The Capacity of Mixed and One-Sided
Gaussian Interference Channels

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

Abstract

This paper adds to the understanding of the capacity region of the Gaussian interference channel. To this end, the capacity region of the one-sided Gaussian interference channel is first fully characterized. This is accomplished by introducing a new representation of the Han-Kobayashi region and providing a new outer bound on the capacity region of this channel which is tight both in the weak and strong interference regimes. In light of this capacity result, the capacity region of the degraded Gaussian interference channel is established immediately. Next, by combining the capacity regions of the one-sided interference channels in the weak and strong interference regimes, new outer bounds on the capacity region of the interference channel are introduced, in various interference regimes. It is then proved that the outer bound corresponding to the mixed interference regime, in which one of the receivers is subject to strong interference while the other one suffers from weak interference, is tight in a broad range of this regime.

The new capacity results, on the whole, confirm the optimality of decoding part of the interference and treating the rest as noise. The optimum amount to be decoded varies from 0 to 100% of the interfering signal, depending on the relative importance of the users’ rates (i.e., the ratio of weights in the weighted sum-rate), their transmission powers, and the gain of the weak interference link. Optimal values are found explicitly, based on the above parameters.

Index Terms

Capacity, Gaussian interference channel, one-sided interference channel, degraded interference channel, mixed interference, Han-Kobayashi scheme, weighted sum-capacity.

I. INTRODUCTION

The interference channel models communication networks consisting of two or more pairs of transmitters and receivers in which each transmitter communicates with its respective receiver while interfering with other receivers. The two-user interference channel is a two-transmitter two-receiver network, in which each transmitter has a message for its respective receiver [1]–[5]. A basic information theoretic model to study the two-user Gaussian interference channel is...
A Gaussian interference channel in the standard form, with inputs $X_1$ and $X_2$, outputs $Y_1$ and $Y_2$, noises $Z_1$ and $Z_2$, and interference gains $a$ and $b$.

Fig. 1. A Gaussian interference channel in the standard form, with inputs $X_1$ and $X_2$, outputs $Y_1$ and $Y_2$, noises $Z_1$ and $Z_2$, and interference gains $a$ and $b$.

channel is shown in Fig. 1. Understanding the capacity region of this channel is one of the long-standing important problems in network information theory.

A limiting expression for the capacity region of the discrete memoryless interference channel was derived by Ahlswede [2]. However, this limiting expression is often considered to have little value because it is computationally excessively complex. Moreover, in general, one cannot simply restrict the inputs to be Gaussian in the limiting expression to characterize the capacity region of the Gaussian interference channel [6], [7]. A single-letter capacity expression\(^1\) for the discrete memoryless interference channel is unknown, except for the strong interference regime [8]. Likewise, the problem of finding a single-letter capacity expression for the one-sided interference channel, a special case of the two-user interference channel in which only one of the receivers experiences interference, has been open for many years.

Han and Kobayashi introduced an achievable region in 1981 [5], which is still the best known inner bound for the general interference channel. In the Han-Kobayashi (HK) scheme, each user can split its message to be sent into two submessages of smaller rates and power. These are known as private and common messages; the former is intended to be decoded only at the respective receiver whereas the latter can be decoded at both receivers. The rationale behind this coding scheme is to decode part of interference (the common message) and treat the rest as noise. The optimal input distributions are not known for the HK region. As such, commonly a subset of the HK region with Gaussian codebooks is used to represent the HK region for the Gaussian channel; see e.g. [9]–[12].

Flexibility in splitting each user’s transmission power into the common/private portions of information and time-sharing between them make the HK scheme very strong, but complicated. Not surprisingly, though, the optimal HK strategy is not well-understood, in general. Nevertheless, the HK scheme is known to be within \(\frac{1}{2}\) bit of the

---

\(^1\) In contrast to a limiting expression, a single-letter capacity expression includes only the channel input and output random variables, and possibly a few auxiliary random variables, involved in one use of the channel.
capacity region of the two-user Gaussian interference channel \cite{9}. Although this gap could be due to a suboptimal scheme, loose outer bounds, or both, the outer bounds seem to be the most crucial.

Characterizing the capacity region of the interference channel 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 receivers \cite{13}, \cite{14} and more capable receivers in which part of the interfering signals is removed. A simple but very important example of the latter is the fact that the capacity region of the general two-user interference channel is within the capacity regions of the corresponding one-sided interference channels. That being said, the capacity region of this apparently simple channel has been unknown in the weak interference regime, i.e., when the gain of the interference link is less than one.

We establish the capacity region of several classes of the Gaussian interference channel in this work. First and most importantly, we fully characterize the capacity region of the one-sided interference channel. This is realized by introducing a new representation of the HK inner bound for this channel and developing a novel outer bound which is tight for the whole range of interference. Second, we use this new capacity result to establish new outer bounds on the capacity region of the general two-user interference channel, in various interference regimes. More importantly, we show that the outer bounds for the mixed interference case are tight when $a < 1$ and $b \geq \frac{1 + P_2}{1 + \alpha P_2}$, or $b < 1$ and $a \geq \frac{1 + P_1}{1 + b P_1}$, and thus we establish the first known capacity region of the interference channel in the mixed interference regime. This capacity region includes the capacity region of the degraded interference channel as a special case. Last but not least, in order to prove the optimality of Gaussian distributions for the outer bound, we develop a few extremal inequalities that may be useful in other network information theory problems.

The optimal encoding and decoding strategies for all capacity results we mentioned are the same. The receiver suffering from weak interference decodes part of interference and treats the rest as noise. The optimum amount to be decoded varies from 0 to 100% of the interfering signal and depends on the relative importance of the users in the weighted sum-rate, their total transmission powers, and the gain of the weak interference. We explicitly express the optimal value in terms of the above parameters. Since the achievability is based on the HK scheme, this identifies the optimal splitting of the transmission power of the user causing weak interference into the common/private portions of information. The other user, i.e., the one causing no interference (in the one-sided channel) or strong interference do not need a power split. It will transmit only private information in the former case and common information in the latter case. The receiver suffering from the weak interference determines the rate of the common message of the user who needs rate splitting, because the other receiver is always capable of decoding its common message at a higher rate, within the aforementioned interference regimes.

The capacity regions of the one-sided interference channel and degraded interference channel, characterized in this work, are among the few cases for which the capacity region is known for an interference channel. Moreover, the capacity of the one-sided interference channel is the first ever capacity region for the interference channel in the weak interference regime. This channel is a building block of many interference networks, and its capacity region is important in understanding and establishing the capacity of more complex networks.

The paper is organized as follows. We review the channel model and previously known results in Section \cite{11} and
we develop new extremal inequalities in Section III. We introduce the new representation of the HK region, and
develop the new outer bound in Sections IV and V, respectively. Main results, including the capacity regions for the
one-sided, degraded, and general interference channels are presented in Section VI. This is followed by concluding
remarks in Section VII.

Regarding notation, we will use uppercase boldface letters (e.g., \( \mathbf{X} \)) for random vectors, uppercase letters (e.g., \( X \)) for random variables and matrices, lowercase boldface letters (e.g., \( \mathbf{x} \)) for deterministic vectors, lowercase letters (e.g., \( x \)) for scalars, \( (\cdot)^t \) for transpose, and \( \gamma(x) \) as an abbreviation for \( \frac{1}{2} \log(1 + x) \). We use \( h(\cdot) \) to denote the
differential entropy, \( I(\cdot; \cdot) \) to denote mutual information, \( I \) to denote the identity matrix of size \( n \), and \( \preceq \) to denote
less than or equal to in the positive semidefinite partial ordering between real symmetric matrices. Vectors are
column vectors and all logarithms are to the base 2.

II. CHANNEL MODEL AND PRELIMINARIES

In this section, we describe the model we use in this paper and summarize the previously known results for the
Gaussian interference channel, as well as 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 trans-
mitter communicates with its corresponding receivers while interfering with the other receiver. Without loss of
generality, we use the standard form of the Gaussian interference channel [1], [15], in which the channel is expressed,
for a single channel use, by

\[
\begin{align*}
Y_1 &= X_1 + \sqrt{a}X_2 + Z_1, \\
Y_2 &= \sqrt{b}X_1 + X_2 + Z_2,
\end{align*}
\]

where \( a \) and \( b \) are two non-negative real numbers representing the crossover gains; and, for \( j \in \{1, 2\} \), \( X_j \), \( Y_j \),
and \( Z_j \), respectively, represent the transmitted signal, received signal, and the channel noise, and \( Z_1 \) and \( Z_2 \) are
independent and identically distributed (i.i.d.) Gaussian random variables with zero means and unit variances. Let
\( w_1 \) and \( w_2 \) be two independent messages uniformly distributed over \( \mathcal{W}_1 = [1, \ldots, 2^{nR_1}] \) and \( \mathcal{W}_2 = [1, \ldots, 2^{nR_2}] \),
respectively. Transmitter \( j \) wishes to transmit message \( w_j \) to receiver \( j \) in \( n \) channel use at rate \( R_j \), and \( X_j \) is
subject to an average power constraint \( P_j \), i.e.,

\[
\frac{1}{n} \sum_{i=1}^{n} \|X_{ji}\|^2 \leq P_j, \quad j = 1, 2.
\]

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

In the following, we consider two other special cases of the interference channel, namely, the one-sided inter-
ference channel and the degraded interference channel. The former is fundamentally important in establishing our
results in this paper, and will be studied first.
Depending on the values of $a$ and $b$ the Gaussian interference channel is classified into the following classes:

- **Weak interference:** $a < 1$ and $b < 1$
- **Strong interference:** $a \geq 1$ and $b \geq 1$
- **Mixed interference:** $a < 1$ and $b \geq 1$ or $a \geq 1$ and $b < 1$
- **One-sided interference:** $a = 0$ or $b = 0$

We study the following two cases particularly.

1) **One-sided interference channel:** 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 $b = 0$. With this, the channel model described by (1) simplifies to

\[
Y_1 = X_1 + \sqrt{a}X_2 + Z_1, \quad (3a)
\]
\[
Y_2 = X_2 + Z_2. \quad (3b)
\]

Depending on the value of $a$, the gain of the interference link, the above channel is classified either as a weak or a strong one-sided interference channel. Specifically, the channel is in the weak interference regime if $a < 1$ and the strong interference regime if $a \geq 1$.

2) **Degraded interference channel:** The Gaussian interference channel described by (1) is called degraded if $ab = 1$. A degraded interference channel is in the mixed interference regime, except when $a = b = 1$.

B. Previous Results

In this subsection, we present the most important previously know results, including all existing capacity results of the two-user Gaussian interference channel, and two of its special cases, namely, the one-sided interference channel and degraded interference channel.

1) **Interference channel:** The Han-Kobayashi [5], achievable scheme is the best known achievable scheme for the two-user Gaussian interference channel, to date. The idea behind the HK scheme is to decode part of interference and treat the rest as noise. With this, the HK scheme splits the information of both users into private and common parts. The transmitted message, for each user, is then the superposition of its submessages and has the total power of that user. The former is intended to be decoded only at the respective receiver whereas the common information can be decoded by both receivers. Arbitrary power allocation to the common and private portion of information besides time-sharing between such splits makes the HK strategy very strong, yet difficult to optimize and fully understand. The fact that the optimal input is not known for the HK strategy makes the matter even more complicated.

If we use jointly Gaussian inputs and do not apply time-sharing then the HK inner bound can be stated as follows.

**Proposition 1** [5]. The Han-Kobayashi achievable region for the two-user interference channel with Gaussian inputs excluding the trivial interference-free case where both $a$ or $b$ are zero.

\[\text{We exclude the trivial interference-free case where both } a \text{ or } b \text{ are zero.}\]
inputs is the set of \((R_1, R_2)\) satisfying

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

\[
R_2 \leq \gamma \left( \frac{P_2}{1 + b \alpha P_1} \right),
\]

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

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

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

\[
2R_1 + R_2 \leq \gamma \left( \frac{\alpha P_1}{1 + a \beta P_2} \right) + \gamma \left( \frac{P_1 + a \beta P_2}{1 + a \beta P_2} \right) + \gamma \left( \frac{\beta P_2 + b \alpha P_1}{1 + b \alpha P_1} \right),
\]

\[
R_1 + 2R_2 \leq \gamma \left( \frac{\beta P_2}{1 + b \alpha P_1} \right) + \gamma \left( \frac{P_2 + b \alpha P_1}{1 + b \alpha P_1} \right) + \gamma \left( \frac{\alpha P_1 + a \beta P_2}{1 + a \beta P_2} \right),
\]

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

In the above region, \(\alpha\) and \(\beta\) control the power allocation for the private and common part of information for
the first and second users, respectively. Specifically, \(\alpha P_1\) and \(\bar{\alpha} P_1\) represent the power allocated to the private and
common information of user 1. Similarly, \(\beta P_2\) and \(\bar{\beta} P_2\) represent the power allocated to the private and common
information of user 2, respectively.

The Han-Kobayashi region is known to be optimal for all capacity results established thus far for the Gaussian
interference channel. We present those results in what follows.

**Proposition 2** [4]. The set of rate pairs \((R_1, R_2)\) satisfying

\[
R_1 \leq \gamma (P_1),
\]

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

\[
R_1 + R_2 \leq \gamma (P_1 + a P_2),
\]

\[
R_1 + R_2 \leq \gamma (b P_1 + P_2),
\]

provides the capacity region of the interference channel in the strong interference regime.

The above result is the only case in which the capacity region is fully characterized for the two-user interference
channel. To get this region both receivers decode both messages [1]. In terms of the HK scheme, the above capacity
region is obtained when both users transmit only the common messages, i.e., in [4], \(\alpha = 0\) and \(\beta = 0\).

**Proposition 3** [10]. The sum-capacity of the interference channel in the mixed interference regime with \(a < 1\) and \(b \geq 1\) is given by

\[
C_{\text{sum}} = \gamma (P_2) + \min \left\{ \gamma \left( \frac{P_1}{1 + a P_2} \right), \gamma \left( \frac{b P_1}{1 + P_2} \right) \right\}.
\]
Proposition 4 [10]–[12]. The sum-capacity of the interference channel in the very weak interference regime where \( \sqrt{a}(bP_1 + 1) + \sqrt{b}(aP_2 + 1) < 1 \) is given by
\[
C_{\text{sum}} = \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma \left( \frac{P_2}{1 + bP_1} \right),
\]
(7)

The above result indicates that treating interference as noise achieves the sum-capacity of the interference channel in the “very” weak interference regime, also called the noisy interference regime [11].

2) One-sided interference channel: Again we start with the best inner bound, i.e., the Han-Kobayashi achievable region. It is easy to show that the representation of the HK region in Proposition 1 for \( b = 0 \) simplifies as below

Proposition 5. 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
\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right),
\]
(8a)
\[
R_2 \leq \gamma (P_2),
\]
(8b)
\[
R_1 + R_2 \leq \gamma \left( \frac{P_1 + a\bar{\beta} P_2}{1 + a\beta P_2} \right) + \gamma (P_2),
\]
(8c)
for some \( \beta \in [0, 1] \) and \( \bar{\beta} = 1 - \beta \).

Note that, there is no reason to have common information for user 1 because \( b = 0 \) implies that receiver 2 is not able to receive it, even if there was any. Hence, \( \alpha = 1 \). In [10] Lemma 7, it is shown that the representation of the HK region in Proposition 5 can be further simplified as

Proposition 6. The HK achievable region for the one-sided interference channel with Gaussian inputs can be further simplified to the set of \((R_1, R_2)\) such that
\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right),
\]
(9a)
\[
R_2 \leq \gamma \left( \frac{a\bar{\beta} P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (P_2),
\]
(9b)
for some \( \beta \in [0, 1] \) and \( \bar{\beta} = 1 - \beta \).

The above region is convex, and it is easy to prove that time-sharing region with \( \beta = 0 \) and \( \beta = 1 \) is inside the above region. Figure ?? 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.

Proposition 7 [17]. The capacity region of the one-sided interference channel in the strong interference regime \((a \geq 1)\) is the set of rate pairs \((R_1, R_2)\) satisfying
\[
R_1 \leq \gamma (P_1),
\]
(10a)
\[
R_2 \leq \gamma (P_2),
\]
(10b)
\[
R_1 + R_2 \leq \gamma (P_1 + aP_2),
\]
(10c)
The above result is the only case where the capacity region is known for the one-sided interference channel. It is achieved by decoding interference at the interfered-with receiver.

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

**Proposition 8** [18]. The sum-capacity of the one-sided interference channel in the weak interference case is given by

\[
C_{\text{sum}} = \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma (P_2). \tag{11}
\]

The above proposition reveals that maximum sum-rate for this channel is achieved by treating the interference as noise. The following proposition shows that such a simple decoding is also optimum for a certain range of weighted sum-rates.

**Proposition 9** [10]–[12]. For \(1 \leq \mu \leq \frac{P_2 + 1/a}{P_2 + 1} \) and \(a \leq 1\), the capacity region of the one-sided interference channel is outer bounded by

\[
\mu R_1 + R_2 \leq \mu \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma (P_2). \tag{12}
\]

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

3) Degraded interference channel: The capacity region of this channel is not known except for \(a = b = 1\). However, Costa [17] has proved that the capacity region of this channel with \(0 < a = \frac{1}{b} < 1\) is equal to the capacity region of the one-sided interference channel with \(0 < a < 1\).

### III. New Extremal Inequalities

In this section, we develop a few extremal inequalities that are used to prove a new tight outer bound on the capacity region of the Gaussian interference channel in Section [V] and establish the main capacity results of this paper in Section [VI]. These extremal inequalities can be useful in solving other network information theory problems.

Before introducing the extremal inequalities, we note that for a Gaussian random \(n\)-vector \(X \sim \mathcal{N}(\mu_X, K_X)\) we have

\[
h(X) = \frac{1}{2} \log \left( (2\pi e)^n |K_X| \right),
\]

in which \(|K_X|\) is the determinant of \(K_X\). In addition, we use the following generalization of [19] Lemma 13 in proving the other results of this section.

**Lemma 1.** Let \(Z_1, \ldots, Z_m, m \geq 2,\) be mutually independent Gaussian \(n\)-vectors with positive definite covariance matrices, and \(X_1, \ldots, X_m\) be random \(n\)-vectors with bounded covariance matrices, where \(X = [X^t_1, \ldots, X^t_m]^t\) is
independent of $Z = [Z_1^t, \ldots, Z_m^t]^t$. Also, let $\sigma_{j_1}^2, \ldots, \sigma_{j_n}^2$ be the diagonal elements of $K_{Z_j}$, $\sigma_{j_{\min}}^2 = \min\{\sigma_{ji}\}$, $i = 1, \ldots, n$, and $\sigma_{j_{\min}}^2 = \min\{\sigma_{j_{\min}}^2\}$ for $j = 2, \ldots, m$, then we have

$$\lim_{\sigma_{j_{\min}}^2 \to \infty} I(X; X + Z) = I(X_1; X_1 + Z_1).$$

Proof: See Appendix A.

We now introduce several results that will be useful in the sequel.

**Lemma 2.** Consider the optimization problem

$$\max_{p(x)} \sum_{k=0}^K \mu_k h(X + Z_k)$$

subject to $K_X \preceq S,$

where $Z_k, k = 0, \ldots, K$ represent $K+1$ Gaussian random $n$-vectors with positive definite covariance matrices $K_{Z_k}$, and $\mu_0 = -1$ and $\mu_k \geq 0$ for $k > 0$ are a total of $K + 1$ nonzero scalars that sum up to zero (i.e., $\sum_{k=0}^K \mu_k = 0$). $X$ is a random $n$-vector independent of $Z_k$s, $K_X$ is the covariance matrix of $X$, and $S$ is a positive semidefinite matrix. A Gaussian $X$ is an optimal solution of (14).

Proof: This lemma is a special case of [20, Theorem 2] for $U = u$. Moreover, we have removed the unnecessary ordering between the covariance matrices of the $Z_k$s.

**Lemma 3.** Let $v_0, v_1, \ldots, v_K$ be $K+1$ deterministic, nonzero vectors in $\mathbb{R}^m$, $Z_0, Z_1, \ldots, Z_K$ be $K+1$ Gaussian noises with positive variances, $X$ be a random $m$-vector independent of the $Z_k$s, $K_X$ be the covariance matrix of $X$, and $S$ be a positive semidefinite matrix. A Gaussian $X$ is an optimal solution of the optimization problem

$$\max_{p(x)} \sum_{k=0}^K \mu_k h(v_k^tX + Z_k)$$

subject to $K_X \preceq S,$

where $\sum_{k=0}^K \mu_k = 0$ and $\mu_k \geq 0$ for $k > 0$.

Proof: See Appendix B.

What we proved in Lemma 3 for scalar $Z_k$ can be generalized to the vector case, as shown in the following lemma.

**Lemma 4.** Let $v_0, v_1, \ldots, v_K$ be $K+1$ deterministic, nonzero vectors in $\mathbb{R}^m$, $Z_0, Z_1, \ldots, Z_K \in \mathbb{R}^n$ be Gaussian noises with positive definite covariance matrices, $X_1, \ldots, X_m$ be random $n$-vectors independent of the $Z_k$s, $K_{X_1}, \ldots, K_{X_m}$ be the covariance matrices of $X_1, \ldots, X_m$, and $S_1, \ldots, S_m$ be positive semidefinite matrices. A jointly Gaussian $(X_1, \ldots, X_m)$ is an optimal solution of the optimization problem

$$\max_{p(x_1, \ldots, x_m)} \sum_{k=0}^K \mu_k h(\sum_{j=1}^m v_{kj}X_j + Z_k)$$

subject to $K_{X_j} \preceq S_j, j = 1, \ldots, m$.
where $\sum_{k=0}^{K} \mu_k = 0$ and $\mu_k \geq 0$ for $k > 0$.

Proof: See Appendix C.

Remark 1. As can be understood from the proof of Lemma 4 this lemma is valid even if instead of $m$ constraints $K_{X_j} \preceq S_j$, $j = 1, \ldots, m$ there is only one constraint $K_X \preceq S$ in which $X = [X_1^t, \ldots, X_m^t]^t$.

It is worth mentioning that the argument we used to prove Lemma 3 and Lemma 4 can be applied to generalize other similar problems. Generally speaking, if $\sum_{k=0}^{K} \mu_k h(X + Z_k)$ for arbitrary $\mu_k$s, is maximized/minimized by $p(x)$, the same $p(x)$ will maximize/minimize $\sum_{k=0}^{K} \mu_k (\sum_{j=1}^{m} v_k j X_j + Z_k)$. In particular, we can generalize [19] Theorem 1 to get the following.

Lemma 5. Let $u = [u_1, \ldots, u_m]^t \neq 0$ and $v = [v_1, \ldots, v_m]^t$ be deterministic vectors in $\mathbb{R}^n$, $Z_1$ and $Z_2 \in \mathbb{R}^n$ be Gaussian random vectors with positive definite covariance matrices $K_{Z_1}$ and $K_{Z_2}$, and $X_1, \ldots, X_m$ be random $n$-vectors independent of $Z_1$ and $Z_2$. For any $\mu \geq 1$ and any positive semidefinite matrices $S_1, \ldots, S_m$, a jointly Gaussian $(X_1, \ldots, X_m)$ is an optimal input distribution for

$$\begin{align*}
\text{maximize} & \quad h(\sum_{j=1}^{m} u_j X_j + Z_1) - \mu h(\sum_{j=1}^{m} v_j X_j + Z_2) \\
\text{subject to} & \quad K_{X_j} \preceq S_j, \ j = 1, \ldots, m,
\end{align*}$$

(17)

in which $K_{X_j}$ is the covariance matrix of $X_j$.

Proof: For $v \neq 0$ the proof is very similar to the proof of Lemma 4. The difference is in the starting point where, in (91), we start the argument with

$$\mathcal{N}(0, K_X^*) = \arg\max_{p(x)} [h(X + Z_1) - \mu h(X + Z_2)],$$

(18)

which comes from [19] Theorem 1. All other arguments are very similar, and thus are not included here. For $v = 0$ again we can use the same arguments but starting with

$$\mathcal{N}(0, K_X^*) = \arg\max_{p(x)} h(X + Z_1),$$

(19)

which is trivial.

It is straightforward to check that Lemma 5 generalizes both Theorem 1 and Corollary 6 of [19]. The former is obtained for $m = 1$ and $u_1 = v_1 = 1$. The latter, i.e., [19] Corollary 6, is obtained for $m = 2$, $n = 1$, and $Z_1 = Z_2$. The only difference here is the fact that we have separate covariance constraints $K_{X_1} \preceq S_1$ and $K_{X_2} \preceq S_2$. However, looking at the proof of Lemma 4 and consequently that of Lemma 5 it is clear they are valid for $K_X \preceq S$, where $X = [X_1, X_2]^t$.

IV. NEW REPRESENTATION OF THE HK REGION FOR THE ONE-SIDED INTERFERENCE CHANNEL

Consider the Han-Kobayashi rate region in Proposition 3 for the one-sided interference channel depicted in Fig. 2. Figure 3 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 [18]. The above region is convex, and it is easy to prove that time-sharing between these two points is inside the HK region.

We are interested in finding the optimal value of $\beta$ such that the weighted sum-rate $\mu R_1 + R_2$, also called as the $\mu$-sum rate, is maximized for any $\mu \geq 1$ [18]. To this end, using (9a)–(9b), it is seen that

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

To determine the optimal value of $\beta$ that maximizes $R_{\mu-\text{sum}}$ for different values of channel parameters, we find the critical point of the bound by evaluating the first-order partial derivative of the right-hand side of (20) with respect to $\beta$ and setting it to zero, which proceeds as

$$R_{\mu-\text{sum}} \triangleq \mu R_1 + R_2 \leq \mu \gamma \left( \frac{P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{a\beta P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (\beta P_2) = \mu \gamma \left( \frac{P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{a\beta P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (\beta P_2) \leq \mu \gamma \left( \frac{P_1}{1 + a\beta P_2} \right) + \gamma \left( \frac{a\beta P_2}{1 + P_1 + a\beta P_2} \right) + \gamma (\beta P_2) \leq \frac{\partial R_{\mu-\text{sum}}}{\partial \beta} = 0 \Rightarrow \mu = \frac{1 + a\beta P_2}{1 + \beta P_2} \frac{1 + P_1 - a}{aP_1}. \quad (22)$$

Let us define

$$\mu^* = \frac{1 + a\beta P_2}{1 + \beta P_2} \frac{1 + P_1 - a}{aP_1}. \quad (23)$$

Fig. 2. A one-sided Gaussian interference channel in the standard form.

---

3 We focus on $\mu \geq 1$ because from Proposition 8 it is clear that the capacity region is inside the sum capacity upper bound ($\mu = 1$). This is visualized in Fig. ??.
The Han-Kobayashi achievable points A ($\beta = 1$) and B ($\beta = 0$) with the corresponding rate region (black solid region), the sum-capacity upper bound (blue dotted lines), and the weighted sum-rate upper 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 the interference. In particular, any point on the line segment AB can be achieved by time-sharing. Note that the line segment AB then is achieved by naive time-sharing, or time-sharing with fixed power. Time sharing with power control achieves a larger region [13]. Also, the HK region strictly includes this line segment.

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

$$
\mu_0^* \triangleq \frac{1 + P_1 - a}{aP_1},
$$

(24a)

$$
\mu_1^* \triangleq \frac{1 + aP_2}{1 + P_2} \mu_0^*.
$$

(24b)

Now, one can check that the optimal value of $\beta$ to maximize (21) is given by

$$
\beta^* = \begin{cases} 
1, & \text{if } 1 \leq \mu \leq \mu_1^* \\
\frac{\mu_0^* - \mu}{\mu - a\mu_0^*} \frac{1}{P_2}, & \text{if } \mu_1^* < \mu < \mu_0^* \\
0, & \text{if } \mu \geq \mu_0^*
\end{cases}
$$

(25)

Consequently, we obtain

$$
\mu R_1 + R_2 \leq \begin{cases} 
\mu \gamma \left( \frac{P_1}{1 + aP_2} \right) + \gamma(P_2), & \text{if } 1 \leq \mu \leq \mu_1^* \\
f(P_1, P_2, a, \mu), & \text{if } \mu_1^* < \mu < \mu_0^* \\
\mu \gamma(P_1) + \gamma \left( \frac{aP_2}{1 + P_1} \right), & \text{if } \mu \geq \mu_0^*
\end{cases}
$$

(26)
where
\[ f(P_1, P_2, a, \mu) = \mu \gamma \left( \frac{P_1}{1 + a P_2} \right) + \gamma \left( \frac{a P_2 - a \frac{\mu - \mu_0}{\mu - a \mu_0}}{1 + P_1 + a \frac{\mu - \mu_0}{\mu - a \mu_0}} \right) + \gamma \left( \frac{\mu_0 - \mu}{\mu - a \mu_0} \right). \]  
(27)

In light of the above optimization, for \( a < 1 \) we have a new representation of the HK inner bound in the following lemma.

**Lemma 6.** The Han-Kobayashi achievable region for the one-sided interference channel in the weak interference regime can be represented by the set of \((R_1, R_2)\) such that
\[
R_1 \leq \gamma(P_1), \\
R_2 \leq \gamma(P_2),
\]
\[
\mu R_1 + R_2 \leq \mu \gamma \left( \frac{P_1}{1 + a P_2} \right) + \gamma(P_2), \quad \text{if} \quad 1 \leq \mu \leq \mu_1^* \tag{28a}
\]
\[
\mu R_1 + R_2 \leq f(P_1, P_2, a, \mu), \quad \text{if} \quad \mu_1^* < \mu < \mu_0^* \tag{28b}
\]
\[
\mu R_1 + R_2 \leq \mu \gamma(P_1) + \gamma \left( \frac{a P_2}{1 + P_1} \right), \quad \text{if} \quad \mu \geq \mu_0^* \tag{28c}
\]
where \( \mu_0^*, \mu_1^*, \) and \( f(P_1, P_2, a, \mu) \) are given in (24a), (24b), and (27), respectively.

As can be seen, the power allocation parameter \( \beta \) does not appear in this representation. This is because the optimal value of \( \beta \) for different ranges of \( \mu \) is found in (25). More importantly, the optimum weighted sum-rate of the HK scheme, for the one-sided interference channel, is revealed for any \( \mu \geq 1 \).

**V. A Tight Outer Bound**

In consideration of the new representation of the HK region in Lemma 6 and seeing that the first two bounds are trivial, a natural question is whether or not we can establish a similar upper bound on the \( \mu \)-sum rate of the one-sided interference channel. To this end, we make use of the below bound for the \( \mu \)-sum rate. Using Fano’s inequality (21), for any codebook of block length \( n \) we can write
\[
n(\mu R_1 + R_2) \leq \mu I(X_1^n; Y_1^n) + I(X_2^n; Z_2^n) + n \epsilon_n
\]
\[
= \mu h(Y_1^n) - \mu h(Y_1^n | X_1^n) + h(Y_2^n) - h(Y_2^n | X_2^n) + n \epsilon_n
\]
\[
= \mu h(X_1^n + \sqrt{a} X_2^n + Z_1^n) - \mu h(\sqrt{a} X_2^n + Z_1^n) + h(X_2^n + Z_2^n) - h(Z_2^n) + n \epsilon_n, \tag{29}
\]
in which \( \epsilon_n \to 0 \) as \( n \to \infty \).

To maximize this bound we need to determine both the input distributions that maximize (29) and the optimal power at each transmitter. We observe that in the previous works, e.g., [9]–[11], the term \( \mu h(X_1^n + \sqrt{a} X_2^n + Z_1^n) \) has been maximized separately from \( h(X_2^n + Z_2^n) - \mu h(\sqrt{a} X_2^n + Z_1^n) \). With this arrangement, the latter terms are in the form of \( h(X + Z_1) - \mu h(X + Z_2) \) which is proved, in [19] Theorem 1], to have a Gaussian \( X \) as an optimal solution, for \( \mu \geq 1 \). Furthermore, in [10] and [11], it is shown that for \( 0 \leq \mu \leq \frac{P_2 + 1/a}{P_2 + 1} \) the optimal covariance matrix of \( X_2^n \) is \( P_2 I \). Besides, the term \( h(X_1^n + \sqrt{a} X_2^n + Z_1^n) \) is maximized by Gaussian \( X_1^n \) and \( X_2^n \), in which
$P_1I$ and $P_2I$ are the optimal covariance matrices, respectively. Thus, the upper bound can be represented as in Proposition 9.

Although such an optimization provides a valid upper 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 simply, can Gaussian inputs be an optimal solution in maximizing all three terms in (29)\(^1\)? If not, can we arrange those terms in a way that this arrangement leads to a tighter upper 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 subsection.

A. New Formulation for the Upper Bound

It can be proved that Gaussian inputs cannot be optimal for (29)\(^2\). Nevertheless, we can write it in the following form

$$n(\mu R_1 + R_2) \leq \mu I(X^n_1; Y^n_1) + I(X^n_2; Y^n_2) + n\epsilon_n$$

$$= \mu h(Y^n_1) - \mu h(Y^n_1 | X^n_1) + h(Y^n_2) - h(Y^n_2 | X^n_2) + n\epsilon_n$$

$$= \mu h(X^n_1 + Z^n_1) + h(X^n_2 + Z^n_2) - h(X^n_1 + Z^n_1) - h(X^n_2 + Z^n_2) + n\epsilon_n$$

$$\leq (\mu - 1)h(X^n_1 + \sqrt{\alpha}X^n_2 + Z^n_1) + \frac{n}{2} \log [2\pi e(P_1 + aP_2 + 1)] + h(X^n_2 + Z^n_2)$$

$$- \mu h(\sqrt{\alpha}X^n_2 + Z^n_1) - \frac{n}{2} \log 2\pi e + n\epsilon_n$$

$$= (\mu - 1)h(X^n_1 + \sqrt{\alpha}X^n_2 + Z^n_1) + h(X^n_2 + Z^n_2) - \mu h(\sqrt{\alpha}X^n_2 + Z^n_1) + n\epsilon_n,$$  

(30)

where the second inequality follows from the fact that the Gaussian distribution maximizes the differential entropy, under a covariance constraint. We next maximize $h(X^n_2 + Z^n_2) - \mu h(\sqrt{\alpha}X^n_2 + Z^n_1) + (\mu - 1)h(X^n_1 + \sqrt{\alpha}X^n_2 + Z^n_1)$ altogether. To begin with, let us define the objective function as

$$W_o \triangleq h(X_2 + Z_2) - \mu h(\sqrt{\alpha}X_2 + Z_1) + (\mu - 1)h(X_1 + \sqrt{\alpha}X_2 + Z_1),$$  

(31)

in which $X_1$ and $X_2$ are independent random $n$-vectors.\(^4\) Then, maximizing the right-hand side of (30) is equivalent to the following optimization problem:

$$W = \maximize_{p(x)} W_o$$

subject to  

$$\text{tr}(K_{X_1}) \leq nP_1,$$  

$$\text{tr}(K_{X_2}) \leq nP_2,$$  

(32)

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).

---

\(^4\)To simplify the notation, $X$ is used instead of $X^n$. That is, both $X$ and $X^n$ represent a random $n$-vector.
The first step in solving (32) is to determine the optimal distributions of \( X_1 \) and \( X_2 \). Then, we need to determine the optimal covariance matrices. These two are the subjects of the following subsections.

### B. Optimal Input Distributions

In light of the extremal inequality developed in Lemma 4, it is straightforward to prove that a jointly Gaussian \((X_1, X_2)\) is an optimal solution of (32). Comparing the optimization problem in (32) with Lemma 4, it can be seen that \( W \) is an special case of Lemma 4 where \( K = 2, \mu_0 = -\mu, \mu_1 = 1, \mu_2 = \mu - 1, m = 2, \) and

\[
\begin{align*}
v_0 &= \begin{bmatrix} 0 \\ \sqrt{a} \end{bmatrix}, \quad v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ \sqrt{a} \end{bmatrix}.
\end{align*}
\]

(33)

Note that only \( \mu_0 < 0 \), since \( \mu \geq 1 \). Also, for \( a \neq 0 \) all \( v_k \)'s are nonzero. Therefore, all conditions of Lemma 4 are satisfied. Thus, a jointly Gaussian \((X_1, X_2)\) is an optimal solution of the optimization problem in (32).

Again, we highlight that the covariance constraint subsumes the average power constraint and thus the constraints of the optimization problem in Lemma 4 are more general than those in (32). Thus, Lemma 4 can be easily applied to (32), for any \( \mu \geq 1 \) and \( 0 < a < 1 \). With this in mind, we find the optimal covariance matrices in the following subsection.

### C. Evaluation of the Upper Bound

Up to now, we have proved that Gaussian distributions are optimal both for \( X_1^n \) and \( X_2^n \), in (30). In this subsection, we find the optimal covariance matrices for \( X_1^n \) and \( X_2^n \) and evaluate the upper bound in (30), for those inputs. Equivalently, we can find the optimum value of \( W_o \) in (32), and substitute it into (30). We choose the latter course.

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

\[
W^G = \max \left\{ \frac{1}{2} \log |(2\pi e)^n|K_{X_2} + I| - \frac{\mu}{2} \log |(2\pi e)^n|aK_{X_2} + I| \\
+ \frac{\mu - 1}{2} \log |(2\pi e)^n|K_{X_1} + aK_{X_2} + I| \right\}
\]

s.t.  \( K_{X_1} \succeq 0, \text{tr}(K_{X_1}) \leq nP_1, \)
\( K_{X_2} \succeq 0, \text{tr}(K_{X_2}) \leq nP_2. \)

Note that the covariance matrices of \( Z_1, Z_2 \) and \( Z_3 \) are equal to \( I \), the identity matrix of size \( n \). The objective function can be simplified as

\[
\frac{1}{2} \log |K_{X_2} + I| - \frac{\mu}{2} \log |aK_{X_2} + I| + \frac{\mu - 1}{2} \log |K_{X_1} + aK_{X_2} + I|.
\]

(34)

Besides, since \( K_{X_1} \) and \( K_{X_2} \) are positive semidefinite, we can decompose them as \( K_{X_1} = U_1 \Lambda_1 U_1^T \) and \( K_{X_2} = U_2 \Lambda_2 U_2^T \), where \( U_1 \) and \( U_2 \) are unitary matrices and \( \Lambda_1 \) and \( \Lambda_2 \) are diagonal matrices with nonnegative entries. Let \( \lambda_{i1} \) and \( \lambda_{i2} \), \( i = 1, \ldots, n \), be the diagonal elements of \( \Lambda_1 \) and \( \Lambda_2 \), respectively. Note that these are the eigenvalues of \( K_{X_1} \) and \( K_{X_2} \), respectively. Also, without loss of generality, let \( \lambda_{11} \leq \lambda_{12} \leq \ldots \leq \lambda_{1n} \) but \( \lambda_{21} \geq \lambda_{22} \geq \ldots \geq \lambda_{2n} \). From the fact that \( U_2 U_2^T = I \) and knowing that for any matrices \( A \) and \( B \) we have
However, \( |AB + I| = |BA + I| \), we obtain \( |KX_2 + I| = |\Lambda_2 + I| \) and \( |aKX_2 + I| = |a\Lambda_2 + I| \).

However, \( |KX_1 + aKX_2 + I| = |\Lambda_1 + a\Lambda_2 + I| \) is not correct in general, and all we can have is
\[
W^G = \max \left\{ \frac{1}{2} \log (|\Lambda_2 + I|) - \frac{\mu}{2} \log (|a\Lambda_2 + I|) + \frac{\mu - 1}{2} \log (|U_1\Lambda_1U_1^* + aU_2\Lambda_2U_2^* + I|) \right\}
\]
\[
\text{s.t. } \Lambda_1 \geq 0, \; \text{tr} (\Lambda_1) \leq nP_1,
\]
\[
\Lambda_2 \geq 0, \; \text{tr} (\Lambda_2) \leq nP_2.
\]

To further simplify the objective function, we can write it based on the eigenvalues of \( KX_1 \) and \( KX_2 \), using
\[
|KX_1 + aKX_2 + I| = \prod_{i=1}^n (\lambda_{2i} + 1) \quad \text{and} \quad |aKX_2 + I| = \prod_{i=1}^n (a\lambda_{2i} + 1).\]

In addition, the last term can be upper bounded as\(^5\)
\[
|KX_1 + aKX_2 + I| \leq \prod_{i=1}^n (\lambda_{1i} + a\lambda_{2i} + 1).
\]

As a result, we will get the new optimization problem
\[
\hat{W}^G = \max \left\{ \frac{1}{2} \sum_{i=1}^n \left[ \log (\lambda_{2i} + 1) - \mu \log (a\lambda_{2i} + 1) + (\mu - 1) \log (\lambda_{1i} + a\lambda_{2i} + 1) \right] \right\}
\]
\[
\text{s.t. } \lambda_{1i} \geq 0, \; \lambda_{2i} \geq 0, \; \forall i,
\]
\[
\sum_{i=1}^n \lambda_{1i} \leq nP_1,
\]
\[
\sum_{i=1}^n \lambda_{2i} \leq nP_2,
\]
whose optimum value is an upper bound for \( W^G \). That is,
\[
W^G \leq \hat{W}^G.
\]

To solve this new optimization problem, we use Lagrange multipliers \( u_1, u_2, v_1 = \{v_{11}, \ldots, v_{1n}\} \) and \( v_2 = \{v_{21}, \ldots, v_{2n}\} \), and study the Lagrangian defined by
\[
L(\lambda_1, \lambda_2, u_1, u_2, v_1, v_2) = \frac{1}{2} \sum_{i=1}^n \left[ \log (\lambda_{2i} + 1) - \mu \log (a\lambda_{2i} + 1) + (\mu - 1) \log (\lambda_{1i} + a\lambda_{2i} + 1) \right]
\]
\[
+ u_1(nP_1 - \sum_{i=1}^n \lambda_{1i}) + u_2(nP_2 - \sum_{i=1}^n \lambda_{2i}) + \sum_{i=1}^n v_{1i}\lambda_{1i} + \sum_{i=1}^n v_{2i}\lambda_{2i}.
\]

Since the constraints are inequalities, we examine the Karush-Kuhn-Tucker (KKT) conditions. The KKT stationary

\(^5\) Let \( \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_n \) and \( \beta_1 \leq \beta_2 \leq \ldots \leq \beta_n \) be the eigenvalues of Hermitian \( n \times n \) matrices \( A \) and \( B \). Then, if \( \alpha_1 + \beta_1 + \theta \geq 0 \), where \( \theta \) is a real number, we have\(^{[22]}\)
\[
|A + B + \theta I| \leq \prod_{i=1}^n (\alpha_i + \beta_{n+1-i} + \theta).
\]
and complementary slackness constraints for \( L(\lambda_1, \lambda_2, u_1, u_2, v_1, v_2) \) are given by

\[
\begin{align*}
\frac{\mu-1}{\lambda_1 a + a \lambda_2 + 1} - u_1 + v_1 &= 0, \quad \text{(39a)} \\
\frac{1}{\lambda_2 + 1} - \frac{\alpha \mu}{a \lambda_2 + 1} + \frac{a (\mu - 1)}{\lambda_1 + a \lambda_2 + 1} - u_2 + v_2 &= 0, \quad \text{(39b)} \\
u_1 (nP_1 - \sum_{i=1}^{n} \lambda_{1i}) &= 0, \quad \text{(39c)} \\
u_2 (nP_2 - \sum_{i=1}^{n} \lambda_{2i}) &= 0, \quad \text{(39d)} \\
v_1 \lambda_{1i} &= 0, \quad \text{(39e)} \\
v_2 \lambda_{2i} &= 0, \quad \text{(39f)}
\end{align*}
\]

for \( i = 1, \ldots, n \). Next, we observe that all \( \lambda_{1i} \)s have the same role in the optimization problem; thus, \( \lambda_{1i} = \cdots = \lambda_{1n} \triangleq \lambda_1 \). With the same reasoning, \( \lambda_{2i} = \cdots = \lambda_{2n} \triangleq \lambda_2 \), and (39) simplifies to

\[
\begin{align*}
\frac{\mu-1}{\lambda_1 a + a \lambda_2 + 1} - u_1 + v_1 &= 0, \quad \text{(40a)} \\
\frac{1}{\lambda_2 + 1} - \frac{\alpha \mu}{a \lambda_2 + 1} + \frac{a (\mu - 1)}{\lambda_1 + a \lambda_2 + 1} - u_2 + v_2 &= 0, \quad \text{(40b)} \\
u_1 (P_1 - \lambda_1) &= 0, \quad \text{(40c)} \\
u_2 (P_2 - \lambda_2) &= 0, \quad \text{(40d)} \\
v_1 \lambda_1 &= 0, \quad \text{(40e)} \\
v_2 \lambda_2 &= 0, \quad \text{(40f)}
\end{align*}
\]

To solve (40), we consider the cases \( \mu = 1 \) and \( \mu > 1 \) separately. For \( \mu = 1 \), we only need to find \( \lambda_2 \). From (40b), we can see that, for \( i = 1, \ldots, n \), \( u_2 - v_{2i} = \frac{1}{\lambda_2 + 1} - \frac{\alpha}{a \lambda_2 + 1} > 0 \), as \( 0 < a < 1 \); hence, \( u_2 > v_{2i} \). But, \( v_{2i} \geq 0 \) and thus \( u_2 > 0 \). This, together with (40d), implies \( \lambda_2^* = P_2 \).

Next, for \( \mu > 1 \), from (40a) we have \( u_1 - v_{1i} = \frac{\mu - 1}{\lambda_1 a + a \lambda_2 + 1} \). Since \( \mu > 1 \), \( \lambda_1 \geq 0 \), and \( \lambda_2 \geq 0 \), from (40a) we conclude that \( u_1 - v_{1i} > 0 \); on the other hand, \( v_{1i} \geq 0 \) for \( i = 1, \ldots, n \), and this implies \( u_1 > 0 \). Then, from (40c) we can see that

\[
\lambda_1^* = P_1,
\]

and therefore \( K_{X_1} = P_1 I \). This, immediately, implies that

\[
\tilde{W}^G = \tilde{W}^G,
\]

since for \( K_{X_1} = P_1 I \), the inequality in (36) becomes an equality, and thus (38) simplifies to \( W^G = \tilde{W}^G \).
Having found the optimal value of $\lambda_1$, the set of equations $[40]$ reduces to

\[
\begin{align*}
\frac{1}{\lambda_2 + N_1} - & \frac{\mu}{\lambda_2 + N_2} + \frac{\mu - 1}{\lambda_2 + N_3} - u_2 + v_{2i} = 0, \\
u_2(P_2 - \lambda_2) = & \ 0, \\
v_{2i}\lambda_2 = & \ 0,
\end{align*}
\]  

(43a) (43b) (43c)

where $N_1 = 1$, $N_2 = \frac{1}{a}$, and $N_3 = \frac{1 + P_1}{a}$, for $i = 1, \ldots, n$. Since $0 < a < 1$, we can see that $N_1 < N_2 \leq N_3$. Next, from (43a), for $i = 1, \ldots, n$, we have $u_2 - v_{2i} = \frac{1}{\lambda_2 + N_1} - \frac{\mu}{\lambda_2 + N_2} + \frac{\mu - 1}{\lambda_2 + N_3}$. The right-hand side of $u_2 - v_{2i}$ becomes zero for

\[
\mu^*(\lambda_2) = \frac{\frac{1}{\lambda_2 + N_1} - \frac{1}{\lambda_2 + N_2}}{\frac{\frac{1}{\lambda_2 + N_1} - \frac{1}{\lambda_2 + N_3}}{\lambda_2 + N_1 N_3 - N_2}},
\]  

(44)

or equivalently for

\[
\lambda_2 = \frac{(N_2 - N_1)\mu}{\mu - \frac{N_3 - N_1}{N_3 - N_2}} - N_2.
\]  

(45)

But, we know that $0 \leq \lambda_2 \leq P_2$. Thus, we first determine the values of $\mu$ corresponding to $\lambda_2 = 0$ and $\lambda_2 = P_2$. These are basically the maximum and minimum of $\mu^*$ for $0 \leq \lambda_2 \leq P_2$. For $N_1 < N_2 \leq N_3$, $\mu^*$ is a decreasing function of $\lambda_2$ and we have

\[
\begin{align*}
\mu^*_\text{min} &= \mu^*(P_2) = \frac{P_2 + N_2 N_3 - N_1}{P_2 + N_1 N_3 - N_2}, \\
\mu^*_\text{max} &= \mu^*(0) = \frac{N_2 N_3 - N_1}{N_1 N_3 - N_2}.
\end{align*}
\]  

(46a) (46b)

Then, for $\mu > \mu^*_\text{max}$ we have $u_2 - v_{2i} < 0$, i.e., $v_{2i} > u_2$. But, $u_2 \geq 0$ and thus $v_{2i}$ cannot be 0. This implies that $\lambda_2 = 0$ for $\mu > \mu^*_\text{max}$. Similarly, if $\mu < \mu^*_\text{min}$ then $u_2 - v_{2i} > 0$; i.e., $u_2 > v_{2i} \geq 0$. Thus, $u_2$ cannot be 0. This in turn implies $\lambda_2 = P_2$. Otherwise, (for $0 < \lambda_2 < P_2$), the optimal value of $\lambda_2$ is given by [45]. Hence,

\[
\lambda_2^* = \begin{cases} 
P_2, & \text{if } 1 \leq \mu \leq \mu^*_\text{min} \\
\frac{(N_2 - N_1)\mu}{\mu - \frac{N_3 - N_1}{N_3 - N_2}} - N_2 & \text{if } \mu^*_\text{min} < \mu < \mu^*_\text{max} \\
0, & \text{if } \mu \geq \mu^*_\text{max}
\end{cases}
\]  

(47)

Finally, since $N_1 = 1$, $N_2 = \frac{1}{a}$, $N_3 = \frac{1 + P_1}{a}$, the optimum objective function of $\hat{W}^G$ in [37] (and equivalently $W^G$) will be

\[
W^* = \begin{cases} 
(\mu - 1)n\gamma\left(\frac{P_1}{1 + aP_2}\right) + \frac{n}{2} \log \left(\frac{1 + P_1}{1 + aP_2}\right), & \text{if } 1 \leq \mu \leq \mu^*_\text{min} \\
n f(P_1, P_2, a, \mu) - n\gamma(P_1 + aP_2), & \text{if } \mu^*_\text{min} < \mu < \mu^*_\text{max} \\
(\mu - 1)n\gamma(P_1), & \text{if } \mu > \mu^*_\text{max}
\end{cases}
\]  

(48)

in which $f(P_1, P_2, a, \mu)$ is defined in [27] 6. From (46a) and (46b), it is also easy to check that $\mu^*_\text{max} = \mu_0^*$ and

6 It is instructive to compare $\lambda_2^*$ with $\beta^* P_2$, where $\beta^*$ is given by in [25].
\[ \mu_{\text{min}}^* = \mu_1^*, \text{ defined in (24a) and (24b)}. \] Eventually, from (30), when \( n \to \infty \), we obtain
\[
\mu R_1 + R_2 \leq \frac{1}{n} W^* + \gamma (P_1 + a P_2) \\
= \begin{cases} 
\mu \gamma \left( \frac{P_1}{1 + a P_2} \right) + \gamma (P_2), & \text{if } 1 \leq \mu \leq \mu_{\text{min}}^* \\
\tilde{f}(P_1, P_2, a, \mu), & \text{if } \mu_{\text{min}}^* < \mu < \mu_{\text{max}}^* \\
\mu \gamma (P_1) + \gamma \left( \frac{a P_2}{1 + P_1} \right), & \text{if } \mu \geq \mu_{\text{max}}^* 
\end{cases} 
\] (49)

D. The Outer Bound for \( a \geq 1 \)

Although the capacity region of the one-sided interference channel is known for \( a \geq 1 \), it is instructive to evaluate the upper bound we developed in (30) for this range too. This is very similar to what we did in the previous subsection, except that \( N_1 \leq N_2 \leq N_3 \) is no longer valid. Instead, when solving (43a)-(43c) the following two cases arise:

1) \( N_2 \leq N_1 \leq N_3 \) if \( 1 \leq a \leq 1 + P_1 \), and
2) \( N_2 \leq N_3 \leq N_1 \) if \( a \geq 1 + P_1 \).

In both cases, it is straightforward to show that the optimal value of \( \lambda_2 \) is zero. Hence, similar to the case \( \lambda_2 = 0 \) in (47)-(49), we end up with the following bound for any \( a \geq 1 \):
\[
\mu R_1 + R_2 \leq \mu \gamma (P_1) + \gamma \left( \frac{a P_2}{1 + P_1} \right). 
\] (50)

This bound, for \( \mu = 1 \), together with the trivial single-user bounds \( R_1 \leq \gamma (P_1) \) and \( R_2 \leq \gamma (P_2) \) make an outer bound which is the same as the capacity region in Proposition 7. Hence, it provides an alternative proof of the capacity region in the strong interference regime. We know that this region is achievable by decoding interference and cancelling it at the interfered-with receiver.

VI. MAIN RESULTS

Based on the optimization problems in Sections IV and V, we are able to fully characterize the capacity region of the one-sided interference channel. In addition, we establish the capacity region of the interference channel for a large range of the mixed interference regime, including the degraded interference case.
A. The Capacity of the One-Sided Interference Channel

Theorem 1. The capacity region of the one-sided interference channel in the weak interference regime is the set of \((R_1, R_2)\) such that

\[
R_1 \leq \gamma(P_1), \quad (51a)
\]

\[
R_2 \leq \gamma(P_2), \quad (51b)
\]

\[
\mu R_1 + R_2 \leq \mu \gamma\left(\frac{P_1}{1 + aP_2}\right) + \gamma(P_2), \quad \text{if } 1 \leq \mu \leq \mu_1^* \quad (51c)
\]

\[
\mu R_1 + R_2 \leq f(P_1, P_2, a, \mu), \quad \text{if } \mu_1^* < \mu < \mu_0^* \quad (51d)
\]

\[
\mu R_1 + R_2 \leq \mu \gamma(P_1) + \gamma\left(\frac{aP_2}{1 + P_1}\right), \quad \text{if } \mu \geq \mu_0^* \quad (51e)
\]

where \(\mu_0^* = \frac{1 + P_1 - a}{aP_1}, \mu_1^* = \frac{1 + aP_2}{1 + P_2} \mu_0^* \) and \(f(P_1, P_2, a, \mu) = \mu \gamma\left(\frac{P_1}{1 + a \frac{P_2}{\mu - \mu_0^*}}\right) + \gamma\left(\frac{aP_2 - a \frac{\mu - \mu_0^*}{\mu - a \mu_0^*}}{1 + P_1 + a \frac{P_2}{\mu - \mu_0^*}}\right) + \gamma\left(\frac{\mu \frac{P_2}{\mu - a \mu_0^*}}{1 + a \frac{P_2}{\mu - \mu_0^*}}\right)\).

Proof: We have already proved the achievability and converse in Sections IV and V, respectively; we summarize the proof here. The achievability is proved in Lemma 6, which is an alternative representation of the Han-Kobayashi region. The converse for (51c)-(51e) is established by formulating a new upper bound on the weighted sum-rate in (30), proving that Gaussian inputs are optimal for that, finding the optimal covariance matrices for the inputs, and evaluating the outer bound which results in (49). 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 6. This completes the proof.

This is the first capacity result for the Gaussian interference channel in the weak interference regime. The reader will have noticed that the capacity region introduced in Theorem 1 has two other representations, given by Propositions 5 and 6. The equivalency of these three regions has already been established and used to obtain the main result in Theorem 1. Each of these representations has its advantages. An advantage of the representations in Theorem 1 is the fact that it eliminates \(\beta\), and thus the need for convexification, if needed, and unification. In addition, it reveals the optimum weighted sum-rates, for any \(\mu \geq 1\). 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 a single set of equations. This is stated in the following theorem.

Theorem 2. The union of the set of rate pairs \((R_1, R_2)\) satisfying

\[
R_1 \leq \gamma\left(\frac{P_1}{1 + a \beta P_2}\right), \quad (52a)
\]

\[
R_2 \leq \gamma(P_2), \quad (52b)
\]

\[
R_1 + R_2 \leq \gamma\left(\frac{P_1 + a \beta P_2}{1 + a \beta P_2}\right) + \gamma(\beta P_2), \quad (52c)
\]

over \(\beta \in [0, 1]\) provides the capacity region of the one-sided interference channel for any \(a \in \mathbb{R}\).

Proof: By Lemma 6 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 7 we know
that setting $\beta = 0$ in (52a)-(52c) is optimal. This has also been proved in Section VI-A. Thus, the same set of inequalities can be used to represent the capacity of the one-sided interference channel for the whole range of $a$.

\section*{B. New Outer Bounds for the Interference Channel}

Consider the interference channel represented in Fig 1. The interference links, whose gains are $a$ and $b$, belong to weak or strong interference regimes, making four different combinations. Removing the interference links one at a time, and considering that each link represents either weak or strong interference, we obtain four types of one-sided interference channels as shown in Fig. 4. Hence, we can develop four outer bounds on the capacity region of the interference channel each of which is valid for one of the above ranges of $a$ and $b$, and collectively cover all ranges of $a$ and $b$. The key to making these outer bounds is the fact that the capacity region of the interference channel is outer-bounded by the capacity region of the corresponding one-sided interference channels, obtained by removing the interference links one at a time. With this, we are ready to state our main results for the interference channel,
in the remainder of this section.

We first remove the interference link with gain \( a \), and use the capacity of the corresponding one-sided channel (type I or II, depending on the value of \( a \)) as an outer bound for the original channel. We repeat this process with the other interference link to get a one-sided channel of type III or IV. The combination of the two outer bounds results in the following outer bound.

**Theorem 3.** The capacity region of the interference channel with \( a < 1 \) and \( b \geq 1 \) is outer bounded by the convex hull of the union of rate pairs \((R_1, R_2)\) satisfying

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

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

\[
R_1 + R_2 \leq \gamma (bP_1 + P_2),
\]

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

over \( \beta \in [0, 1] \).

**Proof:** The above outer bound is made of the combination of the capacity regions of the corresponding one-sided interference channels, i.e., type I and type IV, provided in Fig. 4. Note that we have removed the redundant constraints.

By symmetry, the outer bound for the interference channel in the other mixed interference case \((a \geq 1 \) and \( b < 1 \)) is obtained by swapping the index 1 with 2, \( a \) with \( b \), and \( \alpha \) with \( \beta \) in Theorem 3 and is omitted here.

**Theorem 4.** The capacity region of the interference channel in the weak interference regime is outer bounded by the convex hull of the union of rate pairs \((R_1, R_2)\) satisfying

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

\[
R_2 \leq \gamma \left( \frac{P_2}{1 + b\alpha P_1} \right),
\]

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

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

over \( \alpha, \beta \in [0, 1] \).

**Proof:** This outer bound is made up of the combination of the capacity regions of type I and type III one-sided interference channels. Again, we have removed the redundant constraints.

Obviously, with the same rationale, we can make an outer bound for the interference channel in the strong interference regime. The resulting outer bound is tight, but it is not new as seen in Proposition 2.
C. Capacity Results for the Interference Channel

**Theorem 5.** The capacity region of the interference channel with $b \geq 1 + P_2$ and $a \leq 1$ is the union of the following rate region for $\beta \in [0, 1]$:

\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (55a)
\]
\[
R_2 \leq \gamma(P_2), \quad (55b)
\]
\[
R_1 + R_2 \leq \gamma(bP_1 + P_2), \quad (55c)
\]
\[
R_1 + R_2 \leq \gamma \left( \frac{P_1 + a\beta P_2}{1 + a\beta P_2} \right) + \gamma(P_2). \quad (55d)
\]

**Proof:** The converse follows readily from Theorem 3. To prove the achievability, we note that receiver 1 can first decode the common part of interference at a rate of $\gamma \left( \frac{P_1}{1 + a\beta P_2} \right)$. On the other hand, we see that if $\frac{bP_1}{1 + P_2} \geq \frac{P_1}{1 + a\beta P_2}$, then receiver 2 is capable of simultaneously decoding both messages, without incurring any penalty on $R_1$ or $R_2$. This is because receiver 2 can decode the message of user 1, without any loss, before its own message. The above condition is equivalent to $b + ab\beta P_2 \geq 1 + P_2$, which is valid for any $\beta$ if $b \geq 1 + P_2$. This completes the proof. The reader should have noticed that (55c) is redundant, and can be removed, in the above capacity region. 

**Theorem 6.** The capacity region of the interference channel satisfying $b \geq \frac{1 + P_2}{1 + aP_2}$ and $a \leq 1$ is the same as the capacity region of the one-sided interference channel with the same $a$ (but $b = 0$), and is given by the union of the set of rate pairs satisfying

\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (56a)
\]
\[
R_2 \leq \gamma(P_2), \quad (56b)
\]
\[
R_1 + R_2 \leq \gamma \left( \frac{P_1 + a\beta P_2}{1 + a\beta P_2} \right) + \gamma(P_2), \quad (56c)
\]

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

**Proof:** Consider the outer bound given by Theorem 3. We add one more constraint to this outer bound to obtain

\[
R_1 \leq \gamma \left( \frac{P_1}{1 + a\beta P_2} \right), \quad (57a)
\]
\[
R_2 \leq \gamma(P_2), \quad (57b)
\]
\[
R_1 + R_2 \leq \gamma(P_2) + \gamma \left( \frac{bP_1}{1 + P_2} \right), \quad (57c)
\]
\[
R_1 + R_2 \leq \gamma(P_2) + \gamma \left( \frac{P_1}{1 + aP_2} \right), \quad (57d)
\]
\[
R_1 + R_2 \leq \gamma \left( \frac{P_1 + a\beta P_2}{1 + a\beta P_2} \right) + \gamma(P_2), \quad (57e)
\]
where the new constraint (57d) comes from the sum capacity of the one-sided interference channel, given by Proposition 8. With this, it can be seen that (57c) is redundant for \( b \geq 1 + \frac{P_2}{1 + aP_2} \). Meanwhile, (57d) is also redundant in the presence of (57e) because, at best, for \( \beta = 1 \), (57e) becomes equal to (57d). The remaining constraints, namely, (57a), (57b), and (57e) provide the capacity region of the one-sided interference channel with \( b = 0 \). This completes the proof of the converse part. In the meantime, it suggests the achievability scheme too, because the outer bound in Theorem 3 can be achieved by HK encoding where transmitter 1 uses only the private message while transmitter 2 uses both the private and common messages.

Remark 2. This is the first capacity region for the Gaussian interference channel in which both private and common messages are required to achieve the capacity. This is not however astonishing after seeing that to establish the capacity region of the one-sided interference channel, with weak interference, both of the messages are used.

As a corollary of Theorem 6 it can be seen that the sum-capacity of the interference channel in the mixed interference regime with \( a < 1 \) and \( b \geq 1 + \frac{P_2}{1 + aP_2} \) is given by

\[
C_{\text{sum}} = \gamma(P_2) + \gamma\left(\frac{P_1}{1 + aP_2}\right).
\]

\[(58)\]

D. The Capacity of the Degraded Interference Channel

Recall that \( ab = 1 \) for the degraded interference channel. With this, for \( a \leq 1 \) we can see that \( b = \frac{1}{a} \geq \frac{1 + P_2}{1 + aP_2} \) falls within the range for which the capacity region is known from Theorem 6. Therefore, Theorem 6 includes the capacity region of the degraded channel. Alternatively, as mentioned in Section II-B, the capacity region of the degraded interference channel with \( 0 < a = \frac{1}{b} < 1 \) is equal to the capacity region of the one-sided interference channel with \( 0 < a < 1 \). Therefore the capacity region of the degraded interference channel is immediately established in favor of our new capacity result in Section VI-A.

It is interesting to see how the capacity region looks for different channel parameters. In Fig. 5 the capacity regions of the one-sided interference channel is plotted for \( a = 0.1 \), \( a = 0.4 \), and \( a = 1 \). As expected from the illustration of the HK achievable region for the one-sided interference channel in Fig. ??, as \( a \) increases, the corner point A shifts left while the corner point B shifts up. The capacity regions with the same channel gains but different transmission powers can be found in Figs. 6 and 7. Remember that the above regions also provide the capacity regions of the interference channel with the same \( a \) and any \( b \geq \frac{1 + P_2}{1 + aP_2} \).

E. Corner Points of the Capacity Region

The corner points of the capacity region of the interference channel are defined as the points on the capacity region where one of the users achieves its maximum rate under the constraint that the other user achieves its own interference-free capacity. Mathematically, we need to find \( R_1^* \) and \( R_2^* \) such that

\[
R_1^* = \max R_1 \quad \text{s.t.} \quad R_2 = \gamma(P_2),
\]

\[(59)\]
Fig. 5. Capacity regions of the one-sided inference channel ($b = 0$) for different channel gains with $P_1 = 1$ and $P_2 = 7$. It should be highlighted that the above regions also provide the capacity region of the interference channel with the same $\alpha$ and any $b \geq \frac{1 + P_2}{1 + aP_2}$.

Fig. 6. Capacity regions of the one-sided inference channel ($b = 0$) for different channel gains with $P_1 = 7$ and $P_2 = 1$. Again, the above regions also provide the capacity region of the interference channel with the same $\alpha$ and any $b \geq \frac{1 + P_2}{1 + aP_2}$.
Fig. 7. Capacity regions of the one-sided inference channel \((b = 0)\) for different channel gains with \(P_1 = 7\) and \(P_2 = 7\). The above regions also provide the capacity region of the interference channel with the same \(a\) and any \(b \geq \frac{1+P_2}{1+aP_2}\).

and

\[
R_2^* = \max R_2 \quad \text{s.t.} \quad R_1 = \gamma(P_1),
\]

(60)
to determine the corner points \((R_1^*, \gamma(P_2))\) and \((\gamma(P_1), R_2^*))\).

1) Corner points of the one-sided interference channel: In the strong interference regime the corner points of this channel are known from Proposition 7. In the weak interference case, as we discussed in Section IV and visualized in Fig. ??, the points A and B, correspond to the Han-Kobayashi region for \(\beta = 1\) and \(\beta = 0\), respectively. Then, in light of the capacity region in Theorem 1 these two points are the corner points of the capacity region for \(a < 1\), since from (51c) for \(R_2 = \gamma(P_2)\) we get \(R_1^* = \gamma\left(\frac{P_1}{1+aP_2}\right)\), and from (51c) we have \(R_2^* = \gamma\left(\frac{aP_2}{1+P_1}\right)\).

It should be mentioned that the point A has been known to be a corner point [16], [17], as is evident from Proposition 8 too. Also, the point B was stated to be the second corner point, by Costa [17]. Nevertheless, a problem was found in his proof [18], and since then this has been dubbed the “missing corner point.” Independent works in [23] and [24] also report that the point B is the second corner point.

2) Corner points of the interference channel: It is obvious that where the capacity region is known the corner points are known too. Hence, from Proposition 2 the corner points are known for \(a \geq 1\) and \(b \geq 1\). Additionally, from Proposition 3 it is seen that for \(a < 1\) and \(b \geq 1\) we have

\[
R_1^* = \min \left\{ \gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+P_2}\right) \right\},
\]

(61)
Next, to find $R^*_1$ for $a < 1$ and $b < 1$ we use the argument that the capacity region of the interference channel with $a < 1$ and $b < 1$ is outer bounded by the capacity regions of the two corresponding one-sided interference channels, i.e., types I and III in Fig. 4. This is because removing interference (letting $b = 0$ or $a = 0$) cannot reduce the capacity region. Then, from Theorem 1 for $R_2 = \gamma(P_2)$, we have $R_1 \leq \gamma\left(\frac{P_1}{1+aP_2}\right)$. Besides, by symmetry, from the capacity region of type III one-sided interference channel we have $R_1 \leq \gamma\left(\frac{bP_1}{1+aP_2}\right)$, for $R_2 = \gamma(P_2)$. Therefore, for $a < 1$ and $b < 1$, we get

$$R^*_1 = \min\left\{ \gamma\left(\frac{P_1}{1+aP_2}\right), \gamma\left(\frac{bP_1}{1+aP_2}\right) \right\} = \gamma\left(\frac{bP_1}{1+aP_2}\right).$$

(62)

Observe that $R^*_1$ is achievable by treating the signal of user 2 as noise, at both receivers. Then, receiver 2 can remove the signal of user 1, and decode its own signal free of interference; thus, $R_2 = \gamma(P_2)$ is achievable. Summarizing the results from (61) and (62) we have

$$R^*_1 = \begin{cases} \gamma\left(\frac{bP_1}{1+aP_2}\right), & \text{if } a < 1, \ 0 < b < \frac{1+aP_2}{1+aP_2} \\ \gamma\left(\frac{P_1}{1+aP_2}\right), & \text{if } a < 1, \ b \geq \frac{1+aP_2}{1+aP_2}. \end{cases}$$

(63)

The other corner point is characterized by

$$R^*_2 = \gamma\left(\frac{aP_2}{1+aP_1}\right),$$

(64)

for $0 < a < 1$ and any $b$. The converse is, again, due to (51e). To prove the achievability, we consider the following cases: 1) For $b < 1$, we can obtain the result from (62), by symmetry. 2) For $b \geq 1$, receiver 1 can decode the interference first and its own signal, free of interference, after that. Also, receiver 2 can decode the interference and cancel it. Thus, $R_2 = \gamma\left(\frac{aP_2}{1+aP_1}\right)$ and $R_1 = \gamma(P_1)$ are achievable. This completes the proof and thus the picture of corner points for the Gaussian interference channel.

VII. SUMMARY AND CONCLUDING REMARKS

In this section, we present a more intuitive description of the main results derived in this paper. In particular, we provide intuition about the HK scheme used in this paper, and optimal encoding and decoding for the cases in which we have established the capacity region.

A. Simple Han-Kobayashi Scheme

We have considered a communication scheme in which each user can split its message into two submessages of smaller rates and power. The transmitted message, for each user, is the superposition of its submessages and has total power of that user. The above communication scheme is the celebrated Han-Kobayashi scheme and the two submessages are known as the private and common messages. Since the general HK scheme is complicated, we have made the below simplifying assumptions:

1) We use Gaussian codebooks for the HK scheme, that is, we fixed the general input distribution to be a Gaussian distribution.

2) We exclude time-sharing.
With the above assumptions, the HK inner bound reduces to Proposition 1 in which only $\alpha$ and $\beta$ are to be optimized. However, still optimization over all possible HK strategies to get the largest achievable region is not an easy task. To make it even less complicated, we start with the HK scheme for the one-sided interference channel with $b = 0$.

B. Summary of the Main Results

By substituting $b = 0$ into the HK scheme in Proposition 1 we obtain a simpler region. With this, it is straightforward to verify that the optimal value of $\alpha$ is 1, and the number of active inequalities drops to three, as presented in Proposition 5. Letting the non-interfering user have a private message only ($\alpha = 1$) makes sense because even if there were a common message the other receiver could not exploit it, as there is no link between this user and the other receiver. The optimization of this simpler region reveals the optimal value of $\beta$, as presented in (25). From this equation it can be seen that $\beta^*$, the optimal percentage of power to be allocated to the private message of user 2, is a function of the relative importance of the users’ rates (e.g., $\mu$ in the weighted sum-rate $\mu R_1 + R_2$), their transmission powers, and the gain of the weak interference, which is $a$ here.

The above $\alpha$ and $\beta$ result in the largest achievable region, but are they optimal in terms of the capacity region? Favorably, the answer is positive, as we have proved this through long the optimization process in Sections IV and III and summarized the result in Theorem 1. Further, we have been able to prove the optimality of the above region for the interference channel with the same $a$ and $b \geq \frac{1 + P_2}{1 + a P_1}$. With these ranges of $a$ and $b$, the channel falls into the mixed interference regime. In view of this result, we divide the mixed interference regime into two parts:

- Mixed interference type I: $a < 1$ and $1 < b < \frac{1 + P_2}{1 + a P_1}$ or $b < 1$ and $1 \leq a < \frac{1 + P_2}{1 + b P_1}$
- Mixed interference type II: $a < 1$ and $b \geq \frac{1 + P_2}{1 + a P_1}$ or $b < 1$ and $a \geq \frac{1 + P_1}{1 + b P_1}$

These two regions are identified in Fig. 8(a). We have established the capacity region for the mixed interference type II, in Theorem 6. From Fig. 8(a) it is clear that the mixed interference type II contains the degraded interference regime, plotted in dashed blue lines. The capacity region is still open for the mixed interference type I. Owing to this work, the total ranges of $a$ and $b$ for which the capacity region is not known has shrunk to the inside of the thick solid lines in Fig. 8(a), i.e., only the weak interference and the mixed interference type I. Mathematically, this region can be characterized by

$$D = \left\{ (a, b) \mid 0 < a < \frac{1 + P_1}{1 + b P_1}, \quad 0 < b < \frac{1 + P_2}{1 + a P_2} \right\}.$$ (65)

C. Optimal Encoding and Decoding

The new capacity results established in this paper, on the whole, are based on decoding part of the interference and treating the rest of as noise. The optimum amount to be decoded varies from 0 to 100% of the interfering signal, depending on the relative importance of the users’ rates (i.e., ratio of weights in the weighted sum-rate),

\footnote{Recall that $\alpha$ controls what fraction of the total transmission power of the user 1 is allocated to its private information. $\beta$ has the same role for user 2. Then, $\alpha P_1$ and $\beta P_2$ represent the powers allocated to the common information of those users, respectively.}
their transmission powers, and the gain of the link with weak interference. We have explicitly found the optimal values of $\alpha$ and $\beta$ based on the above parameters. We depict them in Fig. [8(b)] in which $\beta^*$ is given by (25) and $\alpha^*$ is obtained by symmetry, i.e., by swapping $P_1$ with $P_2$, $a$ with $b$, and $\alpha$ with $\beta$ in (25).

1) One-sided interference channel: In terms of the one-sided interference channel with $b = 0$, Theorem 1 indicates that when $a \leq 1$ treating interference as noise achieves the weighted sum-capacity $\mu R_1 + R_2$ of the one-sided interference channel for any $1 \leq \mu \leq \frac{1 + P_2}{1 + P_2} \frac{1 + P_2 - a}{P_1}$. Besides, from (51e) we can see that for $\mu \geq \frac{1 + P_2 - a}{a P_2}$, decoding interference, and cancelling 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 for $a > 1$, simply 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 [17], as stated in Proposition 7.

A closer look at the representation of the capacity region in Proposition 5 reveals that the receiver under weak interference first decodes the common part of the interference and subtracts it from the received signal, and then decodes its own signal by treating the private part of the interference as noise. The best rate with which the common message is decoded at receiver 1 is equal to $\gamma\left(\frac{a \beta P_2}{1 + P_1 + a \beta P_2}\right)$; the remaining part of the interference, i.e., $a \beta P_2$, is treated as noise. Consequently, 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 $\gamma\left(\frac{\beta P_2}{1 + \beta P_2}\right)$, 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 $\gamma\left(\frac{a \beta P_2}{1 + P_1 + a \beta P_2}\right) \leq \gamma\left(\frac{\beta P_2}{1 + \beta P_2}\right)$ for $a \leq 1$. Therefore, the overall rate at receiver 2 is $R_2 \leq \gamma\left(\frac{a \beta P_2}{1 + P_1 + a \beta P_2}\right) + \gamma(\beta P_2)$, and the capacity region can be represented by the union of the following constraints:

\begin{align}
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).
\end{align}

2) Interference channel: When $b \geq \frac{1 + P_2}{1 + a P_2}$, receiver 1 will have the same decoding strategy as for $b = 0$. That is, it first decodes the common message of transmitter 2 at a rate of $\gamma\left(\frac{a \beta P_2}{1 + P_1 + a \beta P_2}\right)$, and then, it decodes its respective message at a rate of $R_1 = \gamma\left(\frac{P_1}{1 + a \beta P_2}\right)$, treating the remaining part of the interference, i.e., $a \beta P_2$, as noise.

For receiver 2 we may treat the following two cases differently:

Case 1 ($b \geq 1 + P_2$): Receiver 2 is able to first decode its interfering signal at a rate of $\gamma\left(\frac{b P_2}{1 + P_2}\right)$, which is greater than $\gamma\left(\frac{P_2}{1 + a \beta P_2}\right)$ for $b \geq 1 + P_2$. Hence, it can subtract $X_1$ from $Y_2$ and obtain the cleaner output $Y_2 = X_2 + Z_2$. Consequently, the capacity of the interference channel for $a < 1$, $b \geq 1 + P_2$ is the same as the capacity of the one-sided interference channel with $a < 1$, $b = 0$.

Case 2 ($b \geq \frac{1 + P_2}{1 + a P_2}$): Receiver 2 is capable of decoding its respective common message at a rate of $\gamma\left(\frac{a \beta P_2}{1 + P_1 + \beta P_2}\right)$. This is larger than the rate of the common message at receiver 1, i.e., $\gamma\left(\frac{a \beta P_2}{1 + a \beta P_2}\right)$, for $a < 1$. After that, it can
(a) State and classes of the interference channel

(b) Optimal values of $\alpha$ and $\beta$

Fig. 8. State of the interference channel in different classes. While before this work, the capacity region was known only for the strong interference case, now we know the capacity for a large part of the mixed interference regime (denoted by “Mixed II” in Fig. 8(a), as well as for the one-sided interference channels ($a = 0$ or $b = 0$). The region for which the capacity region is not known has shrunk to the inside of the thick solid lines only, i.e., the weak interference case and the mixed interference type I. All capacity regions established so far are achievable by the Han-Kobayashi scheme. The optimal splits between the private and common messages are however different, as shown in Fig. 8(b) in which $\beta^*$ is given by (25) and $\alpha^*$ is obtained by symmetry (swapping $P_1$ with $P_2$, and $a$ with $b$). The above figures are for $P_1 = 4$ and $P_2 = 2$.

subtract this part from the received signal and decode its interfering signal at a rate of $\gamma \left( \frac{bP_1}{1 + \alpha P_2} \right)$. For $b \geq \frac{1 + P_2}{1 + \alpha P_2}$, this is greater than $R_1 = \gamma \left( \frac{P_1}{1 + \beta P_2} \right)$, the maximum rate at which receiver 1 can decode its own message. That is, for this range of $b$, if receiver 1 can decode $X_1$ with arbitrarily small probability of error, receiver 2 will perform at least as well in decoding $X_1$. This receiver can thus subtract $X_1$ from $Y_2$ and obtain the cleaner output.
This now illuminates why for $b \geq \frac{1 + P_1}{1 + a P_2}$ the capacity of the interference channel is the same as the capacity of the one-sided interference channel with $b = 0$.

Remark 3. The above argument is very similar to Costa’s proof [17] in establishing the equality between the capacity region of the degraded and one-sided interference channels with $0 < a < 1$. It is however more general than that result, as instead of $b = \frac{1}{a}$ we have $b \geq \frac{1 + P_1}{1 + a P_2}$, which includes the condition $ab = 1$ as a special case, as graphically shown in Fig. 8(a).

Remark 4. It is obvious that case 2 includes case 1, because $b \geq \frac{1 + P_1}{1 + a P_2}$ implies $b \geq 1 + P_2$ for any $a \leq 1$. We differentiate them to highlight that in these two cases the order of decoding can be different, even though the ultimate rate regions are the same. Recall that in case 1, receiver 2 first decodes interference, then its respective common message, and finally its private message. In case 2 however, receiver 2 first decodes its respective common message, then interference, and, in the end, its private message. This indicates that for $b \geq 1 + P_2$ both sequences of decoding are optimal.

The above discussions on optimal decoding imply that the achievability of all capacity regions so far established for the interference channel are based on the HK scheme. The optimal splits between the private and common messages are however different, as shown in Fig. 8(b). Obviously, for the region with unknown capacity, we do not know whether or not the HK scheme is optimal. What is more, we even do not know the choices of $\alpha$ and $\beta$ to optimize the HK scheme. Optimizing over all possible HK strategies to get the largest achievable region within (65) turns out to be a complicated task but it appears to be crucial in fully answering the long-standing questions of the interference channel.

We conclude this paper by remarking the following general strategies about the optimal HK scheme:

- a transmitter causing no interference is not required to have a common message;
- a transmitter causing strong interference is not required to have a private message; and,
- a transmitter causing weak interference needs to have both private and common messages.

The above power-splitting strategies are proven to be optimal for the capacity results established thus far. This can be seen from Fig. 8(b) where the capacity results for $a \geq 1$ have $\beta = 0$ and those for $b \geq 1$ have $\alpha = 0$, whereas those for $a = 0$ and $b = 0$ have $\beta = 1$ and $\alpha = 1$, respectively. Furthermore, $0 < \beta < 1$ for $0 < a < 1$ and $0 < \alpha < 1$ for $0 < b < 1$.

APPENDIX A

Proof of Lemma 1

For $m = 2$ and $n = 1$ this lemma has been proved in [19] Lemma 13. Hence, Lemma 1 is a generalization of [19] Lemma 13 for arbitrary $m$ and $n$. We divide the proof into two parts. We first provide the prove for $m = 2$ and arbitrary $n$, and we next extend this proof to arbitrary $m$. For $m = 2$, by applying the chain rule of mutual information we get

$$I(X; X + Z) = I(X_1; X_1 + Z_1) + I(X_2; X_2 + Z_2 | X_1 + Z_1) + I(X_2; X + Z | X_1).$$

(67)
We prove that
\[
\lim_{\sigma_{\text{min}}^2 \to \infty} I(X_1; X_2 + Z_2| X_1 + Z_1) = 0, \quad (68a)
\]
\[
\lim_{\sigma_{\text{min}}^2 \to \infty} I(X_2; X + Z| X_1) = 0. \quad (68b)
\]
To this end, first note that the following \textit{Markov chain} relations hold:
\[
X_1 + Z_1 \to X_1 \to X_2 + Z_2, \quad (69a)
\]
\[
X_1 \to X_2 \to X_2 + Z_2. \quad (69b)
\]
Hence, we can write
\[
I(X_1; X_2 + Z_2| X_1 + Z_1) \leq I(X_1; X_2 + Z_2) \leq I(X_2; X_2 + Z_2), \quad (70)
\]
where the first and second inequalities are due to (69a) and (69b), respectively. Likewise, we can write
\[
I(X_2; X + Z| X_1) = I(X_2; X_2 + Z_2| X_1) + I(X_2; Z_1| X_1, X_2 + Z_2)
\]
\[\overset{(a)}{=} I(X_2; X_2 + Z_2| X_1) \]
\[\overset{(b)}{\leq} I(X_2; X_2 + Z_2), \quad (71)
\]
in which (a) is due to \(I(X_2; Z_1| X_1, X_2 + Z_2) = 0\), which in turn is correct because \(Z_1\) is independent of \(X\) and \(Z_2\), and (b) is due to (69b). Next, we see that
\[
\lim_{\sigma_{\text{min}}^2 \to \infty} I(X_2; X_2 + Z_2) \leq \lim_{\sigma_{\text{min}}^2 \to \infty} \frac{1}{2} \log |I + K_{Z_2}^{-1} K_{X_2}|
\]
\[= 0, \quad (72)
\]
where the inequality is due to the fact that under a covariance constraint, Gaussian input maximizes the entropy. Also, note that \(\sigma_{\text{min}}^2 \to \infty\) implies \(\sigma_{2,\text{min}}^2 \to \infty\). Then, (68a) and (68b) follow from (70), (71), and (72). This completes the proof for \(m = 2\).

The proof for \(m > 2\) is very similar to that for \(m = 2\). In this case, we first rewrite \(X\) and \(Z\) as
\[
X = [X_1^t, \bar{X}_1^t]^t, \quad \bar{X}_1^t = [X_2^t, \ldots, X_m^t], \quad (73a)
\]
\[
Z = [Z_1^t, \bar{Z}_1^t]^t, \quad \bar{Z}_1^t = [Z_2^t, \ldots, Z_m^t]. \quad (73b)
\]
Then, we apply the proof of the previous case \((m = 2)\), where we replace \(X_2\) by \(\bar{X}_1\) and \(Z_2\) by \(\bar{Z}_1\). Note that all equations and relations are still valid, and the claim readily follows.
Suppose $Z_0, Z_1, \ldots, Z_K$ are Gaussian vectors as defined in Lemma 2. Then, we know from Lemma 2 that for $\mu_0 = -1$

$$\sum_{k=0}^{K} \mu_k h(X + Z_k) \leq \sum_{k=0}^{K} \mu_k \log |K_X^{-1} + K_{Z_k}|,
$$

(74)

in which $K_X$ is the covariance matrix corresponding to the optimal solution described in Lemma 2, which is a Gaussian $m$-vector. Since scaling the objective function by a positive scalar does not change the optimum solution, it is clear that (74) holds for any $\mu_0 < 0$ as long as $\sum_{k=0}^{K} \mu_k = 0$ and $\mu_k \geq 0$ for $k > 0$.

We next add the constant term $-\sum_{k=0}^{K} \mu_k h(Z_k)$ to both sides of (74) to get

$$\sum_{k=0}^{K} \mu_k I(X; X + Z_k) \leq \sum_{k=0}^{K} \mu_k \log |K_{Z_k}^{-1}K_X^{-1} + I|.
$$

(75)

Since for any nonsingular $m \times m$ matrix $A$ we have $h(AX) = h(X) + \log |\det(A)|$, we can write

$$\sum_{k=0}^{K} \mu_k I(X; X + Z_k) = \sum_{k=0}^{K} \mu_k I(V_kX; V_kX + V_kZ_k) + c
$$

(76)

where $\{V_k\}_{k=0}^{K}$ are nonsingular $m \times m$ matrices, and the constant $c$ is given by

$$c = -\sum_{k=0}^{K} \mu_k \log |\det(V_k)|.
$$

(77)

Again, since adding a constant will not change the optimal input distribution, we note that $\sum_{k=0}^{K} \mu_k I(V_kX; V_kX + V_kZ_k)$ and $\sum_{k=0}^{K} \mu_k I(X; X + Z_k)$ will have the same optimal input distributions.

Now, for $k = 0, \ldots, K$, let

$$V_k = \left[ \begin{array}{c} v_{t_k}^t \\ E_k^t \end{array} \right],
$$

(78)

where $E_k$ is $m \times (m - 1)$ matrix such that $\frac{V_k}{\|v_k\|}$ is orthonormal. To put it simply, $V_k$ can be completed by applying Gram-Schmidt orthonormalization to any $m \times m$ nonsingular matrix whose first row is fixed to be $v_{t_k}^t$. Let us now choose the $K_{Z_k}$s as follows:

$$K_{Z_k} = U_k \Sigma_k U_k^t
$$

(79)

in which $U_k = \frac{V_k^t}{\|v_k\|}$ is orthonormal and

$$\Sigma_k = \left[ \begin{array}{cccc} \sigma^2_{k1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{k2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{km} \end{array} \right],
$$

(80)
where \( \sigma^2_{k_1} \) equals the variance of \( Z_k \) divided by \( \|v_k\|^2 \) and \( \sigma^2_{k_2}, \ldots, \sigma^2_{km} \) are arbitrary positive numbers\(^8\). To simplify the notation, let us define

\[
\begin{align*}
\tilde{X}_k &= V_k X = [\tilde{X}_{k1}, \ldots, \tilde{X}_{km}]^t, \\
\tilde{Z}_k &= V_k Z_k = [\tilde{Z}_{k1}, \ldots, \tilde{Z}_{km}]^t.
\end{align*}
\]

It is easy to show that \( \tilde{Z}_{k1}, \ldots, \tilde{Z}_{km} \) are mutually independent. To this end, note that \( \tilde{Z}_k \) is a Gaussian vector because its elements are linear combinations of a Gaussian vector. Hence, to prove its elements are independent, it suffices to show that they are uncorrelated. The latter is rather simple because

\[
E\{\tilde{Z}_k \tilde{Z}_k^t\} = V_k K Z_k V_k^t = \|v_k\|^2 \Sigma_k,
\]

where the last step follows from (79) and the fact that \( U_k = \frac{V_k}{\|v_k\|} \) is orthonormal. Moreover, it is seen that the covariance matrix of \( \tilde{Z}_k \) is given by

\[
\tilde{\Sigma}_k = \|v_k\|^2 \Sigma_k.
\]

Next, let us define

\[
\sigma^2_{\min} \triangleq \min\{\sigma^2_{k_j}, k = 0, \ldots, K, j = 2, \ldots, m\}.
\]

Then, using Lemma [1] we can see that

\[
\begin{align*}
limit_{\sigma^2_{\min} \to \infty} I(X_k; X_k + Z_k) &= I(\tilde{X}_{k1}; \tilde{X}_{k1} + \tilde{Z}_{k1}).
\end{align*}
\]

It should be noted that, from (83), we know that the elements of \( \tilde{Z}_k \) are mutually independent.

Recalling that \( X_k = V_k X, \tilde{X}_{k1} = v_k^t X, \text{ and } \tilde{Z}_{k1} = v_k^t Z_k \), from (85) we can see that for any non-zero \( v_k \) in \( \mathbb{R}^m \), by choosing \( V_k \) and \( K Z_k \) as in (78) and (79) and letting \( \sigma^2_{\min} \to \infty \), \( I(V_k X; V_k X + V_k Z_k) \) approaches \( I(v_k^t X; v_k^t X + v_k^t Z_k) \). Consequently, when \( \sigma^2_{\min} \to \infty \), the first term on the right-hand side of (76) simplifies as

\[
\begin{align*}
\lim_{\sigma^2_{\min} \to \infty} K \mu_k I(V_k X; V_k X + V_k Z_k) &= \sum_{k=0}^{K} K \mu_k I(v_k^t X; v_k^t X + v_k^t Z_k).
\end{align*}
\]

Now let us add the constant term \( \sum_{k=0}^{K} \mu_k h(v_k^t Z_k) \) to (86) and denote

\[
\hat{Z}_k = v_k^t Z_k, \quad 0 \leq k \leq K.
\]
It should be noted that the variance of $\hat{Z}_k$ is equal to the variance of $Z_k$; thus these random variables are statistically identical and so the maximization of (86) over $X$ is equivalent to the maximization of (15) over $X$.

We now complete the proof of Lemma 3. From Lemma 2 we know that $\sum_{k=0}^K \mu_k h(X + Z_k)$ is maximized by a Gaussian $X$. We show that for Gaussian noises $Z_k$, which have strictly positive variances, $\sum_{k=0}^K \mu_k h(v_t^k X + Z_k)$ is also maximized by a Gaussian input. We basically show that

$$N(0, K \hat{X}) = \arg\max_{p(x)} \sum_{k=0}^K \mu_k h(X + Z_k)$$

$$= \arg\max_{p(x)} \sum_{k=0}^K \mu_k I(X; X + Z_k)$$

$$= \arg\max_{p(x)} \sum_{k=0}^K \mu_k I(V_k X; V_k X + V_k Z_k)$$

$$= \arg\max_{p(x)} \sum_{k=0}^K \mu_k I(v_t^k X; v_t^k X + v_t^k Z_k)$$

$$= \arg\max_{p(x)} \sum_{k=0}^K \mu_k h(v_t^k X + Z_k)$$

in which

(a) follows by the fact that adding a constant term does not change the optimal solution (input distribution), where we add $\sum_{k=0}^K \mu_k h(Z_k)$,

(b) is due to the identity $h(A X) = h(X) + \log |\det(A)|$ and the fact that a constant does not change the optimal input distribution,

(c) is proved in (86) by defining $V_k = [v_k \mid E_k]^t$ for a given $v_k$ in $\mathbb{R}^m$, constructing $E_k$ in a way that $V_k / ||v_k||$ is orthonormal, and applying Lemma 1 and

(d) follows by adding the constant $\sum_{k=0}^K \mu_k h(v_t^k Z_k)$ and then defining $\hat{Z}_k = v_t^k Z_k$, for $0 \leq k \leq K$, and noting the statistical equivalence of $\hat{Z}_k$ and $Z_k$.

**APPENDIX C**

**PROOF OF LEMMA 4**

We note that Lemma 4 is a generalization of Lemma 3 where $X_1, \ldots, X_m$ and $Z_0, \ldots, Z_K$ are $n$-dimensional vectors rather than being one dimensional. As such, we can follow a very similar line of proof. Let $X$ be the vector resulting from concatenating $X_1$ through $X_m$, i.e.,

$$X = [X_1^t, \ldots, X_m^t]^t.$$  

Obviously, the length of $X$ is $mn$. At this point, instead of having $m$ constraints $K X_j \preceq S_j$, we consider a single constraint $K X \preceq S$, where $S$ is a positive semidefinite matrix. Note that this constraint subsumes all those $m$ constraints because any principal submatrix\(^9\) of a positive semidefinite matrix is positive semidefinite \([25]\). This

---

\(^9\)An $m \times m$ principal submatrix of an $n \times n$ matrix $A$ is obtained by removing any $n-m$ rows and the “same” $n-m$ columns from $A$.  

---
A single constraint is considered to enable us to use the proof of Lemma 3. Also, let

\[ Z'_k = [Z'_{k1}, \ldots, Z'_{km}]^t, \quad 0 \leq k \leq K \]  

in which \( Z_{kj} \), for any \( 0 \leq k \leq K \) and \( 1 \leq j \leq m \), is a Gaussian \( n \)-vector such that \( Z'_k \)'s, \( k = 0, \ldots, K \), are \( K + 1 \) Gaussian random vectors of length \( mn \) with positive definite covariance matrices. Then, similar to (88) we can write

\[ N'(0, K'X) = \arg\max_{p(x)} K \sum_{k=0}^K \mu_k h(X + Z'_k) \]

\[ \overset{(a)}{=} \arg\max_{p(x)} K \sum_{k=0}^K \mu_k I(X; X + Z'_k) \]

\[ \overset{(b)}{=} \arg\max_{p(x)} K \sum_{k=0}^K \mu_k I(V'_kX; V'_kX + V'_kZ'_k) \]

\[ \overset{(c)}{=} \arg\max_{p(x)} K \sum_{k=0}^K \mu_k I\left( \sum_{j=1}^m v_{kj}X_j; \sum_{j=1}^m v_{kj}(X_j + Z_{kj}) \right) \]

\[ \overset{(d)}{=} \arg\max_{p(x)} K \sum_{k=0}^K \mu_k h\left( \sum_{j=1}^m v_{kj}X_j + Z_k \right) \]  

in which (a) to (d) follow the same arguments as those in (88), with the following differences. Here, \( V'_k \) is an \( mn \times mn \) matrix, defined as a block matrix whose elements are obtained by replacing \( (V_k)_{ij} \) by \( (V_k)_{ij}I \), where \( V_k \) is a square matrix of size \( m \) defined in (78) and \( I \) is the identity matrix of size \( n \). In other words, for \( k = 0, \ldots, K \), we define

\[ V'_k = V_k \otimes I, \]  

where \( \otimes \) denotes the Kronecker product \[26]. Knowing that \( \det(A \otimes B) = \det(A)^{\text{rank}(B)} \det(B)^{\text{rank}(A)} \) \[26], it is easy to see that \( V'_k \) is nonsingular when \( V_k \) is so. We next choose

\[ KZ'_k = U'_k \Sigma'_k U'^{t}_k \]  

in which \( U'_k = \frac{V'_k}{\|v_k\|} \) is orthonormal and

\[ \Sigma'_k = \begin{bmatrix} \Sigma_{k1} & 0 & \cdots & 0 \\ 0 & \Sigma_{k2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{km} \end{bmatrix}, \]  

where \( \Sigma_{k1} \) equals the covariance matrix of \( Z_k \) divided by \( \|v_k\|^2 \) and \( \Sigma_{k2}, \ldots, \Sigma_{km} \) are the covariance matrices corresponding to independent \( Z_{k2}, \ldots, Z_{km} \) in (90), and

\[ \Sigma_{kj} = \sigma^2 I, \quad \forall \ k = 0, \ldots, K, \ j = 2, \ldots, m. \]
Hence, similar to (86), (c) is obtained from Lemma 1, because
\[
\lim_{\sigma^2 \to \infty} \sum_{k=0}^{K} \mu_k I(V_k' X; V_k' X + V_k' Z_k') = \sum_{k=0}^{K} \mu_k I \left( \sum_{j=1}^{m} v_{kj} X_j; \sum_{j=1}^{m} v_{kj} (X_j + Z_k) \right). \tag{96}
\]
Finally, in (d) we add the constant \(\sum_{k=0}^{K} \mu_k h(\sum_{j=1}^{m} v_{kj} Z_k')\) and then define
\[
\hat{Z}_k = \sum_{j=1}^{m} v_{kj} Z_k', \quad 0 \leq k \leq K, \tag{97}
\]
and note that \(\hat{Z}_k\) and \(Z_k\) are statistically equivalent. This is because similar to (82), in general, we have
\[
E\{V_k' Z_k' Z_k' V_k'\} = V_k' K Z_k' V_k' = \|v_k\|^2 \Sigma_k',
\]
thus
\[
K \hat{Z}_k = \|v_k\|^2 \Sigma_k = K Z_k, \quad k = 0, \ldots, K. \tag{98}
\]
This completes the proof.

VIII. ACKNOWLEDGEMENT

Throughout the development of this paper we have had fruitful discussions with several scholars. Specifically, the authors would like to thank Sina Hamidi Ghalehjegh, Ronit Bustin, and Xiaohu Shang. The first author is also indebted to Fabrice Labeau for his constant support during and after his transition to Princeton.

REFERENCES

[1] A. Carleial, “Interference channels,” IEEE Transactions on Information Theory, vol. 24, pp. 60–70, January 1978.
[2] R. Ahlswede, “Multi-way communication channels,” in Proc. Second International Symposium on Information Theory, pp. 23–52, Sept. 1971.
[3] A. Carleial, “A case where interference does not reduce capacity,” IEEE Transactions on Information Theory, vol. 21, no. 5, pp. 569–570, 1975.
[4] 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.
[5] 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.
[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] M. Vaezi and H. V. Poor, “On limiting expressions for the capacity region of Gaussian interference channels,” to appear in Proc. 49th Asilomar Conference on Signals, Systems and Computers, Nov. 2015.
[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] 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.
[10] 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.
[11] 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.
[12] 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.

[13] A. Carleial, “Outer bounds on the capacity of interference channels,” IEEE Transactions on Information Theory, vol. 29, no. 4, pp. 602–606, 1983.

[14] G. Kramer, “Outer bounds on the capacity of Gaussian interference channels,” IEEE Transactions on Information Theory, vol. 50, no. 3, pp. 581–586, 2004.

[15] G. Kramer, “Review of rate regions for interference channels,” in Proc. International Zurich Seminar on Communications, pp. 162–165, February 2006.

[16] H. Sato, “On degraded Gaussian two-user channels,” IEEE Transactions on Information Theory, vol. 24, no. 5, pp. 637–640, 1978.

[17] M. H. M. Costa, “On the Gaussian interference channel,” IEEE Transactions on Information Theory, vol. 31, pp. 607–615, September 1985.

[18] I. Sason, “On achievable rate regions for the Gaussian interference channel,” IEEE Transactions on Information Theory, vol. 50, no. 6, pp. 1345–1356, 2004.

[19] 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.

[20] 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.

[21] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: John Wiley & Sons, 2006.

[22] 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.

[23] R. Bustin, H. V. Poor, and S. Shamai, “The effect of maximal rate codes on the interfering message rate,” arXiv:1404.5699v4, 2015.

[24] Y. Polyanskiy and Y. Wu, “Wasserstein continuity of entropy and outer bounds for interference channels,” arXiv:1504.04419v2, 2015.

[25] A. Berman and N. Shaked-Monderer, Completely Positive Matrices. World Scientific, 2003.

[26] K. B. Petersen and M. S. Pedersen, The Matrix Cookbook. 2012.