Article

Quantum Information as a Non-Kolmogorovian Generalization of Shannon's Theory

Federico Holik [1,2,*], Gustavo M. Bosyk [1,2] and Guido Bellomo [1,2]

[1] Instituto de Física La Plata (IFLP), CONICET, 115 y 49, 1900 La Plata, Argentina

[2] Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 67, 1900 La Plata, Argentina; E-Mails: gbosyk@fisica.unlp.edu.ar (G.M.B.); gbellomo@fisica.unlp.edu.ar (G.B.)

* Author to whom correspondence should be addressed; E-Mail: holik@fisica.unlp.edu.ar.

Abstract: In this article, we discuss the formal structure of a generalized information theory based on the extension of the probability calculus of Kolmogorov to a (possibly) non-commutative setting. By studying this framework, we argue that quantum information can be considered as a particular case of a huge family of non-commutative extensions of its classical counterpart. In any conceivable information theory, the possibility of dealing with different kinds of information measures plays a key role. Here, we generalize a notion of state spectrum, allowing us to introduce a majorization relation and a new family of generalized entropic measures.

Keywords: quantum probability; generalized majorization; information theory; quantum information

1. Introduction

Quantum information theory is not only interesting because of its promising technological applications, but also because of its impact at the very heart of physics, giving place to a new way of studying quantum mechanics [1] and other possible theories, as well. In particular, it has given rise to a quest for the foundational principles that singularize quantum mechanics among a vast family of possible statistical theories [2–5]. The study and characterization of quantum correlations play a central role in this quest [6], entanglement [7–9] and discord [10] being the most important ones. As is well

known, probabilities and correlations are essential concepts in both classical and quantum information theories. However, it turns out that the probabilities involved are fundamentally different on each of these theories. In this work, we will argue that, due to quantum contextuality and the non-Kolmogorovian nature of the underlying probabilities, quantum information theory can be correctly characterized as a non-commutative version of its classical counterpart.

For a statistical theory, it is very illuminating to look at the geometrical aspects of the set of possible states. This has been done extensively for the quantum case [11–19]. However, the geometry of a set of quantum states differs radically from that of a classical one. While the set of states of classical and quantum systems shares the characteristic of being a convex set, like quantum ones, classical models are simplexes. This difference expresses itself also at the level of the axiomatization. While Kolmogorov's axioms suffice for describing classical probabilistic models, the Boolean structure of a sigma algebra must be generalized to an orthomodular lattice of projection operators for the quantal case [14,15,20–31].

One may wonder if there are probabilistic models more general than quantum and classical ones. This is indeed the case, and we must not go too far from standard quantum mechanics in order to find them. For example, in algebraic relativistic quantum field theory, states may be defined as measures over Type III factors [32], a special kind of von Neumann algebra [33–38], which differs from the Type I factors appearing in standard quantum mechanics [22,39–41]. Type II factors can also be found in algebraic statistical mechanics (quantum mechanics with infinite degrees of freedom) [39,42]. Thus, it is clear that states defining measures that go beyond standard quantum mechanics exist, and they appear in examples of interest for physics. These new measures, which go beyond the distributive (or equivalently, commutative) case of the Boolean sigma algebra, are sometimes called non-Kolmogorovian or non-commutative probabilities. The fact that these non-commutative probabilities are involved is responsible for the emergence of the peculiar features of quantum information theory.

However, then, one could also imagine a setting where more general probabilistic theories can be conceived of in order to study its general features and compare them to already known ones. This approach has been developed by many authors, and it is fair to say that it is based on the study of convex sets of states, which define measures on certain algebras of observables. These are usually called events or, more generally, effects (see, for example, [43–53]). The origins of this approach could be traced to the works of G. Ludwig [12,13] and G. Mackey [54], but also to von Neumann. See also [39,54–57] for other axiomatizations of non-Kolmogorovian probabilities and their relationships with lattice theory. Non-linear generalizations of quantum mechanics were studied using a similar approach in [16–18].

It is possible to study many important notions of information theory, such as entanglement, discord and many information protocols in generalized probabilistic models (see, for example, [45–53,58]). We will argue in favor of the existence of a generalized information theory, continuing the lines of previous works [49,50,59] (see also [60,61], where non-commutative versions of many statistical techniques are studied). By focusing on the study of the formal aspects of the probabilities involved in different models, we show that the non-Kolmogorovian character of the probabilities underlying the quantum formalism is responsible for the emergence of quantum information theory (see [62] for a discussion in which the role played by the formal structure of quantum probabilities in the beginnings of quantum information theory is stressed). This allows us to claim: Kolmogorovian probabilities imply

Shannon's information theory; the non-commutative probability calculus of quantum theory, implies quantum information theory. Quantum and classical information theories appear as particular instances of a formalism based on generalized probabilistic measures.

Any information theory depends strongly on our capability of dealing with different information measures. This is the case in classical information theory, where the Shannon [63], Tsallis [64] and Rényi [65] entropies (among other measures) are used for different purposes. A similar diversity of measures should be available in the generalized probabilistic setting. Previous works have focused on some entropic measures in the setting of generalized probabilities [49,66,67]. In this paper, we extend a new family of entropies based on the (h, ϕ)-entropies to the general probabilistic setting [68,69]. These measures include the previous ones studied in the literature as particular cases. Other important notions introduced in this article are the definition of the generalized spectrum for states in general models and the relationship of the generalized majorization between states. These are shown to be useful for defining functions of states and studying the properties of the entropic measures.

The paper is organized as follows. In Section 2, we review classical probabilities in the Kolmogorov approach. Next, we turn to important aspects of the quantum formalism and the formal structure of probability measures in quantum mechanics in Section 3. In Section 4, we discuss the formal aspects of a generalized information theory in the operational approach. In Section 5, we introduce our new family of information measures and the notion of the *generalized spectrum*, which allow us to introduce the concept of the *generalized majorization*. Finally, in Section 6, we draw our conclusions.

2. Classical Probabilities

One of the most used axiomatizations of classical probability theory is the one of A.N. Kolmogorov [70]. If the possible outcomes of an experiment are represented by a set Ω, subsets of it can be considered as representing events. It is usual to restrict events to a σ-algebra Σ of subsets of Ω. Thus, Kolmogorov defines probability measures as functions μ, such that:

$$\mu : \Sigma \to [0, 1] \tag{1a}$$

satisfying:

$$\mu(\Omega) = 1 \tag{1b}$$

and, for any pairwise disjoint denumerable family $\{A_i\}_{i \in I}$,

$$\mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i) \tag{1c}$$

In this way, Kolmogorov's approach puts probability theory in a direct connection with measure theory. From this axiomatic, it is straightforward to see that $\mu(\emptyset) = 0$ and $\mu(A^c) = 1 - \mu(A)$, where $(\cdot)^c$ means the set-theoretical complement.

There exist many approaches to classical probabilities (see, e.g., [71] for a complete review). This subject is too vast to cover here and goes beyond the scope of this work. We only mention the Bayesian school because of its importance and many physical applications [72–74] (see also [75]).

3. Quantum Probabilities

In this section, we will discuss the special features of the probabilities involved in quantum theory. The most salient feature is that, like the classical case, the algebra of events of a quantum system is non-Boolean. This is related to the complementarity principle, for which incompatible experiments are needed to fully describe quantum phenomena.

3.1. Elementary Tests in Quantum Mechanics

Propositions, such as "the value of the energy lies in the interval (a, b)" or "the particle is detected between the interval (a, b)", are examples of how the results of experiments can be expressed in quantum mechanics. Elementary propositions of that form are usually called events, and they are represented by projection operators as follows. A projective valued measure (PVM) is a map M, such that:

$$M : B(\mathbb{R}) \to \mathcal{P}(\mathcal{H}) \tag{2a}$$

where $B(\mathbb{R})$ is any Borel set on $\mathbb{R}$ and $\mathcal{P}(\mathcal{H})$ is the space of projections on a Hilbert space $\mathcal{H}$, satisfying:

$$M(\emptyset) = \mathbf{0} \text{ and } M(\mathbb{R}) = \mathbf{1} \tag{2b}$$

where $\mathbf{0}$ is the null space and $\mathbf{1}$ the identity operator, and:

$$M(\bigcup_{i \in I} B_i) = \sum_{i \in I} M(B_i) \tag{2c}$$

for any disjoint denumerable family $\{B_i\}_{i \in I}$. As in the classical case, from this axiomatic, it results that $M(B^c) = \mathbf{1} - M(B) = M(B)^\perp$ (where $(\cdot)^\perp$ stands for orthogonal complement).

The spectral theorem allows one to assign a PVM to any self-adjoint operator representing a physical observable O [21,22]. We denote by M_O its corresponding PVM. Thus, for any Borel set $(a, b) \in \mathbb{R}$ representing an interval of possible values of O, $M_O((a, b)) = \mathbf{P}_{(a,b)}$ is a projection operator that represents the elementary event "the value of O lies in the interval (a, b)".

The state of a quantum mechanical system is represented by a density operator ρ, which is semi-definite positive and of trace one [76]. Given ρ, the probability that the event represented by $\mathbf{P}_{(a,b)}$ occurs is given by Born's rule:

$$p(\mathbf{P}_{(a,b)}; \rho) = \text{Tr}(\rho \mathbf{P}_{(a,b)}) \tag{3}$$

A generalization of the above mechanism for computing probabilities is given by the notion of quantal effects and positive operator-valued measures (POVM) [77–83]. In quantum mechanics, a POVM is represented by a mapping:

$$E : B(\mathbb{R}) \to \mathcal{B}(\mathcal{H}) \tag{4a}$$

where $\mathcal{B}(\mathcal{H})$ stands for bounded operator, such that:

$$E(\mathbb{R}) = \mathbf{1} \tag{4b}$$

$$E(B) \geq 0, \text{ for all } B \in B(\mathbb{R}) \tag{4c}$$

$$E(\bigcup_{i \in I} B_i) = \sum_{i \in I} E(B_i), \text{ for any disjoint family } \{B_i\}_{i \in I} \tag{4d}$$

Then, the probability of effect **E** given that the system is prepared in state ρ is given by:

$$p(\mathbf{E}; \rho) = \mathrm{Tr}(\rho\mathbf{E}) \tag{5}$$

3.2. von Neumann's Axioms

Is there an analog of Kolmogorov's axioms in quantum theory? As we have seen, events of a classical probabilistic theory can be represented as subsets of a given outcome set, yielding a Boolean σ-algebra. Consequently, classical states can be considered as measures over Boolean algebras. However, as we have seen, the complementarity principle forces the non-commutativity of certain observables. This makes the algebra of projection operators (*i.e.*, the algebra of possible events) non-distributive and, thus, non-Boolean. In this way, quantum states can be characterized as measures over non-Boolean algebras as follows:

$$s : \mathcal{P}(\mathcal{H}) \to [0, 1] \tag{6a}$$

such that:

$$s(\mathbf{1}) = 1 \tag{6b}$$

and, for a denumerable and pairwise orthogonal family of projections $\{\mathbf{P}_i\}_{i \in I}$,

$$s(\sum_{i \in I} \mathbf{P}_i) = \sum_{i \in I} s(\mathbf{P}_i) \tag{6c}$$

We will refer to the above axioms as Kolmogorov's axioms. Gleason's theorem [84] asserts that the family of measures obeying von Neumann's axioms is in bijective correspondence to the set of positive trace class operators of trace one, which is nothing but the set of all possible quantum states. Thus, von Neumann's axioms relate quantum states with the non-Boolean (or non-commutative) measure theory defined by Equations (6a)–(6c). As remarked in the Introduction, this fact lies behind the distinctive features of quantum information theory. Another important remark is that both the collection of all possible measures obeying von Neumann's axioms and the ones obeying Kolmogorov's form convex sets. This geometrical feature can be endowed with a natural physical interpretation: given two probability distributions, one can always form a mixture of them (and this will be represented mathematically by the corresponding convex combination in the state space).

3.3. Quantum Correlations

The non-abelian character of the quantum algebraic setting gives rise to a variety of new possibilities regarding correlations. So far, the most important of these novel quantum features has been the so-called entanglement. First recognized by Schrödinger and Einstein, Podolsky and Rosen in 1935, entanglement had remained at the center of debate, inspiring discussions around the completeness of the formalism, the reality and locality of the theory or, more recently, about its status as a resource for quantum information processing tasks (see, e.g., [85] for a complete review).

In the bipartite scenario, a quantum state is said to be non-entangled if and only if it can be approximated by convex linear combinations of product states. As Werner put it in his 1989 seminal

paper, given a joint AB-bipartite state ρ, the state is separable if there exist a probability distribution $\{p_k\}$ and marginal states $\{\rho_k^A\}$, $\{\rho_k^B\}$, such that [9,86]:

$$\rho = \sum_k p_k \rho_k^A \otimes \rho_k^B \tag{7}$$

Then, ρ is entangled if it is not separable. This definition can be rephrased in more general algebraic terms. Let $\mathcal{N}_A$ and $\mathcal{N}_B$ be von Neumann algebras acting on a common Hilbert space, associated with the A and B subsystems. A state $\omega_\rho : \mathcal{N} \to \mathbb{C}$ is an expectation value functional, where $\omega_\rho(n) = \text{Tr}(n\rho)$ for any observable $n \in \mathcal{N}$. Then, ω_ρ on $\mathcal{N}_A \vee \mathcal{N}_B$ (the smallest von Neumann algebra generated by $\mathcal{N}_A$ and $\mathcal{N}_B$) is a product state with respect to $\mathcal{N}_A$ and $\mathcal{N}_B$ iff $\omega_\rho(ab) = \omega_\rho(a)\omega_\rho(b)$ for any $a \in \mathcal{N}_A$ and any $b \in \mathcal{N}_B$. If a and b are projectors, ω_ρ's being a product state implies that the probability of measuring ab factorizes, the usual criterion for uncorrelation. Moreover, the state ω_ρ on $\mathcal{N}_A \vee \mathcal{N}_B$ is separable with respect to $\mathcal{N}_A$ and $\mathcal{N}_B$ iff it can be approximated by convex linear combinations of product states. Else, it is entangled.

As claimed before, the non-abelian nature of $\mathcal{N}_A$ ($\mathcal{N}_B$) is essential here. No entanglement is possible if the algebras are generated only by commutative observables [86,87]. In other words: probabilities must be non-Kolmogorovian as a condition of the possibility for true entanglement. This fact has important consequences for quantum information processing, because entanglement plays a key role in the most useful protocols.

The non-commutativity is also responsible for the perturbation of the joint state when measuring over one of its parts. This fact can be quantified by the difference between the pre- and post-measurement mutual information after a local (non-selective) measurement, a quantity known as discord [10,88]. A non-discordant or classically-correlated state ρ is one that can be written as Lemma 8.12 in [89] (see also [90]):

$$\rho = \sum_{ij} p_{ij} \Pi_i^A \otimes \Pi_j^B \tag{8}$$

where $\{\Pi_i^A\}$ ($\{\Pi_j^B\}$) is a basis of orthogonal projectors on the Hilbert space of A (B) and $\{p_{ij}\}$ is the corresponding probability distribution. Notice that the projective measurement given by $\{\Pi_i^A \otimes \Pi_j^B\}$ does not perturb the state Equation (8) and provides maximal information about the joint state. Furthermore, one can define states that are classically correlated with respect to one of the parts only. For example, $\rho = \sum_{ij} p_{ij} \Pi_i^A \otimes \rho_j^B$ would be a classical-quantum state. In the last decade, quantum discord was also identified with the quantum advantage for some informational tasks (see [91] for a complete review). Notice that in order to have non-null discord, non-orthogonal (*i.e.*, incompatible) projections must be involved: this is another way in which the non-Boolean character of the event algebra is expressed.

As we explain below, the notions of entanglement and discord are susceptible to be extended upon general probabilistic theories.

Finally, it is worth noting that there are many other ways to assess the quantum peculiarities. For example, steering, first proposed by Schrödinger [7], and which has recently attracted a lot of attention [92–96], concerns the perturbation of a distant part through the manipulation of local degrees of freedom and is closely related to the notion of non-locality.

4. Generalized Setting

The lattice of projection operators of a separable Hilbert space and that of σ-algebras are special instances of orthomodular lattices [27–29,97,98]. Orthomodular lattices are a suitable framework for describing contextual theories: given an orthomodular lattice $\mathcal{L}$, each possible context will be represented by a maximal Boolean subalgebra. If the maximal Boolean subalgebra coincides with the original lattice, then the theory will be non-contextual. In order to describe theories more general than quantum mechanics, one could generalize the above axioms for probability theory to arbitrary orthomodular lattices as follows. Given $\mathcal{L}$, define a measure ν satisfying:

$$\nu : \ \mathcal{L} \to [0, 1] \tag{9a}$$

such that:

$$\nu(\mathbf{1}) = 1 \tag{9b}$$

and, for a denumerable and pairwise orthogonal family of events $\{E_i\}_{i \in I}$,

$$\nu(\sum_{i \in I} E_i) = \sum_{i \in I} \nu(E_i) \tag{9c}$$

See, e.g., [23] for conditions under which these measures exist. It is important to remark that Equations (1a)–(1c) and Equations (6a)–(6c) are just particular examples of the above axioms. However, these are much more general: in algebraic relativistic quantum field theory and in algebraic statistical mechanics, more general orthomodular lattices appear [32,38,39]. Many of the informational notions that can be described in quantum mechanics can be generalized to this formal setting (see, for example, [67,99,100], where the maximum entropy principle is analyzed). It is also important to mention that other types of non-Kolmogorovian probabilistic theories can be conceived of (we will not deal with them here, but see, for example, [101,102]).

In Section 3.2, we have mentioned that both quantum and classical state spaces are convex sets. This has to do with the fact that the collection of measures over an orthomodular lattice can be always endowed with a convex set structure (it is straightforward to show this for measures obeying axiom Equations (9a)–(9c)). The convex structure of the state space will play a key role in probabilistic theories.

Is it possible to describe a generalized probabilistic theories using convex sets as the starting point? The answer is affirmative (see, for example, [16–18] and [45–48,58]). Let us denote by $\mathcal{C}$ the set of all possible states of an arbitrary model. It is reasonable to assume that $\mathcal{C}$ is convex, given the fact that we should be allowed to make mixtures of states. Given an observable quantity, denote by X the set of its possible measurement outcomes. Given an arbitrary state $\nu \in \mathcal{C}$ and any outcome $x \in X$, a number $\nu(x) \in [0, 1]$ should be assigned, representing the probability of obtaining the outcome x given that the system is prepared in state ν. Using this, for outcome x, we can define an affine evaluation-functional $E_x \mathcal{C} \to [0, 1]$ in a canonical way by $E_x(\nu) = \nu(x)$.

As $\mathcal{C}$ is convex, it can be naturally embedded in a vector space $V(\mathcal{C})$. Thus, any affine functional acting on $\mathcal{C}$ belongs to a dual space $V^*(\mathcal{C})$. It is natural to assume that the probabilities for the outcomes of a generalized measurement are represented by functionals $E : \mathcal{C} \to [0, 1]$, usually called generalized effects. It is important to remark that the model will not be fully specified until the geometries of

the convex set of states and that of the set of functionals are both determined. We also assume that there exists a normalization functional $u_\mathcal{C}$, such that $u_\mathcal{C}(\nu) = 1$ for all $\nu \in \mathcal{C}$ (e.g., this functional is represented by the trace operation in the quantum case). Thus, the state ν is completely characterized by the probabilities $E(\nu)$. Furthermore, a (discrete) observable is represented by a set of effects $\{E_i\}$, such that $\sum_i E_i = u_\mathcal{C}$, and therefore, $\sum_i E_i(\nu) = 1$ for $\nu \in \mathcal{C}$.

$\mathcal{C}$ will be said to be finite dimensional if and only if $V(\mathcal{C})$ is finite dimensional. In this paper, we will restrict for simplicity to this case and to compact sets of states. These conditions imply that $\mathcal{C}$ will be expressed as the convex hull of its extreme points. As in the quantum and classical cases, extreme points of the convex set of states will represent pure states.

Define a finite dimensional simplex as the convex hull of $d + 1$ linearly-independent points. A system is said to be classical if and only if it is a simplex. It is a well-known fact that in a simplex, a point may be expressed as a unique convex combination of its extreme points. This characteristic feature of classical theories no longer holds in quantum models. Indeed, in the case of quantum mechanics, there are infinite ways in which one can express a mixed state as a convex combination of pure states (for a graphical representation, think about the maximally-mixed state in the Bloch sphere).

Interestingly enough, there is also a connection between the faces of the convex set of states of a given model and its lattice of properties (in the quantum-logical sense), providing an unexpected connection between geometry, lattice theory and statistical theories [11,23,103]. F is a face if for all x satisfying:

$$x = \lambda x_1 + (1 - \lambda)x_2, \quad 0 \leq \lambda \leq 1 \tag{10}$$

then $x \in F$ if and only if $x_1 \in F$ and $x_2 \in F$ [11]. Thus, faces of a convex set can be interpreted geometrically as subsets that are stable under mixing and purification. It is possible to show that the set of faces of any convex set can be endowed with a lattice structure in a canonical way. For a classical model (*i.e.*, described by a simplex), it turns out that the lattice is Boolean. Thus, probabilities defined by classical state spaces are Kolmogorovian. On the other hand, in QM, the lattice of faces of the convex set of states (defined as the set of positive trace class Hermitian operators of trace one) is isomorphic to the von Neumann lattice of closed subspaces $\mathcal{P}(\mathcal{H})$ [11,23]. This is nothing but saying that quantum states obey von Neumann axioms. In this way, a clear connection can be made between the approach based on orthomodular lattices and the approach based on convex sets. A similar result holds for more general (but not all) state spaces, but we will not deal with this problem here (see [23] and [11] for more discussion on this subject).

It is very important to remark that general probabilistic models will fail to be Kolmogorovian in general. This has important consequences for the possible correlations that can be defined between different systems and, thus, for information theoretical purposes.

We mention finally an important remark about the different degrees of generality that can be attained using different frameworks. It is very reasonable to start with measures over orthomodular lattices, mainly because this framework includes an important family of physical examples (such as classical statistical theories, quantum mechanics, quantum statistics and relativistic quantum field theory), but also because it allows one to represent complementarity in a very direct way. However, more general models of interest can be constructed. For example, σ-orthomodular posets can be used as events algebras (by defining measures similarly as those defined by axiom Equation (9)). All

orthomodular lattices are σ-orthomodular posets, but the last ones are more general, because they can fail to be lattices [57]. Finally, the approach that uses convex sets as a starting point is more general than the one provided by orthomodular lattices (this is so because it is possible to find models for which no ortho-complementarity relation can be defined [23], and thus, their lattice of faces fails to be orthomodular). This notwithstanding, in order to illustrate the most salient features of non-Kolmogorovian probabilistic models, it is sometimes sufficient to stay in the orthomodular lattices setting. This is what we will do mostly in this paper (but we will consider some more general examples in Section 5).

4.1. Non-Kolmogorovian Information Theory and Contextuality

Complementarity and contextuality [104–106] are salient features of quantum theory. The role of the complementarity principle in quantum information theory was discussed in [107], where it is shown that it is crucial for understanding the main features of quantum information protocols. One of the most important formal expressions of the complementarity principle is that of the non-commutativity of operators representing physical observables. Additionally, this is intimately connected with the non-Boolean structure of the lattice of projection operators. Furthermore, the success of the most important quantum computation algorithms is explained in light of the projective geometry underlying the formalism of quantum theory in [108].

To see how this contextual structure reappears in a more general setting, consider an orthomodular lattice $\mathcal{L}$. A maximal Boolean subalgebra is a subset $\mathcal{B} \subseteq \mathcal{L}$, such that: (1) $\mathcal{B}$ is closed and is a Boolean algebra with respect to the operations inherited from $\mathcal{L}$ (*i.e.*, it is a Boolean subalgebra); and (2) if $\mathcal{B}'$ is another Boolean subalgebra, such that $\mathcal{B} \subseteq \mathcal{B}'$, then $\mathcal{B} = \mathcal{B}'$ (*i.e.*, it is maximal). The important thing for us is that maximal Boolean subalgebras can be considered as representing particular experiments to perform on the system. To illustrate this point, think of a spin $\frac{1}{2}$ system. If we want to measure the spin component along axis $\hat{z}$, this will be represented by operator $\hat{\sigma}_z$. Then, this operator is associated with the Boolean subalgebra $\{\mathbf{0}, |+\rangle\langle+|, |-\rangle\langle-|^{\perp}, \mathbf{1}\}$, representing all possible events defined by the experiment, which consists of measuring that quantity: spin up in direction $\hat{z}$ ("$|+\rangle\langle+|$"), spin down in direction $\hat{z}$ ("$|-\rangle\langle-|^{\perp}$"), the contradiction "$\mathbf{0}$" and the tautology "$\mathbf{1}$" (which are the analogs of "$\emptyset$" and the whole outcome set "Ω" in the classical case, respectively).

Denote by **B** the set of all possible Boolean subalgebras of an orthomodular lattice $\mathcal{L}$. It is possible to show that $\mathcal{L}$ can be written as the sum of its maximal Boolean subalgebras [109],

$$\mathcal{L} = \bigvee_{\mathcal{B} \in \mathbf{B}} \mathcal{B} \tag{11}$$

What is the meaning of this technical result for generalized probabilistic theories? If $\mathcal{L}$ is Boolean, the result is trivial: the system can be described by using a single probability distribution over a single experimental setup. If it is not Boolean, this means that the event algebra of our theory may present mutually-complementary contexts. In other words, we will need to perform incompatible experiments (each one represented by a maximal Boolean subalgebra) in order to fully describe phenomena. Notice that each generalized state s, when restricted to a maximal Boolean subalgebra $\mathcal{B}$, gives a Kolmogorovian probability measure $s|_{\mathcal{B}}$. Taken together with Equation (11), this implies that a generalized state on a

contextual model can be considered as a collection of classical probabilities indexed by each empirical setup. The generalized measure obeying axiom Equations (9a)–(9c) provides a coherent pasting of this collection of Kolmogorovian measures. In the quantum case, this role is played by the density matrix representing the state of the system.

These features can be taken as a starting point in the convex sets approach. For example, in [50] (see also [110,111]), a state s is considered as a list of probability distributions: $s = (p(i, W))_{i=0,...,n-1;W=X_0,....,X_{m-1}}$. The possible W's represent a set of fiducial measurements, and the i's label the outcomes of each measurement. Fiducial measurements represent sets of measurements out of which the state can be determined. To fix ideas, let us look in detail at the qubit. In this case, each state can be specified as $s = (p(i, W))_{i=0,1;W=\hat{\sigma}_x,\hat{\sigma}_y,\hat{\sigma}_z}$. The observables represented by $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ are sufficient to determine the state completely (*i.e.*, they form a fiducial set). Notice that from this perspective, a state is considered again as a collection of classical probability distributions.

Non-Kolmogorovian probabilities are a condition of possibility for a departure from Shannon's classical information theory. This can be understood in a simple way following a generalization of the R. T. Cox approach to probability theory as follows:

- R.T. Cox [72,73] showed that if a rational agent is confronted with a Boolean algebra representing empirical events, then any function measuring his degree of belief of the occurrence of any event must be equivalent to a classical probability calculus. In [112], it is shown that if a rational agent is confronted with a non-distributive algebra of physical events, then the consistent probabilities must be those of Equation (9).

- In a similar way, in the Cox approach (see [59] and the references therein), it is shown that Shannon's entropy:

$$H(p_i) = -\sum_i p_i \ln(p_i) \tag{12}$$

  is the most natural information measure for an agent confronted with a Boolean algebra of events. In other words, when only one empirical context is involved and the probabilities are measures defined in a Boolean algebra, such as those of Equation (1), then Shannon's measure is a natural choice. In the quantum case, the probability distribution $\{p_i\}$ must be replaced by a state represented as a density matrix ρ. Now, as explained above, we have a classical probability distribution for each empirical context. This implies that the information measure has to be modified in such a way that Shannon's measure is obtained when the quantum measure is restricted to a Boolean subalgebra of $\mathcal{P}(\mathcal{H})$. As is well known, the von Neumann entropy:

$$S(\rho) = -\text{tr}(\rho \ln(\rho)) \tag{13}$$

  does the job. These ideas can be made more precise, both from the formal and the operational points of view. Indeed, in [59], it is shown that, if the Boolean algebra is replaced by a non-Boolean one (such as $\mathcal{P}(\mathcal{H})$ or more general ones), then the von Neumann and measurement entropies must be used (we refer the reader to [59] for the details of the derivation). In other words: if the event algebras are non-Boolean, then probabilities will be non-Kolmogorovian, and then, the information measures will depart from Shannon's entropy and more general classical ones (we will discuss the specific form of this departure in Section 5).

This is expressed clearly in the formal structure of classical and quantum information theories as follows. In Shannon's theory, a source emits different messages x of an outcome set X with probabilities p_x: this means that the probabilities involved are nothing but a Kolmogorovian measure over the Boolean algebra generated by the possible outcomes of the source. This implies that Shannon's entropy will play a key role in the formalism. For example, in the noiseless channel coding theorem, the value of Shannon's entropy of the source $H(X)$ measures the optimal compression for the source messages [113]. What changes in the quantum setting? Due to the fact that the final output of the source is now represented by a density matrix $\rho = \sum p_x \rho_x$ (*i.e.*, by a non-Kolmogorovian measure), then von Neumann's entropy comes onto stage. This is expressed, for example, in Schumacher's quantum coding theorem, in which the optimal bound for coding is expressed in terms of this quantity [114,115].

The role of the non-Kolmogorovian probability involved in the quantum state emitted by the source is also expressed in the existence of the Holevo bound: the mutual information between emitter and receiver will be bounded from above by a quantity depending on von Neumann's entropy $S(\rho)$:

$$I(X:Y) \leqslant S(\rho) - \sum p_x S(\rho_x) \tag{14}$$

where $I(X{:}Y)$ represents the classical mutual information between random variables X and Y. The above bound means that there is an intrinsic limit to the information accessible to the receiver. For example, it can be shown that if the original mixture is formed by non-orthogonal states, the Holevo bound implies that $I(X{:}Y)$ is strictly less than $H(X)$ (Shannon's measure of the source), and then, it is impossible for the receiver to determine X perfectly if he measures the observable Y [116]. This implies that if the states prepared by the emitter are non-orthogonal, it will not be possible for the receiver to determine the emitted state with certainty. This impossibility is directly related to the complementarity principle and, thus, to the non-Kolmogorovian character of the emitted quantum state.

4.2. Communication and Correlations in the Generalized Setting

Communication is a central aspect of any possible kind of information theory. However, communication involves more than one party: a message (or something) must be sent from one party to another. This is why the study of correlations is so important in order to account for the probabilistic aspects of a source. In order to show that informational notions can be studied in the general setting described above, a suitable description of multipartite states and correlations is needed. This has indeed been done quite extensively [45–48], and many notions essential to quantum information processing (such as entanglement, no-cloning, no-broadcasting and teleportation) can be generalized and studied in arbitrary statistical models. A departure of classical information theory will be found in state spaces for which non-classical probabilities and correlations are involved, and we will review how this is directly related to the non-Kolmogorovian structure of the state space.

Let us consider a compound system, formed of parties A and B, with state spaces $\mathcal{C}_A$ and $\mathcal{C}_B$, respectively. The joint system will also have a state space; let us denote it by $\mathcal{C}_A \otimes \mathcal{C}_B$ (the meaning of this notation is clarified below). In order to study its mathematical features, let us suppose that $\mathcal{C}_A \otimes \mathcal{C}_B$ can be included in the linear span of $V(\mathcal{C}_A) \otimes V(\mathcal{C}_B)$ (this assumption is discussed in [47]). Consider the set that contains all bilinear functionals $\varphi{:}V(\mathcal{C}_A) \times V(\mathcal{C}_B) \longrightarrow \mathbb{R}$ satisfying $\varphi(E, E') \geq 0$ for all effects E and E' and $\varphi(u_A, u_B) = 1$. It is very reasonable, then, to call this set a maximal tensor product

state space $\mathcal{C}_A \otimes_{max} \mathcal{C}_B$ for A and B. $\mathcal{C}_A \otimes_{max} \mathcal{C}_B$ has the property of being the biggest set of states in $(V(\mathcal{C}_A) \otimes V(\mathcal{C}_B))^*$, which assigns probabilities to all product measurements.

Analogously, a minimal tensor product state space $\mathcal{C}_A \otimes_{min} \mathcal{C}_B$ can be defined as the convex hull of all product states. This will be the analog of the convex set of separable states in quantum mechanics (see [53] for more discussion on this). We will write a product state as $\nu_A \otimes \nu_B$ satisfying:

$$\nu_A \otimes \nu_B(E, E') = \nu_A(E)\nu_B(E') \tag{15}$$

for all pairs $(E, E') \in V^*(\mathcal{C}_A) \times V^*(\mathcal{C}_B)$. Given these two extreme possibilities (maximal and minimal tensor product state spaces), the set of states $\mathcal{C}_A \otimes \mathcal{C}_B$ of an actual model lies somewhere "in between":

$$\mathcal{C}_A \otimes_{min} \mathcal{C}_B \subseteq \mathcal{C}_A \otimes \mathcal{C}_B \subseteq \mathcal{C}_A \otimes_{max} \mathcal{C}_B \tag{16}$$

For classical compound systems (for which state spaces are simplices representing Kolmogorovian probabilities), the set of compound states equals the minimal tensor product (and is again a classical state space). This means that if both subsystems are classical, we recover the equality: $\mathcal{C}_A \otimes_{min} \mathcal{C}_B = \mathcal{C}_A \otimes_{max} \mathcal{C}_B$. It can be shown that for quantum mechanics, we have the strict inclusions $\mathcal{C}_A \otimes_{min} \mathcal{C}_B \subseteq \mathcal{C}_A \otimes \mathcal{C}_B \subseteq \mathcal{C}_A \otimes_{max} \mathcal{C}_B$.

With this formal setting, it is now very natural to introduce a general definition of separable state in an arbitrary convex operational model. This is done in an analogous way to that of [9] (see, for example, [48,51]:

Definition 1. A state $\nu \in \mathcal{C}_A \otimes \mathcal{C}_B$ will be called separable if there exist $p_i \in \mathbb{R}_{\geq 0}$, $\nu_A^i \in \mathcal{C}_A$ and $\nu_B^i \in \mathcal{C}_B$, such that:

$$\nu = \sum_i p_i \nu_A^i \otimes \nu_B^i, \quad \sum_i p_i = 1 \tag{17}$$

Entangled states are thus defined as those that are not separable. It can be easily checked that entangled states exist if and only if $\mathcal{C}_A \otimes \mathcal{C}_B$ is strictly greater than $\mathcal{C}_A \otimes_{min} \mathcal{C}_B$. Thus, no entangled states exist for classical theories. In this way, non-classical correlations will not be allowed, and no departure of classical information theory will be found.

It is worth noting that this generalization of entanglement, although natural, is by no way unique, neither the most general possibility. In [117–119], Barnum *et al.* propose a subsystem-independent concept of entanglement, where the focus is on the relation between the convex set of states and a preferred (relevant or prescribed by any means) set of effects. Then, entanglement becomes a relative notion of the purity of the states with respect to the relevant effects (see [117,119,120] for details). Being independent of a certain subsystem decomposition, this notion becomes substantially more general than the usual one, even in the quantum scenario (see, e.g., [121–124]).

Regarding discord, Perinotti studied a possible introduction of the notion in general probabilistic theories [51]. As the original definitions of quantum discord rely on the information content of the states and because the information measures are not uniquely defined for general probabilistic theories, Perinotti prefers to give an operational definition of discord. He starts by defining the set of null-discord states and proves that they can be expressed as:

$$\omega_{nd} = \sum_{i \in I} q_k(\psi_A^i \otimes \sigma_B^i) \tag{18}$$

where $\{\psi_A^i\}_{i \in I}$ is a set of jointly perfectly distinguishable pure states, $\{\omega_B^i\}_{i \in I}$ is a set of arbitrary states of B and $\{q_i\}_{i \in I}$ is a probability distribution (see [51] for details). Then, the discord of a state ν is defined as the minimal operational distance to the set of null-discord states Ω_{nd}:

$$\mathcal{D}(\nu) := \min_{\omega_{nd} \in \Omega_{nd}} ||\nu - \omega_{nd}||_{op} \tag{19}$$

The operational distance is defined through the minimum error probability in the discrimination of both states [125].

The fact that correlations between different parties can be studied using information measures in the generalized setting allows one to pose the problem of communication in a suitable mathematical form. Given that the probabilistic models involved can be non-Kolmogorovian, the departure from Shannon's formalism is unavoidable in most cases.

5. Generalized Entropies and Majorization

In this section, we extend the definition of classical and quantum Salicrú entropies to the case of general probabilistic theories. In addition, we introduce definitions of spectra of states and majorization in those theories.

5.1. Entropies and Majorization in Classical and Quantum Theories

Inspired by [126], Salicrú *et al.* have introduced a very general expression for entropies [68], which we call classical (h, ϕ)-entropies, as follows:

Definition 2. For an N-dimensional probability vector $p = \{p_i\}$ with $p_i \geq 0$ and $\sum_{i=1}^{N} p_i = 1$, the classical (h, ϕ)-entropies are defined as:

$$H_{(h,\phi)}(p) = h\left(\sum_{i=1}^{N} \phi(p_i)\right) \tag{20}$$

where entropic functionals $h : R \mapsto R$ and $\phi : [0, 1] \mapsto R$ are continuous with $\phi(0) = 0$ and $h(\phi(1))=0$ and are such that either: (i) h is increasing and ϕ is concave; or (ii) h is decreasing and ϕ is convex.

It is straightforward to see that this definition yields the most renowned entropies, namely the Shannon [63], Tsallis [64] and Rényi ones [65], as particular cases. Indeed, one key property that all of these entropies share is related to the concept of majorization [127]. Majorization gives a partial order between probability vectors, and it is defined as follows: for given probability vectors p and q of length N sorted in decreasing order, it is said that p is majorized by q, denoted as $p \prec q$, when:

$$\sum_{i=1}^{n} p_i \leq \sum_{i=1}^{n} q_i \text{ for all } n = 1, \ldots, N-1 \text{ and } \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} q_i \tag{21}$$

In [69,128], it has been shown that classical (h, ϕ)-entropies are Schur-concave, that is preserving the majorization relation: if $p \prec q \Rightarrow H_{(h,\phi)}(p) \geq H_{(h,\phi)}(q)$. Many properties of Salicrú entropies can be proven using majorization, e.g., the lower and upper bounds: $0 \leq H_{(h,\phi)}(p) \leq h\left(N\phi\left(\frac{1}{N}\right)\right)$.

On the other hand, it is quite natural to define quantum (h, ϕ)-entropies replacing the probability vector by the density operator and the sum by the trace in Definition 2, as follows [69]:

Definition 3. Let us consider a quantum system described by a density operator ρ acting on an N-dimensional Hilbert space $\mathcal{H}$. The quantum (h, ϕ)-entropies (under the same assumptions for h and ϕ in Definition 2) are defined as follows:

$$\mathbf{H}_{(h,\phi)}(\rho) = h \left(\mathrm{Tr} \, \phi(\rho) \right) \tag{22}$$

As in the classical counterpart, the quantum (h, ϕ)-entropies include as particular cases the von Neumann [129] and the quantum versions of Rényi and Tsallis entropies. It can be shown that if the probability vector p is formed by the eigenvalues of ρ, then:

$$\mathbf{H}_{(h,\phi)}(\rho) = H_{(h,\phi)}(p) \tag{23}$$

In other words, quantum (h, ϕ)-entropies are nothing more than classical (h, ϕ)-entropies of the probability vectors formed by eigenvalues of density operators.

Let us consider two density operators ρ and σ with p and q vectors formed by eigenvalues sorted in decreasing order, respectively. Now, ρ is majorized by σ, denoted as $\rho \prec \sigma$, meaning that $p \prec q$ in the sense of Equation (21). It can be shown that quantum (h, ϕ)-entropies are also Schur-concave [69].

Let $p(E; \rho)$ be the probability vector whose components are given by the Born rule for a rank-one POVM E and state ρ, that is $p_i(\mathbf{E}_i; \rho) = \mathrm{Tr} \, \rho \mathbf{E}_i$. An alternative definition of quantum (h, ϕ)-entropies, which is equivalent to Definition 3, but with more physical meaning related to the probability of measurement, is the following [69]:

Definition 4. Under the same assumptions in Definition 3, the quantum (h, ϕ)-entropies are also defined as:

$$\mathbf{H}_{(h,\phi)}(\rho) = \min_{E \in \mathbb{E}} H_{(h,\phi)}(p(E; \rho)) \tag{24}$$

where $\mathbb{E}$ is the set of all rank-one POVMs.

Further properties of classical and quantum (h, ϕ)-entropies are given in [69] (and the references therein).

5.2. Entropies and Majorization in General Probabilistic Theories

Now, we aim to extend the definition of (h, ϕ)-entropies to more general probabilistic theories. It is possible to do this at least in two different ways. First, one could start with an atomic orthomodular lattice $\mathcal{L}$ defining an algebra of events. A frame in $\mathcal{L}$ will be an orthogonal set $\{a_i\}_{i \in I}$ of atoms, such that $\bigvee_{i \in I} a_i = \mathbf{1}$. Frames represent maximal experiments. For example, in quantum mechanics, any orthonormal basis (or rank-one PVM) is a frame. Thus, for each frame $\mathcal{F} = \{a_i\}_{i \in I}$ and each state ν, we have $p_i = \nu(a_i)$. Then, $\{p_i\}$ defines a probability vector, and this allows us to define the (h, ϕ)-entropies relative to that frame:

$$H_{(h,\phi)}|_{\mathcal{F}}(\nu) = h \left(\sum_{i \in I} \phi(\nu(a_i)) \right) \tag{25}$$

In order to give a definition independent of the frame, we have to take the minimum over all possible frames:

Definition 5. Let us consider a state $\nu \in \mathcal{C}$. The general (h, ϕ)-entropies (under the same assumptions for h and ϕ in Definition 2) are defined as follows:

$$\mathrm{H}_{(h,\phi)}(\nu) = \inf_{\mathcal{F} \in \mathbb{F}} H_{(h,\phi)}|_{\mathcal{F}}(\nu) \tag{26}$$

where $\mathbb{F}$ is the set of all frames. This is the canonical way in which entropies can be defined in general probabilistic theories. We observe that this approach resembles Definition 4 for the quantum case. Measurement entropy given in [49,66] is a particular case of this approach. However, it also includes other quantities, such as the Rényi and Tsallis ones in the case of general probabilistic theories. Notice that by taking the minimum over all possible frames, the contextual structure of the probability measures involved is made explicit.

There is another possible way in which (h, ϕ)-entropies in quite general probabilistic theories can be defined: we will provide a generalization of Definition 3. For this task, we have to define the notions of generalized spectrum and majorization. We restrict to arbitrary (compact) convex sets of states in finite dimensions: for these spaces, each element can be written as a convex combination of its pure states (as is the case in quantum and classical mechanics). In other words, there exist pure states $\{\nu_i\}$, such that every state ν can be written as:

$$\nu = \sum_i p_i \nu_i \tag{27}$$

However, this decomposition is not, in general, unique. For instance, the maximally-mixed state in quantum mechanics has infinite decompositions even in terms of orthogonal pure states. Notwithstanding, the probability vectors defined by the coefficients of these decompositions are all the same. Notice that this uniqueness property need not be true for arbitrary models, as we will discuss below.

We introduce now our notion of the generalized spectrum inspired by the Schrödinger mixture theorem (see, e.g., Theorem 8.2 in [11]). Using this theorem, it can be shown that the probability vector formed by the coefficients of any convex pure decomposition of a quantum state is majorized by the one formed by its eigenvalues. In other words, the spectrum of a quantum state has the distinctive property of being the majorant of all possible probability vectors originated in convex decompositions in terms of pure states. We will abstract this property and use it for defining a generalized spectrum for generalized states as follows. Given a probabilistic model described by a compact convex set, let M_ν be the set of probability vectors of all possible convex decompositions of a state ν in terms of pure states, that is:

$$M_\nu := \{p(\nu) = \{p_i\} \mid \nu = \sum_i p_i \nu_i \text{ for pure } \nu_i\} \tag{28}$$

Then, we propose the following:

Definition 6. Given a state ν, if the majorant of the set M_ν (partially ordered by majorization) exists, it is called the spectrum of ν, and it is denoted by $\bar{p}(\nu)$.

Accordingly, the corresponding generalized spectral decomposition is:

$$\nu = \sum_i \bar{p}_i \bar{\nu}_i \tag{29}$$

Notice that our definition reduces to the usual one for classical theories (where the sets of states are simplexes) and also for quantum mechanics. In the former case, equivalence can be checked easily, because there is only one convex decomposition in terms of pure states. In the latter case, as noted above, equivalence is a consequence of the Schrödinger mixture theorem. Notice however, that for a general statistical theory described by a compact convex set, it could be that the supremum $\bar{p}(\nu)$ does not exist for all possible states.

We observe that an alternative definition of the generalized spectrum has been recently introduced by Barnum *et al.* in [130]. The authors define the spectrum of a state as the unique (up to permutations) convex decomposition into perfectly distinguishable pure states. Distinguishability has the following operational meaning: a set of states $\{\nu_i\}$ is perfectly distinguishable if there is a measurement $\{\mathbf{E}_i\}$, such that $\mathbf{E}_i(\nu_j) = \delta_{ij}$. It is important to remark that their definition of spectrum cannot be used in arbitrary state spaces. This is due to the fact that for certain spaces, the decomposition of a state into perfectly distinguishable pure states can fail to be unique, and different decompositions can yield different probability vectors. Spaces for which decomposition into perfectly distinguishable states always exist are said to satisfy the weak spectrality axiom (WS-spaces). In spaces satisfying strong spectrality (S-spaces), the probability vectors of the convex pure decomposition into perfectly distinguishable states are unique (up to permutations). It can be shown that there are WS-spaces that are not S-spaces, and then, the definition of spectrality presented in [130] does not work in those cases. The definition presented in [130] and ours yield the same result for classical and quantum state spaces. However, they are expected to be non-equivalent in the general case. There could be spaces for which certain states admit different probability vectors for distinct decompositions into perfectly distinguishable pure states, but for which it is still possible to find a maximum according to our definition (see, for example, Figure 1). It is an interesting open question to determine under which conditions both definitions are equivalent and, specially, the range of validity of Definition 6. This last task can be rephrased as follows: which are the spaces for which a generalized version of the Schrödinger mixture theorem is valid? We will not deal with this problem here; we will only restrict to show how our definition can be used to define generalized majorization, functions over states and, in particular, entropic measures.

Definition 6 can be used to introduce the concept of generalized majorization naturally as follows.

Definition 7. Given two states μ and ν, one has that μ is majorized by ν, denoted by $\mu \prec \nu$, if and only if:

$$\bar{p}(\mu) \prec \bar{p}(\nu) \tag{30}$$

where $\bar{p}(\mu)$ and $\bar{p}(\nu)$ are the corresponding generalized spectra from Definition 6.

Moreover, our definition of the generalized spectrum can be also used to evaluate a function ϕ in a generalized state as follows. For any possible mixture $\{p_i, \nu_i\}$ of ν, we define the application of a functional ϕ to the state given the mixture as:

$$\phi(\nu)|_{\{p_i, \nu_i\}} := \sum_i \phi(p_i)\nu_i \tag{31}$$

In particular, we are interested in the mixture $\{\bar{p}_i, \bar{\nu}_i\}$, which leads to the definition:

$$\phi(\nu) := \phi(\nu)|_{\{\bar{p}_i, \bar{\nu}_i\}} \tag{32}$$

We have seen in Section 4 that the partial trace of the quantum formalism can be extended to the general setting by using the normalization functional $u_{\mathcal{C}}$. This allow us to define alternative generalized (h, ϕ)-entropies.
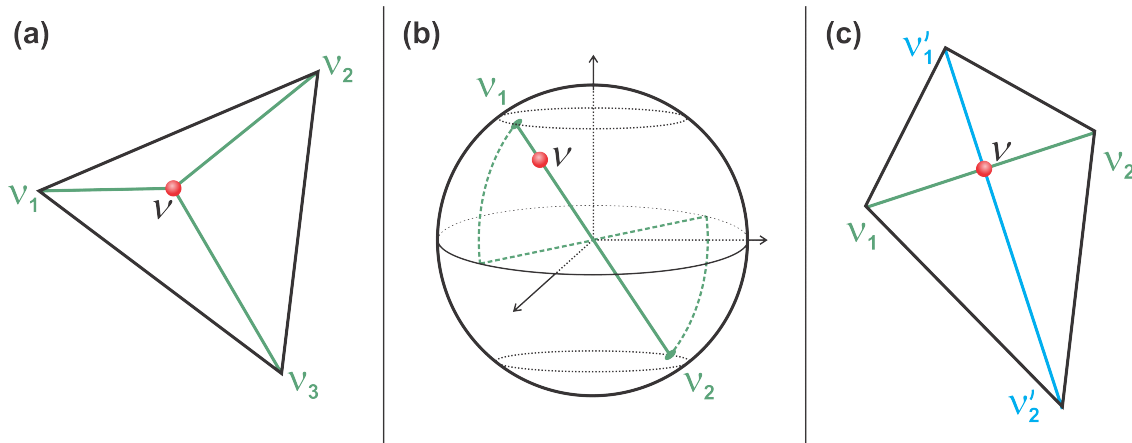

Figure 1. The generalized spectral decomposition, Equation (29), can be computed in a variety of probabilistic theories. (**a**) When the convex set is a simplex, the decomposition in terms of pure states is unique, and so, it determines the spectrum of ν. In the triangle above, ν can be written in a unique way as a mixture of ν_1, ν_2 and ν_3. (**b**) For the state ν of a qubit, the spectrum is given by the eigendecomposition of its density matrix in terms of the orthogonal pure states ν_1 and ν_2. The same happens for any other quantum mechanical model. (**c**) For a general probabilistic theory, there are, *a priori*, many decompositions of a state in terms of pure ones, and we have to look for the majorant one. For example, for the non-regular polygon with four vertices, the state in the barycenter is $\nu = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_2 = x\nu_1' + (1 - x)\nu_2'$, with $x > \frac{1}{2}$. The second set of coefficients majorize the first one, so $\bar{p}(\nu) = \{x, 1 - x\}$ constitute the spectrum of ν. Note, however, that in both decompositions, the pure states are perfectly distinguishable.

Definition 8. Under the same assumptions as in Definition 5, we define the (h, ϕ)-entropies:

$$\widetilde{\mathrm{H}}_{(h,\phi)}(\nu) = h\left(u_{\mathcal{C}}\left(\phi(\nu)\right)\right) \tag{33}$$

In other words, these generalized entropies are are equal to the classical ones evaluated on the probability vector $\bar{p}(\nu)$, that is:

$$\widetilde{\mathrm{H}}_{(h,\phi)}(\nu) = H_{(h,\phi)}(\bar{p}(\nu))$$

In principle, it can be shown that all of the properties of classical (and quantum) (h, ϕ)-entropies that are based on majorization and Schur-concavity hold in this general case (further properties are under investigation).

6. Conclusions

In summary, our main contribution is two-fold. On the one hand, we have provided a discussion of the formal aspects of classical, quantum and generalized probabilistic theories to show that classical and quantum information theories are particular instances of a general informational framework, relying on a rather exhaustive revision of the literature (see Sections 2–4). In other words, when the algebra of events is a Boolean one, we recover Shannon's formalism. Moreover, when the probabilities involved are measures over projection lattices of Hilbert spaces, we obtain quantum information theory. Thus, quantum information theory arises as a non-Kolmogorovian version of Shannon's information theory, and both theories appear as particular cases of a generalized non-Kolmogorovian probabilistic calculus.

On the other hand, we have shown that the Salicrú entropies can be defined in two different ways in this generalized non-Kolmogorovian setting, extending the catalog of extant entropic measures available in the literature. The first one, Definition 5, is based on the frame representation for maximal experiments, and therefore, the measurement entropy appears as a particular case of a general family of entropies. The second one involves a novel definition of the spectrum for generalized measures (Definition 6) that relies on an essential property derived from the Schrödinger mixture theorem. In addition, we have provided some examples to show the scope of our definition of the generalized spectrum with respect to another recently introduced one. Furthermore, our definition of the generalized spectrum has allowed us to introduce a new notion of generalized majorization (Definition 7), from which we have derived another way to obtain generalized entropies, Definition 8. Both definitions of generalized entropies coincide for the classical and quantum cases. It remains open to study whether these definitions are equal in probabilistic theories beyond the classical and quantum cases.

Author Contributions

Federico Holik provided the main idea of the research. Gustavo M. Bosyk and Guido Bellomo commented on the manuscript at all stages. All authors equally performed the research, discussed the results and wrote the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bub, J. Quantum mechanics is about quantum information. *Found. Phys.* **2005**, *35*, 541–560.
2. Popescu, S. Nonlocality beyond quantum mechanics. *Nat. Phys.* **2014**, *10*, 264–270.
3. Amaral, B.; Cunha, M.T.; Cabello, A. Exclusivity principle forbids sets of correlations larger than the quantum set. *Phys. Rev. A* **2014**, *89*, 030101.
4. Cabello, A. Exclusivity principle and the quantum bound of the Bell inequality. *Phys. Rev. A* **2014**, *90*, doi:10.1103/PhysRevA.90.062125.
5. Clifton, R.; Bub, J.; Halvorson, H. Characterizing quantum theory in terms of information-theoretic constraints. *Found. Phys.* **2003**, *33*, 1561–1591.
6. Brunner, N.; Cavalcanti, D.; Pironio, S.; Scarani, V.; Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **2014**, *86*, 419–478.
7. Schrödinger, E. Discussion of probability relations between separated systems. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1935; Volume 31, pp. 555–563.
8. Schrödinger, E. Probability relations between separated systems. *Math. Proc. Camb. Philos. Soc.* **1936**, *32*, 446–452.
9. Werner, R.F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **1989**, *40*, 4277–4281.
10. Ollivier, H.; Zurek, W.H. Quantum discord: A measure of the quantumness of correlations. *Phys. Rev. Lett.* **2001**, *88*, doi:10.1103/PhysRevLett.88.017901.
11. Bengtsson, I.; Zyczkowski, K. *Geometry of Quantum States: An Introduction to Quantum Entanglement*; Cambridge University Press: Cambridge, UK, 2006.
12. Ludwig, G. *Foundations of Quantum Mechanics I*; Springer: Berlin/Heidelberg, Germany, 1983.
13. Ludwig, G. *Foundations of Quantum Mechanics II*; Springer: Berlin/Heidelberg, Germany, 1985.
14. Varadarajan, V.S. *Geometry of Quantum Theory*; Springer: Berlin, Germany; Heidelberg, Germany, 1965.
15. Varadarajan, V.S. *Geometry of Quantum Theory: Volume 2: Quantum Theory of Covariant Systems*; Van Nostrand Reinhold Company: New York, NY, USA, 1970.
16. Mielnik, B. Geometry of quantum states. *Commun. Math. Phys.* **1968**, *9*, 55–80.
17. Mielnik, B. Theory of filters. *Commun. Math. Phys.* **1969**, *15*, 1–46.
18. Mielnik, B. Generalized quantum mechanics. *Commun. Math. Phys.* **1974**, *37*, 221–256.
19. Kibble, T.W.B. Geometrization of quantum mechanics. *Commun. Math. Phys.* **1979**, *65*, 189–201.
20. Birkhoff, G.; von Neumann, J. The logic of quantum mechanics. *Ann. Math.* **1936**, *37*, 823–843.
21. Dalla Chiara, M.L.; Giuntini, R.; Greechie, R. *Reasoning in Quantum Theory: Sharp and Unsharp Quantum Logics*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 22.
22. Rédei, M. *Quantum Logic in Algebraic Approach*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998.
23. Beltrametti, E.G.; Cassinelli, G. *The Logic of Quantum Mechanics*; Cambridge University Press: Cambridge, UK, 2010; Volume 15.

24. Svozil, K. *Quantum Logic*; Springer: Berlin/Heidelberg, Germany, 1998.

25. Engesser, K.; Gabbay, D.M.; Lehmann, D. *Handbook of Quantum Logic and Quantum Structures: Quantum Structures*; Elsevier: Amsterdam, The Netherlands, 2011.

26. Jauch, J.M. *Foundations of Quantum Theory*; Addison-Wesley Educational Publishers Inc.: Boston, MA, USA, 1968.

27. Piron, C. *On the Foundations of Quantum Physics*; Springer: Berlin/Heidelberg, Germany, 1976.

28. Diederik, A. The one and the many: towards a unification of the quantum and the classical description of one and many physical entities. PhD thesis, Free University of Brussels, Brussels, Belgium, 1981.

29. Kalmbach, G. *Orthomodular Lattices*; Academic Press: Waltham, MA, USA, 1983; Volume 18.

30. Kalmbach, G. *Measures and Hilbert Lattices*; World Scientific: Singapore, Singapore, 1986.

31. Dvurecenskij, A.; Pulmannová, S. *New Trends in Quantum Structures*; Kluwer Academic Publisher: Dordrecht, 2000.

32. Yngvason, J. The role of type III factors in quantum field theory. *Rep. Math. Phys.* **2005**, *55*, 135–147.

33. Murray, F.J.; von Neumann, J. On rings of operators. *Ann. Math.* **1936**, *37*, 116–229.

34. Murray, F.J.; von Neumann, J. On rings of operators. II. *Trans. Am. Math. Soc.* **1937**, *41*, 208–248.

35. Von Neumann, J. On rings of operators. III. *Ann. Math.* **1940**, *41*, 94–161.

36. Von Neumann, J.; Murray, F.J. On rings of operators IV. *Ann. Math.* **1943**, *44*, 716–808.

37. Von Neumann, J. Zur algebra der funktionaloperationen und theorie der normalen operatoren. *Math. Ann.* **1930**, *102*, 370–427.

38. Halvorson, H.; Müger, M. Algebraic Quantum Field Theory. In *Philosophy of Physics*; Butterfield, J.B., Earman, J.E., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; pp. 731–922.

39. Rédei, M.; Summers, S.J. Quantum probability theory. *Stud. Hist. Philos. Sci. B* **2007**, *38*, 390–417.

40. Rédei, M. Some historical and philosophical aspects of quantum probability theory and its interpretation. In *Probabilities, Laws, and Structures*; Springer: Berlin/ Heidelberg, Germany, 2012; pp. 497–506.

41. Rédei, M. The Birkhoff-von Neumann Concept of quantum Logic. In *Handbook of Quantum Logic and Quantum Structures*; Engesser, K., Gabbay, D., Lehmann, D., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; p. 1–22.

42. Bratteli, O.; Robinson, D.W. *Operator Algebras and Quantum Statistical Mechanics: Volumes 1 and 2*; Springer: Berlin, Germany; Heidelberg, Germany, 2012.

43. Beltrametti, E.; Bugajski, S.; Varadarajan, V. Extensions of convexity models. *J. Math. Phys.* **2000**, *41*, 2500–2514.

44. Wilce, A. Quantum logic and probability theory. The Stanford Encyclopedia of Philosophy (Winter 2013 Edition); Zalta, E.N., Ed.; Available online: http://plato.stanford.edu/archives/win2013/entries/qt-quantlog/ (accessed on 26 October 2015).

45. Barnum, H.; Wilce, A. Information processing in convex operational theories. *Electron. Notes Theor. Comput. Sci.* **2011**, *270*, 3–15.

46. Barnum, H.; Duncan, R.; Wilce, A. Symmetry, compact closure and dagger compactness for categories of convex operational models. *J. Philos. Log.* **2013**, *42*, 501–523.

47. Barnum, H.; Barrett, J.; Leifer, M.; Wilce, A. Generalized no-broadcasting theorem. *Phys. Rev. Lett.* **2007**, *99*, doi:10.1103/PhysRevLett.99.240501.

48. Barnum, H.; Dahlsten, O.C.; Leifer, M.; Toner, B. Nonclassicality without entanglement enables bit commitment. In Proceedings of the IEEE Information Theory Workshop, Porto, Portugal, 5–9 May 2008; pp. 386–390.

49. Short, A.J.; Wehner, S. Entropy in general physical theories. *New J. Phys.* **2010**, *12*, doi:10.1088/1367-2630/12/3/033023.

50. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **2007**, *75*, doi:10.1103/PhysRevA.75.032304.

51. Perinotti, P. Discord and nonclassicality in probabilistic theories. *Phys. Rev. Lett.* **2012**, *108*, doi:10.1103/PhysRevLett.108.120502.

52. Barnum, H.; Barrett, J.; Leifer, M.; Wilce, A. Teleportation in general probabilistic theories. In *Mathematical Foundations of Information Flow*; Proceedings of Symposia in Applied Mathematics; Abramsky, S.; Milsove, M., Eds.; American Mathematical Society: Providence, RI, 2012; Volume 71, pp. 25–48.

53. Holik, F.; Massri, C.; Plastino, A. Generalizing Entanglement via Informational Invariance for Arbitrary Statistical Theories. *SOP Trans. Thoer. Phys.* **2014**, *1*, 138–153.

54. Mackey, G. *The Mathematical Foundations Of Quantum Mechanics*; W.A. Benjamin, Inc: New York, NY, USA, 1963.

55. Davies, E.B.; Lewis, J.T. An operational approach to quantum probability. *Commun. Math. Phys.* **1970**, *17*, 239–260.

56. Srinivas, M. Foundations of a quantum probability theory. *Commun. Math. Phys.* **1975**, *16*, 1672–1685.

57. Gudder, S.P. Stochastic Methods in Quantum Mechanics. In *Series in Probability and Applied Mathematics*; Bharucha, A.T., Ed.; Dover Publications: Mineola, NY, USA, 1979.

58. Holik, F.; Plastino, A. Quantum Mechanics: A New Turn in Probability Theory. In *Contemporary Research in Quantum Systems*; Zoheir, E., Ed.; Nova Publishers: New York, NY, USA, 2014.

59. Holik, F.; Plastino, A.; Sáenz, M. Natural information measures in Cox' approach for contextual probabilistic theories. *Quant. Inform. Process.* **2015**, submitted.

60. Holevo, A.S. *Probabilistic and Statistical Aspects of Quantum Theory*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 1.

61. Holevo, A. *Statistical Structure of Quantum Theory*; Springer: Berlin, Germany; Heidelberg, Germany, 2001; Volume 67.

62. Ingarden, R.S. Quantum information theory. *Rep. Math. Phys.* **1976**, *10*, 43–72.

63. Shannon, C.E. The mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.

64. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.

65. Rényi, A. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Statistical Laboratory of the University of

California, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.

66. Barnum, H.; Barrett, J.; Clark, L.O.; Leifer, M.; Spekkens, R.; Stepanik, N.; Wilce, A.; Wilke, R. Entropy and information causality in general probabilistic theories. *New J. Phys.* **2010**, *12*, doi:10.1088/1367-2630/12/3/033024.

67. Hein, C.A. Entropy in operational statistics and quantum logic. *Found. Phys.* **1979**, *9*, 751–786.

68. Salicru, M.; Menendez, M.; Morales, D.; Pardo, L. Asymptotic distribution of (h, φ)-entropies. *Commun. Stat. Theory Methods* **1993**, *22*, 2015–2031.

69. Bosyk, G.M.; Zozor, S.; Holik, F.; Portesi, M.; Lamberti, P.W. A family of generalized quantum entropies: Definition and properties. 2015, arXiv:1506.02090.

70. Kolmogorov, A.N. *Foundations of the Theory of Probability*; Chelsea Publishing Company: New York, NY, USA, 1956.

71. Rocchi, P. *Janus-Faced Probability*; Springer: Berlin/Heidelberg, Germany, 2014.

72. Cox, R.T. *Algebra of Probable Inference*; The Johns Hopkins Press: Baltimore, MD, USA, 1961.

73. Cox, R.T. Probability, frequency and reasonable expectation. *Am. J. Phys.* **1946**, *14*, 1–13.

74. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, UK, 2003.

75. De Finetti, B. *Teoria Della Probabilita*; Einaudi: Torino, Italy, 1970.

76. Holik, F.; Massri, C.; Plastino, A.; Zuberman, L. On the lattice structure of probability spaces in quantum mechanics. *Int. J. Theor. Phys.* **2013**, *52*, 1836–1876.

77. Busch, P.; Kiukas, J.; Lahti, P. On the notion of coexistence in quantum mechanics. *Math. Slovaca* **2010**, *60*, 665–680.

78. Heinonen, T. Imprecise Measurements in Quantum Mechanics. PhD. Thesis, University of Turku, Finland, 2005.

79. Ma, Z.; Zhu, S. Topologies on quantum effects. *Rep. Math. Phys.* **2009**, *64*, 429–439.

80. Foulis, D.J.; Gudder, S.P. Observables, calibration, and effect algebras. *Found. Phys.* **2001**, *31*, 1515–1544.

81. Cattaneo, G.; Gudder, S. Algebraic structures arising in axiomatic unsharp quantum physics. *Found. Phys.* **1999**, *29*, 1607–1637.

82. Busch, P.; Lahti, P.J.; Mittelstaedt, P. *The Quantum Theory of Measurement*; Springer: Berlin/Heidelberg, Germany, 1991.

83. Foulis, D.J.; Bennett, M.K. Effect algebras and unsharp quantum logics. *Found. Phys.* **1994**, *24*, 1331–1352.

84. Gleason, A.M. Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.* **1957**, *6*, 885–893.

85. Horodecki, R.; Horodecki, P.; Horodecki, M.; Horodecki, K. Quantum entanglement. *Rev. Mod. Phys.* **2009**, *81*, 865–942.

86. Raggio, G.A. A remark on Bell's inequality and decomposable normal states. *Lett. Math. Phys.* **1988**, *15*, 27–29.

87. Earman, J. Some Puzzles and Unresolved Issues About Quantum Entanglement. *Erkenntnis* **2014**, *80*, 303–337.

88. Henderson, L.; Vedral, V. Classical, quantum and total correlations. *J. Phys. A Math. Gen.* **2001**, *34*, doi:10.1088/0305-4470/34/35/315.

89. Hayashi, M. *Quantum Information*; Springer: Berlin/Heidelberg, Germany, 2006.

90. Datta, A. A condition for the nullity of quantum discord. 2010. arXiv:1003.5256.

91. Modi, K.; Brodutch, A.; Cable, H.; Paterek, T.; Vedral, V. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.* **2012**, *84*, 1655, doi:10.1103/RevModPhys.84.1655.

92. Wittmann, B.; Ramelow, S.; Steinlechner, F.; Langford, N.K.; Brunner, N.; Wiseman, H.M.; Ursin, R.; Zeilinger, A. Loophole-free Einstein-Podolsky-Rosen experiment via quantum steering. *New J. Phys.* **2012**, *14*, doi:10.1088/1367-2630/14/5/053030.

93. Branciard, C.; Cavalcanti, E.G.; Walborn, S.P.; Scarani, V.; Wiseman, H.M. One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering. *Phys. Rev. A* **2012**, *85*, doi:10.1103/PhysRevA.85.010301.

94. Smith, D.H.; Gillett, G.; de Almeida, M.P.; Branciard, C.; Fedrizzi, A.; Weinhold, T.J.; Lita, A.; Calkins, B.; Gerrits, T.; Wiseman, H.M.; *et al*. Conclusive quantum steering with superconducting transition-edge sensors. *Nat. Commun.* **2012**, *3*, doi:10.1038/ncomms1628.

95. Reid, M. Monogamy inequalities for the Einstein-Podolsky-Rosen paradox and quantum steering. *Phys. Rev. A* **2013**, *88*, doi:10.1103/PhysRevA.88.062108.

96. Jevtic, S.; Pusey, M.; Jennings, D.; Rudolph, T. Quantum steering ellipsoids. *Phys. Rev. Lett.* **2014**, *113*, doi:10.1088/1367-2630/16/8/083017.

97. Aerts, D.; Daubechies, I. A characterization of subsystems in physics. *Lett. Math. Phys.* **1979**, *3*, 11–17.

98. Aerts, D.; Daubechies, I. A mathematical condition for a sublattice of a propositional system to represent a physical subsystem, with a physical interpretation. *Lett. Math. Phys.* **1979**, *3*, 19–27.

99. Holik, F.; Plastino, A. Quantal effects and MaxEnt. *J. Math. Phys.* **2012**, *53*, doi:10.1063/1.4731769.

100. Holik, F.; Massri, C.; Plastino, A. The introduction of symmetry constraints within MaxEnt Jaynes's methodology. 2014, arXiv:1501.00074.

101. De Barros, J. Decision Making for Inconsistent Expert Judgments Using Negative Probabilities. In *Quantum Interaction*; Atmanspacher, H., Haven, E., Kitto, K., Raine, D., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8369, pp. 257–269.

102. Bueno-Soler, J.; Carnielli, W. May be and may be not: Paraconsistent probabilities from the LFI viewpoint. *CLE e-Prints* **2015**; *15*, Available online: http://www.cle.unicamp.br/e-prints/vol_15, n_2,2015.html (accessed on 26 October 2015).

103. Holik, F. Logic, Geometry and Probability Theory. *SOP Trans. Thoer. Phys.* **2014**, *1*, 128–137.

104. De Barros, J.A.; Oas, G. Some Examples of Contextuality in Physics: Implications to Quantum Cognition. In *Contextuality from Quantum Physics to Psychology*; Dzhafarov, E., Zhang, R., Jordan, S.M., Eds.; World Scientific: Singapore, Singapore, 2015.

105. De Barros, J.A.; Dzhafarov, E.N.; Kujala, J.V.; Oas, G. Measuring observable quantum contextuality. 2014, arXiv:1406.3088.

106. Cabello, A. Proposal for revealing quantum nonlocality via local contextuality. *Phys. Rev. Lett.* **2010**, *104*, doi:10.1103/PhysRevLett.104.220401.

107. Renes, J.M. The physics of quantum information: Complementarity, uncertainty, and entanglement. *Int. J. Quantum Inf.* **2013**, *11*, doi:10.1142/S0219749913300027.

108. Bub, J. Quantum computation from a quantum logical perspective. *Quant. Inform. Process.* **2007**, *7*, 281–296.

109. Navara, M.; Rogalewicz, V. The pasting constructions for orthomodular posets. *Math. Nachr.* **1991**, *154*, 157–168.

110. Garner, A.J.; Dahlsten, O.C.; Nakata, Y.; Murao, M.; Vedral, V. A framework for phase and interference in generalized probabilistic theories. *New J. Phys.* **2013**, *15*, 093044.

111. Brukner, U.; Zeilinger, A. Information Invariance and Quantum Probabilities. *Found. Phys.* **2009**, *39*, 677–689.

112. Holik, F.; Sáenz, M.; Plastino, A. A discussion on the origin of quantum probabilities. *Ann. Phys.* **2014**, *340*, 293–310.

113. Lombardi, O.; Holik, F.; Vanni, L. What is Shannon information? *Synthese* **2015**, doi:10.1007/s11229-015-0824-z.

114. Schumacher, B. Quantum coding. *Phys. Rev. A* **1995**, *51*, doi:10.1103/PhysRevA.51.2738.

115. Lombardi, O.; Holik, F.; Vanni, L. What is quantum information? 2014, Available online: http://philsci-archive.pitt.edu/11159/ (accessed on 26 October 2015).

116. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.

117. Barnum, H.; Knill, E.; Ortiz, G.; Viola, L. Generalizations of entanglement based on coherent states and convex sets. *Phys. Rev. A* **2003**, *68*, doi:10.1103/PhysRevA.68.032308.

118. Barnum, H.; Knill, E.; Ortiz, G.; Somma, R.; Viola, L. A subsystem-independent generalization of entanglement. *Phys. Rev. Lett.* **2004**, *92*, doi:10.1103/PhysRevLett.92.107902.

119. Barnum, H.; Ortiz, G.; Somma, R.; Viola, L. A generalization of entanglement to convex operational theories: Entanglement relative to a subspace of observables. *Int. J. Theor. Phys.* **2005**, *44*, 2127–2145.

120. Viola, L.; Barnum, H. Entanglement and subsystems, entanglement beyond subsystems, and all that. In *Philosophy of Quantum Information and Entanglement*; Bokulich, A., Jaeger, G., Eds.; Cambridge University Press: Cambridge, UK, 2010; pp. 16–43.

121. Somma, R.; Ortiz, G.; Barnum, H.; Knill, E.; Viola, L. Nature and measure of entanglement in quantum phase transitions. *Phys. Rev. A* **2004**, *70*, doi:10.1103/PhysRevA.70.042311.

122. Benatti, F.; Floreanini, R.; Marzolino, U. Sub-shot-noise quantum metrology with entangled identical particles. *Ann. Phys.* **2010**, *325*, 924–935.

123. Balachandran, A.; Govindarajan, T.; de Queiroz, A.R.; Reyes-Lega, A. Entanglement and Particle Identity: A Unifying Approach. *Phys. Rev. Lett.* **2013**, *110*, doi:10.1103/PhysRevLett.110.080503.

124. Balachandran, A.; Govindarajan, T.; de Queiroz, A.R.; Reyes-Lega, A. Algebraic approach to entanglement and entropy. *Phys. Rev. A* **2013**, *88*, doi:10.1103/PhysRevA.88.022301.

125. Chiribella, G.; D'Ariano, G.M.; Perinotti, P. Informational derivation of quantum theory. *Phys. Rev. A* **2011**, *84*, doi:10.1103/PhysRevA.84.012311.

126. Csiszàr, I. Information-Type Distance Measure and Indirect Observations. *Stud. Sci. Math. Hung.* **1967**, *2*, 299–318.

127. Marshall, A.W.; Olkin, I.; Arnold, B. *Inequalities: Theory of Majorization and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2010.

128. Zozor, S.; Bosyk, G.M.; Portesi, M. General entropy-like uncertainty relations in finite dimensions. *J. Phys. A Math. Theor.* **2014**, *47*, doi:10.1088/1751-8113/47/49/495302.

129. Von Neumann, J. Thermodynamik quantenmechanischer Gesamtheiten. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1927**, *1927*, 273–291.

130. Barnum, H.; Barrett, J.; Krumm, M.; Mueller, M.P. Entropy, majorization and thermodynamics in general probabilistic theories. 2015, arXiv:1508.03107v1.