The Capacity Region of the One-Sided Gaussian Interference Channel

Mojtaba Vaezi and H. Vincent Poor

Department of Electrical Engineering
Princeton University, Princeton, NJ 08544, USA
Email: mvaezi@princeton.edu, poor@princeton.edu

Abstract—The capacity region of the one-sided Gaussian interference channel is established in the weak interference regime. To characterize this region, a new representation of the Han-Kobayashi inner bound for the one-sided Gaussian interference channel is first given. Next, a new outer bound on the capacity region of this channel is introduced which is tight in the weak interference regime. This is the first capacity region for any variant of the interference channel in the weak interference regime.

Index Terms—Capacity region, one-sided interference channel, Z-interference channel, weighted sum-rate.

I. INTRODUCTION

The two-user interference channel is a two-transmitter two-receiver network, in which each transmitter has a message for its respective receiver [1]. The one-sided interference channel, also known as the Z-interference channel, depicted in Fig. 1 is a special case of the interference channel in which only one of the receivers suffers from interference. Despite extensive studies, the problem of characterizing the capacity region of this channel has been open for 40 years. Our understanding of the capacity of this channel has been limited to the strong interference case, i.e., when the gain of the interference link is no less than one [2]–[4]. In the weak interference case, when that gain is less than one, only the sum capacity of this channel is known [5].

The Han-Kobayashi (HK) inner-bound [6] is the best known achievable region for the interference channel. It divides the information of each user into two parts: private and common information. The former is to be decoded only at the intended receiver whereas the latter can be decoded at both receivers. The rationale behind this coding scheme is to decode part of interference (the common information) and treat the remaining as noise. The HK region is, however, complicated and it has not been fully characterized, because the optimum input distributions are not known for it. Commonly, a subset of the HK region with Gaussian input distributions is used to represent the HK region for the Gaussian channel; see for example [7]–[9]. Flexibility in the split of each user’s transmission power to the common/private portions of information makes the HK scheme very strong, yet complicated. As such, the optimal HK strategy is not well-understood in general, and for the one-sided interference channel, in particular.

The problem of finding the capacity region of interference networks is hard mainly because it is not straightforward to come up with tight capacity upper bounds. Several techniques have been developed for upper bounding the capacity region of the interference channel, among them are genie-aided and more capable receivers, where some interfering signals are removed. A well-known example of the latter case is to use the capacity region of the one-sided interference channel, or in general any outer bound on its capacity, as an outer bound for the capacity region of the interference channel.

In this paper, we completely characterize the capacity region of the one-sided interference channel. To this end, we first find a new simple representation of the HK region for the one-sided interference channel, in which the optimal power split between the private/common information to achieve the border of the region is identified. The new representation is based on the constraints on the weighted sum-rate and explicitly shows the optimal power split as a function of the weight.

Next, we develop a tight outer bound on the capacity region of the one-sided interference channel. To prove the outer bound, we develop a new extremal inequality that can be useful in other network information theory problems.

Observing that the capacity region of the one-sided interference channel can be used as an outer bound for the capacity...
region of the interference channel, we can use the introduced capacity region to find a better bound on the capacity region of the interference channel.

The paper is organized as follows. We review the channel model and existing results in Section II. We introduce a new representation of the HK inner bound for the one-sided Gaussian interference channel in Section III. We develop a new outer bound in Section IV and introduce the capacity result in Section V. We conclude the paper in Section VI.

Regarding notation, we will use uppercase boldface letters (e.g., $X$) for random vectors, uppercase letters (e.g., $X$) for random variables and matrices, lowercase boldface letters (e.g., $x$) for deterministic vectors, lowercase letters (e.g., $x$) for scalars, and $\gamma(x)$ as an abbreviation for $\frac{1}{2} \log_2 (1 + x)$. We use $h(\cdot)$ to denote the differential entropy and $I(\cdot; \cdot)$ to denote mutual information. All vectors are column vectors.

II. CHANNEL MODEL AND PRELIMINARIES

In this section, we describe the model we use in this paper and state the previously established capacity results for the one-sided Gaussian interference channel.

A. Channel Model

The two-user Gaussian interference channel is composed of two transmitter-receiver pairs in which each transmitter communicates with its respective receiver while interfering with the other receiver. Without loss of generality, we can consider the standard form of the Gaussian interference channel [2], in which the channel is expressed, for a single channel use, by

$$Y_1 = X_1 + \sqrt{a}X_2 + Z_1,$$  \hspace{1cm} (1a)
$$Y_2 = \sqrt{b}X_1 + X_2 + Z_2,$$  \hspace{1cm} (1b)

where $a$ and $b$ are two non-negative real numbers representing the crossover gains; $X_i$, $Y_i$, and $Z_i$, for $i = 1, 2$, represent the transmitted signal, received signal, and the channel noise, respectively; and, $Z_1$ and $Z_2$ are independent Gaussian random variables with zero means and unit variances. Let $w_1$ and $w_2$ be two independent messages which are uniformly distributed over $\mathcal{W}_1 = \{1, \cdots, 2^nR_1\}$ and $\mathcal{W}_2 = \{1, \cdots, 2^nR_2\}$, respectively. Transmitter $i$ wishes to transmit message $w_i$ to receiver $i$ in $n$ channel uses at rate $R_i$, and $X_i$ is subject to an average power constraint $P_i$, i.e.,

$$\frac{1}{n} \sum_{j=1}^{n} \|X_{ij}\|^2 \leq P_i, \hspace{1cm} i = 1, 2. \hspace{1cm} (2)$$

The capacity region of this channel is defined as the closure of the set of rate pairs $(R_1, R_2)$ for which each receiver is able to decode its own message with arbitrarily small probability of error.

The one-sided Gaussian interference channel is a two-user Gaussian interference channel in which either $a$ or $b$ is equal to zero. Since the analysis of the capacity results in either case is the same, without loss of generality we assume $b = 0$.

This channel is represented in Fig. I. With this, the channel model described by (1) simplifies to

$$Y_1 = X_1 + \sqrt{a}X_2 + Z_1,$$  \hspace{1cm} (3a)
$$Y_2 = X_2 + Z_2.$$  \hspace{1cm} (3b)

Depending on the value of the gain $a$ of the interfering link, the above channel is classified as either weak or strong one-sided interference channel. Specifically, the channel is in the weak interference regime if $a < 1$ and the strong interference regime if $a \geq 1$. In the rest of this paper, we use the above channel model. Since we focus on the Gaussian channel only, we may simply use the one-sided interference channel to refer to the above channel.

B. Previous Results

The capacity region of the one-sided interference channel is not fully known, to date. There are however certain capacity results which are summarized in the following propositions:

**Proposition 1.** [3] [4] The capacity region of the one-sided interference channel in the strong interference regime $(a \geq 1)$ is the set of $(R_1, R_2)$ satisfying

$$R_1 \leq \gamma(P_1), \hspace{1cm} (4a)$$
$$R_2 \leq \gamma(P_2), \hspace{1cm} (4b)$$
$$R_1 + R_2 \leq \gamma(P_1 + aP_2). \hspace{1cm} (4c)$$

Dating back to the early 1980s, the above result is the only case for which the capacity region is known for the one-sided interference channel. The above region is achieved by decoding and removing the interference.

The capacity region is not characterized in the weak interference regime $(a < 1)$. However, as another notable result, we know the sum-capacity of this channel in the weak interference regime, as stated below.

**Proposition 2.** [5] The sum-capacity of the one-sided interference channel in the weak interference regime is the set of rate pairs $(R_1, R_2)$ satisfying

$$R_1 + R_2 \leq \gamma\left(\frac{P_1}{1 + aP_2}\right) + \gamma(P_2). \hspace{1cm} (5)$$

The above proposition reveals that the maximum sum-rate for this channel is achieved by treating interference as noise.

The following proposition shows that such a simple decoding is also optimum for a certain range of weighted sum-rates.

**Proposition 3.** [3] [9] For $1 \leq \mu \leq \frac{P_1 + 1/a}{P_1 + 1+a}$ and $a \leq 1$, the capacity region of the one-sided interference channel is outer bounded by

$$\mu R_1 + R_2 \leq \mu \gamma\left(\frac{P_1}{1 + aP_2}\right) + \gamma(P_2). \hspace{1cm} (6)$$

The above result indicates that treating interference as noise is optimal for a certain range of weighted sum-rates. As we
prove in this work, treating interference as noise is optimal for a larger range of weighted sum-rates.

III. NEW REPRESENTATION OF THE HAN-KOBAYASHI REGION FOR THE ONE-SIDED INTERFERENCE CHANNEL

The best known achievable scheme for the two-user Gaussian interference channel, including the one-sided interference channel, is due to Han and Kobayashi [6]. This scheme splits the information of both users into private and common parts. The former is meant to be decoded only at its respective receiver while the common information can be decoded by both receivers. Flexibility of power allocation to the common and private portions of information besides time-sharing between such splits makes the HK strategy very strong, yet difficult to be optimized and fully understood. The fact that the optimal input distributions are not known for the HK strategy makes the matter even more complicated.

We consider a simple version of the HK scheme, in which time-sharing is not used and the input distributions are fixed to be Gaussian. With these assumptions, the HK inner bound for the one-sided interference channel simplifies as follows.

**Proposition 4.** The Han-Kobayashi achievable region for the one-sided interference channel with Gaussian inputs is given by the set of \((R_1, R_2)\) such that

\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right),
\]

\[
R_2 \leq \gamma (P_2),
\]

\[
R_1 + R_2 \leq \gamma \left( \frac{P_1 + a\beta P_2}{1 + a\beta P_2} \right) + \gamma (\beta P_2),
\]

for some \(\beta \in [0, 1]\) and \(\bar{\beta} = 1 - \beta\).

In the above rate region, \(\beta\) controls the power allocation for the private and common parts of information for the second user. Specifically, \(\beta P_2\) and \(\bar{\beta} P_2\) represent the power allocated to the common and private information, respectively. User 1 is not transmitting common information because there is no link from that to the second receiver. Hence, there is no power split for that user. The above representation of the HK region can be further simplified [8, Lemma 7], as below.

**Proposition 5.** The HK achievable region for the one-sided interference channel with Gaussian inputs is further simplified to the set of \((R_1, R_2)\) such that

\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right),
\]

\[
R_2 \leq \gamma \left( \frac{a\beta P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (\beta P_2).
\]

Figure 2 depicts the achievable points corresponding to \(\beta = 0\) (point B) and \(\beta = 1\) (point A). While the corner point A is known to be the sum-capacity of this channel, it is not known whether B is the other corner point of the capacity region or not. The above region is convex, and it is easy to prove that time-sharing between A and B is inside the above HK region.

![Figure 2. The Han-Kobayashi achievable points A (\(\beta = 1\)) and B (\(\beta = 0\)) with the corresponding rate region (black solid region), the sum-capacity outer bound (blue dotted lines), and the weighted sum-rate outer bound (red dashed lines) for the one-sided interference channel. The point A is achieved by treating the interference as noise whereas the point B is achieved by decoding and removing the interference. In particular, any point on the line segment AB then can be achieved by time-sharing. Note that the line segment AB is achieved by naïve time-sharing (time-sharing with fixed power). Time sharing with power control achieves a larger region [5]. As we show in this work, the HK region strictly includes this line segment.](image-url)

We are interested in finding the optimal value of \(\beta\) such that the weighted sum-rate \(\mu R_1 + R_2\), also known as the \(\mu\)-sum-rate, is maximized for any \(\mu \geq 1\)\(^2\). To this end, using (8a)-(8b), we know that the \(\mu\)-sum rate is upper bounded by

\[
R_{\mu\text{-sum}} \triangleq \mu R_1 + R_2 \leq \mu \gamma \left( \frac{P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{a\beta P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (\beta P_2).
\]

To determine the optimal values of \(\beta\) that maximize the \(R_{\mu\text{-sum}}\), we find the critical point of the bound by evaluating the first-order partial derivative of the right-hand side of (9) with respect to \(\beta\) and setting it to zero, which proceeds as

\[
\frac{\partial R_{\mu\text{-sum}}}{\partial \beta} = 0 \Rightarrow \mu = \frac{1 + a\beta P_2}{1 + \beta P_2} \frac{1 + P_1 - a P_1}{a P_1}.
\]

\(^2\) We focus on \(\mu \geq 1\) because from Proposition 4 it is clear that the capacity region is inside the sum capacity (\(\mu = 1\)). This is visualized in Fig. 2.
Let us define
\[
\mu^* = \frac{1 + \beta P_2}{1 + \beta P_2} + \frac{1}{P_1 - a}.
\] (12)
For \(a < 1\), it is straightforward to see that the maximum and minimum values of \(\mu^*\) respectively correspond to \(\beta = 0\) and \(\beta = 1\), and are given by
\[
\mu_0^* \triangleq \frac{1 + P_1 - a}{a P_1}, \quad (13a)
\]
\[
\mu_1^* \triangleq \frac{1 + a P_2}{1 + P_2} \mu_0^*. \quad (13b)
\]
Now, one can check that the optimal value of \(\beta\) to maximize (10) is given by
\[
\beta^* = \begin{cases} 
1, & \text{if } 1 \leq \mu \leq \mu_1^*, \\
\frac{\mu_1^* - \mu}{\mu - a \mu_0^*}, & \text{if } \mu_1^* < \mu < \mu_0^*, \\
0, & \text{if } \mu \geq \mu_0^*.
\end{cases} \quad (14)
\]
Consequently, we obtain
\[
\mu R_1 + R_2 \leq \begin{cases} 
\gamma \left( \frac{P_1}{1 + a \mu_1^*} \right) + \gamma (P_2), \quad \text{if } 1 \leq \mu \leq \mu_1^*, \\
f(P_1, P_2, a, \mu), \quad \text{if } \mu_1^* < \mu < \mu_0^*, \\
\gamma (P_1) + \gamma \left( \frac{a P_2}{1 + P_1 + a \mu_0^*} \right), \quad \text{if } \mu \geq \mu_0^*
\end{cases} \quad (15)
\]
in which
\[
f(P_1, P_2, a, \mu) = \mu \gamma \left( \frac{P_1}{1 + a \mu - a \mu_0^*} \right) + \frac{a P_2 - a \mu_1^*}{1 + P_1 + a \mu_0^*} \gamma \left( \frac{\mu_1^* - \mu}{\mu - a \mu_0^*} \right). \quad (16)
\]
IV. A Tight Outer Bound

In consideration of the new representation of the HK region in Lemma 1 a natural question is whether or not we can prove a similar outer bound on the \(\mu\)-sum rate of the one-sided interference channel. To this end, for any codebook of block length \(n\), we make use of the following bound for the \(\mu\)-sum rate, which is based on the Fano’s inequality [10],
\[
n(\mu R_1 + R_2) \leq \mu I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n) + ne_n
\]
\[
= \mu h(Y_1^n) - \mu h(Y_1^n | X_1^n) + h(Y_2^n)
- h(Y_2^n | X_2^n) + ne_n
- \mu h(X_1^n + \sqrt{a} X_2^n + Z_1^n) - \mu h(\sqrt{a} X_2^n + Z_1^n)
+ h(X_2^n + Z_2^n) - h(Z_2^n) + ne_n, \quad (18)
\]
in which \(\epsilon_n \to 0\) as \(n \to \infty\).
To maximize this bound, we need to determine both the input distributions that maximize (18) and the optimum power at each transmitter. We observe that in the previous works, e.g., [7]–[9], the term \(\mu h(X_1^n + \sqrt{a} X_2^n + Z_1^n)\) has been maximized separately from \(h(X_2^n + Z_2^n) - h(\sqrt{a} X_2^n + Z_1^n)\). With this arrangement, the latter terms are in the form of \(h(X + Z_2) - \mu h(X + Z_1)\) which is proved in [11] Theorem 1, to have a Gaussian \(X\) as an optimal solution, for \(\mu \geq 1\). Furthermore, in [8] and [9], it is shown that for \(0 \leq \mu \leq P_2 + 1/\alpha\) the optimal power of \(X_2^n\) is \(n P_2\), and thus the outer bound can be represented as in Proposition 3.
Although such an optimization provides a valid outer bound, it may result in a looser one when compared to the case in which those terms are maximized altogether, simply because \(\max x + \max y \geq \max (x + y)\). An important issue to be addressed is to find the optimal input distributions. Put is simply, can Gaussian inputs be an optimal solution in maximizing all three terms in (18)? If not, can we arrange those terms in a way that this arrangement leads to a tighter outer bound? These are the questions we seek to address in the following. In pursuit of this, we formulate a new optimization problem in the following.
A. New Formulation for the Outer Bound
It can be proved that Gaussian inputs cannot be optimal for (18) [12]. Nevertheless, we can write it in the following form
\[
n(\mu R_1 + R_2) \leq \mu I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n) + ne_n
\]
\[
= \mu h(Y_1^n) - \mu h(Y_1^n | X_1^n) + h(Y_2^n)
- h(Y_2^n | X_2^n) + ne_n
- \mu h(X_1^n + \sqrt{a} X_2^n + Z_1^n) - \mu h(\sqrt{a} X_2^n + Z_1^n)
+ h(X_2^n + Z_2^n) - h(Z_2^n) + ne_n
\leq (\mu - 1) h(X_1^n + \sqrt{a} X_2^n + Z_1^n)
+ \frac{n}{2} \log \left[ 2 \epsilon (P_1 + a P_2 + 1) \right] + h(X_2^n + Z_2^n)
- \mu h(\sqrt{a} X_2^n + Z_1^n) - \frac{n}{2} \log 2 \pi e + ne_n
= (\mu - 1) h(X_1^n + \sqrt{a} X_2^n + Z_1^n) + h(X_2^n + Z_2^n)
- \mu h(\sqrt{a} X_2^n + Z_1^n) + n \gamma (P_1 + a P_2) + ne_n, \quad (19)
\]
Lemma 2. \(W_o \triangleq h(X_2 + Z_2) - h(X_1 + Z_1)\) altogether. To begin with, let us define
\[
W_o \triangleq h(X_2 + Z_2) - \mu h(X_2 + Z_1) + (\mu - 1) h(X_1 + \sqrt{\mu} X_2 + Z_1),
\]
in which \(X_1\) and \(X_2\) are independent random \(n\)-vectors\(^3\). Then, maximizing (19) is equivalent to finding \(W_o\) in
\[
W = \max_{p(x)} W_o \quad \mathrm{subject \ to} \quad \text{tr}(K_{X_1}) \leq np_1, \quad \text{tr}(K_{X_2}) \leq np_2,
\]
where \(K_{X_1}\) and \(K_{X_2}\) are the covariance matrices of \(X_1\) and \(X_1\), respectively. The two constraints are basically the power constraints defined in (2).

The first step in solving (21) is to determine the optimal distributions for \(X_1\) and \(X_2\). Thereafter, we need to determine the optimum powers. These are the subjects of the following subsections.

B. Optimal Input Distributions

In this section, we prove that jointly Gaussian \((X_1, X_2)\) maximize \(W_o\) in (20). With this purpose, we first consider a more general version of the optimization problem in (21), i.e.,
\[
\max_{p(x)} W_o \quad \mathrm{subject \ to} \quad 0 \preceq K_X \preceq S,
\]
in which \(K_X\) is the covariance matrix of \(X = [X_1, X_2]\)^4, and \(S\) is a positive semidefinite matrix. The covariance constraint subsumes the average power constraint and thus the optimization problem in (22) is more general than that in (21). We prove that a Gaussian \(X\) is an optimal solution for (22), by setting the ground in the following lemmas.

Let \(Z_0, Z_1,\) and \(Z_2\) be Gaussian \(n\)-vectors with strictly positive definite covariance matrices \(K_{Z_1}, K_{Z_2}\), and \(K_{Z_2}\), respectively, and let \(X\) be a random \(n\)-vector independent of \(Z_0, Z_1,\) and \(Z_2\). Denote the covariance matrix of \(X\) by \(K_X\), and define
\[
\tilde{W}_o = \sum_{k=1}^{2} \mu_k h(X + Z_k) - h(X + Z_0),
\]
where \(\sum_{k=1}^{2} \mu_k = 1\) and \(\mu_k \geq 0\) for \(k = 1, 2\). Now, consider the following optimization problem:
\[
\max_{p(x)} \tilde{W}_o \quad \mathrm{subject \ to} \quad 0 \preceq K_X \preceq S.
\]

Theorem 3. Let \(v_0, v_1\) and \(v_2\) be deterministic, nonzero vectors of length two, \(Z_0, Z_1,\) and \(Z_2\) be Gaussian noises with positive variances, \(X = [X_1, X_2]\) be independent of \(Z_0, Z_1,\) and \(Z_2\), and \(S\) be a positive semidefinite matrix. A Gaussian \(X\) is an optimal solution of the optimization problem
\[
\max_{p(x)} \sum_{k=0}^{2} \mu_k h(v_k^t X + Z_k) \quad \mathrm{subject \ to} \quad K_X \preceq S,
\]
where \(\sum_{k=0}^{2} \mu_k = 0\) and \(\mu_k \geq 0\) for \(k > 0\).

Proof: From Lemma 2 we know that a Gaussian \(X\) is an optimal solution for maximizing \(\sum_{k=0}^{2} \mu_k h(X + Z_k) - h(X + Z_0)\). Since multiplying the objective function by a positive scalar \(\mu_0\) does not change the optimal solution, it is clear that the optimal input distribution of \(\sum_{k=0}^{2} \mu_k h(X + Z_k)\) is the same for any \(\mu_0 + \mu_1 + \mu_2 = 0\) and \(\mu_1 \geq 0, \mu_2 \geq 0\).

From the above lemma, we can see that the normal distribution \(N(0, K_X)\) maximizes \(\sum_{k=0}^{2} \mu_k h(X + Z_k)\), where we assume that \(K_X\) is the optimal covariance matrix. Note that translation does not change the differential entropy, i.e., \(h(X + c) = h(X)\) for any constant vector \(c\); hence, the mean could be any constant without affecting the optimal solution, so we take it 0. From this, we can write
\[
N(0, K_X) = \arg\max_{p(x)} \sum_{k=0}^{2} \mu_k h(X + Z_k)
\]
\[
= \arg\max_{p(x)} \sum_{k=0}^{2} \mu_k I(X; X + Z_k)
\]
\[
= \arg\max_{p(x)} \sum_{k=0}^{2} \mu_k I(V_k X; V_k X + V_k Z_k)
\]
\[
= \arg\max_{p(x)} \sum_{k=0}^{2} \mu_k I(v_k^t X; v_k^t X + v_k^t Z_k)
\]
\[
= \arg\max_{p(x)} \sum_{k=0}^{2} \mu_k h(v_k^t X + Z_k),
\]
in which

(a) follows because adding a constant term to the objective function has no effect on the optimal solution; thus, we can add \(-\sum_{k=0}^{2} \mu_k h(Z_k)\); (b) is due to the identity \(h(AX) = h(X) + \log |\det(\mathbf{A})|\) for any nonsingular matrix \(\mathbf{A}\); (c) follows by considering a specific structure for the covariance matrix of \(Z_1\) and \(V_k\), for a given \(v_k\) and \(k = 1, 2, 3\), as described in details in the following:

\(^3\)To simplify the notation, \(X\) is used instead of \(X^n\); they both represent a random \(n\)-vector.

\(^4\) We have however removed the unnecessary condition \(K_{Z_1} \preceq K_{Z_2}\).
Let, for $k = 1, 2, 3$,
\[ V_k = \begin{bmatrix} \mathbf{v}_k^t & \mathbf{v}_k \end{bmatrix}, \]
where $\mathbf{v}_k = [v_{k1} \ v_{k2}]$ is the given nonzero vector of length two and $\mathbf{v}_k^t = [v_{k1}^t \ v_{k2}^t]$ is such that $\frac{\mathbf{v}_k^t}{\|v_k\|}$ is an orthonormal matrix. Therefore, we must have $v_{k1}v_{k1} + v_{k2}v_{k2} = 0$ and $\|v_k\|^2 = \|v_k\|^2$. Then, $\mathbf{v}_k = [-v_{k2} \ v_{k1}]$ is an answer. Also, let
\[ K\mathbf{Z}_k = U_k \Sigma_k U_k^t \]
in which $U_k = \frac{\mathbf{v}_k^t}{\|v_k\|}$ and
\[ \Sigma_k = \begin{bmatrix} \sigma_{k1}^2 & 0 \\ 0 & \sigma_{k2}^2 \end{bmatrix}, \]
where
\[
\sigma_{k1}^2 = \frac{\sigma_{Z_k}^2}{\|v_k\|^2},
\]
and $\sigma_{k2}^2$ is an arbitrary positive numbers.\(^5\)
To simplify the notation, let us define $\mathbf{X}_k = V_k\mathbf{X}$ and $\mathbf{Z}_k = V_k\mathbf{Z}_k$. Thus,
\[ \tilde{\mathbf{X}}_k = V_k\mathbf{X} = \begin{bmatrix} \mathbf{X}_{k1} \\ \mathbf{X}_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_k^t \mathbf{X} \\ \mathbf{v}_k \mathbf{X} \end{bmatrix}, \]
\[ \tilde{\mathbf{Z}}_k = V_k\mathbf{Z}_k = \begin{bmatrix} \mathbf{Z}_{k1} \\ \mathbf{Z}_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_k^t \mathbf{Z}_k \\ \mathbf{v}_k \mathbf{Z}_k \end{bmatrix}. \]
Now, we show that $\tilde{Z}_{k1}$ and $\tilde{Z}_{k2}$ are independent. First, note that $\tilde{Z}_{k1}$ and $\tilde{Z}_{k2}$ are jointly Gaussian because they are linear combinations of two jointly Gaussian random variables. Hence, to prove $\tilde{Z}_{k1}$ and $\tilde{Z}_{k2}$ are independent, it suffices to show that they are uncorrelated. The latter is correct because
\[
E\{\tilde{Z}_{k1}\tilde{Z}_{k2}\} = E\{V_k\mathbf{Z}_k\mathbf{Z}_k^t V_k^t\} = V_k E\mathbf{Z}_k\mathbf{Z}_k^t V_k^t = \|v_k\|^2 \Sigma_k
\]
in which the last step is achieved by replacing $K\mathbf{Z}_k$ from (28) and $\mathbf{v}_k^t = \|v_k\|U_k$ where $U_k$ is orthonormal. Since $\Sigma_k$ is orthogonal, it is clear that $E\{\tilde{Z}_{k1}\tilde{Z}_{k2}^t\} = 0$. Moreover, it is seen that $\sigma_{Z_k1}^2 = \|v_k\|^2 \sigma_{Z_k1}^2 = \sigma_{Z_k2}^2$ and $\sigma_{Z_k2}^2 = \|v_k\|^2 \sigma_{Z_k2}^2$.

Next, from (11) Lemma 13, we obtain
\[
\lim_{\sigma_{Z_k2}^{-2} \to \infty} I(\tilde{\mathbf{X}}_k; \mathbf{X}_k + \tilde{\mathbf{Z}}_k) = I(\mathbf{X}_{k1}; \mathbf{X}_k + \tilde{Z}_{k1}). \tag{33}
\]

\(^5\) It should be highlighted that any covariance matrix $K$ can be decomposed as $K = U\Sigma U^t$, where $U$ and $\Sigma$ are orthonormal and diagonal matrices, respectively. In addition, since in (26) $Z_k$ is an arbitrary Gaussian vector, we take it to have a covariance matrix $K_{Z_k} = U_k \Sigma_k U_k^t$, as described above.

\(^6\) For completeness, we restate (11) Lemma 13 in the following proposition:

Proposition 6. For any random vector $\mathbf{X} = [X_1 \ X_2]^t$, with finite variances, independent of $\mathbf{Z} = [Z_1 \ Z_2]^t$, where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are two independent Gaussian random variables with $\sigma_{Z_1}^2$ and $\sigma_{Z_2}^2$, we have
\[
\lim_{\sigma_{Z_2}^{-2} \to \infty} I(\mathbf{X}; \mathbf{X} + \mathbf{Z}) = I(X_1; X_1 + Z_1). \tag{34}
\]
That is, for non-zero vectors $v_k$, we can choose $V_k$ and $K\mathbf{Z}_k$, as in (27)-(30), such that in the limit when $\sigma_{Z_2}^{-2} \to \infty$, $I(V_k\mathbf{X}; V_k\mathbf{X} + V_k\mathbf{Z}_k)$ approaches $I(v_k^t \mathbf{X}; v_k^t \mathbf{X} + v_k^t \mathbf{Z}_k)$.

(d) follows by subtracting the constant $\sum_{k=0}^2 \mu_k h(v_k^t \mathbf{Z}_k)$ and noting the statistical equivalence of $v_k^t \mathbf{Z}_k$ and $Z_k$. Recall from (32) that $\tilde{Z}_{k1} = v_k^t \mathbf{Z}_k$ has the same variance as $Z_k$, for any $k = 0, 1, 2$.

Lemma 4. Let $v_0$, $v_1$, and $v_2$ be deterministic, nonzero vectors of length two, $\mathbf{Z}_0$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$ be Gaussian $n$-vectors with positive definite covariance matrices, $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2]^t$, with random $n$-vectors $\mathbf{X}_1$, $\mathbf{X}_2$, be independent of $\mathbf{Z}_0$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$, and $S$ be a positive semidefinite matrix. A Gaussian $\mathbf{X}$ is an optimal solution of the optimization problem
\[
\begin{align*}
\text{maximize}_{\mathbf{X}} & \quad \sum_{k=0}^2 \mu_k h(v_k^t \mathbf{X} + \mathbf{Z}_k) \\
\text{subject to} & \quad K\mathbf{X} \preceq S,
\end{align*}
\]
where $\sum_{k=0}^2 \mu_k = 0$ and $\mu_k \geq 0$ for $k > 0$.

Proof: This lemma is the generalization of Lemma 3 where $\mathbf{X}_1$, $\mathbf{X}_2$, and thus $\mathbf{Z}_0$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are $n$-dimensional vectors rather than being one dimensional, and a similar proof can be applied. Here, we only highlight the differences. First, $V_k$ will be a block matrix defined as
\[
V_k = \begin{bmatrix} v_{k1}^t I & v_{k2}^t I \\ \bar{v}_{k1}^t I & \bar{v}_{k2}^t I \end{bmatrix}, \tag{35}
\]
where $I$ is the identity matrix of size $n$. Similar to the proof of Lemma 3 we choose $\bar{v}_{k2} = -v_{k2}$ and $\bar{v}_{k1} = v_{k1}$. Then again we define $U_k = \frac{V_k}{\|v_k\|}$. It is straightforward to check that $U_k V_k^t$ is an identity matrix; i.e., $U_k$ is orthonormal. Also, let
\[
\Sigma_k = \begin{bmatrix} \Sigma_{k1} & 0 \\ 0 & \Sigma_{k2} \end{bmatrix}, \tag{36}
\]
where $\Sigma_{k2}$ is an arbitrary $n \times n$ positive definite matrix, but
\[
\Sigma_{k1} = \frac{K\mathbf{Z}_k}{\|v_k\|^2}. \tag{37}
\]
Again, let us define
\[ \tilde{\mathbf{X}}_k = V_k\mathbf{X} = \begin{bmatrix} \mathbf{X}_{k1} \\ \mathbf{X}_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_k^t \mathbf{X} \\ \mathbf{v}_k \mathbf{X} \end{bmatrix}, \tag{38a} \]
\[ \tilde{\mathbf{Z}}_k = V_k\mathbf{Z}_k = \begin{bmatrix} \mathbf{Z}_{k1} \\ \mathbf{Z}_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_k^t \mathbf{Z}_k \\ \mathbf{v}_k \mathbf{Z}_k \end{bmatrix}. \tag{38b} \]

The only difference between $\tilde{\mathbf{X}}_k$ (and $\tilde{\mathbf{Z}}_k$) in (38) and (31) is on their dimension. Specifically, $\tilde{\mathbf{X}}_k$ and $\tilde{\mathbf{X}}_k$, the elements of $\mathbf{X}_k$ in (38), are $n$-vectors while they were scalar in (31). Note that here $\mathbf{X}_{k1} = \mathbf{v}_k^t \mathbf{X} = \sum_{i=1}^n v_{ki} X_i$ and $\mathbf{X}_1$ is an

\(^7\) With a little abuse of notation, $v_k^t \mathbf{X} = [v_{k1}^t \ v_{k2}^t] [X_1 \ X_2]^t$ refers to $v_{k1}^t X_1 + v_{k2}^t X_2$, in this lemma. A more accurate notation would replace $v_k^t$ by $[v_{k1}, \ldots, v_{k1}, v_{k2}, \ldots, v_{k2}].$
n-vector. Similarly, \( \tilde{Z}_{k1} = v_k^k Z_k = \sum_{k=1}^{2} X_k, \) for \( k = 0, 1, 2, \) and \( \tilde{Z}_{k2} \) is a Gaussian n-vector.

Again, similar to (15), we have \( E(\tilde{Z}_k, X_k) = \| v_k \| 2 \Sigma_k, \) and \( \Sigma_k \) is given by (6). A quick look at \( \Sigma_k \) reveals that \( \tilde{Z}_{k1} \) and \( \tilde{Z}_{k2} \) are uncorrelated. With this, and the fact that \( \tilde{Z}_{k1} \) and \( \tilde{Z}_{k2} \) are Gaussian vectors, we conclude that they are independent.

Next, it is straightforward to generalize Lemma 13, stated in Proposition 6 to the case where \( X_1 \) and \( X_2 \) are vectors. From this, we obtain

\[
\lim_{\sigma_{min} \to \infty} I(\tilde{X}_k; \tilde{X}_k + \tilde{Z}_k) = I(\tilde{X}_k; \tilde{X}_k + \tilde{Z}_k),
\]

in which \( \sigma_{min} \) is the smallest diagonal element of \( \Sigma_k. \) Finally, subtract the constant \( \sum_{k=0}^{2} \mu_k h(\tilde{Z}_k) \) and recall the statistical equivalence of \( \tilde{Z}_{k1} \) and \( \tilde{Z}_{k2} \) to complete the proof.

Eventually, with Lemma 4, we are in the position to see that a jointly Gaussian \( (X_1, X_2) \) is an optimal solution of the optimization problem in (22). Then again, since the covariance constraint includes the trace constraint as a special case, it is straightforward to see that a jointly Gaussian \( (X_1, X_2) \) is an optimal solution of the optimization problem in (21), where the objective function \( W_0 \) in (20) is obtained for

\[
v_0 = \left[ \begin{array}{c} 0 \\ \sqrt{\alpha} \end{array} \right], \quad v_1 = \left[ \begin{array}{c} 0 \\ 1 \end{array} \right], \quad v_2 = \left[ \begin{array}{c} 1 \\ \sqrt{\alpha} \end{array} \right],
\]

and \( \mu_0 = -\mu, \mu_1 = 1, \) and \( \mu_2 = \mu - 1. \)

C. Evaluation of the Outer Bound

Up to now, we have proved that Gaussian distributions are optimal both for \( X_1^a \) and \( X_2^a, \) in (18). In this subsection, we find the optimal covariance matrices for \( X_1^a \) and \( X_2^a, \) and evaluate the outer bound (18), for those inputs. Equivalently, we can find the optimum value of \( W_0 \) in the optimization problem (21), and substitute it into (18). We choose the latter course.

Knowing that \( X_2, Z_1, \) and \( Z_2 \) are all Gaussian, a quick look at \( W_0 \) in (20), and the constraints in (21), reveals that \( W_0 \) is maximized when \( \text{tr}(K_{X_1}) = n P_2. \) Then, \( X_1 = N(0, P_2 I) \) is an optimal solution for \( X_1 \) in (21), and it simplifies as

\[
W = \max_{p(x)} h(X_2 + Z_1) - \mu h(X_2 + Z_2) + (\mu - 1) h(X_2 + Z_3 - n \frac{1}{2} \log a)
\]

s.t. \( K_{X_2} \succeq 0, \text{tr}(K_{X_2}) \leq n P_2, \)

where \( Z_1 = Z_2, Z_2 = Z_2^a, \) and \( Z_3 = Z_3^a + X_1, \) and we have used the identity \( h(aX) = h(X) + \frac{n}{2} \log |a|. \)

Restricting the solution space within Gaussian distributions, the following maximization problem is obtained:

\[
W^G = \max_{p(x)} \frac{1}{2} \log \left( (\pi^n)^n |K_{X_2} + N_1 I| \right) + \frac{n}{2} \log a - \mu \log \left( (\pi^n)^n |K_{X_2} + N_2 I| \right) + \frac{\mu - 1}{2} \log \left( (\pi^n)^n |K_{X_2} + N_3 I| \right)
\]

s.t. \( K_{X_2} \succeq 0, \text{tr}(K_{X_2}) \leq n P_2, \)

where \( N_1, N_2, \) and \( N_3 \) are the covariance matrices of \( Z_1, Z_2, \) and \( Z_3, \) respectively, and \( I \) is the identity matrix of size \( n. \) One can verify that \( N_1 = 1, N_2 = \frac{1}{\alpha}, \) and \( N_3 = \frac{1 + P_2}{\alpha}, \)

before continuing. The objective function can be simplified as

\[
\frac{1}{2} \log |K_{X_2} + N_1 I| - \frac{\mu}{2} \log |K_{X_2} + N_2 I| + \frac{\mu - 1}{2} \log |K_{X_2} + N_3 I| - \frac{n}{2} \log a.
\]

Finally, let \( \lambda_i, i = 1, \ldots, n, \) be the diagonal elements of \( \Lambda; \) then we can write

\[
W^G = \max_{p(x)} \frac{1}{2} \sum_{i=1}^{n} \log (\lambda_i + N_1) - \mu \log (\lambda_i + N_2) + (\mu - 1) \log (\lambda_i + N_3) + \frac{n}{2} \log a
\]

s.t. \( \lambda_i \geq 0, \forall i, \)

\[
\sum_{i=1}^{n} \lambda_i = n P_2.
\]

To solve this last optimization problem, we use Lagrange multipliers \( u \) and \( v = \{v_1, \ldots, v_i\} \) and study the Lagrangian defined by

\[
L(\Lambda, u, v) = \frac{1}{2} \sum_{i=1}^{n} \left[ \log (\lambda_i + N_1) - \mu \log (\lambda_i + N_2) + (\mu - 1) \log (\lambda_i + N_3) + u(n P_2 - \sum_{i=1}^{n} \lambda_i) + \sum_{i=1}^{n} v_i \lambda_i \right].
\]

Since the constraints are inequalities, we examine the Karush-Kuhn-Tucker (KKT) conditions. The KKT stationary and inequality constraints for the above \( L(\Lambda, u, v) \) are given by

\[
\frac{1}{\lambda_i + N_1} - \frac{\mu}{\lambda_i + N_2} + \frac{\mu - 1}{\lambda_i + N_3} = u + v_i = 0,
\]

\[
u_i = 0,
\]

\[
v_i = 0,
\]

for \( i = 1, \ldots, n. \) Recall that \( N_1 = 1, N_2 = \frac{1}{\alpha}, N_3 = \frac{1 + P_2}{\alpha}, \) and \( 0 \leq a < 1; \) hence, \( N_1 < N_2 \leq N_3. \) Further, since all \( \lambda_i \)s have the same role in the optimization problem then \( \lambda_1 = \ldots = \lambda_n. \)
\[
\begin{align*}
\cdots &= \lambda_n \triangleq \lambda, \text{ and (43) simplifies to} \\
&= \frac{1}{\lambda + N_1} - \frac{\mu}{\lambda + N_2} + \frac{\mu - 1}{\lambda + N_3} - u + v_i = 0, \\
&= u(P_2 - \lambda) = 0, \\
&= v_i \lambda = 0. 
\end{align*}
\]

From (44a), for \( i = 1, \ldots, n \), we have \( u - v_i = \frac{1}{\lambda + N_1} - \frac{\mu}{\lambda + N_2} + \frac{\mu - 1}{\lambda + N_3} \). The right-hand side of \( u - v_i \) becomes zero for
\[
\mu^*(\lambda) = \frac{1}{\lambda + N_1} - \frac{1}{\lambda + N_2} = \lambda + N_2 N_3 - N_1 \\
\frac{\lambda + N_2 N_3 - N_1}{\lambda + N_1 N_3 - N_2}. 
\]

or equivalently for
\[
\lambda = \frac{(N_2 - N_1)\mu}{N_2 N_3 - N_1} - N_2. 
\]

But, we know that \( 0 \leq \lambda \leq P_2 \). Thus, we need to determine the values of \( \mu \) corresponding to \( \lambda = 0 \) and \( \lambda = P_2 \). These are basically the maximum and minimum of \( \mu^* \) for \( 0 \leq \lambda \leq P_2 \). For \( N_1 < N_2 \leq N_3 \), \( \mu^* \) is a decreasing function of \( \lambda \) and we have
\[
\mu_{\min}^* = \mu^*(P_2) = \frac{P_2 + N_2 N_3 - N_1}{P_2 + N_1 N_3 - N_2}, \\
\mu_{\max}^* = \mu^*(0) = \frac{N_2 N_3 - N_1}{N_1 N_3 - N_2}. 
\]

Then, for \( \mu > \mu_{\max}^* \), we have \( u - v_i < 0 \), i.e., \( v_i > u \). But, \( u \geq 0 \) and thus \( v_i \) cannot be 0. This implies that \( \lambda = 0 \) for \( \mu > \mu_{\max}^* \). Similarly, if for \( \mu < \mu_{\min}^* \) then \( u - v_i > 0 \); i.e., \( u > v_i \). Thus, \( u \) cannot be 0. This in turn implies \( \lambda = P_2 \). Otherwise, (for \( 0 < \lambda < P_2 \)), the optimal value of \( \lambda \) is given by (46). Hence,
\[
\lambda^* = \begin{cases} 
P_2, & \text{if } 1 \leq \mu \leq \mu_{\min}^*, \\
\frac{(N_2 - N_1)\mu}{N_2 N_3 - N_1} - N_2, & \text{if } \mu_{\min}^* < \mu < \mu_{\max}^*, \\
0, & \text{if } \mu \geq \mu_{\max}^*. 
\end{cases} 
\]

Finally, for our specific problem, we have \( N_1 = 1, N_2 = \frac{1}{a}, N_3 = \frac{1}{a} + P_2 \); therefore, the optimum objective function of \( W^G \) in (42) will be
\[
W^* = \begin{cases} 
(\mu - 1)n\gamma(P_1 + a P_2) + \frac{1}{n} \log \left( 1 + \frac{P_1 + a P_2}{1 + a P_2} \right), & \text{if } 1 \leq \mu \leq \mu_{\min}^*, \\
nf(P_1, P_2, a, \mu) - n\gamma(P_1 + a P_2), & \text{if } \mu_{\min}^* \leq \mu < \mu_{\max}^*, \\
(\mu - 1)n\gamma(P_1), & \text{if } \mu \geq \mu_{\max}^*. 
\end{cases} 
\]

in which \( nf(P_1, P_2, a, \mu) \) is defined in (16). Eventually, when \( n \to \infty \), from (19) we obtain
\[
\mu R_1 + R_2 \leq \frac{1}{n} W^* + \gamma(P_1 + a P_2) \\
= \begin{cases} 
\mu\gamma(P_1 + a P_2) + \gamma(P_2), & \text{if } 1 \leq \mu \leq \mu_{\min}^*, \\
\gamma(P_1) + \gamma\left( \frac{P_1 + a P_2}{1 + a P_2} \right), & \text{if } \mu_{\min}^* \leq \mu < \mu_{\max}^*, \\
\mu\gamma(P_1) + \gamma\left( \frac{P_1 + a P_2}{1 + a P_2} \right), & \text{if } \mu \geq \mu_{\max}^*. 
\end{cases} 
\]

V. MAIN RESULT

Based on the optimization problems in Sections III and IV we are able to characterize the capacity region of the one-sided interference channel for \( a < 1 \), as stated below.

Theorem 1. The capacity region of the one-sided interference channel for \( a < 1 \) is the set of \( (R_1, R_2) \) such that
\[
\begin{align*}
R_1 &\leq \gamma\left( \frac{P_1}{1 + a P_2} \right), \\
R_2 &\leq \gamma(P_2), \\
\mu R_1 + R_2 &\leq \mu\gamma\left( \frac{P_1}{1 + a P_2} \right) + \gamma(P_2), \quad \text{if } 1 \leq \mu \leq \mu_{\max}^*, \\
\mu R_1 + R_2 &\leq f(P_1, P_2, a, \mu), \quad \text{if } \mu_{\min}^* \leq \mu < \mu_{\max}^*, \\
\mu R_1 + R_2 &\leq \gamma\left( \frac{a P_2}{1 + a P_2} \right) + \gamma\left( \frac{P_1 + a P_2}{1 + a P_2} \right) + \gamma\left( \frac{P_1 + a P_2}{\mu - a P_0} \right), \quad \text{if } \mu \geq \mu_{\max}^*.
\end{align*}
\]

Proof: We have already proved the achievability and converse in Sections III and IV; we summarize the proof here. The achievability is proved in Lemma I which is an alternative representation of the Han-Kobayashi region. The converse for (51c)-(51e) is established by formulating a new outer bound on the weighted sum-rate in (19), proving that Gaussian inputs are optimal for that, finding the optimal covariance matrices for the inputs, and evaluate the outer bound in (50). This bound on \( \mu R_1 + R_2 \) in association with the trivial bounds \( R_1 \leq \gamma(P_1) \) and \( R_2 \leq \gamma(P_2) \) make an outer bound which is the same as the achievable region in Lemma I. This completes the proof.

The above result is the first capacity region for any form of the Gaussian interference channel in the weak interference regime. The reader should have noticed that the capacity region introduced in Theorem I has two other representations, given by Proposition 4 and Proposition 5. The equivalency of these three regions has already been established and used to obtain the main result in Theorem I. Each of these representations has its advantages. An advantage of the representation in Theorem I is the fact that it eliminates the need for convexification. However, for its more tractable presentation, the original Han-Kobayashi region may still be preferred to represent the capacity region. In addition, we can express the capacity region for the whole range of \( a \) using the same set of equations. This is stated in the following theorem.

Theorem 2. The convex hull of the union of the set of rate pairs \( (R_1, R_2) \) satisfying
\[
\begin{align*}
R_1 &\leq \gamma\left( \frac{P_1}{1 + a \beta P_2} \right), \\
R_2 &\leq \gamma(P_2), \\
R_1 + R_2 &\leq \gamma\left( \frac{P_1 + a \beta P_2}{1 + a \beta P_2} \right) + \gamma(\beta P_2),
\end{align*}
\]

over \( \beta \in [0, 1] \) provides the capacity region of the one-sided interference channel for any \( a \in \mathbb{R} \).
Proof: By Lemma 1 we know that (51a)-(51e) is an alternative representation of (52a)-(52c); hence, for \( a < 1 \) the proof of this Theorem immediately follows from that of Theorem 1. For \( a \geq 1 \), from Proposition 1 we know that setting \( \beta = 0 \) in (52a)-(52c) is optimal. This completes the proof for all range of \( a \).

From the above capacity result for the one-sided interference channel we observe that a transmitter creating (weak) interference uses both private and common messages in the HK achievable region. So far, this is the first capacity result for any form of the Gaussian interference channel in which both private and common messages are used to achieve the capacity region. From the capacity region of the one-sided interference channel at strong interference, see Fig. 3, it can be seen that a transmitter causing strong interference requires only the common message. Furthermore, when there is no link from a transmitter to the other receiver, that transmitter will have only the private message.

A closer look at the representation of the capacity region in Proposition 3 reveals that the interfered-with receiver (here, receiver 1) first decodes the common part of the interference, and cancels it out; it then treats the private part of the interference as noise. The best rate with which the common message is decoded at receiver 1 is equal to \( \gamma \left( \frac{a P_2}{1 + a \beta P_2} \right) \); the remaining part of the interference, i.e., \( a \beta P_2 \), is treated then as noise. As a result, receiver 1 can decode its respective message at a maximum rate of \( R_1 = \gamma \left( \frac{P_1}{1 + a \beta P_2} \right) \). On the other hand, the interference-free receiver (receiver 2) can decode the common message of its corresponding transmitter at a rate of \( \gamma \left( \frac{P_2}{1 + a \beta P_2} \right) \), and then the private part of the message at a rate of \( \gamma (\beta P_2) \). Next, it is easy to see that receiver 1 imposes the maximum rate with which the common part of the interfering user’s message can be decoded, simply because \( \gamma \left( \frac{a \beta P_2}{1 + a \beta P_2} \right) \leq \gamma \left( \frac{\beta P_2}{1 + a \beta P_2} \right) \) for\( a \leq 1 \). Therefore, the overall rate at the second receiver is \( R_2 \leq \gamma \left( \frac{a \beta P_2}{1 + a \beta P_2} \right) + \gamma (\beta P_2) \).

The approach described in the previous paragraph is basically the nature of the HK scheme. Theorem 1 sheds more light on that region in several ways: 1) by proving that such a scheme is optimal, 2) by determining the fraction of power required for the private and common messages to achieve the capacity region, and 3) by exposing the optimal decoding required for the weighted sum-rate. This last statement is further discussed below.

In terms of the optimal decoding, from Theorem 1 it can be seen that treating interference as noise achieves the weighted sum-capacity \( \mu R_1 + R_2 \) of the one-sided interference channel for any \( 1 \leq \mu \leq \frac{1 + a P_2}{1 + P_2} - \frac{a}{P_2} \). This implies that when \( a \to 0 \) or \( P_1 \to 0 \) treating interference as noise is optimal for any \( \mu \geq 1 \). Besides, from (51c) we can see that for \( \mu \geq \frac{1 + a P_2}{a P_2} \), decoding interference, and canceling it, is optimal for the weighted sum-rate \( \mu R_1 + R_2 \). For \( a = 1 \), the latter implies that decoding interference is optimal for any \( \mu \geq 1 \), and thus is capacity-achieving. This also indicates that such a scheme is optimal \( a > 1 \), because as \( a \) increases the interference becomes stronger and thus can be decoded without incurring any penalty on \( R_2 \). The optimality of decoding interference for \( a \geq 1 \) has been established in 3, as stated in Proposition 1.

VI. Summary

This paper has fully characterized the capacity region of the one-sided Gaussian interference channel. This result is due to introducing a tight outer bound and a new representation of the Han-Kobayashi inner bound for the above channel. The new representation of the Han-Kobayashi region not only helps to characterize the capacity region but sheds more light on the optimal encoding and decoding techniques. The capacity region established in this paper is the first capacity region for the Gaussian interference channel in the weak interference, and the first one to use both private and common messages to achieve a capacity region. This basic approach together with the extremal inequality we developed to prove it can be useful in many related network information theory problems.

REFERENCES

[1] A. Carleial, “Interference channels,” IEEE Transactions on Information Theory, vol. 24, pp. 60–70, January 1978.
[2] A. Carleial, “A case where interference does not reduce capacity,” IEEE Transactions on Information Theory, vol. 21, no. 5, pp. 569–570, 1975.
Fig. 4. Capacity region of the one-sided inference channel for different channel gains in various transmission powers.

[3] H. Sato, “The capacity of the Gaussian interference channel under strong interference,” IEEE Transactions on Information Theory, vol. 27, no. 6, pp. 786–788, 1981.
[4] M. H. M. Costa, “On the Gaussian interference channel,” IEEE Transactions on Information Theory, vol. 31, pp. 607–615, September 1985.
[5] I. Sason, “On achievable rate regions for the Gaussian interference channel,” IEEE Transactions on Information Theory, vol. 50, no. 6, pp. 1345–1356, 2004.
[6] 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.
[7] R. H. Etkin, D. N. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” IEEE Transactions on Information Theory, vol. 54, no. 12, pp. 5534–5562, 2008.
[8] 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.
[9] 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.
[10] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: John Wiley & Sons, 2006.
[11] T. Liu and P. Viswanath, “An extremal inequality motivated by multi-terminal information-theoretic problems,” IEEE Transactions on Information Theory, vol. 53, no. 5, pp. 1839–1851, 2007.
[12] M. Vaezi and H. V. Poor, “On limiting expressions for the capacity regions of Gaussian interference channels,” in Proc. 49th Asilomar Conference on Signals, Systems and Computers (to appear), Nov. 2015.
[13] 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.