Physical requirements for models of consciousness

Rodolfo Gambini$^1$ and Jorge Pullin$^2$

1. Instituto de Física, Facultad de Ciencias, Iguá 4225, esq. Mataojo, 11400 Montevideo, Uruguay.
2. Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803-4001.

Consciousness presents a series of characteristics that have been observed throughout the years: unity, continuity, richness and robustness are some of them. It manifests itself in regions of the brain capable of processing a huge quantity of integrated information with a level of neural activity close to criticality. We argue that the physics of consciousness cannot be exclusively based on classical physics. Consciousness unity cannot be explained classically as the classical properties are always Humean like a mosaic. One needs an entangled quantum system that can at least satisfy part of the functions of a quantum computer to allow to generate an inner aspect with the unity of consciousness and to couple with a classical system that gives it simultaneous access to preprocessed information at the neural level and to produce events that generate neural firings.

I. INTRODUCTION

Among the many questions that await answers in order to understand the relationship between mind and brain, the most important one is what Chalmers [1] calls the hard problem of consciousness. That is, how does our subjective experience—which Chalmers also calls phenomenal—emerge from the matter that composes our brain? The question has been posed for centuries without us having been able to exorcise the subjective space, internal to our conscious experiences. Nagel [2] summarizes the problem in the following way: “the fact that an organism has conscious experience at all means, basically, that there is something it is like to be that organism.” One of the approaches to the perplexities that the problem produces that appears as the most promising is the one that recognizes in quantum mechanics a model of a physical theory that can make important contributions to the solution of this problem. Briefly, there exist three reasons to consider the quantum alternative to explain conscious phenomena: a) The random (probabilistic) character of the laws of quantum mechanics [3]; b) The nature of the fundamental objects of the quantum theory, its ontology, that as we shall see, and as others have noted too, is more neutral towards the dichotomy between the material and the phenomenal; c) The nonlocal properties and entanglement characteristic of the quantum states that could contribute to solve one of the central aspects of the “hard problem of consciousness” that Chalmers calls “combination problem”. Basically, it consists in understanding how do the experiences of microscopic fundamental physical entities “combine to yield the familiar sort of human conscious experience that we know and love.” The unified and component-less way in which conscious phenomena appear to us, which has been in the center of philosophical reflection since Descartes and inspired in Leibniz his monad model. Seager [4] considers that the hard problem has two aspects and may be described as follows: “Now, we’ve seen that there are at least two problems that enter into the hard problem: what makes some state an ‘intrinsic representation’ as well as
the more traditional generation problem of what makes one of these representational states conscious. Perhaps these two questions are not as distinct as they appear.” We will address the two aspects mentioned by Seager in what follows.

The main obstacle to quantum treatments of the problem of consciousness arises from the difficulties that, at first sight, biological systems like the brain encounter to undergo quantum behavior. Indeed, it is not easy in a hot and highly interactive physical system to find relevant degrees of freedom in the brain that can be sufficiently isolated to retain their quantum coherence. Some years ago Tegmark [5] presented a quantitative analysis of the effects of environmental decoherence in the brain that suggests that such processes occur very rapidly and the decoherence timescales are typically much shorter than the relevant timescales for regular neuron firing. In the second part of this paper we will discuss some ideas to deal with that objection.

Let us first analyze each of the three reasons that lead to consider that quantum mechanics can have a significant role in conscious phenomena.

II. QUANTUM MECHANICS IN CONSCIOUS PHENOMENA

A. The statistical meaning of quantum laws: regularism vs physicalism

Classical physics is deterministic, which implies that a precise knowledge of the present allows to determine any future evolution without ambiguities using the evolution equations. Mechanist determinism is necessarily physicalist. Indeed, in a classical world if anything existed that transcended what the physical laws establish, it could only have an observable effect by contradicting such laws. In other words, if the classical laws are universally valid: “No physical event has a cause outside the physical domain” [6]. Butterfield [7] lays out physicalism in these terms “all empirical subject-matters, such as the biological, the mental, the social, are literally a part of the subject-matter of physics.”

Quantum theory introduced an element of randomness that was absent in classical physics. In deterministic theories randomness is a product of our ignorance of part of the required information for a complete determination of the behaviors of the system. Randomness has an epistemic character in classical physics. In the latter, given initial conditions one can determine with arbitrary precision the future evolution. As the initial data cannot be given with arbitrary precision due to quantum fluctuations, the evolution of a chaotic system becomes truly unpredictable. In classical physics any apparent novelty is a consequence of our ignorance about the initial conditions. See for instance [8].

In quantum mechanics the probabilistic character of the decay of an atom is fundamental and cannot be eliminated with any further knowledge of the system. It is a fundamental feature of nature that manifests itself each time an individual quantum event occurs. What is statistically determined is the behavior of ensembles of such events. Quantum mechanics not only allows effective forms of top down causation where there exist behaviors of the whole that are not dictated by the quantum states of the parts. It also allows in principle the emergence of new behaviors compatible with the laws of physics, but not predictable by them. The universal validity of the quantum physical laws allows in principle the existence of physical events without causes in the physical domain. There exist two positions with respect to the meaning of physical laws. For the regularist the laws describe the regularities that, without exception, hold in a world that transcends them. For the necessitarian the laws are the ultimate description of the physical world. Our view is that quantum theory
favors regularism as it establishes only probabilistic laws. As Swartz [9] observes, “At some point explanations must come to an end. Regularists place that stopping point at the way-the-world-is. Necessitarians place it one, inaccessible, step beyond, at the way-the-world-must-be.” Wittgenstein [10] considers the laws as descriptions that allow to deduce from the knowledge of the behavior of certain fundamental objects, laws and empirical properties that describe a wide range of phenomena. Einstein considers them a product of our free creations that make intelligible our empirical observations. Following Wittgenstein and Einstein we will adopt a regularist position with respect to the laws of nature.

B. Quantum ontology

Quantum mechanics, like all modern physical theories, is composed of three fundamental elements: a mathematical formalism, correspondence rules and an interpretation. It provides mathematical descriptions of physical systems that initially are in certain states and evolve giving rise to events to which we can associate certain physical magnitudes. The notions of event, state and system are basic concepts of the theory associated to elements of its formalism and through the correspondence rules to experimental operations and observations. The states define the possible behaviors of a quantum object and are described by vectors. The systems are associated to the set of possible behaviors, in mathematical terms to the vector space. A physical object is no more than a system in a certain state, for instance a given hydrogen atom. The interaction of a quantum system with others, for instance its measurement, ends up producing events in the measuring device to which certain physical magnitudes can be associated. Whereas in classical physics the knowledge of the state determines completely all the physical magnitudes, in quantum mechanics the knowledge of the state vector only allows to make probabilistic predictions about the events that may be produced and of the values that the associated physical magnitudes may take. In general, if one adopts most of the realistic interpretations of quantum mechanics available like the Modal Interpretation [11], Many Worlds Interpretation [12], the Ghirardi–Rimini–Weber Interpretation, or the Montevideo Interpretation [14], the quantum states describe the disposition of a certain system to produce events in its interaction with other systems, for instance, when it is measured. In other interpretations, like the Copenhagen one, the states only characterize the dispositions to produce events in a measuring device. Some realistic interpretations as Bohmian quantum mechanics do not have events as a primitive concept and are formulated in terms of particle-like fundamental objects. We will adopt a realistic point of view, taking as basic elements of the quantum ontology the ones we already mentioned; systems, states and events.

A first, easily identifiable, element of physical theories that can be considered as an element of reality is that of event (some of these ideas have been discussed in our book [15]). Attempts to base the ontology on events have a long history that go back to Hume and that was reinforced by relativity and the quantum theory. In relativistic physics events are considered points in space-time. We agree with Dorato [16] that states that “If spacetime physics presupposes some concepts coming from ordinary language, then the meaning of such concept must be taken into account. Now, “events” occur by definition, their “being” is their happening, and they are an essential ontological component of any spacetime theory.”

In quantum mechanics, the measurement process does not end until an event takes place. An event can be the formation of a dot produced by the impact of a particle in a photographic plate. In that moment one can say that the measurement has occurred. Without events there
are no measurements. The events involved in the detection are always the product of an interaction with measuring systems that are macroscopic. In the case of the bubble chamber used for the detection of particles emerging from a nuclear reaction, the drops of liquid are vastly larger than the particles that leave the trace and behave macroscopically. Any event is associated with a possible measurement. The measurement is nothing but the assignment of quantitative properties to the event. In the example: the position of the dot. The dot and its properties have, a mathematical counterpart in the formalism of quantum mechanics corresponding to projectors in the Hilbert space of the detecting plate. We are thinking in this kind of events as the building blocks of the apparent reality. Projectors characterize both the event—the appearance of a dot—and its properties. Thus quantum mechanics provides an exhaustive and mathematically precise description of events and their properties. Notice that no reference to a particular interpretation of quantum mechanics is required for speaking about events. The program of accounting for physical reality in terms of events has a long tradition that goes back to Russell, and in different terms to Hume. According to Russell [17] “the enduring thing or object of common sense and the old physics must be interpreted as a world-line, a causally related sequence of events, and ... it is events and not substances that we perceive.” To put it differently: for Russell an object is nothing more than a set of events that are causally connected. Although we consider this point of view a step in the right direction, we think it is incomplete for a foundation of physical reality based on events, particularly in the light of quantum mechanics. Note that for Russell an atom cannot be considered an object as long as it does not interact yielding events, whereas our definition naturally includes any concrete microscopic system as an object, given that its disposition to produce events is always defined by its state (some of these statements are present in our book [15]).

But what is the ontological nature of the state? People move between two attitudes: a) The states are real and exist in the world independently of any observer and therefore must be included in quantum mechanics’ ontology or b) they only describe our knowledge about the world and therefore have a mere epistemological character. We argue that the states should have the same ontological status as events. Indeed, quantum mechanics, or for that matter any physical theory, must have predictive power be it either deterministic or probabilistic. The state of the universe in an instant—or in the case of general relativity in a Cauchy surface—determines via the equations of motion the future behavior. In particular, all the events that can occur and their probability of effective occurrence in the future.

Let us suppose that the state in a certain instant only has an epistemological character, that is, the initial state codifies the information needed to determine the possible future behaviors. In that case, the only elements of reality would be the events. For this to be possible, the initial state should be totally determined by the events that occurred in the past of the initial Cauchy surface or otherwise the future behavior would be unpredictable starting from past events and would not admit a predictive description even at a probabilistic level. But it is generally admitted that the universe started with a Big Bang, so we know that it must have passed in the first instants through a purely quantum regime described in terms of states where the production of events was not possible. We also know from quantum cosmology models that the behaviors of the universe in later times depend on the initial state. What kind of ontological nature we would assign to the initial universe if we do not accept that the states are real?

Given that more or less developed forms of sensitiveness and goal oriented behavior are found in all living creatures, it is natural to explore the possibility that certain forms of
interiority are possible in very primitive stages of development. With philosophers like Spinoza, and more recently with Russell, Whitehead, Feigl and Smart and with physicists like Bohr, Pauli, Bohm, d’Espagnat and the current defenders of the double aspect theories we believe that a suitable ontology for the physical world should include in addition to the acknowledged third person perspective some form of direct inner acquaintance.

Since there is a growing consensus that we live in a purely quantum universe, the adoption of a quantum ontology appears rather inevitable. In a quantum universe the phenomena described by classical physics are particular cases of quantum processes where decoherence is particularly relevant and they should have the same ontological character.

It has been noted some time ago by Mach [18] that an ontology of events allows a promising approach to the problem of consciousness. The position known as neutral monism holds basically that events can be associated both to physical and phenomenal aspects. It is a particular case of the double aspect theories in which one class of entities has both physical or third person manifestations and phenomenal or first person manifestations. Although our position differs significantly from neutral monism, it is important to recognize that it stems from an elaboration on its ideas more adjusted to the quantum ontology that we have previously discussed.

Firstly and most important not only events are required for a complete version of an identity theory. A double aspect conception of the fundamental entities of the quantum ontology allows us attributing to events a phenomenal or internal aspect. In evolved biological systems like human beings certain events in our brain to which we have a first-person access, would be associated to sensations. More explicitly, we only have an indirect access to the events of the systems we study physically, that is we know of them ultimately by their effects on other objects and eventually on us. However, we have direct access to the events in our brain as long as we are aware of them. In that sense we can say that we only have idea of how things are in themselves through our consciousness. If the events have a phenomenal aspect that manifests itself as a sensation when we have a first person access, does something similar occur with the states? While the inner aspect of events has been considered by many, the inner aspect of systems in given states has not been discussed in detail.

As we have emphasized, states characterize the disposition of the system to act producing certain effects on other systems. To characterize the disposition of a state to act on any other system is to give its most complete description. States are private in the sense that one cannot determine in which state is a system by measuring it. One needs an ensemble of identical states and measurements on each member of the ensemble in order to have a complete determination of the states. As it does not make sense to have an ensemble of identical mental states, they are inaccessible to external observers. While sensations, concepts and beliefs could be related with events, goals, desires or intentions would be possibly related with states given their dispositional character. It is this assignment of an inner aspect which is in the basis of the solution of the second question posed by Seager about the generation problem. The first issue raised by the generation problem is closely related with what Seager defines as the combination problem: “Even if we grant that all elements of reality have some kind of mental, conscious aspect to them, how is it that some groups of such elements form higher level and unified states of consciousness? Isn’t this just the generation problem all over again?” The question may be translated here in the following terms: even if we grant that entangled states have an unified inner aspect, does the high level and unified states of consciousness only occur in some of these states? Our answer here
will be, as we shall see, that the high levels of consciousness only occur in entangled systems coupled to neural systems able to produce representations and working close to criticality.

Let us consider an electron that has been diffracted by a small hole and arrives to a photographic plate. The state defines the disposition to produce a dot in different places of the plate, at the arrival to the plate the interaction of the photon with the plate forces the state to choose among the different possible positions of the dot producing the dissociation of silver bromide in a photographic plate leading to a cascade effect that produces the accumulation of millions of atoms of silver in a dot.

For an electron these “choices” cannot be anything that random, following the rules of quantum mechanics. At the level of isolated physical systems like particles, atoms or molecules, the inner aspect of events or states cannot be more than strictly epiphenomenic.

If one sends many electrons through the hole they end up reproducing a diffraction pattern. Different interpretations of the quantum formalism assign different meanings to these “choices” of the state that end up producing the events. For instance in the Many Worlds interpretation all possible histories and futures corresponding to different choices of the state—different measurement results—are real, each representing an actual “world”. The modal interpretation relies on the distinction between “dynamical states” and “value states” or “actual definite-valued” observables [11]. While the dynamical state is the usual state of quantum mechanics, the “value state” represents the physical properties that the system actually has at a given instant. Instead of a choice one has in this case the value of the certain hidden properties that are instantiated at certain instant. Systems at all times possess within this interpretation a number of well-defined physical properties, these properties can be represented by the system’s value state and become manifest during measurements. In the Montevideo Interpretation the use of physical quantum mechanical clocks leads to exact statistical mixtures during measurements. The resulting event and measurement outcome results from a choice of the state in agreement with the statistical prediction of quantum mechanics. A deeper discussion of this point in the various interpretations would require an entire paper devoted to it. An interesting analysis of how some of the existing interpretations can change the form of treating the mind body problem can be found in a paper by Butterfield [7].

C. The unity of consciousness and the non separability of quantum states

Our conscious experiences are extraordinarily rich and complex. Right now I can see the text that I am writing, I listen to an unknown piece of music at a distance, feel a slight itch in my left hand and feel the moist heat of a summer afternoon near the sea. These experiences do not present themselves isolated from each other but as part of a single perspective. The conscious perspective in each case not only has the unity of the various perceptions that coexist at each instant, it also has continuity in time. [19].

Hume seems to have been the first to observe that an ontology of events does not appear enough to explain the unity of conscious phenomena. Indeed, his atomist vision of the world like a mosaic of events must have led him to claim that there does not exist a unity of consciousness but “just “a bundle of different perceptions””, [20]. In the concept of composite system in classical physics, where every observable of the system is a function of the state of motion of the particles that compose it, the notion of supervenience can be stated as follows: “A set of properties A supervenes upon another set B just in case no two things can differ with respect to A-properties without also differing with respect to
their B-properties. In slogan form, “there cannot be an A-difference without a B-difference” [21]. If one adopts a physicalist standpoint based on classical physics, as it is frequently done, the whole is the sum of the parts. The lower level contains all the information. In classical physics this point of view is inescapable. Any property of the whole is a function of the properties of the parts, for instance positions and velocities of particles. The causal properties of the whole result from the properties of the parts, including their relations and interactions. These properties may be extended to include classical fields. The higher level, if it is going to be understood as a physical system, is nothing but a system of particles and eventually of classical fields and its properties should supervene to them. They are examples of what is usually called mereological supervenience [21]. In classical physics there is no room at the bottom. Given a complete specification of the state of each of the parts of the complete system at instant $t$ the state of the whole at $t$ is completely specified, in other words we can predict with certainty any behavior of the whole in terms of the behaviors of their parts. The classical realm is causally closed. Hume himself seems to have had some reservations with respect to this conception of consciousness as a mosaic of perceptions. As Brook and Raymond put it: “Yet, in a famous appendix [of his Treatise] he says that there is something he cannot render consistent with his atomism. He never tells us what it is but it may have been that consciousness strongly appears to be more than a bundle of independent ‘perceptions’” [19]. In spite of the reservations that Hume presents more or less explicitly and of the manifest unity of conscious phenomena, the most extended way of understanding the world from a third person perspective inspired in the British empiricism—and classical physics—conceives it as composed of events. In his article “Why be Humean?” [22], Maudlin carried out a lucid analysis of how quantum mechanics altered the beliefs, both philosophical and physical, with respect to the local nature of reality. We will follow it here partially as we do not share all of his premises. Lewis [23] has described the notion of Humean supervenience in the following terms: “Humean supervenience is named in honor of the great denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another... We have geometry: a system of external relations of spatio-temporal distances between points... And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short, we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that.” This vision was completely superseded by quantum mechanics. The importance of the non-separability or entanglement of the quantum states was not immediately recognized. Two papers appeared in 1935 dealing with the phenomenon. In the first one, Einstein, Podolsky and Rosen [24] argued using an entangled system that quantum mechanics was apparently incomplete. Shortly thereafter Schrödinger [25] published a seminal paper discussing the notion of entanglement. He realized the crucial role of the concept: “[it] is not one, but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.”

The interactions between particles that form an arbitrary composite system in general yield entangled systems. Let us take the typical case of an entangled system composed of two particles with spin, for instance, two electrons. Whereas for the spin of an isolated system there always exists a direction for which it is always up, and one can always assign a definite component of the spin to a particle, in the case of an entangled system, the component particles do not have defined spins. In fact the particles that compose it do not
have any definite property. As a consequence, no state of the spin of particle 1 will yield the answers that the particle 1 yields when it is entangled, and the same occurs with particle 2. Only the composite system has well defined properties. For instance, in the example, the total angular momentum of the composite particle has a vanishing component along the $z$ direction. Entangled states cannot be analyzed as if there were composed by two independent particles.

In fact, there exist two entangled states with vanishing $z$ component of the angular momentum. They are the singlet state and the vanishing component of the triplet state. They are called $|\psi_S\rangle$ and $|\psi_T\rangle$. Whereas the magnitude of the total spin of $|\psi_S\rangle$ is zero, that of $|\psi_T\rangle$ is $\sqrt{2}\hbar$. The behavior of each of the spins that compose the pair is identical for both, whatever component is measurement for one of the particles one half of the time the answer will be up and one half down. But although no local measurement can distinguish them, there exist global measurements that do. The differences can be discovered studying how simultaneous measurements on both particles are correlated.

No theory that considers the states as part of reality can be separable. The quantum description of the world tell us that the latter is not separable. This “holistic” behavior of many quantum systems that are non-separable is not something exceptional but actually it is the most common occurrence in composite systems whose components interact or have interacted. For instance, the electrons in a multi-electron atom are entangled. Or two particles that collide in a particle accelerator may become entangled after the collision. It is a generic property: interactions typically yield entangled states for multi component systems. As we will see, the emergent properties of systems of this kind are key to open the possibility of explaining chemical or biological properties in physical terms.

To summarize the properties of entangled states, Healey [26] introduced the notion of “physical property holism”. It establishes that there are physical objects “not all of whose qualitative intrinsic physical properties and relations supervene on qualitative intrinsic physical properties and relations in the supervenience basis of their basic physical parts.” The emergence of new properties of the whole in a quantum world where states, events and properties play a fundamental role, is a crucial manifestation of the ontological novelty of the quantum mechanical objects like composite systems in entangled states.

As Healey [26], Teller [27], and other philosophers have noticed and we have emphasized elsewhere, entangled systems show strong emergence of properties and top-down causation. Seager rejects emergence. He says: “Radical emergence goes further, asserting that the emergent properties make a real difference to the workings of the world. Radical emergentism claims that the simulation based only on fundamental physical law would simply fail to simulate the world accurately (if thunderstorms were radically emergent, the atmospheric simulation, no matter how perfect, would go on and on, but never generate anything like a thunderstorm).” This view results from a confusion between epistemological and ontological emergence. What is claimed by the previously mentioned authors is that the laws of quantum theory predict that there are properties of composite systems that do not supervene from their parts and that the state of wholes and therefore some of their behaviors are not given by the total knowledge of the states of its parts.

Not only one observes new properties of the complete entangled system, there also exist behaviors of the parts that are not ruled by their states but by the state of the whole. In this case one says that the system has top-down causation [28]. We have seen that in the previously discussed example of entangled spinning particles the state of the whole has more information that the state of the parts, in particular information about correlation
between the events observed in each component. For instance the spin measurements of both particles will be correlated. Whenever both observers measure the spin in the same direction $z$ their results would be opposite, —up for Alice and down for Bob or vice versa—If they measured the correlation in a different direction $x$ then the correlations for $|\psi_S\rangle$ and $|\psi_T\rangle$ are different, for the second state both measurements may lead to up results while for the first the results are always opposite. The local observers of each particle could never have figured out these correlations by looking at the individual systems in isolation without comparing their measurements.

In $n$ particle systems the number of entangled states of the system grows exponentially with $n$. For example, for $n$ spins the number of entangled states goes as $2^n$. Whereas the state of a classical system of $n$ particles grows linearly with $n$ the state of the quantum system grows exponentially. The non-separability, non supervenience and top-down causation types of behaviors become more and more important as the systems involve more particles. At the basis of the appearance of novel properties and top-down behaviors of emergent systems it is this exponential growth in the possible behaviors of the quantum systems whose philosophical implications can be recognized when the appropriate ontology is put into action.

Since a large fraction of quantum systems in interaction are composed by entangled sub-systems, the emergence of new properties of the complete system with top-down causation is part of the typical behaviors observed. Examples of downward causation may be found in the molecular behavior. For instance, it is downward causation of the whole molecular system state including nuclei and electrons that determines the precise vibrational behavior of the nuclei. The individuality of the atoms that contribute to the formation of a given molecule is partially dissolved and the emergent electronic behavior is new and crucial for the understanding of the behavior of the molecular system.

### III. REQUIREMENTS FOR QUANTUM MODELS OF CONSCIOUSNESS

When we analyze conscious processes the first thing that strikes us is the unity and degree of organization of them. Each conscious experience allows to access a cloud of information organized in perceptions. We know when we watch television that it is enough to have a relatively small number of pixels to recover and interpret very complex images. The current mechanisms of transmission of information exploit these capabilities to minimize the amount of information to transmit. On the other hand, conscious experience extends enough to understand the meaning of a word or recognize a melody although the phoneme or notes that compose it are presented sequentially.

The formation of a conscious experience takes time. We cannot determine with precision how long but we can set maxima and minima. A sensorial experience, for instance viewing an image, requires of the order of a tenth of a second for us to start to recognize some of its properties and about a second to have a more complete vision of shapes and colors. Consciousness, beyond how information manifests itself in sensations of color, shape, etc., is clear that it requires the ability to integrate a great quantity of information in a certain period of time. We could say that conscious processes require: a) the integration of information and b) its translation in a phenomenic experience where shapes, colors and other sensory data, memories or emotions manifest themselves with their peculiar qualities. Simplifying, there would be a “computing” processing at neural level that prepares and integrates the information. Then this information would be accessed by a quantum physical system whose
states would have a double quality: a “sensitive” one in which the state expresses its internal content as individual instances of our experience, and a second one analyzable in terms of psycho-physical processes accessible in third person perspective whose character, localization and meaning have to be identified. Both stages form part of the more developed types of consciousness. Without the integration of the information carried out in the first stage, the second would only have an epiphenomenic character. With the first, the second appears to acquire the possibility to orient the process in ways that we do not understand completely yet, but where forms of top-down causation must play a relevant role.

If there is no quantum second stage, then what we call consciousness should be a mere byproduct of classical neural systems that exchange classical information. Even if a classical ontology of particles and electromagnetic fields were able to support consciousness, this kind of consciousness, as we have already discussed, would not have causal powers and would be merely epiphenomenic. This is due to the causal closure of classical physics. Only if the second stage takes place in a quantum system that is in an entangled state with the ability to interact with the neurons involved in the preparatory process there would be some possibility of top-down causation and ontological novelty. In other words, ontological novelty is only possible because the fundamental nature of the universe is quantum mechanical and classical physics is only an approximate behavior in certain regimes.

To understand each of these stages, let us summarize some recognizable characteristics of consciousness in addition to the already mentioned fundamental one of providing an integrated, unitary and continuous access to information.

At a classical level, that is, based on the coordinated functioning of neurons, the unity of consciousness manifests itself in information integrated, at a quantum level, in the holistic behavior of the entangled state. We will give in what follows ways of measuring both effects. But in addition to the unity we have to account for a set of properties of consciousness:

a) It is a robust phenomenon; b) It is a phenomenon associated with the thalamocortical functions and do not appear in other strongly interacting neural systems like the cerebellum; c) In spite of the unity there exist brain damages that lead to kinds of divided consciousness; d) The physical substrate of consciousness is capable to provide unitary experiences, very rich and varied, and therefore the corresponding quantum system should be able to access a strongly integrated neural network and occupy a very large number of states [29].

Let us elaborate these points in more detail:

a) Consciousness is a robust phenomenon, a definitive loss of consciousness only occurs if a severe brain damage occurs, either physical or chemical. For instance, the axonal diffuse damage due to trauma that produce the disruption of the axons, which permit the intercommunication among neurons, is one of the most common encephalic lesions, and among the most devastating ones, since the damage occurs over a wide area rather than in a focal point in the brain —local damage does not seem to affect consciousness—. It is one of the main causes of loss of consciousness and the development of a vegetative state after brain trauma. The result of the damage is frequently a coma out of which about 90% of patients will never regain consciousness. Two anatomical constituents of the brain are needed to underlie consciousness. The first is the cerebral cortex, the gray matter that forms the outer layers of the brain. The other is a structure located in the brain stem, called the reticular activating system. The existence of a system capable of maintaining quantum states for sufficiently long times without decoherence must be based in sub-cellular mechanisms that are robust enough to work when the neural substrate does. One should also understand better whether decoherence is enough to make quantum effects irrelevant. We shall discuss
this point in some detail towards the end of this work.

b) It is not enough to have a strongly interacting set of neurons for consciousness to appear. The cerebellum has more than 80% of the brain’s neurons forming a strongly integrated network, interconnected and very compact, that occupies the rear part of the brain. It receives signals from our senses and carries out a fine control over our movements. When we walk, play tennis or play the piano, we make use of our cerebellum. However, it does not seem to have much to do with consciousness. People who had their cerebellum removed exhibit important limitations, in particular concerning motion, but do not show signs of consciousness impairment. Why does 80% of the brain not involve conscious activity is a mystery that might be related, although we are not certain, to its lack of capacity of sustaining subsystems in entangled quantum states like those needed for the second stage of conscious processes. Another possibility is that the cerebellum is composed by many specialized sub-circuits, as suggested by the types of functions that it carries out. That is, the lack of consciousness would be associated with the lack of integration.

c) There exist brain damages that lead to divided consciousness states: consciousness seems to split within one brain and body. When, for therapeutic reasons, the main communication channel between both cerebral hemispheres—the corpus callosum—must be cut, interesting changes happen in consciousness. Patients that have undergone that procedure do not appreciate major changes in their conscious state nor do they exhibit significant behavioral changes. However, in laboratory conditions, those patients behave as if two conscious centers associated with both hemispheres had been created. Whereas the verbal expression of the patient is controlled by the left hemisphere, the written communication—using the left hand—is controlled by the right one. If they are queried about observations that only one hemisphere has access to, such hemisphere will manifest using the means at its disposal, verbal or written, of the observed object. One can also ask using only information that can be accessed by the right hemisphere to carry out a computation. Whereas the left would insist that it is not doing anything the right one will provide a result [19]. There is no agreement on how should these studies be interpreted. Some insist in the existence of two centers of consciousness. Bayne [30] and others suggest that there exists a single center that alternatively switches material from one hemisphere to the other. This possibility cannot be easily neglected given the fact that subjects do not present major behavioral changes. Let us assume that certain situation of life pose a problem that requires of certain computation or complex reasoning for reaching to the correct answer. While the left hemisphere would know the correct answer the right would probably guess a different one. How would the behavior of the subject respond to these contradictory conclusions?

d) The physical substrate of consciousness is capable of supplying an extremely rich set of experiences and therefore must be able to occupy a large number of states. Tononi has defined a measure of the integrated information handled by neural systems, based on the effective information that bipartitions A and B of a system S may interchange. The effective information is high if “the connections between A and B are strong and specialized, such that different outputs from A will induce different firing patterns in B” [31]. Quantitatively, the effective information EI(A → B) is defined in terms of the mutual information or entropy shared between a source A and a target B. The effective bidirectional information EI(A ↔ B) = EI(A → B) + EI(A ← B) measures the repertoire of possible causal effects of one part on the other.

In order to measure the information integration capacity of a neural system S, we should search for the bipartition(s) of S for which EI(A ↔ B) reaches a minimum (the informational
Indeed, to compare the effective informations of different partitions, they need to be normalized by the maximum information capacity for each bipartition. The minimum information bipartition of a certain system of neurons MIB(S) is the bipartition for which the normalized effective information reaches a minimum. The fundamental quantity introduced by Tononi is the information integration for the subset S, or Φ(S). It is the (non-normalized) value of EI(A ↔ B) for the minimum information bipartition: Φ(S) = EI(MIB(S)). Tononi’s proposal of choosing Φ as a necessary and sufficient measure of the conscious capability of the system has implications that appear highly contentious, as Tononi himself admits. Basically, it is equivalent to assuming that each set of the integrated network with some lighted up connections and others dark would correspond to a conscious state. The only condition is that it forms a sufficiently integrated network. According to Tononi [32], “Information integration theory makes the highly counterintuitive neuro-physiological prediction that consciousness should be present even when most neurons comprising the substrate of consciousness are inactive. What would it feel like? Perhaps, like the pure ‘consciousness state’ or ‘core experience’ documented over many centuries, and within many different cultural contexts and spiritual traditions.” An integrated network turned off would have the same Φ and therefore should give rise to a conscious state. If consciousness were a purely classical phenomenon, the same would happen with a sufficiently complex electrical circuit without any biological component. With our description with two well differentiated stages, Φ would characterize the information that can be integrated and prepared to produce a conscious perspective through the coupling of S with an entangled quantum system Q. It should be noticed that we consider that having an integrated neural network with large Phi could not be a sufficient condition [33] to ensure good “representational system standing in certain relations to things entirely external to the system itself”, as put by Seager. The events, neuron firings, of S would prepare the quantum system in a state ψ_Q whose internal aspect would give rise to the conscious perspective of the case and to possible processes of decision or quantum choices. The second stage of quantum character endows physical objects with the appropriate ontology to support consciousness. The entangled quantum system presents properties that are not supervenient from its parts and top-down causation allowing for efficient actions of the mind compatible with the natural laws that could help resolve the mentioned problem. The mechanisms of formation of entangled systems coupled to the integrated neural structure account for the uniqueness of consciousness in normal conditions and of the other properties here mentioned. On the other hand, as we have indicated, and we will analyze quantitatively later on, quantum information that may be present in an entangled system can widely exceed the classical one for a system with the same number of elements.

Small quantum structures, for instance a set of entangled spinning particles, can encode a huge quantity of classical information, each isolated spinning particle encodes a unit of quantum information: a qubit. Qubits are contained by any quantum system that has two states. One can consider that a single qubit contains an infinite amount of information because a pure state of a qubit is specified by two continuous complex parameters defined up to a common real factor. A string of zeros and ones of any length, capable of codifying an arbitrary amount of classical information, can be codified in the state of a quantum spin, for example using the decimal expansion of the amplitude of the state. However, since the quantum states are not accessible with individual measurements, as the latter destroys the state, it is not possible to recover all the information codified in the state. The Holevo bound establishes that only a single bit of classical information can be reliably retrieved from a
Leifer [34] has emphasized this remarkable property of quantum states: “If the quantum state truly exists in reality, it is puzzling that we cannot detect all of this extra information. Hardy has coined the term ‘ontological excess baggage’ to refer to this phenomenon... The excess baggage problem is exacerbated by considering how the state space scales with the number of qubits. A pure state of \( n \) qubits is specified by \( 2^n \) complex parameters, but only \( n \) bits can be reliably encoded according to the Holevo bound. However, the number of parameters required to specify a probability distribution over \( n \) bits also scales exponentially...” It should be emphasized that in the last few years different protocols have been developed in order to transmit an amount of classical information larger than the Holevo bound using suitable prepared entangled states, a procedure known as superdense coding [35]. If conscious experiences arise by direct inner access to the quantum states their excess baggage is reminiscent of the kind of ineffability that one encounter when one attempts to describe exhaustively a conscious experience using ordinary language, that is, using a classical information channel. It would seem that an entangled system with a few atoms or molecules per neuron of the talamocortical region could be enough to execute that function. It should be pointed out that in the same way that there exist measures of the classical information integration there exist measurements of entanglement of a system of \( N \) quantum particles [36] using the von Neumann entropy. In a system of \( N \) particles the state of the particles may either not be entangled, be partially entangled or totally entangled. A system that is not totally entangled is partially separable and can be factorized. The essential condition to have holistic behaviors of the quantum system is that it is in a totally entangled state, it does not appear as important that it be maximally entangled. That is, entangled systems where the entanglement measure takes its maximum value. An important property of entanglement is that it cannot be incremented by local operations and classical communications (LOCC). This implies that although the entangled state will be modified by LOCCs like the neural firings during the information integration process the measurement of the entanglement of the final state cannot be larger than the initial one. The initial entanglement is the product of quantum interactions produced previously to the diffusion of the particles that constitute the entangled system. Advances in quantum information are allowing now to understand how to encode classical bits of information into qubits, which is a primary stage for accessing the information processed by the brain’s neurons. The formation of multiparticle entanglement is a complex process under active study. Up to now it has been possible to entangle up to fifty atoms using a laser light for their excitation [36].

IV. SCENARIOS FOR QUANTUM MODELS OF CONSCIOUSNESS

A. Are quantum effects compatible with decoherence?

There have been attempts to explain consciousness arguing that in the end, since the world described by classical physics is given by approximate behaviors when there exists decoherence, the world is fundamentally quantum and therefore the quantum ontology is always applicable. In other words, it is argued that it does not matter that the brain is a warm strongly interacting environment where quantum decoherence is lost in microseconds, the brain is in all instances a quantum object. Schlosshauer [37] summarizes the meaning of decoherence as follows “decoherence provides us with an explanation of the appearance of the classical world around us that is ‘as good as it gets,’ if quantum theory is assumed to be
universally valid — and there exists no compelling experimental evidence to the contrary — and classicality is to be explained from within this theory (which constitutes a highly desirable goal). There certainly exist some open conceptual questions related to the ‘ultimate physical reality’ underlying the effective classicality induced by decoherence. The particular form of these questions tends to vary according to one’s interpretive stance toward quantum mechanics. But such questions are unavoidable if the emergence of classicality is to be completely derived from quantum mechanics. After all, at some global level all the ‘strange’ non-classical phenomena of quantum theory, such as superposition states, nonlocal quantum correlations, etc., must still persist, since they have not been exorcised by postulates but have rather been used to derive the classical world of our experience.”

Let us follow this line of reasoning by analyzing if it can account for the properties of consciousness. The starting point is the observation that the nervous system behaves like a complex non-linear system and therefore very sensitive to small variations in the initial conditions that are amplified in its evolution. As in chaotic phenomena, the sensitivity to initial conditions can be considered a kind of amplification of quantum fluctuations [38]. This quantum action can be seen as advantageous with respect to mechanisms that require quantum entanglement, superposition or tunneling, that seem to be mechanisms that are difficult to justify in a biological system like the brain. “Non-linear chaotic dynamics can amplify lowest-level quantum fluctuations upward, modulating even larger-scale mesoscopic and maybe also macroscopic neural activity” [38]. This type of approach has been reinforced by a set of recent observations that show that the neural functioning of the brain is close to criticality [39–42], which is confirmed by the distributions of avalanches and neural activity with a scaling that fits this regime and therefore its behavior is very sensitive to small perturbations. The states of the brain that exhibit critical behavior are the most suitable for more efficient neural processing information. The critical point characterizes the maximum capacity of exchanging information without bottlenecks, and as in every second order phase transition, it exhibits long range correlations of neural function. But the key property that hints that critical behaviors of the brain are key to give rise to consciousness is that close to the critical point the system is maximally sensitive to small fluctuations, for instance a change of one neuron can trigger an avalanche of activity [38, 39]. As Aaronson observes [33] “[B]rains seem “balanced on a knife-edge” between order and chaos: were they as orderly as a pendulum, they could not support interesting behavior; were they as chaotic as the weather, they couldn’t support rationality” It is expected that dynamical non-linear and chaotic phenomena could amplify underlying quantum fluctuations in such a way that the latter could modulate neural activity [43]. Both the information integrated in the complex neural network and criticality seem to play essential roles to underlie consciousness. However, in our understanding, they are not enough to explain the properties observed in conscious phenomena. Indeed, if decoherence prevents from having coherent effects in entangled systems sustained in time, the underlying information in neural activity ends up distributing itself in the environment as the quote by Schlosshauer we started this section with suggested. But the environment that interacts with the central nervous system includes the glial cells, the blood vessels and, microscopically, every molecule that constitute them. Summarizing, there is in this case no more structure that has the information of the neural activity than the complete brain, which on the other hand, is an open system. Think of the permanent blood flow needed to sustain its activity. To state that the brain, as any quantum system, has internal states and admits a first person access, does not help when the different structures would have incoherent internal perspectives. How to explain the unity
of consciousness, how to explain that consciousness refers only to information processed by neurons, what property of the brain would eliminate the background noise of the infinite macroscopic components that according to this hypothesis would be involved in an act of perception? What is the kind of action of an incoherent system allowing to fire a neuron, an event that needs hundreds of atoms? Criticality seem to be important to ensure large effects from the firing of a single network, but not from the action of isolated atoms or molecules. In spite of the difficulties that the existence of an entangled micro system protected from decoherence and interacting with the neural system would present, there exists no other apparent solution for the problem of consciousness.

**B. Fisher’s model of quantum cognition**

Fisher [44] has recently proposed a model of entangled micro systems, protected from decoherence effects, based on nuclear spins capable of interacting with the neural system. The model describes the entanglement of phosphorous spins that are part of Posner molecules that are transported to presynaptic neurons (a neuron from the axon terminal of which an electrical impulse is transmitted across a synaptic cleft to the cell body or one or more dendrites of a postsynaptic neuron by the release of a chemical neurotransmitter.) The Posner molecules can be decomposed producing postsynaptic neuron firings. Multiple entangled Posner molecules, triggering non-local quantum correlations of neuron firing rates, would provide the key mechanism for neural quantum processing. In Fisher’s own words: “It has long been presumed that quantum mechanics cannot play an important (functional) role in the brain, since maintaining quantum coherence on macroscopic time scales (seconds, minutes, hours, etc.) is exceedingly unlikely in a wet environment [45] (although see [46, 47] and references therein). Small molecules, or even individual ions, while described in principle by quantum mechanics, rapidly entangle with the surrounding environment, which causes de-phasing of any putative quantum coherent phenomena. However, there is one exception: Nuclear spins are so weakly coupled to the environmental degrees of freedom that, under some circumstances, phase coherence times of five minutes or perhaps longer are possible [48]. Putative quantum processing with nuclear spins in the wet environment of the brain as proposed by Hu and Wu in Ref. [46] would seemingly require fulfillment of many unrealizable conditions: for example, a common biological element with a long nuclear-spin coherence time to serve as a qubit, a mechanism for transporting this qubit throughout the brain and into neurons, a molecular scale quantum memory for storing the qubits, a mechanism for quantum entangling multiple qubits, a chemical reaction that induces quantum measurements on the qubits which dictates subsequent neuron firing rates, among others.”

Fisher starts by noticing that in the biochemical setting electric fields are the primary source of decoherence for nuclei with \( I > 1/2 \), while \( I = 1/2 \) spins are more weakly decohered only by magnetic fields. Thus, the element hosting the neural qubit must have nuclear spin \( I = 1/2 \). Besides hydrogen, only phosphorus has a nucleus with spin \( I = 1/2 \) among the most frequent biological elements. It is the unique biological element that can serve as a qubit while the phosphate ion works as a qubit transporter.

As Fisher puts it, “We identify the “Posner molecule”, Ca9(PO4)6, as the unique molecule that can protect the neural qubits on very long times and thereby serve as a (working) quantum-memory. A central requirement for quantum-processing is quantum entanglement. It is argued that the enzyme catalyzed chemical reaction which breaks a pyrophosphate ion into two phosphate ions can quantum entangle pairs of qubits. Posner molecules, formed by
binding such phosphate pairs with extracellular calcium ions, will inherit the nuclear spin entanglement. A mechanism for transporting Posner molecules into presynaptic neurons during vesicle endocytosis is proposed. Quantum measurements can occur when a pair of Posner molecules chemically bind and subsequently melt, releasing a shower of intra-cellular calcium ions that can trigger further neurotransmitter release and enhance the probability of postsynaptic neuron firing.”

This model provides most of the fundamental requirements for a quantum computer coupled to an information-integrated system working close to criticality. It identifies decoherence protected qubits that last for sufficiently long times. Recent studies give results which are consistent with the Fisher’s quantum cognition mechanism, but additional investigations are necessary [49]. It provides a mechanism of production of quantum entangled pairs and multiple entangled systems. A mechanism for transporting components of this multiple entangled systems to different regions of the brain is proposed. Neuron firings can be triggered via quantum event measurements. The net result would be having non-local quantum correlations of neuron firings.

Halpern and Crosson [50] have shown that Fisher’s model could satisfy most of the DiVincenzo [51] criteria for quantum computation. Besides the properties mentioned above the molecular mechanism proposed by Fisher’s model allows to perform a family of unitary transformations, to prepare a fiducial initial state, and measure the output result. Whether Posner quantum computation is universal remains an open question.

V. CONCLUSIONS

A conscious mind requires the efficient exchange of classical information: integration and criticality are preconditions for it. But these preconditions are not enough: consciousness’ unity cannot be explained classically. Classical properties are always Humean. They are mosaic-like. Classical laws are closed, there exists causal closure, which prevents any type of top-down causation and therefore the mental phenomena would not be more than epiphenomena. One needs an entangled quantum system that at least satisfies part of the functions of quantum computer to allow to generate an inner aspect with the unity of consciousness and to be coupled to a classical system that allows simultaneous access to the preprocessed information at the neural level. This allows for the production of events that generate neural firings. If such a system is not identified or allowed in a strongly interacting system like the brain, the problem of the identification of the physical basis of consciousness would be unsolvable without some form of new physics. Recent models offer hope that there exist ways of implementing these conditions in the brain.

VI. ACKNOWLEDGMENTS

This work was supported in part by Grants NSF-PHY-1603630, NSF-PHY-1903799, funds of the Hearne Institute for Theoretical Physics, CCT-LSU, Pedeciba and Fondo Clemente Estable FCE_1_2014_1_103803.


[34] M. Leifer, Quanta 3, 68 (2014).


