EPJ Web of Conferences **182**, 02039 (2018) https://doi.org/10.1051/epjconf/201818202039
ICNFP 2017

Introduction to Quantum Logical Information Theory: Talk

David Ellerman[1,a]

[1] *University of California Riverside*

Abstract. Logical information theory is the quantitative version of the logic of partitions just as logical probability theory is the quantitative version of the dual Boolean logic of subsets. The resulting notion of information is about distinctions, differences, and distinguishability, and is formalized using the distinctions ("dits") of a partition (a pair of points distinguished by the partition). All the definitions of simple, joint, conditional, and mutual entropy of Shannon information theory are derived by a uniform transformation from the corresponding definitions at the logical level. The purpose of this talk is to outline the direct generalization to quantum logical information theory that similarly focuses on the pairs of eigenstates distinguished by an observable, i.e., "qudits" of an observable. The fundamental theorem for quantum logical entropy and measurement establishes a direct quantitative connection between the increase in quantum logical entropy due to a projective measurement and the eigenstates (cohered together in the pure superposition state being measured) that are distinguished by the measurement (decohered in the post-measurement mixed state). Both the classical and quantum versions of logical entropy have simple interpretations as "two-draw" probabilities for distinctions. The conclusion is that quantum logical entropy is the simple and natural notion of information for quantum information theory focusing on the distinguishing of quantum states.

1 Introduction

Subsets and partitions are category-theoretic dual concepts. The elements of subsets are dual to the distinctions (dits) of a partition, i.e., an ordered pair of elements in different blocks. Thus the usual

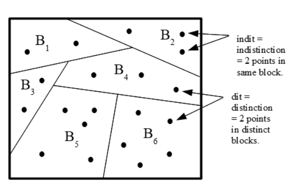

Figure 1. Partition $\pi = \{B_1, ..., B_6\}$ on set $U = \{u_1, ..., u_n\}$

Boolean logic of subsets (often taken as "propositional logic") has a dual logic of partitions [2] where elements of subsets and distinctions of partition have corresponding roles. Gian-Carlo Rota said

[a] e-mail: david@ellerman.org

Table 1	Subset Logic	Partition Logic
Logic of...	Subsets S ⊆ U	Partitions π on U
Elements (its or dits)	*Elements* u of a subset S	*Distinctions* (u,u') of a partition π
All elements	Universe set U (all elements)	Discrete partition **1** (all dits)
No elements	Empty set ∅ (no elements)	Indiscrete partition **0** (no dits)
Partial order on...	S ⊆ T = Inclusion of elements	Refinement σ ≼ π of partitions = inclusion of distinctions: dit(σ) ⊆ dit(π)
Formula variables	Subsets of U	Partitions on U
Logical operations ∪, ∩, ⇒,...	Operations on subsets	Operations on partitions
Propositional interp. of Φ(π,σ,...)	Subset Φ(π,σ,...) contains an element u.	Partition Φ(π,σ,...) makes a distinction (u,u').
Valid formula Φ(π,σ,...)	Φ(π,σ,...) = U for any subsets π,σ,... of any U (|U| ≥ 1), i.e., contains all elements u.	Φ(π,σ,...) = **1** for any partitions π,σ,... on any U (|U| ≥ 2), i.e., makes all distinctions (u,u').

Figure 2. Dual Logics: Boolean subset logic and partition logic.

"Probability is a measure on the Boolean algebra of events" that gives quantitatively the "intuitive idea of the size of a set", so we may ask by "analogy" for some measure to capture a property for a partition like "what size is to a set." Rota goes on to ask:

> How shall we be led to such a property? We have already an inkling of what it should be: it should be a measure of information provided by a random variable. Is there a candidate for the measure of the amount of information? [7, p. 67]

Since elements are to subsets as distinctions are to partitions, the "size" of a partition may be taken as the number of distinctions.

2 Logical Information Theory

The new logical foundations for information theory starts with sets, not probabilities, as suggested by Andrei Kolmogorov.

> Information theory must precede probability theory, and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character. [4, p. 39]

The notion of information-as-distinctions thus starts with the *set of distinctions*, the *information set*, of a partition $\pi = \{B, B', ...\}$ on a finite set U where that set of distinctions (dits) is:

$$dit(\pi) = \{(u, u') : \exists B, B' \in \pi, B \neq B', u \in B, u' \in B'\}. \tag{1}$$

The ditset of a partition is the complement in $U \times U$ of the equivalence relation associated with the partition π. Given any probability measure $p : U \to [0, 1]$ on $U = \{u_1, ..., u_n\}$ which defines $p_i = p(u_i)$ for $i = 1, ..., n$, the *product measure* $p \times p : U \times U \to [0, 1]$ has for any $S \subseteq U \times U$ the value of:

$$p \times p(S) = \sum_{(u_i,u_j)\in S} p(u_i)\, p\left(u_j\right) = \sum_{(u_i,u_j)\in S} p_i p_j. \tag{2}$$

The *logical entropy* of π is the product measure of its ditset:

$$h(\pi) = p \times p(dit(\pi)) = \sum_{(u_i,u_j)\in dit(\pi)} p_i p_j = 1 - \sum_{B\in\pi} p(B)^2. \tag{3}$$

Table 2	Logical Probability Theory	Logical Information Theory
'Outcomes'	Elements $u \in U$ finite	Distinctions $(u,u') \in U \times U$ finite
'Events'	Subsets $S \subseteq U$	Ditsets $dit(\pi) \subseteq U \times U$
Equiprobable outcomes	$Pr(S) = \lvert S\rvert / \lvert U\rvert$ = logical probability of event S	$h(\pi) = \lvert dit(\pi)\rvert / \lvert U \times U\rvert$ = logical entropy of partition π
Point probabilities	$Pr(S) = \Sigma\{p_j : u_j \in S\} = p(S)$ = logical prob. of event S	$h(\pi) = \Sigma\{p_j p_k : (u_j,u_k) \in dit(\pi)\}$ = logical entropy of π
Interpretation	Pr(S) = one draw probability of getting an element from S	$h(\pi)$ = two draw probability (w/replacement) of getting a distinction of π

Figure 3. $\frac{\text{Probability}}{\text{Subset Logic}} = \frac{\text{Information}}{\text{Partition Logic}}$

Given partitions $\pi = \{B_1, ..., B_I\}, \sigma = \{C_1, ..., C_J\}$ on U, the *information set* or *ditset for their join* is:

$$dit(\pi \vee \sigma) = dit(\pi) \cup dit(\sigma) \subseteq U \times U. \tag{4}$$

Given probabilities $p = \{p_1, ..., p_n\}$, the *joint logical entropy* is:

$$h(\pi, \sigma) = h(\pi \vee \sigma) = p \times p(dit(\pi) \cup dit(\sigma)) = 1 - \sum_{i,j} p\left(B_i \cap C_j\right)^2. \tag{5}$$

The infoset for the *conditional logical entropy* $h(\pi|\sigma)$ is the difference of ditsets, and thus: $h(\pi|\sigma) = p \times p(dit(\pi) - dit(\sigma))$.

The infoset for the *logical mutual information* $m(\pi, \sigma)$ is the intersection of ditsets, so: $m(\pi, \sigma) = p \times p(dit(\pi) \cap dit(\sigma))$. The *Information algebra* $I(\pi, \sigma)$ = Boolean subalgebra of $\wp(U \times U)$ generated by ditsets and their complements.

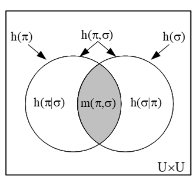

Figure 4. Venn Diagram for Logical Entropies

3 Deriving the Shannon entropies from the logical entropies

The simple and compound definitions for Shannon entropy are often presented as satisfying a Venn diagram, but they are not a measure (in the sense of measure theory). However, all the Shannon entropies can be derived from the definitions of logical entropy (which is a measure) by a uniform transformation that preserves the Venn diagram relationships. That "dit-bit transformation" starts with logical entropy measured in dits and requantifies in terms of bits to obtain the corresponding Shannon entropy.

In the canonical case of n equiprobable elements, $p_i = \frac{1}{n}$, the logical entropy of $p = \left\{\frac{1}{n}, ..., \frac{1}{n}\right\}$ is:

$$\frac{|U \times U - \Delta|}{|U \times U|} = \frac{n^2 - n}{n^2} = 1 - \frac{1}{n} = 1 - p_i. \tag{6}$$

In the general case $p = (p_1, ..., p_n)$, the logical entropy is the average of this dit-count $1 - p_i$:

$$h(p) = \sum_{i=1}^{n} p_i (1 - p_i). \tag{7}$$

In the canonical case of 2^n equiprobable elements so $p_i = \frac{1}{2^n}$, the minimum number of binary partitions ("yes-or-no questions") or "bits" it takes to uniquely determine or *encode* each distinct element or block is n, so the Shannon-Hartley entropy is:

$$n = \log_2 (2^n) = \log_2 \left(\frac{1}{1/2^n} \right) = \log_2 \left(\frac{1}{p_i} \right). \tag{8}$$

In the general case of $p = (p_1, ..., p_{2^n})$, the Shannon entropy is the average of this bit-count $\log_2 \left(\frac{1}{p_i} \right)$:

$$H(p) = \sum_{i=1}^{2^n} p_i \log_2 \left(\frac{1}{p_i} \right). \tag{9}$$

Hence the *Dit-Bit Transform* is: express any logical entropy concept (joint, conditional, or mutual) as average of dit-counts $1 - p_i$, and then substitute the bit-count $\log \left(\frac{1}{p_i} \right)$ to obtain the corresponding formula as defined by Shannon. For the corresponding definitions for random variables and their probability distributions, consider a random variable (x, y) taking values on the product $X \times Y$ of finite sets with the joint probability distribution $\{p_{xy}\}$, and thus with the marginal distributions $\{p_x\}$ and $\{p_y\}$ where $p_x = \sum_{y \in Y} p_{xy}$ and $p_y = \sum_{x \in X} p_{xy}$. The entropies can be considered as functions of the random variables or of their probability distributions, e.g., $h(\{p_{xy}\}) = h(x, y)$, $h(\{p_x\}) = h(x)$, and $h(\{p_y\}) = h(y)$. The dit-bit transform preserves the same Venn diagram formulas for the Shannon entropies in spite of them not being a measure (in the sense of measure theory).

Table 3	The Dit-Bit Transform: $1-p_i \rightarrow \log(1/p_i)$
Entropy	$h(p) = \Sigma_i p_i(1-p_i)$ $H(p) = \Sigma_i p_i(\log(1/p_i))$
Joint Entropy	$h(x,y) = \Sigma_{x,y} p_{xy}(1-p_{xy})$ $H(x,y) = \Sigma_{x,y} p_{xy}(\log(1/p_{xy}))$
Conditional Entropy	$h(x\|y) = \Sigma_{x,y} p_{xy}[(1-p_{xy}) - (1-p_y)]$ $H(x\|y) = \Sigma_{x,y} p_{xy}[(\log(1/p_{xy})) - (\log(1/p_y))]$
Mutual Information	$m(x,y) = \Sigma_{x,y} p_{xy}[(1-p_x) + (1-p_y) - (1-p_{xy})]$ $I(x,y) = \Sigma_{x,y} p_{xy}[(\log(1/p_x) + (\log(1/p_y) - (\log(1/p_{xy})]$

Figure 5. Venn-diagram-preserving dit-bit transform.

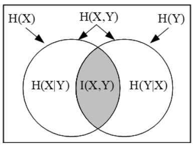

Figure 6. Venn diagram ‘mnemonic’ for Shannon entropies.

4 Logical entropy via density matrices

All this will carry over to quantum logical entropy using density matrices. 'Classically,' the *density matrix* representing the event S is the $n \times n$ symmetric real matrix:

$$\rho(S) = |S\rangle\langle S| = \begin{cases} \frac{1}{p(S)}\sqrt{p_j p_k} \text{ for } u_j, u_k \in S \\ 0 \text{ otherwise} \end{cases} \tag{10}$$

Then $\rho(S)^2 = |S\rangle\langle S|S\rangle\langle S| = \rho(S)$ so borrowing language from QM, $|S\rangle$ is said to be a *pure* state or event. Given any partition $\pi = \{B_1, ..., B_I\}$ on U, its density matrix is the average of the block density matrices:

$$\rho(\pi) = \sum_i p(B_i)\rho(B_i). \tag{11}$$

Then $\rho(\pi)$ represents the *mixed* state, experiment, or lottery where the event B_i occurs with probability $p(B_i)$.

Example 1 *For the throw of a fair die, $U = \{u_1, u_3, u_5, u_2, u_4, u_6\}$ (where u_j represents the number j coming up), the density matrix $\rho(\mathbf{0}_U)$ is the "pure state" 6×6 matrix with each entry being $\frac{1}{6}$ (note the odd states listed first)*

$$\rho(\mathbf{0}_U) = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix} \begin{matrix} u_1 \\ u_3 \\ u_5 \\ u_2 \\ u_4 \\ u_6 \end{matrix}. \tag{12}$$

The nonzero off-diagonal entries represent indistinctions or indits of partition $\mathbf{0}_U$, or in quantum terms as "coherences" where all 6 "eigenstates" cohere together in a pure "superposition" state. All pure states have logical entropy of zero, i.e., $h(\mathbf{0}_U) = 0$ (i.e., no dits).

Example 2 (continued) *Now classify or "measure" the elements by parity (odd or even) partition (observable) $\pi = \{B_{odd}, B_{even}\} = \{\{u_1, u_3, u_5\}, \{u_2, u_4, u_6\}\}$. Mathematically, this is done by the Lüders mixture operation where P_{odd} and P_{even} are the projections to the odd or even components:*

$$P_{odd}\rho(\mathbf{0}_U)P_{odd} + P_{even}\rho(\mathbf{0}_U)P_{even} = \sum_{i=1}^{m} p(B_i)\rho(B_i) = \rho(\pi). \tag{13}$$

$$\rho(\mathbf{0}_U) = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 \end{bmatrix} = \rho(\pi) \tag{14}$$

Theorem 3 (Basic) *The increase in logical entropy due to a Lüders mixture operation is the sum of amplitudes squared of the non-zero off-diagonal entries of the beginning density matrix that are zeroed in the final density matrix.*

Proof. Since for any density matrix ρ, $tr\left[\rho^2\right] = \sum_{i,j}\left|\rho_{ij}\right|^2$ [3, p. 77], $h(\rho(\pi)) - h(\rho(\mathbf{0}_U)) = \left(1 - tr\left[\rho(\pi)^2\right]\right) - \left(1 - tr\left[\rho(\mathbf{0}_U)^2\right]\right) = tr\left[\rho(\mathbf{0}_U)^2\right] - tr\left[\rho(\pi)^2\right] = \sum_{i,j}\left|\rho_{ij}(\mathbf{0}_U)\right|^2 - \sum_{i,j}\left|\rho_{ij}(\pi)\right|^2$. If $(u_i, u_{i'}) \in dit(\pi)$, then and only then are the off-diagonal terms corresponding to u_i and $u_{i'}$ zeroed by the Lüders operation. ■

Example 4 (continued) *In comparison with the matrix $\rho(\mathbf{0}_U)$ of all entries $\frac{1}{6}$, the entries that got zeroed in $\rho(\mathbf{0}_U) \rightsquigarrow \rho(\pi)$ correspond to the distinctions created in the transition $\mathbf{0}_U = \{U\} \rightsquigarrow \pi = \{\{u_1, u_3, u_5\}, \{u_2, u_4, u_6\}\}$. Increase in logical entropy $= h(\pi) - h(\mathbf{0}_U) = 2 \times 9 \times \left(\frac{1}{6}\right)^2 = \frac{18}{36} = \frac{1}{2}$. Usual calculations: $h(\pi) = 1 - 2 \times \left(\frac{1}{2}\right)^2 = \frac{1}{2}$ and $h(\mathbf{0}_U) = 1 - 1^2 = 0$.*

Since a projective measurement's effect on a density matrix in QM *is* the Lüders mixture operation, that means that the effects of the measurement is the above-described "making distinctions" by decohering or zeroing certain coherence terms in the density matrix, and the sum of the absolute squares of the coherences that were decohered is the change in the logical entropy.

5 Generalization to quantum logical information theory

The logical notion of information-as-distinctions generalizes to quantum information theory.

> [Information] is the notion of distinguishability abstracted away from what we are distinguishing, or from the carrier of information. ...And we ought to develop a theory of information which generalizes the theory of distinguishability to include these quantum properties... . [1, p. 155]

A *qubit* is a pair of states definitely distinguishable in the sense of being orthogonal. In general, a qubit, or rather a *qudit*, needs to be *relativized to an observable*–just as a dit is a dit of a partition.

A *qubit of an observable F* is a pair $(u_k, u_{k'})$ in the eigenbasis definitely distinguishable by F, i.e., $\phi(u_k) \neq \phi(u_{k'})$, distinct eigenvalues. Since the quantum version of logical entropy is a straight forward generalization of sets to vector spaces, we give the generalization in a table form–first before the introduction of probabilities.

Table 4a (w/o probs.)	'Classical' Logical Info. Theory	Quantum Logical Info. Theory
Universe	$U = \{u_1, \ldots, u_n\}$	Orthonormal basis $\{u_i\}$ Hilbert space V
Attribute/Observable	Real-valued 'random' variables $f, g: U \to \mathbb{R}$	Commuting self-adjoint operators F, G $\{u_i\}$ O.N. basis of simult. eigenvectors
Values	Image values $\{\phi_i\}_{i\in I}$ of f Image values $\{\gamma_j\}_{j\in J}$ of g	Eigenvalues $\{\phi_i\}_{i\in I}$ of F Eigenvalues $\{\gamma_j\}_{j\in J}$ of G
Partitions / Direct-sum decompositions	Inverse-image $\pi = \{f^{-1}(\phi_i)\}_{i\in I}$ Inverse-image $\sigma = \{g^{-1}(\gamma_j)\}_{j\in J}$	Eigenspace Direct-sum Decomp. F Eigenspace Direct-sum Decomp. G
Distinctions	Dits of π: $(u_k,u_{k'})\in U^2$, $f(u_k) \neq f(u_{k'})$ Dits of σ: $(u_k,u_{k'})\in U^2$, $g(u_k) \neq g(u_{k'})$	Qudits of F: $u_k \otimes u_{k'} \in V\otimes V$, $\phi(u_k) \neq \phi(u_{k'})$ Qudits of G: $u_k \otimes u_{k'} \in V\otimes V$, $\gamma(u_k) \neq \gamma(u_{k'})$
Information sets/spaces	$dit(\pi) \subseteq U\times U$ $dit(\sigma) \subseteq U\times U$	[qudit(F)] = subspace gen. by qudits of F [qudit(G)] = subspace gen. by qudits of G
Joint = Conditional = Mutual =	$dit(\pi) \cup dit(\sigma) \subseteq U\times U$ $dit(\pi) - dit(\sigma) \subseteq U\times U$ $dit(\pi) \cap dit(\sigma) \subseteq U\times U$	$[qudit(F) \cup qudit(G)] \subseteq V\otimes V$ $[qudit(F) - qudit(G)] \subseteq V\otimes V$ $[qudit(F) \cap qudit(G)] \subseteq V\otimes V$

Figure 7. The generalization of classical to quantum logical entropies before probabilities.

The fundamental theorem connecting measurement to logical entropy carries over to the quantum case. The nonzero off-diagonal terms in density matrix $\rho(\psi)$ are called "coherences"–like indistinctions in classical case. Measurement creates distinctions, i.e., turn coherences into 'decoherences'–classically, turn indistinctions into distinctions.

Table 4b (w/ probs.)	'Classical' Logical Info. Theory	Quantum Logical Info. Theory
Probability dist.	Pure state density matrix, e.g., $\rho(\mathbf{0}_U)$	Pure state density matrix $\rho(\psi)$
Product prob. dist.	$p \times p$ on $U \times U$	$\rho(\psi) \otimes \rho(\psi)$ on $V \otimes V$
Logical entropies	$h(\mathbf{0}_U) = 1 - tr[\rho(\mathbf{0}_U)^2] = 0$ $h(\pi) = p \times p(dit(\pi))$ $h(\pi,\sigma) = p \times p(dit(\pi) \cup dit(\sigma))$ $h(\pi\|\sigma) = p \times p(dit(\pi) - dit(\sigma))$ $m(\pi,\sigma) = p \times p(dit(\pi) \cap dit(\sigma))$	$h(\rho(\psi)) = 1 - tr[\rho(\psi)^2] = 0$ $h(F:\psi) = tr[P_{[qudit(F)]} \rho(\psi) \otimes \rho(\psi)]$ $h(F,G:\psi) = tr[P_{[qudit(F) \cup qudit(G)]} \rho(\psi) \otimes \rho(\psi)]$ $h(F\|G:\psi) = tr[P_{[qudit(F) - qudit(G)]} \rho(\psi) \otimes \rho(\psi)]$ $m(F,G:\psi) = tr[P_{[qudit(F) \cap qudit(G)]} \rho(\psi) \otimes \rho(\psi)]$
Venn diagram from being prob. measure	$h(\pi,\sigma) = h(\pi\|\sigma) + h(\sigma\|\pi) + m(\pi,\sigma)$ $h(\pi) = h(\pi\|\sigma) + m(\pi,\sigma)$	$h(F,G) = h(F\|G) + h(G\|F) + m(F,G)$ $h(F) = h(F\|G) + m(F,G)$
Interpretation	$h(\pi)$ = two-draw prob. of getting a dit of π, i.e., different f values.	$h(F:\psi)$ = prob. in two indep. F meas. of ψ in getting different eigenvalues.
Lüders Mixture	$\rho(\pi) = \Sigma_i P_{B_i} \rho(\mathbf{0}_U) P_{B_i}$ and $h(\pi) = p \times p(dit(\pi)) = 1 - tr[\rho(\pi)^2]$	$\rho'(\psi) = \Sigma_i P_{\phi_i} \rho(\psi) P_{\phi_i}$ and $h(F:\psi) = 1 - tr[\rho'(\psi)^2]$
Thm. on L-entropy and measurement.	$h(\pi)$ = sum of squares of terms zeroed in measurement operation: $\rho(\mathbf{0}_U) \to \rho(\pi)$.	$h(F:\psi)$ = sum of absol. squares of terms zeroed in the measurement operation: $\rho(\psi) \to \rho'(\psi)$.

Figure 8. The generalization of classical to quantum logical entropies with probabilities.

Basic Theorem: Measure of distinctions created in measuring pure state ψ by F = sum of absolute squares of off-diagonal terms zeroed (i.e., coherences that were decohered) in measurement = logical entropy increase, e.g., $h(F : \psi) = h(\rho'(\psi)) - h(\rho(\psi))$ = probability that two independent measurements of ψ will yield a qudit of F.

This treatment can be generalized to the case of two non-commuting observables and a given state.

There is also the case where we are only given density operators ρ and τ. Then the *logical entropy* of a state ρ is $h(\rho) = 1 - tr\left[\rho^2\right]$ and the *quantum logical cross-entropy* of ρ and τ is: $h(\rho||\tau) = 1 - tr[\rho\tau]$. Then the classical notion of Hamming distance can be generalized to the *quantum logical Hamming distance* between the two states:

$$d(\rho, \tau) = 2h(\rho||\tau) - h(\rho) - h(\tau). \tag{15}$$

This treatment avoids the problem suggested by Nielsen and Chuang.

> "Unfortunately, the Hamming distance between two objects is simply a matter of labeling, and *a priori* there aren't any labels in the Hilbert space arena of quantum mechanics!" [6, p. 399]

Then the main result is that the quantum logical Hamming distance is same as the usual Hilbert-Schmidt distance [8] (some authors include a factor of 1/2) whose square root is the trace-distance.

Quantum logical Hamming distance $d(\rho, \tau) = tr[(\rho - \tau)^2]$ Hilbert-Schmidt distance.

References

[1] Bennett, Charles H. Quantum Information: Qubits and Quantum Error Correction. International Journal of Theoretical Physics, **42** 153–76 (2003).

[2] Ellerman, David, An Introduction of Partition Logic. Logic Journal of the IGPL. **22**, 94–125 (2014).

[3] Fano, U. Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. Reviews of Modern Physics **29**, 74–93 (1957).

[4] Kolmogorov, Andrei N. Combinatorial Foundations of Information Theory and the Calculus of Probabilities. Russian Math. Surveys **38**, 29–40 (1983).

[5] Kung, Joseph P. S., Gian-Carlo Rota, and Catherine H. Yan. *Combinatorics: The Rota Way*. (New York, Cambridge University Press, 2009).

[6] Nielsen, M., and I. Chuang. *Quantum Computation and Quantum Information*. (Cambridge UK, Cambridge University Press, 2000).

[7] Rota, Gian-Carlo. 2001. Twelve Problems in Probability No One Likes to Bring up. In *Algebraic Combinatorics and Computer Science*, edited by Henry Crapo and Domenico Senato, (Milano, Springer, 2001), 57–93.

[8] Tamir, Boaz, and Eliahu Cohen. A Holevo-Type Bound for a Hilbert Schmidt Distance Measure. Journal of Quantum Information Science **5**, 127–33 (2015).