Quantum Information & Computation
ISSN 1533-7146 Volume 25 (2025) pp. 82–96
©Cerebration Science Publishing

Article

A New Logical Measure for Quantum Information

David Ellerman

Independent Researcher, Ljubljana, Slovenia; david@ellerman.org

Received date: 20 December 2024; Accepted date: 11 February 2025; Published online: 25 February 2025

Abstract: Starting at the logical level of the logic of partitions, dual to the usual Boolean logic of subsets, the notion of logical entropy, i.e., information as distinctions, is developed as the quantification of the distinctions of partitions—just as probability theory starts with the quantification of elements of subsets. Logical entropy is compared and contrasted with the usual notion of Shannon entropy. Then a semi-algorithmic procedure (from the mathematical folklore) is used to translate the notion of logical entropy at the set level to the corresponding notion of quantum logical entropy at the (Hilbert) vector space level. Examples and some important propositions relate quantum logical entropy to projective measurement and to the Hilbert-Schmidt distance. Overall, the approach demonstrates the logical basis and naturality of this notion of entropy for quantum mechanics.

Keywords: partition logic; duality of subsets and partitions; logical entropy; Shannon entropy; linearization of set concepts to vector space concepts; quantum logical entropy

1. Introduction

The purpose of this paper is to derive the relatively new notion of quantum logical entropy ([1–4]) from the relatively new logic of partitions ([5–7]) that is category-theoretically dual to the usual Boolean logic of subsets. The "classical" notion of logical entropy is derived starting as the quantitative version of the distinctions of partitions just as probability is derived starting as the quantitative version of the elements of subsets [8]. The notion of logical entropy is compared and contrasted with the usual Shannon entropy. Then the notion of logical entropy is linearized using a semi-algorithmic procedure to translate set-based concepts into the corresponding vector space concepts. The concept of logical entropy linearized to Hilbert space gives the concept of quantum logical entropy. The linearization procedure also allows results proven at the set level with logical entropy to be extended in a straightforward manner to quantum logical entropy. For instance, logical entropy is a probability measure with a simple probability interpretation (i.e., the two-draw probability of getting a distinction of a partition) and that interpretation extends to quantum logical entropy. That is, given an observable and a quantum state vector, their quantum logical entropy is the probability in two independent projective measurements of the observable on the prepared state that different eigenvalues are obtained. As a measure, the compound notions of difference (or conditional) and mutual logical entropy are immediately defined and their relationships are illustrated in the usual Venn diagrams for measures. The derivation of the quantum logical entropy allows the definitions of the difference and mutual quantum logical entropies which satisfy the corresponding relationships. This method of deriving the concepts of quantum logical entropy makes it a logic-based and natural measure of information for quantum mechanics (QM). Our purpose is to show that naturality and fundamentality of this notion of quantum information so the contrasts and comparisons with other notions such as the von Neumann entropy are left to the reader.

2. Subset Logic and Partition Logic

Subsets and partitions are category-theoretic dual concepts. That is, in the turn-around-the-arrows duality of category theory, a subset is also called a "part" and "The dual notion (obtained by reversing the arrows) of 'part' is the notion of partition" [9, p. 85]. A partition $\pi = \{B_1, ..., B_m\}$ on a universe set $U = \{u_1, ..., u_n\}$ ($|U| \geq 2$) is a set of nonempty subsets B_j called "blocks" such that the blocks are disjoint and their union is U. The partitions on U form a lattice $\Pi(U)$. The partial order (PO) for the lattice is *refinement* where a partition $\sigma = \{C_1, ..., C_{m'}\}$ is *refined* by π, written $\sigma \precsim \pi$, if for every block $B_j \in \pi$, there is a block $C_{j'} \in \sigma$ such that $B_j \subseteq C_{j'}$.

DOI: 10.2478/qic-2025-0005

At a more atomic or granular level, the *elements* of a subset are dual to the *distinctions* (dits) of a partition which are ordered pairs of elements in different blocks of the partition. The set of distinctions or *ditset* of a partition π is $\text{dit}(\pi) \subseteq U \times U$ and the complementary set,

$$\text{indit}(\pi) = U \times U - \text{dit}(\pi) = \cup_{j=1}^{m} B_j \times B_j$$

of indistinctions is the equivalence relation on U associated with π where the equivalence classes are the blocks of π. The refinement PO on partitions is the same as the inclusion PO on ditsets: $\sigma \precsim \pi$ if and only if (iff) $\text{dit}(\sigma) \subseteq \text{dit}(\pi)$.

The *join* $\sigma \vee \pi$ is the partition whose blocks are the nonempty intersections $B_j \cap C_{j'}$, and it is the least upper bound of π and σ for the refinement partial order. The ditset of the join is the union of the ditsets: $\text{dit}(\sigma \vee \pi) = \text{dit}(\sigma) \cup \text{dit}(\pi)$. Since the arbitrary intersection of equivalence relations is an equivalence relation, the *meet* $\sigma \wedge \pi$ can be defined as the partition whose ditset is the complement of the smallest equivalence relation containing $\text{indit}(\sigma) \cup \text{indit}(\pi)$, and it is the greatest lower bound of σ and π. The top of the lattice is the *discrete partition* $\mathbf{1}_U = \{\{u_i\}\}_{u_i \in U}$ of singletons of the elements of U and the bottom of the lattice is the *indiscrete partition* $\mathbf{0}_U = \{U\}$ whose only block is all of U.

The lattice of partitions was known in the 19th century (e.g., Dedekind and Schröder). However, throughout the 20th century "the only operations on the family of equivalence relations fully studied, understood and deployed are the binary join $\vee$ and meet $\wedge$ operations" [10, p. 445]. To go from a lattice of partitions to a *logic* of partitions comparable to the usual Boolean logic of subsets, there needs to be at least an implication operation defined on partitions. That operation would be the parallel to the *subset implication* or *conditional* $S \supset T = S^c \cup T$ for $S, T \subseteq U$ in the powerset Boolean algebra $\wp(U)$ of subsets of U where the partial order is inclusion, the join and meet are union and intersection respectively, and the top and bottom are U and $\emptyset$ respectively. The *implication* $\sigma \Rightarrow \pi$ is the partition on U that is like π except that whenever a block $B_j \in \pi$ is contained in some block of σ, then B_j is replaced by the singletons of its elements. If we denote a block $B \in \pi$ when it has been "discretized" as $\mathbf{1}_B$ and when it remains whole as $\mathbf{0}_B$, then the implication $\sigma \Rightarrow \pi$ functions like a characteristic or indicator function for inclusion of π-blocks in σ-blocks. Thus when they are all included, i.e., when refinement holds, then the implication is the discrete partition $\mathbf{1}_U$:

$$\sigma \Rightarrow \pi = \mathbf{1}_U \text{ iff } \sigma \precsim \pi$$

which is just the partition logic version of the subset logic relation:

$$S \supset T = U \text{ iff } S \subseteq T.$$

Thus, the usual Boolean logic of subsets (often presented in only the special case of "propositional logic") has a dual logic of partitions. The elements (Its) of subsets and the distinctions (Dits) of partitions have dual corresponding roles as illustrated in Table 1.

Table 1. Elements-distinctions duality between the two dual logics.

	Logic $\wp(U)$ of subsets of U	Logic of partitions $\Pi(U)$ on U
Its/Dits	Elements of subsets	Distinctions of partitions
P.O.	Inclusion of subsets	Inclusion of ditsets
Join	Union of subsets	Union of ditsets
Meet	Subset of common elements	Ditset of common dits
Impl.	$S \supset T = U$ iff $S \subseteq T$	$\sigma \Rightarrow \pi = 1_U$ iff $\sigma \precsim \pi$
Top	Subset U with all elements	Partition $\mathbf{1}_U$ with all distinctions
Bottom	Subset $\emptyset$ with no elements	Partition $\mathbf{0}_U$ with no distinctions

For the simplest non-trivial case, the two lattices are illustrated in Figure 1 for $U = \{a, b, c\}$.

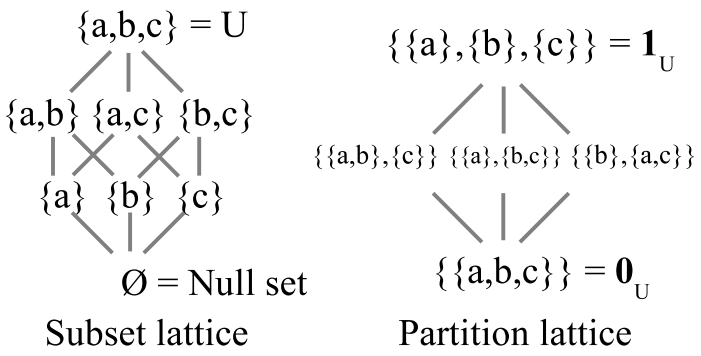

Figure 1. The lattices of the dual subsets and partitions.

3. The New Logical Measure of Information

Gian-Carlo Rota made the crucial connection between the dual notions of subsets and partitions. "The lattice of partitions plays for information the role that the Boolean algebra of subsets plays for size or probability" [11, p. 30].

$$\frac{\text{Subsets}}{\text{Probability}} \approx \frac{\text{Partitions}}{\text{Information}}.$$

In his Fubini Lectures, Rota said "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? [12, p. 67]

We have seen the duality between elements of a subset and dits of a partition, i.e.,

$$\frac{\text{Elements}}{\text{Subset}} \approx \frac{\text{Distinctions}}{\text{Partition}}$$

so the "size" of a partition may be taken as the number of distinctions.

The new logical foundations for information theory [4] start 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. [13, p. 39]

Since logical probability theory [8] starts as the normalized size of a subset, i.e., $\Pr(S) = \frac{|S|}{|U|}$, the notion of information-as-distinctions starts with the normalized size of a partition's ditset. This gives the *logical entropy* (with equiprobable points of U) as:

$$h(\pi) = \frac{|\text{dit}(\pi)|}{|U\times U|} = \frac{|U\times U - \text{indit}(\pi)|}{|U\times U|} = 1 - \frac{\left|\cup_{j=1}^{m} B_j \times B_j\right|}{[U\times U]} = 1 - \sum_{j=1}^{m}\left(\frac{|B_j|}{|U|}\right)^2$$
$$= 1 - \sum_{j=1}^{m} \Pr(B_j)^2 = \sum_{j\neq j'} \Pr(B_j)\Pr(B_{j'}).$$

Given any (always positive) 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:

$$p \times p(S) = \sum_{(u_i,u_k)\in S} p_i p_k.$$

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

$$h(\pi) = p \times p(\text{dit}(\pi)) = \sum_{(u_i,u_j)\in \text{dit}(\pi)} p_i p_j = \sum_{j\neq j'} \Pr(B_j)\Pr(B_{j'})$$

where $\Pr(B_j) = \sum_{u_i \in B_j} p_i$.

Interpretation of logical entropy

The logical entropy $h(\pi)$ of a partition π is the probability
that in two draws from U (with replacement),
one gets a distinction of the partition π.

Similarly, $\Pr(S)$ is the probability that in one draw from U, one gets an element of the subset $S \subseteq U$. Thus the duality between subsets and partitions in their quantitative versions gives a duality between probability theory and information theory illustrated in Table 2.

Table 2. Duality of quantitative subsets and partitions.

	Logical Probability Theory	Logical Information Theory
Outcomes	Elements of S	Distinctions of π
Events	Subsets $S \subseteq U$	Ditsets $\text{dit}(\pi) \subseteq U \times U$
$p_i = \frac{1}{n}$	$\Pr(S) = \frac{\lvert S \rvert}{\lvert U \rvert}$	$h(\pi) = \frac{\lvert \text{dit}(\pi) \rvert}{\lvert U \times U \rvert}$
Probs. p	$\Pr(S) = \sum_{u_i \in S} p_i$	$h(\pi) = \sum_{(u_i, u_k) \in \text{dit}(\pi)} p_i p_k$
Interpretation	1-draw prob. of S-element	2-draw prob. of π-distinction

Given partitions $\pi = \{B_1, ..., B_m\}, \sigma = \{C_1, ..., C_{m'}\}$ on U, the *ditset for their join* is:

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

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

$$h(\pi, \sigma) = h(\pi \vee \sigma) = p \times p(\text{dit}(\pi) \cup \text{dit}(\sigma)) = 1 - \sum_{j,j'} p\big(B_j \cap C_{j'}\big)^2.$$

The ditset for the *difference (or conditional) logical entropy* $h(\pi|\sigma)$ is the difference of ditsets, and thus: $h(\pi|\sigma) = p \times p(\text{dit}(\pi) - \text{dit}(\sigma))$. The ditset for the *logical mutual information* $m(\pi, \sigma)$ is the intersection of ditsets, so: $m(\pi, \sigma) = p \times p(\text{dit}(\pi) \cap \text{dit}(\sigma))$. Venn diagrams apply to measures and since logical entropy is a probability measure on $U \times U$, Figure 2 illustrates its Venn diagram for the compound notions of logical entropy.

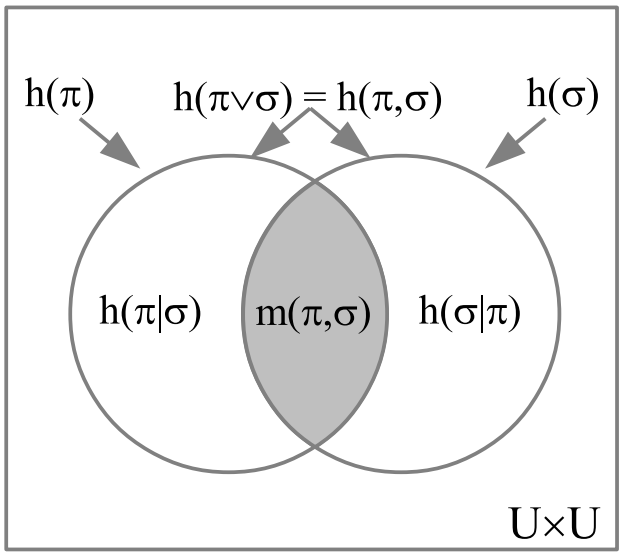

Figure 2. Venn diagram for compound logical entropies.

As in any Venn diagram for values of a measure, certain relationships hold such as:

$$h(\pi \vee \sigma) = h(\pi, \sigma) = h(\pi) + h(\sigma) - m(\pi, \sigma) = h(\pi|\sigma) + m(\pi, \sigma) + h(\sigma|\pi).$$

4. Deriving the Shannon Entropies from the Logical Entropies

The simple and compound definitions for Shannon entropy $H(\pi) = \sum_{j=1}^{m} \Pr(B_j) \log\left(\frac{1}{\Pr(B_j)}\right)$ were defined so that the Venn diagram relationships hold ([14,15]), but they are not defined in terms of a measure (in the sense of measure theory) [16]. However, all the Shannon entropies can be derived from the definitions of logical entropy (which is a measure) by a uniform monotonic transformation that preserves the Venn diagram relationships. It is easiest to work with the entropies of a probability distribution $p = (p_1, ..., p_n)$ on U where:

$$h(p) = h(\mathbf{1}_U) = 1 - \sum_{i=1}^{n} p_i^2 = \sum_{i \neq k} p_i p_k = \sum_{i=1}^{n} p_i(1 - p_i)$$
$$H(p) = H(\mathbf{1}_U) = \sum_{i=1}^{n} p_i \log\left(\tfrac{1}{p_i}\right).$$

Intuitively, if $p_i = 1$, then there is no information in the occurrence of u_i, so information is measured by the 1-complement. But there are two 1-complements, the additive 1-complement of $1 - p_i$ and the multiplicative 1-complement of $\frac{1}{p_i}$.

- The additive probability average of the additive 1-complements is the logical entropy: $h(p) = \sum_{i=1}^{n} p_i(1 - p_i)$.
- The multiplicative probability average of the multiplicative 1-complements is the log-free or anti-log version of Shannon entropy $\prod_{i=1}^{n} \left(\frac{1}{p_i}\right)^{p_i} = \log^{-1}(H(p))$.

Then the particular log is chosen according to the application, e.g., $\log_2$ in coding theory and ln in statistical mechanics. Since taking the log of the log-free version of Shannon entropy turns the multiplicative average into an additive average, we can then see how to directly transform the logical formulas into the Shannon formulas by the *dit-bit transform*:

$$1 - p_i \rightsquigarrow \log\left(\tfrac{1}{p_i}\right).$$

When the compound logical entropy formulas are formulated in terms of the additive 1-complements, then the dit-bit transform gives the corresponding compound formula for the Shannon entropies. This is illustrated in Table 3 for the probability distribution $p : U \to [0, 1]$ and a joint distribution $p : X \times Y \to [0, 1]$.

Table 3. The dit-bit transform from logical entropy to Shannon entropy.

	The Dit-Bit Transform: $1 - p_i \rightsquigarrow \log\left(\frac{1}{p_i}\right)$
$h(p) =$	$\sum_i p_i(1 - p_i)$
$H(p) =$	$\sum_i p_i \log(1/p_i)$
$h(X, Y) =$	$\sum_{x,y} p(x, y)[1 - p(x, y)]$
$H(X, Y) =$	$\sum_{x,y} p(x, y) \log\left(\frac{1}{p(x,y)}\right)$
$h(X\|Y) =$	$\sum_{x,y} p(x, y)[(1 - p(x, y) - (1 - p(y))]$
$H(X\|Y) =$	$\sum_{x,y} p(x, y)\left[\log\left(\frac{1}{p(x,y)}\right) - \log\left(\frac{1}{p(y)}\right)\right]$
$m(X, Y) =$	$\sum_{x,y} p(x, y)[[1 - p(x)] + [1 - p(y)] - [1 - p(x, y)]]$
$I(X, Y) =$	$\sum_{x,y} p(x, y)\left[\log\left(\frac{1}{p(x)}\right) + \log\left(\frac{1}{p(y)}\right) - \log\left(\frac{1}{p(x,y)}\right)\right]$

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) so those relationships, illustrated in Figure 3, are normally termed a "mnemonic" [16, p. 112].

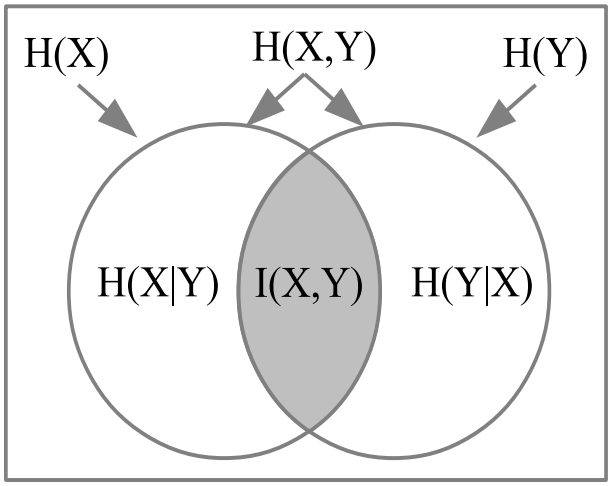

Figure 3. Venn diagram "mnemonic" for compound Shannon entropies.

5. Logical Entropy via Density Matrices

All this will carry over to the quantum version of logical entropy by using density matrices. First, the "classical" treatment of logical entropy is restated using density matrices over the reals. Then that will extend immediately to the quantum case of density matrices over the complex numbers by making the appropriate changes such as replacing the square with the absolute square.

Let's do the density matrix version of $p \times p(\text{dit}(\pi))$.The density matrix associated with each block $B_j \in \pi$ is the projection matrix $\rho(B_j) = |b_j\rangle\langle b_j|$ where $|b_j\rangle$ is the $n \times 1$ column vector with entries $\sqrt{\frac{p_i}{\Pr(B_j)}}$ if $u_i \in B_j$, else 0. Thus the entries are $\rho(B_j)_{i,k} = \frac{\sqrt{p_i p_k}}{\Pr(B_j)}$ if $u_i, u_k \in B_j$, else 0. The density matrix for the partition is $\rho(\pi) = \sum_{j=1}^{m} \Pr(B_j)\rho(B_j)$ where $\rho(\pi)_{ik} = \sqrt{p_i p_k}$ if $(u_i, u_k) \in \text{indit}(\pi)$, else 0. To recover the logical entropy $h(\pi) = p \times p(\text{dit}(\pi))$ using density matrices, it can be calculated as: $1 - \text{tr}\left[\rho(\pi)^2\right]$. A basic result about any density matrix ρ is: $\text{tr}\left[\rho^2\right] = \sum_{i,k} |\rho_{ik}|^2$ [17, p. 77]
so

$$\text{tr}\left[\rho(\pi)^2\right] = \sum_{(u_i,u_k)\in \text{indit}(\pi)} p_i p_k = 1 - \sum_{(u_i,u_k)\in \text{dit}(\pi)} p_i p_k = 1 - p \times p(\text{dit}(\pi))$$

and thus:

$$h(\pi) = p \times p(\text{dit}(\pi)) = 1 - \text{tr}\left[\rho(\pi)^2\right].$$

Example: Consider $U = \{a, b, c\}$ with $p : U \to [0, 1]$ where $p_a = \frac{1}{2}$, $p_b = \frac{1}{3}$, and $p_c = \frac{1}{6}$ and $\pi = \{B_1, B_2\} = \{\{a, b\}, \{c\}\}$. The usual calculation of the logical entropy is $h(\pi) = 1 - \left(\frac{5}{6}\right)^2 - \left(\frac{1}{6}\right)^2 = 1 - \frac{26}{36} = \frac{5}{18}$. Then the density matrix calculation is:

$$\rho(B_1) = |b_1\rangle\langle b_1| = \begin{bmatrix} \sqrt{\frac{1/2}{5/6}} \\ \sqrt{\frac{1/3}{5/6}} \\ 0 \end{bmatrix} \begin{bmatrix} \sqrt{\frac{1/2}{5/6}} & \sqrt{\frac{1/3}{5/6}} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1/2}{5/6} & \frac{\sqrt{1/6}}{5/6} & 0 \\ \frac{\sqrt{1/6}}{5/6} & \frac{1/3}{5/6} & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$\rho(B_2) = |b_2\rangle\langle b_2| = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so that:

$$\rho(\pi) = \sum_{j=1}^{2} \Pr(B_j)\rho(B_j) = \begin{bmatrix} 1/2 & \sqrt{1/6} & 0 \\ \sqrt{1/6} & 1/3 & 0 \\ 0 & 0 & 1/6 \end{bmatrix} \text{ and } \rho(\pi)^2 = \begin{bmatrix} \frac{5}{12} & \frac{5\sqrt{6}}{36} & 0 \\ \frac{5\sqrt{6}}{36} & \frac{5}{18} & 0 \\ 0 & 0 & \frac{1}{36} \end{bmatrix}.$$

Then $\text{tr}\left[\rho(\pi)^2\right] = \frac{15}{36} + \frac{10}{36} + \frac{1}{36} = \frac{26}{36}$ so $1 - \text{tr}\left[\rho(\pi)^2\right] = \frac{5}{18} = h(\pi)$✓.

Borrowing the language of QM, $\rho(B_j)$ as a projection matrix represents a *pure* state, i.e., $\rho(B_j)^2 = \rho(B_j)$. Then $\rho(\pi)$ represents a *mixed* state where the pure states $\rho(B_j)$ occur with the probabilities $\Pr(B_j)$. The dictionary giving the reformulation of set partition concepts in terms of density matrices is given in Table 4.

Table 4. Dictionary translating set partitions into density matrices.

Set concept with probabilities	Set level density matrix concept
Partition π with point probs. p	Density matrix $\rho(\pi) = \sum_{j=1}^{m} \Pr(B_j)\|b_j\rangle\langle b_j\|$
Point probabilities $\{p_1, ..., p_n\}$	Value of diagonal entries of $\rho(\pi)$
Trivial indits (u_i, u_i) of π	Diagonal entries of $\rho(\pi)$
Non-trivial indits of π	Non-zero off-diagonal entries of $\rho(\pi)$
Dits of π	Zero entries of $\rho(\pi)$
Sum $\Pr(B_j) = \sum_{u_i \in B_j} p_i$	Trace $\text{tr}[P_{B_j}\rho(\pi)]$
Block probabilities $\Pr(B_j)$ in π	Eigenvalues $\neq 0$ of $\rho(\pi)$
Block prob. 1 of U in $\mathbf{0}_U = \{U\}$	Non-zero eigenvalue of 1 for $\rho(\mathbf{0}_U)$

We also need to give the set version of the "measurement" of a state $\rho(\pi)$ by an "observable" given by a real-valued numerical attribute $g : U \to \mathbb{R}$ which defines the inverse-image partition $g^{-1} = \left\{g^{-1}(s)\right\}_{s \in g(U)}$. In QM, the transformation of a density matrix $\rho(\pi)$ in the projective measurement by an observable is given by the Lüders mixture operation [18, p. 279]. For each block $g^{-1}(s)$ in the observable partition g^{-1}, the diagonal projection matrix P_s has the diagonal entries $\chi_{g^{-1}(s)}(u_i)$, i.e., $(P_s)_{ii} = 1$ if $g(u_i) = s$, else 0. Then the Lüders mixture operation gives the post-measurement density matrix $\hat{\rho}(\pi)$ as:

$$\hat{\rho}(\pi) = \sum_{s \in g(U)} P_s \rho(\pi) P_s.$$

Proposition 1. $\hat{\rho}(\pi) = \rho\left(\pi \vee g^{-1}\right)$.

Proof. A nonzero entry in $\rho(\pi)$ has the form $\rho(\pi)_{ik} = \sqrt{p_i p_k}$ iff there is some block $B_j \in \pi$ such that $(u_i, u_k) \in B_j \times B_j$, i.e., if $u_i, u_k \in B_j$ and otherwise 0. The matrix operation $P_s\rho(\pi)$ will preserve the entry $\sqrt{p_i p_k}$ if $u_i \in g^{-1}(s)$, otherwise the entry is zeroed. And if the entry was preserved, then the further matrix operation $(P_s\rho(\pi))P_s$ will preserve the entry $\sqrt{p_i p_k}$ if $u_k \in g^{-1}(s)$, otherwise it is zeroed. Hence the entries $\sqrt{p_j p_k}$ in $\rho(\pi)$ that are preserved in $P_s\rho(\pi)P_s$ are the entries where both $u_i, u_k \in B_j$ for some $B_j \in \pi$ and $u_i, u_k \in g^{-1}(s)$. These are the entries in $\rho\left(\pi \vee g^{-1}\right)$ corresponding to the blocks $B_j \cap g^{-1}(s)$ for some $B_j \in \pi$, so summing over the blocks $g^{-1}(s) \in g^{-1}$ gives the result: $\hat{\rho}(\pi) = \sum_{s \in g(U)} P_s \rho(\pi) P_s = \rho\left(\pi \vee g^{-1}\right)$. □

Proposition 2 (Measuring measurement)**.** *In the "projective measurement" $\rho(\pi) \rightsquigarrow \rho\left(\pi \vee g^{-1}\right)$, the sum of the squares of the non-zero off-diagonal entries of $\rho(\pi)$ that were zeroed in $\hat{\rho}(\pi) = \rho\left(\pi \vee g^{-1}\right)$ is the difference in their logical entropies $h\left(\pi \vee g^{-1}\right) - h(\pi) = h\left(\pi \vee g^{-1} | \pi\right)$.*

Proof. Since for any density matrix ρ, $\text{tr}\left[\rho^2\right] = \sum_{i,k} |\rho_{ik}|^2$,

$$h\left(\pi \vee g^{-1} | \pi\right) = h\left(\pi \vee g^{-1}\right) - h(\pi) = \left(1 - \text{tr}\left[\rho\left(\pi \vee g^{-1}\right)^2\right]\right) - \left(1 - \text{tr}\left[\rho(\pi)^2\right]\right) = \sum_{i,k} |\rho(\pi)_{ik}|^2 - \sum_{i,k} \left|\rho\left(\pi \vee g^{-1}\right)_{ik}\right|^2$$

since the action of the projection operators in the Lüders mixture operation is either to zero an entry or leave it the same. □

Example (continued): Let $g(a) = 1, g(b) = g(c) = 0$.Then $g^{-1} = \{\{a\}, \{b, c\}\}$ so $\pi \vee g^{-1} = \{\{a\}, \{b\}, \{c\}\} = \mathbf{1}_U$ and thus $h\left(\pi \vee g^{-1}\right) = h(\mathbf{1}_U) = 1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{3}\right)^2 - \left(\frac{1}{6}\right)^2 = 1 - \frac{9}{36} - \frac{4}{36} - \frac{1}{36} = 1 - \frac{14}{36} = \frac{11}{18}$ so that $h\left(\pi \vee g^{-1}\right) - h(\pi) = \frac{11}{18} - \frac{5}{18} = \frac{1}{3}$. The density matrix for the discrete partition is:

$$\rho\left(\pi \vee g^{-1}\right) = \rho(\mathbf{1}_U) = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/6 \end{bmatrix} \text{ and } \rho(\pi) = \begin{bmatrix} 1/2 & \sqrt{1/6} & 0 \\ \sqrt{1/6} & 1/3 & 0 \\ 0 & 0 & 1/6 \end{bmatrix} \text{ so the sum of the squares of}$$

the zeroed elements is $\left(\sqrt{1/6}\right)^2 + \left(\sqrt{1/6}\right)^2 = \frac{1}{3}$.✓

The measuring measurement result deals with the non-zero off-diagonal terms in a density matrix.

> [T]he off-diagonal terms of a density matrix... are often called *quantum coherences* because they are responsible for the interference effects typical of quantum mechanics that are absent in classical dynamics. [18, p. 177]

Since a projective measurement's effect on a density matrix in QM is given by 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. This is a foretaste of the results for quantum logical entropy.

In coding theory, the Hamming distance between two n-ary $0,1$-vectors is the number of places where they differ. The partition version of this idea is the measure of where two partitions π and σ on U differ [19] which in terms of logical entropy is the *logical Hamming distance between partitions* (see Figure 2):

$$d(\pi,\sigma) := h(\pi|\sigma) + h(\sigma|\pi) = h(\pi \vee \sigma) - m(\pi,\sigma) = 2h(\pi \vee \sigma) - h(\pi) - h(\sigma).$$

Intuitively, it is the logical information that is in each partition but not in the other, so it is a measure of how they differ, i.e., how "far apart" they are.

Lemma 1. $h(\pi \vee \sigma) = 1 - \text{tr}[\rho(\pi)\rho(\sigma)]$.

Proof. The k^{th} diagonal entry in $\rho(\pi)\rho(\sigma)$ is the scalar product $\sum_i \rho(\pi)_{ki}\rho(\sigma)_{ik}$ with $\rho(\pi)_{ki} = \sqrt{p_k p_i}$ if $(u_k, u_i) \in \text{indit}(\pi)$ and otherwise 0, and similarly for $\rho(\sigma)_{ik}$. Hence the only non-zero terms in that sum are for $(u_k, u_i) \in \text{indit}(\pi) \cap \text{indit}(\sigma) = \text{indit}(\pi \vee \sigma)$. Hence $\text{tr}[\rho(\pi)\rho(\sigma)] = \sum_{(u_i,u_k)\in \text{indit}(\pi\vee\sigma)} p_i p_k = 1 - \sum_{(u_i,u_k)\in \text{dit}(\pi\vee\sigma)} p_i p_k$ so $h([\ ]\pi \vee \sigma) = 1 - \text{tr}[\rho(\pi)\rho(\sigma)]$ and similarly for $\text{tr}[\rho(\sigma)\rho(\pi)]$. □

The quantity $\text{tr}\left[(\rho(\pi) - \rho(\sigma))^2\right]$ is usually termed the *Hilbert-Schmidt distance* between two density matrices ([20,21]) (sometimes with a $1/2$ coefficient). It should be noted that the Hilbert-Schmidt distance is defined quite independently of the logical entropy and yet it is equal to the logical distance.

Proposition 3. $\text{tr}\left[(\rho(\pi) - \rho(\sigma))^2\right] = h(\pi|\sigma) + h(\sigma|\pi) = d(\pi,\sigma)$.

Proof. $\text{tr}\left[(\rho(\pi) - \rho(\sigma))^2\right] = \text{tr}\left[\rho(\pi)^2\right] - \text{tr}[\rho(\pi)\rho(\sigma)] - \text{tr}[\rho(\sigma)\rho(\pi)] + \text{tr}\left[\rho(\sigma)^2\right]$ so:

$$\begin{aligned}\text{tr}\left[(\rho(\pi) - \rho(\sigma))^2\right] &= 2[1 - \text{tr}[\rho(\pi)\rho(\sigma)]] - \left(1 - \text{tr}\left[\rho(\pi)^2\right] - \left(1 - \text{tr}\left[\rho(\sigma)^2\right]\right)\right) \\ &= 2h(\pi \vee \sigma) - h(\pi) - h(\sigma) = h(\sigma|\pi) + h(\pi|\sigma) = d(\pi,\sigma).\end{aligned}$$

□

Corollary 1. $\text{tr}\left[(\hat{\rho}(\pi) - \rho(\pi))^2\right] = h(\hat{\rho}(\pi)|\rho(\pi))$.

Proof. Taking $\sigma = g^{-1}$ as the inverse-image partition of the numerical attribute, $\hat{\rho}(\pi) = \rho(\pi \vee \sigma)$. Since $\pi \precsim \pi \vee \sigma$, $\text{dit}(\pi) \subseteq \text{dit}(\pi \vee \sigma)$ so $\text{dit}(\pi) - \text{dit}(\pi \vee \sigma) = \emptyset$ and thus $h(\pi|\pi \vee \sigma) = 0$. □

6. Linearization from Sets to Vector Spaces

Our goal is to systematically derive the quantum logical entropy starting from the logic of partitions. We have so far developed the notion of logical entropy at the set level and given a number of results. There is a semi-algorithmic procedure or "yoga" [22, p. 271] to transform set concepts into the corresponding vector space concepts. The yoga is, in general, part of the mathematical folklore but parts have been stated explicitly [23, pp. 355–361].

Yoga of Linearization

Apply a set concept to a basis set of a vector space, and
whatever is linearly generated is the corresponding vector space concept.

This yoga or procedure shows how the logical entropy concepts developed so far can be transformed into the corresponding concepts and results in the Hilbert vector spaces of quantum mechanics. Indeed, the previous results formulated using density matrices extend, *mutatis mutandis* (e.g., using the absolute square instead of the ordinary square), to the corresponding results about quantum logical entropy.

Hence we need to develop the dictionary to translate set concepts into vector space concepts. A subset of a basis set generates a subspace and the cardinality of the subset is the dimension of the subspace. Without assuming any probability distribution on U, a (real-valued) *numerical attribute* (e.g., weight, height, or age of persons) is a function $f : U \to \mathbb{R}$. In any vector space V over a field containing the reals where U is now a basis set, the numerical attribute generates a linear operator $F : V \to V$ with real eigenvalues by the definition $Fu_i = f(u_i)u_i$. If we let $f \upharpoonright S = rS$ mean that the numerical attribute f restricted to S has the constant value of r, then the vector space version is the eigenvector-eigenvalue equation $Fv = rv$. This means that the set-version of an eigenvector is a constant set of f and the constant value is a set-version of an eigenvalue. The numerical attribute's inverse-image is a partition $f^{-1} = \{f^{-1}(r)\}_{r\in f(U)}$ and each block $f^{-1}(r)$ generates a subspace V_r which is the eigenspace of the induced F for the eigenvalue r. Thus the partition f^{-1} generates a set of subspaces $\{V_r\}_{r\in f(U)}$ such that every vector v can be uniquely expressed as a sum of non-zero vector $v_r \in V_r$, i.e., the partition f^{-1} generates a *direct-sum decomposition* (DSD) $\{V_r\}_{r\in f(U)}$ of the vector space. When the numerical attribute is just a characteristic function $\chi : U \to 2 = \{0, 1\}$ of some attribute on U, then the induced operator $P_1 : V \to V$ is the projection operator to the subspace generated by $\chi^{-1}(1)$. Hence for a general numerical attribute f, we have projection operators P_r to the subspace V_r generated by $f^{-1}(r)$. The spectral decomposition of the induced operator is: $F = \sum_{r\in f(U)} rP_r$ which works backwards gives the spectral decomposition in the set case: $f = \sum_{r\in f(U)} r\chi_{f^{-1}(r)}$ where $\chi_{f^{-1}(r)}$ is the characteristic function for the subset $f^{-1}(r)$. And finally, the direct product $U \times U$ of a basis set U for V will (bi)linearly generate the tensor product $V \otimes V$ (where the ordered pair (u_i, u_k) is written $u_i \otimes u_k$). These linearizations are summarized in Table 5.

Table 5. Linearization dictionary to translate set concepts into corresponding vector space concepts.

Set concept	Vector-space concept
Subset $S \subseteq U$	Subspace $[S] \subseteq V$
Partition $\{f^{-1}(r)\}_{r\in f(U)}$	DSD $\{V_r\}_{r\in f(U)}$
Disjoint union $U = \uplus_{r\in f(U)} f^{-1}(r)$	Direct sum $V = \oplus_{r\in f(U)} V_r$
Numerical attribute $f : U \to \mathbb{R}$	Observable $Fu_i = f(u_i)u_i$
$f \upharpoonright S = rS$	$Fu_i = ru_i$
Constant set S of f	Eigenvector u_i of F
Value r on constant set S	Eigenvalue r of eigenvector u_i
Characteristic fcn. $\chi_S : U \to \{0, 1\}$	Projection operator $P_{[S]}u_i = \chi_S(u_i)u_i$
$\sum_{r\in f(U)} \chi_{f^{-1}(r)} = \chi_U$	$\sum_{r\in f(U)} P_r = I : V \to V$
Spectral Decomp. $f = \sum_{r\in f(U)} r\chi_{f^{-1}(r)}$	Spectral Decomp. $F = \sum_{r\in f(U)} rP_r$
Set of r-constant sets $\wp(f^{-1}(r))$	Eigenspace V_r of r-eigenvectors
Direct product $U \times U$	Tensor product $V \otimes V$

7. Generalization to Quantum Logical Entropy

We have developed the notion of logical entropy as the quantitative version of partitions. The mathematics used was at the level of sets, e.g., numerical attributes and probability distributions on a set U. We have also outlined the semi-algorithmic yoga of linearization to translate set concepts into the corresponding vector space concepts. Hence the new concept of *quantum logical entropy* can be developed in a straightforward manner by linearizing the definition of logical entropy to Hilbert space.

The logical notion of information-as-distinctions generalizes to quantum information theory. A *qubit* is a pair of states definitely distinguishable in the sense of being orthogonal. In general, a *qudit* needs to be *relativized to an observable*—just as a dit is a dit of a partition such as the inverse-image partition f^{-1} of a numerical attribute $f : U \to \mathbb{R}$. Given such a numerical attribute f defined on an orthonormal (ON) basis for a (finite-dimensional) Hilbert space V, a Hermitian (or self-adjoint) operator $F : V \to V$ is defined by $Fu_i = f(u_i)u_i$. The definition can be reversed. Given a Hermitian operator F on V, there is an ON basis of eigenvectors U and a real-valued numerical attribute f, the eigenvalue function, is defined on U by taking each element u_i to its eigenvalue.

A *qudit of an observable* F is a pair (u_i, u_k) in the eigenbasis definitely distinguishable by F, i.e., $f(u_i) \neq f(u_k)$, distinct eigenvalues. Let $qudit(F)$ be the set of tensor product basis elements $u_i \otimes u_k$ for $f(u_i) \neq f(u_k)$. Since the quantum version of logical entropy is a straightforward generalization from sets to vector spaces, we give the generalization in Table 6. Numerical attributes f, g on U generate commuting observables F, G and commuting

observables F, G generate eigenvalue functions f, g on the ON basis U of simultaneous eigenvectors. We follow Kolmogorov's dictum by first giving the basic machinery without probabilities.

Table 6. Yoga of Linearization without probabilities case.

Logical entropy	Quantum logical entropy
$U = \{u_1, ..., u_n\}$	ON basis U for Hilbert space V
$f, g : U \to \mathbb{R}$	Commuting $F, G : V \to V$
$\{r\}_{r\in f(U)}, \{s\}_{s\in g(U)}$	Eigenvalues of F and G
$\pi = \{f^{-1}(r)\}_{r\in f(U)}, \sigma = \{g^{-1}(s)\}_{s\in g(U)}$	DSDs of eigenspaces of F, G
Dits of $\pi : (u_i, u_k), f(u_i) \neq f(u_k)$	Qudits F: $u_i \otimes u_k, f(u_i) \neq f(u_k)$
Dits of $\sigma : (u_i, u_k), g(u_i) \neq g(u_k)$	Qudits G: $u_i \otimes u_k, g(u_i) \neq g(u_k)$
Ditset of π: dit(π)	$[qudit(F)]$: Subspace generated in $V \otimes V$
Ditset of σ: dit(σ)	$[qudit(G)]$: Subspace generated in $V \otimes V$
Join: dit$(\pi) \cup$ dit$(\sigma) \subseteq U \times U$	$[qudit(F) \cup qudit(G)] \subseteq V \otimes V$
Difference: dit$(\pi) -$ dit$(\sigma) \subseteq U \times U$	$[qudit(F) - qudit(G)] \subseteq V \otimes V$
Mutual: dit$(\pi) \cap$ dit$(\sigma) \subseteq U \times U$	$[qudit(F) \cap qudit(G)] \subseteq V \otimes V$

In quantum mechanics, the probability information is carried by the state to be measured. Hence Table 6 deals with the set and quantum version of quantum observables, not quantum states. The next step is to apply linearization to the set and vector space versions of the quantum state which carries the probability information. At the set level, the universal set U is equipped with a probability distribution $p : U \to [0, 1]$. Table 7 gives the translation dictionary to give the quantum logical entropy.

Table 7. Logical entropy + Linearization = quantum logical entropy.

Logical entropy	Quantum logical entropy
$\rho(\mathbf{0}_U) = \rho(U) = \rho(U)^2$	Pure state $\rho(\psi) = \rho(\psi)^2$
$p \times p$ on $U \times U$	$\rho(\psi) \otimes \rho(\psi)$ on $V \otimes V$
$h(\mathbf{0}_U) = 1 - \text{tr}\left[\rho(\mathbf{0}_U)^2\right] = 0$	$h(\rho(\psi)) = 1 - \text{tr}\left[\rho(\psi)^2\right] = 0$
$\pi = f^{-1}, h(\pi) = p \times p(\text{dit}(\pi))$	$h(F : \psi) = \text{tr}\left[P_{[qudit(F)]}\rho(\psi) \otimes \rho(\psi)\right]$
$h(\pi, \sigma) = p \times p(\text{dit}(\pi) \cup \text{dit}(\sigma))$	$h(F, G : \psi) =$ $\text{tr}\left[P_{[qudit(F)\cup qudit(G)]}\rho(\psi) \otimes \rho(\psi)\right]$
$h(\pi\|\sigma) = p \times p(\text{dit}(\pi) - \text{dit}(\sigma))$	$h(F\|G : \psi) =$ $\text{tr}\left[P_{[qudit(F)-qudit(G)]}\rho(\psi) \otimes \rho(\psi)\right]$
$m(\pi, \sigma) = p \times p(\text{dit}(\pi) \cap \text{dit}(\sigma))$	$m(F, G : \psi) =$ $\text{tr}\left[P_{[qudit(F)\cap qudit(G)]}\rho(\psi) \otimes \rho(\psi)\right]$
$h(\pi) = h(\pi\|\sigma) + m(\pi, \sigma)$	$h(F : \psi) = h(F\|G : \psi) + m(F, G : \psi)$
$\rho(\pi) = \widehat{\rho}(\mathbf{0}_U) = \sum_{r\in f(U)} P_r\rho(\mathbf{0}_U)P_r$	$\widehat{\rho}(\psi) = \sum_{r\in f(U)} P_r\rho(\psi)P_r$
$h(\pi) = 1 - \text{tr}\left[\rho(\pi)^2\right]$	$h(F : \psi) = 1 - \text{tr}\left[\widehat{\rho}(\psi)^2\right]$

For an observable F, let $f : U \to \mathbb{R}$ be the F-eigenvalue function assigning the real eigenvalue $f(u_i)$ for each u_i in the ON basis $U = \{u_1, \dots, u_n\}$ of F-eigenvectors. The image $f(U)$ is the set of F-eigenvalues $\{r_1, \dots, r_m\}$. Let $P_r : V \to V$ be the projection matrix in the U-basis to the eigenspace of r. The projective F-measurement of the state ψ transforms the pure state density matrix $\rho(\psi)$ (represented in the ON basis U of F-eigenvectors) to yield the Lüders mixture density matrix $\widehat{\rho}(\psi) = \sum_{r\in f(U)} P_r\rho(\psi)P_r$ [18, p. 279]. The off-diagonal elements of $\rho(\psi)$ that are zeroed in $\widehat{\rho}(\psi)$ are the coherences (quantum indistinctions or *quindits*) that are turned into "decoherences" (quantum distinctions or qudits of the observable being measured).

For any observable F and a pure state ψ, a quantum logical entropy was defined as $h(F : \psi) = \text{tr}\left[P_{[qudit(F)]}\rho(\psi) \otimes \rho(\psi)\right]$. That definition was the quantum generalization of the "classical" logical entropy defined as $h(\pi) = p \times p(\text{dit}(\pi))$. When a projective F-measurement is performed on ψ, the pure state density matrix $\rho(\psi)$ is transformed into the mixed state density matrix by the quantum Lüders mixture operation, which then defines the quantum logical entropy $h(\widehat{\rho}(\psi)) = 1 - \text{tr}\left[\widehat{\rho}(\psi)^2\right]$.

The first result is to show that these two entropies are the same: $h(F:\psi) = h(\widehat{\rho}(\psi))$. The proof proceeds by showing that they are both equal to *classical* logical entropy of the partition $\pi(F:\psi)$ defined on the ON basis $U = \{u_1, ..., u_n\}$ of F-eigenvectors by the F-eigenvalues with the point probabilities $p_i = \alpha_i^* \alpha_i$ where $|\psi\rangle = \sum_{i=1}^n \alpha_i |u_i\rangle$.[1] That is, the inverse images $B_j = f^{-1}(r_j)$ for $j = 1, ..., m$ of the eigenvalue function $f : U \to \mathbb{R}$ define the eigenvalue partition $\pi(F:\psi) = \{B_1, ..., B_m\}$ on the ON basis $U = \{u_1, ..., u_n\}$ with the point probabilities $p_i = \alpha_i^* \alpha_i = |\alpha_i|^2$ provided by the state ψ for $i = 1, ..., n$. The classical logical entropy of that partition is: $h(\pi(F:\psi)) = 1 - \sum_{j=1}^m p(B_i)^2$ where $p(B_j) = \sum_{u_i \in B_j} p_i$.

$$\begin{aligned} h(F:\psi) = \text{tr}\Big[P_{[qudit(F)]}\rho(\psi) \otimes \rho(\psi)\Big] &= \textstyle\sum_{i.k=1}^n \{p_i p_k : f(u_i) \neq f(u_k)\} \\ = \textstyle\sum_{j \neq j'} \sum \{p_i p_k : u_i \in B_j, u_k \in B_{j'}\} &= \textstyle\sum_{j \neq j'} p(B_j) p(B_{j'}) \\ = 1 - \textstyle\sum_{j=1}^m p(B_j)^2 &= h(\pi(F:\psi)). \end{aligned}$$

To show that $h(\widehat{\rho}(\psi)) = 1 - \text{tr}\Big[\widehat{\rho}(\psi)^2\Big] = h(\pi(F:\psi))$ for $\widehat{\rho}(\psi) = \sum_{r \in f(U)} P_r \rho(\psi) P_r$, we need to compute $\text{tr}\Big[\widehat{\rho}(\psi)^2\Big]$. An off-diagonal element in $\rho_{ik}(\psi) = \alpha_i \alpha_k^*$ of $\rho(\psi)$ survives (i.e., is not zeroed and has the same value) the Lüders operation if and only if $f(u_i) = f(u_k)$. Hence, the i-th diagonal element of $\widehat{\rho}(\psi)^2$ is:

$$\textstyle\sum_{k=1}^n \{\alpha_i^* \alpha_k \alpha_i \alpha_k^* : f(u_i) = f(u_k)\} = \sum_{k=1}^n \{p_i p_k : f(u_i) = f(u_k)\} = p_i p(B_j)$$

where $u_i \in B_j$. Then, grouping the i-th diagonal elements for $u_i \in B_j$ gives $\sum_{u_i \in B_j} p_i p(B_j) = p(B_j)^2$. Hence, the whole trace is: $\text{tr}\Big[\widehat{\rho}(\psi)^2\Big] = \sum_{j=1}^m p(B_j)^2$, and thus:

$$h(\widehat{\rho}(\psi)) = 1 - \text{tr}\Big[\widehat{\rho}(\psi)^2\Big] = 1 - \textstyle\sum_{j=1}^m p(B_j)^2 = h(F:\psi).$$

This finishes the proof of the following proposition.

Proposition 4. $h(F:\psi) = h(\pi(F:\psi)) = h(\widehat{\rho}(\psi))$.

This shows how the quantum case is so closely related to the set case that, in many instances, we can compute results in the quantum case by converting to the set case where computations are simpler.

Measurement creates distinctions, i.e., turns coherences into "decoherences," which, classically, is the operation of distinguishing elements by classifying them according to some attribute like classifying the faces of a die by their parity. The fundamental theorem about quantum logical entropy and projective measurement, in the density matrix version, shows how the quantum logical entropy created (starting with $h(\rho(\psi)) = 0$ for the pure state ψ) by the measurement can be computed directly from the coherences of $\rho(\psi)$ that are decohered in $\widehat{\rho}(\psi)$.

Proposition 5 (Measuring measurement)**.** *The increase in quantum logical entropy, $h(F:\psi) = h(\widehat{\rho}(\psi))$ due to the F-measurement of the pure state ψ is the sum of the absolute squares of the non-zero off-diagonal terms (coherences) in $\rho(\psi)$(represented in an ON basis of F-eigenvectors) that are zeroed ("decohered") in the post-measurement Lüders mixture density matrix $\widehat{\rho}(\psi) = \sum_{r \in f(U)} P_r \rho(\psi) P_r$.*

Proof. $h(\widehat{\rho}(\psi)) - h(\rho(\psi)) = \Big(1 - \text{tr}\Big[\widehat{\rho}(\psi)^2\Big]\Big) - \Big(1 - \text{tr}\Big[\rho(\psi)^2\Big]\Big) = \sum_{i,k} \Big(|\rho_{ik}(\psi)|^2 - |\widehat{\rho}_{ik}(\psi)|^2\Big)$. Now u_i and u_k are a qudit of F iff they are the corresponding off-diagonal terms zeroed by the Lüders mixture operation $\sum_{r \in f(U)} P_r \rho(\psi) P_r$ to obtain $\widehat{\rho}(\psi)$ from $\rho(\psi)$. □

Since $h(F:\psi) = h(\pi(F:\psi))$ we can carry over the probability interpretation in the classical case $h(\pi(F:\psi))$ to the quantum case.

Interpretation of quantum logical entropy

The quantum logical entropy $h(F:\psi)$ is the probability,
in two independent F-measurements of a prepared pure state ψ,
that different eigenvalues will be obtained—

[1] Whenever possible, we ignore the ket notation $|u_i\rangle$ and just write u_i.

just as the logical entropy $h(f^{-1})$ is the probability
in two independent draws from U
that different f-values will be obtained.

Example (continued): It might be helpful to carry out a quantum version of the numerical example. In $V = \mathbb{C}^3$, let $|\psi\rangle = \alpha|a\rangle + \beta|b\rangle + \gamma|c\rangle$ be a normalized state vector so that $p_a = \alpha\alpha^* = |\alpha|^2 = \frac{1}{2}$, $p_b = \beta\beta^* = |\beta|^2 = \frac{1}{3}$, and $p_c = \gamma\gamma^* = |\gamma|^2 = \frac{1}{6}$ are the point probabilities on the ON basis $U = \{a, b, c\}$ as in our running numerical example. Let $F : V \to V$ be a Hermitian operator with the eigenvalue function $f : U \to \mathbb{R}$ so that $f^{-1}(r_1) = \{a, b\}$ and $f^{-1}(r_2) = \{c\}$. In the quantum case, the complex density matrix $\rho_C(\psi)$ is:

$$\rho_C(\psi) = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \begin{bmatrix} \alpha^* & \beta^* & \gamma^* \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \alpha\beta^* & \alpha\gamma^* \\ \beta\alpha^* & \frac{1}{3} & \beta\gamma^* \\ \gamma\alpha^* & \gamma\beta^* & \frac{1}{6} \end{bmatrix}.$$

The corresponding real density matrix for those probabilities is:

$$\rho(\psi) = \rho(\mathbf{0}_U) = \begin{bmatrix} 1/2 & \sqrt{1/6} & \sqrt{1/12} \\ \sqrt{1/6} & 1/3 & \sqrt{1/18} \\ \sqrt{1/12} & \sqrt{1/18} & 1/6 \end{bmatrix}.$$

The set partition on the ON basis set U is $\pi = f^{-1} = \{\{a, b\}, \{c\}\}$ so the ditset is $\text{dit}(\pi) = \{(a, c), (b, c), (c, a), (c, b)\} = \{(a, b), (b, c), ...\}$ (where the ellipsis ... stands for the reversed cases of the previously listed ordered pairs). Hence the qudits in $V \otimes V$ are $qudit(F) = \{a \otimes c, b \otimes c, ...\}$. The quantum logical entropy resulting from the F-measurement (always projective) of ψ is $h(F : \psi) = \text{tr}\left[P_{[qudit(F)]}\rho_C(\psi) \otimes \rho_C(\psi)\right]$ where $\rho_C(\psi) \otimes \rho_C(\psi)$ is a 9×9 complex matrix which could be written with some shorthand (since each cell is a 3×3 matrix) as:

$$\rho_C(\psi) \otimes \rho_C(\psi) = \begin{bmatrix} \frac{1}{2}\rho_C(\psi) & \alpha\beta^*\rho_C(\psi) & \alpha\gamma^*\rho_C(\psi) \\ \beta\alpha^*\rho_C(\psi) & \frac{1}{3}\rho_C(\psi) & \beta\gamma^*\rho_C(\psi) \\ \gamma\alpha^*\rho_C(\psi) & \gamma\beta^*\rho_C(\psi) & \frac{1}{6}\rho_C(\psi) \end{bmatrix}.$$

The diagonal elements of $\rho_C(\psi) \otimes \rho_C(\psi)$ are real products of probabilities. The projection operator $P_{[qudit(F)]}$ to the subspace generated by $qudit[F]$ in $V \otimes V$ is a 9×9 diagonal matrix whose non-zero diagonal entries are ones corresponding to $qudits(F) = \{a \otimes c, b \otimes c, ...\}$. Thus the product $P_{[qudit(F)]}\rho_C(\psi) \otimes \rho_C(\psi)$ just picks out (along its diagonal) those pairs of probabilities corresponding to $\text{dit}(f^{-1})$, namely $\{p_a p_c, p_b p_c, ...\}$ and then taking the trace sums them up to yield $h(F : \psi) = h(\pi(F : \psi))$. It sums the four entries corresponding to the qudits, $\frac{1}{2}\frac{1}{6}$ for $a \otimes c$, $\frac{1}{3}\frac{1}{6}$ for $b \otimes c$, and the two reverse products for a total of:

$$2(\tfrac{1}{12} + \tfrac{1}{18}) = 2(\tfrac{3}{36} + \tfrac{2}{36}) = \tfrac{10}{36} = \tfrac{5}{18}.\checkmark$$

Quantum logical entropy also has natural connections with other quantum notions such as the Hilbert-Schmidt distance $\text{tr}[(\rho - \tau)^2]$ [20] between two density matrices ρ and τ. And, as usual, the quantum case is developed as the quantum version of the "classical" logical entropy. We previously defined the logical Hamming distance between two partitions.

$$d(\pi, \sigma) := h(\pi|\sigma) + h(\sigma|\pi) = h(\pi \vee \sigma) - m(\pi, \sigma) = 2h(\pi \vee \sigma) - h(\pi) - h(\sigma).$$

Nielsen and Chuang are skeptical about developing the Hamming distance in the quantum context.

> 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! [24, p. 399]

We have seen how to define a density matrix $\rho(\pi)$ from a set partition U, and this provides a bridge to quantum logical entropy for any density matrix ρ:

$$h(\pi) = 1 - \text{tr}\left[\rho(\pi)^2\right] \text{ extends to } h(\rho) = 1 - \text{tr}\left[\rho^2\right].$$

The formula $h(\rho) = 1 - \text{tr}[\rho^2]$ is not new; only the whole development of logical entropy from partition logic is new (and hence the new name). Indeed, $\text{tr}[\rho^2]$ is usually called the *purity* of the density matrix since a state ρ is *pure* if and only if $\text{tr}[\rho^2] = 1$, so $h(\rho) = 0$, and otherwise, $\text{tr}[\rho^2] < 1$, so $h(\rho) > 0$. The complement $1 - \text{tr}[\rho^2]$ has been called the "mixedness" [25, p. 5] of the state ρ. The seminal paper of Manfredi and Feix [3] approaches the same formula $1 - \text{tr}[\rho^2]$ (which they denote as S_2) from the advanced viewpoint of Wigner functions, and they present strong arguments for this notion of quantum entropy (which resulted in Manfredi editing a special issue of the journal *4Open* on logical entropy [26]). This notion of quantum logical entropy is also called by the misnomer "linear entropy" [27] even though it is quadratic in ρ, so we will not continue that usage.

Using these density matrices, there is also the notion of *logical cross-entropy of π and σ*:

$$h(\pi||\sigma) = 1 - \text{tr}[\rho(\pi)\rho(\sigma)]$$

where $h(\pi||\sigma) = h(\pi, \sigma) = h(\pi \vee \sigma)$. Since there is no join defined for density matrices ρ and τ, we can nevertheless define the *quantum logical cross-entropy of ρ and τ* as (where the dagger is conjugate transpose):

$$h(\rho||\tau) := 1 - \text{tr}[\rho^\dagger \tau].$$

This provides the path to then define the *quantum logical Hamming distance between ρ and τ*:

$$d(\rho, \tau) := 2h(\rho||\tau) - h(\rho) - h(\tau).$$

Since ρ and τ are also Hermitian matrices, each has an ON basis of eigenvectors and this approach to Hamming distance avoids the Nielsen-Chuang misgivings by using the amplitudes of all possible relations between the two ON bases [4, pp. 89–90] so no arbitrary labeling of the bases is involved. Then we have the theorem connecting quantum logical Hamming distance to an important existing notion in quantum information theory.

Proposition 6 (Hamming = Hilbert-Schmidt distance)**.** $\text{tr}\left[(\rho - \tau)^2\right] = d(\rho, \tau)$.

Proof. $\text{tr}\left[(\rho - \tau)^2\right] = \text{tr}[\rho^2 + \tau^2 - 2\rho^\dagger\tau] = \text{tr}[\rho^2] + \text{tr}[\tau^2] - 2\,\text{tr}[\rho^\dagger\tau] = 2h(\rho||\tau) - h(\rho) - h(\tau) = d(\rho, \tau)$. □

8. Concluding Remarks

In that manner, the computation of the quantum logical entropies can be reduced to the computations in the corresponding "classical" case of logical entropies. Moreover, the definitions are made in Table 7 so that we have all the usual compound notions of quantum logical entropy that satisfy the usual Venn diagram relationships as illustrated in Figure 4.

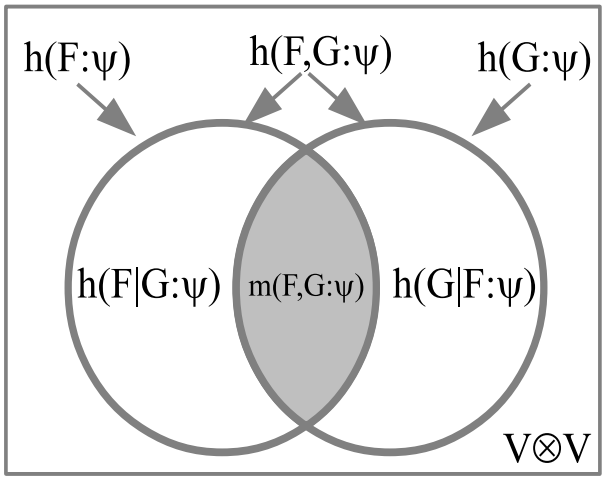

Figure 4. Venn diagram relationships for quantum logical entropy.

The overall purpose of this paper has been to develop quantum logical entropy starting from the logic of partitions at the set level, developing the quantitative version of partitions as logical entropy, and then the development of that

corresponding quantum notion using the yoga of linearization to translate the set concepts into the corresponding (Hilbert) vector space concepts.

There are a number of other results in the literature ([4,20,28]) about quantum logical entropy such as its concavity, subadditivity, non-decreasing value under projective measurement, a Holevo-type bound for quantum logical Hamming distance, and the extension of quantum logical entropy to post-selected quantum systems. More results are sure to come as more researchers are familiar with the logical entropy concepts starting with the quantitative treatment of partitions in terms of distinctions.

> We find this framework of partitions and distinction most suitable (at least conceptually) for describing the problems of quantum state discrimination, quantum cryptography and in general, for discussing quantum channel capacity. In these problems, we are basically interested in a distance measure between such sets of states, and this is exactly the kind of knowledge provided by logical entropy [Reference to [1]]. [2, p. 1]

In conventional information theory or in what Claude Shannon called the "Mathematical Theory of Communication" [14], he noted that "no concept of information itself was defined" [29, p. 458]. The extension of Shannon entropy to the quantum notion of von Neumann entropy did not solve that problem of defining quantum information or the problem of interpretation. Logical entropy as the quantification of partitions defines the notion of information-as-distinctions, and quantum logical entropy extends that notion to the quantum realm as the quantification of quantum distinctions or qudits. This answers the vision of Charles Bennett, one of the founders of quantum information theory.

> So information really is a very useful abstraction. It 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.... [30, pp. 155–157]

Funding
This research received no external funding.

Conflicts of Interest Statement
The author declares no conflict of interest.

Data Availability Statement
The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. D. Ellerman (2009). "Counting distinctions: On the conceptual foundations of shannon's information theory." *Synthese*, 168: 1 May, 119–149.
2. B. Tamir and E. Cohen (2014). "Logical entropy for quantum states." *ArXiv.org*. http://de.arxiv.org/abs/1412.0616v2.
3. G. Manfredi and M. R. Feix (2000). "Entropy and wigner functions." *Physical Review E*, 62, 4665–4674. https://doi.org/10.1103/PhysRevE.62.4665.
4. D. Ellerman (2021). *New Foundations for Information Theory: Logical Entropy and Shannon Entropy*. Cham, Switzerland: SpringerNature. https://doi.org/10.1007/978-3-030-86552-8.
5. D. Ellerman (2010). "The logic of partitions: Introduction to the dual of the logic of subsets." *Review of Symbolic Logic*, 3: 2 June, 287–350.
6. D. Ellerman (2014). "An introduction of partition logic." *Logic Journal of the IGPL*, 22: 1, 94–125.
7. D. Ellerman (2023). *The Logic of Partitions: With Two Major Applications. Studies in Logic 101*. London: College Publications. https://www.collegepublications.co.uk/logic/?00052.
8. G. Boole (1854). *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*. Cambridge: Macmillan and Co.
9. F. William Lawvere and R. Rosebrugh (2003). *Sets for Mathematics*. Cambridge: Cambridge University Press.
10. T. Britz, M. Mainetti and L. Pezzoli (2001). "Some operations on the family of equivalence relations", in H. Crapo and D. Senato (eds), *Algebraic Combinatorics and Computer Science: A Tribute to Gian-Carlo Rota*. Milano: Springer, 445–459.

11. J. P. S. Kung, G.-C. Rota, and C. H. Yan. (2009). *Combinatorics: The Rota Way*. New York: Cambridge University Press.
12. G.-C. Rota (2001). "Twelve problems in probability no one likes to bring up", in H. Crapo and D. Senato (eds), *Algebraic Combinatorics and Computer Science*. Milano: Springer, 57–93.
13. A. N. Kolmogorov (1983). "Combinatorial foundations of information theory and the calculus of probabilities." *Russian Mathematical Surveys*, 38: 4, 29–40.
14. C. E. Shannon (1948). "A mathematical theory of communication." *Bell System Technical Journal*, 27: 379–423; 623–656.
15. C. E. Shannon and W. Weaver. (1964). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
16. L. L. Campbell (1965). "Entropy as a measure." *IEEE Transactions on Information Theory*, IT-11: January, 112–114.
17. U. Fano (1957). "Description of states in quantum mechanics by density matrix and operator techniques." *Reviews of Modern Physics*, 29: 1, 74–93.
18. G. Auletta, M. Fortunato, and G. Parisi (2009). *Quantum Mechanics*. Cambridge: Cambridge University Press.
19. G. Rossi (2011). "Partition distances." *arXiv:1106.4579v1*.
20. B. Tamir and E. Cohen (2015). "A Holevo-type bound for a Hilbert Schmidt distance measure." *Journal of Quantum Information Science*, 5, 127–133.
21. V. Vedral and M. B. Plenio (1998). "Entanglement measures and purification procedures." *Physical Review A*, 57: 3, 1619–1633.
22. G.-C. Rota (1997). *Indiscrete Thoughts*. Boston: Birkhäuser.
23. S. Roman (2008). *Advanced Linear Algebra 3rd. Ed*. New York: Springer Science+Business Media.
24. M. Nielsen and I. Chuang (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
25. G. Jaeger (2007). *Quantum Information: An Overview*. New York: Springer Science+Business Media.
26. G. Manfredi (ed.) (2022). "Logical entropy – special issue." *4Open*, no. 5, E1. https://doi.org/10.1051/fopen/2022005.
27. F. Buscemi, P. Bordone, and A. Bertoni (2007). "Linear entropy as an entanglement measure in two-fermion systems." *ArXiv.org*, March 2. http://arxiv.org/abs/quant-ph/0611223v2.
28. B. Tamir, I. L. Piava, Z. Schwartzman-Nowik, and E. Cohen (2022). "Quantum logical entropy: Fundamentals and general properties." *4Open Special Issue: Logical Entropy*, 5: 2, 1–14. https://doi.org/10.1051/fopen/2021005.
29. C. E. Shannon (1993). "Some topics in information theory", in N. J. A. Sloane and A. D. Wyner (eds), *Claude E. Shannon: Collected Papers*. Piscataway NJ: IEEE Press, 458–459.
30. C. H. Bennett (2003). "Quantum information: Qubits and quantum error correction." *International Journal of Theoretical Physics*, 42: 2 February, 153–176.