Chapter 5

Entropy and Mutual Information on a Quantum Logic

5.1 Introduction

The present chapter deals with the study of entropy of subsystems of a particular quantum dynamical system \((L, s, \phi)\), where \(L\) is an orthomodular lattice and \(s\) is a Bayesian state on it. Relative entropy, mutual information and their properties in a quantum space \((L, s)\), is also investigated. A mathematical model for noncompatible random events in quantum statistical mechanics, termed as quantum logic, was proposed by Birkhoff and von Neumann [16]. Quantum logic is a set of axioms that arose from the specific quantum logic of projection operators on a Hilbert space where the distributive law associated with a Boolean algebra fails to hold. Indeed, the lattice of projection operators on a Hilbert space always satisfies the orthomodular law which is a weaker form of the distributive law. The notion of entropy of partitions in the context of Boolean algebras is a useful tool in studying dynamical systems and has been applied on many others structures [7, 112, 120]. A study of the entropy of partitions on orthomodular lattice (OML) using the notion of a state (or a probability measure) can be seen in [143], and that of conditional
entropy of partitions on the same structure is done in [144]. Relative entropy plays an important role, as a mathematical device, in the stability analysis of master equations [128] and Fokker-Planck equations [126], and in isothermal equilibrium fluctuations and transient nonequilibrium deviations [113] (see also [36, 126]).

The present chapter is organized as follows: Prerequisites for the chapter are collected in Section 5.2. In Section 5.3, quantum dynamical systems of orthomodular lattices with Bayesian state are considered, and using the theory of entropy developed in [63, 64], it is proved that the entropy of \( \Phi_1 = (L_1, s_1, \phi_1) \) which is a subsystem of \( \Phi = (L, s, \phi) \), can be calculated in terms of those partitions of \((L, s)\) that belong to \( \varphi(L_1) \), where \( \varphi \) is corresponding embedding. Finally, it is proved that the entropy of the supremum of a separable spectrum of quantum dynamical systems is same as the limit of entropies of the subsystems under suitable conditions. In Section 5.4 we introduce and study the notion of relative entropy for a given partition of a quantum space. Various useful properties of relative entropy are proved and its relation with mutual information is explored. In particular, convexity of relative entropy with respect to Bayesian states, chain rules for entropy and that for mutual information are established. The data processing inequality in the framework of quantum logic is established, that may form foundation for the corresponding theory of sufficient statistics.

5.2 Preliminaries

5.2.1 [137, 143] Let \( L \) be an orthomodular lattice (or OML), i.e. \( L \) is an orthomodular poset (OMP) which is also a lattice. A map \( s : L \rightarrow [0, 1] \) is called state on \( L \) if it satisfies following conditions:

(i) \( s(1) = 1; \)
(ii) \( s(a \lor b) = s(a) + s(b) \), for all \( a, b \in L \) with \( a \perp b \).

From the definition of state we obtain that (i) \( s(0) = 0 \), (ii) \( s(a) = 1 - s(a') \) and (iii) \( s \) is monotone: Let \( a, b \in L \) with \( a \leq b \). Then by the orthomodular law, we get \( b = a \lor (a' \land b) \), since \( a \perp a' \land b \), therefore \( s(b) = s(a) + s(a' \land b) \), which yields that \( s(a) \leq s(b) \).

5.2.2 [63, 64, 143] A collection \( \mathcal{A} = \{a_i : i = 1, 2, \ldots, n\} \) \( (n \in \mathbb{N}) \) in \( L \) is called a partition of couple \((L, s)\) if

(i) \( \left( \bigvee_{i=1}^{k} a_i \right) \perp a_{k+1} \) for \( k = 1, 2, \ldots, n - 1 \),

(ii) \( s(\bigvee_{i=1}^{n} a_i) = 1 \).

If \( \mathcal{A} \) satisfies condition (i), then \( \mathcal{A} \) is called an orthogonal system.

A state \( s \) on an OML \( L \) is called Bayesian if for any partition of \( \mathcal{A} = \{a_i : i = 1, 2, \ldots, n\} \) \( (n \in \mathbb{N}) \)

\[
\frac{\bigvee_{i=1}^{n} (b \land a_i)}{s} = s(a).
\]

holds for all \( b \in L \) and in this case the couple \((L, s)\) is termed as quantum space.

Let \( \mathcal{A} = \{a_1, a_2, \ldots, a_n\} \) and \( \mathcal{B} = \{b_1, b_2, \ldots, b_m\} \) be partitions of a quantum space \((L, s)\). The common refinement of \( \mathcal{A} \) and \( \mathcal{B} \) defined by \( \mathcal{A} \lor \mathcal{B} := \{a_i \land b_j : a_i \in \mathcal{A}, b_j \in \mathcal{B}; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\} \) is also a partition of \((L, s)\). The partition \( \mathcal{B} \) is called an \( s \)-refinement of \( \mathcal{A} \) (written as \( \mathcal{A} \leq_s \mathcal{B} \)) if, for each \( b_j \in \mathcal{B}, j = 1, 2, \ldots, m \), there exists an \( a_i \in \mathcal{A}, i = 1, 2, \ldots, n \), such that \( s(b_j \land a_i) = s(b_j) \).

5.2.3 An OML-homomorphism \( \phi \) is a map \( \phi : L \to L \) on an OML \( L \) that preserves bounds, orthocomplementation and finite joins, i.e. for all \( a, b \in L \)

(i) \( \phi(1) = 1 \);

(ii) \( \phi(a') = (\phi(a))' \);
(iii) $\phi(a \lor b) = \phi(a) + \phi(b)$.

An OML-homomorphism $\phi$ is called state preserving if $s(a) = s(\phi(a))$ for all $a \in L$, where $s$ is a state on OML $L$. A quantum space together with state preserving OML-homomorphism is referred to as a quantum dynamical system (notice that here in the case of an OML, an additional condition of being Bayesian on $s$ is assumed, cf. Definition 4.3.1).

5.2.4 [63, 64, 143] Let $A = \{a_i : i = 1, 2, \ldots, n\}$ be a partition of a couple $(L, s)$. Then the entropy $H_s(A)$ of $A$ is

$$H_s(A) := -\sum_{i=1}^{n} g(s(a_i)),$$

where $g : [0, \infty) \rightarrow \mathbb{R}$ is the convex function, popularly known as the Shannon’s function, given by $g(x) = x \log x$, if $x > 0$ and $g(0) = 0$.

Let $\Phi = (L, s, \phi)$ be a quantum dynamical system. The entropy $h_s(\phi, A)$ of $\phi$ with respect to $A = \{a_i : i = 1, 2, \ldots, n\}$ is given by

$$h_s(\phi, A) := \lim_{n \to \infty} \frac{1}{n} H_s(\bigvee_{i=0}^{n-1} \phi^i(A)),$$

and the entropy $h_s(\Phi)$ (or $h_s(\phi)$) of $\Phi$ is

$$h_s(\Phi) := \sup h_s(\phi, A),$$

where supremum is taken over all finite partitions $A$ of $(L, s)$.

5.3 Entropy of subsystems

Throughout this section we consider particular quantum dynamical systems of orthomodular lattices. Let us recall the following results:

† Contents of Section 5.3 (along with a major portion of Chapter 4) with title “Spectrum of quantum dynamical systems: Subsystems and entropy” have appeared in Adv. Pure Appl. Math., de Gruyter GmbH, Germany 5(2) (2014), 75-83. doi 10.1515/apam-2013-0034
Chapter 5  Entropy and Mutual Information on a Quantum Logic

Proposition 5.3.1 [64] Let \((L, s, \phi)\) be a quantum dynamical system and let \(\mathcal{A}\) and \(\mathcal{B}\) be partitions of \((L, s)\). Then following statements hold:

(i) \(\mathcal{A} \leq \mathcal{B}\) implies \(h_s(\phi, \mathcal{A}) \leq h_s(\phi, \mathcal{B})\);

(ii) \(h_s(\phi, \mathcal{A}) = h_s(\phi, \bigvee_{i=0}^n \phi^i(\mathcal{A}))\). Furthermore, if \(\phi\) is invertible, then \(h_s(\phi, \mathcal{A}) = h_s(\phi, \bigvee_{i=-n}^n \phi^i(\mathcal{A}))\), \(n \in \mathbb{N}\).

Proposition 5.3.2 If \(\Phi_2 = (L_2, s_2, \phi_2) \sqsubseteq_{\varphi} \Phi_1 = (L_1, s_1, \phi_1)\), then we have the following:

(i) \(\mathcal{A}\) is a partition of \((L_2, s_2)\) implies \(\varphi(\mathcal{A})\) is a partition of \((\varphi(L_2), s_1)\);

(ii) \(H_{s_2}(\mathcal{A}) = H_{s_1}(\varphi(\mathcal{A}))\);

(iii) \(h_{s_2}(\phi_2, \mathcal{A}) = h_{s_1}(\phi_1, \varphi(\mathcal{A}))\);

(iv) \(h_{s_2}(\Phi_2) = \sup \{h_{s_1}(\phi, \varphi(\mathcal{A})) : \mathcal{A}\) is a partition of \((L_2, s_2)\}\);

(v) \(h_s(\Phi_2) \leq h_s(\Phi_1)\);

(vi) \(\Phi_1\) is weakly isomorphic to \(\Phi_2\) implies \(h_{s_1}(\Phi_1) = h_{s_2}(\Phi_2)\).

Proof. We shall prove only (i), (ii) and (iii). Since \(\Phi_2 \sqsubseteq_{\varphi} \Phi_1\) so for all \(a \in L_2\) we have \(s_2(a) = s_1(\varphi(a))\); and \((\varphi \circ \phi_2)(a) = (\phi_1 \circ \varphi)(a)\).

(i) Let \(\mathcal{A} = \{a_1, a_2, \ldots, a_n\}\) be a partition of \((L_2, s_2)\). Then \(\varphi(\mathcal{A}) = \{\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)\}\) is an orthogonal system in \(L_1\) as \(\varphi\) is a morphism. Since \(\varphi\) is a state preserving map, we have \(s_2(a_i) = s_1(\varphi(a_i)), i = 1, 2, \ldots, n\). Then \(s_1(\bigvee_{i=1}^n \varphi(a_i)) = \sum_{i=1}^n s_1(\varphi(a_i)) = \sum_{i=1}^n s_2(a_i) = s_2(\bigvee_{i=1}^n a_i) = 1\). Thus \(\varphi(\mathcal{A})\) is a partition of \((L_1, s_1)\).

(ii) If \(\mathcal{A} = \{a_1, a_2, \ldots, a_n\}\) is a partition of \((L_2, s_2)\), then \(H_{s_2}(\mathcal{A}) = -\sum_{i=1}^n s_2(a_i) \log s_2(a_i) = -\sum_{i=1}^n s_1(\varphi(a_i)) \log s_1(\varphi(a_i)) = H_{s_1}(\varphi(\mathcal{A}))\).

(iii) Let \(\mathcal{A}\) and \(\mathcal{B}\) be partitions of \((L_2, s_2)\). Since \(s_1\) is a Bayesian state, therefore \(\varphi(\mathcal{A}) \vee \varphi(\mathcal{B})\) is a partition of \((L_1, s_1)\) and it is easy to check that \(\varphi(\mathcal{A} \vee \mathcal{B}) = \varphi(\mathcal{A}) \vee \varphi(\mathcal{B})\). By (i), (ii) and the fact that \((\varphi \circ \phi_2) = (\phi_1 \circ \varphi)\), we have
\[
\begin{align*}
    h_{S_2}(\phi_2, A) &= \lim_{n \to \infty} \frac{1}{n} H_{S_2} \left( \bigvee_{i=0}^{n-1} \phi_2^i(A) \right) \\
    &= \lim_{n \to \infty} \frac{1}{n} H_{S_1} \left( \varphi \left( \bigvee_{i=0}^{n-1} \phi_2^i(A) \right) \right) = \lim_{n \to \infty} \frac{1}{n} H_{S_1} \left( \bigvee_{i=0}^{n-1} \varphi^i(\phi_2(A)) \right) \\
    &= \lim_{n \to \infty} \frac{1}{n} H_{S_1} \left( \bigvee_{i=0}^{n-1} \phi_1^i(\varphi(A)) \right) = h_{S_1}(\phi_1, \varphi(A)).
\end{align*}
\]

Let \( \varphi : (L_1, s_1) \to (L_2, s_2) \) be an isomorphism. Then every partition of \((L_2, s_2)\) can be expressed as \( \varphi(A) \), where \( A \) is a partition of \((L_1, s_1)\). Let \( \Phi_1 = (L_1, s_1, \phi_1) \sqsubseteq \varphi \Phi = (L, s, \varphi) \). So, Proposition 5.3.2(iv) shows that the entropy \( h_{S_1}(\Phi_1) \) of \( \Phi_1 \) can be calculated in terms of those partitions of \((L, s)\) which belong to \( \varphi(L_1) \). Thus we obtain the following result:

**Proposition 5.3.3** Let \( \Phi_1 = (L_1, s_1, \phi_1) \sqsubseteq \varphi \Phi = (L, s, \varphi) \). Then

\[
    h_{S_1}(\Phi_1) = \sup \{ h_s(\phi, A) : A \text{ is a partition of } (\varphi(L_1), s) \}.
\]

**Definition 5.3.1** A subsystem \( \Phi_1 \) of a quantum dynamical system \( \Phi = (L, s, \phi) \) is called dense in \( \Phi \) if there exists a partition \( A \) of \( \Phi_1 \) such that, for any partition \( B \) of \( \Phi \), there exists \( k \in \mathbb{N} \) such that

\[
    \begin{cases}
    B \leq_s \bigvee_{i=0}^{k} \phi^i(A), & \text{if } \phi \text{ is noninvertible}, \\
    B \leq_s \bigvee_{i=-k}^{k} \phi^i(A), & \text{if } \phi \text{ is invertible}.
    \end{cases}
\]

In this case we write \( \Phi_1 \sqsubseteq_A \Phi \).

A bounded spectrum \((P, \Phi_\alpha, \varphi_{\alpha\beta})\) of quantum dynamical systems is called separable if there exists \( \alpha_0 \in P \) such that \( \Phi_{\alpha_0} \) is dense in \( \Phi = \sup_{\alpha \in P} \Phi_\alpha \).

**Proposition 5.3.4** If \( \Phi_1 \sqsubseteq_A \Phi \), where \( \Phi = (L, s, \phi) \), then \( h_s(\Phi) = h_s(\phi, A) \).

**Proof.** Let \( B \) be a partition of \((L, s)\) and let \( \phi \) be noninvertible. Then by hypothesis, there exists \( k \in \mathbb{N} \) such that \( B \leq_s \bigvee_{i=0}^{k} \phi^i(A) \). By Proposition
5.3.1(i), (ii), we get
\[ h_s(\phi, B) \leq h_s(\phi, \bigvee_{i=0}^{k} \phi^i(A)) = h_s(\phi, A). \]

Taking supremum over all partitions \( B \) of \((L, s)\), we deduce that
\[ h_s(\Phi) = h_s(\phi, A). \]

In case \( \phi \) is invertible, the proof is analogous. \( \square \)

**Proposition 5.3.5** Let \((J, \Phi_\alpha, \varphi_{\alpha\beta})\) be a separable spectrum of quantum dynamical systems and \( \Phi = \sup_{\alpha \in J} \Phi_\alpha; \Phi = (L, s, \phi) \). Then
\[ h_s(\Phi) = \sup_{\alpha \in J} h_s(\Phi_\alpha) = \lim_{\alpha \in J} h_s(\Phi_\alpha). \]

**Proof.** By Remark 4.3.3 and Proposition 5.3.2(iv), we obtain that \( h(\Phi_\alpha) \) is a monotone net and therefore \( \sup_{\alpha \in J} h_s(\Phi_\alpha) \) exists and is equal to \( \lim_{\alpha \in J} h_s(\Phi_\alpha) \).

Moreover, \( h_s(\Phi_\alpha) \leq h_s(\Phi) \) for all \( \alpha \in J \) and so \( \sup_{\alpha \in J} h_s(\Phi_\alpha) \leq h_s(\Phi) \).

Since \((J, \Phi_\alpha, \varphi_{\alpha\beta})\) is a separable spectrum, so there exists \( \alpha_0 \in J \) such that \( \Phi_{\alpha_0} \sqsubseteq A \Phi \). By Proposition 5.3.4 we get \( h_s(\Phi) = h_s(\phi, A) \).

But
\[ h_s(\phi, A) \leq h_s(\Phi_{\alpha_0}) \leq \sup_{\alpha \in J} h_s(\Phi_\alpha). \]

Hence \( h_s(\Phi) \leq \sup_{\alpha \in J} h_s(\Phi_\alpha) \).

Thus we get \( h_s(\Phi) = \sup_{\alpha \in J} h_s(\Phi_\alpha) \). \( \square \)

5.4 Quantum relative entropy and mutual information†

Let \( A = \{a_1, a_2, \ldots, a_n\} \) and \( B = \{b_1, b_2, \ldots, b_m\} \) be partitions of a couple

† Contents of Section 5.4 with title “Relative entropy and mutual information on a quantum logic” have appeared in *Mathematica Aeterna*, 3(7) (2013), 555-563.
(L, s). Then \( \mathcal{A} \) and \( \mathcal{B} \) are called independent if, \( s(a_i \wedge b_j) = s(a_i)s(b_j) \), where \( i = 1, 2, \ldots, n; \ j = 1, 2, \ldots, m \).

The conditional entropy \( H_s(\mathcal{A}|\mathcal{B}) \) is defined by

\[
H_s(\mathcal{A}|\mathcal{B}) := -\sum_{j=1}^{m} \sum_{i=1}^{n} s(b_j) g(s(a_i|b_j)),
\]

where \( g \) is the Shannon’s function. Notice that \( H_s(\mathcal{A}|\mathcal{B}) \geq 0 \), and \( H_s(\mathcal{A}|\mathcal{A}) = 0 \).

We refer to [63, 64] where a study of entropy and conditional entropy of partitions of a couple (L, s) is made and its relation with the theory of commutators, Boolean quotients in orthomodular lattices [14, 89, 90, 107], and Bell inequalities [109, 111], is discussed.

Let us recall the following log sum inequality, which we shall use in the sequel to establish various results:

for nonnegative real numbers, \( x_1, x_2, \ldots, x_n \) and \( y_1, y_2, \ldots, y_n \),

\[
\sum_{i=1}^{n} x_i \log \frac{x_i}{y_i} \geq \left( \sum_{i=1}^{n} x_i \right) \log \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}; \quad (1)
\]

equality holds if and only if \( \frac{x_i}{y_i} \) corresponding to nonzero \( y_i \) are equal, i.e. \( \frac{x_i}{y_i} \) is constant. (Here we follow the convention that \( x \log \frac{x}{0} = \infty \) for \( x > 0 \), and \( 0 \log \frac{0}{0} = 0 \).

We would like to mention here that the log sum inequality can be obtained from the Jensen’s inequality (see [127]).

The following results from [63] are used in the sequel.

**Proposition 5.4.1** [63] Let \( \mathcal{A}, \mathcal{B} \) and \( \mathcal{C} \) be partitions of a quantum space (L, s). Then following statements hold:

(i) \( H_s(\mathcal{A} \vee \mathcal{B}) \leq H_s(\mathcal{A}) + H_s(\mathcal{B}) \); equality holds if \( \mathcal{A} \) and \( \mathcal{B} \) are independent partitions of (L, s).
(ii) $H_s(A|B) = H_s(A)$ if and only if $A$ and $B$ are independent partitions of $(L, s)$.

(iii) $H_s(A \lor B) = H_s(A) + H_s(B|A)$, and hence $H_s(A \lor B) \geq \max\{H_s(A), H_s(B)\}$.

(iv) $H_s(A \lor B|C) = H_s(A|C) + H_s(B|A \lor C)$.

(v) $H_s(A|B \lor C) \leq H_s(A|B)$.

**Proposition 5.4.2** (Concavity of entropy). Let $L$ be an OML and $r$ and $s$ be states on it. If $A$ is a partition of $L$ corresponding to $r$ and $s$, then for $\alpha \in [0, 1]$, we have

$$\alpha H_s(A) + (1 - \alpha) H_r(A) \leq H_{\alpha s + (1 - \alpha) r}(A),$$

showing that $H_s(A)$ is a concave function of $s$.

**Proof.** Let $A = \{a_1, a_2, \ldots, a_n\}$ be a partition of $(L, r)$ and $(L, s)$. Since $g$ is convex, we have

$$\alpha H_s(A) + (1 - \alpha) H_r(A) = -\alpha \sum_{i=1}^{n} g(s(a_i)) - (1 - \alpha) \sum_{i=1}^{n} g(r(a_i))$$

$$= -\sum_{i=1}^{n} [\alpha g(s(a_i)) + (1 - \alpha) g(r(a_i))] \leq -\sum_{i=1}^{n} g(\alpha s(a_i) + (1 - \alpha) r(a_i))$$

$$= -\sum_{i=1}^{n} g(\alpha s + (1 - \alpha) r)(a_i) = H_{\alpha s + (1 - \alpha) r}(A). \quad \square$$

**Proposition 5.4.3** (Chain rules for entropy). Let $A_1, A_2, \ldots, A_n$ (n $\in$ N), and $C$ be partitions of a quantum space $(L, s)$. Then

(i) $H_s(A_1 \lor A_2 \lor \cdots \lor A_n) = \sum_{i=1}^{n} H_s(A_i | (A_{i-1} \lor \cdots \lor A_1))$.

(ii) $H_s(\bigvee_{i=1}^{n} A_i | C) = \sum_{i=1}^{n} H_s(A_i | \bigvee_{k=1}^{i-1} A_k \lor C)$.

(iii) $H_s(A_1 \lor A_2 \lor A_3) \leq H_s(A_1 \lor A_2) + H_s(A_1 \lor A_3) - H_s(A_1)$. 
Proof. (i). By Proposition 5.4.1(iii) and (iv), we have $H_s(A_1 \lor A_2) = H_s(A_1) + H_s(A_2 | A_1)$. Now suppose that the result is true for a specific value of $n \in \mathbb{N}$. Then

$$H_s(A_1 \lor A_2 \lor \cdots \lor A_n \lor A_{n+1})$$

$$= H_s(A_1 \lor A_2 \lor \cdots \lor A_n) + H_s(A_{n+1} | (A_1 \lor A_2 \lor \cdots \lor A_n))$$

$$= \sum_{i=1}^{n} H_s(A_i | (A_{i-1} \lor \cdots \lor A_1)) + H_s(A_{n+1} | (A_n \lor \cdots \lor A_1))$$

$$= \sum_{i=1}^{n+1} H_s(A_i | (A_{i-1} \lor \cdots \lor A_1)).$$

Proof of (ii) follows similarly, using Proposition 5.4.1(iv) inductively.

(iii) By Proposition 5.4.1(iii) and (v), we get

$$H_s(A_1 \lor A_2 \lor A_3) = H_s(A_1 \lor A_2) + H_s(A_3 | A_1 \lor A_2)$$

$$\leq H_s(A_1 \lor A_2) + H_s(A_3 | A_1)$$

$$= H_s(A_1 \lor A_2) + H_s(A_1 \lor A_3) - H_s(A_1). \quad \square$$

**Definition 5.4.1** Let $s_1$ and $s_2$ be states on an OML $L$, and let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ be a partition of $L$ corresponding to both $s_1$ and $s_2$. Then the relative entropy $D_{\mathcal{A}}(s_1 \parallel s_2)$ is defined by

$$D_{\mathcal{A}}(s_1 \parallel s_2) := \sum_{i=1}^{n} s_1(a_i) \log \frac{s_1(a_i)}{s_2(a_i)}.$$  

The following result suggests interpretation of relative entropy as a distance between two states, i.e. a measure of how different the two states are. Due to non-availability of symmetry and the triangle inequality, it is not a metric in a true sense.
Proposition 5.4.4 If $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ is a partition of $L$ corresponding to states $s_1$ and $s_2$, then $D_\mathcal{A}(s_1 \parallel s_2) \geq 0$, with equality if and only if $s_1(a_i) = s_2(a_i)$, for each $i \in \{1, 2, \ldots, n\}$.

Proof. In the log sum inequality, let $x_i = s_1(a_i)$ and $y_i = s_2(a_i)$ for $i \in \{1, 2, \ldots, n\}$. Then $\sum_{i=1}^n x_i = \sum_{i=1}^n s_1(a_i) = s_1(\bigvee_{i=1}^n a_i) = 1$. Similarly, $\sum_{i=1}^n y_i = 1$. Hence $D_\mathcal{A}(s_1 \parallel s_2) \geq 0$. Also, $D_\mathcal{A}(s_1 \parallel s_2) = 0$ if and only if $s_1(a_i) = \alpha s_2(a_i)$ for all $i$, where $\alpha$ is a constant. Summing over all $i$, we get $\alpha = 1$. Thus $D_\mathcal{A}(s_1 \parallel s_2) = 0$ if and only if $s_1(a_i) = s_2(a_i)$, for all $i$. $\square$

Proposition 5.4.5 Let $\mathcal{A}$ be a partition of $(L, s)$. The relative entropy $D_\mathcal{A}(s_1 \parallel s_2)$ is convex in the pair $(s_1, s_2)$, i.e. if $(s'_1, s'_2)$, $(s''_1, s''_2)$ are pairs of states on $L$, then

$$D_\mathcal{A}((\alpha s'_1 + (1-\alpha)s''_1) \parallel (\alpha s'_2 + (1-\alpha)s''_2)) \leq \alpha D_\mathcal{A}(s'_1 \parallel s'_2) + (1-\alpha)D_\mathcal{A}(s''_1 \parallel s''_2),$$

for all $\alpha \in [0, 1]$.

Proof. Fix $i \in \{1, 2, \ldots, n\}$. Putting $x_1 = \alpha s'_1(a_i)$, $x_2 = (1-\alpha)s''_1(a_i)$, $y_1 = \alpha s'_2(a_i)$, and $y_2 = (1-\alpha)s''_2(a_i)$ in the log sum inequality, we get

$$(\alpha s'_1(a_i) + (1-\alpha)s''_1(a_i)) \log \frac{\alpha s'_1(a_i) + (1-\alpha)s''_1(a_i)}{\alpha s'_2(a_i) + (1-\alpha)s''_2(a_i)} \leq \alpha s'_1(a_i) \log \frac{\alpha s'_1(a_i)}{\alpha s'_2(a_i)} + (1-\alpha)s''_1(a_i) \log \frac{(1-\alpha)s''_1(a_i)}{(1-\alpha)s''_2(a_i)}.$$ 

Summing these inequalities over all $i$, the result follows. $\square$

Definition 5.4.2 Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ and $\mathcal{B} = \{b_1, b_2, \ldots, b_m\}$ be partitions of $(L, s)$. Define mutual information as

$$I(\mathcal{A} : \mathcal{B}) := \sum_{j=1}^m \sum_{i=1}^n s(a_i \land b_j) \log \frac{s(a_i \land b_j)}{s(a_i)s(b_j)}.$$
Notice that $I(A : B) = I(B : A)$. Also, if $A$ and $B$ are independent, then $I(A : B) = 0$.

**Proposition 5.4.6** Let $A$ and $B$ be partitions of a quantum space $(L, s)$. Then

$$I(A : B) = H_s(A) - H_s(A | B) = H_s(A) + H_s(B) - H_s(A \vee B).$$

Consequently, $I(A : B) \geq 0$, and $I(A : A) = 0$.

**Proof.** By Proposition 5.4.1(iii), we have

$$I(A : B) = \sum_{j=1}^{m} \sum_{i=1}^{n} s(a_i \land b_j) \log \frac{s(a_i \land b_j)}{s(a_i)s(b_j)}$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} s(a_i \land b_j) \log \frac{s(a_i \land b_j)}{s(b_j)} - \sum_{j=1}^{m} \sum_{i=1}^{n} s(a_i \land b_j) \log s(a_i)$$

$$= -H_s(A | B) - \sum_{j=1}^{m} \sum_{i=1}^{n} s(a_i \land b_j) \log s(a_i)$$

$$= -H_s(A | B) - \sum_{i=1}^{n} s(a_i) \log s(a_i)$$

$$= H_s(A) - H_s(A \vee B).$$

$$= H_s(A) + H_s(B) - H_s(A \vee B). \quad \square$$

**Definition 5.4.3** Let $A, B$ and $C$ be partitions of a quantum space $(L, s)$. The conditional mutual information of $A$ and $B$ given $C$ is defined by

$$I(A : B | C) := H_s(A | C) - H_s(A | (B \lor C)).$$
Proposition 5.4.8 (Chain rule for mutual information). If $A_1, A_2, \ldots, A_n$ ($n \in \mathbb{N}$), and $B$ are partitions of a quantum space $(L, s)$, then

$$I(\bigvee_{i=1}^{n} A_i : B) = \sum_{i=1}^{n} I(A_i : B | \bigvee_{k=1}^{i-1} A_k).$$

Proof. By Proposition 5.4.3 and Proposition 5.4.6, we have

$$I(\bigvee_{i=1}^{n} A_i : B) = H_s(\bigvee_{i=1}^{n} A_i) - H_s(\bigvee_{i=1}^{n} A_i | B)$$

$$= \sum_{i=1}^{n} H_s(A_i | \bigvee_{k=1}^{i-1} A_k) - \sum_{i=1}^{n} H_s(A_i | \bigvee_{k=1}^{i-1} A_k \vee B)$$

$$= \sum_{i=1}^{n} \left( H_s(A_i | \bigvee_{k=1}^{i-1} A_k) - H_s(A_i | \bigvee_{k=1}^{i-1} A_k \vee B) \right)$$

$$= \sum_{i=1}^{n} I(A_i : B | \bigvee_{k=1}^{i-1} A_k). \quad \Box$$

Definition 5.4.4 Let $A, B$ and $C$ be partitions of $(L, s)$. Then $A$ is called conditionally independent to $B$ given $C$ (written as $A \rightarrow B \rightarrow C$) if $I(A : C | B) = 0$.

Proposition 5.4.9 $A \rightarrow B \rightarrow C$ if and only if $C \rightarrow B \rightarrow A$.

Proof. Let $A \rightarrow B \rightarrow C$. Then $0 = I(A : C | B) = H_s(A | B) - H_s(A | B \vee C)$. Then by Proposition 5.4.1(iii) we have, $H_s(A | B) = H_s(A | B \vee C) = H_s(A \vee B \vee C) - H_s(B \vee C)$.

Now, again by Proposition 5.4.1(iii), $I(C : A | B) = H_s(C | B) - H_s(C | A \vee B) = H_s(B \vee C) - H_s(B) - H_s(A \vee B \vee C) + H_s(A \vee B) = H_s(A \vee B) - H_s(B) - H_s(A | B) = 0$, i.e. $C \rightarrow B \rightarrow A. \quad \Box$

Remark 5.3.1 In view of the above theorem we may write $A \leftrightarrow B \leftrightarrow C$ for $A \rightarrow B \rightarrow C$ and we may say that $A$ and $C$ are conditionally independent given $B$. 
Proposition 5.4.10 For any partition $A, B, C$ of $(L, s)$, we have

$$I(A : B \lor C) = I(A : B) + I(A : C|B) = I(A : C) + I(A : B|C).$$

Proof. By Proposition 5.4.6 we get,

$$I(A : B) + I(A : C|B) = H_s(A) - H_s(A|B) + H_s(A|B) - H_s(A|B \lor C) = H_s(A) - H_s(A|B \lor C) = I(A : B \lor C).$$

Proposition 5.4.11 Let $A \rightarrow B \rightarrow C$. Then

(i) $I(A \lor B : C) = I(B : C)$;

(ii) $I(B : C) = I(A : C) + I(C : B | A)$;

(iii) $I(A : B|C) \leq I(A : B)$. (Data Processing Inequality)

Proof. (i) Let $A \rightarrow B \rightarrow C$, i.e. $I(A : C|B) = 0$. So, by the chain rule for mutual information, we have $I(A \lor B : C) = I(B \lor A : C) = I(B : C) + I(A : C|B) = I(B : C)$.

(ii) By Proposition 5.4.10, we have $I(A \lor B : C) = I(C : B \lor C) = I(C : A) + I(C : B|A) = I(A : C) + I(C : B|A)$. Using (i), it follows that $I(B : C) = I(A : C) + I(C : B | A)$.

(iii) It follows from (ii) that, $I(C : B|A) \leq I(B : C) = I(C : B)$. In view of Proposition 5.4.9, interchanging $A$ and $C$, we get $I(A : B|C) \leq I(A : B)$. □