Chapter 1

Some Examples of Contextuality in Physics: Implications to Quantum Cognition

J. Acacio de Barros

School of Humanities and Liberal Studies
San Francisco State University
1600 Holloway Ave.
San Francisco, CA 94132

Gary Oas

Stanford Pre-Collegiate Studies
Stanford University
Ventura Hall
Stanford, CA 94305-4115

Contextuality, the impossibility of assigning a single random variable to represent the outcomes of the same measurement procedure under different experimental conditions, is a central aspect of quantum mechanics. Thus defined, it appears in well-known cases in quantum mechanics, such as the double-slit experiment, the Bell-EPR experiment, and the Kochen-Specker theorem. Here we examine contextuality in such cases, and discuss how each of them bring different conceptual issues when applied to quantum cognition. We then focus on the shortcomings of using quantum probabilities to describe social systems, and explain how negative quasi-probability distributions may address such limitations.

1. Introduction

Contextuality is defined as the impossibility of assigning a single random variable to represent the outcomes of a measurement under different experimental conditions (thought of as contexts) [1–5]. More precisely, say you have a series of experimental conditions where you measure a property P, whose outcome (out of several runs) is represented by the random variable $\mathbf{P}$. Imagine that, for one of those conditions, P is measured together with other properties $A_1, A_2, \ldots, A_n$ (whose outcomes are also

2   *J. A. de Barros, G. Oas*

represented by random variables $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$), but also imagine that, for another condition, P is measured with $B_1, B_2, \ldots, B_{n'}$, and finally assume that it is not possible to create an experiment where all properties $P, A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_{n'}$ are measured simultaneously. Contextuality, as defined informally above, is thus the impossibility of finding a probability space $(\Omega, \mathcal{F}, p)$ for $\mathbf{P}, \mathbf{A}_1, \ldots, \mathbf{A}_n, \mathbf{B}_1, \ldots, \mathbf{B}_{n'}$ compatible with the distributions observed experimentally.

As an example, take a simple situation where we have three properties X, Y, and Z corresponding to true or false statements. Observing such properties is modeled by ± 1-valued random variables, $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$. Assume now that we only observe two properties at a time, but never all three together. Assume additionally that they are seen as perfectly anti-correlated to each other for each experimental condition, i.e.

$$E\left(\mathbf{XY}\right) = E\left(\mathbf{XZ}\right) = E\left(\mathbf{YZ}\right) = -1. \tag{1}$$

Clearly no probability space giving those correlations exists, since a $\omega \in \Omega$ giving $\mathbf{X}\left(\omega\right) = 1$ implies, from the first and second expectations in (1), $\mathbf{Y}\left(\omega\right) = -1$ and $\mathbf{Z}\left(\omega\right) = -1$, which contradicts the third expectation. However, if we relabel the variables, in the spirit of References [3, 6–9], making the fact that they were measured in a pairwise way explicit, it is possible to construct a $(\Omega, \mathcal{F}, p)$. For instance, if we have a new set of random variables $\mathbf{X}_\mathbf{Y}$, $\mathbf{X}_\mathbf{Z}$, $\mathbf{Y}_\mathbf{X}$, $\mathbf{Y}_\mathbf{Z}$, $\mathbf{Z}_\mathbf{X}$, and $\mathbf{Z}_\mathbf{Y}$, such that correlations in (1) are now

$$E\left(\mathbf{X}_\mathbf{Y}\mathbf{Y}_\mathbf{X}\right) = E\left(\mathbf{X}_\mathbf{Z}\mathbf{Z}_\mathbf{X}\right) = E\left(\mathbf{Y}_\mathbf{Z}\mathbf{Z}_\mathbf{Y}\right) = -1, \tag{2}$$

the sampling of an $\omega \in \Omega$ leads to no contradictions. It is straightforward that the contradiction from (1) comes from assuming that the value of, say, $\mathbf{X}$ when measured with $\mathbf{Y}$ is the same as its value when measured with $\mathbf{Z}$, i.e., it does not depend on the experimental context provided by the simultaneous measurement of the pairs.

Contextuality as defined here is ubiquitous in quantum mechanics, and may be at the heart of what defines quantum systems, as opposed to classical ones (see [10] in this volume). Examples of contextual quantum systems, some of them discussed in more details in Section 2, are successive measurements of spin [6, 11], the double-slit experiment [12], the Leggett-Garg experiment [13], the EPR-Bell experiment [14, 15], and the Kochen-Specker system of observables [16]. So, it should come as no surprise that the mathematical formalism developed to describe quantum systems is suitable to describe (at least certain) contextual systems.

It is perhaps for this reason that such formalism was successfully applied to social systems with a certain degree of success [17, 18]. Social systems, because of their contextuality, lack a joint probability distribution, and the contextual calculus of probability of quantum mechanics seems to offer a suitable framework for situations in which standard probability theory fails. For example, Savage's famous Sure-Thing-Principle (STP) [19], a consequence of classical probability theory, are violated by human decision-makers [20, 21]. Khrennikov and Haven [22] applied principles of quantum interference and showed that certain superpositions of quantum-like states representing mental processes could be used to describe the experimentally verified violation of STP.

So, it is possible the relationship between quantum mechanics, its formalism, and social phenomena goes beyond a simple analogy, but points to a deeper relationship between the Hilbert space formalism and the description of determinate contextual systems. However, arguments exist that certain social processes may not be describable by the quantum formalism [23–25], but instead by other contextual probability theories. Be that as it may, given the increasing importance of the quantum formalism in the social sciences [17, 18, 26], we believe that a distinction of the different quantum process that exhibit contextuality should be fruitful. Here we analyze three different quantum systems that are contextual, and show that each have contextuality that present different conceptual features.

This paper is organized the following way. First, in Section 2, we examine contextuality in quantum mechanics, starting with the famous double-slit experiment, and then moving to the Bell-EPR experiment and the Kochen-Specker theorem. In Section 3, we discuss contextuality in quantum cognition, and how it relates to the examples discussed from physics. Finally, in Section 4 we present a particular alternative model of extended probabilities that is suitable for some of the situations discussed in quantum cognition.

2. Contextuality in QM

A fundamental question in physics is what makes quantum systems different from classical ones (see our other contribution to this volume [10]). At the bottom of it seems to be the apparent impossibility of describing quantum systems in terms of concepts from classical physics, such as particles and fields.

In the early days of the quantum revolution, physicists attempted to

describe the microscopic phenomena observed in terms of what we know nowadays as classical ideas. But soon many realized that causality, one of the main tenets of classical physics, was not compatible with the quantum world. For instance, Rutherford's radioactive decay formula was seen as corresponding to a memoryless Poissonian process, and that therefore atoms who were about to decay had the same state as those whose decay would happen much later. Subsequently, shortly after Bohr published his theory explaining the spectrum of the Hydrogen atom, Rutherford himself remarked that Bohr's theory had a problem with causality [27]. Classical physics, a causal theoretical structure with its description in terms of phase-space states and Hamiltonian dynamics, was thought to not be able to account for what were essentially probabilistic processes. As such, the unavoidable probabilistic character of quantum mechanics became a topic of intense discussion and research.

Connected to this discussion was the idea that quantum probabilities could have their origins in the impossibility of simultaneously observing two complementary quantities, such as momentum and position of a particle. For instance, for Heisenberg, a measurement of position would cause a random disturbance on the momentum in such a way that knowing the position of a particle at a time t_0 would make a prediction of its position at time $t > t_0$ an impossible task. This perspective evolved into a viewpoint some physicists put forth that two incompatible properties could not *exist* at the same time, brought about by the discovery of the spatial quantization (spin).

To see this, imagine that spin is represented by a three-dimensional vector random variable $\mu(\omega)$ (here we follow [6]). If a measurement of spin, say, in the direction $\hat{\mathbf{z}}$, simply reveals the value of such random variable in such direction, without disturbing it, then its result would be $\mu \cdot \hat{\mathbf{z}}$, which experimentally can take only values ± 1 (here we use units where $\hbar/2 = 1$). However, there is nothing special about the direction $\hat{\mathbf{z}}$: the Stern-Gerlach (SG) apparatus measuring spin could be pointing in any direction of our choice. Let us assume two other possible measurement directions, $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$, such that they are each at 120 degrees from each other, i.e.

$$\hat{\mathbf{z}} + \hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2 = 0. \tag{3}$$

Since any direction of the SG apparatus will result in quantized spin (we assume the source is a proper mixture), we have at once that $\mu \cdot \hat{\mathbf{z}}$, $\mu \cdot \hat{\mathbf{e}}_1$, and $\mu \cdot \hat{\mathbf{e}}_2$ have values ± 1. But

$$\mu(\omega) \cdot \hat{\mathbf{z}} + \mu(\omega) \cdot \hat{\mathbf{e}}_1 + \mu(\omega) \cdot \hat{\mathbf{e}}_2 = \mu \cdot (\hat{\mathbf{z}} + \hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2), \tag{4}$$

and we reach an apparent contradiction, since the left hand side can only take values $\pm 3, \pm 1$ but the right hand side is zero because of (3). It is clear that the contradiction comes from the assumption that the ω for each experiment (for the three different directions) is the same, since this is what is required to go from the left to the right hand side of (4). Thus, if we assume that spin $\mu(\omega)$ exists before the measurement, the act of measuring it in one direction, say $\hat{\mathbf{x}}$, changes it to a new $\mu(\omega)$, and the process of measurement does not "reveal" the actual state, but instead changes it to a new state with different properties from the original one. The relationship between μ and $\omega \in \Omega$ can be thought as a hidden variable theory of the outcomes of spin, using the terminology of physics [28], and the dependency of ω on the experimental setup makes this theory contextual. So, even in the simple case of having consecutive measurements of spin, we see the impossibility of treating outcomes of experiments as context-independent.

But perhaps the best-known example of contextuality (again, in the sense used above) in quantum mechanics is the famous double-slit experiment [12]. In classical physics, the double-slit experiment, attributed to Thomas Young [29], was used to demonstrate interference, thus "falsifying" the corpuscular theory of Newton and supporting the wave theory of light. In it, light impinges on a barrier where two small and parallel slits are cut, allowing a small amount of the light to go through and reach a screen at the other side (see Figure 1). Because of its wave character, an interference

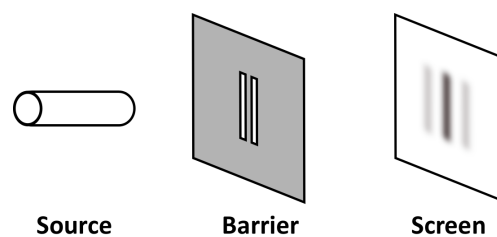**Source**   **Barrier**   **Screen**

Fig. 1.   Double-slit experiment. A source on the left sends light onto a barrier with two slits cut close to each other. An interference pattern appears on the screen to the right.

pattern emerges at the screen, due to differences of phase at each screen location in a way consistent with the geometry of the setup. This interference pattern seems to be incompatible with the Newtonian particle theory of light, since particles arriving at a point on the screen did not interact with both slits, and therefore a concept of phase difference for particles would be meaningless.

J. A. de Barros, G. Oas

In its quantum mechanical version, the double slit needs to be examined in a new light (no pun intended). As is well-known, Einstein used the notion that light was made of particles to explain the photo-electric effect, and evidenced mounted in the early days that light was actually made of particles, called photons [27]. But if light is made of particles, what should we conclude from Young's experiment? An initial hypothesis was that interference was a collective effect of many photons, similar to sound waves being a collective effect of many atoms. However, this idea showed to be inconsistent with experiments where the light intensity was so low that effectively only one photon at a time was present between the slits and the screen (or photographic plate, in this case). So, photons seem to present self-interference, a quite mysterious property. In fact, in his famous Lectures on Physics, Richard Feynman stated that the double-slit experiment contained the "only mystery" of quantum mechanics [30].

To see how the double-slit experiment provides contextuality, let us examine it in a simplified version. In the double-slit experiment, photons are detected on a screen, thus providing a continuum of locations, with a corresponding complicated mathematical description (see [31, 32] and references therein). But what makes the double slit mysterious to Feynman are interference effects, and interference can be studied without resorting to such continuum, in a setup called the Mach-Zehnder interferometer (MZI). So, here we analyze in more detail the contextuality of quantum systems in the MZI.

In the MZI, a light source is directed toward a beam splitter that divides the beam into two distinct beams of equal intensity (see Figure 2). Beam splitters have the important characteristic that light beams reflected by them are phase-shifted by $\pi/2$, whereas the beam's phase going through is not affected, and therefore a $\pi/2$ phase difference exists between beams in arm A and B of Figure 2. After the first beam splitter, some mirrors redirect the beams to another beam splitter, and the beams are recombined, adding once again a $\pi/2$ phase to the reflected beam. Mathematically, we can describe each beam impinging on the second BS with a sine function

$$\psi_A = \frac{A}{2} \sin \omega t,$$
$$\psi_B = \frac{A}{2} \sin \left(\omega t + \frac{\pi}{2} \right) = \frac{A}{2} \cos \omega t,$$

where A is the amplitude of the source S and ω its frequency. After the
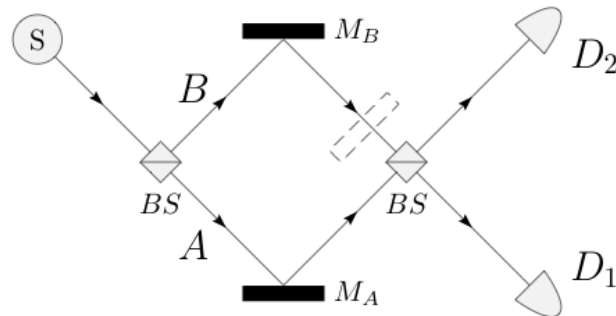
Fig. 2.   Mach Zehnder Interferometer (MZI). A light source S sends a light beam that impinges on the first beam splitter BS. The beam is them divided by BS into equal-intensity beams that travel to both arms (paths) A and B of the interferometer, reflecting on mirrors M_A and M_B, respectively. The beams from arms A and B are then recombined in the second beam splitter. The outcomes are the two beams detected at D_1 and D_2.

second beam splitter, we have

$$
\begin{aligned}
\psi_{D_1} &= \frac{A}{2}\sin\left(\omega t + \frac{\pi}{2}\right) + \frac{A}{2}\cos\omega t \\
&= A, \\
\psi_{D_2} &= \frac{A}{2}\sin\omega t + \frac{A}{2}\cos\left(\omega t + \frac{\pi}{2}\right) \\
&= 0.
\end{aligned}
$$

We then see the main characteristic of the MZI: interference effects give a zero amplitude at D_1 and amplitude A at D_2. We remark that there is an underlying assumption in our derivation above, namely that the length of the interferometer arms A and B were exactly the same. Were they any different, and the phase relations would not match exactly, interference would not be perfect as above (with 0 and A amplitudes)[a].

The above description is the classical one for light waves. But what

---

[a]That is why interferometers are very useful for measuring distances accurately.

J. A. de Barros, G. Oas

happens if we have classical (non-quantum) particles in a similar setup to the MZI? Imagine we send one particle at a time through the MZI. Since no concept of phase or phase relation exists for a classical particle, the first beam splitter would simply reflect it to A or B, with probability $1/2$ to go to A and $1/2$ to B. Once in the second beam splitter, the particle (coming from *either* A or B) would be randomly sent to either detector. Thus, for classical particles we should expect probability $1/2$ of observing the particle in D_1 or D_2 (but never on both detectors!).

Now, what to make of the MZI for quantum particles? If we send a single photon through a MZI, a photodetector placed on either arm of it will reveal the characteristic of a particle: a click on either A or B, but never both. Furthermore, if a photodetector is placed on either A or B, the outcomes of a measurement on D_1 and D_2 are exactly what we expect from a particle: the photon shows up on each of those detectors with probability $1/2$. However, if no detectors are placed on A or B, the photon shows zero probability of detection on D_2 and probability 1 on D_1. In other words, in the absence of detectors on A or B, the photon behaves as if it were a wave, carrying information about the relative phases of the MZI's geometry.

The disturbing aspects of this wave/particle duality for photons has been discussed at length for almost one hundred years, and the interested readers are referred to the many excellent sources (we can particularly recommend the historical account found in Abraham Pais's volume [27]). Here we focus only on the contextual aspects of it. To see them, we start with two ± 1-valued random variables, $\mathbf{P}$ and $\mathbf{D}$, representing the which-path information and detection, respectively. $\mathbf{P}$ is defined such that $\mathbf{P} = 1$ if the particle is detected on A and -1 otherwise, whereas $\mathbf{D}$ is defined such that $\mathbf{D} = 1$ if the photon was detected in D_1 and -1 if detected in D_2.

We have for the MZI two different experimental conditions: one in which no detector is placed on A or B and another where a detector is placed on either A or B (or both), thus yielding which-path information encoded in the outcomes of $\mathbf{P}$. Measuring $\mathbf{D}$ under the no-which-path condition results in $E(\mathbf{D}) = 1$, whereas measuring it together with $\mathbf{P}$ gives as marginal expectation the result $E(\mathbf{D}) = 0$. Thus, according to the above definition of contextuality, the random variable $\mathbf{D}$ is contextual[b].

The double-slit experiment provides a dramatic type of contextuality,

---

[b]Dzhafarov and Kujala refer to this type of strong contextuality as *direct influences*, meaning that the measurement of $\mathbf{P}$ directly influences the outcomes of random variable $\mathbf{D}$ [8, 9]. They, on the other hand, reserve the label *contextual* to refer to other cases where the influences of $\mathbf{P}$ over $\mathbf{D}$ cannot be accounted for by direct influences.

but it is not as surprising as Feynman makes it seem. In fact, not only can we "solve" the mystery of quantum mechanics for this case, if we were to accept a (contextual) hidden-variable theory such as Bohm's [33, 34][c], but we can also clearly understand the possibility of a direct influence from **P** to **D**. In fact, that is exactly what Heisenberg tried to do with his analysis of the wave/particle duality in the double-slit experiment by using what is now known as the Heisenberg microscope [35]. Thus, even though it has at its core one of the main characteristics of QM, i.e. the interference of particles, it is far from containing its only mystery.

Perhaps a deeper mystery comes from a variation in the experiment proposed by Einstein, Podolsky, and Rosen's seminal 1935 paper on the completeness of quantum mechanics [14], now simply known as EPR. In Bohm's version of the EPR argument [15], two correlated photons[d] in the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|++\rangle + |--\rangle \right), \qquad (5)$$

where $|++\rangle$ corresponds to the state where both photons have vertical polarization and $|--\rangle$ horizontal, are sent in two different directions. In one direction, an experimenter, Alice, chooses whether to measure the linear polarization of the photon or not, and if she does, she either observes + or − (see Figure 3). In the other direction, Bob can also measure polarization

Alice   **Source**   **Bob**

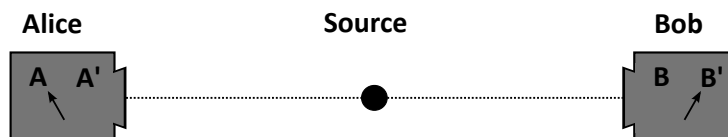A A'   ●   B B'

Fig. 3.   Bell-EPR experiment. A source emits two photons, one toward Alice's lab and another toward Bob's. Each experimenter can make a decision on which direction of spin to measure, represented in the figure by the settings A and A' for Alice and B and B' for Bob. Outcomes of measurements are ± 1, with equal probabilities.

in the same direction as Alice. However, if Alice already did, Bob does not need to, because if Alice measures − Bob knows for sure that the photon getting to him will also be − (due to the correlations contained in the state (5)).

---

[c]Such hidden variable theory provides a mechanism that accounts for the experimental outcomes of the double-slit experiment. As expected, Bohm's theory cited here is a contextual theory, in the sense that the hidden variable needs to be context-dependent.
[d]Bohm's version actually used spin 1/2 particles, not photons, but for our purposes we can use photons.

J. A. de Barros, G. Oas

Up to now there is no mystery from the correlated outcomes of Alice and Bob. What we know is that for state (5) the outcomes of experiments for Alice are the same as for Bob's. But that, according to EPR, presents a problem for QM, since Alice and Bob can perform their measurements as far away from each other as possible. If they do, argue EPR, then it is possible to determine the value of the spin for Bob's particle, say, without ever affecting its state, since all we need to do is use Alice's value. Alice's measurement does not affect the state at Bob's because special relativity forbids faster-than-light interactions. So, continue EPR, the values of the polarization for both particles must come from some state of the system that is not represented in $|\psi\rangle$, and therefore the QM description of nature is incomplete.

A theory that completes QM, in the sense given by EPR, is called a hidden-variable theory, and it so happens that (local) hidden-variable theories are not compatible with the predictions of QM. The first person to point out an empirical incompatibility between QM and (local) hidden variable theories was John Bell. In a seminal paper [36], Bell derived a set of inequalities that were necessary for local hidden variable theories, and proceeded to show that for certain quantum states those inequalities were violated. More than a decade later, Aspect, Grangier, and Gérard [37] showed, in a *tour de force* experiment where for the first time correlations between spacelike separated events were recorded, that the quantum mechanical predictions were correct.

To understand Bell's results, let us examine the setup he discussed. It can be shown that a hidden-variable λ explaining the experimental outcomes of polarization for Alice and Bob exists if and only if a joint probability distribution exits for all possible outcomes [38, 39]. To show that a hidden-variable theory is not compatible with QM, we need to show an QM example that does not allow a joint probability distribution. However, this setup needs to have some constraints, since we saw that the double-slit experiment does not have a joint probability but is also compatible with a (contextual) hidden-variable theory. A suggestion for this constraint is present in EPR's example: the outcomes of a variable **A** cannot be (superluminally) influenced by what happens at another far away location. In other words, from EPR's point of view relativity theory is incompatible with events that are contextual and spacelike separated.

We now proceed to show Bell's argument and setup. Imagine that we have now two possible experimental (and incompatible) measurements of polarization for Alice in two different directions, and the same for Bob (not

necessarily the same directions as Alice). Let us represent the outcomes of measurements with ± 1-valued random variables, namely $\mathbf{A}$ and $\mathbf{A}'$ for both of Alice's directions, and $\mathbf{B}$ and $\mathbf{B}'$ for Bob's (see Figure 3). We can construct a random variable $\mathbf{S}$ defined simply by

$$\mathbf{S} = \mathbf{AB} + \mathbf{A}'\mathbf{B} + \mathbf{AB}' - \mathbf{A}'\mathbf{B}'. \tag{6}$$

If a joint probability exists, for each $\omega \in \Omega$ there are associated outcomes for $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{B}$, and $\mathbf{B}'$ and a corresponding probability. Since $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{B}$, and $\mathbf{B}'$ are ± 1-valued random variables, it follows that for all possible combinations of values for $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{B}$, and $\mathbf{B}'$ the value of $\mathbf{S}$ is either 2 or -2 (for example, if $\mathbf{A} = 1$, $\mathbf{A}' = -1$, $\mathbf{B} = -1$, and $\mathbf{B}' = 1$, then $\mathbf{S} = -1 + 1 + 1 + 1 = 2$). Therefore, the expected value of $\mathbf{S}$ must be a number between -2 and 2, i.e.

$$-2 \le \langle \mathbf{AB} \rangle + \langle \mathbf{A}'\mathbf{B} \rangle + \langle \mathbf{AB}' \rangle - \langle \mathbf{A}'\mathbf{B}' \rangle \le 2. \tag{7}$$

This inequality is one of the Clauser-Horne-Shimony-Holt (CHSH) inequalities [40] (the others are obtained simply by moving the "$-$" sign in (6) to other terms), and a joint probability distribution exits if and only if they are satisfied [41]. It is possible to show that, for a proper choice of angles of measurement for Alice and Bob, their observed correlations result in an $\langle \mathbf{S} \rangle = 2\sqrt{2}$, which violates (7). Therefore, no joint probability distribution exists, and as consequence, no hidden-variable theory exists that explains the correlations between the observables. Furthermore, because no joint exists, the assumption that an $\mathbf{A}$ under experimental condition where it is measured together with $\mathbf{B}$ (as in $\langle \mathbf{AB} \rangle$) is the same as when it is measured with $\mathbf{B}'$ is not correct: the system $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{B}$, and $\mathbf{B}'$ is contextual.

It is hard to overplay the importance of Bell's results. For example, Henry Stapp famously stated it to be "the most profound discovery of science" [42]. The reason is that Bell's theorem clearly shows that far away measurements indeed affect the outcomes of a nearby measurement for entangled systems. In other words, quantum mechanics is non-local, which seems to be incompatible with the principles behind relativity [43]. Furthermore, both quantum mechanics and relativity are tremendously successful theories, from an empirical point of view.

As we saw, Bell's setup differs significantly from the double-slit experiment. First, it does not allow for direct influences between the observable quantities, which if allowed would be in direct conflict with special relativity. Instead, the absence of a joint probability distribution (and therefore of a non-contextual hidden-variable theory) comes from the non-trivial

correlations imposed by QM (and experimental observations). But, more importantly, Bell's setup provides a case where a system with two (or more) parts can have such non-trivial correlations even when the measurements of those parts are spacelike separated.

We end this section with one last example of contextuality, the Kochen-Specker paradigm [16]. In quantum mechanics, projection operators[e] constitute the "simplest" type of measurement possible. For example, for a spin 1/2 particle, its Hilbert space is two dimensional. In this Hilbert space, the projector $P_z = |+\rangle\langle+|$ is an observable with eigenvalues 0 and 1, corresponding to not having and having the property "spin + in direction $\hat{\mathbf{z}}$." So, projection operators correspond to measurements whose outcomes tell you whether the system has the property measured (0) or not (1).

In their paper, Kochen and Specker asked whether it is possible to assign values 0 or 1 to projection operators in a way that is consistent. To show that this is not possible, they used a Hilbert space of dimension three and a total of 117 projectors. However, to understand how their results come about, we show here a simpler version of 18 projectors in a four dimensional Hilbert space, due to Cabello, Estebaranz, and Alcaine [44]. In this version, we have the set of projection operators P_i with corresponding dicotomic random variables $\mathbf{V}_i$ taking values 0 or 1 depending on whether the property is false or true[f]. The index i in P_i corresponds to a vector in the four dimensional space where P_i projects onto. Below is a list of combinations of random variables, and we see that each line corresponds to a set of projectors that commute, and can therefore be measured simultaneously, though this is not true for projectors in different lines.

---

[e]A Hermitian operator P is a projection operator if $P^2 = P$.

[f]Technically, it is not necessary to use random variables, since the Kochen-Specker example uses deterministic (probability one) events. The use of random variables extends this setup to more realistic situations, where probability one events are never observed [45, 46].

$$\mathbf{V}_{0,0,0,1} + \mathbf{V}_{0,0,1,0} + \mathbf{V}_{1,1,0,0} + \mathbf{V}_{1,-1,0,0} = 1,$$
$$\mathbf{V}_{0,0,0,1} + \mathbf{V}_{0,1,0,0} + \mathbf{V}_{1,0,1,0} + \mathbf{V}_{1,0,-1,0} = 1,$$
$$\mathbf{V}_{1,-1,1,-1} + \mathbf{V}_{1,-1,-1,1} + \mathbf{V}_{1,1,0,0} + \mathbf{V}_{0,0,1,1} = 1,$$
$$\mathbf{V}_{1,-1,1,-1} + \mathbf{V}_{1,1,1,1} + \mathbf{V}_{1,0,-1,0} + \mathbf{V}_{0,1,0,-1} = 1,$$
$$\mathbf{V}_{0,0,1,0} + \mathbf{V}_{0,1,0,0} + \mathbf{V}_{1,0,0,1} + \mathbf{V}_{1,0,0,-1} = 1,$$
$$\mathbf{V}_{1,-1,-1,1} + \mathbf{V}_{1,1,1,1} + \mathbf{V}_{1,0,0,-1} + \mathbf{V}_{0,1,-1,0} = 1,$$
$$\mathbf{V}_{1,1,-1,1} + \mathbf{V}_{1,1,1,-1} + \mathbf{V}_{1,-1,0,0} + \mathbf{V}_{0,0,1,1} = 1,$$
$$\mathbf{V}_{1,1,-1,1} + \mathbf{V}_{-1,1,1,1} + \mathbf{V}_{1,0,1,0} + \mathbf{V}_{0,1,0,-1} = 1,$$
$$\mathbf{V}_{1,1,1,-1} + \mathbf{V}_{-1,1,1,1} + \mathbf{V}_{1,0,0,1} + \mathbf{V}_{0,1,-1,0} = 1.$$

Now, since $\mathbf{V}_i$ is either 0 or 1, we can sum all the random variables on the left hand side of the previous set of equations, and because each random variable appears exactly twice, this must be an even number. However, because we have only 9 equations, the sum of the right hand side yields an odd number. This is clearly a contradiction, and as with the spin case at the beginning, the problem comes from identifying a random variable (say, $\mathbf{V}_{0,0,0,1}$) in a given experimental context (say, when measured together with $\mathbf{V}_{0,0,1,0}$, $\mathbf{V}_{1,1,0,0}$, $\mathbf{V}_{1,-1,0,0}$, as in line one) with the same variable in a different experimental context (in our example, $\mathbf{V}_{0,1,0,0}$, $\mathbf{V}_{1,0,1,0}$, $\mathbf{V}_{1,0,-1,0}$, as in line two). In other words, the Kochen-Specker theorem shows that the algebra of observables in QM is such that it is impossible to assign non-contextual values to certain properties of a system, independent of what the system is.

To summarize, in this section we presented several different examples of contextual quantum systems, per our definition of contextuality. We see that each example presents its own subtle issues. In the double-slit experiment, contextuality comes mainly from direct influences on the detection random variable due to a measurement that leads to which-path information[g]. The direct influence arises from the choice of successively measuring incompatible observables, D and P. In one case, observable D is measured first (i.e., without P), and the wave function arriving to it leads to interference. In the other case, P is measured first, collapsing the wave function, and leading to a different quantum state reaching D. So, P directly influ-

---

[g]Though direct influences may not account for the totality of contextual effects. See [47] for an example in the context of the Leggett-Garg experiment, which is formally related to the double-slit experiment.

ences D. To distinguish it from the other cases, we call this *contextuality by direct influences.*

In the Bell-EPR experiment, direct influences are forbidden by special relativity, and are not observed. Because two observables, say A and B', commute, the changes to the wave function made by A do not affect the outcomes of B'. However, the *correlations* between A and B' are affected by their simultaneous measurement. So, the most striking feature of Bell's setup was the contextual dependency of outcomes of experiments in systems that are *spacelike* separated, suggesting some type of superluminal influence [43], or, as Einstein called it, "*spukhafte Fernwirkung*" (spooky action at a distance). The contextuality for observables that are spacelike separated are called *nonlocal contextuality,* or simply *nonlocality.*

Finally, the Kochen-Specker theorem showed that the algebra of observables is such that the random variables representing the outcomes of any state of a measurable experimental system will present contextuality. This is a fact that comes simply from the observables themselves and, contrary to the Bell-EPR experiment, which requires an entangled state, has nothing to do with the state of the system being measured (i.e., is state independent). Consistently with the physics literature, we refer to this as *state-independent contextuality* or simply *contextuality.*

We emphasize that all the physical systems discussed here satisfy the criteria of contextuality put forth in the Introduction. From a *device independent* framework [48, 49], where experiments involve black boxes with local inputs (settings for a measurement) and outputs (measurement outcomes), we look only at random variables. Then, in this framework, the two types of contextuality would be contextuality by direct influences and contextuality not by direct influences. So, the double-slit experiment would fall into the category of contextuality by direct influences, whereas both Bell-EPR and Kochen-Specker would be contextual not by direct influences. The distinction between Bell-EPR and Kochen-Specker (i.e., local vs. nonlocal, state-dependent vs. state-independent) is based on physical principles, and not on probabilistic principles alone.

We saw that exemplary physical systems exhibit contextuality, but from a physical point of view their contextuality is different. This leads to the definition of at least three types of contextuality: contextuality by direct influences, non-locality, and state-independent contextuality. In the next section we examine quantum cognition, and discuss how quantum contextuality shows up in them.

3. Contextuality in quantum cognition

In this section we will examine some examples from the burgeoning field of quantum cognition. Quantum cognition is the use of the mathematical formalism of quantum mechanics to model cognitive processes. As such, it should not be confused with the idea that in order to describe cognition (or consciousness) we need to use quantum mechanics, as is espoused by Penrose and Hammerof [50–52], or by Stapp [53–56], among others. To researchers in quantum cognition, the quantum comes from the use of the contextual probability theory given by the Hilbert space formalism to describe cognitive systems, but the underlying processes that govern it can be classical [57–59].

This section is not intended to be an exhaustive review of this field, but instead to provide an example to illustrate some of the main features of quantum cognition models. Quantum models were used to model the conjunction and disjunction paradoxes [60–63], the Ellsberg paradox [22, 64, 65], order effects [66–68], similarity effects [69], and the Guppy effect [70, 71], to name a few. The interested reader is referred to many of the useful reviews on the subject, such as Khrennikov [26], Busemeyer and Bruza [17], Khrennikov and Haven [18], and Ashtiani and Azgomi [72].

We focus on one of the first applications of the quantum formalism, the modeling of the violation of Savage's Sure-Thing Principle (STP). Simply put, Savage's STP states that if a person holds the subjective view that A is preferred over $\neg A$ if B is true[h], but is also preferred if B is false, then A should be preferred regardless of whether the person knows which is true, B or $\neg B$. In Savage's own words [19, pg. 21], "A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he should buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say. It is all too seldom that a decision can be arrived at on the basis of the principle used by this businessman, but, except possibly for the assumption of simple ordering, I know of no other extralogical principle governing decisions that finds such ready acceptance." Formally, let us

---

[h]The notation $\neg A$ means "not A."

imagine a ± 1-valued random variable $\mathbf{A}$ ($\mathbf{A} = 1$ is buy and $\mathbf{A} = -1$ is not buy), and let another ± 1-valued variable $\mathbf{B}$ be 1 if a Democrat wins -1 if a Republican wins. If $\mathbf{A} = 1$ is preferred over $\mathbf{A} = -1$ when $\mathbf{B} = 1$ and also when $\mathbf{B} = -1$, then $\mathbf{A} = 1$ is always preferred.

Even though most people would agree with Savage that his principle "finds such ready acceptance," in some experimental conditions, human decision makers violate it. For example, in Tversky and Shafir's 1992 paper [20], participants were told about a two-step gamble. In the first step, which was compulsory for all players, there was a 50/50 probability of winning $200 or loosing $100. The second step was not compulsory, and the person could choose whether or not to make a second gamble with the same odds and payoffs. After winning the first bet, 69% of participants chose to place a second gamble, and after loosing 59% also chose to gamble a second time. In terms of probabilities, we have

$$P(\text{``gamble again''}|\text{``won''}) = 0.69 > P(\text{``not gamble again''}|\text{``won''}) = 0.31,$$

$$P(\text{``gamble again''}|\text{``lost''}) = 0.59 > P(\text{``not gamble again''}|\text{``lost''}) = 0.41.$$

Since "gamble again" is preferred over "not gamble again" for both situations, "won" or "lost" the first step, STP tells us that "gamble again" should be preferred over "not gamble again." However, in a later time, the same participants were asked about the second gamble, but this time they were not told whether they won or lost the first step. Under this unknown condition, 64% of the participants rejected the second gamble. But this corresponds to

$$P(\text{``gamble again''}) = 0.36 < P(\text{``not gamble again''}) = 0.64,$$

a clear violation of the STP.

Violations of STP are violations of the standard calculus of probability. To see this, imagine you have two sets, A and B, and we define the conditional probability of A given B as

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

for $P(B) \neq 0$. In this notation, the STP conditions are equivalent to

$$P(A \cap B) > P\left(\overline{A} \cap B\right), \quad P\left(A \cap \overline{B}\right) > P\left(\overline{A} \cap \overline{B}\right), \tag{8}$$

where $\overline{A}$ denotes the complement of A. But from the calculus of probability,

$$P(A \cap B) + P\left(A \cap \overline{B}\right) = P(A)$$

and

$$P\left(\overline{A} \cap B\right) + P\left(\overline{A} \cap \overline{B}\right) = P\left(\overline{A}\right),$$

and adding each term in (8) we obtain

$$P\left(A\right) > P\left(\overline{A}\right).$$

Notice that the violation of the law of probabilities happens because we assume there is a joint probability distribution for A and B, which of course in this case is clear there should be, since we are dealing with the conditional probability of A given B as something given subjectively by the decision maker.

How can we model such violations of classical probability theory with a quantum formalism? The answer to this particular case is given by the quantum description of the double-slit experiment [22, 63, 70] in the MZI paradigm. In the MZI, let us have the statement "detector D_1 is preferred over D_2" as corresponding to a higher probability of detecting a particle in D_1 instead of D_2. In the notation of Section 2, this corresponds to

$$p\left(\mathbf{D} = 1\right) > p\left(\mathbf{D} = -1\right). \tag{9}$$

The which-path information is given by $\mathbf{P}$, and for the MZI we have

$$\begin{aligned}
p\left(\mathbf{D} = 1|\mathbf{P} = 1\right) &= p\left(\mathbf{D} = -1|\mathbf{P} = 1\right) \\
&= p\left(\mathbf{D} = 1|\mathbf{P} = -1\right) \\
&= p\left(\mathbf{D} = -1|\mathbf{P} = -1\right) \\
&= \frac{1}{2}.
\end{aligned}$$

Similarly to violations of STP, because of the symmetry of the probabilities[i], we have no reason to prefer $\mathbf{D} = 1$ over $\mathbf{D} = -1$ or *vice versa*, as from the symmetry and classical probability theory we have that $p\left(\mathbf{D} = 1\right) = p\left(\mathbf{D} = -1\right)$, in disagreement with (9). So, to model a violation of STP for human decision makers, all we need to do is map the MZI, with responses "gamble again" or "not gamble again" corresponding to $\mathbf{D} = 1$ and $\mathbf{D} = -1$, and which path information corresponding to "won" or "lost." The "won" and "lost" intermediate states are though as mental states that "collapse" once the decision maker becomes knowledgeable of the outcome of the first gamble.

---

[i]This symmetry is not necessary, and we put it here to make the argument simpler and to have a direct connection to the experimental setup shown in Section 2. The MZI can be modified to introduce biases that would make the probabilistic system non-symmetric, but would still lead to a violation of STP (using the same mapping as we have here).

We remark that, as far as we know, all the quantum cognition models have similar characteristics to the MZI. What we mean by this is the following. In the MZI the violation of classical probability theory comes from two incompatible sets of experimental data due to different contexts. In one context, which-path information is not present, and the quantum state reaching detectors D_1 and D_2 is in a superposition with components from both paths. In the other context, a which-path measurement is performed, and the wave function collapses, with a corresponding loss of quantum superposition. So, differences in probabilities usually come from collapse/no-collapse of the wave function due to a measurement.

Some researchers have suggested that other quantum-like effects may exist in cognitive systems, such as entanglement [70, 73, 74]. Entanglement comes from states such as (5), where for N-partite systems ($N \geq 2$), the outcomes of a property of a subsystem are connected to another property of another subsystem in a way that cannot be explained by common causes (i.e., hidden variables), as it was the case with the Bell-EPR setup. We have argued elsewhere that, because we cannot rule out other physical mechanisms, such types of entanglement are not as unexpected as the quantum mechanical ones [75], and in fact can be derived by classical-like models [59]. Exactly because of this reason, there are no principles denying violations of the no-signaling condition (corresponding to direct influences between different subsystems), as we have in actual quantum systems [76]. Furthermore, it seems that most of the cases of violations of inequalities such as (7) also violate a form of the no-signaling condition, and no contextuality from entanglement is detectable, suggesting that direct influences are more important in quantum cognition [77].

We can also emphasize that, in quantum cognition, violations of the CHSH inequalities do not necessarily mean nonlocality. To demonstrate non-locality, one needs to not only show violations of the CHSH of events that are spacelike separated (a seemingly impossible task for cognitive events), but also that such violations are not subject to standard loopholes [78]. To see how difficult this task is, attempts to create a loophole free test of nonlocality for quantum systems have yet to be successful, even after many decades of intensive research [79]. One can only imagine the technical and conceptual difficulties that would make it hopeless to show nonlocality for cognitive systems. Other difficulties are also present in the Kochen-Specker system [80]. So, from an empirical point of view, it seems that the predominant "quantum" effect in cognition is related to the MZI.

Let us end this Section with one important example. To motivate it, let

us recall that quantum cognition relies on using the quantum mechanical mathematical apparatus to social systems. However, this seems too constraining, leaving out many situations that would not be describable by the formalism. As an example, mentioned in the previous paragraph, there are no reasons to require cognitive systems to satisfy the no-signaling condition, and we even have evidence that it is violated for some cognitive systems. However, the no-signaling condition is a direct consequence of the quantum formalism: we can derive it from the structure of the Hilbert space. Furthermore, other forms of "superluminal" communications are strictly forbidden by quantum mechanics. This is the case with the no-cloning theorem [81]; if cloning were possible, one could devise a method of sending communications between Alice and Bob in the EPR setup discussed above. But we have no *a priori* reason to rule out state cloning for social systems. So, is the quantum mechanical apparatus too constraining?

As a toy example, we refer back to the ± 1-valued random variables, $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$, discussed in the Introduction. It is possible to concoct artificial (but reasonable) cases where those random variables have no joint probability distribution, presenting contextuality [24]. Furthermore, it is also possible to show that, under reasonable assumptions, neural models that lead to similar outcomes described by quantum mechanics [59, 82, 83], may also generate correlated variables $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ that have no joint [23, 25]. However, such simple example cannot be described by quantum mechanics (unless we create a contrived model of it [24]), since the existence of observed correlations corresponds to pairwise commutations between the quantum operators representing $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$, and from the algebra of operators it follows that all three variables $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are simultaneously observable. Since they are all simultaneously observable, a joint probability distribution must exist. So, the quantum formalism rules out situations such as those described in [23–25][j].

To summarize, we sketched in this section how the formalism of quantum mechanics is often used in quantum cognition. We claimed that among the many different cases where contextuality shows up in quantum mechanics, it seems that the only relevant case may be the double-slit experiment. We also saw that the quantum formalism may present too many restrictions to certain contextual situations. In the next section, we discuss an alternative formalism that we have proposed in previous papers: negative probabilities.

---

[j]For a more detailed discussion of the neural model and its connection to quantum cognition and to the issues mentioned in this paragraph, the reader is referred to [84].

4. Describing contextual probabilities

We saw that contextuality is a key factor in quantum cognition, and the main push for using the quantum mechanical mathematical apparatus was the better fit it provided for certain experiments. This should not come too much as a surprise, as this apparatus was developed to deal with systems that are contextual, such as the double-slit experiment. But we also saw that there may be cases where quantum mechanics imposes too many restrictions that would make its Hilbert space formalism inadequate to represent them. So, the question is how to develop a theory of probabilities that have the same ability to describe contextual systems as quantum mechanics, but also has the flexibility of describing the systems discussed above.

There are many attempts to describe quantum contextual systems, such as contextuality by default [3, 7, 8] or upper and lower probabilities [85–88]. Here we present a possible theory, first appearing in Physics in the works of Wigner [89], but later on considered more seriously by Dirac [90] and Feynman [91] (for a historical but not up-to-date survey of negative probabilities in physics, the reader is referred to [92]).

As mentioned, negative probabilities showed up in quantum mechanics, when Wigner asked which joint probability distributions for momentum and position would result in the same outcomes predicted by quantum statistical mechanics [89]. When such joint probability distributions were computed for some physical systems (see, e.g. [93]), it became clear that they could take negative values, and were therefore discarded as non-physical probabilities (Wigner called them quasi-probability distributions). Though agreeing with Wigner's claim of no physical meaning, Dirac thought negative probabilities could be as useful as negative numbers were in mathematics, and attempted to apply them to the description of quantized fields [90], with no success. Decades later, Feynman also tried to use negative probabilities, but, to his disappointment, thought that they did not offer any new insights or results in quantum-field theory [91]. However, since then, some researchers have been using negative probabilities to help understand certain physical systems, mainly when violations of classical probabilities occur because of contextuality [76, 94–105].

Let us start by defining negative probabilities (we follow [84, 106]). We start with a preliminary definition.

Definition 1 (compatibility). Let Ω be a finite set, $\mathcal{F}$ an algebra over Ω,

and let $(\Omega_i, \mathcal{F}_i, p_i)$, $i = 1, \ldots, n$, a set of n probability spaces, $\mathcal{F}_i \subseteq \mathcal{F}$ and $\Omega_i \subseteq \Omega$. Then $(\Omega, \mathcal{F}, p)$, where p is a real-valued function, $p : \mathcal{F} \to [0, 1]$, $p(\Omega) = 1$, is *compatible* with the probabilities p_i's iff

$$\forall (x \in \mathcal{F}_i) (p_i(x) = p(x)).$$

The marginals p_i are called *viable* iff p is a probability measure.

The idea of the previous definition is that for contextual systems, our observations are always in subspaces of a larger sample space Ω. If it is not possible to put all the observed marginals in those systems in a single space, then the marginals are not *viable*[k], i.e. there does not exists a joint probability distribution over Ω that explains all correlations.

In QM (and, perhaps, social sciences), the marginals are not always viable. This means that no proper joint probability distribution exists, but perhaps a real-valued function (but sometimes negative) p exists that provides all the correct marginals. This p, if normalized, would be a negative probability.

Definition 2 (negative probabilities). Let Ω be a finite set, $\mathcal{F}$ an algebra over Ω, P and P' real-valued functions, $P : \mathcal{F} \to \mathbb{R}$, $P' : \mathcal{F} \to \mathbb{R}$, and let $(\Omega_i, \mathcal{F}_i, p_i)$, $i = 1, \ldots, n$, a set of n probability spaces, $\mathcal{F}_i \subset \mathcal{F}$ and $\Omega_i \subseteq \Omega$. Then $(\Omega, \mathcal{F}, P)$ is a negative probability space, and P a negative probability, if and only if $(\Omega, \mathcal{F}, P)$ is compatible with the probabilities p_i's and

N1. $\forall (P') \left(\sum_{\omega_i \in \Omega} |P(\{\omega_i\})| \leq \sum_{\omega_i \in \Omega} |P'(\{\omega_i\})| \right)$

N2. $\sum_{\omega_i \in \Omega} P(\{\omega_i\}) = 1$

N3. $P(\{\omega_i, \omega_j\}) = P(\{\omega_i\}) + P(\{\omega_j\}), \quad i \neq j.$

In the above definition, the standard axiom of nonnegativity [107] is replaced with a minimization of the L1 norm of P (we use P for negative joint probability distributions, and p for proper probability distributions). Intuitively, we minimize the L1 norm to find a quasi-probability distribution that is as close as possible to a proper probability distribution, since relaxing the nonnegativity axiom leads to an infinite number of quasi-probabilities consistent with the marginals. The value of the *minimum L1 probability norm* is denoted M^*, and is mathematically given by

---

[k]A term coined in Reference [100].

$M^* = \sum_{\omega_i \in \Omega} |P(\{\omega_i\})|$. This is a useful quantity, since P is a proper probability (and therefore $(\Omega, \mathcal{F}, P)$ is a probability space) if and only if $M^* = 1$ [106]. Furthermore, we can think of M^* as a measure of contextuality: the larger its value, the more contextual the system [5, 76].

Not all contextual systems allow for negative probabilities. For instance, in references [76, 97, 99, 105] it was independently proven that negative probabilities exist if and only if the marginals p_i do not allow for direct influences (we called such systems contextually biased in [84]). An example of a system that allows for direct influences are the MZI and the double slit.

So, we see that negative probabilities are a possible extension of standard probability. It is not clear how this extension can be used to describe, in general, random variables that are directly influenced by others. But, in some of the cases treated in quantum cognition, it is possible, by reasoning in a counterfactual way as to preserve the possibility of identifying random variables in different contexts (see [106, 108] for a detailed analysis of the MZI with negative probabilities). But the question remains as to whether negative probabilities provide any advantages over other approaches.

Before we continue our exposition of negative probabilities, we should address the issue of interpretation, which surely is being asked by many at this point. There are many different ways to interpret negative probabilities, such as Khrennikov's quasi-stochastic p-adic processes [109–111], Abramsky and Brandembuerger's negative and positive types [103], or Szekely's square-root of a coin [112], to mention a few. Here we take a more subjective (and pragmatic) approach, where negative probabilities are seem as a computational device to help establishing truth values to propositions (say, the proposition "the random variable $\mathbf{A}$ has value 1"). As such, the minimization of the L1 norm is nothing but a requirement that this computation should give us a quasi-probability distribution that is as close as possible to a (non-existent) proper probability distribution. In other words, as Feynman and Dirac, we see negative probabilities as a computational device, without necessarily having a meaning.

We now turn to an example, first analyzed in [24], and based on the three ± 1-valued random variables $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$. Imagine a decision maker, here named Deanna, who wants to invest in the stocks of three companies, X, Y, and Z. Knowing nothing about the stock market, Deanna hires three experts, Alice, Bob, and Carlos, to give her advice. Their range of expertise overlaps, but are not the same: Alice knows only about X and Y, but knows nothing about Z; Bob knows about X and Z, but not about Y; and Carlos

knows about Y and Z, but not X. All experts agree that the chances of X, Y, and Z going up are the same as them going down. Alice also tells Deanna that she thinks that whenever X goes up, Y is sure to go down, and vice versa. Bob tells her that whenever X goes one way, Z goes the other way with probability $3/4$, and they go the same way with probability $1/4$. Finally, Carlos tells her that he sees no relationship whatsoever between Y and Z. Associating a $+1$ value of a random variable with stock-values going up, and -1 with going down, Deanna has the following expectations for $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$.

$$E\left(\mathbf{X}\right) = E\left(\mathbf{Y}\right) = E\left(\mathbf{Z}\right) = 0, \tag{10}$$

$$E\left(\mathbf{XY}\right) = -1, \tag{11}$$

$$E\left(\mathbf{XZ}\right) = -\frac{1}{2}, \tag{12}$$

and

$$E\left(\mathbf{YZ}\right) = 0. \tag{13}$$

It is not hard to see, using Suppes and Zanotti's inequalities [38], that there is no joint probability distribution for $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ consistent with the expectations (10)–(13).

How should Deanna proceed? There are some alternatives in the literature, but perhaps the most common one would be the Bayesian approach, where Deanna starts with a prior probability distribution which is updated with the experts opinions [113, 114]. However, as we pointed out elsewhere, this approach has problems. First, the triple moment $E\left(\mathbf{XYZ}\right)$ is invariant under Bayesian updates of the pairwise moments. This means that whatever values of triple moments Deanna starts with, those values are not updated [24]. This lack of update presents problems when we expect Deanna to get information about the triple moment in cases where the three experts agree. For instance, if Alice, Bob, and Carlos all say that the stocks are perfectly correlated, Deanna's update should lead to $E\left(\mathbf{XYZ}\right) = 1$. Furthermore, for certain "inconsistent" (with a joint) correlations given by Alice, Bob, and Carlos, weakening them would lead to a restricted value of the triple moment [84]. But how to extract such information from the inconsistent beliefs given by the experts? When a joint negative probability distribution is constructed, not any distribution is allowed, but only those minimizing the L1 norm. We can show that, for the

24   *J. A. de Barros, G. Oas*

above example, minimizing the L1 norm constrains the values of the triple moment to be in the range

$$-\frac{1}{4} \le E\left(\mathbf{XYZ}\right) \le \frac{1}{2}.$$

Therefore, negative probabilities provide information about the range of values of the triple moment that is not part of the standard Bayesian update. With such range, it should be possible, under certain conditions, to formulate a Dutch book.

5. Final Remarks

In this paper we examined different examples of contextuality in physics, namely the double-slit experiment, in the form of the Mach-Zehnder interferometer (MZI), the Bell-EPR entanglement experiment, and the Kochen-Specker theorem. We argued that among those cases, the one that relates more closely to what is usually done in quantum cognition is the MZI, and that both the Bell-EPR and the Kochen-Specker paradigms have only marginal interest (from an empirical point of view).

We also discussed the restrictions on types of systems that can be modeled by the Hilbert space formalism of quantum mechanics. Such formalism cannot model certain systems that are not, in principle, forbidden by any cognitive or behavioral principle. It also implies constraints such as the no-cloning theorem [81], the no-signaling condition [115], and the monogamy of quantum correlations (from entanglement) [116], to name a few cases. Such constraints are almost necessary for Physical systems, or there would be severe conflicts with the (empirically verified) theory of relativity, but they are not at all necessary for cognitive systems. In fact, the Hilbert space formalism is so restrictive that it even forbids the simple three-random-variable example we showed in Section 4.

With the examples presented, we are not attempting here to discourage the use of the quantum formalism in cognition. We believe that the quantum mathematical structure inspired many interesting results in quantum cognition, and the large volume of papers in the subject attest to its importance. Our goal is instead to point out that there are other tools, such as negative probabilities, that should not be neglected, and perhaps studied side-by-side with the quantum structures (this is the subject of our other paper in this book [10]), and to challenge the quantum interactions community to think about cases where the quantum formalism may be inadequate or (at least) cumbersome. For instance, the examples we presented raise

Contextuality in Physics and Quantum Cognition   25

some questions about the quantum models. Can the (apparent) advantage of negative probabilities, in certain examples, over the Bayesian approach be reproduced with the quantum formalism? What principles would have to be added to them? Do human-decision makers follow a process similar to the minimization of the L1 norm for inconsistent situations? If so, how would such process be described in the quantum formalism? Those are open questions that would need to be addressed.

References

[1]   R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, *Physical Review A.* **71**(5), 052108 (May, 2005).

[2]   J.-Å. Larsson, Necessary and sufficient detector-efficiency conditions for the Greenberger-Horne-Zeilinger paradox, *Physical Review A.* **57**(5), R3145–R3149 (May, 1998).

[3]   E. Dzhafarov and J. Kujala, Contextuality is About Identity of Random Variables, *Physica Scripta.* **T163**, 014009 (2014).

[4]   E. N. Dzhafarov and J. V. Kujala, Contextuality in Generalized Klyachko-type, Bell-type, and Leggett-Garg-type Systems, *arXiv:1411.2244 [physics, physics:quant-ph]* (Nov., 2014). arXiv: 1411.2244.

[5]   J. A. de Barros, E. Dzhafarov, J. Kujala, and G. Oas, Unifying Two Methods of Measuring Quantum Contextuality, *arXiv:1406.3088 [quant-ph]* (June, 2014).

[6]   A. Peres, *Quantum Theory: Concepts and Methods.* Springer Science & Business Media (1995).

[7]   E. Dzhafarov and J. Kujala, Random Variables Recorded under Mutually Exclusive Conditions: Contextuality-by-Default, *arXiv:1309.0962 [quant-ph]* (Sept., 2013).

[8]   E. Dzhafarov and J. Kujala, All-Possible-Couplings Approach to Measuring Probabilistic Context, *PLoS ONE.* **8**(5), e61712 (May, 2013).

[9]   E. Dzhafarov and J. Kujala, A qualified Kolmogorovian account of probabilistic contextuality, *Lecture Notes in Computer Science.* **8369**, 201–212 (2014).

[10]   G. Oas and J. A. de Barros. A Survey of Physical Principles Attempting to Define Quantum Mechanics. In eds. E. Dzhafarov, R. Zhang, and S. M. Jor-

dan, *Contextuality From Quantum Physics to Psychology.* World Scientific (2015).

[11] W. Gerlach and O. Stern, Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld, *Zeitschrift für Physik.* **9**(1), 349–352 (Dec., 1922).

[12] P. A. M. Dirac, *The principles of quantum mechanics.* Clarendon Press, Oxford, UK (1947).

[13] A. Leggett and A. Garg, Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?, *Physical Review Letters.* **54**(9), 857–860 (Mar., 1985).

[14] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, *Physical Review.* **47**(10), 777–780 (May, 1935).

[15] D. Bohm, *Quantum Theory.* Courier Corporation (Apr., 2012).

[16] S. Kochen and E. P. Specker, The Problem of Hidden Variables in Quantum Mechanics, *Journal of Mathematics and Mechanics.* **17**, 59–87 (1967).

[17] J. Busemeyer and P. Bruza, *Quantum models of cognition and decision.* Cambridge Univ. Press, Cambridge (2012).

[18] E. Haven and A. Khrennikov, *Quantum Social Science.* Cambridge Univ. Press, Cambridge (2013).

[19] L. J. Savage, *The foundations of statistics*, 2nd edn. Dover Publications Inc., Mineola, New York (1972).

[20] A. Tversky and E. Shafir, The Disjunction Effect in Choice Under Uncertainty, *Psychological Science.* **3**(5), 305–309 (Sept., 1992).

[21] E. Shafir and A. Tversky, Thinking through uncertainty: Nonconsequential reasoning and choice, *Cognitive Psychology.* **24**(4), 449–474 (Oct., 1992).

[22] A. Khrennikov and E. Haven, Quantum mechanics and violations of the Sure-Thing Principle: The use of probability interference and other concepts, *Journal of Mathematical Psychology.* **53**(5), 378–388 (Oct., 2009).

[23] J. A. de Barros. Joint probabilities and quantum cognition. In eds. A. Khrennikov, A. L. Migdall, S. Polyakov, and H. Atmanspacher, *AIP Conference Proceedings*, vol. 1508, pp. 98–107, American Institute of Physics, Vaxjo, Sweden (Dec., 2012).

[24] J. A. de Barros. Decision Making for Inconsistent Expert Judgments Using Negative Probabilities. Lecture Notes in Computer Science, pp. 257–269. Springer, Berlin/Heidelberg (2014).

[25] J. A. de Barros. Beyond the Quantum Formalism: Consequences of a Neural-Oscillator Model to Quantum Cognition. In ed. H. Liljenström, *Advances in Cognitive Neurodynamics (IV)*, Advances in Cognitive Neurodynamics, pp. 401–404. Springer Netherlands (2015).

[26] A. Khrennikov, *Ubiquitous Quantum Structure.* Springer Verlag, Heidelberg (2010).

[27] A. Pais, *Inward bound: of matter and forces in the physical world.* Oxford University Press, Oxford, UK (1986).

[28] J. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy.* Cambridge University Press (June, 2004).

[29] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of*

Propagation, Interference and Diffraction of Light. Cambridge University Press (Oct., 1999).

[30] R. Feynman, R. Leighton, and M. Sands, *The Feynman lectures on physics: Mainly mechanics, radiation, and heat.* vol. 3, Basic Books, New York, NY (2011).

[31] P. Suppes and J. A. de Barros, A random-walk approach to interference, *International Journal of Theoretical Physics.* **33**(1), 179–189 (1994).

[32] P. Suppes and J. A. de Barros, Diffraction with well-defined photon trajectories: A foundational analysis, *Foundations of Physics Letters.* **7**(6), 501–514 (Dec., 1994).

[33] D. Bohm, A suggested interpretation of the quantum theory in terms of" hidden" variables. I, *Physical Review.* **85**(2), 166 (1952).

[34] D. Bohm, A suggested interpretation of the quantum theory in terms of" hidden" variables. II, *Physical Review.* **85**(2), 180 (1952).

[35] W. Heisenberg, *The Physical Principles of the Quantum Theory.* Courier Corporation (1949).

[36] J. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics.* **1**(3), 195–200 (1964).

[37] A. Aspect, P. Grangier, and G. Roger, Experimental Tests of Realistic Local Theories via Bell's Theorem, *Physical Review Letters.* **47**(7), 460–463 (Aug., 1981).

[38] P. Suppes and M. Zanotti, When are probabilistic explanations possible?, *Synthese.* **48**(2), 191–199 (1981).

[39] P. Suppes, J. A. de Barros, and G. Oas. A Collection of Probabilistic Hidden-Variable Theorems and Counterexamples. In eds. R. Pratesi and L. Ronchi, *Waves, Information, and Foundations of Physics: a tribute to Giuliano Toraldo di Francia on his 80th birthday*, Italian Physical Society, Florence, Italy (Oct., 1996).

[40] J. Clauser, M. Horne, A. Shimony, and R. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, *Physical Review Letters.* **23**(15), 880–884 (Oct., 1969).

[41] A. Fine, Hidden Variables, Joint Probability, and the Bell Inequalities, *Physical Review Letters.* **48**(5), 291–295 (Feb., 1982).

[42] H. P. Stapp, Bell's theorem and world process, *Il Nuovo Cimento B Series 11.* **29**(2), 270–276 (Oct., 1975).

[43] T. Maudlin, *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics.* John Wiley & Sons (May, 2011).

[44] A. Cabello, J. Estebaranz, and G. Alcaine, Bell-Kochen-Specker theorem: A proof with 18 vectors, *Physics Letters A.* **212**(4), 183–187 (Mar., 1996). arXiv:quant-ph/9706009.

[45] J.-Å. Larsson, A Kochen-Specker inequality, *EPL (Europhysics Letters).* **58**(6), 799 (June, 2002).

[46] J. A. de Barros and P. Suppes, Inequalities for Dealing with Detector Inefficiencies in Greenberger-Horne-Zeilinger Type Experiments, *Physical Review Letters.* **84**(5), 793–797 (2000).

[47] G. Bacciagaluppi, Leggett-Garg Inequalities, Pilot Waves and Contextu-

J. A. de Barros, G. Oas

ality, *International Journal of Quantum Foundations.* **1**(1), 1–17 (2014). arXiv: 1409.4104.

[48] Scarani V, The device-independent outlook on quantum physics, *Acta Phys. Slovaca Acta Physica Slovaca.* **62**(4), 1–60 (2013).

[49] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Reviews of Modern Physics.* **86**(2), 419–478 (Apr., 2014).

[50] R. Penrose, *Emperor's New Mind.* Oxford University Press, New York, NY (1989).

[51] S. Hameroff and R. Penrose, Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness, *Mathematics and Computers in Simulation.* **40**(3–4), 453–480 (Apr., 1996).

[52] S. Hameroff, Quantum computation in brain microtubules? The Penrose–Hameroff 'Orch OR' model of consciousness, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences.* **356**(1743), 1869–1896 (Aug., 1998).

[53] H. P. Stapp. Mind, Matter, and Quantum Mechanics. In *Mind, Matter and Quantum Mechanics*, The Frontiers Collection, pp. 81–118. Springer-Verlag, Berlin Heidelberg (Jan., 2004).

[54] H. P. Stapp. Mind, Matter, and Quantum Mechanics. In ed. H. P. Stapp, *Mind, Matter and Quantum Mechanics*, The Frontiers Collection, pp. 81–118. Springer Berlin Heidelberg (Jan., 2009).

[55] H. P. Stapp, Mind, brain, and neuroscience, *Cosmos and History.* **10**(1), 227–231 (2014).

[56] J. A. de Barros, On a Model of Quantum Mechanics and the Mind, *arXiv:1404.0714 [quant-ph, q-bio]* (Apr., 2014).

[57] A. Khrennikov, Quantum-like brain: "Interference of minds", *Biosystems.* **84**(3), 225–241 (June, 2006).

[58] A. Khrennikov, Quantum-like model of processing of information in the brain based on classical electromagnetic field, *Biosystems.* **105**(3), 250–262 (2011).

[59] J. A. de Barros, Quantum-like model of behavioral response computation using neural oscillators, *Biosystems.* **110**(3), 171–182 (Dec., 2012).

[60] D. Aerts, Quantum Interference and Superposition in Cognition: Development of a Theory for the Disjunction of Concepts, *arXiv:0705.0975 [physics, physics:quant-ph]* (May, 2007). arXiv: 0705.0975.

[61] E. M. Pothos and J. R. Busemeyer, A quantum probability explanation for violations of 'rational' decision theory, *Proceedings of the Royal Society B: Biological Sciences.* **276**(1665), 2171–2178 (2009).

[62] R. Franco, The conjunction fallacy and interference effects, *Journal of Mathematical Psychology.* **53**(5), 415–422 (2009).

[63] A. Khrennikov, Quantum-like model of cognitive decision making and information processing, *Biosystems.* **95**(3), 179–187 (Mar., 2009).

[64] D. Aerts, B. D'Hooghe, and S. Sozzo. A Quantum Cognition Analysis of the Ellsberg Paradox. In eds. D. Song, M. Melucci, I. Frommholz, P. Zhang, L. Wang, and S. Arafat, *Quantum Interaction*, number 7052 in Lecture Notes in Computer Science, pp. 95–104. Springer Berlin Heidelberg (2011).

[65] D. Aerts, S. Sozzo, and J. Tapia. A Quantum Model for the Ellsberg and Machina Paradoxes. In eds. J. R. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, and M. Melucci, *Quantum Interaction*, number 7620 in Lecture Notes in Computer Science, pp. 48–59. Springer Berlin Heidelberg (2012).

[66] J. S. Trueblood and J. R. Busemeyer, A Quantum Probability Account of Order Effects in Inference, *Cognitive Science*. **35**(8), 1518–1552 (Nov., 2011).

[67] H. Atmanspacher and H. Römer, Order effects in sequential measurements of non-commuting psychological observables, *Journal of Mathematical Psychology*. **56**(4), 274–280 (Aug., 2012).

[68] Z. Wang and J. R. Busemeyer, A Quantum Question Order Model Supported by Empirical Tests of an A Priori and Precise Prediction, *Topics in Cognitive Science*. **5**(4), 689–710 (Oct., 2013).

[69] J. M. Yearsley, E. M. Pothos, J. A. Hampton, and A. B. Duran. Towards a Quantum Probability Theory of Similarity Judgments. In eds. H. Atmanspacher, C. Bergomi, T. Filk, and K. Kitto, *Quantum Interaction*, number 8951 in Lecture Notes in Computer Science, pp. 132–145. Springer International Publishing (June, 2014).

[70] D. Aerts, Quantum structure in cognition, *Journal of Mathematical Psychology*. **53**(5), 314–348 (Oct., 2009).

[71] D. Aerts, J. Broekaert, L. Gabora, and T. Veloz. The Guppy Effect as Interference. In eds. J. R. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, and M. Melucci, *Quantum Interaction*, number 7620 in Lecture Notes in Computer Science, pp. 36–47. Springer Berlin Heidelberg (2012).

[72] M. Ashtiani and M. A. Azgomi, A survey of quantum-like approaches to decision making and cognition, *Mathematical Social Sciences* (2015).

[73] P. Bruza, K. Kitto, D. Nelson, and C. McEvoy, Is there something quantum-like about the human mental lexicon?, *Journal of Mathematical Psychology*. **53**(5), 362–377 (Oct., 2009).

[74] D. Aerts and S. Sozzo. Quantum Entanglement in Concept Combinations. arXiv e-print 1302.3831 (Feb., 2013).

[75] J. A. de Barros and P. Suppes, Quantum mechanics, interference, and the brain, *Journal of Mathematical Psychology*. **53**(5), 306–313 (Oct., 2009).

[76] G. Oas, J. A. de Barros, and C. Carvalhaes, Exploring non-signalling polytopes with negative probability, *Physica Scripta*. **T163**, 014034 (2014).

[77] E. Dzhafarov, R. Zhang, and J. Kujala, Is there contextuality in behavioral and social systems?, *arXiv:1504.07422 [quant-ph, q-bio]* (Apr., 2015). arXiv: 1504.07422.

[78] J.-Å. Larsson, Loopholes in Bell inequality tests of local realism, *Journal of Physics A: Mathematical and Theoretical*. **47**(42), 424003 (Oct., 2014).

[79] M. Stobińska, F. Töppel, P. Sekatski, and A. Buraczewski, Towards loophole-free Bell inequality test with preselected unsymmetrical singlet states of light, *Physical Review A*. **89**(2), 022119 (Feb., 2014).

[80] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. F. Roos, State-independent experimental test

of quantum contextuality, *Nature.* **460**(7254), 494–497 (July, 2009).

[81] D. Dieks, Communication by EPR devices, *Physics Letters A.* **92**(6), 271–272 (Nov., 1982).

[82] J. A. de Barros and G. Oas. Response Selection Using Neural Phase Oscillators. In eds. C. E. Crangle, A. Garcia de la Sienra, and H. E. Logino, *Foundations and Methods from Mathematics to Neuroscience: Essays Inspired by Patrick Suppes.* CSLI Publications, Stanford University, Stanford, California (2014).

[83] P. Suppes, J. A. de Barros, and G. Oas, Phase-oscillator computations as neural models of stimulus–response conditioning and response selection, *Journal of Mathematical Psychology.* **56**(2), 95–117 (Apr., 2012).

[84] J. A. de Barros and G. Oas. Quantum Cognition, Neural Oscillators, and Negative Probabilities. In eds. E. Haven and A. Khrennikov, *The Palgrave Handbook of quantum models in social science: applications and grand challenges.* Palgrave MacMillan (2015).

[85] J. A. de Barros and P. Suppes, Probabilistic Inequalities and Upper Probabilities in Quantum Mechanical Entanglement, *Manuscrito.* **33**(1), 55–71 (2010).

[86] S. Hartmann and P. Suppes. Entanglement, Upper Probabilities and Decoherence in Quantum Mechanics. In eds. M. Suárez, M. Dorato, and M. Rédei, *EPSA Philosophical Issues in the Sciences*, pp. 93–103. Springer Netherlands (Jan., 2010).

[87] F. Holik, M. Saenz, and A. Plastino, A discussion on the origin of quantum probabilities, *Annals of Physics.* **340**(1), 293–310 (Jan., 2014).

[88] P. Suppes and M. Zanotti, Existence of hidden variables having only upper probabilities, *Foundations of Physics.* **21**(12), 1479–1499 (1991).

[89] E. Wigner, On the Quantum Correction For Thermodynamic Equilibrium, *Physical Review.* **40**(5), 749–759 (June, 1932).

[90] P. Dirac, {B}akerian Lecture. {T}he Physical Interpretation of Quantum Mechanics, *Proceedings of the Royal Society of London B.* **A180**, 1–40 (1942).

[91] R. Feynman. Negative probability. In eds. B. Hiley and F. Peat, *Quantum implications: essays in honour of David Bohm*, pp. 235–248. Routledge, London and New York (1987).

[92] G. Mückenheim, A review of extended probabilities, *Physics Reports.* **133**(6), 337–401 (1986).

[93] P. Suppes, Probability concepts in quantum mechanics, *Philosophy of Science.* **28**(4), 378–389 (1961).

[94] M. O. Scully, H. Walther, and W. Schleich, Feynman's approach to negative probability in quantum mechanics, *Physical Review A.* **49**(3), 1562–1566 (Mar., 1994).

[95] R. W. Spekkens, Negativity and Contextuality are Equivalent Notions of Nonclassicality, *Physical Review Letters.* **101**(2), 020401 (July, 2008).

[96] J. B. Hartle, Quantum mechanics with extended probabilities, *Physical Review A.* **78**(1), 012108 (July, 2008).

[97] S. Abramsky and A. Brandenburger, The sheaf-theoretic structure of non-

locality and contextuality, *New Journal of Physics.* **13**(11), 113036 (Nov., 2011).

[98] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, Negative quasi-probability as a resource for quantum computation, *New Journal of Physics.* **14**(11), 113011 (Nov., 2012).

[99] S. Al-Safi and A. Short, Simulating all Nonsignaling Correlations via Classical or Quantum Theory with Negative Probabilities, *Physical Review Letters.* **111**(17), 170403 (Oct., 2013).

[100] J. J. Halliwell and J. M. Yearsley, Negative probabilities, Fine's theorem, and linear positivity, *Physical Review A.* **87**(2), 022114 (Feb., 2013).

[101] V. Veitch, Negative Quasi-Probability in the Context of Quantum Computation (Aug., 2013).

[102] X. Zhu, Q. Wei, Q. Liu, and S. Wu, Negative probabilities and information gain in weak measurements, *Physics Letters A.* **377**(38), 2505–2509 (Nov., 2013).

[103] S. Abramsky and A. Brandenburger. An Operational Interpretation of Negative Probabilities and No-Signalling Models. In eds. F. van Breugel, E. Kashefi, C. Palamidessi, and J. Rutten, *Horizons of the Mind. A Tribute to Prakash Panangaden*, number 8464 in Lecture Notes in Computer Science, pp. 59–75. Springer Int. Pub. (2014).

[104] A. J. Bracken and G. F. Melloy, Waiting for the quantum bus: The flow of negative probability, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.* **48, Part A**, 13–19 (Nov., 2014).

[105] E. R. Loubenets, Context-invariant quasi hidden variable (qHV) modelling of all joint von Neumann measurements for an arbitrary Hilbert space, *Journal of Mathematical Physics.* **56**(3), 032201 (Mar., 2015).

[106] J. A. de Barros, G. Oas, and P. Suppes. Negative probabilities and Counterfactual Reasoning on the double-slit Experiment. In eds. J.-Y. Beziau, D. Krause, and J. Arenhart, *Conceptual Clarification: Tributes to Patrick Suppes (1992-2014)*. College Publications, London (2015).

[107] A. Kolmogorov, *Foundations of the theory of probability*, 2nd edn. Chelsea Publishing Co., Oxford, England (1956).

[108] J. A. de Barros and G. Oas, Negative probabilities and counter-factual reasoning in quantum cognition, *Physica Scripta.* **T163**, 014008 (2014).

[109] A. Khrennikov, p-Adic probability theory and its applications. {T}he principle of statistical stabilization of frequencies, *Theoretical and Mathematical Physics.* **97**(3), 1340–1348 (Dec., 1993).

[110] A. Khrennikov, p-adic statistical models, *Doklady Akademii Nauk.* **330**(3), 300–304 (1993).

[111] A. Khrennikov, *Interpretations of probability*. Walter de Gruyter, Berlin (2009).

[112] G. J. Székely, Half of a coin: negative probabilities, *Wilmott Magazine.* **50**, 66–68 (2005).

[113] P. A. Morris, Decision Analysis Expert Use, *Management Science.* **20**(9), 1233–1241 (May, 1974).

32   *J. A. de Barros, G. Oas*

[114] P. A. Morris, Combining Expert Judgments: A Bayesian Approach, *Management Science.* **23**(7), 679–693 (Mar., 1977).

[115] M. L. G. Redhead, Relativity and Quantum Mechanics—Conflict or Peaceful Coexistence?, *Annals of the New York Academy of Sciences.* **480**(1), 14–20 (Dec., 1986).

[116] M. P. Seevinck, Monogamy of correlations versus monogamy of entanglement, *Quantum Information Processing.* **9**(2), 273–294 (Jan., 2010).

34   *J. A. de Barros, G. Oas*

Index