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

Max Costa[1] Chandra Nair[2] David Ng[2] Yan Nan Wang[2]

[1]University of Campinas

[2]The Chinese University of Hong Kong

June 2020

Gaussian Z-interference channel

- An interference channel (IC) models the situation of two mutually interfering point-to-point communications over a shared medium. Formally, a memoryless interference channel is defined by a stochastic map $W(y_1, y_2|x_1, x_2)$.

Gaussian Z-interference channel

- An interference channel (IC) models the situation of two mutually interfering point-to-point communications over a shared medium. Formally, a memoryless interference channel is defined by a stochastic map $W(y_1, y_2|x_1, x_2)$.
- We are in particular interested to the Gaussian Z-interference channel (GZIC) setting as depicted below, where $a \in (0, 1)$. (If $a \geq 1$ this setting becomes strong interference and the capacity region is known [Costa '85])

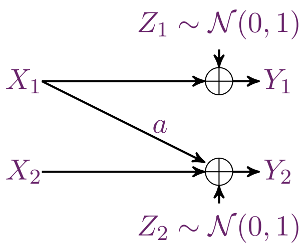
Gaussian Z-interference channel

- An interference channel (IC) models the situation of two mutually interfering point-to-point communications over a shared medium. Formally, a memoryless interference channel is defined by a stochastic map $W(y_1, y_2|x_1, x_2)$.
- We are in particular interested to the Gaussian Z-interference channel (GZIC) setting as depicted below, where $a \in (0, 1)$. (If $a \geq 1$ this setting becomes strong interference and the capacity region is known [Costa '85])

Gaussian Z-interference channel

$$Y_1 = X_1 + Z_1$$
$$Y_2 = X_2 + aX_1 + Z_2$$

where $Z_1, Z_2 \sim \mathcal{N}(0, 1)$ independent of X_1 and X_2.



Gaussian Z-interference channel

- We consider an equivalent formulation of the k-letter form of GZIC:

Gaussian Z-interference channel

- We consider an equivalent formulation of the k-letter form of GZIC:

Gaussian Z-interference channel

$$\mathbf{Y}_1 = \mathbf{X}_1 + \mathbf{Z}_1$$
$$\mathbf{Y}_2 = \mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0, I)$, $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2 I)$ (where $N_2 := \frac{1}{a^2} - 1$, $a \in (0,1)$) 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

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

where $P_1, P_2 \geq 0$.

Han–Kobayashi region

- For general two-user interference channel Han–Kobayashi (HK) region [Han, Kobayashi '81] is the best achievable region known.

Han–Kobayashi region

- For general two-user interference channel Han–Kobayashi (HK) region [Han, Kobayashi '81] is the best achievable region known.
- HK region is known to be strictly suboptimal for some discrete channels. In particular the HK region is strictly improved by multi-letter extensions. [Nair, Xia, Yazdanpanah '15]

Han–Kobayashi region

- For general two-user interference channel Han–Kobayashi (HK) region [Han, Kobayashi '81] is the best achievable region known.
- HK region is known to be strictly suboptimal for some discrete channels. In particular the HK region is strictly improved by multi-letter extensions. [Nair, Xia, Yazdanpanah '15]
- For GZIC *with Gaussian inputs* multi-letter extension does not improve. [Nair, Ng '19]

Han–Kobayashi region

- For general two-user interference channel Han–Kobayashi (HK) region [Han, Kobayashi '81] is the best achievable region known.
- HK region is known to be strictly suboptimal for some discrete channels. In particular the HK region is strictly improved by multi-letter extensions. [Nair, Xia, Yazdanpanah '15]
- For GZIC *with Gaussian inputs* multi-letter extension does not improve. [Nair, Ng '19]
- Question: Is k-letter HK region for GZIC $\stackrel{?}{=}$ k-letter HK region for GZIC *with Gaussian inputs*? i.e.

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

Han–Kobayashi region

- For general two-user interference channel Han–Kobayashi (HK) region [Han, Kobayashi '81] is the best achievable region known.
- HK region is known to be strictly suboptimal for some discrete channels. In particular the HK region is strictly improved by multi-letter extensions. [Nair, Xia, Yazdanpanah '15]
- For GZIC *with Gaussian inputs* multi-letter extension does not improve. [Nair, Ng '19]
- Question: Is k-letter HK region for GZIC $\stackrel{?}{=}$ k-letter HK region for GZIC *with Gaussian inputs*? i.e.

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

- If so then the single-letter HK region with Gaussian inputs $\mathcal{R}_{\text{HK-GS}}^{(1)}(P_1, P_2)$ is the capacity.

Han–Kobayashi region for GZIC

- For our GZIC the HK region reads:

k-letter HK region for GZIC

$$\begin{aligned} kR_1 &\le h(\mathbf{X}_1 + \mathbf{Z}_1 | \mathbf{Q}) - h(\mathbf{Z}_1) \\ kR_2 &\le h(\mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) - h(\mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) \\ k(R_1 + R_2) &\le h(\mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{Q}) - h(\mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) \\ &\quad + h(\mathbf{X}_1 + \mathbf{Z}_1 | \mathbf{U}_1, \mathbf{Q}) - h(\mathbf{Z}_1) \end{aligned}$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0, I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2 I)$.

Han–Kobayashi region for GZIC

- For our GZIC the HK region reads:

k-letter HK region for GZIC

$$
\begin{aligned}
kR_1 &\le h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{Q}) - h(\mathbf{Z}_1)\\
kR_2 &\le h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q}) - h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q})\\
k(R_1+R_2) &\le h(\mathbf{X}_2+\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{Q}) - h(\mathbf{X}_1+\mathbf{Z}_1+\mathbf{Z}_2|\mathbf{U}_1,\mathbf{Q})\\
&\quad + h(\mathbf{X}_1+\mathbf{Z}_1|\mathbf{U}_1,\mathbf{Q}) - h(\mathbf{Z}_1)
\end{aligned}
$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0,I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0,N_2 I)$.

- $(R_1,R_2) \in \mathcal{R}_{\mathrm{HK}}^{(k)}(P_1,P_2)$ if there exists $p(\mathbf{q})p(\mathbf{u}_1,\mathbf{x}_1|\mathbf{q})p(\mathbf{x}_2|\mathbf{q})$ satisfying the above three inequalities, along with $\mathrm{E}[\|\mathbf{X}_i\|^2] \le kP_i$ $(i=1,2)$.

Han–Kobayashi region for GZIC

- For our GZIC the HK region reads:

k-letter HK region for GZIC

$$\begin{aligned} kR_1 &\le h(\mathbf{X}_1 + \mathbf{Z}_1 | \mathbf{Q}) - h(\mathbf{Z}_1) \\ kR_2 &\le h(\mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) - h(\mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) \\ k(R_1 + R_2) &\le h(\mathbf{X}_2 + \mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{Q}) - h(\mathbf{X}_1 + \mathbf{Z}_1 + \mathbf{Z}_2 | \mathbf{U}_1, \mathbf{Q}) \\ &\quad + h(\mathbf{X}_1 + \mathbf{Z}_1 | \mathbf{U}_1, \mathbf{Q}) - h(\mathbf{Z}_1) \end{aligned}$$

where $\mathbf{Z}_1 \sim \mathcal{N}(0, I)$ and $\mathbf{Z}_2 \sim \mathcal{N}(0, N_2 I)$.

- $(R_1, R_2) \in \mathcal{R}_{\text{HK}}^{(k)}(P_1, P_2)$ if there exists $p(\mathbf{q})p(\mathbf{u}_1, \mathbf{x}_1|\mathbf{q})p(\mathbf{x}_2|\mathbf{q})$ satisfying the above three inequalities, along with $\mathrm{E}[\|\mathbf{X}_i\|^2] \le kP_i$ $(i = 1, 2)$.
- $(R_1, R_2) \in \mathcal{R}_{\text{HK-GS}}^{(k)}(P_1, P_2)$ if there exists $p(\mathbf{q})p(\mathbf{u}_1, \mathbf{x}_1|\mathbf{q})p(\mathbf{x}_2|\mathbf{q})$ satisfying the above three inequalities, along with $\mathrm{E}[\|\mathbf{X}_i\|^2] \le kP_i$ $(i = 1, 2)$, with $\mathbf{U}_1$, $\mathbf{X}_1 - \mathbf{U}_1$, $\mathbf{X}_2$ being independent zero-mean Gaussians conditioned on $\mathbf{Q}$ (i.e. with Gaussian inputs).

Han–Kobayashi region for GZIC: Remarks

- It is known that the time-sharing variable $\mathbf{Q}$ strictly improves on the HK region without time-sharing, i.e. under power constraints the optimal distribution for $(\mathbf{X}_1, \mathbf{X}_2)$ when $\mathbf{Q}$ is constant is not Gaussian but *mixture of Gaussian* instead.

Han–Kobayashi region for GZIC: Remarks

- It is known that the time-sharing variable $\mathbf{Q}$ strictly improves on the HK region without time-sharing, i.e. under power constraints the optimal distribution for $(\mathbf{X}_1, \mathbf{X}_2)$ when $\mathbf{Q}$ is constant is not Gaussian but *mixture of Gaussian* instead.

- The traditional "monotonicity along a path" approach [Stam '59] for proving Gaussian optimality of non-convex functionals hence fails.

Han–Kobayashi region for GZIC: Remarks

- It is known that the time-sharing variable $\mathbf{Q}$ strictly improves on the HK region without time-sharing, i.e. under power constraints the optimal distribution for $(\mathbf{X}_1, \mathbf{X}_2)$ when $\mathbf{Q}$ is constant is not Gaussian but *mixture of Gaussian* instead.

- The traditional "monotonicity along a path" approach [Stam '59] for proving Gaussian optimality of non-convex functionals hence fails.

- This motivates us to consider certain Fenchel dual functional which we conjectured to be optimized by Gaussian.

Gaussian optimality conjecture

- In this paper we propose the following conjecture concerning Gaussian optimality of certain functional, which if true would imply that $\mathcal{R}_{\text{HK}}^{(k)}(P_1, P_2) = \mathcal{R}_{\text{HK-GS}}^{(k)}(P_1, P_2)$ and hence would solve the capacity of GZIC.

Gaussian optimality conjecture

- In this paper we propose the following conjecture concerning Gaussian optimality of certain functional, which if true would imply that $\mathcal{R}^{(k)}_{\text{HK}}(P_1, P_2) = \mathcal{R}^{(k)}_{\text{HK-GS}}(P_1, P_2)$ and hence would solve the capacity of GZIC.

Conjecture 1

For $\beta \geq 1$, $N_2 \geq 0$ and $k \times k$ matrices $\Sigma_1, A_2 \succeq 0$, the maximum

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

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 $\mathbf{X}_2$.

Gaussian optimality conjecture

- In this paper we propose the following conjecture concerning Gaussian optimality of certain functional, which if true would imply that $\mathcal{R}^{(k)}_{\text{HK}}(P_1, P_2) = \mathcal{R}^{(k)}_{\text{HK-GS}}(P_1, P_2)$ and hence would solve the capacity of GZIC.

Conjecture 1

For $\beta \geq 1$, $N_2 \geq 0$ and $k \times k$ matrices $\Sigma_1, A_2 \succeq 0$, the maximum

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

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 $\mathbf{X}_2$.

- In the following sessions we will show how does this conjecture imply the optimality of Gaussian inputs.

Dual functional

With a duality argument as in [Costa, Nair '16] for $\beta \geq 1$ and $Q_1, Q_2 \geq 0$ we have

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

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

Dual functional

With a duality argument as in [Costa, Nair '16] for $\beta \geq 1$ and $Q_1, Q_2 \geq 0$ we have

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

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

where

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

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

Dual functional

With a duality argument as in [Costa, Nair '16] for $\beta \geq 1$ and $Q_1, Q_2 \geq 0$ we have

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

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

where

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

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

and

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

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

Dual functional

- Making use of the above dual functional characterization one sees that to prove $\mathcal{R}^{(k)}_{\text{HK}}(P_1, P_2) = \mathcal{R}^{(k)}_{\text{HK-GS}}(P_1, P_2)$ it suffices to show

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

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

Dual functional

- Making use of the above dual functional characterization one sees that to prove $\mathcal{R}^{(k)}_{\text{HK}}(P_1, P_2) = \mathcal{R}^{(k)}_{\text{HK-GS}}(P_1, P_2)$ it suffices to show

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

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

- Our proposed conjecture instead implies that

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

Dual functional

- Making use of the above dual functional characterization one sees that to prove $\mathcal{R}^{(k)}_{\text{HK}}(P_1, P_2) = \mathcal{R}^{(k)}_{\text{HK-GS}}(P_1, P_2)$ it suffices to show

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

  with $K_i = \text{E}[\mathbf{X}_i \mathbf{X}_i^T]$ $(i = 1, 2)$.
- Our proposed conjecture instead implies that

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

- It is *not* in general true for all functionals ϕ that

$$\mathcal{C}_{K_1}[\phi(K_1)] \leq \max_{0 \preceq \hat{K}_1 \preceq K_1} \phi(\hat{K}_1)$$

  (although "$\geq$" always holds under certain regularity conditions). However we can show that the functional $K_1 \mapsto \psi_{\text{G}}(K_1, K_2)$ has such property and this implies the sufficiency of our conjecture.

Main theorem

Theorem 1

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

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

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

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

for any $K_1, K_2 \succeq 0$.

Main theorem: Proof for "$\geq$" part

- The $\geq$ part follows easily from the dual characterization of upper concave envelope.

Main theorem: Proof for "$\geq$" part

- The $\geq$ part follows easily from the dual characterization of upper concave envelope.

$$\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1, K_2)\right] = \inf_{\substack{\Sigma_1 \\ \Sigma_1 = \Sigma_1^T}} \Bigg[\underbrace{\sup_{\hat{K}_1 \succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1, K_2) - \mathrm{tr}(\Sigma_1 \hat{K}_1)\right]}_{= +\infty \text{ if } \Sigma_1 \not\succeq 0} + \mathrm{tr}(\Sigma_1 K_1) \Bigg]$$

Main theorem: Proof for "$\geq$" part

- The $\geq$ part follows easily from the dual characterization of upper concave envelope.

$$\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1,K_2)\right] = \inf_{\substack{\Sigma_1 \\ \Sigma_1=\Sigma_1^T}} \Bigg[\underbrace{\sup_{\hat{K}_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right]}_{=+\infty \text{ if } \Sigma_1 \not\succeq 0} + \mathrm{tr}(\Sigma_1 K_1)\Bigg]$$

$$= \inf_{\Sigma_1\succeq 0} \left[\sup_{\hat{K}_1\succeq 0}\left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right] + \mathrm{tr}(\Sigma_1 K_1)\right]$$

Main theorem: Proof for "$\geq$" part

- The $\geq$ part follows easily from the dual characterization of upper concave envelope.

$$
\begin{aligned}
\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1, K_2)\right] &= \inf_{\substack{\Sigma_1 \\ \Sigma_1 = \Sigma_1^T}} \Bigg[\underbrace{\sup_{\hat{K}_1 \succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1, K_2) - \mathrm{tr}(\Sigma_1 \hat{K}_1)\right]}_{= +\infty \text{ if } \Sigma_1 \not\succeq 0} + \mathrm{tr}(\Sigma_1 K_1) \Bigg] \\
&= \inf_{\Sigma_1 \succeq 0} \left[\sup_{\hat{K}_1 \succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1, K_2) - \mathrm{tr}(\Sigma_1 \hat{K}_1)\right] + \mathrm{tr}(\Sigma_1 K_1) \right] \\
&\geq \sup_{\hat{K}_1 \succeq 0} \underbrace{\inf_{\Sigma_1 \succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1, K_2) - \mathrm{tr}(\Sigma_1 \hat{K}_1) + \mathrm{tr}(\Sigma_1 K_1)\right]}_{= -\infty \text{ if } K_1 - \hat{K}_1 \not\succeq 0}
\end{aligned}
$$

Main theorem: Proof for "$\geq$" part

- The $\geq$ part follows easily from the dual characterization of upper concave envelope.

$$
\begin{aligned}
\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1,K_2)\right] &= \inf_{\substack{\Sigma_1 \\ \Sigma_1=\Sigma_1^T}} \Bigg[\underbrace{\sup_{\hat{K}_1\succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right]}_{=+\infty \text{ if } \Sigma_1 \not\succeq 0} + \mathrm{tr}(\Sigma_1 K_1)\Bigg] \\
&= \inf_{\Sigma_1\succeq 0} \left[\sup_{\hat{K}_1\succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1)\right] + \mathrm{tr}(\Sigma_1 K_1)\right] \\
&\geq \sup_{\hat{K}_1\succeq 0} \underbrace{\inf_{\Sigma_1\succeq 0} \left[\psi_{\mathrm{G}}(\hat{K}_1,K_2) - \mathrm{tr}(\Sigma_1\hat{K}_1) + \mathrm{tr}(\Sigma_1 K_1)\right]}_{=-\infty \text{ if } K_1-\hat{K}_1 \not\succeq 0} \\
&\geq \max_{\substack{\hat{K}_1\succeq 0 \\ \hat{K}_1\preceq K_1}} \psi_{\mathrm{G}}(\hat{K}_1,K_2)
\end{aligned}
$$

Main theorem: Proof for "$\leq$" part

- Recall that we have defined

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

  and so $\psi(\mathbf{X}_1, \mathbf{X}_2) = \psi_{\mathrm{G}}(K_1, K_2)$ for $\mathbf{X}_i \sim \mathcal{N}(0, K_i)$ $(i = 1, 2)$.

Main theorem: Proof for "$\leq$" part

- Recall that we have defined

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

  and so $\psi(\mathbf{X}_1, \mathbf{X}_2) = \psi_{\mathrm{G}}(K_1, K_2)$ for $\mathbf{X}_i \sim \mathcal{N}(0, K_i)$ $(i = 1, 2)$.
- Fixing $K_2 \succeq 0$ and $\mathbf{X}_2 \sim \mathcal{N}(0, K_2)$ we have

$$\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1, K_2)\right] \leq \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

  since the right hand side is a concave functional in K_1 that upper bounds $\psi_{\mathrm{G}}(K_1, K_2)$.

Main theorem: Proof for "$\le$" part

- Recall that we have defined

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

  and so $\psi(\mathbf{X}_1, \mathbf{X}_2) = \psi_{\mathrm{G}}(K_1, K_2)$ for $\mathbf{X}_i \sim \mathcal{N}(0, K_i)$ $(i = 1, 2)$.
- Fixing $K_2 \succeq 0$ and $\mathbf{X}_2 \sim \mathcal{N}(0, K_2)$ we have

$$\mathcal{C}_{K_1}\left[\psi_{\mathrm{G}}(K_1, K_2)\right] \le \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

  since the right hand side is a concave functional in K_1 that upper bounds $\psi_{\mathrm{G}}(K_1, K_2)$.
- It remains the establish the following:

Proposition 1

Let $K_1 \succeq 0$ and let $\mathbf{X}_2$ be Gaussian. Then the maximum

$$\max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

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

Proof of Proposition 1

- We will show Proposition 1 by a subadditivity argument, applying the ”doubling trick” developed in [Geng, Nair ’14].

Proof of Proposition 1

- We will show Proposition 1 by a subadditivity argument, applying the "doubling trick" developed in [Geng, Nair '14].
- Take a maximizer $p^*(\mathbf{x}_1, \mathbf{u}_1)$ (existence of which can be justified by Prokhorov theorem through techniques in Appendix II of [Geng, Nair '14]) for

$$v := \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

We can assume without loss of generality $p^*(\mathbf{x}_1|\mathbf{u}_1)$ is zero-mean or otherwise replace $\mathbf{X}_1$ by $\mathbf{X}_1 - \mathrm{E}[\mathbf{X}_1|\mathbf{U}_1]$.

Proof of Proposition 1

- We will show Proposition 1 by a subadditivity argument, applying the "doubling trick" developed in [Geng, Nair '14].
- Take a maximizer $p^*(\mathbf{x}_1, \mathbf{u}_1)$ (existence of which can be justified by Prokhorov theorem through techniques in Appendix II of [Geng, Nair '14]) for

$$v := \max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

  We can assume without loss of generality $p^*(\mathbf{x}_1|\mathbf{u}_1)$ is zero-mean or otherwise replace $\mathbf{X}_1$ by $\mathbf{X}_1 - \mathrm{E}[\mathbf{X}_1|\mathbf{U}_1]$.
- Doubling: Take two independent copies $(\mathbf{X}_{11}^*, \mathbf{U}_{11}^*)$, $(\mathbf{X}_{12}^*, \mathbf{U}_{12}^*)$ of the maximizer. Let

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

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

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

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

Proof of Proposition 1

Then

$2v$

Proof of Proposition 1

Then

$$2v = \mathrm{E}_{\mathbf{U}_{11}^*}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*}, \mathbf{X}_2)] + \mathrm{E}_{\mathbf{U}_{12}^*}[\psi(\mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*}, \mathbf{X}_2)]$$

Proof of Proposition 1

Then

$$
\begin{aligned}
2v &= \mathrm{E}_{\mathbf{U}_{11}^*}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*}, \mathbf{X}_2)] + \mathrm{E}_{\mathbf{U}_{12}^*}[\psi(\mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*}, \mathbf{X}_2)] \\
&= \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_1}, \mathbf{X}_2) + \psi(\mathbf{X}_{12}^*|_{\mathbf{U}_1}, \mathbf{X}_2)]
\end{aligned}
$$

Proof of Proposition 1

Then

$$\begin{aligned} 2v &= \mathrm{E}_{\mathbf{U}_{11}^*}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*}, \mathbf{X}_2)] + \mathrm{E}_{\mathbf{U}_{12}^*}[\psi(\mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*}, \mathbf{X}_2)] \\ &= \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_1}, \mathbf{X}_2) + \psi(\mathbf{X}_{12}^*|_{\mathbf{U}_1}, \mathbf{X}_2)] \\ &= (\beta - 1)h(\mathbf{Y}_{31}, \mathbf{Y}_{32}|\mathbf{U}_1) + h(\mathbf{Y}_{11}, \mathbf{Y}_{12}|\mathbf{U}_1) - \beta h(\mathbf{Y}_{21}, \mathbf{Y}_{22}|\mathbf{U}_1) \end{aligned}$$

Proof of Proposition 1

Then

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

Proof of Proposition 1

Then

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

Proof of Proposition 1

Then

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

Proof of Proposition 1

Then

$$
\begin{aligned}
2v &= \mathrm{E}_{\mathbf{U}_{11}^*}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_{11}^*}, \mathbf{X}_2)] + \mathrm{E}_{\mathbf{U}_{12}^*}[\psi(\mathbf{X}_{12}^*|_{\mathbf{U}_{12}^*}, \mathbf{X}_2)] \\
&= \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{11}^*|_{\mathbf{U}_1}, \mathbf{X}_2) + \psi(\mathbf{X}_{12}^*|_{\mathbf{U}_1}, \mathbf{X}_2)] \\
&= (\beta - 1)h(\mathbf{Y}_{31}, \mathbf{Y}_{32}|\mathbf{U}_1) + 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) + I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1)] \\
&\qquad + [h(\mathbf{Y}_{11}|\mathbf{Y}_{32}, \mathbf{U}_1) + h(\mathbf{Y}_{12}|\mathbf{Y}_{11}, \mathbf{U}_1) + I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1)] \\
&\qquad - \beta[h(\mathbf{Y}_{21}|\mathbf{Y}_{32}, \mathbf{U}_1) + h(\mathbf{Y}_{22}|\mathbf{Y}_{11}, \mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{U}_1) \\
&\qquad\qquad - I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{U}_1)] \\
&= \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{11}|_{\mathbf{Y}_{32},\mathbf{U}_1}, \mathbf{X}_{21})] + \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{12}|_{\mathbf{Y}_{11},\mathbf{U}_1}, \mathbf{X}_{22})] \\
&\qquad + \beta[I(\mathbf{Y}_{11}; \mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21}; \mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{U}_1) + I(\mathbf{Y}_{21}; \mathbf{Y}_{22}|\mathbf{U}_1)] \\
&= \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{11}|_{\mathbf{Y}_{32},\mathbf{U}_1}, \mathbf{X}_{21})] + \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_{12}|_{\mathbf{Y}_{11},\mathbf{U}_1}, \mathbf{X}_{22})] \\
&\qquad - \beta I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1) \\
&\le 2v - \beta I(\mathbf{Y}_{11}; \mathbf{Y}_{22}|\mathbf{Y}_{21}, \mathbf{Y}_{32}, \mathbf{U}_1)
\end{aligned}
$$

Proof of Proposition 1

where the orange terms

$$I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32} \to (\mathbf{Y}_{22},\mathbf{U}_1) \to \mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32} \to (\mathbf{Y}_{22},\mathbf{U}_1) \to \mathbf{Y}_{21}]}$$

Proof of Proposition 1

where the orange terms

$$I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{21}]}$$

$$= I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1)$$

Proof of Proposition 1

where the orange terms

$$I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{21}]}$$

$$= I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1)$$

$$= -\underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)}_{[\mathbf{Y}_{21}\to(\mathbf{Y}_{11},\mathbf{Y}_{32},\mathbf{U}_1)\to\mathbf{Y}_{22}]} + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)$$

Proof of Proposition 1

where the orange terms

$$I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{21}]}$$

$$= I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1)$$

$$= -\underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)}_{[\mathbf{Y}_{21}\to(\mathbf{Y}_{11},\mathbf{Y}_{32},\mathbf{U}_1)\to\mathbf{Y}_{22}]} + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)$$

$$= -I(\mathbf{Y}_{11},\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)$$

Proof of Proposition 1

where the orange terms

$$I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{21}]}$$

$$= I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1)$$

$$= -\underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)}_{[\mathbf{Y}_{21}\to(\mathbf{Y}_{11},\mathbf{Y}_{32},\mathbf{U}_1)\to\mathbf{Y}_{22}]} + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)$$

$$= -I(\mathbf{Y}_{11},\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)$$

$$= -I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{21},\mathbf{Y}_{32},\mathbf{U}_1)$$

Proof of Proposition 1

where the orange terms

$$
\begin{aligned}
& I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{11}]} + \underbrace{I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{U}_1)}_{[\mathbf{Y}_{32}\to(\mathbf{Y}_{22},\mathbf{U}_1)\to\mathbf{Y}_{21}]} \\
&= I(\mathbf{Y}_{11};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{21};\mathbf{Y}_{32}|\mathbf{U}_1) - I(\mathbf{Y}_{11};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22},\mathbf{Y}_{32}|\mathbf{U}_1) \\
&= - \underbrace{I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1)}_{[\mathbf{Y}_{21}\to(\mathbf{Y}_{11},\mathbf{Y}_{32},\mathbf{U}_1)\to\mathbf{Y}_{22}]} + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1) \\
&= -I(\mathbf{Y}_{11},\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1) + I(\mathbf{Y}_{21};\mathbf{Y}_{22}|\mathbf{Y}_{32},\mathbf{U}_1) \\
&= -I(\mathbf{Y}_{11};\mathbf{Y}_{22}|\mathbf{Y}_{21},\mathbf{Y}_{32},\mathbf{U}_1)
\end{aligned}
$$

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.

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Recall that we have

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

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Recall that we have

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

Since we also have

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

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Recall that we have

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

Since we also have

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

By Lemma 1 we get

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

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Recall that we have

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

Since we also have

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

By Lemma 1 we get

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

Again we also have

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

Proof of Proposition 1

We need the following lemma to proceed:

Lemma 1 (Double Markovity)

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 $\mathbf{X} \to (\mathbf{Y}, \mathbf{Q}) \to \mathbf{Z}$ and $\mathbf{Y} \to (\mathbf{X}, \mathbf{Q}) \to \mathbf{Z}$ form Markov chains. Then $(\mathbf{X}, \mathbf{Y}) \to \mathbf{Q} \to \mathbf{Z}$ forms a Markov chain.

Recall that we have

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

Since we also have

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

By Lemma 1 we get

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

Again we also have

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

and hence by Lemma 1 we obtain a Markov chain

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

Proof of Proposition 1

We will invoke the following lemma, which can be shown by considering the characteristic functions:

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 of Proposition 1

We will invoke the following lemma, which can be shown by considering the characteristic functions:

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 $\mathbf{X}_1 \perp \mathbf{X}_2$.

Recall that we have

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

Proof of Proposition 1

We will invoke the following lemma, which can be shown by considering the characteristic functions:

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 $\mathbf{X}_1 \perp \mathbf{X}_2$.

Recall that we have

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

In particular

$$\mathbf{Y}_{11} \to \mathbf{U}_1 \to \mathbf{Y}_{22}$$

Proof of Proposition 1

We will invoke the following lemma, which can be shown by considering the characteristic functions:

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 $\mathbf{X}_1 \perp \mathbf{X}_2$.

Recall that we have

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

In particular

$$\mathbf{Y}_{11} \to \mathbf{U}_1 \to \mathbf{Y}_{22}$$

That is,

$$\mathbf{X}_{11} + \mathbf{Z}_{11} \to \mathbf{U}_1 \to \mathbf{X}_{12} + \mathbf{Z}_{12} + \mathbf{Z}_{22}$$

Proof of Proposition 1

We will invoke the following lemma, which can be shown by considering the characteristic functions:

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 $\mathbf{X}_1 \perp \mathbf{X}_2$.

Recall that we have

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

In particular

$$\mathbf{Y}_{11} \to \mathbf{U}_1 \to \mathbf{Y}_{22}$$

That is,

$$\mathbf{X}_{11} + \mathbf{Z}_{11} \to \mathbf{U}_1 \to \mathbf{X}_{12} + \mathbf{Z}_{12} + \mathbf{Z}_{22}$$

Hence by Lemma 2,

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

forms a Markov chain.

Proof of Proposition 1

Equivalently we have

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

Proof of Proposition 1

Equivalently we have

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

and the following lemma 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, where $\mathbf{u}_{11}^*$, $\mathbf{u}_{12}^*$ are arbitrary.

Proof of Proposition 1

Equivalently we have

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

and the following lemma 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, where $\mathbf{u}_{11}^*$, $\mathbf{u}_{12}^*$ are arbitrary.

Lemma 3 (Corollary 3 of [Geng, Nair '14])

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 of Proposition 1

Equivalently we have

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

and the following lemma 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, where $\mathbf{u}_{11}^*$, $\mathbf{u}_{12}^*$ are arbitrary.

Lemma 3 (Corollary 3 of [Geng, Nair '14])

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.

This means that the maximizing distribution $(\mathbf{X}_1, \mathbf{U}_1) \sim p^*(\mathbf{x}_1, \mathbf{u}_1)$ for

$$\max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

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 $\hat{K}_1 \succeq 0$.

Proof of Proposition 1

Equivalently we have

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

and the following lemma 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, where $\mathbf{u}_{11}^*$, $\mathbf{u}_{12}^*$ are arbitrary.

Lemma 3 (Corollary 3 of [Geng, Nair '14])

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.

This means that the maximizing distribution $(\mathbf{X}_1, \mathbf{U}_1) \sim p^*(\mathbf{x}_1, \mathbf{u}_1)$ for

$$\max_{\substack{p(\mathbf{x}_1)p(\mathbf{u}_1|\mathbf{x}_1) \\ \mathrm{E}[\mathbf{X}_1\mathbf{X}_1^T] \preceq K_1}} \mathrm{E}_{\mathbf{U}_1}[\psi(\mathbf{X}_1|_{\mathbf{U}_1}, \mathbf{X}_2)]$$

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. This concludes the proof of Proposition 1 and Theorem 1.

Conclusion

- We establish some properties concerning the maximizer of some non-convex matrix functionals using information-theoretic methods.

Conclusion

- We establish some properties concerning the maximizer of some non-convex matrix functionals using information-theoretic methods.
- Based on such result we propose a Gaussian optimality conjecture which if true would imply the capacity region of the Gaussian Z-interference channel.

Conclusion

- We establish some properties concerning the maximizer of some non-convex matrix functionals using information-theoretic methods.

- Based on such result we propose a Gaussian optimality conjecture which if true would imply the capacity region of the Gaussian Z-interference channel.

Thank you