On the Scalar Gaussian Interference Channel

Chandra Nair, &

David Ng

The Chinese University of Hong Kong

ITA 2018
13 Feb, 2018
Question

Does Han-Kobayashi achievable region with Gaussian signaling exhaust the capacity region of the scalar Gaussian interference channel?
Question
Does Han-Kobayashi achievable region with Gaussian signaling exhaust the capacity region of the scalar Gaussian interference channel?

This talk
Perhaps it may
- We establish some evidence towards this end
- Conjecture an information inequality, which if true, would establish the optimality for the Z-interference channel
(Some) known results about the capacity region

- Determined $a \geq 1, b \geq 1$ (Sato ’79)
- *Corner Points* (Sato ’81, Costa ’85, Sason ’02, Polyanskiy-Wu ’15)
- Maximum rate-sum $a(1 + b^2P_2) + b(1 + a^2P_1) \leq 1$ (3 groups ’09)
- Han–Kobayashi region within 0.5 bits per dimension (Etkin, Tse, Wang ’07)
As a side note

Investigations on this problem have led to

- Costa’s discovery: concavity of entropy power
- Use of HWI to establish converses (Polyanskiy-Wu ’15)
- Use of "genies"
  - To establish converses/bounds (Kramer, Etkin-Tse-Wang, ...)
  - As a tool for proving sub-additivity/tensorization
On the Han–Kobayashi achievable region

Background

- 1981: Han and Kobayashi proposed an achievable region (HK-IB) for memoryless interference channels
- 2015: HK-IB was shown to be strictly sub-optimal for some channels (with: Xia, Yazdanpanah)
  - Result: 2-letter extension of HK-IB outperformed HK-IB
  - Difficulty: Evaluating HK-IB (1-letter and 2-letter)
  - Channels: Clean Z-interference channels
On the Han–Kobayashi achievable region

Background

▶ 1981: Han and Kobayashi proposed an achievable region (HK-IB) for memoryless interference channels
▶ 2015: HK-IB was shown to be strictly sub-optimal for some channels (with: Xia, Yazdanpanah)
  ● Result: 2-letter extension of HK-IB outperformed HK-IB
  ● Difficulty: Evaluating HK-IB (1-letter and 2-letter)
  ● Channels: Clean Z-interference channels

Natural Questions

How about if one restricts to the special case: scalar Gaussian interference channels?

▶ Is HK-IB (with Gaussian signaling) optimal?
▶ Or does $k$-letter extensions (with Gaussian signaling), or in other words do correlated Gaussian input vectors improve the region?

● Remark: There is a paper (2016) that claims such an improvement but it ignores the role of "power control" (which was known to improve on naive region since 1985; see also Costa - ITA 2010)
On the Han–Kobayashi achievable region

Background

- 1981: Han and Kobayashi proposed an achievable region (HK-IB) for memoryless interference channels
- 2015: HK-IB was shown to be strictly sub-optimal for some channels (with: Xia, Yazdanpanah)
  - Result: 2-letter extension of HK-IB outperformed HK-IB
  - Difficulty: Evaluating HK-IB (1-letter and 2-letter)
  - Channels: Clean Z-interference channels

Natural Questions

How about if one restricts to the special case: scalar Gaussian interference channels?

- Is HK-IB (with Gaussian signaling) optimal?
- Or does \(k\)-letter extensions (with Gaussian signaling), or in other words do correlated Gaussian input vectors improve the region?
  - **Remark**: There is a paper (2016) that claims such an improvement but it ignores the role of "power control" (which was known to improve on naive region since 1985; see also Costa - ITA 2010)
- **Main Result**: No improvement in going to correlated Gaussians
  - Cheng and Verdu had such a result for \(\alpha I(X_1^k; Y_1^k) + I(X_2^k; Y_2^k)\) (1993)
  - We had a similar result for Z-interference \((b = 0)\) last year.
H–K IB with Gaussian signaling ($k$-letter)

Non-negative rate pairs $R_1, R_2$ satisfying

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

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

$$R_1 + R_2 \leq \frac{1}{2k} E_Q \left( \log \frac{|I + (K_{U_1}^Q + K_{V_1}^Q) + b^2 (K_{U_2}^Q + K_{V_2}^Q)|}{|I + b^2 K_{V_2}^Q|} + \log \frac{|I + K_{V_2}^Q + a^2 K_{V_1}^Q|}{|I + a^2 K_{V_1}^Q|} \right)$$

$$R_1 + R_2 \leq \frac{1}{2k} E_Q \left( \frac{1}{2k} \log \frac{|I + (K_{U_2}^Q + K_{V_2}^Q) + a^2 (K_{U_1}^Q + K_{V_1}^Q)|}{|I + a^2 K_{V_1}^Q|} + \log \frac{|I + K_{V_2}^Q + b^2 K_{V_2}^Q|}{|I + b^2 K_{V_2}^Q|} \right)$$

$$R_1 + R_2 \leq \frac{1}{2k} E_Q \left( \log \frac{|I + K_{V_1}^Q + b^2 (K_{U_2}^Q + K_{V_2}^Q)|}{|I + b^2 K_{V_2}^Q|} + \log \frac{|I + K_{V_2}^Q + a^2 (K_{U_1}^Q + K_{V_1}^Q)|}{|I + a^2 K_{V_1}^Q|} \right)$$

$$2R_1 + R_2 \leq \frac{1}{2k} E_Q \left( \log \frac{|I + (K_{U_1}^Q + K_{V_1}^Q) + b^2 (K_{U_2}^Q + K_{V_2}^Q)|}{|I + b^2 K_{V_2}^Q|} + \log \frac{|I + K_{V_1}^Q + b^2 K_{V_2}^Q|}{|I + b^2 K_{V_2}^Q|} + \log \frac{|I + K_{V_2}^Q + a^2 (K_{U_1}^Q + K_{V_1}^Q)|}{|I + a^2 K_{V_1}^Q|} \right)$$

$$R_1 + 2R_2 \leq \frac{1}{2k} E_Q \left( \log \frac{|I + (K_{U_2}^Q + K_{V_2}^Q) + a^2 (K_{U_1}^Q + K_{V_1}^Q)|}{|I + a^2 K_{V_1}^Q|} + \log \frac{|I + K_{V_1}^Q + a^2 K_{V_1}^Q|}{|I + a^2 K_{V_1}^Q|} + \log \frac{|I + K_{V_2}^Q + b^2 (K_{U_2}^Q + K_{V_2}^Q)|}{|I + b^2 K_{V_2}^Q|} \right)$$

for some $K_{U_1}^q, K_{V_1}^q, K_{U_2}^q, K_{V_2}^q \geq 0$ satisfying $E_Q \left( \text{tr} \left( K_{U_1}^Q + K_{V_1}^Q \right) \right) \leq kP_1$ and $E_Q \left( \text{tr} \left( K_{U_2}^Q + K_{V_2}^Q \right) \right) \leq kP_2$, and some “time-sharing" variable $Q$. 
Result: $k$-letter region is identical to $1$-letter region

**Note:** Dealing with optimizers of a **non-convex** optimization problem
Result: $k$-letter region is identical to $1$-letter region

**Note**: Dealing with optimizers of a **non-convex** optimization problem

**Proof**: Define

$$
\hat{K}^q_{V_1} := \text{diag}\left(\{\lambda_i(K^q_{V_1})\}\right)
$$

$$
\hat{K}^q_{U_1} := \text{diag}\left(\{\lambda_i(K^q_{U_1} + K^q_{V_1}) - \lambda_i(K^q_{V_1})\}\right)
$$

$$
\hat{K}^q_{V_2} := \text{diag}\left(\{\lambda_{n+1-i}(K^q_{V_2})\}\right)
$$

$$
\hat{K}^q_{U_2} := \text{diag}\left(\{\lambda_{n+1-i}(K^q_{U_2} + K^q_{V_2}) - \lambda_{n+1-i}(K^q_{V_1})\}\right).
$$

where $\lambda_1(A) \leq \cdots \leq \lambda_k(A)$ denote the eigenvalues of a $k \times k$ Hermitian matrix $A$, and diag($\{a_i\}$) indicates a diagonal matrix with diagonal entries $a_1, \ldots, a_k$.

These choices dominate the inequalities term-by-term.
This "observation" and feasibility of these choices relies on two well-known results.

**Difficulty**: Making this guess (came after a few months of failed other approaches)
There were multiple solutions to KKT conditions, for instance.
Two results

**Theorem (Courant-Fischer min-max theorem)**

Let $A$ be a $k \times k$ Hermitian matrix. Then we have

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

where $V$ denotes subspaces of the indicated dimension.

**Corollary**

Let $A, B$ be $k \times k$ Hermitian matrices with $B \succeq 0$. Then $\lambda_i(A + B) \geq \lambda_i(A)$ for $i = 1, \ldots, k$. 
Two results

Theorem (Courant-Fischer min-max theorem)

Let $A$ be a $k \times k$ Hermitian matrix. Then we have

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

where $V$ denotes subspaces of the indicated dimension.

Corollary

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

Theorem (Fiedler ’71)

Let $A, B$ be $k \times k$ Hermitian matrices. Suppose $\lambda_k(A) + \lambda_k(B) \geq 0$. Then

$$
\prod_{i=1}^{k} (\lambda_i(A) + \lambda_i(B)) \leq |A + B| \leq \prod_{i=1}^{k} (\lambda_i(A) + \lambda_{k+1-i}(B)).
$$
What next?

**Obvious**: Do Gaussian inputs optimize HK-IB?

**Observations**

- Timesharing variable $Q$ is a cause of trouble
- Without $Q$, there are $P_1, P_2$ for which non-Gaussian distributions outperform Gaussian distribution
  - Using perturbations based on Hermite Polynomials (Abbe-Zhang 09)
Obvious: Do Gaussian inputs optimize HK-IB?

Observations

- Timesharing variable $Q$ is a cause of trouble
- Without $Q$, there are $P_1, P_2$ for which non-Gaussian distributions outperform Gaussian distribution
  - Using perturbations based on Hermite Polynomials (Abbe-Zhang 09)
- What is $Q$ doing?
  - Answer: $Q$ is used to compute the upper concave envelope of a functional defined on $(P_1, P_2)$
  - Observation: Since the dual of the dual (in the sense of Fenchel) yields the concave envelope, we just need to check that Gaussians optimize the dual functional
Let $\alpha, \beta \geq 0$, and $\lambda \geq 1$ be constants.

**Conjecture (main)**

The maximum of

$$(\lambda - 1)h(X_2 + aX_1 + Z) + h(X_1 + Z) - \lambda h(aX_1 + Z) - \alpha E(\|X_1\|^2) - \beta E(\|X_2\|^2)$$

over independent variables $X_1$ and $X_2$ taking values in $\mathbb{R}^k$ is attained by Gaussians $X_1 \sim \mathcal{N}(0, aI), X_2 \sim \mathcal{N}(0, bI)$.
A conjecture

Let $\alpha, \beta \geq 0$, and $\lambda \geq 1$ be constants.

Conjecture (main)

The maximum of

$$(\lambda - 1)h(X_2 + aX_1 + Z) + h(X_1 + Z) - \lambda h(aX_1 + Z) - \alpha E(\|X_1\|^2) - \beta E(\|X_2\|^2)$$

going over independent variables $X_1$ and $X_2$ taking values in $\mathbb{R}^k$ is attained by Gaussians $X_1 \sim \mathcal{N}(0, aI), X_2 \sim \mathcal{N}(0, bI)$.

Why should one care about this

- If true, this establishes the capacity region of the Gaussian Z-interference channel
- Let $\alpha = 0$. Suppose you show that, $\forall \beta > 0, \exists \lambda^* < \infty$ such that the conjecture is true $\forall \lambda \geq \lambda^*$, then we improve on the outer bound obtained using HWI
- Flavor of a reverse-entropy-power inequality
Easy regimes

The conjecture is true when either

\[ \beta \geq \frac{\lambda - 1}{2} \]
\[ \alpha \geq \frac{1-a^2}{2} \]

**Proof**: Consequence of data-processing and Entropy-Power-Inequality (or doubling trick)
Appears that Stam’s path (O-U semigroup) may work

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

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

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

$$f(t) := (\lambda - 1) h(X_{2t} + a X_{1t} + Z) + h(X_{1t} + Z) - \lambda h(a X_{1t} + Z) - \alpha E(X_{1t}^2) - \beta E(X_{2t}^2)$$

where

$$X_{1t} := \sqrt{1 - t} X_1 + \sqrt{t} N(0, Q_1^*)$$

$$X_{2t} := \sqrt{1 - t} X_2 + \sqrt{t} N(0, Q_2^*).$$

Then $f(t)$ is increasing and concave.
Increasing along the path

Conjecture

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

$$\frac{\lambda - 1}{2} \log(1 + a^2 Q_1 + Q_2) + \frac{1}{2} \log(1 + Q_1) - \frac{\lambda}{2} \log(1 + a^2 Q_1) - \alpha Q_1 - \beta Q_2$$

Then

$$(\lambda - 1)(Q_2^* + a^2 Q_1^* + 1)I(X_2 + aX_1 + Z) + (Q_1^* + 1)I(X_1 + Z) - \lambda(a^2 Q_1^* + 1)I(aX_1 + Z) - 2\alpha(Q_1^* - E(X_1^2)) - 2\beta(Q_2^* - E(X_2^2))$$

$$\geq 0,$$

where $I(X)$ is the Fisher information of $X$. 
Conjecture

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

$$\frac{\lambda - 1}{2} \log(1 + a^2 Q_1 + Q_2) + \frac{1}{2} \log(1 + Q_1) - \frac{\lambda}{2} \log(1 + a^2 Q_1) - \alpha Q_1 - \beta Q_2$$

Then

$$(\lambda - 1)(Q_2^* + a^2 Q_1^* + 1)I(X_2 + aX_1 + Z) + (Q_1^* + 1)I(X_1 + Z)$$

$$- \lambda(a^2 Q_1^* + 1)I(aX_1 + Z) - 2\alpha(Q_1^* - E(X_1^2)) - 2\beta(Q_2^* - E(X_2^2))$$

$$\geq 0,$$

where $I(X)$ is the Fisher information of $X$.

Can establish this for some subset of the parameter space (involving $E(X_1^2), E(X_2^2)$).
Remarks about doubling trick

Does the "doubling trick" work to show Gaussian optimality?

**Remarks**

- There are interference channels for which
  \[ C_{X_1 \perp X_2}[(\lambda - 1)H(Y_2) + H(Y_1) - \lambda H(Y_2|X_2)] \]
  is not sub-additive

- To make this approach work, one needs to show the sub-additivity of the above functional for Gaussian interference channel
  - The proof of subadditivity needs to use the channel structure
Remarks about doubling trick

Does the "doubling trick" work to show Gaussian optimality?

Remarks

► There are interference channels for which

\[ C_{X_1 \perp X_2}[(\lambda - 1)H(Y_2) + H(Y_1) - \lambda H(Y_2|X_2)] \]

is not sub-additive

► To make this approach work, one needs to show the sub-additivity of the above functional for Gaussian interference channel

  • The proof of subadditivity needs to use the channel structure

► Of course, there are various arguments in literature that does rely on channel structure

  • Genie based converses
  • Injective deterministic interference channel
Recap

- Correlated Gaussians do not improve the HK-IB

- Conjectured an entropy-variance inequality
  - Motivated by considering the Fenchel dual form of a functional arising in the H–K region for the Z-interference channel
  - Presented possible attack strategies
Recap

- Correlated Gaussians do not improve the HK-IB

- Conjectured an entropy-variance inequality
  - Motivated by considering the Fenchel dual form of a functional arising in the H–K region for the Z-interference channel
  - Presented possible attack strategies
  - Hope it will be resolved one-way or the other soon by someone (perhaps one of you)

Thank You