Capacity Bounds for the Gaussian Interference Channel

by

Abolfazl S. Motahari and Amir K. Khandani

Electrical and Computer Engineering Department
University of Waterloo
Waterloo, Ontario, Canada
Email: {abolfazl, khandani}@cst.uwaterloo.ca

Technical Report UW-ECE #2007-26

August 2007

2007-26
(569)

COPY

Legal Deposit # 968763

PUBLICATION FOR LISTING IN CANADIANA - CANADA'S NATIONAL BIBLIOGRAPHY
PUBLICATION À INSCRIRE DANS CANADIANA - BIBLIOGRAPHIE NATIONALE DU CANADA

The information requested is used for cataloguing Canadian publications. If this title was submitted for Canadian Cataloguing in Publication (CIP) Data, only show changes or information not previously submitted.
Les renseignements servent au catologage des publications canadiennes. Si le titre a déjà été soumis en vue du programme CIP (Catalogage avant publication) canadien, n'indiquer que les changements ou les renseignements nouveaux.

Authors' names in full - Nom et prénom des auteurs	Birthdate - Date de naissance	Canadian Citizen / Citoyen canadien	Landed immigrant / Immigrant reçu	Non-Canadian / Étranger
Motahari, Abolfazl S.	September 2, 1977			X
Khandani, Amir Keyvan	December 2, 1959	X		

Title of publication - Titre de la publication	Publisher - Maison d'édition
Capacity Bounds for the Gaussian Interference Channel	University of Waterloo

Where can booksellers and libraries obtain copies?
Où les librairies et les bibliothèques peuvent-elles se procurer des exemplaires?

University of Waterloo
Dept. of Electrical and Computer Engineering
200 University Avenue West
Waterloo, ON N2L 3G1
CANADA

ISBN number - Numéro d'ISBN	Retail sale price - Prix de détail

Date published - Date d'édition	Print run - Tirage
August 2007	1st run

Is distribution limited? La distribution est-elle restreinte? ☐ Yes (*to whom?*) Oui (*à qui?*) ☒ No - Non

Signature

Date August 24, 2007

Telephone no. - Nº de téléphone 519-885-1211
Fax no. - Nº de télécopieur 519-746-3077
Internet www.ece.uwaterloo.ca

This form should be completed and attached to the copies sent for legal deposit purposes.
Annexer la formule dûment remplie aux exemplaires envoyés conformément au règlement sur le dépôt légal.

NL/BN 15 (95-03)

Sent to National
Library Aug 24/07
Betty Slowinski

Library and Archives Canada
Bibliothèque et Archives Canada

RECEIPT FOR LEGAL DEPOSIT
REÇU POUR DÉPÔT LÉGAL

Legal deposit number
Numéro de dépôt légal
D 968763

ISBN / ISSN

Motahari, Abolfazl S. (Abolfazl Sayed), 1977-
Capacity bounds for the Gaussian interference channel / by Abolfazl S. Motahari and A. K. Khandani
[Waterloo, Ont.] : University of Waterloo. Dept. of Electrical & Computer Engineering, 2007.

Quantity received:
Quantité reçue : 2

Price:
Prix : 0

Date received:
Date de réception : 2007-09-06

University of Waterloo. Dept. of
Electrical & Computer Engineering
Betty Slowinski
200 University AVE W
WATERLOO ON N2L 3G1

LAC-BAC 0054 (2006/06)

Canadä

THAUVA
John Stegenga

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 that outperforms previously known outer bounds is obtained. The sum capacity of the channel for some certain range of channel's parameters is derived. It is shown that the full Han-Kobayashi achievable rate region when Gaussian codebooks are used 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. The full Han-Kobayashi achievable rate region is derived. For the mixed Gaussian IC, a new outer bound that again outperforms previously known outer bounds is obtained. The sum capacity of the channel for all ranges of parameters is derived. It is proved that the full Han-Kobayashi achievable rate region is equivalent to that of the one-sided Gaussian IC for some range of 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. In this paper, we study the capacity region of the two-user Gaussian IC.

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 (GIC), 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 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 stil 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 number of inequalities for the Han-Kobayashi achievable rate region. Having less number of inequalities decreases the cardinality of the time-sharing parameter, 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 outperforms 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 with a certain degraded broadcast channel that its 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 is based on the Genie aided technique. 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, Etkin, Tse, and Wang's genie provides information about the intended signal to the receiver. Remarkably, they have shown that the new outer bound outperforms Kramer's one for some range of parameters. Moreover, using 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 outer bounds tighter than previously proposed outer bounds for both weak and mixed Gaussian ICs. More importantly, the sum capacity of the Gaussian weak IC for some certain range of channel's 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 two methods, time-sharing and concavification, that enlarge the basic Han-Kobayashi achievable rate region. We investigate conditions for which the two regions obtained from time-sharing and concavification coincide. Finally, we consider an optimization problem (extremal inequality) and derive optimum solutions of the problem. In fact, the extremal inequality is used thought the paper.

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

In Section IV, we study capacity region of the weak Gaussian IC. We first derive the sum capacity of this channel for some range 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 obtain 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 concavification. 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 is equivalent to that of the one-sided case. Finally, in Section VII, we conclude the paper.

II. PRELIMINARIES

A. Notations

Throughout this paper, we use the following notations. Vectors are represented by bold faced letters. However, for vectors representing codewords we use the usual notation x^n. Random variables, matrices, and sets are denoted by capital letters where the difference is clear from the context. $|A|$, $tr\{A\}$, and A^t represent respectively the determinant, trace, and transpose of the square matrix A. I denotes the identity matrix. $\mathbb{N}$ and $\Re$ 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 $0.5 \log_2(1 + x)$.

B. The Two-user Interference Channel

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

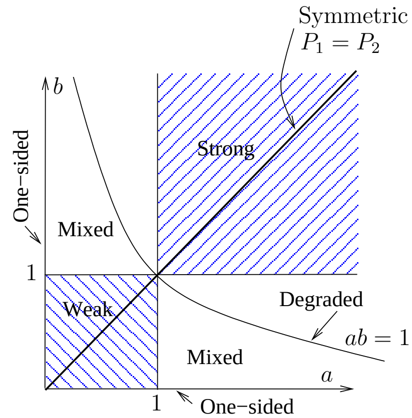

Fig. 1. Classes of the two-user ICs.

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

1) A uniform distributed message set $\mathcal{M}_i \in [1,2,...,2^{nR_i}]$.
2) A codebook $\mathcal{X}_i = \{\mathbf{x}_i(1), \mathbf{x}_i(2), ..., \mathbf{x}_i(2^{nR_1})\}$, where $\mathbf{x}_i(\cdot) \in \mathscr{X}_i^n$.
3) An encoding function $F_i : [1,2,...,2^{nR_i}] \to \mathcal{X}_i$.
4) A decoding function $G_i : \mathbf{y}_i \to [1,2,...,2^{nR_i}]$.
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 $\mathscr{C}_{IC}$ denote the capacity region of the two-user IC. The limiting expression of the capacity region can be stated as [5]

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

In this paper, we focus on the two-user Gaussian IC which can be represented in standard form as [14]

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

where x_i and y_i denote the input and output alphabets, respectively, of user $i \in \{1,2\}$. The $z_1 \sim \mathcal{N}(0,1)$ and $z_2 \sim \mathcal{N}(0,1)$ are standard Gaussian random variables. The 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

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

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

Remark 1: Since the capacity region of the general IC only depends on the marginal distributions [14], the ICs can be classified in equivalent classes in which channels within a class have the same capacity region. In particular, for the Gaussian IC (2), assuming any joint distribution for the pair (z_1, z_2) does not change the capacity region as long as the marginal distributions remain Gaussian with zero mean and unit variance.

Depending on 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 $a-b$ 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 $1 \leq a$ the channel is called mixed Gaussian IC. Finally, when $a = b$ and $P_1 = P_2$, then the channel is called symmetric IC.

Among all, the capacity region of the strong IC is fully characterized []. In this case, the capacity region can be stated as collection of rate pairs (R_1, R_2) satisfying

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

C. Han-Kobayashi Achievable Region

The best inner bound, to date, is the full Han-Kobayashi region denoted by $\mathscr{C}_{HK}$ [8]. In their scheme the message of each user is split into two independent parts, the common part and the private part. The common part is encoded 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. The transmit signal is a deterministic function of common and private parts of the message. In [9], a new description of $\mathscr{C}_{HK}$ with less number of inequalities is obtained. In this paper, we use this new description for characterizing the Han-Kobayashi achievable rate region of the two-user Gaussian IC.

Let us denote the random variables involved in characterizing $\mathscr{C}_{HK}$ as X_{1p}, X_{1c}, X_{2p}, X_{2c}, and Q. X_{ip} and X_{ic} are random variables that carry User i's private and common messages, respectively, for $i \in \{1,2\}$. Q is the time sharing parameter. Even though the Han-Kobayashi scheme has a single letter formula, characterizing the full Han-Kobayashi region is still difficult. In fact, the optimality of the Gaussian distributions for the Han-Kobayashi scheme are not proved yet.

We define $\overline{\mathscr{G}}$ as the special case of the Han-Kobayashi scheme where X_{1p}, X_{1c}, X_{2p}, and X_{2c} are all Gaussian and transmitted signal from Transmitter i is $X_{ip}+X_{ic}$. Moreover, we assume that the cardinality of the time sharing parameter is one. This scheme is call the naive Han-Kobayashi scheme. In the naive Han-Kobayashi scheme αP_1 and $(1-\alpha)P_1$ portion of the first user's power are used for transmitting the private and the common part of the first user's data, respectively, for all $\alpha \in [0,1]$. Similarly, βP_2 and $(1-\beta)P_1$ portion of the second user's power are used for transmitting the private and the common part of the second user's data, respectively, for all $\beta \in [0,1]$. In this case, $\overline{\mathscr{G}}$ is the union of all rate pairs (R_1,R_2) satisfying

$$R_1 \le \rho_1 = \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{4}$$

$$R_2 \le \rho_2 = \gamma\left(\frac{P_2}{1+b\alpha P_1}\right), \tag{5}$$

$$R_1+R_2 \le \rho_{31} = \gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma\left(\frac{\beta P_2}{1+b\alpha P_1}\right), \tag{6}$$

$$R_1+R_2 \le \rho_{32} = \gamma\left(\frac{\alpha P_1}{1+a\beta P_2}\right)+\gamma\left(\frac{P_2+b(1-\alpha)P_1}{1+b\alpha P_1}\right), \tag{7}$$

$$R_1+R_2 \le \rho_{33} = \gamma\left(\frac{\alpha P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma\left(\frac{\beta P_2+b(1-\alpha)P_1}{1+b\alpha P_1}\right), \tag{8}$$

$$2R_1+R_2 \le \rho_4 = \gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma\left(\frac{\alpha P_1}{1+a\beta P_2}\right)+\gamma\left(\frac{\beta P_2+b(1-\alpha)P_1}{1+b\alpha P_1}\right), \tag{9}$$

$$R_1+2R_2 \le \rho_5 = \gamma\left(\frac{\beta P_2}{1+b\alpha P_1}\right)+\gamma\left(\frac{P_2+b(1-\alpha)P_1}{1+b\alpha P_1}\right)+\gamma\left(\frac{\alpha P_1+a(1-\beta)P_2}{1+a\beta P_2}\right), \tag{10}$$

for all $\alpha \in [0,1]$ and $\beta \in [0,1]$. We define ρ_3 as

$$\rho_3 = \min\left\{\rho_{31},\rho_{32},\rho_{33}\right\}. \tag{11}$$

One can use the time sharing parameter to enlarge $\overline{\mathscr{G}}$ to $\mathscr{G}$. Clearly, the relation $\overline{\mathscr{G}} \subseteq \mathscr{G} \subseteq \mathscr{C}_{HK} \subseteq \mathscr{C}$ always holds.

D. Concavification Versus Time-Sharing

Consider an achievable scheme $\mathscr{S}$ for a multiple-user channel with the power constraint $\mathbf{P}=[P_1,P_2,\ldots,P_M]$ is given. We assume that the achievable region associated with $\mathscr{S}$ can be represented as

$$D_0 = \left\{\mathbf{R} | A\mathbf{R} \le \Theta(\mathbf{P})\right\}. \tag{12}$$

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, convex hull operation does not lead to a new enlarged region. However, if extreme points of the region's are not a concave function of $\mathbf{P}$, it is possible to enlarge D_0 by using two different methods which we will explain it now. The first method is to make use of the time sharing parameter. Let us denote this new region as D_1 which can be written as

$$D_1 = \left\{\mathbf{R} | A\mathbf{R} \le \sum_{i=1}^{q}\lambda_i\Theta(\mathbf{P}_i), \sum_{i=1}^{q}\lambda_i\mathbf{P}_i \le \mathbf{P}, \sum_{i=1}^{q}\lambda_i = 1, \lambda_i \ge 0\ \forall i\right\}. \tag{13}$$

In the second method, we split the total power $\mathbf{P}$ as $\sum_{i=1}^{q'}\lambda_i\mathbf{P}_i \le \mathbf{P}$ for some q', $\mathbf{P_i}$s, and λ_is such that $\sum_{i=1}^{q'}\lambda_i\mathbf{P}_i \le \mathbf{P}$ and $\sum_{i=1}^{q'}\lambda_i = 1$. Then, for each power constraint $\mathbf{P}_i$ we use D_0^i as the achievable region obtained from Equation (12) which is

$$D_0^i = \left\{\mathbf{R}_i | A\mathbf{R}_i \le \Theta(\mathbf{P}_i)\right\}. \tag{14}$$

Now, we define the new achievable region D_2 as $D_2 = \sum_{i=1}^{q'} \lambda_i D_0^i$. This region can be stated as

$$D_2 = \left\{ \mathbf{R} = \sum_{i=1}^{q'} \lambda_i \mathbf{R}_i | A\mathbf{R_i} \leq \Theta(\mathbf{P}_i), \sum_{i=1}^{q'} \lambda_i \mathbf{P}_i \leq \mathbf{P}, \sum_{i=1}^{q'} \lambda_i = 1, \lambda_i \geq 0 \ \forall i \right\}. \tag{15}$$

We call this new method as concavification. In fact, the concavification method is equivalent to dividing the available space into subspaces, for example by using TD or FD, and using the given method in each subspace.

It can be readily shown that D_1 and D_2 are closed and convex, and $D_2 \subseteq D_1$. We are interested in situations where the other inclusion holds. To this end, we need 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 $D \in \Re^m$ is a function $\sigma_D : \Re^m \to \Re$ defined as

$$\sigma_D(\mathbf{c}) = \sup\{\mathbf{c}^t\mathbf{R} | \mathbf{R} \in D\}. \tag{16}$$

Clearly, if the set D is compact then the sup is attained and can be replaced by max. In this case, the solutions of (16) correspond to the boundary points of D [15]. The following relation is the dual of (16) and holds when D is closed and convex

$$D = \{\mathbf{R} | \mathbf{c}^t\mathbf{R} \leq \sigma_D(\mathbf{c}), \forall\ \mathbf{c}\}. \tag{17}$$

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

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

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}) = \max\{\mathbf{c}^t\mathbf{R} | A\mathbf{R} \leq \Theta(\mathbf{P})\}. \tag{18}$$

For a fixed $\mathbf{P}$, (18) is a linear program. Using the strong duality of the linear programming, we obtain

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}) = \min\{\mathbf{y}^t\Theta(\mathbf{P}) | A^t\mathbf{y} = \mathbf{c}, 0 \leq \mathbf{y}\}. \tag{19}$$

In general, $\hat{\mathbf{y}}$, the minimizer of (19), is a function of $\mathbf{P}$ and $\mathbf{c}$. We say D_0 satisfies the active extreme points condition if $\hat{\mathbf{y}}$ is only a function of $\mathbf{c}$ for all $\mathbf{c}$. In this case, we have

$$\sigma_{D_0}(\mathbf{c}, \mathbf{P}) = \hat{\mathbf{y}}^t(\mathbf{c})\Theta(\mathbf{P}), \tag{20}$$

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 which is not a function of P. A necessary condition for D_0 to satisfy the active extreme point condition is that each inequality in describing D_0 is either redundant or active for all $\mathbf{P}$.

Theorem 1: If D_0 satisfies the active extreme points condition, then $D_1 = D_2$.

Proof: Since $D_2 \subseteq D_1$ always hold, we need only to show $D_1 \subseteq D_0$. Equivalently, we can show $\sigma_{D_1} \leq \sigma_{D_2}$. The support function of D_1 can be written as

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) = \max_{\mathbf{R} \in D_1} \mathbf{c}^t\mathbf{R} \tag{21}$$

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

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) \leq \min_{A^t\mathbf{y}=\mathbf{c}, 0 \leq \mathbf{y}} \mathbf{y}^t \sum_{i=1}^{q} \lambda_i \Theta(\mathbf{P}_i). \tag{22}$$

Clearly, $\hat{\mathbf{y}}(\mathbf{c})$, the solution of (19), is a feasible point for (22) and we have

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) \leq \hat{\mathbf{y}}^t(\mathbf{c}) \sum_{i=1}^{q} \lambda_i \Theta(\mathbf{P}_i) \tag{23}$$

Using (20), we obtain

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^{q} \lambda_i \sigma_{D_0}(\mathbf{c}, \mathbf{P}_i) \tag{24}$$

Let us assume $\hat{\mathbf{R}}_i$ is the maximizer of (18). In this case, we have

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) \leq \sum_{i=1}^{q} \lambda_i \mathbf{c}^t \hat{\mathbf{R}}_i. \tag{25}$$

Hence, we have

$$\sigma_{D_1}(\mathbf{c}, \mathbf{P}) \leq \mathbf{c}^t \sum_{i=1}^{q} \lambda_i \hat{\mathbf{R}}_i. \tag{26}$$

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

$$\sigma_{D_1}(\mathbf{c},\mathbf{P})\leq\sigma_{D_2}(\mathbf{c},\mathbf{P}). \tag{27}$$

This completes the proof. □

Corollary 1 (Han [16]): If D_0 is a polymatroid then D_1=D_2.

Proof: It is easy to show that D_0 satisfies the active extreme points condition. In fact, for given $\mathbf{c}$, $\hat{\mathbf{y}}$ can be obtained in a greedy fashion which is independent of $\mathbf{P}$. □

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

Theorem 2: The cardinality of the time sharing parameter q in (13) is less than $M+K+1$, where M and K are the dimensions of $\mathbf{P}$ and $\Theta(\mathbf{P})$, respectively. Moreover, if $\Theta(\mathbf{P})$ is a continuous function of $\mathbf{P}$, then $q\leq M+K$.

Proof: Let us define E as

$$E=\left\{\sum_{i=1}^{q}\lambda_i\Theta(\mathbf{P}_i)|\sum_{i=1}^{q}\lambda_i\mathbf{P}_i\leq\mathbf{P},\sum_{i=1}^{q}\lambda_i=1,\lambda_i\geq0\ \forall i\right\}. \tag{28}$$

In fact, E is the collection of all possible bounds for D_1. To prove $q\leq M+K+1$, we define another region D as

$$E_1=\{(\mathbf{P}',\mathbf{S}')|0\leq\mathbf{P}',\mathbf{S}'=\Theta(\mathbf{P}')\}. \tag{29}$$

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

$$\hat{E}=\{\mathbf{S}'|(\mathbf{P}',\mathbf{S}')\in\text{conv } E_1,\mathbf{P}'\leq\mathbf{P}\}. \tag{30}$$

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

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 $D_0=\{\mathbf{R}|\mathbf{R}\leq\Theta(\mathbf{P})\}$. Therefore, both $\mathbf{P}$ and $\Theta(\mathbf{P})$ have dimension M. On the other hand $\Theta(\mathbf{P})$ is a continuous function of $\mathbf{P}$. Now, by applying Theorem 2 we obtain the desired result. □

To upper bound q' in (15), we need some extra definitions and theorems. In fact, in this case D_0 can be viewed as a set valued map and D_2 as another set valued map which is obtained from concavification of D_0. Appendix I summarizes all results regarding concavification of a set valued map. Using these results, we can state the following theorem.

Theorem 3: To characterize boundary points of D_2, it suffices to set $q'\leq N+1$ where N is the dimension of $\mathbf{R}$.

Proof: See . □

Surprising fact about Theorem 3 is that 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, to obtain any point on the boundary of D_1, the cardinality of the time sharing parameter is less than $2M$.

E. Extremal Inequality

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

$$W=\max_{Q_{\mathbf{X}}\leq S}h(\mathbf{X}+\mathbf{Z}_1)-\mu h(\mathbf{X}+\mathbf{Z}_2), \tag{31}$$

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 constrain $Q_{\mathbf{X}}\leq S$, where S is a positive definite matrix. It is shown that for all $\mu\geq1$, 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 [18][19].

In this paper, we consider a special case of (31) where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ have the covariance matrices N_1I and N_2I, respectively, and the constraint is the trace constraint, i.e.,

$$W=\max_{tr\{Q_{\mathbf{X}}\}\leq nP}h(\mathbf{X}+\mathbf{Z}_1)-\mu h(\mathbf{X}+\mathbf{Z}_2). \tag{32}$$

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 (32) has an Gaussian optimal solution for all $0\leq\mu$ with iid components. More precisely, we have

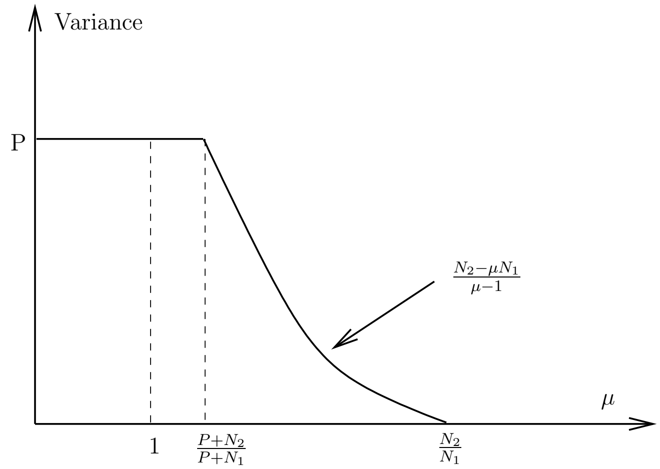

Fig. 2. Optimum variance versus μ.

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

$$W = \frac{n}{2}\log\left[(2\pi e)(P+N_1)\right] - \frac{\mu n}{2}\log\left[(2\pi e)(P+N_2)\right]. \tag{33}$$

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

$$W = \frac{n}{2}\log\left[(2\pi e)\frac{N_2-N_1}{\mu-1}\right] - \frac{\mu n}{2}\log\left[\frac{\mu(2\pi e)(N_2-N_1)}{\mu-1}\right] \tag{34}$$

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

$$W = \frac{n}{2}\log(2\pi e N_1) - \frac{\mu n}{2}\log(2\pi e N_2). \tag{35}$$

Proof: See Appendix II for the proof. □

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 (32) has an Gaussian optimal solution for all $1 \leq \mu$ with iid components. In this case, the optimum variance is 0 and the optimum solution is

$$W = \frac{n}{2}\log(2\pi e N_1) - \frac{\mu n}{2}\log(2\pi e N_2). \tag{36}$$

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

Corollary 4: For $\mu = 1$, the optimization problem (32) has an Gaussian optimal solution with iid components. The optimum solution in this case is

$$W = \begin{cases} \frac{n}{2}\log\left(\frac{P+N_1}{P+N_2}\right), & \text{if } N_1 \leq N_2 \\ \frac{n}{2}\log\left(\frac{N_1}{N_2}\right), & \text{if } N_1 > N_2 \end{cases} \tag{37}$$

We repeatedly use the following optimization problem throughout the paper:

$$f_h = \max_{tr\{Q_{\mathbf{X}}\}\leq nP} h(\mathbf{X}+\mathbf{Z}_1) - \mu h(\sqrt{a}\mathbf{X}+\mathbf{Z}_2), \tag{38}$$

where $N_1 \leq N_2/a$. Using the identity $h(A\mathbf{X}) = \log(|A|) + h(\mathbf{X})$, (38) can be written as

$$f_h = \frac{n}{2}\log a + \max_{tr\{Q_{\mathbf{X}}\}\leq nP} h(\mathbf{X}+\mathbf{Z}_1) - \mu h(\mathbf{X}+\frac{\mathbf{Z}_2}{\sqrt{a}}), \tag{39}$$

Now, by applying Lemma 1, we obtain

$$f_h(P,N_1,N_2,a,\mu) = \begin{cases} \frac{1}{2}\log\left[(2\pi e)(P+N_1)\right] - \frac{\mu}{2}\log\left[(2\pi e)(aP+N_2)\right] & \text{if } 0 \leq \mu \leq \frac{P+N_2/a}{P+N_1} \\ \frac{1}{2}\log\left[(2\pi e)\frac{N_2/a-N_1}{\mu-1}\right] - \frac{\mu}{2}\log\left[\frac{a\mu(2\pi e)(N_2/a-N_1)}{\mu-1}\right] & \text{if } \frac{P+N_2/a}{P+N_1} < \mu \leq \frac{N_2}{aN_1} \\ \frac{1}{2}\log(2\pi e N_1) - \frac{\mu}{2}\log(2\pi e N_2) & \text{if } \frac{N_2}{aN_1} < \mu \end{cases} \tag{40}$$

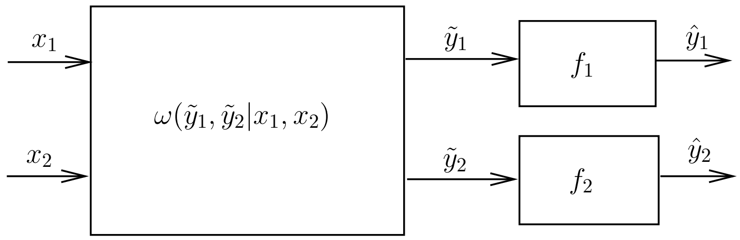

Fig. 3. An admissible channel. f_1 and f_2 are two deterministic functions of their inputs.

III. Admissible Channels

In this section, we aim at building ICs whose capacity regions contain the capacity region of the two-user Gaussian IC, i.e., $\mathscr{C}$. Since ultimately we use them to outer bound $\mathscr{C}$, these ICs need to possess some properties regarding characterization of the capacity region. 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 $\mathscr{C}$ and output letters $\tilde{y}_1$ and $\tilde{y}_2$ for Users 1 and 2, respectively. The capacity region of this channel, say $\mathscr{C}'$, contains $\mathscr{C}$ if

$$I(x_1^n; y_1^n) \leq I(x_1^n; \tilde{y}_1^n), \tag{41}$$

$$I(x_2^n; y_2^n) \leq I(x_2^n; \tilde{y}_2^n), \tag{42}$$

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

One way to satisfy (41) and (42) is to provide some extra information to the one or both receivers. This scheme is called Genie aided outer bounding scheme. 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 exist an outer bound due to Sato and Costa, see [20] and [11], that can be used to outer bound the original channel. In all previous works, the genie's task was to reveal some information about the interfering signal to the receiver(s). In [21], Etkin, Tse, and Wang changed the direction. Their genie 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.

The way that we rely on to satisfy (41) and (42) is to find 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 (see Figure 3)

$$I(x_1^n; y_1^n) \leq I(x_1^n; \hat{y}_1^n), \tag{43}$$

$$I(x_2^n; y_2^n) \leq I(x_2^n; \hat{y}_2^n). \tag{44}$$

for all $p(x_1^n)p(x_2^n)$ and for all $n \in \mathbb{N}$. By using the data processing inequality, it is easy to show that (43) and (44) imply (41) and (42), respectively.

Definition 3 (Admissible Channel): An IC $\mathscr{C}'$ 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 (43) and (44) hold for all $p(x_1^n)p(x_2^n)$ and for all $n \in \mathbb{N}$. $\mathscr{E}$ denotes the collection of all admissible channels.

Clearly, Genie aided channels are among admissible channels. To see this, let us assume a genie provides s_1 and s_2 as side information for User 1 and 2, respectively. In this case, $\tilde{y}_i = (y_i, s_i)$ for $i \in \{1, 2\}$. By choosing $f_i(y_i, s_i) = y_i$, we observe that $\hat{y}_i = y_i$ and hence (43) and (44) 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 that possess two important properties. First, their capacity regions are close to $\mathscr{C}$. Second, either their exact capacity regions are computable or there exist good outer bounds on their capacity regions. Let $\mathscr{F}$ denote the subset of $\mathscr{E}$ containing all appropriate admissible channels. Clearly, we have

$$\mathscr{C} \subseteq \bigcap_{\mathscr{F}} \mathscr{C}'. \tag{45}$$

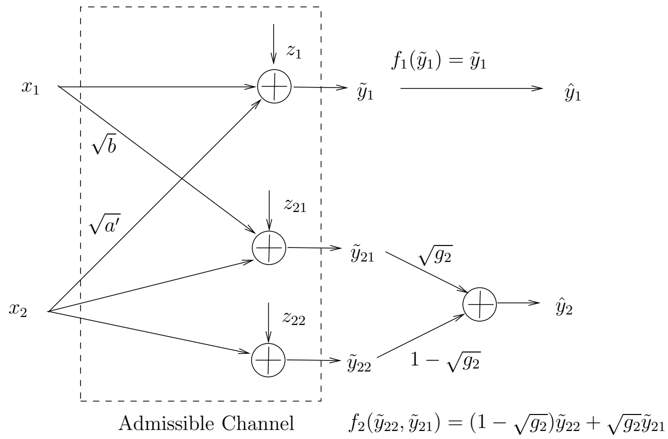

Fig. 4. Class A1 admissible channels.

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

$$\sigma_{\mathscr{C}}(c_1, c_2) = \max_{(R_1,R_2)\in\mathscr{C}} c_1R_1 + c_2R_2 \tag{46}$$

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 $\mathscr{C} \subseteq \mathscr{C}'$, we have

$$\sigma_{\mathscr{C}}(c_1, c_2) \leq \sigma_{\mathscr{C}'}(c_1, c_2). \tag{47}$$

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

$$\sigma_{\mathscr{C}}(c_1, c_2) \leq \min_{\mathscr{C}'\in\mathscr{F}} \sigma_{\mathscr{C}'}(c_1, c_2), \tag{48}$$

which can be written as

$$\sigma_{\mathscr{C}}(c_1, c_2) \leq \min_{\mathscr{C}'\in\mathscr{F}} \max_{(R_1,R_2)\in\mathscr{C}'} c_1R_1 + c_2R_2. \tag{49}$$

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

$$\sigma_{\mathscr{C}}(\mu, 1) = \max_{(R_1,R_2)\in\mathscr{C}} \mu R_1 + R_2, \tag{50}$$

$$\sigma_{\mathscr{C}}(1, \mu) = \max_{(R_1,R_2)\in\mathscr{C}} R_1 + \mu R_2, \tag{51}$$

where $1 \leq \mu$. It is easy to show that solutions of (50) and (51) 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 $\sigma_{\mathscr{C}'}(\mu, 1)$ and $\sigma_{\mathscr{C}'}(1, \mu)$.

A. Classes of Admissible Channels

1) Class A1: This class is specially designed to upper bound $\sigma_{\mathscr{C}}(\mu, 1)$. Therefore, we need to find a tight upper bound for $\sigma_{\mathscr{C}'}(\mu, 1)$. 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

$$\begin{cases} \tilde{y}_1 = & x_1 + \sqrt{a}x_2 + z_1, \\ \tilde{y}_{21} = & x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} = & x_2 + z_{22}, \end{cases} \tag{52}$$

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_{11} and z_{12} are additive Gaussian noises 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.

To investigate admissibility conditions (43) and (44), we need to introduce two deterministic functions. Let us consider two linear functions f_1 and f_2 as follows (see Figure 4)

$$f_1(\tilde{y}_1^n) = \tilde{y}_1^n, \tag{53}$$

$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1-\sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n, \tag{54}$$

where $0 \le g_2$. Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a} x_2^n + z_1^n, \tag{55}$$

$$\hat{y}_2^n = \sqrt{b' g_2} x_1^n + x_2^n + (1-\sqrt{g_2}) z_{22}^n + \sqrt{g_2} z_{21}^n. \tag{56}$$

Hence, this channel is admissible if the channel's parameters satisfy

$$\begin{aligned} b' g_2 &= b, \\ (1-\sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \tag{57}$$

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

$$\begin{aligned} b' &\le N_{21}, \\ a N_{22} &\le 1. \end{aligned} \tag{58}$$

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 $\sigma_{\mathscr{C}'}(\mu, 1)$. In the following lemma, we obtain the required upper bound.

Lemma 3: For the channels modeled by (52) and satisfying (58), we have

$$\begin{aligned} \sigma_{\mathscr{C}'}(\mu,1) \le \min_{\substack{\mu_1,\mu_2 \ge 0 \\ \mu_1+\mu_2 = \mu}} & \frac{\mu_1}{2}\log\left[2\pi e(P_1+aP_2+1)\right] - \frac{\mu_2}{2}\log(2\pi e) + \frac{1}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2+N_{22}}\right) \\ & + \mu_2 f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right) + f_h(P_2, N_{22}, 1, a, \mu_1). \end{aligned} \tag{59}$$

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

$$\begin{aligned} n(\mu R_1 + R_2) \le & \mu I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ \le & \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ \overset{(a)}{=} & \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ = & \mu_1 I(x_1^n; \tilde{y}_1^n) + \mu_2 I(x_1^n; \tilde{y}_1^n | x_2^n) + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n,) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\ = & \mu_1 h(\tilde{y}_1^n) - \mu_1 h(\tilde{y}_1^n | x_1^n) + \mu_2 h(\tilde{y}_1^n | x_2^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) \\ & + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \end{aligned} \tag{60}$$

$$\begin{aligned} = & \left[\mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n)\right] + \left[\mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)\right] \\ & + \left[h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n)\right] + \left[h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_1^n | x_1^n)\right] + n\epsilon_n, \end{aligned} \tag{61}$$

where (a) follows from the fact x_1^n and x_1^n are independent. Now, we separately upper bound the terms within each bracket in (61).

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

$$\begin{aligned} \mu_1 h(\tilde{y}_1^n) - \mu_2 h(\tilde{y}_1^n | x_1^n, x_2^n) &= \mu_1 h(x_1^n + \sqrt{a} x_2^n + z_1^n) - \mu_2 h(z_1^n) \\ &\le \frac{\mu_1 n}{2}\log\left[2\pi e(P_1 + aP_2 + 1)\right] - \frac{\mu_2 n}{2}\log(2\pi e). \end{aligned} \tag{62}$$

Since $b' \le N_{21}$, we can make use of Lemma 4 to upper bound the second bracket. In this case, we have

$$\begin{aligned} \mu_2 h(\tilde{y}_1^n | x_2^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) &= \mu_2\left(h(x_1^n + z_1^n) - \frac{1}{\mu_2} h(\sqrt{b'} x_1^n + z_{21}^n)\right) \\ &\le \mu_2 n f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right), \end{aligned} \tag{63}$$

where f_h is defined in (40).

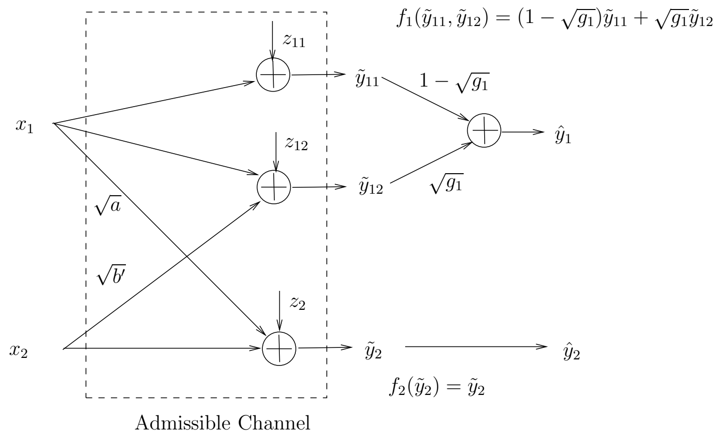

Fig. 5. Class A2 admissible channels.

We upper bound the terms within the third bracket as follows:

$$\begin{aligned}
h(\tilde{y}_{21}^n|\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n|x_2^n) &\overset{(a)}{\leq} \sum_{i=1}^{n} h(\tilde{y}_{21}[i]|\tilde{y}_{22}[i]) - h(z_{22}^n)\\
&\overset{(b)}{\leq} \sum_{i=1}^{n} \frac{1}{2}\log\left[2\pi e\left(N_{21} + b'P_1[i] + \frac{P_2[i]N_{22}}{P_2[i]+N_{22}}\right)\right] - \frac{n}{2}\log\left(2\pi e N_{22}\right)\\
&\overset{(c)}{\leq} \frac{n}{2}\log\left[2\pi e\left(N_{21} + \frac{1}{n}\sum_{i=1}^{n} b'P_1[i] + \frac{\frac{1}{n}\sum_{i=1}^{n} P_2[i]N_{22}}{\frac{1}{n}\sum_{i=1}^{n} P_2[i]+N_{22}}\right)\right] - \frac{n}{2}\log\left(2\pi e N_{22}\right)\\
&\leq \frac{n}{2}\log\left[2\pi e\left(N_{21} + b'P_1 + \frac{P_2N_{22}}{P_2+N_{22}}\right)\right] - \frac{n}{2}\log\left(2\pi e N_{22}\right)\\
&\leq \frac{n}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2+N_{22}}\right)
\end{aligned} \tag{64}$$

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 optimized conditional entropy for given covariance matrix, and (c) follows form Jenson's inequality.

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

$$\begin{aligned}
h(\tilde{y}_{22}^n) - \mu_1 h(\tilde{y}_1^n|x_1^n) &= h(x_2^n + z_{22}^n) - \mu_1 h(\sqrt{a}x_2^n + z_1^n)\\
&\leq nf_h(P_2, N_{22}, 1, a, \mu_1).
\end{aligned} \tag{65}$$

Adding all inequalities, we obtain

$$\begin{aligned}
\mu R_1 + R_2 \leq &\frac{\mu_1}{2}\log\left[2\pi e(P_1 + aP_2 + 1)\right] - \frac{\mu_2}{2}\log(2\pi e) + \frac{1}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2+N_{22}}\right)\\
&+\mu_2 f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right) + f_h(P_2, N_{22}, 1, a, \mu_1),
\end{aligned} \tag{66}$$

where the fact that $\epsilon_n \to 0$ as $n \to \infty$ is used to eliminate ϵ_n form the right hand side of the inequality. Now, by taking the minimum of the right hand sid of (66) over all μ_1 and μ_2, we obtain the desired result. This completes the proof. □

2) Class A2: This class is essentially the complement of Class A1 in a sense that we use it to upper bound $\sigma_{\mathscr{C}}(1, \mu)$. 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 (see Figure 5). The channel model can be written as

$$\begin{cases}
\tilde{y}_{11} = & x_1 + z_{11},\\
\tilde{y}_{12} = & x_1 + \sqrt{a'}x_2 + z_{12},\\
\tilde{y}_2 = & x_2 + \sqrt{b}x_1 + z_2,
\end{cases} \tag{67}$$

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_{22} and z_{21} are additive Gaussian noises with variances N_{22} and N_{21}, respectively, and transmitter 1 and 2 are subject to the average power constraints P_1 and P_2, respectively.

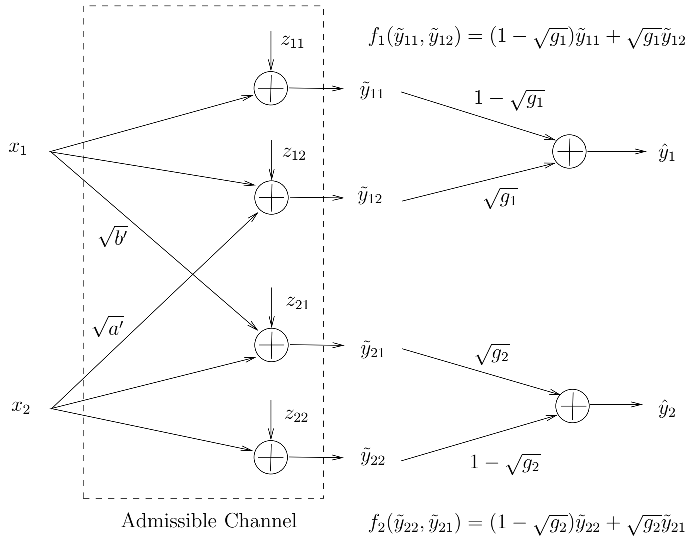

Fig. 6. Class B admissible channels.

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

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1-\sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n, \tag{68}$$

$$f_2(\tilde{y}_2^n) = \tilde{y}_2^n. \tag{69}$$

Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a' g_1}x_2^n + (1-\sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n, \tag{70}$$

$$\hat{y}_2^n = \sqrt{b}x_1^n + x_2^n + z_2^n. \tag{71}$$

We deduce that the channel modeled by (67) is admissible if the channel's parameters satisfy

$$\begin{aligned} a' g_1 &= a, \\ (1-\sqrt{g_1})^2 N_{11} + g_1 N_{12} &= 1. \end{aligned} \tag{72}$$

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

$$\begin{aligned} a' &\leq N_{12}, \\ bN_{11} &\leq 1. \end{aligned} \tag{73}$$

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

Lemma 4: For the channels modeled by (67) and satisfying (73), we have

$$\begin{aligned} \sigma_{\mathscr{C}'}(\mu, 1) \leq \min_{\substack{\mu_1, \mu_2 \geq 0 \\ \mu_1 + \mu_2 = \mu}} & \frac{\mu_1}{2}\log\left[2\pi e(P_1 + aP_2 + 1)\right] - \frac{\mu_2}{2}\log(2\pi e) + \frac{1}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}}\right) \\ & + \mu_2 f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right) + f_h(P_2, N_{22}, 1, a, \mu_1). \end{aligned} \tag{74}$$

Proof: The proof is similar to the proof of Lemma 3 and we omit it here. □

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)

$$\begin{cases} \tilde{y}_{11} = & x_1 + z_{11}, \\ \tilde{y}_{12} = & x_1 + \sqrt{a'}x_2 + z_{12}, \\ \tilde{y}_{21} = & x_2 + \sqrt{b'}x_1 + z_{21}, \\ \tilde{y}_{22} = & x_2 + z_{22}, \end{cases} \tag{75}$$

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 $\sigma_{\mathscr{C}}(\mu, 1)$ and $\sigma_{\mathscr{C}}(1, \mu)$.

Here, we investigate admissibility of this channel and, as a result, the required conditions that must be imposed on the channel's 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)

$$f_1(\tilde{y}_{11}^n, \tilde{y}_{12}^n) = (1-\sqrt{g_1})\tilde{y}_{11}^n + \sqrt{g_1}\tilde{y}_{12}^n \tag{76}$$
$$f_2(\tilde{y}_{22}^n, \tilde{y}_{21}^n) = (1-\sqrt{g_2})\tilde{y}_{22}^n + \sqrt{g_2}\tilde{y}_{21}^n. \tag{77}$$

Therefore, we have

$$\hat{y}_1^n = x_1^n + \sqrt{a' g_1}x_2^n + (1-\sqrt{g_1})z_{11}^n + \sqrt{g_1}z_{12}^n \tag{78}$$
$$\hat{y}_2^n = \sqrt{b' g_2}x_1^n + x_2^n + (1-\sqrt{g_2})z_{22}^n + \sqrt{g_2}z_{21}^n. \tag{79}$$

To satisfy (43) and (44), it suffices to have

$$\begin{aligned} a' g_1 &= a, \\ b' g_2 &= b, \\ (1-\sqrt{g_1})^2 N_{11} + g_1 N_{12} &= 1, \\ (1-\sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \tag{80}$$

Hence, a channel modeled in (75) is admissible if there are two nonnegative numbers g_1 and g_2 such that the set of equalities in (80) holds. We further add the following two constraints to the equality conditions in (80:

$$\begin{aligned} b' N_{11} &\leq N_{21}, \\ a' N_{22} &\leq N_{12}. \end{aligned} \tag{81}$$

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

Lemma 5: For the channels modeled by (75) and satisfying (81), we have

$$\begin{aligned} \sigma_{\mathscr{C}'}(\mu, 1) \leq & \mu\gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}}\right) + \gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}}\right) \\ & + f_h(P_2, N_{22}, N_{12}, a', \mu) + \frac{\mu}{2}\log((2\pi e)(a'P_2 + N_{12})) - \frac{1}{2}\log((2\pi e)(P_2 + N_{22})), \end{aligned} \tag{82}$$

$$\begin{aligned} \sigma_{\mathscr{C}'}(1, \mu) \leq & \mu\gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}}\right) + \gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}}\right) \\ & + f_h(P_1, N_{11}, N_{21}, b', \mu) + \frac{\mu}{2}\log((2\pi e)(b'P_1 + N_{21})) - \frac{1}{2}\log((2\pi e)(P_1 + N_{11})). \end{aligned} \tag{83}$$

Proof: We only upper bound $\sigma_{\mathscr{C}'}(\mu, 1)$ and an upper bound for $\sigma_{\mathscr{C}'}(1, \mu)$ 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

$$\begin{aligned} n(\mu R_1 + R_2) \leq & \mu I(x_1^n; \tilde{y}_{11}^n, \tilde{y}_{12}^n) + I(x_2^n; \tilde{y}_{22}^n, \tilde{y}_{21}^n) + n\epsilon_n \\ = & \mu I(x_1^n; \tilde{y}_{12}^n | \tilde{y}_{11}^n) + \mu I(x_1^n; \tilde{y}_{11}^n) \\ & + I(x_2^n; \tilde{y}_{21}^n | \tilde{y}_{22}^n,) + I(x_2^n; \tilde{y}_{22}^n) + n\epsilon_n \\ = & \mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n) + \mu h(\tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n) \\ & + h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n) + h(\tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n) + n\epsilon_n \\ = & \left[\mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n)\right] + \left[h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) - h(\tilde{y}_{22}^n | x_2^n)\right] \\ & + \left[\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n | x_2^n, \tilde{y}_{22}^n)\right] + \left[h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n | x_1^n, \tilde{y}_{11}^n)\right] + n\epsilon_n \end{aligned} \tag{84}$$

Now, we upper bound different the terms within each bracket in (84) separately.

For the first bracket, we have

$$\begin{aligned} \mu h(\tilde{y}_{12}^n | \tilde{y}_{11}^n) - \mu h(\tilde{y}_{11}^n | x_1^n) & \stackrel{(a)}{\leq} \mu \sum_{i=1}^n h(\tilde{y}_{12}[i] | \tilde{y}_{11}[i]) - \frac{\mu n}{2}\log(2\pi e N_{11}) \\ & \stackrel{(b)}{\leq} \mu \sum_{i=1}^n \frac{1}{2}\log\left[2\pi e\left(N_{12} + a'P_2[i] + \frac{P_1[i]N_{11}}{P_1[i] + N_{11}}\right)\right] - \frac{\mu n}{2}\log(2\pi e N_{11}) \\ & \stackrel{(c)}{\leq} \frac{\mu n}{2}\log\left[2\pi e\left(N_{12} + \frac{1}{n}\sum_{i=1}^n a'P_2[i] + \frac{\frac{1}{n}\sum_{i=1}^n P_1[i]N_{11}}{\frac{1}{n}\sum_{i=1}^n P_1[i] + N_{11}}\right)\right] - \frac{\mu n}{2}\log(2\pi e N_{11}) \\ & \leq \frac{\mu n}{2}\log\left[2\pi e\left(N_{12} + a'P_2 + \frac{P_1 N_{11}}{P_1 + N_{11}}\right)\right] - \frac{\mu n}{2}\log(2\pi e N_{11}) \\ & = \frac{\mu n}{2}\log\left(\frac{N_{12}}{N_{11}} + \frac{a'P_2}{N_{11}} + \frac{P_1}{P_1 + N_{11}}\right) \end{aligned} \tag{85}$$

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 optimized conditional entropy for given covariance matrix, and (c) follows form Jenson's inequality.

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

$$h(y_{21}^n|y_{22}^n) - h(y_{22}^n|x_2^n) \leq \frac{n}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2+N_{22}}\right). \tag{86}$$

By making use of Lemma 4 and using the fact that $N_{11} \leq N_{21}/b'$, the terms within the third bracket can be upper bounded as

$$\begin{aligned}\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n|x_2^n,\tilde{y}_{22}^n) &= \mu\left(h(x_1^n+z_{11}^n) - \frac{1}{\mu}h(\sqrt{b'}x_1^n+z_{21}^n)\right)\\ &\leq \mu n f_h\left(P_1,N_{11},N_{21},b',\frac{1}{\mu}\right).\end{aligned} \tag{87}$$

Since $1 \leq \mu$, we obtain

$$\mu h(\tilde{y}_{11}^n) - h(\tilde{y}_{21}^n|x_2^n,\tilde{y}_{22}^n) \leq \frac{\mu n}{2}\log((2\pi n)(P_1+N_{11})) - \frac{n}{2}\log((2\pi e)(b'P_1+N_{21})). \tag{88}$$

For the last bracket, again we make use of Lemma 4 to obtain

$$\begin{aligned}h(\tilde{y}_{22}^n) - \mu h(\tilde{y}_{12}^n|x_1^n,\tilde{y}_{11}^n) &= h(x_2^n+z_{22}^n) - \mu h(\sqrt{a'}x_2^n+z_{12}^n)\\ &\leq n f_h(P_2,N_{22},N_{12},a',\mu).\end{aligned} \tag{89}$$

Adding all inequalities, we obtain

$$\begin{aligned}\mu R_1 + R_2 \leq &\frac{\mu}{2}\log\left(\frac{N_{12}}{N_{11}} + \frac{a'P_2}{N_{11}} + \frac{P_1}{P_1+N_{11}}\right) + \frac{1}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2+N_{22}}\right)\\ &+\frac{\mu}{2}\log((2\pi e)(P_1+N_{11})) - \frac{1}{2}\log((2\pi e)(b'P_1+N_{21})) + f_h(P_2,N_{22},N_{12},a',\mu),\end{aligned} \tag{90}$$

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

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

This completes the proof. □

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 (82) and (83) become, respectively,

$$\mu R_1 + R_2 \leq \mu\gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2+N_{12}}\right) + \gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1+N_{21}}\right) \tag{91}$$

and

$$R_1 + \mu R_2 \leq \gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2+N_{12}}\right) + \mu\gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1+N_{21}}\right). \tag{92}$$

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

$$R_1 = \gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2+N_{12}}\right) \tag{93}$$

and

$$R_2 = \gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1+N_{21}}\right) \tag{94}$$

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, Tse, and Wang in [21]. 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

$$\mathscr{C}'_{\text{sum}} = \gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2+N_{12}}\right) + \gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1+N_{21}}\right). \tag{95}$$

Proof: Simply by substituting $\mu = 1$ in (91), we obtain the desired result. □

4) Class C: Class C is designed to upper bound $\sigma_{\mathscr{C}}(\mu, 1)$ 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 new method.

For channels in Class C, we use the same model that is given in (52). Therefore, similar to channels in Class A1, this channel is admissible if the channel's parameters satisfy

$$\begin{aligned} b' g_2 &= b, \\ (1-\sqrt{g_2})^2 N_{22} + g_2 N_{21} &= 1. \end{aligned} \tag{96}$$

Here, we change the constraints in (58) to new constraints

$$\begin{aligned} b' &\geq N_{21}, \\ aN_{22} &\leq 1. \end{aligned} \tag{97}$$

In fact, the second condition is unchanged and only the inequality sign for the first one is reversed. By this simple change of constraints, we see that the second receiver after decoding its own signal has a cleaner 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

$$\begin{aligned} \sigma_{\mathscr{C}'}(\mu, 1) \leq & \frac{\mu-1}{2} \log\left(2\pi e(P_1 + aP_2 + 1)\right) + \frac{1}{2}\log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b' P_1 + N_{21}\right)\right) \\ & - \frac{1}{2}\log(2\pi e N_{21}) - \frac{1}{2}\log(2\pi e N_{22}) + f_h(P_2, N_{22}, 1, a, \mu - 1) \end{aligned} \tag{98}$$

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

$$R_1 \leq \frac{1}{n} I(x_1^n; \tilde{y}_1^n) \tag{99}$$

$$R_1 \leq \frac{1}{n} I(x_1^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_2^n) \tag{100}$$

$$R_2 \leq \frac{1}{n} I(x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n | x_1^n) \tag{101}$$

$$R_1 + R_2 \leq \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n) \tag{102}$$

$$\mu R_1 + R_2 \leq \frac{\mu-1}{n} I(x_1^n; \tilde{y}_1^n) + \frac{1}{n} I(x_1^n, x_2^n; \tilde{y}_{21}^n, \tilde{y}_{22}^n) \tag{103}$$

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

$$\frac{\mu-1}{n} h(\tilde{y}_1^n) \leq \frac{\mu-1}{2} \log\left(2\pi e(P_1 + aP_2 + 1)\right) \tag{104}$$

$$\frac{1}{n} h(\tilde{y}_{21}^n | \tilde{y}_{22}^n) \leq \frac{1}{2} \log\left(2\pi e\left(\frac{P_2 N_{22}}{P_2 + N_{22}} + b' P_1 + N_{21}\right)\right) \tag{105}$$

$$\frac{1}{n} h(\tilde{y}_{21}^n, \tilde{y}_{21}^n | x_1^n, x_2^n) = \frac{1}{2}\log(2\pi e N_{21}) + \frac{1}{2}\log(2\pi e N_{22}). \tag{106}$$

$$\begin{aligned} \frac{1}{n} h(\tilde{y}_{22}^n) - \frac{\mu-1}{n} h(\tilde{y}_1^n | x_1^n) &= \frac{1}{n} h(x_2^n + z_{22}^n) - \frac{\mu-1}{n} h(\sqrt{a} x_2^n + z_1) & (107) \\ &\leq f_h(P_2, N_{22}, 1, a, \mu - 1) & (108) \end{aligned}$$

□

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 that previous outer bounds. Finally, we show that using time-sharing parameter and concavification 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:

$$
\begin{aligned}
W = & \min \gamma\left(\frac{P_1}{N_{11}}+\frac{P_1}{a'P_2+N_{12}}\right)+\gamma\left(\frac{P_2}{N_{22}}+\frac{P_2}{b'P_1+N_{21}}\right) \\
& \text{subject to:} \\
& \qquad a'g_1 = a \\
& \qquad b'g_2 = b \\
& \qquad b'N_{11} \leq N_{21} \\
& \qquad a'N_{22} \leq N_{12} \\
& \qquad (1-\sqrt{g_1})^2 N_{11} + g_1 N_{12} = 1 \\
& \qquad (1-\sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1 \\
& \qquad 0 \leq [a', b', g_1, g_2, N_{11}, N_{12}, N_{22}, N_{21}]
\end{aligned} \tag{109}
$$

The objective function in (109) is the sum capacity of Class B channels obtained in Theorem 4. The constraints are the combination of (80) and (81) 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 $\mathscr{C}_{sum} \leq W$. By changing the variables as $S_1 = g_1 N_{12}$ and $S_2 = g_2 N_{21}$, we obtain

$$
\begin{aligned}
W = & \min \gamma\left(\frac{(1-\sqrt{g_1})^2 P_1}{1-S_1}+\frac{g_1 P_1}{aP_2+S_1}\right)+\gamma\left(\frac{(1-\sqrt{g_2})^2 P_2}{1-S_2}+\frac{g_2 P_2}{bP_1+S_2}\right) \\
& \text{subject to:} \\
& \qquad \frac{b(1-S_1)}{(1-\sqrt{g_1})^2} \leq S_2 < 1, \\
& \qquad \frac{a(1-S_2)}{(1-\sqrt{g_2})^2} \leq S_1 < 1, \\
& \qquad 0 < [g_1, g_2],
\end{aligned} \tag{110}
$$

We first minimize the objective in (110) with respect to g_1 and g_2. In this case, 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

$$
\begin{aligned}
W_1 = & \min \gamma\left(\frac{(1-\sqrt{g_1})^2 P_1}{1-S_1}+\frac{g_1 P_1}{aP_2+S_1}\right) \\
& \text{subject to:} \\
& \qquad \frac{b(1-S_1)}{S_2} \leq (1-\sqrt{g_1})^2, \\
& \qquad 0 < g_1,
\end{aligned} \tag{111}
$$

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

$$
W_1 = \begin{cases} \gamma\left(\frac{P_1}{1+aP_2}\right) & \text{if } \sqrt{b}(1+aP_2) \leq \sqrt{S_2(1-S_1)} \\ \gamma\left(\frac{bP_1}{S_2}+\frac{(1-\sqrt{b(1-S_1)/S_2})^2 P_1}{aP_2+S_1}\right) & \text{Otherwise} \end{cases} \tag{112}
$$

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

$$
\begin{aligned}
W_2 = & \min \gamma\left(\frac{(1-\sqrt{g_2})^2 P_2}{1-S_2}+\frac{g_2 P_2}{bP_1+S_2}\right) \\
& \text{subject to:} \\
& \qquad \frac{a(1-S_2)}{S_1} \leq (1-\sqrt{g_2})^2, \\
& \qquad 0 < g_2,
\end{aligned} \tag{113}
$$

The solution to the above optimization problem is

$$
W_2 = \begin{cases} \gamma\left(\frac{P_2}{1+bP_1}\right) & \text{if } \sqrt{a}(1+bP_1) \leq \sqrt{S_1(1-S_2)} \\ \gamma\left(\frac{aP_2}{S_1}+\frac{(1-\sqrt{a(1-S_2)/S_1})^2 P_2}{bP_1+S_2}\right) & \text{Otherwise} \end{cases} \tag{114}
$$

Combining (112) and (114), we obtain

$$\begin{aligned} W =& \min W_1 + W_2 \\ & \text{subject to:} \\ & \quad 0 < S_1 < 1, \\ & \quad 0 < S_2 < 1, \end{aligned} \tag{115}$$

From (112) and (114), 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

$$W = \gamma\left(\frac{P_1}{aP_2 + 1}\right) + \gamma\left(\frac{P_2}{bP_1 + 1}\right), \tag{116}$$

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

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

$$\mathscr{C}_{sum} = \gamma\left(\frac{P_1}{aP_2 + 1}\right) + \gamma\left(\frac{P_2}{bP_1 + 1}\right), \tag{117}$$

for all channel's parameters satisfying

$$\sqrt{a}P_2 + \sqrt{b}P_1 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}}. \tag{118}$$

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

$$D = \left\{(P_1, P_2) | P_1 \leq \frac{\sqrt{S_1(1 - S_2)}}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{\sqrt{S_2(1 - S_1)}}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1, 0 < S_2 < 1\right\}, \tag{119}$$

$$D' = \left\{(P_1, P_2) | \sqrt{b}P_1 + \sqrt{a}P_2 \leq \frac{1 - \sqrt{a} - \sqrt{b}}{\sqrt{ab}}\right\}. \tag{120}$$

To show $D' \subseteq D$, we set $S_1 = 1 - S_2$ in (119) to get

$$\left\{(P_1, P_2) | P_1 \leq \frac{S_1}{b\sqrt{a}} - \frac{1}{b}, P_2 \leq \frac{1 - S_1}{a\sqrt{b}} - \frac{1}{a}, 0 < S_1 < 1\right\} \subseteq D. \tag{121}$$

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

$$J = \max_{(P_1, P_2) \in D} \sqrt{b}P_1 + \sqrt{a}P_2, \tag{122}$$

which can be written as

$$J = \max_{(S_1, S_2) \in (0,1)^2} \frac{\sqrt{S_1(1 - S_2)} + \sqrt{S_2(1 - S_1)}}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \tag{123}$$

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

$$J = \frac{1}{\sqrt{ab}} - \frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}. \tag{124}$$

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

As an example, let us consider the symmetric Gaussian IC. In this case, the constraint in (118) becomes

$$P \leq \frac{1 - 2\sqrt{a}}{2a\sqrt{a}}. \tag{125}$$

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 (109) and the lower bound is sketched for a fixed P and all $0 \leq a \leq 1$. We observe that up to some certain a, the upper bound coincides with the lower bound.

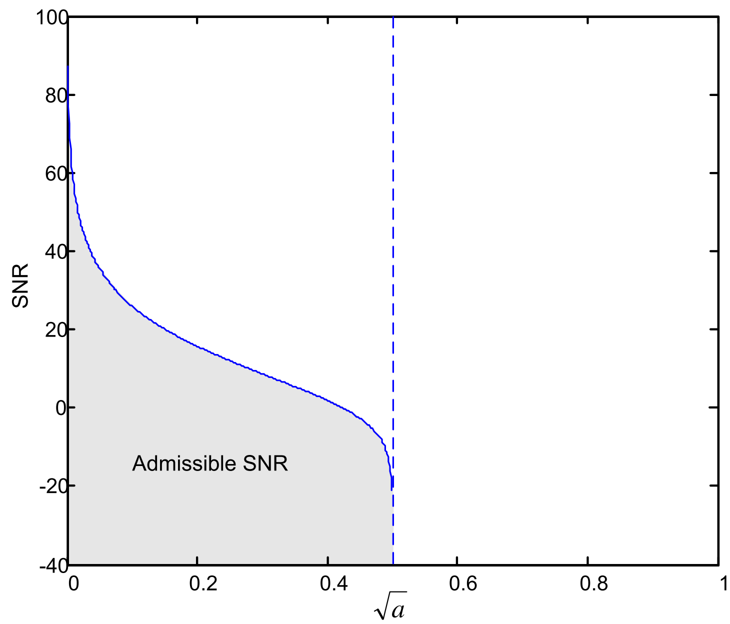

Fig. 7. Admissible region for optimality of treating interference as Gaussian noise.

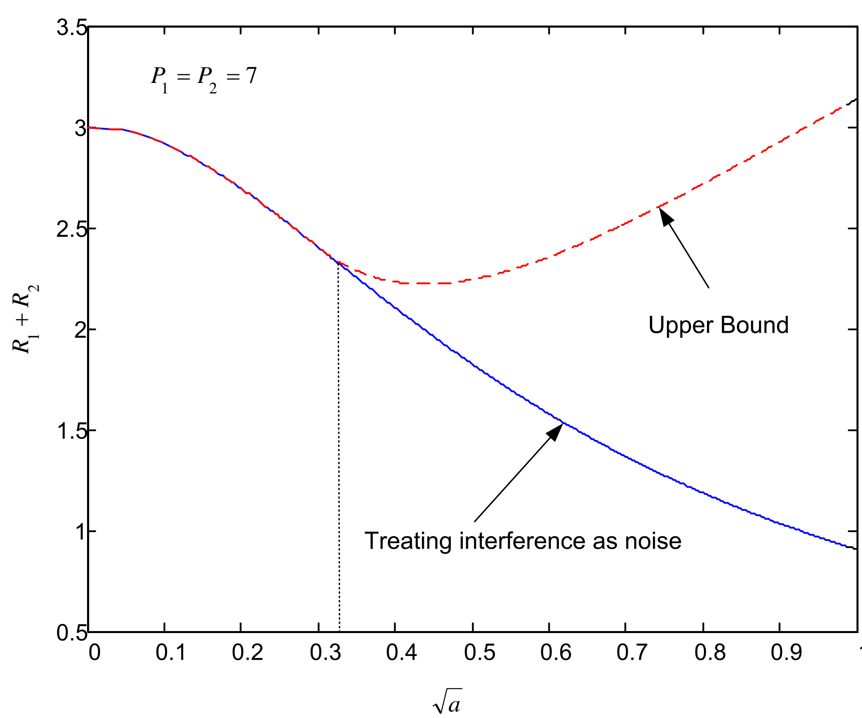

Fig. 8. The upper bound obtained by solving (109). The lower bound is obtained by using the simple scheme of considering the interference as Gaussian noise.

B. New Outer Bound

For the weak Gaussian IC, there are two outer bounds that are tighter than other bounds. The firs 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 exist 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

$$R_1 \leq \gamma\left(\frac{(1-\beta)P'}{\beta P' + 1/a}\right) \tag{126}$$

$$R_2 \leq \gamma(\beta P') \tag{127}$$

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

$$R_1 \leq \gamma(\alpha P'') \tag{128}$$

$$R_2 \leq \gamma\left(\frac{(1-\alpha)P''}{\alpha P'' + 1/b}\right) \tag{129}$$

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

The second outer bound, due to Etkin, Tse, and Wang [21], 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

$$R_1 \leq \gamma(P_1) \tag{130}$$

$$R_2 \leq \gamma(P_2) \tag{131}$$

$$R_1 + R_2 \leq \gamma(P_1) + \gamma\left(\frac{P_2}{1+bP_1}\right) \tag{132}$$

$$R_1 + R_2 \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1+aP_2}\right) \tag{133}$$

$$R_1 + R_2 \leq \gamma\left(aP_2 + \frac{P_1}{1+bP_1}\right) + \gamma\left(bP_1 + \frac{P_2}{1+aP_2}\right) \tag{134}$$

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma\left(bP_1 + \frac{P_2}{1+aP_2}\right) + 0.5\log\left(\frac{1+P_1}{1+bP_1}\right) \tag{135}$$

$$R_1 + 2R_2 \leq \gamma(bP_1 + P_2) + \gamma\left(aP_2 + \frac{P_1}{1+bP_1}\right) + 0.5\log\left(\frac{1+P_2}{1+aP_2}\right). \tag{136}$$

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 $\mathscr{C}$ it suffices to calculate $\sigma_{\mathscr{C}}(\mu, 1)$ and $\sigma_{\mathscr{C}}(1, \mu)$ for all $1 \leq \mu$. To this end, we make use of channels in A1 and B Classes and channels in A2 and B to obtain upper bounds for $\sigma_{\mathscr{C}}(\mu, 1)$ and $\sigma_{\mathscr{C}}(1, \mu)$, respectively.

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

$$\begin{aligned} W_1(\mu) = \min \ & \frac{\mu_1}{2}\log\left[2\pi e(P_1 + aP_2 + 1)\right] - \frac{\mu_2}{2}\log(2\pi e) + \frac{1}{2}\log\left(\frac{N_{21}}{N_{22}} + \frac{b'P_1}{N_{22}} + \frac{P_2}{P_2 + N_{22}}\right) \\ & + \mu_2 f_h\left(P_1, 1, N_{21}, b', \frac{1}{\mu_2}\right) + f_h(P_2, N_{22}, 1, a, \mu_1) \end{aligned} \tag{137}$$

subject to:

$$\begin{aligned} & \mu_1 + \mu_2 = \mu \\ & b' g_2 = b \\ & b' \leq N_{21} \\ & aN_{22} \leq 1 \\ & (1 - \sqrt{g_2})^2 N_{22} + g_2 N_{21} = 1 \\ & 0 \leq [\mu_1, \mu_2, b', g_2, N_{22}, N_{21}] \end{aligned}$$

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 (57) and (58) where applied to confirm the admissibility of the channel and to validate the upper bound obtained in Lemma 3. Hence, $\sigma_{\mathscr{C}}(\mu,1) \le W_1(\mu)$. By using new variable $S = (1-\sqrt{g_2})^2 N_{22}$, we obtain

$$\begin{aligned} W_1(\mu) =& \min \frac{\mu_1}{2}\log\left[2\pi e(P_1 + aP_2 + 1)\right] + \frac{1}{2}\log\left[(1-\sqrt{g_2})^2\left(\frac{1-S+bP_1}{g_2 S} + \frac{P_2}{(1-\sqrt{g_2})^2 P_2 + S}\right)\right] \\ & + \mu_2 f_h\left(P_1, 1, \frac{1-S}{g_2}, \frac{b}{g_2}, \frac{1}{\mu_2}\right) + f_h(P_2, \frac{S}{(1-\sqrt{g_2})^2}, 1, a, \mu_1) - \frac{\mu_2}{2}\log(2\pi e) \end{aligned} \tag{138}$$

subject to:

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

The second optimization problem is written as

$$\begin{aligned} W_2(\mu) =& \min \mu\gamma\left(\frac{P_1}{N_{11}} + \frac{P_1}{a'P_2 + N_{12}}\right) + \gamma\left(\frac{P_2}{N_{22}} + \frac{P_2}{b'P_1 + N_{21}}\right) + f_h(P_2, N_{22}, N_{12}, a', \mu) \\ & + \frac{\mu}{2}\log((2\pi e)(a'P_2 + N_{12})) - \frac{1}{2}\log((2\pi e)(P_2 + N_{22})) \end{aligned} \tag{139}$$

subject to:

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

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 $\sigma_{\mathscr{C}}(\mu,1) \le W_2(\mu)$. By using new variables $S_1 = g_1 N_{12}$ and $S_2 = g_2 N_{21}$, we obtain

$$\begin{aligned} W_2(\mu) =& \min \gamma\left(\frac{(1-\sqrt{g_1})^2 P_1}{1-S_1} + \frac{g_1 P_1}{aP_2 + S_1}\right) + \gamma\left(\frac{(1-\sqrt{g_2})^2 P_2}{1-S_2} + \frac{g_2 P_2}{bP_1 + S_2}\right) \\ & + f_h\left(P_2, \frac{1-S_1}{(1-\sqrt{g_1})^2}, \frac{S_1}{g_1}, \frac{a}{g_1}, \mu\right) + \frac{\mu}{2}\log\left((2\pi e)\left(\frac{aP_2 + S_1}{g_1}\right)\right) - \frac{1}{2}\log\left((2\pi e)\left(P_2 + \frac{1-S_2}{(1-\sqrt{g_2})^2}\right)\right) \end{aligned} \tag{140}$$

subject to:

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

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

$$\mu R_1 + R_2 \le W(\mu) = \min\{W_1(\mu), W_2(\mu)\} \tag{141}$$

$$R_1 + \mu R_2 \le \tilde{W}(\mu) = \min\{\tilde{W}_1(\mu), \tilde{W}_2(\mu)\} \tag{142}$$

hold for all $1 \le \mu$.

$$\mathscr{C}_{\text{sum}} \le \min \frac{(\mu_2 - 1)W(\mu_1) + (\mu_1 - 1)\tilde{W}(\mu_2)}{\mu_1\mu_2 - 1} \tag{143}$$

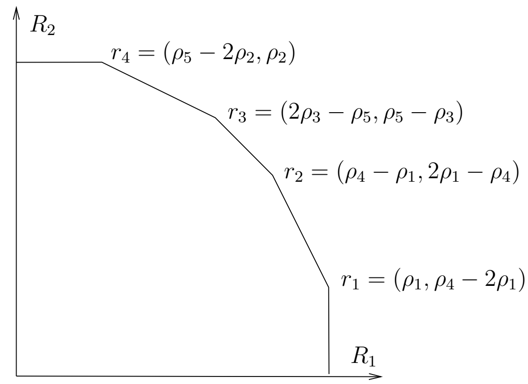

Fig. 9. M-user Interference Channel.

C. Han-Kobayashi Achievable region

Let $D_0(P_1, P_2, \alpha, \beta)$ denote a subset of $\overline{\mathscr{G}}$ where parameters P_1, P_2, α, and β are fixed. In fact, D_0 is a polytope represented by $A\mathbf{R} \leq \Theta$ where $\mathbf{R} = (R_1, R_2)^t$, $\Theta = (\rho_1, \rho_2, \rho_3, \rho_4, \rho_5)^t$, and

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

Let us define D_1 and D_2 as two regions that enlarge D_0 by time-sharing and concavification, respectively. In fact, D_1 that we are interested in its characterization is equivalent to $\mathscr{G}$. To this end, we aim to show $D_1 = D_2$ and make use of the characterization of D_2 to characterize D_1. From Theorem 1, it suffices to show that D_0 possesses the active extreme points condition. The support function of D_0 can be written as

$$\sigma_{D_0}(c_1, c_2, P_1, P_2, \alpha, \beta) = \max\{c_1 R_1 + c_2 R_2 | A\mathbf{R} \leq \Theta(P_1, P_2, \alpha, \beta)\}. \tag{144}$$

To prove D_0 possesses the active extreme points condition, we need to show that $\hat{\mathbf{y}}$, the minimizer of the optimization problem

$$\sigma_{D_0}(c_1, c_2, P_1, P_2, \alpha, \beta) = \min\{\mathbf{y}^t \Theta(\mathbf{P}) | A^t \mathbf{y} = (c_1, c_2)^t, 0 \leq \mathbf{y}\}, \tag{145}$$

is independent of parameters P_1, P_2, α, and β and only depends on c_1 and c_2. We can verify that this indeed valid for the weak Gaussian IC. Since the Han-Kobayashi achievable rate region is symmetrical with respect to R_1 and R_2, we only need to prove it for $(c_1, c_2) = (\mu, 1)$ for all $1 \leq \mu$. However, it is easy to show that $D_0(P_1, P_2, \alpha, \beta)$, a polytope in the first quadrant, has always four extreme points, namely r_1, r_2, r_3, and r_4 (see figure 9). On the other hand, thanks to the results obtained for linear programs, we know that the maximum of (144) is attained at one of its extreme points. It can be shown that for $2 < \mu$, the maximum of (144) is attained at r_1 and we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (\mu - 2)\rho_1 + \rho_4. \tag{146}$$

By comparing (146) with (145), the dual variable that minimizes (145) is $\hat{\mathbf{y}} = (\mu - 2, 0, 0, 1, 0)^t$ which is clearly independent of P_1, P_2, α, and β. For $1 \leq \mu \leq 2$, the maximum of (144) is attained at r_2 and we have

$$\sigma_{D_0}(\mu, 1, P_1, P_2, \alpha, \beta) = (2 - \mu)\rho_3 + (\mu - 1)\rho_4. \tag{147}$$

Again by comparison, we deduce that the dual variable that minimizes (145) is $\hat{\mathbf{y}} = (0, 0, 2 - \mu, \mu - 1, 0)^t$ which is clearly independent of P_1, P_2, α, and β. Hence, D_0 has the active extreme points condition. Using this, we can state the following theorem.

Theorem 7: For the two-user weak Gaussian IC, time-sharing and concavification result in the same region. In other words, $\mathscr{G}$ can fully be characterized by splitting the available space into three subspaces (For example by using FD or TD) and allocating power over subspaces.

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 []. In this case, the capacity region is the union of all rate pairs (R_1, R_2) satisfying

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

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, 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. Finally in [14], Sason used the result of Sato to characterize 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., $\mathscr{G}$.

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 $\sigma_{\mathscr{C}}(\mu, 1) = \max\{\mu R_1 + R_2 | (R_1, R_2) \in \mathscr{C}\}$ for all $1 \leq \mu$.

B. Outer Bound

In [10], Sato derived an outer bound for the capacity of the degraded IC. On the other hand, due to Costa's result, 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

$$\begin{aligned} R_1 &\leq \gamma\left(\tfrac{(1-\beta)P}{1/a+\beta P}\right) \\ R_2 &\leq \gamma(\beta P) \end{aligned} \tag{148}$$

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

Proof: Using Fano's inequality, we have

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

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 (38).

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

$$\mu R_1 + R_2 \leq \mu\gamma\left(\frac{P_1}{1+aP_2}\right) + \gamma(P_2). \tag{149}$$

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 (149) 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

$$\mu R_1 + R_2 \leq \frac{\mu}{2}\log\left(P_1 + aP_2 + 1\right) + \frac{1}{2}\log\left(\frac{1/a - 1}{\mu - 1}\right) - \frac{\mu}{2}\log\left(\frac{\mu a(1/a - 1)}{\mu - 1}\right). \tag{150}$$

Equivalently, we have

$$\mu R_1 + R_2 \leq \frac{\mu}{2}\log\left(\frac{(aP+1)(\mu-1)}{\mu(1-a)}\right) + \frac{1}{2}\log\left(\frac{1/a - 1}{\mu - 1}\right), \tag{151}$$

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

$$E_1 = \left\{(R_1, R_2) | R_1 \leq \gamma\left(\frac{(1-\beta)P}{1/a + \beta P}\right), R_2 \leq \gamma(\beta P),\ \forall \beta \in [0, 1]\right\} \tag{152}$$

and

$$E_2=\left\{(R_1,R_2)|\mu R_1+R_2\leq \frac{\mu}{2}\log\left(\frac{(aP+1)(\mu-1)}{\mu(1-a)}\right)+\frac{1}{2}\log\left(\frac{1/a-1}{\mu-1}\right),\ \forall \frac{P_2+1/a}{P_2+1}<\mu\leq\frac{1}{a}\right\}. \tag{153}$$

In fact, E_2 is the dual representation of E_1, see (17). To show this, we evaluate the support function of E_1 as

$$\sigma_{E_1}(\mu,1)=\max\left\{\mu R_1+R_2|(R_1,R_2)\in E_1\right\}. \tag{154}$$

It is easy to show that

$$\sigma_{E_1}(\mu,1)=\frac{\mu}{2}\log\left(\frac{(aP+1)(\mu-1)}{\mu(1-a)}\right)+\frac{1}{2}\log\left(\frac{1/a-1}{\mu-1}\right). \tag{155}$$

Since E_1 is a closed convex set, we can make use of (17) to obtain the dual representation of it which is indeed equivalent to (153). This completes the proof. □

C. Han-Kobayashi Achievable Region

Since there is no link between Transmitter 1 and Receiver 2, User 1's message in Han-Kobayashi achievable rate region is private message. In this case, we have

$$\rho_1=\gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{156}$$

$$\rho_2=\gamma(P_2), \tag{157}$$

$$\rho_{31}=\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{158}$$

$$\rho_{32}=\gamma\left(\frac{P_1}{1+a\beta P_2}\right)+\gamma(P_2), \tag{159}$$

$$\rho_{33}=\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{160}$$

$$\rho_4=\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma\left(\frac{P_1}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{161}$$

$$\rho_5=\gamma(\beta P_2)+\gamma(P_2)+\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right), \tag{162}$$

It is easy to see that $\rho_3=\rho_{31}$, $\rho_{31}+\rho_1=\rho_4$, $\rho_{31}+\rho_1=\rho_4$. Hence, $\overline{\mathscr{G}}$ can be represented as all rate pairs (R_1,R_2) satisfying

$$R_1\leq\gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{163}$$

$$R_2\leq\gamma(P_2), \tag{164}$$

$$R_1+R_2\leq\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{165}$$

for all $\beta\in[0,1]$. For a fixed β, the region is a pentagon with two extreme points in the interior of the first quadrant, namely r_1 and r_2. The first extreme point which is $r_1=\left(\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2)-\gamma(P_2),\gamma(P_2)\right)$ lies on the boundary of the capacity region. The second extreme point can potentially be a point on the boundary of $\overline{\mathscr{G}}$. This is indeed the case and we prove it in the following lemma.

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

$$R_1\leq\gamma\left(\frac{P_1}{1+\beta aP_2}\right) \tag{166}$$

$$R_2\leq\gamma(\beta P_2)+\gamma\left(\frac{(1-\beta)aP_2}{1+P_1+\beta aP_2}\right) \tag{167}$$

Moreover, $\overline{\mathscr{G}}$ 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 $E\subseteq\overline{\mathcal{G}}$. Indeed, the extreme point of E for a fixed β corresponds to r_2. Hence, we need to show the other inclusion. To this end, let us select an arbitrary point inside $\overline{\mathcal{G}}$, say (R_1',R_2'). Hence, there exist a β' such that R_1' and R_2' satisfy (163), (164), and (165). Since $\gamma\left(\beta P_2\right)+\gamma\left(\frac{(1-\beta)aP_2}{1+P_1+\beta aP_2}\right)$ in (167) is a continuous function over a compact set, there is a $\beta'\leq\beta\leq 1$ such that

$$R_2'=\gamma\left(\beta P_2\right)+\gamma\left(\frac{(1-\beta)aP_2}{1+P_1+\beta aP_2}\right). \tag{168}$$

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 (163), (164), and (165), we have

$$R_1' \leq \min\left\{\gamma\left(\frac{P_1}{1+\beta' a P_2}\right), 0.5\log\left(\frac{1+\beta' P_2}{1+\beta P_2}\right) + 0.5\log\left(\frac{1+P_1+\beta a P_2}{1+\beta' a P_2}\right)\right\}. \tag{169}$$

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, $E = \overline{\mathscr{G}}$.

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

Let us denote D_0 as the collection of all (R_1, R_2) satisfying

$$R_1 \leq \gamma\left(\frac{P_1}{1+\beta a P_2}\right) \tag{170}$$

$$R_2 \leq \gamma(\beta P_2) + \gamma\left(\frac{(1-\beta) a P_2}{1+P_1+\beta a P_2}\right) \tag{171}$$

for fixed β, P_1, and P_2. Clearly, D_0 possesses the active extreme points condition and hence, time-sharing and concavification over D_0 result in the same region. On the other hand, from Lemma 7 we can deduce that $\mathscr{G} = D_1 = D_2$. As a result, boundary points of the full Han-Kobayashi achievable rate region when Gaussian codebooks are used for data transmission can be obtained from the following optimization problem:

$$W = \max \sum_{i=1}^{3} \lambda_i \left[\mu\gamma\left(\frac{\lambda_i P_{1i}}{1+\beta_i a \lambda_i P_{2i}}\right) + \gamma(\beta_i \lambda_i P_{2i}) + \gamma\left(\frac{(1-\beta_i) a \lambda_i P_{2i}}{1+\lambda_i P_{1i} + \beta a \lambda_i P_{2i}}\right)\right] \tag{172}$$

subject to:

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

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

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

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

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

VI. Mixed Gaussian Interference Channels

In this section, we focus on the mixed Gaussian Interference channel. We firs 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.

A. Sum Capacity

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

$$\mathscr{C}_{sum} = \gamma(P_2) + \min\left\{\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+P_2}\right)\right\}. \tag{173}$$

Proof: We need to prove the achievaility 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

$$R_1 = \min\left\{\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+P_2}\right)\right\} \tag{174}$$

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

$$R_2 = \gamma(P_2). \tag{175}$$

By adding (174) and (175), 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

$$\mathscr{C}_{sum} \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1+aP_2}\right). \tag{176}$$

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

$$\mathscr{C}_{sum} \leq \gamma\left(bP_1 + P_2\right) \tag{177}$$

which equivalently can be written as

$$\mathscr{C}_{sum} \leq \gamma(P_2) + \gamma\left(\frac{bP_1}{1+P_2}\right). \tag{178}$$

By taking the minimum of the right hand sides of Inequalities (176) and (178), we obtain

$$\mathscr{C}_{sum} \leq \gamma\left(P_2\right) + \min\left\{\gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+P_2}\right)\right\}. \tag{179}$$

This completes the proof. □

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.

B. Outer Bound

The second outer bound, due to Etkin, Tse, and Wang [21], 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

$$R_1 \leq \gamma(P_1) \tag{180}$$

$$R_2 \leq \gamma(P_2) \tag{181}$$

$$R_1 + R_2 \leq \gamma(P_2) + \gamma\left(\frac{P_1}{1+aP_2}\right) \tag{182}$$

$$R_1 + R_2 \leq \gamma(P_2 + bP_2) \tag{183}$$

$$2R_1 + R_2 \leq \gamma(P_1 + aP_2) + \gamma\left(bP_1 + \frac{P_2}{1+aP_2}\right) + \gamma\left(\frac{P_1}{1+bP_1}\right) \tag{184}$$

Even though, the capacity region of the one-sided Gaussian IC is unknown, there exist 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

$$R_1 \leq \gamma\left(\frac{(1-\beta)P'}{\beta P' + 1/a}\right) \tag{185}$$

$$R_2 \leq \gamma(\beta P') \tag{186}$$

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

$$R_1 \leq \gamma(bP_1) \tag{187}$$

$$R_2 \leq \gamma\left(P_2\right) \tag{188}$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2) \tag{189}$$

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

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 $\mathscr{C}$ it suffices to calculate $\sigma_{\mathscr{C}}(\mu, 1)$ and $\sigma_{\mathscr{C}}(1, \mu)$ for all $1 \leq \mu$. To this end, we make use of channels in A1 and B Classes and channels in A2 and B to obtain upper bounds for $\sigma_{\mathscr{C}}(\mu, 1)$ and $\sigma_{\mathscr{C}}(1, \mu)$, respectively.

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

$$\begin{aligned}W(\mu) =&\min \frac{\mu-1}{2}\log\left(2\pi e(P_1+aP_2+1)\right)+\frac{1}{2}\log\left(2\pi e\left(\frac{P_2N_{22}}{P_2+N_{22}}+b'P_1+N_{21}\right)\right)\\&-\frac{1}{2}\log(2\pi eN_{21})-\frac{1}{2}\log(2\pi eN_{22})+f_h(P_2,N_{22},1,a,\mu-1)\\&\text{subject to:}\\&\qquad b'g_2=b\\&\qquad b'\geq N_{21}\\&\qquad aN_{22}\leq 1\\&\qquad (1-\sqrt{g_2})^2N_{22}+g_2N_{21}=1\\&\qquad 0\leq[b',g_2,N_{22},N_{21}]\end{aligned} \tag{190}$$

$$\begin{aligned}W(\mu) =&\min \frac{\mu-1}{2}\log\left(2\pi e(P_1+aP_2+1)\right)+\frac{1}{2}\log\left(2\pi e\left(\frac{P_2(1-S)}{(1-\sqrt{g_2})^2P_2+1-S}+\frac{bP_1+S}{g_2}\right)\right)\\&-\frac{1}{2}\log\left(\frac{2\pi eS}{g_2}\right)-\frac{1}{2}\log\left(\frac{2\pi e(1-S)}{(1-\sqrt{g_2})^2}\right)+f_h\left(P_2,\frac{1-S}{(1-\sqrt{g_2})^2},1,a,\mu-1\right)\\&\text{subject to:}\\&\qquad S<1\\&\qquad a(1-S)\leq(1-\sqrt{g_2})^2\\&\qquad 0\leq[S,g_2]\end{aligned} \tag{191}$$

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

$$\mu R_1+R_2\leq W(\mu) \tag{192}$$

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 where $a<1$ and $b\geq 1$. Since Receiver 2 can always decode the message of the first user, User 1 associates all its power to the common message. However, User 2 allocates βP_2 and $(1-\beta)P_2$ to the private and common messages, respectively, where $\beta\in[0,1]$. Let R_{2c} and R_{2p} denote the common and private rates of User2. Hence, the rate of the second user can be represented as $R_2=R_{2c}+R_{2p}$. In this case, $\mathcal{G}_{\overline{HK}}$ is the union of all (R_1,R_2) satisfying

$$\rho_1=\gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{193}$$

$$\rho_2=\gamma(P_2), \tag{194}$$

$$\rho_{31}=\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{195}$$

$$\rho_{32}=\gamma(P_2+bP_1), \tag{196}$$

$$\rho_{33}=\gamma\left(\frac{a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2+bP_1), \tag{197}$$

$$\rho_4=\gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2+bP_1), \tag{198}$$

$$\rho_5=\gamma(\beta P_2)+\gamma(P_2+bP_1)+\gamma\left(\frac{a(1-\beta)P_2}{1+a\beta P_2}\right), \tag{199}$$

for all $\beta\in[0,1]$. It is easy to verify that the inequality $\rho_{31}+\rho_1\leq\rho_4$ holds. This means that Inequalities (9) and (10) are redundant for all range of parameters and can be removed. To obtain ρ_3, we need to take the minimum of ρ_{31}, ρ_{32}, and ρ_{33}. By comparison, we can show that the following conditions are sufficient to obtain ρ_3.

C1 $0\leq(b-1)P_1+(1-a)(1-\beta)P_2+\beta P_1P_2(ab-1)$ to have $\rho_{31}\leq\rho_{32}$.
C2 $0\leq(b-1)+(ab-\beta)P_2$ to have $\rho_{31}\leq\rho_{33}$.
C3 $1-a\leq abP_1$ to have $\rho_{32}\leq\rho_{33}$.

Indeed, if C1 and C2 hold then $\rho_3=\rho_{31}$. Surprisingly, the condition $1+P_2\leq b+abP_2$ suffices to satisfy both C1 and C2. To proceed, we consider two cases, namely the case where $1+P_2\leq b+abP_2$ and the case where $1+P_2>b+abP_2$.

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 $1-a>abP_1$.

In what follows, we investigate the Han-Kobayashi achievable rate region for each cases.

Case I ($1+P_2\leq b+abP_2$): In this case, $\rho_3=\rho_{31}$. Moreover, It is easy to verify that $\rho_{31}+\rho_1\leq\rho_4$ and $\rho_{31}+\rho_2\leq\rho_5$ hold. This means that Inequalities (9) and (10) are redundant for all range of parameters and can be removed. Hence, $\overline{\mathscr{G}}$ consists of all rate pairs (R_1,R_2) satisfying

$$R_1\leq \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{200}$$

$$R_2\leq \gamma\left(P_2\right), \tag{201}$$

$$R_1+R_2\leq \gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2), \tag{202}$$

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

$$R_1\leq \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{203}$$

$$R_2\leq \gamma(\beta P_2)+\gamma\left(\frac{a(1-\beta)P_2}{1+P_1+a\beta P_2}\right) \tag{204}$$

$$\tag{205}$$

for all $\beta\in[0,1]$. Now, we can state the following theorem.

Theorem 12: $\mathscr{G}$ of the mixed Gaussian IC satisfying $1+P_2\leq b+abP_2$ is equivalent to that of the one sided Gaussian IC obtained from removing the interfering link between Transmitter 1 and Receiver 2.

Proof: By comparing (203), (204) with () and (), we see that $\overline{\mathscr{G}}$ of this channel is exactly the same as that of the one sided Gaussian IC obtained from removing the interfering link between Transmitter 1 to Receiver 2. Hence, we can deduce that $\mathscr{G}$ is equivalent for both channels. □

Case II ($1+P_2>b+abP_2$ and $1-a\leq abP_1$): In this case, $\rho_3=\min\{\rho_{31},\rho_{32}\}$. Hence, we need to investigate different situations arising from choosing different β.

$\overline{\mathscr{G}}=E_1\bigcup E_2\bigcup E_3$.

E_1 is the union of all rate pairs (R_1,R_2) satisfying

$$R_1\leq \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{206}$$

$$R_2\leq \gamma(\beta P_2)+\gamma\left(\frac{a(1-\beta)P_2}{1+P_1+a\beta P_2}\right). \tag{207}$$

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

$$R_1\leq \gamma\left(\frac{bP_1}{1+\beta P_2}\right), \tag{208}$$

$$R_2\leq \gamma\left(\frac{P_1+a(1-\beta)P_2}{1+a\beta P_2}\right)+\gamma(\beta P_2)-\gamma\left(\frac{bP_1}{1+\beta P_2}\right). \tag{209}$$

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

$$R_1\leq \gamma\left(\frac{bP_1(1+\frac{(1-ab)P_1}{1-a})}{1+bP_1+P_2}\right) \tag{210}$$

$$R_2\leq \gamma\left(P_2\right) \tag{211}$$

$$R_1+R_2\leq \gamma(bP_1+P_2) \tag{212}$$

Case III ($1+P_2>b+abP_2$ and $1-a>abP_1$): In this case, $\rho_3=\min\{\rho_{31},\rho_{32}\}$. Hence, we need to investigate different situations arising from choosing different β.

$\overline{\mathscr{G}}=E_1\bigcup E_2\bigcup E_3$.

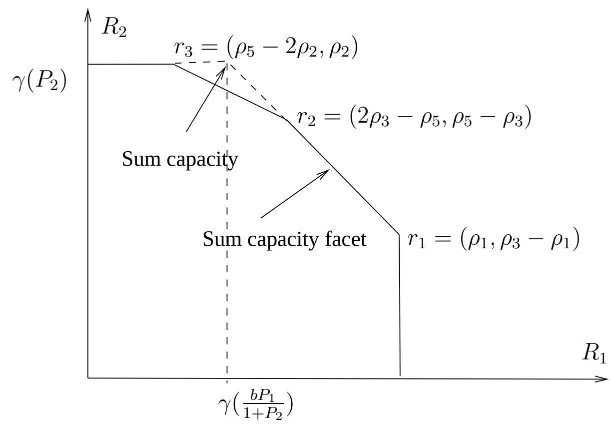

Fig. 10. M-user Interference Channel.

E_1 is the union of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{213}$$

$$R_2 \leq \gamma(\beta P_2) + \gamma\left(\frac{a(1-\beta)P_2}{1+P_1+a\beta P_2}\right). \tag{214}$$

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

$$R_1 \leq \gamma\left(\frac{P_1}{1+a\beta P_2}\right), \tag{215}$$

$$R_2 \leq \gamma\left(\frac{a(1-\beta)P_2}{1+P_1+a\beta P_2}\right) + \gamma(\beta P_2 + bP_1) - \gamma\left(\frac{P_1}{1+a\beta P_2}\right). \tag{216}$$

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

$$R_1 \leq \gamma\left(\frac{P_1}{1+aP_2}\right) \tag{217}$$

$$R_2 \leq \gamma\left(P_2\right) \tag{218}$$

$$R_1 + R_2 \leq \gamma(bP_1 + P_2) \tag{219}$$

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.

APPENDIX I
CONCAVIFICATION OF SET VALUED MAPS

Suppose there exists a strategy S such that for every power constraint $\mathbf{p} \in \Re^n$, one can obtain an achievable rate region $f_S(\mathbf{p}) \subseteq \Re^m$. It is possible, however, to consider the achievable rate region $f_S : \Re^n \rightrightarrows \Re^m$ as a set-valued map that maps points in $\Re^n$ to subsets of $\Re^m$, c.f. [22]. A strategy S is called concave if its associated set-valued map $f_S(\mathbf{p})$ is concave, where a concave set-valued map is defined as follows.

Definition 4 (concave set-valued map): A set-valued map $f(\mathbf{p})$ is concave if it satisfies

$$\lambda f(\mathbf{p}_1) + (1-\lambda) f(\mathbf{p}_2) \subseteq f(\lambda \mathbf{p}_1 + (1-\lambda)\mathbf{p}_2), \tag{220}$$

for every $\mathbf{p}_1, \mathbf{p}_2 \in \Re^n$ and $\lambda \in [0,1]$.
By substituting $\mathbf{p} = \mathbf{p}_1 = \mathbf{p}_2$ into (220), we observe that the image of every point under a concave set-valued map is a convex subset of the range.

For a set-valued map f, we define its concave hull $\text{conch}(f)$ as the least set-valued map minorized by f, i.e., $\text{conch}(f) \leq g$ for every concave set-valued map $g \geq f$, where we say $f_1 \leq f_2$ if $f_1(\mathbf{p}) \subseteq f_2(\mathbf{p})$ for every $\mathbf{p} \in \Re^n$.

Lemma 8: Pointwise intersection of any collection of concave set-valued maps is concave, i.e., $f(\mathbf{p}) = \bigcap_{i \in I} f_i(\mathbf{p})$ is concave if each map f_i is concave.

Proof: It is easy to show that $\lambda f(\mathbf{p}_1) + (1-\lambda) f(\mathbf{p}_2) \subseteq \lambda f_i(\mathbf{p}_1) + (1-\lambda) f_i(\mathbf{p}_2)$, for every $\mathbf{p}_1, \mathbf{p}_2 \in \Re^n$, $\lambda \in [0,1]$, and $i \in I$. By applying (220) for each f_i, we obtain $\lambda f(\mathbf{p}_1) + (1-\lambda) f(\mathbf{p}_2) \subseteq f_i(\lambda \mathbf{p}_1 + (1-\lambda)\mathbf{p}_2)$. Hence $\lambda f(\mathbf{p}_1) + (1-\lambda) f(\mathbf{p}_2) \subseteq \bigcap_{i \in I} f_i(\lambda \mathbf{p}_1 + (1-\lambda)\mathbf{p}_2)$ which completes the proof. □

Lemma 9: $\text{conch}(f)$ is the pointwise intersection of all concave set-valued maps minorized by f.

Proof: Assume $g = \bigcap_{i \in I} f_i$, where f_i's are all concave maps greater than f. Clearly, $g \leq f_i$ for every i. The map g is concave by Lemma 8. Hence $g = \text{conch}(f)$. □

Theorem 13 (concavification of a set-valued map): For $f : \Re^n \rightrightarrows \Re^m$

$$(\text{conch} f)(\mathbf{p}) = \left\{ \sum_{i=0}^{n+m} \lambda_i f(\mathbf{p}_i) \mid \sum_{i=0}^{n+m} \lambda_i = 1, \sum_{i=0}^{n+m} \lambda_i \mathbf{p}_i = \mathbf{p}, \lambda_i \geq 0 \forall i \right\}. \tag{221}$$

Moreover, if

Proof: We define $g(\mathbf{p}) = \left\{ \sum_{i=0}^{q} \lambda_i f(\mathbf{p}_i) \mid \sum_{i=0}^{q} \lambda_i = 1, \sum_{i=0}^{q} \lambda_i \mathbf{p}_i = \mathbf{p}, \lambda_i \geq 0 \ \forall i, q \geq 0 \right\}$. It is easy to show that g satisfies (220) and hence is a concave set-valued map. To show that $g = \text{conch} f$, it suffices to prove that for every concave set-valued map h minorized by f we have $g \leq h$. To this end, we fix a point in $g(\mathbf{p})$, say $\mathbf{x}$, and prove that this point belongs to $h(\mathbf{p})$. From the definition of g, $\mathbf{x}$ can be written as $\mathbf{x} = \sum_{i=0}^{q} \lambda_i \mathbf{x}_i$ for some $q \geq 0$, where $\mathbf{x}_i \in f(\mathbf{p}_i)$ and $\sum_{i=0}^{q} \lambda_i \mathbf{p}_i = \mathbf{p}$. By hypothesis $f(\mathbf{p}_i) \subseteq h(\mathbf{p}_i)$ for all i, hence we have $\sum_{i=0}^{q} \lambda_i f(\mathbf{p}_i) \subseteq \sum_{i=0}^{q} \lambda_i h(\mathbf{p}_i)$. Since h is concave, we can apply property (220) to obtain $\sum_{i=0}^{q} \lambda_i f(\mathbf{p}_i) \subseteq h(\sum_{i=0}^{q} \lambda_i \mathbf{p}_i) = h(\mathbf{p})$. Therefore, $\mathbf{x} \in h(\mathbf{p})$.

By applying Caratheodory theorem, we can prove that the size of q can be proved and its extension []. □

We can define the concave hull of a non-concave strategy S, as a new strategy $\text{conch}(S)$ such that $f_{\text{conch}(S)} = \text{conch}(f_S)$

Corollary 5: For a given non-concave strategy S, the concave hull of f_S is a new strategy which its associated set-valued map is achievable.

We are interested in characterizing boundary points of $(\text{conch} f)(\mathbf{p})$ for a fixed $\mathbf{p}$. In the following theorem, we characterize the boundary points of the region.

Theorem 14: The boundary points of $(\text{conch} f)(\mathbf{p})$ in (16) can be written as

$$\mathbf{bd}((\text{conch} f)(\mathbf{p})) = \left\{ \sum_{i=0}^{n+1} \lambda_i x_i | x_i \in f(\mathbf{p}) \forall i, \sum_{i=0}^{n+m} \lambda_i = 1, \sum_{i=0}^{n+m} \lambda_i \mathbf{p}_i = \mathbf{p}, \lambda_i \geq 0 \forall i \right\}. \tag{222}$$

Proof: we start with □

Due to convexity of $f_{\text{conch}(S)}(\mathbf{p})$ for a fixed $\mathbf{p}$, we can define the following optimization problem to achieve points on the boundary of the region.

$$\sigma_{conchf}(\mathbf{y}, \mathbf{p}) = \sup\{\mathbf{y}^T \mathbf{x} | \mathbf{x} \in f_{\text{conch}(S)}(\mathbf{p})\} \tag{223}$$

$$\sigma_f(\mathbf{y}, \mathbf{p}) = \sup\{\mathbf{y}^T \mathbf{x} | \mathbf{x} \in f(\mathbf{p})\} \tag{224}$$

$$\sigma_{conchf}(\mathbf{y}, \mathbf{p}) = \max \left\{ \sum_{i=0}^{n+m} \lambda_i \sigma_f(\mathbf{y}, \mathbf{p}_i) | \sum_{i=0}^{n+m} \lambda_i = 1, \lambda_i \geq 0 \ \forall \ i, \sum_{i=0}^{n+m} \lambda_i \mathbf{p}_i = \mathbf{p} \right\} \tag{225}$$

APPENDIX II
PROOF OF LEMMA 4

From the general result for (31), we know that the optimum input distribution is a Gaussian vector. Hence, we need to solve the following maximization problem:

$$\begin{aligned} W =& \max \frac{1}{2}\log\left((2\pi e)^n |Q_{\mathbf{X}} + N_1 I|\right) - \frac{\mu}{2}\log\left((2\pi e)^n |Q_{\mathbf{X}} + N_2 I|\right) \\ & \text{subject to:} \\ & \quad 0 \le Q_{\mathbf{X}} \\ & \quad tr\{Q_{\mathbf{X}}\} \le nP \end{aligned} \tag{226}$$

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 (226) and using the identities $tr\{AB\} = tr\{BA\}$ and $|AB + I| = |BA + I|$, we obtain

$$\begin{aligned} W =& \max \frac{1}{2}\log\left((2\pi e)^n |\Lambda + N_1 I|\right) - \frac{\mu}{2}\log\left((2\pi e)^n |\Lambda + N_2 I|\right) \\ & \text{subject to:} \\ & \quad 0 \le \Lambda \\ & \quad tr\{\Lambda\} \le nP \end{aligned} \tag{227}$$

This optimization problem can be simplified as

$$\begin{aligned} W =& \max \frac{n}{2}\sum_{i=1}^{n}\left[\log(2\pi e)(\lambda_i + N_1) - \mu\log(2\pi e)(\lambda_i + N_2)\right] \\ & \text{subject to:} \\ & \quad 0 \le \lambda_i \ \forall i \\ & \quad \sum_{i=1}^{n}\lambda_i \le nP \end{aligned} \tag{228}$$

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

$$L(\Lambda, \psi, \Phi) = \max \frac{n}{2}\sum_{i=1}^{n}\left[\log(2\pi e)(\lambda_i + N_1) - \mu\log(2\pi e)(\lambda_i + N_2)\right] + \psi\left(nP - \sum_{i=1}^{n}\lambda_i\right) + \sum_{i=1}^{n}\phi_i\lambda_i. \tag{229}$$

The first order KKT necessary conditions for the optimum solutions of (229) can be written as

$$\frac{1}{\lambda_i + N_1} - \frac{\mu}{\lambda_i + N_2} - \psi + \phi_i = 0, \ \forall i \in \{1, 2, \ldots, n\} \tag{230}$$

$$\psi\left(nP - \sum_{i=1}^{n}\lambda_i\right) = 0, \tag{231}$$

$$\phi_i\lambda_i = 0, \ \forall i \in \{1, 2, \ldots, n\} \tag{232}$$

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

$$\lambda = \begin{cases} P, & \text{if} \quad 0 \le \mu \le \frac{N_2+P}{N_1+P} \\ \frac{N_2 - \mu N_1}{\mu - 1}, & \text{if} \quad \frac{N_2+P}{N_1+P} < \mu \le \frac{N_2}{N_1} \\ 0, & \text{if} \quad \frac{N_2}{N_1} < \mu \end{cases} \tag{233}$$

Now, by substituting into the objective function, we obtain the desired result.

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 communnication 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. 49–6o, 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 Internatinal Symposiun 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] 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.

[19] S. Ihara, "On the capacity of channels with additive non-Gaussian noise." *Info. Ctrl.*, vol. 37, no. 1, pp. 34–39, Apr. 1978.

[20] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 374–377, May 1978.

[21] 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.

[22] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer-Verlag, Berlin Heidelberg., 1998.