EDIC Candidacy Exam

Majorization Techniques for Entropy Bounds

Anuj Kumar Yadav
LINX
EPFL

EPFL Information Processing Group

15th June 2023

Outline

- **Majorization** [Cicalese et. al '02]
  - Majorization Partial Order
  - It is a lattice!
  - Properties of Entropy on the Majorization Lattice
- **Applications of Majorization** [Sason '18 & Cicalese et. al '19]
  - Lower Bound on Entropy of Random Variables
  - Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$
  - Probability Mass Function Truncation
- **Future Work**

Prelude

Notations

- $\mathscr{P}_n$: set of all PMFs of size n
- $\mathscr{P}_n' \subset \mathscr{P}_n$: set of all ordered PMFs (non-increasing order) of size n
- $H_\alpha(X) \equiv H_\alpha(\mathbf{p})$: Rényi entropy of $X \sim \mathbf{p}$

Prelude

Majorization $`\preceq'$

- Let $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \mathscr{P}_n$. Sort $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ in the non-increasing order, say $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$.

Prelude

Majorization ` $\preceq$ '

- Let $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \mathscr{P}_n$. Sort $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ in the non-increasing order, say $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$.
- Then, $\hat{\mathbf{q}}$ majorizes $\hat{\mathbf{p}}$, i.e., $\hat{\mathbf{p}} \preceq \hat{\mathbf{q}}$ if

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \qquad \forall k \in \{1, \cdots, n\}$$

$$p_1 \leq q_1$$
$$p_1 + p_2 \leq q_1 + q_2$$
$$p_1 + p_2 + p_3 \leq q_1 + q_2 + q_3$$
$$\vdots$$

Prelude

Majorization \`$\preceq$'

- Let $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \mathscr{P}_n$. Sort $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ in the non-increasing order, say $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n'$.
- Then, $\hat{\mathbf{q}}$ majorizes $\hat{\mathbf{p}}$, i.e., $\hat{\mathbf{p}} \preceq \hat{\mathbf{q}}$ if

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \qquad \forall k \in \{1, \cdots, n\}$$

- For PMFs of different sizes, pad extra zeros to the smaller one.

Prelude

Majorization '$\preceq$'

- Let $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \mathscr{P}_n$. Sort $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ in the non-increasing order, say $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$.
- Then, $\hat{\mathbf{q}}$ majorizes $\hat{\mathbf{p}}$, i.e., $\hat{\mathbf{p}} \preceq \hat{\mathbf{q}}$ if

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \qquad \forall k \in \{1, \cdots, n\}$$

- Majorization is a **partial order** on $\mathscr{P}_n^{'}$.

Binary Relation which is:

- **Reflexive.**
- **Anti-symmetric.**
- **Transitive.**

Prelude

Majorization '$\preceq$'

- Let $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \mathscr{P}_n$. Sort $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ in the non-increasing order, say $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n'$.
- Then, $\hat{\mathbf{q}}$ majorizes $\hat{\mathbf{p}}$, i.e., $\hat{\mathbf{p}} \preceq \hat{\mathbf{q}}$ if

$$\sum_{i=1}^{k} p_i \leq \sum_{i=1}^{k} q_i \qquad \forall k \in \{1, \cdots, n\}$$

- Majorization is a **partial order** on $\mathscr{P}_n'$.
- $(\mathscr{P}_n', \preceq)$ is called a POSET.

Prelude

Majorization Partial Order (glb)

Greatest lower bound w.r.t Majorization, i.e., $\wedge$:

- given any $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$, $\mathbf{p} \wedge \mathbf{q}$ is that PMF (If exists !):
  - $\mathbf{p} \wedge \mathbf{q} \preceq \mathbf{p}$.
  - $\mathbf{p} \wedge \mathbf{q} \preceq \mathbf{q}$.
  - $\forall \mathbf{r} \in \mathscr{P}_n^{'}$ s.t. $\mathbf{r} \preceq \mathbf{p}$ & $\mathbf{r} \preceq \mathbf{q}$,

    we also have: $\mathbf{r} \preceq \mathbf{p} \wedge \mathbf{q}$.

Prelude

Majorization Partial Order (lub)

Least Upper Bound w.r.t Majorization, i.e., $\vee$:

- given any $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$, $\mathbf{p} \vee \mathbf{q}$ is that PMF (If exists !):
  - $\mathbf{p} \preceq \mathbf{p} \vee \mathbf{q}$.
  - $\mathbf{q} \preceq \mathbf{p} \vee \mathbf{q}$.
  - $\forall \mathbf{r} \in \mathscr{P}_n^{'}$ s.t. $\mathbf{p} \preceq \mathbf{r}$ & $\mathbf{q} \preceq \mathbf{r}$,

    we also have $\mathbf{p} \vee \mathbf{q} \preceq \mathbf{r}$.

Prelude

Schur concave / convex functions

- A function f: $\mathscr{P}_n \to \mathbb{R}$ is **Schur convex** if it is **order-preserving,** i.e.,

$$\forall \mathbf{p}, \mathbf{q} \in \mathscr{P}_n \text{ s.t. } \mathbf{p} \preceq \mathbf{q} \implies f(\mathbf{p}) \leq f(\mathbf{q})$$

Prelude

Schur concave / convex functions

- A function f: $\mathscr{P}_n \to \mathbb{R}$ is **Schur convex** if it is **order-preserving**, i.e.,

$$\forall \mathbf{p}, \mathbf{q} \in \mathscr{P}_n \text{ s.t. } \mathbf{p} \preceq \mathbf{q} \implies f(\mathbf{p}) \leq f(\mathbf{q})$$

- Rényi entropy ($H_\alpha(\cdot)$) is a Schur-concave function, for every $\alpha \geq 0$. [MO' 79]

$$\forall \mathbf{p}, \mathbf{q} \in \mathscr{P}_n \text{ s.t. } \mathbf{p} \preceq \mathbf{q} \implies H_\alpha(\mathbf{p}) \geq H_\alpha(\mathbf{q})$$

Marshall, A.W. and Olkin, I. (1979) Inequalities: Theory of Majorization and Its Applications. Mathematics in Science and Engineering, Vol. 143, Academic Press, New York.

Majorization Lattice

Majorization Partial Order is a Lattice !

Theorem [also Bapat '91]:

The POSET $(\mathscr{P}_n^{'}, \preceq)$ with majorization partial order is a Lattice $(\mathscr{P}_n^{'}, \preceq, \vee, \wedge)$

Special class of POSET
s.t. $\vee$ and $\wedge$ exist and are unique.

Bapat, Ravindra B.. "Majorization and singular values. III." Linear Algebra and its Applications 145 (1991): 59-70.

Majorization Lattice

Majorization Partial Order is a Lattice !

Theorem:

The POSET $(\mathscr{P}_n^{'}, \preceq)$ with majorization partial order is a Lattice $(\mathscr{P}_n^{'}, \preceq, \vee, \wedge)$

Proof Idea:

- $\mathbf{p} \wedge \mathbf{q}$ exists in $\mathscr{P}_n^{'}$ for every $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$
- $\mathbf{p} \vee \mathbf{q}$ exists in $\mathscr{P}_n^{'}$ for every $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$

Majorization Lattice

Majorization Partial Order is a Lattice !

Theorem (Extension) [Bapat '91]:

The POSET $(\mathscr{P}_n', \preceq)$ with majorization partial order is a Lattice $(\mathscr{P}_n', \preceq, \vee, \wedge)$.
Indeed, its a complete lattice.

Bapat, Ravindra B.. "Majorization and singular values. III." Linear Algebra and its Applications 145 (1991): 59-70.

Majorization Lattice

Majorization Partial Order is a Complete Lattice !

Theorem (Extension):

The majorization partial order $(\mathscr{P}_n', \preceq)$ is a lattice $(\mathscr{P}_n', \preceq, \vee, \wedge)$. **Indeed its a complete lattice.**

Proof Idea:

- $\wedge \mathscr{Q}$ exists in $\mathscr{P}_n'$ for every $\mathscr{Q} \subseteq \mathscr{P}_n'$
- $\vee \mathscr{Q}$ exists in $\mathscr{P}_n'$ for every $\mathscr{Q} \subseteq \mathscr{P}_n'$

Majorization Lattice

Properties of Entropy on Majorization Lattice — Supermodularity

- A real-valued function f defined on a lattice $(\mathscr{P}, \preceq, \vee, \wedge)$ is called **supermodular** if $\forall a, b \in \mathscr{P}$:

$$f(a \vee b) + f(a \wedge b) \geq f(a) + f(b)$$

Majorization Lattice

Properties of Entropy on Majorization Lattice — Supermodularity

- A real-valued function f defined on a lattice $(\mathscr{P}, \preceq, \vee, \wedge)$ is called **supermodular** if $\forall a, b \in \mathscr{P}$:

$$f(a \vee b) + f(a \wedge b) \geq f(a) + f(b)$$

Theorem:

The Shannon entropy is **supermodular** on the majorization lattice $(\mathscr{P}_n', \preceq, \vee, \wedge)$, i.e., $\forall \mathbf{p}, \mathbf{q} \in \mathscr{P}_n'$

$$H(\mathbf{p} \vee \mathbf{q}) + H(\mathbf{p} \wedge \mathbf{q}) \geq H(\mathbf{p}) + H(\mathbf{q})$$

Majorization Lattice

Properties of Entropy on Majorization Lattice — Subadditivity

- A real-valued function f defined on a lattice ($\mathscr{P}$, $\preceq$, $\vee$, $\wedge$) is called subaddditive if $\forall a, b \in \mathscr{P}$:

$$f(a \vee b) \leq f(a) + f(b) \qquad \text{(w.r.t lub)}$$

$$f(a \wedge b) \leq f(a) + f(b) \qquad \text{(w.r.t glb)}$$

Majorization Lattice

Properties of Entropy on Majorization Lattice — Subadditivity

- A real-valued function f defined on a lattice ($\mathscr{P}, \preceq, \vee, \wedge$) is called subadddititive if $\forall a, b \in \mathscr{P}$:

$$f(a \vee b) \leq f(a) + f(b) \qquad \text{(w.r.t lub)}$$

$$f(a \wedge b) \leq f(a) + f(b) \qquad \text{(w.r.t glb)}$$

Theorem:

The Shannon entropy is subadditive on the majorization lattice ($\mathscr{P}_n^{'}, \preceq, \vee, \wedge$) w.r.t both glb as well as lub i.e., $\forall \mathbf{p}, \mathbf{q} \in \mathscr{P}_n^{'}$

$$H(\mathbf{p} \vee \mathbf{q}) \leq H(\mathbf{p} \wedge \mathbf{q}) \leq H(\mathbf{p}) + H(\mathbf{q})$$

Outline

- **Majorization Lattice** [Cicalese et. al '02]
  - Majorization Partial Order
  - It is a lattice!
  - Properties of Entropy on the Majorization Lattice
- **Applications of Majorization**
  - Lower Bound on Entropy of Random Variables [Sason '18]
  - Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$ [Sason '18]
  - Probability Mass Function Truncation [Cicalese et. al '19]
- **Future Work**

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathcal{X}_n$

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathcal{X}_n$
- Given δ, where $\frac{P_{\text{max}}}{P_{\text{min}}} \leq \delta \in [1,\infty)$

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$
- Given δ, where $\frac{P_{\max}}{P_{\min}} \leq \delta \in [1,\infty)$
- Comments on **Rényi entropy** of X

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$.
- Given δ, where $\frac{P_{\text{max}}}{P_{\text{min}}} \leq \delta \in [1, \infty)$.
- Comments on **Rényi entropy** of X

$$\mathscr{P}_n(\delta) := \left\{ (p_1, p_2, \cdots, p_n) \in \mathscr{P}_n : \frac{p_{max}}{p_{min}} \leq \delta \right\}$$

Similarly, $\mathscr{P}_n^{'}(\delta)$

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$.
- Given δ, where $\frac{P_{\max}}{P_{\min}} \leq \delta \in [1,\infty)$.
- Comments on **Rényi entropy** of X

Upper Bound: $\max_{\mathbf{p}\in\mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$

Lower Bound: $\min_{\mathbf{p}\in\mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$.
- Given δ, where $\frac{P_{\text{max}}}{P_{\text{min}}} \leq \delta \in [1, \infty)$.
- Comments on **Rényi entropy** of X

Upper Bound: $\max\limits_{\mathbf{p} \in \mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$ = log n

Lower Bound: $\min\limits_{\mathbf{p} \in \mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$

Applications

Bounds on Entropy of Random Variables

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$.
- Given δ, where $\frac{P_{\text{max}}}{P_{\text{min}}} \leq \delta \in [1,\infty)$.
- Comments on **Rényi entropy** of X

Upper Bound: $\max_{\mathbf{p}\in\mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$ = log n

Lower Bound: $\min_{\mathbf{p}\in\mathscr{P}_n(\delta)} H_\alpha(\mathbf{p})$

Applications

Bounds on Entropy of Random Variables

The Solution (open-form expression):

$$\min_{\mathbf{p}\in\mathcal{P}_n(\delta)} H_\alpha(\mathbf{p}) = \min_{\beta\in\Gamma_n^\delta} H_\alpha(\mathbf{q}_\beta)$$

$$\Gamma_n^\delta := \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$$

- Where $\mathbf{q}_\beta \in \mathscr{P}_n^{'}(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1, \cdots, i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2, \cdots, n\} \end{cases} \quad \text{and} \quad i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$

Applications

Bounds on Entropy of Random Variables

<u>**The Solution:**</u>

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$

Applications

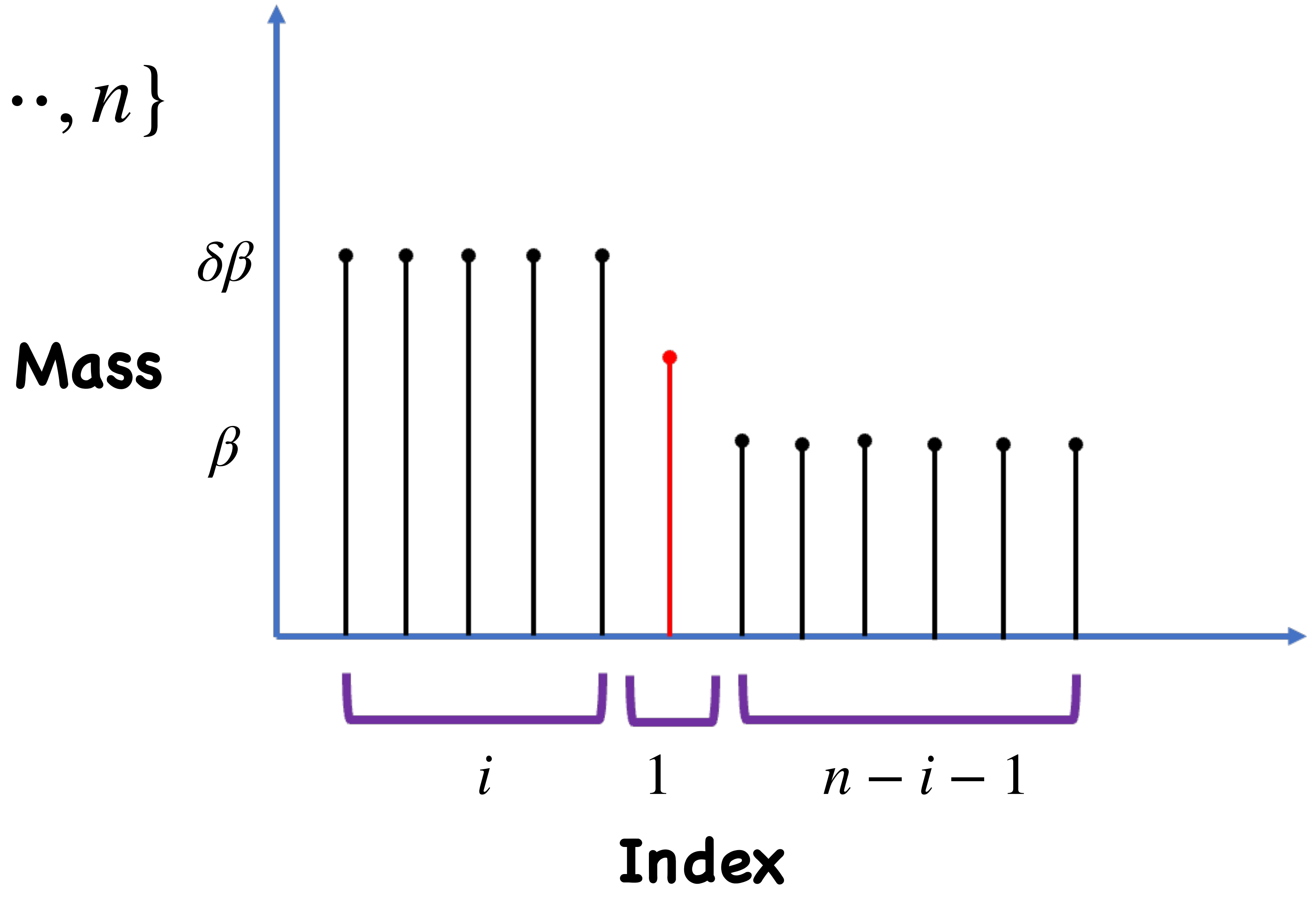
Bounds on Entropy of Random Variables

The Solution:

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\text{min}} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$
- Define a PMF $\mathbf{q}_\beta \in \mathscr{P}_n(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1, \cdots, i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2, \cdots, n\} \end{cases}$$

$$i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$



Applications

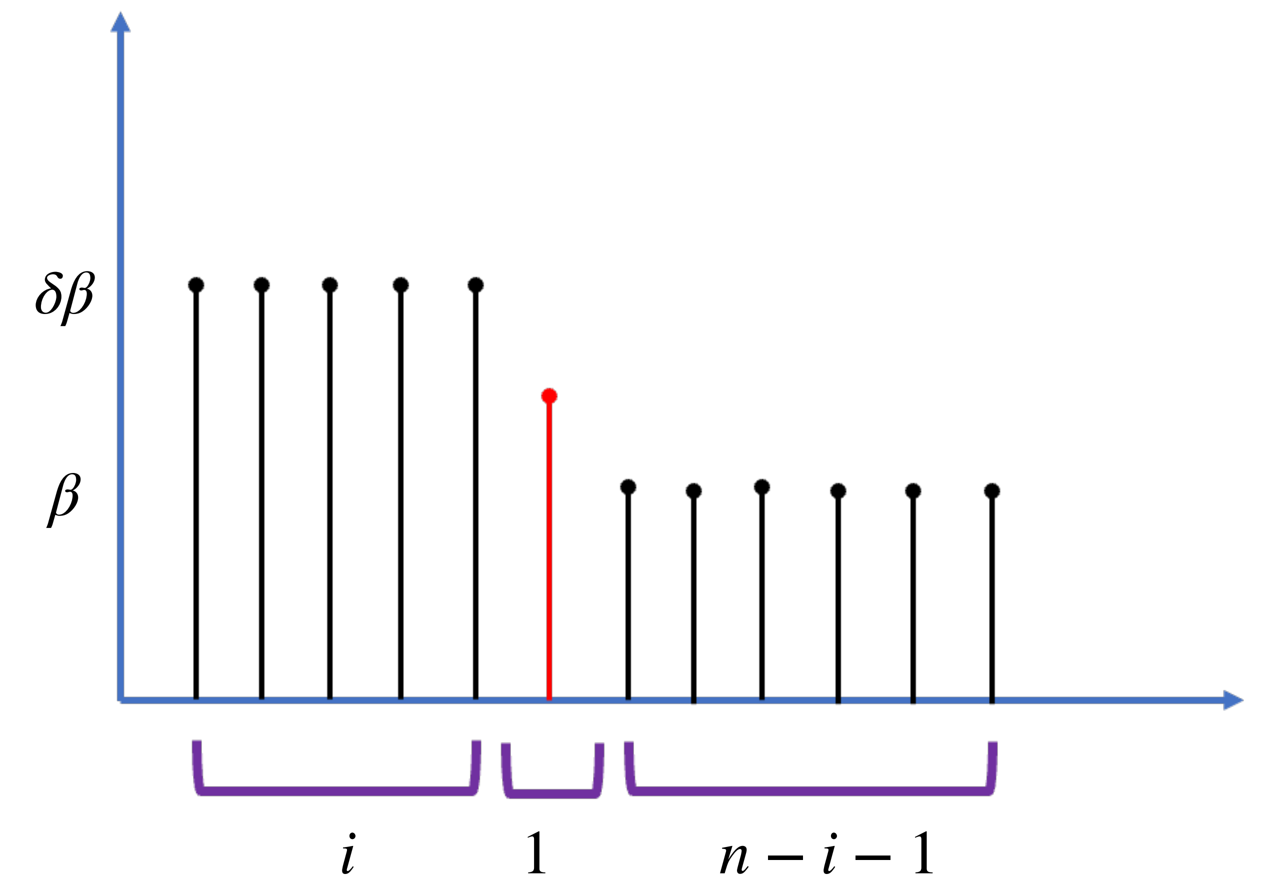
Bounds on Entropy of Random Variables

The Solution:

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$.
- Define a PMF $\mathbf{q}_\beta \in \mathscr{P}_n(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1, \cdots, i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2, \cdots, n\} \end{cases} \quad \text{and} \quad i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$

- $\mathbf{p} \preceq \mathbf{q}_\beta$



Applications

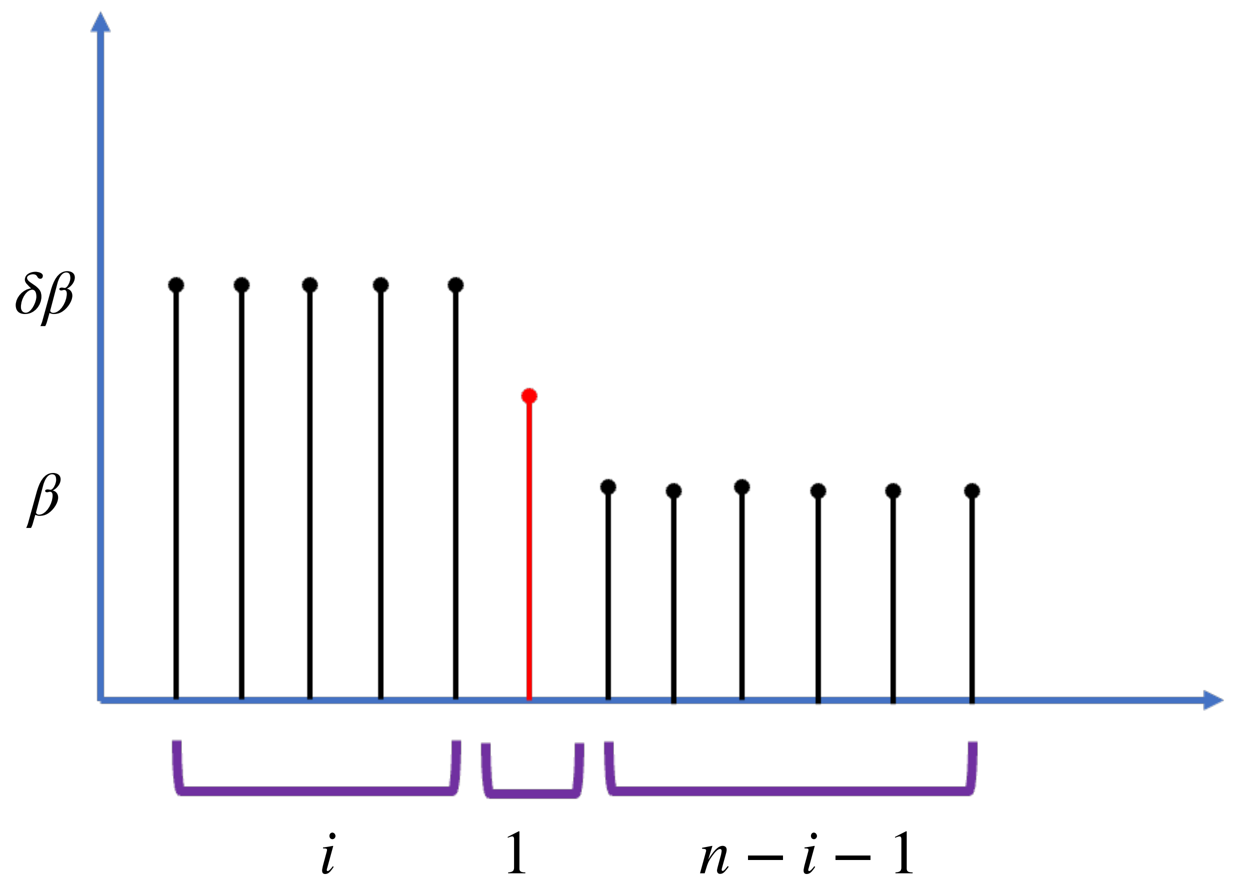
Bounds on Entropy of Random Variables

The Solution:

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\text{min}} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$.
- Define a PMF $\mathbf{q}_\beta \in \mathscr{P}_n(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1, \cdots, i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2, \cdots, n\} \end{cases} \quad \text{and} \quad i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$

- $\mathbf{p} \preceq \mathbf{q}_\beta$
- $\mathbf{p} \preceq \mathbf{q}_\beta$ for every $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\text{min}} := \beta$



Applications

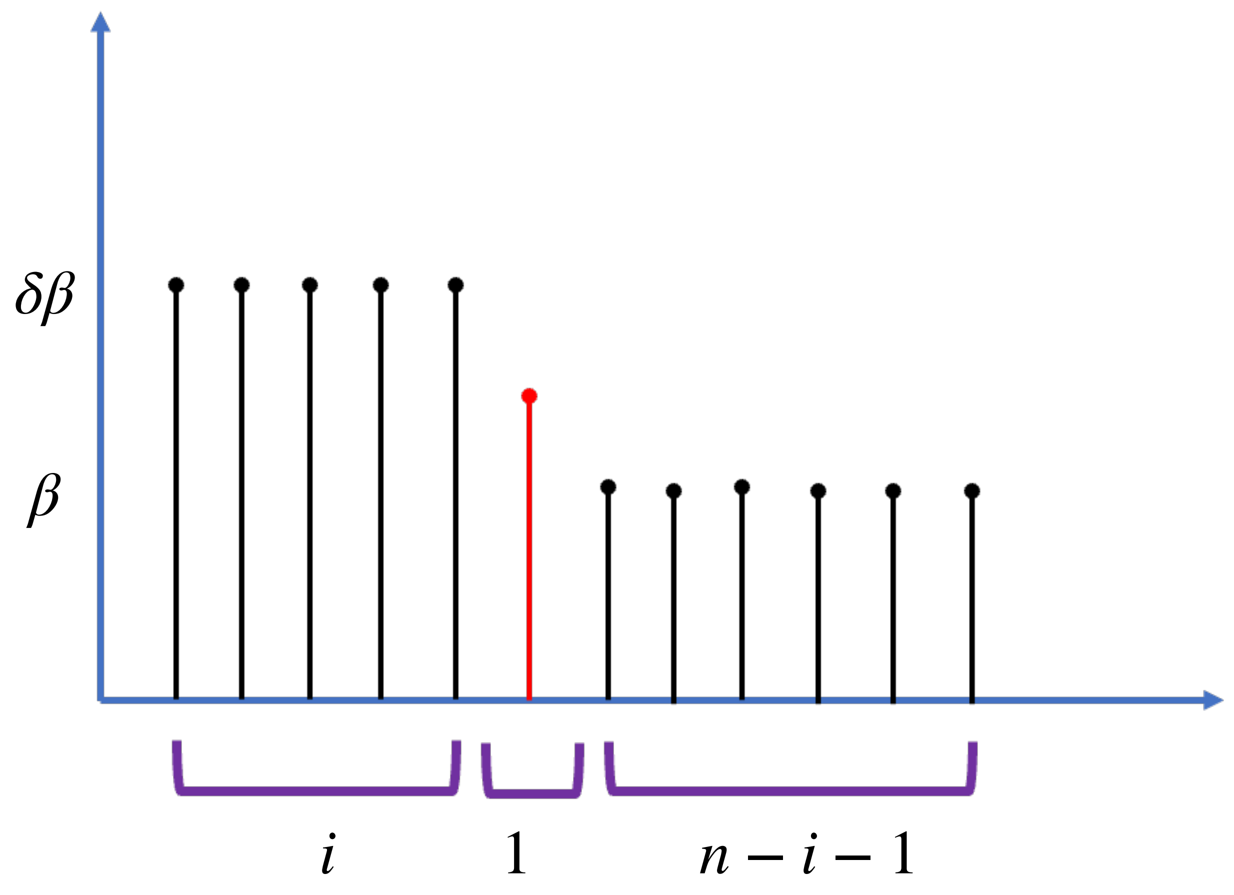
Bounds on Entropy of Random Variables

The Solution:

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$.
- Define a PMF $\mathbf{q}_\beta \in \mathscr{P}_n(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1,\cdots,i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2,\cdots,n\} \end{cases} \quad \text{and} \quad i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$

- $\mathbf{p} \preceq \mathbf{q}_\beta$
- $\mathbf{p} \preceq \mathbf{q}_\beta$ for every $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta$
- $\min_{\substack{\mathbf{p}\in\mathcal{P}_n(\delta) \\ \text{s.t.} \\ p_{\min}=\beta}} H_\alpha(\mathbf{p}) = H(\mathbf{q}_\beta)$



Applications

Bounds on Entropy of Random Variables

The Solution:

- Given a $\delta > 1$. Fix a $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta \in \left[\frac{1}{1+(n-1)\delta}, \frac{1}{n}\right]$ — Γ_n^δ
- Define a PMF $\mathbf{q}_\beta \in \mathscr{P}_n(\delta)$ such that:

$$q_j = \begin{cases} \delta\beta & j \in \{1,\cdots,i\} \\ 1-(n+i\delta-i-1)\beta & j = i+1 \\ \beta & j \in \{i+2,\cdots,n\} \end{cases} \quad \text{and} \quad i := \left\lfloor \frac{1-n\beta}{(\delta-1)\beta} \right\rfloor$$

- $\mathbf{p} \preceq \mathbf{q}_\beta$
- $\mathbf{p} \preceq \mathbf{q}_\beta$ for every $\mathbf{p} \in \mathscr{P}_n(\delta)$ with $p_{\min} := \beta$
- $\min_{\substack{\mathbf{p}\in\mathcal{P}_n(\delta) \\ \text{s.t.} \\ p_{\min}=\beta}} H_\alpha(\mathbf{p}) = H(\mathbf{q}_\beta)$ $\Longrightarrow$ $\min_{\mathbf{p}\in\mathcal{P}_n(\delta)} H_\alpha(\mathbf{p}) = \min_{\beta\in\Gamma_n^\delta} H_\alpha(\mathbf{q}_\beta)$

Applications

Bounds on Entropy of Random Variables

The Solution (closed form expression):

- For every $\mathbf{p} \in \mathscr{P}_n(\delta)$, for $\alpha > 0$ and $\delta > 1$, we have

$$\min_{\mathbf{p} \in \mathscr{P}_n(\delta)} H_\alpha(\mathbf{p}) = \min_{\beta \in \Gamma_n^\delta} H_\alpha(\mathbf{q}_\beta)$$
$$\geq \log n - c_\alpha^\infty(\delta)$$

$$c_\alpha^\infty(\delta) = \frac{1}{\alpha - 1} \log\left(1 + \frac{1 + \alpha(\delta - 1) - \delta^\alpha}{(1-\alpha)(\delta - 1)}\right) - \frac{\alpha}{\alpha - 1} \log\left(1 + \frac{1 + \alpha(\delta - 1) - \delta^\alpha}{(1-\alpha)(\delta^\alpha - 1)}\right)$$

Where $c_\alpha^\infty(\delta) \leq \log(\delta)$

Outline

- **Majorization Lattice** [Cicalese et. al '02]
  - Majorization Partial Order
  - It is a lattice!
  - Properties of Entropy on the Majorization Lattice
- **Applications of Majorization**
  - Lower Bound on Entropy of Random Variables [Sason '18]
  - Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$ [Sason '18]
  - Probability Mass Function Truncation [Cicalese et. al '19]
- **Future Work**

Applications

Upper and Lower Bounds on $H_\alpha(f(X))$

The Problem:

- Given a discrete random variable $X \in \mathcal{X}_n$ with PMF $\mathbf{p}$

Applications

Upper and Lower Bounds on $H_\alpha(f(X))$

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$ with PMF $\mathbf{p}$
- Let f be a **deterministic** and **surjective** function s.t. $f : \mathscr{X}_n \to \mathscr{Y}_m$ $(m < n)$

Applications

Upper and Lower Bounds on $H_\alpha(f(X))$

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$ with PMF $\mathbf{p}$
- Let f be a **deterministic** and **surjective** function s.t. $f : \mathscr{X}_n \to \mathscr{Y}_m$ $(m < n)$
- $\mathscr{F}_m$: set of all such functions

Applications

Upper and Lower Bounds on $H_\alpha(f(X))$

The Problem:

- Given a discrete random variable $X \in \mathscr{X}_n$ with PMF $\mathbf{p}$
- Let f be a **deterministic** and **surjective** function s.t. $f : \mathscr{X}_n \to \mathscr{Y}_m$ $(m < n)$
- $\mathscr{F}_m$: set of all such functions
- Comments on **Rényi entropy** of $f(X)$, i.e., $H_\alpha(f(X))$

Applications

Upper and Lower Bounds on $H_\alpha(f(X))$

The Problem:

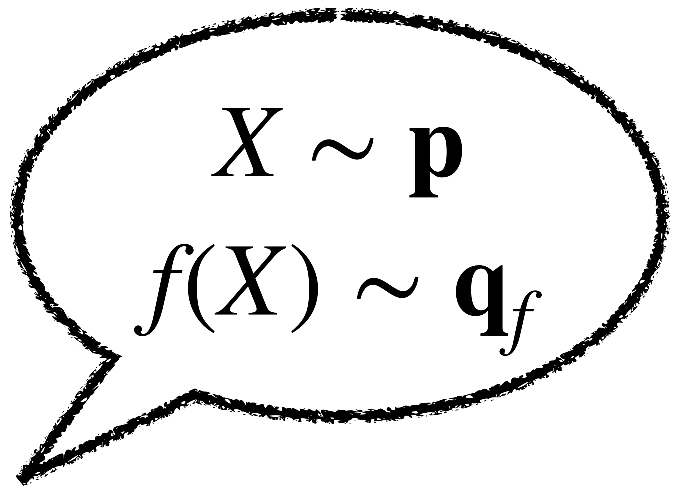
- Given a discrete random variable $X \in \mathcal{X}_n$ with PMF $\mathbf{p}$
- Let f be a **deterministic** and **surjective** function s.t. $f : \mathcal{X}_n \to \mathcal{Y}_m$ $(m < n)$
- $\mathcal{F}_m$: set of all such functions
- Comments on **Rényi entropy** of $f(X)$, i.e., $H_\alpha(f(X))$

Upper Bound: $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

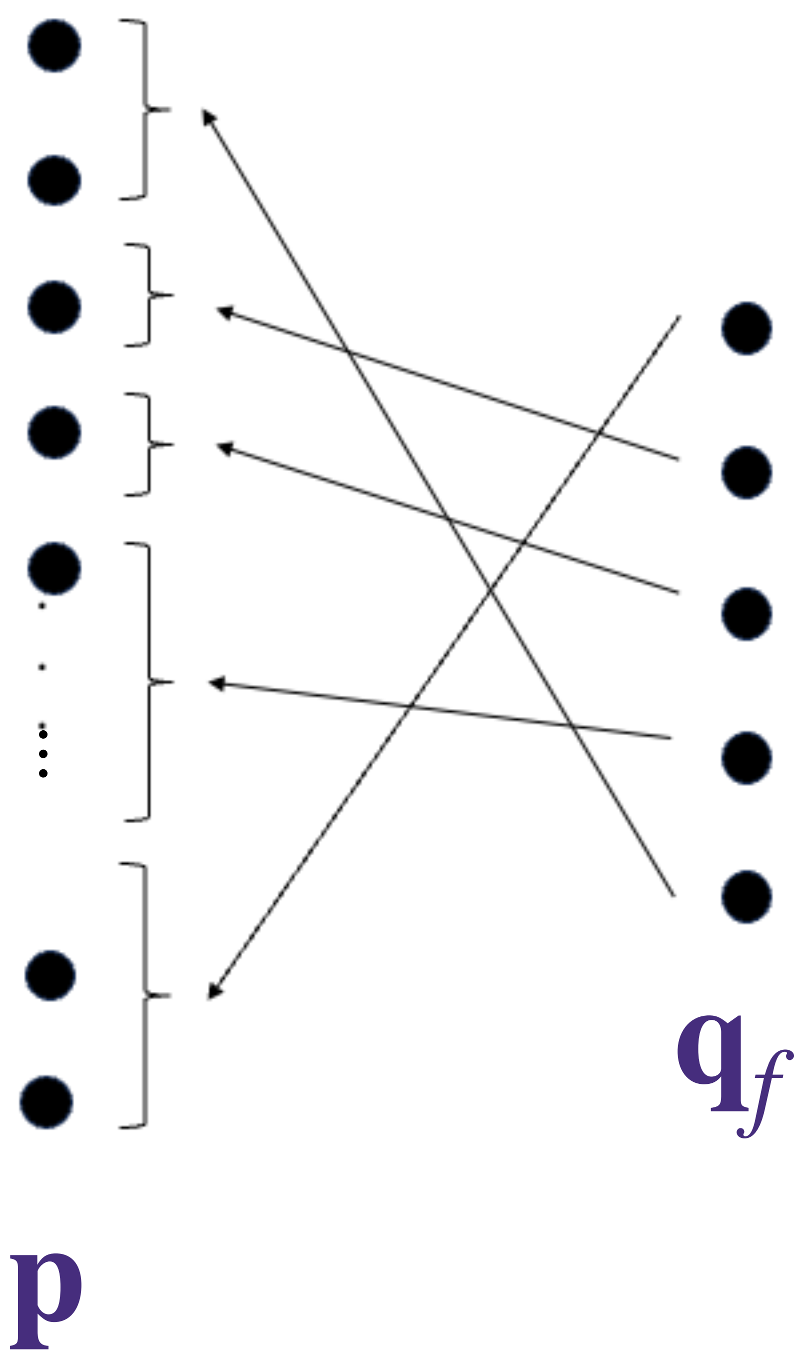
Lower Bound: $\min_{f \in \mathcal{F}_m} H_\alpha(f(X))$

Applications

Proving $H_\alpha(f(X)) \leq H_\alpha(X)$



- For every $f \in \mathscr{F}_m$, $\mathbf{q}_f$ is an aggregation of $\mathbf{p}$, i.e., $\mathbf{p} \sqsubseteq \mathbf{q}_f$



Applications

Proving $H_\alpha(f(X)) \leq H_\alpha(X)$

$X \sim \mathbf{p}$
$f(X) \sim \mathbf{q}_f$

- For every $f \in \mathscr{F}_m$, $\mathbf{q}_f$ is an aggregation of $\mathbf{p}$, i.e., $\mathbf{p} \sqsubseteq \mathbf{q}_f$
- Aggregation implies majorization, i.e., $\mathbf{p} \preceq \mathbf{q}_f$ [Cicalese et. al '17]

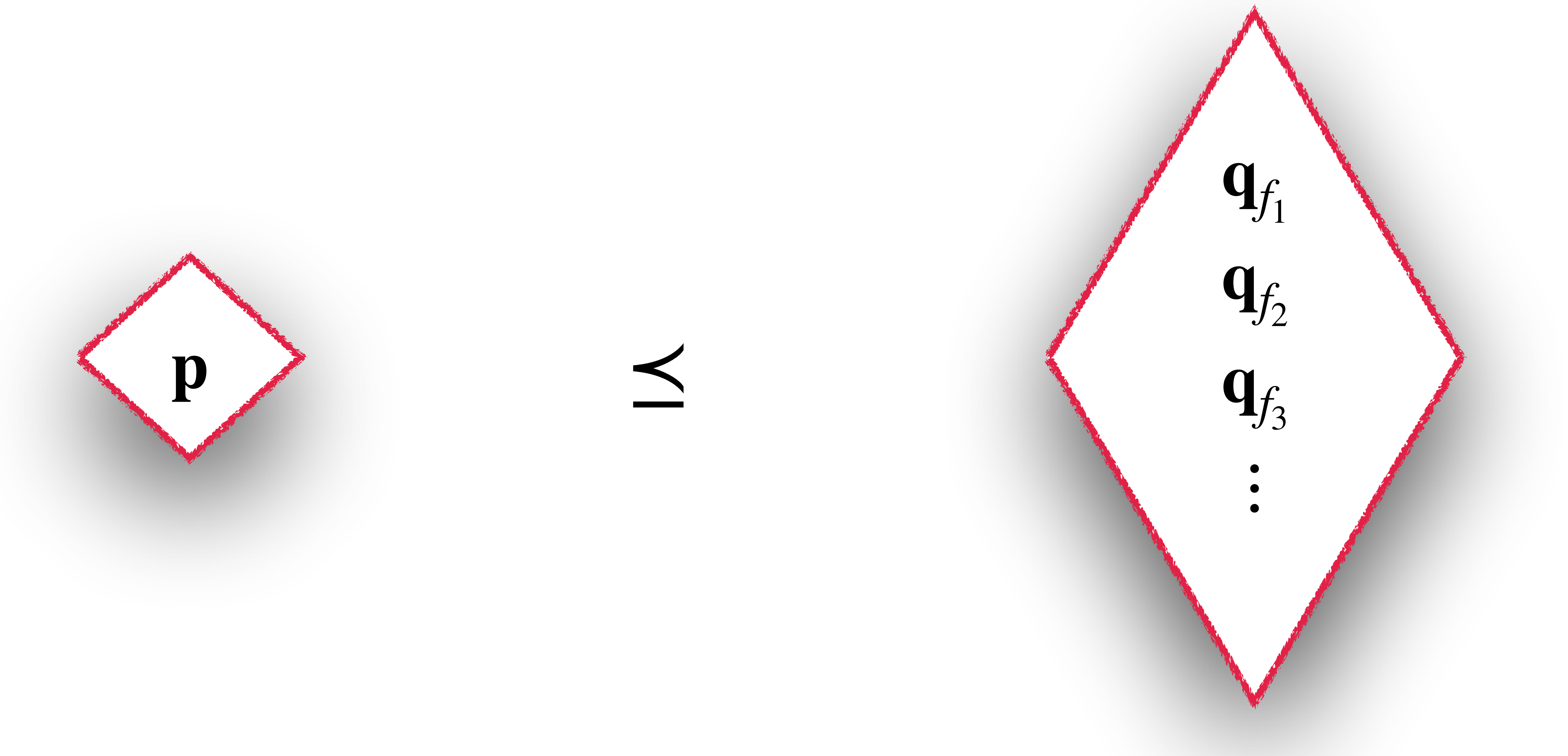
Cicalese, Ferdinando et al. "Bounds on the Entropy of a Function of a Random Variable and Their Applications." IEEE Transactions on Information Theory 64 (2017): 2220-2230.

Applications

Proving $H_\alpha(f(X)) \leq H_\alpha(X)$

$X \sim \mathbf{p}$
$f(X) \sim \mathbf{q}_f$

- For every $f \in \mathscr{F}_{m'}$, $\mathbf{q}_f$ is an aggregation of $\mathbf{p}$, i.e., $\mathbf{p} \sqsubseteq \mathbf{q}_f$
- Aggregation implies majorization, i.e., $\mathbf{p} \preceq \mathbf{q}_f$

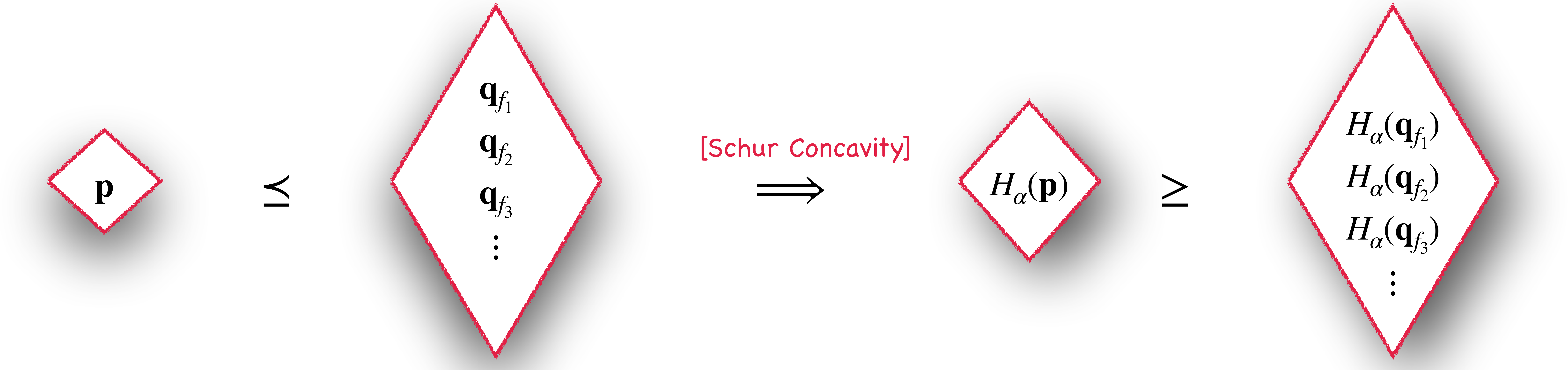


Applications

Proving $H_\alpha(f(X)) \leq H_\alpha(X)$

$X \sim \mathbf{p}$
$f(X) \sim \mathbf{q}_f$

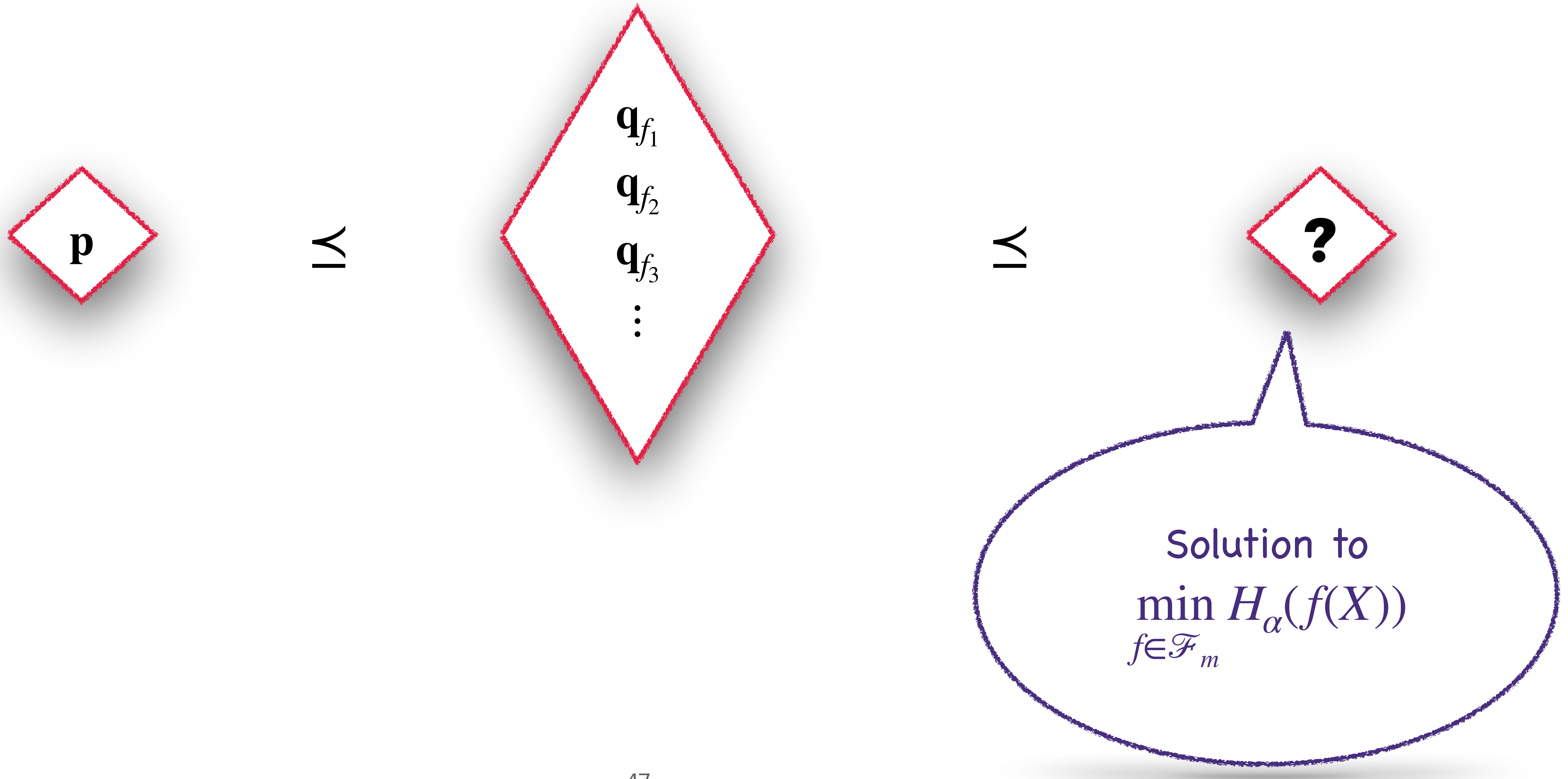
- For every $f \in \mathscr{F}_m$, $\mathbf{q}_f$ is an aggregation of $\mathbf{p}$, i.e., $\mathbf{p} \sqsubseteq \mathbf{q}_f$
- Aggregation implies majorization, i.e., $\mathbf{p} \preceq \mathbf{q}_f$



Applications

Lower Bound on $H_\alpha(f(X))$

Approach for Lower bound:



Applications

Lower Bound on $H_\alpha(f(X))$

The Solution for : $\min_{f \in \mathscr{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathscr{P}_n$

Applications

Lower Bound on $H_\alpha(f(X))$

The Solution for : $\min_{f \in \mathscr{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathscr{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathscr{P}_n'$

Applications

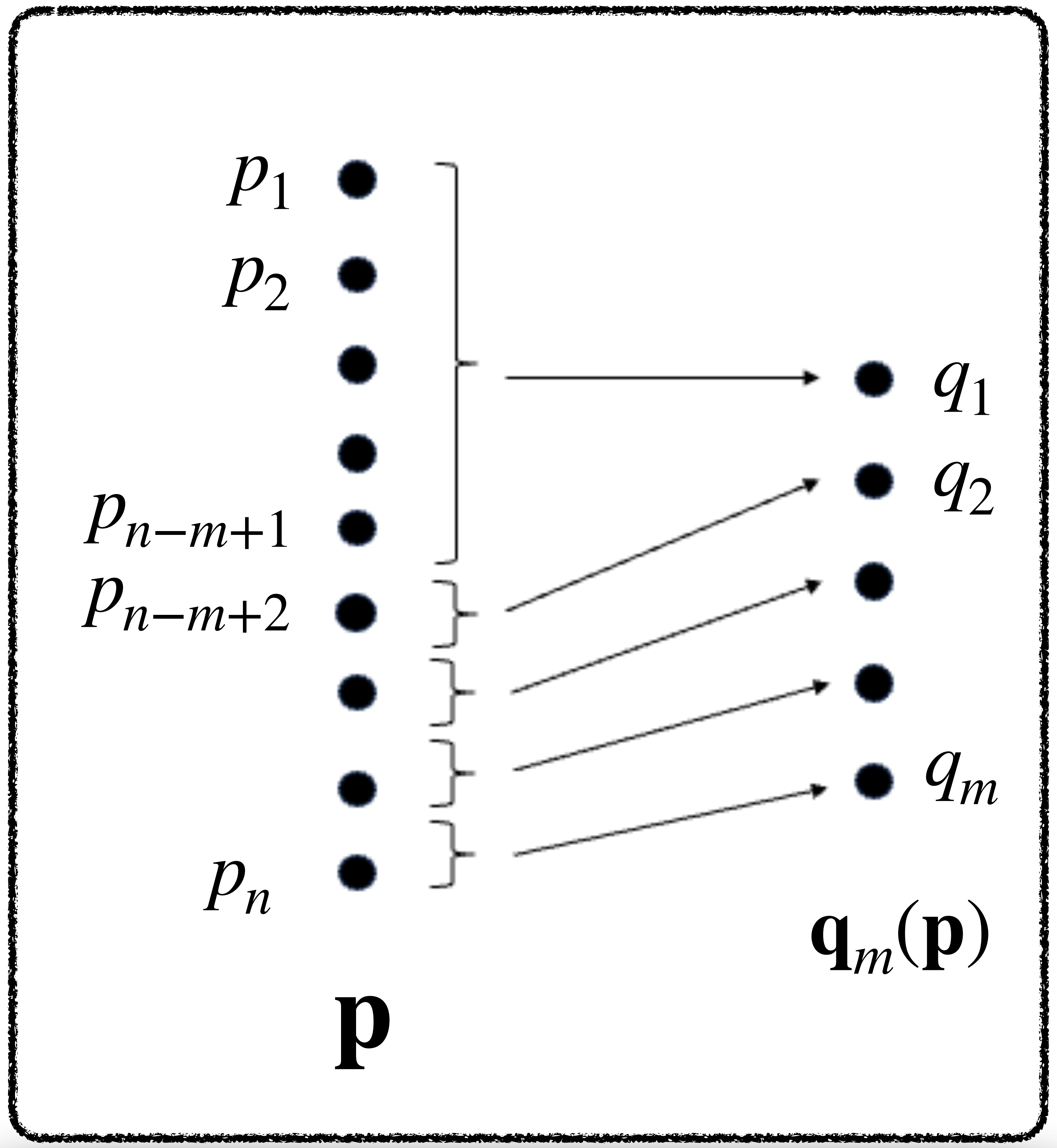
Lower Bound on $H_\alpha(f(X))$

<u>The Solution for</u> : $\min_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathcal{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathcal{P}'_n$
- Define $\mathbf{q}_m(\mathbf{p})$ as:

$$q_i = \begin{cases} \sum_{k=1}^{n-m+1} p_k & i = 1 \\ p_{n-m+i} & i = 2,3,\cdots,m \end{cases}$$

$(m-1)$



Applications

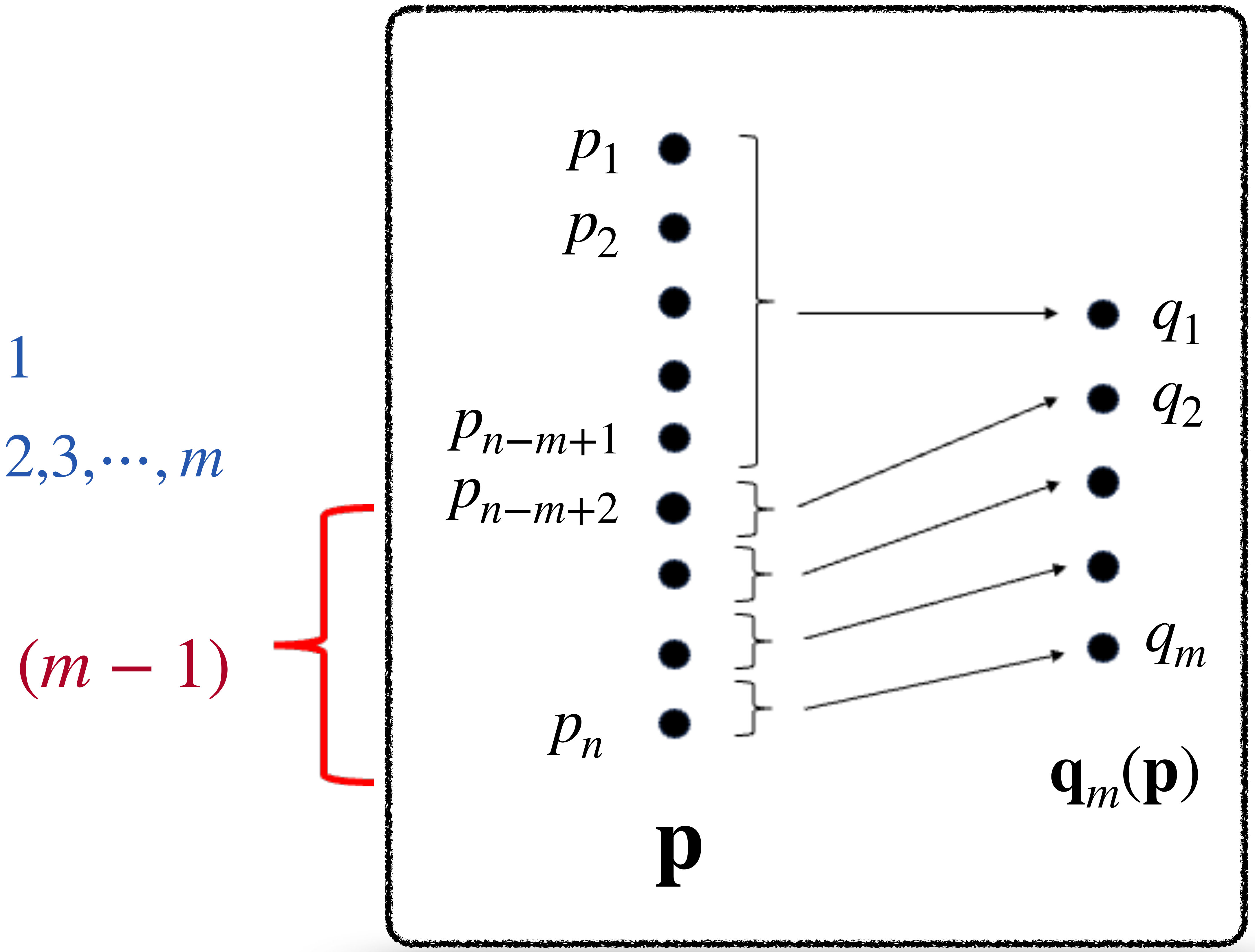
Lower Bound on $H_\alpha(f(X))$

<u>**The Solution for**</u> : $\min_{f \in \mathscr{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathscr{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathscr{P}'_n$
- Define $\mathbf{q}_m(\mathbf{p})$ as:

$$q_i = \begin{cases} \sum_{k=1}^{n-m+1} p_k & i = 1 \\ p_{n-m+i} & i = 2,3,\cdots,m \end{cases}$$

- For every $f \in \mathscr{F}_m$, $\mathbf{q}_f \preceq \mathbf{q}_m(\mathbf{p})$.



Applications

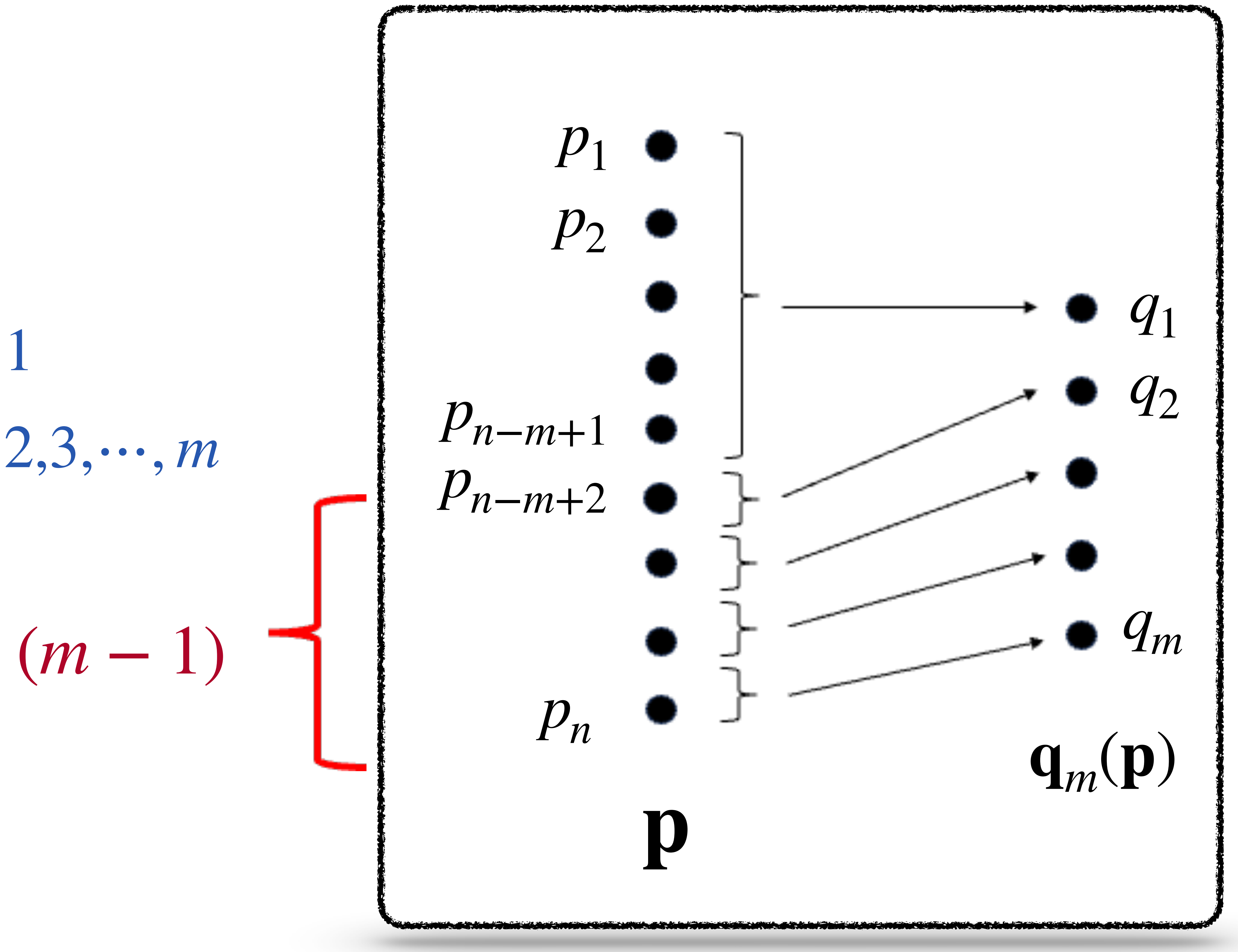
Lower Bound on $H_\alpha(f(X))$

The Solution for : $\min_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathcal{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathcal{P}'_n$
- Define $\mathbf{q}_m(\mathbf{p})$ as:

$$q_i = \begin{cases} \sum_{k=1}^{n-m+1} p_k & i = 1 \\ p_{n-m+i} & i = 2,3,\cdots,m \end{cases}$$

- For every $f \in \mathcal{F}_m$, $\mathbf{q}_f \preceq \mathbf{q}_m(\mathbf{p})$
- $\mathbf{q}_m(\mathbf{p})$ is an aggregation of $\mathbf{p}$



Applications

Lower Bound on $H_\alpha(f(X))$

<u>**The Solution for**</u> : $\min_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathcal{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathcal{P}'_n$
- Define $\mathbf{q}_m(\mathbf{p})$ as:

$$q_i = \begin{cases} \sum_{k=1}^{n-m+1} p_k & i = 1 \\ p_{n-m+i} & i = 2,3,\cdots,m \end{cases}$$

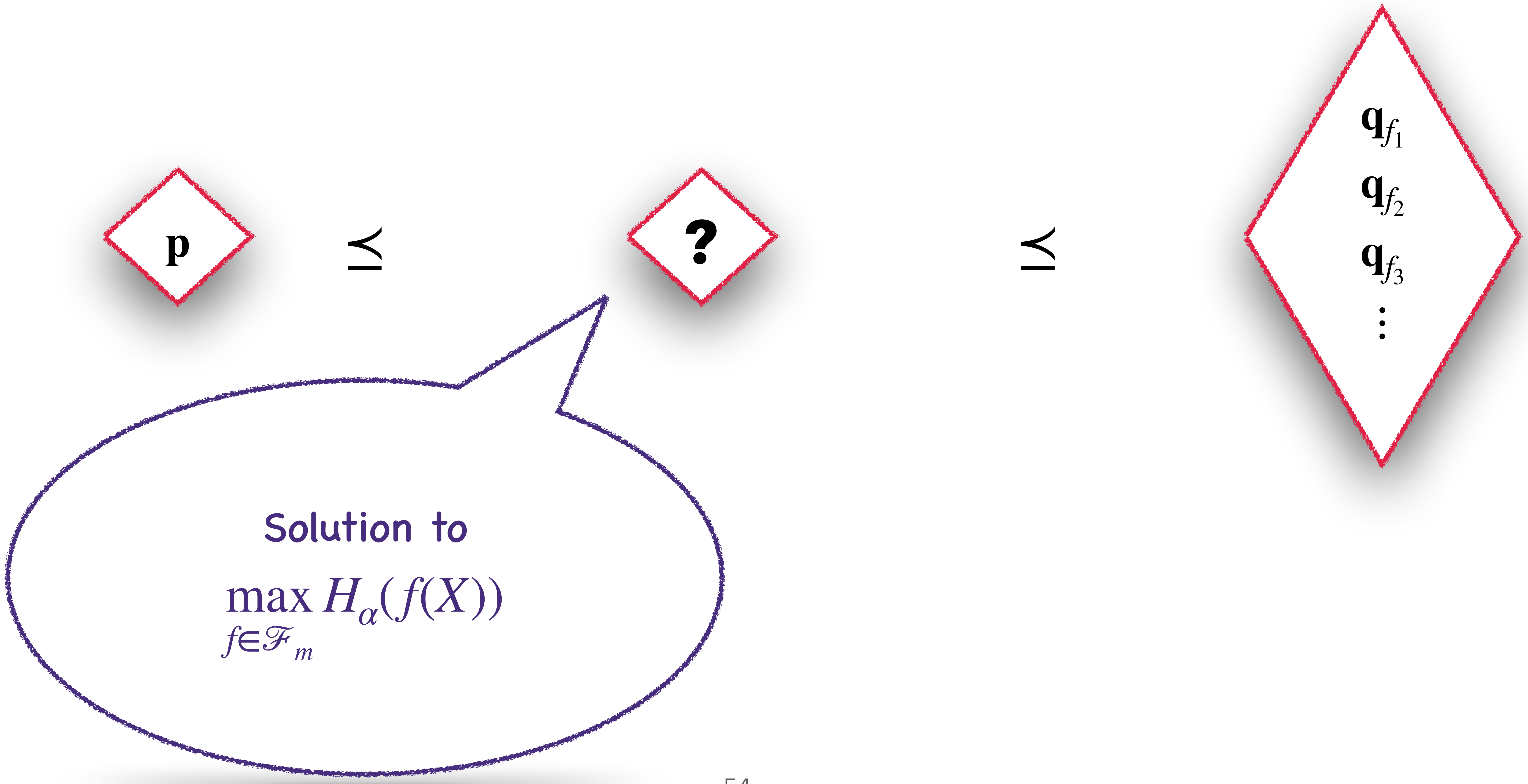
- For every $f \in \mathcal{F}_m$, $\mathbf{q}_f \preceq \mathbf{q}_m(\mathbf{p})$
- $\mathbf{q}_m(\mathbf{p})$ is an aggregation of $\mathbf{p}$ $\Longrightarrow$

$$\min_{f \in \mathcal{F}_m} H_\alpha(f(X)) = H_\alpha(\mathbf{q}_m(\mathbf{p}))$$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

Approach for Upper bound:



Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Finding $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$ is NP-Hard [Cicalese et. al '17]

Cicalese, Ferdinando et al. "Bounds on the Entropy of a Function of a Random Variable and Their Applications." IEEE Transactions on Information Theory 64 (2017): 2220-2230.

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Finding $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$ is NP-Hard
- Upper bound and Lower bound on $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Finding $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$ is NP-Hard
- Upper bound and Lower bound on $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

Upper Bound

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathcal{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathcal{P}_n^{'}$

Applications

Strengthening $H_\alpha(f(X)) \le H_\alpha(X)$

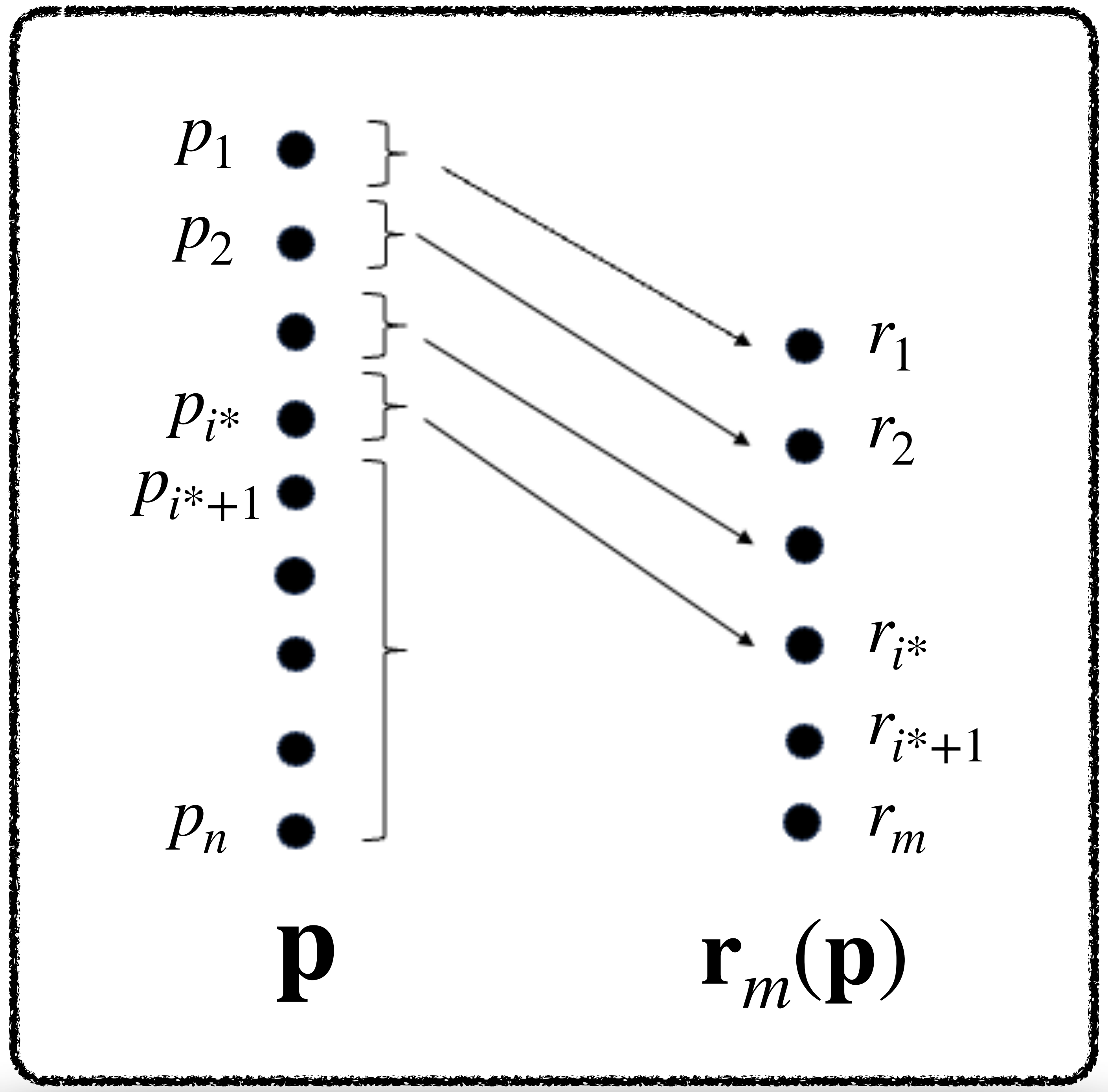
<u>**The Solution for**</u> **:** $\max_{f \in \mathscr{F}_m} H_\alpha(f(X))$

- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathscr{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathscr{P}_n'$
- Define $\mathbf{r}_m(\mathbf{p}) \in \mathscr{P}_m'$ as:
  - If $p_1 < 1/m$: $\mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m).$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathscr{F}_m} H_\alpha(f(X))$



- Given the PMF of $X \sim \hat{\mathbf{p}} \in \mathscr{P}_n$
- Sort $\hat{\mathbf{p}}$ in non-increasing order, say $\mathbf{p} \in \mathscr{P}'_n$
- Define $\mathbf{r}_m(\mathbf{p}) \in \mathscr{P}'_m$ as:
  - If $p_1 < 1/m$: $\mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m)$.
  - If $p_1 \geq 1/m$:

$$r_i = \begin{cases} p_i & i = 1, 2, \cdots, i^* \\ \left(\sum_{j=i^*+1}^{n} p_j\right)/(m - i^*) & i = i^* + 1, \cdots, m \end{cases}$$

where i^* is the maximum index i such that $p_i \geq \dfrac{\sum_{j=i+1}^{n} p_j}{m - i}$.

Applications

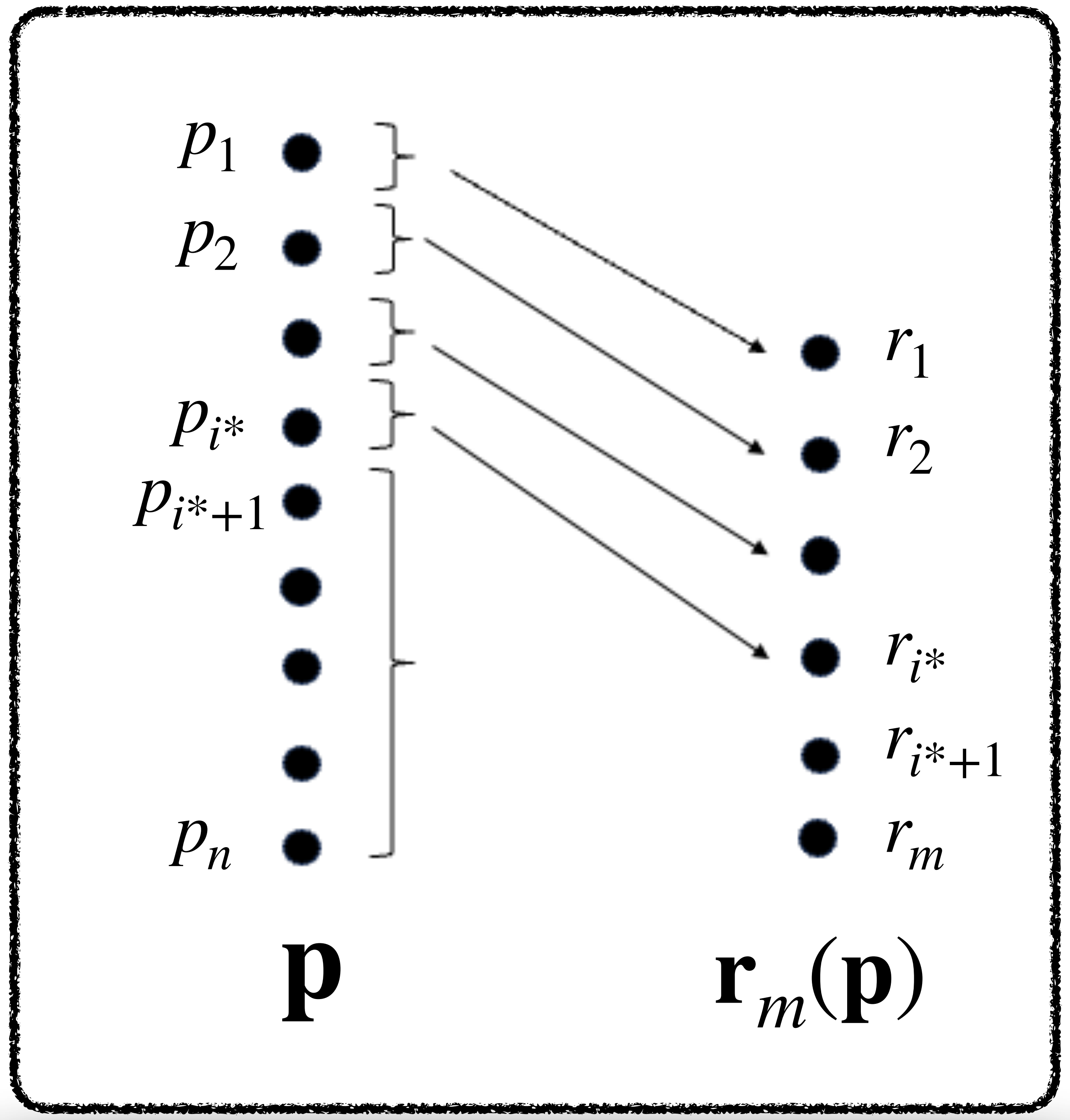
Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Define $\mathbf{r}_m(\mathbf{p}) \in \mathscr{P}_m$ as:
  - If $p_1 < 1/m$: $\quad \mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m).$
  - If $p_1 \geq 1/m$:

$$r_i = \begin{cases} p_i & i = 1,2,\cdots,i^* \\ \left(\sum_{j=i^*+1}^{n} p_j\right)/(m - i^*) & i = i^*+1,\cdots,m \end{cases}$$

where i^* is the maximum index i such that $p_i \geq \dfrac{\sum_{j=i+1}^{n} p_j}{m-i}$.



- Closest to $\mathbf{p}$ than any other $\mathbf{r} \in \mathscr{P}'_m$, w.r.t majorization

$$\mathbf{p} \preceq \mathbf{r}_m(\mathbf{p}) \preceq \mathbf{r} \in \mathscr{P}'_m$$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Define $\mathbf{r}_m(\mathbf{p}) \in \mathcal{P}_m$ as:
  - If $p_1 < 1/m$: $\mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m)$.
  - If $p_1 \geq 1/m$:

$$r_i = \begin{cases} p_i & i = 1, 2, \cdots, i^* \\ \left(\sum_{j=i^*+1}^{n} p_j\right)/(m - i^*) & i = i^* + 1, \cdots, m \end{cases}$$

where i^* is the maximum index i such that $p_i \geq \dfrac{\sum_{j=i+1}^{n} p_j}{m - i}$.

- Closest to $\mathbf{p}$ than any other $\mathbf{r} \in \mathcal{P}'_m$, w.r.t majorization

$$\mathbf{p} \preceq \mathbf{r}_m(\mathbf{p}) \preceq \mathbf{r} \in \mathcal{P}'_m \implies \max_{f \in \mathcal{F}_m} H_\alpha(f(X)) \leq H_\alpha(\mathbf{r}_m(\mathbf{p}))$$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$

- Define $\mathbf{r}_m(\mathbf{p}) \in \mathcal{P}_m$ as:
  - If $p_1 < 1/m$: $\mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m)$.
  - If $p_1 \geq 1/m$:

$$r_i = \begin{cases} p_i & i = 1,2,\cdots,i^* \\ \left(\sum_{j=i^*+1}^{n} p_j\right)/(m-i^*) & i = i^*+1,\cdots,m \end{cases}$$

where i^* is the maximum index i such that $p_i \geq \dfrac{\sum_{j=i+1}^{n} p_j}{m-i}$.

$$\mathbf{r}_m(\mathbf{p}) := \text{argmax}_{\mathbf{r} \in \mathcal{P}_m \text{ and } \mathbf{p} \preceq \mathbf{r}} H_\alpha(\mathbf{r})$$

- Closest to $\mathbf{p}$ than any other $\mathbf{r} \in \mathcal{P}'_m$, w.r.t majorization

$$\mathbf{p} \preceq \mathbf{r}_m(\mathbf{p}) \preceq \mathbf{r} \in \mathcal{P}'_m \implies \max_{f \in \mathcal{F}_m} H_\alpha(f(X)) \leq H_\alpha(\mathbf{r}_m(\mathbf{p}))$$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for : $\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$ **Lower Bound**

- Construct f^* via Huffman algorithm such that $f^*(X) \sim q$
- We have,

$$H_\alpha(\mathbf{q}) \geq H_\alpha(\mathbf{r}_m(\mathbf{p})) - c_\alpha^\infty(2)$$

$$c_\alpha^\infty(2) = \log\left(\frac{\alpha - 1}{2^\alpha - 2}\right) - \frac{\alpha}{\alpha - 1}\log\left(\frac{\alpha}{2^\alpha - 1}\right) \leq 1$$

Applications

Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$

The Solution for :

$$\max_{f \in \mathcal{F}_m} H_\alpha(f(X))$$

Upper Bound

$$\max_{f \in \mathcal{F}_m} H_\alpha(f(X)) \leq H_\alpha(\mathbf{r}_m(\mathbf{p}))$$

Lower Bound

$$H_\alpha(f^*) \geq H_\alpha(\mathbf{r}_m(\mathbf{p})) - c_\alpha^\infty(2)$$

$$\max_{f \in \mathcal{F}_m} H_\alpha(f(X)) \in \left[H_\alpha(\mathbf{r}_m(\mathbf{p})) - c_\alpha^\infty(2), H_\alpha(\mathbf{r}_m(\mathbf{p})) \right]$$

Outline

- **Majorization Lattice** [Cicalese et. al '02]
  - Majorization Partial Order
  - It is a lattice!
  - Properties of Entropy on the Majorization Lattice
- **Applications of Majorization**
  - Lower Bound on Entropy of Random Variables [Sason '18]
  - Strengthening $H_\alpha(f(X)) \leq H_\alpha(X)$ [Sason '18]
  - Probability Mass Function Truncation [Cicalese et. al '19]
- **Future Work**

Applications

Probability Mass Function Truncation

- Let $X \sim \mathbf{p} := (p_1, p_2, \cdots, p_n)$ be a discrete random variable s.t. $X \in \mathscr{X}_n$
- Restrict $X \in \mathscr{Y}_m \subset \mathscr{X}_n$
- Resulting conditional PMF, say $\mathbf{q} := (q_1, q_2, \cdots, q_m)$, is **Truncated PMF** of $\mathbf{p}$

Applications

Probability Mass Function Truncation

- Let $X \sim \mathbf{p} := (p_1, p_2, \cdots, p_n)$ be a random variable s.t. $X \in \mathscr{X}_n$
- Restrict $X \in \mathscr{Y}_m \subset \mathscr{X}_n$
- Resulting conditional PMF, say $\mathbf{q} := (q_1, q_2, \cdots, q_m)$, is **Truncated PMF** of $\mathbf{p}$

- Explore criterions to truncate a PMF,
- **Condition: Truncated PMF and Original PMF are close !**

Applications

Common Examples of PMF Truncation

- Let $\mathbf{p} := (p_1, p_2, \cdots, p_n)$ denote the Original PMF
- $\mathbf{q} := (q_1, q_2, \cdots, q_m)$, where $m < n$, denote the truncated PMF

Applications

Common Examples of PMF Truncation

- Let $\mathbf{p} := (p_1, p_2, \cdots, p_n)$ denote the Original PMF
- And $\mathbf{q} := (q_1, q_2, \cdots, q_m)$, where $m < n$, denote the truncated PMF

Operator $\mathbf{t}_m$

$$\mathbf{t}_m(\mathbf{p}) := (q_1, \cdots, q_m) = \left(\frac{p_1}{\sum_{i=1}^{m} p_i}, \cdots, \frac{p_m}{\sum_{i=1}^{m} p_i} \right)$$

$$\mathbf{t}_m(\mathbf{p}) = \text{argmin}_{\mathbf{q} \in \mathscr{P}_m} D_{\mathsf{KL}}(\mathbf{q} \| \mathbf{p})$$

Applications

Common Examples of PMF Truncation

- Let $\mathbf{p} := (p_1, p_2, \cdots, p_n)$ denote the Original PMF
- And $\mathbf{q} := (q_1, q_2, \cdots, q_m)$, where $m < n$, denote the truncated PMF

Operator $\mathbf{t}_m$

$$\mathbf{t}_m(\mathbf{p}) := (q_1, \cdots, q_m) = \left(\frac{p_1}{\sum_{i=1}^{m} p_i}, \cdots, \frac{p_m}{\sum_{i=1}^{m} p_i} \right)$$

$$\mathbf{t}_m(\mathbf{p}) = \text{argmin}_{\mathbf{q} \in \mathscr{P}_m} D_{\mathrm{KL}}(\mathbf{q} \| \mathbf{p})$$

Operator $\mathbf{s}_m$

$$\mathbf{s}_m(\mathbf{p}) := (q_1, \cdots, q_m) = \left(p_1 + \Delta, \cdots, p_m + \Delta\right)$$

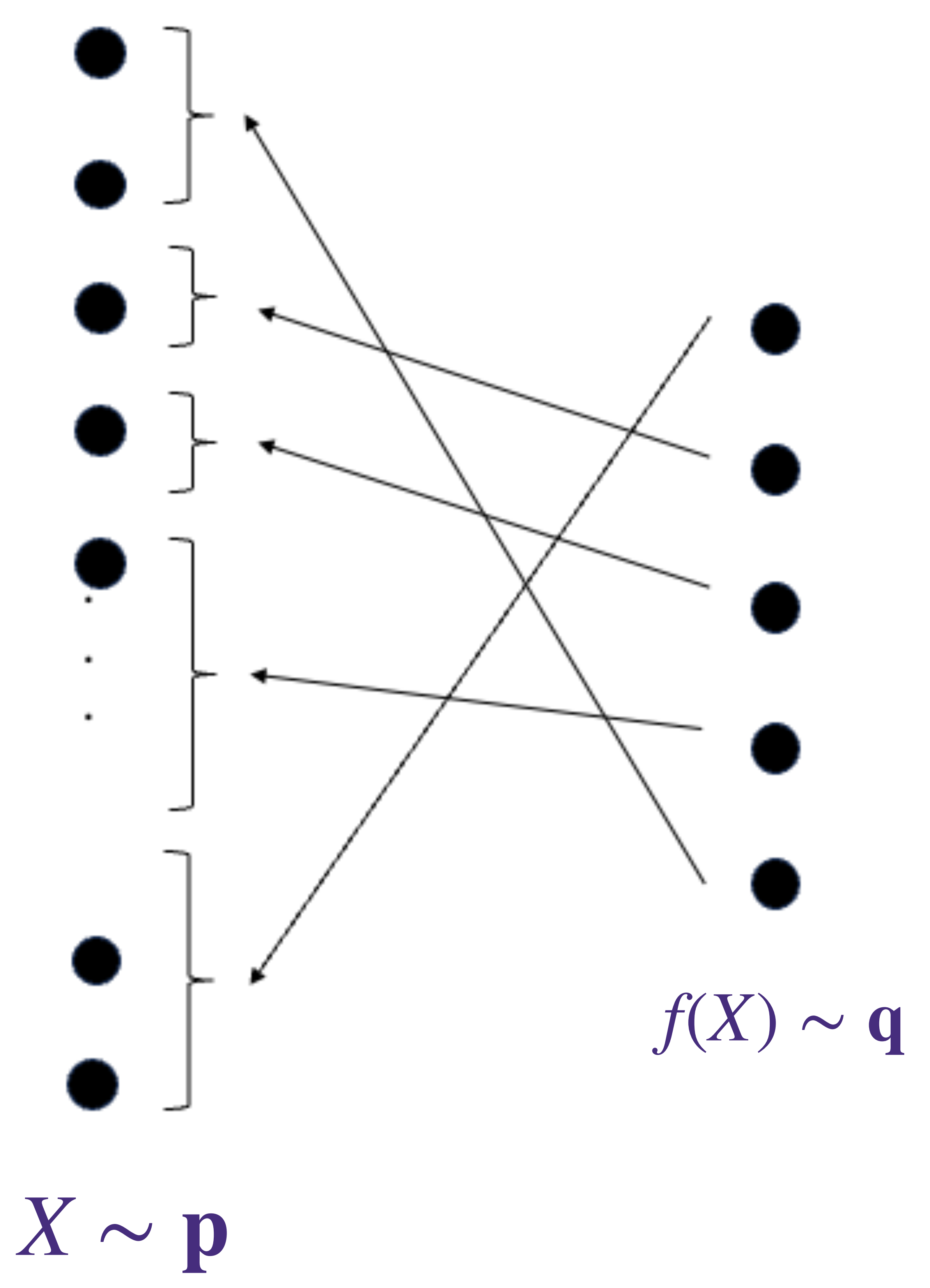
$$\Delta = \sum_{i=m+1}^{n} p_i / m$$

$$\mathbf{S}_m(\mathbf{p}) = \text{argmin}_{\mathbf{q} \in \mathscr{P}_m} \ell_\alpha(\mathbf{q}, \mathbf{p}), \qquad \alpha > 1$$

Applications

Aggregation as Truncation

- Aggregation is a truncation !



Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathscr{X}_n$

Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathscr{X}_n$
- $f : \mathscr{X}_n \to \mathscr{Y}_m \; (m < n)$

Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathcal{X}_n$
- $f : \mathcal{X}_n \to \mathcal{Y}_m \ (m < n)$
- $f(X) \sim \mathbf{q}, \ f(X) \in \mathcal{Y}_m$

Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathcal{X}_n$
- $f : \mathcal{X}_n \rightarrow \mathcal{Y}_m \ (m < n)$
- $f(X) \sim \mathbf{q}, \ f(X) \in \mathcal{Y}_m$
- Metric of closeness: $I(X; f(X))$

Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathscr{X}_n$
- $f : \mathscr{X}_n \to \mathscr{Y}_m \ (m < n)$
- $f(X) \sim \mathbf{q}, \ f(X) \in \mathscr{Y}_m$
- Metric of closeness: $I(X; f(X))$

Best Choice: $\max_{f \in \mathscr{F}_m} I(X; f(X))$

Applications

Aggregation as Truncation

- Aggregation is a truncation !
- $X \sim \mathbf{p}, \quad X \in \mathcal{X}_n$
- $f : \mathcal{X}_n \to \mathcal{Y}_m \ (m < n)$
- $f(X) \sim \mathbf{q}, \ f(X) \in \mathcal{Y}_m$
- Metric of closeness: $I(X; f(X))$

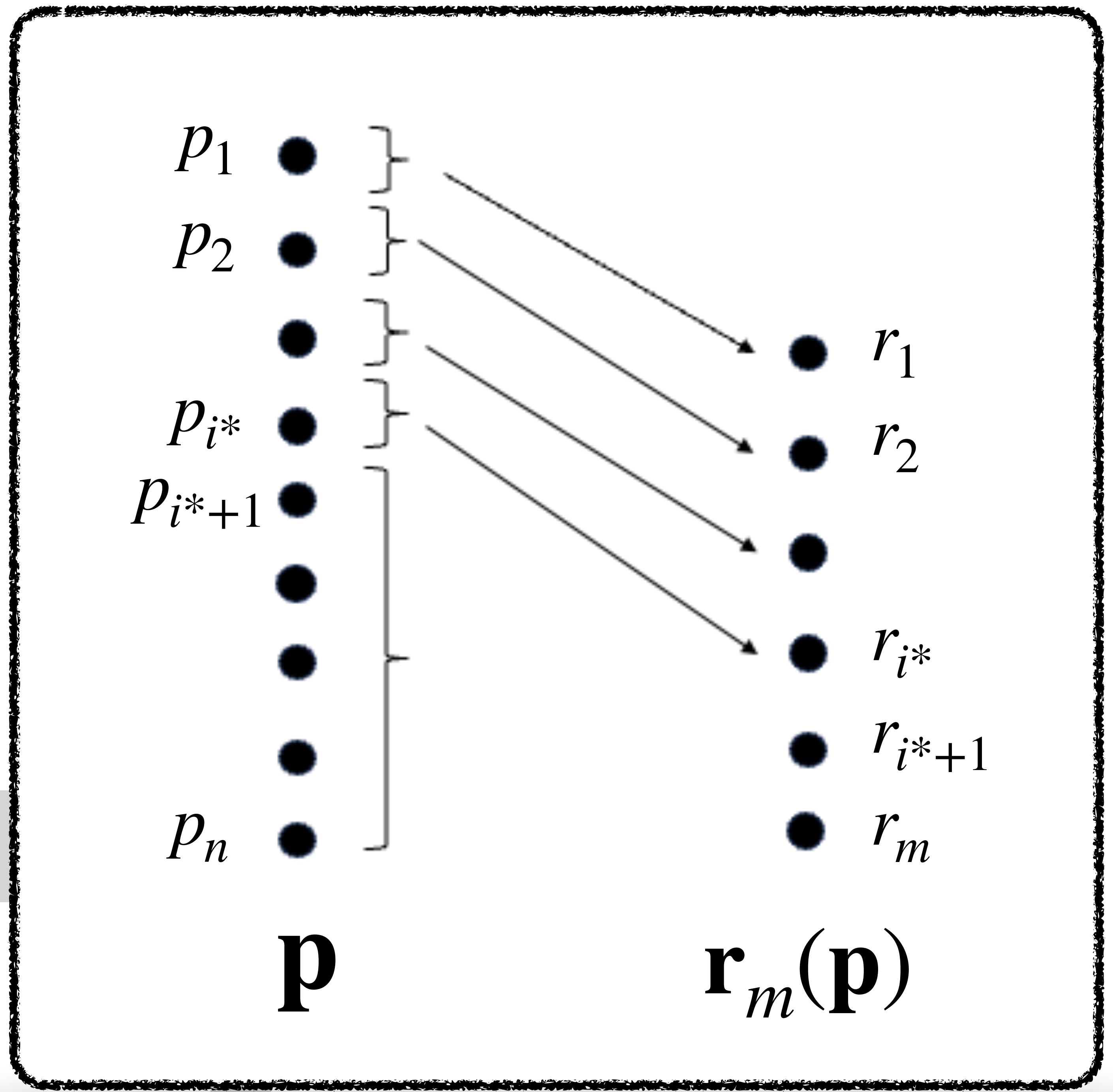
Best Choice: $\max_{f \in \mathcal{F}_m} I(X; f(X)) \equiv \max_{f \in \mathcal{F}_m} H(f(X))$

Construction of f^* via Huffman algorithm !! $\quad H(f^*) \geq H(\mathbf{r}_m(\mathbf{p})) - c_1^{\infty}(2)$

Applications

Recall the Operator $\mathbf{r}_m$

- It's a truncation Operator !!!
- Preserves the components of original PMF $\mathbf{p}$



- If $p_1 < 1/m$: $\mathbf{r}_m(\mathbf{p}) := (1/m, 1/m, \cdots, 1/m)$.
- If $p_1 \geq 1/m$:

$$r_i = \begin{cases} p_i & i = 1,2,\cdots,i^* \\ \left(\sum_{j=i^*+1}^{n} p_j\right)/(m-i^*) & i = i^*+1,\cdots,m \end{cases}$$

where i^* is the maximum index i such that $p_i \geq \dfrac{\sum_{j=i+1}^{n} p_j}{m-i}$.

Applications

On Operator $\mathbf{r}_m$

- Preserves the majorization partial order

Theorem:

For every $\mathbf{p}, \mathbf{q} \in \mathscr{P}_n$, and any $m < n$, it holds that:

$$\mathbf{p} \preceq \mathbf{q} \implies \mathbf{r}_m(\mathbf{p}) \preceq \mathbf{r}_m(\mathbf{q})$$

Applications

On Operator $\mathbf{r}_m$

- Preserves the majorization partial order
- Closest w.r.t ℓ_1 distance:

Theorem:

For any $m < n$, $\mathbf{p} \in \mathscr{P}_n$, and any $\mathbf{q} \in \mathscr{P}_m$, we have:

$$\ell_1(\mathbf{p}, \mathbf{r}_m(\mathbf{p})) \leq \ell_1(\mathbf{p}, \mathbf{q})$$

Applications

Information-theoretic distance $d(\cdot, \cdot)$

- Information-theoretic distance $d(\cdot, \cdot)$: [Cicalese et. al '13]

$$\forall \mathbf{x}, \mathbf{y} \in \mathscr{P}_n \qquad d(\mathbf{x}, \mathbf{y}) := H(\mathbf{x}) + H(\mathbf{y}) - 2H(\mathbf{x} \vee \mathbf{y})$$

Distance metric on the
Majorization Lattice.

Cicalese, Ferdinando et al. "Information theoretic measures of distances and their econometric applications." 2013 IEEE International Symposium on Information Theory (2013): 409-413.

Applications

Information-theoretic distance $d(\cdot, \cdot)$

- Information-theoretic distance $d(\cdot, \cdot)$:

$$\forall \mathbf{x}, \mathbf{y} \in \mathscr{P}_n \qquad d(\mathbf{x}, \mathbf{y}) := H(\mathbf{x}) + H(\mathbf{y}) - 2H(\mathbf{x} \vee \mathbf{y})$$

Generalizes the "Theil Index":

$$d(\mathbf{x}, u_n) := \log n - H(\mathbf{x}) .$$

Applications

On Operator $\mathbf{r}_m$

- Preserves the majorization partial order
- Closest w.r.t ℓ_1 distance
- Closest w.r.t information-theoretic distance $d(\cdot,\cdot)$ defined as:

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{P}_n \qquad d(\mathbf{x}, \mathbf{y}) := H(\mathbf{x}) + H(\mathbf{y}) - 2H(\mathbf{x} \vee \mathbf{y})$$

Theorem:

For any $m < n$, $\mathbf{p} \in \mathcal{P}_n$, and any $\mathbf{q} \in \mathcal{P}_m$, we have:

$$d(\mathbf{p}, \mathbf{r}_m(\mathbf{p})) \leq d(\mathbf{p}, \mathbf{q})$$

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice)
- Properties of Shannon Entropy on the Majorization Lattice
  - Schur Concavity
  - Supermodularity
  - Subadditivity

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
  - Schur Concavity
  - Supermodularity
  - Subadditivity
- Applications in Information Theory & Econometrics
  - Lower Bound on Renyi Entropy $\xrightarrow{\forall \mathbf{p} \in \mathscr{P}_n(\delta)}$ $H_\alpha(\mathbf{p}) \geq \min_{\beta \in \Gamma_n^\delta} H_\alpha(\mathbf{q}_\beta)$

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
    - Schur Concavity
    - Supermodularity
    - Subadditivity
- Applications in Information Theory & Econometrics
    - Lower Bound on Renyi Entropy

$\forall \mathbf{p} \in \mathscr{P}_n(\delta)$

$$H_\alpha(\mathbf{p}) \geq \min_{\beta \in \Gamma_n^\delta} H_\alpha(\mathbf{q}_\beta)$$

$$\geq \log n - c_\alpha^\infty(\delta)$$

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
    - Schur Concavity
    - Supermodularity
    - Subadditivity
- Applications in Information Theory & Econometrics
    - Lower Bound on Renyi Entropy
    - Bounds on $H_\alpha(f(X))$ → $\min_{f \in \mathcal{F}_m} H_\alpha(f(X)) = H_\alpha(\mathbf{q}_m(\mathbf{p}))$

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
  - Schur Concavity
  - Supermodularity
  - Subadditivity
- Applications in Information Theory & Econometrics
  - Lower Bound on Renyi Entropy
  - Bounds on $H_\alpha(f(X))$

$$\min_{f\in\mathcal{F}_m} H_\alpha(f(X)) = H_\alpha(\mathbf{q}_m(\mathbf{p}))$$

$$\max_{f\in\mathcal{F}_m} H_\alpha(f(X)) \in \left[H_\alpha(\mathbf{r}_m(\mathbf{p})) - c_\alpha^\infty(2), H_\alpha(\mathbf{r}_m(\mathbf{p}))\right]$$

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
  - Schur Concavity
  - Supermodularity
  - Subadditivity
- Applications in Information Theory & Econometrics
  - Lower Bound on Renyi Entropy
  - Bounds on $H_\alpha(f(X))$
  - PMF truncation → Aggregation as Truncation!

Conclusion

- Majorization Partial Order is a Lattice (Complete Lattice).
- Properties of Shannon Entropy on the Majorization Lattice
  - Schur Concavity
  - Supermodularity
  - Subadditivity
- Applications in Information Theory & Econometrics
  - Lower Bound on Renyi Entropy
  - Bounds on $H_\alpha(f(X))$
  - PMF truncation → Aggregation as Truncation!
    → Operator $\mathbf{r}_m$ as the truncation operator!

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?
- Minimum Entropy Couplings/ Functional Representations

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?
- Minimum Entropy Couplings/ Functional Representations
  - Majorization in an `information-spectrum' ($\preceq_\imath$) sense. [Shkel & Yadav '23]

$$\imath_X(x) = \log \frac{1}{P_X(x)}$$

Let $U \sim \mathbf{p}$ and $V \sim \mathbf{q}$ be random variables. Then, we say $\mathbf{p} \preceq_\imath \mathbf{q}$ if:

$$\mathbb{P}\left[\imath_U(U) \le t\right] \le \mathbb{P}\left[\imath_V(V) \le t\right]$$

For all $t \in [0, \infty)$.

Shkel Y. Y* and Yadav A. K*, "Information spectrum converse for minimum entropy couplings and functional representations", 2023 IEEE International Symposium on Information Theory (ISIT), (To appear) (* : equal contributions).

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?
- Minimum Entropy Couplings/ Functional Representations
  - Majorization in an `information-spectrum' ($\preceq_\imath$) sense.

$$\imath_X(x) = \log \frac{1}{P_X(x)}$$

> Let $U \sim \mathbf{p}$ and $V \sim \mathbf{q}$ be random variables. Then, we say $\mathbf{p} \preceq_\imath \mathbf{q}$ if:
>
> $$\mathbb{P}\left[\imath_U(U) \leq t\right] \leq \mathbb{P}\left[\imath_V(V) \leq t\right]$$
>
> For all $t \in [0,\infty)$.

  - $\mathbf{p} \preceq_\imath \mathbf{q} \implies \mathbf{p} \preceq \mathbf{q}$.

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?
- Minimum Entropy Couplings/ Functional Representations
    - Majorization in an `information-spectrum' ($\preceq_\imath$) sense.

$$\imath_X(x) = \log \frac{1}{P_X(x)}$$

> Let $U \sim \mathbf{p}$ and $V \sim \mathbf{q}$ be random variables. Then, we say $\mathbf{p} \preceq_\imath \mathbf{q}$ if:
>
> $$\mathbb{P}\left[\imath_U(U) \leq t\right] \leq \mathbb{P}\left[\imath_V(V) \leq t\right]$$
>
> For all $t \in [0,\infty)$.

- $\mathbf{p} \preceq_\imath \mathbf{q} \implies \mathbf{p} \preceq \mathbf{q}$.
- Constructions for minimum entropy couplings?

Applications

Future Work

- Sub(Super)additivity and Super(Sub)modularity properties for Rényi Entropy?
- Minimum Entropy Couplings/ Functional Representations
  - Majorization in an `information-spectrum' ($\preceq_\imath$) sense.

    Let $U \sim \mathbf{p}$ and $V \sim \mathbf{q}$ be random variables. Then, we say $\mathbf{p} \preceq_\imath \mathbf{q}$ if:

    $$\mathbb{P}\left[\imath_U(U) \le t\right] \le \mathbb{P}\left[\imath_V(V) \le t\right]$$

    For all $t \in [0,\infty)$.
  - $\mathbf{p} \preceq_\imath \mathbf{q} \implies \mathbf{p} \preceq \mathbf{q}$.
  - Constructions for minimum entropy couplings?
- $\alpha-$**strong majorization** [Compton `22] to strengthen upper and lower bounds on $\max_{f \in \mathcal{F}_m} H(f(X))$?

Compton, Spencer. "A Tighter Approximation Guarantee for Greedy Minimum Entropy Coupling." 2022 IEEE International Symposium on Information Theory (ISIT) (2022): 168-173.

Thank you!