On the Scalar Gaussian Interference Channel

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

***Abstract*—In this paper we show that colored Gaussian inputs to any k-letter extension of the standard scalar Gaussian interference channel do not improve the 1-letter region with Gaussian signaling. Further, we conjecture an inequality, which if true, would establish the capacity of the scalar Gaussian Z-interference channel.**

***Index Terms*—Interference channel, Multi-letter Gaussians**

I. Introduction

Determining a computable characterization of the capacity region of a scalar Gaussian interference channel is a central open question in network information theory. In particular, it is not known whether the Han–Kobayashi region [1] with Gaussian auxiliaries (and power control) yields the capacity region or not. Recently, it was shown [2] that the multi-letter extension of the Han–Kobayashi region for some discrete memoryless interference channels strictly improves on the single-letter region, thus demonstrating the sub-optimality of the Han–Kobayashi scheme.

Motivated by this result, it is natural to ask the same question for the Gaussian interference channel: do the multi-letter extensions of the Han–Kobayashi region with Gaussian auxiliaries (and power control) improve on the single-letter region. In this paper, we answer this question in the negative. In the second part we conjecture an optimality result that would imply that the Han–Kobayashi achievable region would match the capacity region for the Gaussian Z-interference channel.

A. Preliminaries

A scalar Gaussian interference channel is defined by

$$
\begin{aligned}
Y_1 &= X_1 + bX_2 + Z_1 \\
Y_2 &= X_2 + aX_1 + Z_2
\end{aligned}
$$

where $Z_1, Z_2 \sim \mathcal{N}(0,1)$ are independent unit-power Gaussians. We assume power constraints P_1 and P_2 for the inputs X_1 and X_2, respectively. This channel setting has been actively studied in the literature since mid 70s, so a complete literature survey is beyond the scope of this paper. In the next paragraph we summarize some known results.

The capacity region has been established for the case $a, b \geq 1$ [3]. The capacity region has two pareto-optimal points, also called "corner" points, of the form: (C_1, R_2^*) and (R_1^*, C_2) where $C_1 = \frac{1}{2}\log(1+P_1)$ and $C_2 = \frac{1}{2}\log(1+P_2)$ denote the interference-free point-to-point capacities to the two receivers. The above corner points have been determined, see [3]–[5], for all ranges of parameters. Additionally, the Pareto-optimal point that maximizes the rate sum $R_1 + R_2$ under the condition: $a(1+b^2P_2) + b(1+a^2P_1) \leq 1$ has been established independently in [6]–[8]. The result in [9] establishes that the Hausdorff distance (under L^1-norm) between true capacity region and the Han–Kobayashi region is at most 1, for all ranges of parameters.

There have been as yet unsuccessful attempts to improve on the Han–Kobayashi rate region using ideas such as perturbations using Hermite polynomials [10], as well as correlated coding schemes [11].

There has been some instances in network information theory, including work by the authors, where multi-letter Gaussian schemes have been shown to match the single-letter scheme, such as [12]–[14]. This work is a natural extension of such results; however the optimization problem that occurs in this instance has non-trivial local maximizers and yet one can obtain the global maximizers using some structural results.

It is rather immediate to see that the Han–Kobayashi inner bound (Theorem 6.4 in [15]) for the k-letter extension, when evaluated with Gaussian random variables, reduces to the set of rate pairs (R_1, R_2) that satisfy

$$
R_1 \leq \mathrm{E}_Q\left(\frac{1}{2k}\log\frac{\left|I + K_{U_1}^Q + K_{V_1}^Q + b^2K_{V_2}^Q\right|}{\left|I + b^2K_{V_2}^Q\right|}\right) \tag{1a}
$$

$$
R_2 \leq \mathrm{E}_Q\left(\frac{1}{2k}\log\frac{\left|I + K_{U_2}^Q + K_{V_2}^Q + a^2K_{V_1}^Q\right|}{\left|I + a^2K_{V_1}^Q\right|}\right) \tag{1b}
$$

$$
\begin{aligned}
R_1 + R_2 \leq \mathrm{E}_Q\Bigg(&\frac{1}{2k}\log\frac{\left|I + K_{U_1}^Q + K_{V_1}^Q + b^2K_{U_2}^Q + b^2K_{V_2}^Q\right|}{\left|I + b^2K_{V_2}^Q\right|} \\
&+\frac{1}{2k}\log\frac{\left|I + K_{V_2}^Q + a^2K_{V_1}^Q\right|}{\left|I + a^2K_{V_1}^Q\right|}\Bigg)
\end{aligned} \tag{1c}
$$

$$
\begin{aligned}
R_1 + R_2 \leq \mathrm{E}_Q\Bigg(&\frac{1}{2k}\log\frac{\left|I + K_{U_2}^Q + K_{V_2}^Q + a^2K_{U_1}^Q + a^2K_{V_1}^Q\right|}{\left|I + a^2K_{V_1}^Q\right|} \\
&+\frac{1}{2k}\log\frac{\left|I + K_{V_1}^Q + b^2K_{V_2}^Q\right|}{\left|I + b^2K_{V_2}^Q\right|}\Bigg)
\end{aligned} \tag{1d}
$$

$$
R_1 + R_2 \leq \mathrm{E}_Q\Bigg(\frac{1}{2k}\log\frac{\left|I + K_{V_1}^Q + b^2K_{U_2}^Q + b^2K_{V_2}^Q\right|}{\left|I + b^2K_{V_2}^Q\right|}
$$

$$+\frac{1}{2k}\log\frac{\left|I+K_{V_2}^Q+a^2K_{U_1}^Q+a^2K_{V_1}^Q\right|}{\left|I+a^2K_{V_1}^Q\right|}\Bigg) \tag{1e}$$

$$2R_1+R_2\le \mathrm{E}_Q\Bigg(\frac{1}{2k}\log\frac{\left|I+K_{U_1}^Q+K_{V_1}^Q+b^2K_{U_2}^Q+b^2K_{V_2}^Q\right|}{\left|I+b^2K_{V_2}^Q\right|}$$
$$+\frac{1}{2k}\log\frac{\left|I+K_{V_1}^Q+b^2K_{V_2}^Q\right|}{\left|I+b^2K_{V_2}^Q\right|}$$
$$+\frac{1}{2k}\log\frac{\left|I+K_{V_2}^Q+a^2K_{U_1}^Q+a^2K_{V_1}^Q\right|}{\left|I+a^2K_{V_1}^Q\right|}\Bigg) \tag{1f}$$

$$R_1+2R_2\le \mathrm{E}_Q\Bigg(\frac{1}{2k}\log\frac{\left|I+K_{U_2}^Q+K_{V_2}^Q+a^2K_{U_1}^Q+a^2K_{V_1}^Q\right|}{\left|I+a^2K_{V_1}^Q\right|}$$
$$+\frac{1}{2k}\log\frac{\left|I+K_{V_2}^Q+a^2K_{V_1}^Q\right|}{\left|I+a^2K_{V_1}^Q\right|}$$
$$+\frac{1}{2k}\log\frac{\left|I+K_{V_1}^Q+b^2K_{U_2}^Q+b^2K_{V_2}^Q\right|}{\left|I+b^2K_{V_2}^Q\right|}\Bigg), \tag{1g}$$

for $K_{U_1}^q, K_{V_1}^q, K_{U_2}^q, K_{V_2}^q \in \mathbb{R}^{k\times k}$ being symmetric positive semi-definite matrices satisfying $\mathrm{E}_Q\left(\mathrm{tr}\left(K_{U_1}^Q+K_{V_1}^Q\right)\right)\le kP_1$ and $\mathrm{E}_Q\left(\mathrm{tr}\left(K_{U_2}^Q+K_{V_2}^Q\right)\right)\le kP_2$, and some "time-sharing" variable Q. By a standard application of cardinality-bounding techniques, it suffices to consider $|Q|\le 9$ (not needed in this note). Let $\mathcal{R}_k^{GS}$ denote the above region.

The main result of this note is the following:

Theorem 1. *$\mathcal{R}_k^{GS}=\mathcal{R}_1^{GS}$ for all $k\ge 1$.*

We will prove this theorem in the next section.

II. Main

For a $k\times k$ Hermitian matrix A, let $\lambda_1(A)\le\cdots\le\lambda_k(A)$ denote its eigenvalues. The proof uses a couple of standard technical results that we state at the outset.

Theorem 2 (Fiedler [16])**.** *Let A, B be $k\times k$ Hermitian matrices. Suppose $\lambda_k(A)+\lambda_k(B)\ge 0$. Then*

$$\prod_{i=1}^k(\lambda_i(A)+\lambda_i(B))\le|A+B|\le\prod_{i=1}^k(\lambda_i(A)+\lambda_{k+1-i}(B))$$

Theorem 3 (Courant-Fischer min-max theorem)**.** *Let A be a $k\times k$ Hermitian matrix. Then we have*

$$\lambda_i(A)=\inf_{\substack{V\subseteq\mathbb{R}^k\\ \dim V=i}}\sup_{\substack{x\in V\\ \|x\|=1}}x^TAx=\sup_{\substack{V\subseteq\mathbb{R}^k\\ \dim V=n-i+1}}\inf_{\substack{x\in V\\ \|x\|=1}}x^TAx,$$

where V denotes subspaces of the indicated dimension.

Corollary 1. *Let A, B be $k\times k$ Hermitian matrices with $B\succeq 0$. Then $\lambda_i(A+B)\ge\lambda_i(A)$ for $i=1,\cdots,k$.*

Proof. Theorem 3 and $B\succeq 0$ implies that

$$\begin{aligned}\lambda_i(A+B)&=\inf_{\substack{V\le\mathbb{R}^k\\ \dim V=i}}\sup_{\substack{x\in V\\ \|x\|=1}}x^T(A+B)x\\ &\ge\inf_{\substack{V\le\mathbb{R}^k\\ \dim V=i}}\sup_{\substack{x\in V\\ \|x\|=1}}x^TAx\\ &=\lambda_i(A)\end{aligned}$$

□

Given any collection of symmetric positive semi-definite matrices $K_{U_1}^q, K_{V_1}^q, K_{U_2}^q, K_{V_2}^q\in\mathbb{R}^{k\times k}$, define

$$\begin{aligned}\hat{K}_{V_1}^q&:=\mathrm{diag}\left(\{\lambda_i(K_{V_1}^q)\}\right),\\ \hat{K}_{U_1}^q&:=\mathrm{diag}\left(\{\lambda_i(K_{U_1}^q+K_{V_1}^q)-\lambda_i(K_{V_1}^q)\}\right)\succeq 0,\\ \hat{K}_{V_2}^q&:=\mathrm{diag}\left(\{\lambda_{n+1-i}(K_{V_2}^q)\}\right),\\ \hat{K}_{U_2}^q&:=\mathrm{diag}\left(\{\lambda_{n+1-i}(K_{U_2}^q+K_{V_2}^q)-\lambda_{n+1-i}(K_{V_1}^q)\}\right)\succeq 0.\end{aligned}$$

where $\mathrm{diag}(\{a_i\})$ indicates a diagonal matrix with diagonal entries $a_1,..,a_k$. The positive semi-definiteness of $\hat{K}_{U_1},\hat{K}_{U_2}$ follows from Corollary 1. Note that these are trace preserving operations, i.e. $\mathrm{tr}(\hat{K}_{U_1}^q+\hat{K}_{V_1}^q)=\mathrm{tr}(K_{U_1}^q+K_{V_1}^q)$ and $\mathrm{tr}(\hat{K}_{U_2}^q+\hat{K}_{V_2}^q)=\mathrm{tr}(K_{U_2}^q+K_{V_2}^q)$. Further

$$\left|I+a^2K_{V_1}^q\right|=\left|I+a^2\hat{K}_{V_1}^q\right|=\prod_{i=1}^n\left(1+a^2\lambda_i(K_{V_1}^q)\right), \tag{2a}$$

$$\left|I+b^2K_{V_2}^q\right|=\left|I+b^2\hat{K}_{V_2}^q\right|=\prod_{i=1}^n\left(1+b^2\lambda_i(K_{V_2}^q)\right). \tag{2b}$$

Lemma 1. *For any $c_1, c_2\ge 0$, let $(A_1,\hat{A}_1)=(K_{V_1}^q,\hat{K}_{V_1}^q)$ or $(K_{U_1}^q+K_{V_1}^q,\hat{K}_{U_1}^q+\hat{K}_{V_1}^q)$, and let $(A_2,\hat{A}_2)=(K_{V_2}^q,\hat{K}_{V_2}^q)$ or $(K_{U_2}^q+K_{V_2}^q,\hat{K}_{U_2}^q+\hat{K}_{V_2}^q)$. Then*

$$|I+c_1A_1+c_2A_2|\le\left|I+c_1\hat{A}_1+c_2\hat{A}_2\right|.$$

Proof.

$$\begin{aligned}|I+c_1A_1+c_2A_2|&\le\prod_{i=1}^k(1+c_1\lambda_i(A_1)+c_2\lambda_{n+1-i}(A_2))\\ &=\left|I+c_1\hat{A}_1+c_2\hat{A}_2\right|\end{aligned}$$

where the inequality follows from Theorem 2. □

Proof of Theorem 1: From (2) and Lemma 1, observe that replacing $(K_{U_1}^q,K_{V_1}^q,K_{U_2}^q,K_{V_2}^q)$ by $(\hat{K}_{U_1}^q,\hat{K}_{V_1}^q,\hat{K}_{U_2}^q,\hat{K}_{V_2}^q)$ cannot decrease any of the right-hand-sides of (1). This shows that $\mathcal{R}_k^{GS}$ can be attained by diagonal covariance matrices. Now the inclusion $\mathcal{R}_k^{GS}\subseteq\mathcal{R}_1^{GS}$ is immediate, thus establishing Theorem 1.

III. Gaussian Z-interference channel

A scalar Gaussian Z-interference channel is defined by

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

where $Z_1, Z_2\sim\mathcal{N}(0,1)$ are independent unit-power Gaussians. We assume power constraints P_1 and P_2 for the inputs

X_1 and X_2, respectively. Motivated by the results in the previous section, we study the optimality of the Han–Kobayashi achievable region.

For this channel [3], [4] determined that $R_1 + R_2$ is maximized at the "corner-point" (C_1, R_2^*) where $C_1 = \frac{1}{2}\log(1 + P_1)$ and $R_2^* = \frac{1}{2}\log\left(1 + \frac{P_2}{1+a^2P_1}\right)$. Therefore it suffices to maximize $\lambda R_2 + R_1$ for $\lambda \geq 1$, to compute the capacity region.

For $\lambda \geq 1$ the weighted sum-rate of the k-letter extension of the Han–Kobayashi region for the Z-interference channel is given by

$$\begin{aligned}&\lambda R_2 + R_1\\&= \max_{p(q)p_1(u_1,\mathbf{x}_1|q)p_2(\mathbf{x}_2|q)} I(U_1, \mathbf{X}_2; \mathbf{Y}_2|Q)\\&\quad + I(\mathbf{X}_1; \mathbf{Y}_1|U_1, Q) + (\lambda - 1)I(\mathbf{X}_2; \mathbf{Y}_2|U_1, Q).\end{aligned}$$

Define the function

$$\begin{aligned}&f(Q_1, Q_2)\\&:= \max_{p_1(u_1,\mathbf{x}_1)p_2(\mathbf{x}_2)} I(U_1, \mathbf{X}_2; \mathbf{Y}_2)\\&\quad + I(\mathbf{X}_1; \mathbf{Y}_1|U_1) + (\lambda - 1)I(\mathbf{X}_2; \mathbf{Y}_2|U_1).\end{aligned}$$

where the maximum is over independent $\mathbf{X}_1$ and $\mathbf{X}_2$ with power kQ_1 and kQ_2. It is immediate that

$$\begin{aligned}f(Q_1, Q_2) &= \max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_2) + \mathcal{C}_{\mathbf{X}_1}\left[I(\mathbf{X}_1; \mathbf{Y}_1)\right.\\&\qquad \left.+(\lambda - 1)I(\mathbf{X}_2; \mathbf{Y}_2) - I(\mathbf{X}_1; \mathbf{Y}_2|\mathbf{X}_2)\right]\\&= \max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_2)\\&\quad + \mathcal{C}_{\mathbf{X}_1}\left[(\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z})\right.\\&\quad \left.-\lambda h(a\mathbf{X}_1 + \mathbf{Z})\right]\end{aligned}$$

where, as before, the maximum is over independent $\mathbf{X}_1$ and $\mathbf{X}_2$ with power kQ_1 and kQ_2, and $\mathcal{C}_{\mathbf{X}_1}$ denotes the concave envelope taken over the distributions of $\mathbf{X}_1$.

Then the Han–Kobayashi region with power constraints P_1, P_2 is given by the concave envelope of the function $f(Q_1, Q_2)$ evaluated at the pair (P_1, P_2).

Therefore the key is to compute the function $f(Q_1, Q_2)$. We are going to upper bound $f(Q_1, Q_2)$ as follows:

$$\begin{aligned}f(Q_1, Q_2) &\leq \frac{k}{2}\log(1 + Q_2 + a^2Q_1)\\&\quad + \max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} \mathcal{C}_{\mathbf{X}_1}\left[(\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z})\right.\\&\qquad \left.+ h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\right]\\&= \frac{k}{2}\log(1 + Q_2 + a^2Q_1)\\&\quad + \mathcal{C}_{Q_1}\left[\max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} (\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z})\right.\\&\qquad \left.+ h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\right].\end{aligned}$$

Remark 1. We note the following

- The maximization inside the concave envelope needs to be clarified. Define the function

$$\begin{aligned}g(\hat{Q}_1, Q_2) := \max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} &(\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z})\\&+ h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\end{aligned}$$

  where the maximum is over independent $\mathbf{X}_1$ and $\mathbf{X}_2$ with power $k\hat{Q}_1$ and kQ_2. Then we are computing the concave envelope of $g(\hat{Q}_1, Q_2)$ with respect to the first coordinate $\hat{Q}_1$ at the point Q_1.
- If the maximizers are Gaussians, then the upper bound is achievable.

For $\alpha \geq 0$, let us define the "Fenchel-dual" function

$$\begin{aligned}\hat{g}(\alpha, Q_2) = \max_{p_1(\mathbf{x}_1)p_2(\mathbf{x}_2)} &(\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z})\\&+ h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\\&- \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2)\end{aligned}$$

Now there is no power constraint on $\mathbf{X}_1$. We still require $\mathrm{E}(\mathbf{X}_2^2) \leq Q_2$.

Note that the dual of the dual yields the concave envelope, i.e. the concave envelope of $g(\hat{Q}_1, Q_2)$ at Q_1 is given by

$$\min_{\alpha\geq 0}\left\{\hat{g}(\alpha, Q_2) + \alpha Q_1\right\}.$$

Observation: In summary, if Gaussians yield $\hat{g}(\alpha, Q_2)$ we are done. Taking Lagrange multiplier β for the power constraint on $\mathbf{X}_2$ we arrive at the sufficiency of Conjecture 1 below.

A. A conjecture

Let $\alpha, \beta \geq 0$ be constants.

Conjecture 1. *The maximum of*

$$\begin{aligned}(\lambda - 1)&h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\\&- \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2)\end{aligned}$$

over independent variables $\mathbf{X}_1$ and $\mathbf{X}_2$ taking values in $\mathbb{R}^k$ is attained by Gaussians $\mathbf{X}_1 \sim \mathcal{N}(0, a\mathbf{I}), \mathbf{X}_2 \sim \mathcal{N}(0, b\mathbf{I})$.

Remark 2. The following is worth noting: from the preliminary development in this section, if the conjecture is true then Han–Kobayashi scheme with gaussian signaling achieves the capacity region of the Gaussian Z-interference channel.

Proposition 1. *Conjecture 1 holds if $\beta \geq \frac{\lambda-1}{2}$ or $\alpha \geq \frac{1-a^2}{2}$.*

Proof. The proofs are rather immediate as seen below.
Case 1: $\beta \geq \frac{\lambda-1}{2}$.

Note that

$$\begin{aligned}&(\lambda - 1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z})\\&\quad - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2)\\&= (\lambda - 1)\big(h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z})\big)\\&\quad + h(\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z})\\&\quad - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2)\\&\overset{(a)}{\leq} (\lambda - 1)\big(h(\mathbf{X}_2 + \mathbf{Z}) - h(\mathbf{Z})\big) + h(\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z})\\&\quad - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2)\\&\leq h(\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z}) - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2),\end{aligned}$$

where (a) follows by data-processing. The last inequality follows as $(\lambda - 1)\big(h(\mathbf{X}_2 + \mathbf{Z}) - h(\mathbf{Z})\big) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2)$ is maximized when $\mathrm{E}(\|\mathbf{X}_2\|^2) = 0$. (For a fixed $\mathrm{E}(\|\mathbf{X}_2\|^2)$, the first part is maximized by Gaussian $\mathbf{X}_2$ and now differentiate and note that the maximizing power is 0.)

The final inequality is maximized by Gaussians for any fixed $\mathrm{E}(\|\mathbf{X}_1\|^2)$ as an immediate consequence of entropy-power-inequality (EPI).

Case 2: $\alpha \geq \frac{1-a^2}{2}$. Similar to the previous case note that

$$\begin{aligned}
&(\lambda-1)h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) + h(\mathbf{X}_1 + \mathbf{Z}) - \lambda h(a\mathbf{X}_1 + \mathbf{Z}) \\
&\quad - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2) \\
&\quad = (\lambda-1)\big(h(\mathbf{X}_2 + a\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z})\big) \\
&\quad + h(\mathbf{X}_1 + \mathbf{Z}) - h(a\mathbf{X}_1 + \mathbf{Z}) \\
&\quad - \alpha\,\mathrm{E}(\|\mathbf{X}_1\|^2) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2) \\
&\leq (\lambda-1)\big(h(\mathbf{X}_2 + \mathbf{Z}) - h(\mathbf{Z})\big) - \beta\,\mathrm{E}(\|\mathbf{X}_2\|^2).
\end{aligned}$$

As before $h(\mathbf{X}_1+\mathbf{Z})-h(a\mathbf{X}_1+\mathbf{Z})$ is maximized by Gaussians for any fixed $\mathrm{E}(\|\mathbf{X}_1\|^2)$ and further if $\alpha \geq \frac{1-a^2}{2}$, then $\mathrm{E}(\|\mathbf{X}_1\|^2) = 0$ is the maximizer. Clearly the final inequality is maximized by a Gaussian $\mathbf{X}_2$. The fact that the covariances can be assumed to be multiples of identity matrix is a simple exercise in both cases. □

A natural way to prove the above conjecture is to adopt a variational approach along traditional lines and move to the Gaussian maximizers along the "Stam-path". Numerical simulations indicate that this technique holds promise. Therefore, we present the numerical observation as a conjecture below.

Conjecture 2. *Let X_1, X_2 be independent random variables. Suppose Q_1^*, Q_2^* maximizes*

$$\begin{aligned}
&\frac{\lambda-1}{2}\log(1 + a^2Q_1 + Q_2) + \frac{1}{2}\log(1 + Q_1) \\
&\qquad - \frac{\lambda}{2}\log(1 + a^2Q_1) - \alpha Q_1 - \beta Q_2.
\end{aligned}$$

For $t \in [0,1]$ define

$$\begin{aligned}
f(t) := &(\lambda-1)h(X_{2t} + aX_{1t} + Z) + h(X_{1t} + Z) \\
&- \lambda h(aX_{1t} + Z) - \alpha\,\mathrm{E}(X_{1t}^2) - \beta\,\mathrm{E}(X_{2t}^2)
\end{aligned}$$

where

$$\begin{aligned}
X_{1t} &:= \sqrt{1-t}X_1 + \sqrt{t}\mathcal{N}(0, Q_1^*) \\
X_{2t} &:= \sqrt{1-t}X_2 + \sqrt{t}\mathcal{N}(0, Q_2^*).
\end{aligned}$$

Then $f(t)$ is increasing and concave.

A simple calculation yields that

$$\begin{aligned}
f'(t) = &\frac{1}{2(1-t)}(\lambda-1)(Q_2^* + a^2Q_1^* + 1)I(X_{2t} + aX_{1t} + Z) \\
&+ \frac{1}{2(1-t)}(Q_1^* + 1)I(X_{1t} + Z) \\
&- \frac{1}{2(1-t)}\lambda(a^2Q_1^* + 1)I(aX_{1t} + Z) \\
&- \alpha(Q_1^* - \mathrm{E}\left(X_1^2\right)) - \beta(Q_2^* - \mathrm{E}\left(X_2^2\right)),
\end{aligned}$$

where

$$I(X) := \frac{d}{dt}h(X + \sqrt{2s}\mathcal{N}(0,1))\Big|_{s\downarrow 0^+},$$

denotes the Fisher information.

Conjecture 2 stipulates that $f'(t) \geq 0$. Note that it suffices to show $f'(0) \geq 0$ for any independent X_1, X_2. Reason: if we map $(X_1, X_2) \leftarrow (X_{1t}, X_{2t})$, the value $f'(0)$ becomes $(1-t)f'(t)$.

Hence the first part of Conjecture 2, i.e. that $f(t)$ is increasing, is equivalent to the following conjecture.

Conjecture 3. *Let X_1, X_2 be independent random variables. Suppose Q_1^*, Q_2^* maximizes*

$$\begin{aligned}
&\frac{\lambda-1}{2}\log(1 + a^2Q_1 + Q_2) + \frac{1}{2}\log(1 + Q_1) \\
&\qquad - \frac{\lambda}{2}\log(1 + a^2Q_1) - \alpha Q_1 - \beta Q_2
\end{aligned}$$

Then

$$\begin{aligned}
&(\lambda-1)(Q_2^* + a^2Q_1^* + 1)I(X_2 + aX_1 + Z) \\
&\quad + (Q_1^* + 1)I(X_1 + Z) - \lambda(a^2Q_1^* + 1)I(aX_1 + Z) \\
&\quad - 2\alpha(Q_1^* - \mathrm{E}\left(X_1^2\right)) - 2\beta(Q_2^* - \mathrm{E}\left(X_2^2\right)) \\
&\geq 0
\end{aligned}$$

Remark 3. From existing bounds on Fisher information the above conjecture can be easily established for X_1, X_2 satisfying some power (second-moment) constraints on X_1, X_2.

B. On the doubling trick

Another potential method for establishing Conjecture 1 is to use the so-called "doubling trick" employed in [17]. This may indeed work in this scenario, but it must be noted that the results in [2] imply that there are discrete memoryless Z-interference channels for which the function

$$\mathfrak{C}_{X_1}[(\lambda-1)H(Y_2) + H(Y_1) - \lambda H(Y_2|X_2)]$$

does not satisfy sub-additivity or equivalently the "doubling property". Hence one has to first establish sub-additivity for a sub-class containing the Gaussian Z-interference channels to use this trick. This necessitates making use of the channel structure rather than generic channel-oblivious manipulations (such arguments do exist in literature). It is conceivable that some ideas such as those used in [18] might turn out to be useful, given the particular channel structure, to establish the doubling property.

IV. Discussion

There has been some interest (for instance [11]) in using multivariate Gaussians to improve on the Han–Kobayashi achievable region. However the result in this note says that such an improvement is not possible. The authors in [11] do not consider the effect of power control using Q, and hence their conclusion is not general enough. The need for power control was noted as early as [4], but more recently was a central theme of [19]. A similar result had already been established by the authors and Costa for Z-interference and mixed interference regimes in [14]. The argument in this note is more general and works for all regimes; on the other hand the proof ideas in [14] yield more insight into the single-letter optimizers via water-filling operation.

The result in this note may be viewed as evidence (perhaps) to the optimality of the Han–Kobayashi achievable region for this setting. There is an inherent rotational invariance to the optimizers of Han–Kobayashi expression, and the k-letter Han–Kobayashi region goes to capacity. Hence it is not inconceivable that the result in this note, along with a proof of optimality of Gaussian distributions (along any of the lines outlined here) would settle this long standing open problem.

Acknowldegements

The authors wish to thank Max Costa for stimulating exchanges as well as inspiring the authors to believe in the optimality of Gaussian signaling for the Han–Kobayashi region for the scalar Gaussian interference channel (which is still open). The authors also wish to thank Salman Beigi for his contributions to a similar result (as Theorem 1) in [13], which laid the foundations for subsequent research.

References

[1] T. 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.

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

[3] 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, Nov 1981.

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

[5] Y. Polyanskiy and Y. Wu, "Wasserstein continuity of entropy and outer bounds for interference channels," *CoRR*, vol. abs/1504.04419, 2015. [Online]. Available: http://arxiv.org/abs/1504.04419

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

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

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

[9] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *Information Theory, IEEE Transactions on*, vol. 54, no. 12, pp. 5534–5562, dec. 2008.

[10] E. Abbe and L. Zheng, "A coordinate system for gaussian networks," *Information Theory, IEEE Transactions on*, vol. 58, no. 2, pp. 721–733, Feb 2012.

[11] W. Huleihel and N. Merhav, "Codewords with memory improve achievable rate regions of the memoryless Gaussian interference channel," *CoRR*, vol. abs/1508.05726, 2015. [Online]. Available: http://arxiv.org/abs/1508.05726

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

[13] S. Beigi, S. Liu, C. Nair, and M. Yazdanpanah, "Some results on the scalar Gaussian interference channel," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 2199–2203.

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

[15] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.

[16] 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. [Online]. Available: http://www.jstor.org/stable/2038212

[17] 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.

[18] A. Gamal and M. Costa, "The capacity region of a class of deterministic interference channels (corresp.)," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 343–346, Mar 1982.

[19] M. H. M. Costa, "Noisebergs in Z Gaussian interference channels," *Information Theory and Applications Workshop (ITA)*, 2011.