Entanglement in Sets and Vector Spaces

David Ellerman

UCR

April 2012

Comparison between set and quantum case

	Set Case	Quantum Case
Product	$X \times Y$	$H^A \otimes H^B$
Given state	$p(x,y)$	$\rho^{AB} = \vert\psi\rangle\langle\psi\vert$
Marginals	$p(x), p(y)$	ρ^A, ρ^B
Independent	$p(x,y) = p(x)\,p(y)$	$\rho^{AB} = \rho^A \otimes \rho^B$
Entangled	$p(x,y) \neq p(x)\,p(y)$	$\rho^{AB} \neq \rho^A \otimes \rho^B$
Bijection	$\{x_i\} \longleftrightarrow \{y_i\}$	$\{\vert i_A\rangle\} \longleftrightarrow \{\vert i_B\rangle\}$
Schmidt p_i	$p(x_i, y_i) = p_i$	$\vert\psi\rangle = \sum_i \sqrt{p_i}\,\vert i_A\rangle\,\vert i_B\rangle$
Ent. Meas.	$d(p(x,y)\,\Vert p(x)\,p(y))$	$d\left(\rho^{AB}\Vert\rho^A \otimes \rho^B\right)$
Formula	$\sum_i p_i^2 - 2\sum_i p_i^3 + \sum_{i,j} p_i^2 p_j^2$	$1 - 2\sum_i p_i^3 + \sum_{i,j} p_i^2 p_j^2$
Max Entang.	$p_i = p_j$	$p_i = p_j$

What is entanglement?: I

- The research program of providing an objective indistinctness interpretation of QM uses, in part, the device of looking at the relevant mathematics of partitions at the level of sets, and then lifting the math to vector spaces where QM lives.
- Hence the beginning of understanding entanglement from this perspective is to understand it at the level of sets.
- Given a set X (thinking of the singletons as the set of completely distinct or eigen-elements), a subset S_X is the set-analogue of a superposition of its elements. Thus a block S_X in a partition on X is a "state" that is indistinct between its elements but has been distinguished from the elements outside of it.

What is entanglement?: II

- As more distinctions are made, the block gets refined eventually into singletons, the fully distinct or eigen-elements.
- To visualize this refinement of blocks, consider the powerset $\wp(X)$ partially ordered by inclusion and then flip it over (and throw away the null set) to get the *block refinement partial ordering* $\wp(X)^{op} - \emptyset$.
- Then you have the picture of the blob or ur-block X at the bottom and then all the other blocks that can result from making more and more distinctions until you arrive at the maximally distinguished atoms or singletons in the reverse-inclusion partial order.

What is entanglement?: III

- Each block or subset of X is understood as a mini-blob or objectively indistinct "element" which with more distinctions could be eventually refined into one its eigen-elements. It is indistinct between those eigen-alternatives, but it has been distinguished from all the other eigen-elements (outside the subset).
- Consider the reverse question of what happens if we go the other way of grouping the fully distinct eigen-elements together (i.e., superposing them) to get indistinct elements or blocks? Do we get anything new? In this case, no. All the blocks that can be obtained by forming new indistinct elements are just subsets of X already in the order $\wp(X)^{op} - \emptyset$.

What is entangement?: IV

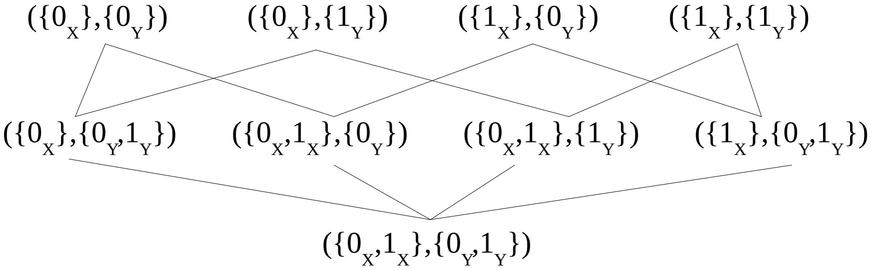
- Suppose we start with two sets X and Y and consider the two block-refinement orderings $\wp(X)^{op} - \emptyset$ and $\wp(Y)^{op} - \emptyset$. In each ordering by itself, nothing new appears if we superpose distinct eigen-elements to get blocks indistinct between their elements.
- Now we take the product of the two orderings $\left[\wp(X)^{op} - \emptyset\right] \times \left[\wp(Y)^{op} - \emptyset\right]$. The bottom of the ordering is $X \times Y$, the blob or ur-block in the combined system, and the fully distinct maximal elements in the product ordering are the eigen-pairs of singletons $(\{x\}, \{y\})$ or less pedantically $(x, y) \in X \times Y$, i.e., all the fully distinct alternatives that can be developed out of the ur-block $X \times Y$ by making distinctions.

What is entangleme nt?: V

- Note that there is nothing new in the product eigen-elements; they are just pairs of eigen-elements from X and Y. The fully distinct eigen-elements of $\wp(X \times Y)^{op} - \emptyset$ are all there in $\left[\wp(X)^{op} - \emptyset\right] \times \left[\wp(Y)^{op} - \emptyset\right]$.
- We ask: if we group together or superpose some of these fully distinct eigen-pairs, $(\{x\}, \{y\})$ or simply (x, y) for $x \in X$ and $y \in Y$, do we get anything new or just the products of blocks of elements of X and of Y?
- We can get all the products of X-blocks and Y-blocks in this way, but we also get new indistinct blocks that are not product blocks.
- Those new blocks are the entangled states which represent the new ways to make indistinct elements due to the "interaction" of X and Y.

- Consider the set version of two qubit space where $X = \{0_X, 1_X\}$ and $Y = \{0_Y, 1_Y\}$. The two block refinement orders are simple:

- Trivially, each superposition of the maximally distinct elements gives nothing new within each ordering.
- Then we take the product of the two orders to obtain:



- A product block such as $(\{0_X\},\{0_Y,1_Y\})$ represents the subset $\{(0_X,0_Y),(0_X,1_Y)\} \subseteq X \times Y$ so this gives a suborder of the block refinement ordering for $X \times Y$.
- But not all subsets of $X \times Y$ can be obtained as products of blocks from X and Y.

Example: III

- By "interacting" X and Y, new types of indistinct "states" can be formed by grouping together or "superposing" some of the fully distinct eigen-elements.
- For instance $\{(0_X, 0_Y), (1_X, 1_Y)\}$ and $\{(0_X, 1_Y), (1_X, 0_Y)\}$ are new types of "entangled" indistinct states as well as $\{(0_X, 0_Y), (0_X, 1_Y), (1_X, 1_Y)\}$ and three others (see continuation of the example below).
- This set example illustrates the **basic point** that by interacting two systems, no new fully distinct states are created but new types of indistinct states become possible and they are the entangled states. In spite of all the talk of entanglement as a uniquely quantum phenomenon, we see that it arises already at the level of ordinary sets. Now to measure it.

Entanglement of sets: I

- Given two finite sets X and Y, a subset $S \subseteq X \times Y$ of their Cartesian product is a *product* subset if there are subsets $S_X \subseteq X$ and $S_Y \subseteq Y$ such that $S = S_X \times S_Y$.
- A subset $S \subseteq X \times Y$ that is not a product subset might be called an *entangled* subset.
- For any subset $S \subseteq X \times Y$, a natural measure of its entanglement can be constructed by first viewing S as the support of the equiprobable or Laplacian joint probability distribution on S.
- If $|S| = N$, then define $p(x,y) = \frac{1}{N}$ if $(x,y) \in S$ and $p(x,y) = 0$ otherwise.
- The marginal distributions are defined in the usual way:
  - $p(x) = \sum_y p(x,y)$

Entanglement of sets: II

- $p(y) = \sum_x p(x,y)$.

- A joint probability distribution $p(x,y)$ on $X \times Y$ is *independent* if for all $(x,y) \in X \times Y$,

$$p(x,y) = p(x)\,p(y).$$

Independent distribution

Theorem

A subset $S \subseteq X \times Y$ is entangled iff the equiprobable distribution on S is not independent.

Proof: Let S_X be the support or projection of S on X, i.e., $S_X = \{x : \exists y \in Y, (x,y) \in S\}$ and similarly for S_Y. If S is not entangled, i.e., $S = S_X \times S_Y$, then $p(x) = |S_Y|/N$ for $x \in S_X$ and

$p(y) = |S_X| / N$ for $y \in S_Y$ where $|S_X| |S_Y| = N$. Then for $(x, y) \in S$,

$$p(x, y) = \frac{1}{N} = \frac{N}{N^2} = \frac{|S_X||S_Y|}{N^2} = p(x) p(y)$$

and $p(x, y) = 0 = p(x) p(y)$ for $(x, y) \notin S$ so the equiprobable distribution is independent. If $S \neq S_X \times S_Y$, then $S \subsetneq S_X \times S_Y$ so let $(x, y) \in S_X \times S_Y - S$. Then $p(x), p(y) > 0$ but $p(x, y) = 0$ so it is not independent. $\square$

- Hence for sets, a measure of entanglement of a subset $S \subseteq X \times Y$ would the measure of the logical divergence between the equiprobable distribution $p(x, y)$ on S and the product of marginals distribution $p(x) p(y)$:

Entanglement of sets: IV

$$d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) =$$
$$2h\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) - h\left(p\left(x,y\right)\right) - h\left(p\left(x\right)p\left(y\right)\right).$$

- $d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) > 0$ iff S is an entangled subset and
- $d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) = 0$ iff S is a product subset.
- The logical entropy of the equiprobable distribution is $h\left(p\left(x,y\right)\right) = 1 - N\frac{1}{N^2} = 1 - \frac{1}{N}$ where $|S| = N$.
- For any $x \in S_X$, let $s_x = |\{y : (x,y) \in S\}|$ and similarly for s_y. Then $p\left(x\right) = \frac{s_x}{N}$ and $p\left(y\right) = \frac{s_y}{N}$ so the logical entropy of the product distribution is

$$h\left(p\left(x\right)p\left(y\right)\right) = 1 - \sum_{x\in S_X, y\in S_Y} \frac{s_x^2 s_y^2}{N^4}.$$

- The cross-entropy is:

$$h\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) = 1 - \sum_{(x,y)\in S}\frac{s_x s_y}{N^3}.$$

- Hence putting it together:

$$d\left(p||p\left(x\right)p\left(y\right)\right) = 2h\left(p||p\left(x\right)p\left(y\right)\right) - h\left(p\right) - h\left(p\left(x\right)p\left(y\right)\right)$$
$$= 2\left[1 - \sum_{(x,y)\in S}\frac{s_x s_y}{N^3}\right] - \left[1 - \frac{1}{N}\right] - \left[1 - \sum_{x\in S_X, y\in S_Y}\frac{s_x^2 s_y^2}{N^4}\right]$$
$$\boxed{d\left(p||p\left(x\right)p\left(y\right)\right) = \sum_{x\in S_X, y\in S_Y}\frac{s_x^2 s_y^2}{N^4} + \frac{1}{N} - 2\sum_{(x,y)\in S}\frac{s_x s_y}{N^3}}.$$

Entanglement of sets: VI

- In the relevant range, $1 \leq s_x \leq |S_X|$ and $1 \leq s_y \leq |S_Y|$, the divergence is inversely related to the s_x and s_y so the maximum entanglement (by this measure) occurs when $s_x = s_y = 1$ which means that S is the graph of a bijection between a subset of X and a subset of Y (which in combinatorics is a partial matching between X and Y).
- In that maximum entanglement case, the value of the divergence is:

$$MaxDiv = \sum_{x \in S_X, y \in S_Y} \frac{s_x^2 s_y^2}{N^4} + \frac{1}{N} - 2 \sum_{(x,y) \in S} \frac{s_x s_y}{N^3}$$
$$= |S_X| \, |S_Y| \frac{1}{N^4} + \frac{1}{N} - 2 \frac{1}{N^2} = \frac{1}{N^4} \left[|S_X| \, |S_Y| + N^3 - 2N^2 \right].$$

Since S is the graph of a bijection between a subset of X and of Y, we might as well throw away the rest of X and Y so that $|S_X| = |S_Y| = N$, and then we have:

Entanglement of sets: VII

$$MaxDiv = \frac{1}{N^4}\left[|S_X|\,|S_Y| + N^3 - 2N^2\right] = \frac{1}{N^4}\left[N^3 - N^2\right] \text{ so}$$

$$\boxed{MaxDiv = \frac{1}{N}\left[1 - \frac{1}{N}\right]}.$$

- Alternatively, instead of cutting down X and Y so that $|S_X| = |S_Y| = N$, we can increase the *MaxDiv* by increasing N (since the derivative of $\frac{1}{N}\left[1 - \frac{1}{N}\right]$ is positive for positive N) until $N = \min\left(|X|, |Y|\right)$. Then we could throw away the excess in X or Y and use the above *MaxDiv* formula where a bijection, that is a *complete* matching between X and Y, is the maximum entanglement subset.

Example continued: I

- Consider again the set version of two qubit space where $X = \{0_X, 1_X\}$ and $Y = \{0_Y, 1_Y\}$.
- The product space $X \times Y$ has 15 nonempty subsets. Each factor X and Y has 3 nonempty subsets so $3 \times 3 = 9$ of the 15 subsets are product subsets leaving 6 entangled subsets.
- The 6 entangled subsets and their divergences are:

$\{(0,0),(1,1)\}$	$\{(0,1),(1,0)\}$
$\frac{1}{4}$	$\frac{1}{4}$
$\{(0,0),(0,1),(1,0)\}$	$\{(0,0),(0,1),(1,1)\}$
$\frac{4}{81}$	$\frac{4}{81}$
$\{(0,1),(1,0),(1,1)\}$	$\{(0,0),(1,0),(1,1)\}$
$\frac{4}{81}$	$\frac{4}{81}$

Example continued: II

- The first two "Bell subsets" are the two graphs of bijections $X \longleftrightarrow Y$ and have the maximum entanglement which can be calculated by the formula for the maximum divergence where $N = 2$, $MaxDiv = \frac{1}{2}\left[1 - \frac{1}{2}\right] = \frac{1}{4}$.
- The entanglement of say $\{(0,0), (0,1), (1,0)\}$ is calculated using $N = 3$, $s_{0X} = 2 = s_{0Y}$ and $s_{1X} = 1 = s_{1Y}$.
- All the 9 product states have zero entanglement. For instance, for $S = \{(0,0), (0,1)\}$, we have $N = 2$, $s_{0X} = 2$, $s_{1X} = 0$, and $s_{0Y} = s_{1Y} = 1$ so that:

$$d(p||p(x)p(y)) = \sum_{x \in S_X, y \in S_Y} \frac{s_x^2 s_y^2}{N^4} + \frac{1}{N} - 2\sum_{(x,y) \in S} \frac{s_x s_y}{N^3}$$
$$= \left[\frac{4}{16} + \frac{4}{16}\right] + \frac{1}{2} - 2\left[\frac{2}{8} + \frac{2}{8}\right] = \frac{1}{2} + \frac{1}{2} - 1 = 0.$$

Digression on probabilities as random variables: I

- Any finite probability distribution $p = \{p_1, ..., p_n\}$ can be viewed as a random variable taking the value p_i with the probability p_i.
- The *expectation* of p is $E_p(p) = \sum_i p_i^2$ so the logical entropy $h(p) = 1 - \sum_i p_i^2$ is the complement of the expectation of p.
- Given another distribution $q = \{q_1, ..., q_n\}$ over the same index set, the *cross-expectation* is:

$$E(p||q) = E_p(q) = E_q(p) = \sum_i p_i q_i$$

- The logical cross-entropy $h(p||q) = 1 - \sum_i p_i q_i$ is the complement of the cross-expectation.

Digression on probabilities as random variables: II

- The logical divergence is:

$$d(p||q) = \sum_i (p_i - q_i)^2 = 2h(p||q) - h(p) - h(q)$$
$$= 2[1 - E(p||q)] - \left[1 - E_p(p)\right] - \left[1 - E_q(q)\right]$$
$$= E_p(p) + E_q(q) - 2E(p||q) = E_p(p) + E_q(q) - E_p(q) - E_q(p)$$
$$\boxed{d(p||q) = \left(E_p - E_q\right)(p - q)}$$

Divergence in terms of linear expectation operators

- The logical information inequality that $d(p||q) \geq 0$ can then be written as:

$$E_p(p) + E_q(q) \geq E_p(q) + E_q(p)$$

Sum of self-expectations $\geq$ sum of cross-expectations.

Probabilities on bijections: I

- In the set case, a subset $S \subseteq X \times Y$ that is the graph of a bijection is the set analogue of the Schmidt decomposition of a pure state on a tensor product which is always available when working over Hilbert spaces. The different pairs of orthogonal basis states in a Schmidt decomposition $|\psi\rangle = \sum_i \sqrt{p_i} |i_A\rangle \otimes |i_B\rangle$ may have different Schmidt coefficients $\sqrt{p_i}$. Hence to develop the set analogue, we assume S is a bijection graph but allow an arbitrary probability distribution on S.
- A bijective S has the form $\{(x_i, y_i) : i = 0, ..., N-1\}$ so we assume a probability distribution with $p(x_i, y_i) = p_i$ and 0 otherwise.

Probabilities on bijections: II

- Since the set S is a bijection, the marginals are $p(x_i) = p_i = p(y_i)$, so that $h(p(x)) = h(p(y)) = 1 - \sum_i p_i^2$. This is the set analogue of the reduced density matrices having the same eigenvalues p_i in the quantum case.
- The logical entropy of $p(x,y)$ is: $h(p(x,y)) = 1 - \sum_i p_i^2$.
- The logical entropy of the product distribution is: $h(p(x)p(y)) = 1 - \sum_{i,j} p_i^2 p_j^2 = 1 - \left(\sum_i p_i^2\right)^2$.
- The cross-entropy is: $h(p(x,y) \| p(x)p(y)) = 1 - \sum_i p_i^3$.
- Hence the measure of entanglement in this case is:

Probabilities on bijections: III

$$d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) =$$
$$2h\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) - h\left(p\left(x,y\right)\right) - h\left(p\left(x\right)p\left(y\right)\right)$$
$$= 2\left[1 - \sum_i p_i^3\right] - \left[1 - \sum_i p_i^2\right] - \left[1 - \left(\sum_i p_i^2\right)^2\right] \text{ so}$$
$$\boxed{d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) = \sum_i p_i^2 - 2\sum_i p_i^3 + \left(\sum_i p_i^2\right)^2}.$$

- In terms of the expectations,
  - $E_p\left(p\right) = \sum_i p_i^2$,
  - $E_{p(x)p(y)}\left(p\left(x\right)p\left(y\right)\right) = \sum_{i,j} p_i^2 p_j^2 = \left(\sum_i p_i^2\right)^2$,
  - $E\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) = \sum_i p_i^3$.
- The variance of the random variable p is:

$$Var\left(p\right) = E_p\left(p^2\right) - E_p\left(p\right)^2 = \sum_i p_i^3 - \left(\sum_i p_i^2\right)^2.$$

Probabilities on bijections: IV

- Hence the divergence formula in this special case is:

$$d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right) = E_p\left[p\left(x,y\right) - p\left(x\right)p\left(y\right)\right] - Var\left(p\right).$$

- Here again, the maximum divergence & entanglement is the equiprobable case, $p_i = \frac{1}{N}$, where $E_p\left(p\right) = \frac{1}{N}$, $E_p\left(p\left(x\right)p\left(y\right)\right) = \frac{1}{N^2}$, and $Var\left(p\right) = 0$ so the formula gives the previous result:

$$\frac{1}{N} - \frac{1}{N^2} = \frac{1}{N}\left[1 - \frac{1}{N}\right].$$

- Note how the variance of p takes away from the divergence in this bijective case, and the variance is 0 for both the extreme cases: $p_i = \frac{1}{N}$ and $p_1 = 1$. The first case ($p_i = \frac{1}{N}$) is the maximum entanglement and the second case ($p_1 = 1$) is zero entanglement (product state).

Quantum case: I

- Given two systems A, B represented in Hilbert spaces H^A and H^B, let $\rho^{AB} = |\psi\rangle\langle\psi|$ be a *pure* state in the tensor product $H^A \otimes H^B$.
- If $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ are orthonormal bases for the component spaces, let $[\alpha]$ be the matrix of coefficients for ψ, i.e.,

$$[\alpha] = [\alpha_{ij}] \text{ where } |\psi\rangle = \sum_{i,j} \alpha_{ij} |a_i\rangle \otimes |b_j\rangle.$$

- Then $[\alpha][\alpha]^\dagger = \rho^A$ is the reduced density matrix on H^A and $[\alpha]^\dagger [\alpha] = \rho^B$ is the reduced density matrix on H^B, where $[]^\dagger$ is the Hermitian transpose.

Quantum case: II

- The Schmidt decomposition of ψ is $|\psi\rangle = \sum_i \sqrt{p_i}\,|i_A\rangle \otimes |i_B\rangle$ where $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ are orthonormal bases for the component spaces.
- Where $p_i \neq 0$, the Schmidt decomposition establishes a bijection between a subset of the basis $\{|i_A\rangle\}$ and a subset of the basis $\{|i_B\rangle\}$ which is the "lift" of such a bijection in the set case.
- The reduced density matrices can then be expressed as: $\rho^A = \sum_i p_i\,|i_A\rangle\langle i_A|$ and $\rho^B = \sum_i p_i\,|i_B\rangle\langle i_B|$ so the p_i are the non-negative eigenvalues for both reduced density matrices.
- Since the trace is invariant under similarity transformations and since each density matrix could be diagonalized to its diagonal matrix of eigenvalues, the traces of the squares are:

Quantum case: III

$$\mathrm{tr}\left[\left(\rho^A\right)^2\right] = \mathrm{tr}\left[\left(\rho^B\right)^2\right] = \sum_i p_i^2$$

so that $h\left(\rho^A\right) = h\left(\rho^B\right) = 1 - \sum_i p_i^2$.

- Since ρ^{AB} is assumed to be a pure state, $\mathrm{tr}\left[\left(\rho^{AB}\right)^2\right] = 1$ so $h\left(\rho^{AB}\right) = 1 - \mathrm{tr}\left[\left(\rho^{AB}\right)^2\right] = 0$.
- The logical entropy of the product state $\rho^A \otimes \rho^B$ is:

$$h\left(\rho^A \otimes \rho^B\right) = 1 - \mathrm{tr}\left[\left(\rho^A\right)^2\right] \mathrm{tr}\left[\left(\rho^B\right)^2\right] = 1 - \left(\sum_i p_i^2\right)^2.$$

Quantum case: IV

- The Schmidt number is the number of non-zero p_i, and it is 1 with $p_1 = 1$ iff ρ^{AB} is a product state, i.e., $\rho^{AB} = \rho^A \otimes \rho^B$. Then $\sum_i p_i^2 = 1$ and $h\left(\rho^A\right) = h\left(\rho^B\right) = 0$ so that $d\left(\rho^{AB} || \rho^A \otimes \rho^B\right) = 0$ as well.
- In the Schmidt basis for the case where both H^A and H^B are three dimensional, then:

$$|\psi\rangle = \sqrt{p_0}\,|0_A\rangle \otimes |0_B\rangle + \sqrt{p_1}\,|1_A\rangle \otimes |1_B\rangle + \sqrt{p_2}\,|2_A\rangle \otimes |2_B\rangle.$$

- Then the matrix for ρ^{AB} in the $\{|i_A\rangle \otimes |j_B\rangle\}$ basis is:

Quantum case: V

$$\rho^{AB} = \begin{bmatrix} p_0 & 0 & 0 & 0 & \sqrt{p_0 p_1} & 0 & 0 & 0 & \sqrt{p_0 p_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \sqrt{p_1 p_0} & 0 & 0 & 0 & p_1 & 0 & 0 & 0 & \sqrt{p_1 p_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \sqrt{p_2 p_0} & 0 & 0 & 0 & \sqrt{p_2 p_1} & 0 & 0 & 0 & p_2 \end{bmatrix}.$$

- The matrix for $\rho^A \otimes \rho^B$ is diagonal:

$$\rho^A \otimes \rho^B = \begin{bmatrix} p_0^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_0p_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_0p_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_1p_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_1p_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_2p_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2p_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_2^2 \end{bmatrix}.$$

- The cross-entropy $h\left(\rho^{AB}||\rho^A \otimes \rho^B\right) = 1 - \mathrm{tr}\left[\rho^{AB}\left(\rho^A \otimes \rho^B\right)\right]$ will in general just pick out the cubic terms:

$$h\left(\rho^{AB}||\rho^A \otimes \rho^B\right) = 1 - \sum_i p_i^3.$$

- The suggested measure of entanglement is the logical divergence $d\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)$ which can be computed as:

$$d\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)=2h\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)-h\left(\rho^{AB}\right)-h\left(\rho^{A}\otimes\rho^{B}\right)$$
$$=2\left[1-\operatorname{tr}\left[\rho^{AB}\left(\rho^{A}\otimes\rho^{B}\right)\right]\right]-\left[1-\operatorname{tr}\left[\left(\rho^{AB}\right)^{2}\right]\right]-$$
$$\left[1-\operatorname{tr}\left[\left(\rho^{A}\otimes\rho^{B}\right)^{2}\right]\right]$$
$$=\operatorname{tr}\left[\left(\rho^{AB}\right)^{2}\right]-2\operatorname{tr}\left[\rho^{AB}\left(\rho^{A}\otimes\rho^{B}\right)\right]+\operatorname{tr}\left[\left(\rho^{A}\otimes\rho^{B}\right)^{2}\right]$$

which is the lift of the set version:

$$d\left(p\left(x,y\right)||p\left(x\right)p\left(y\right)\right)=\sum_i p_i^2-2\sum_i p_i^3+\sum_{i,j}p_i^2p_j^2.$$

- Since we have assumed that ρ^{AB} is a pure state (in order to use the Schmidt decomposition), $\operatorname{tr}\left[\left(\rho^{AB}\right)^2\right] = 1$ so the final formula for the entanglement measure in terms of the Schmidt coefficients is:

$$\boxed{d\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right) = 1 - 2\sum_i p_i^3 + \left(\sum_i p_i^2\right)^2}.$$

Entanglement measure for pure ρ^{AB}
in terms of Schmidt coefficients

- Here again, it can be shown that the entanglement measure is a maximum when the non-zero Schmidt coefficients are equal so if there are n non-zero p_i's, then $p_1 = ... = p_n = \frac{1}{n}$. Such a case is often called "maximally entangled" so the entanglement measure agrees.

- When all the Schmidt coefficients are equal, the value of the maximum divergence is:

$$d\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)=1-2\sum_i p_i^3+\left(\sum_i p_i^2\right)^2$$
$$=1-2\frac{n}{n^3}+\left(\frac{n}{n^2}\right)^2=1-2\frac{n^2}{n^4}+\frac{n^2}{n^4}=1-\frac{n^2}{n^4}=1-\frac{1}{n^2}$$

which differs from the set formula in the first term which is $\operatorname{tr}\left[\left(\rho^{AB}\right)^2\right]=1$ instead of $\sum_i p_i^2$ since there is no set analogue of a non-trivial pure state. In the set case, a "pure state" is the trivial case $p_1=1$ and then indeed $\sum_i p_i^2=1$.

- The Schmidt coefficients (squared) give a probability distribution $p=\{p_1,...,p_n\}$ so we may restate the divergence formula using the expectations and variance:

Quantum case: X

$$d\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)=1-\sum_i p_i^3-Var\left(p\right)=$$
$$h\left(\rho^{AB}||\rho^{A}\otimes\rho^{B}\right)-Var\left(p\right)$$

where the variance in the Schmidt coefficients is 0 in both the extreme cases of maximum entanglement and zero entanglement.

- For all the Bell basis vectors in two qubit space, $p_1=p_2=\frac{1}{2}$ and their maximal measure of entanglement is $1-\left(\frac{1}{2}\right)^2=\frac{3}{4}$.

Comparison between set and quantum case

	Set Case	Quantum Case
Product	$X \times Y$	$H^A \otimes H^B$
Given state	$p(x,y)$	$\rho^{AB} = \|\psi\rangle\langle\psi\|$
Marginals	$p(x), p(y)$	ρ^A, ρ^B
Independent	$p(x,y) = p(x)p(y)$	$\rho^{AB} = \rho^A \otimes \rho^B$
Entangled	$p(x,y) \neq p(x)p(y)$	$\rho^{AB} \neq \rho^A \otimes \rho^B$
Bijection	$\{x_i\} \longleftrightarrow \{y_i\}$	$\{\|i_A\rangle\} \longleftrightarrow \{\|i_B\rangle\}$
Schmidt p_i	$p(x_i, y_i) = p_i$	$\|\psi\rangle = \sum_i \sqrt{p_i}\,\|i_A\rangle\,\|i_B\rangle$
Ent. Meas.	$d(p(x,y)\|\|p(x)p(y))$	$d(\rho^{AB}\|\|\rho^A \otimes \rho^B)$
Formula	$\sum_i p_i^2 - 2\sum_i p_i^3 + \sum_{i,j} p_i^2 p_j^2$	$1 - 2\sum_i p_i^3 + \sum_{i,j} p_i^2 p_j^2$
Max Entang.	$p_i = p_j$	$p_i = p_j$