}$ . Applying Theorem 2 yields the desired result.  $\square$

In the following theorem, we obtain an upper bound on  $q'$ .

*Theorem 3:* To characterize boundary points of  $D_2$ , it suffices to set  $q' \leq M + 1$ .

*Proof:* Let us assume  $\hat{\mathbf{R}}$  is a boundary point of  $D_2$ . Hence, there exists  $\mathbf{c}$  such that

$$\sigma_{D_2}(\mathbf{c}, \mathbf{P}) = \max_{\mathbf{R} \in D_2} \mathbf{c}^t \mathbf{R} = \mathbf{c}^t \hat{\mathbf{R}}, \quad (41)$$

where  $\hat{\mathbf{R}} = \sum_{i=1}^{q'} \hat{\lambda}_i \hat{\mathbf{R}}_i$  and the optimum is achieved for the set of parameters  $\hat{\Theta}_i$ ,  $\hat{\lambda}_i$ , and  $\hat{\mathbf{P}}_i$ . The optimization problem in (41) can be written as

$$\begin{aligned} \sigma_{D_2}(\mathbf{c}, \mathbf{P}) = \max & \sum_{i=1}^{q'} \lambda_i g(\mathbf{c}, \mathbf{P}_i) \\ \text{subject to:} & \sum_{i=1}^{q'} \lambda_i = 1, \quad \sum_{i=1}^{q'} \lambda_i \mathbf{P}_i \leq \mathbf{P}, \\ & 0 \leq \lambda_i, 0 \leq \mathbf{P}_i, \quad \forall i \in \{1, 2, \dots, q'\}, \end{aligned} \quad (42)$$

where  $g(\mathbf{c}, \mathbf{P})$  is defined as

$$\begin{aligned} g(\mathbf{c}, \mathbf{P}) = \max & \mathbf{c}^t \mathbf{R} \\ \text{subject to:} & \mathbf{A} \mathbf{R} \leq \Psi(\mathbf{P}, \Theta), \quad 0 \leq \Theta \leq 1, \end{aligned} \quad (43)$$

In fact,  $\sigma_{D_2}(\mathbf{c}, \mathbf{P})$  in (42) can be viewed as the result of the concavification of  $g(\mathbf{c}, \mathbf{P})$  [19]. Hence, using Theorem 2.16 in [19], we conclude that  $q' \leq M + 1$ .  $\square$

Remarkable point about Theorem 3 is that the upper bound on  $q'$  is independent of the number of inequalities involved in the description of the achievable rate region.

*Corollary 3:* For the  $M$ -user Gaussian IC where users use Gaussian codebooks and treat the interference as noise, we have  $D_2 = D$  and  $q = q' = M + 1$ .

### E. Extremal Inequality

In [14], the following optimization problem is studied:

$$W = \max_{Q_{\mathbf{X}} \leq S} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\mathbf{X} + \mathbf{Z}_2), \quad (44)$$

where  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are  $n$ -dimensional Gaussian random vectors with the strictly positive definite covariance matrices  $Q_{\mathbf{Z}_1}$  and  $Q_{\mathbf{Z}_2}$ , respectively. The optimization is over all random vectors  $\mathbf{X}$  independent of  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$ .  $\mathbf{X}$  is also subject to the covariance matrix constraint  $Q_{\mathbf{X}} \leq S$ , where  $S$  is a positive definite matrix. In [14], it is shown that for all  $\mu \geq 1$ , this optimization problem has a Gaussian optimal solution for all positive definite matrices  $Q_{\mathbf{Z}_1}$  and  $Q_{\mathbf{Z}_2}$ . However, for  $0 \leq \mu < 1$  this optimization problem has a Gaussian optimal solution provided  $Q_{\mathbf{Z}_1} \leq Q_{\mathbf{Z}_2}$ , i.e.,  $Q_{\mathbf{Z}_2} - Q_{\mathbf{Z}_1}$  is a positive semi-definite matrix. It is worth noting that for  $\mu = 1$  this problem when  $Q_{\mathbf{Z}_1} \leq Q_{\mathbf{Z}_2}$  is studied under the name of the worse additive noise [21], [22].

In this paper, we consider a special case of (44) where  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  have the covariance matrices  $N_1 I$  and  $N_2 I$ , respectively, and the trace constraint is considered, i.e.,

$$W = \max_{\text{tr}\{Q_{\mathbf{X}}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\mathbf{X} + \mathbf{Z}_2). \quad (45)$$

In the following lemma, we provide the optimal solution for the above optimization problem when  $N_1 \leq N_2$ .

*Lemma 1:* If  $N_1 \leq N_2$ , the optimal solution of (45) is iid Gaussian for all  $0 \leq \mu$  and we have

1) For  $0 \leq \mu \leq \frac{N_2 + P}{N_1 + P}$ , the optimum covariance matrix is  $PI$  and the optimum solution is

$$W = \frac{n}{2} \log [(2\pi e)(P + N_1)] - \frac{\mu n}{2} \log [(2\pi e)(P + N_2)]. \quad (46)$$

2) For  $\frac{N_2+P}{N_1+P} < \mu \leq \frac{N_2}{N_1}$ , the optimum covariance matrix is  $\frac{N_2-\mu N_1}{\mu-1}I$  and the optimum solution is

$$W = \frac{n}{2} \log \left[ (2\pi e) \frac{N_2 - N_1}{\mu - 1} \right] - \frac{\mu n}{2} \log \left[ \frac{\mu(2\pi e)(N_2 - N_1)}{\mu - 1} \right]. \quad (47)$$

3) For  $\frac{N_2}{N_1} < \mu$ , the optimum covariance matrix is 0 and the optimum solution is

$$W = \frac{n}{2} \log(2\pi e N_1) - \frac{\mu n}{2} \log(2\pi e N_2). \quad (48)$$

*Proof:* From the general result for (44), we know that the optimum input distribution is Gaussian. Hence, we need to solve the following maximization problem:

$$\begin{aligned} W = \max \frac{1}{2} \log((2\pi e)^n |Q_{\mathbf{X}} + N_1 I|) - \frac{\mu}{2} \log((2\pi e)^n |Q_{\mathbf{X}} + N_2 I|) \\ \text{subject to: } 0 \leq Q_{\mathbf{X}}, \text{tr}\{Q_{\mathbf{X}}\} \leq nP. \end{aligned} \quad (49)$$

Since  $Q_{\mathbf{X}}$  is a positive semi-definite matrix, it can be decomposed as  $Q_{\mathbf{X}} = U\Lambda U^t$ , where  $\Lambda$  is a diagonal matrix with nonnegative entries and  $U$  is a unitary matrix, i.e.,  $UU^t = I$ . Substituting  $Q_{\mathbf{X}} = U\Lambda U^t$  in (49) and using the identities  $\text{tr}\{AB\} = \text{tr}\{BA\}$  and  $|AB + I| = |BA + I|$ , we obtain

$$\begin{aligned} W = \max \frac{1}{2} \log((2\pi e)^n |\Lambda + N_1 I|) - \frac{\mu}{2} \log((2\pi e)^n |\Lambda + N_2 I|) \\ \text{subject to: } 0 \leq \Lambda, \text{tr}\{\Lambda\} \leq nP. \end{aligned} \quad (50)$$

This optimization problem can be simplified as

$$\begin{aligned} W = \max \frac{n}{2} \sum_{i=1}^n [\log(2\pi e)(\lambda_i + N_1) - \mu \log(2\pi e)(\lambda_i + N_2)] \\ \text{subject to: } 0 \leq \lambda_i \forall i, \sum_{i=1}^n \lambda_i \leq nP. \end{aligned} \quad (51)$$

By introducing Lagrange multipliers  $\psi$  and  $\Phi = \{\phi_1, \phi_2, \dots, \phi_n\}$ , we obtain

$$L(\Lambda, \psi, \Phi) = \max \frac{n}{2} \sum_{i=1}^n [\log(2\pi e)(\lambda_i + N_1) - \mu \log(2\pi e)(\lambda_i + N_2)] + \psi \left( nP - \sum_{i=1}^n \lambda_i \right) + \sum_{i=1}^n \phi_i \lambda_i. \quad (52)$$

The first order KKT necessary conditions for the optimum solution of (52) can be written as

$$\frac{1}{\lambda_i + N_1} - \frac{\mu}{\lambda_i + N_2} - \psi + \phi_i = 0, \forall i \in \{1, 2, \dots, n\}, \quad (53)$$

$$\psi \left( nP - \sum_{i=1}^n \lambda_i \right) = 0, \quad (54)$$

$$\phi_i \lambda_i = 0, \forall i \in \{1, 2, \dots, n\}. \quad (55)$$

It is easy to show that when  $N_1 \leq N_2$ ,  $\lambda = \lambda_1 = \dots = \lambda_n$  and the only solution for  $\lambda$  is

$$\lambda = \begin{cases} P, & \text{if } 0 \leq \mu \leq \frac{N_2+P}{N_1+P} \\ \frac{N_2-\mu N_1}{\mu-1}, & \text{if } \frac{N_2+P}{N_1+P} < \mu \leq \frac{N_2}{N_1} \\ 0, & \text{if } \frac{N_2}{N_1} < \mu \end{cases} \quad (56)$$

Substituting  $\lambda$  into the objective function gives the desired result.  $\square$

In Figure 2, the optimum variance as a function of  $\mu$  is plotted. This figure shows that for any value of  $\mu \leq \frac{P+N_2}{P+N_1}$ , we need to use the maximum power to optimize the objective function, whereas for  $\mu > \frac{P+N_2}{P+N_1}$ , we use less power than what is permissible.

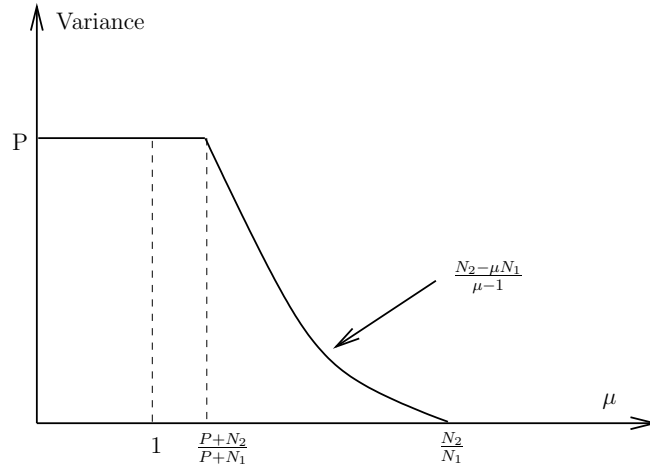


Fig. 2. Optimum variance versus  $\mu$ .

*Lemma 2:* If  $N_1 > N_2$ , the optimal solution of (45) is iid Gaussian for all  $1 \leq \mu$ . In this case, the optimum variance is 0 and the optimum  $W$  is

$$W = \frac{n}{2} \log(2\pi e N_1) - \frac{\mu n}{2} \log(2\pi e N_2). \quad (57)$$

*Proof:* The proof is similar to that of Lemma 1 and we omit it here.  $\square$

*Corollary 4:* For  $\mu = 1$ , the optimal solution of (45) is iid Gaussian and the optimum  $W$  is

$$W = \begin{cases} \frac{n}{2} \log\left(\frac{P+N_1}{P+N_2}\right), & \text{if } N_1 \leq N_2 \\ \frac{n}{2} \log\left(\frac{N_1}{N_2}\right), & \text{if } N_1 > N_2. \end{cases} \quad (58)$$

We frequently apply the following optimization problem in the rest of the paper:

$$f_h(P, N_1, N_2, a, \mu) = \max_{\text{tr}\{Q_{\mathbf{X}}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\sqrt{a}\mathbf{X} + \mathbf{Z}_2), \quad (59)$$

where  $N_1 \leq N_2/a$ . Using the identity  $h(A\mathbf{X}) = \log(|A|) + h(\mathbf{X})$ , (59) can be written as

$$f_h(P, N_1, N_2, a, \mu) = \frac{n}{2} \log a + \max_{\text{tr}\{Q_{\mathbf{X}}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h\left(\mathbf{X} + \frac{\mathbf{Z}_2}{\sqrt{a}}\right). \quad (60)$$

Now, Lemma 1 can be applied to obtain

$$f_h(P, N_1, N_2, a, \mu) = \begin{cases} \frac{1}{2} \log[(2\pi e)(P + N_1)] - \frac{\mu}{2} \log[(2\pi e)(aP + N_2)] & \text{if } 0 \leq \mu \leq \frac{P+N_2/a}{P+N_1} \\ \frac{1}{2} \log\left[(2\pi e) \frac{N_2/a - N_1}{\mu - 1}\right] - \frac{\mu}{2} \log\left[\frac{a\mu(2\pi e)(N_2/a - N_1)}{\mu - 1}\right] & \text{if } \frac{P+N_2/a}{P+N_1} < \mu \leq \frac{N_2}{aN_1} \\ \frac{1}{2} \log(2\pi e N_1) - \frac{\mu}{2} \log(2\pi e N_2) & \text{if } \frac{N_2}{aN_1} < \mu \end{cases} \quad (61)$$

### III. ADMISSIBLE CHANNELS

In this section, we aim at building ICs whose capacity regions contain the capacity region of the two-user Gaussian IC, i.e.,  $\mathcal{C}$ . Since we ultimately use these to outer bound  $\mathcal{C}$ , these ICs need to have a tractable expression (or a tractable outer bound) for their capacity regions.

Let us consider an IC with the same input letters as that of  $\mathcal{C}$  and the output letters  $\tilde{y}_1$  and  $\tilde{y}_2$  for Users 1 and 2, respectively. The capacity region of this channel, say  $\mathcal{C}'$ , contains  $\mathcal{C}$  if

$$I(x_1^n; y_1^n) \leq I(x_1^n; \tilde{y}_1^n), \quad (62)$$

$$I(x_2^n; y_2^n) \leq I(x_2^n; \tilde{y}_2^n), \quad (63)$$

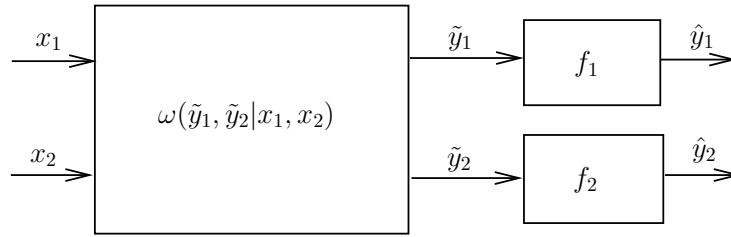


Fig. 3. An admissible channel.  $f_1$  and  $f_2$  are deterministic functions.

for all  $p(x_1^n)p(x_2^n)$  and for all  $n \in \mathbb{N}$ .

One way to satisfy (62) and (63) is to provide some extra information to either one or to both receivers. This technique is known as *Genie aided outer bounding*. In [12], Kramer has used such a genie to provide some extra information to both receivers such that they can decode both users' messages. Since the capacity region of this new interference channel is equivalent to that of the *Compound Multiple Access Channel* whose capacity region is known, reference [12] obtains an outer bound on the capacity region. To obtain a tighter outer bound, reference [12] further uses the fact that if a genie provides the exact information about the interfering signal to one of the receivers, then the new channel becomes the one-sided Gaussian IC. Although the capacity region of the one-sided Gaussian IC is unknown for all ranges of parameters, there exists an outer bound for it due to Sato and Costa [23], [11] that can be applied to the original channel. In [13], Etkin *et al.* use a different genie that provides some extra information about the intended signal. Even though at first glance their proposed method appears to be far from achieving a tight bound, remarkably they show that the corresponding bound is tighter than the one due to Kramer for certain ranges of parameters.

Next, we introduce the notion of admissible channels to satisfy (62) and (63).

*Definition 3 (Admissible Channel):* An IC  $\mathcal{C}'$  with input letter  $x_i$  and output letter  $\tilde{y}_i$  for User  $i \in \{1, 2\}$  is an admissible channel if there exist two deterministic functions  $\hat{y}_1^n = f_1(\tilde{y}_1^n)$  and  $\hat{y}_2^n = f_2(\tilde{y}_2^n)$  such that

$$I(x_1^n; y_1^n) \leq I(x_1^n; \hat{y}_1^n), \quad (64)$$

$$I(x_2^n; y_2^n) \leq I(x_2^n; \hat{y}_2^n) \quad (65)$$

hold for all  $p(x_1^n)p(x_2^n)$  and for all  $n \in \mathbb{N}$ .  $\mathcal{C}$  denotes the collection of all admissible channels (see Figure 3).

*Remark 2:* Genie aided channels are among admissible channels. To see this, let us assume a genie provides  $s_1$  and  $s_2$  as side information for User 1 and 2, respectively. In this case,  $\tilde{y}_i = (y_i, s_i)$  for  $i \in \{1, 2\}$ . By choosing  $f_i(y_i, s_i) = y_i$ , we observe that  $\hat{y}_i = y_i$ , and hence, (64) and (65) trivially hold.

To obtain the tightest outer bound, we need to find the intersection of the capacity regions of all admissible channels. Nonetheless, it may happen that finding the capacity region of an admissible channel is as hard as that of the original one (in fact, based on the definition, the channel itself is one of its admissible channels). Hence, we need to find classes of admissible channels, say  $\mathcal{F}$ , which possess two important properties. First, their capacity regions are close to  $\mathcal{C}$ . Second, either their exact capacity regions are computable or there exist good outer bounds for them. Since  $\mathcal{F} \subseteq \mathcal{C}$ , we have

$$\mathcal{C} \subseteq \bigcap_{\mathcal{F}} \mathcal{C}'. \quad (66)$$

Recall that there is a one to one correspondence between a closed convex set and its support function. Since  $\mathcal{C}$  is closed and convex, there is a one to one correspondence between  $\mathcal{C}$  and  $\sigma_{\mathcal{C}}$ . In fact, boundary points of  $\mathcal{C}$  correspond to the solutions of the following optimization problem

$$\sigma_{\mathcal{C}}(c_1, c_2) = \max_{(R_1, R_2) \in \mathcal{C}} c_1 R_1 + c_2 R_2. \quad (67)$$

Since we are interested in the boundary points excluding the  $R_1$  and  $R_2$  axes, it suffices to consider  $0 \leq c_1$  and  $0 \leq c_2$  where  $c_1 + c_2 = 1$ .

Since  $\mathcal{C} \subseteq \mathcal{C}'$ , we have

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \sigma_{\mathcal{C}'}(c_1, c_2). \quad (68)$$

Taking the minimum of the right hand side, we obtain

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \min_{\mathcal{C}' \in \mathcal{F}} \sigma_{\mathcal{C}'}(c_1, c_2), \quad (69)$$

which can be written as

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \min_{\mathcal{C}' \in \mathcal{F}} \max_{(R_1, R_2) \in \mathcal{C}'} c_1 R_1 + c_2 R_2. \quad (70)$$

For convenience, we use the following two optimization problems

$$\sigma_{\mathcal{C}}(\mu, 1) = \max_{(R_1, R_2) \in \mathcal{C}} \mu R_1 + R_2, \quad (71)$$

$$\sigma_{\mathcal{C}}(1, \mu) = \max_{(R_1, R_2) \in \mathcal{C}} R_1 + \mu R_2, \quad (72)$$

where  $1 \leq \mu$ . It is easy to show that the solutions of (71) and (72) correspond to the boundary points of the capacity region.

In the rest of this section, we introduce classes of admissible channels and obtain upper bounds on  $\sigma_{\mathcal{C}'}(\mu, 1)$  and  $\sigma_{\mathcal{C}'}(1, \mu)$ .

#### A. Classes of Admissible Channels

1) *Class A1*: This class is designed to obtain an upper bound on  $\sigma_{\mathcal{C}}(\mu, 1)$ . Therefore, we need to find a tight upper bound on  $\sigma_{\mathcal{C}'}(\mu, 1)$ . A member of this class is a channel in which User 1 has one transmit and one receive antenna whereas User 2 has one transmit antenna and two receive antennas (see Figure 4). The channel model can be written as

$$\begin{aligned} \tilde{y}_1 &= x_1 + \sqrt{a}x_2 + z_1, \\ \tilde{y}_{21} &= x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} &= x_2 + z_{22}, \end{aligned} \quad (73)$$

where  $\tilde{y}_1$  is the signal at the first receiver,  $\tilde{y}_{21}$  and  $\tilde{y}_{22}$  are the signals at the second receiver,  $z_1$  is additive Gaussian noise with unit variance,  $z_{21}$  and  $z_{22}$  are additive Gaussian noise with variances  $N_{21}$  and  $N_{22}$ , respectively. Transmitters 1 and 2 are subject to the power constraints of  $P_1$  and  $P_2$ , respectively.

To investigate admissibility conditions in (64) and (65), we introduce two deterministic functions  $f_1$  and  $f_2$  as follows (see Figure 4)

$$f_1(\tilde{y}_1^n) = \tilde{y}_1^n, \quad (74)$$

$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1 - \sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n, \quad (75)$$

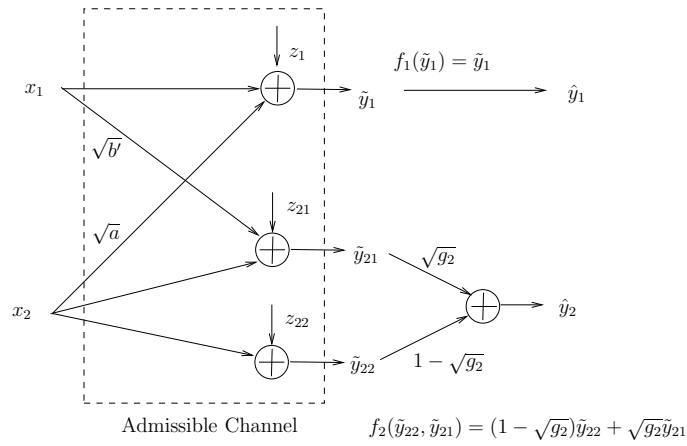


Fig. 4. Class A1 admissible channels.

where  $0 \leq g_2$ . For  $g_2 = 0$ , the channel can be converted to the one-sided Gaussian IC by letting  $N_{21} \rightarrow \infty$  and  $N_{22} = 1$ . Hence, Class A1 contains the one-sided Gaussian IC obtained by removing the link between Transmitter 1 and Receiver 2. Using  $f_1$  and  $f_2$ , we obtain

$$\hat{y}_1^n = x_1^n + \sqrt{a}x_2^n + z_1^n, \quad (76)$$

$$\hat{y}_2^n = \sqrt{b'g_2}x_1^n + x_2^n + (1 - \sqrt{g_2})z_{22}^n + \sqrt{g_2}z_{21}^n. \quad (77)$$

Hence, this channel is admissible if the corresponding parameters satisfy

$$\begin{aligned} b'g_2 &= b, \\ (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \quad (78)$$

We further add the following constraints to the conditions of the channels in Class A1:

$$\begin{aligned} b' &\leq N_{21}, \\ aN_{22} &\leq 1. \end{aligned} \quad (79)$$

Although these additional conditions reduce the number of admissible channels within the class, they are needed to get a closed form formula for an upper bound on  $\sigma_{\mathcal{C}'}(\mu, 1)$ . In the following lemma, we obtain the required upper bound.

*Lemma 3:* For the channels modeled by (73) and satisfying (79), we have

$$\begin{aligned} \sigma_{\mathcal{C}'}(\mu, 1) &\leq \min \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left( \frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\ &\quad + \mu_2 f_h \left( P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right) + f_h(P_2, N_{22}, 1, a, \mu_1) \\ &\text{subject to: } \mu_1 + \mu_2 = \mu, \mu_1, \mu_2 \geq 0. \end{aligned} \quad (80)$$

*Proof:* Let us assume  $R_1$  and  $R_2$  are achievable rates for User 1 and 2, respectively. Furthermore, we split  $\mu$  into  $\mu_1 \geq 0$

and  $\mu_2 \geq 0$  such that  $\mu = \mu_1 + \mu_2$ . Using Fano's inequality, we obtain

$$\begin{aligned}
n(\mu R_1 + R_2) &\leq \mu I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\
&= \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\
&\stackrel{(a)}{\leq} \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\
&= \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\
&= \mu_1 h(\tilde{y}_1^n) - \mu_1 h(\tilde{y}_1^n | x_1^n) + \mu_2 h(\tilde{y}_1^n | x_2^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) \\
&\quad + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \\
&= [\mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n)] + [\mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)] \\
&\quad + [h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n)] + [h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_1^n | x_1^n)] + n\epsilon_n, \tag{81}
\end{aligned}$$

where (a) follows from the fact that  $x_1^n$  and  $x_2^n$  are independent. Now, we separately upper bound the terms within each bracket in (81).

To maximize the terms within the first bracket, we use the fact that Gaussian distribution maximizes the differential entropy subject to a constraint on the covariance matrix. Hence, we have

$$\begin{aligned}
\mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) &= \mu_1 h(x_1^n + \sqrt{a}x_2^n + z_1^n) - \mu_2 h(z_1^n) \\
&\leq \frac{\mu_1 n}{2} \log [2\pi e (P_1 + aP_2 + 1)] - \frac{\mu_2 n}{2} \log (2\pi e). \tag{82}
\end{aligned}$$

Since  $b' \leq N_{21}$ , we can make use of Lemma 1 to upper bound the second bracket. In this case, we have

$$\begin{aligned}
\mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) &= \mu_2 \left( h(x_1^n + z_1^n) - \frac{1}{\mu_2} h(\sqrt{b'}x_1^n + z_{21}^n) \right) \\
&\leq \mu_2 n f_h \left( P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right), \tag{83}
\end{aligned}$$

where  $f_h$  is defined in (61).

We upper bound the terms within the third bracket as follows [13]:

$$\begin{aligned}
h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) &\stackrel{(a)}{\leq} \sum_{i=1}^n h(\tilde{y}_{21}[i] | \tilde{y}_{22}[i]) - h(z_{22}^n) \\
&\stackrel{(b)}{\leq} \sum_{i=1}^n \frac{1}{2} \log \left[ 2\pi e \left( N_{21} + b'P_1[i] + \frac{P_2[i]N_{22}}{P_2[i] + N_{22}} \right) \right] - \frac{n}{2} \log (2\pi e N_{22}) \\
&\stackrel{(c)}{\leq} \frac{n}{2} \log \left[ 2\pi e \left( N_{21} + \frac{1}{n} \sum_{i=1}^n b'P_1[i] + \frac{\frac{1}{n} \sum_{i=1}^n P_2[i]N_{22}}{\frac{1}{n} \sum_{i=1}^n P_2[i] + N_{22}} \right) \right] - \frac{n}{2} \log (2\pi e N_{22}) \\
&\leq \frac{n}{2} \log \left[ 2\pi e \left( N_{21} + b'P_1 + \frac{P_2 N_{22}}{P_2 + N_{22}} \right) \right] - \frac{n}{2} \log (2\pi e N_{22}) \\
&\leq \frac{n}{2} \log \left( \frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right), \tag{84}
\end{aligned}$$

where (a) follows from the chain rule and the fact that removing independent conditions does not decrease differential entropy, (b) follows from the fact that Gaussian distribution maximizes the conditional entropy for a given covariance matrix, and (c) follows from Jensen's inequality.

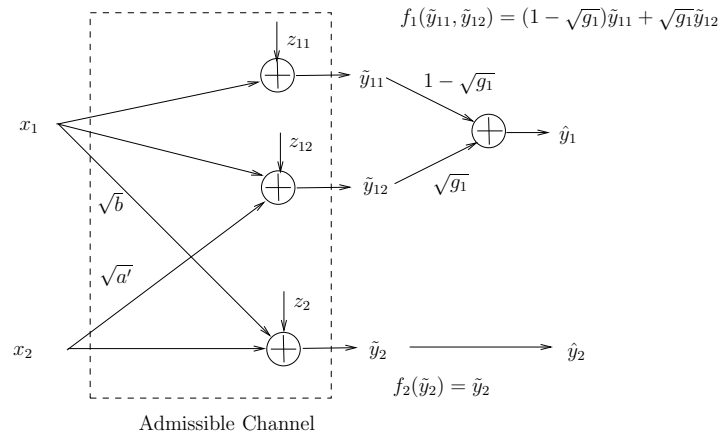


Fig. 5. Class A2 admissible channels.

For the last bracket, we again make use of the definition of  $f_h$ . In fact, since  $aN_{22} \leq 1$ , we have

$$\begin{aligned} h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_{11}^n | x_1^n) &= h(x_2^n + z_{22}^n) - \mu_1 h(\sqrt{a}x_2^n + z_1^n) \\ &\leq n f_h(P_2, N_{22}, 1, a, \mu_1). \end{aligned} \quad (85)$$

Adding all inequalities, we obtain

$$\begin{aligned} \mu R_1 + R_2 &\leq \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left( \frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\ &\quad + \mu_2 f_h \left( P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right) + f_h(P_2, N_{22}, 1, a, \mu_1), \end{aligned} \quad (86)$$

where the fact that  $\epsilon_n \rightarrow 0$  as  $n \rightarrow \infty$  is used to eliminate  $\epsilon_n$  from the right hand side of the inequality. Now, by taking the minimum of the right hand side of (86) over all  $\mu_1$  and  $\mu_2$ , we obtain the desired result. This completes the proof.  $\square$

2) *Class A2*: This class is the complement of Class A1 in the sense that we use it to upper bound  $\sigma_{\mathcal{E}}(1, \mu)$ . A member of this class is a channel in which User 1 is equipped with one transmit and two receive antennas, whereas User 2 is equipped with one antenna at both transmitter and receiver sides (see Figure 5). The channel model can be written as

$$\begin{aligned} \tilde{y}_{11} &= x_1 + z_{11}, \\ \tilde{y}_{12} &= x_1 + \sqrt{a'}x_2 + z_{12}, \\ \tilde{y}_2 &= x_2 + \sqrt{b}x_1 + z_2, \end{aligned} \quad (87)$$

where  $\tilde{y}_{11}$  and  $\tilde{y}_{12}$  are the signals at the first receiver,  $\tilde{y}_2$  is the signal at the second receiver,  $z_2$  is additive Gaussian noise with unit variance,  $z_{11}$  and  $z_{12}$  are additive Gaussian noise with variances  $N_{11}$  and  $N_{12}$ , respectively. Transmitter 1 and 2 are subject to the power constraints  $P_1$  and  $P_2$ , respectively.

For this class, we consider two linear functions  $f_1$  and  $f_2$  as follows (see Figure 5):

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1 - \sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n, \quad (88)$$

$$f_2(\tilde{y}_2^n) = \tilde{y}_2^n. \quad (89)$$

Similar to Class A1, when  $g_1 = 0$ , the admissible channels in Class A2 become the one-sided Gaussian IC by letting  $N_{12} \rightarrow \infty$

and  $N_{11} = 1$ . Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a'g_1}x_2^n + (1 - \sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n, \quad (90)$$

$$\hat{y}_2^n = \sqrt{b}x_1^n + x_2^n + z_2^n. \quad (91)$$

We conclude that the channel modeled by (87) is admissible if the corresponding parameters satisfy

$$\begin{aligned} a'g_1 &= a, \\ (1 - \sqrt{g_1})^2N_{11} + g_1N_{12} &= 1. \end{aligned} \quad (92)$$

Similar to Class A1, we further add the following constraints to the conditions of Class A2 channels:

$$\begin{aligned} a' &\leq N_{12}, \\ bN_{11} &\leq 1. \end{aligned} \quad (93)$$

In the following lemma, we obtain the required upper bound.

*Lemma 4:* For the channels modeled by (87) and satisfying (93), we have

$$\begin{aligned} \sigma_{\mathcal{E}'}(1, \mu) &\leq \min \frac{\mu_1}{2} \log [2\pi e(bP_1 + P_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left( \frac{N_{12}}{N_{11}} + \frac{a'P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right) \\ &\quad + \mu_2 f_h \left( P_2, 1, N_{12}, a', \frac{1}{\mu_2} \right) + f_h(P_1, N_{11}, 1, b, \mu_1) \\ &\text{subject to: } \mu_1 + \mu_2 = \mu, \mu_1, \mu_2 \geq 0. \end{aligned} \quad (94)$$

*Proof:* The proof is similar to that of Lemma 3 and we omit it here.  $\square$

3) *Class B:* A member of this class is a channel with one transmit antenna and two receive antennas for each user modeled by (see Figure 6)

$$\begin{aligned} \tilde{y}_{11} &= x_1 + z_{11}, \\ \tilde{y}_{12} &= x_1 + \sqrt{a'}x_2 + z_{12}, \\ \tilde{y}_{21} &= x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} &= x_2 + z_{22}, \end{aligned} \quad (95)$$

where  $\tilde{y}_{11}$  and  $\tilde{y}_{12}$  are the signals at the first receiver,  $\tilde{y}_{21}$  and  $\tilde{y}_{22}$  are the signals at the second receiver, and  $z_{ij}$  is additive Gaussian noise with variance  $N_{ij}$  for  $i, j \in \{1, 2\}$ . Transmitter 1 and 2 are subject to the power constraints  $P_1$  and  $P_2$ , respectively. In fact, this channel is designed to upper bound both  $\sigma_{\mathcal{E}}(\mu, 1)$  and  $\sigma_{\mathcal{E}}(1, \mu)$ .

Next, we investigate admissibility of this channel and the conditions that must be imposed on the underlying parameters. Let us consider two linear deterministic functions  $f_1$  and  $f_2$  with parameters  $0 \leq g_1$  and  $0 \leq g_2$ , respectively, as follows (see Figure 6)

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1 - \sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n, \quad (96)$$

$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1 - \sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n. \quad (97)$$

Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a'g_1}x_2^n + (1 - \sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n, \quad (98)$$

$$\hat{y}_2^n = \sqrt{b'g_2}x_1^n + x_2^n + (1 - \sqrt{g_2})z_{22}^n + \sqrt{g_2}z_{21}^n. \quad (99)$$

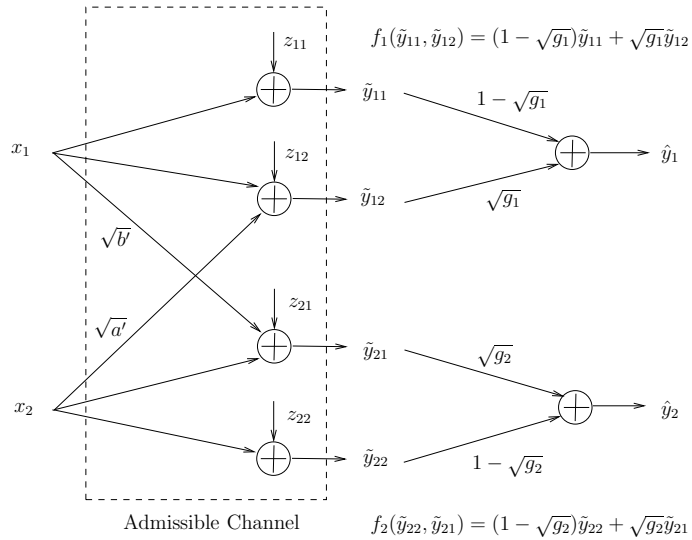


Fig. 6. Class B admissible channels.

To satisfy (64) and (65), it suffices to have

$$\begin{aligned}
 a'g_1 &= a, \\
 b'g_2 &= b, \\
 (1 - \sqrt{g_1})^2 N_{11} + g_1 N_{12} &= 1, \\
 (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1.
 \end{aligned} \tag{100}$$

Hence, a channel modeled by (95) is admissible if there exist two nonnegative numbers  $g_1$  and  $g_2$  such that the equalities in (100) are satisfied. We further add the following two constraints to the equality conditions in (100):

$$\begin{aligned}
 b'N_{11} &\leq N_{21}, \\
 a'N_{22} &\leq N_{12}.
 \end{aligned} \tag{101}$$

Although adding more constraints reduces the number of the admissible channels, it enables us to compute an outer bound on  $\sigma_{\mathcal{C}'}(\mu, 1)$  and  $\sigma_{\mathcal{C}'}(1, \mu)$ .

*Lemma 5:* For the channels modeled by (95) and satisfying (101), we have

$$\begin{aligned}
 \sigma_{\mathcal{C}'}(\mu, 1) &\leq \mu\gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) \\
 &\quad + f_h(P_2, N_{22}, N_{12}, a', \mu) + \frac{\mu}{2} \log((2\pi e)(a'P_2 + N_{12})) - \frac{1}{2} \log((2\pi e)(P_2 + N_{22})),
 \end{aligned} \tag{102}$$

$$\begin{aligned}
 \sigma_{\mathcal{C}'}(1, \mu) &\leq \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \mu\gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) \\
 &\quad + f_h(P_1, N_{11}, N_{21}, b', \mu) + \frac{\mu}{2} \log((2\pi e)(b'P_1 + N_{21})) - \frac{1}{2} \log((2\pi e)(P_1 + N_{11})).
 \end{aligned} \tag{103}$$

*Proof:* We only upper bound  $\sigma_{\mathcal{C}'}(\mu, 1)$  and an upper bound on  $\sigma_{\mathcal{C}'}(1, \mu)$  can be similarly obtained. Let us assume  $R_1$

and  $R_2$  are achievable rates for User 1 and User 2, respectively. Using Fano's inequality, we obtain

$$\begin{aligned}
n(\mu R_1 + R_2) &\leq \mu I(x_1^n; \tilde{y}_{11}^n, \tilde{y}_{12}^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\
&= \mu I(x_1^n; \tilde{y}_{12}^n | \tilde{y}_{11}^n) + \mu I(x_1^n; \tilde{y}_{11}^n) \\
&\quad + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\
&= \mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n) + \mu h(\tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n) \\
&\quad + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \\
&= [\mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{12}^n | x_1^n)] + [h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n)] \\
&\quad + [\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)] + [h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n)] + n\epsilon_n.
\end{aligned} \tag{104}$$

Next, we upper bound the terms within each bracket in (104) separately. For the first bracket, we have

$$\begin{aligned}
\mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{12}^n | x_1^n) &\stackrel{(a)}{\leq} \mu \sum_{i=1}^n h(\tilde{y}_{12}[i] | \tilde{y}_{11}[i]) - \frac{\mu n}{2} \log(2\pi e N_{11}) \\
&\stackrel{(b)}{\leq} \mu \sum_{i=1}^n \frac{1}{2} \log \left[ 2\pi e \left( N_{12} + a' P_2[i] + \frac{P_1[i] N_{11}}{P_1[i] + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\
&\stackrel{(c)}{\leq} \frac{\mu n}{2} \log \left[ 2\pi e \left( N_{12} + \frac{1}{n} \sum_{i=1}^n a' P_2[i] + \frac{\frac{1}{n} \sum_{i=1}^n P_1[i] N_{11}}{\frac{1}{n} \sum_{i=1}^n P_1[i] + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\
&\leq \frac{\mu n}{2} \log \left[ 2\pi e \left( N_{12} + a' P_2 + \frac{P_1 N_{11}}{P_1 + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\
&= \frac{\mu n}{2} \log \left( \frac{N_{12}}{N_{11}} + \frac{a' P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right),
\end{aligned} \tag{105}$$

where (a) follows from the chain rule and the fact that removing independent conditions increases differential entropy, (b) follows from the fact that Gaussian distribution optimizes conditional entropy for a given covariance matrix, and (c) follows from Jensen's inequality.

Similarly, the terms within the second bracket can be upper bounded as

$$h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n) \leq \frac{n}{2} \log \left( \frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right). \tag{106}$$

Using Lemma 1 and the fact that  $N_{11} \leq N_{21}/b'$ , the terms within the third bracket can be upper bounded as

$$\begin{aligned}
\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) &= \mu \left( h(x_1^n + z_{11}^n) - \frac{1}{\mu} h(\sqrt{b'} x_1^n + z_{21}^n) \right) \\
&\leq \mu n f_h \left( P_1, N_{11}, N_{21}, b', \frac{1}{\mu} \right).
\end{aligned} \tag{107}$$

Since  $1 \leq \mu$ , from (61) we obtain

$$\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) \leq \frac{\mu n}{2} \log((2\pi e)(P_1 + N_{11})) - \frac{n}{2} \log((2\pi e)(b' P_1 + N_{21})). \tag{108}$$

For the last bracket, again we use Lemma 1 to obtain

$$\begin{aligned}
h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n) &= h(x_2^n + z_{22}^n) - \mu h(\sqrt{a'} x_2^n + z_{12}^n) \\
&\leq n f_h(P_2, N_{22}, N_{12}, a', \mu).
\end{aligned} \tag{109}$$

Adding all inequalities, we have

$$\begin{aligned} \mu R_1 + R_2 \leq & \frac{\mu}{2} \log \left( \frac{N_{12}}{N_{11}} + \frac{a' P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right) + \frac{1}{2} \log \left( \frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\ & + \frac{\mu}{2} \log((2\pi e)(P_1 + N_{11})) - \frac{1}{2} \log((2\pi e)(b' P_1 + N_{21})) + f_h(P_2, N_{22}, N_{12}, a', \mu), \end{aligned} \quad (110)$$

where the fact that  $\epsilon_n \rightarrow 0$  as  $n \rightarrow \infty$  is used to eliminate  $\epsilon_n$  from the right hand side of the inequality. By rearranging the terms, we obtain

$$\begin{aligned} \mu R_1 + R_2 \leq & \mu \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right) \\ & + f_h(P_2, N_{22}, N_{12}, a', \mu) + \frac{\mu}{2} \log((2\pi e)(a' P_2 + N_{12})) - \frac{1}{2} \log((2\pi e)(P_2 + N_{22})). \end{aligned}$$

This completes the proof.  $\square$

A unique feature of the channels within Class B is that for  $1 \leq \mu \leq \frac{P_2 + N_{12}/a'}{P_2 + N_{22}}$  and  $1 \leq \mu \leq \frac{P_1 + N_{21}/b'}{P_1 + N_{11}}$ , the upper bounds in (102) and (103) become, respectively,

$$\mu R_1 + R_2 \leq \mu \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right) \quad (111)$$

and

$$R_1 + \mu R_2 \leq \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \mu \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right). \quad (112)$$

On the other hand, if the receivers treat the interference as noise, it can be shown that

$$R_1 = \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) \quad (113)$$

and

$$R_2 = \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right) \quad (114)$$

are achievable. Comparing upper bounds and achievable rates, we conclude that the upper bounds are indeed tight. In fact, this property is first observed by Etkin *et al.* in [13]. We summarize this result in the following theorem:

*Theorem 4:* The sum capacity in Class B is attained when transmitters use Gaussian codebooks and receivers treat the interference as noise. In this case, the sum capacity is

$$\mathcal{C}'_{\text{sum}} = \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right). \quad (115)$$

*Proof:* By substituting  $\mu = 1$  in (111), we obtain the desired result.  $\square$

4) *Class C:* Class C is designed to upper bound  $\sigma_{\mathcal{C}}(\mu, 1)$  for the mixed Gaussian IC where  $1 \leq b$ . Class C is similar to Class A1 (see Figure 4), however we impose different constraints on the parameters of the channels within Class C. These constraints assist us in providing upper bounds by using the fact that at one of the receivers both signals are decodable.

For channels in Class C, we use the same model that is given in (73). Therefore, similar to channels in Class A1, this channel is admissible if the corresponding parameters satisfy

$$\begin{aligned} b' g_2 &= b, \\ (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \quad (116)$$

Next, we change the constraints in (79) as

$$\begin{aligned} b' &\geq N_{21}, \\ aN_{22} &\leq 1. \end{aligned} \quad (117)$$

Through this change of constraints, the second receiver after decoding its own signal will have a less noisy version of the first user's signal, and consequently, it is able to decode the signal of the first user as well as its own signal. Relying on this observation, we have the following lemma.

*Lemma 6:* For a channel in Class C, we have

$$\begin{aligned} \sigma_{\mathcal{C}'}(\mu, 1) &\leq \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21}\right)\right) \\ &\quad - \frac{1}{2} \log(2\pi e N_{21}) - \frac{1}{2} \log(2\pi e N_{22}) + f_h(P_2, N_{22}, 1, a, \mu - 1). \end{aligned} \quad (118)$$

*Proof:* Since the second user is able to decode both users' messages, we have

$$R_1 \leq \frac{1}{n} I(x_1^n; \tilde{y}_1^n), \quad (119)$$

$$R_1 \leq \frac{1}{n} I(x_1^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_2^n), \quad (120)$$

$$R_2 \leq \frac{1}{n} I(x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n), \quad (121)$$

$$R_1 + R_2 \leq \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n). \quad (122)$$

From  $aN_{22} \leq 1$ , we have  $I(x_1^n; \tilde{y}_1^n) \leq I(x_1^n; \tilde{y}_{21}^n | x_2^n) = I(x_1^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_2^n)$ . Hence, (120) is redundant. It can be shown that

$$\mu R_1 + R_2 \leq \frac{\mu-1}{n} I(x_1^n; \tilde{y}_1^n) + \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n). \quad (123)$$

Hence, we have

$$\begin{aligned} \mu R_1 + R_2 &\leq \frac{\mu-1}{n} h(\tilde{y}_1^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) + \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n) - \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) \\ &= \frac{\mu-1}{n} h(\tilde{y}_1^n) + \frac{1}{n} h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) \\ &\quad + \left[ \frac{1}{n} h(\tilde{y}_{22}^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) \right] \end{aligned} \quad (124)$$

Next, we bound the different terms in (124). For the first term, we have

$$\frac{\mu-1}{n} h(\tilde{y}_1^n) \leq \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)). \quad (125)$$

The second term can be bounded as

$$\frac{1}{n} h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) \leq \frac{1}{2} \log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21}\right)\right). \quad (126)$$

The third term can be bounded as

$$\frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) = \frac{1}{2} \log(2\pi e N_{21}) + \frac{1}{2} \log(2\pi e N_{22}). \quad (127)$$

The last terms can be bounded as

$$\frac{1}{n} h(\tilde{y}_{22}^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) = \frac{1}{n} h(x_2^n + z_{22}^n) - \frac{\mu-1}{n} h(\sqrt{a}x_2^n + z_1) \quad (128)$$

$$\leq f_h(P_2, N_{22}, 1, a, \mu - 1). \quad (129)$$

Adding all inequalities, we obtain the desired result.  $\square$

#### IV. WEAK GAUSSIAN INTERFERENCE CHANNEL

In this section, we focus on the weak Gaussian IC. We first obtain the sum capacity of this channel for a certain range of parameters. Then, we obtain an outer bound on the capacity region which is tighter than the previously known outer bounds. Finally, we show that time-sharing and concavification result in the same achievable region for Gaussian codebooks.

##### A. Sum Capacity

In this subsection, we use the Class B channels to obtain the sum capacity of the weak IC for a certain range of parameters. To this end, let us consider the following minimization problem:

$$\begin{aligned}
 W = \min \gamma & \left( \frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) & (130) \\
 \text{subject to:} & \\
 & a'g_1 = a \\
 & b'g_2 = b \\
 & b'N_{11} \leq N_{21} \\
 & a'N_{22} \leq N_{12} \\
 & (1 - \sqrt{g_1})^2 N_{11} + g_1 N_{12} = 1 \\
 & (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1 \\
 & 0 \leq [a', b', g_1, g_2, N_{11}, N_{12}, N_{22}, N_{21}].
 \end{aligned}$$

The objective function in (130) is the sum capacity of Class B channels obtained in Theorem 4. The constraints are the combination of (100) and (101) where applied to confirm the admissibility of the channel and to validate the sum capacity result. Since every channel in the class is admissible, we have  $\mathcal{C}_{sum} \leq W$ . Substituting  $S_1 = g_1 N_{12}$  and  $S_2 = g_2 N_{21}$ , we have

$$\begin{aligned}
 W = \min \gamma & \left( \frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) + \gamma \left( \frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) & (131) \\
 \text{subject to:} & \\
 & \frac{b(1 - S_1)}{(1 - \sqrt{g_1})^2} \leq S_2 < 1 \\
 & \frac{a(1 - S_2)}{(1 - \sqrt{g_2})^2} \leq S_1 < 1 \\
 & 0 < [g_1, g_2].
 \end{aligned}$$

By first minimizing with respect to  $g_1$  and  $g_2$ , the optimization problem (131) can be decomposed as

$$\begin{aligned}
 W = \min W_1 + W_2 & & (132) \\
 \text{subject to: } & 0 < S_1 < 1, \quad 0 < S_2 < 1.
 \end{aligned}$$

where  $W_1$  is defined as

$$\begin{aligned}
 W_1 = \min_{g_1} \gamma & \left( \frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) & (133) \\
 \text{subject to: } & \frac{b(1 - S_1)}{S_2} \leq (1 - \sqrt{g_1})^2, \quad 0 < g_1.
 \end{aligned}$$

Similarly,  $W_2$  is defined as

$$W_2 = \min_{g_2} \gamma \left( \frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) \quad (134)$$

subject to:  $\frac{a(1 - S_2)}{S_1} \leq (1 - \sqrt{g_2})^2$ ,  $0 < g_2$ .

The optimization problems (133) and (134) are easy to solve. In fact, we have

$$W_1 = \begin{cases} \gamma \left( \frac{P_1}{1 + aP_2} \right) & \text{if } \sqrt{b}(1 + aP_2) \leq \sqrt{S_2(1 - S_1)} \\ \gamma \left( \frac{bP_1}{S_2} + \frac{(1 - \sqrt{b(1 - S_1)/S_2})^2 P_1}{aP_2 + S_1} \right) & \text{Otherwise} \end{cases} \quad (135)$$

$$W_2 = \begin{cases} \gamma \left( \frac{P_2}{1 + bP_1} \right) & \text{if } \sqrt{a}(1 + bP_1) \leq \sqrt{S_1(1 - S_2)} \\ \gamma \left( \frac{aP_2}{S_1} + \frac{(1 - \sqrt{a(1 - S_2)/S_1})^2 P_2}{bP_1 + S_2} \right) & \text{Otherwise} \end{cases} \quad (136)$$

From (135) and (136), we observe that for  $S_1$  and  $S_2$  satisfying  $\sqrt{b}(1 + aP_2) \leq \sqrt{S_2(1 - S_1)}$  and  $\sqrt{a}(1 + bP_1) \leq \sqrt{S_1(1 - S_2)}$ , the objective function becomes independent of  $S_1$  and  $S_2$ . In this case, we have

$$W = \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma \left( \frac{P_2}{1 + bP_1} \right), \quad (137)$$

which is achievable by treating interference as noise. In the following theorem, we prove that it is possible to find a certain range of parameters such that there exist  $S_1$  and  $S_2$  yielding (137).

*Theorem 5:* The sum capacity of the two-user Gaussian IC is

$$\mathcal{C}_{sum} = \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma \left( \frac{P_2}{1 + bP_1} \right), \quad (138)$$

for the range of parameters satisfying

$$\sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}}. \quad (139)$$

*Proof:* Let us fix  $a$  and  $b$ , and define  $D$  as

$$D = \left\{ (P_1, P_2) \mid P_1 \leq \frac{\sqrt{S_1(1 - S_2)}}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{\sqrt{S_2(1 - S_1)}}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1, 0 < S_2 < 1 \right\}. \quad (140)$$

In fact, if  $D$  is feasible then there exist  $0 < S_1 < 1$  and  $0 < S_2 < 1$  satisfying  $\sqrt{b}(1 + aP_2) \leq \sqrt{S_2(1 - S_1)}$  and  $\sqrt{a}(1 + bP_1) \leq \sqrt{S_1(1 - S_2)}$ . Therefore, the sum capacity of the channel for all feasible points is attained due to (137).

We claim that  $D = D'$ , where  $D'$  is defined as

$$D' = \left\{ (P_1, P_2) \mid \sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}} \right\}. \quad (141)$$

To show  $D' \subseteq D$ , we set  $S_1 = 1 - S_2$  in (140) to get

$$\left\{ (P_1, P_2) \mid P_1 \leq \frac{S_1}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{1 - S_1}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1 \right\} \subseteq D. \quad (142)$$

It is easy to show that the left hand side of the above equation is another representation of the region  $D'$ . Hence, we have  $D' \subseteq D$ .

To show  $D \subseteq D'$ , it suffices to prove that for any  $(P_1, P_2) \in D$ ,  $\sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}}$  holds. To this end, we introduce the following maximization problem:

$$J = \max_{(P_1, P_2) \in D} \sqrt{b}P_1 + \sqrt{a}P_2, \quad (143)$$

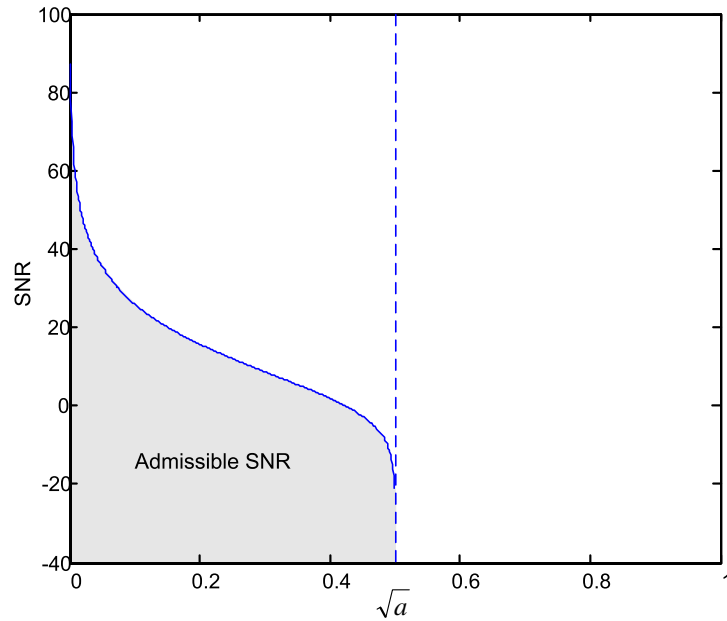


Fig. 7. The shaded area is the region where treating interference as noise is optimal for obtaining the sum capacity of the symmetric Gaussian IC.

which can be written as

$$J = \max_{(S_1, S_2) \in (0,1)^2} \frac{\sqrt{S_1(1-S_2)} + \sqrt{S_2(1-S_1)}}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \quad (144)$$

It is easy to show that the solution to the above optimization problem is

$$J = \frac{1}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \quad (145)$$

Hence, we deduce that  $D \subseteq D'$ . This completes the proof.  $\square$

*Remark 3:* The above sum capacity result for the weak Gaussian IC (see also [24]) has been established independently in [25] and [26].

As an example, let us consider the symmetric Gaussian IC. In this case, the constraint in (139) becomes

$$P \leq \frac{1 - 2\sqrt{a}}{2a\sqrt{a}}. \quad (146)$$

In Figure 7, the admissible region for  $P$ , where treating interference as noise is optimal, versus  $\sqrt{a}$  is plotted. For a fixed  $P$  and all  $0 \leq a \leq 1$ , the upper bound in (130) and the lower bound when receivers treat the interference as noise are plotted in Figure 8. We observe that up to a certain value of  $a$ , the upper bound coincides with the lower bound.

### B. New Outer Bound

For the weak Gaussian IC, there are two outer bounds that are tighter than the other known bounds. The first one, due to Kramer [12], is obtained by relying on the fact that the capacity region of the Gaussian IC is inside the capacity regions of the two underlying one-sided Gaussian ICs. Even though the capacity region of the one-sided Gaussian IC is unknown, there exists an outer bound for this channel that can be used instead. Kramers' outer bound is the intersection of two regions  $E_1$

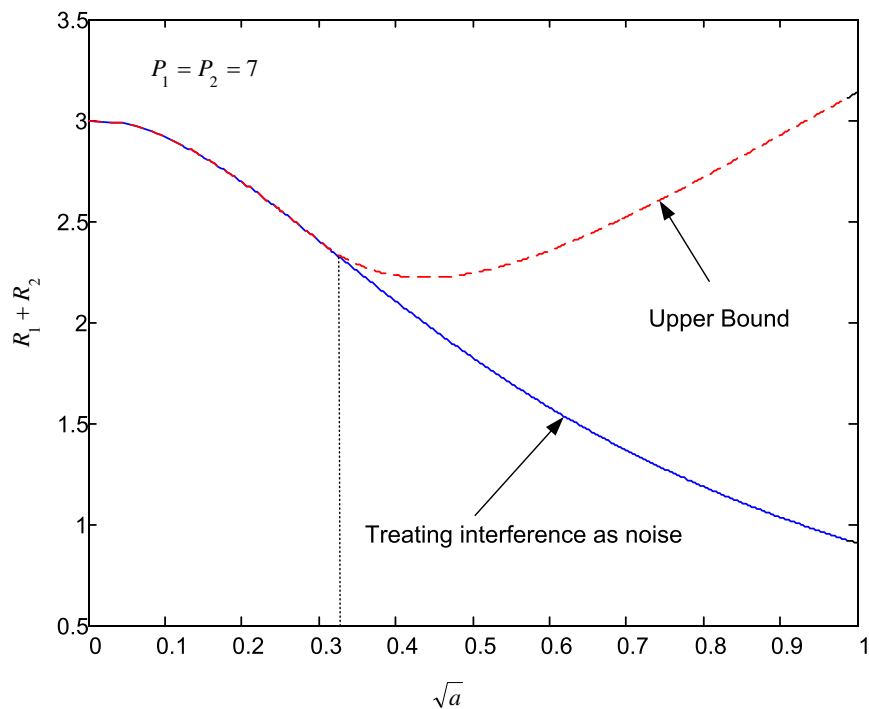


Fig. 8. The upper bound obtained by solving (130). The lower bound is obtained by treating the interference as noise.

and  $E_2$ .  $E_1$  is the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{(1-\beta)P'}{\beta P' + 1/a} \right), \quad (147)$$

$$R_2 \leq \gamma(\beta P'), \quad (148)$$

for all  $\beta \in [0, \beta_{\max}]$ , where  $P' = P_1/a + P_2$  and  $\beta_{\max} = \frac{P_2}{P'(1+P_1)}$ . Similarly,  $E_2$  is the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma(\alpha P''), \quad (149)$$

$$R_2 \leq \gamma \left( \frac{(1-\alpha)P''}{\alpha P'' + 1/b} \right), \quad (150)$$

for all  $\alpha \in [0, \alpha_{\max}]$ , where  $P'' = P_1 + P_2/b$  and  $\alpha_{\max} = \frac{P_1}{P''(1+P_2)}$ .

The second outer bound, due to Etkin *et al.* [13], is obtained by using Genie aided technique to upper bound different linear combinations of rates that appear in the HK achievable region. Their outer bound is the union of all rate pairs  $(R_1, R_2)$

satisfying

$$R_1 \leq \gamma(P_1), \quad (151)$$

$$R_2 \leq \gamma(P_2), \quad (152)$$

$$R_1 + R_2 \leq \gamma(P_1) + \gamma\left(\frac{P_2}{1 + bP_1}\right), \quad (153)$$

$$R_1 + R_2 \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1 + aP_2}\right), \quad (154)$$

$$R_1 + R_2 \leq \gamma\left(aP_2 + \frac{P_1}{1 + bP_1}\right) + \gamma\left(bP_1 + \frac{P_2}{1 + aP_2}\right), \quad (155)$$

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma\left(bP_1 + \frac{P_2}{1 + aP_2}\right) + 0.5 \log\left(\frac{1 + P_1}{1 + bP_1}\right), \quad (156)$$

$$R_1 + 2R_2 \leq \gamma(bP_1 + P_2) + \gamma\left(aP_2 + \frac{P_1}{1 + bP_1}\right) + 0.5 \log\left(\frac{1 + P_2}{1 + aP_2}\right). \quad (157)$$

In the outer bound proposed here, we derive an upper bound on all linear combinations of the rates. Recall that to obtain the boundary points of the capacity region  $\mathcal{C}$ , it suffices to calculate  $\sigma_{\mathcal{C}}(\mu, 1)$  and  $\sigma_{\mathcal{C}}(1, \mu)$  for all  $1 \leq \mu$ . To this end, we make use of channels in A1 and B classes and channels in A2 and B classes to obtain upper bounds on  $\sigma_{\mathcal{C}}(\mu, 1)$  and  $\sigma_{\mathcal{C}}(1, \mu)$ , respectively.

In order to obtain an upper bound on  $\sigma_{\mathcal{C}}(\mu, 1)$ , we introduce two optimization problems as follows. The first optimization problem is written as

$$W_1(\mu) = \min \frac{\mu_1}{2} \log[2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}}\right) + \mu_2 f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right) + f_h(P_2, N_{22}, 1, a, \mu_1) \quad (158)$$

subject to:

$$\mu_1 + \mu_2 = \mu$$

$$b'g_2 = b$$

$$b' \leq N_{21}$$

$$aN_{22} \leq 1$$

$$(1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1$$

$$0 \leq [\mu_1, \mu_2, b', g_2, N_{22}, N_{21}].$$

In fact, the objective of the above minimization problem is an upper bound on the support function of a channel within Class A1 which is obtained in Lemma 3. The constraints are the combination of (78) and (79) which are applied to guarantee the admissibility of the channel and to validate the upper bound obtained in Lemma 3. Hence,  $\sigma_{\mathcal{C}}(\mu, 1) \leq W_1(\mu)$ . By using a new variable  $S = (1 - \sqrt{g_2})^2 N_{22}$ , we obtain

$$W_1(\mu) = \min \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] + \frac{1}{2} \log \left[ (1 - \sqrt{g_2})^2 \left( \frac{1 - S + bP_1}{g_2 S} + \frac{P_2}{(1 - \sqrt{g_2})^2 P_2 + S} \right) \right] \quad (159)$$

$$+ \mu_2 f_h \left( P_1, 1, \frac{1 - S}{g_2}, \frac{b}{g_2}, \frac{1}{\mu_2} \right) + f_h(P_2, \frac{S}{(1 - \sqrt{g_2})^2}, 1, a, \mu_1) - \frac{\mu_2}{2} \log(2\pi e)$$

subject to:

$$\begin{aligned} \mu_1 + \mu_2 &= \mu \\ S &\leq 1 - b \\ S &\leq \frac{(1 - \sqrt{g_2})^2}{a} \\ 0 &\leq [\mu_1, \mu_2, S, g_2]. \end{aligned}$$

The second optimization problem is written as

$$W_2(\mu) = \min \mu \gamma \left( \frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left( \frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) + f_h(P_2, N_{22}, N_{12}, a', \mu) \quad (160)$$

$$+ \frac{\mu}{2} \log((2\pi e)(a'P_2 + N_{12})) - \frac{1}{2} \log((2\pi e)(P_2 + N_{22}))$$

subject to:

$$\begin{aligned} a'g_1 &= a \\ b'g_2 &= b \\ b'N_{11} &\leq N_{21} \\ a'N_{22} &\leq N_{12} \\ (1 - \sqrt{g_1})^2 N_{11} + g_1 N_{12} &= 1 \\ (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1 \\ 0 &\leq [a', b', g_1, g_2, N_{11}, N_{12}, N_{22}, N_{21}]. \end{aligned}$$

For this problem, Class B channels are used. In fact, the objective is the upper bound on the support function of channels within the class obtained in Lemma 5 and the constraints are defined to obtain the closed form formula for the upper bound and to confirm that the channels are admissible. Hence, we deduce  $\sigma_{\mathcal{C}}(\mu, 1) \leq W_2(\mu)$ . By using new variables  $S_1 = g_1 N_{12}$  and  $S_2 = g_2 N_{21}$ , we obtain

$$W_2(\mu) = \min \mu \gamma \left( \frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) + \gamma \left( \frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) \quad (161)$$

$$+ f_h \left( P_2, \frac{1 - S_2}{(1 - \sqrt{g_2})^2}, \frac{S_1}{g_1}, \frac{a}{g_1}, \mu \right) + \frac{\mu}{2} \log \left( (2\pi e) \left( \frac{aP_2 + S_1}{g_1} \right) \right) - \frac{1}{2} \log \left( (2\pi e) \left( P_2 + \frac{1 - S_2}{(1 - \sqrt{g_2})^2} \right) \right)$$

subject to:

$$\begin{aligned} \frac{b(1 - S_1)}{(1 - \sqrt{g_1})^2} &\leq S_2 < 1 \\ \frac{a(1 - S_2)}{(1 - \sqrt{g_2})^2} &\leq S_1 < 1 \\ 0 &< [g_1, g_2]. \end{aligned}$$

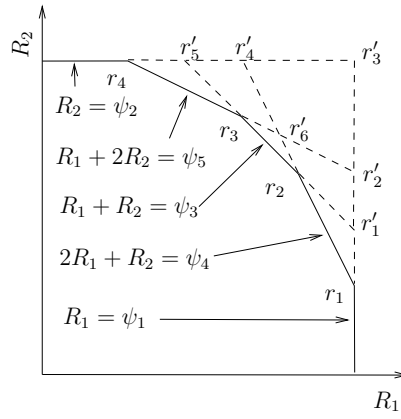


Fig. 9.  $\mathcal{G}_0$  for the weak Gaussian IC.  $r_1, r_2, r_3,$  and  $r_4$  are extreme points of  $\mathcal{G}_0$  in the interior of the first quadrant.

In a similar fashion, one can introduce two other optimization problems, say  $\tilde{W}_1(\mu)$  and  $\tilde{W}_2(\mu)$ , to obtain upper bounds on  $\sigma_{\mathcal{G}}(1, \mu)$  by using the upper bounds on the support functions of channels in Class A2 and Class B.

*Theorem 6 (New Outer Bound):* For any rate pair  $(R_1, R_2)$  achievable for the two-user weak Gaussian IC, the inequalities

$$\mu_1 R_1 + R_2 \leq W(\mu_1) = \min\{W_1(\mu_1), W_2(\mu_1)\}, \quad (162)$$

$$R_1 + \mu_2 R_2 \leq \tilde{W}(\mu_2) = \min\{\tilde{W}_1(\mu_2), \tilde{W}_2(\mu_2)\}, \quad (163)$$

hold for all  $1 \leq \mu_1, \mu_2$ .

To obtain an upper bound on the sum rate, we can apply the following inequality:

$$\mathcal{C}_{\text{sum}} \leq \min_{1 \leq \mu_1, \mu_2} \frac{(\mu_2 - 1)W(\mu_1) + (\mu_1 - 1)\tilde{W}(\mu_2)}{\mu_1 \mu_2 - 1}. \quad (164)$$

### C. Han-Kobayashi Achievable region

In this sub-section, we aim at characterizing  $\mathcal{G}$  for the weak Gaussian IC. To this end, we first investigate some properties of  $\mathcal{G}_0(P_1, P_2, \alpha, \beta)$ . First of all, we show that none of the inequalities in describing  $\mathcal{G}_0$  is redundant. In Figure 9, all possible extreme points are shown. It is easy to prove that  $r'_i \notin \mathcal{G}_0$  for  $i \in \{1, 2, \dots, 6\}$ . For instance, we consider  $r'_6 = \left(\frac{2\psi_4 - \psi_5}{3}, \frac{2\psi_5 - \psi_4}{3}\right)$ . Since  $\psi_{31} + \psi_{32} + \psi_{33} = \psi_4 + \psi_5$  (see Section II.C), we have

$$\begin{aligned} \psi_3 &= \min\{\psi_{31}, \psi_{32}, \psi_{33}\} \\ &\leq \frac{1}{3}(\psi_{31} + \psi_{32} + \psi_{33}) \\ &= \frac{1}{3}(\psi_4 + \psi_5). \end{aligned}$$

However,  $\frac{1}{3}(\psi_4 + \psi_5)$  is the sum of the components of  $r'_6$ . Therefore,  $r'_6$  violates (7) in the definition of the HK achievable region. Hence,  $r'_6 \notin \mathcal{G}_0$ . As another example, let us consider  $r'_1 = (\psi_1, \psi_3 - \psi_1)$ . We claim that  $r'_1$  violates (8). To this end, we need to show that  $\psi_4 \leq \psi_3 + \psi_1$ . However, it is easy to see that  $\psi_4 \leq \psi_{31} + \psi_1$ ,  $\psi_4 \leq \psi_{32} + \psi_1$ , and  $\psi_4 \leq \psi_{33} + \psi_1$  reduce to  $0 \leq (1 - \alpha)(1 - b + \beta(1 - ab)P_2)$ ,  $0 \leq (1 - \beta)(1 - a + (1 - ab)P_1)$ , and  $0 \leq (1 - \alpha)(1 - \beta)aP_2$ , respectively. Therefore,  $r'_1 \notin \mathcal{G}_0$ .

We conclude that  $\mathcal{G}$  has four extreme points in the interior of the first quadrant, namely

$$r_1 = (\psi_1, \psi_4 - 2\psi_1), \quad (165)$$

$$r_2 = (\psi_4 - \psi_3, 2\psi_3 - \psi_4), \quad (166)$$

$$r_3 = (2\psi_3 - \psi_5, \psi_5 - \psi_3), \quad (167)$$

$$r_4 = (\psi_5 - 2\psi_2, \psi_2). \quad (168)$$

Most importantly,  $\mathcal{G}_0$  possesses the unique minimizer property. To prove this, we need to show that  $\hat{\mathbf{y}}$ , the minimizer of the optimization problem

$$\begin{aligned} \sigma_{D_0}(c_1, c_2, P_1, P_2, \alpha, \beta) &= \max\{c_1 R_1 + c_2 R_2 \mid \mathbf{AR} \leq \Psi(P_1, P_2, \alpha, \beta)\} \\ &= \min\{\mathbf{y}^t \Psi(P_1, P_2, \alpha, \beta) \mid A^t \mathbf{y} = (c_1, c_2)^t, \mathbf{y} \geq 0\}, \end{aligned} \quad (169)$$

is independent of the parameters  $P_1, P_2, \alpha,$  and  $\beta$  and only depends on  $c_1$  and  $c_2$ . We first consider the case  $(c_1, c_2) = (\mu, 1)$  for all  $1 \leq \mu$ . It can be shown that for  $2 < \mu$ , the maximum of (169) is attained at  $r_1$  regardless of  $P_1, P_2, \alpha,$  and  $\beta$ . Therefore, the dual program has the minimizer  $\hat{\mathbf{y}} = (\mu - 2, 0, 0, 1, 0)^t$  which is clearly independent of  $P_1, P_2, \alpha,$  and  $\beta$ . In this case, we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (\mu - 2)\psi_1 + \psi_4, \quad 2 < \mu. \quad (170)$$

For  $1 \leq \mu \leq 2$ , one can show that  $r_2$  and  $\hat{\mathbf{y}} = (0, 0, 2 - \mu, \mu - 1, 0)^t$  are the maximizer and the minimizer of (169), respectively. In this case, we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (2 - \mu)\psi_3 + (\mu - 1)\psi_4, \quad 1 \leq \mu \leq 2. \quad (171)$$

Next, we consider the case  $(c_1, c_2) = (1, \mu)$  for all  $1 \leq \mu$ . Again, it can be shown that for  $2 < \mu$  and  $1 \leq \mu \leq 2$ ,  $\hat{\mathbf{y}} = (0, \mu - 2, 0, 0, 1)^t$  and  $\hat{\mathbf{y}} = (0, 0, 2 - \mu, 0, \mu - 1)^t$  minimizes (169), respectively. Hence, we have

$$\sigma_{D_0}(1, \mu, P_1, P_2, \alpha, \beta) = (\mu - 2)\psi_2 + \psi_5, \quad \text{if } 2 < \mu, \quad (172)$$

$$\sigma_{D_0}(1, \mu, P_1, P_2, \alpha, \beta) = (2 - \mu)\psi_3 + (\mu - 1)\psi_5, \quad \text{if } 1 \leq \mu \leq 2. \quad (173)$$

We conclude that the solutions of the dual program are always independent of  $P_1, P_2, \alpha,$  and  $\beta$ . Hence,  $\mathcal{G}_0$  possesses the unique minimizer property.

*Theorem 7:* For the two-user weak Gaussian IC, time-sharing and concavification result in the same region. In other words,  $\mathcal{G}$  can be fully characterized by using TD/FD and allocating power over three different dimensions.

*Proof:* Since  $\mathcal{G}_0$  possesses the unique minimizer property, from Theorem 1, we deduce that  $\mathcal{G} = \mathcal{G}_2$ . Moreover, using Theorem 3, the number of frequency bands is at most three.  $\square$

To obtain the support function of  $\mathcal{G}_2$ , we need to obtain  $g(c_1, c_2, P_1, P_2, \alpha, \beta)$  defined in (43). Since  $\mathcal{G}_0$  possesses the unique minimizer property, (43) can be simplified. Let us consider the case where  $(c_1, c_2) = (\mu, 1)$  for  $\mu > 2$ . It can be shown that for this case

$$g = \max_{(\alpha, \beta) \in [0, 1]^2} (\mu - 2)\psi_1(P_1, P_2, \alpha, \beta) + \psi_4(P_1, P_2, \alpha, \beta). \quad (174)$$

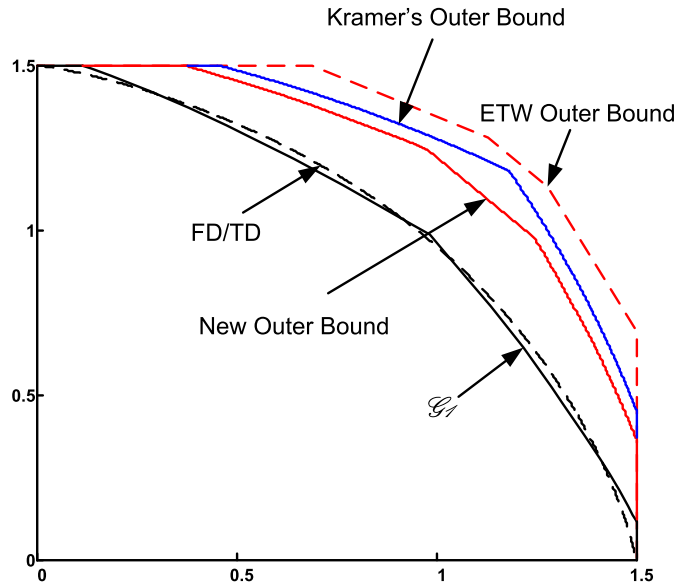


Fig. 10. Comparison between different bounds for the symmetric weak Gaussian IC when  $P = 7$  and  $a = 0.2$ .

Substituting into (42), we obtain

$$\sigma_{\mathcal{G}_2}(\mu, 1, P_1, P_2) = \max \sum_{i=1}^3 \lambda_i [(\mu - 2)\psi_1(P_{1i}, P_{2i}, \alpha_i, \beta_i) + \psi_4(P_{1i}, P_{2i}, \alpha_i, \beta_i)] \quad (175)$$

subject to:

$$\sum_{i=1}^3 \lambda_i = 1$$

$$\sum_{i=1}^3 \lambda_i P_{1i} \leq P_1$$

$$\sum_{i=1}^3 \lambda_i P_{2i} \leq P_2$$

$$0 \leq \lambda_i, 0 \leq P_{1i}, 0 \leq P_{2i}, \forall i \in \{1, 2, 3\}$$

$$0 \leq \alpha_i \leq 1, 0 \leq \beta_i \leq 1, \forall i \in \{1, 2, 3\}.$$

For other ranges of  $(c_1, c_2)$ , a similar optimization problem can be formed. It is worth noting that even though the number of parameters in characterizing  $\mathcal{G}$  is reduced, it is still prohibitively difficult to characterize boundary points of  $\mathcal{G}$ . In Figures (10) and (11), different bounds for the symmetric weak Gaussian IC are plotted. As shown in these figures, the new outer bound is tighter than the previously known bounds.

## V. ONE-SIDED GAUSSIAN INTERFERENCE CHANNELS

Throughout this section, we consider the one-sided Gaussian IC obtained by setting  $b = 0$ , i.e., the second receiver incurs no interference from the first transmitter. One can further split the class of one-sided ICs into two subclasses: the *strong one-sided IC* and the *weak one-sided IC*. For the former,  $a \geq 1$  and the capacity region is fully characterized [16]. In this case, the

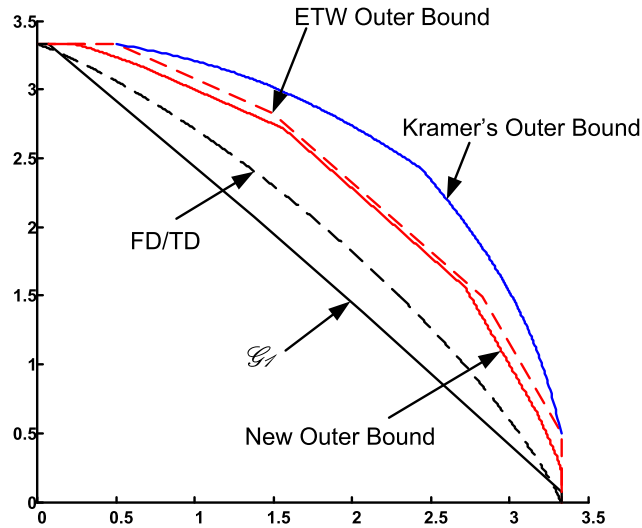


Fig. 11. Comparison between different bounds for the symmetric weak Gaussian IC when  $P = 100$  and  $a = 0.1$ .

capacity region is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$\begin{aligned} R_1 &\leq \gamma(P_1), \\ R_2 &\leq \gamma(P_2), \\ R_1 + R_2 &\leq \gamma(P_1 + aP_2). \end{aligned}$$

For the latter,  $a < 1$  and the full characterization of the capacity region is still an open problem. Therefore, we always assume  $a < 1$ . Three important results are proved for this channel. The first one, proved by Costa in [11], states that the capacity region of the weak one-sided IC is equivalent to that of the degraded IC with an appropriate change of parameters. The second one, proved by Sato in [10], states that the capacity region of the degraded Gaussian IC is outer bounded by the capacity region of a certain degraded broadcast channel. The third one, proved by Sason in [16], characterizes the sum capacity by combining Costa's and Sato's results.

In this section, we provide an alternative proof for the outer bound obtained by Sato. We then characterize the full HK achievable region where Gaussian codebooks are used, i.e.,  $\mathcal{G}$ .

#### A. Sum Capacity

For the sake of completeness, we first state the sum capacity result obtained by Sason.

*Theorem 8 (Sason):* The rate pair  $\left(\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma(P_2)\right)$  is an extreme point of the capacity region of the one-sided Gaussian IC. Moreover, the sum capacity of the channel is attained at this point.

#### B. Outer Bound

In [10], Sato derived an outer bound on the capacity of the degraded IC. This outer bound can be used for the weak one-sided IC as well. This is due to Costa's result which states that the capacity region of the degraded Gaussian IC is equivalent to that of the weak one-sided IC with an appropriate change of parameters.

*Theorem 9 (Sato):* If the rate pair  $(R_1, R_2)$  belongs to the capacity region of the weak one-sided IC, then it satisfies

$$\begin{aligned} R_1 &\leq \gamma\left(\frac{(1-\beta)P}{1/a+\beta P}\right), \\ R_2 &\leq \gamma(\beta P), \end{aligned} \quad (176)$$

for all  $\beta \in [0, 1]$  where  $P = P_1/a + P_2$ .

*Proof:* Since the sum capacity is attained at the point where User 2 transmits at its maximum rate  $R_2 = \gamma(P_2)$ , other boundary points of the capacity region can be obtained by characterizing the solutions of  $\sigma_{\mathcal{C}}(\mu, 1) = \max\{\mu R_1 + R_2 | (R_1, R_2) \in \mathcal{C}\}$  for all  $1 \leq \mu$ . Using Fano's inequality, we have

$$\begin{aligned} n(\mu R_1 + R_2) &\leq \mu I(x_1^n; y_1^n) + I(x_2^n; y_2^n) + n\epsilon_n \\ &= \mu h(y_1^n) - \mu h(y_1^n | x_1^n) + h(y_2^n) - h(y_2^n | x_2^n) + n\epsilon_n \\ &= [\mu h(x_1^n + \sqrt{a}x_2^n + z_1^n) - h(z_1^n)] + [h(x_2^n + z_2^n) - \mu h(\sqrt{a}x_2^n + z_1^n)] + n\epsilon_n \\ &\stackrel{(a)}{\leq} \frac{\mu n}{2} \log[2\pi e(P_1 + aP_2 + 1)] - \frac{n}{2} \log(2\pi e) + [h(x_2^n + z_2^n) - \mu h(\sqrt{a}x_2^n + z_1^n)] + n\epsilon_n \\ &\stackrel{(b)}{\leq} \frac{\mu n}{2} \log[2\pi e(P_1 + aP_2 + 1)] - \frac{n}{2} \log(2\pi e) + n f_h(P_2, 1, 1, a, \mu) + n\epsilon_n, \end{aligned}$$

where (a) follows from the fact that Gaussian distribution maximizes the differential entropy for a given constraint on the covariance matrix and (b) follows from the definition of  $f_h$  in (59).

Depending on the value of  $\mu$ , we consider the following two cases:

1- For  $1 \leq \mu \leq \frac{P_2+1/a}{P_2+1}$ , we have

$$\mu R_1 + R_2 \leq \mu \gamma\left(\frac{P_1}{1+aP_2}\right) + \gamma(P_2). \quad (177)$$

In fact, the point  $\left(\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma(P_2)\right)$  which is achievable by treating interference as noise at Receiver 1, satisfies (177) with equality. Therefore, it belongs to the capacity region. Moreover, by setting  $\mu = 1$ , we deduce that this point corresponds to the sum capacity of the one-sided Gaussian IC. This is in fact an alternative proof for Sason's result.

2- For  $\frac{P_2+1/a}{P_2+1} < \mu \leq \frac{1}{a}$ , we have

$$\mu R_1 + R_2 \leq \frac{\mu}{2} \log(P_1 + aP_2 + 1) + \frac{1}{2} \log\left(\frac{1/a - 1}{\mu - 1}\right) - \frac{\mu}{2} \log\left(\frac{a\mu(1/a - 1)}{\mu - 1}\right). \quad (178)$$

Equivalently, we have

$$\mu R_1 + R_2 \leq \frac{\mu}{2} \log\left(\frac{(aP+1)(\mu-1)}{\mu(1-a)}\right) + \frac{1}{2} \log\left(\frac{1/a-1}{\mu-1}\right), \quad (179)$$

where  $P = P_1/a + P_2$ . Let us define  $E_1$  as the set of all rate pairs  $(R_1, R_2)$  satisfying (179), i.e.

$$E_1 = \left\{ (R_1, R_2) | \mu R_1 + R_2 \leq \frac{\mu}{2} \log\left(\frac{(aP+1)(\mu-1)}{\mu(1-a)}\right) + \frac{1}{2} \log\left(\frac{1/a-1}{\mu-1}\right), \forall \frac{P_2+1/a}{P_2+1} < \mu \leq \frac{1}{a} \right\}. \quad (180)$$

We claim that  $E_1$  is the dual representation of the region defined in the statement of the theorem, see (4). To this end, we define  $E_2$  as

$$E_2 = \left\{ (R_1, R_2) | R_1 \leq \gamma\left(\frac{(1-\beta)P}{1/a+\beta P}\right), R_2 \leq \gamma(\beta P), \forall \beta \in [0, 1] \right\} \quad (181)$$

We evaluate the support function of  $E_2$  as

$$\sigma_{E_2}(\mu, 1) = \max\{\mu R_1 + R_2 | (R_1, R_2) \in E_2\}. \quad (182)$$

It is easy to show that  $\beta = \frac{1/a-1}{P(\mu-1)}$  maximizes the above optimization problem. Therefore, we have

$$\sigma_{E_2}(\mu, 1) = \frac{\mu}{2} \log \left( \frac{(aP+1)(\mu-1)}{\mu(1-a)} \right) + \frac{1}{2} \log \left( \frac{1/a-1}{\mu-1} \right). \quad (183)$$

Since  $E_2$  is a closed convex set, we can use (4) to obtain its dual representation which is indeed equivalent to (180). This completes the proof.  $\square$

### C. Han-Kobayashi Achievable Region

In this subsection, we characterize  $\mathcal{G}_0$ ,  $\mathcal{G}_1$ ,  $\mathcal{G}_2$ , and  $\mathcal{G}$  for the weak one-sided Gaussian IC.  $\mathcal{G}_0$  can be characterized as follows. Since there is no link between Transmitter 1 and Receiver 2, User 1's message in the HK achievable region is only the private message, i.e.,  $\alpha = 1$ . In this case, we have

$$\psi_1 = \gamma \left( \frac{P_1}{1+a\beta P_2} \right), \quad (184)$$

$$\psi_2 = \gamma(P_2), \quad (185)$$

$$\psi_{31} = \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (186)$$

$$\psi_{32} = \gamma \left( \frac{P_1}{1+a\beta P_2} \right) + \gamma(P_2), \quad (187)$$

$$\psi_{33} = \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (188)$$

$$\psi_4 = \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma \left( \frac{P_1}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (189)$$

$$\psi_5 = \gamma(\beta P_2) + \gamma(P_2) + \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right), \quad (190)$$

It is easy to show that  $\psi_3 = \min\{\psi_{31}, \psi_{32}, \psi_{33}\} = \psi_{31}$ ,  $\psi_{31} + \psi_1 = \psi_4$ ,  $\psi_{31} + \psi_2 = \psi_5$ . Hence,  $\mathcal{G}_0$  can be represented as all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1+a\beta P_2} \right), \quad (191)$$

$$R_2 \leq \gamma(P_2), \quad (192)$$

$$R_1 + R_2 \leq \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2). \quad (193)$$

We claim that  $\mathcal{G}_2 = \mathcal{G}$ . To prove this, we need to show that  $\mathcal{G}_0$  possesses the unique minimizer property.  $\mathcal{G}_0$  is a pentagon with two extreme points in the interior of the first quadrant, namely  $r_1$  and  $r_2$  where

$$r_1 = \left( \gamma \left( \frac{P_1}{1+a\beta P_2} \right), \gamma \left( \frac{(1-\beta)aP_2}{1+P_1+\beta aP_2} \right) + \gamma(\beta P_2) \right), \quad (194)$$

$$r_2 = \left( \gamma \left( \frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2) - \gamma(P_2), \gamma(P_2) \right). \quad (195)$$

Using above, it can be verified that  $\mathcal{G}_0$  possesses the unique minimizer property.

Next, we can use the optimization problem in (42) to obtain the support function of  $\mathcal{G}$ . However, we only need to consider  $(c_1, c_2) = (\mu, 1)$  for  $\mu > 1$ . Therefore, we have

$$g(\mu, 1, P_1, P_2, \beta) = \max_{0 \leq \beta \leq 1} \mu \gamma \left( \frac{P_1}{1+\beta a P_2} \right) + \gamma(\beta P_2) + \gamma \left( \frac{(1-\beta)aP_2}{1+P_1+\beta a P_2} \right). \quad (196)$$

Substituting into (42), we conclude that boundary points of  $\mathcal{G}$  can be characterized by solving the following optimization problem:

$$W = \max \sum_{i=1}^3 \lambda_i \left[ \mu \gamma \left( \frac{P_{1i}}{1 + \beta_i a P_{2i}} \right) + \gamma(\beta_i P_{2i}) + \gamma \left( \frac{(1 - \beta_i) a P_{2i}}{1 + P_{1i} + \beta_i a P_{2i}} \right) \right] \quad (197)$$

subject to:

$$\sum_{i=1}^3 \lambda_i = 1$$

$$\sum_{i=1}^3 \lambda_i P_{1i} \leq P_1$$

$$\sum_{i=1}^3 \lambda_i P_{2i} \leq P_2$$

$$0 \leq \beta_i \leq 1, \forall i \in \{1, 2, 3\}$$

$$0 \leq [P_{1i}, P_{2i}, \lambda_i], \forall i \in \{1, 2, 3\}.$$

For the sake of completeness, we provide a simple description for  $\mathcal{G}_1$  in the next lemma.

*Lemma 7:* The region  $\mathcal{G}_1$  can be represented as the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta' P_2} \right), \quad (198)$$

$$R_2 \leq \gamma(\beta' P_2) + \gamma \left( \frac{a(1 - \beta') P_2}{1 + P_1 + a\beta' P_2} \right), \quad (199)$$

for all  $\beta' \in [0, 1]$ . Moreover,  $\mathcal{G}_1$  is convex and any point that lies on its boundary can be achieved by using superposition coding and successive decoding.

*Proof:* Let  $E$  denote the set defined in the above lemma. It is easy to show that  $E$  is convex and  $E \subseteq \mathcal{G}_1$ . To prove the inverse inclusion, it suffices to show that the extreme points of  $\mathcal{G}_0$ ,  $r_1$  and  $r_2$  (see (194) and (195)) are inside  $E$  for all  $\beta \in [0, 1]$ . By setting  $\beta' = \beta$ , we see that  $r_1 \in E$ . To prove  $r_2 \in E$ , we set  $\beta' = 1$ . We conclude that  $r_2 \in E$  if the following inequality holds

$$\gamma \left( \frac{P_1 + a(1 - \beta) P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2) - \gamma(P_2) \leq \gamma \left( \frac{P_1}{1 + aP_2} \right), \quad (200)$$

for all  $\beta \in [0, 1]$ . However, (200) reduces to  $0 \leq (1 - a)(1 - \beta)P_2$  which holds for all  $\beta \in [0, 1]$ . Hence,  $\mathcal{G}_1 \subseteq E$ . Using these facts, it is straightforward to show that the boundary points  $\mathcal{G}_1$  are achievable by using superposition coding and successive decoding.  $\square$

Figure 12 compares different bounds for the one-sided Gaussian IC.

## VI. MIXED GAUSSIAN INTERFERENCE CHANNELS

In this section, we focus on the mixed Gaussian Interference channel. We first characterize the sum capacity of this channel. Then, we provide an outer bound on the capacity region. Finally, we investigate the HK achievable region. Without loss of generality, we assume  $a < 1$  and  $b \geq 1$ .

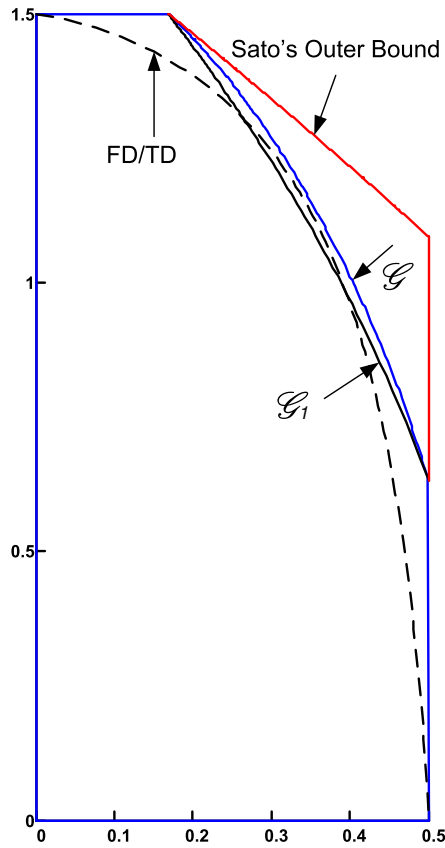


Fig. 12. Comparison between different bounds for the one-sided Gaussian IC when  $P_1 = 1$ ,  $P_2 = 7$ , and  $a = 0.4$ .

### A. Sum Capacity

*Theorem 10:* The sum capacity of the mixed Gaussian IC with  $a < 1$  and  $b \geq 1$  can be stated as

$$\mathcal{C}_{sum} = \gamma(P_2) + \min \left\{ \gamma \left( \frac{P_1}{1 + aP_2} \right), \gamma \left( \frac{bP_1}{1 + P_2} \right) \right\}. \quad (201)$$

*Proof:* We need to prove the achievability and converse for the theorem.

**Achievability part:** Transmitter 1 sends a common message to both receivers, while the first user's signal is considered as noise at both receivers. In this case, the rate

$$R_1 = \min \left\{ \gamma \left( \frac{P_1}{1 + aP_2} \right), \gamma \left( \frac{bP_1}{1 + P_2} \right) \right\} \quad (202)$$

is achievable. At Receiver 2, the signal from Transmitter 1 can be decoded and removed. Therefore, User 2 is left with a channel without interference and it can communicate at its maximum rate which is

$$R_2 = \gamma(P_2). \quad (203)$$

By adding (202) and (203), we obtain the desired result.

**Converse part:** The sum capacity of the Gaussian IC is upper bounded by that of the two underlying one-sided Gaussian ICs. Hence, we can obtain two upper bounds on the sum rate. We first remove the interfering link between Transmitter 1 and Receiver 2. In this case, we have a one-sided Gaussian IC with weak interference. The sum capacity of this channel is known

[16]. Hence, we have

$$\mathcal{C}_{sum} \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1+aP_2}\right). \quad (204)$$

By removing the interfering link between Transmitter 2 and Receiver 1, we obtain a one-sided Gaussian IC with strong interference. The sum capacity of this channel is known. Hence, we have

$$\mathcal{C}_{sum} \leq \gamma(bP_1 + P_2), \quad (205)$$

which equivalently can be written as

$$\mathcal{C}_{sum} \leq \gamma(P_2) + \gamma\left(\frac{bP_1}{1+P_2}\right). \quad (206)$$

By taking the minimum of the right hand sides of (204) and (206), we obtain

$$\mathcal{C}_{sum} \leq \gamma(P_2) + \min\left\{\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+P_2}\right)\right\}. \quad (207)$$

This completes the proof.  $\square$

*Remark 4:* In an independent work [25], the sum capacity of the mixed Gaussian IC is obtained for a certain range of parameters, whereas in the above theorem, we characterize the sum capacity of this channel for the entire range of its parameters (see also [24]).

By comparing  $\gamma\left(\frac{P_1}{1+aP_2}\right)$  with  $\gamma\left(\frac{bP_1}{1+P_2}\right)$ , we observe that if  $1+P_2 \leq b+abP_2$ , then the sum capacity corresponds to the sum capacity of the one-sided weak Gaussian IC, whereas if  $1+P_2 > b+abP_2$ , then the sum capacity corresponds to the sum capacity of the one-sided strong IC. Similar to the one-sided Gaussian IC, since the sum capacity is attained at the point where User 2 transmits at its maximum rate  $R_2 = \gamma(P_2)$ , other boundary points of the capacity region can be obtained by characterizing the solutions of  $\sigma_{\mathcal{C}}(\mu, 1) = \max\{\mu R_1 + R_2 | (R_1, R_2) \in \mathcal{C}\}$  for all  $1 \leq \mu$ .

### B. New Outer Bound

The best outer bound to date, due to Etkin *et al.* [13], is obtained by using the Genie aided technique. This bound is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma(P_1), \quad (208)$$

$$R_2 \leq \gamma(P_2), \quad (209)$$

$$R_1 + R_2 \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1+aP_2}\right), \quad (210)$$

$$R_1 + R_2 \leq \gamma(P_2 + bP_1), \quad (211)$$

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma\left(bP_1 + \frac{P_2}{1+aP_2}\right) + \gamma\left(\frac{P_1}{1+bP_1}\right). \quad (212)$$

The capacity region of the mixed Gaussian IC is inside the intersection of the capacity regions of the two underlying one-sided Gaussian ICs. Removing the link between Transmitter 1 and Receiver 2 results in a weak one-sided Gaussian IC whose outer bound  $E_1$  is the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma\left(\frac{(1-\beta)P'}{\beta P' + 1/a}\right), \quad (213)$$

$$R_2 \leq \gamma(\beta P'), \quad (214)$$

for all  $\beta \in [0, \beta_{\max}]$ , where  $P' = P_1/a + P_2$  and  $\beta_{\max} = \frac{P_2}{P'(1+P_1)}$ . On the other hand, removing the link between Transmitter 2 and Receiver 1 results in a strong one-sided Gaussian IC whose capacity region  $E_2$  is fully characterized as the collection of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma(bP_1), \quad (215)$$

$$R_2 \leq \gamma(P_2), \quad (216)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2). \quad (217)$$

Using the channels in Class C, we upper bound  $\sigma_{\mathcal{C}}(\mu, 1)$  based on the following optimization problem:

$$W(\mu) = \min \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log \left( 2\pi e \left( \frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21} \right) \right) \quad (218)$$

$$- \frac{1}{2} \log(2\pi e N_{21}) - \frac{1}{2} \log(2\pi e N_{22}) + f_h(P_2, N_{22}, 1, a, \mu - 1)$$

subject to:

$$b'g_2 = b$$

$$b' \geq N_{21}$$

$$aN_{22} \leq 1$$

$$(1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1$$

$$0 \leq [b', g_2, N_{22}, N_{21}].$$

By substituting  $S = g_2 N_{21}$ , we obtain

$$W(\mu) = \min \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log \left( 2\pi e \left( \frac{P_2(1-S)}{(1-\sqrt{g_2})^2 P_2 + 1 - S} + \frac{bP_1 + S}{g_2} \right) \right) \quad (219)$$

$$- \frac{1}{2} \log \left( \frac{2\pi e S}{g_2} \right) - \frac{1}{2} \log \left( \frac{2\pi e(1-S)}{(1-\sqrt{g_2})^2} \right) + f_h \left( P_2, \frac{1-S}{(1-\sqrt{g_2})^2}, 1, a, \mu - 1 \right)$$

subject to:

$$S < 1$$

$$a(1-S) \leq (1-\sqrt{g_2})^2$$

$$0 \leq [S, g_2].$$

Hence, we have the following theorem that provides an outer bound on the capacity region of the mixed Gaussian IC.

*Theorem 11:* For any rate pair  $(R_1, R_2)$  achievable for the two-user mixed Gaussian IC,  $(R_1, R_2) \in E_1 \cap E_2$ . Moreover, the inequality

$$\mu R_1 + R_2 \leq W(\mu) \quad (220)$$

holds for all  $1 \leq \mu$ .

### C. Han-Kobayashi Achievable Region

In this subsection, we study the HK achievable region for the mixed Gaussian IC. Since Receiver 2 can always decode the message of the first user, User 1 associates all its power to the common message. User 2, on the other hand, allocates  $\beta P_2$

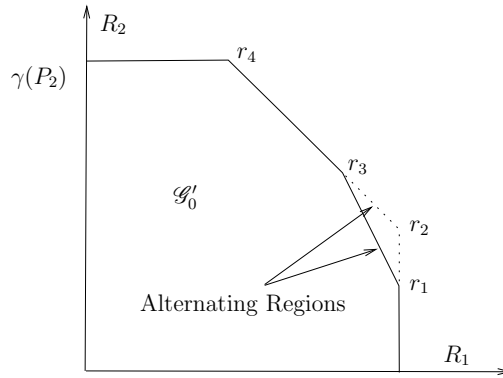


Fig. 13. The new region  $\mathcal{G}'_0$  which is obtained by enlarging  $\mathcal{G}_0$ .

and  $(1 - \beta)P_2$  of its total power to its private and common messages, respectively, where  $\beta \in [0, 1]$ . Therefore, we have

$$\psi_1 = \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (221)$$

$$\psi_2 = \gamma(P_2), \quad (222)$$

$$\psi_{31} = \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2), \quad (223)$$

$$\psi_{32} = \gamma(P_2 + bP_1), \quad (224)$$

$$\psi_{33} = \gamma \left( \frac{a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1), \quad (225)$$

$$\psi_4 = \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1), \quad (226)$$

$$\psi_5 = \gamma(\beta P_2) + \gamma(P_2 + bP_1) + \gamma \left( \frac{a(1 - \beta)P_2}{1 + a\beta P_2} \right). \quad (227)$$

Due to the fact that the sum capacity is attained at the point where the second user transmits at its maximum rate, the last inequality in the description of the HK achievable region can be removed. Although the point  $r'_5 = (\psi_3 - \gamma(P_2), \gamma(P_1))$  in Figure 9 may not be in  $\mathcal{G}_0$ , this point is always achievable due to the sum capacity result. Hence, we can enlarge  $\mathcal{G}_0$  by removing  $r_3$  and  $r_4$ . Let us denote the resulting region as  $\mathcal{G}'_0$ . Moreover, one can show that  $r'_2, r'_3, r'_4$ , and  $r'_6$  are still outside  $\mathcal{G}'_0$ . However, for the mixed Gaussian IC, it is possible that  $r'_1$  belongs to  $\mathcal{G}'_0$ . In Figure 13, two alternative cases for the region  $\mathcal{G}'_0$  along with the new labeling of its extreme points are plotted. The new extreme points can be written as

$$r_1 = (\psi_1, \psi_4 - 2\psi_1),$$

$$r_2 = (\psi_1, \psi_3 - \psi_1),$$

$$r_3 = (\psi_4 - \psi_3, 2\psi_3 - \psi_4),$$

$$r_4 = (\psi_3 - \psi_2, \psi_2).$$

In fact, we have either  $\mathcal{G}'_0 = \text{conv}\{r_1, r_3, r_4\}$  or  $\mathcal{G}'_0 = \text{conv}\{r_2, r_4\}$ .

To simplify the characterization of  $\mathcal{G}_1$ , we consider three cases:

Case I:  $1 + P_2 \leq b + abP_2$ .

Case II:  $1 + P_2 > b + abP_2$  and  $1 - a \leq abP_1$ .

Case III:  $1 + P_2 > b + abP_2$  and  $1 - a > abP_1$ .

**Case I** ( $1 + P_2 \leq b + abP_2$ ): In this case,  $\psi_3 = \psi_{31}$ . Moreover, it is easy to verify that  $\psi_{31} + \psi_1 \leq \psi_4$  which means (8) is redundant for the entire range of parameters. Hence,  $\mathcal{G}'_0 = \text{conv}\{r_2, r_4\}$  consists of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (228)$$

$$R_2 \leq \gamma(P_2), \quad (229)$$

$$R_1 + R_2 \leq \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2), \quad (230)$$

where  $\beta \in [0, 1]$ . Using a reasoning similar to the one used to express boundary points of  $\mathcal{G}_1$  for the one-sided Gaussian IC, we can express boundary points of  $\mathcal{G}_1$  as

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (231)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left( \frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right), \quad (232)$$

for all  $\beta \in [0, 1]$ .

*Theorem 12:* For the mixed Gaussian IC satisfying  $1 \leq ab$ , region  $\mathcal{G}$  is equivalent to that of the one sided Gaussian IC obtained from removing the interfering link between Transmitter 1 and Receiver 2.

*Proof:* If  $1 \leq ab$ , then  $1 + P_2 \leq b + abP_2$  holds for all  $P_1$  and  $P_2$ . Hence,  $\mathcal{G}'_0(P_1, P_2, \beta)$  is a pentagon defined by (228), (229), and (230). Comparing with the corresponding region for the one-sided Gaussian IC, we see that  $\mathcal{G}'_0$  is equivalent to  $\mathcal{G}_0$  obtained for the one-sided Gaussian IC. This directly implies that  $\mathcal{G}$  is the same for both channels.  $\square$

**Case II** ( $1 + P_2 > b + abP_2$  and  $1 - a \leq abP_1$ ): In this case,  $\psi_3 = \min\{\psi_{31}, \psi_{32}\}$ . It can be shown that  $\mathcal{G}_1$  is the union of three regions  $E_1$ ,  $E_2$ , and  $E_3$ , i.e.,  $\mathcal{G}_0 = E_1 \cup E_2 \cup E_3$ . Region  $E_1$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (233)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left( \frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right). \quad (234)$$

for all  $\beta \in [0, \frac{b-1}{(1-ab)P_2}]$ . Region  $E_2$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{bP_1}{1 + \beta P_2} \right), \quad (235)$$

$$R_2 \leq \gamma \left( \frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2) - \gamma \left( \frac{bP_1}{1 + \beta P_2} \right). \quad (236)$$

for all  $\beta \in [\frac{b-1}{(1-ab)P_2}, \frac{(b-1)P_1 + (1-a)P_2}{(1-ab)P_1 P_2 + (1-a)P_2}]$ . Region  $E_3$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{bP_1(1 + \frac{(1-ab)P_1}{1-a})}{1 + bP_1 + P_2} \right), \quad (237)$$

$$R_2 \leq \gamma(P_2), \quad (238)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2). \quad (239)$$

**Case III** ( $1 + P_2 > b + abP_2$  and  $1 - a > abP_1$ ): In this case,  $\psi_3 = \min\{\psi_{31}, \psi_{32}\}$ . Similar to Case II, we have  $\mathcal{G}_1 = E_1 \cup E_2 \cup E_3$ , where regions  $E_1$ ,  $E_2$ , and  $E_3$  are defined as follows. Region  $E_1$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (240)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left( \frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right). \quad (241)$$

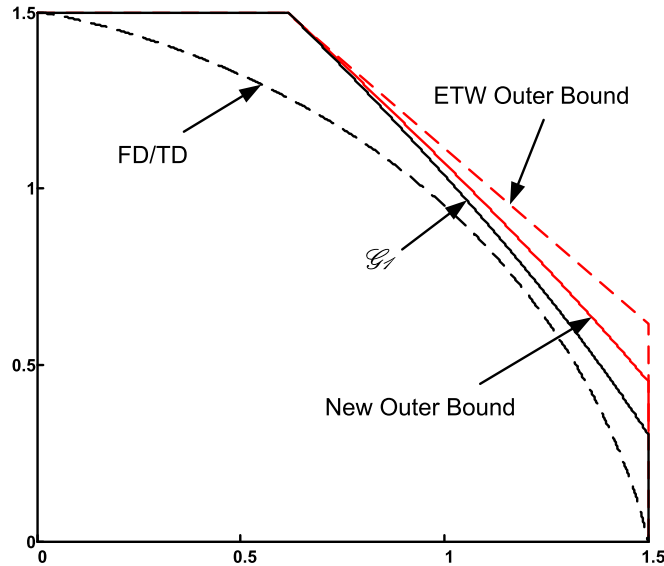


Fig. 14. Comparison between different bounds for the mixed Gaussian IC when  $1 + P_2 \leq b + abP_2$  (Case I) for  $P_1 = 7$ ,  $P_2 = 7$ ,  $a = 0.6$ , and  $b = 2$ .

for all  $\beta \in [0, \frac{b-1}{(1-ab)P_2}]$ . Region  $E_2$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (242)$$

$$R_2 \leq \gamma \left( \frac{a(1-\beta)P_2}{1 + P_1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1) - \gamma \left( \frac{P_1}{1 + a\beta P_2} \right). \quad (243)$$

for all  $\beta \in [\frac{b-1}{(1-ab)P_2}, 1]$ . Region  $E_3$  is the union of all rate pairs  $(R_1, R_2)$  satisfying

$$R_1 \leq \gamma \left( \frac{P_1}{1 + aP_2} \right), \quad (244)$$

$$R_2 \leq \gamma(P_2), \quad (245)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2). \quad (246)$$

*Remark 5:* Region  $E_3$  in Case II and Case III represents a facet that belongs to the capacity region of the mixed Gaussian IC. It is important to note that, surprisingly, this facet is obtainable when the second transmitter uses both the common message and the private message.

Different bounds are compared for the mixed Gaussian IC for Cases I, II, and III in Figures 14, 15, and 16, respectively.

## VII. CONCLUSION

We have studied the capacity region of the two-user Gaussian IC. The sum capacities, inner bounds, and outer bounds have been considered for three classes of channels: weak, one-sided, and mixed Gaussian IC. We have used admissible channels as the main tool for deriving outer bounds on the capacity regions.

For the weak Gaussian IC, we have derived the sum capacity for a certain range of channel parameters. In this range, the sum capacity is attained when Gaussian codebooks are used and interference is treated as noise. Moreover, we have derived a new outer bound on the capacity region. This outer bound is tighter than the Kramer's bound and the ETW's bound. Regarding inner bounds, we have reduced the computational complexity of the HK achievable region. In fact, we have shown that when

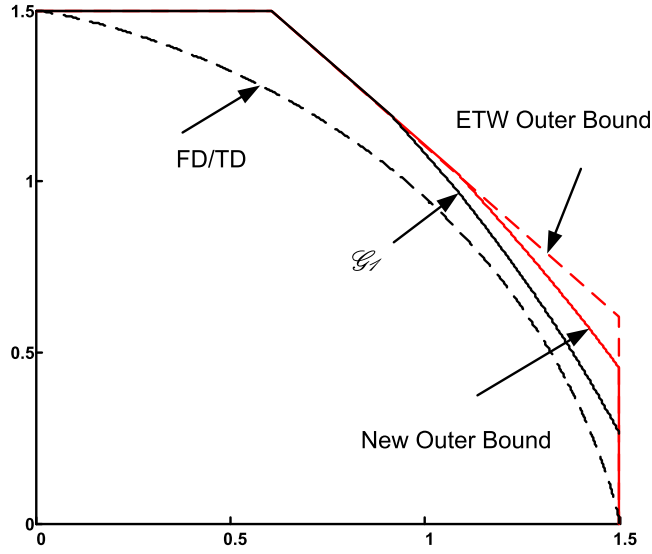


Fig. 15. Comparison between different bounds for the mixed Gaussian IC when  $1 + P_2 > b + abP_2$  and  $1 - a \leq abP_1$  (Case II) for  $P_1 = 7$ ,  $P_2 = 7$ ,  $a = 0.4$ , and  $b = 1.5$ .

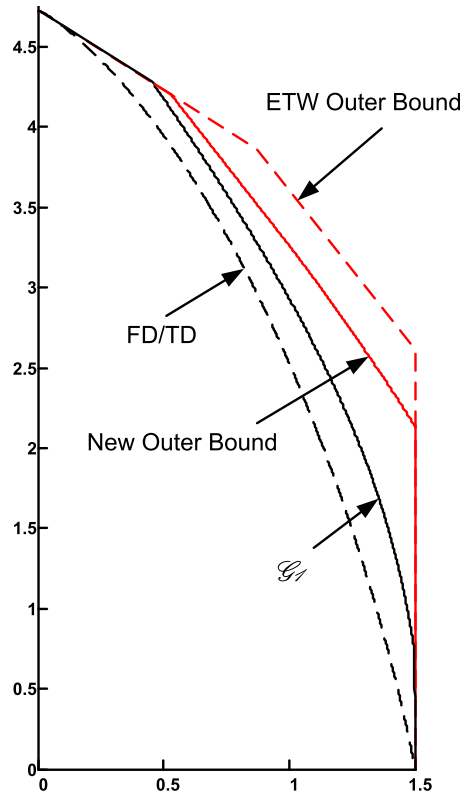


Fig. 16. Comparison between different bounds for the mixed Gaussian IC when  $1 + P_2 > b + abP_2$  and  $1 - a > abP_1$  (Case III) for  $P_1 = 7$ ,  $P_2 = 700$ ,  $a = 0.01$ , and  $b = 1.5$ .

Gaussian codebooks are used, the full HK achievable region can be obtained by using the naive HK achievable scheme over three frequency bands.

For the one-sided Gaussian IC, we have presented an alternative proof for the Sato's outer bound. We have also derived the full HK achievable region when Gaussian codebooks are used.

For the mixed Gaussian IC, we have derived the sum capacity for the entire range of its parameters. Moreover, we have presented a new outer bound on the capacity region that outperforms ETW's bound. We have proved that the full HK achievable region using Gaussian codebooks is equivalent to that of the one-sided Gaussian IC for a particular range of channel gains. We have also derived a facet that belongs to the capacity region for a certain range of parameters. Surprisingly, this facet is obtainable when one of the transmitters uses both the common message and the private message.

## REFERENCES

- [1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. on Mathematical Statistics and Probability*, vol. 1, 1961, pp. 611–644.
- [2] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 569–570, Sept. 1975.
- [3] H. Sato, "The capacity of the gaussian interference channel under strong interference," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 786–788, Nov. 1981.
- [4] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd International Symp. on Information theory*, U. Tsahkadzor, Armenia, Ed., Sep 2-8 1971, pp. 23–52.
- [5] I. Csiszár and J. Körner, *Information Theory: Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Hungarian Acad. Sci., 1981.
- [6] R. S. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE Trans. Inform. Theory*, vol. 39, pp. 609–612, Mar. 1993.
- [7] T. Cover and J. Thomas, *Elements of information theory*. NY, John Wiley, 1991.
- [8] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 49–60, Jan. 1981.
- [9] H. Chong, M. Motani, H. Garg, and H. E. Gamal, "On the Han-Kobayashi region for the interference channel," *Submitted to the IEEE Trans. on Inf.*, Aug. 2006.
- [10] H. Sato, "On degraded gaussian two-user channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 637–640, Sept. 1978.
- [11] M. H. M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 607–615, Sept. 1985.
- [12] G. Kramer, "Outer bounds on the capacity of gaussian interference channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 581–586, Mar. 2004.
- [13] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit." *submitted to the IEEE Transactions on Information Theory*. Available at <http://www.eecs.berkeley.edu/~dtse/pub.html>, Feb. 2007.
- [14] T. Liu and P. Viswanath, "An extremal inequality motivated by multi terminal information theoretic problems," in *2006 International Symposium on Information Theory (ISIT)*, Seattle, WA, July 2006, pp. 1016–1020.
- [15] A. B. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 60–70, Jan. 1978.
- [16] I. Sason, "On achievable rate regions for the gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1345–1356, June 2004.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [18] T. Han, "The capacity region of general multiple-access channel with certain correlated sources," *Inform. Contr.*, vol. 40, no. 1, pp. 37–60, 1979.
- [19] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer-Verlag, Berlin Heidelberg., 1998.
- [20] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE Journal of Selected Area of Comm.*, vol. 52, pp. 1813–1827, April 2007.
- [21] S. N. Diggavi and T. Cover, "The worst additive noise under a covariance constraint." *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.
- [22] S. Ihara, "On the capacity of channels with additive non-Gaussian noise." *Info. Ctrl.*, vol. 37, no. 1, pp. 34–39, Apr. 1978.
- [23] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 374–377, May 1978.
- [24] A. S. Motahari and A. K. Khandani, "Capacity bounds for the gaussian interference channel," *Library Archives Canada Technical Report UW-ECE 2007-26* (Available at [http://www.cst.uwaterloo.ca/pub\\_tech\\_rep.html](http://www.cst.uwaterloo.ca/pub_tech_rep.html)), Aug. 2007.
- [25] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for gaussian interference channels (available at: <http://arxiv.org/abs/0712.1987>)," *Submitted to IEEE Trans. Inform. Theory*, Dec. 2007.

- [26] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of the gaussian interference channel in the low interference regime (available at: <http://arxiv.org/abs/0801.0452>)," *Proceedings of ITA Workshop, San Diego, CA*, Jan. 2008.