On the structure of certain non-convex functionals and the Gaussian Z-interference channel

Max Costa
Department of Communications
University of Campinas
São Paulo, Brasil
max@fee.unicamp.br

Chandra Nair, David Ng, and Yannan Wang
Department of Information Engineering
The Chinese University of Hong Kong
Hong Kong
{chandra,david,dustin}@ie.cuhk.edu.hk

***Abstract*—In this paper we establish that a maximizer of a non-convex problem in positive semidefinite matrices has a certain property using information-theoretic methods. Further, we propose a Gaussian extremality conjecture, which if true, would imply that Gaussian signaling achieves the capacity region of the Gaussian Z-interference channel. The non-convex problem mentioned above arose naturally in the reduction from the conjecture to the optimality of Gaussian signaling.**

I. Introduction

From the celebrated entropy power inequality due to Shannon [1] (say pictured in Lieb's equivalent formulation [2]) to Brascamp–Lieb [3] inequalities, Gaussian distributions have often turned up as extremizers of non-convex functionals over spaces of probability distributions. The proofs of Gaussian optimality have traditionally been obtained by showing monotonicity along a path [4], or more recently using transport maps [5], and most recently using subadditivity of an associated functional [6] and invoking the technique developed in [7] that employs the Skitovič-Darmois characterization of Gaussians [8], [9].

Clearly the "monotonicity along a path" argument [4] shows that any local maximizer of the functional is also a global maximizer. In this paper we employ an observation that the Gaussian optimality technique developed in [7] can be applied to derive properties of the maximizers of concave envelopes of functionals. While these properties were implicitly observable in previous instances where the technique was employed, here the subadditivity technique is used primarily to obtain the desired property. We then present a conjecture that Gaussians are extremizers of a certain functional and show that the capacity region of the Gaussian Z-interference channel can be deduced as a corollary if the conjecture is true. A key element of this reduction involves the use of the property of the maximizer derived using the Gaussian optimality technique.

Recently there was a result [10] making use of geodesic convexity saying that a functional on the space of Gaussian distributions related to the Brascamp–Lieb constant has the property that the local maximizers of which are also global maximizers. Motivated by this and certain other non-convex optimization problems in network information theory, we consider a particular functional and show that it has such property on the space of distributions through information-theoretic methods. We believe such observations can help develop gradient descent based algorithms to efficiently compute the optimizers of these non-convex functionals.

A. Preliminaries

The *upper concave envelope* $\mathcal{C}_x[f(x)]$ of a functional f defined on a convex subset of some Hilbert space can be defined by one of the many equivalent ways (cf. [11]):

- (i) $\mathcal{C}_x[f(x)] := \inf_g g(x)$, where the infimum runs over all concave functional g that upper bounds f,
- (ii) $\mathcal{C}_x[f(x)] := \sup_\mu \int f d\mu$, where the supremum runs over all Borel probability measures μ with mean x,
- (iii) $\mathcal{C}_x[f(x)] := \inf_\alpha[\sup_{\hat{x}}[f(\hat{x}) - \langle \alpha, \hat{x}\rangle] + \langle \alpha, x\rangle]$, which is the *dual characterization* of upper concave envelope using Fenchel duality.

Remark 1. The dual characterization plays an important role in the formulation of Conjecture 1.

II. Main result

Theorem 1. *Let $\beta \geq 1$ and $N_2 \geq 0$. Define*

$$\psi_G(K_1, K_2) := \frac{1}{2}\Big[(\beta-1)\log|K_2 + K_1 + I + N_2 I| + \log|K_1 + I| - \beta \log|K_1 + I + N_2 I|\Big] \quad (1)$$

for $k \times k$ ($k \geq 1$) matrices $K_1, K_2 \succeq 0$. Then it holds that

$$\mathcal{C}_{K_1}\big[\psi_G(K_1, K_2)\big] = \max_{\substack{\hat{K}_1 \succeq 0 \\ \hat{K}_1 \preceq K_1}} \psi_G(\hat{K}_1, K_2)$$

for any $K_1, K_2 \succeq 0$.

Remark 2. For the scalar case this result can be shown directly without much effort by showing the following properties of the function $\psi_{\mathrm{G}}(Q_1, Q_2)$ for scalar $Q_1, Q_2 \geq 0$.

- (i) $Q_1 \mapsto \psi_{\mathrm{G}}(Q_1, Q_2)$ either is increasing on $[0, +\infty)$, is decreasing on $[0, +\infty)$, or is increasing on $[0, Q_1^*)$ and decreasing on $[Q_1^*, +\infty)$ for some $Q_1^* > 0$,
- (ii) $\frac{\partial}{\partial Q_1}\psi_{\mathrm{G}}(Q_1, Q_2) \geq 0$ implies $\frac{\partial^2}{\partial Q_1^2}\psi_{\mathrm{G}}(Q_1, Q_2) \leq 0$,
- (iii) $\lim_{Q_1 \to +\infty} \psi_{\mathrm{G}}(Q_1, Q_2) = 0$.

However for high-dimensional spaces it does not seem to admit a simple extension of the above argument especially since K_1, K_2 may not be simultaneously diagonalizable. This necessitated us to come up with the different argument presented below. This technique adds to the list of linear algebra inequalities that have information-theoretic proofs [12].

Proof of Theorem 1. We first show the "$\le$" side. For independent random variables $\mathbf{X}_1, \mathbf{X}_2$ in $\mathbb{R}^k$ we denote

$$\psi(\mathbf{X}_1,\mathbf{X}_2) := (\beta-1)h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) + h(\mathbf{X}_1+\mathbf{Z}_1) - \beta h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) \quad (2)$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0,I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0,N_2 I)$ are independent with $(\mathbf{X}_1,\mathbf{X}_2)$. Notice that when $\mathbf{X}_i \sim (0,K_i)$ for $i=1,2$ one has $\psi(\mathbf{X}_1,\mathbf{X}_2) = \psi_{\mathrm{G}}(K_1,K_2)$. Now fix $K_2 \succeq 0$ and $\mathbf{X}_2 \sim \mathcal{N}(0,K_2)$, and we have

$$\begin{aligned}\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1,K_2)\right] &\overset{\text{(a)}}{\le} \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1)\\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]\preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1},\mathbf{X}_2)]\\ &\overset{\text{(b)}}{=} \max_{\substack{\hat{K}_1\succeq 0\\ \hat{K}_1\preceq K_1}} \psi_{\mathrm{G}}(\hat{K}_1,K_2)\end{aligned}$$

for any $K_1 \succeq 0$, where (a) holds since the right hand side is concave in K_1 and upper bounds $\psi_{\mathrm{G}}(K_1,K_2)$ (by taking $\mathbf{U}_1$ to be constant and $\mathbf{X}_1 \sim \mathcal{N}(0,K_1)$), and (b) follows from Proposition 1.

Now we show the "$\ge$" side. Using the dual characterization of upper concave envelope we get

$$\begin{aligned}&\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1,K_2)\right]\\ &= \inf_{\substack{\Sigma_1\\ \Sigma_1=\Sigma_1^T}} \left[\sup_{\hat{K}_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right] + \mathrm{tr}(\Sigma_1 K_1)\right]\\ &\overset{\text{(a)}}{=} \inf_{\Sigma_1\succeq 0}\left[\sup_{\hat{K}_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right] + \mathrm{tr}(\Sigma_1 K_1)\right]\\ &\ge \sup_{\hat{K}_1\succeq 0}\inf_{\Sigma_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1) + \mathrm{tr}(\Sigma_1 K_1)\right]\\ &\ge \max_{\substack{\hat{K}_1\succeq 0\\ \hat{K}_1\preceq K_1}} \psi_{\mathrm{G}}(\hat{K}_1,K_2)\end{aligned}$$

for $K_1, K_2 \succeq 0$, where (a) follows from the fact that

$$\sup_{\hat{K}_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right] = +\infty$$

for any symmetric Σ_1 with $\Sigma_1 \not\succeq 0$. $\square$

Proposition 1. *Let $\beta \ge 1$, $K_1 \succeq 0$ and let $\mathbf{X}_2, \mathbf{Z}_1, \mathbf{Z}_2$ be independent Gaussian random variables in $\mathbb{R}^k$. Then the maximum*

$$\max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1)\\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]\preceq K_1}}\Big[(\beta-1)h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1) + h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{U}_1) - \beta h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1)\Big]$$

is attained by some zero-mean Gaussian $\mathbf{X}_1$ and constant random variable $\mathbf{U}_1$.

We need a few lemmas for establishing Proposition 1. Lemma 1 is a well-known property called double Markovity (cf. Problem 16.25 of [13]). Lemma 2 relies on the fact that the characteristic function of a Gaussian random variable vanishes nowhere. Lemma 3 follows from a characterization of Gaussian random variables given by Ghurye and Olkin [14], which was shown by solving a functional equation, generalizing an earlier functional equation of Skitovič [8], satisfied by the characteristic functions.

Lemma 1. *Let $\mathbf{Q}$ be a random variable and let $(\mathbf{X},\mathbf{Y},\mathbf{Z})$ be random variables on $\mathbb{R}^k$ such that for any $\mathbf{q}$ the conditional distribution $p(\mathbf{x},\mathbf{y},\mathbf{z}|\mathbf{q})$ has everywhere non-zero density. Suppose*

$$\mathbf{X}\to(\mathbf{Y},\mathbf{Q})\to\mathbf{Z} \quad \textit{and} \quad \mathbf{Y}\to(\mathbf{X},\mathbf{Q})\to\mathbf{Z}$$

form Markov chains. Then

$$(\mathbf{X},\mathbf{Y})\to\mathbf{Q}\to\mathbf{Z}$$

forms a Markov chain.

Proof. For any $\mathbf{q},\mathbf{x},\mathbf{y},\mathbf{z}$ the Markov chains give

$$p(\mathbf{z}|\mathbf{q},\mathbf{x}) = p(\mathbf{z}|\mathbf{q},\mathbf{x},\mathbf{y}) = p(\mathbf{z}|\mathbf{q},\mathbf{y})$$

and hence

$$\begin{aligned}p(\mathbf{z}|\mathbf{q}) &= \mathrm{E}_{\mathbf{X}}[p(\mathbf{z}|\mathbf{q},\mathbf{X})|\mathbf{Q}=\mathbf{q}]\\ &= \mathrm{E}_{\mathbf{X}}[p(\mathbf{z}|\mathbf{q},\mathbf{y})|\mathbf{Q}=\mathbf{q}]\\ &= p(\mathbf{z}|\mathbf{q},\mathbf{y})\\ &= p(\mathbf{z}|\mathbf{q},\mathbf{x},\mathbf{y})\end{aligned}$$

as required. $\square$

Lemma 2. *Let $\mathbf{X}_1,\mathbf{X}_2$ be random variables in $\mathbb{R}^k$ and $\mathbf{Z}_1,\mathbf{Z}_2$ be k-dimensional Gaussian random variables such that $(\mathbf{X}_1,\mathbf{X}_2)$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are independent. Then $\mathbf{X}_1+\mathbf{Z}_1 \perp \mathbf{X}_2+\mathbf{Z}_2$ implies $\mathbf{X}_1 \perp \mathbf{X}_2$.*

Proof. See Proposition 2 of [7]. $\square$

Lemma 3. *Let $\mathbf{X}_1,\mathbf{X}_2$ be random variables in $\mathbb{R}^k$ such that $\mathbf{X}_1 \perp \mathbf{X}_2$ and $(\mathbf{X}_1+\mathbf{X}_2) \perp (\mathbf{X}_1-\mathbf{X}_2)$. Then $\mathbf{X}_1,\mathbf{X}_2$ are Gaussians having the same covariance matrix.*

Proof. See Corollary 3 of [7]. $\square$

Proof of Proposition 1. By the translation-invariance of entropy we can without loss of generality assume $\mathbf{X}_2,\mathbf{Z}_1,\mathbf{Z}_2$ are zero-mean Gaussians. For any distribution $p(\mathbf{x}_1)$ on $\mathbb{R}^k$ denote

$$\phi(p(\mathbf{x}_1)) := (\beta-1)h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) + h(\mathbf{X}_1+\mathbf{Z}_1) - \beta h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2)$$

where $\mathbf{X}_1 \sim p(\mathbf{x}_1)$. Let $p^*(\mathbf{x}_1,\mathbf{u}_1)$ be a maximizer (existence of which can be justified by Prokhorov theorem through techniques in Appendix II of [7]) for

$$v := \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1)\\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]\preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\phi(p(\mathbf{x}_1|\mathbf{U}_1))].$$

Assume without loss of generality that $p^*(\mathbf{x}_1|\mathbf{u}_1)$ has mean zero, or otherwise replace $\mathbf{X}_1$ by $\mathbf{X}_1 - \mathrm{E}[\mathbf{X}_1|\mathbf{U}_1]$, which indeed satisfies the constraint. To prove our proposition it suffices to show that $p^*(\mathbf{x}_1|\mathbf{u}_1)$ is a Gaussian distribution with covariance matrix independent of choice of $\mathbf{u}_1$. We shall show this by a subadditivity argument.

Define the random variables

$$(\mathbf{X}_{11}^*, \mathbf{U}_{11}^*, \mathbf{X}_{12}^*, \mathbf{U}_{12}^*) \sim p^*(\mathbf{x}_{11}^*, \mathbf{u}_{11}^*)p^*(\mathbf{x}_{12}^*, \mathbf{u}_{12}^*).$$

as well as

$$\mathbf{X}_{11} := \frac{\mathbf{X}_{11}^* + \mathbf{X}_{12}^*}{\sqrt{2}}, \qquad \mathbf{X}_{12} := \frac{\mathbf{X}_{11}^* - \mathbf{X}_{12}^*}{\sqrt{2}}$$

and $\mathbf{U}_1 := (\mathbf{U}_{11}^*, \mathbf{U}_{12}^*)$. For $i = 1, 2$ we write

$$\begin{aligned}
\mathbf{Y}_{1i} &:= \mathbf{X}_{1i} + \mathbf{Z}_{1i} \\
\mathbf{Y}_{2i} &:= \mathbf{X}_{1i} + \mathbf{Z}_{1i} + \mathbf{Z}_{2i} \\
\mathbf{Y}_{3i} &:= \mathbf{X}_{1i} + \mathbf{Z}_{1i} + \mathbf{Z}_{2i} + \mathbf{X}_{2i}
\end{aligned}$$

where $(\mathbf{X}_{2i}, \mathbf{Z}_{1i}, \mathbf{Z}_{2i})$ are identically distributed with $(\mathbf{X}_2, \mathbf{Z}_1, \mathbf{Z}_2)$. We have

$$\begin{aligned}
2v &= \mathrm{E}_{\mathbf{U}_{11}^*}[\phi(p(\mathbf{x}_{11}^*|\mathbf{U}_{11}^*))] + \mathrm{E}_{\mathbf{U}_{12}^*}[\phi(p(\mathbf{x}_{12}^*|\mathbf{U}_{12}^*))] \\
&= \mathrm{E}_{\mathbf{U}_1}[\phi(p(\mathbf{x}_{11}^*|\mathbf{U}_1)) + \phi(p(\mathbf{x}_{12}^*|\mathbf{U}_1))] \\
&\overset{(a)}{=} (\beta - 1)h(\mathbf{Y}_{31}, \mathbf{Y}_{32}|\mathbf{U}_1) \\
&\quad + h(\mathbf{Y}_{11}, \mathbf{Y}_{12}|\mathbf{U}_1) - \beta h(\mathbf{Y}_{21}, \mathbf{Y}_{22}|\mathbf{U}_1) \\
&= (\beta - 1)[h(\mathbf{Y}_{31}|\mathbf{Y}_{32}, \mathbf{U}_1) + h(\mathbf{Y}_{32}|\mathbf{Y}_{11}, \mathbf{U}_1) \\
&\qquad + I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1)] \\
&\quad + [h(\mathbf{Y}_{11}|\mathbf{Y}_{32}, \mathbf{U}_1) + h(\mathbf{Y}_{12}|\mathbf{Y}_{11}, \mathbf{U}_1) \\
&\qquad + I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1)] \\
&\quad - \beta[h(\mathbf{Y}_{21}|\mathbf{Y}_{32}, \mathbf{U}_1) + h(\mathbf{Y}_{22}|\mathbf{Y}_{11}, \mathbf{U}_1) \\
&\qquad + I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{U}_1) \\
&\qquad - I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{U}_1)] \\
&= \mathrm{E}[\phi(p(\mathbf{x}_{11}|\mathbf{Y}_{32}, \mathbf{U}_1))] + \mathrm{E}[\phi(p(\mathbf{x}_{22}|\mathbf{Y}_{11}, \mathbf{U}_2))] \\
&\quad + \beta[I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) \\
&\qquad - I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{U}_1)] \\
&\overset{(b)}{\leq} 2v - \beta I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1)
\end{aligned}$$

where (a) can be shown by the rotation-invariance of entropy, and (b) follows from

$$\begin{aligned}
&I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) \\
&\quad - I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{U}_1) \\
&\overset{(c)}{=} I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) \\
&\quad - I(\mathbf{Y}_{11}; \mathbf{Y}_{22}, \mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}, \mathbf{Y}_{32}|\mathbf{U}_1) \\
&= -I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{32}, \mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{Y}_{32}, \mathbf{U}_1) \\
&\overset{(d)}{=} -I(\mathbf{Y}_{11}, \mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{Y}_{32}, \mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{Y}_{32}, \mathbf{U}_1) \\
&= -I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1)
\end{aligned}$$

where (c) holds since $\mathbf{Y}_{32} \to (\mathbf{Y}_{22}, \mathbf{U}_1) \to \mathbf{Y}_{11}$ and $\mathbf{Y}_{32} \to (\mathbf{Y}_{22}, \mathbf{U}_1) \to \mathbf{Y}_{21}$ form Markov chains, and (d) holds since $\mathbf{Y}_{21} \to (\mathbf{Y}_{11}, \mathbf{Y}_{32}, \mathbf{U}_1) \to \mathbf{Y}_{22}$ forms a Markov chain. Hence we have $I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1) = 0$ and so

$$\mathbf{Y}_{11} \to (\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1) \to \mathbf{Y}_{22}$$

forms a Markov chain. Since we also have the Markov chain

$$\mathbf{Y}_{21} \to (\mathbf{Y}_{11}, \mathbf{Y}_{32}, \mathbf{U}_1) \to \mathbf{Y}_{22}$$

by Lemma 1 we obtain a Markov chain

$$(\mathbf{Y}_{11}, \mathbf{Y}_{21}) \to (\mathbf{Y}_{32}, \mathbf{U}_1) \to \mathbf{Y}_{22}.$$

Again we also have the Markov chain

$$(\mathbf{Y}_{11}, \mathbf{Y}_{21}) \to (\mathbf{Y}_{22}, \mathbf{U}_1) \to \mathbf{Y}_{32}$$

and hence by Lemma 1 we obtain a Markov chain

$$(\mathbf{Y}_{11}, \mathbf{Y}_{21}) \to \mathbf{U}_1 \to (\mathbf{Y}_{22}, \mathbf{Y}_{32}).$$

Now Lemma 2 implies that

$$\mathbf{X}_{11} \to \mathbf{U}_1 \to \mathbf{X}_{12}$$

forms a Markov chain, which means that for any $\mathbf{u}_{11}^*, \mathbf{u}_{12}^*$ we have

$$(\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*=\mathbf{u}_{11}^*} + \mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*=\mathbf{u}_{12}^*}) \perp (\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*=\mathbf{u}_{11}^*} - \mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*=\mathbf{u}_{12}^*})$$

which, by Lemma 3, implies that $p(\mathbf{x}_{11}^*|\mathbf{u}_{11}^*)$ and $p(\mathbf{x}_{12}^*|\mathbf{u}_{12}^*)$ are Gaussian distributions having the same covariance matrix. Thus we can conclude that the maximizing distribution $(\mathbf{X}_1, \mathbf{U}_1) \sim p^*(\mathbf{x}_1, \mathbf{u}_1)$ must satisfy

$$\mathbf{X}_1|_{\mathbf{U}_1=\mathbf{u}_1} \sim \mathcal{N}(\mu_{\mathbf{u}_1}, \hat{K}_1)$$

for some $\mu_{\mathbf{u}_1} \in \mathbb{R}^k$ and $\hat{K}_1 \succeq 0$. Finally $\mu_{\mathbf{u}_1} = 0$ since $p^*(\mathbf{x}_1|\mathbf{u}_1)$ is zero-mean. □

III. Relation to Gaussian Z-Interference Channel

The *Gaussian Z-Interference Channel* (GZIC) is a two-user one-sided interference channel defined by

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

where $a > 0$, $Z_1, Z_2 \sim \mathcal{N}(0, 1)$ and X_i, Y_i, Z_i $(i = 1, 2)$ are real random variables. Under power constraints on the inputs X_1, X_2, the capacity of this channel in the case of *strong interference* (i.e. $a \geq 1$) is known [15], while that for $0 < a < 1$ remains open. In this paper we shall consider an equivalent formulation of the channel

$$\begin{aligned}
\mathbf{Y}_1 &= \mathbf{X}_1 + \mathbf{Z}_1 \\
\mathbf{Y}_2 &= \mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2
\end{aligned}$$

where $0 < a < 1$, $\mathbf{Z}_1 \sim \mathcal{N}(0, I)$, $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2 I)$ (where $N_2 := \frac{1}{a^2} - 1$) and $\mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i$ $(i = 1, 2)$ are random variables in $\mathbb{R}^k$ $(k \geq 1)$, under the power constraints

$$\mathrm{E}[\|\mathbf{X}_1\|^2] \leq kP_1 \quad \text{and} \quad \mathrm{E}[\|\mathbf{X}_2\|^2] \leq kP_2$$

where $P_1, P_2 \geq 0$.

Determining the capacity region of the GZIC has been a fundamental yet open problem in network information theory. In this section we propose the following conjecture concerning Gaussian optimality of a functional, which, if true, would imply the capacity region of the GZIC:

Conjecture 1. *For $\beta \geq 1$, $N_2 \geq 0$ and $\Sigma_1, A_2 \succeq 0$, the maximum*

$$\max_{\substack{p(\mathbf{x}_1)p(\mathbf{x}_2) \\ \mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T] \preceq A_2}} \Big[(\beta-1)h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2)+h(\mathbf{X}_1+\mathbf{Z}_1) - \beta h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) - \mathrm{tr}(\Sigma_1 \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T])\Big]$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0,I)$, $\mathbf{Z}_2 \sim \mathcal{N}(0,N_2 I)$ and $\mathbf{X}_i, \mathbf{Z}_i$ $(i=1,2)$ are random variables in $\mathbb{R}^k$ $(k \geq 1)$, is attained by Gaussian $\mathbf{X}_1$ and $\mathbf{X}_2$.

For a general two-user interference channel the Han–Kobayashi (HK) region [16] is the best achievable region known, which, however, has been shown [17] to be sub-optimal for some discrete channels. In particular the HK region is strictly improved by multi-letter extensions. Yet some of the authors have recently shown [18] that for the Gaussian interference channel, and in particular the GZIC, the multi-letter extension of the HK scheme with inputs restricted to be Gaussian does not improve on the scalar one. This result motivates a natural question: whether the k-letter HK region is the same as the k-letter HK region *with Gaussian inputs*, i.e. whether $\mathcal{R}_{\text{HK}}^{(k)}(P_1,P_2) = \mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)$. A positive answer to this question would imply that the single-letter HK region with Gaussian inputs $\mathcal{R}_{\text{HK-GS}}^{(1)}(P_1,P_2)$ is the capacity.

The set of inequalities characterizing the HK region of GZIC simplifies (cf. [19]) to

$$kR_1 \leq h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{Q}) - h(\mathbf{Z}_1) \tag{3}$$

$$kR_2 \leq h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q}) - h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q}) \tag{4}$$

$$k(R_1+R_2) \leq h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{Q}) - h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q}) + h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{U}_1,\mathbf{Q}) - h(\mathbf{Z}_1) \tag{5}$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0,I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0,N_2 I)$. Now we define the k-letter HK region $\mathcal{R}_{\text{HK}}^{(k)}(P_1,P_2)$ as well as that with Gaussian inputs $\mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)$, where GS stands for Gaussian signaling, with power control, as follows.

Definition 1. Let $\mathcal{R}_{\text{HK}}^{(k)}(P_1,P_2)$ be the set of $(R_1,R_2) \in \mathbb{R}^2_{\geq 0}$ satisfying the inequalities (3), (4), (5) for some $p(\mathbf{q})p(\mathbf{u}_1,\mathbf{x}_1|\mathbf{q})p(\mathbf{x}_2|\mathbf{q})$ with $\mathrm{E}[\|\mathbf{X}_i\|^2] \leq kP_i$ $(i=1,2)$.

Definition 2. Let $\mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)$ be the set of $(R_1,R_2) \in \mathbb{R}^2_{\geq 0}$ satisfying the inequalities (3), (4), (5) for some $p(\mathbf{q})p(\mathbf{u}_1,\mathbf{x}_1|\mathbf{q})p(\mathbf{x}_2|\mathbf{q})$ with $\mathrm{E}[\|\mathbf{X}_i\|^2] \leq kP_i$ $(i=1,2)$ and $\mathbf{U}_1$, $\mathbf{X}_1 - \mathbf{U}_1$, $\mathbf{X}_2$ being independent zero-mean Gaussians conditioned on $\mathbf{Q}$.

It is easy to see (with an argument similar to [20]) that for $\beta \geq 1$ and $Q_1, Q_2 \geq 0$ we have

$$\max_{\mathcal{R}_{\text{HK}}^{(k)}(Q_1,Q_2)} k(R_1+\beta R_2) = \mathcal{C}_{Q_1,Q_2}\Big[\max_{\substack{p(\mathbf{x}_1)p(\mathbf{x}_2) \\ \mathrm{E}[\|\mathbf{X}_1\|^2]\leq kQ_1 \\ \mathrm{E}[\|\mathbf{X}_2\|^2]\leq kQ_2}} f_\beta(\mathbf{X}_1,\mathbf{X}_2)\Big] \tag{6}$$

$$\max_{\mathcal{R}_{\text{HK-GS}}^{(k)}(Q_1,Q_2)} k(R_1+\beta R_2) = \mathcal{C}_{Q_1,Q_2}\Big[\max_{\substack{K_1,K_2\succeq 0 \\ \mathrm{tr}(K_1)\leq kQ_1 \\ \mathrm{tr}(K_2)\leq kQ_2}} f_{\beta,\text{GS}}(K_1,K_2)\Big] \tag{7}$$

where

$$f_\beta(\mathbf{X}_1,\mathbf{X}_2) := h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) - h(\mathbf{Z}_1) + \mathcal{C}_{\mathbf{X}_1}\big[\psi(\mathbf{X}_1,\mathbf{X}_2)\big] \tag{8}$$

$$f_{\beta,\text{GS}}(K_1,K_2) := \frac{1}{2}\log|K_2+K_1+I+N_2 I| + \max_{\substack{\hat{K}_1\succeq 0 \\ \hat{K}_1 \preceq K_1}} \psi_{\text{G}}(\hat{K}_1,K_2) \tag{9}$$

with ψ, ψ_{G} defined by (2), (1) respectively.

Remark 3. One of the main difficulties in making a Gaussian extremality conjecture directly for the expression in (8) is that previous work has shown that Gaussian signaling with non-trivial Q (or power control) can improve on Gaussian signaling with a constant Q. Hence, the main conjecture (Conjecture 1) in this paper is obtained by utilizing a carefully constructed dual functional.

As the reader will see the main difficulty in proving $\mathcal{R}_{\text{HK}}^{(k)}(P_1,P_2) = \mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)$ is to establish the upper bound

$$\mathcal{C}_{\mathbf{X}_1}\big[\psi(\mathbf{X}_1,\mathbf{X}_2)\big] \leq \max_{\substack{\hat{K}_1\succeq 0 \\ \hat{K}_1 \preceq K_1}} \psi_{\text{G}}(\hat{K}_1,K_2)$$

with $K_i = \mathrm{E}[\mathbf{X}_i\mathbf{X}_i^T]$ $(i=1,2)$. While Conjecture 1 implies

$$\mathcal{C}_{\mathbf{X}_1}\big[\psi(\mathbf{X}_1,\mathbf{X}_2)\big] \leq \mathcal{C}_{K_1}\big[\psi_{\text{G}}(K_1,K_2)\big]$$

as one can see in the proof of Proposition 3, there is still a missing link as it is in general not true for all functionals ϕ that $\mathcal{C}_{K_1}[\phi(K_1)] = \max_{0\preceq \hat{K}_1 \preceq K_1} \phi(\hat{K}_1)$. However Theorem 1 says that $K_1 \mapsto \psi_{\text{G}}(K_1,K_2)$ has such property, constituting the key step towards Proposition 3.

Proposition 2. *Let $k \geq 1$. The following are equivalent:*

(i) *For any $P_1, P_2 \geq 0$ it holds that*

$$\mathcal{R}_{HK}^{(k)}(P_1,P_2) = \mathcal{R}_{HK\text{-}GS}^{(k)}(P_1,P_2).$$

(ii) *For any $\beta \geq 1$ and $\alpha_1, \alpha_2 \geq 0$ it holds that*

$$\sup_{p(\mathbf{x}_1)p(\mathbf{x}_2)} \big[f_\beta(\mathbf{X}_1,\mathbf{X}_2) - \alpha_1 \mathrm{E}[\|\mathbf{X}_1\|^2] - \alpha_2 \mathrm{E}[\|\mathbf{X}_2\|^2]\big] \leq \sup_{K_1,K_2\succeq 0} \big[f_{\beta,GS}(K_1,K_2) - \alpha_1 \mathrm{tr}(K_1) - \alpha_2 \mathrm{tr}(K_2)\big]$$

where $\mathbf{X}_1, \mathbf{X}_2$ are in $\mathbb{R}^k$, K_1, K_2 are $k \times k$ matrices, and f_β, $f_{\beta,GS}$ are defined by (8), (9) respectively.

Proof. By (6) one has that for $\beta \geq 1$ and $P_1, P_2 \geq 0$,

$$\begin{aligned}
&\max_{\mathcal{R}_{\text{HK}}^{(k)}(P_1,P_2)} k(R_1+\beta R_2) \\
&= \inf_{\alpha_1,\alpha_2\in\mathbb{R}} \Big[\sup_{\substack{Q_1,Q_2\geq 0,\, p(\mathbf{x}_1)p(\mathbf{x}_2) \\ \mathrm{E}[\|\mathbf{X}_1\|^2]\leq kQ_1 \\ \mathrm{E}[\|\mathbf{X}_2\|^2]\leq kQ_2}} \big[f_\beta(\mathbf{X}_1,\mathbf{X}_2) \\
&\qquad - \alpha_1 kQ_1 - \alpha_2 kQ_2 \big] + \alpha_1 kP_1 + \alpha_2 kP_2 \Big] \\
&= \inf_{\alpha_1,\alpha_2\geq 0} \Big[\sup_{p(\mathbf{x}_1)p(\mathbf{x}_2)} \big[f_\beta(\mathbf{X}_1,\mathbf{X}_2) \\
&\qquad - \alpha_1 \mathrm{E}[\|\mathbf{X}_1\|^2] - \alpha_2 \mathrm{E}[\|\mathbf{X}_2\|^2] \big] + \alpha_1 kP_1 + \alpha_2 kP_2 \Big]
\end{aligned}$$

and similarly by (7),

$$\begin{aligned}
&\max_{\mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)} k(R_1+\beta R_2) \\
&= \inf_{\alpha_1,\alpha_2\geq 0} \Big[\sup_{K_1,K_2\succeq 0} \big[f_{\beta,\text{GS}}(K_1,K_2) \\
&\qquad - \alpha_1 \operatorname{tr}(K_1) - \alpha_2 \operatorname{tr}(K_2) \big] + \alpha_1 kP_1 + \alpha_2 kP_2 \Big].
\end{aligned}$$

Note that this already gives that (ii) implies (i). Now assuming (i) we have

$$\begin{aligned}
&\sup_{K_1,K_2\succeq 0} \big[f_{\beta,\text{GS}}(K_1,K_2) - \alpha_1 \operatorname{tr}(K_1) - \alpha_2 \operatorname{tr}(K_2) \big] \\
&\geq \max_{\mathcal{R}_{\text{HK-GS}}^{(k)}(P_1,P_2)} k(R_1+\beta R_2) - \alpha_1 kP_1 - \alpha_2 kP_2 \\
&\geq \max_{\substack{p(\mathbf{x}_1)p(\mathbf{x}_2) \\ \mathrm{E}[\|\mathbf{X}_1\|^2]\leq kP_1 \\ \mathrm{E}[\|\mathbf{X}_2\|^2]\leq kP_2}} f_\beta(\mathbf{X}_1,\mathbf{X}_2) - \alpha_1 kP_1 - \alpha_2 kP_2
\end{aligned}$$

for any $\beta \geq 1$, $P_1, P_2 \geq 0$ and $\alpha_1, \alpha_2 \geq 0$. Finally, taking supremum over $P_1, P_2 \geq 0$ on the last step gives (ii). $\square$

Proposition 3. *Conjecture 1 implies that*

$$\mathcal{R}_{\textit{HK}}^{(k)}(P_1,P_2) = \mathcal{R}_{\textit{HK-GS}}^{(k)}(P_1,P_2).$$

for any $k \geq 1$ and $P_1, P_2 \geq 0$.

Proof. We shall prove the proposition by showing that Conjecture 1 implies (ii) of Proposition 2.

Conjecture 1 implies that for any $\Sigma_1, A_2 \succeq 0$ we have

$$\begin{aligned}
&\sup_{\substack{p(\mathbf{x}_1)p(\mathbf{x}_2) \\ \mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T]\preceq A_2}} \Big[\psi(\mathbf{X}_1,\mathbf{X}_2) - \operatorname{tr}(\Sigma_1 \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]) \Big] \\
&= \sup_{\substack{K_1,K_2\succeq 0,\, \mu_1,\mu_2\in\mathbb{R}^k \\ K_2+\mu_2\mu_2^T\preceq A_2}} \Big[\psi_{\text{G}}(K_1,K_2) - \operatorname{tr}(\Sigma_1 K_1) - \mu_1^T\Sigma_1\mu_1 \Big] \\
&= \sup_{\substack{K_1,K_2\succeq 0 \\ K_2\preceq A_2}} \Big[\psi_{\text{G}}(K_1,K_2) - \operatorname{tr}(\Sigma_1 K_1) \Big] \\
&= \sup_{K_1\succeq 0} \Big[\psi_{\text{G}}(K_1,A_2) - \operatorname{tr}(\Sigma_1 K_1) \Big] \qquad (10)
\end{aligned}$$

where the last equality is a consequence of the monotonicity of $\psi_{\text{G}}(K_1,K_2)$ in K_2. Then for any $p(\mathbf{x}_1)p(\mathbf{x}_2)$ it holds that

$$\begin{aligned}
&\mathcal{C}_{\mathbf{X}_1}\big[\psi(\mathbf{X}_1,\mathbf{X}_2)\big] \\
&\overset{\text{(a)}}{\leq} \inf_{\Sigma_1\succeq 0} \Big[\sup_{p(\hat{\mathbf{x}}_1)} \big[\psi(\hat{\mathbf{X}}_1,\mathbf{X}_2) - \mathrm{E}[\hat{\mathbf{X}}_1^T\Sigma_1\hat{\mathbf{X}}_1] \big] + \mathrm{E}[\mathbf{X}_1^T\Sigma_1\mathbf{X}_1] \Big] \\
&\leq \inf_{\Sigma_1\succeq 0} \Big[\sup_{\substack{p(\hat{\mathbf{x}}_1)p(\hat{\mathbf{x}}_2) \\ \mathrm{E}[\hat{\mathbf{X}}_2\hat{\mathbf{X}}_2^T]\preceq \mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T]}} \big[\psi(\hat{\mathbf{X}}_1,\hat{\mathbf{X}}_2) - \operatorname{tr}(\Sigma_1 \mathrm{E}[\hat{\mathbf{X}}_1\hat{\mathbf{X}}_1^T]) \big] \\
&\qquad + \operatorname{tr}(\Sigma_1 \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]) \Big] \\
&\overset{\text{(b)}}{=} \inf_{\Sigma_1\succeq 0} \Big[\sup_{K_1\succeq 0} \Big[\psi_{\text{G}}(K_1,\mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T]) - \operatorname{tr}(\Sigma_1 K_1) \Big] \\
&\qquad + \operatorname{tr}(\Sigma_1 \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]) \Big] \\
&\overset{\text{(c)}}{=} \inf_{\substack{\Sigma_1 \\ \Sigma_1=\Sigma_1^T}} \Big[\sup_{K_1\succeq 0} \Big[\psi_{\text{G}}(K_1,\mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T]) - \operatorname{tr}(\Sigma_1 K_1) \Big] \\
&\qquad + \operatorname{tr}(\Sigma_1 \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]) \Big] \\
&\overset{\text{(d)}}{=} \mathcal{C}_{K_1}\big[\psi_{\text{G}}(K_1,\mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T])\big]\Big|_{K_1=\mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T]}
\end{aligned}$$

where (a) holds since the right hand side is a concave functional in $p(\mathbf{x})$ that upper bounds $\psi(\mathbf{X}_1,\mathbf{X}_2)$, (b) follows from (10), (c) holds since the inner supremum equals $+\infty$ for any symmetric Σ_1 with $\Sigma_1 \not\succeq 0$, and (d) is the dual characterization of upper concave envelope. Now note also that for any $p(\mathbf{x}_1)p(\mathbf{x}_2)$,

$$\begin{aligned}
&h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2) - h(\mathbf{Z}_1) \\
&\qquad \leq \frac{1}{2}\log|\mathrm{E}[\mathbf{X}_2\mathbf{X}_2^T] + \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] + I + N_2 I|
\end{aligned}$$

and hence for any $\beta \geq 1$ and $\alpha_1, \alpha_2 \geq 0$ we have

$$\begin{aligned}
&\sup_{p(\mathbf{x}_1)p(\mathbf{x}_2)} \big[f_\beta(\mathbf{X}_1,\mathbf{X}_2) - \alpha_1 \mathrm{E}[\|\mathbf{X}_1\|^2] - \alpha_2 \mathrm{E}[\|\mathbf{X}_2\|^2] \big] \\
&\leq \sup_{K_1,K_2\succeq 0} \Big[\frac{1}{2}\log|K_2+K_1+I+N_2 I| + \mathcal{C}_{K_1}\big[\psi_{\text{G}}(K_1,K_2)\big] \\
&\qquad - \alpha_1 \operatorname{tr}(K_1) - \alpha_2 \operatorname{tr}(K_2) \Big] \\
&\overset{\text{(e)}}{=} \sup_{K_1,K_2\succeq 0} \big[f_{\beta,\text{GS}}(K_1,K_2) - \alpha_1 \operatorname{tr}(K_1) - \alpha_2 \operatorname{tr}(K_2) \big]
\end{aligned}$$

as required, where (e) follows from (9) and Theorem 1. $\square$

IV. Conclusion

In this paper we use information-theoretic arguments to establish some properties of the optimizers of some matrix functionals. We then propose a Gaussian extremality conjecture using the dual formulation of the upper concave envelope that, if proved, would establish the capacity region of the Gaussian Z-interference channel. The properties established in the first part were crucially used to complete the link between the conjecture to the capacity region for the Gaussian Z-interference channel.

REFERENCES

[1] C. E. Shannon, “A mathermatical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October, 1948.

[2] E. Lieb, “Proof of an entropy conjecture of wehrl,” *Communications in Mathematical Physics*, vol. 62, August 1978.

[3] H. J. Brascamp and E. H. Lieb, “Best constants in young’s inequality, its converse, and its generalization to more than three functions,” *Advances in Mathematics*, vol. 20, no. 2, pp. 151 – 173, 1976. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0001870876901845

[4] 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. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019995859903481

[5] O. Rioul, “Optimal transportation to the entropy-power inequality,” in *2017 Information Theory and Applications Workshop (ITA)*, Feb 2017, pp. 1–5.

[6] V. Anantharam, V. Jog, and C. Nair, “Unifying the brascamp-lieb inequality and the entropy power inequality,” *CoRR*, vol. abs/1901.06619, 2019. [Online]. Available: http://arxiv.org/abs/1901.06619

[7] 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, April 2014.

[8] V. P. Skitovic, “Linear combinations of independent random variables and the normal distribution law,” *Select Transl Math Stat Probab*, vol. 2, pp. 211–228, 1962.

[9] G. Darmois, “Analyse générale des liaisons stochastiques: etude particulière de l’analyse factorielle linéaire,” *Revue de l’Institut International de Statistique / Review of the International Statistical Institute*, vol. 21, no. 1/2, pp. pp. 2–8, 1953. [Online]. Available: http://www.jstor.org/stable/1401511

[10] S. Sra, N. K. Vishnoi, and O. Yildiz, “On Geodesically Convex Formulations for the Brascamp-Lieb Constant,” in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), E. Blais, K. Jansen, J. D. P. Rolim, and D. Steurer, Eds., vol. 116. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, pp. 25:1–25:15. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2018/9429

[11] C. Nair, “Upper concave envelopes and auxiliary random variables,” *International Journal of Advances in Engineering Sciences and Applied Mathematics*, vol. 5, no. 1, pp. 12–20, 2013. [Online]. Available: http://dx.doi.org/10.1007/s12572-013-0081-7

[12] A. Dembo, T. M. Cover, and J. A. Thomas, “Information theoretic inequalities,” *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1501–1518, Nov 1991.

[13] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, FL, USA: Academic Press, Inc., 1982.

[14] S. G. Ghurye and I. Olkin, “A characterization of the multivariate normal distribution,” *The Annals of Mathematical Statistics*, vol. 33, no. 2, pp. pp. 533–541, 1962. [Online]. Available: http://www.jstor.org/stable/2237532

[15] M. H. M. Costa, “On the Gaussian interference channel,” *Information Theory, IEEE Transactions on*, vol. 31, no. 5, pp. 607–615, Sep 1985.

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

[17] C. Nair, L. Xia, and M. Yazdanpanah, “Sub-optimality of the Han–Kobayashi Achievable Region for Interference Channels,” *ArXiv e-prints*, Feb. 2015.

[18] C. Nair and D. Ng, “Invariance of the han–kobayashi region with respect to temporally-correlated gaussian inputs,” *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1372–1374, March 2019.

[19] M. Costa, C. Nair, and D. Ng, “On the Gaussian Z-interference channel,” *Information Theory and Applications Workshop*, 2017.

[20] M. H. M. Costa and C. Nair, “Gaussian z-interference channel: around the corner,” *Information Theory and Applications Workshop (ITA)*, 2016.