What is entropy? A perspective from games of chance

Sarah Brandsen, Isabelle Jianing Geng, and Gilad Gour

Phys. Rev. E **105**, 024117 — Published 11 February 2022

DOI: 10.1103/PhysRevE.105.024117

What is Entropy?
A new perspective from games of chance

Sarah Brandsen,[1,2,*] Isabelle Jianing Geng,[2] and Gilad Gour[2]

[1]*Department of Physics, Duke University, Durham, NC, USA 27708*
[2]*Department of Mathematics and Statistics, Institute for Quantum Science and Technology, University of Calgary, AB, Canada T2N 1N4*

The crucial role of channels in physics and information theory motivates the task of characterizing the entropy, or uncertainty, of a channel. Games of chance become a natural candidate for this task, as a system's performance in a gambling game depends solely on the uncertainty of its output. In this work, we construct families of games which induce pre-orders corresponding to majorization, conditional majorization, and channel majorization. Finally, we provide operational interpretations for all pre-orders, show the relevance of these results to dynamical resource theories, and find the only asymptotically continuous classical channel entropy.

I. INTRODUCTION

Entropy plays a central role in many areas of physics and science including statistical mechanics, thermodynamics, information theory, black hole physics, cosmology, chemistry, and even economics [1–3]. Consequently, there are multiple approaches to understanding entropy: in thermodynamics it can be understood as a measure of energy dispersal at a given temperature, whereas in information theory entropy is a compression rate. Other properties related to entropy such as disorder, chaos, randomness of a system, and the arrow of time [4], have also been studied extensively in literature. Despite the many different definitions of entropy, one unifying theme is the idea of uncertainty.

The diverse roles of entropy would benefit from a more systematic and unifying approach, in which entropy is defined rigorously and in a way that is independent of the physical context. In this work, we introduce games of chance as a means of characterizing the uncertainty of a physical system. Games of chance are ideal candidates for studying uncertainty, as the performance in such games depends solely on the certainty about the outcome of the game such that "more uncertain" systems will have a lower expected reward. Thus, for any physical object such as a system, measurement, or channel, we define its degree of uncertainty in terms of the probability of winning a game of chance. Since we construct a family of games, uncertainty cannot be quantified with one function that is induced by a single game, but rather is characterized with a partial order where system A is said to be "less uncertain" than system B if system A performs at least as well as system B for *all* games of chance. Equivalently, we state that A majorizes B.

We then apply our framework to the task of characterizing the entropy of a channel. The classical channel is a fundamental concept in classical information theory, as classical states (which can be viewed as a probability distribution), random relabelings, and stochastic evolutions can all be viewed as a certain type of classical channels. Likewise, quantum states, unitary evolutions, quantum measurements, state preparations and marginalization over any subsystem of a larger quantum system can be regarded as special cases of quantum channels [5]. Over the past few years, several works have aimed to characterise and order channels based on their performance for certain tasks [6–11]. The seminal work by John Kelly [12] in 1956 developed the Kelly criterion which yields the optimal protocol for a player to maximize the growth rate of their gambling profit via N uses of a communication channel in the limit as $N \to \infty$.

The idea of using majorization to study entropy could be found in [13], while the notion of conditional majorization was first introduced in [14], and extended in [15]. However, channel majorization had not been considered in previous literature, nor had a unifying operational interpretation across different types of majorization. In our work, we construct gambling games that give rise to three types of partial orders: majorization, conditional majorization, and channel majorization.The first two partial orders characterize the degree of uncertainty and conditional uncertainty in (possibly composite) physical systems, while the last characterizes the uncertainty associated with a channel. Critically, we provide operational interpretations for each type of majorization and demonstrate that the definition of conditional majorization coincides with the definition provided in [15].

Our main result is demonstrating that there exists a unique channel entropy which reduces to the Shannon entropy on states. The definition of classical entropy based on different sets of axioms has been well studied [16–18], and the Shannon entropy is the most well known measure of the uncertainty of a given probability distribution [19]. The Shannon entropy allows for a reinterpretation of the thermodynamic entropy as consistent with information-theoretic entropy [20]. Additionally the Shannon entropy is the *unique* entropy with the desirable property of asymptotic continuity such that the difference between the entropy of two states approaches zero as the distance between the two states approaches zero. Thus, we would expect that any reasonable entropy of

* sarah.brandsen@duke.edu

a channel would be equivalent to the Shannon entropy whenever the channel is a classical state. Some groundbreaking works which characterize properties of channels are thus excluded from providing reasonable measures of entropy, as there is no possibility of this reduction to states. For example, Kelly's work characterizes channels based on the probability of predicting the channel input given the output state, and as such offers no possibility for reduction to the Shannon entropy on states (in fact, no way to order classical states at all.)

We find the unique asymptotically continuous classical channel entropy function which reduces to Shannon entropy on classical states. Based on these results, one key application is to investigate whether a unique asymptotically continuous channel entropy also exists in the quantum case. Additionally, we provide a first operational interpretation to the complete family of dynamical monotones introduced in [21]. Finally, we expect our framework to have a broad applicability given the key role of uncertainty and entropy in multiple fields. One key extension of our work is determining the entropy of a quantum channel.

Notation We denote classical channels by calligraphic letters $\mathcal{P}$, $\mathcal{Q}$, etc and use capital text letters to denote the corresponding transition probability matrices as $\mathrm{P} = \{p_{y|x}\}$ where $p_{y|x} \coloneqq \Pr(Y = y \big| X = x)$. The set of all classical channel that takes inputs from a m-dimensional system and produces outputs in am n-dimensional system is the set of all $n \times m$ column stochastic matrices and we denote this set as $\mathrm{Stoch}(m, n)$.

We likewise use capital text letters to denote joint probability distributions. For example, $P = \{p_{yx}\}$, $Q = \{q_{y'x'}\}$, $T = \{t_{y''x''}\}$, where $\{p_{yx}\}$ is the joint distribution for X and Y such that $p_{yx} \coloneqq \Pr(Y = y, X = x)$ and similarly for $\{q_{y'x'}\}$ and $\{t_{y''x''}\}$.

II. DICE GAMES AND MAJORIZATION

Gambling games are games in which a player is provided with partial information and use statistical inference to take their best guess in the face of incomplete information. We first consider a gambling game in which the host rolls a biased dice, and the player has to guess its outcome. Denote by $\mathbf{p} = (p_1, ..., p_n)^T$ the probability vector corresponding to the n possible outcomes, and denote by $\mathbf{p}^\downarrow = (p_1^\downarrow, ..., p_n^\downarrow)^T$ the vector obtained from $\mathbf{p}$ by rearranging its components in non-increasing order.

A *w-gambling game* occurs when the player is allowed to provide a set with w outcomes as guesses prior to rolling the dice. The player then wins if the outcome from the dice roll belongs to the set of guesses. For example, if $w = 2$, then the player will choose to provide numbers $\{1, 2\}$ (as these have the highest probability of occurring), and will win the game with probability $p_1^\downarrow + p_2^\downarrow$. In general, the maximum probability of winning a w-game with dice $\mathbf{p}$ can be denoted as:

$$\mathrm{Prob}_w(\mathbf{p}) = \|\mathbf{p}\|_{(w)} \coloneqq \sum_{x=1}^{w} p_x^\downarrow \tag{1}$$

where $\|\cdot\|_{(w)}$ denotes the Ky-Fan norm. For simplicity, in the remainder of the paper we will assume that $\mathbf{p}$ is ordered with components arranged in non-decreasing order such $\mathbf{p} = \mathbf{p}^\downarrow$ unless stated otherwise.

Suppose that at the beginning of each game, the player is allowed to choose between two dice with corresponding probabilities $\mathbf{p}$ and $\mathbf{q}$. Clearly, the player will choose the dice which gives better odds of winning the game, and so will choose the $\mathbf{p}$-dice if $\|\mathbf{p}\|_{(w)} \geqslant \|\mathbf{q}\|_{(w)}$. In general, the player's choice will depend on the value of w- for example, if $\mathbf{p} = (\frac{1}{2}, \frac{1}{2}, 0)$ and $\mathbf{q} = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ then the player will choose $\mathbf{q}$ when $w = 1$ and $\mathbf{p}$ when $w = 2$.

We now consider a more general game where w itself is no longer predetermined, but rather is drawn from some distribution $\{t_w\}_{w=0}^m$. If the player knows the distribution from which the w-gambling game is determined, then the probability that the player wins such a w-game is given by

$$\mathrm{Prob}_{\mathbf{t}}(\mathbf{p}) = \sum_{w=1}^{m} t_w \|\mathbf{p}\|_{(w)} \coloneqq \sum_{x=1}^{m} \sum_{w=x}^{m} t_w p_x \ , \tag{2}$$

In general, we allow $t_0 > 0$, such that there is a non-zero probability that the player loses the game irrespective of the dice outcome. We likewise set $p_x^\downarrow \coloneqq 0$ if $x > n$.

Finally, we say that $\mathbf{p}$ majorizes $\mathbf{q}$ and write $\mathbf{q} \precsim \mathbf{p}$ if and only if

$$\mathrm{Prob}_{\mathbf{t}}(\mathbf{q}) \leqslant \mathrm{Prob}_{\mathbf{t}}(\mathbf{p}) \ . \tag{3}$$

for all (possibly incomplete) distributions $\{t_w\}$. This can be interpreted as stating that $\mathbf{q} \precsim \mathbf{p}$ if the player will *always* choose the $\mathbf{p}$-dice over the $\mathbf{q}$-dice for any gambling game.

III. CONDITIONAL MAJORIZATION: GAMES WITH A CORRELATED SOURCE

Here we consider a game in which the host rolls a dice with two outcomes x and y. The host sends the value of x to the player, and the value y is kept hidden from the player. The player knows the distribution $\{p_{xy}\}$ from which x and y are sampled, and the player's goal is to guess the value of y. Our goal is to construct all possible gambling games that incorporate a correlated source, so we allow the player to choose a value z and then have the host select w from a conditional distribution $\mathcal{T}$ with conditional probability matrix $\mathrm{T} \coloneqq \{t_{w|z}\}$ after receiving the value z from the player. We denote the player's choice of z with a function $z = f(x)$. In general, the player will choose z based on their knowledge of x, as well as the fixed distributions $\{p_{xy}\}$ and $\{t_{w|z}\}$. In Fig. 1 we depict such a $\mathcal{T}$-gambling game.

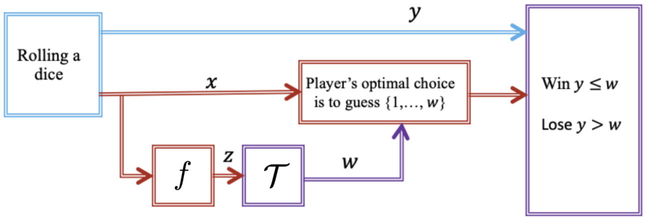

FIG. 1: A classical gambling game with a correlated source. The player is provided with the value x. Based on this value, the player chooses z (or the function f) and sends it to the host. The host then chooses the w game based on a (possibly incomplete) distribution matrix T$=(t_{w|z})$. The player's optimal choice is then to guess $\{1,...,w\}$ such that the player will win the game if $y \leqslant w$.

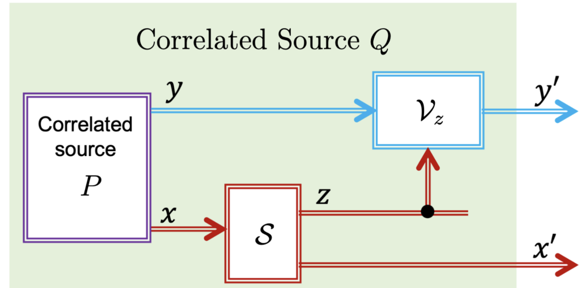

FIG. 2: The action of conditional random relabeling map on a correlated source $P=(p_{xy})$ yields the correlated source $Q=(q_{x'y'})=\sum_z \mathrm{S}_z P \mathrm{V}_z$. When channel $\mathcal{S}$ takes an input x, there is a probability of $s_{x'z|x}$ to output a pair (x',z), where $s_{x'z|x}$ is the matrix component of S_z.

Let $P=(p_{xy})$ be the $m\times n$ probability matrix, and w.l.o.g. suppose $P=P^{\downarrow}$ such that

$$p_{x1} \geqslant p_{x2} \geqslant \cdots \geqslant p_{xn} \quad \forall\, x=1,...,m. \tag{4}$$

For a given x and z, the probability to win the game can be expressed as $\mathbf{r}_z\cdot\mathbf{p}_x$, where $\{\mathbf{p}_x\}$ are the rows of P and $\mathbf{r}_z := U\mathbf{t}_z$, where $\{\mathbf{t}_z\}$ are the columns of the $\ell\times q$ matrix T$=(t_{w|z})$ and where $U=\begin{pmatrix} 1 & 1 & \dots & 1\\ 0 & 1 & \dots & 1\\ & & \dots & \\ 0 & 0 & \dots & 1\end{pmatrix}$ is an upper triangular matrix. Therefore, the optimal probability to win a $\mathcal{T}$-game is given by

$$\mathrm{Prob}_{\mathcal{T}}(P)=\sum_{x=1}^{m}\max_z \mathbf{r}_z\cdot\mathbf{p}_x\ . \tag{5}$$

We are now ready to compare between two dice, a P-dice and a Q-dice, and call this comparison conditional majorization.

Definition 1. *Let $P=(p_{xy})$ be an $m\times n$ probability matrix, and $Q=(q_{x'y'})$ be an $m'\times n'$ probability matrix. We say that P conditionally majorizes Q and write*

$$Q \precsim_c P \quad \textit{if and only if} \quad \mathrm{Prob}_{\mathcal{T}}(Q)\leqslant \mathrm{Prob}_{\mathcal{T}}(P) \tag{6}$$

for all classical channels $\mathcal{T}$ with corresponding (column) stochastic transition matrices T.

Theorem 1. *Let P and Q be two $m\times n$ joint probability matrices. Then,*

$$Q\precsim_c P \quad\iff\quad Q=\sum_z S_z P V_z \tag{7}$$

where each S_z is a sub-stochastic matrix such that $\sum_z S_z$ is a column stochastic matrix (i.e. a classical channel), and each V_z is a permutation matrix.

Remark 1. This theorem states that conditional majorization is equivalent to a relation induced by a conditional random relabeling map; see Fig 2.

Proof. The proof is contained in appendix A [29]. ∎

IV. CHANNEL MAJORIZATION: GAMES WITH A CLASSICAL CHANNEL

We now identify the class of gambling games corresponding to the uncertainty of a channel. Roughly speaking, our aim is to identify games in which the player has a lower probability of winning the game with a noisier channel. We will consider a classical channel $\mathcal{P}$ with transition probability matrix $\mathrm{P}=(p_{y|x})\in\mathrm{Stoch}(m,n)$ which takes inputs $x\in\{1,...,m\}$ and gives outputs $y\in\{1,...,n\}$. The goal of the game is for the player to correctly guess the value y at the output of the channel.

In the most general settings, the host does not provide the player the full information about w at the early stage of the game. Instead, the player receives a number z that is sampled from a joint distribution $\{t_{wz}\}$ with $w=1,...,m$ and $z=1,...,\ell$. Denote by $\{\mathbf{t}_z\}$ the columns of T. Then $\Pr(Z=z)=|\mathbf{t}_z|$, where $|\cdot|$ is the 1-norm, and one can write the conditional probability for w given z as $t_{w|z}=\frac{t_{wz}}{|\mathbf{t}_z|}$ for all w, z.

The player knows the $m\times\ell$ joint probability matrix $T=(t_{wz})$. While $\mathcal{T}$ in the previous section represented a classical channel which takes z as an input and outputs w, here T represents a correlated bipartite source which generates w and z. Based on this partial information about which w-game will be played later on, the player will choose the optimal value of x to send through the channel. Finally, the host draws the value of w and the player wins if $y\leqslant w$.

Such a T-gambling game with channel $\mathcal{P}$ is then depicted in Fig. 3.

Let $\mathrm{P}^{\downarrow}$ be the ordered transition matrix of the channel $\mathcal{P}$, in which the columns of P are arranged in non-increasing order. For a given choice of x and z the probability that the player wins the game is given by

$$\sum_{w=1}^{m} t_{w|z}\sum_{y=1}^{w} p_{y|x}=\sum_{y=1}^{m}\sum_{w=y}^{m} t_{w|z}p_{y|x} := \frac{\mathbf{r}_z\cdot\mathbf{p}_x}{|\mathbf{t}_z|} \tag{8}$$

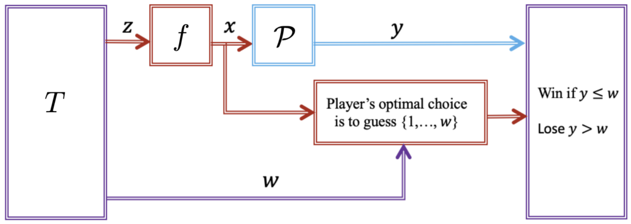

FIG. 3: A classical gambling game with a channel. The host provides the player with a value z that is drawn from $T = (t_{wz})$, then the player chooses channel input $x = f(z)$. The host selects w, and the player is allowed w guesses. Given that the player's optimal choice is to always guess the first w numbers with the highest corresponding probabilities, the player wins if $y \leqslant w$.

where $\mathbf{r}_z := U\mathbf{t}_z$ and $\{\mathbf{p}_x\}$ are the columns of the transition matrix P; and in particular, each $\mathbf{p}_x$ is a probability vector. The player will choose x (i.e. $f(z)$) such that $\mathbf{r}_z \cdot \mathbf{p}_x = \max_{x'} \mathbf{r}_z \cdot \mathbf{p}_{x'}$. Thus, the optimal probability to win a T-gambling game with a classical channel $\mathcal{P}$ can be written as

$$\text{Prob}_T(\mathcal{P}) = \sum_{z=1}^{\ell} \max_x \ \mathbf{r}_z \cdot \mathbf{p}_x \tag{9}$$

Note that the above quantity is the dual of (5) in the sense that the maximum is over x instead of z. Unlike (5), here $\mathbf{p}_x$ is a probability vector for each x.

Definition 2. *Let $\mathcal{Q}$ and $\mathcal{P}$ be two classical channels. We say that the channel $\mathcal{P}$ majorizes $\mathcal{Q}$ and write*

$$\mathcal{Q} \precsim \mathcal{P} \quad \textit{if and only if} \quad \text{Prob}_T(\mathcal{Q}) \leqslant \text{Prob}_T(\mathcal{P}) \tag{10}$$

for all (possibly incomplete) probability matrices T.

We now provide the following characterization of channel majorization, which demonstrates that $\mathrm{Q} \precsim \mathrm{P}$ if and only if Q can be simulated using P, an arbitrary pre-processing channel, and random isometry post-processing channels which may in general be correlated with the input to the pre-processing channel.

Theorem 2. *Let $\mathrm{Q} = (q_{y|x}) \in Stoch(m,n)$ and $\mathrm{P} = (p_{y'|x'}) \in Stoch(m',n')$ correspond to two classical channels $\mathcal{Q}$ and $\mathcal{P}$. Then, $\mathcal{Q} \precsim \mathcal{P}$ if and only if there exists a preprocessing channel with transition matrices $S = \{s_{x'w'|x}\}|_{x',w',x} \in Stoch(m, m' \times \ell)$, and postprocessing channels with transition matrices $\{V_{w'}\}|_{w'}$ described by permutation matrices in $Stoch(n',n)$, such that*

$$\mathrm{Q} = \sum_{w'} V_{w'} \mathrm{P} S_{w'} \tag{11}$$

where $S_{w'} = \{s_{x'w'|x}\}|_{x',x}$ for all w' (see Fig. 4).

Proof. See appendix B of the supplementary material for the complete proof. We note here that the proof additionally provides the first operational interpretation of the dynamical monotones introduced in [21]. ∎

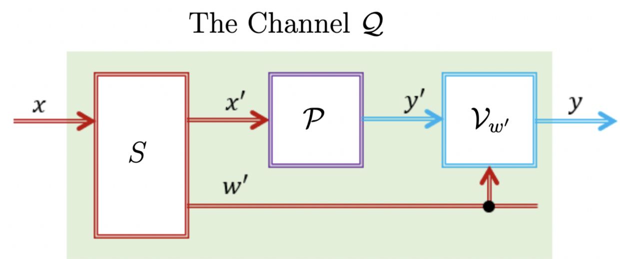

FIG. 4: The simulation of $\mathcal{Q}$ with $\mathcal{P}$ in the case that $\mathcal{Q} \precsim \mathcal{P}$.

We now provide a lemma showing that it is sufficient to consider only games with a fixed value of z.

Lemma 1. *The pre-order induced by the set of classical gambling games is unchanged if we restrict to the set of gambling games such that $T = \mathbf{p}$ where $\mathbf{p}$ is a vector probability distribution. Equivalently, $\mathcal{Q} \precsim \mathcal{P}$ if and only if $Prob_{\mathbf{p}}(\mathcal{Q}) \leqslant Prob_{\mathbf{p}}(\mathcal{P})$ for all $\mathbf{p}$.*

Proof. See appendix C of the supplementary material for the complete proof. ∎

Equipped with the above pre-order, we now provide an operational definition of the family of entropy functions which take the channel's transition probability matrix as an input.

Definition 3. *(cf. [6, 22]) A non-zero function,*

$$H : \bigcup_{m,n \in \mathbb{N}^+} Stoch(m,n) \to \mathbb{R}$$

where the union is over all finite $m \times n$ stochastic probability matrices, is a channel entropy if it satisfies the following two conditions:

1. *It is monotonic under channel majorization; i.e. given classical channels $\mathcal{Q}$ and $\mathcal{P}$ with corresponding transition matrices Q and P, then*

$$H(\mathrm{Q}) \geqslant H(\mathrm{P}) \quad \textit{if} \quad \mathcal{Q} \precsim \mathcal{P}$$

2. *It is additive under tensor products; i.e.*

$$H(\mathrm{P} \otimes \mathrm{Q}) = H(\mathrm{P}) + H(\mathrm{Q})$$

for all finite stochastic matrices P and Q.

Remark 2. In appendix D, we demonstrate that games of chance provide an operational motivation for the definition of entropy for classical channels previously outlined in [6, 22].

Finally, we prove that there is only one channel entropy function that reduces to the Shannon entropy on classical states. Suppose that $\mathcal{P}$ is a classical state preparation channel which takes a trivial system as an input and prepares one of n possible outputs according to distribution $\mathbf{p}$. A channel entropy function H is said to reduce to to the Shannon entropy on states if and only if for *all* state preparation channels with transition probability matrix $\mathrm{P} = \mathbf{p}$:

$$H(\mathrm{P}) = H_S(\mathbf{p})$$

where H_S is the Shannon entropy.

Theorem 3. *Suppose that H is a channel entropy as given by Definition 3 and that H reduces to the Shannon entropy on states. Then for any channel $\mathcal{P}$ with transition probability matrix $\mathrm{P} = \{p_{y|x}\}|_{x,y}$,*

$$H(\mathrm{P}) = \min_x H_S(\mathbf{p}_x).$$

where $\mathbf{p}_x = \{p_{y|x}\}|_y$ are the column vectors of P.

Proof. See appendix E of the supplementary material for the complete proof. ∎

Remark 3. It follows from [23] that $H(\mathrm{P})$ is asymptotically continuous (see appendix E of supplementary material for details).

V. CONCLUSIONS

In this work, we introduce a new method for characterizing uncertainty via payoff functions from games of chance. From this, we introduce families of games of chance which give rise to three different partial orders: majorization, conditional majorization, and channel majorization. We provide an operational interpretation for each ordering. Finally, we find the only asymptotically continuous channel entropy.

One natural extension of this work is to characterise the uncertainty of *quantum* channels. Unlike the classical case, we expect the ordering induced by quantum gambling games to also take into account resources such as entanglement. An open question is whether there also exists a unique asymptotically continuous channel entropy in the quantum case. Finally, previous works on majorization have led to a variety of applications such as finding the capacity of bosonic Gaussian channels, computing quantum discord, characterizing allowed thermodynamic transformations, and developing entanglement detection protocols [24–28]. Likewise, previous entropy definitions such as the Shannon entropy have had wide-reaching applications in thermodynamics and data compression [16, 17]. Thus, given the central role of channels in multiple areas of physics and our results for characterizing channel entropy via majorization, we expect broad potential applications in thermodynamics and information theory.

Acknowledgments. The authors would like to thank Henry Pfister and Mark Wilde for helpful discussions. GG and IG acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC). SB acknowledges support from the National Science Foundation (NSF) under Grant No. 1908730 and 1910571. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

Author Contributions. G.G. conceived the presented idea. S.B. extended the results. S.B. and I.G. verified all results and finalized the manuscript.

[1] A. Jakimowicz, "The Role of Entropy in the Development of Economics," *Entropy, 22, 452*, 2020.

[2] J. D. Bekenstein, "Black holes and the second law," *Lettere al Nuovo Cimento, 4, 737*, 1972.

[3] J. D. Bekenstein, "Black holes and entropy," *Physical Review D, 7, 2333*, 1973.

[4] M. Mackey, "Time's Arrow: The Origins of Thermodynamic Behavior," *Berlin Heidelberg New York: Springer*, 1992.

[5] G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, "The Resource Theory of Informational Nonequilibrium in Thermodynamics," *Physics Reports 583, 1-58*, 2015.

[6] G. Gour and M. M. Wilde, "Entropy of a quantum channel," *arXiv:1808.06980v2*, 2018.

[7] X. Yuan, "Hypothesis testing and entropies of quantum channels," *Phys. Rev. A 99, 032317*, 2019.

[8] K. Fang, O. Fawzi, R. Renner, and D. Sutter, "A chain rule for the quantum relative entropy," *Phys. Rev. Lett. 124, 100501*, 2020.

[9] I. Devetak, C. King, M. Junge, and M. B. Ruskai, "Multiplicativity of completely bounded p-norms implies a new additivity result," *Communications in Mathematical Physics*, 2006.

[10] M. Dall'Arno, F. Buscemi, A. Bisio, and A. Tosini "Data-driven inference and observational completeness of quantum devices," *Phys. Rev. A 102, 062407*, 2020.

[11] M. Dall'Arno, S. Brandsen, and F. Buscemi, "Device-independent tests of quantum channels," *Proc. R. Soc. A, 473, 20160721*, 2017.

[12] Kelly, J. L."A New Interpretation of Information Rate". Bell System Technical Journal. 35 (4): 917–926. 1956.

[13] M. H. Partovi, "Majorization Formulation of Uncertainty in Quantum Mechanics,"*Phys. Rev. A 84, 052117*, 2011.

[14] S. Friedland, V. Gheorghiu, and G. Gour, "Universal Uncertainty Relations,"*Phys. Rev. Lett. 111, 230401*, 2013.

[15] G. Gour, A. Grudka, M. Horodecki, W. Klobus, J. Lodyga, and V. Narasimhachar, "The Conditional Uncertainty Principle," *Phys. Rev. A 97, 042130*, 2018.

[16] C. G.Chakrabarti and I. Chakrabarty, "Shannon Entropy: Axiomatic Characterization and Application,"*International Journal of Mathematics and Mathematical Sciences 2005:17, 2847-2854*, 2005.

[17] J. C. Baez, T. Fritz, and T. Leinster, "A Characteriza-

tion of Entropy in Terms of Information Loss, "*Entropy , 13(11), 1945-1957*, 2011.

[18] A. Renyi, "On measures of information and entropy," *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, 1960.

[19] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656*, 1948.

[20] E. T. Jaynes, "Information Theory and Statistical Mechanics,"*Physical Review, Vol. 106, No. 4, pp. 620-630*, 1957.

[21] G. Gour and C. M. Scandolo, "Dynamical Resources," *arXiv:2101.01552v1*, 2020.

[22] G. Gour, "Comparison of quantum channels by superchannels," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5880–5904, 2019.

[23] G. Gour and A. Winter, "How to quantify a dynamical quantum resource," *Physical Review Letters*, vol. 123, Oct 2019.

[24] A. Mari, V. Giovannetti, and A. Holevo, "Quantum state majorization at the output of bosonic Gaussian channels," *Nat Commun 5, 3826*, 2014.

[25] V. Giovannetti, A.S. Holevo, and A. Mari, "Majorization and additivity for multimode bosonic Gaussian channels," *Theor Math Phys 182, 284–293*, 2015.

[26] M. Horodecki, J. Oppenheim, "Fundamental limitations for quantum and nanoscale thermodynamics," *Nat Commun 4, 2059*, 2015.

[27] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, "Quantum majorization and a complete set of entropic conditions for quantum thermodynamics," *Nat Commun 9, 5352*, 2018.

[28] K. Wang, N. Wu, and F. Song, "Entanglement Detection via Direct-Sum Majorization Uncertainty Relations," *Sci Rep 10, 452*, 2020.

[29] See Supplemental Material at URL for proof details.

[30] G. Gour and M. Tomamichel, "Optimal extensions of resource measures and their applications," *Physical Review A*, vol. 102, Dec 2020.

[31] G. Gour, "Uniqueness and Optimality of Dynamical Extensions of Divergences," *PRX Quantum 2, 010313*, 2021.