

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)
{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 ICs. For the weak Gaussian IC, a new outer bound on the capacity region is obtained that outperforms previously known outer bounds. The channel sum capacity for some certain range of the channel parameters is derived. 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, a new proof for Sato's outer bound is presented. We derive the full Han-Kobayashi achievable rate region when Gaussian code books are utilized. For the mixed Gaussian IC, a new outer bound is obtained that again outperforms previously known outer bounds. For this case, the channel sum capacity for all ranges of 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 the channel gains.

Index Terms

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

I. INTRODUCTION

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

A limiting expression for the capacity region is obtained in [4], c.f. [5]. Unfortunately, due to computational complexity, this kind of expressions does not give any tractable approach to fully characterize the capacity region of the Gaussian IC. 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 only relying on Gaussian distributions [6]. However, it is worth noting that there is a point on the boundary of the capacity region of the MAC that can be obtained directly from the limiting expression. This point indeed is achievable by using simple scheme of FD/TD.

One reason is that, in the limiting expression, the encoding and decoding strategies are the simplest one possible. 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 from the other user as noise. In contrast, using the more sophisticated encoders and decoders may result in collapsing the limiting expression into a single letter formula for the capacity region of the IC. As an evidence, it is known that the joint typical decoder for the MAC indeed achieves the capacity region [7]. Moreover, there are some special cases, such as strong ICs, where the exact characterization of the capacity region has been derived, c.f. [2] and [3], and decoding the interference is the main part of the proof.

In their pioneering work [8], Han and Kobayashi proposed a new encoding and decoding strategy in which the receivers are allowed to decode some part of the interfering user's data as well as its own data. Their achievable rate 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 of data is encoded in such a way that both users can successfully decode it. The private part, on the other hand, can be decoded only by the corresponding receiver and the other user treats it as noise. Briefly, the resulting region of this scheme is the intersection of the capacity region of two three-user MACs, projected to a two-dimensional space.

The Han-Kobayashi scheme can be directly applied to the Gaussian IC. Nonetheless, there are two sources of difficulties in characterizing the full Han-Kobayashi achievable rate region. First, the optimal distributions are unknown. Second, even if we confine the distributions to be Gaussian, computation of the full Han-Kobayashi region under Gaussian distributions is still difficult due to numerous degrees of freedom involved in the problem. The parameter which is the main cause of the difficulty for characterizing the Han-Kobayashi region with Gaussian distributions is the time-sharing parameter.

Recently in [9], Chong et.al have obtained a simpler expression with less inequalities for the Han-Kobayashi achievable rate region. In this case, the cardinality of the time-sharing parameter is decreased, since the cardinality of the time-sharing

parameter is directly related to the number of inequalities appearing in the achievable rate region. However, finding the full Han-Kobayashi achievable rate region is still prohibitively difficult.

Regarding outer bounds on the capacity region, there are three results that outperform other outer bounds. The first one obtained by Sato in [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 broadcast channel is equivalent to that of the one-sided weak Gaussian IC. Hence, Sato's 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's outer bound is based on the fact that removing one of the interfering links in the channel increases the capacity region. Therefore, the capacity region of the two-user Gaussian IC is inside the intersection of the capacity regions of two underlying one-sided Gaussian ICs. For the case of weak Gaussian IC, the underlying one-sided Gaussian IC is weak and the capacity region is unknown. However, Kramer has used the outer bound obtained by Sato to obtain an outer bound for the Gaussian IC.

The third outer bound due to Etkin, Tse, and Wang (ETW) is based on the Genie aided technique [22]. 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 so that the receiver can decode its own signal more easily by removing the interference. But, ETW's genie provides information about the intended signal to the receiver. Remarkably, they have shown that the new outer bound outperforms Kramer's bound for some ranges of parameters. Moreover, using a similar method, they have obtained an outer bound for the mixed Gaussian IC.

In this paper, by introducing the notion of admissible ICs, we propose a new outer bounding scheme for the two-user Gaussian IC. This scheme relies on an extremal inequality recently proved by Liu and Viswanath [13]. We show that by using this method, one can obtain tighter outer bounds for both weak and mixed Gaussian ICs. More importantly, the sum capacity of the Gaussian weak IC for some certain range of the channel parameters is derived by using this scheme.

The rest of this paper is organized as follows. In Section II, we rewrite some basic definitions and review Han-Kobayashi achievable rate region when Gaussian codebooks are used. We study the time-sharing and the convexification methods, that both enlarge the basic Han-Kobayashi achievable rate region. We investigate conditions for which the two regions obtained from time-sharing and convexification coincide. Finally, we consider an optimization problem (extremal inequality) and derive optimum solutions of the problem. In fact, the extremal inequality is used throughout the paper.

In Section III, admissible channels are introduced. Some classes of admissible channels for the two-user Gaussian IC is considered. Moreover, outer bounds on the capacity regions of these classes are obtained.

In Section IV, we study the capacity region of the weak Gaussian IC. We first derive the sum capacity of this channel for some ranges of parameters. It is shown that for this range of parameters, it suffices that users treat the interference as Gaussian noise and transmit at their highest rate. We then derive an outer bound on the capacity region which is the best known upper bound to date. We finally prove that the basic Han-Kobayashi achievable rate region possesses the desired property of having the same enlarged region by using time-sharing or convexification. This reduces the complexity of characterization of the full Han-Kobayashi achievable rate region when Gaussian codebooks are used.

In Section V, we study capacity region of the one-sided Gaussian IC. We present a new proof on Sato's outer bound using the extremal inequality. Then, we simplify the Han-Kobayashi achievable rate region so that the full region can be characterized.

In Section VI, we study capacity region of the mixed Gaussian IC. We first obtain the sum capacity of this channel. Then, we derive an outer bound which outperforms other existing outer bounds. Finally, by investigating the Han-Kobayashi achievable rate region for different cases, we prove that for some range of channel parameters the full Han-Kobayashi achievable rate region using Gaussian codebooks is equivalent to that of the one-sided case. 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. $|A|$, $\text{tr}\{A\}$, 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 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 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 $\mathcal{M}_i \in [1, 2, \dots, 2^{nR_i}]$.

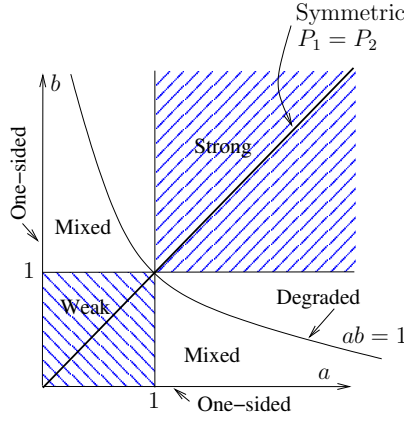


Fig. 1. Classes of the two-user ICs.

- 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 $G_i : \mathbf{y}_i \rightarrow [1, 2, \dots, 2^{nR_i}]$.
- 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 said to be achievable if there is a sequence of codes $(2^{nR_1}, 2^{nR_2}, n, \lambda_1^n, \lambda_2^n)$ with vanishing average probability of errors. 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 of the capacity region 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 [14]

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

where x_i and y_i denote respectively the input and output alphabets of user $i \in \{1, 2\}$. $z_1 \sim \mathcal{N}(0, 1)$ and $z_2 \sim \mathcal{N}(0, 1)$ are standard Gaussian random variables. Constants $a \geq 0$ and $b \geq 0$ represent the interference link gains. Furthermore, transmitter i is subject to the average power constraint P_i for $i \in \{1, 2\}$. Achievable rates and the capacity region of the Gaussian IC can be defined in a similar fashion as that of the general IC except the codewords must satisfy the following power constraints

$$\|\mathbf{x}_i(m)\|^2 \leq nP_i; \quad \forall m \in [1, 2, \dots, 2^{nR_i}] \text{ and } i \in \{1, 2\} \quad (3)$$

where $\|\cdot\|$ denotes the Euclidean norm. The capacity region of the two-user Gaussian IC is denoted by \mathcal{C} . Clearly, \mathcal{C} is a function of the channel's parameters P_1 , P_2 , a , and b . To emphasize this relationship, we may write \mathcal{C} as $\mathcal{C}(P_1, P_2, a, b)$.

Remark 1: Since the capacity region of the general IC only depends on the marginal distributions [14], 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 channel's 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, for the purpose of exposition, the symmetric Gaussian IC when $a = b$ and $P_1 = P_2$ is used.

Among all, the capacity region of the strong 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

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

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 [15]. The support function of any set $D \in \mathbb{R}^m$ is a function $\sigma_D : \mathbb{R}^m \rightarrow \mathbb{R}$ defined as

$$\sigma_D(\mathbf{c}) = \sup\{\mathbf{c}^t \mathbf{R} | \mathbf{R} \in D\}. \quad (4)$$

Clearly, if the set D is compact then the sup is attained and can be replaced by max. In this case, the solutions of (4) correspond to the boundary points of D [15]. The following relation is the dual of (4) and holds when D is closed and convex

$$D = \{\mathbf{R} | \mathbf{c}^t \mathbf{R} \leq \sigma_D(\mathbf{c}), \forall \mathbf{c}\}. \quad (5)$$

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 Han-Kobayashi achievable rate 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

$$R_1 \leq \psi_1 = \gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \quad (6)$$

$$R_2 \leq \psi_2 = \gamma \left(\frac{P_2}{1 + b\alpha P_1} \right), \quad (7)$$

$$R_1 + R_2 \leq \psi_3 = \min\{\psi_{31}, \psi_{32}, \psi_{33}\}, \quad (8)$$

$$2R_1 + R_2 \leq \psi_4 = \gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left(\frac{\alpha P_1}{1 + a\beta P_2} \right) + \gamma \left(\frac{\beta P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right), \quad (9)$$

$$R_1 + 2R_2 \leq \psi_5 = \gamma \left(\frac{\beta P_2}{1 + b\alpha P_1} \right) + \gamma \left(\frac{P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right) + \gamma \left(\frac{\alpha P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right), \quad (10)$$

for fixed $\alpha \in [0, 1]$ and $\beta \in [0, 1]$.¹ ψ_3 is the minimum of ψ_{31} , ψ_{32} , and ψ_{33} defined as

$$\psi_{31} = \gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left(\frac{\beta P_2}{1 + b\alpha P_1} \right), \quad (11)$$

$$\psi_{32} = \gamma \left(\frac{\alpha P_1}{1 + a\beta P_2} \right) + \gamma \left(\frac{P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right), \quad (12)$$

$$\psi_{33} = \gamma \left(\frac{\alpha P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma \left(\frac{\beta P_2 + b(1 - \alpha)P_1}{1 + b\alpha P_1} \right). \quad (13)$$

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

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

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 vectors $\mathbf{R} = (R_1, R_2)^t$ satisfying

$$A\mathbf{R} \leq \sum_{i=1}^q \lambda_i \Psi(P_{1i}, P_{2i}, \alpha_i, \beta_i), \quad (14)$$

¹In the Han-Kobayashi scheme two independent messages are encoded at each transmitter. One message is the common message and the other one is the private message. α (β) is the parameter that determines the amount of power allocated to the common and private messages, i.e., αP_1 and $(1 - \alpha)P_1$ (βP_2 and $(1 - \beta)P_2$) of the total power is used for transmission of the private and common messages, respectively

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

$$\sum_{i=1}^q \lambda_i P_{1i} \leq P_1, \quad (15)$$

$$\sum_{i=1}^q \lambda_i P_{2i} \leq P_2, \quad (16)$$

$$\sum_{i=1}^q \lambda_i = 1, \quad (17)$$

$$\lambda_i \geq 0, (\alpha_i, \beta_i) \in [0, 1]^2; \forall i \in \{1, \dots, q\}. \quad (18)$$

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 points of both \mathcal{C} and \mathcal{G} . Hence, in order 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 is valid 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 two regions \mathcal{G}_1 and \mathcal{G}_2 that enlarge \mathcal{G}_0 in different ways. \mathcal{G}_1 is defined as

$$\mathcal{G}_1(P_1, P_2) = \bigcup_{(\alpha, \beta) \in [0, 1]^2} \mathcal{G}_0(P_1, P_2, \alpha, \beta). \quad (19)$$

\mathcal{G}_1 is not necessarily a convex region. Hence, it can be further enlarged by the convex hull operation, i.e., $\text{conv } \mathcal{G}_1$.

\mathcal{G}_2 is defined as the collection of all rate vectors $\mathbf{R} = (R_1, R_2)^t$ satisfying

$$\mathbf{R} = \sum_{i=1}^{q'} \lambda_i \mathbf{R}_i \quad (20)$$

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

$$A\mathbf{R}_i \leq \Psi(P_{1i}, P_{2i}, \alpha_i, \beta_i), \quad (21)$$

$$\sum_{i=1}^{q'} \lambda_i P_{1i} \leq P_1, \quad (22)$$

$$\sum_{i=1}^{q'} \lambda_i P_{2i} \leq P_2, \quad (23)$$

$$\sum_{i=1}^{q'} \lambda_i = 1, \quad (24)$$

$$\lambda_i \geq 0, (\alpha_i, \beta_i) \in [0, 1]^2; \forall i \in \{1, \dots, q'\}. \quad (25)$$

It is easy to show that \mathcal{G}_2 is a closed, bounded and convex region. In fact, \mathcal{G}_2 is enlarged by using the simple method of TD/FD. To see this, let us divide the available frequency band into q' sub-bands where λ_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 band, respectively. Therefore, all rate pairs in $\mathcal{G}_0(P_{1i}, P_{2i}, \alpha_i, \beta_i)$ are achievable in the i th 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 $\sum_{i=1}^{q'} \lambda_i P_{1i} \leq P_1$ 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. Convexification Versus Time-Sharing

In this subsection, we have two objectives. Firstly, we aim at providing some necessary conditions such that $\mathcal{G}_2 = \mathcal{G}$. Secondly, 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, consider an achievable scheme for an M -user channel with the given power constraint $\mathbf{P} = [P_1, P_2, \dots, P_M]$ is given. We assume that the achievable region can be represented as

$$D_0(\mathbf{P}, \Theta) = \{\mathbf{R} | A\mathbf{R} \leq \Psi(\mathbf{P}, \Theta)\}, \quad (26)$$

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 as follows. The first method is to make use of the time sharing parameter. Let us denote this new region as D which can be written as

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

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

In the second method, we make use of TD/FD to enlarge the achievable rate region. This results in a new achievable region D_2 represented as

$$D_2 = \left\{ \mathbf{R} = \sum_{i=1}^{q'} \lambda_i \mathbf{R}_i | A\mathbf{R}_i \leq \Psi(\mathbf{P}_i, \Theta_i), \sum_{i=1}^{q'} \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^{q'} \lambda_i = 1, \lambda_i \geq 0, \Theta_i \in [0, 1]^M \forall i \right\}, \quad (28)$$

where $q' \in \mathbb{N}$. We call this new method convexification.

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} , Θ , and \mathbf{c} . Hence, we have

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \max\{\mathbf{c}^t \mathbf{R} | A\mathbf{R} \leq \Psi(\mathbf{P}, \Theta)\}. \quad (29)$$

For fixed \mathbf{P} and Θ , (29) is a linear program. Using the strong duality of the linear programming, we obtain

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \min\{\mathbf{y}^t \Psi(\mathbf{P}, \Theta) | A^t \mathbf{y} = \mathbf{c}, \mathbf{y} \geq 0\}. \quad (30)$$

In general, $\hat{\mathbf{y}}$, the minimizer of (30), is a function of \mathbf{P} , Θ , 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

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}, \Theta) = \hat{\mathbf{y}}^t(\mathbf{c}) \Psi(\mathbf{P}, \Theta), \quad (31)$$

where $A^t \hat{\mathbf{y}} = \mathbf{c}$. This condition essentially means that for any \mathbf{c} the extreme point of D_0 maximizing the objective $\mathbf{c}^t \mathbf{R}$ is a certain 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 Θ .

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

Proof: Since $D_2 \subseteq D$ always holds, we need only 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

$$\sigma_D(\mathbf{c}, \mathbf{P}) = \max_{\mathbf{R} \in D} \mathbf{c}^t \mathbf{R} \quad (32)$$

By fixing \mathbf{P} , \mathbf{P}_i 's, Θ_i 's, and λ_i 's, the above maximization becomes a linear program. Hence, by making use of the weak duality of the linear programming, we obtain

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \min_{A^t \mathbf{y} = \mathbf{c}, \mathbf{y} \geq 0} \mathbf{y}^t \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i). \quad (33)$$

Clearly, $\hat{\mathbf{y}}(\mathbf{c})$, the solution of (30), is a feasible point for (33) and we have

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \hat{\mathbf{y}}^t(\mathbf{c}) \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i) \quad (34)$$

Using (31), we obtain

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^q \lambda_i \sigma_{D_0}(\mathbf{c}, \mathbf{P}_i, \Theta_i) \quad (35)$$

Let us assume $\hat{\mathbf{R}}_i$ is the maximizer of (29). In this case, we have

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^q \lambda_i \mathbf{c}^t \hat{\mathbf{R}}_i. \quad (36)$$

Hence, we have

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \mathbf{c}^t \sum_{i=1}^q \lambda_i \hat{\mathbf{R}}_i. \quad (37)$$

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

$$\sigma_D(\mathbf{c}, \mathbf{P}) \leq \sigma_{D_2}(\mathbf{c}, \mathbf{P}). \quad (38)$$

This completes the proof. \square

Corollary 1 (Han [16]): 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 Θ . \square

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

Theorem 2: The cardinality of the time sharing parameter q in (27) 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

$$E = \left\{ \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i) \mid \sum_{i=1}^q \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^q \lambda_i = 1, \lambda_i \geq 0, \Theta_i \in [0, 1]^M \forall i \right\}. \quad (39)$$

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

$$E_1 = \{(\mathbf{P}', \mathbf{S}') \mid 0 \leq \mathbf{P}', \mathbf{S}' = \Psi(\mathbf{P}', \Theta'), \Theta' \in [0, 1]^M\}. \quad (40)$$

From the direct consequence of the Caratheodory's theorem [18], 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 extension of the Caratheodory's theorem [18]. Now, we define the region \hat{E} as

$$\hat{E} = \{\mathbf{S}' \mid (\mathbf{P}', \mathbf{S}') \in \text{conv } E_1, \mathbf{P}' \leq \mathbf{P}\}. \quad (41)$$

Clearly, $\hat{E} \subseteq E$. To show the other inclusion, we take a point in E , say $\mathbf{S} = \sum_{i=1}^q \lambda_i \Psi(\mathbf{P}_i, \Theta_i)$. Since $(\mathbf{P}_i, \Psi(\mathbf{P}_i, \Theta_i))$ is 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 [17]): For the M -user Gaussian IC where users use Gaussian codebooks for data transmission and treat the interference as Gaussian 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

$$\sigma_{D_2}(\mathbf{c}, \mathbf{P}) = \max_{\mathbf{R} \in D_2} \mathbf{c}^t \mathbf{R} = \mathbf{c}^t \hat{\mathbf{R}}, \quad (42)$$

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 (42) can be written as

$$\begin{aligned} \sigma_{D_2}(\mathbf{c}, \mathbf{P}) = & \max \sum_{i=1}^{q'} \lambda_i g(\mathbf{c}, \mathbf{P}_i) \\ & \text{subject to:} \\ & \sum_{i=1}^{q'} \lambda_i = 1 \\ & \sum_{i=1}^{q'} \lambda_i \mathbf{P}_i \leq \mathbf{P} \\ & 0 \leq \lambda_i, 0 \leq \mathbf{P}_i \forall i \in \{1, 2, \dots, q'\} \end{aligned} \quad (43)$$

$$\quad (44)$$

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

$$\begin{aligned} g(\mathbf{c}, \mathbf{P}) = & \max \mathbf{c}^t \mathbf{R} \\ & \text{subject to:} \\ & A\mathbf{R} \leq \Psi(\mathbf{P}, \Theta) \\ & 0 \leq \Theta \leq 1 \end{aligned} \quad (45)$$

In fact, $\sigma_{D_2}(\mathbf{c}, \mathbf{P})$ in (43) can be viewed as the convexification of the function $g(\mathbf{c}, \mathbf{P})$ [18]. Hence, by Theorem 2.16 in [18], we conclude that $q' \leq M + 1$. \square

Surprising fact about Theorem 3 is that the upper bound for q' is independent of the number of inequalities in the description of the achievable rate region.

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

E. Extremal Inequality

In [13], the following optimization problem is studied:

$$W = \max_{Q_{\mathbf{X}} \leq S} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\mathbf{X} + \mathbf{Z}_2), \quad (46)$$

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 [13], 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 [19][20].

In this paper, we consider a special case of (46) where \mathbf{Z}_1 and \mathbf{Z}_2 have the covariance matrices $N_1 I$ and $N_2 I$, respectively, and the constraint is the trace constraint, i.e.,

$$W = \max_{tr\{Q_{\mathbf{X}}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\mathbf{X} + \mathbf{Z}_2). \quad (47)$$

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 optimization problem (47) has a Gaussian optimal solution for all $0 \leq \mu$ with iid components. More precisely, 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

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

- 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

$$W = \frac{n}{2} \log \left[(2\pi e) \frac{N_2 - N_1}{\mu - 1} \right] - \frac{\mu n}{2} \log \left[\frac{\mu(2\pi e)(N_2 - N_1)}{\mu - 1} \right] \quad (49)$$

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

$$W = \frac{n}{2} \log(2\pi e N_1) - \frac{\mu n}{2} \log(2\pi e N_2). \quad (50)$$

Proof: From the general result for (46), we know that the optimum input distribution is a Gaussian vector. Hence, we need to solve the following maximization problem:

$$W = \max \frac{1}{2} \log ((2\pi e)^n |Q_{\mathbf{X}} + N_1 I|) - \frac{\mu}{2} \log ((2\pi e)^n |Q_{\mathbf{X}} + N_2 I|) \quad (51)$$

subject to:

$$0 \leq Q_{\mathbf{X}}$$

$$tr\{Q_{\mathbf{X}}\} \leq nP$$

Since $Q_{\mathbf{X}}$ is a positive semi-definite matrix, it can be decomposed as $Q_{\mathbf{X}} = U\Lambda U^t$, where Λ 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 (51) and using the identities $tr\{AB\} = tr\{BA\}$ and $|AB + I| = |BA + I|$, we obtain

$$W = \max \frac{1}{2} \log ((2\pi e)^n |\Lambda + N_1 I|) - \frac{\mu}{2} \log ((2\pi e)^n |\Lambda + N_2 I|) \quad (52)$$

subject to:

$$0 \leq \Lambda$$

$$tr\{\Lambda\} \leq nP$$

This optimization problem can be simplified as

$$W = \max \frac{n}{2} \sum_{i=1}^n [\log(2\pi e)(\lambda_i + N_1) - \mu \log(2\pi e)(\lambda_i + N_2)] \quad (53)$$

subject to:

$$0 \leq \lambda_i \quad \forall i$$

$$\sum_{i=1}^n \lambda_i \leq nP$$

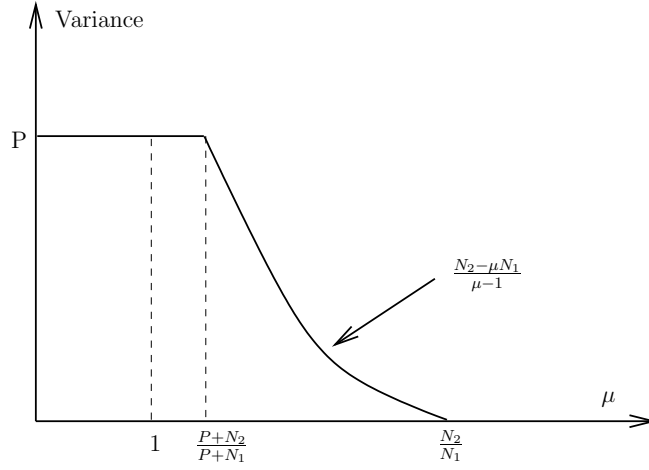


Fig. 2. Optimum variance versus μ .

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

$$L(\Lambda, \psi, \Phi) = \max \frac{n}{2} \sum_{i=1}^n [\log(2\pi e)(\lambda_i + N_1) - \mu \log(2\pi e)(\lambda_i + N_2)] + \psi \left(nP - \sum_{i=1}^n \lambda_i \right) + \sum_{i=1}^n \phi_i \lambda_i. \quad (54)$$

The first order KKT necessary conditions for the optimum solutions of (54) can be written as

$$\frac{1}{\lambda_i + N_1} - \frac{\mu}{\lambda_i + N_2} - \psi + \phi_i = 0, \quad \forall i \in \{1, 2, \dots, n\} \quad (55)$$

$$\psi \left(nP - \sum_{i=1}^n \lambda_i \right) = 0, \quad (56)$$

$$\phi_i \lambda_i = 0, \quad \forall i \in \{1, 2, \dots, n\} \quad (57)$$

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

$$\lambda = \begin{cases} P, & \text{if } 0 \leq \mu \leq \frac{N_2 + P}{N_1 + P} \\ \frac{N_2 - \mu N_1}{\mu - 1}, & \text{if } \frac{N_2 + P}{N_1 + P} < \mu \leq \frac{N_2}{N_1} \\ 0, & \text{if } \frac{N_2}{N_1} < \mu \end{cases} \quad (58)$$

Now, substituting λ into the objective function yields the desired result. \square

In Figure 2, the optimum variance as a function of μ is sketched. 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 obtain the maximum of the objective, whereas for $\mu > \frac{P+N_2}{P+N_1}$ we use less power than the given power constraint.

Lemma 2: If $N_1 > N_2$, the optimization problem (47) has a Gaussian optimal solution for all $1 \leq \mu$ with iid components. In this case, the optimum variance is 0 and the optimum solution is

$$W = \frac{n}{2} \log(2\pi e N_1) - \frac{\mu n}{2} \log(2\pi e N_2). \quad (59)$$

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

Corollary 4: For $\mu = 1$, the optimization problem (47) has a Gaussian optimal solution with iid components. The optimum solution in this case is

$$W = \begin{cases} \frac{n}{2} \log \left(\frac{P+N_1}{P+N_2} \right), & \text{if } N_1 \leq N_2 \\ \frac{n}{2} \log \left(\frac{N_1}{N_2} \right), & \text{if } N_1 > N_2 \end{cases} \quad (60)$$

We repeatedly use the following optimization problem throughout the paper:

$$f_h = \max_{\text{tr}\{\mathbf{Q}_\mathbf{X}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\sqrt{a}\mathbf{X} + \mathbf{Z}_2), \quad (61)$$

where $N_1 \leq N_2/a$. Using the identity $h(A\mathbf{X}) = \log(|A|) + h(\mathbf{X})$, (61) can be written as

$$f_h = \frac{n}{2} \log a + \max_{\text{tr}\{\mathbf{Q}_\mathbf{X}\} \leq nP} h(\mathbf{X} + \mathbf{Z}_1) - \mu h(\mathbf{X} + \frac{\mathbf{Z}_2}{\sqrt{a}}). \quad (62)$$

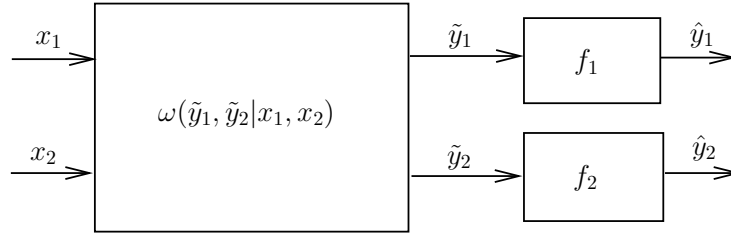


Fig. 3. An admissible channel. f_1 and f_2 are two deterministic functions of their inputs.

Now, Lemma 1 can be applied to obtain

$$f_h(P, N_1, N_2, a, \mu) = \begin{cases} \frac{1}{2} \log [(2\pi e)(P + N_1)] - \frac{\mu}{2} \log [(2\pi e)(aP + N_2)] & \text{if } 0 \leq \mu \leq \frac{P+N_2/a}{P+N_1} \\ \frac{1}{2} \log \left[(2\pi e) \frac{N_2/a - N_1}{\mu - 1} \right] - \frac{\mu}{2} \log \left[\frac{a\mu(2\pi e)(N_2/a - N_1)}{\mu - 1} \right] & \text{if } \frac{P+N_2/a}{P+N_1} < \mu \leq \frac{N_2}{aN_1} \\ \frac{1}{2} \log(2\pi e N_1) - \frac{\mu}{2} \log(2\pi e N_2) & \text{if } \frac{N_2}{aN_1} < \mu \end{cases} \quad (63)$$

III. ADMISSIBLE CHANNELS

In this section, we aim at building IC's whose capacity regions contain the capacity region of the two-user Gaussian IC, i.e., \mathcal{C} . Since we ultimately use them to outer bound \mathcal{C} , these IC's need to possess some properties regarding the capacity region characterization. In other words, if characterizing the capacity regions or obtaining tight upper bounds of these channels are as hard as the original one, then the new channels are useless.

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

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

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

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

One way to satisfy (64) and (65) is to provide some extra information to either one or both receivers. This technique is known as Genie aided outer bounding. In [12], Kramer used a genie to provide some extra information to both receivers so that they can decode both users' messages. Since the capacity region of this new interference channel is equivalent to the capacity of the Compound Multiple Access Channel whose capacity region is known, he managed to obtain an outer bound on the capacity region. In order to obtain a tighter outer bound, he also used 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 due to Sato and Costa, see [21] and [11], that can be used to outer bound the original channel. In [22], Etkin et al. used a different genie that provides some extra information about the intended signal. Even though it seems that their channel is far from having a tight capacity region with respect to that of the original channel, they showed that their channel is tighter than Kramer's outer bound for some ranges of parameters.

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 for the two-user Gaussian IC 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

$$I(x_1^n; y_1^n) \leq I(x_1^n; \hat{y}_1^n), \quad (66)$$

$$I(x_2^n; y_2^n) \leq I(x_2^n; \hat{y}_2^n). \quad (67)$$

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

Clearly, an admissible channel also satisfies (64) and (65).

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. Hence, $\tilde{y}_i = (y_i, s_i)$ 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 (66) and (67) hold with equality sign.

To obtain the tightest outer bound, we need to take 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 the admissible channels. Hence, we need to find classes of admissible channels, say \mathcal{F} , that 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 on their capacity regions. Since $\mathcal{F} \subseteq \mathcal{E}$, we have

$$\mathcal{C} \subseteq \bigcap_{\mathcal{F}} \mathcal{C}'. \quad (68)$$

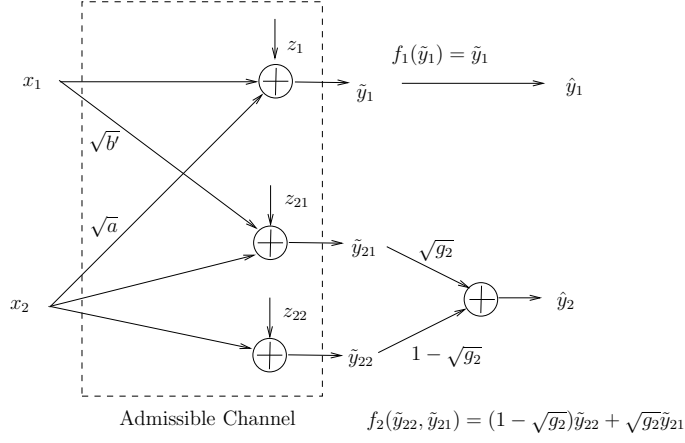


Fig. 4. Class A1 admissible channels.

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

$$\sigma_{\mathcal{C}}(c_1, c_2) = \max_{(R_1, R_2) \in \mathcal{C}} c_1 R_1 + c_2 R_2 \quad (69)$$

Since we are interested in boundary points not including 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

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \sigma_{\mathcal{C}'}(c_1, c_2). \quad (70)$$

Hence, taking the minimum of the right hand side, we obtain

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \min_{\mathcal{C}' \in \mathcal{F}} \sigma_{\mathcal{C}'}(c_1, c_2), \quad (71)$$

which can be written as

$$\sigma_{\mathcal{C}}(c_1, c_2) \leq \min_{\mathcal{C}' \in \mathcal{F}} \max_{(R_1, R_2) \in \mathcal{C}'} c_1 R_1 + c_2 R_2. \quad (72)$$

For the sake of convenience, we make use of the following two optimization problems

$$\sigma_{\mathcal{C}}(\mu, 1) = \max_{(R_1, R_2) \in \mathcal{C}} \mu R_1 + R_2, \quad (73)$$

$$\sigma_{\mathcal{C}}(1, \mu) = \max_{(R_1, R_2) \in \mathcal{C}} R_1 + \mu R_2, \quad (74)$$

where $1 \leq \mu$. It is easy to show that solutions of (73) and (74) correspond to the boundary points of the capacity region that we are interested in.

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 upper bound $\sigma_{\mathcal{C}}(\mu, 1)$. Therefore, we need to find a tight upper bound for $\sigma_{\mathcal{C}'}(\mu, 1)$. 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

$$\begin{aligned} \tilde{y}_1 &= x_1 + \sqrt{a}x_2 + z_1, \\ \tilde{y}_{21} &= x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} &= x_2 + z_{22}, \end{aligned} \quad (75)$$

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

To investigate admissibility conditions (66) and (67), we need to introduce two deterministic functions. Let us consider two linear functions f_1 and f_2 as follows (see Figure 4)

$$f_1(\tilde{y}_1^n) = \tilde{y}_1^n, \quad (76)$$

$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1 - \sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n, \quad (77)$$

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 channel obtained by removing the link between Transmitter 1 and Receiver 2. We have

$$\hat{y}_1^n = x_1^n + \sqrt{a}x_2^n + z_1^n, \quad (78)$$

$$\hat{y}_2^n = \sqrt{b'g_2}x_1^n + x_2^n + (1 - \sqrt{g_2})z_{22}^n + \sqrt{g_2}z_{21}^n. \quad (79)$$

Hence, this channel is admissible if the corresponding parameters satisfy

$$\begin{aligned} b'g_2 &= b, \\ (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \quad (80)$$

We further add the following constraints to the required conditions of the class A1 channels:

$$\begin{aligned} b' &\leq N_{21}, \\ aN_{22} &\leq 1. \end{aligned} \quad (81)$$

Although they reduce the number of admissible channels within the class, these latter constraints help us to provide 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 (75) and satisfying (81), we have

$$\begin{aligned} \sigma_{\mathcal{C}'}(\mu, 1) \leq & \min \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\ & + \mu_2 f_h \left(P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right) + f_h(P_2, N_{22}, 1, a, \mu_1) \end{aligned} \quad (82)$$

subject to:

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

$$\mu_1, \mu_2 \geq 0$$

Proof: Let us assume R_1 and R_2 are two rates achievable for User 1 and 2, respectively. Furthermore, we split μ into $\mu_1 \geq 0$ and $\mu_2 \geq 0$ such that $\mu = \mu_1 + \mu_2$. Using Fano's inequalities, we obtain

$$\begin{aligned} n(\mu R_1 + R_2) &\leq \mu I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ &= \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ &\stackrel{(a)}{\leq} \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ &= \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\ &= \mu_1 h(\tilde{y}_1^n) - \mu_1 h(\tilde{y}_1^n | x_1^n) + \mu_2 h(\tilde{y}_1^n | x_2^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) \\ &\quad + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \\ &= [\mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n)] + [\mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)] \\ &\quad + [h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n)] + [h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_1^n | x_1^n)] + n\epsilon_n, \end{aligned} \quad (83)$$

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 (83).

To maximize the terms within the first bracket, we use the fact that Gaussian distribution maximizes the differential entropy for given covariance matrix constraint. Hence, we have

$$\begin{aligned} \mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) &= \mu_1 h(x_1^n + \sqrt{a}x_2^n + z_1^n) - \mu_2 h(z_1^n) \\ &\leq \frac{\mu_1 n}{2} \log [2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2 n}{2} \log(2\pi e). \end{aligned} \quad (84)$$

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

$$\begin{aligned} \mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) &= \mu_2 \left(h(x_1^n + z_1^n) - \frac{1}{\mu_2} h(\sqrt{b'}x_1^n + z_{21}^n) \right) \\ &\leq \mu_2 n f_h \left(P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right), \end{aligned} \quad (85)$$

where f_h is defined in (63).

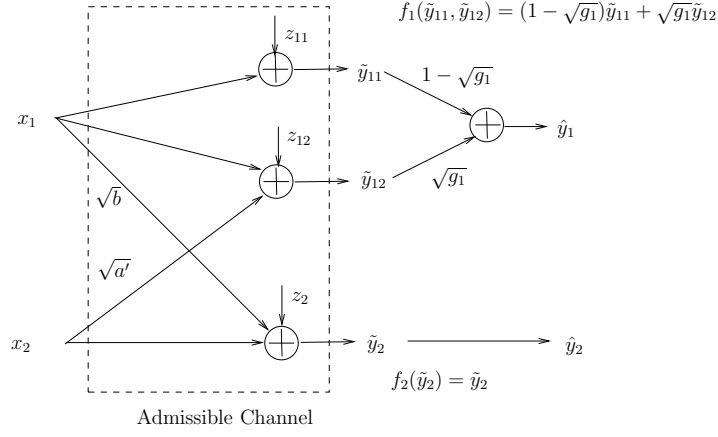


Fig. 5. Class A2 admissible channels.

We upper bound the terms within the third bracket as follows, c.f. [22]:

$$\begin{aligned}
 h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) &\stackrel{(a)}{\leq} \sum_{i=1}^n h(\tilde{y}_{21}[i] | \tilde{y}_{22}[i]) - h(z_{22}^n) \\
 &\stackrel{(b)}{\leq} \sum_{i=1}^n \frac{1}{2} \log \left[2\pi e \left(N_{21} + b' P_1[i] + \frac{P_2[i] N_{22}}{P_2[i] + N_{22}} \right) \right] - \frac{n}{2} \log(2\pi e N_{22}) \\
 &\stackrel{(c)}{\leq} \frac{n}{2} \log \left[2\pi e \left(N_{21} + \frac{1}{n} \sum_{i=1}^n b' P_1[i] + \frac{\frac{1}{n} \sum_{i=1}^n P_2[i] N_{22}}{\frac{1}{n} \sum_{i=1}^n P_2[i] + N_{22}} \right) \right] - \frac{n}{2} \log(2\pi e N_{22}) \\
 &\leq \frac{n}{2} \log \left[2\pi e \left(N_{21} + b' P_1 + \frac{P_2 N_{22}}{P_2 + N_{22}} \right) \right] - \frac{n}{2} \log(2\pi e N_{22}) \\
 &\leq \frac{n}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right), \tag{86}
 \end{aligned}$$

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 optimizes conditional entropy for given covariance matrix, and (c) follows from Jensen's inequality.

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

$$\begin{aligned}
 h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_{11}^n | x_1^n) &= h(x_2^n + z_{22}^n) - \mu_1 h(\sqrt{a} x_2^n + z_{11}^n) \\
 &\leq n f_h(P_2, N_{22}, 1, a, \mu_1). \tag{87}
 \end{aligned}$$

Adding all inequalities, we obtain

$$\begin{aligned}
 \mu R_1 + R_2 &\leq \frac{\mu_1}{2} \log[2\pi e(P_1 + a P_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\
 &\quad + \mu_2 f_h \left(P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right) + f_h(P_2, N_{22}, 1, a, \mu_1), \tag{88}
 \end{aligned}$$

where the fact that $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$ is used to eliminate ϵ_n from the right hand side of the inequality. Now, by taking the minimum of the right hand side of (88) over all μ_1 and μ_2 , we obtain the desired result. This completes the proof. \square

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

$$\begin{aligned}
 \tilde{y}_{11} &= x_1 + z_{11}, \\
 \tilde{y}_{12} &= x_1 + \sqrt{a'} x_2 + z_{12}, \\
 \tilde{y}_2 &= x_2 + \sqrt{b} x_1 + z_2, \tag{89}
 \end{aligned}$$

where \tilde{y}_{11} and \tilde{y}_{12} are received signals at the first user's receiver, \tilde{y}_2 is the received signal at the second user's receiver, z_2 is an additive Gaussian noise with unit variance, z_{11} and z_{12} are additive Gaussian noise with variances N_{11} and N_{12} , respectively, and transmitter 1 and 2 are subject to the average 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)

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1 - \sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n, \quad (90)$$

$$f_2(\tilde{y}_2^n) = \tilde{y}_2^n. \quad (91)$$

Similar to Class A1, when $g_1 = 0$ the admissible channel in Class A2 becomes the one-sided Gaussian IC by letting $N_{12} \rightarrow \infty$ and $N_{11} = 1$. Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a'g_1}x_2^n + (1 - \sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n, \quad (92)$$

$$\hat{y}_2^n = \sqrt{b}x_1^n + x_2^n + z_2^n. \quad (93)$$

We conclude that the channel modeled by (89) is admissible if the corresponding parameters satisfy

$$\begin{aligned} a'g_1 &= a, \\ (1 - \sqrt{g_1})^2N_{11} + g_1N_{12} &= 1. \end{aligned} \quad (94)$$

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

$$\begin{aligned} a' &\leq N_{12}, \\ bN_{11} &\leq 1. \end{aligned} \quad (95)$$

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

Lemma 4: For the channels modeled by (89) and satisfying (95), we have

$$\begin{aligned} \sigma_{\mathcal{C}'}(1, \mu) &\leq \min \frac{\mu_1}{2} \log [2\pi e(bP_1 + P_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left(\frac{N_{12}}{N_{11}} + \frac{a'P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right) \\ &\quad + \mu_2 f_h \left(P_2, 1, N_{12}, a', \frac{1}{\mu_2} \right) + f_h(P_1, N_{11}, 1, b, \mu_1) \end{aligned} \quad (96)$$

subject to:

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

$$\mu_1, \mu_2 \geq 0$$

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)

$$\begin{cases} \tilde{y}_{11} = x_1 + z_{11}, \\ \tilde{y}_{12} = x_1 + \sqrt{a'}x_2 + z_{12}, \\ \tilde{y}_{21} = x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} = x_2 + z_{22}, \end{cases} \quad (97)$$

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

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

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1 - \sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n \quad (98)$$

$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1 - \sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n. \quad (99)$$

Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a'g_1}x_2^n + (1 - \sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n \quad (100)$$

$$\hat{y}_2^n = \sqrt{b'g_2}x_1^n + x_2^n + (1 - \sqrt{g_2})z_{22}^n + \sqrt{g_2}z_{21}^n. \quad (101)$$

To satisfy (66) and (67), it suffices to have

$$\begin{aligned} a'g_1 &= a, \\ b'g_2 &= b, \\ (1 - \sqrt{g_1})^2N_{11} + g_1N_{12} &= 1, \\ (1 - \sqrt{g_2})^2N_{22} + g_2N_{21} &= 1. \end{aligned} \quad (102)$$

Hence, a channel modeled in (97) is admissible if there are two nonnegative numbers g_1 and g_2 such that the set of equalities in (102) holds. We further add the following two constraints to the equality conditions in (102):

$$\begin{aligned} b'N_{11} &\leq N_{21}, \\ a'N_{22} &\leq N_{12}. \end{aligned} \quad (103)$$

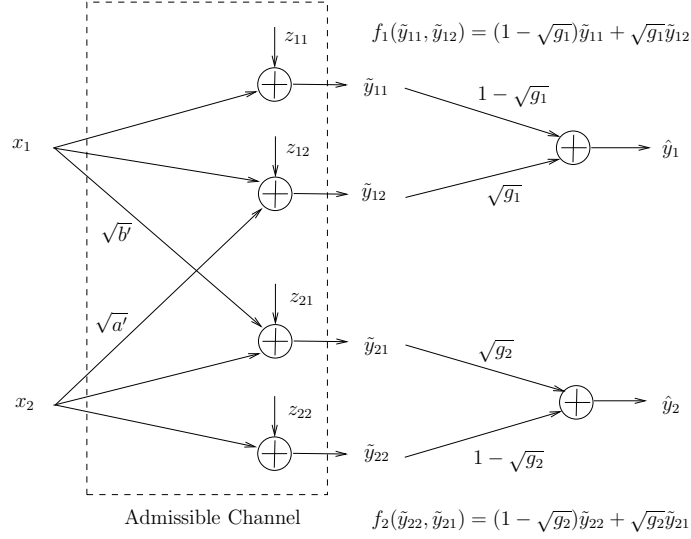


Fig. 6. Class B admissible channels.

Although adding more constraints reduces the number of the admissible channels, it helps us to provide an outer bound on $\sigma_{\mathcal{C}'}(\mu, 1)$ and $\sigma_{\mathcal{C}'}(1, \mu)$ with a closed form formula.

Lemma 5: For the channels modeled by (97) and satisfying (103), we have

$$\begin{aligned} \sigma_{\mathcal{C}'}(\mu, 1) \leq & \mu\gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) \\ & + f_h(P_2, N_{22}, N_{12}, a', \mu) + \frac{\mu}{2} \log((2\pi e)(a'P_2 + N_{12})) - \frac{1}{2} \log((2\pi e)(P_2 + N_{22})), \end{aligned} \quad (104)$$

$$\begin{aligned} \sigma_{\mathcal{C}'}(1, \mu) \leq & \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \mu\gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) \\ & + f_h(P_1, N_{11}, N_{21}, b', \mu) + \frac{\mu}{2} \log((2\pi e)(b'P_1 + N_{21})) - \frac{1}{2} \log((2\pi e)(P_1 + N_{11})). \end{aligned} \quad (105)$$

Proof: We only upper bound $\sigma_{\mathcal{C}'}(\mu, 1)$ and an upper bound for $\sigma_{\mathcal{C}'}(1, \mu)$ can similarly be obtained. Let us assume R_1 and R_2 are achievable rates for User 1 and 2, respectively. Using Fano's inequalities, we obtain

$$\begin{aligned} n(\mu R_1 + R_2) & \leq \mu I(x_1^n; \tilde{y}_{11}^n, \tilde{y}_{12}^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ & = \mu I(x_1^n; \tilde{y}_{12}^n | \tilde{y}_{11}^n) + \mu I(x_1^n; \tilde{y}_{11}^n) \\ & \quad + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\ & = \mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n) + \mu h(\tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n) \\ & \quad + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \\ & = [\mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n)] + [h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n)] \\ & \quad + [\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)] + [h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n)] + n\epsilon_n \end{aligned} \quad (106)$$

Now, we upper bound different the terms within each bracket in (106) separately.

For the first bracket, we have

$$\begin{aligned} \mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n) & \stackrel{(a)}{\leq} \mu \sum_{i=1}^n h(\tilde{y}_{12}[i] | \tilde{y}_{11}[i]) - \frac{\mu n}{2} \log(2\pi e N_{11}) \\ & \stackrel{(b)}{\leq} \mu \sum_{i=1}^n \frac{1}{2} \log \left[2\pi e \left(N_{12} + a'P_2[i] + \frac{P_1[i]N_{11}}{P_1[i] + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\ & \stackrel{(c)}{\leq} \frac{\mu n}{2} \log \left[2\pi e \left(N_{12} + \frac{1}{n} \sum_{i=1}^n a'P_2[i] + \frac{\frac{1}{n} \sum_{i=1}^n P_1[i]N_{11}}{\frac{1}{n} \sum_{i=1}^n P_1[i] + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\ & \leq \frac{\mu n}{2} \log \left[2\pi e \left(N_{12} + a'P_2 + \frac{P_1 N_{11}}{P_1 + N_{11}} \right) \right] - \frac{\mu n}{2} \log(2\pi e N_{11}) \\ & = \frac{\mu n}{2} \log \left(\frac{N_{12}}{N_{11}} + \frac{a'P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right) \end{aligned} \quad (107)$$

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 given covariance matrix, and (c) follows from Jensen's inequality.

The terms within the second bracket can be upper bounded similarly to that of the first one. Hence, we have

$$h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) \leq \frac{n}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right). \quad (108)$$

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

$$\begin{aligned} \mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) &= \mu \left(h(x_1^n + z_{11}^n) - \frac{1}{\mu} h(\sqrt{b'} x_1^n + z_{21}^n) \right) \\ &\leq \mu n f_h \left(P_1, N_{11}, N_{21}, b', \frac{1}{\mu} \right). \end{aligned} \quad (109)$$

Since $1 \leq \mu$, From (63) we obtain

$$\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) \leq \frac{\mu n}{2} \log((2\pi e)(P_1 + N_{11})) - \frac{n}{2} \log((2\pi e)(b' P_1 + N_{21})). \quad (110)$$

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

$$\begin{aligned} h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n) &= h(x_2^n + z_{22}^n) - \mu h(\sqrt{a'} x_2^n + z_{12}^n) \\ &\leq n f_h(P_2, N_{22}, N_{12}, a', \mu). \end{aligned} \quad (111)$$

Adding all inequalities, we obtain

$$\begin{aligned} \mu R_1 + R_2 &\leq \frac{\mu}{2} \log \left(\frac{N_{12}}{N_{11}} + \frac{a' P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}} \right) + \frac{1}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b' P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \\ &\quad + \frac{\mu}{2} \log((2\pi e)(P_1 + N_{11})) - \frac{1}{2} \log((2\pi e)(b' P_1 + N_{21})) + f_h(P_2, N_{22}, N_{12}, a', \mu), \end{aligned} \quad (112)$$

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

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

This completes the proof. \square

The 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 (104) and (105) become, respectively,

$$\mu R_1 + R_2 \leq \mu \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right) \quad (113)$$

and

$$R_1 + \mu R_2 \leq \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \mu \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right). \quad (114)$$

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

$$R_1 = \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) \quad (115)$$

and

$$R_2 = \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right) \quad (116)$$

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 [22]. We summarize this result in the following theorem:

Theorem 4: The sum capacity of Class B channels are attained when receivers treat the interference as noise. In this case, the sum capacity is

$$\mathcal{C}'_{\text{sum}} = \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a' P_2 + N_{12}} \right) + \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b' P_1 + N_{21}} \right). \quad (117)$$

Proof: Simply by substituting $\mu = 1$ in (113), 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 ICs where $1 \leq b$. Structurally, Class C is essentially similar to Class A1 (see Figure 4). However, we impose different constraints on the parameters for the channels within Class C. These constraints assist us in providing new upper bounds by using a different method.

For channels in Class C, we use the same model that is given in (75). Therefore, similar to channels in Class A1, this channel is admissible if the corresponding parameters satisfy

$$\begin{aligned} b'g_2 &= b, \\ (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \quad (118)$$

Here, we change the constraints in (81) to new constraints

$$\begin{aligned} b' &\geq N_{21}, \\ aN_{22} &\leq 1. \end{aligned} \quad (119)$$

By this simple change of constraints, we see that the second receiver after decoding its own signal has a less noisy version of the first user's signal. Therefore, it is able to decode the signal of the first user as well as its own signal. Applying this observation, we have the following lemma.

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

$$\begin{aligned} \sigma_{\mathcal{C}'}(\mu, 1) &\leq \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21}\right)\right) \\ &\quad - \frac{1}{2} \log(2\pi e N_{21}) - \frac{1}{2} \log(2\pi e N_{22}) + f_h(P_2, N_{22}, 1, a, \mu - 1) \end{aligned} \quad (120)$$

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

$$R_1 \leq \frac{1}{n} I(x_1^n; \tilde{y}_1^n), \quad (121)$$

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

$$R_2 \leq \frac{1}{n} I(x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n), \quad (123)$$

$$R_1 + R_2 \leq \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n). \quad (124)$$

Since (122) is redundant, it can be shown that

$$\mu R_1 + R_2 \leq \frac{\mu-1}{n} I(x_1^n; \tilde{y}_1^n) + \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n) \quad (125)$$

Hence, we have

$$\begin{aligned} \mu R_1 + R_2 &\leq \frac{\mu-1}{n} h(\tilde{y}_1^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) + \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n) - \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) \\ &= \frac{\mu-1}{n} h(\tilde{y}_1^n) + \frac{1}{n} h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - \frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) \\ &\quad + \left[\frac{1}{n} h(\tilde{y}_{22}^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) \right] \end{aligned} \quad (126)$$

We now bound different terms of (126). For the first term, we have

$$\frac{\mu-1}{n} h(\tilde{y}_1^n) \leq \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)). \quad (127)$$

The second term can be bounded as

$$\frac{1}{n} h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) \leq \frac{1}{2} \log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21}\right)\right). \quad (128)$$

The third term can be bounded as

$$\frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n, x_2^n) = \frac{1}{2} \log(2\pi e N_{21}) + \frac{1}{2} \log(2\pi e N_{22}). \quad (129)$$

The last terms can be bounded as

$$\frac{1}{n} h(\tilde{y}_{22}^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) = \frac{1}{n} h(x_2^n + z_{22}^n) - \frac{\mu-1}{n} h(\sqrt{a}x_2^n + z_1) \quad (130)$$

$$\leq f_h(P_2, N_{22}, 1, a, \mu - 1) \quad (131)$$

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

IV. WEAK GAUSSIAN INTERFERENCE CHANNELS

In this section, we focus on the weak Gaussian ICs. We first obtain the sum capacity of this channel for some certain range of parameters. Then, we obtain an outer bound on the capacity region which is tighter than previous outer bounds. Finally, we show that using time-sharing parameter and convexification result in the same achievable rate region for this channel when Gaussian distributions are used for generating codebooks.

A. Sum Capacity

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

$$\begin{aligned}
 W = \min \gamma & \left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) \\
 \text{subject to:} \\
 & a'g_1 = a \\
 & b'g_2 = b \\
 & b'N_{11} \leq N_{21} \\
 & a'N_{22} \leq N_{12} \\
 & (1 - \sqrt{g_1})^2 N_{11} + g_1 N_{12} = 1 \\
 & (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1 \\
 & 0 \leq [a', b', g_1, g_2, N_{11}, N_{12}, N_{22}, N_{21}]
 \end{aligned} \tag{132}$$

The objective function in (132) is the sum capacity of Class B channels obtained in Theorem 4. The constraints are the combination of (102) and (103) 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 $\mathcal{C}_{sum} \leq W$. By changing the variables as $S_1 = g_1 N_{12}$ and $S_2 = g_2 N_{21}$, we obtain

$$\begin{aligned}
 W = \min \gamma & \left(\frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) + \gamma \left(\frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) \\
 \text{subject to:} \\
 & \frac{b(1 - S_1)}{(1 - \sqrt{g_1})^2} \leq S_2 < 1 \\
 & \frac{a(1 - S_2)}{(1 - \sqrt{g_2})^2} \leq S_1 < 1 \\
 & 0 < [g_1, g_2]
 \end{aligned} \tag{133}$$

We first minimize the objective in (133) with respect to g_1 and g_2 . To this end, the optimization problem can be decomposed into two separate optimization problems with respect to g_1 and g_2 . The optimization problem with respect to g_1 reads as

$$\begin{aligned}
 W_1 = \min \gamma & \left(\frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) \\
 \text{subject to:} \\
 & \frac{b(1 - S_1)}{S_2} \leq (1 - \sqrt{g_1})^2 \\
 & 0 < g_1
 \end{aligned} \tag{134}$$

It is easy to solve the above optimization problem. In fact, we have

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

Similarly, the optimization problem with respect to g_2 can be written as

$$\begin{aligned}
 W_2 = \min \gamma & \left(\frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) \\
 \text{subject to:} \\
 & \frac{a(1 - S_2)}{S_1} \leq (1 - \sqrt{g_2})^2, \\
 & 0 < g_2,
 \end{aligned} \tag{136}$$

The solution to the above optimization problem is

$$W_2 = \begin{cases} \gamma \left(\frac{P_2}{1+bP_1} \right) & \text{if } \sqrt{a}(1+bP_1) \leq \sqrt{S_1(1-S_2)} \\ \gamma \left(\frac{aP_2}{S_1} + \frac{(1-\sqrt{a(1-S_2)/S_1})^2 P_2}{bP_1+S_2} \right) & \text{Otherwise} \end{cases} \quad (137)$$

Combining (135) and (137), we obtain

$$W = \min W_1 + W_2 \quad (138)$$

subject to:

$$0 < S_1 < 1$$

$$0 < S_2 < 1$$

From (135) and (137), 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 becomes independent of S_1 and S_2 . In this case, we obtain

$$W = \gamma \left(\frac{P_1}{1+aP_2} \right) + \gamma \left(\frac{P_2}{1+bP_1} \right), \quad (139)$$

which is achievable by simple strategy of treating interference as noise. In the following theorem, we prove that it is possible to find appropriate S_1 and S_2 for certain range of parameters.

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

$$\mathcal{C}_{sum} = \gamma \left(\frac{P_1}{1+aP_2} \right) + \gamma \left(\frac{P_2}{1+bP_1} \right), \quad (140)$$

for all channel's parameters satisfying

$$\sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}}. \quad (141)$$

Proof: Let us fix a and b . In order to find all P_1 and P_2 such that we can find $0 < S_1 < 1$ and $0 < S_2 < 1$ satisfying $\sqrt{b}(1+aP_2) \leq \sqrt{S_2(1-S_1)}$ and $\sqrt{a}(1+bP_1) \leq \sqrt{S_1(1-S_2)}$, we define D and D' as follows

$$D = \left\{ (P_1, P_2) \mid P_1 \leq \frac{\sqrt{S_1(1-S_2)}}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{\sqrt{S_2(1-S_1)}}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1, 0 < S_2 < 1 \right\}, \quad (142)$$

$$D' = \left\{ (P_1, P_2) \mid \sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}} \right\}. \quad (143)$$

To show $D' \subseteq D$, we set $S_1 = 1 - S_2$ in (142) to get

$$\left\{ (P_1, P_2) \mid P_1 \leq \frac{S_1}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{1-S_1}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1 \right\} \subseteq D. \quad (144)$$

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

$$J = \max_{(P_1, P_2) \in D} \sqrt{b}P_1 + \sqrt{a}P_2, \quad (145)$$

which can be written as

$$J = \max_{(S_1, S_2) \in (0,1)^2} \frac{\sqrt{S_1(1-S_2)} + \sqrt{S_2(1-S_1)}}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \quad (146)$$

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

$$J = \frac{1}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \quad (147)$$

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

As an example, let us consider the symmetric Gaussian IC. In this case, the constraint in (141) becomes

$$P \leq \frac{1 - 2\sqrt{a}}{2a\sqrt{a}}. \quad (148)$$

In Figure 7, the admissible region for P versus \sqrt{a} , where treating interference as Gaussian noise is optimal, is plotted. In Figure 8, the upper bound in (132) and the lower bound when receivers treat the interference as Gaussian noise are sketched for a fixed P and all $0 \leq a \leq 1$. We observe that up to certain value of a , the upper bound coincides with the lower bound.

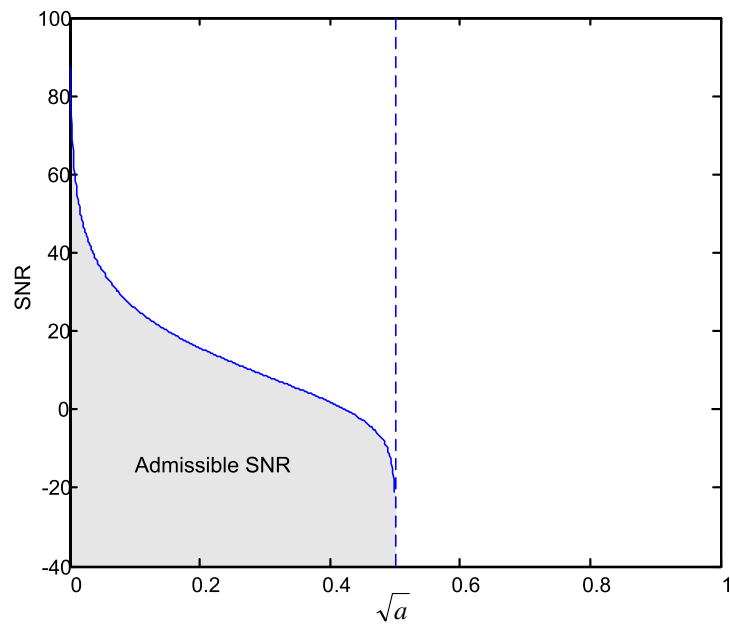


Fig. 7. The shaded area is the region where treating interference as Gaussian noise is optimal for obtaining the sum capacity of the symmetric Gaussian IC.

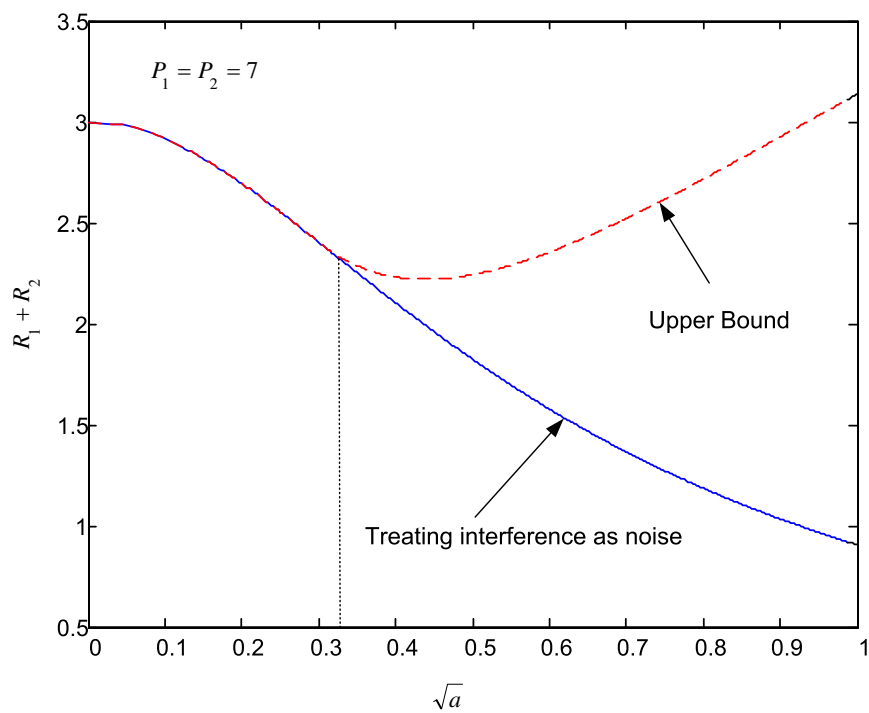


Fig. 8. The upper bound obtained by solving (132). The lower bound is obtained by using the simple scheme of considering the interference as Gaussian noise.

B. New Outer Bound

For the weak Gaussian IC, there are two outer bounds that are tighter than other bounds. The first one, due to Kramer [12], is obtained by considering 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 to derive the outer bound for the original channel. Kramers' outer bound is the intersection of two regions E_1 and E_2 . E_1 is the collection of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma \left(\frac{(1-\beta)P'}{\beta P' + 1/a} \right), \quad (149)$$

$$R_2 \leq \gamma(\beta P'), \quad (150)$$

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

$$R_1 \leq \gamma(\alpha P''), \quad (151)$$

$$R_2 \leq \gamma \left(\frac{(1-\alpha)P''}{\alpha P'' + 1/b} \right), \quad (152)$$

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. [22], is obtained by using the Genie aided technique to upper bound different linear combinations of rates that appear in Han-Kobayashi achievable rate region. Their outer bound is the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma(P_1), \quad (153)$$

$$R_2 \leq \gamma(P_2), \quad (154)$$

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

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

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

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma \left(bP_1 + \frac{P_2}{1 + aP_2} \right) + 0.5 \log \left(\frac{1 + P_1}{1 + bP_1} \right), \quad (158)$$

$$R_1 + 2R_2 \leq \gamma(bP_1 + P_2) + \gamma \left(aP_2 + \frac{P_1}{1 + bP_1} \right) + 0.5 \log \left(\frac{1 + P_2}{1 + aP_2} \right). \quad (159)$$

In the new outer bound that we propose here, an upper bound for each linear combination of rates is derived. 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

$$W_1(\mu) = \min \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] - \frac{\mu_2}{2} \log(2\pi e) + \frac{1}{2} \log \left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}} \right) \quad (160)$$

$$+ \mu_2 f_h \left(P_1, 1, N_{21}, b', \frac{1}{\mu_2} \right) + f_h(P_2, N_{22}, 1, a, \mu_1)$$

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 (80) and (81) where applied to confirm 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 new variable $S = (1 - \sqrt{g_2})^2 N_{22}$, we obtain

$$W_1(\mu) = \min \frac{\mu_1}{2} \log [2\pi e(P_1 + aP_2 + 1)] + \frac{1}{2} \log \left[(1 - \sqrt{g_2})^2 \left(\frac{1 - S + bP_1}{g_2 S} + \frac{P_2}{(1 - \sqrt{g_2})^2 P_2 + S} \right) \right] \quad (161)$$

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

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

$$W_2(\mu) = \min \mu \gamma \left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}} \right) + \gamma \left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}} \right) + f_h(P_2, N_{22}, N_{12}, a', \mu) \quad (162)$$

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

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 optimization problem, the channels in Class B are used. In fact, the objective is the upper bound on the support function 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

$$W_2(\mu) = \min \mu \gamma \left(\frac{(1 - \sqrt{g_1})^2 P_1}{1 - S_1} + \frac{g_1 P_1}{aP_2 + S_1} \right) + \gamma \left(\frac{(1 - \sqrt{g_2})^2 P_2}{1 - S_2} + \frac{g_2 P_2}{bP_1 + S_2} \right) \quad (163)$$

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

subject to:

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

In a similar fashion, one can introduce two optimization problems, say $\tilde{W}_1(\mu)$ and $\tilde{W}_2(\mu)$, to obtain upper bounds on $\sigma_{\mathcal{C}}(1, \mu)$ by using the upper bounds on 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

$$\mu_1 R_1 + R_2 \leq W(\mu_1) = \min\{W_1(\mu_1), W_2(\mu_1)\}, \quad (164)$$

$$R_1 + \mu_2 R_2 \leq \tilde{W}(\mu_2) = \min\{\tilde{W}_1(\mu_2), \tilde{W}_2(\mu_2)\}, \quad (165)$$

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

To obtain an upper bound for the sum rate, we can readily use the following inequality:

$$\mathcal{C}_{\text{sum}} \leq \min_{1 \leq \mu_1, \mu_2} \frac{(\mu_2 - 1)W(\mu_1) + (\mu_1 - 1)\tilde{W}(\mu_2)}{\mu_1 \mu_2 - 1}. \quad (166)$$

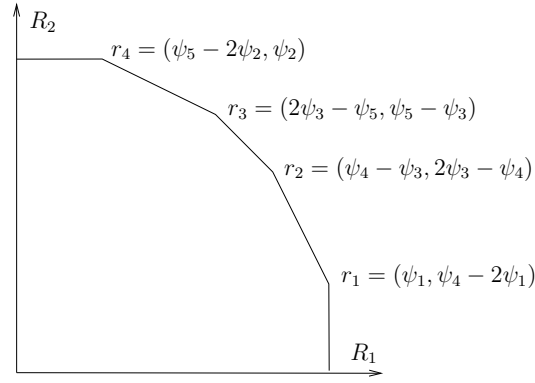


Fig. 9. \mathcal{G}_0 for the weak Gaussian IC. r_1 , r_2 , r_3 , and r_4 are extreme points of \mathcal{G}_0 in the interior of the first quadrant.

C. Han-Kobayashi Achievable region

In this 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, it is easy to show that none of the inequalities in describing \mathcal{G}_0 is redundant. Moreover, \mathcal{G} has at most seven extreme points where among them $\{r_1, r_2, r_3, r_4\}$ are in the interior of the first quadrant (see Figure 9), where

$$r_1 = (\psi_1, \psi_4 - 2\psi_1), \quad (167)$$

$$r_2 = (\psi_4 - \psi_3, 2\psi_3 - \psi_4), \quad (168)$$

$$r_3 = (2\psi_3 - \psi_5, \psi_5 - \psi_3), \quad (169)$$

$$r_4 = (2\psi_5 - 2\psi_2, \psi_2). \quad (170)$$

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

$$\sigma_{D_0}(c_1, c_2, P_1, P_2, \alpha, \beta) = \min\{\mathbf{y}^t \Psi(P_1, P_2, \alpha, \beta) | A^t \mathbf{y} = (c_1, c_2)^t, 0 \leq \mathbf{y}\}, \quad (171)$$

is independent of parameters P_1 , P_2 , α , and β 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 minimum of (171) is attained at $\hat{\mathbf{y}} = (\mu - 2, 0, 0, 1, 0)^t$ which is clearly independent of P_1 , P_2 , α , and β . In this case, we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (\mu - 2)\psi_1 + \psi_4. \quad (172)$$

For $1 \leq \mu \leq 2$ one can show that $\hat{\mathbf{y}} = (0, 0, 2 - \mu, \mu - 1, 0)^t$ minimizes (171), and it is again independent of P_1 , P_2 , α , and β . In this case, we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (2 - \mu)\psi_3 + (\mu - 1)\psi_4. \quad (173)$$

Now, 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 (171), respectively. Hence, we have

$$\sigma_{D_0}(1, \mu, P_1, P_2, \alpha, \beta) = (\mu - 2)\psi_2 + \psi_5, \quad \text{if } 2 < \mu, \quad (174)$$

$$\sigma_{D_0}(1, \mu, P_1, P_2, \alpha, \beta) = (2 - \mu)\psi_3 + (\mu - 1)\psi_5, \quad \text{if } 1 \leq \mu \leq 2, \quad (175)$$

Theorem 7: For the two-user weak Gaussian IC, time-sharing and convexification 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 can be bounded to be 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 (45). Since \mathcal{G}_0 possesses the unique minimizer property, (45) can be simplified easily. Let us consider the case where $(c_1, c_2) = (\mu, 1)$ for $\mu > 2$. It can be shown that for this case

$$g = \max_{(\alpha, \beta) \in [0, 1]^2} (\mu - 2)\psi_1(P_1, P_2, \alpha, \beta) + \psi_4(P_1, P_2, \alpha, \beta). \quad (176)$$

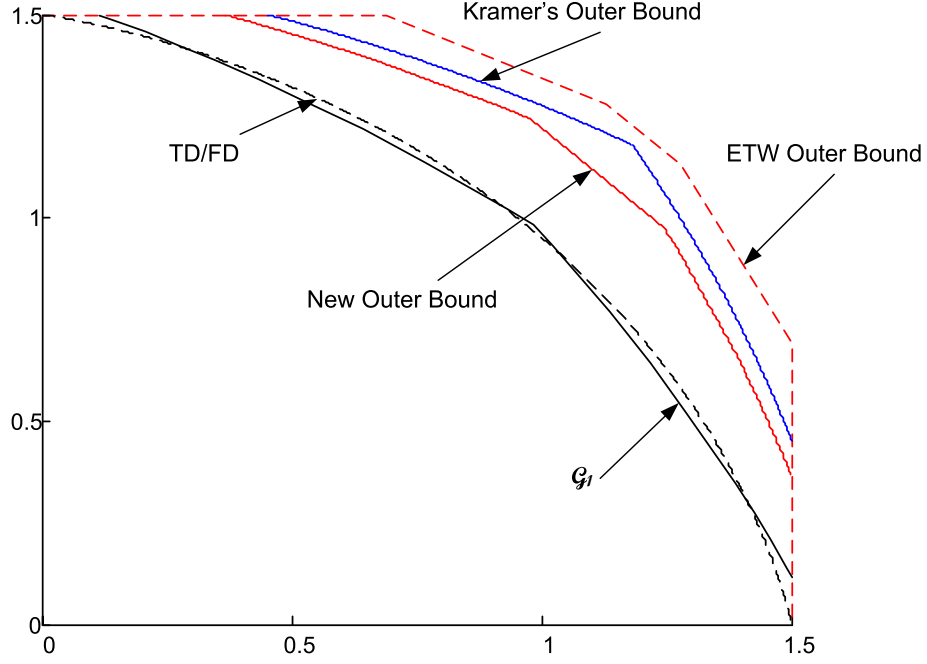


Fig. 10. Comparison between different bounds for the symmetric weak Gaussian IC when $P = 7$ and $a = 0.2$

Substituting into (43), we obtain

$$\sigma_{\mathcal{G}_2}(\mu, 1, P_1, P_2) = \max \sum_{i=1}^3 \lambda_i [(\mu - 2)\psi_1(P_{1i}, P_{2i}, \alpha_i, \beta_i) + \psi_4(P_{1i}, P_{2i}, \alpha_i, \beta_i)] \quad (177)$$

subject to:

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

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

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

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

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

For other range of (c_1, c_2) , a similar optimization problem can be defined. In Figures (10) and (11), different bounds for the symmetric weak Gaussian IC are plotted. As depicted in figures, the new outer bound is tighter than previous outer 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 ICs and the weak one-sided ICs. For the former, $a \geq 1$ and the capacity region is fully characterized, c.f. [14]. In this case, the capacity region is the union of all rate pairs (R_1, R_2) satisfying

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

For the latter, however, $a < 1$ 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 ICs are equivalent to that of the degraded ICs with some appropriate parameter changes. The second one,

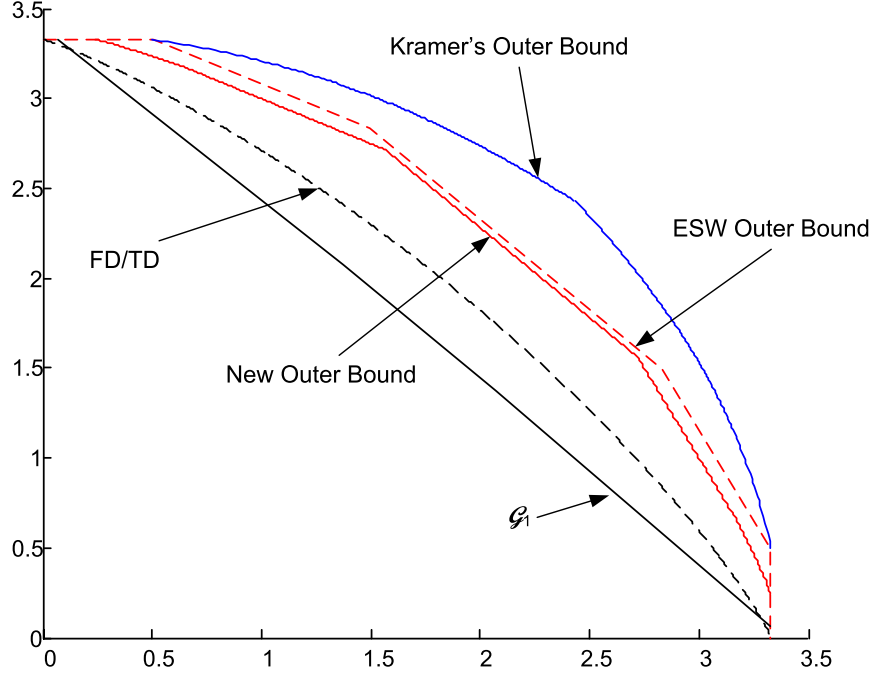


Fig. 11. Comparison between different bounds for the symmetric weak Gaussian IC when $P = 100$ and $a = 0.1$.

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 [14], states the sum capacity of this channel.

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

A. Sum Capacity

For the sake of completeness, we 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.

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. Outer Bound

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

$$\begin{aligned} R_1 &\leq \gamma\left(\frac{(1-\beta)P}{1/a+\beta P}\right) \\ R_2 &\leq \gamma(\beta P) \end{aligned} \quad (178)$$

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

Proof: 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 given covariance matrix constraint and (b) follows from definition of f_h in (61).

Recall that it suffices to consider $1 \leq \mu$. Depending on μ , we consider two cases.

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

$$\mu R_1 + R_2 \leq \mu \gamma \left(\frac{P_1}{1+aP_2} \right) + \gamma(P_2). \quad (179)$$

In fact, the point $\left(\gamma \left(\frac{P_1}{1+aP_2} \right), \gamma(P_2) \right)$ which is achievable by simply treating interference as noise at Receiver 1, satisfies (179) 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.

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

$$\mu R_1 + R_2 \leq \frac{\mu}{2} \log(P_1 + aP_2 + 1) + \frac{1}{2} \log \left(\frac{1/a - 1}{\mu - 1} \right) - \frac{\mu}{2} \log \left(\frac{\mu a(1/a - 1)}{\mu - 1} \right). \quad (180)$$

Equivalently, we have

$$\mu R_1 + R_2 \leq \frac{\mu}{2} \log \left(\frac{(aP+1)(\mu-1)}{\mu(1-a)} \right) + \frac{1}{2} \log \left(\frac{1/a - 1}{\mu - 1} \right), \quad (181)$$

where $P = P_1/a + P_2$. Let us define two sets E_1 and E_2 as

$$E_1 = \left\{ (R_1, R_2) | R_1 \leq \gamma \left(\frac{(1-\beta)P}{1/a + \beta P} \right), R_2 \leq \gamma(\beta P), \forall \beta \in [0, 1] \right\} \quad (182)$$

and

$$E_2 = \left\{ (R_1, R_2) | \mu R_1 + R_2 \leq \frac{\mu}{2} \log \left(\frac{(aP+1)(\mu-1)}{\mu(1-a)} \right) + \frac{1}{2} \log \left(\frac{1/a - 1}{\mu - 1} \right), \forall \frac{P_2+1/a}{P_2+1} < \mu \leq \frac{1}{a} \right\}. \quad (183)$$

In fact, E_2 is the dual representation of E_1 , see (5). To show this, we evaluate the support function of E_1 as

$$\sigma_{E_1}(\mu, 1) = \max \{ \mu R_1 + R_2 | (R_1, R_2) \in E_1 \}. \quad (184)$$

It is easy to show that

$$\sigma_{E_1}(\mu, 1) = \frac{\mu}{2} \log \left(\frac{(aP+1)(\mu-1)}{\mu(1-a)} \right) + \frac{1}{2} \log \left(\frac{1/a - 1}{\mu - 1} \right). \quad (185)$$

Since E_1 is a closed convex set, we can make use of (5) to obtain the dual representation of it which is indeed equivalent to (183). 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 Han-Kobayashi achievable rate region is only private message, i.e., $\alpha = 0$. In this case, we have

$$\psi_1 = \gamma \left(\frac{P_1}{1+a\beta P_2} \right), \quad (186)$$

$$\psi_2 = \gamma(P_2), \quad (187)$$

$$\psi_{31} = \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (188)$$

$$\psi_{32} = \gamma \left(\frac{P_1}{1+a\beta P_2} \right) + \gamma(P_2), \quad (189)$$

$$\psi_{33} = \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (190)$$

$$\psi_4 = \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right) + \gamma \left(\frac{P_1}{1+a\beta P_2} \right) + \gamma(\beta P_2), \quad (191)$$

$$\psi_5 = \gamma(\beta P_2) + \gamma(P_2) + \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1+a\beta P_2} \right), \quad (192)$$

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

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

$$R_2 \leq \gamma(P_2), \quad (194)$$

$$R_1 + R_2 \leq \gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2). \quad (195)$$

In the following lemma, we provide a simple description of \mathcal{G}_1 .

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

$$R_1 \leq \gamma \left(\frac{P_1}{1 + \beta a P_2} \right) \quad (196)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left(\frac{(1 - \beta)a P_2}{1 + P_1 + \beta a P_2} \right) \quad (197)$$

Moreover, \mathcal{G}_1 is convex and any point that lies on the boundary of the region can be achieved by using superposition coding and successive decoding.

Proof: Let E denote the set defined in the lemma. It is easy to prove that $E \subseteq \mathcal{G}_1$. Hence, we need to show that the inverse inclusion holds. To this end, let us select an arbitrary point inside \mathcal{G}_1 , say (R'_1, R'_2) . Hence, there exist a β' such that R'_1 and R'_2 satisfy (193), (194), and (195). Since $\gamma(\beta P_2) + \gamma \left(\frac{(1 - \beta)a P_2}{1 + P_1 + \beta a P_2} \right)$ in (197) is a continuous function with respect to β over a compact set, there is a $\beta' \leq \beta \leq 1$ such that

$$R'_2 = \gamma(\beta P_2) + \gamma \left(\frac{(1 - \beta)a P_2}{1 + P_1 + \beta a P_2} \right). \quad (198)$$

For this β , every point (R_1, R'_2) with $R_1 \leq \gamma \left(\frac{P_1}{1 + \beta a P_2} \right)$ is in E . Hence, we need to show $R'_1 \leq \gamma \left(\frac{P_1}{1 + \beta a P_2} \right)$. From (193), (194), and (195), we have

$$R'_1 \leq \min \left\{ \gamma \left(\frac{P_1}{1 + \beta' a P_2} \right), 0.5 \log \left(\frac{1 + \beta' P_2}{1 + \beta P_2} \right) + 0.5 \log \left(\frac{1 + P_1 + \beta a P_2}{1 + \beta' a P_2} \right) \right\}. \quad (199)$$

It is easy to show that the right hand side of the above inequality is less than $\gamma \left(\frac{P_1}{1 + \beta a P_2} \right)$ when $\beta' \leq \beta$. Hence, $E = \mathcal{G}_1$.

By having a new description, It is straightforward to show that \mathcal{G}_1 is convex and the boundary points are achievable by using superposition coding and successive decoding. \square

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

$$r_1 = \left(\gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \gamma \left(\frac{(1 - \beta)a P_2}{1 + P_1 + \beta a P_2} \right) + \gamma(\beta P_2) \right), \quad (200)$$

$$r_2 = \left(\gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2) - \gamma(P_2), \gamma(P_2) \right). \quad (201)$$

From this, it can be readily verified that \mathcal{G}_0 indeed possesses the unique minimizer property.

Now, we can use the optimization defined in (43) 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

$$g(\mu, 1, P_1, P_2, \beta) = \max_{0 \leq \beta \leq 1} \mu \gamma \left(\frac{P_1}{1 + \beta a P_2} \right) + \gamma(\beta P_2) + \gamma \left(\frac{(1 - \beta)a P_2}{1 + P_1 + \beta a P_2} \right) \quad (202)$$

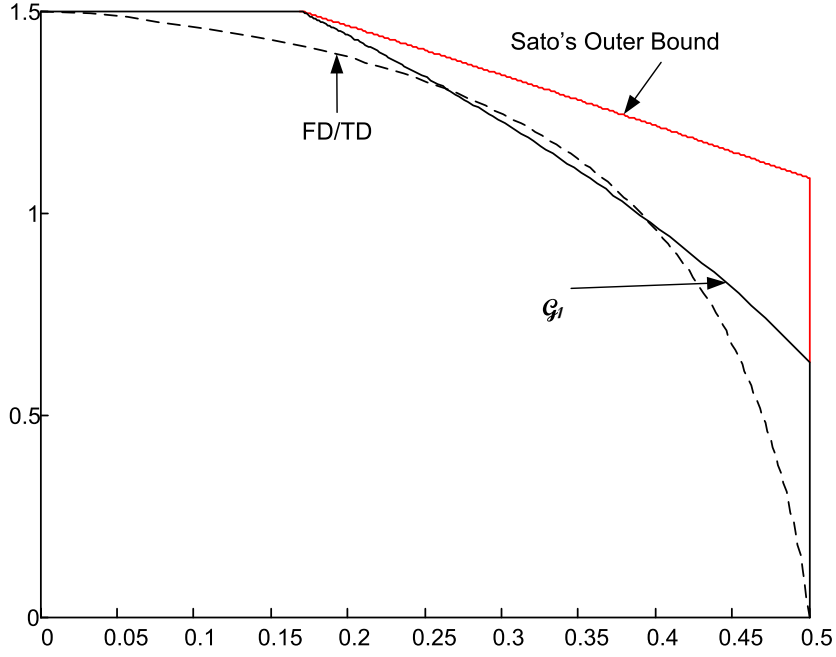


Fig. 12. Comparison between different bounds for the one-sided Gaussian IC when $P_1 = 1$, $P_2 = 7$, and $a = 0.4$.

Substituting into (43), we conclude that boundary points of \mathcal{G} can be characterized by solving the following optimization problem:

$$\begin{aligned}
 W = \max \sum_{i=1}^3 \lambda_i & \left[\mu \gamma \left(\frac{P_{1i}}{1 + \beta_i a P_{2i}} \right) + \gamma(\beta_i P_{2i}) + \gamma \left(\frac{(1 - \beta_i) a P_{2i}}{1 + P_{1i} + \beta_i a P_{2i}} \right) \right] \\
 \text{subject to:} & \\
 \sum_{i=1}^3 \lambda_i &= 1 \\
 \sum_{i=1}^3 \lambda_i P_{1i} &\leq P_1 \\
 \sum_{i=1}^3 \lambda_i P_{2i} &\leq P_2 \\
 0 \leq \beta_i &\leq 1 \quad \forall i \in \{1, 2, 3\} \\
 0 \leq [P_{1i}, P_{2i}, \lambda_i] &\quad \forall i \in \{1, 2, 3\}
 \end{aligned} \tag{203}$$

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 to the capacity region. Finally, we investigate the Han-Kobayashi achievable rate region. Without loss of generality, we assume $a < 1$ and $b \geq 1$.

A. Sum Capacity

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

$$\mathcal{C}_{sum} = \gamma(P_2) + \min \left\{ \gamma \left(\frac{P_1}{1 + a P_2} \right), \gamma \left(\frac{b P_1}{1 + P_2} \right) \right\}. \tag{204}$$

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 Gaussian noise at both receivers. In this case, the rate

$$R_1 = \min \left\{ \gamma \left(\frac{P_1}{1 + aP_2} \right), \gamma \left(\frac{bP_1}{1 + P_2} \right) \right\} \quad (205)$$

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

$$R_2 = \gamma(P_2). \quad (206)$$

By adding (205) and (206), we obtain the desired result.

Converse part: The sum capacity of the Gaussian IC is upper bounded by that of the underlying two one-sided Gaussian ICs. Hence, we can obtain two upper bounds for 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 [14]. Hence, we have

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

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 also known. Hence, we have

$$\mathcal{C}_{sum} \leq \gamma(bP_1 + P_2) \quad (208)$$

which equivalently can be written as

$$\mathcal{C}_{sum} \leq \gamma(P_2) + \gamma \left(\frac{bP_1}{1 + P_2} \right). \quad (209)$$

By taking the minimum of the right hand sides of Inequalities (207) and (209), we obtain

$$\mathcal{C}_{sum} \leq \gamma(P_2) + \min \left\{ \gamma \left(\frac{P_1}{1 + aP_2} \right), \gamma \left(\frac{bP_1}{1 + P_2} \right) \right\}. \quad (210)$$

This completes the proof. \square

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. Outer Bound

The best outer bound to date, due to Etkin et al. [22], is obtained by using the Genie aided technique. This bound is the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma(P_1) \quad (211)$$

$$R_2 \leq \gamma(P_2) \quad (212)$$

$$R_1 + R_2 \leq \gamma(P_2) + \gamma \left(\frac{P_1}{1 + aP_2} \right) \quad (213)$$

$$R_1 + R_2 \leq \gamma(P_2 + bP_2) \quad (214)$$

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma \left(bP_1 + \frac{P_2}{1 + aP_2} \right) + \gamma \left(\frac{P_1}{1 + bP_1} \right) \quad (215)$$

One can obtain an outer bound using a similar approach as that of Kramer's outer bound for the weak Gaussian IC. In fact, the capacity region of the mixed Gaussian IC is inside the intersection of the capacity region 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

$$R_1 \leq \gamma \left(\frac{(1 - \beta)P'}{\beta P' + 1/a} \right) \quad (216)$$

$$R_2 \leq \gamma(\beta P') \quad (217)$$

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

$$R_1 \leq \gamma(bP_1), \quad (218)$$

$$R_2 \leq \gamma(P_2), \quad (219)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2). \quad (220)$$

We also upper bound $\sigma_{\mathcal{C}}(\mu, 1)$ by making use of the channels in Class C. We introduce the following optimization problem

$$W(\mu) = \min \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log \left(2\pi e \left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b'P_1 + N_{21} \right) \right) - \frac{1}{2} \log(2\pi e N_{21}) - \frac{1}{2} \log(2\pi e N_{22}) + f_h(P_2, N_{22}, 1, a, \mu-1) \quad (221)$$

subject to:

$$b'g_2 = b$$

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

$$aN_{22} \leq 1$$

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

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

Clearly, $\sigma_{\mathcal{C}}(\mu, 1) \leq W(\mu)$ for all $\mu > 1$. By substituting $S = g_2 N_{21}$, we obtain

$$W(\mu) = \min \frac{\mu-1}{2} \log(2\pi e(P_1 + aP_2 + 1)) + \frac{1}{2} \log \left(2\pi e \left(\frac{P_2(1-S)}{(1-\sqrt{g_2})^2 P_2 + 1 - S} + \frac{bP_1 + S}{g_2} \right) \right) - \frac{1}{2} \log \left(\frac{2\pi e S}{g_2} \right) - \frac{1}{2} \log \left(\frac{2\pi e(1-S)}{(1-\sqrt{g_2})^2} \right) + f_h \left(P_2, \frac{1-S}{(1-\sqrt{g_2})^2}, 1, a, \mu-1 \right) \quad (222)$$

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 introduces 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

$$\mu R_1 + R_2 \leq W(\mu) \quad (223)$$

holds for all $1 \leq \mu$.

C. Han-Kobayashi Achievable Region

In this subsection, we study the Han-Kobayashi achievable rate 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 βP_2 and $(1-\beta)P_2$ to its private and common messages, respectively, where $\beta \in [0, 1]$. Therefore, we have

$$\psi_1 = \gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \quad (224)$$

$$\psi_2 = \gamma(P_2), \quad (225)$$

$$\psi_{31} = \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2), \quad (226)$$

$$\psi_{32} = \gamma(P_2 + bP_1), \quad (227)$$

$$\psi_{33} = \gamma \left(\frac{a(1-\beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1), \quad (228)$$

$$\psi_4 = \gamma \left(\frac{P_1 + a(1-\beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1), \quad (229)$$

$$\psi_5 = \gamma(\beta P_2) + \gamma(P_2 + bP_1) + \gamma \left(\frac{a(1-\beta)P_2}{1 + a\beta P_2} \right). \quad (230)$$

To characterize \mathcal{G}_1 , we distinguish between 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 $\psi_{31} + \psi_1 \leq \psi_4$ and $\psi_{31} + \psi_2 \leq \psi_5$ hold. This means that Inequalities (9) and (10) are redundant for all range of parameters and can be removed. Hence, \mathcal{G}_0 consists of all rate pairs (R_1, R_2) satisfying

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

$$R_2 \leq \gamma(P_2), \quad (232)$$

$$R_1 + R_2 \leq \gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2), \quad (233)$$

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

$$R_1 \leq \gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \quad (234)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left(\frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right) \quad (235)$$

$$(236)$$

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

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

Proof: When $1 \leq ab$, $1 + P_2 \leq b + abP_2$ holds for any pair of powers (P_1, P_2) . Hence, $\mathcal{G}_0(P_1, P_2, \beta)$ is a pentagon defined by inequalities (231), (232), and (232). Comparing with $\mathcal{G}_0(P_1, P_2, \beta)$ of the one-sided Gaussian IC, we see that the two channels have the same \mathcal{G}_0 region. This directly implies that \mathcal{G} is the same for both channel. \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., $\mathcal{G}_0 = E_1 \cup E_2 \cup E_3$. E_1 is the union of all rate pairs (R_1, R_2) satisfying

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

$$R_2 \leq \gamma(\beta P_2) + \gamma \left(\frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right). \quad (238)$$

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

$$R_1 \leq \gamma \left(\frac{bP_1}{1 + \beta P_2} \right), \quad (239)$$

$$R_2 \leq \gamma \left(\frac{P_1 + a(1 - \beta)P_2}{1 + a\beta P_2} \right) + \gamma(\beta P_2) - \gamma \left(\frac{bP_1}{1 + \beta P_2} \right). \quad (240)$$

for all $\beta \in [\frac{b-1}{(1-ab)P_2}, \frac{(b-1)P_1 + (1-a)P_2}{(1-ab)P_1P_2 + (1-a)P_2}]$. E_3 is the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma \left(\frac{bP_1(1 + \frac{(1-ab)P_1}{1-a})}{1 + bP_1 + P_2} \right) \quad (241)$$

$$R_2 \leq \gamma(P_2) \quad (242)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2) \quad (243)$$

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 E_1 , E_2 , and E_3 is defined as follows. E_1 is the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \quad (244)$$

$$R_2 \leq \gamma(\beta P_2) + \gamma \left(\frac{a(1 - \beta)P_2}{1 + P_1 + a\beta P_2} \right). \quad (245)$$

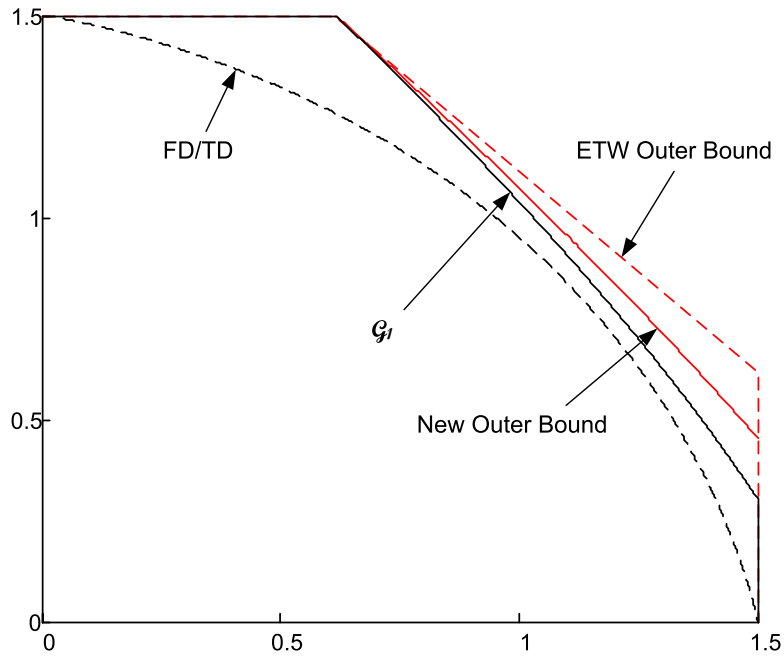


Fig. 13. Comparison between different bounds for the mixed Gaussian IC when $1 + P_2 \leq b + abP_2$ (Case I). Here $P_1 = 7$, $P_2 = 7$, $a = 0.6$, and $b = 2$.

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

$$R_1 \leq \gamma \left(\frac{P_1}{1 + a\beta P_2} \right), \quad (246)$$

$$R_2 \leq \gamma \left(\frac{a(1-\beta)P_2}{1 + P_1 + a\beta P_2} \right) + \gamma(\beta P_2 + bP_1) - \gamma \left(\frac{P_1}{1 + a\beta P_2} \right). \quad (247)$$

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

$$R_1 \leq \gamma \left(\frac{P_1}{1 + aP_2} \right) \quad (248)$$

$$R_2 \leq \gamma(P_2) \quad (249)$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2) \quad (250)$$

Remark 3: Region E_3 for Cases II and III represents a facet that belongs to the capacity region of the mixed Gaussian ICs. Surprisingly, this facet is obtainable when both common and private messages are used at the second transmitter.

Different bounds are compared for the mixed Gaussian ICs for Cases I, II, and III in Figures 13, 14, and 15, respectively.

VII. CONCLUSION

We have investigated data transmission over the M -user interference channel when transmitters use single codebooks for data transmission, and receivers are allowed to decode other users' data. The basic problem of finding the maximum decodable subset of users is addressed. By establishing the main properties of the maximum decodable subset, we have proposed a polynomial time algorithm that separate the interfering users into two disjoint parts, namely the users that the receiver is able to jointly decode them and the rest. We have introduced an optimization problem that gives us an achievable rate for a channel with finite number of interfering users. A polynomial time algorithm for solving this optimization problem has been proposed. The capacity of the additive Gaussian channel with Gaussian interfering users is established and it is shown that the Gaussian distribution is optimal and the proposed achievable rate is the capacity of this channel. Using this result, we have established some points on the capacity region of the generalized Z Gaussian ICs.

For the M -user Gaussian IC, we have characterized some extreme points of the achievable rate region corresponding to successively maximization of users' rates for any permutation of users. We have also established the capacity region for the strong generalized Z Gaussian ICs.

We have studied data transmission over M -user ICs. When there is a rate game between users, we have proven that there exist a fixed point for this game. We have investigated the conditions that the fixed point of the game corresponds to the users' conservative rates.

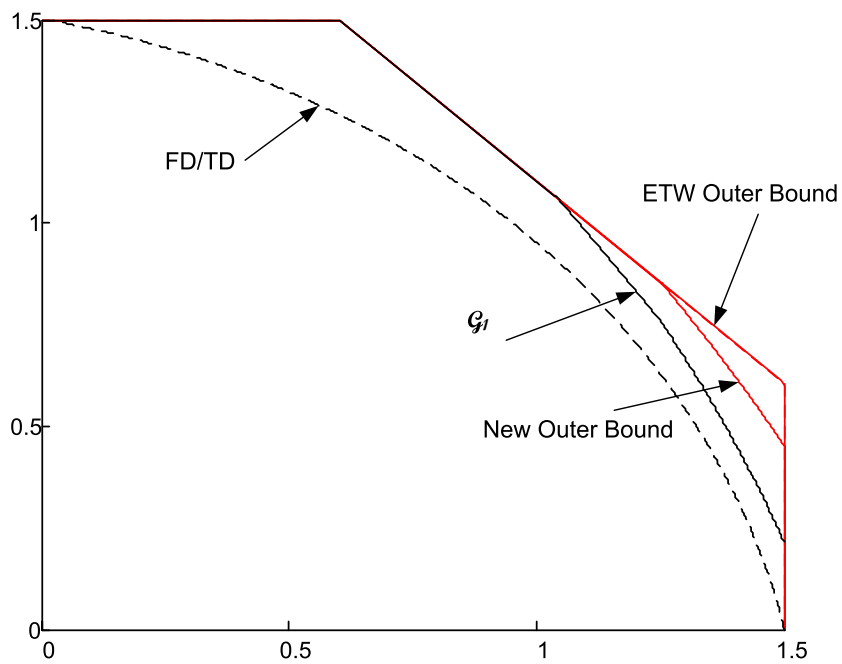


Fig. 14. Comparison between different bounds for the mixed Gaussian IC when $1 + P_2 > b + abP_2$ and $1 - a \leq abP_1$ (Case II). Here $P_1 = 7$, $P_2 = 7$, $a = 0.4$, and $b = 1.5$.

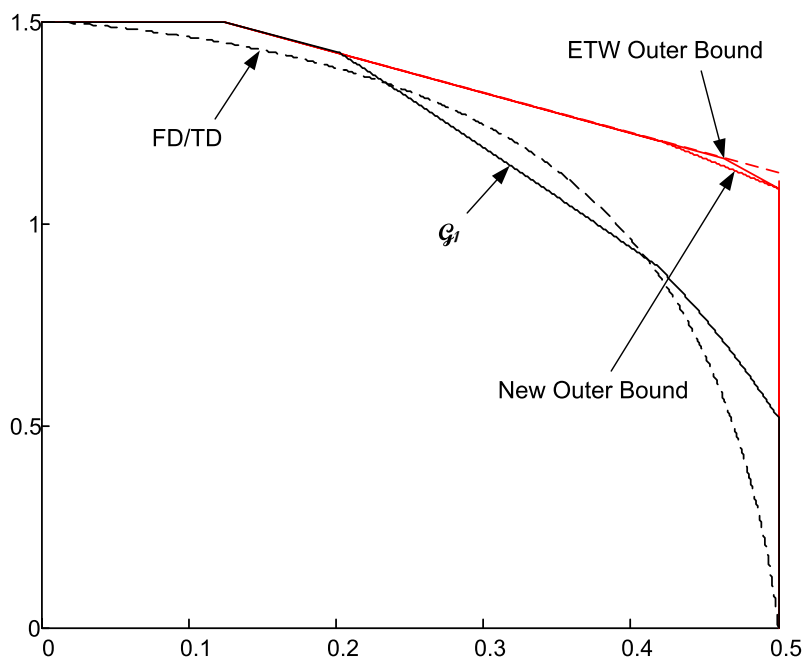


Fig. 15. Comparison between different bounds for the mixed Gaussian IC when $1 + P_2 > b + abP_2$ and $1 - a > abP_1$ (Case III). Here $P_1 = 1$, $P_2 = 7$, $a = 0.3$, and $b = 1.5$.

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. Tsahkadsor, 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] 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.
- [14] I. Sason, "On achievable rate regions for the gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1345–1356, June 2004.
- [15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [16] T. Han, "The capacity region of general multiple-access channel with certain correlated sources," *Inform. Contr.*, vol. 40, no. 1, pp. 37–60, 1979.
- [17] 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.
- [18] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer-Verlag, Berlin Heidelberg, 1998.
- [19] 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.
- [20] S. Ihara, "On the capacity of channels with additive non-Gaussian noise," *Info. Contr.*, vol. 37, no. 1, pp. 34–39, Apr. 1978.
- [21] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 374–377, May 1978.
- [22] 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.