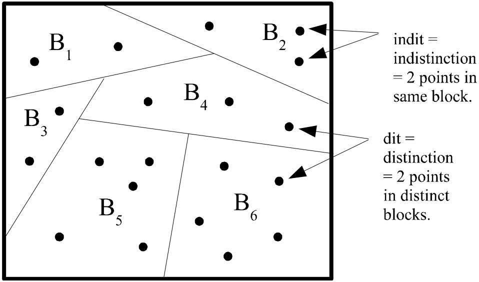
# New Foundations for Information Theory: $\frac{\text{Probability Theory}}{\text{Subset Logic}} = \frac{\text{Information Theory}}{\text{Partition Logic}}$

David Ellerman
University of California-Riverside
University of Ljubljana, Slovenia

- “The dual notion (obtained by reversing the arrows) of ‘part’ is the notion of partition.” [Lawvere]; mono $S \to X$ dualizes to epi $X \to Y$.
- Duality of Elements and Distinctions ("Its" & "Dits")




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

# Duality of Subsets and Partitions: II

| Table 1 | Subset Logic | Partition Logic |
|---|---|---|
| Logic of... | Subsets $S \subseteq U$ | Partitions $\pi$ on U |
| 'Elements' (its or dits) | *Elements* u of a subset S | *Distinctions* (u,u') of a partition $\pi$ |
| All 'elements' | Universe set U (all elements) | Discrete partition **1** (all dits) |
| No 'elements' | Empty set $\varnothing$ (no elements) | Indiscrete partition **0** (no dits) |
| Partial order on... | $S \subseteq T$ = Inclusion of elements | Refinement $\sigma \preceq \pi$ of partitions = inclusion of ditsets: $\operatorname{dit}(\sigma) \subseteq \operatorname{dit}(\pi)$ |
| Formula variables | Subsets of U | Partitions on U |
| Logical operations $\cup, \cap, \Rightarrow, \ldots$ | Operations on subsets | Operations on partitions |
| Propositional interp. of $\Phi(\pi,\sigma,\ldots)$ | Subset $\Phi(\pi,\sigma,\ldots)$ contains an element u. | Partition $\Phi(\pi,\sigma,\ldots)$ makes a distinction (u,u'). |
| Valid formula $\Phi(\pi,\sigma,\ldots)$ | $\Phi(\pi,\sigma,\ldots) = U$ for any subsets $\pi,\sigma,\ldots$ of any U ($\lvert U\rvert \geq 1$), i.e., contains all elements u. | $\Phi(\pi,\sigma,\ldots) = \mathbf{1}$ for any partitions $\pi,\sigma,\ldots$ on any U ($\lvert U\rvert \geq 2$), i.e., makes all distinctions (u,u'). |

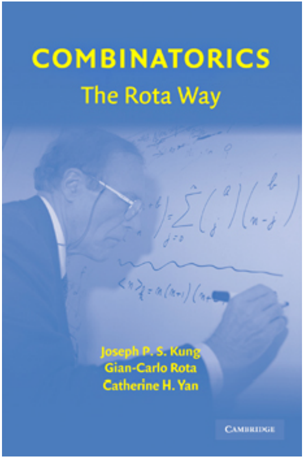
Dual Logics: Boolean subset logic of subsets and partition logic

# Duality of Subsets and Partitions: III

- Published papers on partition logic:
  - The Logic of Partitions: Introduction to the Dual of the Logic of Subsets. *Review of Symbolic Logic*, 3(2 June), 287–350, 2010.
  - An introduction of partition logic. *Logic Journal of the IGPL*, 22(1), 94–125, 2014.
- Birkhoff & von Neumann created quantum logic by linearizing logic of subsets to logic of subspaces of a vector space.
- Hence the dual form of quantum logic created by linearing logic of partitions to logic of direct-sum decompositions of a vector space:
  - The Quantum Logic of Direct-Sum-Decompositions: The Dual to the Quantum Logic of Subspaces. *Logic Journal of the IGPL. Online limbo*.

# Duality of Subsets and Partitions: IV

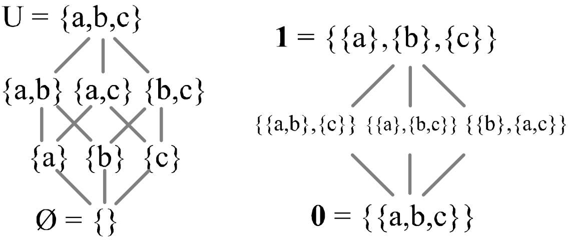
- All papers on: *www.ellerman.org*.




"Lattice of partitions plays for information the role that the Boolean algebra of subsets plays for size or probability."

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

# Duality of Subsets and Partitions: V




Boolean lattice of subsets of U

Partition lattice of partitions on U

- Rota: "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:

# Duality of Subsets and Partitions: VI

*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? (Rota's Fubini Lecture)*

- Elements : Subsets :: Dits : Partitions, so

#elements ("size of subset") $\approx$ #dits ("size of partition").

# The logical theory of information: I

- New foundations of information theory starts with sets, not probabilities.

  *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. [Kolmogorov, A. N. 1983]*

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

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

# The logical theory of information: II

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

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

- The *logical entropy* of $\pi$ is the product measure of its ditset:

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

# The logical theory of information: III

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

# Information is distinctions: I

- The logical theory and, to some extent, the Shannon theory show that information is about:
    - Distinctions,
    - Differences,
    - Distinguishing by classifications, and
    - Symmetry-breaking.

*For in the general we must note, That whatever is capable of a competent Difference, perceptible to any Sense, may be a sufficient Means whereby to express the Cogitations. [John Wilkins 1641]*

# Information is distinctions: II

- James Gleick comments: "That word, differences, must have struck Wilkins readers as an odd choice. . . . Wilkins was reaching for a conception of information in its purest, most general form. . . Here, in this arcane and anonymous treatise of 1641, the essential idea of information theory poked to the surface of human thought, saw its shadow, and disappeared again for [three] hundred years."
- 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.*

# Information is distinctions: III

| Published papers on logical information theory: |
|---|
| Ellerman, David 2009. "Counting Distinctions: On the Conceptual Foundations of Shannon's Information Theory." *Synthese* 168 (1 May): 119-49. |
| Ellerman, David 2013. "An Introduction to Logical Entropy and Its Relation to Shannon Entropy." *International Journal of Semantic Computing* 7 (2): 121—45. |
| Tamir, Boaz, and Eliahu Cohen. 2014. Logical Entropy for Quantum States. *ArXiv.org*. December. |
| Tamir, Boaz, and Eliahu Cohen. 2015. "A Holevo-Type Bound for a Hilbert Schmidt Distance Measure." *Journal of Quantum Information Science* 5: 127-33. |
| Ellerman, David. 2017. "Logical Information Theory: New Foundations for Information Theory." *Logic Journal of the IGPL* 25 (5 Oct.): 806–35. |
| Ellerman, David 2017. "Introduction to Quantum Logical Information Theory." *ArXiv.org*. Aug. |

# History of logical entropy formula: I

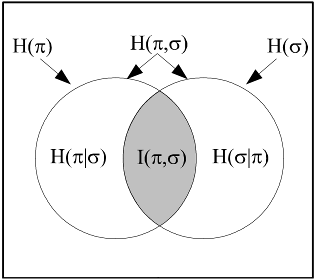
- In 1912, Gini defined $1 - \sum_i p_i^2$ as the index of mutability.
- In 1922, cryptographer William Friedman defined $\sum_i p_i^2$ as index of coincidence.
- Alan Turing worked at Bletchley Park in WWII on crypography and defined $\sum_i p_i^2$ as the repeat rate.
- Turing's assistant, Edward Simpson, published in 1949, $\sum_i p_i^2$ as "index of species concentration" so $1 - \sum_i p_i^2$ is now often called *Gini-Simpson index of diversity* in biodiversity literature.

# History of logical entropy formula: II

- Simpson along with I. J. Good worked at Bletchley during WWII, and, according to Good, “E. H. Simpson and I both obtained the notion [the repeat rate] from Turing.” Simpson (again, according to Good) did not acknowledge Turing “fearing that to acknowledge him would be regarded as a breach of security.”
- For $d_{ij}$ = "distance" between $u_i$ and $u_j$ where $d_{ii} = 0$, C.R. Rao (1982), quadratic entropy: $Q(p) = \sum_{i,j} d_{ij} p_i p_j$.
- "Logical distance" is $d_{ij} = 1 - \delta_{ij}$, and the Rao quadratic entropy with logical distances is the logical entropy $h(p) = \sum_{i \neq j} p_i p_j = 1 - \sum_i p_i^2$.
- Quantum version: $\operatorname{tr}\left[\rho^2\right]$ = *purity* and $1 - \operatorname{tr}\left[\rho^2\right]$ called *mixedness*.

# Compound logical entropies: I

- Shannon entropy:
  - 'looks' like a measure;
  - 'walks' like a measure;
  - 'quacks' like a measure;
  - but is *not* a measure in the sense of measure theory.



# Compound logical entropies: II

- As the eminent information theorist, Lorne Campbell, put it in a 1965 paper:

  *Certain analogies between entropy and measure have been noted by various authors. These analogies provide a convenient mnemonic for the various relations between entropy, conditional entropy, joint entropy, and mutual information. It is interesting to speculate whether these analogies have a deeper foundation. It would seem to be quite significant if entropy did admit an interpretation as the measure of some set.*

- After seeing this paper, Campbell replied: "*on first reading it seems to provide the deeper relationship that I sought 50 years ago.*"

# Compound logical entropies: III

- Hence those who want to represent Shannon entropy as a measure desire:

  *that $H(\alpha)$ and $H(\beta)$ are measures of sets, that $H(\alpha,\beta)$ is the measure of their union, that $I(\alpha,\beta)$ is the measure of their intersection, and that $H(\alpha|\beta)$ is the measure of their difference. The possibility that $I(\alpha,\beta)$ is the entropy of the "intersection" of two partitions is particularly interesting. This "intersection," if it existed, would presumably contain the information common to the partitions $\alpha$ and $\beta$. [Campbell, Lorne 1965. Entropy as a Measure. IEEE Trans. on Information Theory. IT-11 (January): 112-114, p. 113]*

# Compound logical entropies: IV

- Given partitions $\pi = \{B_1, ..., B_I\}$, $\sigma = \{C_1, ..., C_J\}$ on $U$, the *information set* or *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* = product measure $p \times p$ on the ditset:

$$h(\pi, \sigma) = h(\pi \vee \sigma) = p \times p(\text{dit}(\pi) \cup \text{dit}(\sigma)).$$

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

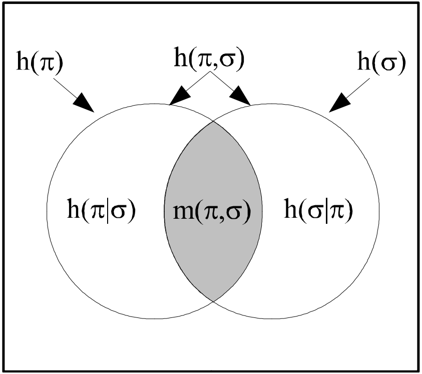
$$h(\pi|\sigma) = p \times p(\text{dit}(\pi) - \text{dit}(\sigma)).$$

- The infoset for the *logical mutual information* $m(\pi, \sigma)$ is the intersection of ditsets, so:

# Compound logical entropies: V

$$m(\pi,\sigma) = p \times p\left(\operatorname{dit}(\pi) \cap \operatorname{dit}(\sigma)\right).$$

- *Information algebra* $\mathcal{I}(\pi,\sigma)$ = Boolean subalgebra of $\wp(U \times U)$ generated by ditsets and their complements.




Logical entropies Venn diagram

# Deriving the Shannon entropies from the logical entropies: I

- All the Shannon entropies can be derived from the logical definitions by a uniform transformation–since they are two different ways to quantify distinctions.
- Canonical case of $n$ equiprobable elements, $p_i = \frac{1}{n}$, the logical entropy is:

$$h(p_i) = \frac{|U \times U - \Delta|}{|U \times U|} = \frac{n^2 - n}{n^2} = 1 - \frac{1}{n} = 1 - p_i.$$

- General case $p = (p_1, ..., p_n)$:

$$h(p) = \sum_i p_i h(p_i) = \sum_i p_i (1 - p_i).$$

# Deriving the Shannon entropies from the logical entropies: II

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

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

- General case $p = (p_1, ..., p_n)$:

$$H(p) = \sum_i p_i H(p_i) = \sum_i p_i \log_2\left(\frac{1}{p_i}\right).$$

# Deriving the Shannon entropies from the logical entropies: III

- *Dit-Bit Transform*: express any logical entropy concept (joint, conditional, or mutual) as average of dit-counts $1 - p_i$, and then substitute the bit-count $\log\left(\frac{1}{p_i}\right)$ to obtain the corresponding formula as defined by Shannon.

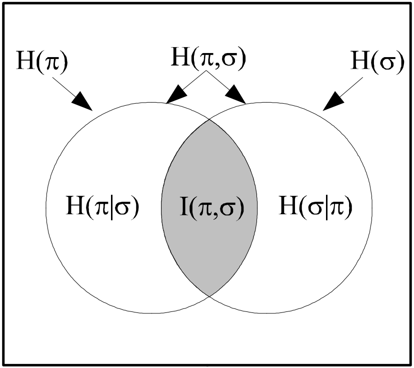
$$(1 - p_i) \rightsquigarrow \log_2\left(\frac{1}{p_i}\right).$$

- The dit-bit transform is linear in the sense of preserving plus and minus, so the Shannon formulas satisfy the same Venn diagram formulas in spite of not being a measure (in the sense of measure theory).

# Deriving the Shannon entropies from the logical entropies: IV

| Table 3 | The Dit-Bit Transform: $1-p(B_i) \rightsquigarrow \log(1/p(B_i))$ |
|---|---|
| Entropy | $h(\pi) = \Sigma_i\, p(B_i)(1-p(B_i))$<br>$H(\pi) = \Sigma_i\, p(B_i)(\log(1/p(B_i)))$ |
| Joint Entropy | $h(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)(1-p(B_i \cap C_j))$<br>$H(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)\log(1/p(B_i \cap C_j))$ |
| Conditional Entropy | $h(\pi\|\sigma) = \Sigma_{i,j} p(B_i \cap C_j)(1-p(B_i \cap C_j)) - \Sigma_j p(C_j)(1-p(C_j))$<br>$H(\pi\|\sigma) = \Sigma_{i,j} p(B_i \cap C_j)\log(1/p(B_i \cap C_j)) - \Sigma_j p(C_j)\log(1/p(C_j))$ |
| Mutual Information | $m(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)[(1-p(B_i)) + (1-p(C_j)) - (1-p(B_i \cap C_j))]$<br>$I(\pi,\sigma) = \Sigma_{i,j} p(B_i \cap C_j)[(\log(1/p(B_i)) + (\log(1/p(C_j)) - (\log(1/p(B_i \cap C_j))]$ |

# Deriving the Shannon entropies from the logical entropies: V




Venn diagram 'mnemonic' for Shannon entropies

# Logical entropy via density matrices: I

- 'Classically,' the *density matrix* representing the event $S$ is the $n \times n$ symmetric real matrix:

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

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

$$\rho(\pi) = \sum_i p(B_i)\,\rho(B_i).$$

# Logical entropy via density matrices: II

- Then $\rho(\pi)$ represents the *mixed* state, experiment, or lottery where the event $B_i$ occurs with probability $p(B_i)$.
- Logical entropy: $h(\pi) = 1 - \sum_i p(B_i)^2 = 1 - \operatorname{tr}\left[\rho(\pi)^2\right]$.

## Example

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

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

# Logical entropy via density matrices: III

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

## Example (continued)

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

# Logical entropy via density matrices: IV

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

$$\rho\left(\mathbf{0}_U\right)=\begin{bmatrix}1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6\end{bmatrix}\rightsquigarrow$$

$$\begin{bmatrix}1/6 & 1/6 & 1/6 & 0 & 0 & 0\\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0\\ 1/6 & 1/6 & 1/6 & 0 & 0 & 0\\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6\\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6\\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6\end{bmatrix}=\rho\left(\pi\right)$$

# Logical entropy via density matrices: V

## Theorem (Basic)

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

## Proof.

Since for any density matrix $\rho$, $\mathrm{tr}\left[\rho^2\right] = \sum_{i,j} \left|\rho_{ij}\right|^2$,
$h\left(\rho\left(\pi\right)\right) - h\left(\rho\left(\mathbf{0}_U\right)\right) = \left(1 - \mathrm{tr}\left[\rho\left(\pi\right)^2\right]\right) - \left(1 - \mathrm{tr}\left[\rho\left(\mathbf{0}_U\right)^2\right]\right) =$
$\mathrm{tr}\left[\rho\left(\mathbf{0}_U\right)^2\right] - \mathrm{tr}\left[\rho\left(\pi\right)^2\right] = \sum_{i,j} \left|\rho_{ij}\left(\mathbf{0}_U\right)\right|^2 - \sum_{i,j} \left|\rho_{ij}\left(\pi\right)\right|^2$. $\square$

# Logical entropy via density matrices: VI

## Example (continued)

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

- In quantum mechanics, projective measurement = Lüders mixture operation.
- Measurement means making distinctions by classifying according to eigenvalues of an observable.

# Logical entropy via density matrices: VII

- In density matrix, making distinctions means zeroing or 'decohering' off-diagonal coherence terms.
- Measure of distinctions created by measurement = sum of absolute squares of decohered terms = logical quantum entropy.

## Example (quantum)

Consider a system with two observable eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ (like spin) where the given normalized pure state is $|\psi\rangle = \alpha_{\uparrow}|\uparrow\rangle + \alpha_{\downarrow}|\downarrow\rangle$ so the pure state density matrix is $\rho(\psi) = \begin{bmatrix} p_{\uparrow} & \alpha_{\uparrow}\alpha_{\downarrow}^{*} \\ \alpha_{\downarrow}\alpha_{\uparrow}^{*} & p_{\downarrow} \end{bmatrix}$ where $p_{\uparrow} = \alpha_{\uparrow}\alpha_{\uparrow}^{*}$, $p_{\downarrow} = \alpha_{\downarrow}\alpha_{\downarrow}^{*}$, and $h(\rho(\psi)) = 0$. The measurement in that observable produces the Lüders mixture $\hat{\rho}(\psi) = P_{\uparrow}\rho(\psi)P_{\uparrow} + P_{\downarrow}\rho(\psi)P_{\downarrow} = \begin{bmatrix} p_{\uparrow} & 0 \\ 0 & p_{\downarrow} \end{bmatrix}$ where $P_{\uparrow}$ and $P_{\downarrow}$ are the projection matrices.

## Example (continued)

The post-measurement logical entropy is $h(\hat{\rho}(\psi)) = 1 - \operatorname{tr}\left[\hat{\rho}(\psi)^2\right] = 1 - \left(p_\uparrow^2 + p_\downarrow^2\right) = p_\uparrow p_\downarrow + p_\downarrow p_\uparrow$ which is the sum of the absolute squares of the off-diagonal entries that are zeroed in the transition $\rho(\psi) \rightsquigarrow \hat{\rho}(\psi)$ due to the projective measurement. As always, the logical quantum entropy $h(\hat{\rho}(\psi)) = p_\uparrow p_\downarrow + p_\downarrow p_\uparrow$ has a simple interpretation: the probability that in two independent measurements, distinct eigenvalues are obtained.