On Limiting Expressions for the Capacity Region of Gaussian Interference Channels

Mojtaba Vaezi and H. Vincent Poor
Department of Electrical Engineering
Princeton University, Princeton, NJ 08544, USA
Email: mvaezi@princeton.edu, poor@princeton.edu

(Invited Paper)

***Abstract*—With an explicit counterexample, it is shown that the restriction to Gaussian distributions in the limiting expression for the capacity region of memoryless Gaussian interference channel falls short of achieving capacity, in general. Techniques that enable one to make use of the limiting expression to obtain valid and possibly tight outer bounds for Gaussian interference channels are also explained.**

I. Introduction

Interference is a fundamental phenomenon in wireless communication networks when multiple users share a common communication medium. A K-user *interference channel*, consisting of $K, K \geq 2$, transmitter and receiver pairs, is used to model such communication networks in general. The two-user interference channel, originating from [1], is the simplest interference channel and a building block of many wireless networks. Fundamental limits of this channel have been explored for many years. A *limiting expression*[1] for the capacity region for the *discrete memoryless* interference channel was derived in [2], but characterization of a *single-letter expression* for the capacity region of this channel is, in general, unknown.

Limiting characterizations of capacity regions first appeared in Shannon's work [1]. Ahlswede [2] showed that the capacity region of the discrete memoryless interference channel, with inputs X_1^n and X_2^n, and outputs Y_1^n and Y_2^n, is given by the following limiting expression:

$$\mathcal{C}_{\rm IC} = \lim_{n\to\infty} \mathrm{co}\left(\bigcup_{p(x_1^n)p(x_2^n)} \left\{ \begin{array}{ll} R_1 & \leq \frac{1}{n} I(X_1^n;Y_1^n) \\ R_2 & \leq \frac{1}{n} I(X_2^n;Y_2^n) \end{array} \right\}\right), \tag{1}$$

where $\mathrm{co}(\cdot)$ denotes the *convex hull* and superscript n denotes the length of the input and output vectors.

Such multiletter expressions can be readily obtained for other multi-user channels. In particular, for $Y_1 = Y_2 = Y$, (1) gives a limiting characterization of the capacity region of the *multiple access channel* (MAC) [3], with an output Y. Although the above limiting characterization of the capacity region is well-defined, it is considered to have little value because [4]–[6]

- it is computationally excessively complex and thus it is not clear how to compute it;
- it does not provide any insight into how to best code for this channel; and,
- it cannot be directly used to compute the capacity region of the Gaussian interference channel by restricting to Gaussian inputs.

In contrast to *multiletter* expressions, a *single-letter* capacity expression includes only the channel input and output random variables, and in some cases a few auxiliary random variables, involved in "one" use of the channel. A *single-letter* capacity expression for the discrete memoryless interference channel is in general unknown, except in some special cases. Similarly, the capacity region of the *Gaussian* interference channel is not known in general. Cheng and Verdú [6] proved that restriction to Gaussian inputs in the limiting expression of the MAC falls short of achieving the capacity region of the Gaussian MAC.

In this paper, by providing an explicit counterexample, we first prove a similar result for the interference channel. We next show that although, in general, one cannot restrict the inputs in (1) to Gaussian inputs to compute the capacity region, the multiletter expression can still be useful. We discuss how the multiletter expression can be exploited to develop valid outer bounds on the capacity region of the Gaussian interference channel. This identifies a general approach for improving the outer bounds for Gaussian interference networks, which in turn gives hints on the structure of the optimal codes, at least in certain cases.

The paper is organized as follows. The system model is described in Section II. Section III provides the counterexample. Section IV discusses techniques that help improve outer bounds based on the multiletter expression. This is followed by conclusion in Section V.

II. Preliminaries

Consider a two-user *Gaussian* interference channel where user 1 and user 2 wish to transmit independent messages M_1 and M_2, respectively. The input-output relationships for this channel, for a single channel use, can be expressed in the

This work was supported in part by a fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC), and in part by the U.S. National Science Foundation (NSF) under Grant CCF–1420575.

[1]As will seen, a *limiting expression* of the capacity region of a channel is a *multiletter expression* that becomes the capacity region of that channel in the limit. The terms *limiting expression* and *multiletter expression* are usually used exchangeably in the literature.

standard form by [7]

$$Y_1 = X_1 + \sqrt{a}X_2 + Z_1, \tag{2a}$$
$$Y_2 = \sqrt{b}X_1 + X_2 + Z_2, \tag{2b}$$

where a and b are two non-negative real numbers, Z_1 and Z_2 are two independent $\mathcal{N}(0,1)$ noise terms which are also independent from channel use to channel use, and input i is subject to an average power constraint P_i, i.e.,

$$\mathbb{E}(\|X_i\|^2) \leq P_i, \ i = 1, 2. \tag{3}$$

A $(2^{nR_1}, 2^{nR_2}, n, \epsilon_1^n, \epsilon_2^n)$ code for this channel consists of two independent messages M_i, $i \in \{1,2\}$, two encoding functions f_i, two decoding functions g_i, and two average probability errors ϵ_i^n, in which

1) M_i is uniformly distributed over $[1, 2, \ldots, 2^{nR_i}]$,
2) encoder i assigns a codeword $x_i^n(m_i)$ to each message m_i
3) decoder i assigns an estimate $\hat{m}_i \in [1, 2, \ldots, 2^{nR_i}]$ to each received sequence y_i^n, and
4) $\epsilon_i^n = p(\hat{M}_i \neq M_i) = \frac{1}{2^{nR_i}} \sum_{i=1}^{2^{nR_i}} p(\hat{m}_i \neq m_i)$.

A rate pair (R_1, R_2) is achievable if there exist a sequence of codes $(2^{nR_1}, 2^{nR_2}, n, \epsilon_1^n, \epsilon_2^n)$ with $\epsilon_1^n \to 0$ and $\epsilon_2^n \to 0$. The capacity region of the interference channel is the closure of the set of achievable rates.

The capacity region of the two-user interference channel is a long-standing open problem that traces back to Shannon [1]. A *limiting expression* for the capacity region of discrete memoryless interference channel is given in (1). Nevertheless, characterization of a *single-letter* expression for the capacity region of this channel is unknown, except in the *strong* interference case [8]; i.e., when $I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2)$ and $I(X_2;Y_2|X_1) \leq I(X_2;Y_1|X_1)$ for all $p(x_1)p(x_2)$. In this case, the limiting characterization of the capacity region simplifies to the single-letter one. This is because from (1) we can write

$$\begin{aligned} R_1 + R_2 &\leq \frac{1}{n}I(X_1^n;Y_1^n) + \frac{1}{n}I(X_2^n;Y_2^n) \\ &\overset{(a)}{\leq} \frac{1}{n}I(X_1^n;Y_1^n|X_2^n) + \frac{1}{n}I(X_2^n;Y_2^n) \\ &\overset{(b)}{\leq} \frac{1}{n}I(X_1^n;Y_2^n|X_2^n) + \frac{1}{n}I(X_2^n;Y_2^n) \\ &\overset{(c)}{=} \frac{1}{n}\sum_{i=1}^{n} I(X_{1i}, X_{2i};Y_{2i}) \\ &\overset{(d)}{=} I(X_1, X_2;Y_2|Q), \end{aligned} \tag{4}$$

in which (a) is due to the independence of X_1^n and X_2^n which follows from the independence of M_1 and M_2, (b) follows from the strong interference condition $I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2)$ [8], (c) follows by the memoryless property of the channel, and (d) follows by a standard argument [9] in which the random variable Q is uniformly distributed over $\{1, 2, \ldots, n\}$ and is independent of all other random variables. With similar arguments, it is easy to check that

$$R_1 \leq \frac{1}{n}I(X_1^n;Y_1^n) \leq I(X_1;Y_1|X_2,Q). \tag{5}$$

Then again, by swapping the indices 1 and 2 in (4) and (5) we obtain two other inequalities and a single-letter capacity expression is characterized in the strong interference case [8].

As positive as this limiting expression may be, in this paper we show that it cannot, in general, be directly used to compute the capacity region of the Gaussian interference channel by restricting the inputs to Gaussian distributions. This expression should, however, reduce to a single-letter characterizations of the capacity region, if one exists. Also, as we show in this paper, it can be useful in finding improved outer bounds on the capacity region of the interference channel.

We end this section with some notational conventions. We will denote deterministic vectors with boldface lowercase letters, random vectors with uppercase boldface letters, and matrices with uppercase letters. All logarithms are to base 2, and $\gamma(x) \triangleq \frac{1}{2}\log(1+x)$. I is the identity matrix of size n, and for a real number $x \in [0,1]$ we define $\bar{x} = 1 - x$.

III. A Counterexample

With an explicit counterexample, in this section we prove that evaluating (1) with Gaussian inputs falls short of achieving the capacity region. We show that, when evaluated using Gaussian inputs, (1) is a proper subset of an inner bound, for certain channel parameters. The counterexample is developed for the one-sided interference channel. Specifically, we let $b = 0$ and $a < 1$ in (2a)-(2b), which results in a one-sided interference channel in the *weak interference* regime.

A. Inner Bound

Consider the Gaussian interference channel defined by (2a)-(2b) for $b = 0$. We use time-sharing with power allocation as an achievable scheme. Let us divide the set of channel-uses into two non-overlapping sets with proportions τ_1 and τ_2 of the total ($\tau_1 + \tau_2 = 1$). Also, let R_{ji}, $i, j \in \{1,2\}$, represent the rate corresponding to user j over the channel uses in subset i. An achievable rate region for this scheme is given by

Lemma 1. *The set of non-negative (R_1, R_2) satisfying*

$$R_1 \leq \tau_1 R_{11}, \tag{6a}$$
$$R_2 \leq \tau_1 R_{21} + \tau_2 R_{22}, \tag{6b}$$

in which

$$R_{11} \leq \gamma\Big(\frac{\frac{P_1}{\tau_1}}{1 + a\beta_1 P_{21}}\Big), \tag{7a}$$
$$R_{21} \leq \gamma\Big(\frac{a\bar{\beta}_1 P_{21}}{1 + \frac{P_1}{\tau_1} + a\beta_1 P_{21}}\Big) + \gamma(\beta_1 P_{21}), \tag{7b}$$
$$R_{22} \leq \gamma(P_{22}), \tag{7c}$$

is achievable for the one-sided Gaussian interference channel where $\tau_1 + \tau_2 = 1$, $\tau_1 P_{21} + \tau_2 P_{22} = P_2$, $0 \leq \beta_1 \leq 1$, and $\bar{\beta}_1 = 1 - \beta_1$.

Proof. To prove this region we use a time-sharing argument. Consider two sequences of codes, namely C_1 and C_2, achieving (R_{11}, R_{21}) and $(0, R_{22})$, respectively. For each block length n, assume that the first and second codes are used during n_1 and n_2 uses of the channel, respectively, where

$n_1+n_2=n$. Let us also define $\tau_1=\frac{n_1}{n}$ and $\tau_2=\frac{n_2}{n}$. Thus, τ_1 and τ_2 are the fraction of the channel uses for the first and second codes, and $\tau_1+\tau_2=1$. Then from [9, Proposition 4.1 and Remark 4.3] the rate pair $(R_1,R_2)=(\tau_1 R_{11},\tau_1 R_{21}+\tau_2 R_{22})$ is achievable. In effect, we transmit C_1 for a fraction τ_1 of channel uses and C_2 a fraction τ_2 of channel uses.

Here we explain how to construct the above codes. We split the message M_2 into independent submessages M_{21} at rate R_{21} and M_{22} at rate R_{22}. Transmitter 2 divides its power P_2 into P_{21} and P_{22} such that $\tau_1 P_{21}+\tau_2 P_{22}=P_2$, i.e., it uses P_{21} and P_{22} for the channel uses corresponding to τ_1 and τ_2, respectively. Transmitter 1, however, consumes all its power in the channel uses corresponding to τ_1. We explain the communication over each fraction of channel uses; i.e., τ_1 and τ_2.

- During the τ_1 fraction of channel uses both transmitters are active. Transmitter 1 transmits M_1 with an average power of $\frac{P_1}{\tau_1}$. Transmitter 2 divides its submessage M_{21} into two parts: the private and common messages. Let $\beta_1 P_{21}$ and $\bar{\beta}_1 P_{21}$, respectively, represent the power allocated to the *private* and *common* part of the message of user 2 in the n_1 uses. Receiver 1 first decodes the common message of transmitter 2 at a rate of $\gamma\big(\frac{a\bar{\beta}_1 P_{21}}{1+\frac{P_1}{\tau_1}+a\beta_1 P_{21}}\big)$ and cancels it from the received signal. After that, receiver 1 is capable of decoding its own message at a rate of $\gamma\big(\frac{P_1/\tau_1}{1+a\beta_1 P_{21}}\big)$. Clearly, receiver 2 which does not experience any interference is capable of reliably decoding the common message at the same rate. This receiver then decodes its private message at a rate of $\gamma(\beta_1 P_{21})$. As a result, user 1 and user 2 can, respectively, achieve (7a) and (7b) in the n_1 channel uses.
- As user 1 has consumed all its power for the n_1 channel uses, it is silent for the n_2 channel uses. That is, these channel uses are reserved for the transmission of user 2. Transmitter 2 simply sends M_{22} with power P_{22}, and receiver 2 decodes it at a rate of $\gamma(P_{22})$. Then, (7c) is achievable, for user 2, in the n_2 channel uses.

With the above strategy, the rate (7a)-(7c) and consequently (6a)-(6b) are achievable. [2] □

Recalling that $\tau_1 P_{21}+\tau_2 P_{22}=P_2$ and $\tau_1+\tau_2=1$, it is easy to check that the weighted sum-rate μR_1+R_2, $\mu\geq 0$, for the above achievable region is given by

$$\begin{aligned}\mu R_1+R_2&=\tau_1(\mu R_{11}+R_{21})+\tau_2 R_{22}\\&\leq \tau_1\Big[\mu\gamma\Big(\frac{\frac{P_1}{\tau_1}}{1+a\beta_1 P_{21}}\Big)+\gamma\Big(\frac{a\bar{\beta}_1 P_{21}}{1+\frac{P_1}{\tau_1}+a\beta_1 P_{21}}\Big)\\&\quad+\gamma(\beta_1 P_{21})\Big]+\bar{\tau}_1\gamma\Big(\frac{P_2-\tau_1 P_{21}}{\bar{\tau}_1}\Big),\end{aligned}\tag{8}$$

in which $0\leq\tau_1\leq 1$, $0\leq\beta_1\leq 1$, and $0\leq P_{21}\leq\frac{P_2}{\tau_1}$.

B. Limiting Capacity Expression with Gaussian Inputs

In [6], it is noted that the convex hull operation is unnecessary in the limiting expression for the capacity region of this interference channel, given by (1). Therefore, from (1), for any $\mu\geq 0$ we can write

$$\mu R_1+R_2\leq\lim_{n\to\infty}\bigcup_{p(x_1^n)p(x_2^n)}\frac{1}{n}\big[\mu I(X_1^n;Y_1^n)+I(X_2^n;Y_2^n)\big],\tag{9}$$

or equivalently,

$$\mu R_1+R_2\leq\frac{1}{n}\big[\mu I(X_1^n;Y_1^n)+I(X_2^n;Y_2^n)\big],\tag{10}$$

for some $p(x_1^n)p(x_2^n)$, when $n\to\infty$. The inequality (10) can be simply proved by using Fano's inequality too, and it gives a multiletter expression for the weighted sum-capacity of the interference channel. If we expand the right-hand side of (10) for the one-sided interference channel, we obtain

$$\begin{aligned}\mu R_1+R_2&\leq\frac{1}{n}\big[\mu I(X_1^n;Y_1^n)+I(X_2^n;Y_2^n)\big]\\&=\frac{1}{n}\big[\mu h(Y_1^n)-\mu h(Y_1^n|X_1^n)+h(Y_2^n)-h(Y_2^n|X_2^n)\big]\\&=\frac{1}{n}\big[\mu h(X_1^n+\sqrt{a}X_2^n+Z_1^n)-\mu h(\sqrt{a}X_2^n+Z_1^n)\\&\quad+h(X_2^n+Z_2^n)-h(Z_2^n)\big]\triangleq\frac{1}{n}U_o.\end{aligned}\tag{11}$$

We now evaluate U_o for Gaussian input distributions. Without loss of generality, we assume the covariance matrices of the noise vectors are equal to the identity matrix of size n, denoted by I. Also, we assume

$$p(x_1^n)\sim\mathcal{N}(\mathbf{0},K_1),\qquad p(x_2^n)\sim\mathcal{N}(\mathbf{0},K_2),\tag{12a}$$

in which K_1 and K_2 are two positive semidefinite covariance matrices with $\mathrm{tr}(K_1)\leq nP_1$ and $\mathrm{tr}(K_2)\leq nP_2$, which are due to the power constraints defined in (3). Now, let us form

$$\begin{aligned}U=\;&\underset{p(x_1^n)p(x_2^n)}{\text{maximize}}\quad U_o\\&\text{subject to}\quad \mathrm{tr}(K_{X_1^n})\leq nP_1,\\&\qquad\qquad\quad\;\mathrm{tr}(K_{X_2^n})\leq nP_2,\end{aligned}\tag{13}$$

in which U_o is defined in (11).

This problem is solved in Appendix A, from which we have

$$\begin{aligned}\mu R_1+R_2\leq\;&\xi\Big[\mu\gamma\Big(\frac{\frac{P_1}{\xi}}{1+aP_2'}\Big)+\gamma(P_2')\Big]\\&+\bar{\xi}\gamma\Big(\frac{P_2-\xi P_2'}{\bar{\xi}}\Big),\end{aligned}\tag{14}$$

in which $0\leq\xi\leq 1$, $\bar{\xi}=1-\xi$, and $0\leq P_2'\leq P_2$.

C. Proof of Counterexample

In light of the inner and outer bounds in (8) and (14), it is now straightforward to verify that with Gaussian inputs the limiting capacity expression (1) falls short of achieving the capacity region of memoryless Gaussian interference channels.

To this end, we show that the right-hand side (RHS) of (8) can be strictly greater than the RHS of (14). Let $\beta_1=1$ in (8). Now, it is clear that, for $\tau_1=\xi$ and $P_{21}=P_2'$, the inner bound is at least as large as the bound in (14). We next show that the inner bound can be strictly larger than the outer bound. Equivalently, we may prove that $\beta_1=1$ is

[2] It can be checked that the above achievable region is a subset of Han-Kobayashi scheme with *time-sharing* [10].

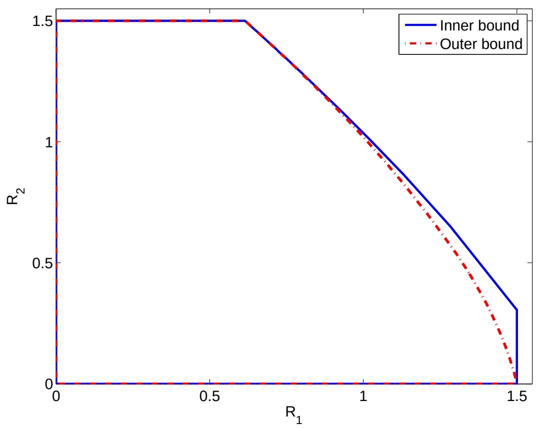

Fig. 1. The inner bound (Lemma 1) and the outer bound assuming that Gaussian inputs are optimal for the multiletter expression (see (14)). The figure clearly shows that such an assumption is not correct. The curves are for the one-sided interference channel with $a = 0.6$, $P_1 = 7$ and $P_2 = 7$.

not the maximizer the RHS of (8) for the whole range of parameters. This is rather easy as, for example, the optimum β_1 is zero for $\mu \geq \frac{\tau_1(1-a)+P_1}{aP_1}$. This completes the proof of the counterexample, indicating that Gaussian inputs cannot maximize U_o, or, equivalently, the restriction to Gaussian inputs in the limiting expression falls short of achieving the capacity of the memoryless Gaussian interference channel. Figure 1 compares the inner and outer bounds numerically.

It should be highlighted that this result applies only to the multiletter capacity expression in (1), or (9) equivalently. It does not imply that Gaussian inputs are not capacity-achieving for the Gaussian interference channel. On the contrary, there might be single-letter capacity expression, or even a different multiletter capacity expression, for the interference channel for which the Gaussian inputs are optimal. To appreciate this, see Cheng and Verdú's discussion [6] in which they have a similar result for the MAC while highlighting the well-known fact that Gaussian inputs are optimal for the Gaussian MAC.

IV. How to Better Use the Multiletter Expression?

In the previous section we have proved that we cannot simply apply Gaussian input distributions to the multiletter expression in (1), and expect to obtain a valid outer bound on the capacity region of the Gaussian interference channel.

Lacking a single-letter expression for the capacity region of the interference channel, besides the fact that for many linear Gaussian channels the capacity region is achieved by Gaussian inputs, motivates us to further examine the limiting expression to obtain valid tight outer bounds for the Gaussian interference channel. To this end, we take another look at (13). Obviously, we can break down the objective function U_o into two or more functions and maximize them separately to get a valid outer bound on $\mu R_1 + R_2$ because $\max(x+y) \leq \max x + \max y$.

A. Tightening Outer Bounds: A Systematic Approach

Consider the multiletter outer bound given by (11). In what follows, by decomposing U_o in three different ways, we study three outer bounds to familiarize the reader with the process of improving the outer bound based on the multiletter expression. With this purpose, we decompose the objective function of the optimization problem (13), defined in (11), in different ways.

1) First Outer Bound: Let $U_o = U_{o11} + U_{o12} + U_{o13} - h(Z_2^n)$ in which

$$U_{o11} = \mu h(X_1^n + \sqrt{a}X_2^n + Z_1^n), \tag{15a}$$
$$U_{o12} = -\mu h(\sqrt{a}X_2^n + Z_1^n), \tag{15b}$$
$$U_{o13} = h(X_2^n + Z_2^n), \tag{15c}$$

and $h(Z_2^n) = \frac{n}{2}\log 2\pi e$ is a constant term. From the fact that a Gaussian distribution maximizes the differential entropy under a covariance constraint, and the power constraints in (3), it is easy to see that both U_{o11} and U_{o13} are maximized when $p(x_1^n) \sim \mathcal{N}(\mathbf{0}, P_1 I)$ and $p(x_2^n) \sim \mathcal{N}(\mathbf{0}, P_2 I)$. Besides, we can see that $-\mu h(\sqrt{a}X_2^n + Z_1^n) \leq -\mu h(\sqrt{a}X_2^n + Z_1^n | X_2^n) = -\mu h(Z_1^n) = \frac{\mu n}{2}\log 2\pi e$. Hence, from (11), we have

$$\begin{aligned}\mu R_1 + R_2 &\leq \frac{1}{n}U_o = \frac{1}{n}\Big[U_{o11} + U_{o12} + U_{o13} - \frac{n}{2}\log 2\pi e\Big] \\ &\leq \frac{1}{n}\big[U_{o11} + U_{o12}\big] + \gamma(P_2) \qquad (16)\\ &\leq \frac{1}{n}U_{o11} + \frac{\mu}{2}\log 2\pi e + \gamma(P_2) \\ &\leq \mu\gamma(P_1 + aP_2) + \gamma(P_2),\end{aligned}$$

where the last three inequalities are due to maximizing U_{o13}, U_{o12}, and U_{o11}, one by one. This gives a valid outer bound, but obviously not a tight one, except for $a = 0$, because from single-user upper bounds $R_1 \leq \gamma(P_1)$ and $R_2 \leq \gamma(P_2)$ it is clear that $\mu R_1 + R_2 \leq \mu\gamma(P_1) + \gamma(P_2)$.

2) Second Outer Bound: As proposed in [11]–[13], a better bound on $\mu R_1 + R_2$ is obtained if we maximize the second and third terms of U_o together. Formally, let

$$U_{o21} = \mu h(X_1^n + \sqrt{a}X_2^n + Z_1^n), \tag{17a}$$
$$U_{o22} = h(X_2^n + Z_2^n) - \mu h(\sqrt{a}X_2^n + Z_1^n). \tag{17b}$$

Again, it is clear that U_{o21} is maximized by Gaussian inputs. Using [14, Corollary 4] it can be checked that, for $a \leq 1$, U_{o22} is maximized by a Gaussian X_2^n, too. Thus, we can write

$$\begin{aligned}\mu R_1 + R_2 &\leq \frac{1}{n}U_o = \frac{1}{n}\Big[U_{o21} + U_{o22} - \frac{n}{2}\log 2\pi e\Big] \qquad (18)\\ &\leq \mu\gamma\big(P_1 + aP_2\big) + \frac{1}{n}U_{o22} - \frac{1}{2}\log 2\pi e \\ &\leq \begin{cases}\mu\gamma\big(\frac{P_1}{1+aP_2}\big) + \gamma(P_2), & \text{if } 0 \leq \mu \leq \frac{P_2+1/a}{P_2+1} \\ f_2(P_1, P_2, a, \mu), & \text{if } \frac{P_2+1/a}{P_2+1} < \mu < \frac{1}{a} \\ \mu\gamma(P_1 + aP_2), & \text{if } \mu \geq \frac{1}{a}\end{cases},\end{aligned}$$

in which the last step is proved in [11], and $f_2(P_1, P_2, a, \mu) = \frac{\mu}{2}\log\frac{(P_1+aP_2+1)(\mu-1)}{\mu(1-a)} + \frac{1}{2}\log\frac{\frac{1}{a}-1}{\mu-1}$. It is clear that the above arrangement results in a better bound on $\mu R_1 + R_2$. Furthermore, this outer bound is tight for $0 \leq \mu \leq \frac{P_2+1/a}{P_2+1}$ since it is achieved by treating interference as noise.

A closer examination of the outer bound (18), reveals that for $0 \leq \mu \leq \frac{P_2+1/a}{P_2+1}$, the Gaussian inputs $X_1^n \sim \mathcal{N}(\mathbf{0}, P_1 I)$ and $X_2^n \sim \mathcal{N}(\mathbf{0}, P_2 I)$ are optimal for U_{o21} and U_{o22}, and thus for U_o. This means that, for this range of μ, Gaussian inputs are sufficient to obtain the capacity region of the one-sided Gaussian interference channel when the limiting expression is used. Such a conclusion cannot be made for $\mu > \frac{P_2+1/a}{P_2+1}$.

3) Third Outer Bound: There is an even more efficient way to decompose U_o. We let $U_o = U_{o31} + U_{o32} - h(Z_2^n)$, where

$$U_{o31} = h(X_1^n + \sqrt{a}X_2^n + Z_1^n), \tag{19a}$$

$$U_{o32} = (\mu - 1)h(X_1^n + \sqrt{a}X_2^n + Z_1^n) - \mu h(\sqrt{a}X_2^n + Z_1^n) + h(X_2^n + Z_2^n), \tag{19b}$$

and maximize U_{o31} and U_{o31} individually. It is clear that U_{o31} is maximized by Gaussian inputs and we can write

$$\begin{aligned} \mu R_1 + R_2 &\leq \frac{1}{n}U_o = \frac{1}{n}\big[U_{o31} + U_{o32} - \frac{n}{2}\log 2\pi e\big] \\ &\leq \gamma\big(P_1 + aP_2\big) + \frac{1}{n}U_{o32}. \end{aligned} \tag{20}$$

From [15, Theorem 2], we conjecture that U_{o32} is also maximized by Gaussian inputs. This is true at least for $\mu \leq \frac{P_2+1/a}{P_2+1}$, in view of (18). If so, (20) will result in a tighter outer bound than (18) and (16).

B. A New Representation of the Limiting Expression

Considering the inner bound in Lemma 1, it is straightforward to show that $X_1^n \sim \mathcal{N}(\mathbf{0}, P_1 I)$ and $X_2^n \sim \mathcal{N}(\mathbf{0}, P_2 I)$ cannot be optimal for the outer bound on $\mu R_1 + R_2$, for all μ. In fact, such a bound is even worse than (14). On the other hand, from (20), we know that the above distributions maximizes $\mu R_1 + R_2$ for $\mu \leq \frac{P_2+1/a}{P_2+1}$. Then, it can be checked that X_2^n cannot not have a fixed covariance for all μ.

In other words, in (1) we cannot restrict X_2^n to have a single Gaussian distribution, if our goal is to get a valid outer bound or to achieve the capacity region of the Gaussian interference channel. Interestingly, however, the optimal X_2^n is Gaussian for a certain range of μ, as discussed above. As a result, some points on the capacity region are achieved by a Gaussian X_2^n. This specific example motivates us to advocate (9) as a new representation of (1), as it gives a better way to use the limiting expression to find the optimal distributions for that.

C. Comparison and Discussion

Section III proves that restriction to Gaussian inputs in the limiting expression for the capacity region of the Gaussian interference channel is insufficient to achieve the capacity region of this channel, in general. While this could diminish the importance of the limiting capacity expression, in Section IV-A it is shown that a judicious use of the multiletter expression can result in the capacity region of the interference channel, for a certain range of parameters.

It should be noted that the fact that restriction to Gaussian inputs in the multiletter expression is not sufficient to achieve the capacity region does not imply that Gaussian inputs are not capacity-achieving for the Gaussian interference channel. Neither does it mean that the multiletter capacity expression is useless. Conversely, the capacity regions thus far established for this channel indicates that by massaging it we can find a single-letter expression for the capacity region, for which Gaussian inputs are optimal.[3] Therefore, the multiletter capacity expression is useful and can be reduced to a single-letter capacity expression, at least in certain cases. How to find this single-letter expression from the multiletter expression is a fundamental challenge. We have tried to address this challenge in Section (IV-A).

V. Conclusion

We have shown that the capacity region of the Gaussian interference channel cannot, in general, be achieved by Gaussian inputs in the multiletter expression of the capacity region. This however does not imply that Gaussian inputs are not capacity-achieving for the Gaussian interference channel. More importantly, this work shows the potential of the multiletter expression in improving the outer bounds of the capacity regions of multi-user Gaussian channels. The key insight is to expand the multiletter expression and decompose it into two or more terms in a way that the overlap between their optimal solutions is maximized, in order to obtain or get closer to the optimal solution of the multiletter expression.

Appendix A

Consider the optimization problem in (13). Since for any Gaussian random n-vector $X^n \sim \mathcal{N}(\boldsymbol{\mu}, K)$ we have $h(X^n) = \frac{1}{2}\log\big((2\pi e)^n\,|K|\big)$, where $|K|$ denotes the determinant of K, by restricting its solution space within Gaussian distributions determined by (12), the problem (13) is equivalent to

$$\begin{aligned} U^G = \underset{K_1,\,K_2}{\text{maximize}} \quad & \Big[\frac{\mu}{2}\log|K_1 + aK_2 + I| - \frac{\mu}{2}\log|aK_2 + I| \\ & + \frac{1}{2}\log|K_2 + I|\Big] \\ \text{subject to} \quad & K_1 \succeq 0,\ \mathrm{tr}(K_1) \leq nP_1, \\ & K_2 \succeq 0,\ \mathrm{tr}(K_2) \leq nP_2, \end{aligned} \tag{21}$$

in which $\succeq$ denotes greater than or equal to in the positive semidefinite partial ordering between real symmetric matrices. We only study the nontrivial case of $\mu > 0$.

Next, since $K_{\boldsymbol{X}_1}$ and $K_{\boldsymbol{X}_2}$ are positive semidefinite, we can decompose them as $K_{\boldsymbol{X}_1} = U_1\Lambda_1 U_1^t$ and $K_{\boldsymbol{X}_2} = U_2\Lambda_2 U_2^t$, where U_1 and U_2 are unitary matrices and Λ_1 and Λ_2 are diagonal matrices with nonnegative entries. Let λ_{1i} and λ_{2i}, $i = 1, \ldots, n$, be the diagonal elements of Λ_1 and Λ_2, respectively. Note that these are the eigenvalues of $K_{\boldsymbol{X}_1}$ and $K_{\boldsymbol{X}_1}$, respectively. From the fact that $U_2 U_2^t = I$ and knowing that for any matrices A and B we have $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$ and $|AB + I| = |BA + I|$, we obtain $|K_{\boldsymbol{X}_2} + I| = |\Lambda_2 + I|$ and $|aK_{\boldsymbol{X}_2} + I| = |a\Lambda_2 + I|$. However, $|K_{\boldsymbol{X}_1} + aK_{\boldsymbol{X}_2} + I| = |\Lambda_1 + a\Lambda_2 + I|$ is not correct in general, and all we can have

[3]The very strong interference case [16] is the only case that the multiletter expression in (1) does not require any manipulation and $p(x_1^n) \sim \mathcal{N}(\mathbf{0}, P_1 I)$ and $p(x_2^n) \sim \mathcal{N}(\mathbf{0}, P_2 I)$ are optimal.

is

$$U^G = \max \quad \Big[\frac{1}{2}\log\left(|\Lambda_2 + I|\right) - \frac{\mu}{2}\log\left(|a\Lambda_2 + I|\right) + \frac{\mu}{2}\log\left(|U_1\Lambda_1U_1^t + aU_2\Lambda_2U_2^t + I|\right)\Big] \quad \text{s.t.} \quad \Lambda_1 \succeq 0,\ \operatorname{tr}(\Lambda_1) \le nP_1,\ \Lambda_2 \succeq 0,\ \operatorname{tr}(\Lambda_2) \le nP_2. \tag{22}$$

To further simplify the objective function, we can write it based on the eigenvalues of $K_{\boldsymbol{X}_1}$ and $K_{\boldsymbol{X}_2}$, using $|K_{\boldsymbol{X}_2}+I| = \prod_{i=1}^n(\lambda_{2i}+1)$ and $|aK_{\boldsymbol{X}_2}+I| = \prod_{i=1}^n(a\lambda_{2i}+1)$. In addition, without loss of generality, we assume $\lambda_{11} \ge \lambda_{12} \ge \ldots \ge \lambda_{1n}$ and $\lambda_{21} \le \lambda_{22} \le \ldots \le \lambda_{2n}$. With these assumptions, the last term can be upper bounded as [4]

$$|K_{\boldsymbol{X}_1} + aK_{\boldsymbol{X}_2} + I| \le \prod_{i=1}^n \big(\lambda_{1i} + a\lambda_{2i} + 1\big). \tag{23}$$

As a result, we will get the new optimization problem

$$\hat{U}^G = \max \quad \frac{1}{2}\sum_{i=1}^n \Big[\log\left(\lambda_{2i}+1\right) - \mu\log\left(a\lambda_{2i}+1\right) + \mu\log\left(\lambda_{1i} + a\lambda_{2i} + 1\right)\Big] \quad \text{s.t.} \quad \lambda_{1i} \ge 0,\ \lambda_{2i} \ge 0,\ \forall i,\ \sum_{i=1}^n \lambda_{1i} \le nP_1, \quad \sum_{i=1}^n \lambda_{2i} \le nP_2, \tag{24}$$

whose optimum value is an upper bound for U^G. That is,

$$U^G \le \hat{U}^G. \tag{25}$$

To solve this new optimization problem, we use Lagrange multipliers u_1, u_2, $\boldsymbol{v}_1 = \{v_{11}, \ldots, v_{1n}\}$ and $\boldsymbol{v}_2 = \{v_{21}, \ldots, v_{2n}\}$, and study the Lagrangian defined by

$$L = \frac{1}{2}\sum_{i=1}^n \big[\log\left(\lambda_{2i}+1\right) - \mu\log\left(a\lambda_{2i}+1\right) + \mu\log\left(\lambda_{1i} + a\lambda_{2i} + 1\right)\big] + u_1(nP_1 - \sum_{i=1}^n \lambda_{1i}) + u_2(nP_2 - \sum_{i=1}^n \lambda_{2i}) + \sum_{i=1}^n v_{1i}\lambda_{1i} + \sum_{i=1}^n v_{2i}\lambda_{2i}. \tag{26}$$

Because of the the inequality constraints, we need to apply the Karush-Kuhn-Tucker (KKT) conditions. The KKT stationary and complimentary slackness constraints are given by

$$\frac{1}{2\ln 2}\left(\frac{\mu}{\lambda_{1i} + a\lambda_{2i} + 1}\right) - u_1 + v_{1i} = 0, \tag{27a}$$

$$\frac{1}{2\ln 2}\left(\frac{1}{\lambda_{2i}+1} - \frac{a\mu}{a\lambda_{2i}+1} + \frac{a\mu}{\lambda_{1i} + a\lambda_{2i} + 1}\right) - u_2 + v_{2i} = 0, \tag{27b}$$

$$u_1\Big(nP_1 - \sum_{i=1}^n \lambda_{1i}\Big) = 0, \tag{27c}$$

$$u_2\Big(nP_2 - \sum_{i=1}^n \lambda_{2i}\Big) = 0, \tag{27d}$$

$$v_{1i}\lambda_{1i} = 0, \tag{27e}$$

$$v_{2i}\lambda_{2i} = 0, \tag{27f}$$

in which $i = 1, \ldots, n$, $u_1 \ge 0, u_2 \ge 0, v_{1i} \ge 0, v_{2i} \ge 0$, and obviously the constraints in (24) must hold too.

Next, for each of the complementary slackness condition we need to check whether its corresponding constraint is binding or the multiplier is zero. Therefore, the following cases arise:

1) Case I: $\lambda_{1i} = \lambda_{1j}, \forall\ i, j$. In words, in this case all λ_{1i}s are equal. Letting $\lambda_{1i} \triangleq \lambda_1$, it is easy to check that $0 \le \lambda_1 \le P_1$. Then, we considers the following two subcases separately:

- Subcase 1: $0 \le \lambda_{1i} < P_1$. If so, $\sum_{i=1}^n \lambda_{1i} < nP_1$ and from (27c) we must have $u_1 = 0$. Substituting this in (27a) we can see that $v_{1i} < 0$ which is contradicting. So, this case is impossible.
- Subcase 2: $\lambda_{1i} = P_1$. In this case, $v_{1i} = 0$ and from (27a) we will have $\frac{1}{2\ln 2}\left(\frac{\mu}{\lambda_1 + a\lambda_{2i} + 1}\right) = u_1, \forall\ i$. This implies that all λ_{2i}s are the same.

2) Case II: $\lambda_{1i} \ne \lambda_{1j}, \lambda_{1i} \ne 0, \lambda_{1j} \ne 0$. First, with $\lambda_{1i} \ne 0$ and $\lambda_{1j} \ne 0$, from (27e) we can see that v_{1i} and v_{1j} must be equal to zero. Then, from (27a), we must have

$$\frac{\mu}{\lambda_{1i} + a\lambda_{2i} + 1} = \frac{\mu}{\lambda_{1j} + a\lambda_{2j} + 1}. \tag{28}$$

Now we consider the following subcases:

- Subcase 1: $\lambda_{2i} = \lambda_{2j}$. In this case, from (28), we get $\lambda_{1i} = \lambda_{1j}$ which contradicts the conditions of Case I. Thus, this case impossible.
- Subcase 2: $\lambda_{2i} \ne \lambda_{2j}, \lambda_{2i} \ne 0, \lambda_{2j} \ne 0$. in this case, from (27f) we have $v_{2i} = 0$ and $v_{2j} = 0$. Then again, from (27b) have

$$\frac{1}{\lambda_{2i}+1} - \frac{a\mu}{a\lambda_{2i}+1} + \frac{a\mu}{\lambda_{1i} + a\lambda_{2i} + 1} = \frac{1}{\lambda_{2j}+1} - \frac{a\mu}{a\lambda_{2j}+1} + \frac{a\mu}{\lambda_{1j} + a\lambda_{2j} + 1}. \tag{29}$$

  Since both (28) and (29) must hold in this subcase, we conclude that $\lambda_{2i} = \lambda_{2j}$ which is contradicting.
- Subcase 3: $\lambda_{2i} \ne \lambda_{2j}, \lambda_{2i} = 0$. in this case, from (27b) we obtain

$$u_2 = \frac{1}{2\ln 2}\left(\frac{1}{\lambda_{2i}+1} - \frac{a\mu}{a\lambda_{2i}+1} + \frac{a\mu}{\lambda_{1i} + a\lambda_{2i} + 1}\right).$$

[4] Let $\alpha_1 \le \alpha_2 \le \ldots \le \alpha_n$ and $\beta_1 \le \beta_2 \le \ldots \le \beta_n$ be the eigenvalues of Hermitian $n \times n$ matrices A and B. Then, if $\alpha_1 + \beta_1 + \theta \ge 0$, where θ is a real number, we have [17]

$$|A + B + \theta I| \le \prod_{i=1}^n \left(\alpha_i + \beta_{n+1-i} + \theta\right).$$

Next, substituting this back into (27b) implies that $v_{2j} < 0$ unless $\lambda_{2j} = 0$. Again, this is contradicting. Considering the above subcases we conclude that Case II is not possible, i.e., we cannot have $\lambda_{1i} \neq \lambda_{1j}, \lambda_{1i} > 0, \lambda_{1j} > 0$. In words, this means that either all λ_{1i}s are equal or they can have two values and one of these values is 0.

3) Case III: $\lambda_{1i} \neq \lambda_{1j}, \lambda_{1i} = 0$. This case is possible as it does not violate any of the KKT conditions. However, the followings should be noted:

- There can only be two different values for the λ_{1i}s. This is obvious in light of Case II, as it proves that there cannot be two different non-zero values for λ_{1i}.
- There can only be two different values for λ_{2i}s. To prove this, we use the above fact that λ_{1i}s can only have two different values. More precisely, we show $\lambda_{1j} = \lambda_{1k}$ implies $\lambda_{2j} = \lambda_{2k}$, $\forall\, j, k \in \{1, \dots, n\}$. Let us consider the following subcases:
  Subcase 1: $\lambda_{1k} = \lambda_{1j} \neq 0$. With this, from (27e), we see that $v_{1j} = v_{1k} = 0$ which implies $\lambda_{2j} = \lambda_{2k}$, in view of (27a).
  Subcase 2: $\lambda_{1k} = \lambda_{1i} = 0$. In this case, we first show that $\lambda_{2j} \neq \lambda_{2k}$, $\lambda_{2j} \neq 0$, and $\lambda_{2k} \neq 0$ is contradicting. This is because $\lambda_{2j} \neq 0$ and $\lambda_{2k} \neq 0$ imply $v_{2i} = v_{2k} = 0$ and then from (27b) we will get $\lambda_{2j} = \lambda_{2k}$. Similarly, we can show that $\lambda_{2j} = 0$ and $\lambda_{2k} \neq 0$ is impossible because from (27b) we can obtain

  $$\frac{1}{2\ln 2} + v_{2j} = \frac{1}{2\ln 2}\frac{1}{\lambda_{2k}+1}$$

  which is contradicting as it implies $v_{2j} < 0$.
  Here, we add that in Case III there exist at least one $\lambda_{1i} = 0$ and $\lambda_{1j} \neq 0$. Hence, from (27b) we will have

  $$\frac{1}{2\ln 2}\frac{1}{\lambda_{2i}+1} - u_2 + v_{2i} = 0. \tag{30}$$

  Now, it is easy to see that $u_2 > 0$. Also, from (27a) it is clear that $u_1 > 0$ for any $\mu > 0$. Thus, using (27c) and (27d) we obtain

  $$\sum_{i=1}^{n} \lambda_{1i} = nP_1, \tag{31a}$$
  $$\sum_{i=1}^{n} \lambda_{2i} = nP_2. \tag{31b}$$

It can be checked that Case I-III are exhaustive, i.e., they include all plausible combinations of λ_{1i} and λ_{2i}. We next show that the two acceptable cases can be unified. By letting $1 \leq m \leq n$, and remembering that $\lambda_{11} \geq \lambda_{12} \geq \dots \geq \lambda_{1n}$, we can write

$$\lambda_{1i} = \begin{cases} \frac{n}{m}P_1, & \text{for} \quad 1 \leq i \leq m \\ 0, & \text{for} \quad m < i \leq n \end{cases}, \tag{32}$$

where $m = n$ corresponds to the subcase 2 of Case I, and $m < n$ corresponds to Case III. Similarly, we can write

$$\lambda_{2i} = \begin{cases} P_2', & \text{for} \quad 1 \leq i \leq m \\ \frac{nP_2 - mP_2'}{n-m}, & \text{for} \quad m < i \leq n \end{cases}, \tag{33}$$

in which $0 \leq P_2' \leq P_2$. The upper bound on P_2' is due to the assumption that $\lambda_{21} \leq \lambda_{22} \leq \dots \leq \lambda_{2n}$. Note that, with these choices of λ_{1i} and λ_{2i} all conditions of the acceptable cases, including (31a) and (31b), are satisfied.

We are now at the position to evaluate the optimal value of the $\hat{U}^G$ in (24), which is given by

$$\begin{aligned}
\hat{U}^G &= \sum_{i=1}^{n} \Big[\gamma(\lambda_{2i}) + \mu\gamma\Big(\frac{\lambda_{1i}}{1+a\lambda_{2i}}\Big)\Big] \\
&= m\Big[\gamma(P_2') + \mu\gamma\Big(\frac{\frac{n}{m}P_1}{1+aP_2'}\Big)\Big] + (n-m)\gamma\Big(\frac{nP_2 - mP_2'}{n-m}\Big) \\
&= n\xi\Big[\gamma(P_2') + \mu\gamma\Big(\frac{\frac{P_1}{\xi}}{1+aP_2'}\Big)\Big] + n\bar{\xi}\gamma\Big(\frac{P_2 - \xi P_2'}{\bar{\xi}}\Big),
\end{aligned}$$

in which $\xi \triangleq \frac{m}{n}$, $0 \leq \xi \leq 1$, $\bar{\xi} = 1 - \xi$, and $0 \leq P_2' \leq P_2$. Finally, recalling that $\hat{U}^G$ is an upper bound for U^G, the proof of outer bound is completed.

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 611–644, 1961.

[2] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd IEEE International Symposium on Information Theory*, pp. 23–52, Sept. 1971.

[3] E. C. Van Der Meulen, "The discrete memoryless channel with two senders and one receiver," in *Proc. 2nd IEEE International Symposium on Information Theory (ISIT)*, pp. 95–102, Sept. 1971.

[4] E. C. van der Meulen, "Some reflections on the interference channel," in *Communications and Cryptography*, pp. 409–421, Springer, 1994.

[5] M. Mandell and R. McEliece, "Some properties of memoryless multi-terminal interference channels," in *Proc. IEEE International Symposium on Information Theory*, pp. 212–212, 1991.

[6] R. S. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 609–612, 1993.

[7] A. Carleial, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, pp. 60–70, January 1978.

[8] M. H. Costa and A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Transactions on Information Theory*, vol. 33, no. 5, pp. 710–711, 1987.

[9] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[10] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, pp. 49–60, January 1981.

[11] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 620–643, 2009.

[12] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum–rate capacity for Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 689–699, 2009.

[13] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3032–3050, 2009.

[14] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Transactions on Information Theory*, vol. 53, no. 5, pp. 1839–1851, 2007.

[15] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "A vector generalization of Costa's entropy-power inequality with applications," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1865–1879, 2010.

[16] A. Carleial, "A case where interference does not reduce capacity," *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 569–570, 1975.

[17] M. Fiedler, "Bounds for the determinant of the sum of hermitian matrices," *Proceedings of the American Mathematical Society*, vol. 30, no. 1, pp. 27–31, 1971.