A Proof of the Noiseberg Conjecture for the Gaussian Z-Interference Channel

Max Costa, Amin Gohari, Chandra Nair, and David Ng

Abstract

We establish the noiseberg conjecture regarding the Han–Kobayashi region of the Gaussian Z-Interference channel with Gaussian signaling. We also provide a refined conjecture for the optimality of the HK inner bound with Gaussian signaling.

1 Introduction

The *Gaussian Z-interference channel* (GZIC), with non-zero parameter $a \in \mathbb{R}$, is a two-user interference channel defined by

$$\begin{aligned} Y_1 &= X_1 + Z_1, \\ Y_2 &= X_2 + aX_1 + Z_2, \end{aligned}$$

where X_i, Y_i, Z_i $(i = 1, 2)$ are real random variables, $Z_1, Z_2 \sim \mathcal{N}(0, 1)$, and X_1, X_2, Z_1, Z_2 are mutually independent. In this paper we assume $0 < |a| < 1$, as the capacity for *strong interference* $(|a| \geq 1)$ is known [1, 2]. The key question of interest is a computable characterization of the set of achievable rate pairs (R_1, R_2), or the capacity region, in terms of the power constraints Q_1, Q_2 imposed on X_1, X_2 respectively.

As shown in [3], from a capacity region perspective, the GZIC can be equivalently formulated as a degraded interference channel given by

$$Y_1 = X_1 + Z_1, \tag{1a}$$
$$Y_2 = X_2 + X_1 + Z_1 + Z_2 = X_2 + Y_1 + Z_2, \tag{1b}$$

where X_i, Y_i, Z_i $(i = 1, 2)$ are real random variables, $Z_1 \sim \mathcal{N}(0, N_1)$, $Z_2 \sim \mathcal{N}(0, N_2)$ (with $N_1 := 1$, $N_2 := \frac{1}{a^2} - 1$), and X_1, X_2, Z_1, Z_2 are mutually independent. Further the power constraints on X_1, X_2 become transformed, in this equivalent model, to P_1, P_2 where $P_1 = Q_1$ and $P_2 = \frac{Q_2}{a^2}$ respectively. For consistency, we will use the latter model in this paper.

The best single-letter achievable region known for a general two-user interference channel is given by Han and Kobayashi (HK) [2]. Though there are instances of discrete interference channels [4] for which the HK achievable region is strictly suboptimal, for the GZIC it is still unknown whether the HK achievable region (restricted to Gaussian inputs and allowing for power control) represents the capacity region or not.

Unfortunately, the existing techniques to prove the optimality of jointly Gaussian input distributions are insufficient to address this question for the HK region (or its multi-letter extensions). In particular, the monotone path argument of [5] and the subadditivity/doubling argument of [6] for showing optimality of Gaussian input distributions are only applicable when time-sharing (*power control*) does not improve a region. For the HK region, it is shown in [7, 8] that HK region with time-sharing between Gaussian distributions strictly improves over the achievable region without time-sharing. This presents a fundamental obstacle in showing optimality of Gaussian distributions (with time-sharing). To tackle this issue, the authors of [9] considered a different functional, namely the Fenchel dual for the weighted sum-rate expression of the HK achievable region, and showed that establishing Gaussian optimality of input distributions for the new functional (if established) would imply Gaussian optimality (with time-sharing) for the original HK region. Gaussian optimality for the new functional is stated as a conjecture in [9]. This conjecture is sufficient to prove optimality of HK region (with Gaussian signaling and power control). The conjecture had been shown to hold for some parameter regimes [10].

However, some counterexamples to the conjecture were recently found in [11]. Nonetheless, the counterexamples are in a certain parameter regime and do not end up disproving the optimality of the HK achievable region with Gaussian inputs (see Theorem 5 and Theorem 6 in [11]). Rather, they suggest the need for a proper refining of the conjecture. Such a refinement is discussed in Section 4. The refined conjecture is based on the exposed sets of the *noiseberg* region in [7].

In [7], it was conjectured that a particular scheme obtained using Gaussian signaling and power control would achieve the capacity region of GZIC. Leaving the optimality of Gaussian signaling aside, built into [7] was the claim that a particular time-sharing between two schemes (in one of the two schemes, message was only transmitted to Receiver 1) achieved the upper concave envelope of Han–Kobayashi bound with Gaussian signaling and it is this claim that we term as the *noiseberg conjecture*. The calculations of the slope of the region at the corner points in [12], [13] shows the non-triviality of computing the Han–Kobayashi region with Gaussian signaling. In this paper, we provide a proof of the noiseberg conjecture. Following up on the intuition in [7], the authors of [14] claimed to have a proof for the correctness of the noiseberg conjecture. However, as we show in Section 3, the proof has a gap.

This paper is organized as follows: some background information is given in Section 1.1. The noiseberg conjecture is formally proven in Section 2. Following that, a refined Gaussian optimality conjecture is discussed in Section 4.

1.1 Background

For a function f we use $\mathfrak{C}f$ to denote the upper concave envelope of f. We use $\mathbb{R}^n_{>0}$ and $\mathbb{R}^n_{\geq 0}$ to denote the positive and non-negative vectors in $\mathbb{R}^n$ respectively.

Definition 1. *The* Han–Kobayashi achievable region *for GZIC, denoted by $\mathcal{R}_{HK}(P_1, P_2, N_1, N_2)$, is the set of all rate pairs $(R_1, R_2) \in \mathbb{R}^2_{\geq 0}$ satisfying*

$$R_1 \leq I(X_1; Y_1|Q), \tag{2a}$$

$$R_2 \leq I(X_2; Y_2|U_1, Q), \tag{2b}$$

$$R_1 + R_2 \leq I(U_1, X_2; Y_2|Q) + I(X_1; Y_1|U_1, Q), \tag{2c}$$

and the power constraints

$$\mathrm{E}[X_1^2] \leq P_1, \tag{3a}$$

$$\mathrm{E}[X_2^2] \leq P_2, \tag{3b}$$

for some $p(q)p(u_1, x_1|q)p(x_2|q)$, where $Y_1 = X_1 + Z_1$, $Y_2 = X_2 + X_1 + Z_1 + Z_2$, and X_1, X_2, Z_1, Z_2 are mutually independent random variables in $\mathbb{R}$ with $Z_1 \sim \mathcal{N}(0, N_1)$ and $Z_2 \sim \mathcal{N}(0, N_2)$.

Definition 2. *The* Han–Kobayashi achievable region with Gaussian signaling and power control *for GZIC, denoted by $\mathcal{R}_{HK\text{-}GS}(P_1, P_2, N_1, N_2)$, is the set of all rate pairs $(R_1, R_2) \in \mathbb{R}^2_{\geq 0}$ such that (2) and (3) hold with $X_1 := U_1 + V_1$ for some $p(q)p(u_1|q)p(v_1|q)p(x_2|q)$, where the conditional distributions $p(u_1|q), p(v_1|q), p(x_2|q)$ are zero-mean scalar Gaussian distributions for each q.*

In order to state the noiseberg conjecture, we need the following proposition:

Proposition 1. *For $\lambda \geq 1$, let the function $f_{\lambda,GS} : \mathbb{R}^2_{\geq 0} \to \mathbb{R}$ be defined by*

$$\begin{aligned} f_{\lambda,GS}(P_1, P_2) := \frac{1}{2}\Big(&-\log(N_1) + \log(P_1 + P_2 + N_1 + N_2) \\ &+ \max_{0 \leq P_{1P} \leq P_1} \big((\lambda - 1)\log(P_{1P} + P_2 + N_1 + N_2) + \log(P_{1P} + N_1) - \lambda \log(P_{1P} + N_1 + N_2)\big)\Big). \end{aligned} \tag{4}$$

Then

$$\sup_{(R_1,R_2) \in \mathcal{R}_{HK\text{-}GS}(P_1,P_2,N_1,N_2)} (R_1 + \lambda R_2) = \mathfrak{C}f_{\lambda,GS}(P_1, P_2),$$

where $\mathfrak{C}f$ denotes the upper concave envelope of f.

Proof. Since Y_2 is a degraded version of Y_1, by the data-processing inequality, we have $I(U_1; Y_1|Q) \geq I(U_1; Y_2|Q)$ for any $p(q)p(u_1, x_1|q)p(x_2|q)$. Therefore, for any $\lambda \geq 1$, we obtain

$$\sup_{(R_1,R_2) \in \mathcal{R}_{\mathrm{HK}}(P_1,P_2,N_1,N_2)} (R_1 + \lambda R_2)$$

$$= \sup_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} \big(I(U_1, X_2; Y_2|Q) + I(X_1; Y_1|U_1, Q) + (\lambda - 1) I(X_2; Y_2|U_1, Q) \big).$$

With the parameterization $U_1|_{Q=q} \sim \mathcal{N}(0, P_{1C}^q)$, $V_1|_{Q=q} \sim \mathcal{N}(0, P_{1P}^q)$, $X_2|_{Q=q} \sim \mathcal{N}(0, P_2^q)$, we have

$$\begin{aligned}
&\sup_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} (I(U_1, X_2; Y_2|Q) + I(X_1; Y_1|U_1, Q) + (\lambda - 1) I(X_2; Y_2|U_1, Q)) \\
&= \sup_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} (-h(Z_1) + h(X_1 + X_2 + Z_1 + Z_2|Q) \\
&\quad + (\lambda - 1) h(V_1 + X_2 + Z_1 + Z_2|Q) + h(V_1 + Z_1|Q) - \lambda h(V_1 + Z_1 + Z_2|Q)) \\
&= \sup_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} \frac{1}{2} \mathrm{E}_Q \Big[- \log(N_1) + \log\Big(P_{1C}^Q + P_{1P}^Q + P_2^Q + N_1 + N_2\Big) \\
&\quad + (\lambda - 1) \log\Big(P_{1P}^Q + P_2^Q + N_1 + N_2\Big) + \log\Big(P_{1P}^Q + N_1\Big) - \lambda \log\Big(P_{1P}^Q + N_1 + N_2\Big)\Big] \\
&= \sup_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} \mathrm{E}_Q \Big[f_{\lambda,\mathrm{GS}}\Big(P_{1C}^Q + P_{1P}^Q, P_2^Q\Big)\Big] \\
&= \mathfrak{C} f_{\lambda,\mathrm{GS}}(P_1, P_2),
\end{aligned}$$

where the suprema are subjected to the constraints $P_{1C}^q, P_{1P}^q, P_2^q \geq 0$, $\mathrm{E}_Q\Big[P_{1C}^Q + P_{1P}^Q\Big] \leq P_1$ and $\mathrm{E}_Q\Big[P_2^Q\Big] \leq P_2$, and in the last equality we have used the fact that $f_{\lambda,\mathrm{GS}}(P_1, P_2)$ is non-decreasing in both P_1 and P_2. $\square$

Recall that

$$\begin{aligned}
f_{\lambda,\mathrm{GS}}(P_1, P_2) &= \frac{1}{2}\Big(- \log(N_1) + \log(P_1 + P_2 + N_1 + N_2) \\
&+ \max_{0 \leq P_{1P} \leq P_1} \big((\lambda - 1) \log(P_{1P} + P_2 + N_1 + N_2) + \log(P_{1P} + N_1) - \lambda \log(P_{1P} + N_1 + N_2)\big)\Big).
\end{aligned}$$

The objective of the maximization over P_{1P} is strictly increasing (respectively, strictly decreasing) if and only if

$$\frac{P_2 + N_2}{P_2}\left(1 + \frac{N_2}{P_{1P} + N_1}\right) > \lambda \text{ (respectively, } < \lambda).$$

Hence the optimal P_{1P} is unique in $[0, P_1]$ and is given by

$$P_{1P}^* := \begin{cases} 0 & \text{if } (P_1, P_2) \in \mathcal{R}_1, \\ P_1 & \text{if } (P_1, P_2) \in \mathcal{R}_2, \\ \frac{N_2(P_2+N_2)}{P_2(\lambda-1)-N_2} - N_1 & \text{if } (P_1, P_2) \in \mathcal{R}_3, \end{cases}$$

where the regions $\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3$ are defined by (5)-(7) given in the next page. This gives an explicit expression for $f_{\lambda,\mathrm{GS}}(P_1, P_2)$ given in (8). Now we can compute the gradient and Hessian of $f_{\lambda,\mathrm{GS}}(P_1, P_2)$ as shown in the next page. By inspecting the values and the gradients of $f_{\lambda,\mathrm{GS}}$ at the boundaries, one can see that $f_{\lambda,\mathrm{GS}}$ is continuously differentiable on $\mathbb{R}_{>0}^2$.

$$\mathcal{R}_1 := \left\{ (P_1, P_2) \in \mathbb{R}^2_{\geq 0} \,:\, \lambda \geq \frac{P_2 + N_2}{P_2} \left(1 + \frac{N_2}{N_1} \right) \right\}, \tag{5}$$

$$\mathcal{R}_2 := \left\{ (P_1, P_2) \in \mathbb{R}^2_{\geq 0} \,:\, \lambda \leq \frac{P_2 + N_2}{P_2} \left(1 + \frac{N_2}{P_1 + N_1} \right) \right\}, \tag{6}$$

$$\mathcal{R}_3 := \left\{ (P_1, P_2) \in \mathbb{R}^2_{\geq 0} \,:\, \frac{P_2 + N_2}{P_2} \left(1 + \frac{N_2}{P_1 + N_1} \right) < \lambda < \frac{P_2 + N_2}{P_2} \left(1 + \frac{N_2}{N_1} \right) \right\}. \tag{7}$$

$$\begin{aligned} & f_{\lambda,\mathrm{GS}}(P_1, P_2) \\ &= \frac{1}{2} \begin{cases} \log(P_1 + P_2 + N_1 + N_2) + (\lambda - 1) \log(P_2 + N_1 + N_2) - \lambda \log(N_1 + N_2) & \text{if } (P_1, P_2) \in \mathcal{R}_1, \\ - \log(N_1) + \lambda \log(P_1 + P_2 + N_1 + N_2) + \log(P_1 + N_1) - \lambda \log(P_1 + N_1 + N_2) & \text{if } (P_1, P_2) \in \mathcal{R}_2, \\ - \log(N_1) + \log(P_1 + P_2 + N_1 + N_2) & \\ \qquad + \lambda \log(P_2 + N_2) - \log(P_2) - (\lambda - 1) \log(N_2) + (\lambda - 1) \log(\lambda - 1) - \lambda \log(\lambda) & \text{if } (P_1, P_2) \in \mathcal{R}_3. \end{cases} \end{aligned} \tag{8}$$

Region $\mathcal{R}_1$

$$\partial_{P_1} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{1}{P_1 + P_2 + N_1 + N_2} \right), \tag{9}$$

$$\partial_{P_2} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{1}{P_1 + P_2 + N_1 + N_2} + \frac{\lambda - 1}{P_2 + N_1 + N_2} \right), \tag{10}$$

$$\mathcal{H} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \begin{pmatrix} \frac{-1}{(P_1+P_2+N_1+N_2)^2} & \frac{-1}{(P_1+P_2+N_1+N_2)^2} \\ \frac{-1}{(P_1+P_2+N_1+N_2)^2} & \frac{-1}{(P_1+P_2+N_1+N_2)^2} - \frac{\lambda-1}{(P_2+N_1+N_2)^2} \end{pmatrix}. \tag{11}$$

Region $\mathcal{R}_2$

$$\partial_{P_1} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{\lambda}{P_1 + P_2 + N_1 + N_2} + \frac{1}{P_1 + N_1} - \frac{\lambda}{P_1 + N_1 + N_2} \right), \tag{12}$$

$$\partial_{P_2} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{\lambda}{P_1 + P_2 + N_1 + N_2} \right), \tag{13}$$

$$\mathcal{H} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \begin{pmatrix} \frac{-\lambda}{(P_1+P_2+N_1+N_2)^2} - \frac{1}{(P_1+N_1)^2} + \frac{\lambda}{(P_1+N_1+N_2)^2} & \frac{-\lambda}{(P_1+P_2+N_1+N_2)^2} \\ \frac{-\lambda}{(P_1+P_2+N_1+N_2)^2} & \frac{-\lambda}{(P_1+P_2+N_1+N_2)^2} \end{pmatrix}. \tag{14}$$

Region $\mathcal{R}_3$

$$\partial_{P_1} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{1}{P_1 + P_2 + N_1 + N_2} \right), \tag{15}$$

$$\partial_{P_2} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \left(\frac{1}{P_1 + P_2 + N_1 + N_2} + \frac{\lambda}{P_2 + N_2} - \frac{1}{P_2} \right), \tag{16}$$

$$\mathcal{H} f_{\lambda,\mathrm{GS}}(P_1, P_2) = \frac{1}{2} \begin{pmatrix} \frac{-1}{(P_1+P_2+N_1+N_2)^2} & \frac{-1}{(P_1+P_2+N_1+N_2)^2} \\ \frac{-1}{(P_1+P_2+N_1+N_2)^2} & \frac{-1}{(P_1+P_2+N_1+N_2)^2} - \frac{\lambda}{(P_2+N_2)^2} + \frac{1}{P_2^2} \end{pmatrix}. \tag{17}$$

The noiseberg conjecture [7] then states:

Conjecture 1 (The Noiseberg Conjecture)**.** *Let* $\lambda \geq 1$ *and* $f_{\lambda,GS}$ *be the function defined by* (4)*. Then for all* $P_1, P_2 \geq 0$*,*

$$\mathfrak{C}f_{\lambda,GS}(P_1,P_2) = \max_{\alpha,\tilde{P}}\left(\alpha f_{\lambda,GS}\left(\tilde{P},\frac{P_2}{\alpha}\right) + (1-\alpha)f_{\lambda,GS}\left(\frac{P_1-\alpha\tilde{P}}{1-\alpha},0\right)\right),$$

where the maximum is subjected to the constraints $\frac{P_2}{P_1+P_2} \leq \alpha \leq 1$ *and* $0 \leq \tilde{P} \leq P_1 + P_2 - \frac{P_2}{\alpha}$*.*

1.2 Preliminaries

We will state some lemmas that will turn out to be useful in our proofs. The proofs are given in Appendix A.

Lemma 1. *Take a continuous function* $f : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ *satisfying*

$$\lim_{\|\vec{x}\|\to\infty} \frac{f(\vec{x})}{\|\vec{x}\|} = 0. \tag{18}$$

Then, for any $\vec{x}$ *one can find* $n+1$ *points* $\vec{z}_1, \cdots, \vec{z}_{n+1}$ *and non-negative weights* ω_i $(1 \leq i \leq n+1)$ *adding up to one such that*

$$\mathfrak{C}f(\vec{x}) = \sum_{j=1}^{n+1} \omega_j f(\vec{z}_j)$$

and

$$\vec{x} - \sum_{j=1}^{n+1} \omega_j \vec{z}_j \in \mathbb{R}^n_{\geq 0}. \tag{19}$$

Moreover, if $f(x_1, \cdots, x_n)$ *is non-decreasing in* x_i *for* $1 \leq i \leq n$*, then* (19) *can be replaced with the stronger condition*

$$\vec{x} = \sum_{j=1}^{n+1} \omega_j \vec{z}_j.$$

Lemma 2. *Take a function* $f : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ *that is differentiable with a continuous derivative. Let* $\mathcal{S}$ *be the set of vectors* $\vec{x} \in \mathbb{R}^n_{>0}$ *for which one can find* $n+1$ *points* $\vec{z}_1, \cdots, \vec{z}_{n+1}$ *(with at least one of* $\vec{z}_i$*'s in the interior of the domain,* $\mathbb{R}^n_{>0}$*) and non-negative weights* ω_i $(1 \leq i \leq n+1)$ *adding up to one such that* $\vec{x} = \sum_{j=1}^{n+1} \omega_j \vec{z}_j$ *and*

$$\mathfrak{C}f(\vec{x}) = \sum_{j=1}^{n+1} \omega_j f(\vec{z}_j).$$

Let $\mathcal{P}_f$ *be the set of all* $x \in \mathbb{R}^n_{\geq 0}$ *where* $f(x) = \mathfrak{C}f(x)$*. Let* $\mathcal{P}'_f$ *be the intersection of* $\mathcal{P}_f$ *with* $\mathbb{R}^n_{>0}$ *(the interior of the domain). Then, for every* $\vec{x} \in \mathcal{S}$ *we have*

$$\inf_{\vec{y}\in\mathcal{P}'_f} [f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y}\rangle] = \mathfrak{C}f(\vec{x}).$$

2 Proof of Conjecture 1

From (4), it is clear that $f_{\lambda,\mathrm{GS}}(P_1,P_2)$ is non-decreasing in P_1 and P_2. An explicit formula for $f_{\lambda,\mathrm{GS}}(P_1,P_2)$ is derived in (8), which shows that this function is continuous in P_1 and P_2 and has logarithmic growth in P_1 and P_2. Then, using Lemma 1 from Section 1.2, we can deduce that every point in the upper concave envelope of $f_{\lambda,\mathrm{GS}}$ can be written as the convex combination of at most three points in the domain.

The conjecture claims that a convex combination of at most two points is needed to yield the upper concave envelope. Moreover, these two points have to be of the form

$$\left(\tilde{P},\frac{P_2}{\alpha}\right), \qquad \left(\frac{P_1-\alpha\tilde{P}}{1-\alpha},0\right)$$

meaning that at most one of the points has $P_2 > 0$. Additionally, the conjecture also imposes some restrictions on the ranges of $\tilde{P}$ and the weight α. This is what forms the crux of our proof below.

Let us take three points (P_1^q, P_2^q) for $q = 1, 2, 3$ where q in P_1^q is just an index (not P_1 raised to the power q). Corresponding to (P_1^q, P_2^q) in the definition of $f_{\lambda,\mathrm{GS}}$, there is some optimizer P_{1P}^q given by (4). Let $P_{1C}^q = P_1^q - P_{1P}^q \geq 0$. Assuming the weight distribution $p_Q(q)$ for $q = 1, 2, 3$, the upper concave envelope can be compactly expressed as

$$\begin{aligned}
&\mathfrak{C}f_{\lambda,\mathrm{GS}}(P_1, P_2)\\
&= \max_{p(q),\, P_{1C}^q, P_{1P}^q, P_2^q} \frac{1}{2}\,\mathrm{E}_Q\Big[-\log(N_1) + \log\Big(P_{1C}^Q + P_{1P}^Q + P_2^Q + N_1 + N_2\Big)\\
&\qquad + (\lambda-1)\log\Big(P_{1P}^Q + P_2^Q + N_1 + N_2\Big) + \log\Big(P_{1P}^Q + N_1\Big) - \lambda\log\Big(P_{1P}^Q + N_1 + N_2\Big)\Big] \qquad (20)
\end{aligned}$$

where the maximum is subject to the constraints $q \in \{1, 2, 3\}$, $P_{1C}^q, P_{1P}^q, P_2^q \geq 0$, $\mathrm{E}_Q\Big[P_{1P}^Q\Big] \leq P_1$, $\mathrm{E}_Q\Big[P_{1C}^Q\Big] \leq P_1 - \mathrm{E}_Q\Big[P_{1P}^Q\Big]$ and $\mathrm{E}_Q\Big[P_2^Q\Big] \leq P_2$. Note that we can also write (20) as

$$\begin{aligned}
&\frac{1}{2}\max_{p(q),\, P_{1P}^q}\Big(-\log(N_1) + \mathrm{E}_Q\Big[\log\Big(P_{1P}^Q + N_1\Big) - \lambda\log\Big(P_{1P}^Q + N_1 + N_2\Big)\Big]\\
&\quad + \sup_{P_{1C}^q, P_2^q}\mathrm{E}_Q\Big[\log\Big(P_{1C}^Q + P_{1P}^Q + P_2^Q + N_1 + N_2\Big) + (\lambda-1)\log\Big(P_{1P}^Q + P_2^Q + N_1 + N_2\Big)\Big]\Big). \qquad (21)
\end{aligned}$$

By considering the inner maximization in (21), Proposition 4 given in Section 2.1 (with the choice of $n = 2$, $x_{1q} = P_{1C}^q + P_2^q$, $x_{2q} = P_2^q$, $c_1 = 1$, $c_2 = \lambda - 1$, $a_q = P_{1P}^q + N_1 + N_2$) yields that any maximizer must satisfy the following equations:

$$P_{1C}^q = \max\{\mu - (P_{1P}^q + N_1 + N_2), 0\} - \max\{\nu - (P_{1P}^q + N_1 + N_2), 0\}, \qquad (22)$$
$$P_2^q = \max\{\nu - (P_{1P}^q + N_1 + N_2), 0\}, \qquad (23)$$

for some $\mu \geq \nu \geq 0$ such that $\mathrm{E}_Q\Big[P_{1C}^Q\Big] = P_1 - \mathrm{E}_Q\Big[P_{1P}^Q\Big]$ and $\mathrm{E}_Q\Big[P_2^Q\Big] = P_2$. This is essentially a layered-water-filling argument.

Next, we show that we can restrict to the case that there is a unique index q such that $P_2^q > 0$. To show this, take a maximizer and assume that there are two indices q_1 and q_2 such that $P_2^{q_1} > 0$ and $P_2^{q_2} > 0$. Then, (23) implies that $P_2^{q_1} + P_{1P}^{q_1} = P_2^{q_2} + P_{1P}^{q_2} = \nu - N_1 - N_2$, and (22) implies that $P_{1C}^{q_1} = P_{1C}^{q_2} = \mu - \nu$. Thus, considering (20), the expression corresponding to q_1 equals

$$-\log(N_1) + \log(\mu) + (\lambda-1)\log(\nu) + r(P_{1P}^{q_1})$$

where $r(x) = \log(x + N_1) - \lambda\log(x + N_1 + N_2)$. A similar statement holds for q_2. Next, using the form of the optimization problem in (20) from Proposition 2 from Section 2.1, we obtain that $P_{1P}^{q_1}$ and $P_{1P}^{q_2}$ belong to $[0, x^*]$ where x^* is given in (24). Since $r(x)$ is concave on $[0, x^*]$ (see Lemma 3), if we replace the two points indexed by q_1 and q_2 by their average (according to the weight associated to them in $p(q_1)$ and $p(q_2)$), the value of the expression in (20) does not decrease. Thus, we can restrict to the case that there is a unique index q such that $P_2^q > 0$.

Next, consider the set of indices q where $P_2^q = 0$. For these indices, (ignoring the factor $\frac{1}{2}$) the expression in (20) becomes

$$-\log(N_1) + \log(P_{1C}^q + P_{1P}^q + N_1 + N_2) + \log\left(\frac{P_{1P}^q + N_1}{P_{1P}^q + N_1 + N_2}\right).$$

Note that the choice of $\tilde{P}_{1C}^q = 0$, $\tilde{P}_{1P}^q = P_{1C}^q + P_{1P}^q$, increases the above expression as $(x+N_1)/(x+N_1+N_2)$ is increasing in x. Thus, without loss of generality we can assume that $P_{1C}^q = 0$ and the expression in (20) becomes $-\log(N_1) + \log(P_{1P}^q + N_1)$ which is concave in P_{1P}^q. Therefore, replacing all points with their average point would not decrease the expression in (20). Thus, we can assume that there is only one index q where $P_2^q = P_{1C}^q = 0$.

To sum this up, we can assume that we have (at most) one index q_1 where $P_2^{q_1} > 0$, and (at most) one index q_2 where $P_2^{q_2} = 0$. In the latter case, we must also have $P_{1C}^{q_2} = 0$. Assume that the weight associated to q_1 is α and the weight associated to q_2 is $1 - \alpha$. Then, from $\alpha P_2^{q_1} + (1-\alpha)P_2^{q_2} = P_2$, we obtain $P_2^{q_1} = \frac{P_2}{\alpha}$. Let us denote $P_1^{q_1} = P_{1P}^{q_1} + P_{1C}^{q_1}$ by $\tilde{P}$. Then, from $\alpha P_1^{q_1} + (1-\alpha)P_1^{q_2} = P_1$, we obtain $P_1^{q_2} = \frac{P_1 - \alpha\tilde{P}}{1-\alpha}$. We can parametrize the two points by

$$(P_1^{q_1}, P_2^{q_1}) = \left(\tilde{P}, \frac{P_2}{\alpha}\right), \qquad (P_1^{q_2}, P_2^{q_2}) = \left(\frac{P_1 - \alpha\tilde{P}}{1-\alpha}, 0\right).$$

It remains to show that we can restrict to $0 \le \tilde{P} \le P_1 + P_2 - \frac{P_2}{\alpha}$. Note that this inequality also implies that $\frac{P_2}{P_1+P_2} \le \alpha$. From (22), (23) and $P_2^{q_2} = P_{1C}^{q_2} = 0$ we deduce that

$$\mu \le P_{1P}^{q_2} + N_1 + N_2 = P_1^{q_2} + N_1 + N_2 = \frac{P_1 - \alpha\tilde{P}}{1-\alpha} + N_1 + N_2.$$

From (22), (23) and $P_2^{q_1} > 0$ we deduce that

$$\mu = P_2^{q_1} + P_{1C}^{q_1} + P_{1P}^{q_1} + N_1 + N_2 = P_2^{q_1} + P_1^{q_1} + N_1 + N_2 = \frac{P_2}{\alpha} + \tilde{P} + N_1 + N_2.$$

Putting these together gives $\frac{P_2}{\alpha} + \tilde{P} \le \frac{P_1 - \alpha\tilde{P}}{1-\alpha}$, or equivalently $\tilde{P} \le P_1 + P_2 - \frac{P_2}{\alpha}$.

2.1 Some useful lemmas

Lemma 3. *Let $\lambda > 1$. The second derivative of the function $r(x) = \log(x + N_1) - \lambda\log(x + N_1 + N_2)$ is strictly negative for xx^* and strictly positive for $x > x^*$ where*

$$x^* = \frac{N_1 + N_2 - N_1\sqrt{\lambda}}{\sqrt{\lambda} - 1}. \tag{24}$$

The proof is immediate by taking the second derivative.

Proposition 2. *Consider any arbitrary maximizer of* (20). *If $P_2^q > 0$ for some $q \in \{1, 2, 3\}$, then $P_{1P}^q \in [0, x^*]$ where x^* is given in* (24).

Proof. Assume that $P_2^q > 0$ and $P_{1P}^q > x^*$ for some q. From Lemma 3, one can find some $0 < \epsilon < P_2^q$ such that $\log(x + N_1) - \lambda\log(x + N_1 + N_2)$ is strictly convex on the interval $[P_{1P}^q - \epsilon, P_{1P}^q + \epsilon]$. Consider the following two points: $P_{1C}^{q_a} = P_{1C}^{q_b} = P_{1C}^q$, $P_{1P}^{q_a} = P_{1P}^q - \epsilon$, $P_{1P}^{q_b} = P_{1P}^q + \epsilon$, $P_2^{q_a} = P_2^q + \epsilon$ and $P_2^{q_b} = P_2^q - \epsilon$, i.e., we are considering two new points by preserving P_{1c}^q and $P_2^q + P_{1P}^q$ and varying P_{1P}^q. Thus, the objective function is preserved, save for $\log(P_{1P}^q + N_1) - \lambda\log(P_{1P}^q + N_1 + N_2)$. If we replace the point $(P_{1P}^q, P_{1C}^q, P_2^q)$ with these two points (with equal weight), we get a strict increase as the objective function is strictly convex on the interval $[P_{1P}^q - \epsilon, P_{1P}^q + \epsilon]$. □

Proposition 3 (Water-filling)**.** *Let Q be a random variable, $a_q \ge 0$, and $u \ge 0$. Then the maximum*

$$\max_{\substack{x_q \ge 0 \\ \mathrm{E}_Q[x_Q] \le u}} \mathrm{E}_Q[\log(x_Q + a_Q)]$$

is attained at $x_q = x_q^* := \max\{\mu^* - a_q, 0\}$ for some $\mu^* \ge 0$ satisfying $\mathrm{E}_Q\left[x_Q^*\right] = u$.

This is rather well-known and follows (with minor modifications as we are considering expected values) the proof of Example 5.2 in [15].

Proposition 4 (Layered Water-filling)**.** *Let n be a positive integer, Q be a random variable, $a_q \ge 0$, and $c_i, u_i \ge 0$ $(i = 1, \ldots, n)$. Suppose $u_1 \ge \cdots \ge u_n$. Then the maximum*

$$\max_{\substack{x_{1,q},\ldots,x_{n,q} \ge 0 \\ x_{1,q} \ge \cdots \ge x_{n,q} \\ \mathrm{E}_Q[x_{i,Q}] \le u_i}} \mathrm{E}_Q\left[\sum_{i=1}^{n} c_i \log(x_{i,Q} + a_Q)\right]$$

is attained at $x_{i,q} = x_{i,q}^* := \max\{\mu_i^* - a_q, 0\}$ for some $\mu_1^*, \ldots, \mu_n^* \ge 0$ such that $\mu_1^* \ge \cdots \ge \mu_n^*$ and $\mathrm{E}_Q\left[x_{i,Q}^*\right] = u_i$ for $i = 1, \ldots, n$.

Proof. In view of Proposition 3, the value of the relaxed maximization problem

$$\max_{\substack{x_{1,q},\dots,x_{n,q}\geq 0\\ \mathrm{E}_Q[x_{i,Q}]\leq u_i}} \mathrm{E}_Q\left[\sum_{i=1}^{n} c_i \log(x_{i,Q}+a_Q)\right]$$

$$=\sum_{i=1}^{n} c_i \max_{\substack{x_{i,q}\geq 0\\ \mathrm{E}_Q[x_{i,Q}]\leq u_i}} \mathrm{E}_Q[\log(x_{i,Q}+a_Q)]$$

is attained at $x_{i,q}=x_{i,q}^* := \max\{\mu_i^*-a_q,0\}$ for some $\mu_1^*,\dots,\mu_n^*\geq 0$ satisfying $\mathrm{E}_Q\left[x_{i,Q}^*\right]=u_i$ for $i=1,\dots,n$. Now it remains to show that $\mu_1^*\geq\dots\geq\mu_n^*$ from which it would follow that $x_{1,q}^*\geq\dots\geq x_{n,q}^*$ for all q. This follows from the assumption that $u_1\geq\dots\geq u_n$, as $\mathrm{E}_Q\left[x_{i,Q}^*\right]=u_i$ implies

$$\mathrm{E}_Q[\max\{\mu_1^*-a_Q,0\}]\geq\dots\geq \mathrm{E}_Q[\max\{\mu_n^*-a_Q,0\}].$$

Combining this with the monotonicity of the function $\mu\mapsto \mathrm{E}_Q[\max\{\mu-a_Q,0\}]$, we are done. □

Remark 1. *A physical interpretation of this proposition is: pouring different immiscible liquids into a decreasing profile and determining the levels of the various liquids in equilibrium.*

3 Gap in the proof of [14]

The authors in [14] argue that `"Since the scenarios illustrated in Fig. 6(a)(b)(c) are equivalent to noiseberg cases, it remains to argue that the power allocation scheme with at spectrum top as in Fig. 6(d) is not optimal. This is because the achievable rates under such a scheme are formed by convex combinations of points on the curve of associated broadcast channel capacity, as the flat top requires` $\frac{P_{1c\bar{\lambda}}}{\bar{\lambda}}=\frac{P_{1c\lambda}}{\lambda}$`. As the broadcast channel capacity curve is convex, we can only achieve the points on the chord, which are dominated by the points on the curve corresponding to the scheme with no frequency division. Thus they are not optimal."` This is rather identical to the intuitive reasoning in [7] while one of the authors proposed the noiseberg region.

The main error with this reasoning is the following: The associated broadcast channel is assumed to have a total power budget of $P_t=P_1+P_2$. The rate points on the capacity region are of the form $R_1=\frac{1}{2}\log\left(1+\frac{\alpha P_t}{N_1}\right)$, $R_2=\frac{1}{2}\log\left(1+\frac{(1-\alpha)P_t}{\alpha P_t+N_1+N_2}\right)$. These rate pairs (since it allows for transmitter cooperation) constitute an outer bound for the Gaussian Z-interference channel. When $(1-\alpha)P_t\geq P_2$, i.e. the power budget for the common message M_2 exceeds the power budget for the second transmitter, one cannot achieve $R_2=\frac{1}{2}\log\left(1+\frac{(1-\alpha)P_t}{\alpha P_t+N_1+N_2}\right)$. (Unfortunately this happens at all points above the maximum sum-rate point, i.e. essentially the entire regime of interest). One achievable rate-pair for the interference channel in this case is obtained by assigning the power difference, $(1-\alpha)P_t-P_2$, to be used by the first transmitter. This leads to the following set of achievable rate pairs when $(1-\alpha)P_t\geq P_2$: $R_1=\frac{1}{2}\log\left(1+\frac{\alpha P_t}{N_1}\right)+\frac{1}{2}\log\left(1+\frac{(1-\alpha)P_t-P_2}{P_2+\alpha P_t+N_1+N_2}\right)$, $R_2=\frac{1}{2}\log\left(1+\frac{P_2}{\alpha P_t+N_1+N_2}\right)$. These set of points (R_1,R_2) are not known to lie on a convex curve and hence the previous argument breakdown.

Remark 2. *A moment's reflection is useful here: were the points on the associated broadcast channel achievable, then it simultaneously would be an inner bound and outer bound (hence the capacity region). This is clearly not true as the capacity region has true non-trivial corner points unlike the smooth curve of the capacity region for the associated broadcast channel.*

4 Refinement of the Gaussian optimality conjecture

The noiseberg conjecture states that

$$\mathfrak{C}f_{\lambda,\mathrm{GS}}(P_1,P_2)=\max_{\alpha,\tilde{P}}\left(\alpha f_{\lambda,\mathrm{GS}}\left(\tilde{P},\frac{P_2}{\alpha}\right)+(1-\alpha)f_{\lambda,\mathrm{GS}}\left(\frac{P_1-\alpha\tilde{P}}{1-\alpha},0\right)\right). \tag{25}$$

Let us define the set $\mathcal{P}_{f_{\lambda,\mathrm{GS}}}$ to be the set of positive pairs (P_1,P_2) such that the maximum in (25) is attained at $\alpha=1$, i.e., $\mathfrak{C}f_{\lambda,\mathrm{GS}}(P_1,P_2)=f_{\lambda,\mathrm{GS}}(P_1,P_2)$ and no time-sharing between two points is needed in this case. We propose a Gaussian optimality conjecture for power pairs in set $\mathcal{P}_{f_{\lambda,\mathrm{GS}}}$. This is an explicit and succinct conjecture about Gaussian optimality consistent with the optimality and sub-optimality results established in literature so far.

Conjecture 2. *Let $\lambda \geq 1$. Take $N_1, N_2 > 0$ and some pair $(P_1, P_2) \in \mathcal{P}_{f_{\lambda,GS}}$. Let $\mathbf{Z}_1 \sim \mathcal{N}(0, N_1 I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2 I)$ be independent random variables in $\mathbb{R}^n$. Then the supremum*

$$\sup_{\substack{p(\mathbf{x}_1)p(\mathbf{x}_2)\\ \mathrm{E}\left[\|\mathbf{X}_1\|^2\right] \leq nP_1 \\ \mathrm{E}\left[\|\mathbf{X}_2\|^2\right] \leq nP_2}} \left((\lambda-1)h(\mathbf{X}_1+\mathbf{X}_2+\mathbf{Z}_1+\mathbf{Z}_2) + h(\mathbf{X}_1+\mathbf{Z}_1) - \lambda h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2)\right),$$

where $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Z}_1, \mathbf{Z}_2$ are mutually independent random variables, is attained by Gaussian $\mathbf{X}_1$ and $\mathbf{X}_2$ such that the covariance of $\mathbf{X}_2$ is a multiple of identity.

Definition 3. *The achievable rate obtained by the convex combination $\alpha f_{\lambda,GS}\big(0, \frac{P_2}{\alpha}\big) + (1-\alpha) f_{\lambda,GS}\Big(\frac{P_1}{1-\alpha}, 0\Big)$, for some $\alpha \in [0,1]$, is defined as the time-division rate.*

Let $\mathcal{S}$ denote the set of model parameters $(P_1, P_2, N_1, N_2, \lambda)$ for which the maximum in (25) is attained at a point $\tilde{P} > 0$. Alternately, time-division strategy is not the maximizer for this $(P_1, P_2, N_1, N_2, \lambda)$.

Remark 3. *Note that the region $\mathcal{S}$ can be explicitly determined by the noiseberg region. For large classes of channel parameters (P_1, P_2, N_1, N_2), one can show that $\mathcal{S}$ contains the entire capacity region, i.e. all $\lambda \geq 1$. However, numerical simulations do indicate that for some special choice of parameters $\tilde{P} = 0$ may occur; one such example seems to be $N_1 = N_2 = 1$, $P_1 = 3(1-t)$, $P_2 = 2t$ and $\lambda = (\ln(1/4) + 3/4)/(\ln(1/2) + 1/2)$ where $t \in (0,1)$ is arbitrary.*

Theorem 1. *If Conjecture 2 holds, Han–Kobayashi weighted sum-rate is equal to the corresponding weighted sum-capacity of GZIC for the points in $\mathcal{S}$, where $\mathcal{S}$ is defined above.*

Proof. Take some arbitrary $\lambda \geq 1$. Let $g(P_1, P_2)$ be the achievable λ-sum rate $\sup_{(R_1,R_2)}(R_1 + \lambda R_2)$, where (R_1, R_2) is taken over the capacity region of the Z-interference channel (1) with power constraints $\mathrm{E}[X_1^2] \leq P_1$ and $\mathrm{E}[X_2^2] \leq P_2$. Note that $g(P_1, P_2)$ is concave in (P_1, P_2); this can be shown by using time-sharing (power control). Moreover, $g(P_1, P_2) \geq \mathfrak{C} f_{\lambda,\mathrm{GS}}(P_1, P_2)$ since $g(P_1, P_2)$ is the λ-sum rate of the actual capacity region and is greater than or equal to the value obtained by the HK inner bound with Gaussian signalling. We need to prove that $g(P_1, P_2) = \mathfrak{C} f_{\lambda,\mathrm{GS}}(P_1, P_2)$.

Recell that the function $f_{\lambda,\mathrm{GS}}(P_1, P_2)$ is differentiable with a continuous derivative. The noiseberg conjecture implies that for any $P_1, P_2 > 0$, $\mathfrak{C} f_{\lambda,\mathrm{GS}}(P_1, P_2)$ is the convex combination of two points for the form $\left(\tilde{P}, \frac{P_2}{\alpha}\right)$ and $\left(\frac{P_1 - \alpha\tilde{P}}{1-\alpha}, 0\right)$. Since the optimal $\tilde{P} > 0$ for points in $\mathcal{S}$, the pair $\left(\tilde{P}, \frac{P_2}{\alpha}\right)$ will be in the interior and Lemma 2 from Section 1.2 implies that for any $\vec{x} = (P_1^*, P_2^*) \in \mathcal{S}$ we have

$$\mathfrak{C} f_{\lambda,\mathrm{GS}}(\vec{x}) = \inf_{\vec{y} \in \mathcal{P}'_{f_{\lambda,\mathrm{GS}}}} \left[f_{\lambda,\mathrm{GS}}(\vec{y}) + \langle \nabla f_{\lambda,\mathrm{GS}}(\vec{y}), \vec{x} - \vec{y}\rangle\right], \tag{26}$$

where $\mathcal{P}'_{f_{\lambda,\mathrm{GS}}}$ is the set of points in $\mathcal{P}_{f_{\lambda,\mathrm{GS}}}$ with strictly positive coordinates.

Lemma 4 (given below in this section) shows that if Conjecture 2 holds, then for any $(P_1, P_2) \in \mathcal{P}_{f_{\lambda,\mathrm{GS}}}$, we have $g(P_1, P_2) = f_{\lambda,\mathrm{GS}}(P_1, P_2)$. Then, Lemma 5 from Appendix A shows that for any for any $(P_1, P_2) \in \mathcal{P}'_{f_{\lambda,\mathrm{GS}}}$ we have

$$\nabla g(P_1, P_2) = \nabla f(P_1, P_2).$$

Finally, (26) implies that for any $\vec{x} = (P_1^*, P_2^*) \in \mathcal{S}$ we have

$$\begin{aligned}
\mathfrak{C} f_{\lambda,\mathrm{GS}}(\vec{x}) &= \inf_{\vec{y} \in \mathcal{P}'_{f_{\lambda,\mathrm{GS}}}} \left[f_{\lambda,\mathrm{GS}}(\vec{y}) + \langle \nabla f_{\lambda,\mathrm{GS}}(\vec{y}), \vec{x} - \vec{y}\rangle\right] \\
&= \inf_{\vec{y} \in \mathcal{P}'_{f_{\lambda,\mathrm{GS}}}} \left[g(\vec{y}) + \langle \nabla g(\vec{y}), \vec{x} - \vec{y}\rangle\right] \\
&\geq g(\vec{x}),
\end{aligned}$$

where the last step follows from concavity of g. Since we also had $g(\vec{x}) \geq \mathfrak{C} f_{\lambda,\mathrm{GS}}(\vec{x})$, this completes the proof. $\square$

Lemma 4. *If Conjecture 2 holds, then for any* $(P_1, P_2) \in \mathcal{P}_{f_{\lambda,GS}}$*, we have* $g(P_1, P_2) = f_{\lambda,GS}(P_1, P_2)$ *where* $g(P_1, P_2)$ *is the achievable* λ*-sum rate* $\sup_{(R_1,R_2)}(R_1 + \lambda R_2)$ *of the capacity region of the Z-interference channel* (1) *with power constraints* $\mathrm{E}[X_1^2] \le P_1$ *and* $\mathrm{E}[X_2^2] \le P_2$*.*

Proof of Lemma 4. From the scalar case of Theorem 1 of [9], it follows that the covariance of the optimal $\mathbf{X}_1$ is also a multiple of the identity. In such case, the covariance of the optimal $\mathbf{X}_1$ and $\mathbf{X}_2$ are $\hat{P}_1 I$ and $P_2 I$, respectively, for some $0 \le \hat{P}_1 \le P_1$. Then a standard application of Fano's inequality gives that for any sequence of codebooks $(\mathbf{X}_1, \mathbf{X}_2)$ of rate (R_1, R_2) satisfying the power constraints $\mathrm{E}[\|\mathbf{X}_1\|^2] \le nP_1$ and $\mathrm{E}[\|\mathbf{X}_2\|^2] \le nP_2$, whose average probability of error goes to zero, there exists a sequence $\epsilon_n \ge 0$ such that $\epsilon_n \to 0$ and

$$\begin{aligned}
&R_1 + \lambda R_2 - \epsilon_n \\
&\le \frac{1}{n}(I(\mathbf{X}_1; \mathbf{X}_1 + \mathbf{Z}_1) + \lambda I(\mathbf{X}_2; \mathbf{X}_1 + \mathbf{X}_2 + \mathbf{Z}_1 + \mathbf{Z}_2)) \\
&= \frac{1}{n}(-h(\mathbf{Z}_1) + h(\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{Z}_1 + \mathbf{Z}_2) \\
&\quad + (\lambda - 1)h(\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{Z}_1 + \mathbf{Z}_2) + h(\mathbf{X}_1 + \mathbf{Z}_1) - \lambda h(\mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2)) \\
&\le \frac{1}{2}\Big(- \log(N_1) + \log(P_1 + P_2 + N_1 + N_2) \\
&\quad + \max_{0 \le \hat{P}_1 \le P_1} \Big((\lambda - 1)\log\Big(\hat{P}_1 + P_2 + N_1 + N_2\Big) + \log\Big(\hat{P}_1 + N_1\Big) - \lambda \log\Big(\hat{P}_1 + N_1 + N_2\Big)\Big)\Big) \\
&= f_{\lambda,\mathrm{GS}}(P_1, P_2).
\end{aligned}$$

□

5 Conclusion

In this paper, we proved the noiseberg conjecture of [7], which simplifies the calculation of the HK achievable region with Gaussian signaling to a two-variable optimization problem involving α and $\tilde{P}$ given in (25). The intuitive notion is that pushing noise above the surface level frees up prime (low SNR) space for the signals that matter in the lower portions of the spectrum. The mathematical reasoning is that time or frequency division does a linear combination of two modes of operation, one of pure superposition and another where the second transmitter is silent. This leads to an achievable point in the concave envelope of the surfaces, otherwise unattainable without the zipline-style connection. Specifically, we considered two classes of model parameters in the conjecture, (i) when the optimizer satisfies $\alpha = 1$ and (ii) when the optimizer satisfies $\tilde{P} > 0$. Then, we considered a refined form of the Gaussian optimality conjecture of [9] when we restrict to class (i) and showed that it implies a characterization of the capacity region for class (ii). Finding a closed-form expression for class (i) and class (ii) of parameters is left as future work.

Acknowledgements

Max Costa thanks for the support of Huawei Hong Kong Research. The work described of Chandra Nair and David Ng was fully supported by the following two grants from the Research Grants Council of the Hong Kong Special Administrative Region, China CUHK 14221822 and CUHK 14210120.

References

[1] H. Sato, "The capacity of the Gaussian interference channel under strong interference (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, 1981.

[2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 49–60, 1981.

[3] M. Costa, "On the Gaussian interference channel," *IEEE Transactions on Information Theory*, vol. 31, no. 5, pp. 607–615, 1985.

[4] C. Nair, L. Xia, and M. Yazdanpanah, "Sub-optimality of Han–Kobayashi achievable region for interference channels," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 2416–2420.

[5] A. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information and Control*, vol. 2, no. 2, pp. 101–112, 1959.

[6] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2087–2104, 2014.

[7] M. H. M. Costa, "Noisebergs in Z Gaussian interference channels," in *Proc. Information Theory and Applications Workshop (ITA)*, 2011, pp. 1–6.

[8] M. H. M. Costa and C. Nair, "On the achievable rate sum for symmetric Gaussian interference channels," in *Proc. Information Theory and Applications Workshop (ITA)*, 2012.

[9] M. Costa, C. Nair, D. Ng, and Y. N. Wang, "On the structure of certain non-convex functionals and the Gaussian Z-interference channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 1522–1527.

[10] A. Gohari, C. Nair, and D. Ng, "An information inequality motivated by the Gaussian Z-interference channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2021.

[11] J. Liu, "Stability of the Gaussian stationary point in the Han-Kobayashi region for Z-interference channels," 2022. [Online]. Available: https://arxiv.org/abs/2209.00163

[12] M. H. M. Costa and C. Nair, "Gaussian Z-interference channel: around the corner," in *Proc. Information Theory and Applications Workshop (ITA)*, 2016, pp. 1–6.

[13] M. Costa, C. Nair, and D. Ng, "On the Gaussian Z-interference channel," in *Proc. Information Theory and Applications Workshop (ITA)*, 2017, pp. 1–15.

[14] Y. Zhao, F. Zhu, and B. Chen, "The Han-Kobayashi region for a class of Gaussian interference channels with mixed interference," in *2012 IEEE International Symposium on Information Theory Proceedings*, 2012, pp. 2791–2795.

[15] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

A Properties of the convex envelope and convex functions

Proof of Lemma 1. Take some point $\vec{x}$. Consider the compact set $\|\vec{r}\| \leq L$ for some $L > \|\vec{x}\|$. Let $\mathfrak{C}_L f(\vec{x})$ be the upper concave envelope of f when we restrict the domain to vectors $\vec{r}$ satisfying $\|\vec{r}\| \leq L$. By Caratheodery's theorem, one can write

$$\mathfrak{C}_L f(\vec{x}) = \sum_{j=1}^{n+1} \omega_{L,j} f(\vec{z}_{L,j})$$

for some points $\vec{z}_{L,j} \in \mathbb{R}^n_{\geq 0}$ and weights $\omega_{L,j}$ for $1 \leq j \leq n+1$ satisfying

$$\vec{x} = \sum_{j=1}^{n+1} \omega_{L,j} \vec{z}_{L,j}. \tag{27}$$

Note that (27) implies that the non-negative vector $\vec{z}_{L,j}$ is coordinatewise less than or equal to $\vec{x}/\omega_{L,j}$. Therefore,

$$\|\vec{z}_{L,j}\| \leq \frac{\|\vec{x}\|}{\omega_{L,j}}. \tag{28}$$

For each L, the tuple $(\omega_{L,1}, \omega_{L,2}, \cdots, \omega_{L,n+1})$ lies in the probability simplex on $n+1$ variables. Therefore as we let L tend to infinity, we can find a convergent subsequence, i.e., for a sequence $(L_1, L_2, \cdots)$ where $L_i \to \infty$, the sequence $(\omega_{L_i,1}, \omega_{L_i,2}, \cdots, \omega_{L_i,n+1})$ converges to some $(\omega_1^*, \omega_2^*, \cdots, \omega_{n+1}^*)$.

Let $\mathcal{J}$ be the set of indices j where $\omega_j^* > 0$. Note that for such a $j \in \mathcal{J}$, the length of $\|\vec{z}_{L_i,j}\|$ will remain bounded from (28). Therefore, by taking a convergent subsequent, without loss of generality we can assume that the sequence $\vec{z}_{L_i,j}$ converges to a limit vector $\vec{z}_j^*$ for any $j \in \mathcal{J}$.

Next, take some j where $\omega_j^* = 0$. We claim that

$$\lim_{i\to\infty} \omega_{L_i,j} f(\vec{z}_{L_i,j}) = 0. \tag{29}$$

If $\vec{z}_{L_i,j}$ remains bounded as i tends to infinity, (29) holds. Otherwise, we have a subsequence where $\|\vec{z}_{L_i,j}\|$ tends to infinity. Thus, without loss of generality we can assume that $\lim_{i\to\infty}\|\vec{z}_{L_i,j}\| = \infty$. In this case,

$$\begin{aligned}
\lim_{i\to\infty} \omega_{L_i,j} f(\vec{z}_{L_i,j}) &= \|\vec{x}\| \cdot \lim_{i\to\infty} \frac{f(\vec{z}_{L_i,j})}{\frac{\|\vec{x}\|}{\omega_{L_i,j}}} && (30)\\
&\leq \|\vec{x}\| \cdot \lim_{i\to\infty} \frac{f(\vec{z}_{L_i,j})}{\|\vec{z}_{L_i,j}\|} && (31)\\
&= 0, && (32)
\end{aligned}$$

where we used (18) and (28).

Therefore, by taking the limit as i goes to infinity and using (29) we obtain

$$\mathfrak{C}f(\vec{x}) = \sum_{j\in\mathcal{J}} \omega_j^* f(\vec{z}_j^*) \tag{33}$$

for some points $\vec{z}_j^* \in \mathbb{R}^n_{\geq 0}$ and positive weights ω_j^* for $1 \leq j \leq r$ satisfying

$$\vec{x} \geq \sum_{j\in\mathcal{J}} \omega_j^* \vec{z}_j^*. \tag{34}$$

This completes the proof for the first part of the lemma. If f is non-decreasing in its coordinates, by increasing z_j^* in (34), we can make it hold with equality. Moreover, the right hand side of (33) will not decrease.

Remark 4. *The above proof is similar to (and motivated by) an argument in the Appendix [6].*

□

Proof of Lemma 2. Since $\mathfrak{C}f = f$ on $\mathcal{P}'_f$, and f is differentiable in the interior of the domain, from Lemma 5 we deduce that $\nabla(\mathfrak{C}f)(\vec{y}) = \nabla f(\vec{y})$ for every $\vec{y} \in \mathcal{P}'_f$.

Therefore, for any arbitrary $\vec{x}$ we have

$$\begin{aligned}
&\inf_{\vec{y}\in\mathcal{P}'_f} [f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y}\rangle]\\
&= \inf_{\vec{y}\in\mathcal{P}'_f} [\mathfrak{C}f(\vec{y}) + \langle \nabla(\mathfrak{C}f)(\vec{y}), \vec{x} - \vec{y}\rangle]\\
&\geq \mathfrak{C}f(\vec{x}). && (35)
\end{aligned}$$

where in the last step we used the concavity of $\mathfrak{C}f$. It remains to show that for any $\vec{x} \in \mathcal{S}$ we have

$$\inf_{\vec{y}\in\mathcal{P}'_f} [f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y}\rangle] \leq \mathfrak{C}f(\vec{x}).$$

Take some arbitrary $\vec{x} \in \mathcal{S}$. It suffices to identify some $\vec{y} \in \mathcal{P}'_f$ such that

$$f(\vec{y}) + \langle \nabla f(\vec{y}), \vec{x} - \vec{y}\rangle \leq \mathfrak{C}f(\vec{x}). \tag{36}$$

Since $\vec{x} \in \mathcal{S}$ one can find $r \leq n+1$ points $\vec{z}_1, \cdots, \vec{z}_r$ and positive weights ω_i $(1 \leq i \leq r)$ adding up to one such that

$$\mathfrak{C}f(\vec{x}) = \sum_{j=1}^{r} \omega_j f(\vec{z}_j) \tag{37}$$

and

$$\vec{x} = \sum_{j=1}^{r} \omega_j \vec{z}_j.$$

Moreover, $\vec{z}_j$ is in the interior of the domain for some j.

We claim that $z_j \in \mathcal{P}_f$ for any $1 \le j \le r$. To show this, observe that

$$\begin{aligned} \sum_{j=1}^{r} \omega_j \mathfrak{C}f(\vec{z}_j) \le \mathfrak{C}f(\vec{x}) &= \sum_{j=1}^{r} \omega_j f(\vec{z}_j) \qquad (38) \\ &\le \sum_{j=1}^{r} \omega_j \mathfrak{C}f(\vec{z}_j) \end{aligned}$$

where (38) follows from concavity of $\mathfrak{C}f$ and (37). Thus, $\mathfrak{C}f(\vec{z}_j) = f(\vec{z}_j)$ for all $1 \le j \le r$.

Next, take some small $\epsilon > 0$. We have

$$\begin{aligned} \mathfrak{C}f(\vec{x}) &\ge \sum_{j=1}^{r} \omega_j f(\vec{z}_j + \epsilon(\vec{x} - \vec{z}_j)) \\ &= \sum_{j=1}^{r} \omega_j \big[f(\vec{z}_j) + \epsilon \langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle + o(\epsilon) \big] \\ &= \mathfrak{C}f(\vec{x}) + \sum_{j=1}^{r} \omega_j \big[\epsilon \langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle + o(\epsilon) \big] \end{aligned}$$

Thus,

$$\sum_{j=1}^{r} \omega_j \big[\langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle \big] \le 0.$$

Consequently,

$$\mathfrak{C}f(\vec{x}) \ge \sum_{j=1}^{r} \omega_j \big[f(\vec{z}_j) + \langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle \big].$$

On the other hand, since $z_j \in \mathcal{P}_f$, from (35), we have that

$$\mathfrak{C}f(\vec{x}) \le f(\vec{z}_j) + \langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle.$$

This implies that for any j

$$\mathfrak{C}f(\vec{x}) = f(\vec{z}_j) + \langle \nabla f(\vec{z}_j), \vec{x} - \vec{z}_j) \rangle.$$

Thus, we can choose the z_j that lies in the interior of the domain as evidence for (36). □

Lemma 5. *For two functions f and g assume that*

1. *$f(\vec{y}) \le g(\vec{y})$ for all vectors $\vec{y}$,*
2. *$g(\cdot)$ is concave,*
3. *$f(\vec{x}) = g(\vec{x})$ for some vector x.*

Then, $\partial g(\vec{x}) \subseteq \partial f(\vec{x})$ where $\partial g(\vec{x})$ and $\partial f(\vec{x})$ are the sets of sub-gradients of f and g respectively at $\vec{x}$.

Proof. Take an arbitrary sub-gradient vector $\vec{v} \in \partial g$ at $\vec{x}$. Then, for any $\vec{y}$ we have

$$\begin{aligned} f(\vec{x}) + \langle \vec{v}, \vec{y} - \vec{x} \rangle &= g(\vec{x}) + \langle \vec{v}, \vec{y} - \vec{x} \rangle \qquad (39) \\ &\ge g(\vec{y}) \qquad (40) \\ &\ge f(\vec{y}), \qquad (41) \end{aligned}$$

where (39) holds because of assumption 3, (40) holds because of assumption 2 and (41) holds because of assumption 1. This shows that $\vec{v} \in \partial f$ at $\vec{x}$. □