Uncertainty and entropies of classical channels

Takla Nateeboon

June 30, 2025

Department of Mathematics and Statistics, University of Calgary

Many results in this presentation are from our published manuscript

Inevitable negativity: Additivity commands negative quantum channel entropy

Authors: Gilad Gour[1], Doyeong Kim[2], Takla Nateeboon[2], Guy Shemesh[1], and Goni Yoeli[1]

Phys. Rev. A 111, 052424

1. Department of Mathematics and Helen Diller Quantum Cente, Technion - Israel Institute of Technology, Haifa, Israel
2. Department of Mathematics and Statistics and Institute for Quantum Science and Technology, University of Calgary, Calgary, AB, Canada

What is uncertainty?

What is uncertainty?

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = \left(\frac{3}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad 0 \quad 0\right)^T$$

What is uncertainty?

Given a random variable X which has n possible values, $\{x_1, x_2, \ldots, x_n\}$.

$$\begin{array}{rcccccc} & \text{⚀} & \text{⚁} & \text{⚂} & \text{⚃} & \text{⚄} & \text{⚅} \\ \mathbf{p} = (& \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0)^T \end{array}$$

What is uncertainty?

Given a random variable X which has n possible values, $\{x_1, x_2, \dots, x_n\}$.

$$\begin{array}{cccccccc} & & ⚀ & ⚁ & ⚂ & ⚃ & ⚄ & ⚅ \\ \mathbf{p} & = & (\frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0)^T \end{array} \longleftarrow \text{How can we quantify its } \underline{\text{uncertainty}}?$$

What is uncertainty?

Given a random variable X which has n possible values, $\{x_1, x_2, \dots, x_n\}$.

	⚀	⚁	⚂	⚃	⚄	⚅	
$\mathbf{p} = ($	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	0	$0)^T$	$\longleftarrow$ How can we quantify its uncertainty?

(Shannon) Entropy

$$H(\mathbf{p}) = \sum_{i \in [n]} -p_i \log_2(p_i)$$

Variance

$$\sigma^2 = \mathbb{E}\left[X^2 - \mu^2\right]$$

Given a random variable X which has n possible values, $\{x_1, x_2, \dots, x_n\}$.

$$\begin{array}{rcccccc} & ⚀ & ⚁ & ⚂ & ⚃ & ⚄ & ⚅ \\ \mathbf{p} = (& \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0)^T \end{array}$$

$\longleftarrow$ How can we quantify its <u>uncertainty</u>?

(Shannon) Entropy

$$H(\mathbf{p}) = \sum_{i \in [n]} -p_i \log_2(p_i)$$

- Dependent on the distribution

Variance

$$\sigma^2 = \mathbb{E}\left[X^2 - \mu^2\right]$$

- Dependent on the distribution <u>and the value of X</u>

What is uncertainty?

Given a random variable X which has n possible values, $\{x_1, x_2, \dots, x_n\}$.

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$\mathbf{p} = (\frac{3}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad \frac{1}{6} \quad 0 \quad 0)^T$ $\longleftarrow$ How can we quantify its <u>uncertainty</u>?

(Shannon) Entropy

$$H(\mathbf{p}) = \sum_{i \in [n]} -p_i \log_2(p_i)$$

- Dependent on the distribution
- Axiomatically defined using information processing task

Variance

$$\sigma^2 = \mathbb{E}\left[X^2 - \mu^2\right]$$

- Dependent on the distribution <u>and the value of X</u>

A family of entropies

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

Shannon entropy: $H_1(\mathbf{p}) = -\sum_{y\in[n]} p_y \log_2(p_y)$

A family of entropies

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

Shannon entropy: $H_1(\mathbf{p}) = -\sum_{y\in[n]} p_y \log_2(p_y)$

Max-entropy: $H_0(\mathbf{p}) = \log_2 |\mathrm{supp}(\mathbf{p})|$

Min-entropy: $H_\infty(\mathbf{p}) = -\log_2(\max_{y\in[n]} p_y)$

Denoted by $\mathrm{supp}(\mathbf{p}) = \{ y : p_y \neq 0 \}$ is a set of outcomes y with non-zero probability.

A family of entropies

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

Shannon entropy: $H_1(\mathbf{p}) = -\sum_{y\in[n]} p_y \log_2(p_y)$
$\leftrightarrow$ probability of obtaining a long string drawn from an i.i.d. source = $2^{-H_1(\mathbf{p})}$.

Max-entropy: $H_0(\mathbf{p}) = \log_2 |\mathrm{supp}(\mathbf{p})|$
$\leftrightarrow$ number of possibilities = $2^{H_0(\mathbf{p})}$.

Min-entropy: $H_\infty(\mathbf{p}) = -\log_2(\max_{y\in[n]} p_y)$
$\leftrightarrow$ probability of giving a correct guess of an outcome = $2^{-H_\infty(\mathbf{p})}$

Denoted by $\mathrm{supp}(\mathbf{p}) = \left\{y : p_y \neq 0\right\}$ is a set of outcomes y with non-zero probability.

Entropy and uncertainty

Shannon entropy

$$H(\mathbf{p}) = -\sum_{x\in\mathcal{I}} p_x \log p_x.$$

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

The least uncertained distribution: If $\mathbf{p} = \mathbf{e}_1 = (1, 0, \dots, 0)^T$, knowing the value of X with certainty,

Entropy and uncertainty

Shannon entropy

$$H(\mathbf{p}) = -\sum_{x\in\mathcal{I}} p_x \log p_x.$$

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

The least uncertained distribution: If $\mathbf{p} = \mathbf{e}_1 = (1, 0, \dots, 0)^T$, knowing the value of X with certainty,

$$H(\mathbf{e}_1) = H_\alpha(\mathbf{e}_1) = 0.$$ ← the *least* entropy can be.

Entropy and uncertainty

Shannon entropy

$$H(\mathbf{p}) = -\sum_{x \in \mathcal{I}} p_x \log p_x.$$

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log \left(\sum_{x \in [n]} p_x^\alpha \right).$$

The least uncertained distribution: If $\mathbf{p} = \mathbf{e}_1 = (1, 0, \dots, 0)^T$, knowing the value of X with certainty,

$H(\mathbf{e}_1) = H_\alpha(\mathbf{e}_1) = 0.$ ← the *least* entropy can be.

The most uncertained distribution: If $\mathbf{p} = \mathbf{u}^{(n)} = \frac{1}{n}(1, 1, \dots, 1)^T$, every outcome is equally likely,

Entropy and uncertainty

Shannon entropy

$$H(\mathbf{p}) = -\sum_{x\in\mathcal{I}} p_x \log p_x.$$

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x\in[n]} p_x^\alpha\right).$$

The least uncertained distribution: If $\mathbf{p} = \mathbf{e}_1 = (1, 0, \dots, 0)^T$, knowing the value of X with certainty,

$H(\mathbf{e}_1) = H_\alpha(\mathbf{e}_1) = 0.$ ← the *least* entropy can be.

The most uncertained distribution: If $\mathbf{p} = \mathbf{u}^{(n)} = \frac{1}{n}(1, 1, \dots, 1)^T$, every outcome is equally likely,

$H(\mathbf{u}^{(n)}) = H_\alpha(\mathbf{u}^{(n)}) = \log(n).$ ← the *most* entropy can be.

Entropy and uncertainty

Shannon entropy

$$H(\mathbf{p}) = -\sum_{x \in \mathcal{I}} p_x \log p_x.$$

α-Rényi entropy ($\alpha \in \mathbb{R}_+$)

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{x \in [n]} p_x^\alpha\right).$$

The least uncertained distribution: If $\mathbf{p} = \mathbf{e}_1 = (1, 0, \dots, 0)^T$, knowing the value of X with certainty,

$$H(\mathbf{e}_1) = H_\alpha(\mathbf{e}_1) = 0. \leftarrow \text{the } \textit{least} \text{ entropy can be.}$$

The most uncertained distribution: If $\mathbf{p} = \mathbf{u}^{(n)} = \frac{1}{n}(1, 1, \dots, 1)^T$, every outcome is equally likely,

$$H(\mathbf{u}^{(n)}) = H_\alpha(\mathbf{u}^{(n)}) = \log(n). \leftarrow \text{the } \textit{most} \text{ entropy can be.}$$

Nice common properties: non-negative, additive, invariant with permutation and adding zero.

Classical states

Classical channels

Classical states

- A die rolled a five,

$$\begin{array}{ccccccc} & \text{⚀} & \text{⚁} & \text{⚂} & \text{⚃} & \text{⚄} & \text{⚅} \\ \mathbf{p} = & (0 & 0 & 0 & 0 & 1 & 0)^T \end{array}$$

Classical channels

Classical states

- A die rolled a five,

$$\begin{array}{cccccc} ⚀ & ⚁ & ⚂ & ⚃ & ⚄ & ⚅ \end{array}$$

$$\mathbf{p} = (0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0)^T$$

- A fair die is tossed,

$$\mathbf{p} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T.$$

Classical channels

Classical states

- A die rolled a five,

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = (0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0)^T$$

- A fair die is tossed,

$$\mathbf{p} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T.$$

Classical channels

- sending a bit through a telephone line,

$$\mathcal{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathcal{N} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$$

Classical states

- A die rolled a five,

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = (0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0)^T$$

- A fair die is tossed,

$$\mathbf{p} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T .$$

Classical channels

- sending a bit through a telephone line,

$$\mathcal{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathcal{N} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$$

- tossing a fair die,

$$\mathcal{P}_{\mathbf{p}} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T .$$

Classical states

- A die rolled a five,

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}^T$$

- A fair die is tossed,

$$\mathbf{p} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T .$$

Statics

Classical channels

- sending a bit through a telephone line,

$$\mathcal{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathcal{N} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$$

- tossing a fair die,

$$\mathcal{P}_{\mathbf{p}} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T .$$

Dynamics

Classical states

- A die rolled a five,

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = (0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0)^T$$

- A fair die is tossed,

$$\mathbf{p} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T.$$

Statics

Classical channels

- sending a bit through a telephone line,

$$\mathcal{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathcal{N} = \begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$$

- tossing a fair die,

$$\mathcal{P}_{\mathbf{p}} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}^T.$$

Dynamics

Question: can we have entropy for classical channels as well?

Question: can we have entropy for classical channels as well?

$$\text{Classical entropy} \xrightarrow{\text{extension}} \text{Classical channel entropy}$$

Question: can we have entropy for <u>classical channels</u> as well?

$$\begin{tikzcd}
\text{Classical entropy} \arrow[rr, "\text{extension}"] & & \text{Classical channel entropy} \\
\text{Probability vector majorization} \arrow[u, "\text{axiomatic definition}"']
\end{tikzcd}$$

Question: can we have entropy for classical channels as well?

$$\begin{tikzcd} \text{Classical entropy} \arrow[r, "\text{extension}"] & \text{Classical channel entropy} \\ \text{Probability vector majorization} \arrow[u, "\text{axiomatic definition}"'] & \text{Channel majorization} \arrow[u, "\text{axiomatic definition}"'] \end{tikzcd}$$

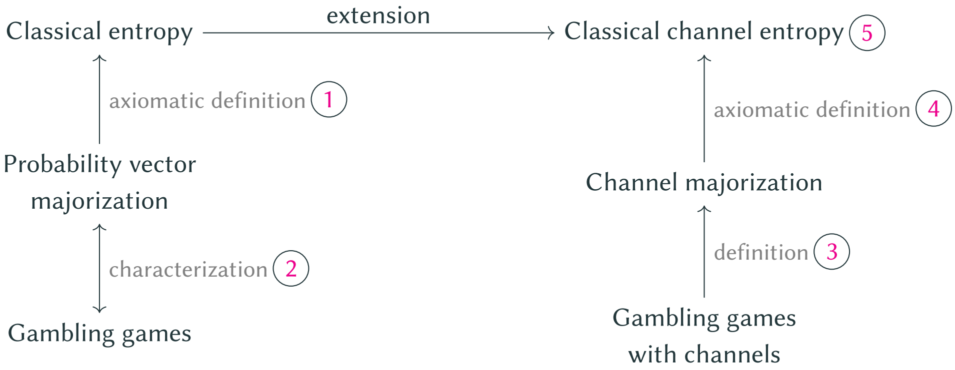
Question: can we have entropy for <u>classical channels</u> as well?

$$
\begin{tikzcd}
\text{Classical entropy} \arrow[r, "\text{extension}"] & \text{Classical channel entropy} \\
\text{Probability vector majorization} \arrow[u, "\text{axiomatic definition}"'] & \text{Channel majorization} \arrow[u, "\text{axiomatic definition}"'] \\
\text{Gambling games} \arrow[u, leftrightarrow, "\text{characterization}"'] &
\end{tikzcd}
$$

Question: can we have entropy for <u>classical channels</u> as well?

$$
\begin{tikzcd}
\text{Classical entropy} \arrow[r, "\text{extension}"] & \text{Classical channel entropy} \\
\text{Probability vector majorization} \arrow[u, "\text{axiomatic definition}"'] & \text{Channel majorization} \arrow[u, "\text{axiomatic definition}"'] \\
\text{Gambling games} \arrow[u, leftrightarrow, "\text{characterization}"'] & \text{Gambling games with channels} \arrow[u, "\text{definition}"']
\end{tikzcd}
$$

Question: can we have entropy for classical channels as well?



Axiomatic definition of classical state entropy

Def. An entropy is a function $H : \mathrm{ProbabilityVectors} \to \mathbb{R}_+$ such that

1. it is an antitone under majorization relation
2. it is additive under a (Kronecker) tensor product

Axiomatic definition of classical state entropy

Def. An entropy is a function $H : \text{ProbabilityVectors} \to \mathbb{R}_+$ such that

1. it is an antitone under majorization relation
2. it is additive under a (Kronecker) tensor product

Majorization ($\succ$) is a preorder on the set of probability vectors.
Def. $\mathbf{p} \succ \mathbf{q}$ if for any $k \in [n]$ the sum of k largest elements of $\mathbf{p}$ is larger or equal to that of $\mathbf{q}$.

Axiomatic definition of classical state entropy

Def. An entropy is a function H : ProbabilityVectors $\to \mathbb{R}_+$ such that

1. it is an antitone under majorization relation
2. it is additive under a (Kronecker) tensor product

Majorization ($\succ$) is a preorder on the set of probability vectors.
Def. $\mathbf{p} \succ \mathbf{q}$ if for any $k \in [n]$ the sum of k largest elements of $\mathbf{p}$ is larger or equal to that of $\mathbf{q}$.

$$\mathbf{p} \succ \mathbf{q}$$

$\mathbf{p}$ is more certain than $\mathbf{q}$

Axiomatic definition of classical state entropy

Def. An entropy is a function H : ProbabilityVectors $\to \mathbb{R}_+$ such that

1. it is an antitone under majorization relation
2. it is additive under a (Kronecker) tensor product

Majorization ($\succ$) is a preorder on the set of probability vectors.
Def. $\mathbf{p} \succ \mathbf{q}$ if for any $k \in [n]$ the sum of k largest elements of $\mathbf{p}$ is larger or equal to that of $\mathbf{q}$.

$$\mathbf{p} \succ \mathbf{q}$$

$$\mathbf{p} \text{ is more certain than } \mathbf{q} \underbrace{\implies H(\mathbf{p}) \leq H(\mathbf{q})}_{\text{being antitone}}$$

Axiomatic definition of classical state entropy

Def. An entropy is a function H : ProbabilityVectors $\to \mathbb{R}_+$ such that

1. it is an antitone under majorization relation
2. it is additive under a (Kronecker) tensor product

Majorization ($\succ$) is a preorder on the set of probability vectors.
Def. $\mathbf{p} \succ \mathbf{q}$ if for any $k \in [n]$ the sum of k largest elements of $\mathbf{p}$ is larger or equal to that of $\mathbf{q}$.

$$\mathbf{p} \succ \mathbf{q}$$

$$\mathbf{p} \text{ is more certain than } \mathbf{q} \underbrace{\implies H(\mathbf{p}) \leq H(\mathbf{q})}_{\text{being antitone}}$$

Additivity under tensor product: two independent random variables $X \sim \mathbf{p}$, $Y \sim \mathbf{q}$ has a joint probability vector $\mathbf{p} \otimes \mathbf{q}$. Additivity means

$$H(\mathbf{p} \otimes \mathbf{q}) = H(\mathbf{p}) + H(\mathbf{q})$$

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = \begin{pmatrix} \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 \end{pmatrix}^T$$

k-**game**: within k guesses, correctly guess the value of a random variable $X \sim \mathbf{p}$.

Majorization and uncertainty

⚀ ⚁ ⚂ ⚃ ⚄ ⚅

$$\mathbf{p} = \left(\tfrac{3}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad 0 \quad 0\right)^T$$

k-**game**: within k guesses, correctly guess the value of a random variable $X \sim \mathbf{p}$.

$\mathbf{p} \succ \mathbf{q}$ is equivalent to $\mathbf{p}$ is at more or equally likely to win any k-game than with $\mathbf{q}$.

Majorization and uncertainty

$$\begin{array}{cccccc} ⚀ & ⚁ & ⚂ & ⚃ & ⚄ & ⚅ \end{array}$$

$$\mathbf{p} = (\tfrac{3}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad 0 \quad 0)^T$$

k-**game**: within k guesses, correctly guess the value of a random variable $X \sim \mathbf{p}$.

> $\mathbf{p} \succ \mathbf{q}$ is equivalent to $\mathbf{p}$ is at more or equally likely to win any k-game than with $\mathbf{q}$.

Examples,

$$(1, 0, 0)^T \succ (\frac{2}{3}, \frac{1}{3}, 0)^T \succ (\frac{1}{2}, \frac{1}{2}, 0)^T \succ \frac{1}{3}(1, 1, 1)^T.$$

Majorization and uncertainty

$$\begin{array}{cccccc} ⚀ & ⚁ & ⚂ & ⚃ & ⚄ & ⚅ \end{array}$$

$$\mathbf{p} = (\tfrac{3}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad \tfrac{1}{6} \quad 0 \quad 0)^T$$

k-**game**: within k guesses, correctly guess the value of a random variable $X \sim \mathbf{p}$.

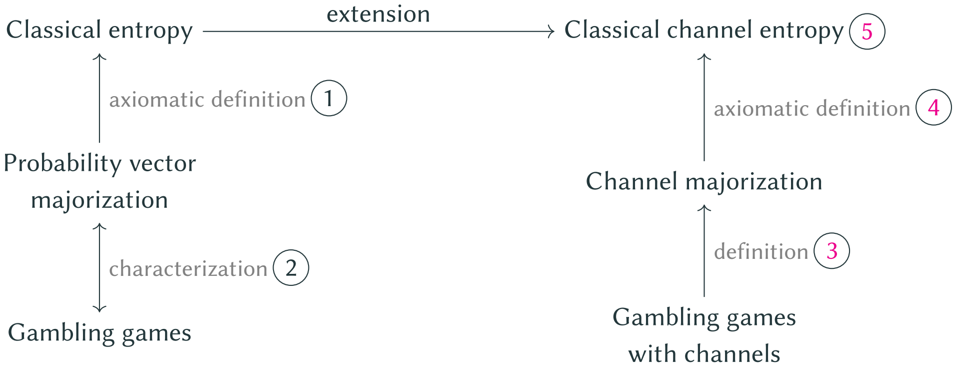
> $\mathbf{p} \succ \mathbf{q}$ is equivalent to $\mathbf{p}$ is at more or equally likely to win any k-game than with $\mathbf{q}$.

Examples,

$$(1, 0, 0)^T \succ (\frac{2}{3}, \frac{1}{3}, 0)^T \succ (\frac{1}{2}, \frac{1}{2}, 0)^T \succ \frac{1}{3}(1, 1, 1)^T.$$

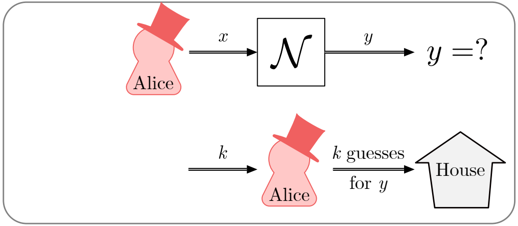
$(1, 0, 0)^T$ is called the maximal element as it *majorizes* everything.
$\frac{1}{3}(1, 1, 1)^T$ is called the minimal element as it *is majorized* by all vector in $\mathrm{Prob}(3)$.

We're going to discuss new results!



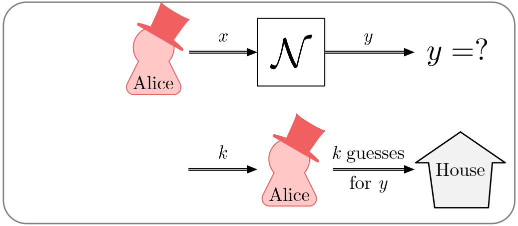
Extending game to classical channel

Goal: guess the value of the output of a channel within k guesses.
The player can pick any input x but k is randomly picked.



Extending game to classical channel

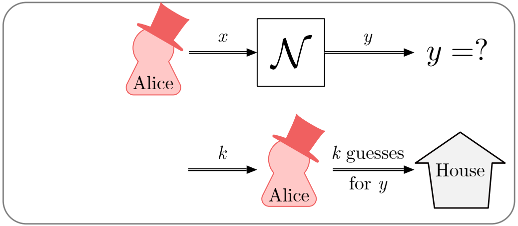
Goal: guess the value of the output of a channel within k guesses.
The player can pick any input x but k is randomly picked.



The player knows prior to giving input x

- transition matrix N associated with the channel $\mathcal{N}$ and
- Probability to have k number of guesses.

Goal: guess the value of the output of a channel within k guesses.

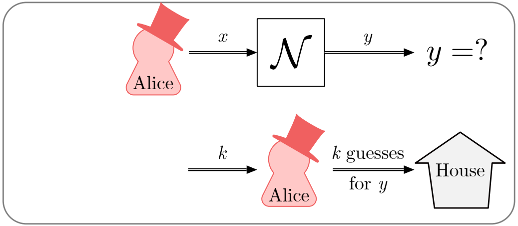


$$N = \begin{array}{cc} \begin{bmatrix} 0.7 & 0.4 \\ 0.15 & 0.1 \\ 0.15 & 0.5 \end{bmatrix} & \begin{array}{l} \leftarrow a \\ \leftarrow b \\ \leftarrow c \end{array} \\ \begin{array}{cc} \uparrow & \uparrow \\ x = 1, & 2 \end{array} & \end{array}$$

k	Pick x	Guess y	Winning Rate

Let's play the game

Goal: guess the value of the output of a channel within k guesses.

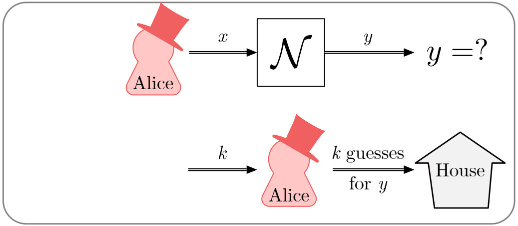


$$N = \begin{array}{c} \begin{bmatrix} 0.7 & 0.4 \\ 0.15 & 0.1 \\ 0.15 & 0.5 \end{bmatrix} \begin{array}{l} \leftarrow a \\ \leftarrow b \\ \leftarrow c \end{array} \\ \begin{array}{cc} \uparrow & \uparrow \\ x = 1, & 2 \end{array} \end{array}$$

k	Pick x	Guess y	Winning Rate
1	1	a	70%

Let's play the game

Goal: guess the value of the output of a channel within k guesses.

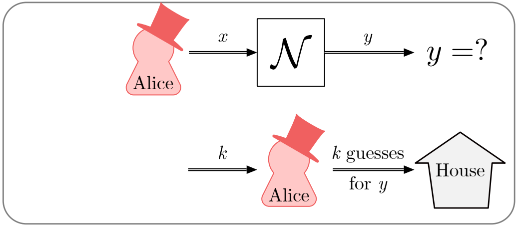


$$N = \begin{bmatrix} 0.7 & 0.4 \\ 0.15 & 0.1 \\ 0.15 & 0.5 \end{bmatrix} \begin{array}{l} \leftarrow a \\ \leftarrow b \\ \leftarrow c \end{array}$$

$$\begin{array}{cc} \uparrow & \uparrow \\ x = 1, & 2 \end{array}$$

k	Pick x	Guess y	Winning Rate
1	1	a	70%
2	2	a and c	90%

Goal: guess the value of the output of a channel within k guesses.

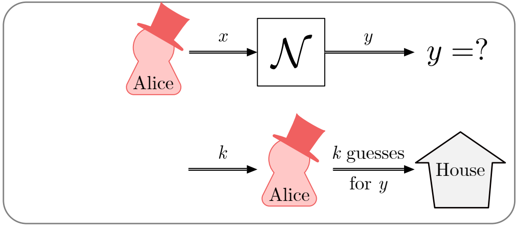


$$N = \begin{array}{cc} \begin{bmatrix} 0.7 & 0.4 \\ 0.15 & 0.1 \\ 0.15 & 0.5 \end{bmatrix} & \begin{array}{c} \leftarrow a \\ \leftarrow b \\ \leftarrow c \end{array} \\ \begin{array}{cc} \uparrow & \uparrow \\ x = 1, & 2 \end{array} & \end{array}$$

k	Pick x	Guess y	Winning Rate
1	1	a	70%
2	2	a and c	90%
1 (50%), 2 (50%)	1	a and b	77.5%

Majorization of classical channels

Goal: guess the value of the output of a channel within k guesses.
The player can pick any input x but k is randomly picked.



Suppose $\mathcal{N}^{X \to Y}$ and $\mathcal{M}^{X' \to Y}$ two classical channels. $\mathcal{N} \succ \mathcal{M}$ if for any distribution of k, the winning chance with channel $\mathcal{N}$ is higher.

Extending majorization to channels

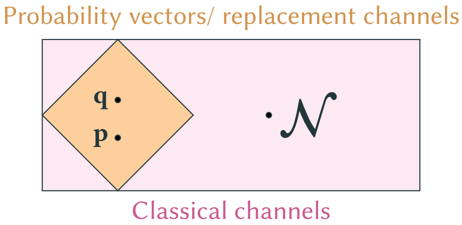
This extended relation

1. defines uncertainty inherent in classical channel,

Extending majorization to channels

This extended relation

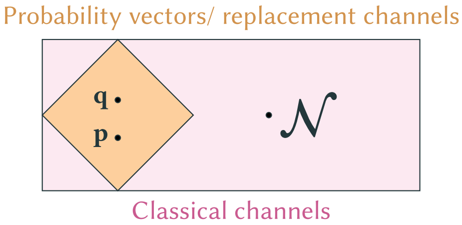
1. defines uncertainty inherent in classical channel,
2. reduces to probability vector majorization on replacement channel,



Extending majorization to channels

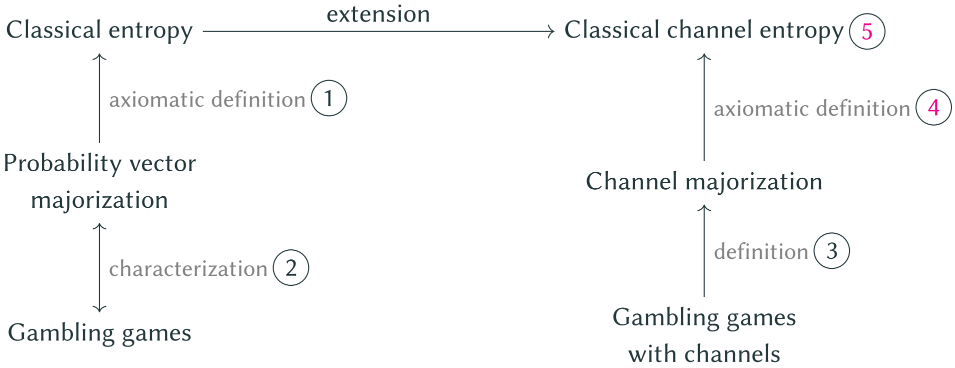
This extended relation

1. defines uncertainty inherent in classical channel,
2. reduces to probability vector majorization on replacement channel,



3. has an identity channel as a maximal element, and
4. has a maximally randomizing channel as a minimal element.

Question: can we have entropy for classical channels as well?



Channel entropy

Definition. An entropy is a function H : ClassicalChannel $\to \mathbb{R}_+$ such that

1. it is an antitone under Channel majorization relation,

$$\mathcal{N} \succ \mathcal{M} \implies H(\mathcal{N}) \leq H(\mathcal{M}),$$

2. it is additive under a tensor product,

$$H(\mathcal{N} \otimes \mathcal{M}) = H(\mathcal{N}) + H(\mathcal{M}).$$

Channel entropy

Definition. An entropy is a function $H : \text{ClassicalChannel} \to \mathbb{R}_+$ such that

1. it is an antitone under Channel majorization relation,

$$\mathcal{N} \succ \mathcal{M} \implies H(\mathcal{N}) \leq H(\mathcal{M}),$$

2. it is additive under a tensor product,

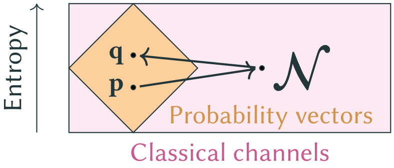
$$H(\mathcal{N} \otimes \mathcal{M}) = H(\mathcal{N}) + H(\mathcal{M}).$$

Existence?

Optimal extensions of an antitone

Maximal extension:

Minimal extension:



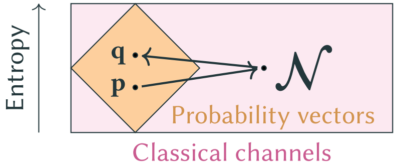
Gour, G. and Tomamichel, M. “Optimal extensions of resource measures and their applications”. In: Phys. Rev. A 102 6 062401 (2020)

Optimal extensions of an antitone

Maximal extension: entropy of the least noisy probability vector that is still more noisy than the channel.

$$\overline{\mathbb{H}}(\mathcal{N}) = \inf_{\substack{\mathbf{q}\in\mathrm{Prob}(m)\\ m\in\mathbb{N}}} \left\{ \mathbb{H}(\mathbf{q}) \,:\, \mathcal{N} \succ \mathbf{q} \right\}$$

Minimal extension:



Gour, G. and Tomamichel, M. "Optimal extensions of resource measures and their applications". In: Phys. Rev. A 102 6 062401 (2020)

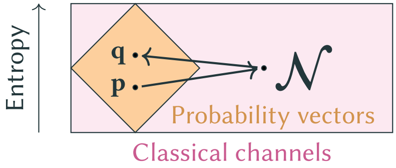
Optimal extensions of an antitone

Maximal extension: entropy of the least noisy probability vector that is still more noisy than the channel.

$$\overline{\mathbb{H}}(\mathcal{N}) = \inf_{\substack{\mathbf{q}\in\mathrm{Prob}(m)\\ m\in\mathbb{N}}} \left\{ \mathbb{H}(\mathbf{q}) \,:\, \mathcal{N} \succ \mathbf{q} \right\}$$

Minimal extension: entropy of the most noisy probability vector that is still less noisy than the channel.

$$\underline{\mathbb{H}}(\mathcal{N}) = \sup_{\substack{\mathbf{p}\in\mathrm{Prob}(m)\\ m\in\mathbb{N}}} \left\{ \mathbb{H}(\mathbf{p}) \,:\, \mathbf{p} \succ \mathcal{N} \right\}$$



Gour, G. and Tomamichel, M. “Optimal extensions of resource measures and their applications”. In: Phys. Rev. A 102 6 062401 (2020)

Examples of classical channel entropy

The maximal extension of a Rényi entropy H_α is additive and it is the *min-entropy output*,

$$\overline{\mathbb{H}}(\mathcal{N}) = \min_{y\in[m]} \mathbb{H}(\mathbf{p}_y) \tag{1}$$

where $\mathbf{p}_y \stackrel{\text{def}}{=} \mathcal{N}(\mathbf{e}_y)$.

α-Rényi entropies

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{y\in[n]} p_y^\alpha\right).$$

Denoted by $\mathrm{supp}(\mathbf{p}) = \left\{ y \,:\, p_y \neq 0 \right\}$ is a set of outcomes y with non-zero probability.

Examples of classical channel entropy

The maximal extension of a Rényi entropy H_α is additive and it is the *min-entropy output*,

$$\overline{\mathbb{H}}(\mathcal{N}) = \min_{y\in[m]} \mathbb{H}(\mathbf{p}_y) \tag{1}$$

where $\mathbf{p}_y \stackrel{\text{def}}{=} \mathcal{N}(\mathbf{e}_y)$.

α-Rényi entropies

$$H_\alpha(\mathbf{p}) = \frac{1}{1-\alpha} \log\left(\sum_{y\in[n]} p_y^\alpha \right).$$

Denoted by $\mathrm{supp}(\mathbf{p}) = \{ y : p_y \neq 0 \}$ is a set of outcomes y with non-zero probability.

Unique entropy extensions

$$\text{Max-entropy} \to H_0(\mathbf{p}) = \log|\mathrm{supp}(\mathbf{p})|,$$

$$\text{Shannon entropy} \to H_1(\mathbf{p}) = -\sum_{y\in[n]} p_y \log(p_y),$$

$$\text{Min-entropy} \to H_\infty(\mathbf{p}) = -\log\left(\max_{y\in[n]} p_y \right).$$

Nonexample: entropy of the Choi state

Suppose that $\mathcal{N} \in \mathrm{CPTP}(X \to Y)$ and $\mathfrak{J}_{\mathcal{N}} \in \mathfrak{L}(XY)$ is its Choi matrix and $\hat{\mathfrak{J}}_{\mathcal{N}}$ to be the normalized Choi matrix. A function f of $\mathcal{N}$ is defined by

$$f(\mathcal{N}) = H\left(\hat{\mathfrak{J}}_{\mathcal{N}}\right) - \log(|X|)$$

where H is a von Neumann entropy. The function f is purposed to be an entropy function[1,2].

Choi state

In probability vector representation,

$$\hat{\mathfrak{J}}_{\mathcal{N}} = \sum_{x \in [n]} \frac{1}{n} \mathbf{e}_x \otimes \mathcal{N}(\mathbf{e}_x).$$

A classical state that correlate each choice of input with its output.

[1] J. Czartowski, D. Braun, and K. Życzkowski. "Trade-off relations for operation entropy of complementary quantum channels". Int. J. Quantum Inf. 17 05 1950046 (2019).
[2] Y. Chu et al. "An entropy function of a quantum channel". Quantum Inf. Process. 22 1 27 (2022)

Nonexample: entropy of the Choi state

Suppose that $\mathcal{N} \in \mathrm{CPTP}(X \to Y)$ and $\mathfrak{J}_{\mathcal{N}} \in \mathfrak{L}(XY)$ is its Choi matrix and $\hat{\mathfrak{J}}_{\mathcal{N}}$ to be the normalized Choi matrix. A function f of $\mathcal{N}$ is defined by

$$f(\mathcal{N}) = H\left(\hat{\mathfrak{J}}_{\mathcal{N}}\right) - \log(|X|)$$

where H is a von Neumann entropy. The function f is purposed to be an entropy function[1,2].

Choi state

In probability vector representation,

$$\hat{\mathfrak{J}}_{\mathcal{N}} = \sum_{x\in[n]} \frac{1}{n} \mathbf{e}_x \otimes \mathcal{N}(\mathbf{e}_x).$$

A classical state that correlate each choice of input with its output.

f is not a channel entropy

$\mathbf{e}_1 \sim (\mathbf{e}_1, \mathbf{p})$ for any probability vector $\mathbf{p}$.

If H is a channel entropy, then

$$H(\mathbf{e}_1) = H(\mathbf{e}_1, \mathbf{p}).$$

However, there is $\mathbf{p}$ such that $f(\mathbf{e}_1) \neq f(\mathbf{e}_1, \mathbf{p})$.

[1] J. Czartowski, D. Braun, and K. Życzkowski. "Trade-off relations for operation entropy of complementary quantum channels". Int. J. Quantum Inf. 17 05 1950046 (2019).
[2] Y. Chu et al. "An entropy function of a quantum channel". Quantum Inf. Process. 22 1 27 (2022)

Conclusion

We defined entropy of a classical channel axiomatically and concretely.

Axiomatically:

1. Axiomatic definition of classical state entropy and majorization.
2. Extension of majorization to classical channels.
3. Axiomatic definition of classical channel entropy follows from extended majorization.

Concretely:

1. Antitone of the extended majorization can be extended from a state entropy.
2. An extension of α-Rényi entropy to a channel entropy.