Inevitable negativity: quantum additivity commands negative quantum channel entropy

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

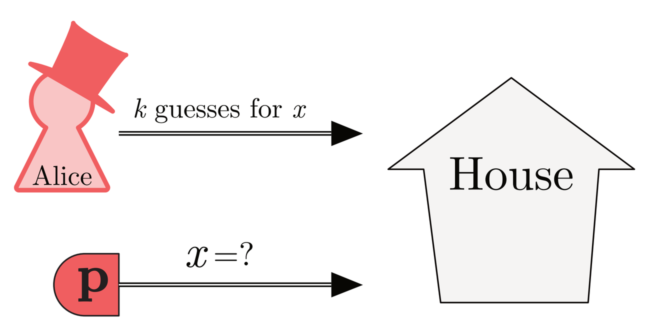
[1]Department of Mathematics, Technion - Israel Institute of Technology, Haifa, Israel

[2]Department of Mathematics and Statistics and Institute for Quantum Science and Technology, University of Calgary, Calgary, AB T2N 1N4, Canada

Introduction

In information theory, entropy is a well-established measure of a system's uncertainty. A deeper understanding of entropy is built on the concept of majorization, a relation that compares uncertainty between different states of a system. As channels have become a central focus in modern information theory, multiple channel entropies have been proposed, but its axiomatic definition has yet to be established. In this study, we provide a strong conceptual foundation for channel entropy by extending majorization to classical and quantum channels. Furthermore, we show that the traditional assumption of additivity in entropy inevitably leads to negative values for certain quantum channels, challenging conventional views on the nature of entropy.

Playing games with probability vectors



Consider the following game of chance, a player, Alice, has a memoryless random source (e.g. a die, a slot machine, or a deck of cards) where she knows the probability distribution $\mathbf{p}$ associated with the random source. To win the game, Alice has to guess an outcome within the k number of guesses correctly. Her chance of winning is the sum of the probability of the largest k outcomes. This is called a k-game.

Majorization and uncertainty

Suppose $\mathbf{p}$ and $\mathbf{q}$ are two probability vectors of n dimensions, we have that $\mathbf{p}$ majorizes $\mathbf{q}$, and write $\mathbf{p} \succ \mathbf{q}$ if $\|\mathbf{p}\|_{(k)} \geq \|\mathbf{q}\|_{(k)}$ for all $k = 1, 2, \ldots, n$, where $\|\mathbf{p}\|_{(k)}$ is the k^{th} Ky Fan norm, the sum of the largest k elements of $\mathbf{p}$.

Alternatively, we can also define majorization via a convertibility relation,

$$\mathbf{q} = M(\mathbf{p}), \tag{1}$$

between two probability vectors via a mixing operation $M : \mathrm{Prob}(n) \to \mathrm{Prob}(n)$, i.e. a stochastic matrix. The mixing operation must not decrease the amount of uncertainty associated with a probability vector. There are three approaches to define mixing operations.

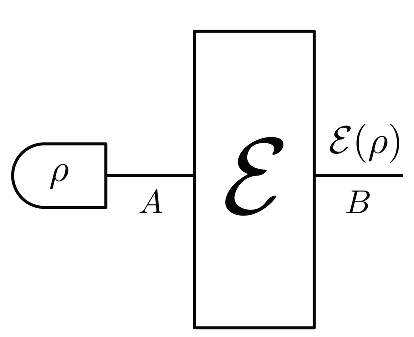
Constructive approach: A mixing operation is an operation that is expressible as a random permutation of outcomes. $M = \sum_i p_i \Pi_i$ where Π_i is a permutation matrix.

Axiomatic approach: A mixing operation is an operation that preserves the most uncertain probability vector, a uniform vector $\mathbf{u}^{(n)}$.

Operational approach: $\mathbf{p} \succ \mathbf{q}$ if and only if the winning chance of the vector $\mathbf{p}$ is greater than that of $\mathbf{q}$ for any k-games.

These three definitions of majorization coincide. It is known that a stochastic matrix that preserves a uniform vector $\mathbf{u}^{(n)}$ is doubly stochastic. From Birkhoff-von Neumann's theorem [1], any doubly stochastic matrix is a convex sum of permutation matrices. The winning chance for a k-game with probability vector $\mathbf{p}$ is a k^{th} Ky Fan norm. Lastly, the majorization relation defined by Ky Fan norm is equivalence with convertability via doubly stochastic matrix [2].

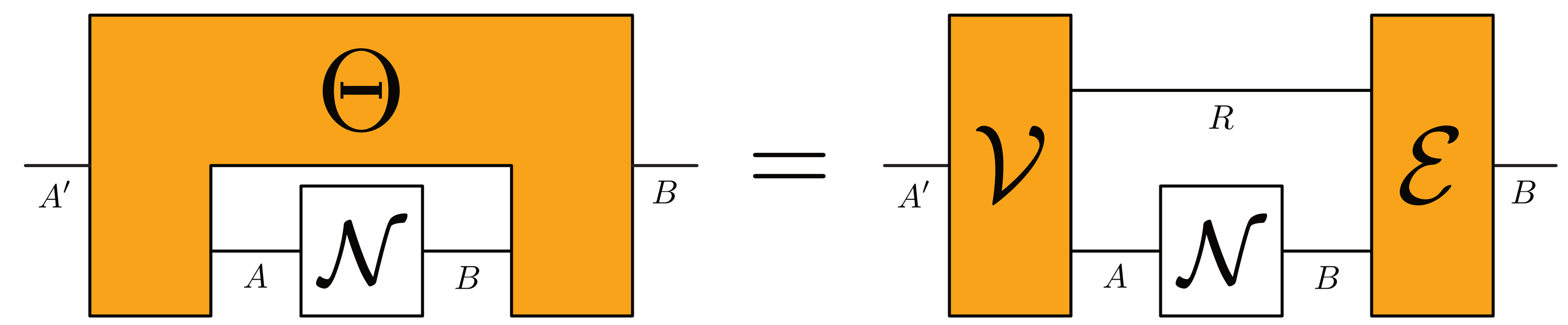
Quantum channels



A quantum channel describes the transformations and evolution of quantum states, capturing both their static properties and dynamic behavior. Mathematically, a quantum channel is represented as a completely positive and trace-preserving (CPTP) linear map. In the left diagram, a quantum channel $\mathcal{N}^{A\to B} \in \mathrm{CPTP}(A \to B)$ maps a quantum state $\rho^A \in \mathfrak{D}(A)$ to $\mathcal{N}^{A\to B}(\rho^A) \in \mathfrak{D}(B)$.

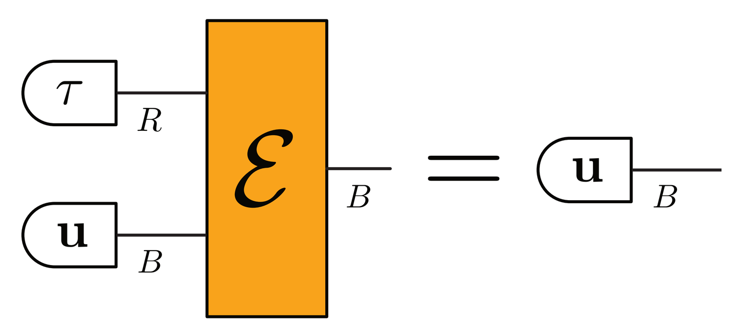
Three approaches to quantum channel majorization

A linear map between quantum channels is called a superchannel, denoted with Θ. An action of a superchannel on a channel $\mathcal{N}$ is realizable as a composition of pre-processing and post-processing channels with the channel $\mathcal{N}$.



We define a majorization on the domain of quantum channels via a convertibility relation between channels. We says a channel $\mathcal{N}$ majorizes a channel $\mathcal{M}$, and write $\mathcal{N} \succ \mathcal{M}$ if there is a superchannel Θ such that $\mathcal{M} = \Theta[\mathcal{N}]$. The challenge was in how to define correctly what are mixing operations on the domain of channels. In our work, we defined mixing operations with three conceptually distinct approaches.

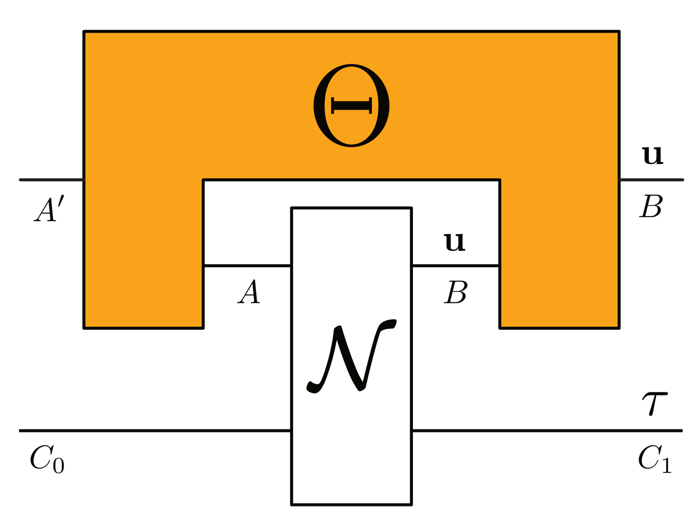
Constructive Approach



A mixing operation is a superchannel realizable with conditionally unital post-processing channel $\mathcal{E}$. We call a channel $\mathcal{E}$ is conditionally unital whenever

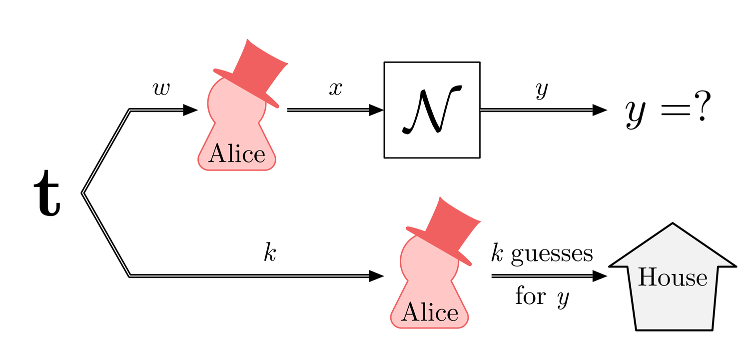
$$\mathcal{E}^{BR\to B}(\mathbf{u}^B \otimes \tau^R) = \mathbf{u}^B \quad \forall \tau \in \mathfrak{D}(R). \tag{2}$$

Axiomatic Approach



A mixing operation has to be completely uniformity-preserving; they must preserve marginally uniform channels. A marginally uniform channel, denoted as $\mathcal{N}^{AC_0\to BC_1}$, is a bipartite channel that always outputs a product state where on the system B the state is maximally mixed, regardless of the input on AC_0.

Operational Approach (classical channels only)
Majorization is defined by comparing the winning chance of all $\mathbf{t}$-games. That is, $\mathcal{N} \succ \mathcal{M}$ if for any $\mathbf{t}$-games, a winning chance with the channel $\mathcal{N}$ is larger than that of the channel $\mathcal{M}$.



In a $\mathbf{t}$-game, the player has full control over the input and aims to correctly predict the output within a randomly selected number of guesses k. Before choosing an input to send through the channel, the player receives a value of w. This value may provide partial, complete, or no information about k, depending on the joint probability distribution of w and k represented by $\mathbf{t}$.

We showed that these three approaches coincide. A superchannel is completely uniformity-preserving if and only if its post-processing channel can be realized using a conditionally unital channel. In the domain of classical channels, comparisons of winning chances relate to the existence of a completely uniformity-preserving channel, which can be expressed as a linear program.

Axiomatic definition of entropy

Similar to the axiomatic definition of quantum state entropy, we defined the entropy of a quantum channel to be an additive monotone on the majorization relation of quantum channels.
That is a mapping $\mathbb{H} : \bigcup_{A,B} \mathrm{CPTP}(A \to B) \to \mathbb{R}$ is a quantum channel entropy if it satisfies the following for any two channels $\mathcal{N}^{A\to B}$ and $\mathcal{M}^{A'\to B'}$.

$$\mathcal{N}^{A\to B} \succ \mathcal{M}^{A'\to B'} \implies \mathbb{H}(B|A)_{\mathcal{N}} \leq \mathbb{H}(B|A')_{\mathcal{M}} \tag{3}$$

$$\mathbb{H}(BB'|AA')_{\mathcal{N}\otimes\mathcal{M}} = \mathbb{H}(B|A)_{\mathcal{N}} \otimes \mathbb{H}(B'|A')_{\mathcal{M}} \tag{4}$$

Examples of channel entropies

Quantum channel entropy can be derived from channel relative entropy $\mathbb{D}$.

$$\mathbb{H}(B|A)_{\mathcal{N}} = \log|B| - \mathbb{D}(\mathcal{N}|\mathcal{R}). \tag{5}$$

The channel relative entropy of channels $\mathcal{N}^{A\to B}$ and $\mathcal{M}^{A\to B}$ is defined by a relative entropy $\mathbb{D}$ [3, 4]

$$\mathbb{D}(\mathcal{N}|\mathcal{M}) := \sup_{\rho\in\mathfrak{D}(RA)} \mathbb{D}(\mathcal{N}(\rho)|\mathcal{M}(\rho)). \tag{6}$$

Another example is a quantum channel entropy derived from the minimization of a conditional entropy of bipartite quantum states,

$$\mathbb{H}(B|A)_{\mathcal{N}} = \min_{\psi\in\mathrm{PURE}(RA)} \mathbb{H}(B|R)_{\mathcal{N}^{A\to B}(\psi^{RA})} \tag{7}$$

where minimization is also over any system R [5].

Some channels have negative entropy

We found that any isometry channel $\mathcal{V} \in \mathrm{CPTP}(A \to B)$ is required to have negative entropy. In particular,

$$\mathbb{H}(B|A)_{\mathcal{V}} = -\log|A|. \tag{8}$$

To put this in context, consider an output system B. Pure states and channels that consistently produce pure states have zero entropy. In contrast, the maximally mixed state $\mathbf{u}^B$ and the maximally randomizing channel $\mathcal{R}^{A\to B}$ possess the highest entropy. Negative channel entropy indicates that the channel is even more predictable than a pure state!

Open questions

- Can the operational definition of channel majorization be extended to quantum channels?
- What are the operational meanings for negative entropies?
- Are all quantum channel entropies extended from a relative entropy?
- Is an extension of a classical state entropy to a classical channel domain unique?

References

[1] Garrett Birkhoff. Three observations on linear algebra. *Univ. Nac. Tacuman, Rev. Ser. A*, 5:147–151, 1946.

[2] G. H. Hardy, J. E. Littlewood, and George Pólya. *Inequalities*. Cambridge University Press, 1934.

[3] Tom Cooney, Milán Mosonyi, and Mark M Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. *Communications in Mathematical Physics*, 344(3):797–829, 2016.

[4] Felix Leditzky, Eneet Kaur, Nilanjana Datta, and Mark M. Wilde. Approaches for approximate additivity of the holevo information of quantum channels. *Phys. Rev. A*, 97:012332, Jan 2018.

[5] Gilad Gour and Mark M. Wilde. Entropy of a quantum channel. *Phys. Rev. Res.*, 3:023096, May 2021.