On Classical and Quantum Logical Entropy: The analysis of measurement

David Ellerman
University of California at Riverside

September 8, 2016

Abstract

The notion of a partition on a set is mathematically dual to the notion of a subset of a set, so there is a logic of partitions dual to Boole's logic of subsets (Boolean subset logic is usually mis-specified as the special case of "propositional" logic). The notion of an *element* of a subset has as its dual the notion of a *distinction* of a partition (a pair of elements in different blocks). Boole developed finite logical probability as the normalized counting measure on elements of subsets so there is a dual concept of *logical entropy* which is the normalized counting measure on distinctions of partitions. Thus the logical notion of information is a measure of distinctions. Classical logical entropy also extends naturally to the notion of *quantum logical entropy* which provides a more natural and informative alternative to the usual von Neumann entropy in quantum information theory. The quantum logical entropy of a post-measurement density matrix has the simple interpretation as the probability that two independent measurements of the same state using the same observable will have different results.

The main result of the paper is that this increase in quantum logical entropy due to a projective measurement of a pure state is the sum of the absolute squares of the off-diagonal entries ('coherences') of the pure state density matrix that are zeroed (decohered) by the measurement, i.e., the measure of the distinctions ('decoherences') created by the measurement. The von Neumann entropy provides no such analysis of measurement. That result is also classically modelled using ordinary partitions and density matrices for such partitions in the pedagogical model of QM/Sets with point probabilities–which for a fixed basis is just sampling a real r.v. in classical finite probability theory.

Contents

1 Introduction and history

The formula for what is here called "classical logical entropy" is not new. Given a finite probability distribution $p = (p_1, ..., p_n)$, the formula $h(p) = 1 - \sum_{i=1}^{n} p_i^2$ was used by Gini in 1912 ([13] reprinted in [14, p. 369]) as a measure of “mutability” or diversity. What is new is the derivation of the formula from recent developments in mathematical logic.

Although usually named after the special case of "propositional" logic, the general case is Boole's logic of subsets of a universe U (the special case of $U = 1$ allows the propositional interpretation since the only subsets are 1 and $\emptyset$ standing for truth and falsity). Category theory shows that is a duality between sub-sets and quotient-sets (or partitions and equivalence relations), and that allowed the recent development of the dual logic of partitions ([7], [10]). As indicated in the title of his book, *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities* [3], Boole also developed the normalized counting measure on subsets of a finite universe U which was finite logical probability theory. When the same mathematical notion of the normalized counting measure is applied to the partitions on a finite universe set U (when the partition is represented as the complement of the corresponding equivalence relation on $U \times U$) then the result is the formula for logical entropy.

The formula in the complementary form, $\sum_i p_i = 1 - h(p)$, was used early in the 20^{th} century in cryptography. The American cryptologist, William F. Friedman, devoted a 1922 book ([12]) to the "index of coincidence" (i.e., $\sum p_i^2$). Solomon Kullback worked as an assistant to Friedman and wrote a book on cryptology which used the index. [21] During World War II, Alan M. Turing worked for a time in the Government Code and Cypher School at the Bletchley Park facility in England. Probably unaware of the earlier work, Turing used $\rho = \sum p_i^2$ in his cryptoanalysis work and called it the *repeat rate* since it is the probability of a repeat in a pair of independent draws from a population with those probabilities.

After the war, Edward H. Simpson, a British statistician, proposed $\sum_{B\in\pi} p_B^2$ as a measure of species concentration (the opposite of diversity) where $\pi = \{B, B', ...\}$ is the partition of animals or plants according to species and where each animal or plant is considered as equiprobable so $p_B = \frac{|B|}{|U|}$. And Simpson gave the interpretation of this homogeneity measure as "the probability that two individuals chosen at random and independently from the population will be found to belong to the same group."[30, p. 688] Hence $1 - \sum_{B\in\pi} p_B^2$ is the probability that a random ordered pair will belong to different species, i.e., will be distinguished by the species partition. In the biodiversity literature [28], the formula $1 - \sum_{B\in\pi} p_B^2$ is known as "Simpson's index of diversity" or sometimes, the *Gini-Simpson index* [27].

However, Simpson along with I. J. Good worked at Bletchley Park during WWII, and, according to Good, "E. H. Simpson and I both obtained the notion [the repeat rate] from Turing." [15, p. 395] When Simpson published the index in 1948, he (again, according to Good) did not acknowledge Turing "fearing that to acknowledge him would be regarded as a breach of security." [16, p. 562] I. J. Good pointed out a certain naturalness:

> If $p_1, ..., p_t$ are the probabilities of t mutually exclusive and exhaustive events, any statistician of this century who wanted a measure of homogeneity would have take about two seconds to suggest $\sum p_i^2$ which I shall call ρ. [16, p. 561]

In view of the frequent and independent discovery and rediscovery of the formula $\rho = \sum p_i^2$ or its complement $h(p) = 1 - \sum p_i^2$ by Gini, Friedman, Turing, and many others [e.g., the Hirschman-Herfindahl index of industrial concentration in economics ([19], [18])], I. J. Good wisely advises that "it is unjust to associate ρ with any one person." [16, p. 562]

2 Duality of subsets and partitions

Logical entropy is to the logic of partitions as logical probability is to the Boolean logic of subsets. Hence we will start with a brief review of the relationship between these two dual forms of logic.

Modern category theory shows that the concept of a subset dualizes to the concept of a quotient set, equivalence relation, or partition. F. William Lawvere called a subset or, in general, a subobject a "part" and then noted: “The dual notion (obtained by reversing the arrows) of ‘part’ is the notion of *partition*.” [23, p. 85] That suggests that the Boolean logic of subsets should have a dual logic of partitions ([7], [10]).

A *partition* $\pi = \{B_1, ..., B_m\}$ on U is a set of subsets or "blocks" B_i that are mutually disjoint and jointly exhaustive ($\cup_i B_i = U$). In the duality between subset logic and partition logic, the dual to the notion of an "element" (an ‘it’) of a subset is the notion of a "distinction" (a ‘dit’) of a partition, where $(u, u') \in U \times U$ is a *distinction* or *dit* of π if the two elements are in different blocks. Let $\operatorname{dit}(\pi) \subseteq U \times U$ be the set of distinctions or *ditset* of π. Similarly an *indistinction* or *indit* of π is a pair $(u, u') \in U \times U$ in the same block of π. Let $\operatorname{indit}(\pi) \subseteq U \times U$ be the set of indistinctions or *inditset* of π. Then $\operatorname{indit}(\pi)$ is the equivalence relation associated with π and $\operatorname{dit}(\pi) = U \times U - \operatorname{indit}(\pi)$ is the complementary binary relation that might be called a *partition relation* or an *apartness relation*.

3 Classical subset logic and partition logic

The algebra associated with the subsets $S \subseteq U$ is, of course, the Boolean algebra $\wp(U)$ of subsets of U with the partial order as the inclusion of elements. The corresponding algebra of partitions π on U is the *partition algebra* $\prod(U)$ defined as follows:

- the *partial order* $\sigma \preceq \pi$ of partitions $\sigma = \{C, C', ...\}$ and $\pi = \{B, B', ...\}$ holds when π refines σ in the sense that for every block $B \in \pi$ there is a block $C \in \sigma$ such that $B \subseteq C$, or, equivalently, using the element-distinction (‘its’ & ‘dits’) pairing, the partial order is the inclusion of distinctions: $\sigma \preceq \pi$ if and only if (iff) $\operatorname{dit}(\sigma) \subseteq \operatorname{dit}(\pi)$;

- the minimum or bottom partition is the *indiscrete partition* (or blob) $\mathbf{0} = \{U\}$ with one block consisting of all of U;

- the maximum or top partition is the *discrete partition* $\mathbf{1} = \{\{u_j\}\}_{j=1,...,n}$ consisting of singleton blocks;

- the *join* $\pi \vee \sigma$ is the partition whose blocks are the non-empty intersections $B \cap C$ of blocks of π and blocks of σ, or, equivalently, using the element-distinction pairing, $\operatorname{dit}(\pi \vee \sigma) = \operatorname{dit}(\pi) \cup \operatorname{dit}(\sigma)$;

- the *meet* $\pi \wedge \sigma$ is the partition whose blocks are the equivalence classes for the equivalence relation generated by: $u_j \sim u_{j'}$ if $u_j \in B \in \pi$, $u_{j'} \in C \in \sigma$, and $B \cap C \neq \emptyset$; and

- $\sigma \Rightarrow \pi$ is the *implication partition* whose blocks are: (1) the singletons $\{u_j\}$ for $u_j \in B \in \pi$ if there is a $C \in \sigma$ such that $B \subseteq C$, or (2) just $B \in \pi$ if there is no $C \in \sigma$ with $B \subseteq C$, so that trivially: $\sigma \Rightarrow \pi = \mathbf{1}$ iff $\sigma \preceq \pi$.[1]

Since the same operations can be defined for subsets and partitions, one can interpret a formula $\Phi(\pi, \sigma, ...)$ either way as a subset or a partition. Given either subsets on or partitions of U substituted for the variables π, σ,..., one can apply, respectively, subset or partition operations to evaluate the

[1] There is a general method to define operations on partitions corresponding to operations on subsets ([7], [10]) but the lattice operations of join and meet, and the implication operation are sufficient to define a partition algebra $\prod(U)$ parallel to the familiar powerset Boolean algebra $\wp(U)$.

whole formula. Since $\Phi(\pi,\sigma,...)$ is either a subset or a partition, the corresponding proposition is "u is an element of $\Phi(\pi,\sigma,...)$" or "(u,u') is a distinction of $\Phi(\pi,\sigma,...)$". And then the definitions of a valid formula are also parallel, namely, no matter what is substituted for the variables, the whole formula evaluates to the top of the algebra. In that case, the subset $\Phi(\pi,\sigma,...)$ contains all elements of U, i.e., $\Phi(\pi,\sigma,...) = U$, or the partition $\Phi(\pi,\sigma,...)$ distinguishes all pairs (u,u') for distinct elements of U, i.e., $\Phi(\pi,\sigma,...) = \mathbf{1}$. The parallelism between the dual logics is summarized in the following table 1.

Table 1	Subset logic	Partition logic
'Elements' (its or dits)	Elements u of S	Distinctions (u,u') of π
Inclusion of 'elements'	Inclusion $S \subseteq T$	Refinement: $\operatorname{dit}(\sigma) \subseteq \operatorname{dit}(\pi)$
Top of order = all 'elements'	U all elements	$\operatorname{dit}(\mathbf{1}) = U^2 - \Delta$, all dits
Bottom of order = no 'elements'	$\emptyset$ no elements	$\operatorname{dit}(\mathbf{0}) = \emptyset$, no dits
Variables in formulas	Subsets S of U	Partitions π on U
Operations: $\vee, \wedge, \Rightarrow, ...$	Subset ops.	Partition ops.
Formula $\Phi(x,y,...)$ holds	u element of $\Phi(S,T,...)$	(u,u') dit of $\Phi(\pi,\sigma,...)$
Valid formula	$\Phi(S,T,...) = U$, $\forall S,T,...$	$\Phi(\pi,\sigma,...) = \mathbf{1}$, $\forall \pi,\sigma,...$

Table 1: Duality between subset logic and partition logic

4 Classical probability and classical logical entropy

George Boole [3] extended his logic of subsets to classical finite probability theory where, in the equiprobable case, the probability of a subset S (event) of a finite universe set (outcome set or sample space) $U = \{u_1,...,u_n\}$ was the number of elements in S over the total number of elements: $\Pr(S) = \frac{|S|}{|U|} = \sum_{u_j \in S} \frac{1}{|U|}$. Laplace's classical finite probability theory [22] also dealt with the case where the outcomes were assigned real point probabilities $p = \{p_1,...,p_n\}$ (where $p_j \geq 0$ and $\sum_j p_j = 1$) so rather than summing the equal probabilities $\frac{1}{|U|}$, the point probabilities of the elements were summed: $\Pr(S) = \sum_{u_j \in S} p_j = p(S)$–where the equiprobable formula is for $p_j = \frac{1}{|U|}$ for $j = 1,...,n$. The conditional probability of an event $T \subseteq U$ given an event S is $\Pr(T|S) = \frac{p(T\cap S)}{p(S)}$. Given a real-valued random variable $f : U \to \mathbb{R}$ on the outcome set U, the possible values of f are $f(U) = \{\phi_1,...,\phi_m\}$ and the probability of getting a certain value given S is: $\Pr(\phi_i|S) = \frac{p\left(f^{-1}(\phi_i)\cap S\right)}{p(S)}$.

Then we may mimic Boole's move going from the logic of subsets to the finite logical probabilities of subsets by starting with the logic of partitions and using the dual relation between elements and distinctions. The dual notion to probability turns out to be "information content" or "entropy" so we define the *logical entropy* of π, denoted $h(\pi)$, as the size of the ditset $\operatorname{dit}(\pi) \subseteq U \times U$ normalized by the size of $U \times U$:

$$h(\pi) = \frac{|\operatorname{dit}(\pi)|}{|U\times U|} = \sum_{(u_j,u_k)\in \operatorname{dit}(\pi)} \frac{1}{|U|}\frac{1}{|U|}$$

Logical entropy of π (equiprobable case).

The inditset of π is $\operatorname{indit}(\pi) = \cup_{i=1}^{m} (B_i \times B_i)$ so where $p(B_i) = \frac{|B_i|}{|U|}$ in the equiprobable case, we have:

$$h(\pi) = \frac{|\operatorname{dit}(\pi)|}{|U\times U|} = \frac{|U\times U| - \sum_{i=1}^{m}|B_i\times B_i|}{|U\times U|} = 1 - \sum_{i=1}^{m}\left(\frac{|B_i|}{|U|}\right)^2 = 1 - \sum_{i=1}^{m} p(B_i)^2.$$

This definition corresponds to Boole's equiprobable case $\Pr(S) = \frac{|S|}{|U|}$ of the normalized number of elements rather than normalized number of distinctions.

The corresponding definition for the case of point probabilities $p = \{p_1, ..., p_n\}$ is to just add up the probabilities of getting a particular distinction, i.e., the count of the probability-weighted distinctions of π:

$$h_p(\pi) = \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k$$

Logical entropy of π with point probabilities p.

This suggests that in the case of point probabilities, we should take $p(B_i) = \sum_{u_j\in B_i} p_j$ and have $h_p(\pi) = 1 - \sum_{i=1}^m p(B_i)^2$. This is confirmed with a little calculation using that definition of $p(B_i)$:

$$1 = [p(B_1) + ... + p(B_m)]\,[p(B_1) + ... + p(B_m)] = \sum_{i=1}^m p(B_i)^2 + \sum_{i\neq i'} p(B_i)\,p(B_{i'})$$

so that:

$$1 - \sum_{i=1}^m p(B_i)^2 = \sum_{i\neq i'} p(B_i)\,p(B_{i'}).^2$$

Moreover, we have:

$$\sum_{i=1}^m p(B_i)^2 = \sum_i \left(\sum_{u_j\in B_i} p_j\right)^2 = \sum_i \sum_{(u_j,u_k)\in B_i\times B_i} p_j p_k = \sum_{(u_j,u_k)\in\operatorname{indit}(\pi)} p_j p_k$$

so that:

$$1 - \sum_{i=1}^m p(B_i)^2 = \sum_{i\neq i'} p(B_i)\,p(B_{i'}) = 1 - \sum_{(u_j,u_k)\in\operatorname{indit}(\pi)} p_j p_k = \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k$$

since:

$$\begin{aligned}1 &= (p_1 + ... + p_n)(p_1 + ... + p_n) = \sum_{(u_j,u_k)\in U\times U} p_j p_k \\ &= \sum_{(u_j,u_k)\in\operatorname{indit}(\pi)} p_j p_k + \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k.\end{aligned}$$

Thus the logical entropy with point probabilities is (using the point probability definition of $p(B_i)$):

$$h_p(\pi) = \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k = \sum_{i\neq i'} p(B_i)\,p(B_{i'}) = 2\sum_{i<i'} p(B_i)\,p(B_{i'}) = 1 - \sum_{i=1}^m p(B_i)^2.$$

One other version of the classical logical entropy might be mentioned. Instead of being given a partition $\pi = \{B_1, ..., B_m\}$ on U with point probabilities p_j defining the finite probability distribution of block probabilities $\{p(B_i)\}_i$, one might be given only a finite probability distribution $p = \{p_1, ..., p_m\}$. The substituting p_i for $p(B_i)$ gives the:

$$h(p) = 1 - \sum_{i=1}^m p_i^2$$

logical entropy of a finite probability distribution.

There are also parallel element $\leftrightarrow$ distinction interpretations:

- $\Pr(S) = p_S$ is the probability that a single draw, sample, or experiment with U gives a *element* u_j of S, and
- $h_p(\pi) = \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k = \sum_{i\neq i'} p(B_i)\,p(B_{i'})$ is the probability that two independent (with replacement) draws, samples, or experiments with U gives a *distinction* (u_j, u_k) of π, or if we interpret the independent experiments as sampling from the set of blocks $\pi = \{B_i\}$, then it is the probability of getting distinct blocks.

The parallelism or duality between logical probabilities and logical entropies is summarized in the following table 2.

[2] A pair $\{i, i'\}$ of distinct indices satisfies $i \neq i'$ both ways so $\sum_{i\neq i'} p_{B_i} p_{B_{i'}} = 2\sum_{i<i'} p_{B_i} p_{B_{i'}}$.

Table 2	Logical Probability Theory	Logical Information Theory
'Outcomes'	Elements $u \in U$ finite	Dits $(u, u') \in U \times U$ finite
'Events'	Subsets $S \subseteq U$	Ditsets $\mathrm{dit}(\pi) \subseteq U \times U$
Equiprobable points	$\Pr(S) = \frac{\lvert S\rvert}{\lvert U\rvert}$	$h(\pi) = \frac{\lvert \mathrm{dit}(\pi)\rvert}{\lvert U \times U\rvert}$
Point probabilities	$\Pr(S) = \sum \{p_j : u_j \in S\}$	$h(\pi) = \sum \{p_j p_k : (u_j, u_k) \in \mathrm{dit}(\pi)\}$
Interpretation	$\Pr(S) =$ 1-draw prob. of S-element	$h(\pi) =$ 2-draw prob. of π-distinction

Table 2: Classical logical probability theory and classical logical information theory

5 Classical logical entropy of a density matrix

Density matrices [25] are usually associated with quantum theory but they can also be used in ordinary or 'classical' logical information theory. In the general case of a finite outcome set $U = \{u_1, ..., u_n\}$ with point probabilities $p = \{p_1, ..., p_n\}$, then for any subset S with $p(S) > 0$ where $\chi_S(u_i)$ is the characteristic function for the subset S, we have a normalized column vector in $\mathbb{R}^n$ (with $[\cdots]^t$ representing the transpose):

$$|S\rangle = \frac{1}{\sqrt{p(S)}} \left[\chi_S(u_1)\sqrt{p_1}, ..., \chi_S(u_n)\sqrt{p_n}\right]^t.$$

If we denote the corresponding row vector by $\langle S|$, then we may define the $n \times n$ density matrix $\rho(S)$ as:

$$\rho(S) = |S\rangle\langle S| =$$
$$\frac{1}{p(S)} \begin{bmatrix} \chi_S(u_1)^2 p_1 & \chi_S(u_1)\chi_S(u_2)\sqrt{p_1p_2} & \cdots & \chi_S(u_1)\chi_S(u_n)\sqrt{p_1p_n} \\ \chi_S(u_2)\chi_S(u_1)\sqrt{p_2p_1} & \chi_S(u_2)^2 p_2 & \cdots & \chi_S(u_2)\chi_S(u_n)\sqrt{p_2p_n} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_S(u_n)\chi_S(u_1)\sqrt{p_np_1} & \chi_S(u_n)\chi_S(u_2)\sqrt{p_np_2} & \cdots & \chi_S(u_n)^2 p_n \end{bmatrix}$$

Density matrix for a subset $S \subseteq U$.

It is then easy to characterize each entry in the matrix:

$$(\rho(S))_{jk} = \begin{cases} \frac{1}{p(S)}\sqrt{p_jp_k} \text{ if } u_j, u_k \in S \\ 0 \text{ otherwise.} \end{cases}$$

Then a little calculation shows that the density matrix for a subset S acts like a "pure state" in QM in the sense that it is idempotent $\rho(S)^2 = \rho(S)$ since:

$$\left(\rho(S)^2\right)_{jk} = \frac{1}{p(S)^2}\sum_{l=1}^{n} \chi_S(u_j)\chi_S(u_l)\sqrt{p_jp_l}\chi_S(u_l)\chi_S(u_k)\sqrt{p_lp_k}$$
$$= \frac{1}{p(S)^2}\chi_S(u_j)\chi_S(u_k)\sqrt{p_jp_k}\sum_{l=1}^{n}\chi_S(u_l)p_l = \frac{1}{p(S)^2}\chi_S(u_j)\chi_S(u_k)\sqrt{p_jp_k}p(S) = (\rho(S))_{jk}.$$

Then given a partition $\pi = \{B_1, ..., B_m\}$ on U, the density matrix for π, like a mixed-state density matrix in QM, is the probability-weighted sum of the density matrices for the blocks:

$$\rho(\pi) = \sum_{i=1}^{m} p(B_i)\rho(B_i).$$

Density matrix for a partition π on U.

Since $p(B_i) \times \frac{1}{p(B_i)}\sqrt{p_jp_k} = \sqrt{p_jp_k}$ for $(u_j, u_k) \in \mathrm{indit}(\pi)$, the entries in $\rho(\pi)$ are easily characterized:

$$\rho(\pi)_{jk} = \begin{cases} \sqrt{p_jp_k} \text{ if } (u_j, u_k) \in \mathrm{indit}(\pi) \\ 0 \text{ otherwise.} \end{cases}$$

In particular, it should be noted how the $n \times n$ pairs (u_j, u_k) of $U \times U$, that are divided into distinctions $\operatorname{dit}(\pi)$ and indistinctions $\operatorname{indit}(\pi)$ of a partition π on U in partition logic, are paired with the $n \times n$ entries $\rho(\pi)_{jk}$ in the density matrix $\rho(\pi)$. That is the simple mathematical basis for using the notions of (classical or quantum) logical entropy to analyze (classical or quantum) density matrices. The Shannon or von Neumann entropies have no such direct pair (u_j, u_k) to entry ρ_{jk} connection with the entries ρ_{jk} in a density matrix ρ.

Then the classical logical entropy formula $h_p(\pi) = 1 - \sum_{i=1}^{m} p(B_i)^2$ easily generalizes to density matrices by replacing the sum by the trace (sum of diagonal elements of a matrix) and the squared probabilities by the square of the density matrix:

$$h(\rho(\pi)) = 1 - \operatorname{tr}\left[\rho(\pi)^2\right]$$

Logical entropy of a classical density matrix.

Since $\rho(S)^2 = \rho(S)$ and the trace of any density matrix is 1, we immediately have $h(\rho(S)) = 0$, i.e., the logical entropy of a "pure state" subset $S \subseteq U$ is 0.

We saw previously that $h_p(\pi) = 1 - \sum_{i=1}^{m} p(B_i)^2$ so we need to check that the density matrix formulation gives the same result–which only requires a little calculation of the diagonal entries of $\rho(\pi)^2$. If $u_j \in B_i$, then

$$\begin{gathered}\left(\rho(\pi)^2\right)_{jj} = \sum_{k=1}^{n} \rho(\pi)_{jk}\, \rho(\pi)_{kj} \\ = p_j \sum_{k=1,(u_j,u_k)\in \operatorname{indit}(\pi)}^{n} p_k = p_j p(B_i)\end{gathered}$$

so the sum of the diagonal entries for j where $u_j \in B_i$ is $p(B_i)^2$ and the sum of all the diagonal entries is:

$$\operatorname{tr}\left[\rho(\pi)^2\right] = \sum_{i=1}^{m} p(B_i)^2 = \sum_{(u_j,u_k)\in \operatorname{indit}(\pi)} p_j p_k.$$

Thus we have the density matrix treatment giving the same logical entropy $h_p(\pi)$:

$$h_p(\pi) = 1 - \sum_{i=1}^{m} p(B_i)^2 = \sum_{(u_j,u_k)\in \operatorname{dit}(\pi)} p_j p_k = 1 - \operatorname{tr}\left[\rho(\pi)^2\right] = h(\rho(\pi)).$$

6 Quantum logical entropy and measurement

Let V be an n-dimensional Hilbert space $\mathbb{C}^n$. If $|\psi_i\rangle$ for $i = 1, ..., m$ is a set of orthogonal normalized vectors from V and $p = \{p_1, ..., p_m\}$ is a probability distribution, then

$$\rho(\psi) = \sum_{i=1}^{m} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$$

is a density matrix, and any positive operator on V of trace 1 has such an orthogonal decomposition [25, p. 101]. Then the *quantum logical entropy* $h(\rho)$ ([9]; [**?**]) of a density matrix ρ is defined as above for the classical logical entropy in terms of a classical density matrix:

$$h(\rho) = 1 - \operatorname{tr}\left[\rho^2\right].$$

The formula $1 - \operatorname{tr}\left[\rho^2\right]$ is not new in quantum information theory. Indeed, $\operatorname{tr}\left[\rho^2\right]$ is usually called the *purity* of the density matrix since a state ρ is *pure* if and only if $\operatorname{tr}\left[\rho^2\right] = 1$ so $h(\rho) = 0$, and otherwise $\operatorname{tr}\left[\rho^2\right] < 1$ so $h(\rho) > 0$ and the state is said to be *mixed*. Hence the complement $1 - \operatorname{tr}\left[\rho^2\right]$ has been called the "mixedness" [20, p. 5] or "impurity" of the state ρ.[3] What is new is not the

[3] It is also called by the misnomer "linear entropy" [4] even though it is obviously a quadratic formula–so we will not continue that usage. The quantum logical entropy is also the quadratic special case of the Tsallis-Havrda-Charvat entropy ([17], [33]).

formula but the whole backstory of partition logic outlined above which gives the logical notion of entropy arising out of partition logic as the normalized counting measure on ditsets–just as logical probability arises out of Boolean subset logic as the normalized counting measure on subsets. The basic idea of information is differences, distinguishability,[4] and distinctions ([6], [8]), so the logical notion of entropy is just the (normalized) counting measure of distinctions of a partition and its various natural generalizations.

Quantum logical entropy $h(\rho)$ is to be compared to the *von Neumann entropy* $S(\rho) = -\operatorname{tr}[\rho \log \rho]$, which is the quantum version of the well-known *Shannon entropy* $H(p) = -\sum_{i=1}^{m} p_i \log(p_i)$ ([25, p.510]; [29]). Along with Boaz and Cohen ([**?**], [**?**]), I would argue that the quantum logical entropy is a more natural and informative entropy concept for quantum mechanics than von Neumann entropy.

> 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 ([6]). [**?**, p. 1]

Our approach here is to compare what each notion of entropy tells about quantum measurement and our results seem to confirm their judgment.

Let $F : V \to V$ be a self-adjoint operator with the real eigenvalues $\phi_1, ..., \phi_m$ and let $U = \{u_1, ..., u_n\}$ be an orthonormal basis of eigenvectors of F. Then there is a set partition $\pi = \{B_i\}_{i=1,...,m}$ on U so that B_i is a basis for the eigenspace of the eigenvalue ϕ_i and $|B_i|$ is the "multiplicity" (dimension of the eigenspace) of the eigenvalue ϕ_i for $i = 1, ..., m$. Note that the real-valued function $f : U \to \mathbb{R}$ that takes each eigenvector in $u_j \in B_i \subseteq U$ to its eigenvalue ϕ_i so that $f^{-1}(\phi_i) = B_i$ contains all the information in the self-adjoint operator $F : V \to V$ since F can be reconstructed by defining it on the basis U as $Fu_j = f(u_j) u_j$.

In a measurement using a self-adjoint operator F, the operator does not provide the point probabilities; they come from the pure (normalized) state ψ being measured. Let $|\psi\rangle = \sum_{j=1}^{n} \langle u_j|\psi\rangle |u_j\rangle = \sum_{j=1}^{n} \alpha_j |u_j\rangle$ be the resolution of $|\psi\rangle$ in terms of the orthonormal basis $U = \{u_1, ..., u_n\}$ of eigenvectors for F. Then $p_j = \alpha_j \alpha_j^*$ (α_j^* is the complex conjugate of α_j) for $j = 1, ..., n$ are the point probabilities on U and the pure state density matrix $\rho(\psi) = |\psi\rangle\langle\psi|$ (expressed in the U-basis) has the entries: $\rho(\psi)_{jk} = \alpha_j \alpha_k^*$ so the diagonal entries $\rho(\psi)_{jj} = \alpha_j \alpha_j^* = p_j$ are the point probabilities. Let $S_0 = \{u_j \in U : p_j > 0\}$ be the *support* consisting of the points of positive probability.

One of our themes is the extent to which quantum calculations can be reformulated or mirrored using classical set-based notions. The classical density matrix $\rho(S_0)$ defined above (using point probabilities) with $\rho(S_0)_{jk} = \frac{1}{p(S_0)}\sqrt{p_j p_k} = \sqrt{p_j p_k}$ (since $p(S_0) = \sum_{p_j>0} p_j = 1$) only has real entries while $\rho(\psi) = \alpha_j \alpha_k^*$ has complex entries, but they both have the same logical entropy of 0:

$$h(\rho(S_0)) = 1 - \operatorname{tr}\left[\rho(S_0)^2\right] = 0 = 1 - \operatorname{tr}\left[\rho(\psi)^2\right] = h(\rho(\psi))$$

since in both cases the density matrices are idempotent $\rho^2 = \rho$ and $\operatorname{tr}[\rho] = 1$.

Measurement turns pure states into mixed states. Let $\hat{P}_i$ for $i = 1, ..., m$ be the projection operator to the eigenspace for ϕ_i so the $n \times n$ projection matrix in the basis U is the diagonal matrix P_i where $(P_i)_{jj} = 1$ if $u_j \in B_i$ and otherwise 0, i.e., $(P_i)_{jj} = \chi_{B_i}(u_j)$, and let $\psi_i = \hat{P}_i(\psi)$. The probability that a (projective) measurement will have the outcome ϕ_i and thus project ψ to ψ_i is:

$$p(B_i) = \sum_{u_j \in B_i} p_j = \sum_{u_j \in B_j} \alpha_j \alpha_j^* = \|\psi_i\|^2 = \operatorname{tr}[P_i \rho(\psi)].$$

[4] As Charles Bennett, one of the founders of quantum information theory, put it: "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." [2, p. 155]

Normalizing $|\psi_i\rangle$ by its norm $\|\psi_i\| = \sqrt{p(B_i)}$ allows us to construct its density matrix as (where the conjugate transpose $P_i^\dagger = P_i$):

$$\rho(\psi_i) = \frac{1}{\sqrt{p(B_i)}} |\psi_i\rangle \langle\psi_i| \frac{1}{\sqrt{p(B_i)}} = \frac{1}{p(B_i)} P_i |\psi\rangle \langle\psi| P_i^\dagger = \frac{1}{p(B_i)} P_i \rho(\psi) P_i.$$

The density matrix for the mixed state resulting from the projective measurement [25, p. 515] is the probability weighted sum of projected density matrices $\rho(\psi_i)$:

$$\hat{\rho}(\psi) = \sum_{i=1}^m p(B_i) \rho(\psi_i) = \sum_{i=1}^m P_i \rho(\psi) P_i.$$

Since the logical entropy of the pure state $\rho(\psi)$ was 0, the increase in logical entropy resulting from the measurement is the logical entropy of $\hat{\rho}(\psi)$:

$$h(\hat{\rho}(\psi)) = 1 - \operatorname{tr}\left[\hat{\rho}(\psi)^2\right].$$

In terms of sets, we have the set partition $\pi = \{B_i\}_{i=1,\ldots,m}$ on the set U with the point probabilities $p = \{p_j\}_{j=1,\ldots,n}$ which has the classical logical entropy $h_p(\pi) = 1 - \sum_{i=1}^m p(B_i)^2$. As noted in the previous section, this can also be obtained as the logical entropy of the classical density matrix $\rho(\pi)$ of that partition:

$$h_p(\pi) = 1 - \sum_{i=1}^m p(B_i)^2 = 1 - \operatorname{tr}\left[\rho(\pi)^2\right] = h(\rho(\pi)).$$

Our first result is that the quantum logical entropy $h(\hat{\rho}(\psi))$ resulting from the projective measurement can be computed classically as $h_p(\pi) = h(\rho(\pi))$.

Proposition 1 $h_p(\pi) = h(\hat{\rho}(\psi))$.

Proof: Pre- and post-multiplying $\rho(\psi)$ by the diagonal projection matrices P_i with $(P_i)_{jj} = \chi_{B_i}(u_j)$ gives:

$$(P_i \rho(\psi) P_i)_{jk} = \begin{cases} \alpha_j \alpha_k^* \text{ if } (u_j, u_k) \in B_i \times B_i \\ 0 \text{ if not.} \end{cases}$$

and since $\operatorname{indit}(\pi) = \cup_{i=1}^m B_i \times B_i$,

$$(\hat{\rho}(\psi))_{jk} = \left(\sum_{i=1}^m P_i \rho(\psi) P_i\right)_{jk} = \begin{cases} \alpha_j \alpha_k^* \text{ if } (u_j, u_k) \in \operatorname{indit}(\pi) \\ 0 \text{ if } (u_j, u_k) \in \operatorname{dit}(\pi) \end{cases}.$$

Hence to compute the quantum logical entropy $h(\hat{\rho}(\psi))$:

$$\left(\hat{\rho}(\psi)^2\right)_{jj} = \sum_k \hat{\rho}(\psi)_{jk} \hat{\rho}(\psi)_{kj} = \sum_{k,(u_j,u_k)\in\operatorname{indit}(\pi)} \|\alpha_j\|^2 \|\alpha_k\|^2$$

so the trace is:

$$\begin{aligned} \operatorname{tr}\left[\hat{\rho}(\psi)^2\right] &= \sum_j \left(\hat{\rho}(\psi)^2\right)_{jj} = \sum_j \sum_{k,(u_j,u_k)\in\operatorname{indit}(\pi)} \|\alpha_j\|^2 \|\alpha_k\|^2 \\ &= \sum_{(u_j,u_k)\in\operatorname{indit}(\pi)} p_j p_k \end{aligned}$$

and the quantum logical entropy is:

$$\begin{aligned} h(\hat{\rho}(\psi)) &= 1 - \operatorname{tr}\left[\hat{\rho}(\psi)^2\right] = 1 - \sum_{(u_j,u_k)\in\operatorname{indit}(\pi)} p_j p_k \\ &= \sum_{(u_j,u_k)\in\operatorname{dit}(\pi)} p_j p_k = h_p(\pi). \ \square \end{aligned}$$

Corollary 1 *The quantum logical entropy $h(\hat{\rho}(\psi))$ of the projective measurement result $\hat{\rho}(\psi)$ is the probability of getting different results ($\phi_i \neq \phi_{i'}$) in two independent measurements with the same observable and the same pure state.* $\square$

The quantum logical entropy $h(\rho(\psi))$ and the von Neumann entropy $S(\rho(\psi))$ for both zero for a pure state $\rho(\psi)^2 = \rho(\psi)$, but the measurement entropies $h(\hat{\rho}(\psi))$ and $S(\hat{\rho}(\psi))$ are different although both increase under projective measurement [**?**]. The question is:

Which entropy concept gives insight into what happens in a quantum measurement?

The last corollary shows firstly that the quantum logical entropy $h(\hat{\rho}(\psi))$ associated with the projective measurement has a very simple interpretation–whereas there seems to be no such simple interpretation for $S(\hat{\rho}(\psi))$.

Our next and main result shows that the quantum logical entropy $h(\hat{\rho}(\psi))$ (which is also the increase in quantum logical entropy due to the projective measurement since $h(\rho(\psi)) = 0$) is precisely related to what happens to the off-diagonal entries in the change in the density matrices $\rho(\psi) \to \hat{\rho}(\psi)$ due to the measurement–whereas no such specific result seems to hold for the von Neumann entropy other than it also increases: $S(\rho(\psi)) < S(\hat{\rho}(\psi))$ [25, p. 515] [when $\rho(\psi) \neq \hat{\rho}(\psi)$].

Theorem 1 (Main) *The quantum logical entropy $h(\hat{\rho}(\psi))$ is the sum of the absolute squares of the nonzero off-diagonal entries in the pure state density matrix $\rho(\psi)$ that are zeroed in the transition $\rho(\psi) \longrightarrow \hat{\rho}(\psi)$ due to a projective measurement.*

Proof: On the support $S_0 = \{u_j \in U : p_j > 0\}$, the nonzero elements of $\rho(\psi)$ are $(\rho(\psi))_{jk} = \alpha_j\alpha_k^*$ for $(u_j, u_k) \in S_0 \times S_0$ with the absolute squares $\alpha_j\alpha_k^*\alpha_j^*\alpha_k = p_j p_k$. The post-measurement density matrix has the nonzero entries $(\hat{\rho}(\psi))_{jk} = \alpha_j\alpha_k^*$ if $(u_j, u_k) \in \operatorname{indit}(\pi) \cap S_0 \times S_0$ where $\operatorname{indit}(\pi) = \cup_{i=1}^m B_i \times B_i$. Hence the nonzero entries of $\rho(\psi)$ that got zeroed in $\hat{\rho}(\psi)$ are precisely the entries $(\rho(\psi))_{jk} = \alpha_j\alpha_k^*$ for $(u_j, u_k) \in \operatorname{dit}(\pi) \cap S_0 \times S_0$. Since the entries for $(u_j, u_k) \notin S_0 \times S_0$ were already zero, we have that the sum of the absolute squares of entries zeroed by the measurement is:

$$\sum_{(u_j,u_k)\in \operatorname{dit}(\pi)\cap S_0\times S_0} p_j p_k = \sum_{(u_j,u_k)\in \operatorname{dit}(\pi)} p_j p_k = h_p(\pi) = h(\hat{\rho}(\psi)). \ \square$$

7 'Measurement' with classical density matrices

Now we can describe a similar process of 'measurement' in the classical case using density matrices for subset $S \subseteq U$ and partitions π on U.[5] In the classical version $\rho(S)$ of $\rho(\psi)$, the nonzero entries $\sqrt{p_j p_k}$ are the indistinction "amplitudes" whose square is the probability $p_j p_k$ of drawing (u_j, u_k) in two independent draws from $S_0 = \{u_j \in U : p_j > 0\}$. In the equiprobable case of $p_j = \frac{1}{n}$, $S_0 = U$ and the only partition with the classical logical entropy of zero is the indiscrete partition (or "blob") $\mathbf{0} = \{U\}$ which has no distinctions. But with point probabilities $\{p_j\}_j$ and $S_0 \neq U$, the "outcomes" in $S_0^c = U - S_0$ have zero probability. The partition $\sigma = \{S_0, S_0^c\}$ has a nonempty dit set $\operatorname{dit}(\sigma) = (S_0 \times S_0^c) \cup (S_0^c \times S_0)$ but clearly:

$$h(\sigma) = \sum_{(u_j,u_k)\in \operatorname{dit}(\sigma)} p_j p_k = 0 = h(\rho(S_0))$$

since σ has no distinctions with positive probability–so σ is effectively like the blob $\mathbf{0}$ as a classical "pure state."

The classical version of the measurement can be stated in classical finite probability theory where the outcome set or sample space $U = \{u_1, ..., u_n\}$ has point probabilities $p' = \{p'_j > 0\}_{j=1,...,n}$ for the

[5] This can be seen as the measurement process in a pedagogical model of quantum mechanics over sets, QM/Sets, but it would take us too far afield to go into that whole model here. Hence we will describe the 'measurement' process in just straight classical terms. For QM/Sets with equiprobable points, see [11].

outcomes. A real random variable $f : U \to \mathbb{R}$ has an image $f(U) = \{\phi_1, ..., \phi_m\}$ in the codomain $\mathbb{R}$ and induces a partition $\pi = \left\{f^{-1}(\phi_i) = B_i\right\}_{i=1,...,m}$ on the domain U. Given any (nonempty) event or 'state' $S \subseteq U$, the point probabilities for $u_j \in S$ can be conditionalized to $p_j = \frac{p'_j}{\sum_{u_k \in S} p'_k} = \frac{p'_j}{p'(S)}$ and set to 0 outside of S (so henceforth $S = S_0$, the support for the p_j)–which gives the initial classical density matrix $\rho(S)$. The experiment to 'measure' f given the 'state' S returns a value ϕ_i with the probability $\Pr(\phi_i|S) = \sum_{u_j \in S \cap B_i} p_j = \frac{p'(S \cap B_i)}{p'(S)}$ and we could even have a "projection postulate" that the state S is projected to the state $S \cap B_i$ with that probability. The join $\pi \vee \sigma$ of the 'observable' partition $\pi = \left\{f^{-1}(\phi_i) = B_i\right\}_{i=1,...,m}$ with $\sigma = \{S, S^c\}$ has as blocks the nonempty intersections $S \cap B_i$ and $S^c \cap B_i$. The post-measurement 'mixed state' density matrix is $\rho(\pi \vee \sigma)$. Since the S^c points have probability 0, we can focus on S. The join-measurement operation has the effect of "chopping up" the mini-blob S into the partition $\left\{S \cap f^{-1}(\phi_i)\right\} = \{S \cap B_i\}_{i=1,...,m}$ of S. The join operation creates distinctions since $\operatorname{dit}(\pi \vee \sigma) = \operatorname{dit}(\pi) \cup \operatorname{dit}(\sigma)$, but taking the point probabilities p_j into account, all of the distinctions in $\operatorname{dit}(\sigma)$ has probability 0 (since one of the points in $(u_j, u_k) \in \operatorname{dit}(\sigma)$ has to be in S^c so $p_j p_k = 0$). Hence $\rho(\pi \vee \sigma) = \rho(\pi)$ and the only distinctions of $\pi \vee \sigma$ with nonzero probabilities are in $\operatorname{dit}(\pi) \cap S \times S$. The sum of those distinction probabilities is the logical entropy:

$$\sum_{(u_j,u_k) \in \operatorname{dit}(\pi) \cap S \times S} p_j p_k = \sum_{(u_j,u_k) \in \operatorname{dit}(\pi)} p_j p_k = h_p(\pi).$$

Corollary 2 *The classical logical entropy $h_p(\pi) = h(\rho(\pi))$ is the sum of the squares of the nonzero off-diagonal entries in $\rho(S)$ that are zeroed in the transition $\rho(S) \to \rho(\pi)$.*

Proof: The proof carries over substituting $\rho(S)$ for $\rho(\psi)$, $\rho(\pi)$ for $\hat{\rho}(\psi)$, and $\sqrt{p_j p_k}$ for $\alpha_j \alpha_k^*$. □

Note that by taking U as an orthonormal basis for the observable in the above quantum case and the $p_j = \alpha_j^* \alpha_j$ as supplied by expressing ψ in that base, we have: $h(\hat{\rho}(\psi)) = h_p(\pi) = h(\rho(\pi))$. Thus we have a pedagogical classical model, using sets with point probabilities, for the projective quantum measurement.

8 Conclusions

This main theorem about quantum logical entropy as well as the connection to classical logical entropy, i.e., information as distinctions, together with the backstory of partition logic allows some elucidation of a projective quantum measurement.

We now see how this quantum measurement transition from $\rho(\psi)$ to $\hat{\rho}(\psi)$ can be precisely modelled or mirrored in completely classical terms as the transition from $\rho(S)$ to $\rho(\pi) = \rho(\pi \vee \sigma)$. The partition $\pi = \left\{f^{-1}(\phi_i)\right\}_{i=1,...,m}$ supplied by the real random variable $f : U \to \mathbb{R}$ 'chops up' the given state S into the distinct blocks $\left\{f^{-1}(\phi_i) \cap S\right\}$ for $i = 1, ..., m$. The previous indistinctions of $\sigma = \{S, S^c\}$ of the form $(u_j, u_k) \in S \times S$ that had distinct 'eigenvalues' ϕ_i and $\phi_{i'}$ were turned into distinctions in the partition $\pi \vee \sigma$, and they correspond precisely to the nonzero off-diagonal entries in $\rho(S)$ that got zeroed in the density matrix transition $\rho(S)$ to $\rho(\pi \vee \sigma) = \rho(\pi)$. In terms of logical entropy, the measurement took the logical entropy $h(\rho(S)) = 0$ of the pure state density matrix $\rho(S)$ to the sum of the squares of all those nonzero off-diagonal entries that got zeroed in the transition which is the logical entropy $h_p(\pi)$ of the mixed state density matrix $\rho(\pi \vee \sigma) = \rho(\pi)$.

In the quantum case, the nonzero off-diagonal entries $\alpha_j \alpha_k^*$ in the pure state density matrix $\rho(\psi)$ are called quantum "coherences" ([5, p. 303], [1, p.177]) because they give the amplitude of the eigenstates $|u_j\rangle$ and $|u_k\rangle$ "cohering" together in the coherent superposition state vector $|\psi\rangle = \sum_j \langle u_j|\psi\rangle |u_j\rangle = \sum_j \alpha_j |u_j\rangle$. They are modelled by the nonzero off-diagonal entries $\sqrt{p_j p_k}$ whose squares are the two-draw probabilities for the indistinctions $(u_j, u_k) \in S \times S$. Coherences are modelled by indistinctions. The off-diagonal elements of $\rho(\psi)$ that are zeroed by the measurement to yield $\hat{\rho}(\psi)$

are the coherences (like quantum indistinctions) that are turned into 'decoherences' (like quantum distinctions). Thus the off-diagonal entries zeroed in the transition from the pure state $\rho(\psi)$ to the mixed state $\hat{\rho}(\psi)$ density matrices are modelled by the disappearing off-diagonal entries in $\rho(S)$ corresponding to the distinctions created in the partitioning of S into $\left\{f^{-1}(\phi_i) \cap S\right\}_{i=1,\ldots,m}$ in the classical model of measurement as the join $\pi \vee \sigma$ of σ with π. And the sum of the two-draw probabilities of those new distinctions created in the measurement is, in each case, the logical entropy.

Table 3	Quantum proj. measurement	Classical model
Basis $U = \{u_1, ..., u_n\}$	$U =$ ON eigen-basis for F	Set $U = \{u_1, ..., u_n\}$
Observable	Self-adjoint operator F	Real r.v. $f : U \to \mathbb{R}$
Eigenvalues	Eigenvalues of F: $\{\phi_1, ..., \phi_m\}$	$f(U) = \{\phi_i\}_{i=1,\ldots,m}$
Partition	Partition π of ON basis U by ϕ_i	$\pi = \left\{f^{-1}(\phi_i)\right\}_i = \{B_i\}_i$
Given state to measure	ψ	$S \subseteq U$
Point probabilities	$p_j = \alpha_j^* \alpha_j$ for $\alpha_j = \langle u_j \vert \psi \rangle$	Same p_j for $u_j \in U$
Support S_0	$S_0 = \{u_j : p_j > 0\}$	Set $S = S_0$
Pure state density matrix	$\rho(\psi)$	$\rho(S)$
Positive prob. indistinction	$(u_j, u_k) \in S_0 \times S_0$	$(u_j, u_k) \in S \times S$
Post-Meas. mixed state	$\hat{\rho}(\psi)$	$\rho(\pi \vee \{S, S^c\}) = \rho(\pi)$
Sum sqs. off-diag. zeroed	$h(\hat{\rho}(\psi))$	$h_p(\pi) = h(\rho(\pi))$

Table 3: Summary of classical modelling of projective quantum measurement.

The positive probability classical or quantum indistinctions that got turned into distinctions by the measurement correspond to those nonzero off-diagonal entries in the pre-measurement pure state density matrices that got zeroed in the post-measurement mixed state density matrices.

What was here called the "classical model" of the projective measurement is that type of measurement in the pedagogical model QM/Sets with point probabilities–where the model QM/Sets for equiprobable points is developed in [11]. That "classical measurement" is just the usual classical finite probability sampling of a real r.v. $f : U \to \mathbb{R}$ given an event $S \subseteq U$ which returns the value ϕ_i with the probability $\Pr(\phi_i|S) = p\left(f^{-1}(\phi_i) \cap S\right)/p(S)$ which for the choices of $U, f, S,$ and p is the same as the quantum probability $\Pr(\phi_i|\psi) = \sum\{p_j : \phi_i \text{ eigenvalue of } u_j\}$. And then we even get the same logical entropies: $h_p(\pi) = h(\rho(\pi)) = h(\hat{\rho}(\psi))$.

Needless to say, the von Neumann entropy gives no such simple, detailed, and precise description of what happens in the quantum projective measurement that turns a pure state density matrix $\rho(\psi)$ to the mixed state density matrix $\hat{\rho}(\psi)$.

Measurement creates distinctions, i.e., turns indistinctions (or coherences) into distinctions (or decoherences), to make an indefinite state (superposition of definite eigenstates) more definite in the measured observable. The off-diagonal coherences in $\rho(\psi)$ are essentially the amplitudes for quantum indistinctions so the ones that are zeroed are turned into distinctions (i.e., 'decoherences') and the sum of squared amplitudes, i.e., the distinction probabilities, is the post-measurement classical and quantum logical entropy.

References

[1] Auletta, Gennaro, Mauro Fortunato, and Giorgio Parisi. 2009. *Quantum Mechanics*. Cambridge UK: Cambridge University Press.

[2] Bennett, Charles H. 2003. Quantum Information: Qubits and Quantum Error Correction. *International Journal of Theoretical Physics* 42 (2 February): 153–76.

[3] Boole, George 1854. *An Investigation of the Laws of Thought on which are founded the Mathematical Theories of Logic and Probabilities*. Cambridge: Macmillan and Co.

[4] Buscemi, Fabrizio, Paolo Bordone, and Andrea Bertoni. 2007. Linear Entropy as an Entanglement Measure in Two-Fermion Systems. *ArXiv.org*. March 2. http://arxiv.org/abs/quant-ph/0611223v2.

[5] Cohen-Tannoudji, Claude, Bernard Diu and Franck Laloë 2005. *Quantum Mechanics Vol. 1*. New York: John Wiley & Sons.

[6] Ellerman, David. 2009. Counting Distinctions: On the Conceptual Foundations of Shannon's Information Theory. *Synthese* 168 (1 May): 119–49.

[7] Ellerman, David 2010. The Logic of Partitions: Introduction to the Dual of the Logic of Subsets. *Review of Symbolic Logic*. 3 (2 June): 287-350.

[8] Ellerman, David. 2013. An Introduction to Logical Entropy and Its Relation to Shannon Entropy. *International Journal of Semantic Computing* 7 (2): 121–45.

[9] Ellerman, David. 2013. The Objective Indefiniteness Interpretation of Quantum Mechanics. *ArXiv.org*. January. http://arxiv.org/abs/1210.7659v2.

[10] Ellerman, David 2014. An Introduction of Partition Logic. *Logic Journal of the IGPL*. 22, no. 1: 94–125.

[11] Ellerman, David. 2016. Quantum Mechanics over Sets: A Pedagogical Model with Non-Commutative Finite Probability Theory as Its Quantum Probability Calculus. *Synthese Online First* (August): 34 pages. doi: 10.1007/s11229-016-1175-0.

[12] Friedman, William F. 1922. *The Index of Coincidence and Its Applications in Cryptography*. Geneva IL: Riverbank Laboratories.

[13] Gini, Corrado 1912. *Variabilità e mutabilità*. Bologna: Tipografia di Paolo Cuppini.

[14] Gini, Corrado 1955. Variabilità e mutabilità. In *Memorie di metodologica statistica*. E. Pizetti and T. Salvemini eds., Rome: Libreria Eredi Virgilio Veschi.

[15] Good, I. J. 1979. A.M. Turing's statistical work in World War II. *Biometrika*. 66 (2): 393-6.

[16] Good, I. J. 1982. Comment (on Patil and Taillie: Diversity as a Concept and its Measurement). *Journal of the American Statistical Association*. 77 (379): 561-3.

[17] Havrda, Jan, and Frantisek Charvat. 1967. Quantification Methods of Classification Processes: Concept of Structural α-Entropy. *Kybernetika* (Prague) 3: 30–35.

[18] Herfindahl, Orris C. 1950. *Concentration in the U.S. Steel Industry*. Unpublished doctoral dissertation, Columbia University.

[19] Hirschman, Albert O. 1964. The Paternity of an Index. *American Economic Review*. 54 (5): 761-2.

[20] Jaeger, Gregg. 2007. *Quantum Information: An Overview*. New York: Springer Science+Business Media.

[21] Kullback, Solomon 1976. *Statistical Methods in Cryptanalysis*. Walnut Creek CA: Aegean Park Press.

[22] Laplace, Pierre-Simon. 1995 (1825). *Philosophical Essay on Probabilities*. Translated by A. I. Dale. New York: Springer Verlag.

[23] Lawvere, F. William and Robert Rosebrugh 2003. *Sets for Mathematics*. Cambridge: Cambridge University Press.

[24] MacArthur, Robert H. 1965. Patterns of Species Diversity. *Biol. Rev.* 40: 510–33.

[25] Nielsen, M., and I. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

[26] Patil, G. P. and C. Taillie 1982. Diversity as a Concept and its Measurement. *Journal of the American Statistical Association*. 77 (379): 548-61.

[27] Rao, C. R. 1982. Diversity and Dissimilarity Coefficients: A Unified Approach. *Theoretical Population Biology*. 21: 24-43.

[28] Ricotta, Carlo and Laszlo Szeidl 2006. Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index. *Theoretical Population Biology*. 70: 237-43.

[29] Shannon, Claude E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*. 27: 379-423; 623-56.

[30] Simpson, Edward Hugh 1949. Measurement of Diversity. *Nature*. 163: 688.

[31] Tamir, Boaz, and Eliahu Cohen. 2014. Logical Entropy for Quantum States. *ArXiv.org*. December. http://de.arxiv.org/abs/1412.0616v2.

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

[33] Tsallis, Constantino 1988. Possible Generalization for Boltzmann-Gibbs Statistics. *J. Stat. Physics* 52: 479–87.