Information Spectrum Converse for Minimum Entropy Couplings and Functional Representations

Anuj Yadav **Yanina Shkel**

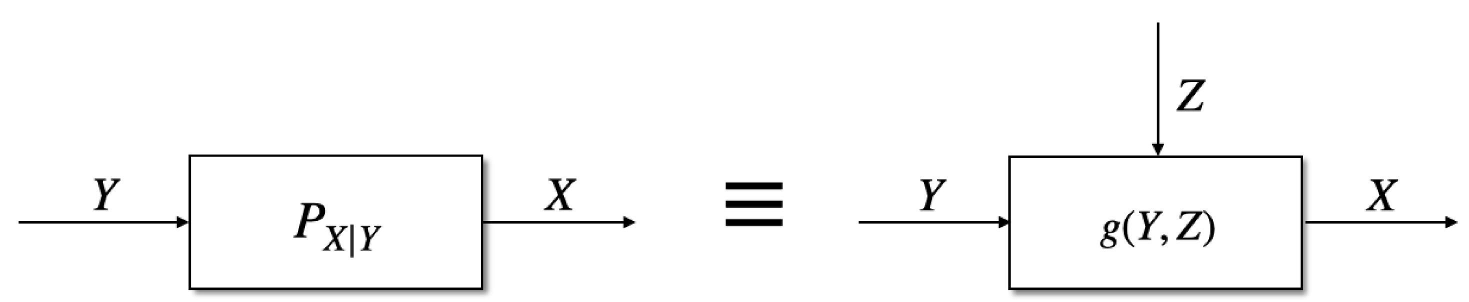
EPFL Information Processing Group

Functional Representation Lemma (FRL)

Given $(X, Y) \sim P_{X,Y}$, there exists a random variable Z ($Z \perp Y$) and a function $g(\cdot,\cdot)$ s.t. $X = g(Y, Z)$ i.e.,

$$I(Y; Z) = 0$$

$$H(X|Y, Z) = 0$$



Minimum Entropy - {Functional Representations (FR), Couplings (C)}

Minimum Entropy - FR

Given: $(X, Y) \sim P_{X,Y}$
Find: random variable Z
Such that: $Y \perp Z$
$X = g(Y, Z)$
Minimize: Rényi entropy $H_\alpha(Z)$
$\forall \alpha \geq 0$

Minimum Entropy - C

Given: marginal distributions $\{P_1, P_2, \cdots, P_m\}$
Find: coupling $(X_1, X_2, \cdots, X_m)$
Such that: $X_i \sim P_i$
$\forall i \in \{1, 2, \cdots, m\}$
Minimize: Rényi entropy $H_\alpha(X_1, X_2, \cdots, X_m)$
$\forall \alpha \geq 0$

The Problem

The two problems are one !

Finding the minimum Rényi entropy of Z in FRL is equivalent to solving the minimum entropy coupling problem for the marginal distributions $\{P_{X|Y=y}\}_{y\in\mathcal{Y}}$.

- However, it is a **NP-Hard** Problem !!!
- **Lower Bounds** - Converse type results.
  Upper Bounds - Achievability type results.

We are only concerned with **lower bounds (converses)** here!

Information and Entropy

Information:

$$\imath_X(x) := \log \frac{1}{P_X(x)} \quad \forall x \in \mathcal{X}$$

Information spectrum of X:

$$\mathbb{F}_{\imath_X}(t) = \mathbb{P}[\imath_X(X) \leq t] \quad \forall t \in [0, \infty)$$

Shannon entropy:

$$H(X) = \mathbb{E}[\imath_X(X)] = \int_0^\infty \left(1 - \mathbb{F}_{\imath_X}(t)\right) dt$$

Rényi entropy:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\mathbb{E}\left[2^{(1-\alpha)\imath_X(X)}\right]\right) \quad \forall \alpha \in [0, 1) \cup (1, \infty)$$

Majorization ($\preceq_m$)

Definition:

given,

$$Q = (q_1, q_2, q_3, \cdots); \;\; q_1 \geq q_2 \geq q_3 \geq \cdots$$
$$P = (p_1, p_2, p_3, \cdots); \;\; p_1 \geq p_2 \geq p_3 \geq \cdots$$

We say $Q \preceq_m P$, if:

$$\sum_{i=1}^k q_i \leq \sum_{i=1}^k p_i; \quad \forall k \geq 1$$

Greatest lower bound w.r.t Majorization:

$$\bigwedge_{j=1}^m P_j \preceq_m P_i; \quad \forall i \in \{1, 2, \cdots, m\}$$

$$Q \preceq_m P_i \implies Q \preceq \bigwedge_{j=1}^m P_j$$

Schur concavity:

$$Q \preceq_m P \implies H_\alpha(Q) \geq H_\alpha(P)$$

Existing Lower bound - 1

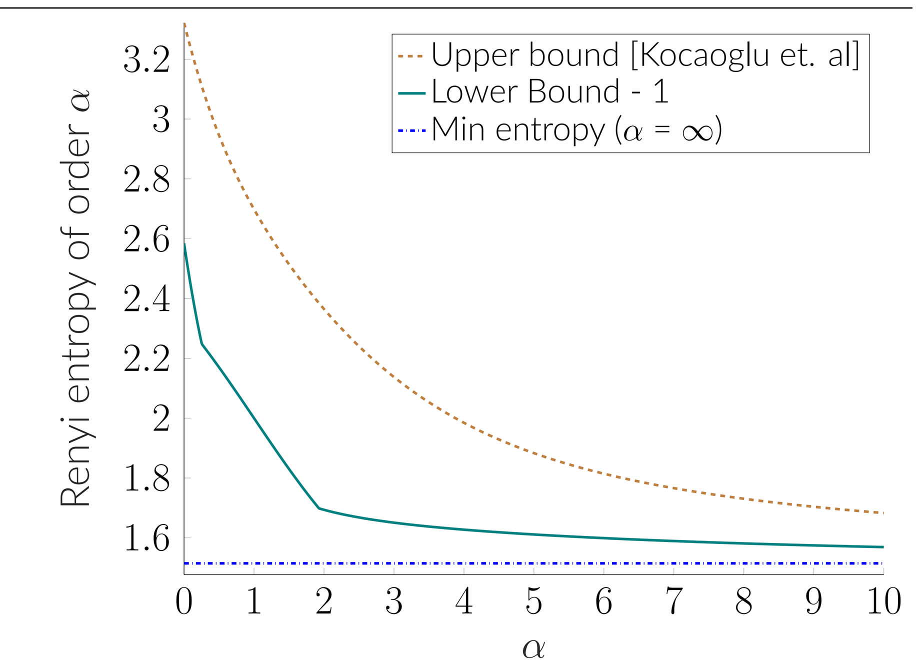
- $P_{\{X_1,\cdots,X_m\}} \sqsubseteq P_{X_i} \quad (\forall i \in \{1, \cdots, m\})$
- Aggregation implies majorization, i.e.,
- $P_{\{X_1,\cdots,X_m\}} \preceq_m P_{X_i} \quad (\forall i \in \{1, \cdots, m\})$

⇓ [Schur concavity]

$$H_\alpha(X_1, X_2, \cdots, X_m) \geq \max_{i \in \{1,2,\cdots,m\}} H_\alpha(X_i)$$

or

$$H_\alpha(Z) \geq \max_{y\in\mathcal{Y}} H_\alpha(X|Y = y)$$



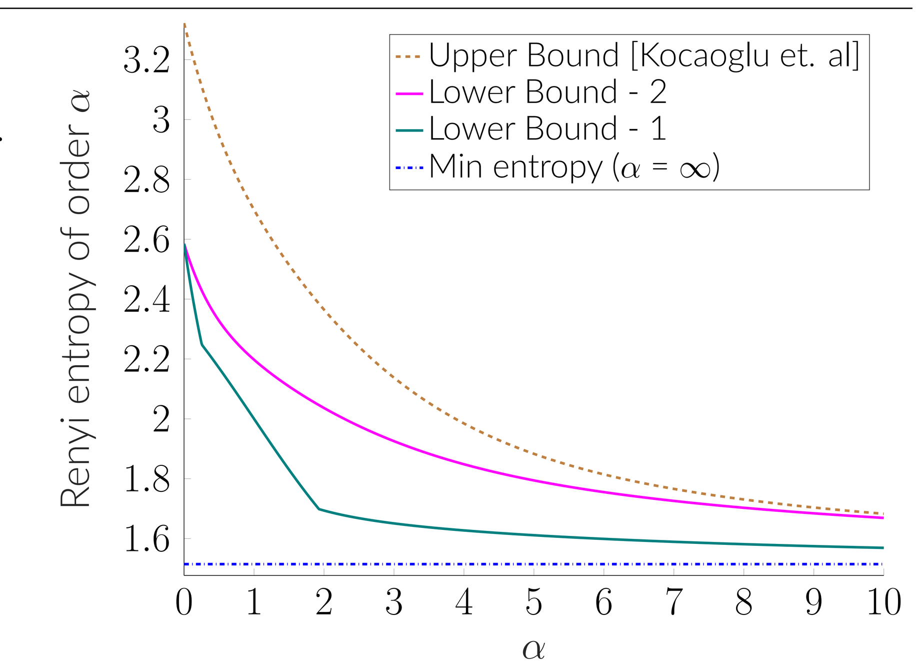
Existing Lower bound - 2

- $P_{\{X_1,\cdots,X_m\}} \preceq_m P_{X_i} \quad (\forall i \in \{1, \cdots, m\})$
- $\preceq_m$ is a partial order and a complete lattice.
- $P_{\{X_1,\cdots,X_m\}} \preceq_m \left(\bigwedge_{j=1}^m P_j\right)$

$$H_\alpha(X_1, \cdots, X_m) \geq H_\alpha\left(\bigwedge_{i=1}^m P_i\right)$$

or

$$H_\alpha(Z) \geq H_\alpha\left(\bigwedge_{y\in\mathcal{Y}} P_{X|Y=y}\right)$$



Majorization - in 'information-spectrum' sense ($\preceq_\imath$)

Let $U \sim Q$ and $V \sim P$. We say $Q \preceq_\imath P$, if

$$F_{\imath_U}(t) \leq F_{\imath_V}(t) \implies \mathbb{P}[\imath_U(U) \leq t] \leq \mathbb{P}[\imath_V(V) \leq t] \quad \forall t \in [0, \infty)$$

Lemma 1: $Q \preceq_\imath P \Rightarrow Q \preceq_m P$

Lemma 2: Let $\mathcal{F} = \{Q : Q \preceq_\imath P_i \quad \forall i \in \{1, \ldots, m\}\}$
Then, $\exists\, Q^* \in \mathcal{F}$ s.t. $Q \preceq_m Q^* \; \forall\, Q \in \mathcal{F}$.

Main Result - 1

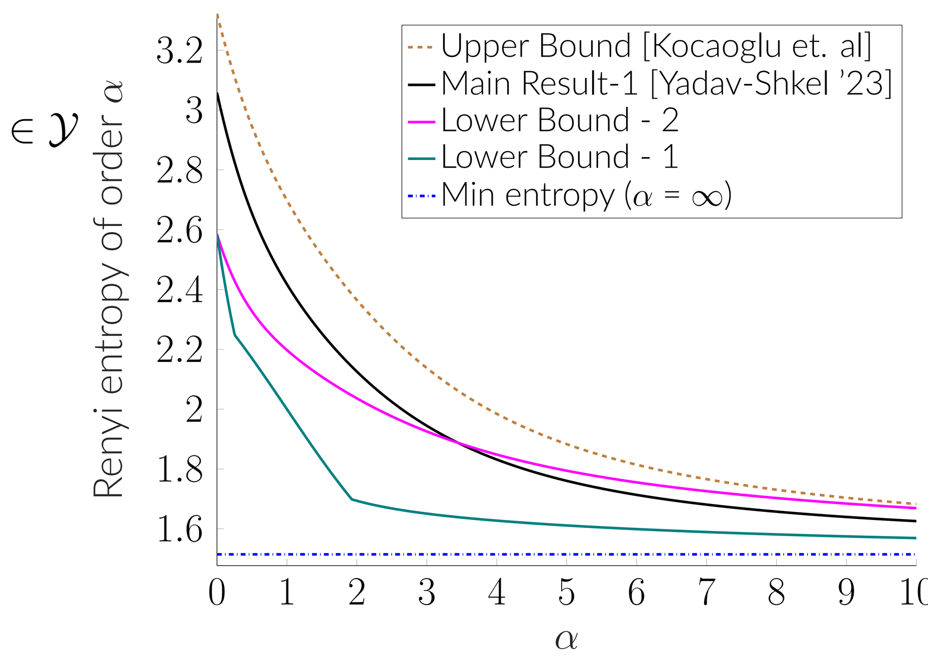
$$\mathbb{P}[\imath_Z(Z) > t] \geq \max_{y\in\mathcal{Y}} \mathbb{P}[\imath_{X|Y}(X|Y) > t | Y = y]$$

$$\implies \mathbb{F}_{\imath_{X|Y=y}}(t) \geq \mathbb{F}_{\imath_Z}(t) \text{ or } P_Z \preceq_\imath P_{X|Y=y} \quad \forall y \in \mathcal{Y}$$

⇓

$$H(Z) = \mathbb{E}[\imath_Z(Z)] = \int_0^\infty \left(1 - \mathbb{F}_{\imath_Z}(t)\right) dt \geq \int_0^\infty \max_{y\in\mathcal{Y}} \left(1 - \mathbb{F}_{\imath_{X|Y=y}}(t)\right) dt$$

$$H(Z) \geq \int_0^\infty \max_{y\in\mathcal{Y}} \left(1 - \mathbb{F}_{\imath_{X|Y=y}}(t)\right) dt$$



Main Result - 2

From above: $P_Z \preceq_\imath P_{X|Y=y} \quad \forall y \in \mathcal{Y}$

Let:

$$\mathcal{S} = \{Q : Q \preceq_\imath P_{X|Y=y} \; \forall y \in \mathcal{Y}\}$$

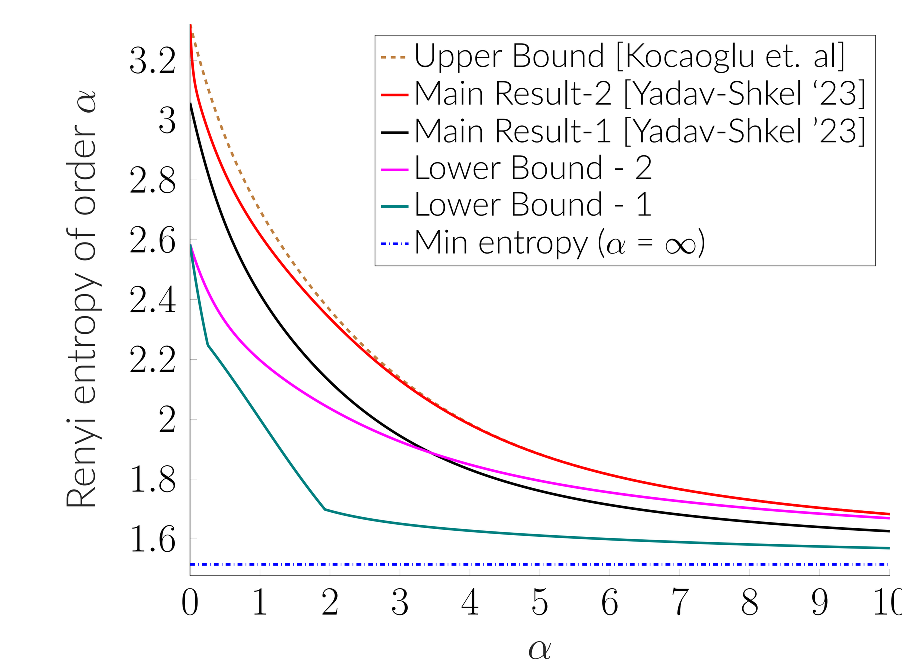
Then, $\exists\, Q^* \in \mathcal{S}$ s.t. $Q \preceq_m Q^*$; $\forall Q \in \mathcal{S}$

$$\implies Z \preceq_m Q^* \preceq_m P_{X|Y=y}$$

Therefore,

$$H_\alpha(Z) \geq H_\alpha(Q^*)$$

- support size of the lower bounding distribution is enlarged !
- Improves on all other existing lower bounds !



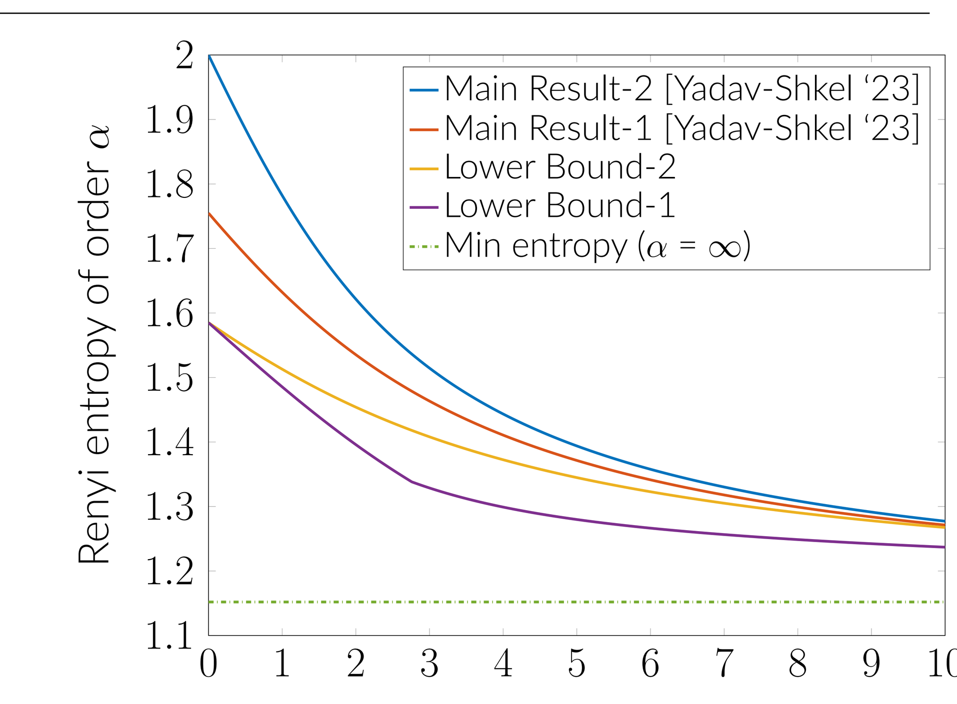
Example:

Given:

$$P_{X|Y}(\cdot|y_1) = (0.45, 0.4, 0.15)$$
$$P_{X|Y}(\cdot|y_2) = (0.5, 0.3, 0.2)$$

Entropy computation for all lower bounds:

- LB - 1: $H(Z) \geq 1.4855$
- LB - 2: $H(Z) \geq 1.5129$
- LB - 3: $H(Z) \geq 1.6325$
- LB - 4: $H(Z) \geq 1.7822$



2023 IEEE European School of Information Theory (ESIT) – University of Bristol, United Kingdom

anuj.yadav@epfl.ch