



City Research Online

City, University of London Institutional Repository

Citation: Asano, M., Basieva, I., Pothos, E. M. & Khrennikov, A. (2018). State entropy and differentiation phenomenon. *Entropy*, 20(6), 394. doi: 10.3390/e20060394

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/19807/>

Link to published version: <https://doi.org/10.3390/e20060394>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Article

State Entropy and Differentiation Phenomenon

Masanari Asano ^{1,*}, Irina Basieva ², Emmanuel M. Pothos ² and Andrei Khrennikov ^{3,4}

¹ Liberal Arts Division, National Institute of Technology, Tokuyama College, Gakuendai, Shunan, Yamaguchi 745-8585, Japan

² Department of Psychology, University of London, London WC1E 7HU, UK; emmanuel.pothos.1@city.ac.uk

³ International Center for Mathematical Modeling in Physics and Cognitive Sciences Linnaeus University, 351 95 Växjö-Kalmar, Sweden; andrei.khrennikov@lnu.se

⁴ National Research University of Information Technologies, Mechanics and Optics, St. Petersburg 197101, Russia

* Correspondence: asano@tokuyama.ac.jp; Tel.: +81-834-29-6200

Received: 5 April 2018; Accepted: 21 May 2018; Published: date

Abstract: In the formalism of quantum theory, a state of a system is represented by a *density operator*. Mathematically, a density operator can be decomposed into a weighted sum of (projection) operators representing an ensemble of pure states (a state distribution), but such decomposition is not unique. Various pure states distributions are mathematically described by the same density operator. These distributions are categorized into classical ones obtained from the Schatten decomposition and other, non-classical, ones. In this paper, we define the quantity called the *state entropy*. It can be considered as a generalization of the *von Neumann entropy* evaluating the diversity of states constituting a distribution. Further, we apply the state entropy to the analysis of non-classical states created at the intermediate stages in the process of *quantum measurement*. To do this, we employ the model of *differentiation*, where a system experiences step by step state transitions under the influence of environmental factors. This approach can be used for modeling various natural and mental phenomena: cell's differentiation, evolution of biological populations, and decision making.

Keywords: density operator; state entropy; von neumann entropy; quantum measurement; differentiation

1. Introduction

In quantum theory, a state of a system is represented by a density operator. A density operator, e.g. ρ , can be decomposed into a weighted sum of (projection) operators representing “pure states”. This linear combination represents a statistical distribution of pure states in an ensemble of systems. However, the same density operator ρ can be decomposed in various ways. Hence, numerous statistical state distributions are mathematically encoded by the same ρ , unless ρ coincides with a pure state.

One class of these statistical distributions, namely, obtained from “Schatten decompositions” of ρ , plays a special role. We remark that, for a density operator with degenerate spectrum, Schatten decomposition is not unique. Any selection of orthogonal bases in eigensubspaces of ρ generates some Schatten decomposition. Each Schatten decomposition corresponds to the statistical distribution of eigenstates of ρ . The crucial point is that these eigenstates may be distinguishable on the basis of measurement of some physical quantity X , because these states are orthogonal to each other. The eigenvalues are interpreted as the frequency probabilities of the measurement outcomes. In this sense, the distribution corresponding to the concrete Schatten decomposition of the density operator ρ is conceptually equivalent to a “classical” or “standard” probability distribution.

On the other hand, other decompositions of the same state ρ are “non-classical” or “non-standard” and represent ensembles of pure states which may be not orthogonal to each other. In Section 2, we discuss these points in more detail.

The main topic of this paper is a quantity that evaluates structural features of various statistical state distributions encoded in the same density operator ρ . It is well-known that the von Neumann entropy [1,2], defined as $-\rho \log \rho$, can evaluate how ρ deviates from a pure state, i.e., the degree of mixture of pure states. In fact, $-\rho \log \rho$ can be rewritten as $\sum_k -\lambda_k \log \lambda_k$, where $\{\lambda_k\}$ are eigenvalues of ρ . It equals to zero if and only if ρ is a pure state. Note that the quantity $\sum_k -\lambda_k \log \lambda_k$ is the Shannon entropy for classical probability distribution $\{\lambda_k\}$. Thus, the von Neumann entropy evaluates only the classical distribution encoded in ρ , but not non-classical ones.

In this paper, we define a quantity such that more detailed information about the structure of statistical state distributions, especially non-classical ones, is reflected. Our discussion is fundamental, but straightforward. First, in Section 3, we mention the “differentiation phenomenon” which an ensemble of pure states experiences under a quantum measurement of some physical observable, say X . Each pure state is stochastically differentiated into an eigenstate of X . If pure states in the statistical ensemble are different, the expectation values of X estimated from each of them are also generally different. In Section 4, we focus on dispersion of these expectation values and discuss its mathematical property reflecting structural features of the state distribution. Finally, in Section 5, we define a “state entropy” (see Equation (17)). This quantity evaluates the “diversity” of pure states constituting an ensemble. It is proportional to the number of pure states and inversely proportional to similarities among them.

We also point to the interrelation between the state entropy and the von Neumann entropy. It can be briefly described in the following way. If a state distribution, which is encoded in ρ , is classical, then its state entropy is equal to the von Neumann entropy. The state entropies of non-classical state distributions do not exceed the latter; see the inequality of Equation (18): The state entropy is a generalization of von Neumann entropy which is extensively used in different types of quantum entropies, e.g. conditional, relative and mutual entropies [3–5].

State entropy evaluates non-classical statistical state distributions. To stress significance of the notion of state entropy, we explain the theoretical context of state distributions. We note that classical state distributions are always identified *after* completion of quantum measurements. Therefore, non-classical distributions may exist at the stages *before* measurements are completed, more generally, in the process of differentiation.

In Section 6, we focus on the model of differentiation that was discussed in Reference [6]. This model describes accumulation of very small state transitions experienced by the system, and each transition is mathematically represented by a map in the state space, i.e., by a “quantum channel” in the terminology of quantum information theory. A quantum channel denoted by Λ^* is given by Equation (28), which is concerned with “environmental elements” around the system. They are weakly interacting with the system causing numerous small state transitions step by step, if differentiations of states occur sequentially. The above picture corresponds to an ideal “open quantum system dynamics”. To describe the process of differentiation in the system, we consider a more complicated model, assuming differentiations not only of the system state, but also in the elements of the environment. The differentiation in each environmental element is similar to the determination of a “pointer basis” in the theory of quantum decoherence proposed by Zurek [7]. In our approach, the Lindblad equation [8,9], which is a traditional way to describe open quantum system dynamics, is not employed directly.

We believe that the described model can be applicable to a variety of natural and mental phenomena (not only in the micro-world). The process of creation of a diversity of states in an ensemble of systems, which were originally prepared in the same pure state Ψ , through mutual interaction with environmental factors is universal. Originally, the formalism of quantum theory was established to describe microscopic phenomena, but now it is widely used in psychology,

decision making, and finance (see [10–38]). It is also applied to model behavior of biological systems, especially the functioning of genetic and epigenetic systems (see [39–46]). We plan to explore the novel mathematical apparatus developed in this paper (based on the state entropy) for such applications elsewhere.

In psychology, there has been extensive interest in employing classical entropy for quantifying uncertainty, e.g., in decision making (entropy minimization was used to model decision biases in [47]), categorization (as a way to formalize intuitions in spontaneous grouping [48]), and learning [49,50]. We plan to apply the apparatus of the quantum state entropy to these problems.

As shown in Figure 1, the accumulation of transitions generated by channel Λ^* represents an ideal differentiation process realized in the system. Further, in this modeling, non-classical state distributions in the intermediate stages are identified (see Equations (29)–(31)). We analyze them by means of the state entropy (see Figures 2 and 3).

2. State Representation by Density Operator

If a physical quantity X is measurable in a system, the frequency probabilities $\{P(x)\}$ for observed values $\{x\}$ may be estimated. Then, the quantity X is a “stochastic variable” in terms of probability theory, and the distribution $\{P(x)\}$ is a “state of the system” which can be analyzed, e.g., by calculating the expectation value $E(X)$ or dispersion $V(X) = E(X^2) - (E(X))^2$, as is usual in statistics.

The mathematical framework of quantum theory includes probability theory, where classical concepts of stochastic variables and probability distribution are expanded using the notion of “operator”. Firstly, a physical quantity is defined in the form of

$$X = \sum_{k=1}^M x_k |x_k\rangle \langle x_k|. \quad (1)$$

This is a Hermitian operator in Hilbert space $\mathcal{H} = \mathbb{C}^M$ with real eigenvalues $x_k \in \mathbb{R}$ ($k = 1, 2, \dots, M$) and eigenvectors $\{|x_k\rangle\}$. (A vector $|x\rangle \in \mathcal{H}$ whose norm is 1 is called ket-vector, and $\langle x|$, which is Hermitian conjugate of $|x\rangle$, i.e., $|x\rangle^\dagger = \langle x|$, is called bra-vector.) The form of Equation (1) implies that after a non-degenerate value x_k is observed, the system under the measurement has the definite (pure) state represented by the operator $|x_k\rangle \langle x_k|$. Note that the trace of product of X and $|x_k\rangle \langle x_k|$ is equal to x_k :

$$\text{Tr}(X |x_k\rangle \langle x_k|) = \langle x_k| X |x_k\rangle = x_k.$$

For the calculation, the orthogonality of vectors, i.e., $\langle x_k | x_{k'} \rangle = 0$ if $k \neq k'$, is used. Next, using the pure states $\{|x_k\rangle \langle x_k|\}$, let us construct the operator:

$$\rho = \sum_{k=1}^M P(x_k) |x_k\rangle \langle x_k|. \quad (2)$$

where $\{P(x_k)\}$ corresponds to the frequency probabilities of the observed values $\{x_k\}$, and, in fact, the trace of $X\rho$ is equal to the expected value $E(X)$:

$$\text{Tr}(X\rho) = E(X). \quad (3)$$

Mathematically, ρ is a Hermitian matrix satisfying $\text{Tr}(\rho) = 1$ and $\langle x | \rho | x \rangle \geq 0$, $\forall x \in \mathcal{H} = \mathbb{C}^M$. Such operator is called density operator and used for representing a statistical mixture of pure states (a mixed state). A density operator may be given in the form of Schatten decomposition, i.e., represented as a diagonal matrix:

$$\sum_{k=1}^M \lambda_k |\phi_k\rangle \langle \phi_k|, \tag{4}$$

where $\{\lambda_k \geq 0\}$ are the eigenvalues of the matrix (the same as probabilities $\{P(x_k)\}$ of ρ), and $\{|\phi_k\rangle \in \mathcal{H} = \mathbb{C}^M\}$ are the corresponding eigenvectors (the same as $|x_k\rangle$ of ρ). From Equation (4), one can obtain a picture of statistical mixture of $\{|\phi_k\rangle \langle \phi_k|\}$. (This mixture is denoted by $\{\phi_k, \lambda_k\}$ hereafter.) As can be seen from the construction of ρ in Equation (2), to give a Schatten decomposition is conceptually equivalent to giving a probability distribution of measurement of some physical quantity. In this sense, the state distribution $\{\phi_k, \lambda_k\}$ is “classical”. We have to point out here that decomposition of density operator is not unique, generally: By considering various linear combinations of $\{|\phi_k\rangle\}$, one can find a set of vectors $\{|\Psi_i\rangle, i = 1, \dots, N\}$, which satisfies

$$\sum_{k=1}^M \lambda_k |\phi_k\rangle \langle \phi_k| = \sum_{i=1}^N P_i |\Psi_i\rangle \langle \Psi_i|, \sum_{i=1}^N P_i = 1. \tag{5}$$

Note that $N \geq M$ and the vectors $\{|\Psi_i\rangle \in \mathcal{H} = \mathbb{C}^M\}$ need not be orthogonal to each other, that is, they need not be eigenstates of a single physical quantity: the state distribution $\{\Psi_i, P_i\}$ is “non-classical”. There exist numerous state distributions corresponding to same density operator, other than $\{\phi_k, \lambda_k\}$ and $\{\Psi_i, P_i\}$, and they are non-classical.

3. Differentiation Phenomenon in Quantum Measurement Process

As shown in Equation (3), for the density operator

$$\rho = \sum_{k=1}^M P(x_k) |x_k\rangle \langle x_k|,$$

where $\{|x_k\rangle \langle x_k|\}$ are the eigenstates of X , $Tr(X\rho) = \sum_{k=1}^M P(x_k)x_k = E(X)$ is satisfied. In this section, noting the non-uniqueness of decomposition of density operator, we mention the meaning of $Tr(X\rho)$ that has not been discussed in the classical theory. Let us consider a different decomposition, $\rho = \sum_{i=1}^N P_i |\Psi_i\rangle \langle \Psi_i|$, that is, we assume the existence of non-classical state distribution $\{\Psi_i, P_i\}$. Then, $Tr(X\rho)$ is described as the statistical average of the averages $\{\langle X \rangle_{\Psi_i} = Tr(X |\Psi_i\rangle \langle \Psi_i|)\}$ of observable X with respect to the pure states $\{\Psi_i\}$:

$$Tr(X\rho) = \sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}. \tag{6}$$

Each term, e.g. $\langle X \rangle_{\Psi}$, in the above is expanded as

$$\langle X \rangle_{\Psi} = \sum_{k=1}^M x_k |\langle \Psi | x_k \rangle|^2. \tag{7}$$

($\sum_{k=1}^M |\langle \Psi | x_k \rangle|^2 = 1$ is satisfied.) The square of inner product $|\langle \Psi | x_k \rangle|^2$ is frequently called “transition probability”. It is related to a problem of measurement that has been discussed in the quantum theory. In the concept of quantum measurement, the existence of the measurement device is considered first, because it is assumed that some interaction between the device and the system realizes the measurement of a physical quantity. Due to the interaction, the initial state of system $|\Psi\rangle \langle \Psi|$ is transferred to one of $\{|x_k\rangle \langle x_k|\}$, and the values of $\{x_k\}$ can be read out from the device. If $\langle X \rangle_{\Psi} = \sum_{k=1}^M x_k |\langle \Psi | x_k \rangle|^2$ means the average of outputs, the value of $|\langle \Psi | x_k \rangle|^2$ corresponds to the probability of transition from $|\Psi\rangle \langle \Psi|$ to $|x_k\rangle \langle x_k|$.

We interpret the process of quantum measurement as a sort of “differentiation”, in which a group of systems in one initial state is divided into groups having different states by means of external

or environmental factors. The expected value of $\langle X \rangle_{\Psi_i}$ comes from one differentiation denoted by $\Psi_i \rightarrow \{x_k\}$, and the value of $Tr(X\rho) = \sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}$ is to be calculated supposing a statistical mixture of M kinds of differentiations, $\{\Psi_i \rightarrow \{x_k\}\} (i = 1, \dots, M)$.

4. Characteristic Quantity of State Distribution

We assume a definitive state distribution denoted by $\{\Psi_i, P_i\}$ is given, and the calculations of $\{\langle X \rangle_{\Psi_i}\}$ are possible. The average of $\{\langle X \rangle_{\Psi_i}\}$, i.e., $\sum_{i=1}^N P_i \langle X \rangle_{\Psi_i} = Tr(X\rho)$ depends only on the density operator ρ , in which $\{\Psi_i, P_i\}$ is encoded. A statistical quantity reflecting more detailed information on the structure of $\{\Psi_i, P_i\}$ is dispersion of $\{\langle X \rangle_{\Psi_i}\}$, formulated as

$$V(\{\langle X \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2 - \left(\sum_{i=1}^N P_i \langle X \rangle_{\Psi_i} \right)^2 = \sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2 - (Tr(X\rho))^2. \tag{8}$$

Below, we prove the inequality

$$V(\{x_i, P(x_i)\}) \geq V(\{\langle X \rangle_{\Psi_i}, P_i\}), \tag{9}$$

where $V(\{x_i, P(x_i)\})$ is the dispersion of observable X , i.e., its dispersion with respect to the probability distribution encoded in the Shatten decomposition (see Equation (1)), corresponding to the spectral decomposition of the observable X (see Equation (2)). Thus, the probability distribution corresponding to the spectral decomposition of X maximizes the dispersions with respect to decompositions in Equation (5). The inequality for dispersions can be interpreted by the theory of weak measurements. The quantities $\langle X \rangle_{\Psi_i}$ can be interpreted as weak values. In this framework, the inequality in Equation (9) simply means that dispersion of a weak measurement is always majorized by dispersion of the “maximally disturbing measurement”, represented by a Hermitian operator. At the same time, we are aware that interpretation of weak values is a complex foundational problem of itself.

To prove the inequality in Equation (9), let us consider the first term given by

$$D(\{\langle X \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2. \tag{10}$$

Let us note the following inequality

$$\sum_{k=1}^M \langle x_k | \rho | x_k \rangle (x_k)^2 \geq D(\{\langle X \rangle_{\Psi_i}, P_i\}) \geq (Tr(X\rho))^2. \tag{11}$$

which follows from the convexity of $y = x^2$, because

$$\sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2 \geq \left(Tr(X \sum_{i=1}^N P_i |\Psi_i\rangle \langle \Psi_i|) \right)^2 = (Tr(X\rho))^2, \tag{12}$$

and since

$$\sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2 = \sum_{i=1}^N P_i \left(\sum_{k=1}^M x_k |\langle x_k | \Psi_i \rangle|^2 \right)^2$$

where $X = \sum_{k=1}^M x_k |x_k\rangle \langle x_k|$, one can see

$$\sum_{i=1}^N P_i \langle X \rangle_{\Psi_i}^2 \leq \sum_{i=1}^N \sum_{k=1}^M P_i |\langle x_k | \Psi_i \rangle|^2 (x_k)^2 = \sum_{k=1}^M \langle x_k | \rho | x_k \rangle (x_k)^2. \tag{13}$$

Such inequality can be derived with the use of other convex functions, not limited to $y = x^2$. Even if the dispersion V is defined as

$$V(\{\langle X \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i f(\langle X \rangle_{\Psi_i}) - f(\text{Tr}(X\rho)), \tag{14}$$

using another convex function, e.g. $f(x)$, the result holds true, that is, the inequality

$$\sum_{k=1}^M \langle x_k | \rho | x_k \rangle f(x_k) - f(\text{Tr}(X\rho)) \geq V(\{\langle X \rangle_{\Psi_i}, P_i\}) \geq 0, \tag{15}$$

is satisfied.

We redefine the first term D of Equation (10) as

$$D(\{\langle X \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i f(\langle X \rangle_{\Psi_i}). \tag{16}$$

As discussed in the next section, we believe that, under proper choices of X and $f(x)$, this D itself becomes a quantity that captures structural features of $\{\Psi_i, P_i\}$.

5. State entropy

In this section, we consider D of Equation (16) in the case of $X = \rho$ and $f(x) = -\log x$:

$$D(\{\langle \rho \rangle_{\Psi_i}, P_i\}) = - \sum_{i=1}^N P_i \log \langle \rho \rangle_{\Psi_i}. \tag{17}$$

Here, we fix the state distribution $\{\Psi_i, P_i\}$ for the density operator ρ , and $y = -\log x$ is our choice of a convex function. What does the above D tell us about $\{\Psi_i, P_i\}$? To discuss this question, we first focus on the term of $\text{Tr}(\rho |\Psi_i\rangle \langle \Psi_i|) = \langle \rho \rangle_{\Psi_i}$. Since

$$\langle \rho \rangle_{\Psi_i} = P_i + \sum_{j \neq i} P_j |\langle \Psi_i | \Psi_j \rangle|^2,$$

$$1 > \langle \rho \rangle_{\Psi_i} \geq P_i,$$

is satisfied. One can see, $\langle \rho \rangle_{\Psi_i} = P_i$ if all the vectors $\{|\Psi_{j \neq i}\rangle\}$ are orthogonal to $|\Psi_i\rangle$, and $\langle \rho \rangle_{\Psi_i} = 1$ if all $\{|\Psi_{j \neq i}\rangle\}$ are parallel to $|\Psi_i\rangle$. Based on this, we interpret $\langle \rho \rangle_{\Psi_i}$ as a degree of “similarity” of $|\Psi_i\rangle \langle \Psi_i|$ and ρ . This interpretation of quantity $\langle \rho \rangle_{\Psi_i}$ as the degree of similarity can also be illustrated by the representation of the operators $|\Psi_i\rangle \langle \Psi_i|$ and ρ as vectors in the Hilbert space of Hilbert–Schmidt operators endowed with the scalar product $\langle A | B \rangle = \text{Tr} A^* B$. We start with the remark that $\langle A | B \rangle = \cos \theta_{AB} \|A\|_2 \|B\|_2$, where $\|\cdot\|_2$ is the Hilbert–Schmidt norm; we also remark that, for a self-adjoint operator A , $\|A\|_2 = \text{Tr} A^2$. In particular, the norm of any pure state and the norm of any projector are equal to one. We have

$$\langle |\Psi_i\rangle \langle \Psi_i| | \rho \rangle = \text{Tr} \left(\sum_j P_j |\Psi_i\rangle \langle \Psi_i| |\Psi_j\rangle \langle \Psi_j| \right) = \langle \rho \rangle_{\Psi_i}.$$

Hence,

$$\langle \rho \rangle_{\Psi_i} = \cos \theta \text{Tr} \rho^2,$$

where θ is the angle between the vectors $|\Psi_i\rangle \langle \Psi_i|$ and ρ . The scaling coefficient $\text{Tr} \rho^2$ is the purity of the state ρ .

Further, noting that $y = -\log x$ is a monotonically decreasing function, we interpret $-\log \langle \rho \rangle_{\Psi_i} = -\log \cos \theta - \log \text{Tr} \rho^2$ as a degree of orthogonality between the vectors $|\Psi_i\rangle \langle \Psi_i|$ and ρ . We note that the following inequality is satisfied: $-\log P_i \geq -\log(\langle \rho \rangle_{\Psi_i}) > 0$.

In general, any convex and monotonically decreasing function is allowed as $f(x)$. The average of orthogonality $-\log \langle \rho \rangle_{\Psi_i}$, i.e., $\sum_{i=1}^N P_i(-\log \langle \rho \rangle_{\Psi_i})$ corresponds to D of Equation (17). Generally, the value of D will increase in proportion to the number of states and decrease in proportion to similarities among them. That is why we call the value D “state diversity” or “state entropy”.

The following inequality shows the significance of state entropy D :

$$\sum_{k=1}^M \lambda_k (-\log \lambda_k) \geq D(\{\langle \rho \rangle_{\Psi_i}, P_i\}) \geq -\log(\text{Tr}(\rho^2)). \tag{18}$$

It can be derived with the use of convexity of $y = -\log x$, in a similar way as derivation of Equation (11). In the above form, $\{\lambda_k\}$ are the eigenvalues of $\rho = \sum_{k=1}^M \lambda_k |\phi_k\rangle \langle \phi_k|$. The term of $\sum_{k=1}^M \lambda_k (-\log \lambda_k)$ in the left-hand side corresponds to von Neumann entropy given by $-\rho \log \rho$. Further, $\text{Tr}(\rho^2)$ in the right-hand side is a well-known quantity in the quantum theory, too. The von Neumann entropy $-\rho \log \rho$ and $\text{Tr}(\rho^2)$ are frequently used to evaluate the degree of “mixing” in ρ : if ρ is pure, then, $-\rho \log \rho = 0$ and $\text{Tr}(\rho^2) = 1$. If ρ is a mixed state, $-\rho \log \rho > 0$ and $\text{Tr}(\rho^2) < 1$, and especially, when $\lambda_1 = \lambda_2 = \dots = \lambda_M = 1/M$, $-\rho \log \rho$ takes the maximum value of $\log M$, and $\text{Tr}(\rho^2)$ takes minimum value of $1/M$. Mathematically, these two quantities have the relation of $-\rho \log \rho \geq -\log(\text{Tr}(\rho^2))$. The inequality in Equation (18) implies that the intermediate values between these two correspond to other kinds of state entropy, which can be estimated for various non-classical state distributions reducing to ρ . In other words, the well-known $-\rho \log \rho$ and $\text{Tr}(\rho^2)$ are newly interpreted as maximum and minimum values of state entropy.

Note that the state entropy D is different from the generalized quantum entropic measures that have been proposed until now. This point is mentioned in the Appendix.

6. Model of Differentiation and Calculation of State Entropy

As mentioned in Section 2, a Schatten decomposition of a density operator such as Equation (2) represents a probabilistic distribution of orthogonal pure states. Such an ensemble of states is postulated to be the resulting state of the system after measurement of some physical quantity, whose eigenstates are orthogonal. On the other hand, using another decomposition of the density operator, a mixture of non-orthogonal pure states may be obtained, and we call such mixture non-classical. In Section 3, we point out that the essence of quantum measurement is state differentiation caused by external or environmental factors. If state distribution corresponding to Schatten decomposition is a goal of differentiation, various non-classical ones will appear in intermediate stages before reaching the goal. Below, we model this mechanism as proposed in [6]. This model mathematically explains what state distribution may occur in the differentiation process. Our aim in this section is to evaluate state structural features by using the state entropy defined in Section 5.

Let us consider a typical state transition caused by a quantum measurement, which is denoted by

$$\Psi \rightarrow \{\psi_k, P_k\}.$$

Ψ means an initial state of system represented by $|\Psi\rangle \langle \Psi|$, and $\{\psi_k, P_k\}$ means a distribution where the states $\{|\psi_k\rangle \langle \psi_k|\}$ exist with probabilities $\{P_k\}$. $\{|\psi_k\rangle\}$ correspond to eigenstates of some physical quantity defined in Hilbert space $\mathcal{H} = \mathbb{C}^M$, and the initial vector $|\Psi\rangle$ is expanded as

$$|\Psi\rangle = \sum_{k=1}^M \sqrt{P_k} |\psi_k\rangle,$$

where $\sqrt{P_k}$ means a complex number satisfying $|\sqrt{P_k}|^2 = P_k$, that is,

$$|\Psi\rangle\langle\Psi| = \sum_{k=1}^M P_k |\psi_k\rangle\langle\psi_k| + \sum_{k \neq k'} \sqrt{P_k} \sqrt{P_{k'}}^* |\psi_k\rangle\langle\psi_{k'}|. \quad (19)$$

The first term $\sum_{k=1}^M P_k |\psi_k\rangle\langle\psi_k|$ corresponds to the distribution $\{\psi_k, P_k\}$, and, therefore, vanishing of the second term, the process called “decoherence” in quantum theory, means accomplishment of the measurement. The relation of Ψ and $\{\psi_k, P_k\}$ is represented as

$$\sum_{k=1}^M M_k |\Psi\rangle\langle\Psi| M_k^* = \sum_{k=1}^M |\langle\psi_k|\Psi\rangle|^2 |\psi_k\rangle\langle\psi_k| = \sum_{k=1}^M P_k |\psi_k\rangle\langle\psi_k|, \quad (20)$$

with the use of projection operator $M_k = |\psi_k\rangle\langle\psi_k|$. (The transition probability $|\langle\psi_k|\Psi\rangle|^2$ is equal to P_k .)

If the above transition is interpreted as a sort of differentiation, its development, i.e., what state distributions occur between Ψ and $\{\psi_k, P_k\}$, becomes a crucial concern. The model of differentiation, which was proposed in [6], presents the picture that the initial state Ψ is differentiated to $\{\psi_k, P_k\}$ step by step through many state transitions. Each state transition is described with use of a map from state to state, which is denoted by Λ^* . The map is called “quantum channel” in quantum information theory. A chain of state transitions given as

$$\rho(0) = |\Psi\rangle\langle\Psi| \rightarrow \rho(1) = \Lambda^*(\rho(0)) \rightarrow \rho(2) = \Lambda^*(\rho(1)) \rightarrow \cdots \rightarrow \rho(n) = \Lambda^*(\rho(n-1)),$$

is regarded as a process of differentiation, if

$$\lim_{n \rightarrow \infty} \rho(n) = \sum_{k=1}^M P_k |\psi_k\rangle\langle\psi_k|, \quad (21)$$

is satisfied. A channel Λ^* is to be defined based on the following: There exist numerous environmental elements around the system. Initially, states of system and these elements are given independently. Let $|\Phi\rangle\langle\Phi|$ be the initial state of one element, which is defined on a space $\mathcal{K}_1 = \mathbb{C}^L$. The initial compound state of the system and the element, on the space $\mathcal{H} \otimes \mathcal{K}_1$, is factorized

$$|\Psi\rangle\langle\Psi| \otimes |\Phi\rangle\langle\Phi|.$$

At the next step, the states of the system and the element become non-separable. Such compound state is generally defined as

$$U |\Psi\rangle\langle\Psi| \otimes |\Phi\rangle\langle\Phi| U^*,$$

using a unitary operator U on $\mathcal{H} \otimes \mathcal{K}_1$. The unitary transformation U specifies a correlation generated between the system and the element, and, in the modeling, the following form is assumed:

$$U = \sum_{k=1}^M |\psi_k\rangle\langle\psi_k| \otimes u_k. \quad (22)$$

where u_k is a unitary on \mathcal{K}_1 . Actually, by this U , the vector $|\Psi\rangle \otimes |\Phi\rangle$ is transformed to

$$U |\Psi\rangle \otimes |\Phi\rangle = \sum_{k=1}^M \sqrt{P_k} |\psi_k\rangle \otimes |\Phi_k\rangle,$$

where $|\Phi_k\rangle = u_k |\Phi\rangle$. Then, the states of the system and the element are “entangled”, since the above form cannot be factorized into two vectors independently defined on \mathcal{H} and \mathcal{K}_1 , if $|\Phi_k\rangle \neq |\Phi_{k'}\rangle$ for some $k \neq k'$. A compound state at the third step is described as

$$\sum_{j=1}^L (I \otimes \bar{M}_j) U |\Psi\rangle \langle \Psi| \otimes |\Phi\rangle \langle \Phi| U^* (I \otimes \bar{M}_j)^*. \tag{23}$$

$\{\bar{M}_j\}$ are projection operators corresponding to the basis set of $\mathcal{K}_1 = \mathbb{C}^L$, say $\{|\phi_j\rangle\}$. As can be seen from Equation (20), a state transition given by a projection operator mathematically represents accomplishment of differentiation. The operation of $\{\bar{M}_j\}$ means that the state of the element is eventually differentiated into $\{|\phi_j\rangle \langle \phi_j|\}$. Note that the states of the system and the element are correlated at the second step. Thus, the state of the system is affected by the differentiation. Actually, Equation (23) may be rewritten to

$$\sum_{j=1}^L E_j |\Psi\rangle \langle \Psi| E_j^* \otimes |\phi_j\rangle \langle \phi_j|, \tag{24}$$

by introducing the operator,

$$E_j = \sum_{k=1}^M \langle \phi_j | \Phi_k \rangle |\psi_k\rangle \langle \psi_k| = \sum_{k=1}^M v_{jk} |\psi_k\rangle \langle \psi_k|. \tag{25}$$

The above form implies that the state of the element of the environment gets transformed to $|\phi_j\rangle \langle \phi_j|$ with probability,

$$P_j = \text{Tr}(E_j |\Psi\rangle \langle \Psi| E_j^* \otimes |\phi_j\rangle \langle \phi_j|) = \langle \Psi | E_j^* E_j | \Psi \rangle, \tag{26}$$

and at the same time the state of the system transits to

$$|\Psi_j\rangle \langle \Psi_j| = \frac{1}{P_j} E_j |\Psi\rangle \langle \Psi| E_j^*. \tag{27}$$

The operator E_j introduced in Equation (25) is called Kraus operator and satisfies $\sum_{j=1}^L E_j^* E_j = I$. (In general, a set of Hermitian positive operators $\{F_i\}$ with $\sum_{i=1}^N F_i = I$ is called positive-operator valued measure (POVM).) With the use of $\{E_j\}$, a quantum channel Λ^* is defined:

$$\Lambda^*(\cdot) = \sum_{j=1}^L E_j \cdot E_j^*. \tag{28}$$

$\Lambda^*(|\Psi\rangle \langle \Psi|) = \sum_{j=1}^L E_j |\Psi\rangle \langle \Psi| E_j^*$ means the density operator obtained from the partial trace of the compound state, $\text{Tr}_{\mathcal{K}}(\sum_{j=1}^L E_j |\Psi\rangle \langle \Psi| E_j^* \otimes |\phi_j\rangle \langle \phi_j|)$.

The other environmental elements are defined in Hilbert spaces denoted by $\mathcal{K}_{2,3,\dots}$. If they interact with the system in a similar way,

$$\rho(n) = \Lambda^*(\rho(n-1)) = \dots = \Lambda^*(\Lambda^*(\dots \Lambda^*(|\Psi\rangle \langle \Psi|) \dots)),$$

is defined as the density operator of the system that is obtained after interacting with n environmental elements. From the definition of Λ^* (see Equation (28)), this $\rho(n)$ is decomposed as

$$\rho(n) = \sum_{\{j_1, j_2, \dots, j_n\}} P_{\{j_1, j_2, \dots, j_n\}} \left| \Psi_{\{j_1, j_2, \dots, j_n\}} \right\rangle \left\langle \Psi_{\{j_1, j_2, \dots, j_n\}} \right|, \tag{29}$$

where

$$P_{\{j_1, j_2, \dots, j_n\}} = \langle \Psi | \mathbf{E}_{\{j_1, j_2, \dots, j_n\}}^* \mathbf{E}_{\{j_1, j_2, \dots, j_n\}} | \Psi \rangle, \tag{30}$$

and

$$\left| \Psi_{\{j_1, j_2, \dots, j_n\}} \right\rangle = \frac{1}{P_{\{j_1, j_2, \dots, j_n\}}} \mathbf{E}_{\{j_1, j_2, \dots, j_n\}} | \Psi \rangle. \tag{31}$$

(The notation $\mathbf{E}_{\{j_1, j_2, \dots, j_n\}}$ means $E_{j_n} \cdots E_{j_2} E_{j_1}$.) $P_{\{j_1, j_2, \dots, j_n\}}$ is the probability that the states of n environmental elements eventually become $\{\phi_{j_1}, \phi_{j_2}, \dots, \phi_{j_n}\}$, and $|\Psi_{\{j_1, j_2, \dots, j_n\}}\rangle \langle \Psi_{\{j_1, j_2, \dots, j_n\}}|$ is a pure state of the system at this event. It should be noted here that the density operator $\rho(n)$ can be expanded as

$$\rho(n) = \sum_{k=1}^M P_k |\psi_k\rangle \langle \psi_k| + \sum_{k \neq k'} \sqrt{P_k} \sqrt{P_{k'}}^* (\langle \Phi_k | \Phi_{k'} \rangle)^n |\psi_k\rangle \langle \psi_{k'}|, \tag{32}$$

by using Equation (19) and the property of

$$\Lambda^*(|\psi_k\rangle \langle \psi_{k'}|) = \sum_{j=1}^L \langle \Phi_k | \psi_j \rangle \langle \psi_j | \Phi_{k'} \rangle |\psi_k\rangle \langle \psi_{k'}| = \langle \Phi_k | \Phi_{k'} \rangle |\psi_k\rangle \langle \psi_{k'}|.$$

Since $|\langle \Phi_k | \Phi_{k'} \rangle| < 1$, the condition of Equation (21), i.e., $\lim_{n \rightarrow \infty} \rho(n) = \sum_{k=1}^M P_k |\psi_k\rangle \langle \psi_k|$ is clearly satisfied. Thus, the state distribution $\{\Psi_{\{j_1, j_2, \dots, j_n\}}, P_{\{j_1, j_2, \dots, j_n\}}\}$, which is encoded in $\rho(n)$, is identified at an intermediate stage in the differentiation process $\Psi \rightarrow \{\psi_k, P_k\}$.

Figure 1 shows the result of computational simulation with $M = L = 2$, $|\Psi\rangle = \sqrt{0.7}|\psi_1\rangle + \sqrt{0.3}|\psi_2\rangle$ ($P_1 = 0.7$, $P_2 = 0.3$), $\nu_{1|1} = \sqrt{0.5}$ ($\nu_{2|1} = \sqrt{0.5}$) and $\nu_{1|2} = \sqrt{0.45}$ ($\nu_{2|2} = \sqrt{0.55}$). The histograms of population rates of states with $\frac{l-1}{20} < |\langle \psi_1 | \Psi_{\{j_1, j_2, \dots, j_n\}} \rangle|^2 \leq \frac{l}{20}$ ($l = 1, 2, \dots, 20$) are calculated in the case of $n = 0, 10, 100, 500$ and 2000 . One can see that with increasing n , the state distribution approaches the goal of differentiation, i.e., $\{\{\psi_1, \psi_2\}, \{0.7, 0.3\}\}$.

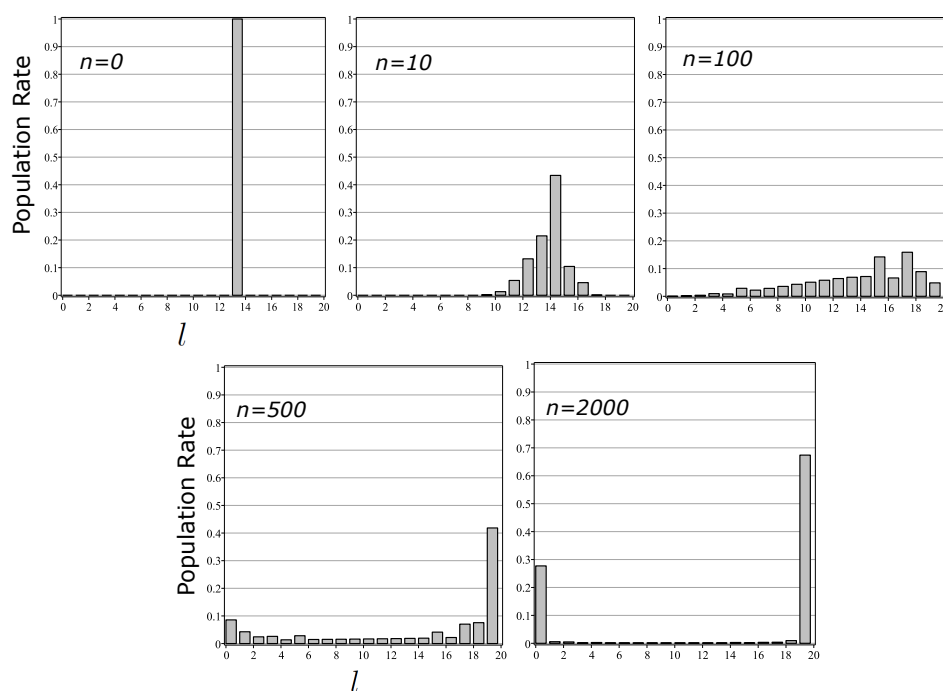


Figure 1. Histograms of population rates of states with $\frac{l-1}{20} < |\langle \psi_1 | \Psi_{\{i_1, i_2, \dots, i_n\}} \rangle|^2 \leq \frac{l}{20}$ ($l = 1, 2, \dots, 20$) in the case of $n = 0, 10, 100, 500$ and 2000 . The parameters are set by $M = L = 2$, $|\Psi\rangle = \sqrt{0.7}|\psi_1\rangle + \sqrt{0.3}|\psi_2\rangle$ ($P_1 = 0.7$, $P_2 = 0.3$), $\nu_{1|1} = \sqrt{0.5}$ ($\nu_{2|1} = \sqrt{0.5}$) and $\nu_{1|2} = \sqrt{0.45}$ ($\nu_{2|2} = \sqrt{0.55}$). If $|\Psi_{\{i_1, i_2, \dots, i_n\}}\rangle \approx |\psi_1\rangle$ ($|\psi_2\rangle$), $|\langle \psi_1 | \Psi_{\{i_1, i_2, \dots, i_n\}} \rangle|^2$ takes a value nearby 1(0). With increasing n , the state distribution approaches to $\{\{\psi_1, \psi_2\}, \{0.7, 0.3\}\}$.

Figure 2 shows the behavior of the state entropy D , von Neumann entropy and $-\log(\text{Tr}(\rho^2))$ for the distribution $\{\Psi_{\{i_1, i_2, \dots, i_n\}}, P_{\{i_1, i_2, \dots, i_n\}}\}$, which are calculated in the same setting of parameters. One can directly see that the inequality of Equation (18) is satisfied at any n . Note that the state

entropy D takes values close to von Neumann entropy at very large n . In fact, as shown in Figure 3, the difference between von Neumann entropy and the state entropy is noticeable mostly at earlier stages. These results imply that state distributions appearing in the differentiation process are non-classical in general.

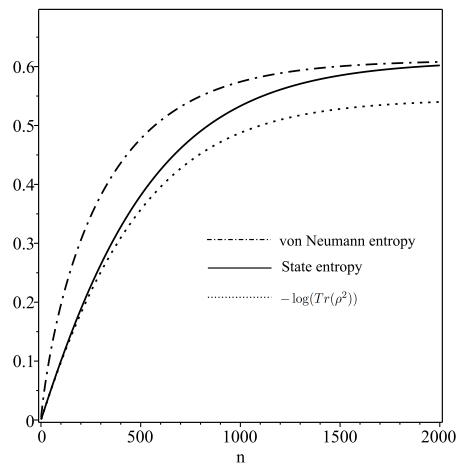


Figure 2. Behaviors of state entropy, von Neumann entropy and $-\log(\text{Tr}(\rho^2))$ at the parameters of $M = L = 2$, $|\Psi\rangle = \sqrt{0.7}|\psi_1\rangle + \sqrt{0.3}|\psi_2\rangle$ ($P_1 = 0.7$, $P_2 = 0.3$), $\nu_{1|1} = \sqrt{0.5}$ ($\nu_{2|1} = \sqrt{0.5}$) and $\nu_{1|2} = \sqrt{0.45}$ ($\nu_{2|2} = \sqrt{0.55}$)

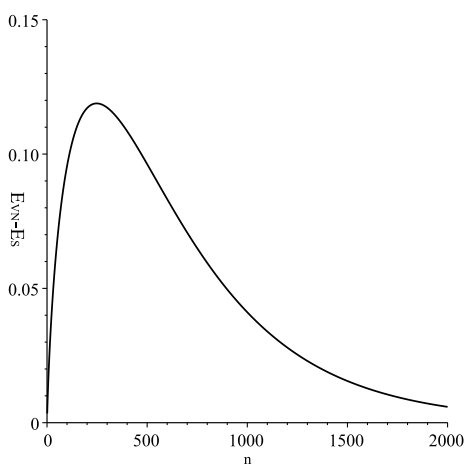


Figure 3. Difference between von Neumann entropy and state entropy.

7. Conclusions

The state entropy is a truly non-classical quantity because it depends not only on statistical probabilities, but also on similarities among states. The differentiation phenomenon is also non-classical, because it is interpreted as dynamics of the probabilities and similarities. Definition of the state entropy and modeling of the differentiation process are impossible in the framework of classical probability theory.

We believe that evaluation of an ensemble of systems by the state entropy fits the empirical reasoning: No matter how many systems are in the ensemble, we may not recognize high diversity if we know that these states are not very different. Further, we believe that various areas of the nature dynamics of character change in the population of individuals is very much like the differentiation phenomena. This makes prospects of the quantum-like formalism grow stronger.

Author Contributions: Conceptualization, M.A., I.B. and A.K.; Methodology, M.A. and A.K.; Validation, I.B., E.M.P. and A.K.; Writing-Original Draft Preparation, M.A.; Writing-Review & Editing, I.B., E.M.P. and A.K.

Acknowledgments: I.B. was supported by Marie Curie Fellowship at City University of London, H2020-MSA-IF-2015, grant N 696331.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. State Entropy and other Quantum Entropies

In this paper, we propose the state entropy,

$$D(\{\langle \rho \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i (-\log(\text{Tr}(\rho |\Psi_i\rangle \langle \Psi_i|))).$$

More generally, it is defined as

$$D(\{\langle \rho \rangle_{\Psi_i}, P_i\}) = \sum_{i=1}^N P_i f(\text{Tr}(\rho |\Psi_i\rangle \langle \Psi_i|)),$$

by using a convex and monotone decreasing function $f(x)$. Mathematically, $D(\{\langle \rho \rangle_{\Psi_i}, P_i\})$ depends on the way of decomposition of ρ . As discussed in Section 5, for $\rho = \sum_{i=1}^N P_i |\Psi_i\rangle \langle \Psi_i|$, $f(\text{Tr}(\rho |\Psi_i\rangle \langle \Psi_i|)) = f(\langle \rho \rangle_{\Psi_i})$ is interpreted as the degree of orthogonality between $|\Psi_i\rangle \langle \Psi_i|$ and ρ . In this sense, D evaluates a sort of diversity in the state distribution $\{\Psi_i, P_i\}$, and it takes the maximal value equivalent to $-\text{Tr}(\rho \log \rho)$ for the Schatten decomposition (see Equation (18)).

On the other hand, there are many mathematical expansions of von Neumann entropy. As examples, the quantum version of Rényi entropy [51],

$$R_\alpha(\rho) = \frac{1}{1-\alpha} \log(\text{Tr}(\rho^\alpha)), \tag{A1}$$

and the one of Tsallis entropy [52],

$$T_\alpha(\rho) = \frac{1}{1-\alpha} (\text{Tr}(\rho^\alpha) - 1), \tag{A2}$$

are well-known. These entropies approach to von Neumann entropy $S(\rho) = -\text{Tr}(\rho \log \rho)$ in the limit $\alpha \rightarrow 1$. The index α , which is called the entropic parameter, is nonnegative and $\alpha \neq 1$. Further, such generalized entropies are uniformly represented in the form of quantum version of Salicrú entropy [53,54], which is given by

$$H_{(h,\phi)}(\rho) = h(\text{Tr}\phi(\rho)), \tag{A3}$$

where the functions $h : \mathbb{R} \rightarrow \mathbb{R}$ and $\phi : [0, 1] \rightarrow \mathbb{R}$ satisfy either of the following conditions: (i) h is increasing and ϕ is concave; or (ii) h is decreasing and ϕ is convex. In the form of Rényi entropy, $h(x) = \frac{\log(x)}{1-\alpha}$ and $\phi(x) = x^\alpha$, and in the form of Tsallis entropy, $h(x) = \frac{x-1}{1-\alpha}$ and $\phi(x) = x^\alpha$. Of course, von Neumann entropy is also recovered at $h(x) = x$ and $\phi(x) = -x \log x$.

Here, we have to point out that $H_{(h,\phi)}(\rho)$ is practically calculated as

$$H_{(h,\phi)}(\rho) = h\left(\sum_{k=1}^M \phi(\lambda_k)\right),$$

by using the eigenvalues of ρ , that is, any quantum entropic measure that is reduced into $H_{(h,\phi)}(\rho)$ does not depend on the way of decomposition of ρ . The state entropy is different in this point. Actually, it is clear that $D(\{\langle \rho \rangle_{\Psi_i}, P_i\})$ is not recovered in the form of $H_{(h,\phi)}(\rho)$.

References

1. Von Neumann, J. Thermodynamik quantummechanischer Gesamtheiten. *Gott. Nach.* **1927**, *1*, 273–291.
2. Von Neumann, J. *Mathematische Grundlagen der Quantenmechanik*; Springer: Berlin, Germany, 1932.
3. Horodecki, M.; Oppenheim, J.; Winter, A. Partial quantum information. *Nature* **2005**, *436*, 673–676.
4. Umegaki, H. Conditional expectations in an operator algebra IV (entropy and information). *Kodai Math. Sem. Rep.* **1962**, *14*, 59–85.
5. Ohya, M. Fundamentals of Quantum Mutual Entropy and Capacity. *Open Syst. Inf. Dyn.* **1999**, *6*, 69–78.
6. Asano, M.; Basieva, I.; Khrennikov, A.; Yamato, I. A model of differentiation in quantum bioinformatics. *Prog. Biophys. Mol. Biol.* **2017**, *130*, 88–98.
7. Zurek, W.H. Decoherence and the Transition from Quantum to Classical. *Phys. Today* **1991**, *44*, 36–44.
8. Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **1976**, *48*, 119.
9. Gorini, V.; Kossakowski, A.; Sudarshan, E.C.G. Completely positive semigroups of N-level systems. *J. Math. Phys.* **1976**, *17*, 821.
10. Khrennikov, A. Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social and anomalous phenomena. *Found. Phys.* **1999**, *29*, 1065–1098.
11. Khrennikov, A. Quantum-like formalism for cognitive measurements. *Biosystems* **2003**, *70*, 211–233.
12. Khrennikov, A. On quantum-like probabilistic structure of mental information. *Open Syst. Inf. Dyn.* **2014**, *11*, 267–275.
13. Khrennikov, A. *Information Dynamics in Cognitive, Psychological, Social, and Anomalous Phenomena*; Ser.: Fundamental Theories of Physics; Kluwer: Dordrecht, The Netherlands, 2004.
14. Busemeyer, J.B.; Wang, Z.; Townsend, J.T. Quantum dynamics of human decision making. *J. Math. Psychol.* **2006**, *50*, 220–241.
15. Haven, E. Private information and the ‘information function’: A survey of possible uses. *Theory Decis.* **2008**, *64*, 193–228.
16. Yukalov, V.I.; Sornette, D. Processing Information in Quantum Decision Theory. *Entropy* **2009**, *11*, 1073–1120.
17. Khrennikov, A. *Ubiquitous Quantum Structure: From Psychology to Finances*; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2010.
18. Asano, M.; Masanori, O.; Tanaka, Y.; Khrennikov, A.; Basieva, I. Quantum-like model of brain’s functioning: Decision making from decoherence. *J. Theor. Biol.* **2011**, *281*, 56–64.
19. Busemeyer, J.R.; Pothos, E.M.; Franco, R.; Trueblood, J. A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* **2011**, *118*, 193–218.
20. Asano, M.; Ohya, M.; Khrennikov, A. Quantum-Like Model for Decision Making Process in Two Players Game - A Non-Kolmogorovian Model. *Found. Phys.* **2011**, *41*, 538–548.
21. Asano, M.; Ohya, M.; Tanaka, Y.; Khrennikov, A.; Basieva, I. Dynamics of entropy in quantum-like model of decision making. *AIP Conf. Proc.* **2011**, *63*, 1327.
22. Bagarello, F. *Quantum Dynamics for Classical Systems: With Applications of the Number Operator*; Wiley: New York, NY, USA, 2012; Volume 90, p. 015203.
23. Busemeyer, J.R.; Bruza, P.D. *Quantum Models of Cognition and Decision*; Cambridge Press: Cambridge, UK, 2012.
24. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y. Quantum-like dynamics of decision-making. *Phys. A Stat. Mech. Appl.* **2010**, *391*, 2083–2099.
25. De Barros, A.J. Quantum-like model of behavioral response computation using neural oscillators. *Biosystems* **2012**, *110*, 171–182.
26. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y. Quantum-like generalization of the Bayesian updating scheme for objective and subjective mental uncertainties. *J. Math. Psychol.* **2012**, *56*, 166–175.
27. De Barros, A.J.; Oas, G. Negative probabilities and counter-factual reasoning in quantum cognition. *Phys. Scr.* **2014**, *T163*, 014008.
28. Wang, Z.; Busemeyer, J.R. A quantum question order model supported by empirical tests of an a priori and precise prediction. *Top. Cogn. Sci.* **2013**, *5*, 689–710.
29. Dzharfarov, E.N.; Kujala, J.V. On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Top. Cogn. Sci.* **2014**, *6*, 121–128.

30. Wang, Z.; Solloway, T.; Shiffrin, R.M.; Busemeyer, J.R. Context effects produced by question orders reveal quantum nature of human judgments. *PNAS* **2014**, *111*, 9431–9436.
31. Khrennikov, A. Quantum-like modeling of cognition. *Front. Phys.* **2015**, *3*, 77.
32. Boyer-Kassem, T.; Duchene, S.; Guerci, E. Testing quantum-like models of judgment for question order effect. *Math. Soc. Sci.*, 2016, *80*, 33–46.
33. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y. A Quantum-like Model of Selection Behavior. *J. Math. Psychol.* **2016**, doi:10.1016/j.jmp.2016.07.006
34. Yukalov, V.I.; Sornette, D. Quantum Probabilities as Behavioral Probabilities. *Entropy* **2017**, *19*, 112.
35. Igamberdiev, A.U.; Shklovskiy-Kordi, N.E. The quantum basis of spatiotemporality in perception and consciousness. *Prog. Biophys. Mol. Biol.* **2017**, *130*, 15–25.
36. De Barros, J.A.; Holik, F.; Krause, D. Contextuality and indistinguishability. *Entropy* **2017**, *19*, 435.
37. Bagarello, F.; Di Salvo, R.; Gargano, F.; Oliveri, F. (H, ρ) -induced dynamics and the quantum game of life. *Appl. Math. Mod.* **2017**, *43*, 15–32.
38. Takahashi, K.S.-J.; Makoto, N. A note on the roles of quantum and mechanical models in social biophysics. *Prog. Biophys. Mol. Biol.* **2017**, *130 (Pt A)*, 103–105.
39. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato, I. Quantum-like model of diauxie in *Escherichia coli*: Operational description of precultivation effect. *J. Theor. Biol.* **2012**, *314*, 130–137.
40. Accardi, L.; Ohya, M. Compound channels, transition expectations, and liftings. *Appl. Math. Optim.* **1999**, *39*, 33–59.
41. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato, I. A model of epigenetic evolution based on theory of open quantum systems. *Syst. Synth. Biol.* **2013**, *7*, 161.
42. Asano, M.; Hashimoto, T.; Khrennikov, A.; Ohya, M.; Tanaka, A. Violation of contextual generalization of the Leggett-Garg inequality for recognition of ambiguous figures. *Phys. Scr.* **2014**, *2014*, T163
43. Asano, M.; Basieva, I.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato, I. Quantum Information Biology: From Information Interpretation of Quantum Mechanics to Applications in Molecular Biology and Cognitive Psychology. *Found. Phys.* **2015**, *45*, 1362.
44. Asano, M.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato, I. Three-body system metaphor for the two-slit experiment and *Escherichia coli* lactose-glucose metabolism. *Philos. Trans. R. Soc. A* **2016**, doi:10.1098/rsta.2015.0243.
45. Ohya, M.; Volovich, I. *Mathematical Foundations of Quantum Information and Computation and its Applications to Nano- and Bio-Systems*; Springer: Berlin, Germany, 2011.
46. Asano, M.; Khrennikov, A.; Ohya, M.; Tanaka, Y.; Yamato, I. *Quantum Adaptivity in Biology: From Genetics to Cognition*; Springer: Berlin, Germany, 2015.
47. Oaksford, M.; Chater, N. A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychol. Rev.* **1994**, *101*, 608–631.
48. Pothos, E.M.; Chater, N. A simplicity principle in unsupervised human categorization. *Cogn. Sci.* **2002**, *26*, 303–343.
49. Miller, G.A. Free Recall of Redundant Strings of Letters. *J. Exp. Psychol.* **1958**, *56*, 485–491.
50. Jamieson, R.K.; Mewhort, D.J.K. The influence of grammatical, local, and organizational redundancy on implicit learning: An analysis using information theory. *J. Exp. Psychol. Learn. Mem. Cogn.* **2005**, *31*, 9–23.
51. Rényi, A. On measures of entropy and information. In Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20–30 July 1960; Volume 1, p. 547.
52. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479.
53. Salicrú, M.; Menéndez, M.L.; Morales, D.; Pardo, L. Asymptotic distribution of (h, ϕ) -entropies. *Commun. Stat. Theory Methods* **1993**, *22*, 2015.
54. Bosyk, G.M.; Zozor, S.; Holik, F.; Portesi, M.; Lamberti, P.W. A family of generalized quantum entropies: definition and properties. *Quantum Inf. Proc.* **2016**, *15*, 3393–3420.

