entropy

MDPI

Article

Entropy of Quantum Measurements

Stanley Gudder

Department of Mathematics, University of Denver, Denver, CO 80208, USA; sgudder@du.edu

Abstract: If a is a quantum effect and ρ is a state, we define the ρ-entropy $S_a(\rho)$ which gives the amount of uncertainty that a measurement of a provides about ρ. The smaller $S_a(\rho)$ is, the more information a measurement of a gives about ρ. In Entropy for Effects, we provide bounds on $S_a(\rho)$ and show that if $a+b$ is an effect, then $S_{a+b}(\rho) \geq S_a(\rho) + S_b(\rho)$. We then prove a result concerning convex mixtures of effects. We also consider sequential products of effects and their ρ-entropies. In Entropy of Observables and Instruments, we employ $S_a(\rho)$ to define the ρ-entropy $S_A(\rho)$ for an observable A. We show that $S_A(\rho)$ directly provides the ρ-entropy $S_{\mathcal{I}}(\rho)$ for an instrument $\mathcal{I}$. We establish bounds for $S_A(\rho)$ and prove characterizations for when these bounds are obtained. These give simplified proofs of results given in the literature. We also consider ρ-entropies for measurement models, sequential products of observables and coarse-graining of observables. Various examples that illustrate the theory are provided.

Keywords: entropy; quantum measurements; effects; observables

check for updates

Citation: Gudder, S. Entropy of Quantum Measurements. *Entropy* **2022**, *24*, 1686. https://doi.org/10.3390/e24111686

Academic Editor: Ignazio Licata

Received: 25 October 2022
Accepted: 14 November 2022
Published: 18 November 2022

1. Introduction

In an interesting article, D. Šafránek and J. Thingna introduce the concept of entropy for quantum instruments [1]. Various important theorems are proved and applications are given. In quantum computation and information theory one of the most important problems is to determine an unknown state by applying measurements on the system [2–5]. Entropy provides a quantification for the amount of information given to solve this so-called state discrimination problem [6–8]. In this article, we first define the entropy for the most basic measurement, namely a quantum effect a [2,3,9,10]. If ρ is a state, we define the ρ-entropy $S_a(\rho)$ which gives the amount of uncertainty (or randomness) that a measurement of a provides about ρ. The smaller $S_a(\rho)$ is, the more information a measurement of a provides about ρ. In Section 2, we give bounds on $S_a(\rho)$ and show that if $a+b$ is an effect then $S_{a+b}(\rho) \leq S_a(\rho) + S_b(\rho)$. We then prove a result concerning convex mixtures of effects. We also consider sequential products of effects and their ρ-entropies.

In Section 3, we employ $S_a(\rho)$ to define the entropy $S_A(\rho)$ for an observable A. Then $S_A(\rho)$ gives the uncertainty that a measurement of A provides about ρ. We show that $S_A(\rho)$ directly gives the ρ-entropy $S_{\mathcal{I}}(\rho)$ for an instrument $\mathcal{I}$. We establish bounds for $S_A(\rho)$ and characterize when these bounds are obtained. These give simplified proofs of results given in [1,5,11]. We also consider ρ-entropies for measurement models, sequential products of observables and coarse-graining of observables. Various examples that illustrate the theory are provided. In this work, all Hilbert spaces are assumed to be finite dimensional. Although this is a restriction, the work applies for quantum computation and information theory [2,3,9,10].

2. Entropy for Effects

Let H be a finite dimensional complex Hilbert space with dimension n. We denote the set of linear operators on H by $\mathcal{L}(H)$ and the set of states on H by $\mathcal{S}(H)$. If $\rho \in \mathcal{S}(H)$ with nonzero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ including multiplicities, the *von Neumann entropy* of ρ is [4,6–8].

$$S(\rho) = -\sum_{i=1}^{m} \lambda_i \ln(\lambda_i) = -\mathrm{tr}\,[\rho \ln(\rho)]$$

We consider $S(\rho)$ as a measure of the randomness or uncertainty of ρ and smaller values of $S(\rho)$ indicate more information content. For example, ρ is the completely random state I/n, where I is the identity operator, if and only if $S(\rho) = \ln(n)$ and ρ is a pure state if and only if $S(\rho) = 0$. Moreover, it is well-known that $0 \leq S(\rho) \leq \ln(n)$ for all $\rho \in \mathcal{S}(H)$. The following properties of S are well-known [4,6,8]:

$$S(U\rho U^*) = S(\rho) \text{ when } U \text{ is unitary}$$
$$S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2)$$
$$\sum \mu_i S(\rho_i) \leq S\left(\sum \mu_i \rho_i\right) \leq \sum \mu_i S(\rho_i) - \sum \mu_i \ln(\mu_i)$$

where $0 \leq \mu_i = 1$ with $\sum \mu_i = 1$.

An operator $a \in \mathcal{L}(H)$ that satisfies $0 \leq a \leq I$ is called an *effect* [2,3,9,10]. We think of an effect a as a two-outcome yes-no measurement. If a measurement of a results in outcome yes we say that *a occurs* and if it results in outcome no then *a does not occur*. The effect $a' = I - a$ is the *complement* of a and a' occurs if and only if a does not occur. We denote the set of effects by $\mathcal{E}(H)$. If $a \in \mathcal{E}(H)$ and $\rho \in \mathcal{S}(H)$ then $0 \leq \mathrm{tr}\,(\rho a) \leq 1$ and we interpret $\mathrm{tr}\,(\rho a)$ as the probability that a occurs when the system is in state ρ. If $a \neq 0$ we define the *ρ-entropy* of a to be

$$S_a(\rho) = -\mathrm{tr}\,(\rho a) \ln\left[\frac{\mathrm{tr}\,(\rho a)}{\mathrm{tr}\,(a)}\right] \tag{1}$$

We interpret $S_a(\rho)$ as the amount of uncertainty that the system is in state ρ resulting from a measurement of a. The smaller $S_a(\rho)$ is, the more information a measurement of a gives about ρ. Such information is useful for state discrimination problems [2–5].

If ρ is the completely random state I/n then (1) becomes

$$S_a(I/n) = -\mathrm{tr}\,(Ia/n) \ln\left[\frac{\mathrm{tr}\,(Ia/n)}{\mathrm{tr}\,(a)}\right] = -\frac{1}{n}\,\mathrm{tr}\,(a) \ln\left(\frac{1}{n}\right) = \frac{\mathrm{tr}\,(a)}{n} \ln(n)$$

Since $\mathrm{tr}\,(a) \leq n$ we conclude that $S_a(I/n) \leq S(I/n)$ for all $a \in \mathcal{E}(H)$. Another extreme case is when $a = \lambda I$ for $0 < \lambda \leq 1$. We then have for any $\rho \in \mathcal{S}(H)$ that

$$S_{\lambda I}(\rho) = -\mathrm{tr}\,(\rho\lambda I) \ln\left[\frac{\mathrm{tr}\,(\rho\lambda I)}{\mathrm{tr}\,(\lambda I)}\right] = -\lambda \ln\left[\frac{\lambda}{\lambda \mathrm{tr}\,(I)}\right] = \lambda \ln(n)$$

Thus, as λ gets smaller, the more information we gain.

A real-valued function with domain $\mathcal{D}(f)$, an interval in $\mathbb{R}$, is *strictly convex* if for any $x_1, x_2 \in \mathcal{D}(f)$ with $x_1 \neq x_2$ and $0 < \lambda < 1$ we have

$$f[\lambda x_1 + (1-\lambda)x_2] < \lambda f(x_1) + (1-\lambda) f(x_2)$$

If the opposite inequality holds, then f is *strictly concave*. It is clear that f is strictly convex if and only if $-f$ is strictly concave. Of special importance in this work are the strictly convex functions $-\ln x$ and $x \ln x$. We shall frequently employ Jensen's theorem which says: if f is strictly convex and $0 \leq \mu_i \leq 1$ with $\sum_{i=1}^{m} \mu_i = 1$, then

$$f\left(\sum_{i=1}^{m} \mu_i x_i\right) \leq \sum_{i=1}^{m} \mu_i f(x_i)$$

Moreover, we have equality if and only if $x_i = x_j$ for all $i, j = 1, 2, \ldots, m$ [1].

Theorem 1. *If $\rho \in \mathcal{S}(H)$ with nonzero eigenvalues λ_i, $i = 1, 2, \ldots, m$, and $a \in \mathcal{E}(H)$ with* $\mathrm{tr}\,(\rho a) \neq 0$, *then*

$$-\sum_i \mathrm{tr}\,(P_i a)\lambda_i \ln(\lambda_i) \leq S_a(\rho) \leq \ln\left[\tfrac{\mathrm{tr}\,(a)}{\mathrm{tr}\,(\rho a)}\right]$$

where $\rho = \sum_i \lambda_i P_i$ is the spectral decomposition of ρ. Moreover, $S_a(\rho) = \ln[\mathrm{tr}\,(a)/\mathrm{tr}\,(\rho a)]$ if and only if $\mathrm{tr}\,(\rho a) = 1$ *in which case $S_a(\rho) = ln[\mathrm{tr}\,(a)]$ and if*

$$S_a(\rho) = -\sum_i \mathrm{tr}\,(P_i a)\lambda_i \ln(\lambda_i) \tag{2}$$

then $\mathrm{tr}\,(P_i a) = \mathrm{tr}\,(P_j a)$ *for all $i, j = 1, 2, \ldots, m$ and $S_a(\rho) = (\mathrm{tr}\,(a)/m)S(\rho)$ while if* $\mathrm{tr}\,(P_i a) = \mathrm{tr}\,(P_j a)$ *for all $i, j = 1, 2, \ldots m$ then $S_a(\rho) = (\mathrm{tr}\,(a)/m)\ln(m)$.*

Proof. Letting $\mu_j = \mathrm{tr}\,(P_j a)/\mathrm{tr}\,(a)$, $j = 1, 2, \ldots, m$, we have that $0 \leq \mu_j \leq 1$ and $\sum_j \mu_j = 1$. Since $-x\ln(x)$ is strictly concave we obtain

$$\begin{aligned}
S_a(\rho) &= -\mathrm{tr}\,(\rho a)\ln\left[\tfrac{\mathrm{tr}\,(\rho a)}{\mathrm{tr}\,(a)}\right] = -\mathrm{tr}\left(\sum_i \lambda_i P_i a\right)\ln\left[\tfrac{\mathrm{tr}\left(\sum_j \lambda_j P_j a\right)}{\mathrm{tr}\,(a)}\right] \\
&= -\sum \lambda_i \mathrm{tr}\,(P_i a)\ln\left(\sum_j \lambda_j \mu_j\right) = \mathrm{tr}\,(a)\left[-\sum_i \lambda_i \mu_i\left(\sum_j \lambda_j \mu_j\right)\right] \\
&\geq -\mathrm{tr}\,(a)\sum_i \mu_i \lambda_i \ln(\lambda_i) = -\mathrm{tr}\,(a)\sum_i \tfrac{\mathrm{tr}\,(P_i a)}{\mathrm{tr}\,(a)}\,\lambda_i \ln(\lambda_i) \\
&= -\sum_i \mathrm{tr}\,(P_i a)\lambda_i \ln(\lambda_i)
\end{aligned}$$

Since

$$\mathrm{tr}\,(\rho a) = \mathrm{tr}\,(a^{1/2}\rho a^{1/2}) \leq \mathrm{tr}\,(\rho) = 1$$

we have that

$$S_a(\rho) = \mathrm{tr}\,(\rho a)\ln\left[\tfrac{\mathrm{tr}\,(a)}{\mathrm{tr}\,(\rho a)}\right] \leq \ln\left[\tfrac{\mathrm{tr}\,(a)}{\mathrm{tr}\,(\rho a)}\right]$$

If $\mathrm{tr}\,(\rho a) = 1$, then

$$S_a(\rho) = -\mathrm{tr}\,(\rho a)\ln\left[\tfrac{\mathrm{tr}\,(\rho a)}{\mathrm{tr}\,(\rho a)}\right] = -\ln\left[\tfrac{1}{\mathrm{tr}\,(a)}\right] = \ln[\mathrm{tr}\,(a)]$$

Conversely, if $S_a(\rho) = \ln[\mathrm{tr}\,(a)/\mathrm{tr}\,(\rho a)]$, then clearly $\mathrm{tr}\,(\rho a) = 1$. If (2) holds, then we have equality for Jensen's inequality. Hence, $\mathrm{tr}\,(P_i a) = \mathrm{tr}\,(P_j a)$ for all $i, j = 1, 2, \ldots, m$. Since

$$\mathrm{tr}\,(a) = \sum_i \mathrm{tr}\,(P_i a) = m\mathrm{tr}\,(P_i a)$$

we conclude that

$$S_a(\rho) = -\mathrm{tr}\,(P_1 a)\sum_i \lambda_i \ln(\lambda_i) = \tfrac{\mathrm{tr}\,(a)}{m}\,S(\rho)$$

Finally, suppose $\mathrm{tr}\,(P_i a) = \mathrm{tr}\,(P_j a)$ for all $i, j = 1, 2, \ldots, m$. Then

$$\mathrm{tr}\,(a) = \sum_i \mathrm{tr}\,(P_i a) = m\mathrm{tr}\,(P_1 a)$$

We conclude that

$$\begin{aligned}
S_a(\rho) &= -\mathrm{tr}\,(P_1 a)\sum_i \lambda_i \ln\left[\sum_j \lambda_j \tfrac{\mathrm{tr}\,(P_1 a)}{\mathrm{tr}\,(a)}\right] = -\mathrm{tr}\,(P_1 a)\sum_i \lambda_i \ln\left(\sum_j \lambda_j \tfrac{1}{m}\right) \\
&= -\mathrm{tr}\,(P_1 a)\sum_i \lambda_i \ln\left(\tfrac{1}{m}\right) = \tfrac{\mathrm{tr}\,(a)}{m}\;\ln(m)e
\end{aligned}$$

□

For $a, b \in \mathcal{E}(H)$ we write $a \perp b$ if $a + b \in \mathcal{E}(H)$.

Theorem 2. *If $a \perp b$, then $S_{a+b}(\rho) \geq S_a(\rho) + S_b(\rho)$ for all $\rho \in \mathcal{S}(H)$. Moreover, $S_{a+b}(\rho) = S_a(\rho) + S_b(\rho)$ if and only if $\mathrm{tr}\,(b)\mathrm{tr}\,(\rho a) = \mathrm{tr}\,(a)\mathrm{tr}\,(\rho b)$.*

Proof. Since $-x \ln x$ is concave, letting $\lambda_1 = \mathrm{tr}\,(a)/[\mathrm{tr}\,(a) + \mathrm{tr}\,(b)]$, $\lambda_2 = \mathrm{tr}\,(b)/[\mathrm{tr}\,(a) + \mathrm{tr}\,(b)]$, $x_1 = \mathrm{tr}\,(\rho a)/\mathrm{tr}\,(a)$, $x_2 = \mathrm{tr}\,(\rho b)/\mathrm{tr}\,(b)$ we obtain

$$\begin{aligned}
S_{a+b}(\rho) &= -\mathrm{tr}\,[\rho(a+b)] \ln\left\{\frac{\mathrm{tr}\,[\rho(a+b)]}{\mathrm{tr}\,(a+b)}\right\} \\
&= -\mathrm{tr}\,(a+b)\left[\frac{\mathrm{tr}\,(\rho a)+\mathrm{tr}\,(\rho b)}{\mathrm{tr}\,(a+b)}\right] \ln\left[\frac{\mathrm{tr}\,(\rho a)+\mathrm{tr}\,(\rho b)}{\mathrm{tr}\,(a+b)}\right] \\
&= -\mathrm{tr}\,(a+b)(\lambda_1 x_1 + \lambda_2 x_2) \ln(\lambda_1 x_1 + \lambda_2 x_2) \\
&\geq -\mathrm{tr}\,(a+b)[\lambda_1 x_1 \ln(x_1) + \lambda_2 x_2 \ln(x_2)] \\
&= -\mathrm{tr}\,(\rho a) \ln\left[\frac{\mathrm{tr}\,(\rho a)}{\mathrm{tr}\,(a)}\right] - \mathrm{tr}\,(\rho b) \ln\left[\frac{\mathrm{tr}\,(\rho b)}{\mathrm{tr}\,(b)}\right] = S_a(\rho) + S_b(\rho)
\end{aligned}$$

We have equality if and only if $x_1 = x_2$ which is equivalent to $\mathrm{tr}\,(b)\mathrm{tr}\,(\rho a) = \mathrm{tr}\,(a)\mathrm{tr}\,(\rho b)$. □

Corollary 1. *$S_a(\rho) + S_{a'}(\rho) \leq \ln(n)$ and $S_a(\rho) + S_{a'}(\rho) = \ln(n)$ if and only if $\mathrm{tr}\,(a) = n\mathrm{tr}\,(\rho a)$.*

Proof. Applying Theorem 2 we obtain

$$S_a(\rho) + S_{a'}(\rho) \leq S_{a+a'}(\rho) = S_I(\rho) = \ln(n)$$

$$\begin{aligned}
\text{We have equality } &\Leftrightarrow \mathrm{tr}\,(a')\mathrm{tr}\,(\rho a) = \mathrm{tr}\,(a)\mathrm{tr}\,(\rho a') \\
&\Leftrightarrow [n - \mathrm{tr}\,(a)]\mathrm{tr}\,(\rho a) = \mathrm{tr}\,(a)[1 - \mathrm{tr}\,(\rho a)] \\
&\Leftrightarrow \mathrm{tr}\,(a) = n\mathrm{tr}\,(\rho a)e
\end{aligned}$$

□

Corollary 2. *$S_{a+b}(\rho) \geq S_a(\rho), S_b(\rho)$.*

Corollary 3. *If $a \leq b$, then $S_a(\rho) \leq S_b(\rho)$ for all $\rho \in \mathcal{S}(H)$.*

Proof. If $a \leq b$, then $b = a + c$ for $c = b - a \in \mathcal{E}(H)$. Hence,

$$S_b(\rho) = S_{a+c}(\rho) \geq S_a(\rho) + S_c(\rho) \geq S_a(\rho)$$

for every $\rho \in \mathcal{S}(H)$. □

Applying Theorem 2 and induction we obtain the following.

Corollary 4. *If $a_1 + a_2 + \cdots + a_m \leq I$, then $S_{\sum a_i}(\rho) \geq \sum S_{a_i}(\rho)$. Moreover, we have equality if and only if $\mathrm{tr}\,(a_j)\mathrm{tr}\,(\rho a_i) = \mathrm{tr}\,(a_i)\mathrm{tr}\,(\rho a_j)$ for all $i, j = 1, 2, \ldots, m$.*

Notice that $\mathcal{E}(H)$ is a convex set in the sense that if $a_i \in \mathcal{E}(H)$ and $0 \leq \lambda_i \leq 1$ with $\sum_{i=1}^m \lambda_i = 1$, then $\sum \lambda_i a_i \in \mathcal{E}(H)$.

Corollary 5. (i) *If $0 < \lambda \leq 1$ and $a \in \mathcal{E}(H)$, then $S_{\lambda a}(\rho) = \lambda S_a(\rho)$ for all $\rho \in \mathcal{S}(H)$.* (ii) *If $0 < \lambda_i \leq 1$, $a_i \in \mathcal{E}(H)$, with $\sum\limits_{i=1}^m \lambda_i = 1$, then $S_{\sum \lambda_i a_i}(\rho) \leq \sum \lambda_i S_{a_i}(\rho)$ for all $\rho \in \mathcal{S}(H)$. We have equality if and only if $\mathrm{tr}\,(a_j)\mathrm{tr}\,(\rho a_i) = \mathrm{tr}\,(a_i)\mathrm{tr}\,(\rho a_j)$ for all $i, j = 1, 2, \ldots, m$.*

Proof. (i) We have that

$$S_{\lambda a}(\rho) = -\text{tr}\,(\rho\lambda a)\ln\left[\frac{\text{tr}\,(\rho\lambda a)}{\text{tr}\,(\lambda a)}\right] = -\text{tr}\,(\rho a)\ln\left[\frac{\lambda\text{tr}\,(\rho a)}{\lambda\text{tr}\,(a)}\right] = \lambda S_a(\rho)$$

(ii) Applying (i) and Corollary 4 gives

$$S_{\sum\lambda_i a_i}(\rho) \geq \sum S_{\lambda_i a_i}(\rho) = \sum \lambda_i S_{a_i}(\rho)$$

together with the equality condition. □

As with $\mathcal{E}(H)$, $\mathcal{S}(H)$ is a convex set and we have the following.

Theorem 3. *If* $0 < \lambda_i \leq 1$ $\rho_i \in \mathcal{S}(H)$, $i = 1, 2, \ldots, m$, *with* $\sum\limits_{i=1}^{m} \lambda_i = 1$, *then*

$$S_a\left(\sum \lambda_i\rho_i\right) \geq \sum \lambda_i S_a(\rho_i)$$

for all $a \in \mathcal{E}(H)$. *We have equality if and only if* $\text{tr}\,(\rho_i a) = \text{tr}\,(\rho_j a)$ *for all* $i, j = 1, 2, \ldots, m$.

Proof. Letting $x_i = \text{tr}\,(\rho_i a)/\text{tr}\,(a)$, since $-x\ln x$ is concave, we obtain

$$\begin{aligned}
S_a\left(\sum\lambda_i\rho_i\right) &= -\text{tr}\,\left(\sum\lambda_i\rho_i a\right)\ln\left[\frac{\text{tr}\,(\sum\lambda_i\rho_i a)}{\text{tr}\,(a)}\right] \\
&= -\text{tr}\,(a)\sum\lambda_i\frac{\text{tr}\,(\rho_i a)}{\text{tr}\,(a)}\ln\left[\frac{\sum\lambda_i\text{tr}\,(\rho_i a)}{\text{tr}\,(a)}\right] \\
&= \text{tr}\,(a)\left[-\sum\lambda_i x_i\ln\left(\sum\lambda_j x_j\right)\right] \geq -\text{tr}\,(a)\sum\lambda_i x_i\ln(x_i) \\
&= -\text{tr}\,(a)\sum\lambda_i\frac{\text{tr}\,(\rho_i a)}{\text{tr}\,(a)}\ln\left[\frac{\text{tr}\,(\rho_i a)}{\text{tr}\,(a)}\right] = -\sum\lambda_i\text{tr}\,(\rho_i a)\ln\left[\frac{\text{tr}\,(\rho_i a)}{\text{tr}\,(a)}\right] \\
&= \sum\lambda_i S_a(\rho_i)
\end{aligned}$$

We have equality if and only if $x_i = x_j$ which is equivalent to $\text{tr}\,(\rho_i a) = \text{tr}\,(\rho_j a)$ for all $i, j = 1, 2, \ldots, m$. □

Theorem 4. *If* $a_i \in \mathcal{E}(H_i)$, $\rho_i \in \mathcal{S}(H_i)$, $i = 1, 2$, *then*

$$S_{a_1\otimes a_2}(\rho_1\otimes\rho_2) = \text{tr}\,(\rho_2 a_2)S_{a_1}(\rho_1) + \text{tr}\,(\rho_1 a_1)S_{a_2}(\rho_2) \leq S_{a_1}(\rho_1) + S_{a_2}(\rho_2).$$

Proof. This follows from

$$\begin{aligned}
S_{a_1\otimes a_2}(\rho_1\otimes\rho_2) &= -\text{tr}\,(\rho_1\otimes\rho_2 a_1\otimes a_2)\ln\left[\frac{\text{tr}\,(\rho_1\otimes\rho_2 a_1\otimes a_2)}{\text{tr}\,(a_1\otimes a_2)}\right] \\
&= -\text{tr}\,(\rho_1 a_1)\text{tr}\,(\rho_2 a_2)\ln\left[\frac{\text{tr}\,(\rho_1 a_1)\text{tr}\,(\rho_2 a_2)}{\text{tr}\,(a_1)\text{tr}\,(a_2)}\right] \\
&= -\text{tr}\,(\rho_1 a_1)\text{tr}\,(\rho_2 a_2)\left\{\ln\left[\frac{\text{tr}\,(\rho_1 a_1)}{\text{tr}\,(a_1)}\right] + \ln\left[\frac{\text{tr}\,(\rho_2 a_2)}{\text{tr}\,(a_2)}\right]\right\} \\
&= \text{tr}\,(\rho_2 a_2)S_{a_1}(\rho_1) + \text{tr}\,(\rho_1 a_1)S_{a_2}(\rho_2) \leq S_{a_1}(\rho_1) + S_{a_2}(\rho_2)e
\end{aligned}$$

□

An *operation* on H is a completely positive linear map $\mathcal{I}\colon \mathcal{L}(H) \to \mathcal{L}(H)$ such that $\text{tr}\,[\mathcal{I}(A)] \leq \text{tr}\,(A)$ for all $A \in \mathcal{L}(H)$ [2,3,6,9,10]. If $\mathcal{I}$ is an operation we define the *dual* of $\mathcal{I}$ to be the unique linear map $\mathcal{I}^*\colon \mathcal{L}(H) \to \mathcal{L}(H)$ that satisfies $\text{tr}\,[\mathcal{I}(A)B] = \text{tr}\,[A\mathcal{I}^*(B)]$ for all $A, B \in \mathcal{L}(H)$. If $a \in \mathcal{E}(H)$ then for any $\rho \in \mathcal{S}(H)$ we have $0 \leq \text{tr}\,[\mathcal{I}(\rho)a] \leq 1$ and it follows that $\mathcal{I}^*(a) \in \mathcal{E}(H)$. We say that $\mathcal{I}$ *measures* $a \in \mathcal{E}(H)$ if $\text{tr}\,[\mathcal{I}(\rho)] = \text{tr}\,(\rho a)$ for all $\rho \in \mathcal{S}(H)$. If $\mathcal{I}$ measures a we define the $\mathcal{I}$-*sequential product* $a \circ b = \mathcal{I}^*(b)$ for all $b \in \mathcal{E}(H)$ [12,13]. Although $a \circ b$ depends on the operation used to measure a we do not include $\mathcal{I}$ in the notation for simplicity. We interpret $a \circ b$ as the effect that results from first measuring a using $\mathcal{I}$ and then measuring b.

Theorem 5. (i) *If $b \perp c$, then $a \circ (b+c) = a \circ b + a \circ c$.* (ii) $a \circ I = a$. (iii) $a \circ b \leq a$ *for all* $b \in \mathcal{E}(H)$. (iv) $S_{a\circ b}(\rho) \leq S_a(\rho)$ *for all* $\rho \in \mathcal{S}(H)$.

Proof. (i) For every $\rho \in \mathcal{S}(H)$ we obtain

$$\begin{aligned}\mathrm{tr}\,[\rho\, a \circ (b+c)] &= \mathrm{tr}\,[\rho \mathcal{I}^*(b+c)] = \mathrm{tr}\,[\mathcal{I}(\rho)(b+c)] = \mathrm{tr}\,[\mathcal{I}(\rho)b] + \mathrm{tr}\,[\mathcal{I}(\rho)c] \\ &= \mathrm{tr}\,[\rho\mathcal{I}^*(b)] + \mathrm{tr}\,[\rho\mathcal{I}^*(c)] = \mathrm{tr}\,[\rho\, a \circ b] + \mathrm{tr}\,[\rho\, a \circ c] \\ &= \mathrm{tr}\,[\rho(a \circ b + a \circ c)]\end{aligned}$$

Hence, $a \circ (b+c) = a \circ b + a \circ c$. (ii) For all $\rho \in \mathcal{S}(H)$ we have

$$\mathrm{tr}\,(\rho\, a \circ I) = \mathrm{tr}\,[\rho\mathcal{I}^*(I)] = \mathrm{tr}\,[\mathcal{I}(\rho)I] = \mathrm{tr}\,[\mathcal{I}(\rho)] = \mathrm{tr}\,(\rho a)$$

Hence, $a \circ I = a$. (iii) By (i) and (ii) we have

$$a \circ b + a \circ b' = a \circ (b+b') = a \circ I = a$$

It follows that $a \circ b \leq a$. (iv) Since $a \circ b \leq a$, by Corollary 3 we obtain $S_{a\circ b}(\rho) \leq S_a(\rho)$ for all $\rho \in \mathcal{S}(H)$. □

Theorem 5(iv) shows that $a \circ b$ gives more information than a about ρ. We can continue this process and make more measurements as follows. If $\mathcal{I}^i$ measures a^i, $i = 1, 2, \ldots, m$, we have

$$a^1 \circ a^2 \circ \cdots \circ a^m = (\mathcal{I}^1)^*(\mathcal{I}^2)^* \cdots (\mathcal{I}^{m-1})^*(a^m)$$

and it follows from Theorem 5(iv) that

$$S_{a^1\circ a^2 \circ \cdots \circ a^m}(\rho) \leq S_{a^1 \circ a^2 \circ \cdots \circ a^{m-1}}(\rho)$$

Notice that the probability of occurrence of the effect $a^1 \circ a^2 \circ \cdot \circ a^m$ in state ρ is

$$\begin{aligned}\mathrm{tr}\,(\rho\, a^1 \circ a^2 \circ \cdots \circ a^m) &= \mathrm{tr}\left[\rho(\mathcal{I}^1)^*(\mathcal{I}^2)^* \cdots (\mathcal{I}^{m-1})^*(a^m)\right] \\ &= \mathrm{tr}\left[\mathcal{I}^{m-1}\mathcal{I}^{m-2}\cdots\mathcal{I}^1(\rho)a^m\right]\end{aligned}$$

Thus, we begin with the input state ρ, then measure a^1 using $\mathcal{I}^1$, then measure a^2 using $\mathcal{I}^2, \ldots$ and finally measuring a^m.

Example 1. *1 For $a \in \mathcal{E}(H)$ we define the Lüders operation $\mathcal{L}^a(A) = a^{1/2}Aa^{1/2}$ [14]. Since*

$$\mathrm{tr}\,[A(\mathcal{L}^a)^*(B)] = [\mathcal{L}^a(A)B] = \mathrm{tr}\left[a^{1/2}Aa^{1/2}B\right] = \mathrm{tr}\,(Aa^{1/2}Ba^{1/2})$$

we have $(\mathcal{L}^a)^*(B) = a^{1/2}Ba^{1/2}$ so $(\mathcal{L}^a)^* = \mathcal{L}^a$. We have that $\mathcal{L}^a$ measures a because

$$\mathrm{tr}\,[\mathcal{L}^a(\rho)] = \mathrm{tr}\,(a^{1/2}\rho a^{1/2}) = \mathrm{tr}\,(\rho a)$$

for every $\rho \in \mathcal{S}(H)$. We conclude that the $\mathcal{L}^a$ sequential product is

$$a \circ b = (\mathcal{L}^a)^*(b) = a^{1/2}ba^{1/2}$$

We also have that

$$\begin{aligned}S_{a\circ b}(\rho) &= -\mathrm{tr}\,(\rho\, a \circ b)\ln\left[\tfrac{\mathrm{tr}\,(\rho\, a\circ b)}{\mathrm{tr}\,(a\circ b)}\right] = -\mathrm{tr}\,(\rho\, a^{1/2}ba^{1/2})\ln\left[\tfrac{\mathrm{tr}\,(\rho\, a^{1/2}ba^{1/2})}{\mathrm{tr}\,(a^{1/2}ba^{1/2})}\right] \\ &= -\mathrm{tr}\,(a \circ \rho\, b)\ln\left[\tfrac{\mathrm{tr}\,(a\circ\rho\, b)}{\mathrm{tr}\,(ab)}\right].\end{aligned}$$

Example 2. *2 For $a \in \mathcal{E}(H)$, $\alpha \in \mathcal{S}(H)$ we define the Holevo operation [15] $\mathcal{H}^{(a,\alpha)}(A) = \mathrm{tr}\,(Aa)\alpha$. Since*

$$\begin{aligned}\mathrm{tr}\left[A\left(\mathcal{H}^{(a,\alpha)}\right)^*(B)\right] &= \mathrm{tr}\left[\mathcal{H}^{(a,\alpha)}(A)B\right] = \mathrm{tr}\,[\mathrm{tr}\,(Aa)\alpha B] = \mathrm{tr}\,(Aa)\mathrm{tr}\,(\alpha B)\\ &= \mathrm{tr}\,[A\mathrm{tr}\,(\alpha B)a]\end{aligned}$$

we have $\left(\mathcal{H}^{(a,\alpha)}\right)^*(B) = \mathrm{tr}\,(\alpha B)a$. We have $\mathcal{H}^{(a,\alpha)}$ measures a because

$$\mathrm{tr}\left[\mathcal{H}^{(a,\alpha)}(\rho)\right] = \mathrm{tr}\,(\rho a)$$

for every $\rho \in \mathcal{S}(H)$. We conclude that the $\mathcal{H}^{(a,\alpha)}$ sequential product is

$$a \circ b = \left(\mathcal{H}^{(a,\alpha)}\right)^*(b) = \mathrm{tr}\,(\alpha b)a$$

We also have that

$$S_{a\circ b}(\rho) = -\mathrm{tr}\,(\alpha b)\mathrm{tr}\,(\rho a)\ln\left[\frac{\mathrm{tr}\,(\rho a)}{\mathrm{tr}\,(a)}\right] = \mathrm{tr}\,(\alpha b)S_a(\rho)$$

If $a_i \in \mathcal{E}(H)$, $i = 1, 2, \ldots, m$, and we measure a_i with operations $\mathcal{H}^{(a_i,\alpha_i)}$, $i = 1, 2, \ldots, m-1$, then

$$\begin{aligned}a_1 \circ a_2 \circ \cdots \circ a_m = a_1 \circ (a_2 \circ \cdots \circ a_m) &= \mathrm{tr}\,(\alpha_1 a_2 \circ \cdots \circ a_m)a_1\\ &= \mathrm{tr}\,[\alpha_1 \mathrm{tr}\,(\alpha_2 a_3 \circ \cdots \circ a_m)a_2]a_1\\ &= \mathrm{tr}\,(\alpha_2 a_3 \circ \cdots \circ a_m)\mathrm{tr}\,(\alpha_1 a_2)a_1\\ &\;\vdots\\ &= \mathrm{tr}\,(\alpha_{m-1}a_m)\mathrm{tr}\,(\alpha_{m-2}a_{m-1})\cdots\mathrm{tr}\,(\alpha_1 a_2)a_1\end{aligned}$$

Moreover, it follows from Corollary 5(i) that

$$S_{a_1\circ\cdots\circ a_m}(\rho) = \mathrm{tr}\,(\alpha_{m-1}a_m)\mathrm{tr}\,(\alpha_{m-2}a_{m-1})\cdots\mathrm{tr}\,(\alpha_1 a_2)S_{a_1}(\rho)$$

for all $\rho \in \mathcal{S}(H)$.

3. Entropy of Observables and Instruments

We now extend our work on entropy of effects to entropy of observables and instruments. An *observable* on H is a finite collection of effects $A = \{A_x : x \in \Omega_A\}$, $A_x \neq 0$, where $\sum_{x\in\Omega_A} A_x = I$ [2,3,9]. The set Ω_A is called the *outcome space* of A. The effect A_x occurs when a measurement of A results in the outcome x. If $\rho \in \mathcal{S}(H)$, then $\mathrm{tr}\,(\rho A_x)$ is the probability that outcome x results from a measurement of A when the system is in state ρ. If $\Delta \subseteq \Omega_A$, then

$$\Phi_\rho^A(\Delta) = \sum_{x\in\Delta}\mathrm{tr}\,(\rho A_x)$$

is the probability that A has an outcome in Δ when the system is in state ρ and Φ_ρ^A is called the *distribution* of A. We also use the notation $A(\Delta) = \sum\{A_x : x \in \Delta\}$ so $\Phi_\rho^A(\Delta) = \mathrm{tr}\,[\rho A(\Delta)]$ for all $\Delta \subseteq \Omega_A$. In this way, an observable is a *positive operation-valued measure* (POVM). We say that an observable A is *sharp* if A_x is a projection on H for all $x \in \Omega_A$ and A is *atomic* if A_x is a one-dimensional projection for all $x \in \Omega_A$.

If A is an observable and $\rho \in \mathcal{S}(H)$ the ρ*-entropy* of A is $S_A(\rho) = \sum S_{A_x}(\rho)$ where the sum is over the $x \in \Omega_A$ such that $\mathrm{tr}\,(\rho A_x) \neq 0$. Then $S_A(\rho)$ is a measure of the information that a measurement of A gives about ρ. The smaller $S_A(\rho)$ is, the more information given. Notice that if A is sharp, then $\mathrm{tr}\,(A_x) = \dim(A_x)$ and if A is atomic, then

$$S_A(\rho) = -\sum_x \mathrm{tr}\,(\rho A_x)\ln[\mathrm{tr}\,(\rho A_x)]$$

There are two interesting extremes for $S_A(\rho)$. If ρ has spectral decomposition $\rho = \sum_{i=1}^{m} \lambda_i P_i$ and A is the observable $A = \{P_i \colon i = 1, 2, \ldots, m\}$, then

$$S_A(\rho) = -\sum_i \operatorname{tr}(\rho P_i) \ln[\operatorname{tr}(\rho P_i)] = -\sum \lambda_i \ln(\lambda_i) = S(\rho)$$

As we shall see, this gives the minimum entropy (most information). For the completely random state I/n and any observable A we obtain

$$\begin{aligned} S_A(I/n) &= -\sum_x \frac{\operatorname{tr}(A_x)}{n} \ln\left[\frac{\operatorname{tr}(A_x)/n}{\operatorname{tr}(A_x)}\right] = -\frac{1}{n}\sum_x \operatorname{tr}(A_x) \ln\left(\frac{1}{n}\right) \\ &= \frac{\ln(n)}{n} \sum_x \operatorname{tr}(A_x) = \frac{\ln(n)}{n} \operatorname{tr}(I) = \ln(n) \end{aligned} \tag{3}$$

We shall also see that this gives the maximum entropy (least information).

Theorem 6. *For any observable A and $\rho \in \mathcal{S}(H)$ we have*

$$S(\rho) \le S_A(\rho) \le \ln(n)$$

Proof. Applying Theorem 1 we obtain

$$\begin{aligned} S_A(\rho) &= \sum_{x \in \Omega_A} S_{A_x}(\rho) \ge -\sum_{x \in \Omega_A} \sum_i \operatorname{tr}(P_i A_x) \lambda_i \ln(\lambda_i) \\ &= -\sum_i \operatorname{tr}\left(P_i \sum_{x \in \Omega_A} A_x\right) \lambda_i \ln(\lambda_i) \\ &= -\sum_i \operatorname{tr}(P_i) \lambda_i \ln(\lambda_i) = -\sum_i \lambda_i \ln(\lambda_i) = S(\rho) \end{aligned}$$

Since $\ln(x)$ is concave and $\operatorname{tr}(\rho A_x) > 0$, $\sum_x \operatorname{tr}(\rho A_x) = 1$ we have by Jensen's inequality

$$\begin{aligned} S_A(\rho) &= \sum_x \operatorname{tr}(\rho A_x) \ln\left[\frac{\operatorname{tr}(A_x)}{\operatorname{tr}(\rho A_x)}\right] \le \ln\left[\sum_x \operatorname{tr}(\rho A_x) \frac{\operatorname{tr}(A_x)}{\operatorname{tr}(\rho A_x)}\right] \\ &= \ln\left[\sum_x \operatorname{tr}(A_x)\right] = \ln[\operatorname{tr}(I)] = \ln(n)e \end{aligned}$$

□

An observable A is *trivial* if $A_x = \lambda_x I$, $0 < \lambda_x \le 1$, $\sum \lambda_x = 1$.

Corollary 6. (i) *$S_A(\rho) = \ln(n)$ if and only if $\operatorname{tr}(A_x)\operatorname{tr}(\rho A_y) = \operatorname{tr}(A_y)\operatorname{tr}(\rho A_x)$ for all $x, y \in \Omega_A$.* (ii) *A is trivial if and only if $S_A(\rho) = \ln(n)$ for all $\rho \in \mathcal{S}(H)$.* (iii) *$\rho = I/n$ if and only if $S_A(\rho) = \ln(n)$ for all observables A.* (iv) *$S(\rho) = \ln(n)$ if and only if $\rho = I/n$.*

Proof. (i) This follows from the proof of Theorem 6 because this is the condition for equality in Jensen's inequality. (ii) Suppose A is trivial with $A_x = \lambda_x I$. Then for every $\rho \in \mathcal{S}(H)$ we have

$$S_A(\rho) = -\sum_x \operatorname{tr}(\rho \lambda_x I) \ln\left[\frac{\operatorname{tr}(\rho \lambda_x I)}{\operatorname{tr}(\lambda_x I)}\right] = -\sum_x \lambda_x \ln\left(\frac{\lambda_x}{n\lambda_x}\right) = \ln(n) \sum_x \lambda_x = \ln(n)$$

Conversely, suppose $S_A(\rho) = \ln(n)$ for all $\rho \in \mathcal{S}(H)$. By (i) we have that $\operatorname{tr}(A_x)\operatorname{tr}(\rho A_y) = \operatorname{tr}(A_y)\operatorname{tr}(\rho A_x)$ for all $\rho \in \mathcal{S}(H)$. It follows that

$$\langle \phi, A_y \phi \rangle = \langle \phi, A_x \phi \rangle \frac{\operatorname{tr}(A_y)}{\operatorname{tr}(A_x)}$$

for every $\phi \in H$, $\phi \neq 0$. Hence, $A_y = (\mathrm{tr}\,(A_y))/(\mathrm{tr}\,(A_x))A_x$ so that

$$I = \sum_y A_y = \sum_y \frac{\mathrm{tr}\,(A_y)}{\mathrm{tr}\,(A_x)} A_x = \frac{n}{\mathrm{tr}\,(A_x)} A_x$$

We conclude that $A_x = (\mathrm{tr}\,(A_x))/n\, I$ for all $x \in \Omega_A$ so A is trivial. (iii) If $\rho = I/n$, we have shown in (3) that $S_A(\rho) = \ln(n)$ for all observables A. Conversely, if $S_A(\rho) = \ln(n)$ for every observable A, as before, we have $\mathrm{tr}\,(A_x)\mathrm{tr}\,(\rho A_y) = \mathrm{tr}\,(A_y)\mathrm{tr}\,(\rho A_x)$ for every observable A. Letting A_x be the observable given by the spectral decomposition $\rho = \sum \lambda_x A_x$ where A is atomic, we conclude that $\lambda_x = \lambda_y$ for all $x, y \in \Omega_A$. Hence, $\lambda_x = 1/n$ and $\rho = \sum(1/n)A_x = I/n$. (iv) If $S(\rho) = \ln(n)$, by Theorem 6, $S_A(\rho) = \ln(n)$ for every observable A. Applying (iii), $\rho = I/n$. Conversely, if $\rho = I/n$, then

$$S(\rho) = -\sum_{i=1}^{n} \frac{1}{n} \ln\left(\frac{1}{n}\right) = -\ln\left(\frac{1}{n}\right) = \ln(n)e$$

□

We now extend Corollary 5(ii) and Theorem 3 to observables. If $A^i = \{A^i_x : x \in \Omega\}$ are observables with the same outcome space Ω, $i = 1, 2, \ldots, m$, and $0 < \lambda_i \leq 1$ with $\sum_{i=1}^{m} \lambda_i = 1$, then the observable $A = \{A_x : x \in \Omega\}$ where $A_x = \sum_{i=1}^{m} \lambda_i A^i_x$ is called a *convex combination* of the A^i [12].

Theorem 7. (i) *If A is a convex combination of A^i, $i = 1, 2, \ldots, m$, then for all $\rho \in \mathcal{S}(H)$ we have*

$$S_A(\rho) \geq \sum_{i=1}^{m} \lambda_i S_{A^i}(\rho)$$

(ii) *If $0 < \lambda_i \leq 1$ with $\sum_{i=1}^{m} \lambda_i = 1$, $\rho_i \in \mathcal{S}(H)$, $i = 1, 2, \ldots, m$, and A is an observable, then*

$$S_A\left(\sum_i \lambda_i \rho_i\right) \geq \sum_i \lambda_i S_A(\rho_i)$$

Proof. (i) Applying Corollary 5(ii) gives

$$\begin{aligned} S_A(\rho) &= \sum_x S_{A_x}(\rho) = \sum_x S_{\sum \lambda_i A^i_x}(\rho) \geq \sum_x \sum_i \lambda_i S_{A^i_x}(\rho) \\ &= \sum_i \lambda_i \sum_x S_{A^i_x}(\rho) = \sum_i \lambda_i S_{A^i}(\rho) \end{aligned}$$

(ii) Applying Theorem 3 gives

$$\begin{aligned} S_A\left(\sum_i \lambda_i \rho_i\right) &= \sum_x S_{A_x}\left(\sum_i \lambda_i \rho_i\right) \geq \sum_x \sum_i \lambda_i S_{A_x}(\rho_i) \\ &= \sum_i \lambda_i \sum_x S_{A_x}(\rho_i) = \sum_i \lambda_i S_A(\rho_i)e \end{aligned}$$

□

We say that an observable B is a *coarse-graining* of an observable A if there exists a surjection $f\colon \Omega_A \to \Omega_B$ such that

$$B_y = \sum\{A_x : f(x) = y\} = A\left[f^{-1}(y)\right]$$

for every $y \in \Omega_B$ [2,12,16].

Theorem 8. *If B is a coarse-graining of A, then $S_B(\rho) \geq S_A(\rho)$ for all $\rho \in \mathcal{S}(H)$.*

Proof. Let $B_y = A\left[f^{-1}(y)\right]$ for all $y \in \Omega_B$ and let $p_y = \mathrm{tr}\,(\rho B_y)$, $p'_x = \mathrm{tr}\,(\rho A_x)$ for all $y \in \Omega_b$, $x \in \Omega_A$. Then

$$p_y = \mathrm{tr}\left(\rho \sum_{f(x)=y} A_x\right) = \sum_{f(x)=y} \mathrm{tr}\,(\rho A_x) = \sum_{f(x)=y} p'_x$$

Let $V_y = \mathrm{tr}\,(B_y)$, $V'_x = \mathrm{tr}\,(A_x)$ so that

$$V_y = \mathrm{tr}\sum\left(\sum_{f(x)=y} A_x\right) = \sum_{f(x)=y} \mathrm{tr}\,(A_x) = \sum_{f(x)=y} V'_x$$

Since $-x\ln(x)$ is concave, we conclude that

$$\begin{aligned}
S_B(\rho) &= -\sum_y p_y \ln\left(\tfrac{p_y}{V_y}\right) = -\sum_y \sum_{f(x)=y} p'_x \ln\left[\frac{\sum_{f(x)=y} p'_x}{V_y}\right] \\
&= -\sum_y V_y \left(\sum_{f(x)=y} \tfrac{p'_x V'_x}{V'_x V_y}\right) \ln\left(\sum_{f(x)=y} \tfrac{p'_x V'_x}{V'_x V_y}\right) \\
&\geq -\sum_y V_y \sum_{f(x)=y} \tfrac{V'_x}{V_y}\left[\tfrac{p'_x}{V'_x} \ln\left(\tfrac{p'_x}{V'_x}\right)\right] = -\sum_y \sum_{f(x)=y} p'_x \ln\left(\tfrac{p'_x}{V'_x}\right) \\
&= -\sum_x p'_x \ln\left(\tfrac{p'_x}{V'_x}\right) = S_A(\rho)e
\end{aligned}$$

□

The equality condition for Jensen's inequality gives the following.

Corollary 7. *An observable A possesses a coarse-graining $B_y = A\left[f^{-1}(y)\right]$ with $S_B(\rho) = S_A(\rho)$ for all $\rho \in \mathcal{S}(H)$ if and only if for every $x_1, x_2 \in \Omega_A$ with $f(x_1) = f(x_2)$ we have*

$$\mathrm{tr}\,(A_{x_2})\mathrm{tr}\,(\rho A_{x_1}) = \mathrm{tr}\,(A_{x_1})\mathrm{tr}\,(\rho A_{x_2})$$

A trace preserving operation is called a *channel*. An *instrument* on H is a finite collection of operations $\mathcal{I} = \{\mathcal{I}_x \colon x \in \Omega\}$ such that $\sum_{x\in\Omega_{\mathcal{I}}} \mathcal{I}_x$ is a channel [2,3,9]. We call $\Omega_{\mathcal{I}}$ the *outcome space* for $\mathcal{I}$. If $\mathcal{I}$ is an instrument, there exists a unique observable A such that $\mathrm{tr}\,(\rho A_x) = \mathrm{tr}\,[\mathcal{I}_x(\rho)]$ for all $x \in \Omega_A = \Omega_{\mathcal{I}}$, $\rho \in \mathcal{S}(H)$ and we say that $\mathcal{I}$ *measures* A. Although an instrument measures a unique observable, an observable is measured by many instruments For example, if A is an observable, the corresponding *Lüders instrument* [14] is defined by

$$\mathcal{L}_x^A(B) = A_x^{1/2} B A_x^{1/2}$$

for all $B \in \mathcal{L}(H)$. Then $\mathcal{L}^A$ is an instrument because

$$\begin{aligned}
\mathrm{tr}\left[\sum_x \mathcal{L}_x^A(B)\right] &= \sum_x \mathrm{tr}\left[\mathcal{L}_x^A(B)\right] = \sum_x \mathrm{tr}\,(A_x^{1/2} B A_x^{1/2}) = \sum_x \mathrm{tr}\,(A_x B) \\
&= \mathrm{tr}\left(\sum_x A_x B\right) = \mathrm{tr}\,(IB) = \mathrm{tr}\,(B)
\end{aligned}$$

for all $B \in \mathcal{L}(H)$. Moreover, $\mathcal{L}^A$ measures A because

$$\mathrm{tr}\left[\mathcal{L}_x^A(\rho)\right] = \mathrm{tr}\,(A_x^{1/2} \rho A_x^{1/2}) = \mathrm{tr}\,(\rho A_x)$$

for all $\rho \in \mathcal{S}(H)$. Of course, this is related to Example 1. Corresponding to Example 2, we have a *Holevo instrument* $\mathcal{H}^{(A,\alpha)}$ where $\alpha_x \in \mathcal{S}(H)$, $x \in \Omega_A$ and

$$\mathcal{H}_x^{(A,\alpha)}(B) = \mathrm{tr}\,(BA_x)\alpha_x$$

for all $B \in \mathcal{L}(H)$ [15]. To show that $\mathcal{H}^{(A,\alpha)}$ is an instrument we have

$$\begin{aligned}\mathrm{tr}\left[\sum_x \mathcal{H}_x^{(A,\alpha)}(B)\right] &= \sum_x \mathrm{tr}\left[\mathcal{H}_x^{(A,\alpha)}(B)\right] = \sum_x \mathrm{tr}\,[\mathrm{tr}\,(BA_x)\alpha_x] \\ &= \sum_x \mathrm{tr}\,(BA_x) = \mathrm{tr}\left(B\sum_x A_x\right) = \mathrm{tr}\,(B)\end{aligned}$$

Moreover, $\mathcal{H}^{(A,\alpha)}$ measures A because

$$\mathrm{tr}\left[\mathcal{H}_x^{A,\alpha}(\rho)\right] = \mathrm{tr}\,[(\rho A_x)\alpha_x] = \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\alpha_x) = \mathrm{tr}\,(\rho A_x)$$

Let A, B be observables and let $\mathcal{I}$ be an instrument that measures A. We define the *$\mathcal{I}$-sequential* product $A \circ B$ [12,13] by $\Omega_{A\circ B} = \Omega_A \times \Omega_B$ and

$$A \circ B_{(x,y)} = \mathcal{I}_x^*(B_y) = A_x \circ B_y$$

Defining $f\colon \Omega_{A\circ B} \to \Omega_A$ by $f(x,y) = x$,we obtain

$$A \circ B\left[f^{-1}(x)\right] = \sum_{f(x,y)=x} A_x \circ B_y = \sum_{y\in\Omega_B} \mathcal{I}_x^*(B_y) = \mathcal{I}_\alpha^*(I) = A_x$$

We conclude that A is a coarse-graining of $A \circ B$. Applying Theorem 8 we obtain the following.

Corollary 8. *If A, B are observables, the $S_{A\circ B}(\rho) \le S_A(\rho)$ for all $\rho \in \mathcal{S}(H)$. Equality $S_{A\circ B}(\rho) = S_A(\rho)$ holds if and only if for every $x \in \Omega_A$, $y_1, y_2 \in \Omega_B$ we have*

$$\frac{\mathrm{tr}\,(\rho A_x \circ B_{y_1})}{\mathrm{tr}\,(A_x \circ B_{y_1})} \ln\left[\frac{\mathrm{tr}\,(\rho A_x \circ B_{y_1})}{\mathrm{tr}\,(A_x \circ B_{y_1})}\right] = \frac{\mathrm{tr}\,(\rho A_x \circ B_{y_2})}{\mathrm{tr}\,(A_x \circ B_{y_2})} \ln\left[\frac{\mathrm{tr}\,(\rho A_x \circ B_{y_2})}{\mathrm{tr}\,(A_x \circ B_{y_2})}\right]$$

Extending this work to more than two observables, let $\mathcal{I}^1, \mathcal{I}^2, \dots, \mathcal{I}^{m-1}$ be instruments that measure the observables $A^1, A^2, \dots, A^{m-1}$, respectively. If A^m is another observable, we have that

$$(A^1 \circ A^2 \circ \cdots \circ A^m)_{(x_1,x_2,\dots,x_m)} = (\mathcal{I}_{x_1}^1)^*(\mathcal{I}_{x_2}^2)^* \cdots (\mathcal{I}_{x_{m-1}}^{m-1})^*(A_{x_m}^m)$$

The next result follows from Corollary 8.

Corollary 9. *If $A^1, A^2, \dots, A^m$ are observables, then*

$$S_{A^1\circ A^2\circ\cdots\circ A^m}(\rho) \le S_{A^1\circ A^2\circ\cdots\circ A^{m-1}}(\rho)$$

for all $\rho \in \mathcal{S}(H)$.

If $\mathcal{I}$ is an instrument, let A be the unique observable that $\mathcal{I}$ measures so $\mathrm{tr}\,[\mathcal{I}_x(\rho)] = \mathrm{tr}\,(\rho A_x)$ for all $x \in \Omega_{\mathcal{I}}$ and $\rho \in \mathcal{S}(H)$. We define the *ρ-entropy* of $\mathcal{I}$ as $S_{\mathcal{I}}(\rho) = S_A(\rho)$. Since $A_x = \mathcal{I}_x^*(I)$ we have

$$\mathrm{tr}\,(A_x) = \mathrm{tr}\,[\mathcal{I}_x^*(I)] = \mathrm{tr}\,[\mathcal{I}_x(I)]$$

Hence,

$$S_{\mathcal{I}}(\rho) = S_A(\rho) = -\sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\frac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right] = -\sum_x \mathrm{tr}\,[\mathcal{I}_x(\rho)] \ln\left\{\frac{\mathrm{tr}\,[\mathcal{I}_x(\rho)]}{\mathrm{tr}\,[\mathcal{I}_x(I)]}\right\}$$

Now let $\mathcal{I}^1, \mathcal{I}^2, \ldots, \mathcal{I}^m$ be instruments and let $A^1, A^2, \ldots, A^m$ be the unique observables they measure, respectively. Denoting the composition of two instruments $\mathcal{I}, \mathcal{J}$ by $\mathcal{I} \circ \mathcal{J}$ we have

$$\begin{aligned}\mathrm{tr}\left[\mathcal{I}^m_{x_m} \circ \mathcal{I}^{m-1}_{x_{m-1}} \circ \cdots \circ \mathcal{I}^1_{x_1}(\rho)\right] &= \mathrm{tr}\left[\rho(\mathcal{I}^1_{x_1})^*(\mathcal{I}^1_{x_2})^* \cdots (\mathcal{I}^m_{x_m})^*(I)\right] \\ &= \mathrm{tr}\,(\rho A^1_{x_1} \circ A^2_{x_2} \circ \cdots \circ A^m_{x_m})\end{aligned}$$

Hence, the observable measured by $\mathcal{I}^m \circ \mathcal{I}^{m-1} \circ \cdots \circ \mathcal{I}^1$ is $A^1 \circ A^2 \circ \cdots \circ A^m$. It follows that

$$S_{\mathcal{I}^m \circ \mathcal{I}^{m-1} \circ \cdots \circ \mathcal{I}^1}(\rho) = S_{A^1 \circ A^2 \circ \cdots \circ A^m}(\rho)$$

We conclude that Theorems 1, 2 and 3 [1] follow from our results. Moreover, our proofs are simpler since they come from the more basic concept of ρ-entropy for effects.

Let A, B be observables on H and let $\mathcal{I}$ be an instrument that measures A. The corresponding sequential product becomes

$$(A \circ B)_{(x,y)} = \mathcal{I}^*_x(B_y) = A_x \circ B_y$$

The ρ-entropy of $A \circ B$ has the form

$$\begin{aligned}S_{A\circ B}(\rho) &= -\sum_{x,y} \mathrm{tr}\left[\rho(A \circ B)_{(x,y)}\right] \ln\left\{\frac{\mathrm{tr}\left[\rho(A\circ B)_{(x,y)}\right]}{\mathrm{tr}\left[(A\circ B)_{(x,y)}\right]}\right\} \\ &= -\sum_{x,y} \mathrm{tr}\left[\rho \mathcal{I}^*_x(B_y)\right] \ln\left\{\frac{\mathrm{tr}\left[\rho \mathcal{I}^*_x(B_y)\right]}{\mathrm{tr}\left[\mathcal{I}^*_x(B_y)\right]}\right\} \\ &= -\sum_{x,y} \mathrm{tr}\left[\mathcal{I}_x(\rho) B_y\right] \ln\left\{\frac{\left[\mathcal{I}_x(\rho) B_y\right]}{\mathrm{tr}\left[\mathcal{I}_x(I) B_y\right]}\right\}\end{aligned}$$

If $\mathcal{L}^A$ is the Lüders instrument $\mathcal{I}^A_x(\rho) = A_x^{1/2}\rho A_x^{1/2}$ we have $(A \circ B)_{(x,y)} = A_x^{1/2} B_y A_x^{1/2}$ and

$$S_{A\circ B}(\rho) = -\sum_{x,y} \mathrm{tr}\,(A_x^{1/2}\rho A_x^{1/2} B_y) \ln\left[\frac{\mathrm{tr}\,(A_x^{1/2}\rho A_x^{1/2} B_y)}{\mathrm{tr}\,(A_x B_y)}\right]$$

If $\mathcal{H}^{(A,\alpha)}$ is the Holevo instrument $\mathcal{H}^{(A,\alpha)}_x(\rho) = \mathrm{tr}\,(\rho A_x)\alpha_x$, $\alpha_x \in \mathcal{S}(H)$ we obtain

$$\begin{aligned}S_{A\circ B}(\rho) &= -\sum_{x,y} \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\alpha_x B_y) \ln\left[\frac{\mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\alpha_x B_y)}{\mathrm{tr}\,(A_x)\mathrm{tr}\,(\alpha_x B_y)}\right] \\ &= -\sum_{x,y} \mathrm{tr}\,(\rho A_x)\mathrm{tr}\,(\alpha_x B_y) \ln\left[\frac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right] \\ &= -\sum_{x} \mathrm{tr}\,(\rho A_x) \ln\left[\frac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right] = S_A(\rho)\end{aligned}$$

This also follows from Corollary 8 because

$$\frac{\mathrm{tr}\,(\rho A_x \circ B_y)}{\mathrm{tr}\,(A_x \circ B_y)} = \frac{\mathrm{tr}\,(\alpha_x B_y)\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(\alpha_x B_y)\mathrm{tr}\,(A_x)} = \frac{(\rho A_x)}{\mathrm{tr}\,(A_x)}$$

If A is an observable on H and B is an observable on K we form the *tensor product observable* $A \otimes B$ on $H \otimes K$ given by $(A \otimes B)_{(x,y)} = A_x \otimes B_y$ where $\Omega_{A\otimes B} = \Omega_A \times \Omega_B$ [12].

Lemma 1. *If $\rho_1 \in \mathcal{S}(H)$, $\rho_2 \in \mathcal{S}(K)$, then*

$$S_{A\circ B}(\rho_1 \otimes \rho_2) = S_A(\rho_1) + S_B(\rho_2)$$

Proof. From the definition of $A \otimes B$ we obtain

$$\begin{aligned}
S_{A\otimes B}(\rho_1 \otimes \rho_2) &= -\sum_{x,y} \mathrm{tr}\,(\rho_1 \otimes \rho_2 A_x \otimes B_y) \ln\left[\tfrac{\mathrm{tr}\,(\rho_1\otimes\rho_2 A_x \otimes B_y)}{\mathrm{tr}\,(A_x \otimes B_y)}\right] \\
&= -\sum_{x,y} \mathrm{tr}\,(\rho_1 A_x)\mathrm{tr}\,(\rho_2 B_y) \ln\left[\tfrac{\mathrm{tr}\,(\rho_1 A_x)\mathrm{tr}\,(\rho_2 B_y)}{\mathrm{tr}\,(A_x)\mathrm{tr}\,(B_y)}\right] \\
&= -\sum_{x,y} \mathrm{tr}\,(\rho_1 A_x)\mathrm{tr}\,(\rho_2 B_y) \ln\left[\tfrac{\mathrm{tr}\,(\rho_1 A_x)}{\mathrm{tr}\,(A_x)}\right] \\
&\quad -\sum_{x,y} \mathrm{tr}\,(\rho_1 A_x)\mathrm{tr}\,(\rho_2 B_y) \ln\left[\tfrac{\mathrm{tr}\,(\rho_2 B_y)}{\mathrm{tr}\,(B_y)}\right] \\
&= -\sum_{x} \mathrm{tr}\,(\rho_1 A_x) \ln\left[\tfrac{\mathrm{tr}\,(\rho_1 A_x)}{\mathrm{tr}\,(A_x)}\right] - \sum_{y} \mathrm{tr}\,(\rho_2 B_y) \ln\left[\tfrac{\mathrm{tr}\,(\rho_2 B_y)}{\mathrm{tr}\,(B_y)}\right] \\
&= S_A(\rho_1) + S_B(\rho_2)e
\end{aligned}$$

□

We conclude that A gives more information about ρ_1 than A and B give about $\rho_1 \otimes \rho_2$ and similarly for B.

A *measurement model* [2,3,9] is a 5-tuple $\mathcal{M} = (H, K, \nu, \sigma, P)$ where H is the *system* Hilbert space, K is the *probe* Hilbert space, ν is the *interaction* channel, $\sigma \in \mathcal{S}(K)$ is the initial *probe state* and P is the *probe observable* on K. We interpret $\mathcal{M}$ as an apparatus that is employed to measure an instrument and hence an observable. In fact, $\mathcal{M}$ measures the unique instrument $\mathcal{I}$ on H given by

$$\mathcal{I}_x(\rho) = \mathrm{tr}_K[\nu(\rho \otimes \sigma)(I \otimes P_x)]$$

In this way, a state $\rho \in \mathcal{S}(H)$ is input into the apparatus and combined with the initial state σ of the probe system. The channel ν interacts the two states and a measurement of the probe P is performed resulting in outcome x. The outcome state is reduced to H by applying the partial trace over K. Now $\mathcal{I}$ measures an unique observable A on H that satisfies

$$\mathrm{tr}\,(\rho A_x) = \mathrm{tr}\,[\mathcal{I}_x(\rho)] = \mathrm{tr}\,[\nu(\rho \otimes \sigma)(I \otimes P_x)] \tag{4}$$

The ρ-entropy of $\mathcal{I}$ becomes

$$S_{\mathcal{I}}(\rho) = S_A(\rho) = -\sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right]$$

where $\mathrm{tr}\,(\rho A_x)$ is given by (4). Of course, $S_{\mathcal{I}}(\rho) = S_A(\rho)$ gives the amount of information that a measurement by $\mathcal{M}$ provides about ρ. A closely related concept is the observable $I \otimes P$ and $S_{I\otimes P}[\nu(\rho \otimes \sigma)]$ also provides the amount of information that a measurement $\mathcal{M}$ provides about ρ. It follows from (4) that the distribution of A in the state ρ equals the distribution of $I \otimes P$ in the state $\nu(\rho \otimes \sigma)$. We now compare $S_A(\rho)$ and $S_{I\otimes P}[\nu(\rho \otimes \sigma)]$. Applying (4) gives

$$\begin{aligned}
&S_{I\otimes P}[\nu(\rho \otimes \sigma)] \\
&\quad = -\sum_x \mathrm{tr}\,[\nu(\rho \otimes \sigma)(I \otimes P_x)] \ln\left\{\tfrac{\mathrm{tr}\,[\nu(\rho\otimes\sigma)(I\otimes P_x)]}{\mathrm{tr}\,(I\otimes P_x)}\right\} \\
&\quad = -\sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(\rho A_x)}{n\mathrm{tr}\,(P_x)}\right] = -\sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(A_x)}{n\mathrm{tr}\,(P_x)}\,\tfrac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right] \\
&\quad = -\sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(\rho A_x)}{\mathrm{tr}\,(A_x)}\right] - \sum \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(A_x)}{n\mathrm{tr}\,(P_x)}\right] \\
&\quad = S_A(\rho) - \sum_x \mathrm{tr}\,(\rho A_x) \ln\left[\tfrac{\mathrm{tr}\,(A_x)}{n\mathrm{tr}\,(P_x)}\right]
\end{aligned}$$

It follows that $S_A(\rho) \leq S_{I\otimes P}[\nu(\rho\otimes\sigma)]$ if and only if

$$\sum_x \operatorname{tr}(\rho A_x) \ln\left[\frac{\operatorname{tr}(A_x)}{n\operatorname{tr}(P_x)}\right] \leq 0 \tag{5}$$

Now (5) may or may not hold depending on A, ρ and P. In many cases, P is atomic [2,9] and then

$$\ln\left[\frac{\operatorname{tr}(A_x)}{n\operatorname{tr}(P_x)}\right] = \ln\left[\frac{\operatorname{tr}(A_x)}{n}\right] < 0$$

so $S_A(\rho) \leq S_{I\otimes P}[\nu(\rho\otimes\sigma)]$ for all $\rho \in \mathcal{S}(H)$. Also, (5) holds if P is sharp.

Funding: This research received no external funding.

Conflicts of Interest: The author declare no conflict of interest.

References

1. Šafránek, D.; Thingna, J. Quantifying information extraction using generalized quantum measurements. *arXiv* **2022**, arXiv:2007.07246.
2. Heinosaari, T.; Ziman, M. *The Mathematical Language of Quantum Theory*; Cambridge University Press: Cambridge, UK, 2012.
3. Nielson, M.; Chuang, I. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2000.
4. Ohya, M.; Petz, D. *Quantum Entropy and It's Uses*; Springer: Berlin/Heidelberg, Germany, 2004.
5. Šafránek, D.; Aguirre, A.; Schindler, J.; Deutsch, J. A brief introduction to observational entropy. *Found. Phys.* **2021**, *51*, 101. [CrossRef]
6. Lindblad, G. Completely positive maps and entropy inequalities. *Comm. Math. Phys.* **1975**, *40*, 147–151. [CrossRef]
7. von Neumann, J. *Mathematical Foundations of Quantum Mechanics*; Princeton University Press: Princeton, NJ, USA, 1955.
8. Wehrl, A. General properties of entropy. *Rev. Mod. Phys.* **1978**, *50*, 221. [CrossRef]
9. Busch, P.; Lahti, P.; Mittlestaedt, P. *The Quantum Theory of Measurement*; Springer: Berlin/Heidelberg, Germany, 1996.
10. Kraus, K. *States, Effects and Operations*; Springer: Berlin/Heidelberg, Germany, 1983.
11. Šafránek, D.; Deutsch, J.; Aguirre, A. Quantum coarse-grained entropy and thermodynamics. *Phys. Rev. A* **2019**, *99*, 010101. [CrossRef]
12. Gudder, S. Combinations of quantum observables and instruments. *J. Phys. A Math. Theor.* **2021**, *54*, 364002. [CrossRef]
13. Gudder, S. Sequential products of Quantum measurements. *arXiv* **2021**, arXiv:2108.07925.
14. Lüders, G. Über due Zustandsänderung durch den Messprozess. *Ann. Physik* **1951**, *6*, 322–328.
15. Holevo, A. *Probabilistic and Statistical Aspects of Quantum Theory*; North-Holland: Amsterdam, The Netherlands, 1982.
16. Gudder, S. Coarse-graining of observables. *Quant. Rep.* **2022**, *4*, 401–417. [CrossRef]